This is to certify that the
dissertation entitled

BACTERIAL DIVERSITY AND NUTRITIONAL SIGNIFICANCE OF THE SURFACE MICROLAYER IN *ANOPHELES GAMBIAE* (DIPTERA: CULICIDAE) LARVAL HABITATS

presented by

Shahnaz Rahim Maknojia

has been accepted towards fulfillment
of the requirements for the

Doctoral degree in Entomology

Edward D. Walker
Major Professor's Signature

3 December 2006
Date

BACTERIAL DIVERSITY AND NUTRITIONAL SIGNIFICANCE OF THE SURFACE MICROLAYER IN *ANOPHELES GAMBIAE* (DIPTERA: CULICIDAE) LARVAL HABITATS

By

Shahnaz Rahim Maknojia

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Entomology

2006

ABSTRACT

BACTERIAL DIVERSITY AND NUTRITIONAL SIGNIFICANCE OF THE SURFACE MICROLAYER IN *ANOPHELES GAMBIAE* (DIPTERA: CULICIDAE) LARVAL HABITATS

By

Shahnaz Rahim Maknojia

Anopheles gambiae and *Anopheles funestus* are well known vectors of human malaria in subSaharan Africa. The abundance of the adult stages of these mosquitoes necessarily depends on the number and productivity of the larval habitats. Even though these habitats are the source of these competent vectors, little is known about the productivity for larva habitat. Larvae are specialized gathering-filtering feeders and feed on materials in the surface microlayer which is enriched with microorganisms and other materials relative to subsurface zones of the water column.

In this study the bacterial composition of the surface microlayer, and its significance to larval nutrition and growth was examined. Removal of the surface microlayer at regular intervals resulted in decreased survival of larvae, prolonged developmental time to pupation, and produced adults with lower body mass. Supplementations of the surface microlayer from habitats with no larval grazing improved larval growth, shortened larval developmental time, and produced adults with higher total mass. Importance of heterotrophic bacteria in relation to larval nutrition was studied and it was observed that larvae grew, molted, and achieve metamorphosis to pupation when heterotrophic bacterial growth was enhanced by addition of glucose, but larval survival rate and total adult emergence was very low compared to sunlit treatments

rich in algae. Glucose addition to sterilized habitats resulted in complete growth failure of larvae.

Effects of larval grazing pressure on bacterial communities was studied with two different soil types using 16S rDNA sequence library construction and Terminal Fragment Length Polymorphism (TRFLP) analysis. Community shifts were observed either by presence or absence of certain taxa, or changes in the frequencies of certain taxa, as represented by the sequence data.

Finally, a culture-independent survey of bacteria present in the surface microlayer of natural *An. gambiae* and *An. funestus* larval habitats in western Kenya was undertaken. Overall both *An. gambiae* and *An. funestus* larval habitats were very diverse and revealed few dominant and many uncommon taxa. LIBSHUFF analysis revealed that these communities were statistically different, but Principal Component Analysis (PCA) on the sequence data and TRFLP analysis did not show any significant clustering of specific habitats. Therefore, there was no clear evidence supporting habitat segregation based on bacterial community structure.

ACKNOWLEDGEMENTS

I would like to thank and acknowledge the NIH grant (AI50703) for funding this project. I am grateful to the Department of Entomology at Michigan State University for all the support and resources.

I thank my major advisor Dr. Edward Walker for his encouragement, support, and patience and for always being available when needed throughout my dissertation research. I would also like to thank the members of my advisory committee, Dr. James Miller for his advice not just pertaining to my dissertation but also surviving graduate school, Dr. Michael Bagdasarian for his comforting conversations, Dr. Orlando Sarnelle for guiding my thoughts in the right direction, and Dr. Frank Dazzo for being critical about every small detail. My deep thanks to Dr. Michael Kaufman for his guidance in setting up experiments, help in analyzing data specially PCA analysis and for always being available for help. I would also like to acknowledge Dr. Terry Marsh for his guidance on TRFLP analysis. I express thanks to Dr. Karim Maredia and Dr. Mywish Maredia for their immense support right from the beginning and for all the encouragement.

I thank Mr. William Morgan and Mr. Blair Bullard for their assistance whenever it was needed. I thank John Wertz for his enormous help in learning ARB, Kristin Huizinga for sharing her experience with DOTUR, Lane LaFleur Frasier and Max Foster for their help with the microscopic analysis. I also thank Shicheng Chen for being a wonderful laboratory companion. Thanks to Susannah Cooper, Alicia Bray and Kirsten Pelz-Stelinski for being great friends.

I would like to extend my earnest and heartfelt gratitude to many people to whom I owe this dissertation. This dissertation would have not been possible without them. I thank my husband Rahim Hasanali Maknojia for his love and encouragement. While working on my dissertation I was blessed with two beautiful children, Zoe Rahim Maknojia and Izaan Rahim Maknojia, fulfilling and enriching my life.

I offer my heartfelt gratitude to my father-in-law Hasanali Raje Maknojia and my grandparents Mr. Hasanali and Mrs. Kulsum Hasanali Maredia, for what I am today I owe it to them. I thank my parents Mr. Dilawar and Mrs. ShahSultana Prasla, my uncles Nooruddin, Pyarali and Imtiyaz Maredia, and my aunt Mrs. Mumtaz Dhuka, for all their encouragement and for always being proud of me. I specially thank my mother for taking care of my daughter, while I was working on my dissertation. I would also like to remember my dear friend Rachana Bheda for believing in me and being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Literature Review

The Malaria Problem

The World Health Organization estimates that, yearly, 300-500 million cases of malaria occur and more than 1 million people die of malaria (Breman et al. 2001). About 1,200 cases of malaria are diagnosed in the United States each year. Most cases in the United States are in immigrants and travelers returning from malaria-risk areas, mostly from sub-Saharan Africa and the Indian subcontinent. Malaria occurs in over 100 countries and territories. More than 40% of the people in the world are at risk. Large areas of North, Central and South America, Hispaniola, Africa, the Indian subcontinent, Europe, South Asia, the Middle East, and Oceania are considered malaria-risk areas, although malaria was formerly a health problem in most of these regions, including temperate areas such as Michigan.

Malaria is a disease syndrome caused by infection with protozoan parasites of the genus *Plasmodium.* Four species of *Plasmodium* can produce the disease: *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale, Plasmodium malaria. P. falciparum* is the most widespread and dangerous of the four; untreated it can lead to fatal cerebral malaria. Unfortunately, it is also the most common in Africa. Malaria parasites are transmitted by bites of the female *Anopheles* mosquitoes. The discovery of this relationship was accomplished by Sir Ronald Ross in 1899, an accomplishment for which he was awarded the Nobel Prize in 1902. Giovanni Battista Grassi simultaneously

discovered the biological role of *Anopheles* mosquitoes as "vectors" (i.e., biological hosts capable of transmission), but ironically he did not share in the award. Infective sporozoite stages are transferred to the human circulatory system when the mosquito salivates into the skin, prior to when it begins to imbibe blood. There are about 760 species of anopheline mosquito, but only 60 or so are competent to transmit the parasite (Budiansky 2002). Their competency derives from innate genetic attributes and biological associations such as host selection that enhance population capacity for transmission. Malaria kills over one million people each year, most of whom are children under 5, and almost 90% of whom live in Africa, south of the Sahara (WHO 1998). So important is malaria that it has become a medical, research and public health specialty.

Malaria is such an enormous public health problem with clear negative affects on individual human health and on socioeconomic conditions that there are active attempts to suppress it. In general, vaccines and drugs are important in prevention and treatment of an illness. So far development of malaria vaccine has achieved little success and there are no malaria vaccines available. There is a wide spectrum of antimalarial drugs (e.g.; chloroquine) used for treatment and prophylaxis; but, these parasites rapidly evolve resistance to them. Drugs such as quinine, sulfadoxine pyrimethamine (Fansidar) and artemether are prescribed for populations but they are expensive and can be toxic if dose is not carefully controlled.

As it has long been said, prevention is better than cure; for malaria, protection can be afforded by minimizing mosquito bites. Mosquito bites can be reduced in frequency by using measures to avoid exposure during peak mosquito activity, using window screens and pyrethroid impregnated bed nets. Mosquitoes can be repelled using devices

or chemical (DEET) repellents. These above efforts should be coupled with measures to lower the mosquito densities to obtain effective malaria control. Mosquito densities can be controlled using either chemical or biological control agents targeted against larvae or adults. For adults insecticides (e.g.; DDT) are applied to the resting surfaces and these surfaces may retain their toxicity for days to months. Other insecticides include organophosphates, carbamates, pyrethrins, etc. But again development of resistance to insecticides (Hemingway et al. 2002) and change in adult feeding and resting behavior may render these chemicals ineffective. The rich and often dramatic history of organized mosquito control to control malaria has been well summarized (Bruce-Chwatt 1988).

Larval control (source reduction) can be obtained by habitat modification preventing oviposition, hatching or larval development. Biological control can be obtained by use of predators (e.g.; fish) or pathogens (fungal, viral and bacterial) but they are still under development or have limited effectiveness, with the exception of *Bacillus thuringiensis israelensi* (Bti). Synthetic chemicals are used as larvicides and these include organophosphates, juvenile hormone mimics. Adult resting and feeding places may in some cases be altered for control, but overall suppressing production of adults itself is an element of control since production of adult *Anopheles* from larval habitats is a key factor in malaria risk.

The world-wide attempt to eradicate malaria by indoor house residual spraying of insecticides showed dramatic results initially, but it lacked sustainability as it relied solely on one method of control. It has been realized and known from failures of eradication programs that only one method cannot be employed for effective malaria control but a more realistic control program which is integrated pest management using

combination of methods to reduce mosquito abundance and disease prevalence is required. For example, in the Indian subcontinent and in Sri Lanka, malaria programs were so successful that malaria cases were brought to levels below detection by public health surveillance; but when the programs were eliminated owing to shifts in funding priorities, epidemics ensued and malaria quickly re-established (Collins and Paskewitz 1995). It is clear that unless socioeconomic changes including increased standard of living and improved health care accompany malaria control programs, then the conditions for transmission remain and programs are not ultimately viable. Economic losses due to malaria in human populations are such that an easy argument can be made to control malaria on the basis of economic productivity alone. However, given that poor and underprivileged populations are often the beneficiaries of malaria control, the political will to direct scarce resources to these populations is often lacking.

Biology of mosquitoes

The biology and public health significance of mosquitoes has recently been reviewed (Foster and Walker 2002). Mosquitoes are a highly diverse group, with some 3,000 species worldwide among 38 genera distributed into 3 subfamilies and 10 tribes. Their higher classification places them in the Nematocera, the most primitive suborder of the Diptera or true flies. Their geographic distribution ranges from the high arctic to the tropics. Most species occur in the tropics, but mosquito abundance can be extremely high in arctic settings and temperate latitudes.

The generalized mosquito life cycle involves both aquatic and terrestrial environments and is one of complete metamorphosis with four separate and distinct

stages - egg, larva, pupa, and adult. The adults are terrestrial whereas the egg, larva, and pupa stages are strictly aquatic, but their habitats are highly variable. The larval stage is characterized by four instars each separated by a molt followed by a period of feeding and growth. The final molt yields the pupa which is a motile but non-feeding stage of short duration. Those of the anophelines of interest mentioned above are typically water habitats, generally soil or mud in substrate, and are often small and transient.

Unlike other mosquito larvae, *Anopheles* larvae do not have a respiratory siphon but instead have one pair of spiracles located at the terminal abdominal segment. Thus the larvae position themselves parallel to the surface of water to breathe. Correspondingly, *Anopheles* larvae feed at the air water interface and are specialized gathering-filtering feeders, rotating the head 180° degrees from the normal position and directing the rapidly beating mouthparts to the air-water interface (Merritt et al. 1992). This mode of feeding differs substantially from that of *Aedes* (Merritt et al. 1992, Walker and Merritt 1991) in that it is largely restricted to a single zone of the water column. The air-water interface (termed by limnologists the 'surface microlayer') of *Anopheles* habitats is a zone that is enriched with microorganisms and other materials relative to subsurface zones of the water column (Walker and Merritt 1993). The specialized surface feeding behavior of these larvae upon the surface microlayer – exhibited by some other aquatic invertebrates as well might very well be an adaptation to exploit this food-rich region.

Each mosquito species has particular environmental requirements for the maintenance of its life cycle, and these requirements define the larval habitat. The adult female mosquitoes show distinct preferences for oviposition sites (i.e.; egg-laying sites),

and this preference may be a major determinant of larval mosquito distribution in nature. Thus, mosquitoes of different species may show habitat overlap or habitat segregation, sometimes along both temporal and spatial scales. In Kenya, studies have shown that the Anophelines differ in their breeding, feeding and resting habits (Gimnig et al. 2001, Mutero and Birley 1987).

Ecology of *Anopheles* mosquitoes of Africa

The anophelines of the world include nearly 760 species of which 60 or so can transmit malaria. Many of the species in Africa are highly efficient vectors of human malaria, in particular species of the *Anopheles gambiae* and *An. funestus* complexes. Within the gambiae complex, *An. gambiae and An. arabiensis* have the widest distribution in sub-Saharan Africa (Coluzzi 1984). *An. gambiae* and *An. arabiensis* breed in fresh water and are associated with small habitats often created by man or animal activity, such as foot or hoof prints, burrow pits, roadside puddles formed by tire tracks, and irrigation ditches. *An. gambiae* habitats are turbid and persist for short periods and lack aquatic vegetation (Gimnig et al. 2001, Gilles and Coetzee 1987). *An. funestus* on the other hand breed in large, semi-permanent bodies of water, characterized by emergent vegetation, such as swamps, river edges and ditches. In addition *An. funestus* is one of the most 'domestic' of African anophelines, preferring to feed and rest inside human houses for most of its adult life. Even in human dwellings shared with cattle, most female *An. funestus* bite humans and congregate in sections occupied by humans. The biogeography and habitat associations of these species are currently rather poorly known and are under intensive study.

The *An. gambiae* complex includes 7 members: *An. gambiae, An. arabiensis, An. quadriannulatu species A and species B, An. merus, An. melas,* and *An. bwambae* (White 1974, Coluzzi 1984). All the species of the so called "*gambiae* complex" are morphologically indistinguishable, but can be distinguished through chromosomal analysis and with molecular methods. The discovery of this complex of species is fascinating because at one time all were considered to be the same species, *Anopheles gambiae* (Giles) sensu lato. However, institution of the malaria eradication campaign in the 1950s and early 1960s revealed subtle variations in ecology and behavior, especially that related to the relative extent to which local populations fed upon humans (anthropophily) or cattle (zoophily) and the location of larval habitats (brackish water, fresh water, or mineral springs). Simultaneously with these observations was the development and application of chromosomal banding and cytology methods to discern so-called chromosomal forms with characteristic inversion morphologies which would create genetic barriers to fertilization, and so function as post-mating isolation barriers. More recent studies have revealed population genetic substructuring within *An. gambiae* s.s. in which chromosomal ecoforms (Toure et al. 1998) and so-called M and S mitochondrial forms have been elucidated in west Africa (della Torre et al. 2001, 2002). Overall, the history of the discovery of closely-related species in the *An. gambiae* complex represents a story in evolutionary biology suggestive of on-going speciation through adaptive radiation to blood hosts and larval environments. The latter context is provided by the human living environment. In Kenya, studies have shown that two of the sibling species that commonly co-occur, *An. gambiae s.s.* and *An. arabiensis*

(Koenraadt et al. 2004), differ (sometimes subtly) in their breeding, feeding and resting habits (Gimnig 2001, Mutero et al. 1987).

A similar story is emerging for the other important African malaria vector, *Anopheles funestus*. Although originally thought to be a single species, it is now clear that it represents a species complex comprised of at least nine members which also show variations in host utilization, larval habitat utilization, and comparative vectorial capacity for malaria (Kamau et al. 2002). The members of this complex include *An. funestus, An. vaneedeni, An. parensis, An. aruni, An. confuses, An. leesoni,* and *An. brucei.* Of the nine species in the complex, *An. funestus* has the widest distribution. It is a highly efficient vector of malaria owing to its dependency on humans for blood and the strong predilection of males and females to rest indoors.

An. gambiae predominate in humid areas with larval production occurring during the rainy periods, whereas *An. arabiensis* are also found in arid areas and are likely to reproduce year round (Takken et al.1998). *An. gambiae* are usually associated with human dwellings whereas *An. arabiensis* are found in habitats near cattle (Gimnig et al. 2001). *An. arabiensis* will feed on animals when available and feed on man when alternative hosts are not available. These two species may occur within the same habitat though survival rates of *An. gambiae* are higher compared to *An. arabiensis*, suggesting interspecific competition between these two closely related species (Schneider et al. 2000). The latter study was confined to laboratory conditions, however, thus the significance of interspecific competition in field populations is poorly known. Gimnig et al. (2002) demonstrated intraspecific competition processes with *An. gambiae* s.s. in artificial field habitats where larval densities were similar to those found in nature.

Observations of cannibalism amongst larvae of this species indicate that competition for food resources may be extreme (Koenraadt and Takken 2003).

Nutrition

Knowledge about mosquito nutrition and diet come from development of fully chemically defined diets developed for *Culex pipiens* and *Aedes aegypti* larvae. *Culex pipiens* larvae could be reared to adults with these defined diets with 90% or more survival and at growth rates comparable to those obtained by unrefined diets (Clements 1992). The ten amino acids (arginine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan and valine) essential for growth of other insects and many vertebrates are also required by mosquitoes although omission of proline in *Aedes aegypti* diet even though a non-essential amino acid retarded larval development (Dadd 1985). *Culex pipiens* requires asparagine in addition to the ten essential amino acids (Dadd 1978).

All insects require C_{20} polyunsaturated fatty acids for various body functions such as use of fat soluble vitamins, as constituents of membrane glycerophospholipids, and as precursors for prostaglandin synthesis. Nearly all insects can synthesize C_{20} from C_{18} polyunsaturated fatty acids but mosquitoes are different than other insects in this regard. They cannot synthesize C_{20} from C_{18} polyunsaturated fatty acids, because they lack the necessary enzyme systems to do so. In mosquitoes, the lack of C_{20} polyunsaturated fatty acids in their diets renders the adults unable to fly (Clements 1992). Inclusion in the mosquito diet of arachidonic acid or other fatty acids (such as eicosapentanoic acid)

which have a series of four double bonds of *cis* configuration, terminating at the n'6 position in a C_{20} or C_{22} positions, remediates this problem.

Vitamins are required by all insects including mosquitoes. Vitamin B6 complex are required by *Ae. aegypti* for survival of first instars and also for pupation. This is true also for vitamin A, though its absence does not have detrimental effects on *Ae. aegypti*; its absence in diet causes abnormalities in the receptor cells of the adult eye. Choline is required by all insects and also mosquitoes because it is a neurotransmitter and a component of some phospholipids, thus it is essential in their diets. Nucleotides are required for producing nucleic acids in all insects. Addition of only bases in their diet does not allow growth. Nucleotides for 'a', 'g', 'c' and 't/u' are required, though omission of 'g' does not alter the growth of the larvae. All insects including mosquitoes lack the ability to synthesize sterols and require an exogenous source of sterols such as cholesterol or various phytosterols (Clements 1992).

Temperature

The influences of temperature on the adult and aquatic stages of several species, including *An. gambiae* have been studied. The rates at which the new individuals are produced determine the growth rate of insect populations. Population growth rate critically depends on growth characteristics of immature stages. Growth rates can be influenced by temperature where food is not limiting. Larvae of *An. gambiae* develop into adults at temperature ranging from 16 to 34°C with reduced survival at temperatures between 38-40°C and 14-20°C (Bayoh and Lindsay 2004). Interestingly, the production of adult mosquitoes is not directly proportional to the rate of development of the aquatic

stages since at temperatures resulting in faster rate of larval development fewer adults were produced. *An. gambiae* developmental rate increased with temperatures between 22-28°C then declined. Even though adult developmental rate was highest between 28-32°C adult emergence was highest between 22-26°C (Bayoh and Lindsay 2003).

Aquatic Microbial Communities: Factors Affecting Structure and Dynamics

Larvae of the *Anopheles* species reviewed above have some obvious and characteristic habitat features and qualities, but the nutrient and microbial facets of these habitats that contribute to primary and secondary (including mosquito) production and to habitat differences among these *Anopheles* species are poorly known. My research proposal aims to address some of the aspects of these qualities and features, as detailed below. It is informed by studies of aquatic microbial community ecology in general, about which much is known and there exists an enormous literature ranging from pelagic ocean to stream and tree hole ecosystems. In general, autotrophic freshwater communities are thought to be organized around three predominant, shaping forces: the intensity of nutrient inputs ("bottom up effects"), the intensity of predation ("top down effects"), and the availability of sun light ("insolation") all of which dictate the flow of energy and structure of the food web (Cohen et al. 2003). These three interacting and sometimes counteracting forces will influence energy flow and biomass accumulation at each trophic level, and are determinants of the extent of primary and secondary production. Indeed, mesocosm studies and whole lake manipulation studies have suggested that aquatic management schemes could be developed on the basis of these ecological concepts. However, little of this research has been applied conceptually to an understanding of

mosquito aquatic habitat production; container habitats are a notable exception to this statement (Kaufman et al. 2002). In particular, there is currently a poor understanding of the interactions of nutrients, energy inputs, microorganisms and mosquito production. An overall theme that emerges from studies of aquatic communities, however, is that strong effects of predation on microbial communities limit secondary production. How relevant is this theme to mud puddle habitats of *Anopheles* mosquitoes in Kenya?

In shallow, turbid, mud-bottomed habitats of the kind occupied by *Anopheles* larvae in Kenya, Cooper et al. (1998) have suggested that algal production and abundance of zooplankton and insects is limited by phosphorus availability. This provides evidence that P could limit algal production, and in turn mosquito production, in Kenyan mud-lined pools of water. A bottom-up effect was clearly identified in these cases. Evidence for top-down effects is provided by more recent studies including some of my own, outlined below. In field microcosm experiments, Gimnig et al. (2002) showed that the most likely important food source for larval *An. gambiae* was algae, which were significantly reduced in the presence of larvae as measured by chlorophyll *a* in surface water samples and by counts of algae in sedimentation chambers, compared to when larvae were absent. In their first experiment, mosquito feeding reduced dry algal biomass by 100 μg/ml, or a total of 7 mg per habitat surface layer. Based upon total adult production of 2.8 and 7.5 mg in treatments with 20 and 60 larvae respectively, the 4.8 mg decline in algal biomass they measured as an indicator of larval grazing pressure on algae explains much larval production. Bacteria may have supplemented or supplanted algae under such feeding pressure as it may have formed a secondary food source, but bacterial densities were not affected significantly by the presence of larvae so the extent of grazing

mosquito aquatic habitat production; container habitats are a notable exception to this statement (Kaufman et al. 2002). In particular, there is currently a poor understanding of the interactions of nutrients, energy inputs, microorganisms and mosquito production. An overall theme that emerges from studies of aquatic communities, however, is that strong effects of predation on microbial communities limit secondary production. How relevant is this theme to mud puddle habitats of *Anopheles* mosquitoes in Kenya?

In shallow, turbid, mud-bottomed habitats of the kind occupied by *Anopheles* larvae in Kenya, Cooper et al. (1998) have suggested that algal production and abundance of zooplankton and insects is limited by phosphorus availability. This provides evidence that P could limit algal production, and in turn mosquito production, in Kenyan mud-lined pools of water. A bottom-up effect was clearly identified in these cases. Evidence for top-down effects is provided by more recent studies including some of my own, outlined below. In field microcosm experiments, Gimnig et al. (2002) showed that the most likely important food source for larval *An. gambiae* was algae, which were significantly reduced in the presence of larvae as measured by chlorophyll *a* in surface water samples and by counts of algae in sedimentation chambers, compared to when larvae were absent. In their first experiment, mosquito feeding reduced dry algal biomass by 100 μg/ml, or a total of 7 mg per habitat surface layer. Based upon total adult production of 2.8 and 7.5 mg in treatments with 20 and 60 larvae respectively, the 4.8 mg decline in algal biomass they measured as an indicator of larval grazing pressure on algae explains much larval production. Bacteria may have supplemented or supplanted algae under such feeding pressure as it may have formed a secondary food source, but bacterial densities were not affected significantly by the presence of larvae so the extent of grazing

on them was uncertain. The levels of nitrogen and phosphorus in the habitats were not affected by the presence of larvae although there was evidence for decreasing nitrogen levels with increasing larval densities suggesting that nitrogen may be a limiting resource in the larval environment. These trends find parallels in studies of other aquatic systems. For example, a series of studies in lake ecosystems in northern Europe has revealed marked predator-prey relationships, as follows. Daphnia fed upon both bacteria and their protozoan predators having the general effect of reducing abundance of both groups and severely dampening bottom-up nutrient effects. A fraction of the bacteria may have possessed some kind of grazing resistance that would buffer seasonal fluctuations and maintain high densities or the main bacteriovores might be top-down controlled. Due to rapid compensation of bacterial grazing losses by the development of resistant forms, the total bacterial biomass may remain less affected and could be maintained on a relatively high level (Jurgens 1994, Jurgens and Gude 1994). The planktonic bacterial communities in these studies responded both phenotypically and genotypically to strong grazing pressure by metazooplankton (Jurgens and Matz 2002), interacted strongly in their population dynamics to carbon and phosphorus limitations (Matz and Jurgens 2003), and were shaped by combined protozoan and metazoan predation (Langenheder and Jurgens 2001). These studies showed how microcosm and mesocosm experiments can support whole ecosystem studies and represented a model for long-term, integrated analysis of these relationships.

Methodological Approaches to Aquatic Microbial Community Structure

Estimates of microbial diversity within an environment can be divided into two broad categories: culture-dependent or independent methods. Culture-dependent methods involving isolation and cultivations are ideal, but they are not representative of the community and are slow and laborious. Very few (0.1%-1%) bacteria in the natural environment have been cultured as pure isolates in the laboratory. Possible reasons why naturally-occurring bacteria and other microbes have been proved difficult to grow in the laboratory include the following: many microbes lack the ability to adapt from severe oligotrophy to the high nutrient concentration of culture media. Many bacteria exist in consortia in natural assemblages, where inter and intra population interactions (including quorum sensing) are important. Our poor understanding of the basic physiology of many microbes makes it difficult to design appropriate culture media and thus it is clear that current isolation procedures will fail to adequately investigate the microbial diversity extant in natural environments (Zoetendal et al. 2004, Nubel et al. 1999). Also classical identification relies heavily on phenotypic characterization, including morphology, growth requirements and characteristics, fermentation profiles, cell wall protein analysis, serology and more recent fatty acid methyl ester (FAME) analysis. From the phenotypic characterization of certain lactic acid bacteria (cell wall profiling) and some species of *Bifidobacterium* (cell wall peptidoglycan analysis), there is general awareness that observation of a similar phenotype does not always relate to similar, closely-related, genotype (McCartney 2002). Also these methods have weaknesses like poor reproducibility, ambiguity in some techniques and poor discriminatory power.

A method for analysis of microbial diversity should involve evaluation of the total population-those that grow on laboratory media plus those that don't, those that have been described and those that never will be. Since the simple task to assess microbial diversity is insurmountable, a reasonable compromise to achieve this ideal objective would be to use a surrogate to assess microbial diversity. This surrogate indicator of microbial diversity would have to occur only in the living cells, vary in a meaningful way in relationship to the overall microbial diversity in the system, and be sufficiently variable that groups could be constructed based on similarity. While several cell components are informative, SSU rRNA genes are highly conserved among organisms and make them the best for studies of phylogenetic relationships of microbes in ecological studies (Hugenholtz et al. 1998). Molecular techniques therefore provide main source of information on the genetic and physiological diversity of environmental populations. While sequencing of the gene for the small subunit ribosomal RNA (16S) from natural assemblages is now routine, the 16S sequences in databases rarely match culturable microbes, instead they usually have the closet affinity with sequences from other uncultured organisms (Lu et al. 2003). The delineation of species on basis of morphologies, even though, the most common practice does not necessarily result in evolutionarily and ecologically coherent entities, particularly when applied to microorganisms. Microbial phylogeneticists in particular depend on molecular sequence characters, because prokaryotes offer relatively little in the way of complex morphology and behavior (Doolittle 1999). Numerous limitations inherent in the various methodologies currently available, microbial diversity has not yet been convincingly reported or it would be at present impossible (Nubel et al. 1999). The determination of

prokaryotic species richness and diversity in nature is impracticable. Depending on the research objectives, it may be more fruitful to take into account the organisms specific identities and their ecologically relevant properties. However molecular techniques are not without limitations, thus a polyphasic approach might often be necessary and is perhaps the most desirable among the range of options (Hugenholtz et al. 1998).

Current molecular techniques used to study microbial ecosystems can be separated on the bases of their wide use. Clone libraries can be sequenced to identify the composition of microbiota down to the species level. Microbial community structure can be analyzed via fingerprinting techniques, while FISH and dot blot hybridization can be used to measure abundance of a particular taxa. Also there are approaches based on functional genes and their expression and the use of stable isotopes and biomarkers that are being optimized to study metabolic activities of groups or individual organisms in situ. (Zoetendal et al. 2004). 16S rDNA sequencing relies on sequences of rRNA genes obtained by cloning directly from environmental DNA or, as in the majority of studies, after amplification by the PCR (Hugenholtz et al. 1998). Although 16S rDNA analysis represents a very useful technique for culture-independent analysis of complex microbial communities, and it gives significant information about the identity of uncultured bacteria, the clone frequencies in the clone libraries do not reflect the in situ quantities of the respective microorganisms (Eschenhagen et al. 2003). This approach is more qualitative than quantitative. Possible reasons are differences in the numbers of rRNA operons in different microbes, different efficacies of cell lyses and DNA extraction, or shifts due to PCR amplification. Sequencing of SSU rRNA genes has become a standard procedure in the identification of isolates. Currently, >79,000 16S rRNA sequences are

available in DNA databases, which is far greater than for any other gene (Zoetendal et al. 2004). Studies have shown that many of the 16S rDNA sequences exhibit low sequence similarity to genes of known cultured bacterial genera (Lu et al. 2003). This method has also helped in developing cultivation strategies for many previously unknown or uncultivated bacteria (butyrate-producing, cellobiose-degrading bacteria) (Zoetendal et al. 2004).

Denaturing gradient gel electrophoresis (DGGE) and terminal-restriction fragment length polymorphism (T-RFLP) are community fingerprinting techniques which are widely used to monitor communities over time or in response to dietary treatments (Juck et al. 2000; Duatre et al. 2001). DGGE is PCR-based and generates profiles representing the sequence diversity within the selected ecosystem. The general principle of DGGE/TGGE is separation of individual rRNA genes based on differences in chemical stability or melting temperature of these genes (McCartney 2002). The total number of bands visualized in a DGGE gel also provides an estimate of the genetic diversity within a given environment. T-RFLP is based on specific target site for restriction enzymes (Eschenhagen et al. 2003). Profiling is based on the banding patterns obtained from DNA restriction digests. The method shows good sensitivity and has been employed for environmental studies including bacterial soil communities and comparative community analysis (McCartney 2002). For T-RFLP, 16S rRNA sequences have been extensively used as markers. Non-16S rRNA profiling approaches based on cellular fatty acid composition or G + C content of DNA have also been used successfully to monitor shifts in bacterial communities in an ecosystem. But the disadvantage to these methods is the lack of phylogenetic databases.

Although TRFLP fingerprints can show the dominant groups and the ones which are rare it provides only semi quantitative information about the abundance of the respective groups in the community. To estimate the abundance of particular taxa a direct approach is more appropriate. Fluorescent in situ hybridization (FISH) is used to quantify bacterial cells in the environmental samples. This method involves use of fluorescent-labeled oligonucleotide probes targeted towards SSU rRNA and visualization using epiflourescent light microscopy. Several probes have been developed to quantify bacteria of various domain and genera. This method is widely used to quantify bacterial cells in an ecosystem like marine arctic sediments (Ravenschlag et al. 2001), for characterizing nitrifying bacteria in biofilms (Gieseke et al. 2001). Limitations of this method are that it is dependent on SSU rDNA sequences available in the databases, and that only a few probes can be used per analysis (Zoetendal et al. 2004). This method also depends on the permeability of the bacterial cells, accessibility of the target and number of ribosomes per cell. This technique is reliable and relatively easy to use, but probe design and detection limits are the main disadvantages. Such protocols are useful for bacterial groups or predominant genera but an extensive list of probes would be necessary for diversity studies. Dot-blot hybridization is used to quantify specific 16S rRNA in a mixture relative to the total rRNA. It involves probing DNA or RNA extracts from bacterial isolates or environmental samples. Total RNA is isolated from the sample, bound to a filter, and hybridized with labeled oligonucleotide probes. The relative concentration is estimated by dividing the concentration of specific probe by the concentration of the universal probe, however it is important to note that such quantification are only relative.

Purpose/Scope of the Project

Adult *An. gambiae* density depends on the number and collective productivity of the larval habitats. Even though these habitats are the source of these competent vectors, little is known about the larval biology of these important insects, in particular, what forms the basis for production of insect biomass in them. According to the study conducted by Gimnig et. al. (2002), the most likely important food source for larval *An. gambiae* seem to be algae, which were significantly reduced in presence of larvae as measured by chlorophyll *a* in the surface water samples and by counts of algae in sedimentation chambers, compared to with when larvae were absent. Other experiments support this study and report that presence of algae does improve the growth and development of the larvae, with shorten developmental time, and increased survival over that when algae were absent, a condition obtained by shading the habitats (Kaufman et al. 2006). But this ignores algal turnover rates, larval consumption and other nutritional sources (i.e.; bacteria and non cellular organics in the surface layer). Algae probably play a key role as food resources in *An. gambiae* habitats and bacteria may likely form a secondary food source. Studies on larval gut analysis reveal that bacteria form the bulk of the food bolus, along with organic particulates, algae, and small invertebrate parts (Walker et al. 1988). Of interest is that larval mosquito feeding on bacteria in their habitats has been estimated to account for production of a large fraction of the insect biomass (Kaufman et al. 2001), but simultaneously, larval mosquito feeding reduces microbial abundance, alters the microbial community qualitatively and quantitatively, and diminishes the quality of the microbes as food (Kaufman et al. 2000, 2002; Xu et al. in preparation). Thus it is essential to conduct studies to determine the contribution of

bacteria to *Anopheles* larval nutrition. Bacteria along with other organisms that occur in larval habitats may form the basis for production of adults, by transforming inorganic nutrients through autotrophic and heterotrophic processes into assimilable forms, by providing food and specific larval nutritional demands (Walker et al 1988, Walker and Merritt 1993, Kaufman et al. 2000, 2001, 2002).

Further, microbial communities in larval mosquito habitats have been postulated to influence ovipositional site selection by mosquitoes owing to volatile organic compounds emitted from these communities, and owing to microbially-derived flavors in water (Trexler et al. 2003). This research proposal cleaves to that hypothesis, and applies it to the *An. gambiae* and *An. funestus* habitats currently under study in western Kenya by my advisor and colleagues (Gimnig et al. 2001, 2002). Location and selection of an ovipositional site is an essential part of the life history of all mosquito species. The location and selection of an ovipostional site involves visual, olfactory, and tactile responses. Intensive field studies have shown that mosquitoes are quite discriminating in selecting sites for egg deposition and considerable evidence points to this site discriminating larval distribution (Bentley and Day 1989). Oviposition site selection is the net result of the interaction of a complex array of both chemical and physical factors. These attracting substances may, in some cases be pheromones produced by the larvae, but, studies have shown that concurrent presence of *A. gambiae* larvae reduced oviposition, while turbid water from natural breeding sites increased oviposition selectivity. Thus other possible attractants such as microbial metabolites must also be considered keeping in mind that mosquito biting on the human body is also mediated by

microbes. *An. gambiae*, a human biter, prefers to bite human feet and is attracted to odors emitted by bacteria from feet, such as isovaleric acid (Braks et al. 1999).

