PARTNERS IN TIME: EVOLUTIONARY AND POPULATION GENETIC PATTERNS OF COEVOLVING ORGANISMS

By

JULIA M. ALLEN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

To my fellow graduate students, my committee and my family for all of their support

ACKNOWLEDGMENTS

I would like to first thank all of the graduate students in the Biology Department for all of their support and encouragement, in particular those in my cohort. I would also like to thank my advisor David Reed for always listening and allowing me to have the freedom to develop my own ideas, but also challenging them, which made me think about things more critically all along the way. In particular, I would like to thank him for his seemingly unlimited amounts of time to discuss ideas. I would also like to thank Mike Miyamoto for spending hours with me explaining coalescent processes and population genetics, and always pushing me to be a better scientist. Marta Wayne for providing true mentorship not only to me, but to all of the graduate students in the department. I would like to thank Matt Gittendanner for his time helping me with the cluster, and Pam and Doug Soltis for giving me unlimited access to their lab, without which, many of the techniques I needed to learn would not have happened. I would also like to thank Rebecca Kimball and Edward Braun for giving me the time to discuss lab work and troubleshooting, usually in last minute frantic meetings for which they always found the time. Also, I would like to thank Gordon Burleigh for all of his work on the bacterial phylogenetics, he is an honorary member of my committee. Lastly, I would like to thank my family for all of their support along the way.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

LIST OF TABLES 8

LIST OF FIGURES 9

ABSTRACT 10

CHAPTER

1 INTRODUCTION 12

Overview 12
Part I: Lice and their Primary Endosymbionts 12
Part II: Sucking Lice and their Mammalian Hosts 13

2 EVOLUTIONARY RELATIONSHIPS OF *CANDIDATUS* RIESIA SPP., ENDOSYMBIOTIC *ENTEROBACTERIACEAE* LIVING WITHIN HEMATOPHAGOUS PRIMATE LICE 15

Introduction 15
Methods 18
Specimen Preparation 18
Bacterial Diversity in *Pediculus Humanus Capitis* Using Molecular Methods 18
Targeted Sequencing of Endosymbionts from Additional Taxa of Lice 19
Fluorescent In-Situ Hybridization 20
Phylogenetic Analysis 21
Results 22
Phylogenetic Analysis and Taxonomic Position of Louse Endosymbiont 22
Fluorescent In-Situ Hybridization 23
Discussion 24
Overview 24
Nomenclature 25
Coevolution and Rates of Substitution 26

3 MUTATIONAL MELTDOWN IN PRIMARY ENDOSYMBIONTS: SELECTION LIMITS MULLER'S RATCHET 31

Introduction 31
Materials and Methods 35
Age of the Riesia/Louse Assemblage 35
Specimen collection and DNA sequencing 35
Phylogenetic analysis 37
Absolute rates of nucleotide evolution in *Riesia* 38

Substitution Rates Among Host/Endosymbiont Lineages 39
Results 40
Age of *Riesia*/Louse Association 40
Absolute Rates 41
Substitution Rates Among Host/Endosymbiont Lineages 42
Discussion 42
Overview 42
Endosymbiosis 44

4 MULTIPLE LINEAGES OF BACTERIA IN ANOPLURA INDICATE A HIGH RATE OF BACTERIAL REPLACEMENT ON A SHORT EVOLUTIONARY TIME-SCALE 53

Introduction 53
Methods 55
Lice Sampling 55
Bacterial Sampling Alignment and Tree Building 56
Results 57
Endosymbionts 57
Phylogenetic Analysis 58
Discussion 59
Endosymbionts 59
Taxonomy 60
Bartonella as an Endosymbiont 61
Number of lineages 61
Evolution of Sucking Lice and their Endosymbionts 62
Blood Feeding as a Source of Endosymbionts 63

5 POPULATION GENETICS OF HABITAT SENSITIVE RED COLOBUS SUGGEST LONG-TERM STABILITY OF KIBALE NATIONAL PARK 70

Introduction 70
Methods 72
DNA Collection and Storage 72
DNA Extraction and Quantification 72
Genotype analysis 73
Number of Populations 74
Analysis 75
Estimates of Theta, Population Size and Bottlenecks 75
Results 76
Genotyping Errors, Linkage and Hardy Weinberg Equilibrium 76
Number of Populations 77
Effective population size, Coalescent Point and Bottleneck 78
Discussion 79
Number and size of populations 79
History of Red Colobus in Kibale National Park 80

6 CONCLUSIONS 87

Part I: Lice and their Primary Endosymbionts 87
Part II: Sucking Lice and their Mammalian Hosts 88

LIST OF REFERENCES 91

BIOGRAPHICAL SKETCH 103

LIST OF TABLES

Table		page
2-1	Pairwise differences for p-endosymbionts from human head and body lice *Pediculus humanus*.	28
3-1	Percent sequence divergence and rate of nucleotide substitution of the *16S rDNA* gene of *Riesia.*	47
3-2	Bacteria taxa, their hosts, and GenBank accession numbers used in the phylogenetic analysis presented in Figure 3-1.	48
4-1	Louse taxa used in this study (arranged by family), host associations, and GenBank accession numbers.	64
4-2	Taxon sampling, and endosymbiont determination.	66
5-1	Microsatellite loci genotyped for Red Colobus (*Piliocolobus tephrosceles*) from Kibale National Park.	82
5-2	Results from RSTCALC.	83

LIST OF FIGURES

Figure		page
2-1	Phylogenetic tree of endosymbiont from sucking lice (Anoplura) and other bacteria..	29
2-2	Fluorescent in situ hybridization microphotograph of thorax and abdomen (ventral view) of a second instar nymph of a head louse (*Pediculus humanus capitis*)..	30
3-1	Maximum likelihood phylogram representing phylogenetic relationships of alll bacteria used in this study..	49
3-2	Maximum likelihood phylogram representing phylogenetic relationships of *Riesia* p-endosymbionts.	50
3-3	Primary endosymbiont nucleotide substitution rates as a function of the age of insect/p-endosymbiont association.	51
3-4	Reduced major axis regression of the log transformed data from Figure 3-3. .	52
4-1	Bacterial tree showing placement of Anopluran endosymbionts.	67
4-2	Phylogenetic tree of Louse taxa from Light et al. (In Review).	68
4-3	Maximum Likelihood tree of 279 *Bartonella* sequences..	69
5-1	Map of Kibale National Park (KNP).	84
5-2	Log likelihood scores for output from Structure.	85
5-3	Log likelihood scores for the program Migrate..	86

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

PARTNERS IN TIME: EVOLUTIONARY AND POPULATION GENETIC PATTERNS OF COEVOLVING ORGANISMS

By

Julia M. Allen

August 2010

Chair: David L. Reed
Major: Zoology

The way organisms evolve is thought to be due partly in response to their interactions with other organisms. An extreme form of interactions between organisms is relationships where one organism is completely dependent on another. My dissertation has focused on how organisms in these types of relationships have evolved over both long time scales (millions of years) and short ones (just a few generations). I have examined two groups of interacting organisms. First, sucking lice have mutualistic bacteria, which live inside the insect and provide nutrients to the louse. The bacteria are passed from mother to offspring to get incorporated into the next generation. I looked at 27 species of lice to see if they share a single genus of bacteria, which would suggest that the lice and endosymbionts have evolved together for millions of years. I found that closely related lice do share a single genus of bacteria but when you look at more distantly related lice they have