Notably, *An. gambiae* and *An. funestus* habitats are segregated and distinct in Kenya (Gimnig et al. 2001). The bacterial communities are postulated to be distinct between these habitat types. These differences may be crucial in differential ovipositional site selection by females of these two species, and the differences may account for habitat production for adult mosquitoes. Since very little information exists on the bacterial community, it is essential to study it and also to determine what comprises the larval food. By conducting a comparative analysis, one can also determine what makes one habitat more favorable for one species (*An. gambiae* and *An. arabiensis*) and the other habitat for other species (*An. funestus*). Further, experimentation in which the surface microlayer and its microbial components are intentionally stripped away may reveal the importance of those components to mosquito feeding and growth. Therefore, the objectives of my study are 1) to analyze the effects of removal of the surface microlayer on mosquito growth. 2) to study the contribution of bacteria to larval nutrition and 3) to determine the bacterial diversity of the surface microlayer in presence and absence of larval feeding pressure, and 4) to determine the bacterial diversity of *An. gambiae* and *An. funestus* habitats using 16S rDNA sequence library construction.

REFERENCES

Bayoh M. N., Linsay S. W. 2004. Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito *Anpheles gambiae* in the laboratory. Medical and Veterinary Entomology. 18: 174-179.

Bayoh M. N., Linsay S. W. 2003. Effects of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera: Culicidae). Bulletin of Entomology Research. 93: 375-381.

Bentley M.D., Day J.F. 1989. Chemical ecology and behavioral aspects of mosquito oviposition. Annual reviews in Entomology. 34:401-421

Braks, M.A.H., Anderson, R.A., Knols, B.G.J. 1999. Infochemicals in Mosquito Host Selection: Human Skin Microflora and *Plasmodium* Parasites. Parasitology Today. 15(10): 409-413.

Breman, J. G., A. Egan and G. T. Keusch. 2001. The intolerable burden of malaria: a new look at the numbers. American Journal of Hygiene and Tropical Medicine. 64 (supplement 1-2): iv-vii.

Bruce-Chwatt, L. J. 1988. History of malaria from pre-history to eradication. p. 1-59 *In* I. A. McGregor and W. Wernsdorfer (ed.), Principles and practice of malariology. Churchill Livingstone, Edinburgh.

Budiansky S. 2002. Creatures of our own making. Science. v298 i5591:80-85.

Clements, A. N. 1992. Larval nutrition, excretion and respiration. p. 100-123 *In* The Biology of Mosquitoes. Volume 1: Development, Nutrition and Reproduction. Chapman & Hall, London.

Cohen J. E., J. T., Carpenter S. R. 2003. Ecological community description using the food web, species abundance, and body size. Proceedings of the National Academy of Sciences. 100(4): 1781-1786.

Collins F. H., Paskewitz S.M., 1995. MALARIA: Current and future prospects for control. Annual reviews in Entomology. 40: 195-219

Coluzzi M. 1984. Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. Bulletin of the world health organization. 62: 107-113.

Cooper S. D., Diehl S., Kratz K., Sarnelle O. 1998. Implications of scale for patterns and processes in freshwater ecology. Australian Journal of Ecology 23:27-40.

Dadd R. H. 1978. Amino acid requirements of the mosquito *Culex pipiens*: asparagine essential. Journal of Insect physiology. 24: 25-30

Dadd R. H. 1985. Nutrition: *In* Comprehensive Insect physiology biochemistry and pharmacology, eds Kerkut G. A. and Gilbert L. I., Pergamon press, Oxford. 4: 313-90

della Torre A, Costantini C, Besansky NJ, Caccone A, Petrarca V, Powell JR, Coluzzi M. 2002. Speciation within Anopheles gambiae--the glass is half full. Science 298(5591):115-117.

della Torre A, Fanello C, Akogbeto M, Dossou-yovo J, Favia G, Petrarca V, Coluzzi M. 2001. Molecular evidence of incipient speciation within *Anopheles gambiae s.s.* in West Africa. Insect Molecular Biology. 10(1): 3-7.

Doolittle, W.F. 1999. Phylogenetic Classification and the Universal Tree. Science. 284(5423): 2124-2128.

Duatre G. F., Rosado A. S., Seldin L., Arujo W., Elsas J. D 2001. Analysis of bacterial community structure in sulfurous-oil-containign soils and detection of species carrying dibenzothiophene desulfurization (*dsz*) gene. Applied and Environmental Microbiology 67(3): 1052-1062.

Eschenhagen M., Schuppler M., Roske I. 2003. Molecular characterization of the microbial community structure in two activated sludge for the advanced treatment of domestic effluents. Water Research 37: 3224-3232.

Foster, W. A. and E. D. Walker. 2002. Mosquitoes (Culicidae), Chapter 12. Pp.204-262 *In* G. Mullen and L. Durden (eds.), Medical and Veterinary Entomology. Academic Press, New York.

Gieseke A., Ulrike P., Wagner M., Amann R., Schramm A. 2001. Community Structure and Activity Dynamics of Nitrifying Bacteria in a Phosphate-Removing Biomfilm. Applied and Environmental Microbiology 67(3): 1351-1362.

Gillies M. T., and Coetzee M. 1987. A supplement to the Anophelines of Africa south of Sahara. South African Institute for Medical Research, No. 55: 143.

Gimnig J., Ombok M., Kamau L., Hawley W. 2001. Characteristics of larval Anopheline (Diptera: Culicidae) Habitats in Western Kenya. Journal of Medical Entomology 38(2): 282-288.

Gimnig, J. E., Ombok M., Otieno S., Kaufman M., Vulule J., Walker E. D. 2002. Density-dependent development of *Anopheles gambiae* larvae in artificial habitats. Journal of Medical Entomology. 39: 162-172.

Hemingway. J., Field L., Vontas J. 2002. An overview of insecticide resisitance. Science 298: 96-97

Hugenholtz P., Goebel B. M., Pace N. R. 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. Journal of Bacteriology 180(18): 4765-4774.

Juck D, Charles T., Whyte LG, Greer CW. 2000. Polyphasic microbial community analysis of petroleum hydrocarbon-contaminated soils from two northern Canadian communities. FEMS Microbiology Ecology. 33(3): 241-249.

Jurgens K. 1994. Impact of Daphnia on planktonic microbial food webs – A review. Marine Microbial Food Webs. 8(1-2): 295-324.

Jurgens K., Gude H. 1994. The potential importance of grazing-resistant bacteria in planktonic systems. Marine ecology progress series 112: 169-188.

Jurgens K., Matz C. 2002. Predation as a shaping force for the phenotypic and genotypic compostion of planktonic bacteria. Antonie van Leeuwenhoek 81: 413-434.

Kamau L., Hunt R., Coetzee M. 2002. Analysis of the population structure of *Anopheles funestus* (Diptera: Culicidae) from Western and Coastal Kenya using paracentric chromosomal inversion frequencies. Entomological Society of America. 39(1): 78-83.

Kaufman, M.G., E.D. Walker, D.A. Odelson, & M.J. Klug. 2000. Microbial community ecology and insect nutrition. American Entomology. 46: 173-184.

Kaufman, M.G., S.J. Bland, M.E. Worthen, E.D. Walker, & M.J. Klug. 2001. Bacterial and fungal contributions to growth of larval *Aedes triseriatus* (Diptera: Culicidae). Journal of Medical Entomology. 38: 711-719.

Kaufman, M.G., Goodfriend W., Kohler-Garrigan A., Walker E.D., Klug M.J. 2002. Soluble nutrient effects on microbial communities and mosquito production in *Ochlerotatus triseriatus* (Say) habitats. Aquatic Microbial Ecology 29: 73-88.

Kaufman M. G., Wanja E., Maknojia S. R., Bayoh N. M., Vulule J. M., Walker E. D. 2006. The importance of algal biomass to the growth and development of *Anopeheles gambiae* larvae. Journal of Medical Entomology. 43: 669-676.

Koenraadt C. J. M., Takken W. 2003. Cannibalism and predation among larvae of the *Anopheles gambiae* complex. Medical and veterinary Entomology. 17: 61-66

Koenraadt, C.J.M., Githeko, A.K., Takken, W. 2004. The effects of rainfall and evapotranspiration on the temporal dynamics of *Anopheles gambiae s.s.* and *Anopheles arabiensis* in a Kenyan village. *Acta Tropica.* 90: 141-153

Langenheder S., Jurgens K. 2001. Regulation of bacterial biomass and community structure by metazoan and protozoan predation. American Society of Limnology and Oceanography 46(1): 121-134.

Lu J., Idris U., Harmon B., Hofacre C., Maurer J. J., Lee M. D. 2003. Diversity and succession of the intestinal bacterial community of the maturing broiler chicken. Applied and Environmental Microbiology 69(11): 6816-6824.

Matz C., Jurgens K. 2003. Interaction of nutrient limitation and protozoan grazing determines the phenotypic structure of a bacterial community. Microbial Ecology 45: 384-398.

McCartney, A. L. 2002. Application of molecular biological methods for stydying probiotics and the gut flora. British Journal of Nutrition. 88(1): S29-S37.

Merritt, R.W., D.A. Craig, E.D. Walker, H.A. Vanderploeg, & R.S. Wotton. 1992. Interfacial feeding behavior and particle flow patterns of *Anopheles quadrimaculatus* larvae (Diptera: Culicidae). Journal Insect Behavior 5: 741-761.

Mutero C. M., Birley M. H. 1987. Estimation of the survival rate and oviposition cycle of the field populations of malaria vectors in Kenya. Journal of Applied Ecology. 24: 853-863.

Nubel U., Garcia-Pichel F., Kuhl M. , Muyzer G. 1999. Quantifying Microbial Diversity: Morphotypes, 16S rRNA Genes, and Carotenoids of Oxygenic Phototrophs in Microbial Mats. Applied and Environmental Microbiology 65(2): 422-430.

Ravenschlag K., Sahm. K., Amann R. 2001. Quantitative Molecular Analysis of the Microbial Community in Marine Arctic Sediments (Svalbard). Applied and Environmental Microbiology 67(1): 387-395.

Schneider P., Takken W., McCall P. J. 2000. Interspecific competition between sibling species larvae of Anopheles arabiensis and An. Gambiae. Medical and Veterinary Entomology. 14(2): 165-170.

Takken W., Charlwood J. D., Billingsley P. F., Gort G. 1998. Dipersal and survival of *An. funestus* and *An. gambiae* s.l. (Diptera: Culicidae) during the rainy season in southeast Tanzania. Bulletin of Entomological Research. 88: 561-566.

Toure Y. T., Petrarca V., Traore S.F., Coulibaly A., Maiga H. M., Sankare O., Sow M., Di Deco M. A., Coluzzi M. 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the Anopheles gambiae complex in Mali, West Africa. Parassitologia 40(4): 477-511.

Trexler, J.D., Apperson C. S., Zurek L., Gemeno C., Schal C., Kaufman M., Walker E., Watson D. W. and Wallace L.. 2003. Role of bacteria in mediating oviposition responses of *Aedes albopictus* (Diptera: Culicidae). Journal of Medical Entomology. 40(6): 841–848.

Walker E. D., Olds E. J., Merritt R. W. 1988. Gut content analysis of mosquito larvae (Diptera: Culicidae) using DAPI stain and Epifluorescence microscopy. Journal of Medical Entomology. 25(6): 551-554.

Walker, E. D., Merritt R.W. 1991. Behavior of *Aedes triseriatus* (Diptera: Culicidae) larvae. Journal of Medical Entomology. 28: 581-589.

Walker E. D., Merritt R. W. 1993. Bacterial enrichment in the surface microlayer of an Anopheles quadrimaculatus (Diptera: Culicidae) larval habitat. Journal of Medical Entomology. 30:1050-1052.

White G. B. 1974. *Anopheles gambiae* complex and disease transmission in Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene. 68(4):278-301.

Zoetendal E. G, Collier C. T., Koike S., Mackie R. I., Gaskins H. R. 2004. Molecular ecological analysis of the gastrointestinal microbiota: A review. Ameican Society for Nutritional Sciences 134: 465-472.

CHAPTER II

Contribution of the surface microlayer and its algal and bacterial biomass to growth of *Anopheles gambiae* larvae (Diptera: Culicidae)

ABSTRACT

The importance of the aquatic surface microlayer and its components on the growth and development of larval *An. gambiae* was studied in two experiments in field and greenhouse microcosms. Removal of surface microlayer at regular intervals in the first experiment decreased survival of larvae, prolonged developmental time to pupation, and lowered body mass. Supplementations of the surface microlayer from sources with no larval grazing improved larval growth. Larval developmental time was shortened and total mass was significantly higher, compared to microcosms from which surface microlayer was deliberately removed. In the second experiment, the importance of bacteria as a food source was studied using a method in which heterotrophic bacterial production was stimulated with addition of glucose as a carbon source, in microcosms that were intentionally shaded to reduce autotrophic algal production. Glucose addition to sterilized habitats did not stimulate growth. In nonsterilized habitats to which glucose was added, larvae grew, molted, and achieved metamorphosis under these conditions, but larval survival rate and total adult emergence was very low compared to sunlit treatments.

INTRODUCTION

Anopheles larvae, unlike mosquito larvae in most other genera in the Culicidae, position their bodies parallel to the air-water interface, allowing them to breathe and feed. These larvae do not have a respiratory siphon but rather have one pair of spiracles located at the terminal abdominal segments. Larvae feed at the air-water interface and feed as specialized collector-gatherer, whereby they rotate the head 180° from the normal position and direct the rapidly beating labral brushes to the air-water interface (Merritt et al. 1992a). This mode of feeding in mosquitoes differs substantially from that of larvae in genera where a respiratory siphon is present (Merritt et al. 1992 a, b, Walker and Merritt 1991) in that *Anopheles* feeding is largely restricted to a single zone of the water column. This feeding mode has been studied in detail mainly in *Anopheles quadrimaculatus*, a species in a different subgenus from *Anopheles gambiae*. Nevertheless, feeding modes are likely so similar that comparisons are useful. In *An. quadrimaculatus* larvae, magnified visual observations of feeding larvae (Renn 1943) and high speed microcinematography of larval feeding (Merritt et al. 1992a, b, Merritt et al. 1996) showed that the lateral palatal brushes beat 7 times per second at room temperature, generating currents that were both linear and curvilinear around the head of the larvae and extending to about one body length away. Particles were entrained in a zone of the surface microlayer equal to the depth of the larval head, ca. 1 mm, and rapidly approached the anterior of the larval head in distinct starts and stops with no inertia, indicating that low Reynolds numbers governed entrained particle movements. As particles neared the anterior of the larval head, they accelerated and entered the preoral

cavity. A plume of uningested materials formed and entered the water column vertically at a 90 degree angle from the original direction of particle movement horizontally. It appeared to originate at the region of the maxillary pads in the larval preoral cavity, a set of structures which has been postulated to form a particle capture mechanism in all mosquito larvae in general (Dahl et al. 1988). Although particle movement could not be traced thereafter in those studies, formation of a food bolus at a rate of ca. one bolus every 4.4 seconds in the pharynx was quantified; indicating that rate of particle entrainment by the lateral palatal brushes was decoupled from rate of food bolus formation and swallowing. Renn (1943) and Schremmer (1949) postulated that the mandibles and maxillae interact to pack the pharynx with food particles, while Dahl et al. (1988) suggest that the pharynx itself functions as an organ which by contraction sucks entrained particles into the true mouth (i.e., the opening of the pharynx itself, which is the most anterior section of the foregut). However, this model of particle capture does not account for the simultaneous role of the pharynx in food bolus formation, nor the mechanism by which excess water is expelled as the bolus is formed. How the pharynx can simultaneously suck in water with food particles, pack the food particles into a bolus, swallow that bolus into the posterior section of the foregut, and expel excess water is a currently unresolved set of contradictory functions, all attributed to it (Dahl et al. 1988, Merritt et al. 1992b, Merritt et al. 1996). Head movements by larvae indicate that not all particles are acceptable; some are examined by the mandibles and are crushed and then swallowed; or are rejected through a 180 degree or less head turn to the normal position after mandibular manipulation (Merritt et al. 1996). This brief review of larval *Anopheles* feeding shows that much remains to be determined about feeding mechanisms

and food gathering processes. For example, later stages of *An. gambiae* larvae are known to capture and ingest neonate larvae and eggs, such that cannibalism has been proposed as a supplementary feeding strategy (Koenraadt and Takken 2003, Juan Huang et al. unpublished). Yet all studies support the notion that these larvae are highly adapted to a particle capture feeding mode in a specialized surface feeding posture.

The surface microlayer is a zone of *Anopheles* habitats enriched with microorganisms and other materials compared to subsurface zones of the water column (Walker and Merritt 1993). The specialized surface feeding behavior of these larvae upon the surface microlayer might very well be an adaptation to exploit this food-rich region. The surface microlayer varies in thickness (or depth) depending upon the way it is sampled. Generally, it consists of an approximately 1 mm deep portion of the uppermost column of any natural, lentic body of water, small or large. It is perhaps best conceived of as an organic film with accumulation of hydrophobic lipids and lipoproteins at the surface. This region of the air-water interface is enriched in both particulate organic matter and microorganisms, the latter comprising the microbial 'neuston' of its living components (Clements 1992). The surface microlayer is highly enriched in chemicals, particles and biota (algae and bacteria). A high surface tension created by hydrostatic forces and the hydrophobic nature of the composition prevents particles and biota from sinking. An epineustonic community of invertebrates exists above the surface microlayer and may in some ways be considered to be part of it, and floating and moving on it. Other materials may deposit from the atmosphere and form part of the nonliving epineustonic matrix and may interact with the surface microlayer and its components.

Experimental ecology requires manipulation of particular factors under conditions that have sufficient realism to allow generalized conclusions about the phenomenon under study. One approach has been to develop and utilize experimental microcosms as suitable models of natural habitats. Recently, development of experimental, field microcosms designed to simulate natural habitats was accomplished for larval *Anopheles gambiae* habitats, in which density-dependent larval growth (suggesting food limitation) and utilization of microbial food was demonstrated for larval growth cohorts (Gimnig et al. 2002). It is a novel use of microcosms and is developed further in my studies here. In the field microcosms, a surface microlayer developed well in the absence of actively feeding larvae, and it could be sampled with a syringe and fine gauge needle (Gimnig et al. 2002). An array of such habitats was established at a secure study site in western Kenya, and similar habitats were established in greenhouses at the Michigan State University campus.

According to the study conducted by Gimnig et al. (2002), the most likely important food source for larval *An. gambiae* was likely algae. Other experiments also reveal that presence of algae improved the growth and development of the larvae with shorter developmental time and greater survival as compared to when algae were absent, a condition established by shading the habitats without reducing water temperature (Kaufman et. al 2006). But this finding ignores algal turnover rates, larval consumption, their digestibility and other nutritional sources (i.e.; bacteria and non cellular organics in the surface layer).

Bacteria are an integral component of planktonic food webs and mediate key pathways in global biogeochemical cycles (Jurgens and Matz 2002, Cole et al. 1988).

Bacterial biomass increases in proportion to bacterial production but there is another phenomenon which not explained by the current models that bacterioplankton show relatively small seasonal fluctuations within a given system (Jurgens and Matz 2002). In contrast phytoplankton biomass vary between blooms and collapses even within one season. One study revealed that phytoplankton did not bloom in all microcosms enriched with N and P and when C was added, the phytoplanktonic groups could not compete with mineral nutrient-limited bacteria. In mesocosm set where no silicate was added, enrichment with phytoplankton nutrients and glucose led to replacement of diatoms not by other algae, but by heterotrophic bacteria (Havskum et al. 2003).

Studies reveal that around 75% of ingested algae pass through the gut of mosquito larvae (Clements 1992, Wotton et al. 1997) and larval gut analyses reveal that bacteria form the bulk of the food bolus, along with organic particulates, algae, and small invertebrate parts, when direct counts but not biovolume were used as the data collection method (Walker et al. 1988). In marine environments, gelatinous zooplankton (larvaceans, salps, doliolids) are potentially important bacterivores although their general impact on bacterial communities is poorly studied (Jurgens and Matz. 2002). Algae might be playing a key role as food resources in *An. gambiae* habitats but bacteria may indeed form a secondary food source. This aspect of the feeding biology of *Anopheles* larvae is poorly understood. Thus, here I propose to determine the contribution of bacteria to larval nutrition. A controlled greenhouse experiment was conducted to study the contribution of bacteria to larval nutrition. A set of artificial microhabitats was shaded in order to reduce algal production and to determine if non-photosynthetic microorganism biomass can compensate for the absence of algae. I used the field habitat microcosms described by

Gimnig et al. (2002) and Kaufman et al. (2006) to conduct studies on the effect of surface microlayer removal, transfer, and supplementation on larval mosquito growth. Because I specifically postulate that formation of the surface microlayer simultaneously provides larvae with food, but larval grazing reduces their food, I conducted an experiment in which this layer was manipulated to reduce or increase its availability to larvae volumetrically. Additionally, I conducted an experiment in which shade was used to reduce experimentally algal biomass, and glucose was added to stimulate bacterial production, to determine if bacterial biomass could compensate for lack of algal biomass and could promote larval growth.

METHODS AND MATERIALS

Experiment 1. Contribution of surface microlayer to larval *Anopheles gambiae* growth.

Habitat setup Microcosms were designed to simulate natural habitats of larval *Anopheles gambiae* (Gimnig et al. 2002). An array of such habitats was established in the green house at Michigan State University. Plastic basins (30cm in diameter) were provided with organically enriched soil (green house potting soil) so as to fill 1/2 of the depth of the basin. 1.5 liters of distilled water was added to fill up the basins. Water levels were maintained throughout the experiment with distilled water. The microcosms were covered with insect nets so as to prevent invasion by other insects and oviposition by local mosquitoes. After the habitats (plastic basins) with soil were filled with water, they were allowed to develop microbial growth in full sun for 3 days, and then were stocked with 40 first instar *An. gambiae* larvae (KISUMU strain) obtained from the colony maintained in our laboratory at Michigan State University (Huang et al. 2005). The experimental design was comprised of four treatments spread in a randomized complete block design with a total of 24 experimental units. Each treatment was replicated 6 times. The treatments were as follows: (1) The surface layer was purposefully removed every other day (a treatment abbreviated T1), (2) The composite surface layer was added from habitats T1 containing larvae (a treatment abbreviated T2), (3) The composite surface layer was added from source habitats without larvae (a treatment abbreviated T3), and (4) An experimental control in which there were no additions or removals of surface microlayer material, but the microcosms contained

larvae (a treatment abbreviated T4). A set of 6 habitats (abbreviated T5) containing no larvae was used as the source of surface layer material added to microcosms assigned to T3, see above. Temperature of the water was air was recorded at 30 minute intervals throughout the experiment. Manipulations were done on day 2 and continued every other day. Twelve ml of the surface microlayer material was harvested from each replicate of T1 (surface harvested) and from the "no larvae" source habitats (T5) using a syringe and 16 gauge needle and was pooled by source. Ten ml of the composite was added to the T2 and T3 microcosms, respectively. On days 5 and 9 of the experiment , 3 ml of surface microlayer was preserved in formalin (10% final conc.) for direct counts of bacteria (see below), 3 ml in Lugol's preservative (few drops) for microscopic counts of algae, and 5 ml filtered for quantification of chlorophyll *a*. The above samples were obtained as a composite by sampling 2 ml from each of the 6 replicate habitats.

Each habitat was monitored daily by visual inspection to collect pupae. The date of pupation was recorded, and individual pupae were held in cups and adults were allowed to emerge. The adults were separated by sex and stored at -80° C. The frozen mosquitoes were later lyophilized and their dry mass (mg) determined by a microbalance. The growth responses were summarized by microcosm as the proportion surviving, mean individual mass of males and females, and development time of males and females.

Microbial and Chemical analysis Direct bacterial counts of formalin preserved samples were performed using 4'6-diamidino-2-phenylindole (DAPI) stain and epifluoresence microscopy as described in Walker and Merritt (1993). Algae were counted using settling chambers and inverted microscopy in the laboratory of Dr. Orlando Sarnelle at Michigan

State University. Counts were converted with standard formulae to numbers of cells per ml of original sample. Chlorophyll *a* content of filtered particulates was determined fluorimetrically by extraction overnight in 95% ethanol (Welschmeyer 1994), yielding chlorophyll *a* in micrograms per liter of original sample.

Statistical Analysis Analysis of variance (ANOVA) was performed using SAS for the growth data. Data were transformed as needed to meet the assumptions of ANOVA.

Experiment 2. Contribution of bacteria to larval *Anopheles gambiae* nutrition

Habitat setup: Microcosms were designed similar to Experiment 1, to simulate natural habitats for *Anopheles gambiae* larvae. Twenty-seven such microcosms were set up and 24 of them were shaded. The shading was accomplished by using aluminum foil-covered, perforated boxes which were placed over each microcosm in such a way as to facilitate air circulation but greatly reduce light. Air over the treatments was well circulated with fans to maintain even temperatures at ambient levels. All the treatments were covered with insect nets to prevent predation and oviposition by other insects. The experimental design was a randomized complete block design with 24 total experimental units. To stimulate heterotrophic microbial growth, 12 of the 24 shaded habitats were supplemented with glucose to obtain 30 ppm of final concentration. Glucose was added beginning two days prior to introduction of the larvae and continuing every 3 days until most of the larvae pupated. The concentration of glucose added to the treatments was determined by conducting a pilot assay with different concentrations of glucose to obtain maximum bacterial productivity. Twelve habitats did not receive any glucose, as controls.

Out of 12 treatments with glucose and no glucose, 6 treatments did not receive any larvae. Thus, 6 habitats with glucose and 6 without glucose did not receive any larvae to determine bacterial counts in absence of the larvae. Microcosms with specific treatments were stocked with 10 larvae four days after the habitats were setup. Each treatment had 6 replications. Three habitats were left in bright sunlight and stocked with larvae to allow normal algal growth to occur and to confirm normal development of the larvae. Pupae collections and data acquisition and summarization were done as in the first experiment.

To ascertain that glucose did not have any direct effect on the growth responses in the above experiment a different experiment was setup. Fifteen experimental units were setup similar to Experiment 1 with the exception that the plastic basins used were 15 cm in diameter containing 250 gm of soil, 600 ml of sterile distilled water and five larvae added to each unit. Five replicates of each of the following three treatments were setup. Treatments included sterile soil with glucose, non sterile soil with glucose and control. Fish food (used in maintaining the laboratory colony) was added to the control units. Glucose was added every other day to all the treatments receiving glucose to final concentration of 30 ppm. Sterile treatments received filter sterilized glucose.

Bacterial counts: 3 ml sample of surface microlayer were collected from the habitats one and three days before the larvae were added for bacterial counts. After the introduction of larvae, the bacterial counts were carried out on days 12 and 18). The first pupa from the glucose treatment was harvested on day 15. The samples were preserved in 4% formalin and stored at 4° C until counts were performed. Counts were done using 4'6-diamidino-2-phenylindole (DAPI) stain and epifluoresence microscopy.

Bacterial productivity: 2 ml samples of the surface microlayer were collected on day 6 (three days after larvae added) from each shaded glucose and no glucose habitat to determine the bacterial productivity within those habitats. The productivity assay was performed by incubating sub-samples with ^{3}H-leucine and measuring the ^{3}H–leucine incorporation into protein using liquid scintillation (Kirchman 2001, Kaufman et al. 2002).

Statistical Analysis: A non-parametric statistical test, the Wilcoxon/Kruskal-Wallis rank sum test, was performed on the growth data.

RESULTS

Experiment 1. Contribution of surface microlayer to larval *Anopheles gambiae* growth.

Pupal productivity was greatest from microcosms receiving surface microlayer (T3) material from source habitats containing no larvae (70.8% mean survival to pupation; 170 out of 240 pupae harvested); this productivity was higher than that in habitats in which surface microlayer was removed (T1) (42.9% mean survival to pupation; 103/240 pupae harvested). Pupal production was lowest in habitats to which surface microlayer material was added from habitats containing larvae (T2) (34.2% mean survival to pupation; 82/240 total pupae harvested) (Fig 2.1). In control habitats (T4) with no additions or removals of surface microlayer 47.5% larvae survived to pupation and 114 from 240 pupae were harvested). The average number of pupae developing from each individual unit was 17.1 (± 7.4) from T1, 13.6 (± 5.6) from T2, 28.3 (± 2.7) from T3, and 19 (± 9.6) from T4.

The total adults that emerged from the pupae collected were 86 (83.5 %) for T1, 47 (57.32 %) for T2, 143 (84.18 %) for T3, and 85 (74.56%) for T4. The average number of adults that emerged from each unit were 14.3 (± 8), 7.8 (± 4.9), 23.8 (± 4.1), and 14.2 (± 9.2). The above results indicate that treatment T3, in which composite surface layer was added from source habitats, had the highest survival rate and yielded the most pupae compared to other treatments, ($F = 5.12$, $P = 0.0086$) as well as adults ($F = 5.43$, $P = 0.0068$). Supplementing T3 microcosms with food from source habitats improved the production efficiency of the habitat but the addition of food source from microcosm with

larvae did not improve the production instead it had lowered number of pupae and adults coming out of this microcosm (T2). There were significant difference in the female development time ($F= 5.29$, $p= 0.0075$) between the treatments (Fig 2.2). Females from T3 treatment had the shortest development time of 18.9 ±1.9 days compared to 23.3 ± 2.7 in T1, 24.7 ± 2.23 in T2 and 23.3 ± 3.6 in T4. The male development time ($F=2.26$, $p= 0.1163$) was not affected by the treatments. The total mass of adults emerging fromT3 was higher significantly higher ($F= 4.85$, $p=0.01$) than T1 and T4 (Fig 2.3). Adults from T2 had significantly lower adult dry weights than the rest of the treatments. Similar results were observed for total male mass ($F= 4.44$, $p= 0.0151$) and total female mass ($F= 3.71$, $p= 0.0286$) between treatments.

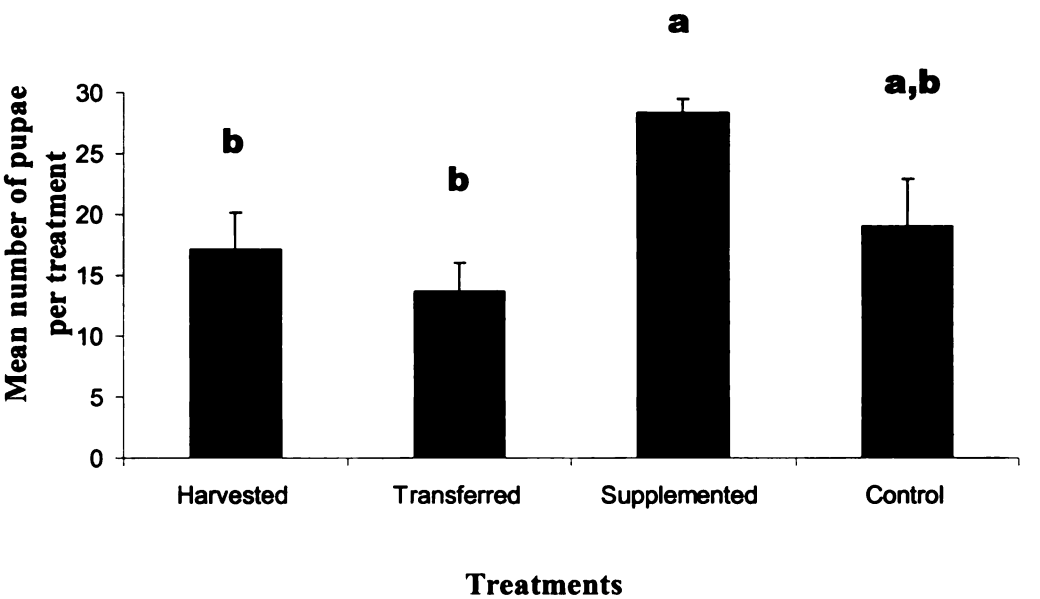


Figure 2.1: Mean number of pupae produced in the four treatments in the Experiment 1. T1, Surface harvested; T2, Surface microlayer added from habitats with LARVAE; T3, Surface microlayer added from source habitats; T4 Control (no harvesting or supplementing). N = 6 replications per treatment

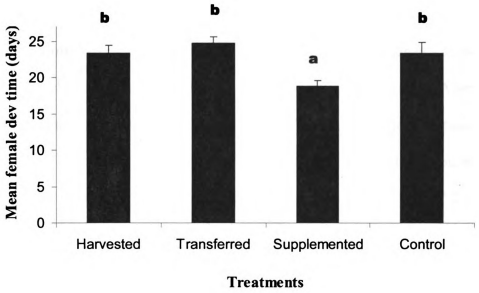


Figure 2.2: Average female development time in different treatments. T1, Surface harvested; T2, Surface microlayer added from habitats with LARVAE; T3, Surface microlayer added from source habitats; T4 Control (no harvesting or supplementing). N = 6 replications per treatment

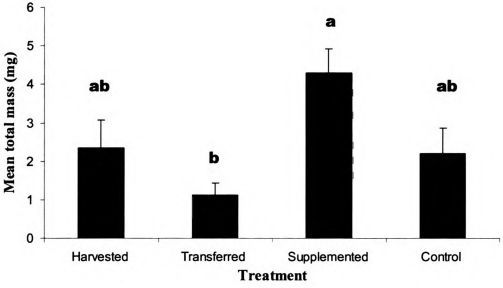


Figure 2.3: Mean total dry mass produced in different treatments. T1, Surface harvested; T2, Surface microlayer added from habitats with LARVAE; T3, Surface microlayer added from source habitats; T4 Control (no harvesting or supplementing). N=6

Chemical and Microbial analysis Chlorophyll *a*, a measure of algae was estimated on day 9. The chlorophyll *a* measure in T1 was 23.74 μg/L, T2 9.29 μg/L (After the manipulation was performed, 9.29 μg/L), T3 23.74 μg/L (after manipulation was performed 53.68 μg/L), T4 20.65 μg/L, T5 609.05 μg/L. The bacterial counts on day 5 did not vary much between the treatments and ranged from 1.8×10^6 to 2.2×10^6 cells/ml in all the four treatments (Fig 2.4). The bacterial count of the source habitats on day 5 was 2.9×10^7 cells/ml. On day 9 the bacterial counts remained low in treatments T1 and T2 being 6.2×10^6 and 9.3×10^6 cells/ml respectively. Bacterial numbers in T3, T4 and T5 were higher (by one order of magnitude from T1 and T2), but fairly similar between them and they were 2×10^7, 1.7×10^7, and 1.9×10^7 cells/ml respectively. The above counts represent the bacterial numbers before the manipulations were done on that day. Algal counts on day 5 were 1.1×10^4 cells/ml for T1; 1.3×10^3 cells/ml for T2; 4.4×10^2 cells/ml for T3; 5.4×10^3 cells/ml for T4 and 2.4×10^5 cells/ml for T5. On day 9 algal counts in T1 was 3.6×10^4 cells/ml, in T2 was 2.2×10^4 cells/ml (after manipulation, 2.8×10^4 cells/ml), in T3 was 6.3×10^4 cells/ml (after manipulation 8.4×10^4 cells/ml), T4 was 5.6×10^4 cells/ml and T5 was 1.6×10^5 cells/ml.

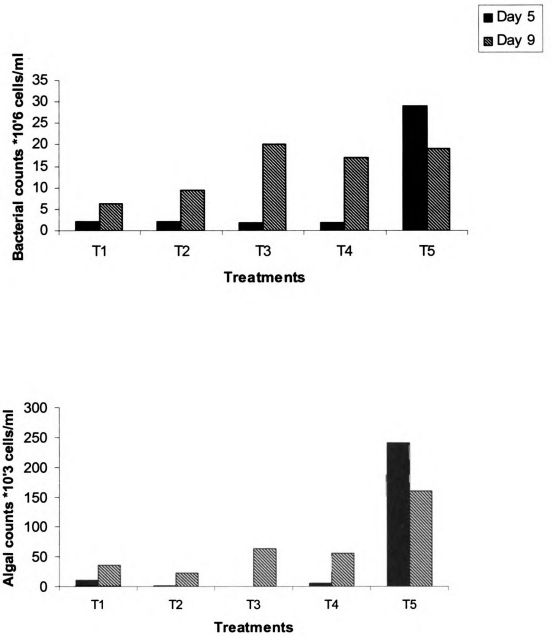


Figure 2.4: Bacterial and Algal counts of the surface microlayer from different treatments. T1, Surface harvested; T2, Surface microlayer added from habitats with LARVAE; T3, Surface microlayer added from source habitats; T4 Control (no harvesting or supplementing)

Experiment 2. Contribution of bacteria to larval *Anopheles gambiae* nutrition.

No adults were produced in the shaded treatments without glucose amendment during the entire duration of the experiment (25 days). The larvae did not develop beyond late second or early third instar. The glucose amended treatments produced pupae that also emerged as adults. A total of 21 out of 60 larvae (35%) pupae developed from glucose amended treatment with an average of 3.5 ± 1.3 per unit (Fig 2.5 A). On the other hand a total of 26/30 (86%) pupae developed from sunlit treatments which were added to the experiment as a control. The average number of pupae per unit was 8.6 ±0.33 in sunlit treatments. Of the pupae produced from glucose treatment 57.14% emerged as adults the remaining died while emerging or prior to emergence. The emergence rate for the pupae collected from the sunlit treatment was very high up to 96.15%. The average development time for *Anopheles* larvae was 7.76 (± 0.03) days in sunlit treatment and 15.38 (± 8) days in glucose treatment. Non-parametric statistical test, Wilcoxon/Kruskal-Wallis test (rank sum) was performed for number of pupae and the probability (Z stat) was 0.0284. The adults produced in glucose treatment had lower body weight with an average of 0.16 mg (± 0.026) than the treatments in open sun which had an average adult dry mass of 0.59 mg (± 0.11). Similar outcome was observed in sterile and non sterile treatments containing glucose (Fig 2.5 B); pupae were produced in non sterile treatments containing glucose (36.7%). No pupae were produced in the sterile treatments containing glucose, the larvae did not grow beyond first or second instar. Larvae developed normally with 86.7 % pupae produced in control treatments containing fish food.

Bacterial counts were performed at different time points before and after first instar larvae were added to the experimental units. The bacterial counts performed at the beginning of the experiment prior to glucose addition in all the treatment were in the range of 5.7 x 10^6 to 43.5 x 10^6 cells/ml. In sun treatments the bacterial numbers ranged from 3 x 10^6 to 7.7 x 10^6 cells/ml. Also counts were performed for unidentified algae in sunlit treatments, which ranged from 4 x 10^5 to 4 x 10^6 cells/ml. No such algae were observed in shaded treatments. One day after the first addition of glucose in the respective treatments, the counts in treatments without glucose ranged from 9.8 x 10^6 to 28.2 x 10^6 cells/ml whereas in treatments with glucose the numbers ranged from 2.6 x 10^6 to 22 x 10^6 cells/ml. Three days prior to pupation the bacterial counts in treatments with glucose/larvae were 4.02 x 10^6 cells/ml and in glucose/no larvae 1.9 x 10^7 cells/ml. In treatments with no glucose/larvae the counts were 5.06 x 10^6 cells/ml and in treatments with no glucose/no larvae the counts were 5.1 x 10^7 cells/ml.

Bacterial productivity determined after two glucose additions, by measuring the ^{3}H–leucine incorporation into bacterial proteins was estimated and was 105.27 μg C/ L /day (SEM = 10.0) in the glucose-amended treatments and 49.16 μg C/L/day (SEM= 2.33) in non amended ones (Fig 2.6).

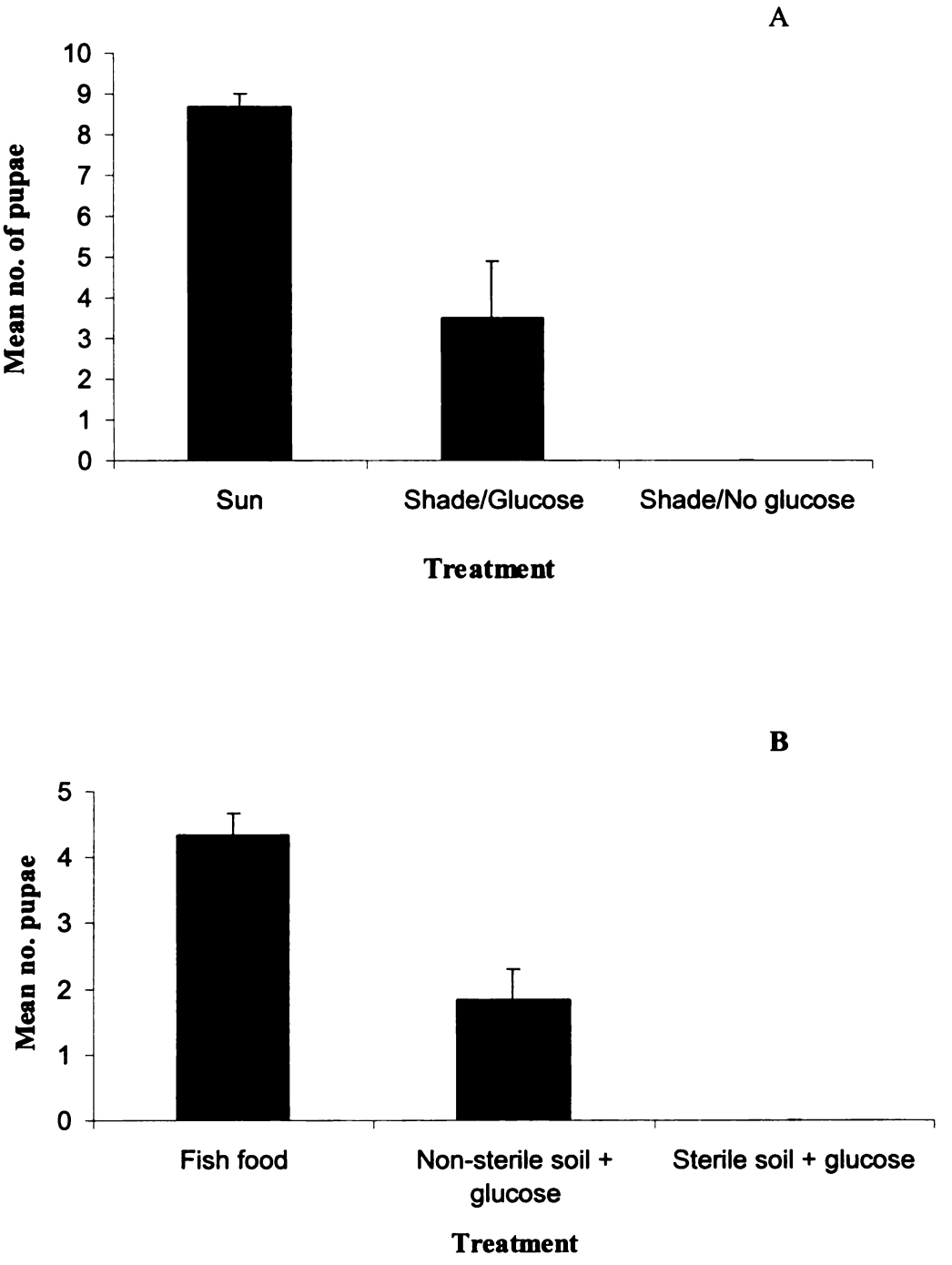


Figure2.5: Average number of pupae produced from the respective treatments in two different set of experiments (refer to text). A. Habitats shaded or not, and glucose supplemented in some shaded habitats but not others. B. Habitats sterile or nonsterile, and glucose added to all. Fish food was added to some nonsterile ones as a control. A. N = 6 and B. N=5 replications per treatment.

Figure 2.6: Bacterial productivity measured as mean µg of carbon/L/day in glucose amended and non amended microcosms. N = 6 replications per treatment.

DISCUSSION

Larval growth processes have strong effects on adult success and, ultimately, determine the adult population abundance. A wide range of factors influence larval success and adult emergence, a primary one being larval mortality which result from food limitation, competition, predation, disease, or physical destruction of the habitat (Olson and Olson 1989). Competition for nutrients when food supply is limited or when larvae are crowded can affect *An. gambiae* by slowing larval development and reducing pupation success (Gimnig et al. 2002). Observations of cannibalism amongst larvae also indicate that competition for food resources may be extreme (Koenraadt and Takken 2003). Also, factors like temperature have a critical effect in larval developmental and adult success when other factors such as food are not limiting (Bayoh and Linsay 2003, 2004). Results of this study (Experiment 1) indicate that removal of surface microlayer results in lengthened developmental time of *An. gambiae* larvae, and correspondingly smaller adults at emergence. Supplementing the larval microcosms with a food source not previously utilized, by contrast, shortened larval development time, increased larval survival, and produced larger adults. Deprivation of food by harvesting the surface microlayer or addition of surface microlayer from microcosm grazed by larvae does not show any positive effect on the larval growth, in fact the opposite was observed. The treatments supplemented with surface microlayer that had larval grazing pressure had the least number of pupae and adults and the total mass was also minimal. The chlorophyll *a* levels as well as algal counts were higher in treatments that received the non-utilized food than that in the other three treatments. This result suggests that algae are indeed a key

nutrient required for the success of these larvae. Bacteria were high in density and similar in density for all treatments except for the harvested treatment and in the treatments receiving previously utilized food. That bacterial counts were lower in these treatments suggests that bacteria were fed upon, and could therefore compensate for a deficiency in algae. Measurements of biovolume instead of cell count would have been a better approach in interpreting these data, but it was not done.

The classic method of assessing the importance of food limitation is to manipulate food supply through enrichment or reduction (Olson and Olson 1989). The glucose amendment study (Experiment 2) aimed to enrich the heterotrophic bacteria in the habitats to compensate for lack of sufficient algae. Bacteria apparently provided some larvae the nutrition required for growth, but few adults were produced even though the density of larvae per microcosm was intentionally kept low so as to reduce competition. Addition of glucose and subsequent increase in bacterial growth rates enhanced pupal production above that seen in non-amended microcosms. Direct effects of glucose on larval growth were ruled out since no larvae grew in treatments with sterile soil and amendments of glucose. No studies on *Anopheles* larvae or other invertebrate larvae have indicated that the larvae can subsist on dissolved organic matter (DOM) alone (Olson and Olson 1989, Clements 1992). This is consistent with the results observed here. Manahan et al. (1989) found that growth of molluscan larvae on the defined nutrient mixture alone was significantly greater than that of larvae which were starved, but not as high as that fed with algae. By 8 days the larvae fed on nutrient mixture showed tissue loss and began dying. Merritt et al. (1996) provided experimental evidence that hatching and survival of *An. quadrimaculatus* was higher with addition of surface microlayer

alone and with combination of DOM but not with DOM alone (Merritt et al.1996, Wotton et al. 1997).

Studies have revealed that *Anopheles gambiae* larvae are commonly associated with algae in the surface microlayer (Gimnig et al. 2002), and that algal biomass measured as chlorophyll *a* and algal counts were significantly reduced in surface water samples in presence of larvae. Other studies reveal that presence of algae improved the growth and development of the larvae with shorter developmental time and greater survival as compared to when algae were absent, a condition established by shading the habitats (Kaufman *et al.* 2006). The microbial components and nutrients available as food for *Anopheles* larvae found within the surface microlayer and in gut contents have been documented in few studies (Walker *et al.* 1988, Walker and Merritt 1993). Larval gut analysis revealed that bacteria form the bulk of the food bolus, along with organic particulates, algae, and small invertebrate parts (Walker *et al.* 1988, Wallace and Merritt 2004). There is no doubt that algae are an important component of *An. gambiae* diet but bacteria can also serve as a food source, although they might not serve as a sole source in natural settings. Addition of glucose in Experiment 2 with subsequent increase in heterotrophic bacterial growth rates (bacterial productivity was enhanced 2-3X) enhanced larval development compared to that seen in treatments without glucose. However the pupation rates were lower than in sunlit treatments. Adult emergence was much lower with either dead pupae or death of pupae while emerging. Bacteria predominate numerically in the surface microlayer (Walker and Merritt 1993), but their biomass may not be able to compensate for biomass produced by primary producers. Even though there is evidence that bacteria can provide nutrition for survival and pupation it is inadequate

food resource by itself in natural habitats. Also, algae can provide certain essential nutrients that the bacteria cannot provide, in particular long chain, polyunsaturated fatty acids required by all mosquitoes in their larval diet (Clements 1992). Lack of the above in the diet causes an inability in the adults to fly. Wotton et al. (1997) showed that surface microlayer bacteria aided growth and survival of *An. gambiae* larvae but they did not study the pupation of the larvae. Overall, my study demonstrates that bacteria in the surface microlayer bacteria promote growth of *An. gambiae* larvae, but are likely to be insufficient for typical cohorts in nature.

REFERENCES

Bayoh M. N., Linsay S. W., 2003. Effects of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera: Culicidae). Bulletin of Entomology Research. 93. 375-381.

Bayoh M. N., Linsay S. W., 2004. Temperature-related duration of aquatic stages of the Afrotropical malaria vector mosquito *Anpheles gambiae* in the laboratory. Medical and Veterinary Entomology. 18. 174-179.

Clements A. N. 1992. Larval feeding. Pg74-99 *In* eds. Chapman & Hall. The biology of mosquitoes. University press, Cambridge. v2.

Cole, J. J., Findlay S., Pace M. L. 1988. Bacterial production in fresh and saltwater ecosystems: a cross overview. Marine Ecology Progress Series. 43: -10.

Dahl C., Widahl L., Nilsson C. 1988. Functional Analysis of the Suspension Feeding System in Mosquitoes (Diptera: Culicidae). Annals of the Entomological Society of America. 81(1): 105-127.

Gimnig, J. E., Ombok M., Kamau L., Hawley W.A. 2001. Characteristics of larval anopheline (Diptera: Culicidae) habitats in western Kenya. Journal of Medical Entomology. 38: 282-288.

Gimnig, J. E., Ombok M., Otieno S., Kaufman M. G., Vulule J. M., and Walker E. D. 2002. Density-dependent development of *Anopheles gambiae* in simulated natural habitats. Journal of Medical Entomology. 39:162-172.

Havskum H., Thingstad T. F., Scharek R., Peters F., Berdalet E., Sala M. M., Alcaraz M., Bangsholtz J. C., Zweifel U. L., Hagstrom A., Perez M., and Dolan J. R., 2003. Silicate and labile DOC interfere in structuring the microbial food web via algal–bacterial competition for mineral nutrients: Results of a mesocosm experiment. Limnology and Oceanography. 48(1): 129-140.

Huang J., Walker E . D., Giroux P. Y., Vulule J. 2005. Ovipositional site selection by *Anopheles gambiae* : influences of substrate moisture and texture. Medical and Veterinary Entomology.19: 442 – 450.

Huang J., Miller R. R., Chen Shi-Cheng., Vulule J. M., and Walker E. D. 2006. *Anopheles gambiae* (Diptera: Culicidae) Oviposition in response to agarose media ad cultured bacterial volatiles. Journal of Medical Entomology. 43(3): 498-504.

Jurgens K., Matz C. 2002. Predation as a shaping force for the phenotypic and genotypic compostion of planktonic bacteria. Antonie van Leeuwenhoek 81: 413-434.

Kaufman, M.G., Goodfriend W., Kohler-Garrigan A., Walker E.D., & Klug M.J. 2002. Soluble nutrient effects on microbial communities and mosquito production in *Ochlerotatus triseriatus* (Say) habitats. Aquatic Microbial Ecology 29: 73-88.

Kaufman M. G., Wanja E., Maknojia S. R., Bayoh N. M., Vulule J. M., Walker E. D. 2006. The importance of algal biomass to the growth and development of *Anopheles gambiae* larvae. Journal of Medical Entomology. 43: 669-676.

Kirchman, D. 2001. Measuring bacterial biomass production and growth rates from leucine incorporation in natural aquatic environments. Methods in Microbiology. 30: 227-237.

Koenraadt C. J. M., Takken W. 2003. Cannibalism and predation among larvae of the *Anopheles gambiae* complex. Medical and veterinary Entomology. 17: 61-66

Manahan DT, Jaeckle WB, Nourizadeh S. 1989. Ontogenetic changes in the rates of amino acid transport from seawater by marine invertebrate larvae (Echinodermata, Echiura, Mollusca). Biological Bulletin. 176: 161–168.

Merritt, R.W., Craig D.A., Walker E.D., Vanderploeg H.A., Wotton R.S. 1992a. Interfacial feeding behavior and particle flow patterns of *Anopheles quadrimaculatus* larvae (Diptera: Culicidae). Journal of Insect Behavior 5: 741-761.

Merritt, R. W., Dadd R. H., Walker E. D. 1992b. Feeding behavior, natural food, and nutritional relationships of larval mosquitoes. Annual Reviews in Entomology. 37: 349-76

Merritt R. W., Craig D. A., Wotton R. S., Walker E. D. 1996. Feeding behaviour of aquatic insects: Case studies on blackfly and mosquito larvae. Invertebrate Biology. 115(3). 206-217.

Olson R. R. and Olson M. H. 1989. Food limitation of planktotrophic marine invertebrate larvae: Does it control recruitment success? Annual Reviews in Ecology and Systematics. 20:225-247

Renn C. E. 1943. Emergent vegetation, mechanical properties of water surface, and distribution of *Anopheles* larvae. National Malaria society journal.2: 47-52

Schremmer F. 1949. Morphologische und funktionelle Analyse der Mundteile und des Pharynx der Larve von *Anopheles maculipennis* Meig. Osterreichische Zoologische Zeitschrift. 2:173-222.

Walker E. D., Olds E. J., Merritt R. W. 1988. Gut content analysis of mosquito larvae (Diptera: Culicidae) using DAPI stain and Epifluorescence microscopy. Journal of Medical Entomology. 25(6): 551-554.

Walker E. D., Merritt R.W. 1991. Behavior of *Aedes triseriatus* (Diptera: Culicidae) larvae. Journal of Medical Entomology. 28: 581-589.

Walker E. D., Merritt R. W. 1993. Bacterial enrichment in the surface microlayer of an *Anopheles quadrimaculatus* (Diptera: Culicidae) larval habitat. Journal of Medical Entomology. 30:1050-1052.

Wallace J. R., Merritt R. W. 2004. Diel feeding periodicity of larval anopheline mosquitoes on microorganisms and microinvertebrates: a spatial and temporal comparison of *Anopheles quadrimaculatus* (Diptera: Culicidae) diets in a Michigan pond. Journal of Medical Entomology. 41(5):853-860.

Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll *a* in the presence of chlorophyll *b* pheopigments. Limnology and Oceanography. 39:1985-1992

Wotton, R. S., Chaloner D.T., Yardley C.A., Merritt R.W. 1997. Growth of *Anopheles* mosquito larvae on dietary microbiota in aquatic surface microlayers. Medical and Veterinary Entomology. 11: 65-70.

CHAPTER III

Effects of larval *Anopheles gambiae* grazing on the bacterial community structure of larval habitats

ABSTRACT

A culture-independent molecular phylogenetic survey was carried out for the bacterial community in microcosms simulating natural habitats of larvae of the mosquito, *Anopheles gambiae* s.s. Two different local (Kisumu, Kenya) soil types, Red Soil and Black Soil, were added to microcosms, some of which received larvae and some did not resulting in a 2 x 2 factorial experiment. A total of 732 sequences were obtained. These sequences were classified using the Ribosomal Database Project II bacterial classification system. The analysis tool called DOTUR was used to define operational taxonomic units and to calculate diversity indices. Overall, the bacterial communities were highly diverse, Rank abundance curves revealed few common and many rare OTUs, a pattern typical of bacterial communities in general. Presence of larvae had strong effects on bacterial community composition. Comparisons using the LIBSHUFF procedure and principal components analysis indicated that the bacterial communities in larvae/no larvae treatments of both soil types (Red Soil with larvae/Black Soil with larvae, and Red Soil without larvae/ Black Soil without larvae) were significantly different. There were a large number of unclassified sequences in no larvae treatments in both soil types, and they were closest to sequences of phototrophic primary producers, likely representing chloroplast rDNA sequences or unclassified cyanobacterial rDNA sequences. Community shifts were observed either by presence or absence of certain taxa, or

changes in the frequencies of certain taxa, as represented by the sequence data. These community shifts could explain variation in productivity of habitats in natural settings and also habitat segregation of *Anopheles* species due to ovipositional site selection.

INTRODUCTION

The thin, liquid zone at the air-water interface (termed by limnologists the 'surface microlayer') of *Anopheles* larval habitats is a zone enriched with microorganisms and other materials compared to subsurface zones of the water column (Walker and Merritt 1993, and see Chapter 2). These larvae position themselves parallel to the surface of the water to breathe, and they feed at the air-water interface. They are specialized collector-gatherer or gatherer-filter feeders, whereby they rotate the head 180° degrees from the normal position and direct the rapidly beating lateral palatal brushes to the air-water interface (Merritt et al. 1992, Merritt et al. 1996). The resultant currents generated by the extension and flexion of the lateral palatal brushes entrain particles within the surface microlayer, drawing them to the anterior of the larval head, where they are processed by a series of mouthpart movements leading to ingestion into the pharynx, followed by food bolus formation and swallowing (Merritt et al. 1992, Merritt et al. 1996). Enrichment of the surface microlayer in both freshwater and saltwater environments with microorganisms, lipids, and other materials is well known (Chapter 1). In an investigation of a larval habitat of *Anopheles quadrimaculatus* larvae, Walker and Merritt (1993) found that the total numbers of bacteria ranged from 18.8 x 10^6/ml to 65.4 x 10^6/ml in surface microlayer samples, and from 3.8 x 10^6/ml to 14.3 x 10^6/ml in subsurface samples. These results suggest that microbial food items of *Anopheles* larvae are more concentrated in this feeding zone, thus their feeding strategy is adaptive in taking advantage of a food rich zone of the water column.

Various studies involving larval gut analysis reveal that bacteria comprise a substantial portion of the food bolus, along with organic particulates, algae, and small invertebrate parts (Walker et al. 1988). Bacteria could account for production of a large fraction of the mosquito biomass produced from tree hole environments (Kaufman et al. 2001), but simultaneously, larval mosquito feeding reduces microbial abundance, alters the microbial community quantitatively, and diminishes the quality of the microbes as food (Kaufman et al. 2000, 2002; Xu et al. in preparation). Bacteria along with other organisms that occur in larval habitats may form the basis for production of adults, by transforming inorganic nutrients through autotrophic and heterotrophic processes into assimilable forms, by providing food, and specific larval nutritional demands (Walker et al. 1988, Walker and Merritt 1993, Kaufman et al. 2000, 2001, 2002).

Adult mosquito density depends on the number and productivity of the larval habitats; larval habitat productivity in turn depends upon nutrient mobilization into the microbial food base for these larvae (Peck and Walton 2005, Peck and Walton 2006, Kaufman and Walker 2006, Kaufman et al. 2006). In water-filled tree holes, nitrogen supplementation resulted in a cascade of effects, in which fungal biomass and metabolic activity increased, thereby increasing rate of leaf decay and availability of larval mosquito food (Kaufman and Walker 2006). Although the amount of dissolved carbohydrate (DOC) increased as well, bacterial production on leaf surfaces did not increase; thus the role of bacteria in supporting mosquito production was diminished compared to that of fungi. Peck and Walton (2005, 2006) showed that variation in ratios of carbon, nitrogen, and phosphorus, or variation in phosphorus content in bacteria, differentially affected growth responses of larvae of two species of *Culex*. Gimnig et al.

(2001) showed in field microcosm experiments that the most likely food source for larval *An. gambiae* was algal biomass, which was significantly reduced in the presence of larvae as measured by chlorophyll *a* in surface water samples and by counts of algae in sedimentation chambers, compared to when larvae were absent. Mosquito feeding reduced dry algal biomass by 100 μg/ml, or a total of 7 mg per habitat surface layer. Based upon total adult production of 2.8 and 7.5 mg in treatments with 20 and 60 larvae respectively, the decline in algal biomass documented effects of larval grazing pressure on algae, and also explained much of the observed larval production. Bacteria may have supplemented or supplanted algae under such feeding pressure as it may have formed a secondary food source, as bacterial densities were not affected significantly by the presence of larvae. The levels of nitrogen and phosphorus in these soil-lined habitats (mud puddles) were not affected by the presence of larvae, although there was evidence for decreasing nitrogen levels with increasing larval densities, suggesting that nitrogen may be a limiting resource in the larval environment. Kaufman et al. (2006) extended on these findings in a set of experiments in which habitats were shaded or left in full sun, such that primary production was nearly eliminated, or was left to occur normally. Under shaded conditions, Kaufman et al. (2006) found that primary production as measured by algal biomass was greatly reduced and *An. gambiae* larval growth was stalled, consistent with the hypothesis that algal biomass is crucial to larval growth. As shown in Chapter 2, bacterial biomass did support some larva growth but could not compensate for the lack of algal biomass in furthering complete development of larval cohorts in those microcosms. Bacteria rich in phosphorus promoted growth of *Culex quinquefasciatus* but, by contrast, *Cu. tarsalis* larvae grew better on bacteria low in concentration of this nutrient (Peck and

Walton 2006). Collectively, these studies suggest that the structure of the microbial communities in a range of larval mosquito habitats (tree holes, mud puddles, waste water lagoons, constructed wetlands) affect larval mosquito growth in positive and negative ways, and that nutrient dynamics, microbial community dynamics, mosquito feeding, and mosquito production interact with each other.

One of the ways in which mosquito growth could be limited is that larval mosquito feeding itself reduces microbial biomass, microbial density, and alters microbial community structure, such that the quality and quantity of microbial food becomes unsupportive for growth. This idea finds parallels in studies of other aquatic systems where invertebrate grazing on microorganisms is a key trophic process. For example, a series of studies in lake ecosystems in northern Europe revealed marked predator-prey relationships, as follows. Planktonic bacterial communities in these studies responded both phenotypically and genotypically to strong grazing pressure by metazooplankton (Jurgens and Matz 2002), interacted strongly in their population dynamics to carbon and phosphorus limitations (Matz and Jurgens 2003), and were shaped by combined protozoan and metazoan predation (Langenheder and Jurgens 2001). *Daphnia* fed upon both bacteria and their protozoan predators, having the general effect of reducing abundance of both groups and severely dampening bottom-up nutrient effects. Due to rapid compensation of bacterial grazing losses by the development of resistant forms, the total bacterial biomass may remain less affected and could be maintained on a relatively high level (Jurgens 1994, Jurgens and Gude 1994). These studies show how mesocosm experiments can support whole ecosystem studies, and can model long-term, integrated analysis of these relationships. These observations represent

in their broadest sense an extension of the analysis of predator–prey interactions, which have been considered as a driving force in population dynamics since the beginning of ecological studies (May 1974).

In this study, I examined bacterial community structure and diversity associated with presence and absence of larval grazing pressure. The second experimental context related to nutrient sources stemming from underlying soil type. The overall hypothesis was that bacterial community structure in the surface microlayer will vary with intensity of mosquito feeding, where the *Anopheles* larvae are predators, and will vary with the nutrient regime as determined by underlying soil type. As shown elsewhere (Kaufman et al. 2006), the dark, alluvial, clay, top soil prevalent in the Lake Victoria basin (locally called black cotton soil, Ranteng, or Anyuong) and herein referred to as Black Soil) was higher in phosphorus content compared to the other locally prevalent soil, which is a red, sandy soil used commonly in road and house wall construction (locally called Luala makuoyo and herein referred to as Red Soil). This observation is confirmed by soil analyses in this region, and variation in soil nutrients has been correlated with agricultural practices and agricultural productivity (Mango 1999). *Anopheles gambiae* larvae have been found in habitats formed in each soil type.

In this study, I utilized molecular methods to study bacterial diversity in these environments. There is virtually no information on the bacterial community structure of *An. gambiae* larval habitats, with the exception of the recent study by Huang et al. (2006), in which 61 bacterial isolates from a single habitat were identified to genus using a 16S rDNA sequencing method, in preparation for oviposition studies. Yet, bacteria may be important as larval food at least in supplementing algal biomass (Gimnig et al.

2002, Kaufman 2006), may be sources of odors influencing oviposition decisions by gravid females (Huang et al. 2006), and may have a role as mediators of habitat segregation between molecular forms of *An. gambiae* (Edillo et al. 2006). Of interest is that Huang et al. (2006) showed that bacterial cultures were typically repellent to ovipositing females, suggesting that those mosquitoes were sensitive to bacterial-derived odors. Accordingly, an investigation of the community structure of bacteria in *An. gambiae* habitats relative to larval feeding pressure and underlying soil structure is highly relevant to elucidate an important missing feature of the biology of this highly dangerous malaria vector.

It has become axiomatic that the majority of bacteria in any given environment are uncultivable for a variety of reasons (Hugenholtz et al. 1998), thus making alternatives to culture-based methods important for studies of bacterial diversity. According to Zwart et al. 2002, on the basis of cultivation techniques, Rheinheimer (1980) concluded that bacteria found in groundwater, spring water and streams also commonly occur in soils, and that there was no clear separation between soil bacteria and aquatic bacteria when soil was in close association with the overlying water. These observations call into question the existence of a unique freshwater bacterial flora (Zwart et al. 2002). The advent of molecular techniques and especially the polymerase chain reaction has made it possible to obtain information on microbial community composition directly, without cultivation (Giovannoni et al. 1990). An environmental sample can be inventoried for taxa present by direct nucleic acid isolation, followed by amplification of particular marker genes and analysis of the sequence of base pairs. The most widely used marker gene is the small subunit rRNA gene (16S rDNA), and the recent application of

molecular techniques in a variety of habitats has produced a large set of sequences from this gene (Maidak et al. 1999, Cole et al. 2005). This growing body of information has shown that the diversity of the bacterial domain is much larger than was thought before the application of molecular methods (Pace 1997, Hugenholtz et al. 1998). For the purpose of this study I have used the small subunit rRNA gene (16S rDNA) as a marker gene for community fingerprinting and sequencing analysis.

METHODS AND MATERIALS

Habitat array In 2002, an array of artificial habitats/field microcosms was designed to simulate natural habitats in which density dependent larval growth (suggesting food limitation) and utilization of microbial food was demonstrated for larval growth cohorts (Gimnig et al. 2002, Kaufman et al. 2006). In these field microcosms, a surface microlayer developed well in the absence of actively feeding larvae, and it could be sampled with a syringe and 16 gauge needle (Gimnig et al. 2002). An array of such habitats was established at a secure study site at the Kenya Medical Research Institute (KEMRI), Vector Biology and Control Research Center in Kisian, Kenya to study growth and development of *Anopheles gambiae* larvae (Kaufman et al. 2006). I utilized this array of naturalistic habitats for my studies on the effects of mosquito feeding and soil type on bacterial community diversity. The artificial habitats were constructed from plastic washtubs (approx. 40 cm diam.) lined with black plastic and set into the ground. Approximately 2000 grams of Black Soil or Red Soil (as described above) were added to randomly assigned habitats within the array of 44 individual treatments. These soils are associated with crops in the area, and there is local knowledge of variation in agricultural productivity and nutrient depletion, as well as scientific data on soil fertility (Mango 1999). *Anopheles gambiae* larvae have been found in habitats formed in each soil type (Gimnig et al. 2001, Mutuku et al. 2006). Four liters of rainwater collected from roof run-off at the facility was added to each habitat 2 days prior to the addition of 40 newly hatched first instar larvae, to allow sufficient time for the formation of surface microlayer. Larvae were obtained from a laboratory colony of *An. gambiae* at the Kisumu

facility. Some habitats were left without larvae. All habitats were covered with aluminum window screen to keep out debris and prevent oviposition by native mosquitoes. Within each soil treatment, half of the habitats were randomly assigned a shade or sun treatment to complete a full factorial design with 8 replicates of each treatment combination. Water levels were maintained with additions of rainwater, as needed. Three habitats (one shaded, one sun plus Black Soil, and one in sun with Red Soil) received temperature recording probes. This experimental setup was used to collect samples to evaluate the changes in bacterial community structure in presence or absence of larval grazing pressure, four habitats of each larval density (larvae, no larvae) x soil type treatment combination were randomly selected for sampling. These habitats were the ones exposed to sunlight. Thirteen ml of surface microlayer was collected on day 5 (larvae added on day 0), and stored on ice until taken to the lab for further analysis. Surface microlayer samples (13 ml) were collected using a syringe and 16 gauge needles, as described in Chapter 2. At the lab, 10 ml was utilized for DNA extraction and 3 ml was preserved in 4% formalin for direct counts of bacteria, using methods described in Chapter 2.

Nucleic Acid Extraction, Amplification, Cloning and Sequencing Surface microlayer samples from the habitats were centrifuged at 6000 rpm for 30 minutes and the pellet was retained for nucleic acid (DNA) extraction. The DNA was extracted using Ultraclean Soil DNA kits (Cat. # 12800-50, MO BIO Laboratories Inc., Carlsbad, California) following the manufacturer's instructions. The presence of DNA was confirmed on 1% agarose gels in TBE buffer. Extractions from each sample were stored at -20° C. An approximately 1.3 kb region of a consensus 16S rRNA gene of bacteria was amplified by PCR

amplification using forward primer 63f 5' -CAG GCC TAA CAC ATG CAA GTC- 3' and reverse primer 1387r 5' – GGG CGG WGT GTA CAA GGC- 3' (Marchesi et al. 1998). The PCR reactions conditions were carried out as per the Taq DNA polymerase instruction manual (New England Biolabs). PCR conditions consisted of initial denaturing at 94°C for 2 min followed by 25 cycles of denaturing at 94°C for 45 sec, annealing at 55°C for 30 sec, extension at 72°C for 1.3 min, and final extension at 72°C for 7 min. The resulting PCR products were subjected to low melting agarose gel electrophoresis; bands were then excised and purified by sodium acetate and ethanol precipitation. Purified 16S rDNA fragments were cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin) and transformed clones were picked and purified. The plasmid DNA from transformed clones was extracted using Wizard Plus SV Miniprep kit (Promega). After confirmation that the plasmids contained an insert of the expected size, these plasmids (or clones for some samples) were then subjected to high throughput sequencing using dideoxy dye terminator chemistry, at the Genomic Technology Support Facility, Michigan State University. The 519R 5'- G(AT)ATTACCGCGGC(GT)GCTG- 3' sequencing primer (Lane et al. 1985) which reads the 16S rDNA was used to obtain partial 16S rDNA sequences.

Classification of sequences Sequence data from each clone were examined for chimera formation using the CHECK_CHIMERA program offered through Ribosomal Database Project II (RDP II; Release 8.1, May 18, 2001 (Maidak et al. 1999) following procedures outlined in Lilburn et al. (1999). Possible chimeras and sequences with fewer than 250 base pair in length were excluded from any further analysis. The 16S rDNA sequences

were then classified to named, operational taxonomic units (OTUs) using the Ribosomal Database Project II (RDP II; Release 9.38, April 03, 2006, http://rdp.cme.msu.edu). The RDP II contained 210,976 aligned 16S rRNA sequences as of April 03, 2006 and this database formed the reference library for the classifications here. The hierarchy model used by the naïve Bayesian rRNA classifier in RDP II comes from the bacterial classification proposed in release 6.0 of the Bergey's Manual of Systematic Bacteriology (http://bergeysoutline.com). The classifier calculates the joint probability of finding eight base subsequences ("words") in the query. When a query sequence is submitted, the joint probability of observing all the words in the query can be calculated separately for each genus from the training set probability values. Using the naïve Bayesian assumption, the query is most likely a member of the genus with the highest probability, given the limitations of the available database (Cole et al. 2005). This classification process puts a name on each sequence.

Sequences from all 8 libraries (2 replicates from each treatment combinations) were aligned separately with existing bacterial sequences in ARB 16S rDNA database (http://www.arb-home.de) and phylogenetic trees were constructed using the ARB software package (Strunk and Ludwig 1997, Lilburn et al. 1999). The selection of the reference sequence for initial sample sequence alignment was based on the RDP II classification of the samples to genus level. A mask was generated by the filtering method in ARB to produce valid columns for comparison of sequences. Sequences with short lengths were excluded so as to generate at least 265 valid columns. Distance matrices were generated using the above filter and Jukes-cantor correction. From these

libraries, 4 new libraries were created, as follows: (1) Black Soil/Larvae Present; (2) Black Soil/Larvae Absent; (3) Red Soil/Larvae Present; and (4) Red Soil/Larvae Absent.

Statistical comparisons Sequence libraries were compared between clone libraries using the LIBSHUFF program, (http://www.arches.uga.edu/~whitman/libshuff.html, Singleton et al. 2001). The program estimates differences between homologous coverage curves $C_X(D)$, and heterologous coverage curves $C_{XY}(D)$, by a Cramér-von Mises-type statistic, $\Delta C_{XY} = \sum(C_X - C_{XY})^2$. The input for these two-way comparisons was the distance matrices of aligned sequences generated in ARB. The distance matrices derived from nucleotide sequence differences generated in ARB were also used as input for diversity analysis using the DOTUR program (Distance based OTU and Richness determination; http://www.plantpath.wisc.edu/fac/joh/dotur.html). A 97% similarity value was used to determine operational taxonomic units (OTUs), (Schloss and Handelsman 2005). Using these new OTU assignments, which were un-named in contrast to the classification in RDP II, DOTUR constructed rarefaction curves for sampling intensity, richness estimators, and diversity indices. To address relative evenness of the bacterial communities reflected by sequence libraries, rank abundance graphs were generated where the abundance (i.e., frequency) of each OTU was plotted on a logarithmic scale against the rank from most abundant to least abundant (Magurran 2004).

Principal Component Analysis (PCA) PCA is a technique for simplifying a dataset where the data are transformed to a new coordinate system such that the greatest variance by any projection of the data lies on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA can

be used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance. PCA analysis was invoked for the sequence dataset with RDP classification as well as OTUs obtained by DOTUR analysis. Factorial ANOVA was performed on the PCA scores and the sequences that explained most of the variance were identified through the PCA loading values.

Terminal Restriction Fragment Length Polymorphism Analysis (TRFLP) The DNA extraction procedure and the PCR condition were the same as above with the exception that the forward 63f primer was labeled with 6-FAM (Integrated DNA Technologies, Coralville IA) at the 5' end. Three 100 µl PCR reactions were performed for each sample to assure maximum coverage of the diversity of 16S rDNA present in the sample. These reactions were combined and purified using Qiaquick PCR purification kit (Qiagen Valencia, CA). To generate TRFLP fragments, restriction enzyme HhaI and MspI (New England Biolabs Inc.) were selected since they offer decent number of restriction fragment and discrimination between different species (Marsh 1999).

The digestions were carried out in a two 10 µl aliquots (Liu et al. 1997, Marsh et al. 2000). 400-600 ng of purified PCR product was digested for 3 hours with either HhaI or MspI at 37° C. The reaction mixture contained approximately 400 ng (1-5 µl) of DNA, 2 µl of 10x restriction enzyme buffer, 1.5 µl of the one of the above restriction enzymes. The reaction mixture was brought to 10 µl with nuclease free water. After 3 hrs of digestion, the reactions were stopped at 60° C for 20 minutes. The T-RF lengths from the digest were determined on an automated DNA sequencer (ABI PRISM 310, Foster City, California) in GeneScan mode. The fragments were separated by electrophoresis using capillary electrophoresis system. The fragment sizes were determined with ABI

Genescan Analysis Software (Applied Biosystems) and the alignment of the community profiles were done in Genotyper software (Genotyper 2.5, Applied Biosystems). The peaks ranging between 50-950 bp were used for further analysis. In case of Red Soil with larvae and without larvae two different samples with similar treatments were amplified by PCR (each 3 x 50 µl) and later pooled before the digestion was performed (same was done for Black Soil larvae and without larvae). Cluster analysis of TRFLP profiles was done using PAUP* 4.0 (http://paup.csit.fsu.edu/).

Nucleotide sequence accession numbers The sequences obtained during this study were deposited to GenBank. The accession numbers are as follows: EF139885 to EF140054 for clones obtained from Red Soil/larvae; EF140055 to EF140206 for clones obtained from Red Soil/no larvae; EF140207 to EF140406 for clones obtained from Black Soil/larvae; EF140407 to EF140586 for clones obtained from Black Soil/no larvae.

FIGURES IN THIS CHAPTER ARE PRESENTED IN COLOR.

RESULTS

Classification of sequences A total of 732 sequences was obtained. Twenty three sequences from Red Soil and 7 sequences from Black Soil libraries were excluded from the analysis as chimeras or sequences with short lengths (< 250 bp). A total of 170 and 152 16S rDNA clones was retained for further analysis from microcosms with Red Soil with larvae (RL) and no larvae (RNL), respectively. Similarly, a total of 200 and 180 sequences was retained from microcosms with Black Soil with larvae (BL) and no larvae (BNL), respectively. The number of base pairs per sequence used for the analysis ranged from 255-520 bp.

Classifications based on the RDP II for each of the four treatment combinations are summarized in Tables 3.1, 3.2, 3.3, and 3.4. The classification process assigned the sequences to 4 phyla, namely, Actinobacteria, Bacteroidetes, Cyanobacteria, and Proteobacteria. The most abundant phylum was Proteobacteria, representing 62.9% in RL, 65.8% RNL, 65.5% in BL and 45.5% in BNL. The second most abundant phylum was Bacteroidetes (RL, 22.9%; RNL, 15.1%; BL, 32.5%; and BNL, 24.4%). The remaining sequences were classified into Actinobacteria (RL, 11.7%; RNL, 1.3%; BL, 2%; BNL, 1.1%) and Cyanobacteria (RL, 0.6%, RNL, 5.9, BL 0.0%, and BNL 1.7%). In RNL and BNL libraries, there was a significant percent of unclasssified sequences (11.8% and 27.2%, respectively). There were a total of 14 and 19 generic classifications observed for RL and BL, and 19 and 14 classifications for RNL and BNL, respectively.

Each of the sequences in the clone libraries was placed into a phylogenetic tree using ARB, using the RDP II classifier as a guide for initial placement. For the sake of

brevity, these phylogenetic trees are shown in the Appendix A, with the exception of a set of sequences which were left "unclassified" by RDP II (see below).

Analysis of treatment effects

Effects of larval presence/absence The most abundant sequences obtained from Red Soil with larvae (RL) at the class level are in the order as follows: Betaproteobacteria (51.76%), Sphingobacteria (15.29%), Actinobacteria (11.18%), Alphaproteobacteria (9.41%), Flavobacteria (5.88%), and Cyanobacteria (0.59%) (Table 3.1). Interestingly, there were no gammproteobacteria in this library. Within the Betaproteobacteria class, most of the sequences belonged to one order, the Burkholderiales. The most dominant genus was *Polynucleobacter* (family Burkholderiaceae) along with unclassified sequences belonging to Alcaligenaceae (matching closest to Achromobacter) and Comamonadaceae. Within Sphingobacteria the most dominant genus was *Pedobacter* (family Sphingobacteriaceae). The sequences in Actinobacteria all belonged to the order Actinomycetales, and could not be classified to lower taxon than order but RDP II found the sequences closest to genus *Arsenicicoccus*. Within the Alphaproteobacteria, the dominant genus was *Novosphingobium* (family Spingomonadaceae) and sequences belonging to unclassified Rhizobiales, and within the class Flavobacteria, the genus *Flavobacterim* (family Flavobacteriaceae) was the most dominant.

In sequences obtained from Red Soil without larvae (RNL), the most dominant classes observed were Betaproteobacteria (45.39%), Flavobacteria (13.16%), Gammaproteobacteria (11.84%) and Cyanobacteria (5.92%). They comprised up to 76.3% of the total sequences. The other classes Alphaproteobacteria, Deltaproteobacteria

and Actinobacteria represented < 5% each (Table 3.2). Within the Betaproteobacteria class, the most common genera were *Limnobacter* (family Burkholderiaceae), unclassified *Incertae sedis* 5, and unclassified Comamonadaceae. Genus *Methylophilus* (family Methylophilaceae) was also present in this class. The genus *Flavobacterium* (family Flavobacteriaceae) was most abundant in the class Flavobacteria. Within Gammaproteobacteria, the genera *Aeromonas* (family Aeromonadaceae) and *Pseudomonas* (family Pseudomonanadaceae) were the most abundant. All the Cyanobacteria fell into unclassified cyanobacterial groups without assignment to lower taxa.

Comparing both RL and RNL communities, the number of unclassified bacteria was higher in RNL (11.8%) as compared to RL (2.4%) treatments. In the RNL treatments, 5% of the total classified sequences were Cyanobacteria, whereas only 0.6% in RL belonged to this class. Notably, there were no sequences from the class Gammaproteobacteria and Deltaproteobacteria in RL, whereas there were no sequences from class Sphingobacteria in RNL. *Pedobacter* was present and there were increased numbers of *Polynucleobacter* in RL. *Pedobacter* was absent in RNL treatments. A comparison of libraries of surface microlayer from RL and RNL indicated that the two libraries were significantly different from each other ($\Delta C_{RL/RNL}$= 1.344, p= 0.001; $\Delta C_{RNL/RL}$= 1.992, p= 0.001) (Fig 3.1).

Red Soil/no larvae treatments exhibited higher bacterial diversity estimated by diversity indices. The Simpson diversity index (1/D) for RL determined by DOTUR was 13.93 and that for RNL was 27.39. Note that the value of the index increases with increasing diversity. The Shannon diversity index was 3.09 for RL and 3.52 for RNL.

The OTU data obtained by DOTUR analysis was used to plot rank/abundance graph where the relative abundance of each OTU was plotted against the OTU (0.03 % dissimilarity) ranked from most abundant to the least abundant (Fig 3.2). The density of bacteria estimated by DAPI ranged from 2.6 x 10^6/ml to 6.1 x 10^6/ml in the surface microlayer of RL, and 9.7 x 10^7/ml to 38 x 10^7/ml in RNL.

Table 3.1. Bacterial community composition of surface microlayer from microcosms with Red Soil having larval grazing pressure (RL) based on 16S rDNA sequences classification with RDP II. The numbers of sequences observed per taxon are in parenthesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (19)	Actinobacteria (19)	Actinomycetales (19)	Microbacteriaceae (1)	unclassified Microbacteriaceae (1)
			unclassified Actinomycetales (18)	
Bacteroidetes (39)	Flavobacteria (10)	Flavobacteriales (10)	Flavobacteriaceae (10)	Flavobacterium (6)
				unclassified Flavobacteriaceae (4)
	Sphingobacteria (26)	Sphingobacteriales (26)	Sphigobacteriaceae (25)	Pedobacter(19)
				unclassified Sphingobacteriaceae (6)
			Crenotrichaceae (1)	Chitinophaga (1)
	unclassified Bacteroidetes (3)			
Cyanobacteria (1)	Cyanobacteria (1)	unclassified Cyanobacteria (1)		
Proteobacteria) (107	Alphaproteobacteria (16)	Rhodobacterales (2)	Rhodobacteraceae (2)	Rhodobacter (1)
				unclassified Rhodobacteraceae (1)
		Sphingomonadales (6)	Sphingomonadaceae (6)	phingomonas (1)
				Sphingopyxis (1)
				Novosphingobium (4)
		Rhizobiales (8)	unclassified Rhizobiales (8)	
	Betaproteobacteria (88)	Burkholderiales (85)	Comamonadaceae (19)	Rhodoferax (1)
				Polaromonas (4)
				Curvibacter (1)
				unclassified Comamonadaceae (13)
			Oxalobacteraceae (2)	Herbaspirillum(2)
			Alcaligenaceae (13)	unclassified Icaligenaceae (13)
			Burkholderiaceae (44)	Limnobacter (1)
				Polynucleobacter (42)
				unclassified Burkholderiaceae (1)
			Incertae sedis 5 (2)	Roseateles (2)
			unclassified Burkholderiales (5)	
		unclassified Betaproteobacteria (3)		
	unclassified Proteobacteria (3)			
unclassified Bacteria (4)				
Total Bacteria 170				

Table 3.2. Bacterial community composition of surface microlayer from microcosms with Red Soil with no larval (RNL) grazing pressure based on 16S rDNA sequences classification with RDP II. The number of sequences observed per taxon are in parenthesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (2)	Actinobacteria (2)	Actinomycetales (2)	Microbacteriaceae (1)	unclassified Microbacteriaceae (1)
			unclassified Actinomycetales (1)	
Bacteroidetes (23)	Flavobacteria (20)	Flavobacteriales (20)	Flavobacteriaceae (20)	Flavobacterium (18)
				unclassified Flavobacteriaceae (2)
	unclassified Bacteroidetes (3)			
Cyanobacteria (9)	Cyanobacteria (9)	unclassified Cyanobacteria (9)		
Proteobacteria (100)	Alphaproteobacteria (7)	Rhodobacterales (4)	Rhodobacteraceae (4)	Rhodobacter (2)
				unclassified Rhodobacteraceae (2)
		Sphingomonadales (2)	Sphingomonadaceae (2)	Novosphingobium (1)
				Unclassified Sphingomonadaceae(1)
		Rhizobiales (1)	unclassified Rhizobiales (1)	
	Betaproteobacteria (69)	Methylophilales (4)	Methylophilaceae (4)	Methylophilus (3)
				unclassified Methylophilaceae (1)
		Burkholderiales (61)	Comamonadaceae (18)	Polaromonas (1)
				Curvibacter (2)
				Ramlibacter (1)
				Hydrogenophaga (2)
				Acidovorax (1)
				unclassified Comamonadaceae(11)
			Oxalobacteraceae (2)	unclassified Oxalobacteraceae (2)
			Burkholderiaceae (17)	Limnobacter (11)
				Polynucleobacter (6)
			Incertae sedis 5 (16)	Roseateles (2)
				Aquabacterium (1)
				unclassified Incertae sedis 5 (13)
			unclassified Burkholderiales (8)	
		Neisseriales (1)	Neisseriaceae (1)	Vogesella (1)
		unclassified Betaproteobacteria (3)		
	Gammaproteobacteria(18)	Xanthomonadales (2)	Xanthomonadaceae (2)	Nevskia (2)
		Aeromonadales (7)	Aeromonadaceae (7)	Aeromonas (7)
		Pseudomonanadales (6)	Pseudomonanadaceae (6)	Pseudomonas (3)
				unclassified Pseudomonadaceae(3)
		Legionellales (1)	Coxiellaceae (1)	Aquicella (1)
		unclassified Gammaproteobacteria (2)		
	Deltaproteobacteria (6)	Bdellovibrionales (3)	Bdellovibrionaceae (3)	Bdellovibrio (3)
		unclassified Deltaproteobacteria (3)		
unclassified Bacteria (18)				
Total Bacteria	152			

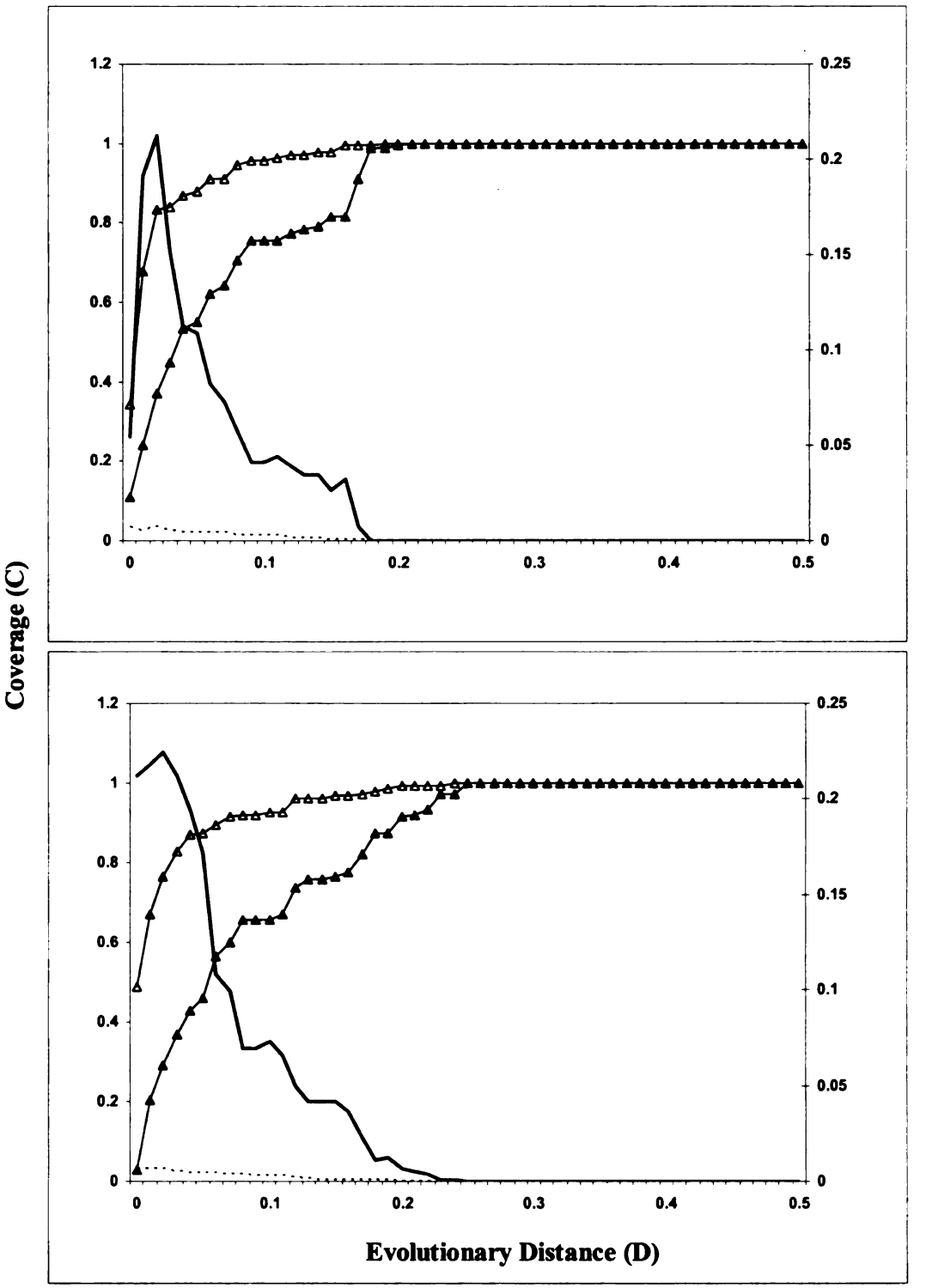


Figure 3.1. Comparison of bacterial 16S rDNA sequence libraries of Red Soil with larval grazing pressure (RL) to Red Soil with no larval grazing pressure (RNL) using LIBSHUFF. Homologous (open triangles) and heterlogous (solid triangles) coverage curves are shown. Solid lines indicate values of $(C_{RL}-C_{RL/RNL})^2$ (panel A) or of $(C_{RNL}-C_{RNL/RL})^2$ (panel B) at each value of evolutionary distance (D). Broken lines indicate the 950th value (or p=0.05) of corresponding $(C_{RL}-C_{RL/RNL})^2$ or $(C_{RNL}-C_{RNL/RL})^2$ for the randomized samples.

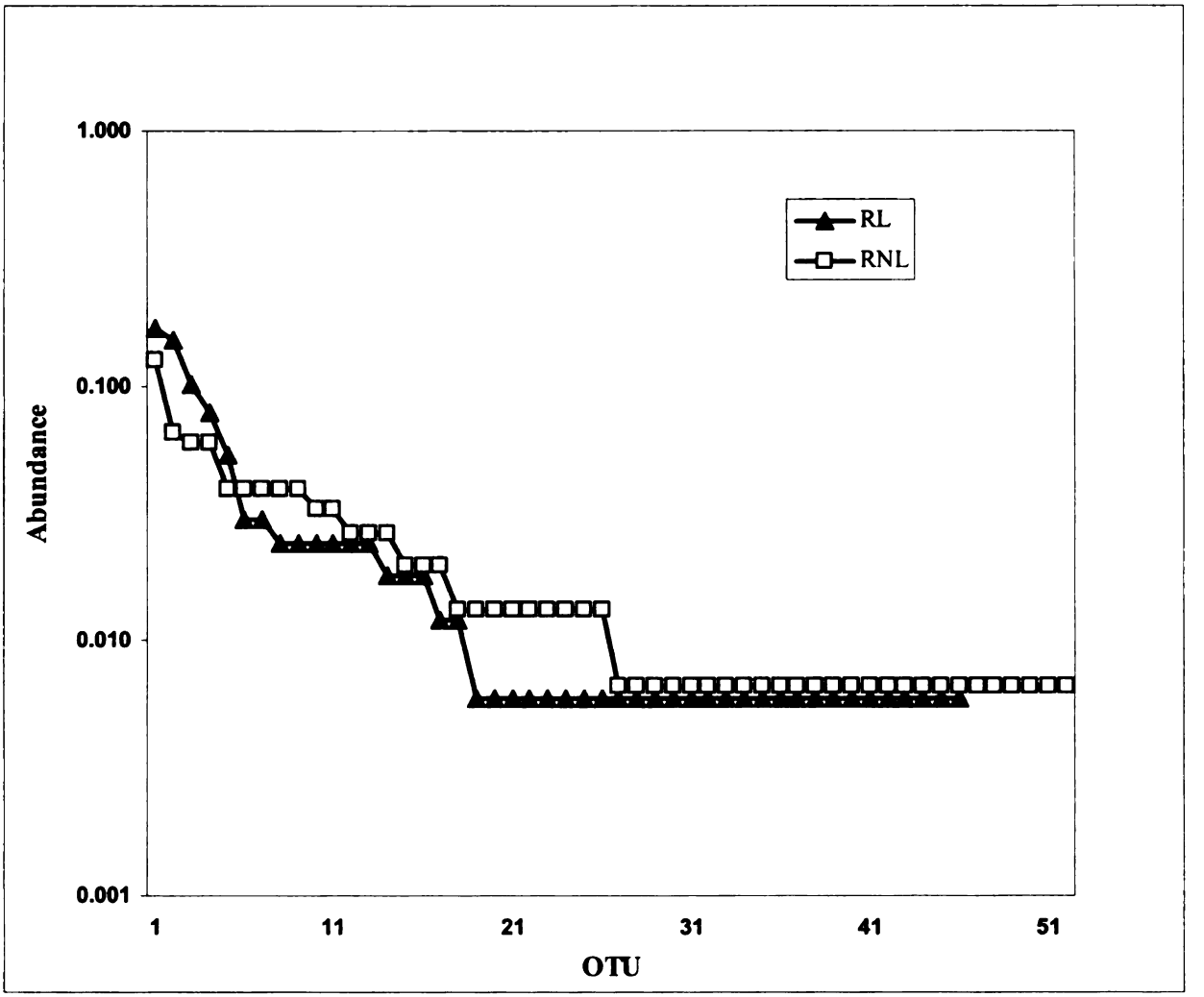


Figure 3.2. Rank abundance plot showing diversity of Red Soil with larvae (RL) and Red Soil with no larvae (RNL) bacterial communities at evolutionary distance of 0.03

T

the

and

Sp

we

W

Bu

we

we

Fla

Al

No

ob

do

Ga

the

un

les

mo

sea

an

as

The most dominant sequences obtained from Black Soil with larvae (BL) were from the class Betaproteobacteria (31.5%), Flavobacteria (28.5%), Alphaproteobacteria (24%), and Gammaproteobacteria (9%) (Table 3.3). The remaining sequences belonged to Sphingobacteria and Deltaproteobacteria comprising less than 5% of each class. There were no sequences assigned to the unclassified bacteria category in BL treatments. Within the Betaproteobacteria class, most of the sequences were assigned to the order Burkholderiales and almost one third remained unclassified. Of the rest the most common were unclassified Comamonadaceae. Most of the sequences from the class Flavobacteria were classified to the genus *Flavobacterium*, or were classified to the family Flavobacteriaceae but not assigned to genus. The most dominant genera in the class Alphaproteobacteria were *Rhodobacter* (family Rhodobacteraceae) and *Novosphingobium* (family Sphingomonadaceae). *Aeromonas* was the only genus observed within Gammaproteobacteria to which sequences were classified.

Within sequences obtained from Black Soil without larvae (BNL), the most dominant classes were Flavobacteria (17.8%), Betaproteobacteria (16.1%), Gammaproteobacteria (16.1%) and Alphaproteobacteria (13.3%) (Table 3.4). Together they comprised 63.33% of the total sequences, while 27.2% of the total sequences were unclassified bacteria, and the remainder Sphingobacteria and Cyanobacteria represented less than 5% of the total each. Sequences assigned to the genus *Flavobacterium* were the most dominant in the class Flavobacteria. Unclassified sequences assigned to *Incertae sedis* 5 were the most common sequences in the family Betaproteobacteria, *Aeromonas* and *Pseudomonas* were the most common in the Gammaproteobacteria, and sequences assigned to *Rhodobacter* were the most common in the class Alphaproteobacteria.

A comparison of libraries of surface microlayer from BL and BNL treatments indicated that the two libraries were significantly different from each other ($\Delta C_{BL/BNL}$= 0.327, p= 0.001; $\Delta C_{BNL/BL}$= 2.934, p= 0.001) (Fig 3.3). The Simpson diversity index (1/D) for BL obtained by DOTUR analysis was 11.88 and for BNL 12.29. The Shannon index was 2.83 and 3.19, respectively. Bacterial diversity in BNL treatments was higher than BL indicated by the diversity indices. The OTU data obtained by DOTUR were used to plot rank/abundance graph where the relative abundance of each OTU was plotted against the OTU (0.03 % dissimilarity) ranked from most abundant to the least abundant (Fig 3.4). The total number of bacteria ranged from 7.7 x 10^6/ml to 16 x 10^6/ml in the surface microlayer of BL treatments, and the sole sample available from BNL had count of 1.9 x 10^7/ml. The rarefaction curves indicated that all four sequence libraries were undersampled, as the curves did not reach an asymptote given the number of sequences in each clone libraries (Fig. 3.5).

Table 3.3. Bacterial community composition of surface microlayer from microcosms with Black Soil with larval grazing pressure (BL) based on 16S rDNA sequences classification with RDP II. The number of sequences observed per taxon are in parenethesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (4)	Actinobacteria (4)	Actinomycetales (4)	Microbacteriaceae (3)	unclassified Microbacteriaceae (3)
			unclassified Actinomycetales (1)	
Bacteroidetes (65)	Flavobacteria (57)	Flavobacteriales (57)	Flavobacteriaceae (57)	Flavobacterium (2) unclassified Flavobacteriaceae (55)
	Sphingobacteria (4) unclassified Bacteroidetes (4)	Sphingobacteriales (4)	Sphingobacteriaceae (4)	Pedobacter(4)
Proteobacteria (131)	Alphaproteobacteria (48)	Rhodobacterales (30)	Rhodobacteraceae (30)	Rhodobacter (21) unclassified Rhodobacteraceae (9)
		Sphingomonadales (14)	Sphingomonadaceae (14)	Sphingopyxis (1) Novosphingobium (13)
		Rhizobiales (4)	unclassified Rhizobiales (4)	
	Betaproteobacteria (63)	Burkholderiales (60)	Comamonadaceae (17)	Polaromonas (1) unclassified Comamonadaceae (16)
			Oxalobacteraceae (7)	Herbaspirillum(3) unclassified Oxalobacteraceae (4)
			Alcaligenaceae (8)	unclassified Alcaligenaceae (8)
			Burkholderiaceae (3)	Polynucleobacter (3)
			Incertae sedis 5 (1)	unclassified Incertae sedis 5 (1)
			unclassified Burkholderiales (24)	
		Neisseriales (1)	Neisseriaceae (1)	Aquaspirillum (1)
		unclassified Betaproteobacteria (2)		
	Gammaproteobacteria (18)	Aeromonadales (18)	Aeromonadaceae (18)	Aeromonas (18)
	Deltaproteobacteria (1)	Bdellovibrionales (1)	Bdellovibrionaceae (1)	Bdellovibrio (1)
	unclassified Proteobacteria (1)			
Total Bacteria	200			

Table 3.4. Bacterial community composition of surface microlayer from microcosms with Black Soil with no larval (BNL) grazing pressure based on 16S rDNA sequences classification with RDP II. The number of sequences observed per taxon are in parenethesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (2)	Actinobacteria (2)	Actinomycetales (1)	Microbacteriaceae (1)	unclassified Microbacteriaceae (1)
		unclassified Actinobacteria (1)		
Bacteroidetes (4)	Flavobacteria (32)	Flavobacteriales (32)	Flavobacteriaceae (32)	Flavobacterium (13)
				unclassified Flavobacteriaceae (19)
	Sphingobacteria (5)	Sphingobacteriales (5)	Sphigobacteriaceae (5)	Pedobacter(3)
				unclassified Sphingobacteriaceae (2)
	unclassified Bacteroidetes (7)			
Cyanobacteria (3)	Cyanobacteria (3)	unclassified Cyanobacteria (3)		
Proteobacteria (82)	Alphaproteobacteria (24)	Rhodobacterales (21)	Rhodobacteraceae (21)	Rhodobacter (9)
				unclassified Rhodobacteraceae(12)
		Rhizobiales (2)	Rhizobiaceae (1)	unclassified Rhizobiaceae (1)
			unclassified Rhizobiales (1)	
		unclassified Alphaproteobacteria (1)		
	Betaproteobacteria (29)	Burkholderiales (27)	Comamonadaceae (6)	Polaromonas (1)
				Caenibacterium (1)
				Hydrogenophaga (1)
				Acidovorax (1)
			unclassified Comamonadaceae (2)	
			Incertae sedis 5 (12)	unclassified Incertae sedis 5 (12)
			unclassified Burkholderiales (9)	
		Neisseriales (2)	Neisseriaceae (2)	Aquaspirillum (2)
	Gammaproteobacteria (29)	Xanthomonadales (5)	Xanthomonadaceae (5)	Silanimonas (1)
				Nevskia (3)
		unclassified Xanthomonadaceae (1)		
		Aeromonadales (11)	Aeromonadaceae (11)	Aeromonas (11)
		Pseudomonanadales (13)	Pseudomonanadaceae (9)	Pseudomonas (4)
				Cellvibrio (1)
				Flavimonas (1)
				unclassified Pseudomonadaceae(3)
			unclassified Pseudomonadales (4)	
unclassified Bacteria (49)				
Total Bacteria 180				

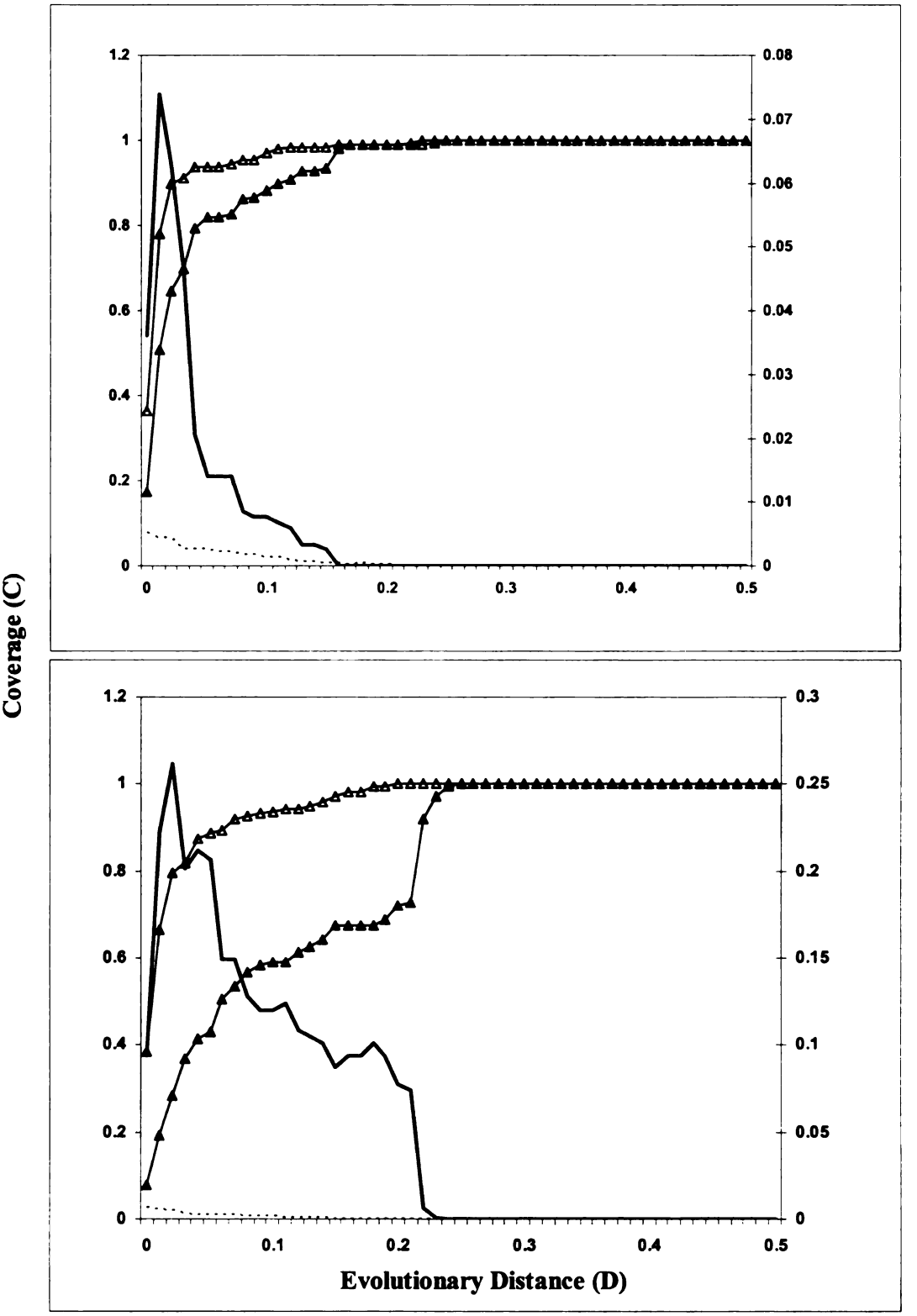


Figure 3.3. Comparison of bacterial 16S rDNA sequence libraries of Black Soil with larval grazing pressure (BL) to Black Soil with no larval grazing pressure (BNL) using LIBSHUFF. Homologous (open triangles) and heterlogous (solid triangles) coverage curves are shown. Solid lines indicate values of $(C_{BL}-C_{BL/BNL})^2$ (panel A) or of $(C_{BNL}-C_{BNL/BL})^2$ (panel B) at each value of evolutionary distance (D). Broken lines indicate the 950th value (or p=0.05) of corresponding $(C_{BL}-C_{BL/BNL})^2$ or $(C_{BNL}-C_{BNL/BL})^2$ for the randomized samples.

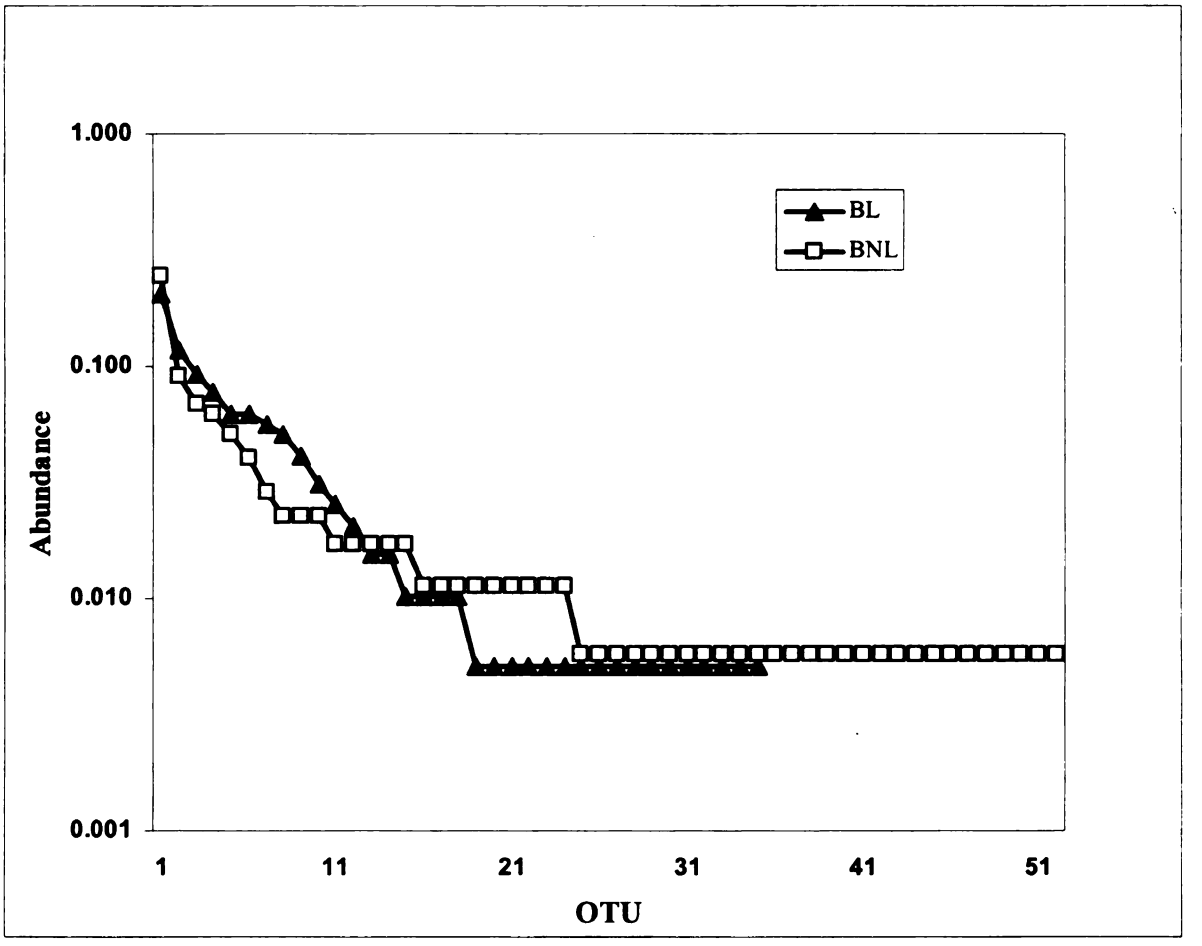


Figure 3.4. Rank abundance plot showing diversity of Black Soil with larvae (BL) and Black Soil with no larvae (BNL) bacterial communities at evolutionary distance of 0.03

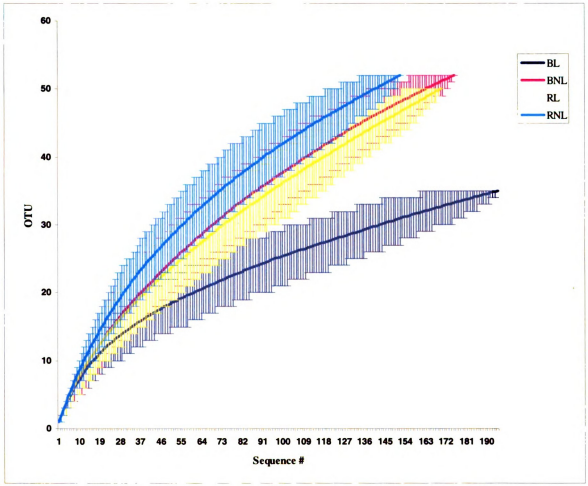


Figure 3.5. Rarefaction curve generated from the OTUs at 3% distance using DOTUR for all four treatments (RL, RNL, BL, BNL).

Effects of soil type LIBSHUFF analysis was also performed to determine if the libraries from larvae/no larvae treatments from both soil types were different statistically. Distance matrices from Red Soil with larvae and Black Soil with larvae were compared ($\Delta C_{BL/RL}$= 1.203, p= 0.001; $\Delta C_{RL/BL}$= 0.856, p= 0.001), and similarly Red Soil and Black Soil without larvae were compared ($\Delta C_{BNL/RNL}$= 1.752, p= 0.001; $\Delta C_{RNL/BNL}$= 1.411, p= 0.001), indicating all four libraries were statistically different from one another and that none of the libraries were subsets of any other library (Fig. 3.5).

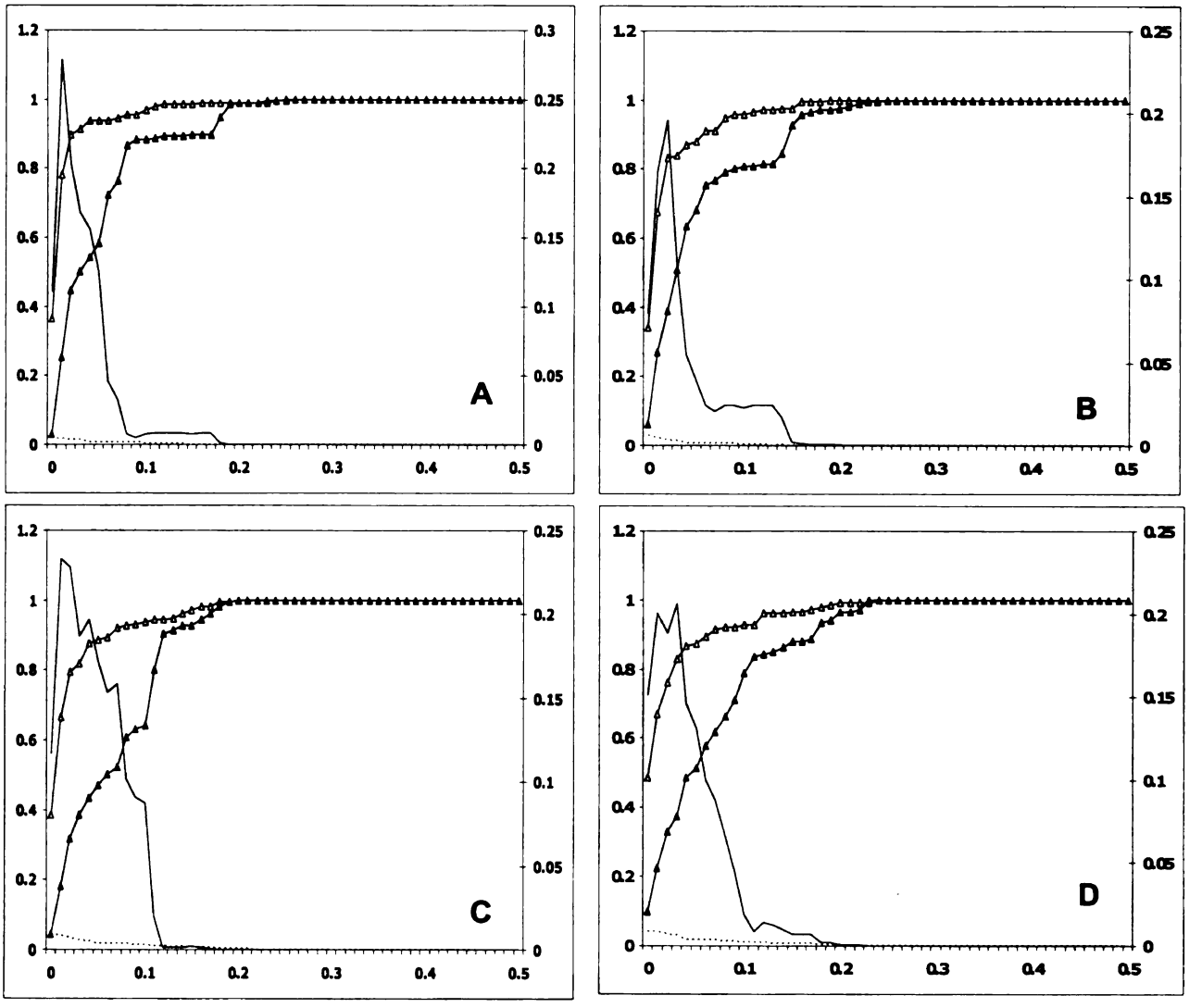


Figure 3.6. Homologous (open triangles) and Heterologous (solid triangles) coverage curves for 16S rDNA sequence libraries from Black Soil and Red Soil with larvae and without larvae are shown. Solid lines indicate the value of $(C_X-C_{XY})^2$ or $(C_Y-C_{YX})^2$ for the original samples at each evolutionary distance. Broken lines indicate the 950th value (or p=0.05). The above comparisons are BLvsRL (Panel A, B) and BNLvsRNL (Panel C,D)

Principal Components Analysis PCA was invoked to identify a pattern with the OTU dataset (0.03 dissimilarity distances) between the larvae/no larvae treatments and to distinguish treatment effects as well. The OTU data obtained from DOTUR using a 0.03 distance criterion was used to perform PCA analysis. 46.17% variance was explained by PC1-2. Factorial ANOVA of PCA Scores for OTUs on PC I showed significant larvae/no larvae treatment effects (Fig. 3.7 A) (Larvae, F = 59.3, P = 0.0015; Soil, F = 0.01, P = 0.91; and Soil*Larvae, F = 2.89, P = 0.16). On PC II ANOVA showed significant soil effects (Larvae, F = 0.14, P = 0.73; Soil F = 63.68, P = 0.0013; and Soil*Larvae F = 2.42, P = 0.19). Using the factor loadings it was possible to explain which sequences caused this variation and in PC 1, and those sequences belonged to unclassified Alcaligenaceae (closest matching genus was *Achromobacter*) from order Burkholderiales, *Flavobacterium* and unclassified Burkholderiales (closest matching genus was *Herbaspirillum*) present in Black Soil larvae treatments. The same analysis was performed on RDP II classification data at the order level for comparison. PCA on taxonomic classification by RDP II classification data indicated significant larvae/ no larvae effect (Fig. 3.7 B). 59.41% variance was explained by PC1 and PC2. Factorial ANOVA of PCA Scores for PC I for taxa from RDP II classification was significant for larvae/ no larvae treatment but not significant for soil neither in PC I nor PC II (Larvae F=24.3976, p= 0.0078, Soil F=0.1819, p=0.6917and Soil*Larvae F=0.1895, p=0.6858). Unclassified bacteria, Pseudomonanadales, Actinomycetales (class Actinobacteria), Sphingobacteriales (class Sphingobacteria) and Sphingomonadales (Alphaproteobacteria) contributed to most of the explained variation.

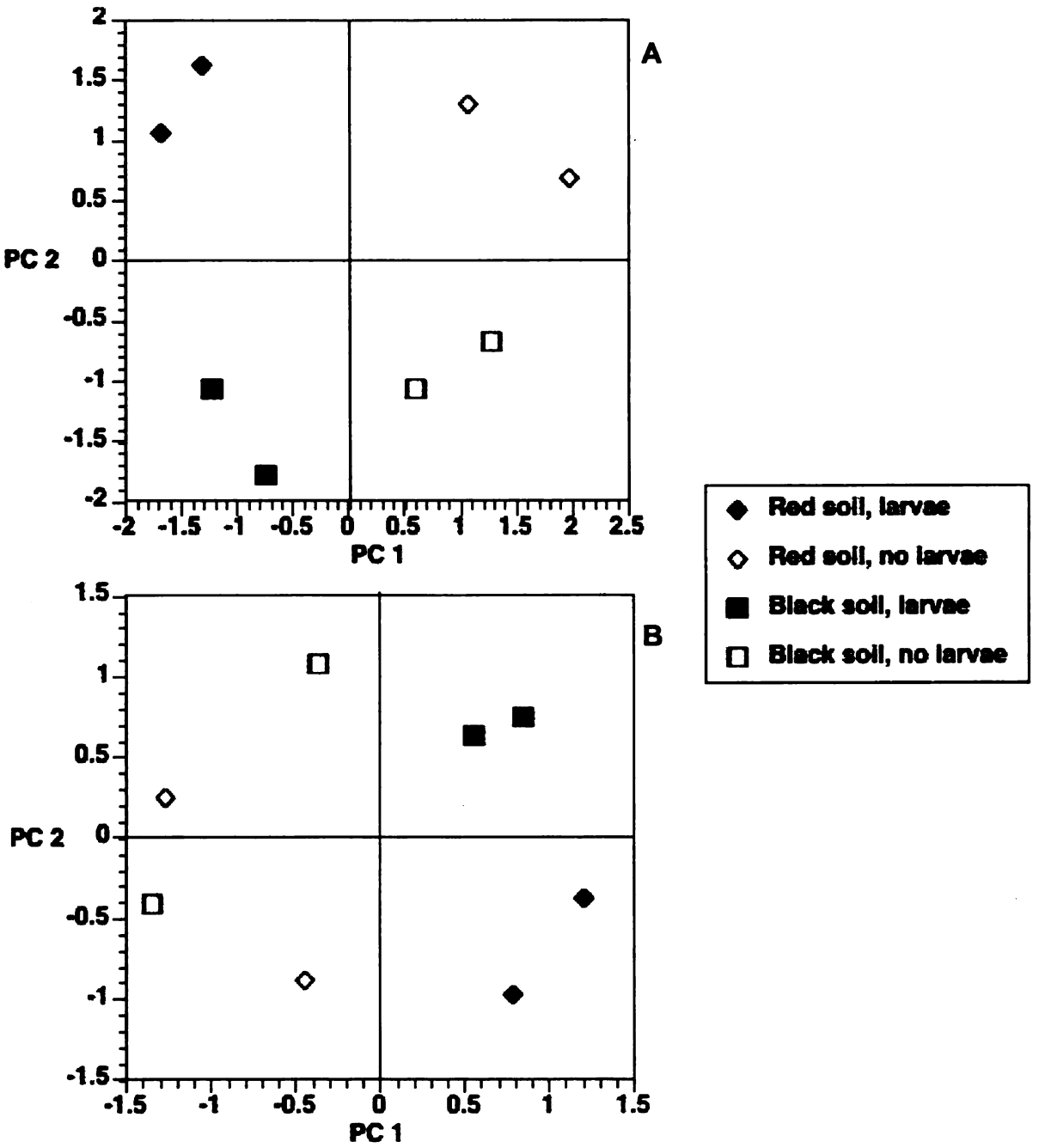


Fig 3.7. A, B: Principal component analysis of the 16S rDNA sequence data from Red Soil with/ without larvae and Black Soil with/ without larvae. OTUs determined by DOTUR analysis at 3% distance, A; OTUs obtained by RDP II classification at order level, B; 46.27% of the variance was explained by PC1-2 for DOTUR OTU data and 59.41% of the variance was explained by PC1-2 for RDP classification data.

RDP classification revealed that there were a total of 75 sequences that fell either in unclassified bacteria or unclassified Cyanobacteria category. A blastn search of these sequences indicated that these sequences bear a close resemblance to chloroplast sequences from algae. Phylogenetic analysis was performed to place these sequences in an evolutionary tree with chloroplast sequences (Fig. 3.8). Notably the Red Soil sequences clustered separately from the Black Soil sequences indicating that there were differences in the community composition of these two soil types.

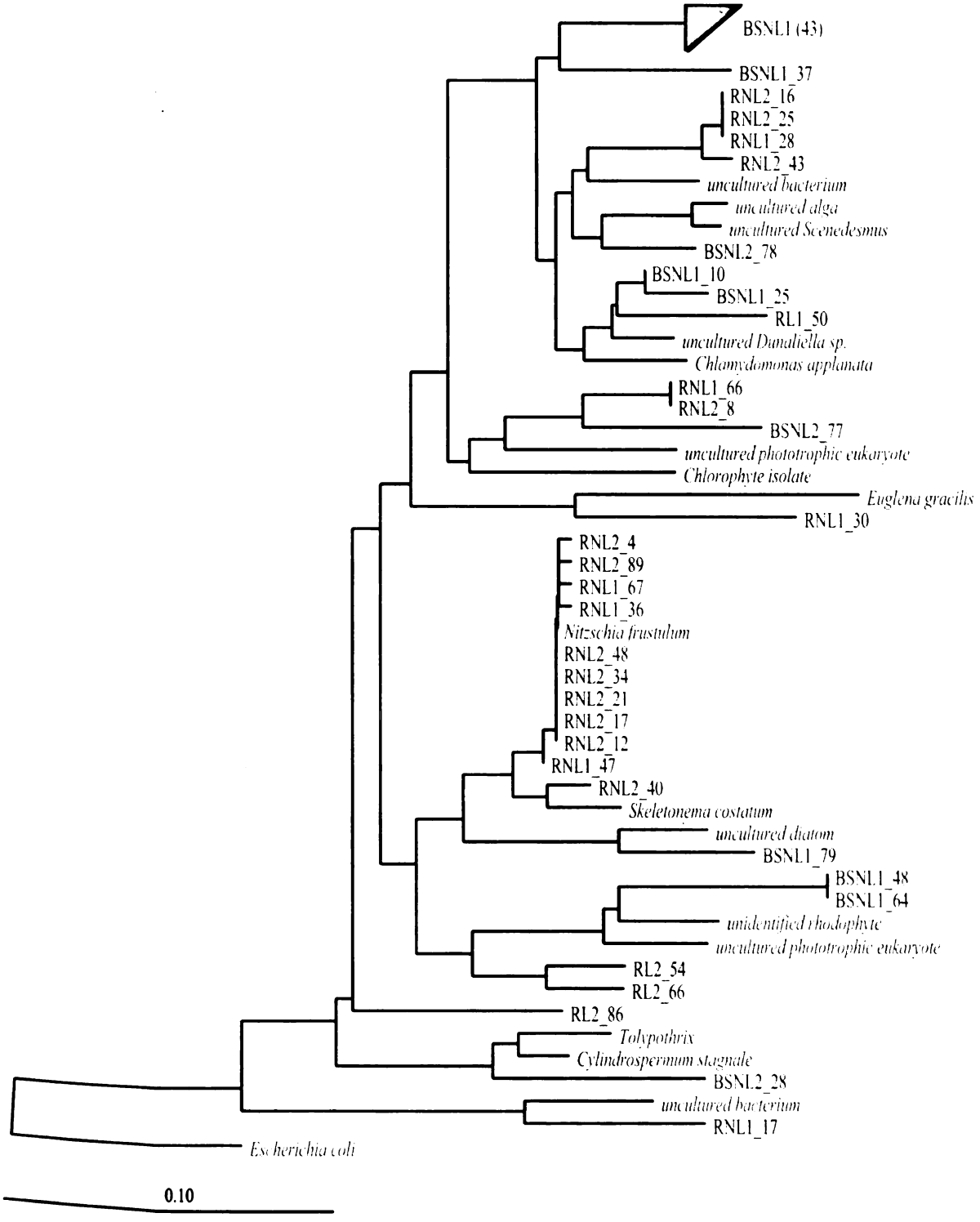


Figure 3.8. Phylogentic analysis of 16S rDNA gene sequences from the unclassified group from various treatments. The dendogram was generated by Neighbor Joining Method using a filter of 241 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequence within each group are in parenthesis. RL=Red Soil/ larvae; RNL=Red Soil/no larvae; BSNL=Black Soil/no larvae

Terminal Restriction Fragment Length Polymorphism Analysis (TRFLP) Results of cluster analysis of TRFLP data are shown in Fig 3.8. The total number of unique differently sized terminal fragments obtained from HhaI digestion for all four treatments was 124. BL produced 36, BNL 50 terminal fragments individually. Sixty-six terminal fragments were obtained in combined BL and BNL profiles, with 16 unique in BL and 30 in BNL. The rest 20 terminals fragments were common in both the treatments. Ninety-seven fragments were obtained in RL and RNL profiles combined, with 20 unique in RL and 46 unique in RNL. Thirty-one fragments were common in both treatment types.

The total number of 65 uniquely sized terminal fragments was obtained from MspI digests of all four treatments. BL produced 21, BNL 32 individually. Thirty-eight terminal fragments were obtained in combined BL and BNL profiles, with 6 unique in BL and 17 in BNL. The remaining 15 terminals fragments were common in both the treatments. Similarly RL produced 31 and RNL 27 fragments individually. Forty-three fragements were obtained in both RL and RNL profiles combined, with 16 unique in RL and 12 unique in RNL. 5 fragments were common in both treatment types.

Cluster analysis based on the UPGMA showed that samples typically segregated by larvae/no larvae treatments within soil types along the trees generated by the PAUP program, for the MspI restriction fragments (Fig. 3.9). A similar result was observed HhaI restriction fragments for Black Soil, however the Red Soil no larvae treatments formed their own branch separately from the other three treatment combinations.

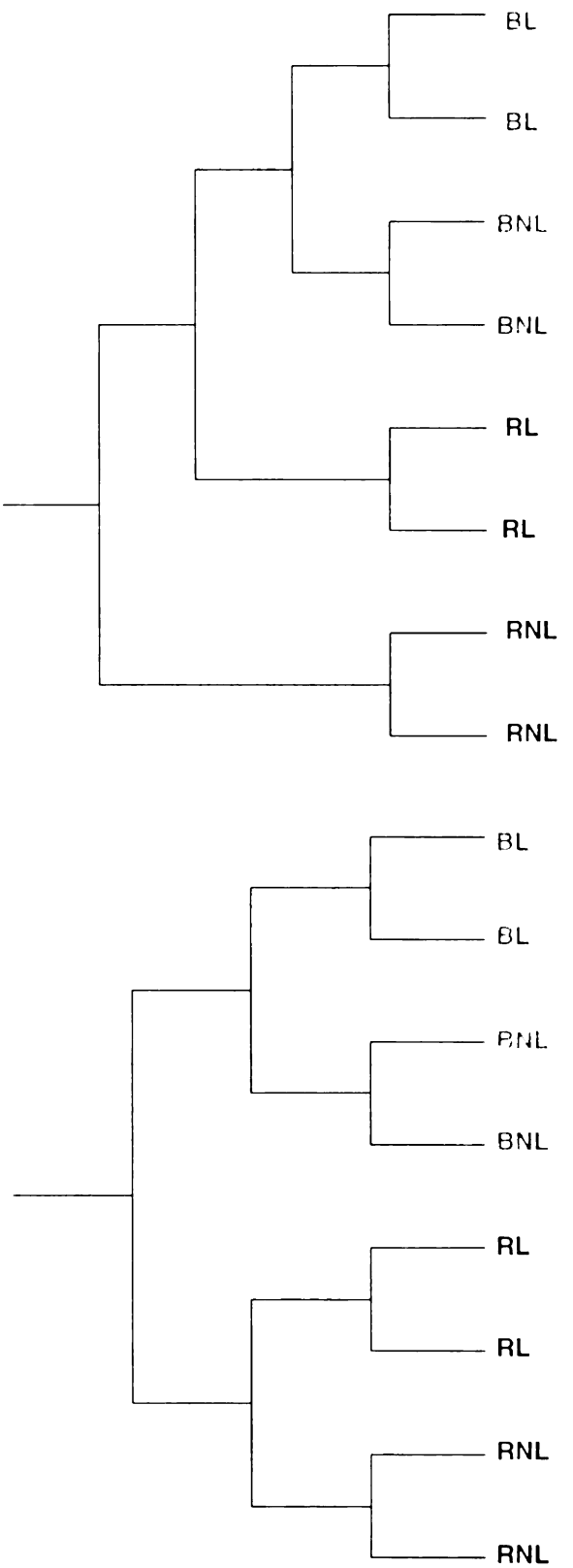


Figure 3.9. TRFLP Cluster Analysis using Paup for presence/absence of the terminal fragements obtained for HhaI and MspI restriction enzyme digestions. (RL, Red Soil with larvae; RNL, Red Soil with no larvae; BL, Black Soil with larvae; BNL, Black Soil with no larvae).

DISCUSSION

Overall, the results of this study demonstrate that the bacterial community was highly diverse in the habitats that were constructed for this study and that experimental treatment effects on community composition were evident. These microcosms very likely replicated the natural habitats of the *An. gambiae* for the following reasons: (1) natural rain water was used; (2) local soil was used; (3) they were established in a natural setting. Accordingly, results from this study would be expected to be similar to experiments done in entirely natural settings. The high diversity here, with a large number of bacterial taxa and some 16S rDNA sequences that were unclassified at various levels of taxonomic group, is typical of investigation of bacterial communities in most environments that have been studied to date (Hugenholtz et al. 1998, Suzuki et al. 1997, Giovannoni et al. 1990). Further, the pattern of diversity as reflected by the shape of rank abundance curves is also typical, with few common and many rare taxa (Magurran 2004, Hughes et al. 2001). Despite this high diversity in this study system there were discernable changes in the bacterial community attributable to the experimentally imposed treatment effects (i.e., predation by mosquito larvae; underlying soil type).

Various methods have been employed to study bacterial community diversity in a wide range of environments. Any given method will influence the way one interprets this diversity. No method is without limitations and depending on the research question, one has to use the best method available to address the relevant hypothesis. In this study, I used two molecular approaches (16S rDNA sequence library construction, and 16S rDNA TRFLP analysis) to generate genotypic data, followed by several statistical procedures

and summary analyses to test the hypotheses stemming from the factorial experimental design. Analysis of 16S rDNA sequence data, or TRFLP frequency data, is daunting due to the tremendous diversity encountered by such methods, but currently they represent the most powerful tools for studies of this kind, particularly when culturing is not feasible. These are genotypic-based methods and are rooted in phylogenetic analysis. Other methods that could be used include culture-based methods on enriched or minimal essential media; direct count methods using various stains and genotypic probes; and phenotypic analyses such as BIOLOG (a system that identifies carbon substrates utilized in the system); and cellular fatty acid analysis (which generates chromatographic signatures representing bacterial community composition). Chapter 1 reviews methods and approaches and indicates their advantages and disadvantages. For example, clone sequence libraries do not provide data on abundance of organisms (Zoetendal et al. 2004); however, one can make inferences about relative abundance of different taxa on the basis of frequency of clones recovered, even with the well known problem of operon copy number variation leading to biased estimates (Farrelly 1995, Klappenbach et al. 2000).

The second challenge in analysis of bacterial community diversity is the choice of statistical analyses of the genotypic data, both for purposes of summarization, comparison, and to test hypotheses in an experimental context. For this reason, I chose several different methods to analyze the sequence libraries in particular, with the aim of determining if these different methods supported a set of general conclusions, which could lead to a strong inference. Here, the use of analysis of variance on factor scores generated by principle components analysis allows such an inference, and the problem

was separately approached with the LIBSHUFF procedure. Both methods provide statistical tests. The principle components analysis *per se*, and the dendrogram generated from the TRFLP data, are data reduction methods which provide support for these general conclusions but do not provide statistical tests. Similarly, diversity, richness, and evenness indices, and rank abundance curves, provide summary statistics or graphical representations.

Despite the high bacterial community diversity, there were clear experimental treatment effects that allow the following general conclusions. First, presence of actively feeding and growing *An. gambiae* larvae resulted in a change in community composition, compared to those microcosms where larvae were absent. This effect is consistent with the phenomenon of "top down" effects of predation on lower trophic level community structure observed in other aquatic ecosystems (Carpenter et al. 1985), and it indicates that larval mosquito feeding within the surface microlayer has a dramatic effect on composition of that layer, a finding consistent with recent studies on algal biomass (Gimnig et al. 2002, Kaufman et al. 2006). This effect was mediated both by changes in frequencies of certain taxa represented by the sequence data, and by presence or absence of certain taxa. In particular, several groups were either absent or decreased in frequency, possibly due to larval grazing. Gammaproteobacteria were completely absent in treatments with larvae and Red Soil and were in reduced frequency in treatments with larvae in Black Soil compared with other combinations. These finding are consistent with Kaufman et al. (1999) in which cultivated pseudomonads decreased in frequency under feeding pressure by larvae of the mosquito *Ochlerotatus triseriatus*, in contrast with cultured Enterobacteriaceae which increased in relative proportion. The number of

sequences belonging to group of unclassified Bacteria, genus *Flavobacterium*, unclassified *Incerta sedis* 5, genus *Limnobacter*, unclassified Cyanobacteria, unclassified Burkholderiales, and genus *Aeromonas* decreased in Red Soil treatments with larvae. By contrast there was a trend towards increase in frequency of sequences belonging to Betaproteobacteria and Alphaproteobacteria in both soil treatments in presence of larvae. Also there was an increase in frequency of sequences of *Polynucleobacter*, unclassified Actinomycetales, unclassified Rhizobiales and unclassified Sphingobacteriaceae. *Pedobacter* and unclassified Alcaligenaceae sequences were present only in samples from Black Soil with larvae. These results also indicate that apart from a marked larval effect, there were differences in bacterial community composition in the soil types.

PCA identified different taxa changing in frequency under larval feeding pressure, depending upon whether RDP II or DOTUR were used to classify the sequence data. However, the overall conclusion that bacterial community shifts occurred is upheld. The differences between RDP II and DOTUR are likely related to the algorithms used to classify sequences and the robustness of the sequence library in RDP II. It is unlikely that variation in ingestibility explains these patterns since the mosquito larvae are generalist collector gatherers and would ingest all the bacteria present in the surface microlayer, given the size range of bacteria (Merritt et al. 1992).

The soil effects on the bacterial community composition were less pronounced than were the larval feeding effects, but they were present and detected by LIBSHUFF and PCA. LIBSHUFF analysis revealed significant differences in clone libraries from RL and RNL; the same was true for BL and BNL. Furthermore, there were significant differences in clone libraries RL and BL and similarly RNL and BNL were different

indicating all four libraries were different than one another. Thus, the larval effect was apparent and strong in both soil types and there were differences between the soil types as well. LIBSHUFF did not point out specific shifts in the community but PCA was useful to resolve this question, indicating the groups associated with the variations. Indeed, results obtained by PCA clearly indicated effects of larval presence in both soil types. TRFLP analysis supports the same conclusion by separating larvae/no larvae treatments within the Red Soil cluster. However, Kaufman et al. (2006) showed that different soil types did not affect the productivity of these habitats, even though nutrient regime was affected.

Shannon and Simpson diversity indices indicated that diversity in no larvae treatments (RNL and BNL) community was higher than that in the treatments with larvae (RL and BL). However, it should be noted that the diversity was not completely assessed, as indicated by the rarefaction curve generated in DOTUR. This finding limits the power of various methods used with regards to the sequence libraries. The diversity revealed here was much higher than that in a single cultivation study published recently (Huang et al. 2006), suggesting that cultivation methods for *An. gambiae* habitats will under-represent the bacterial diversity present. It is important to realize that interpretation is not always simple, and one must bear in mind that a particular analysis provides only a minimum estimate of the diversity (Fuhrman 2002). Methods affect results and different methods used in microbial ecology and bacterial diversity studies should be used to test or examine the same hypothesis. PCA supported by ANOVA on factor scores is rarely used, but looks to be a powerful adjunct tool for analysis of data from sequence libraries,

both when a taxonomic classification is achieved (here, RDP II) or when it is not but OTUs can be reliably derived using a specified similarity value (here, DOTUR).

Sequences that were unclassified by RDP II, when subjected to Blast search revealed that most of the sequences matched close to Cyanobacteria (*Cylindrospermum*, *Tolypothrix*), or chloroplasts of unicellular green algae (*Chlamydomonas*, *Dunaliella*), diatoms (*Nitzchia*, *Skeletonema*), colonial algae (*Scenedesmus*), and *Euglena*. These sequences detected in the treatments without larvae in both soil types support the conclusion of few others studies that these primary producers might be important food for the *Anopheles* mosquito larvae. Absences of these sequences in the treatments with larvae also support findings regarding changes in algal abundance and biomass with larvae present (Kaufman et al. 2006, Gimnig et al. 2001). Also it has been noted in many aquatic ecology studies that predator mediated top down effects shift bacterial communities and make bacteria less useful to the predators as food items (Jurgens and Gude 1994). Thus, changes in the bacterial community in current study might reflect a similar top down effect of the larvae grazing on acceptable forms and shifting community to unacceptable forms (Kaufman et al. 2000, 2002, Xu et al. in preparation). The changes in bacterial community composition in treatments with larvae might also be an indirect effect of the absence of algae. Algal and bacterial productivity are usually tightly linked in freshwater systems and absence of algae would limit bacterial growth (Cole et al. 1988) and planktonic community structure can determine bacterial production (Pace et al. 1990). Increase in particular groups of bacteria may also be due the removal of grazing pressures of other organisms feeding on bacteria. Grazing experiments revealed that *Pedobacter* was preferred as prey by the ciliate over *Brevundimonas* by a factor of four (Becks et al

2005). Thus for instance if ciliates are removed by *Anopheles* larvae, bacteria like *Pedobacter* might increase in numbers. Thus the concept of cascading trophic interactions (Carpenter et al. 1985) might also be applicable for these mosquito habitats. The shifts in the community might also be an effect of predation by other predators like protozoans and zooplanktons that might be present in these habitats. There is increase in number of reports on appearance of grazing resistant bacteria in marine and freshwater habitats aquatic ecosystems (Jurgens and Matz 2002, Langenheder and Jurgens 2001, Jurgens 1994). The importance of graze resistant bacteria increases with increasing grazing pressure exerted by protozoans, whereas decreases with increasing top down control of protozoans by zooplankton. This might reduce the productivity of planktonic systems through decrease in trophic transfer efficiencies and reduced regeneration of bacteria bound nutrients (Jurgens and Gude 1994).

It is essential to be aware of bottom up effects along with top down control of bacterial populations. Nutrient conditions alone can result in differences in the structure of bacterial community and under grazing pressure different graze resistant bacteria can develop under different nutrient conditions (Matz and Jurgens 2003). In their study Matz and Jurgens (2003) noticed that bacteria developed different survival mechanism, highly motile bacteria developed under C limitation and were dominated by filamentous forms in P limitation. In the current study, the underlying soil had effects on P concentrations, with Black Soil releasing higher amounts than the Red Soil (Kaufman et al. 2006). In any of the above cases where the habitat might become less productive for the *Anopheles* larvae possibly resulting in segregation of larval habitats in natural setting and may also explain oviposition selection by the female *Anopheles* mosquito. Further studies on these

interactions might be more useful on understanding the food web in these habitats, in particular, studies which examine nutrient depletion from soil and allow successive cohorts of larvae to establish, feed and grow.

REFERENCES

Becks L., Hilker F. M., Malchow H., Jurgens K., Arndt H. 2005. Experimental demonstration of chaos in a microbial food web. Nature 435: 1226-1229

Carpenter S. R., Kitchell J. F., Hodgson J. R. 1985 Cascading trophic interactions and lake productivity: Fish predation and herbivory can regulate lake ecosystems. BioScience 35(10): 634-638.

Cole, J. J., Findlay S., Pace M. L. 1988. Bacterial production in fresh and saltwater ecosystems: a cross overview. Marine Ecology Progress Series. 43: -10.

Cole, J. R., B. Chai, R. J. Farris, Q. Wang, S. A. Kulam, D. M. McGarrell, G. M. Garrity, and J. M. Tiedje. 2005. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Research. 33:294-6.

Cooper S. D., Diehl S., Kratz K., Sarnelle O. 1998. Implications of scale for patterns and processes in freshwater ecology. Australian Journal of Ecology 23:27-40.

Edillo F. E., Tripét F., Touré Y. T., Lanzaro G. C., Dolo G., Taylor C. E. 2006. Water quality and immatures of the M and S forms of *Anopheles gambiae* s.s. and *An. arabiensis* in a Malian village. Malaria Journal 5:35

Farrelly, V., F.A. Rainey, and E. Stackebrandt. 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Applied Environmental Microbiology. 61:2798-2801.

Fuhrman J. A., Griffith J. F., Schwalbach M. S. 2002. Prokaryotic and viral diversity patterns in marine plankton. Ecological Research. 17: 183–194

Gimnig J., Ombok M., Kamau L., Hawley W. 2001. Characteristics of larval Anopheline (Diptera: Culicidae) Habitats in Western Kenya. Journal of Medical Entomology 38: 282-288.

Gimnig, J.E., M. Ombok, S. Otieno, M. Kaufman, J. Vulule, & E. D. Walker. 2002. Density-dependent development of *Anopheles gambiae* larvae in artificial habitats. Journal of Medical Entomology 39: 162-172.

Giovannoni S. J., Britschgi T. B., Moyer C. L., and Field K. G. 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature 345: 60–63.

Hahn M. W. 2003. Isolation of strains belonging to the cosmopolitan *Polynucleobacter necessarius* cluster from freshwater habitats located in three climatic zones. Applied Environmental Microbiology 69: 5248–5254.

Hahn M.W., Pockl M., Wu Q. L. 2005. Low intraspecific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. Applied Environmental Microbiology 71: 4539–4547.

Heckmann K., Schmidt H. J. 1987. *Polynucleobacter necessaries* gen. nov., sp. nov., an obligately endosymbiotic bacterium living in the cytoplasm of *Euplotes aediculatus*. International Journal of Systematics and Bacteriology 37: 456–457.

Huang J., Miller R. R., Chen Shi-Cheng., Vulule J. M., and Walker E. D. 2006. Anopheles gambiae (Diptera: Culicidae) Oviposition in response to agarose media ad cultured bacterial volatiles. Journal of Medical Entomology. 43(3): 498-504.

Hugenholtz P., Goebel B. M., Pace N. R. 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. Journal of Bacteriology 180(18): 4765-4774.

Hughes J. B., Hellmann J. J, Ricketts T. H., Bohannan B. J. M. 2001. Counting the uncountable: Statistical approaches to estimating microbial diversity. Applied and Environmental Microbiology. 67. 4399–4406.

Jurgens K. 1994. Impact of *Daphnia* on planktonic microbial food webs – A review. Marine Microbial Food Webs. 8(1-2): 295-324.

Jurgens K., Gude H. 1994. The potential importance of grazing-resistant bacteria in planktonic systems. Marine ecology progress series **112**: 169-188.

Jurgens K., Matz C. 2002. Predation as a shaping force for the phenotypic and genotypic compostion of planktonic bacteria. Antonie van Leeuwenhoek 81: 413-434.

Kaufman, M. G., Walker E. D., Smith T. W., Merritt R. W., Klug M J. 1999. The effects of larval mosquitoes *Aedes triseriatus* and stemflow on microbial community dynamics in container habitats. Applied Environmental Microbiology. 65: 2661-2673.

Kaufman, M.G., E.D. Walker, D.A. Odelson, & M.J. Klug. 2000. Microbial community ecology and insect nutrition. American Entomology. 46: 173-184.

Kaufman, M.G., S.J. Bland, M.E. Worthen, E.D. Walker, & M.J. Klug. 2001. Bacterial and fungal contributions to growth of larval *Aedes triseriatus* (Diptera: Culicidae). Journal of Medical Entomology. 38: 711-719.

Kaufman, M.G., Goodfriend W., Kohler-Garrigan A., Walker E.D., & Klug M.J. 2002. Soluble nutrient effects on microbial communities and mosquito production in *Ochlerotatus triseriatus* (Say) habitats. Aquatic Microbial Ecology 29: 73-88.

Kaufman M. G., Walker E. D. 2006. Indirect effects of soluble nitrogen on growth of *Ochlerotatus triseriatus* larvae in container habitats. Journal of Medical Entomology. 43(4): 677–688.

Kaufman M. G., Wanja E., Maknojia S. R., Bayoh N. M., Vulule J. M., Walker E. D. 2006. The importance of algal biomass to the growth and development of *Anopeheles gambiae* larvae. Journal of Medical Entomology 43: 669-676.

Klappenbach J.A., Dunbar J.M., and Schmidt T. 2000. rRNA operon copy number reflects ecological strategies of bacteria. Applied Environmental Microbiology. 66:1328-1333.

Lane D. J., Pace B., Olsen G. J., Stahl D. A., Sogin M. L., Pace N. R. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proceedings of National Academy of Sciences. 82: 6955-6959.

Langenheder S., Jurgens K. 2001. Regulation of bacterial biomass and community structure by metazoan and protozoan predation. American Society of Limnology and Oceanography 46(1): 121-134.

Lilburn T. G., Schmidt T. M., Breznak J. A. 1999. Phylogenetic diversity of termite gut spirochaetes. Environmental Microbiology 1(4): 331-345.

Liu W.T., Marsh T.L., Cheng H, Forney L.J. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16s rRNA. Applied Environmental Microbiology. 63:4516-4522.

Ludwig J. A., Reynolds J. F. 1988. Statistical Ecology, A primer on methods and computing. Chapter 8. John Wiley & Sons, Inc. USA.

Magurran A. E. 2004. Measuring biological diversity. Blackwell Science Limited, 108 Cowley Road, Oxford, OX4 1JF.

Maidak B. L., C. J. R., Parker C. T., Garrity G. M., Larsen N., Li B., Lilburn T. G., McCaughey M. J., Olsen G. J., Overbeek R., Pramanik S., Schmidt T. M., Tiedje J. M., Woese C. R. 1999. A new version of the RDP (Ribosomal Database Project). Nucleic Acids Research 27: 171-173.

Mango N. A. R. 1999. Integrated soil fertility management in Siaya District, Kenya *In* Managing Africa's Soils: No.7

Marchesi J. R., Sato T., Weightman A. J., Martin T. A., Fry J. C., Hiom S. J., Wade W. G. 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Applied and Environmental Microbiology. 64: 795-799.

Marsh T. L. 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Current Opinion in Microbiology. 2:323-327

Marsh T. L., Saxman P., Cole J., Tiedje J. 2000. Terminal Restriction Fragment Length Polymorphism Analysis Program, a Web-Based Research Tool for Microbial Community Analysis. Applied and Environmental Microbiology. 66: 3616-3620

Matz C., Jurgens K. 2003. Interaction of nutrient limitation and protozoan grazing determines the phenotypic structure of a bacterial community. Microbial Ecology 45: 384-398.

May R. M. 1974. Biological populations with non overlapping generations: stable points, stable cycles, and chaos. Science. 186: 645-647.

Merritt, R.W., D.A. Craig, E.D. Walker, H.A. Vanderploeg, & R.S. Wotton. 1992. Interfacial feeding behavior and particle flow patterns of *Anopheles quadrimaculatus* larvae (Diptera: Culicidae). Journal of Insect Behavior 5: 741-761.

Merritt, R.W., Craig D.A., Wotton R.S., Walker E.D. 1996. Feeding behavior of aquatic insects: Case studies on black fly and mosquito larvae. Invertebrate Biology. 115: 206-217.

Mutuku, F. M., Bayoh M. N., Gimnig J. E., Vulule J. M., Kamau L., Walker E. D., Kabiru E. W., Hawley W. A. 2006. Pupal habitat productivity for *Anopheles gambiae* complex mosquitoes in a rural village in western Kenya. American Journal of Tropical Medicine and Hygiene. 74: 54-61.

Pace M. L, McManus G. B., Findlay S. E. G. 1990. Planktonic community structure determines the fate of bacterial production in a temperate lake. Limnology and Oceanography. 35(4): 795-808.

Pace N. R. 1997. A Molecular View of Microbial Diversity and the Biosphere. Science. 276(5313): 734 – 740.

Peck, G. W., Walton W. E. 2005. Effect of natural assemblages of larval foods on *Culex quinquefasciatus* and *Culex tarsalis* (Diptera: Culicidae) growth and whole body stoichiometry. Environmental Entomology: 34: 767-774.

Peck, G. W., Walton W. E. 2006. Effect of bacterial quality and density on growth and whole body stoichiometry of *Culex quinquefasciatus* and *Culex tarsalis* (Diptera: Culicidae). Journal of Medical Entomology. 43(1): 25-33.

Schloss P. D., Handelsman Jo. 2005. Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness. Applied and Environmental Microbiology. 71(3): 1501–1506.

Shivaji S., Chaturvedi P., Reddy G. S. N., Suresh K. 2005. Pedobacter himalayensis sp. nov., from the Hamta glacier located in the Himalayan mountain ranges of India. International Journal of Systematic and Evolutionary Microbiology 55: 1083–1088

Singleton, D. R., Furlong M. A., Rathbun S. L., Whitman W. B. 2001. Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. Applied and Environmental Microbiology. 67: 4374-6.

Steyn P. L., Segers P., Vancanneyt M., Sandra P., Kersters K., Joubert, J. J. 1998. Classification of heparinolytic bacteria into a new genus, *Pedobacter*, comprising four species: *Pedobacter heparinus* comb. nov., *Pedobacter piscium* comb. nov., *Pedobacter africanus* sp. nov. and *Pedobacter saltans* sp. nov. Proposal of the family Sphingobacteriaceae fam. nov. International Journal of Systematic Bacteriology 48: 165–177.

Strunk O. and Ludwig W. 1995. ARB- a software environment for sequence data. Department of Microbiology, Technical University of Munich, Germany.

Suzuki M. T., Rappe M. S., Haimberger Z. W., Winfield H., Adair N., Strobel J., Giovannoni S. J. 1997. Bacterial diversity among small-subunit rrna gene clones and cellular isolates from the same seawater sample. Applied and Environmental Microbiology. 63(3): 983–989.

Walker E. D., Olds E. J., Merritt R. W. 1988. Gut content analysis of mosquito larvae (Diptera: Culicidae) using DAPI stain and Epifluorescence microscopy. Journal of Medical Entomology. 25(6): 551-554.

Walker, E. D. and Merritt R.W. 1991. Behavior of *Aedes triseriatus* (Diptera: Culicidae) larvae. Journal of Medical Entomology. 28: 581-589.

Walker E. D., Merritt R. W. 1993. Bacterial enrichment in the surface microlayer of an *Anopheles quadrimaculatus* (Diptera: Culicidae) larval habitat. Journal of Medical Entomology. 30.1050-1052.

Wu Q. L., Hahn M. W. 2006. Differences in structure and dynamics of Polynucleobacter communities in a temperate and a subtropical lake, revealed at three phylogenetic levels. FEMS Microbiology. Ecol. 57: 67–79

Zwart G., Crump B. C., Kamst-van Agterveld M. P., Hagen F., Suk-Kyun Han S-K. 2002. Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquatic Microbial Ecology. Vol. 28: 141–155

Zoetendal E. G, Collier C. T., Koike S., Mackie R. I., Gaskins H. R. 2004. Molecular ecological analysis of the gastrointestinal microbiota: A review. Ameican Society for Nutritional Sciences 134: 465-472.

CHAPTER IV

Bacterial diversity of the surface microlayer of *Anopheles gambiae* and *Anopheles funestus* habitats using 16S rDNA sequence library construction

ABSTRACT

A descriptive analysis of bacterial diversity of *Anopheles gambiae* and *Anopheles funestus* habitats was carried out using 16R rDNA sequencing and TRFLP analysis. A total of 644 rDNA sequences from *An. gambiae* habitats and 642 sequences from *An. funestus* habitats obtained were classified using Ribosomal Database project II. Both communities were highly diverse with few dominant and many rare taxa. LIBSHUFF analysis revealed that these communities were statistically different but principal component analysis and TRFLP analysis did not show any significant clustering of these habitats. There were a large number of unclassified bacteria and a few other groups like Actinomycetales and Rhodobacterales which were found in high frequencies in *An. gambiae* habitats. Alteromonadales was found in high frequency in *An. funestus* habitats. There was no evidence that would support habitat segregation based on bacterial community structure within this habitat that might play a role in oviposition selection by gravid females.

INTRODUCTION

Many of the *Anopheles* species in Africa are highly efficient vectors of human malaria, in particular species of the *Anopheles gambiae* and the *An. funestus* complex. Within the gambiae complex, *An. gambiae and An. arabiensis* have the widest distribution in sub-Saharan Africa (Coluzzi 1984). *An. gambiae* and *An. arabiensis* breed in fresh water and are associated with small habitats often created by man or animal activity, such as foot or hoof prints, burrow pits, roadside puddles formed by tire tracks, and irrigation ditches. Their habitats are turbid and persist for short periods and lack aquatic vegetation. *An. funestus* on the other hand is distinguished from the gambiae complex by breeding in large, semi-permanent bodies of water, characterized by emergent vegetation, such as swamps, river edges and ditches.

Larvae of the *Anopheles* species reviewed above have some obvious and characteristic habitat features and qualities, but the nutrient and microbial facets of these habitats which contribute to primary and secondary (including mosquito) production and to habitat differences among these *Anopheles* species are poorly known. The adult mosquito density depends on the number and collective productivity of the larval habitats. Even though these habitats are the source of these competent vectors, little is known about the larval biology of these important insects, in particular, what forms the basis for production of insect biomass in them. According to the study conducted by Gimnig et al. 2002 the most likely important food source for larval *An. gambiae* seem to be algae, which were significantly reduced in presence of larvae as measured by chlorophyll *a* in the surface water samples and by counts of algae in sedimentation

chambers, compared to with when larvae were absent. Other experiments support Gimnig et al. study and report that presence of algae does improve the growth and development of the larvae with shorter developmental time and greater survival as compared to when algae were absent, a condition obtained by shading the habitats (Kaufman et al. 2006). But this ignores algal turnover rates, larval consumption and other nutritional sources (i.e.; bacteria and non cellular organics in the surface layer). Algae probably play a key role as food resources in *An. gambiae* habitats and bacteria may likely form a secondary food source. Studies on larval gut analysis reveal that bacteria form the bulk of the food bolus, along with organic particulates, algae, and small invertebrate parts (Walker et al. 1988). Of interest is that larval mosquito feeding on bacteria in their habitats has been estimated to account for production of a large fraction of the insect biomass (Kaufman et al. 2001), but simultaneously, larval mosquito feeding reduces microbial abundance, alters the microbial community qualitatively and quantitatively, and diminishes the quality of the microbes as food (Kaufman et al. 2000, 2002; Xu et al. in preparation).

Further, microbial communities in larval mosquito habitats have been hypothesized to influence oviposition site selection by mosquitoes owing to volatile organic compounds emitted from these communities, and owing to microbially-derived flavors in water (Trexler et al. 2003). Location and selection of an oviposition site is an essential part of the life history of all mosquito species. The location and selection of an ovipostional site involves visual, olfactory, and tactile responses. Intensive field studies have shown that mosquitoes are quite discriminating in selecting sites for egg deposition and considerable evidence points to this site discriminating larval distribution (Bentley and Day 1989). Oviposition site selection is the net result of the interaction of a complex

array of both chemical and physical factors. These attracting substances may, in some cases be pheromones produced by the larvae. Studies have shown that concurrent presence of *A. gambiae* larvae reduced oviposition, while turbid water from natural breeding sites increased oviposition selectivity. Thus other possible attractants such as microbial metabolites must also be considered keeping in mind that mosquito biting on the human body is also mediated by microbes. *An. gambiae*, a human biter, prefers to bite human feet and is attracted to odors emitted by bacteria from feet, such as isovaleric acid (Braks et al 1999).

Notably, *An. gambiae* and *An. funestus* habitats are segregated and distinct in Kenya (Gimnig et al. 2001). The bacterial communities are postulated to be distinct between these habitat types. These differences may be crucial in differential oviposition site selection by females of these two species, and the differences may account for habitat production for adult mosquitoes. Since very little information exists on the bacterial community, it is essential to study it and also to determine what comprises the bacterial community within these habitats. By conducting a comparative analysis one can also determine what makes one habitat more favorable for one species (*An. gambiae* and *An. arabiensis*) and the other habitat for other species (*An. funestus*).

METHODS AND MATERIALS

Sampling the Sites: Samples were obtained from habitats characteristic of *An. gambiae* and *An. funestus* and were identical to sites used in previous studies (Gimnig et al. 2001, Walker unpublished) from the Asembo District near Kisumu, western Kenya. Five habitats of *An. gambiae* and 5 habitats of *An. funestus* were sampled and were designated as G1, G3, G6, G7, G13 and F1 through F5, respectively. The surface microlayer from each habitat was sampled (25ml) using a needle and syringe as previously described (Gimnig et al. 2002). These samples were returned to the laboratory and preserved at -20° C until processed. The surface microlayer sample was divided and preserved for different analysis. A 3 ml aliquot of surface microlayer sample from *An. gambiae* and *An. funestus* habitats were separated and preserved with Lugol's solution for algal counts, 5 ml filtered for chlorophyll *a* analysis, and 3 ml for direct bacterial counts. The remaining 14 ml was used for nucleic acid extraction.

Bacterial and Algal density: Direct bacterial counts of 3 ml formalin preserved samples were performed using 4'6-diamidino-2-phenylindole (DAPI) stain and epifluoresence microscopy. Algae were counted using settling chambers and inverted microscopy. Counts were converted with standard formulae to numbers of cells per ml of original sample.

Chlorophyll *a* Analysis: Surface microlayer samples for chlorophyll *a* analyses were collected from the central region of each habitat using a syringe and 16 gauge needle.

Five ml of the surface microlayer were filtered on site through glass fiber filters (nominal pore size 1 μM). Filters were then kept frozen until analysis. Chlorophyll *a*, a measure of algal biomass in *u*g per ml of sample, was estimated by fluorometric analysis (Welschmeyer, 1994). Chlorophyll *a* was extracted overnight in 95% ethanol and the samples were read against a 95% ethanol blank. The chlorophyll *a* was calculated in μg/L.

Nucleic Acid Extraction, Amplification, Cloning and Sequencing Surface microlayer samples (14 ml) from each of the above mentioned habitats were centrifuged at 6000 rpm for 30 minutes and the pellet was retained for nucleic acid (DNA) extraction. The DNA was extracted using Ultraclean Soil DNA kits (Cat. # 12800-50, MO BIO Laboratories Inc., Carlsbad, California) following the manufacturer's instructions. The presence of DNA was confirmed on 1% agarose gels in TBE buffer. Extractions from each sample were stored at -20° C. An approximately 1.3 kb region of a consensus 16S rRNA gene of bacteria was amplified by PCR amplification using forward primer 63f 5' -CAG GCC TAA CAC ATG CAA GTC- 3' and reverse primer 1387r 5' – GGG CGG WGT GTA CAA GGC- 3' (Marchesi et al. 1998). The PCR reactions conditions were carried out as per the Taq DNA polymerase instruction manual (New England Biolabs). PCR conditions consisted of initial denaturing at 94°C for 2 min followed by 25 cycles of denaturing at 94°C for 45 sec, annealing at 55°C for 30 sec, extension at 72°C for 1.3 min, and final extension at 72°C for 7min. The resultant PCR products were purified using low melting agarose gel electrophoresis, by cutting out the bands and purifying it by sodium acetate and ethanol precipitation. Purified 16S rDNA fragments were cloned

into the pGEM-T Easy vector (Promega, Madison, Wisconsin) and transformed clones were picked and purified. The plasmid DNA from transformed clones was extracted using Wizard Plus SV Miniprep kit (Promega). After confirmation that the plasmids contained an insert of the expected size, these plasmids (or clones for some samples) were then subjected to high throughput sequencing using dideoxy dye terminator chemistry, at the Genomic Technology Support Facility, Michigan State University. The 519R 5'- G(AT)ATTACCGCGGC(GT)GCTG- 3' sequencing primer (Lane et al 1985) was used to obtain partial 16S rDNA sequences.

Classification of sequences Sequence data from each of the 10 clone libraries (5 *An. gambiae* and 5 *An. funestus* habitats) was examined for possible chimeras using the CHECK_CHIMERA program offered through Ribosomal Database Project II (RDP II; Release 8.1, May 18, 2001 (Maidak et al. 1999) following procedures outlined in Lilburn et al. (1999). Possible chimeras along with sequences with fewer than 250 base pair in length were excluded from any further analysis. The 16S rDNA sequences were then classified to named, operational taxonomic units (OTUs) using the Ribosomal Database Project II (RDP II; Release 9.38, April 03, 2006, http://rdp.cme.msu.edu). The RDP II contains 210,976 aligned 16S rRNA sequences as of April 03, 2006, and this database formed the reference library for the classifications here. The hierarchy model used by the naïve Bayesian rRNA classifier in RDP II comes from the bacterial classification proposed in release 6.0 of the Bergey's Manual of Systematic Bacteriology (http://bergeysoutline.com). The classifier calculates the joint probability of finding eight base subsequences ("words") in the query. When a query sequence is submitted, the joint

probability of observing all the words in the query can be calculated separately for each genus from the training set probability values. Using the naïve Bayesian assumption, the query is most likely a member of the genus with the highest probability, given the limitations of the available database (Cole et al. 2005). This classification process puts a name on each sequence.

Sequences from all 10 libraries were aligned with existing bacterial sequences in ARB 16S rDNA database http://www.arb-home.de and phylogenetic trees were constructed using the ARB software package (Strunk and Ludwig 1997, Lilburn et al. 1999). The RDP II classification was used as a guide for initial alignment since RDP places the sequences to genus level which was used select reference sequences for alignments. A mask was generated by the filtering method in ARB to produce valid columns for comparison of sequences. Sequences with short lengths were excluded so as to generate at least 200 valid columns. Distance matrices were generated using the above filter and Jukes-cantor correction.

Statistical comparisons Sequence libraries were compared between clone libraries using the LIBSHUFF program, (http://www.arches.uga.edu/~whitman/libshuff.html, Singleton et al. 2001). The program estimates differences between homologous coverage curves $C_X(D)$, and heterologous coverage curves $C_{XY}(D)$, by a Cramér-von Mises-type statistic, $\Delta C_{XY} = \sum(C_X - C_{XY})^2$. The input for these two-way comparisons was the distance matrices of aligned sequences generated in ARB. The distance matrices derived from nucleotide sequence differences generated in ARB were also used as input for diversity analysis using the DOTUR program (Distance based OTU and Richness determination;

http://www.plantpath.wisc.edu/fac/joh/dotur.html). A 97% similarity value was used to determine operational taxonomic units (OTUs) (Schloss and Handelsman 2005). Using these new OTU assignments, which were un-named in contrast to the classification in RDP II, DOTUR constructed rarefaction curves for sampling intensity, richness estimators, and diversity indices. To address relative evenness of the bacterial communities reflected by sequence libraries, rank abundance graphs were generated where the abundance (i.e., frequency) of each OTU was plotted on a logarithmic scale against the rank from most abundant to least abundant (Magurran 2004).

Principal Component Analysis (PCA) PCA is a technique for simplifying a dataset where the data are transformed to a new coordinate system such that the greatest variance by any projection of the lies on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA can be used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance. PCA analysis was invoked for the sequence dataset with RDP classification as well as OTUs obtained by DOTUR analysis. Factorial ANOVA was performed on the PCA scores and the sequences that explained most of the variance were identified through the PCA loading values.

Terminal Restriction Fragment Length Polymorphism Analysis (TRFLP) The DNA extraction procedure and the PCR condition were the same as above with the exception that the forward 63f primer was labeled with 6-FAM (Integrated DNA Technologies, Coralville IA) at the 5' end. Three 100 µl PCR reactions were performed for each sample to assure maximum coverage of the diversity of 16S rDNA present in the sample. These

reactions were combined and purified using Qiaquick PCR purification kit (Qiagen Valencia, CA). To generate TRFLP fragments, restriction enzyme HhaI and MspI (New England Biolabs Inc.) were selected since they offer decent number of restriction fragment and discrimination between different species (Marsh 1999).

The digestions were carried out in a two 10 µl aliquots (Liu et al. 1997, Marsh et al. 2000). 400-600 ng of purified PCR product was digested for 3 hours with either HhaI or MspI at 37° C. The reaction mixture contained approximately 400 ng (1-5 µl) of DNA, 2 µl of 10x restriction enzyme buffer, 1.5 µl of the one of the above restriction enzyme. The reaction mixture was brought to 10 µl with nuclease free water. After 3 hrs of digestion, the reactions were stopped at 60° C for 20 minutes. The T-RF lengths from the digest were determined on an automated DNA sequencer (ABI PRISM 310, Foster City, California) in GeneScan mode. The fragments were separated by electrophoresis using capillary electrophoresis system. The fragment sizes were determined with ABI Genescan Analysis Software (Applied Biosystems) and the alignment of the community profiles were done in Genotyper software (Genotyper 2.5, Applied Biosystems). The peaks ranging between 50-950 bp were used for further analysis. Cluster analysis of TRFLP profiles was done using PAUP (http://paup.csit.fsu.edu/).

Nucleotide sequence accession numbers The sequences obtained during this study were deposited to GenBank under the accession numbers EF149017 to EF149660 for clones obtained from *An. gambiae* habitats and EF149661 to EF150302 for clones from *An. funestus* habitats.

FIGURES IN THIS CHAPTER ARE PRESENTED IN COLOR.

RESULTS

Microbial and Chemical Analysis The total number of bacteria in the surface microlayer from *An. gambiae* habitats ranged from 0.76 x 10^6/ml to 2.3 x 10^6/ml and in and *An. funestus* habitats the bacterial counts ranged from 1.6 x 10^6/ml to 37 x 10^6/ml. The algal counts in *An. gambiae* ranged from 3.5 x 10^3/ml to 8.6 x 10^3/ml and in *An. funestus* habitats the counts ranged from1.8 x 10^3/ml to 15 x 10^3/ml. Mean ± SE chlorophyll *a* in the surface microlayer of *An. gambiae* and *An. funestus* habitats was 43.89 ± 15.35 µg/L and 42.7 ± 12.52 µg/L.

Classification of sequences: A total of 1362 sequences were obtained for *An. gambiae* and non-gambiae (*An. funestus*) habitats. Individually 673 sequences from *An. gambiae* and 689 sequences from *An. funestus* habitats were obtained. Sequences with short lengths (< 250bp) and chimeras were removed from each library reducing the library to 644 *An. gambiae* and 642 *An. funestus* sequences which were then classified using RDP II.

Classifications based on the RDP II for *An. gambiae* and *An. funestus* sequences libraries are summarized in Tables 4.1 and 4.2 respectively. The classification scheme assigned the sequences from both libraries to 8 phyla, namely, Actinobacteria, Bacteroidetes, Firmicutes, Nitrospira, Cyanobacteria, Proteobacteria, Verrucomicrobia and Acidobacteria. Sequences belonging to Nitrospira, Verrucomicrobia and Acidobacteria were absent in *An. gambiae* habitats. The most dominant Phylum in both libraries was Proteobacteria (72.8% in *An. gambiae* habitats and 70.9 % in *An. funestus*

habitats). The second most common Phylum in *An. gambiae* library was Actinobacteria (11.8%) and Bacteriodetes (9.8%). Within Proteobacteria sequences belonging to class Alphaproteobacteria (22.7%), Betaproteobacteria (27.3%) and Gammaproteobacteria (21.3%) were present in almost equal frequencies. Rhodobacterales and Sphingomonadales were most abundant in class Alphaproteobacteria, Burkholderiales were abundant in class Betaproteobacteria, and within Gammaproteobacteria, Pseudomonadales and Alteromonadales were abundant. The second most abundant phyla in *An. funestus* library were unclassified bacteria (12.6%), Cyanobacteria (7.16%) and Bacteriodetes (6.38%). Within Proteobacteria sequences belonging to class Gammaproteobacteria (38.8%) and Betaproteobacteria (23.4%) were most abundant.

In *An. gambiae* library within class Gammaproteobacteria genus *Acinetobacter* (Alteromonadaceae) and *Alishwanella* (Alteromonadaceae) were most common. Within Betaproteobacteria genus *Rhodobacter* (Rhodobacteraceae) and *Polynucleobacter* (Burkholderiaceae) was most common. Within Actniobacteria the sequences remained unclassified at the family level (Microbacteriaceae). In An. funestus library within class Gammaproteobacteria genus *Alishwanella* (Alteromondaceae) was dominant and Cellvibrio was common in order Pseudomonadales.

The most evident difference between *An. gambiae* and *An. funestus* libraries was as follows. Higher frequencies of clones around 11.65% from order Actinomycetales (Actinobacteria) and 13.2% Rhodobacterales (Alphaproteobacteria) in *An. gambiae* library as compared to 2.17% and 1.55% of the same in *An. funestus* library. On the other hand, in *An. funestus* library Alteromonadales (Gammaproteobacteria) were observed with high frequency, around 16.3% while only 4.97% in *An. gambiae* library.

Table 4.1. Bacterial community composition of surface microlayer from *An. gambiae* habitats based on 16S rDNA sequences classification with RDP II. The numbers of sequences observed per taxon are in parenthesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (76)	Actinobacteria (76)	Actinomycetales (75)	Microbacteriaceae (47)	Unclassified Microbacteriaceae(47)
			Unclassified Actinomycetales (28)	
		Unclassified Actinobacteria (1)		
Bacteroidetes (63)	Flavobacteria (32)	Flavobacteriales (32)	Flavobacteriaceae (32)	Flavobacterium (6)
				Riemerella (1)
				Unclassified Flavobacteriaceae (25)
	Sphingobacteria (20)	Sphingobacteriales (20)	Sphigobacteriaceae (9)	Pedobacter(2)
				Unclassified Sphingobacteriaceae (7)
			Flexibacteraceae (1)	Unclassified Flexibacteraceae (1)
			Crenotrichaceae (3)	Chitinophaga (3)
			Unclassified Sphingobacteriales (7)	
	Bacteroidetes (4)	Bacteroidales (4)	Prevotellaceae (4)	Prevotella (4)
	Unclassified Bacteroidetes (7)			
Cyanobacteria (6)	Cyanobacteria (6)	subsection Subsection 4 (1)	Family 4.1 (1)	Anabena (1)
		Unclassified Cyanobacteria (5)		
Firmicutes (4)	Bacilli (1)	Bacillales (1)	Bacillaceae (1)	Exiguobacterium (1)
	Clostridia (3)	Clostridiales (3)	Clostridiaceae (2)	Acetivibrio (2)
			Peptococcaceae (1)	Unclassified Peptococcaceae (1)
Proteobacteria (469)	Alphaproteobacteria (146)	Rhodobacterales (85)	Rhodobacteraceae (85)	Rhodobacter (20)
				Unclassified Rhodobacteraceae (65)
		Sphingomonadales (36)	Sphingomonadaceae (36)	Porphyrobacter (4)
				Novosphingobium (18)
				Unclassified Sphingomonadaceae (14)
		Rhizobiales (14)	Beijerinckiaceae (1)	Unclassified Beijerinckiaceae (1)
			Unclassified Rhizobiales (13)	
		Rhodospirillales (3)	Rhodospirillaceae (3)	Magnetospirillum (3)
		Caulobacterales (1)	Caulobacteraceae (1)	Caulobacter (1)

Phylum	Class	Order	Family	Genus
		Unclassified Alphaproteobacteria (7)		
	Betaproteobacteria (176)	Burkholderiales (135)	Comamonadaceae (48)	Polaromonas (2)
				Curvibacter (4)
				Comamonas (2)
				Hydrogenophaga (11)
				Acidovorax (1)
				Unclassified Comamonadaceae (28)
			Oxalobacteraceae (2)	Unclassified Oxalobacteraceae (2)
			Alcaligenaceae (1)	Unclassified Alcaligenaceae (1)
			Burkholderiaceae (63)	Polynucleobacter (61)
				Unclassified Burkholderiaceae (2)
			Incertae sedis 5 (14)	Roseateles (1)
				Leptothrix (1)
				Aquabacterium (3)
				Ideonella (2)
				Unclassified Incertae sedis 5 (7)
			Unclassified Burkholderiales (7)	
		Neisseriales (15)	Neisseriaceae (15)	Vogesella (3)
				Chitinibacter (1)
				Aquaspirillum (8)
				Unclassified Neisseriaceae (3)
		Rhodocyclales (14)	Rhodocyclaceae (14)	Dechoromonas (1)
				Unclassified Rhodocyclaceae (13)
		Unclassified Betaproteobacteria (12)		
	Gammaproteobacteria (137)	Xanthomonadales (5)	Xanthomonadaceae (5)	Pseudoxanthomonas (2)
				Unclassified Xanthomonadaceae (3)
		Aeromonadales (11)	Aeromonadaceae (11)	Aeromonas (11)
		Pseudomonanadales (63)	Pseudomonadaceae (21)	Pseudomonas (6)
				Cellvibrio (2)
				Unclassified Pseudomonadaceae (13)
			Moraxellaceae (42)	Acinetobacter (42)
		Alteromonadales (32)	Incertae sedis 7 (31)	Alishewanella (31)
			Shewanellaceae (1)	Shewanella (1)

Phylum	Class	Order	Family	Genus
		Chromatiales (3)	Chromatiaceae (3)	Rheinheimera (3)
		Enterobacteriales (4)	Enterobacteriaceae (4)	Escherichia (1)
				Klebsiella (1)
				Unclassified Enterobacteriaceae (2)
		Unclassified Gammaproteobacteria (19)		
	Deltaproteobacteria (1)	Desulfobacterales (1)	Desulfobacteraceae (1)	Unclassified Desulfobacteraceae (1)
	Epsilonbacteria (6)	Campylobacterales (6)	Campylobacteraceae (6)	Arcobacter (6)
	Unclassified Proteobacteria (3)			
Unclassified Bacteria (26)				
Total Bacteria	644			

Table 4.2 Bacterial community composition of surface microlayer from *An. funestus* habitats based on 16S rDNA sequences classification with RDP II. The numbers of sequences observed per taxon are in parenthesis.

Phylum	Class	Order	Family	Genus
Actinobacteria (15)	Actinobacteria (15)	Actinomycetales (14)	Microbacteriaceae (10)	Unclassified Microbacteriaceae (10)
			Unclassified Actinomycetales (4)	
		Unclassified Actinobacteria (1)		
Bacteroidetes (41)	Flavobacteria (15)	Flavobacteriales (15)	Flavobacteriaceae (14)	Flavobacterium (6)
				Unclassified Flavobacteriaceae (8)
			Unclassified Flavobacteriales (1)	
	Sphingobacteria (3)	Sphingobacteriales (3)	Flexibacteraceae (3)	Flectobacillus (2)
				Unclassified Flexibacteraceae (1)
	Bacteroidetes (3)	Bacteroidales (3)	Porphyromonadaceae (1)	Unclassified Porphyromonadaceae (1)
	Unclassified Bacteroidetes (20)		Unclassified Bacteroidales (2)	
Cyanobacteria (46)	Cyanobacteria (46)	subsection Subsection 3 (1)	Family 3.1 (1)	Oscillatoria (1)
		subsection Subsection 4 (2)	Family 4.1 (2)	Nodularia (1)
		Unclassified Cyanobacteria (43)		Unclassified Family 4.1 (1)
Firmicutes (1)	Bacilli (1)	Bacillales (1)	Bacillaceae (1)	Unclassified Bacillaceae (1)
Nitrospira (1)	Nitrospira (1)	Nitospirales (1)	Nitrospiraceae (1)	Nitrospira (1)
Proteobacteria (455)	Alphaproteobacteria (50)	Rhodobacterales (10)	Rhodobacteraceae (10)	Rhodobacter (3)
				Unclassified Rhodobacteraceae (7)
		Sphingomonadales (14)	Sphingomonadaceae (14)	Sphingomonas (1)
				Porphyrobacter (1)
				Novosphingobium (3)
				Unclassified Sphingomonadaceae (9)
		Rhizobiales (7)	Rhizobiaceae (2)	Unclassified Rhizobiaceae (2)
			Hyphomicrobiaceae (1)	Devosia (1)

Phylum	Class	Order	Family	Genus
			Unclassified Rhizobiales (4)	
		Rhodospirillales (8)	Rhodospirillaceae (8)	Azospirillum (1)
				Unclassified Rhodospirillaceae (7)
		Unclassified Alphaproteobacteria (11)		
	Betaproteobacteria (150)	Methylophilales (1)	Methylophilaceae (1)	Methylophilus (1)
		Burkholderiales (130)	Comamonadaceae (34)	Rhodoferax (5)
				Curvibacter (7)
				Hydrogenophaga (7)
				Acidovorax (2)
				Unclassified Comamonadaceae (13)
			Alcaligenaceae (1)	Unclassified Alcaligenaceae (1)
			Burkholderiaceae (16)	Limnobacter (2)
				Polynucleobacter (10)
				Wautersia (1)
				Unclassified Burkholderiaceae (3)
			Incertae sedis 5 (51)	Roseateles (5)
				Leptothrix (1)
				Aquabacterium (3)
				Ideonella (1)
				Unclassified Incertae sedis 5 (41)
			Unclassified Burkholderiales (28)	
		Neisseriales (1)	Neisseriaceae (1)	Aquaspirillum (1)
		Rhodocyclales (12)	Rhodocyclaceae (12)	Unclassified Rhodocyclaceae (12)
		Unclassified Betaproteobacteria (6)		
	Gammaproteobacteria(249)	Xanthomonadales (5)	Xanthomonadaceae (5)	Nevskia (1)
				Xanthomonas (1)
				Unclassified Xanthomonadaceae (3)
		Methylococcales (1)	Methylococcaceae (1)	Methylosarcina (1)
		Aeromonadales (7)	Aeromonadaceae (7)	Aeromonas (7)
		Pseudomonanadales (62)	Pseudomonanadaceae (33)	Pseudomonas (3)
				Alkanindiges (7)
				Cellvibrio (22)
				Unclassified Pseudomonadaceae (8)

Phylum	Class	Order	Family	Genus
			Unclassified Pseudomonadales (1)	
			Moraxellaceae (28)	Acinetobacter (20)
				Unclassified Moraxellaceae (1)
		Alteromonadales (105)	Incertae sedis 7 (105)	Alishewanella (105)
		Enterobacteriales (4)	Enterobacteriaceae (4)	Enterobacter (1)
				Unclassified Enterobacteriaceae (3)
		Unclassified Gammaproteobacteria (65)		
	Deltaproteobacteria (1)	Bdellovibrionales (1)	Bdellovibrionaceae (1)	Bdellovibrio (1)
	Unclassified Proteobacteria (5)			
Acidobacteria (1)	Acidobacteria (1)	Acidobacteriales (1)	Acidobacteriaceae (1)	Acidobacterium (1)
Verrucomicrobia (1)	Verrucomicrobia (1)	Verrucomicrobiales (1)	Unclassified Verrucomicrobiales (1)	
Unclassified Bacteria (81)				
Total Bacteria	642			

Community Comparison To determine if the *An. gambiae* and *An. funestus* libraries were different statistically distance matrices obtained from ARB were compared using LIBSHUFF and it revealed that both communities were significantly different from each other ($\Delta C_{F/G}$= 0.242, p= 0.001; $\Delta C_{G/F}$= 0.610, p= 0.001). The comparisons indicate that the communities differed greatly at high levels of relatedness but shared many deep phylogenetic taxa at low levels of relatedness (Fig 4.1). Sequences were classified into Operational Taxonomic Units (OTUs) and diversity indices were determined using DOTUR. The Simpson diversity index (1/D) for *An. gambiae* library determined by DOTUR was 63.51 and that for *An. funestus* library was 53.17. Note that the value of the index increases with increasing diversity. The Shannon diversity index was 4.57 for *An. gambiae* library and 4.8 for *An. funestus* library. The OTUs obtained by DOTUR analysis were used to plot rank/abundance graph where the relative abundance was plotted on logarithmic scale against the rank of each OTU (97 % similarity), ranked from most abundant to the least abundant (Fig 4.2). The rarefaction curves indicated that the diversity estimates obtained above might be an underestimation of the total diversity, as the curves do not reach an asymptote given the number of sequences in each clone libraries (Fig 4.3).

Principal Components Analysis PCA was invoked to identify a pattern and to see how different libraries from *An. gambiae* and *An. funestus* habitats clustered. The OTUs obtained from DOTUR using a 0.03 distance (97% similarity) criterion was used to perform PCA analysis (Fig 4.4 A). Similar analysis was performed on RDP II classification data at the order level for comparison (Fig 4.4 B). Analysis of variance

performed on PCA scores for OTU data (PC I: F= 0.72, P= 0.4197 and PC 2: F= 0.66, P= 0.4401) was not significant for libraries from *An. gambiae* and *An. funestus* habitats. Similarly ANOVA on PCA scores for RDP classification data (PC I: F= 2.2, P= 0.1756 and PC 2:F= 1.37, P= 0.2749) was not significant for *An. gambiae* and *An. funestus* libraries.

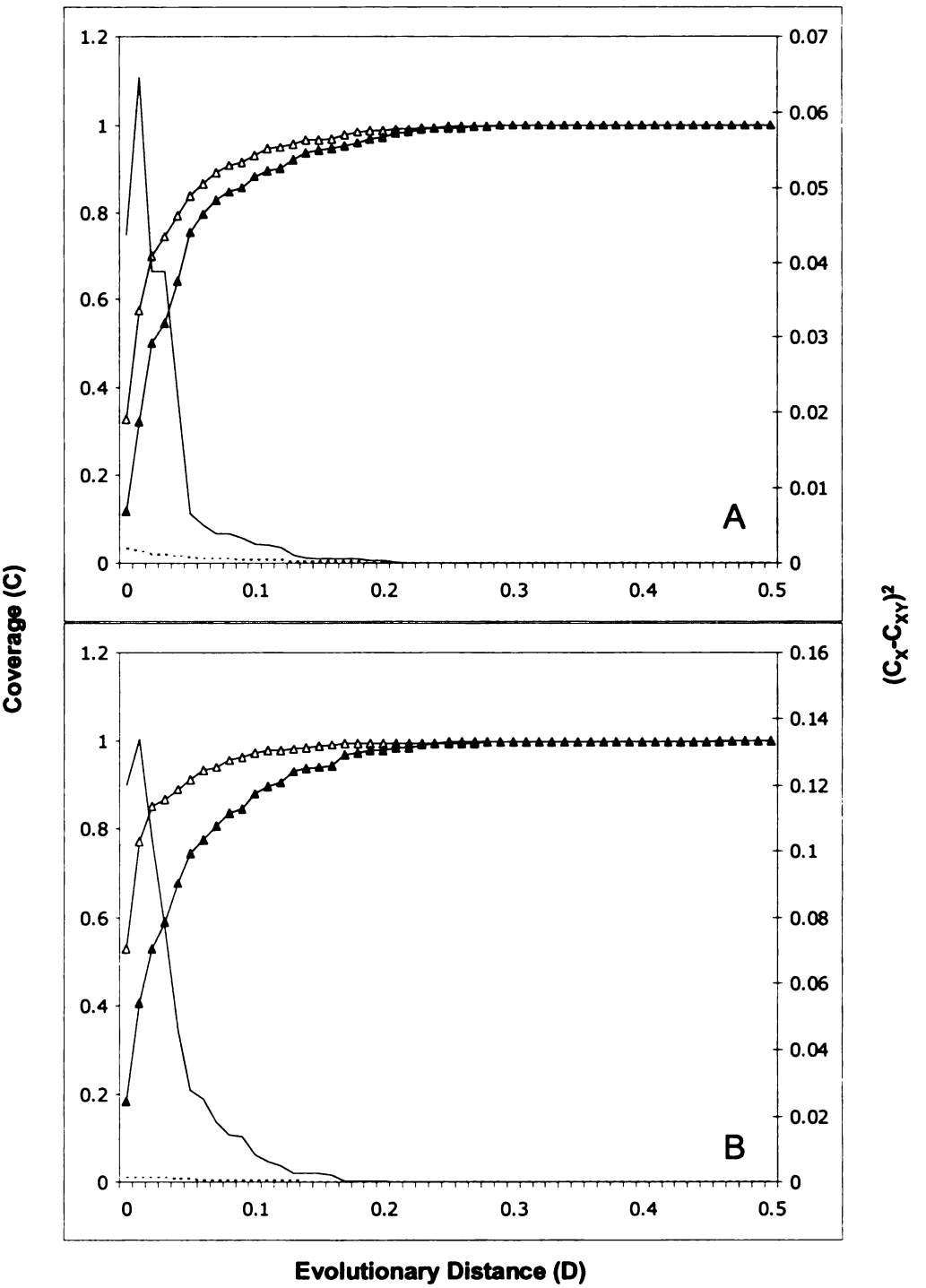


Figure 4.1. Comparison of bacterial 16S rDNA sequence libraries of *An. funestus* habitats to *An. gambiae* using LIBSHUFF. Homologous (open triangles) and heterologous (solid triangles) coverage curves are shown. Solid lines indicate values of $(C_F-C_{F/G})^2$ (panel A) or of $(C_G-C_{G/F})^2$ (panel B) at each value of evolutionary distance (D). Broken lines indicate the 950th value (or p=0.05) of corresponding $(C_F-C_{F/G})^2$or $(C_G-C_{G/F})^2$ for the randomized samples.

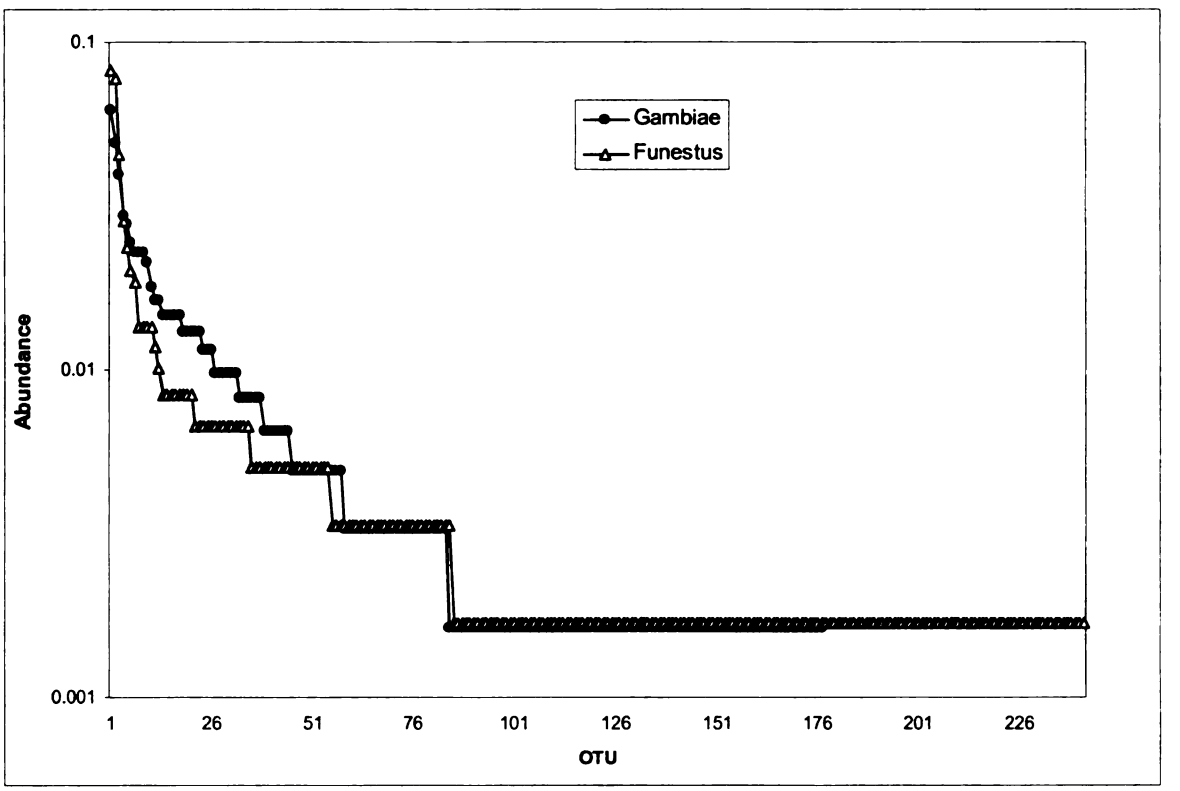
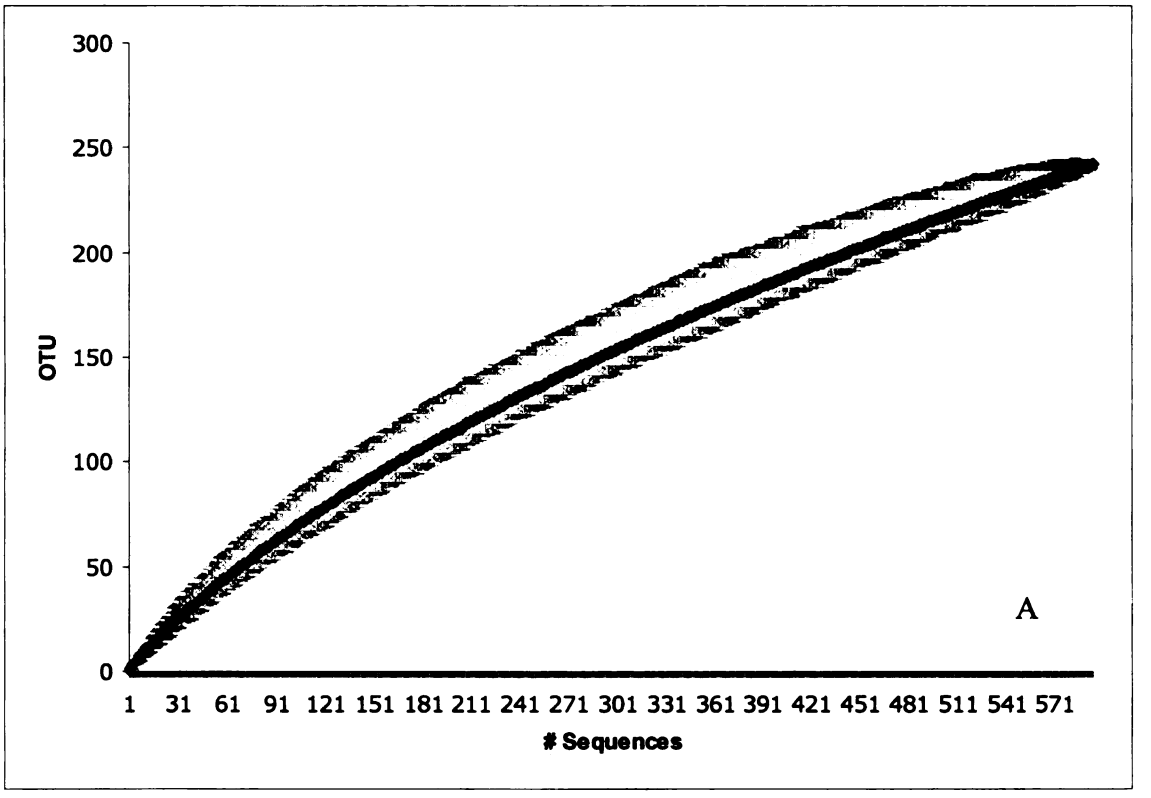


Figure 4.2. Rank abundance plot showing diversity of *An. gambiae* and *An. funestus* bacterial communities at evolutionary distance of 0.03. Relative abundance is plotted on logarithmic scale against species rank from most abundant to least abundant.



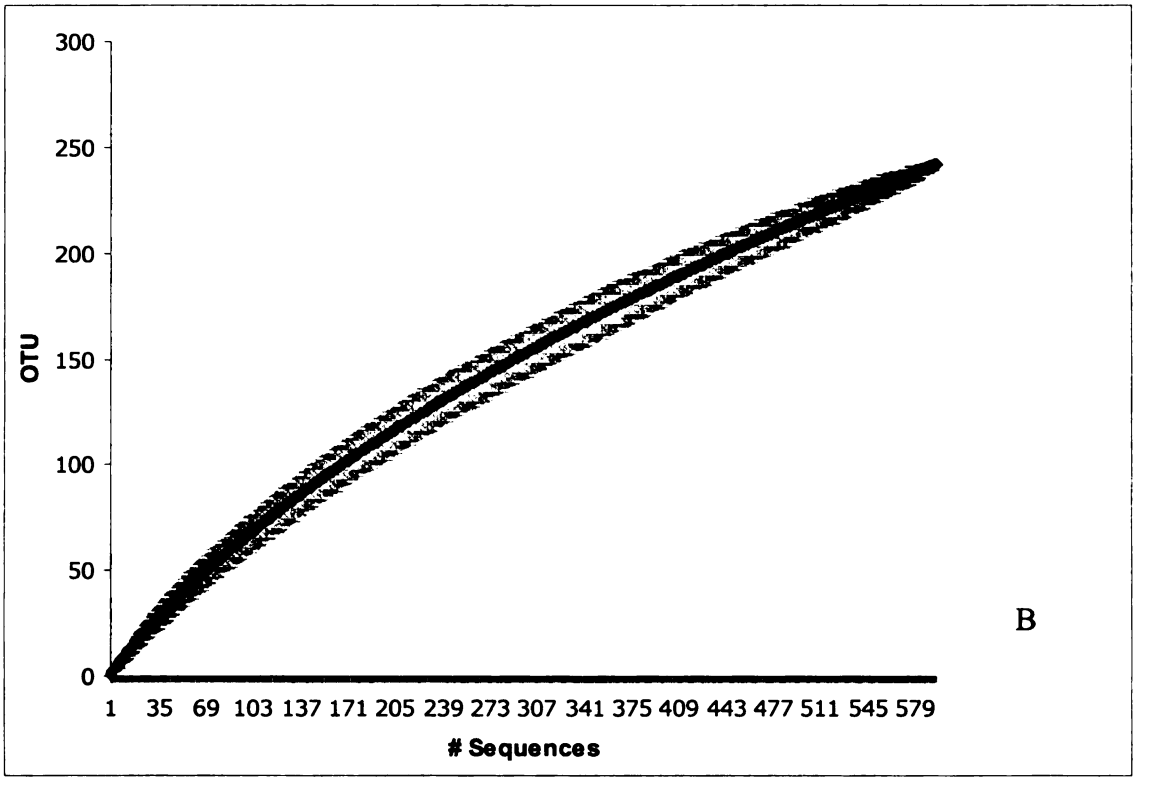


Fig 4.3. Rarefaction curves generated from the OTUs at 0.03 distance (97% similarity) using DOTUR for *An. gambiae* (A) and *An. funestus* (B) habitats.

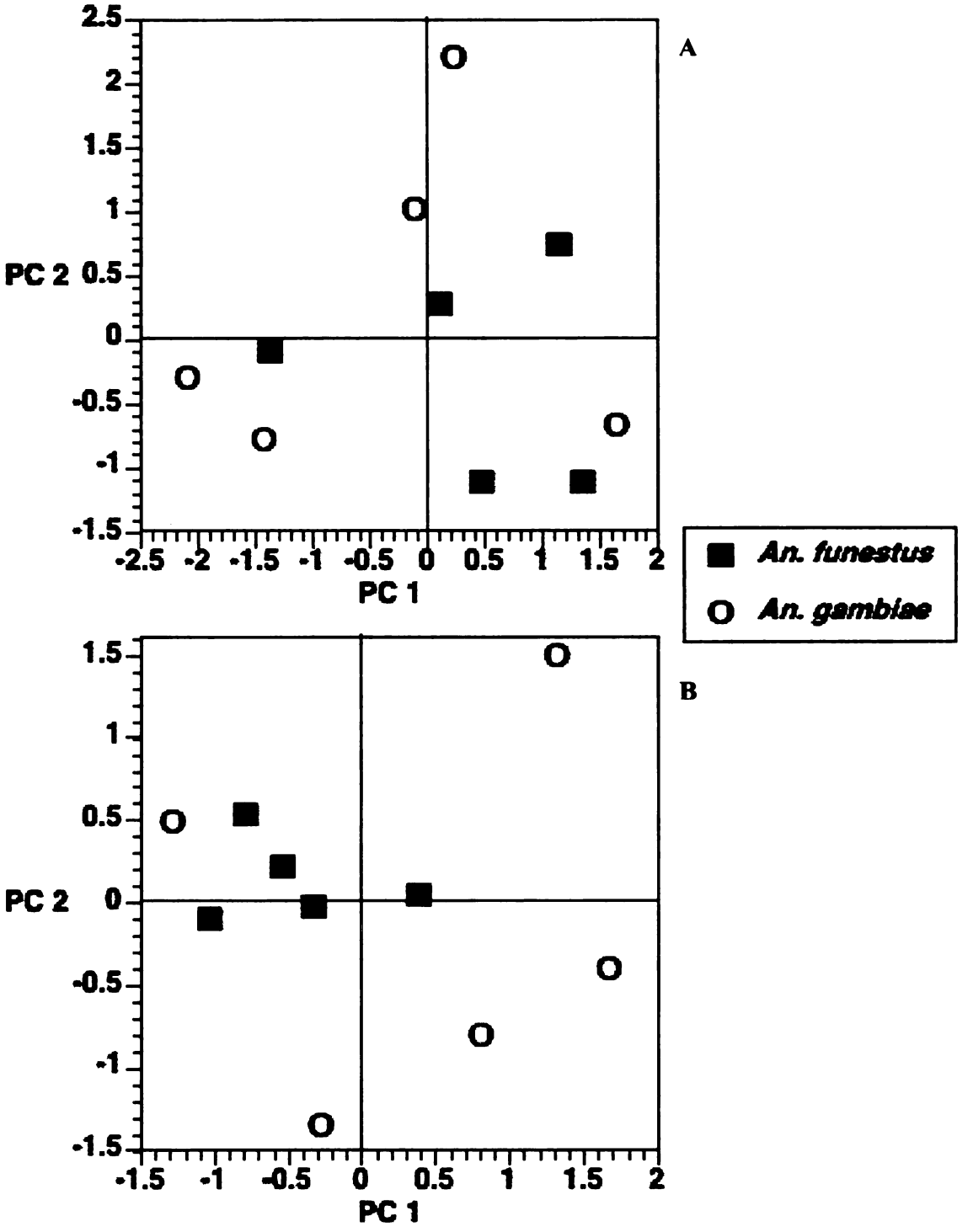


Fig 4.4. A, B: Principal component analysis of the 16S rDNA sequence data from *An. gambiae* and *An. funestus* habitats. OTUs determined by DOTUR analysis at 0.03 distance (97% similarity), A; OTUs obtained by RDP II classification at order level, B; 32.71% of the variance was explained by PC1-2 for DOTUR OTU data and 50.75% of the variance was explained by PC1-2 for RDP classification data.

Terminal Restriction Fragment Length Polymorphism Analysis (TRFLP) Results of cluster analysis of TRFLP data are shown in Fig 4.5. The total number of unique differently sized terminal fragments obtained from HhaI digestion for both *An. gambiae* and *An. funestus* combined were 203. 152 terminal fragments were observed in gambiae profiles and 121 fragments in funestus profiles. Eighty fragments were unique in gambiae profiles which were absent in funestus profiles, similarly 51 fragments were unique in funestus profiles.

The total number of unique differently sized terminal fragments obtained from MspI digestion for both *An. gambiae* and *An. funestus* combined was 178. 103 terminal fragments were observed in *An. gambiae* habitat profiles and 114 fragments in *An. funestus* habitat profiles. Sixty fragments were unique in gambiae profiles which were absent in funestus profiles, similarly 70 fragments were unique in funestus profiles.

Cluster analysis based on the UPGMA algorithm did not show any segregation of *An. gambiae* and *An. funestus* habitats along the trees generated by the PAUP program (Fig. 4.5 A, B). Thus fingerprints of bacterial communities of *An. gambiae* and *An. funestus* habitats obtained by either HhaI or MspI endonucleases digestion were not able to separate *An. gambiae* from *An. funestus* habitats.

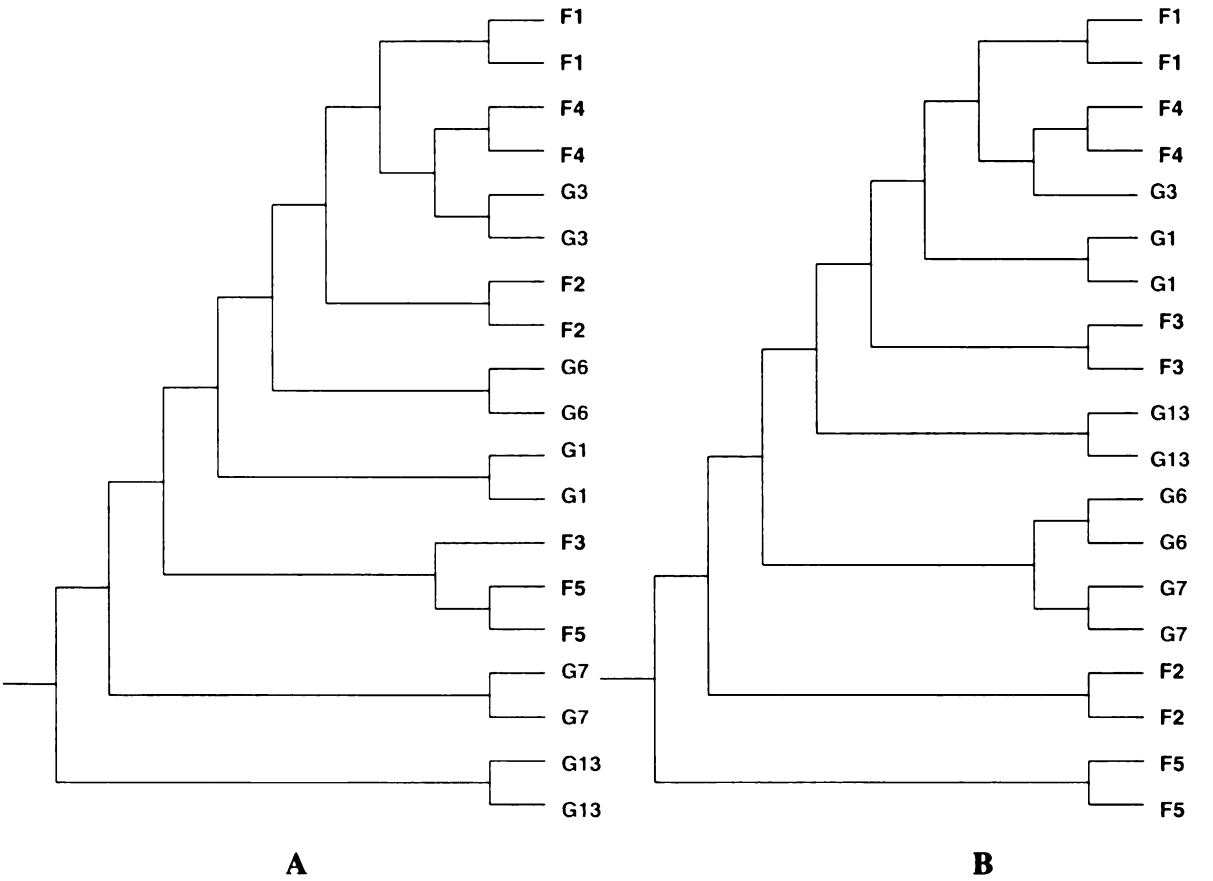


Figure 4.5. TRFLP Cluster Analysis using Paup for presence/absence of the terminal fragements obtained for HhaI (A) and MspI (B) restriction enzyme digestions. (G, *An. gambiae* habitats; F, *An. funestus* habitats).

DISCUSSION

In this study, the bacterial community composition of the surface microlayer from a discrete set of previously characterized *An. gambiae* and *An. funestus* habitats was assessed. Overall, the diversity was high in both the communities, as reflected by the values of the diversity indices, the shape of the rank abundance curves, and the fact that the rarefaction curves indicated that these communities were markedly under sampled for assessment of "true" diversity, which is typical of many microbial diversity studies in a wide range of environments (Hughes et al. 2001, Magurran 2004). The latter point was true, even though the sequence libraries generated from the process were generously large, compared with many studies (Singleton et al. 2001). LIBSHUFF analysis indicated that both of these communities were statistically significantly different in composition; however, the PCA and TRFLP analyses did not show any clear trends towards habitat segregation on the basis of bacterial community composition. These differences are very likely due to the differences of the procedures, LIBSHUFF used sequence data directly (as generated from a distance matrix in ARB), which would maximize the amount of diversity reflected in the sequence library at the various levels of similarity encompassed by the LIBSHUFF program (see Fig. 4.1). LIBSHUFF is more likely to detect differences when there are more operational taxonomic units represented by the original data. By contrast, TRFLP analyses used derivative variables (bins of 16S rDNA amplicon fragments sizes, with frequencies of fragments per bin as the actual variable) that reduce the variation to higher orders compared to LIBSHUFF, thus any true differences would be masked.

As in the research reported in Chapter 3, PCA was utilized here as a tool for reduction of sequence library data, where points representing the individual habitats were displayed in a bivariate plane defined by principle components 1 and 2. Two approaches were used, one in which the input data for PCA were frequencies of OTUs generated by DOTUR. Sequences showing 97% or higher sequence similarity were considered to be the same OTU. Other input data for PCA was the taxonomic classifications of sequences based upon the RDP II at the order level, regardless of the percent similarity of those matches. Results differed between these approaches, with the DOTUR classification system explaining more variation in PCA (ca. 50%) than RDP classification (ca. 33%). These results indicate that, from the standpoint of interpretation, the more precisely one attempts an OTU classification, the more variation will be explained or the more likely a difference in communities will be found, as was the case with the LIBSHUFF comparisons. I am unaware of any study in which the sensitivity, specificity, and error rate of LIBSHUFF have been estimated for sequence libraries of different sizes and compositions (Singleton et al. 2001), so the "alpha error" rate for LIBSHUFF is currently unknown. The LIBSHUFF website (http://whitman.myweb.uga.edu/libshuff.html) indicates as follows:

> "In simulations, the sensitivity of LIBSHUFF increases with the number of sequences in the library. For instance, when the library size n is 50, the introduction of 10-20 novel sequences into one library is frequently sufficient to allow LIBSHUFF to distinguish between the libraries (p=0.05). Similarly, when n = 100, the introduction of 10-20 novel sequences into one library is also frequently sufficient to allow LIBSHUFF to distinguish between the libraries (p=0.05)."

From this statement, one can see that both the sample size of the libraries being compared (i.e., total number of sequences) and the number of novel (i.e., unique) sequences will affect the sensitivity of the statistical test. The libraries compared here had both a large

number of sequences and many novel sequences, indicating that there was a sensitive test.

As discussed in Chapter 3, the use of PCA with classified 16S rDNA sequence data offers several advantages, in particular the use of analysis of variance on factor scores as a means of identifying particular groups contributing to differences in sequence library composition. However, such an approach is limited if a large number of sequences remain unclassified by the reference database, which was true for RDP II; a large fraction of sequences were unclassified here. Here, the ANOVAs did not identify significant factor scores when either RDP II or DOTUR were used as classification systems, although the *An. funestus* habitat samples did form clusters in PCA for the DOTUR classification, whereas *An. gambiae* habitat samples did not (Fig. 4.4).

Despite these findings, in this study there were no strong bacterial community composition differences that would allow inferences about habitat segregation of these mosquito species. Even though LIBSHUFF analysis performed on combined data from *An. gambiae* and *An. funestus* habitats showed that these habitats were statistically different not such indication was given by PCA and TRFLP analysis performed with data on individual habitat samples. Differences that were noted by the LIBSHUFF were at the higher taxonomic level and while communities shared deep taxa at lower taxononmic level. Notable difference between combined data that were observed by RDP classification was presence of several unclassified Cyanobacteria (7.16%) and unclassified bacteria (13.6%) in *An. funestus* habitats as opposed to 0.93 % and 4.03% respectively, in *An. gambiae habitats*. Higher frequencies of Actinomycetales and Rhodobacterales in *An. gambiae* habitats and higher frequencies of Alteromonadales in

An. funestus habitats, in particular a set of 105 sequences branching very closely to *Alishewanella fetalis*, an organism detected in a human fetus (Vogel et al. 2000) (see Appendix A for ARB trees and phylogenetic placement). These groups were present in most of the individual habitat samples and only 31 were found in *An. gambiae* habitats.

One possible explanation for substantial overlap in bacterial community composition of *An. gambiae* and *An. funestus* habitats within close proximity to each other, as here, is that heavy rains result in inundations and movement of water across the landscape, causing mixing of soil, water, and microorganisms into any location where water would settle after the rain subsided. The result would be that many common bacterial sequences would thus occur in both types of habitats.

Intensive field studies have shown that mosquitoes are quite discriminating in selecting sites for egg deposition and considerable evidence points to this site discriminating larval distribution (Bentley and Day 1989, Trexler et al. 2003). True larval habitat segregation in these species may be due to differences in female behavior with regard to egg laying (oviposition), unrelated to bacterial diversity in habitats. For example, Huang et al. (2006) showed that some bacteria were repellent or were not attractive to gravid An. gambiae females, whereas contrast, darkness, and substrate moisture were much strong elicitors of egg laying by this species. Boyd and Foot 1928 found negative association between Anopheline larvae and unicellular Cyanophyceace. No positive correlation found to exist between the distribution of larvae and the distribution of this plankton, due to the fact that these forms were also found in places that had not been utilized as breeding places.

McCrae (1984) also showed that *An. gambiae* preferred turbid water rather than clean water to oviposit their eggs. Maybe turbid waters for these species might be an indication of freshly formed habitats not previously utilized by cohorts and free of predators. In their review Bentley and Day 1989 have noted that concurrent presence of *An. gambiae* larvae reduced oviposition. This maybe either be pheromones produced by larvae or maybe due to fecal material present or microbial metabolites. *An funestus* on the other hand prefer clean water with partial coverage of aquatic vegetation and were not found in open waters (Giminig et al. 2001). For *An. funestus*, there is apparently little or nothing known about oviposition behavior, in part because of the lack of availability of laboratory colonies for research.

It is also possible that microbes other than bacteria might be important in oviposition selection by gravid females *An. gambiae*. Bond et al. 2005 showed two kinds of algae *Spirogyra majuscula* and *Cladophora glomerata* influenced oviposition site selection by *An. pseudopunctipennis* in a positive manner. Geetha et al. (2003) demonstrated that secondary metabolites produced by *Trichoderma viride*, a fungus, showed remarkable attractancy to oviposition by gravid females of *Culex quinquefasciatus* mosquito. It might thus be fruitful to evaluate microbes, otherthan bacteria as oviposition attractants for gravid *An. gambiae* females.

REFERENCES

Bentley M.D., Day J.F. 1989. Chemical ecology and behavioral aspects of mosquito oviposition. Annual reviews in Entomology. 34:401-421

Bond J. G., Arredondo-Jimenez J. I., Rodriguez M. H., Humberto Quiroz-Martinez Q., Williams T. 2005. Oviposition habitat selection for a predator refuge and food source in a mosquito. Ecological Entomology. 30: 255–263

Boyd M. F., Foot H. 1928. Studies on the bionomics of American Anophelines. The alimentation of Anopheline larvae and its relation to their distribution in nature. 219-242.

Braks, M.A.H., Anderson, R.A., Knols, B.G.J. 1999. Infochemicals in Mosquito Host Selection: Human Skin Microflora and *Plasmodium* Parasites. Parasitology Today, 15(10), 409-413.

Cole, J. R., B. Chai, R. J. Farris, Q. Wang, S. A. Kulam, D. M. McGarrell, G. M. Garrity, and J. M. Tiedje. 2005. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Research. 33:294-6.

Coluzzi M. 1984. Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. Bulletin of the world health organization. 62: 107-113.

Geetha I., Paily K.P., Padmanaban V., Balaraman K. 2003. Oviposition Response of the Mosquito, Culex quinquefasciatus to the Secondary Metabolite(s) of the Fungus, Trichoderma viride. Mem Inst Oswaldo Cruz, Rio de Janeiro. 98: 223-226.

Gimnig J., Ombok M., Kamau L., Hawley W. 2001. Characteristics of larval Anopheline (Diptera: Culicidae) Habitats in Western Kenya. Journal of Medical Entomology 38(2): 282-288.

Gimnig, J.E., M. Ombok, S. Otieno, M. Kaufman, J. Vulule, & E. D. Walker. 2002. Density-dependent development of *Anopheles gambiae* larvae in artificial habitats. Journal of Medical Entomology 39: 162-172.

Huang J., Miller J. R., Chen Shi-Cheng., Vulule J. M., Walker E. D. 2006. Anopheles gambiae (Diptera: Culicidae) oviposition in response to agarose media and cultured bacterial volatiles. Journal of Medical Entomology. 43: 498-504.

Hughes J. B., Hellmann J. J, Ricketts T. H., Bohannan B. J. M. 2001. Counting the uncountable: Statistical approaches to estimating microbial diversity. Applied and Environmental Microbiology. 67. 4399–4406

Kaufman, M.G., E.D. Walker, D.A. Odelson, & M.J. Klug. 2000. Microbial community ecology and insect nutrition. American Entomology. 46: 173-184.

Kaufman, M.G., S.J. Bland, M.E. Worthen, E.D. Walker, & M.J. Klug. 2001. Bacterial and fungal contributions to growth of larval *Aedes triseriatus* (Diptera: Culicidae). Journal of Medical Entomology. 38: 711-719.

Kaufman, M.G., Goodfriend W., Kohler-Garrigan A., Walker E.D., & Klug M.J. 2002. Soluble nutrient effects on microbial communities and mosquito production in *Ochlerotatus triseriatus* (Say) habitats. Aquatic Microbial Ecology 29: 73-88.

Kaufman M. G., Wanja E., Maknojia S. R., Bayoh N. M., Vulule J. M., Walker E. D. 2006. The importance of algal biomass to the growth and development of *Anopeheles gambiae* larvae. Journal of Medical Entomology. 43: 669-676.

Lane D. J., Pace B., Olsen G. J., Stahl D. A., Sogin M. L., Pace N. R. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc. Natl. Acad. Sci. USA. 82: 6955-6959.

Lilburn T. G., Schmidt T. M., Breznak J. A. 1999. Phylogenetic diversity of termite gut spirochaetes. Environmental Microbiology 1(4): 331-345.

Liu Wen-Tso, Marsh T. L., Cheng H, Forney L. J. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Applied and Environmental Microbiology. 63. 4516–4522

Magurran A. E. 2004. Measuring biological diversity. Blackwell Science Limited, 108 Cowley Road, Oxford, OX4 1JF.

Maidak B. L., C. J. R., Parker C. T., Garrity G. M., Larsen N., Li B., Lilburn T. G., McCaughey M. J., Olsen G. J., Overbeek R., Pramanik S., Schmidt T. M., Tiedje J. M., Woese C. R. 1999. A new version of the RDP (Ribosomal Database Project). Nucleic_Acids_Research 27(1): 171-173.

Marchesi J. R., Sato T., Weightman A. J., Martin T. A., Fry J. C., Hiom S. J., Wade W. G. 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Applied and Environmental Microbiology. 64(2): 795-799.

Marsh T.L. 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Current. Opinions in Microbiology. 2: 323-327.

Marsh T. L., Saxman P., Cole J., Tiedje J. 2000. Terminal Restriction Fragment Length Polymorphism Analysis Program, a Web-Based Research Tool for Microbial Community Analysis. Applied and Environmental Microbiology. 66: 3616-3620

McCrae A. W. R. 1984. Oviposition by African malaria vector mosquitoes II. Effects of site tone, water type and conspecific immatures on target selection by freshwater *Anohpeles gambiae* Giles *sensu lato*. Annals Tropical Medicine and Parasitology. 78: 307-318.

Singleton D. R., Furlong M. A., Rathbun S. L., Whitman W. B. 2001. Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. Applied Environmental Microbiology. 67: 4374-6.

Strunk O. and Ludwig W. 1995. ARB- a software environment for sequence data. Department of Microbiology, Technical University of Munich, Germany.

Trexler, J.D., Apperson C. S., Zurek L., Gemeno C., Schal C., Kaufman M., Walker E., Watson D. W. and Wallace L.. 2003. Role of bacteria in mediating oviposition responses of *Aedes albopictus* (Diptera: Culicidae). Journal of Medical Entomology. 40(6): 841–848.

Trexler, J.D., Apperson C. S., Zurek L., Gemeno C., Schal C., Kaufman M., Walker E., Watson D. W. and Wallace L.. 2003. Role of bacteria in mediating oviposition responses of *Aedes albopictus* (Diptera: Culicidae). Journal of Medical Entomology. 40(6): 841–848.

Vogel B. F., Venkateswaran K., Christensen H., Falsen E., Christiansen G,. Gram L. 2000. Polyphasic taxonomic approach in the description of *Alishewanella fetalis* gen. nov., sp. nov., isolated from a human foetus. International Journal of Systematic and Evolutionary Microbiology. 50: 1133–1142

Walker E. D., Olds E. J., Merritt R. W. 1988. Gut content analysis of mosquito larvae (Diptera: Culicidae) using DAPI stain and Epifluorescence microscopy. Journal of Medical Entomology. 25(6): 551-554.

Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b pheopigments. Limnol. Oceanogr. 39:1985-1992

SUMMARY

The objectives underlying this research project were to analyze the effects of removal of the surface microlayer on mosquito growth and to study the contribution of bacteria to larval nutrition. Along with the above objectives, I also wanted to determine the bacterial diversity of the surface microlayer of *An. gambiae* habitats in presence and absence of larval feeding pressure and also the bacterial diversity of naturally occuring *An. gambiae* and *An. funestus* habitats using 16S rDNA sequence library construction. Although there is lot of research that has been undertaken with the adults of these species, very little is known about the larval biology and the habitats in which these larvae develop.

Chapter II of this dissertation attempts to address the first two objectives. Results indicate that removal of surface microlayer at regular intervals resulted in decreased survival of *An. gambiae* larvae, prolonged developmental time to pupation, and resultant adults with lower body mass. Supplementations of the surface microlayer from sources with no larval grazing improved larval growth. Larval developmental time in these microcosms was shortened and total mass was significantly greater, compared to microcosms from which surface microlayer was deliberately removed. *An. gambiae* habitats are hydrologically unstable and dry quickly, thus larvae must feed and develop before habitats dry up. Food limitation can affect this critical life history parameter (developmental time) which was the most sensitive variable in the above case. In another experiment, the importance of bacteria as a food source was studied. Heterotrophic bacterial production was stimulated with addition of glucose as a carbon source, in

microcosms that were intentionally shaded to reduce autotrophic algal production. It was observed that larvae grew poorly, few molted, and few achieved metamorphosis to pupation under these conditions, and that larval survival rate and total adult emergence was very low compared to sunlit treatments which were rich in algae. Thus, bacterial biomass does not compensate for algal biomass in larval growth, but it may be important as a larval food supplement under conditions where algae are grazed. Bacteria supported some larval growth but probably did not provide the broad suite of micronutrients required by mosquito larvae for full metamorphosis.

To determine whether larval grazing had any effect on bacterial community compostion in *An. gambiae* habitats, the bacterial diversity within these habitats with presence and/or absence of larvae within two different soil types was studied. Overall, the results in (Chapter III) demonstrate that presence of larvae has strong effects on bacterial community composition and that these bacterial communities were significantly different. There were a large number of unclassified sequences in no larvae treatments in both soil types, and they were closest to sequences of phototrophic primary producers, likely representing chloroplast rDNA sequences or unclassified cyanobacterial rDNA sequences. Community shifts were observed either by presence or absence of certain taxa, or changes in the frequencies of certain taxa, as represented by the sequence data. These community shifts may explain patterns of variation in productivity of habitats in natural settings and also habitat segregation of *Anopheles* species due to oviposition site selection. However, the results in Chapter IV indicate that the bacterial communities in natural *An. gambiae* and *An. funestus* habitats seem to have no specific bacterial

signatures that might explain habitat segregation and selective oviposition site selection among these species.

APPENDIX

APPENDIX A

Phylogenetic trees

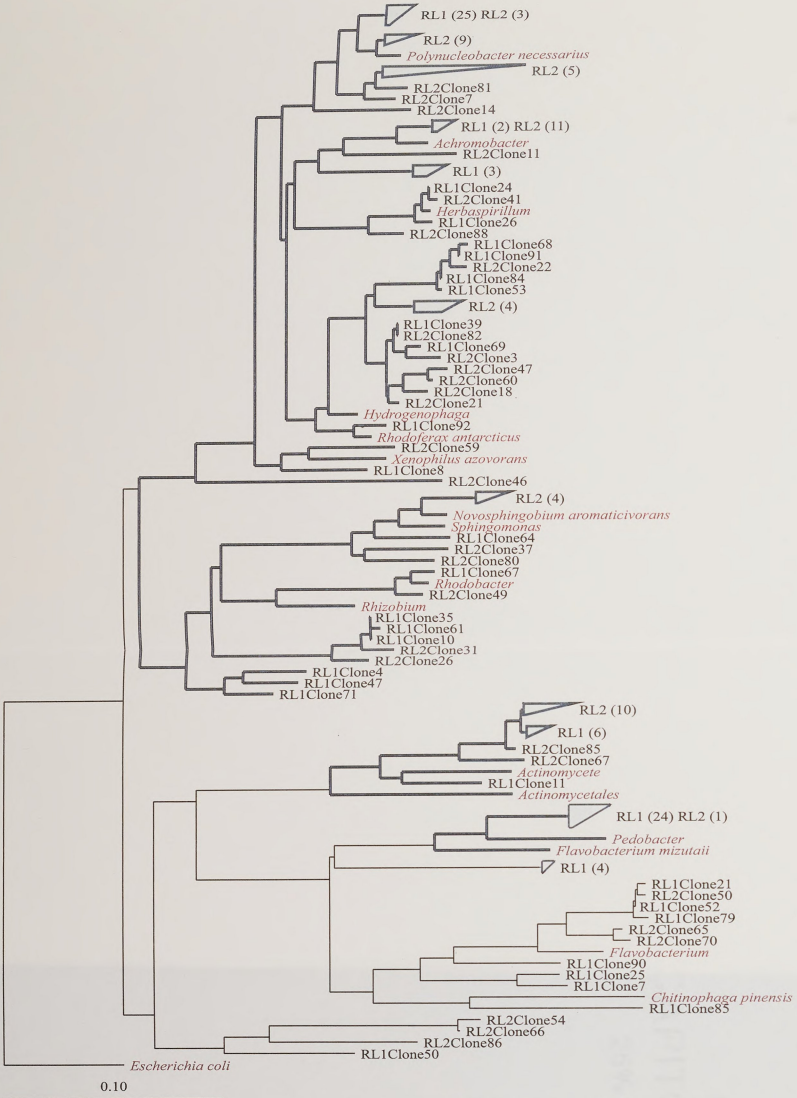


Fig 3.10. Phylogenetic analysis of 16S rDNA gene sequences obtained from Red Soil/larvae. The dendogram was generated by Neighbour Joining Method uaing a filter of 265 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Labels RL1 and RL2 refer to clones retreived from two different replicates of same treatment. Number of sequences within each group are in parenthesis.

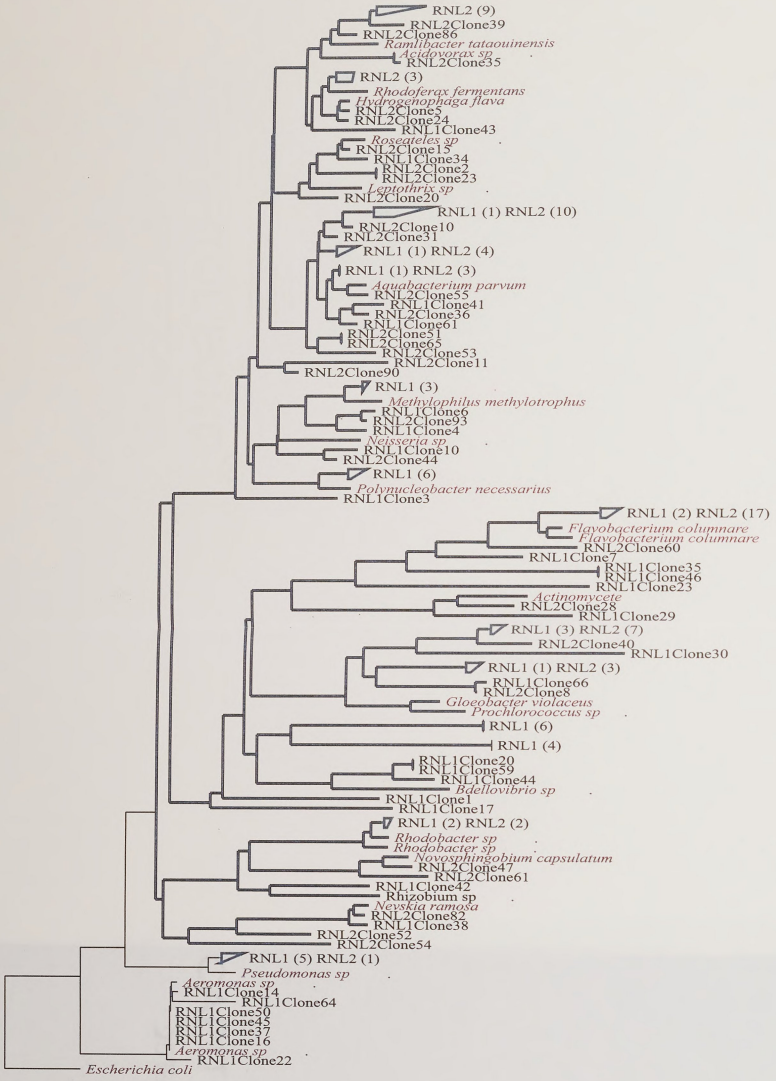


Fig 3.11. Phylogenetic analysis of 16S rDNA gene sequences obtained from Red Soil/no larvae. The dendogram was generated by Neighbour Joining Method uaing a filter of 265 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Labels RNL1 and RNL2 refer to clones retreived from two different replicates of same treatment. Number of sequences within each group are in parenthesis.

Fig 3.12. Phylogenetic analysis of 16S rDNA gene sequences obtained from Black Soil/larvae. The dendogram was generated by Neighbour Joining Method uaing a filter of 265 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Labels BSL1 and BSL2 refer to clones retreived from two different repliates of same treatment. Number of sequences within each group are in parenthesis.

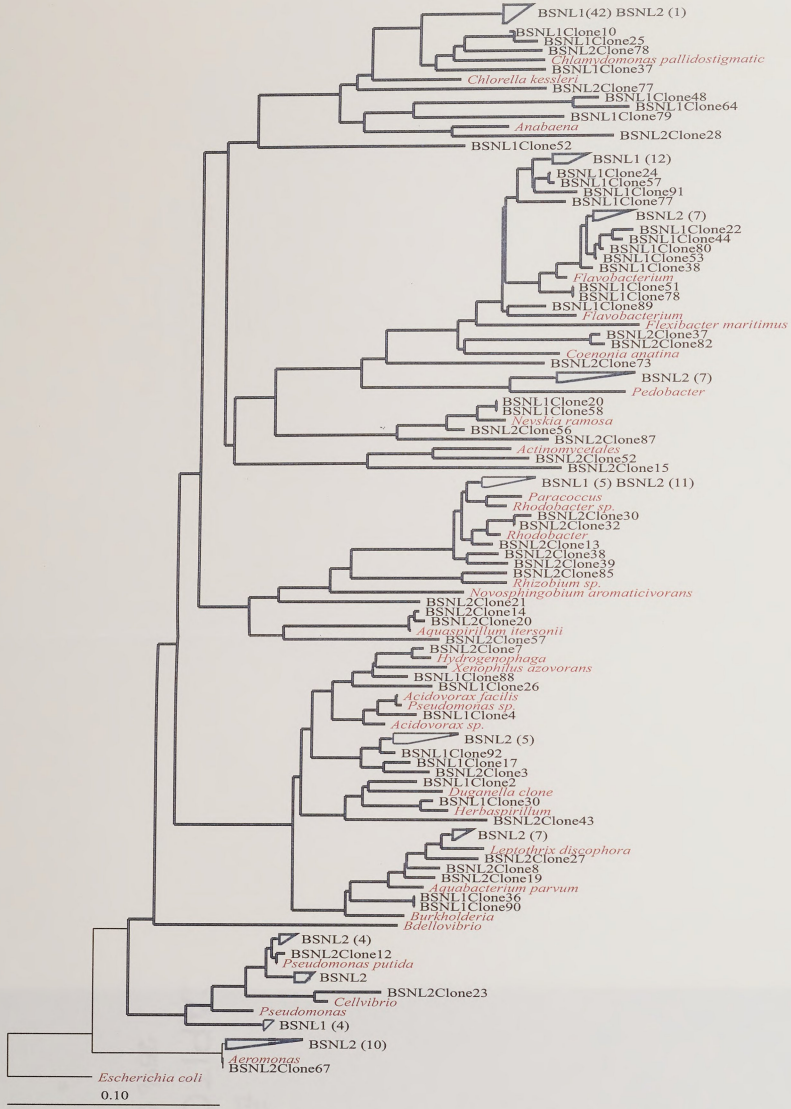


Fig 3.13. Phylogenetic analysis of 16S rDNA gene sequences obtained from Black Soil/no larvae. The dendogram was generated by Neighbour Joining Method uaing a filter of 265 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Labels BSNL1 and BSNL2 refer to clones retreived from two different repliates of same treatment. Number of sequences within each group are in parenthesis.

Figure 4.6. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. gambiae* habitat (G1). The dendogram was generated by Neighbor Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

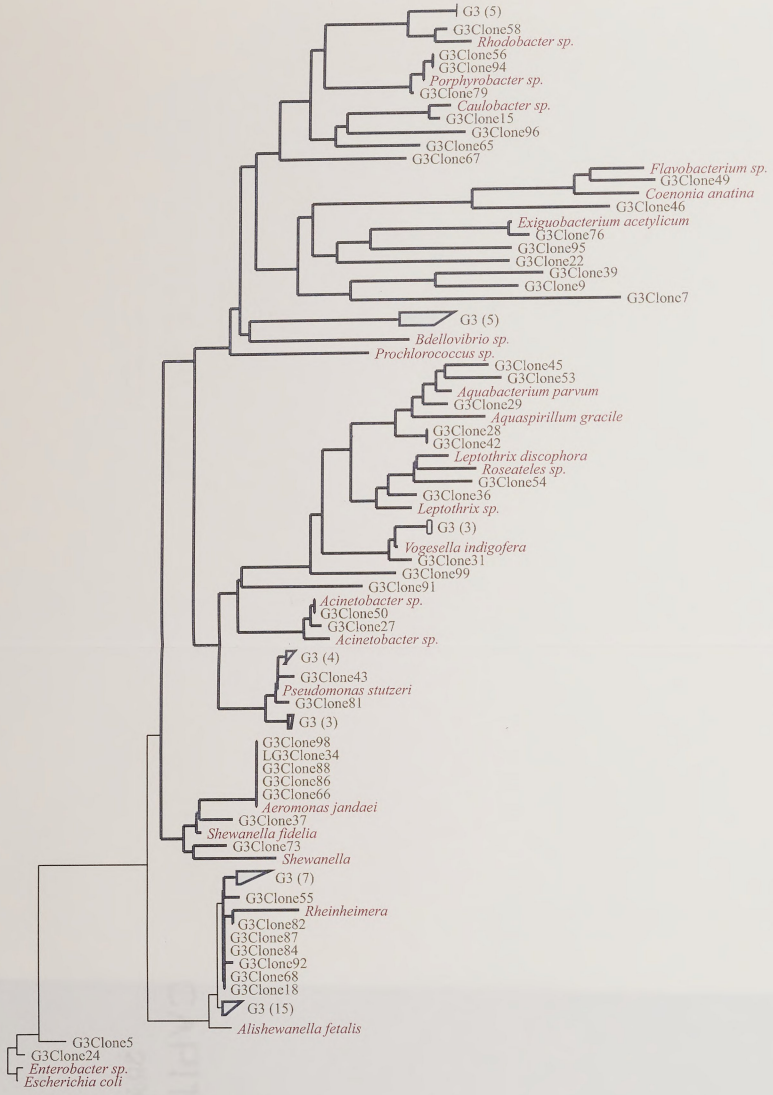


Figure 4.7. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. gambiae* habitat (G3). The dendogram was generated by Neighbor Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

Figure 4.8. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. gambiae* habitat (G6). The dendogram was generated by Neighbor Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

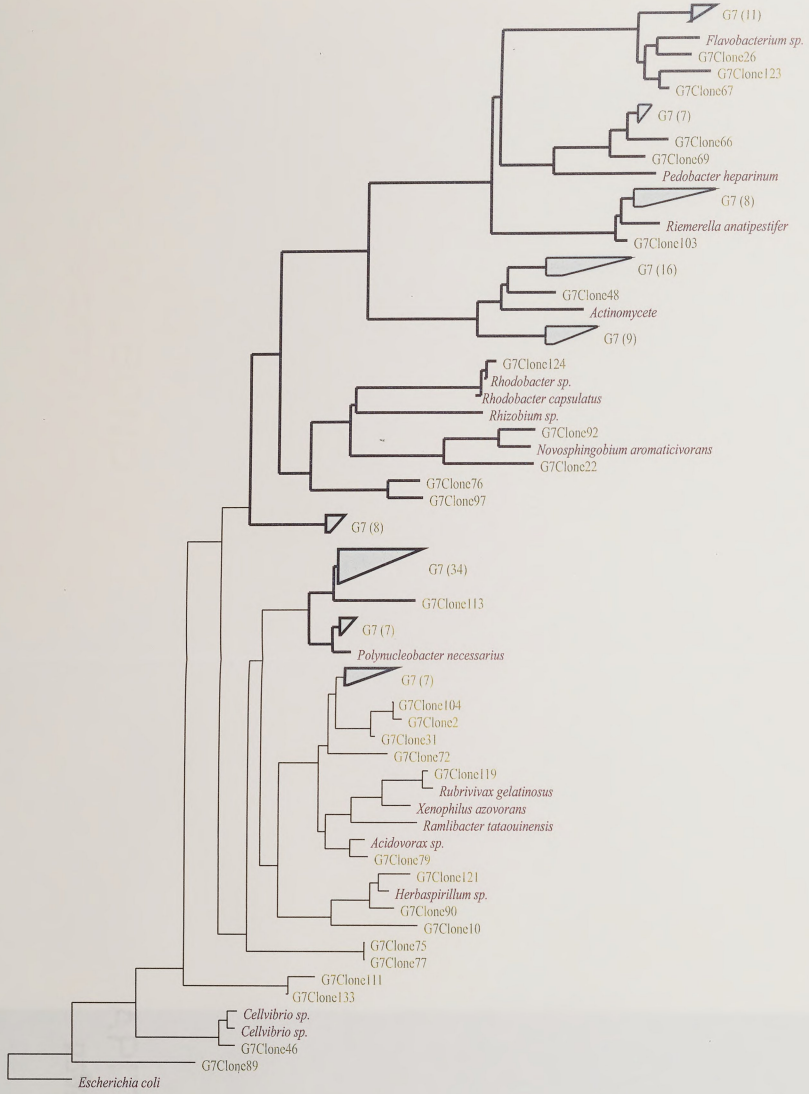


Figure 4.9. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. gambiae* habitat (G7). The dendogram was generated by Neighbor Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

Figure 4.10. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. gambiae* habitat (G13). The dendogram was generated by Neighbor Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

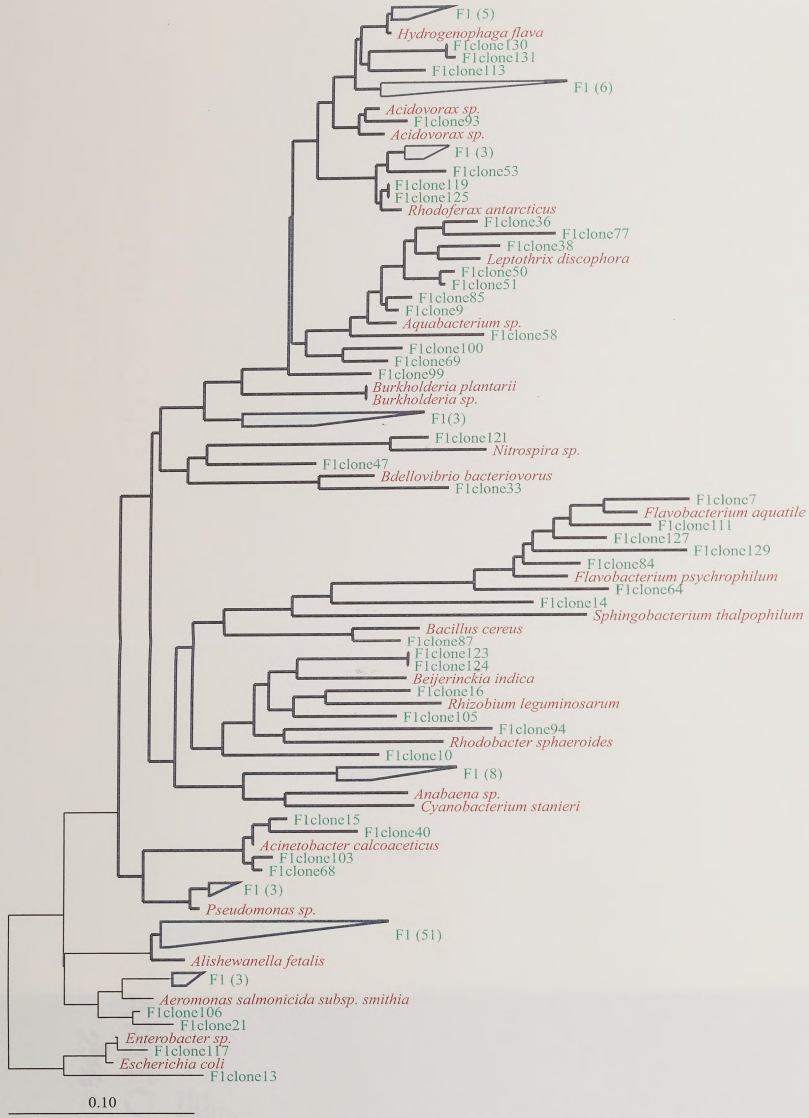


Figure 4.11. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. funestus* habitat (F1). The dendogram was generated by Neighbour Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

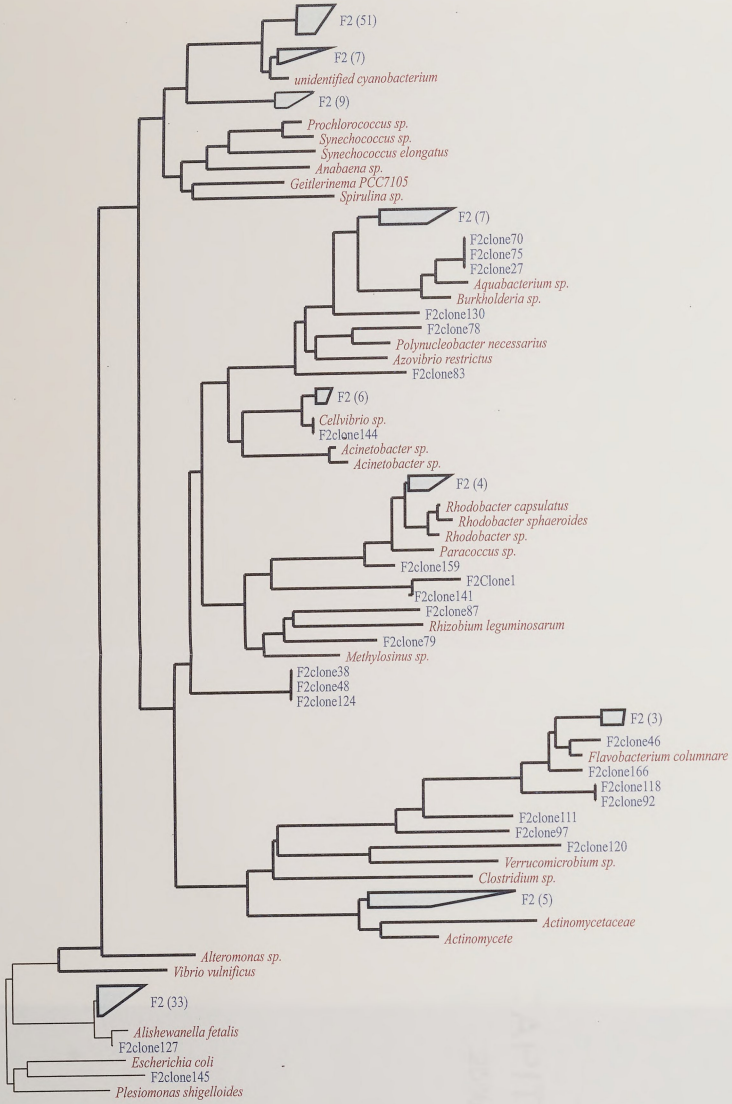


Figure 4.12. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. funestus* habitat (F2). The dendogram was generated by Neighbour Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

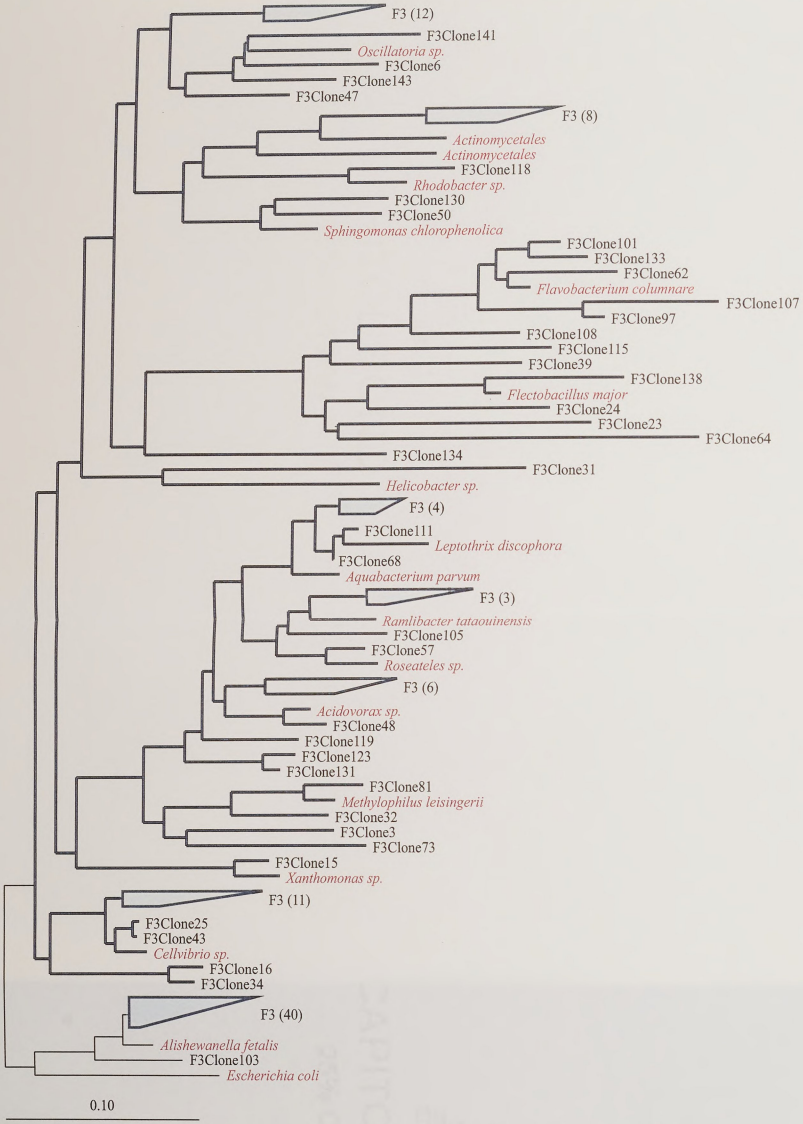


Figure 4.13. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. funestus* habitat (F3). The dendogram was generated by Neighbour Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

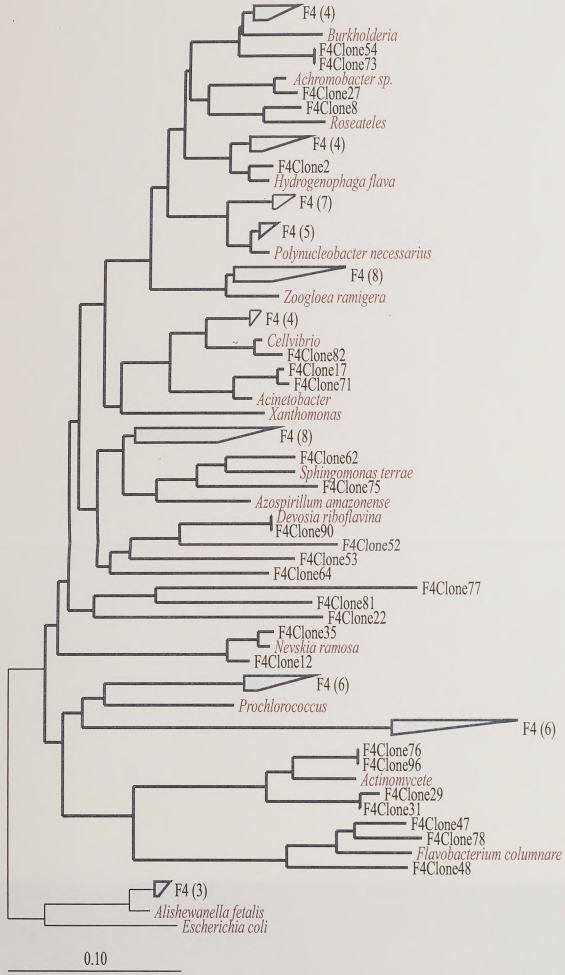


Figure 4.14. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. funestus* (F4) habitat. The dendogram was generated by Neighbour Joining Method uaing a filter of 265 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

F5Clone144
Sphingomonas aromaticivorans

Figure 4.15. Phylogenetic analysis of 16S rDNA gene sequences obtained from *An. funestus* habitat (F5). The dendogram was generated by Neighbour Joining Method uaing a filter of 200 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

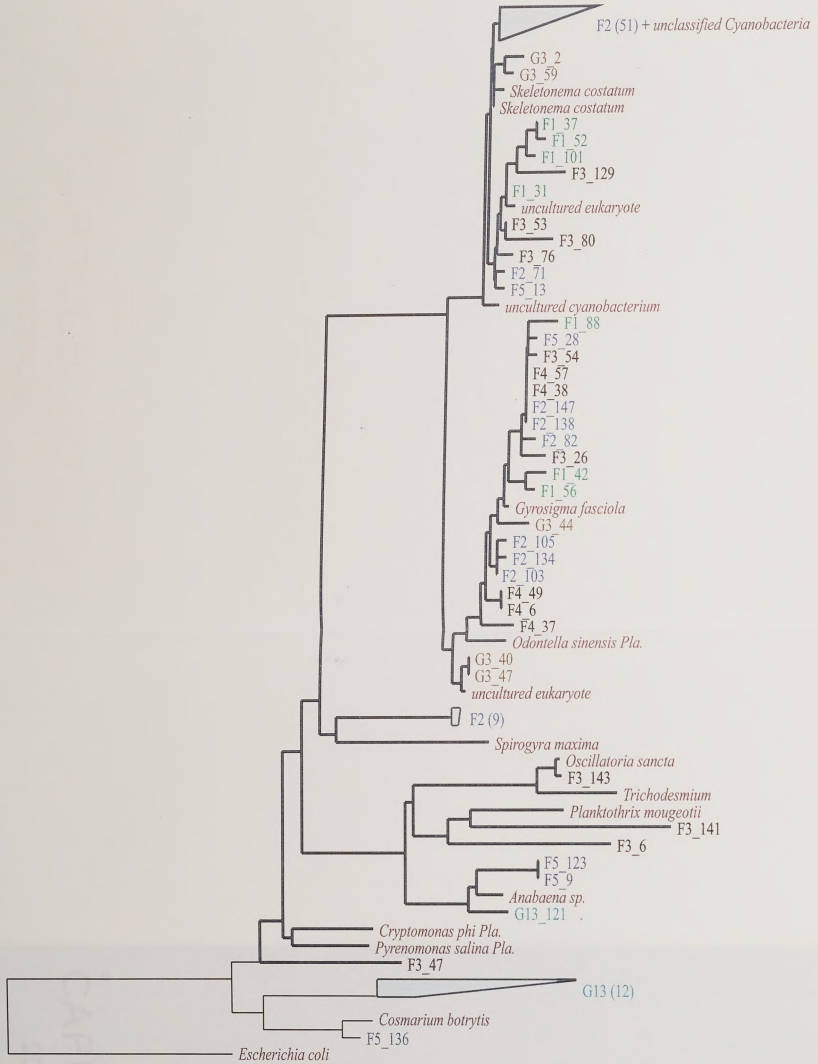


Figure 4.16. Phylogenetic analysis of 16S rDNA gene sequences from the unclassfied group obtained from *An. gambiae* (G1, G3, G6, G7 and G13) and *An. funestus* (F1, F2, F3, F4 and F5) habitat. The dendogram was generated by Neighbour Joining Method uaing a filter of 245 bases and Jukes cantor correction using ARB software. The scale bar represents 10% sequence divergence. Number of sequences within each group are in parenthesis.

APPENDIX B

Record of Deposition of Voucher Specimens

Appendix B

Record of Deposition of Voucher Specimens*

The specimens listed on the following sheet(s) have been deposited in the named museum(s) as samples of those species or other taxa, which were used in this research. Voucher recognition labels bearing the Voucher No. have been attached or included in fluid-preserved specimens.

Voucher No.: 2006-01

Title of thesis or dissertation (or other research projects):

BACTERIAL DIVERSITY AND NUTRITIONAL SIGNIFICANCE OF THE SURFACE MICROLAYER IN ANOPHELES GAMBIAE (DIPTERA: CULICIDAE) LARVAL HABITATS

Museum(s) where deposited and abbreviations for table on following sheets:

Entomology Museum, Michigan State University (MSU)

Other Museums:

Investigator's Name(s) (typed)

Shahnaz Rahim Maknojia

Date 03/23/06

*Reference: Yoshimoto, C. M. 1978. Voucher Specimens for Entomology in North America. Bull. Entomol. Soc. Amer. 24: 141-42.

Deposit as follows:

Original: Include as Appendix B in ribbon copy of thesis or dissertation.

Copies: Include as Appendix B in copies of thesis or dissertation.
Museum(s) files.
Research project files.

This form is available from and the Voucher No. is assigned by the Curator, Michigan State University Entomology Museum.

Voucher Specimen Data

Page ___ of ___ Pages

Species or other taxon	Label data for specimens collected or used and deposited	Number of: Eggs	Larvae	Nymphs	Pupae	Adults ♀	Adults ♂	Other	Museum where deposited
Anopheles gambiae	Michigan Ingham Co. East Lansing - MSU Campus Lab culture (Kisumu strain) 20-Mar-06 S. Maknojia coll.		20			10	10		

(Use additional sheets if necessary)

Investigator's Name(s) (typed)

Shahnaz Rahim Maknojia

Date 3/23/2006

Voucher No. 2006-01

Received the above listed specimens for deposit in the Michigan State University Entomology Museum.

Curator 6 Dec 2006
Date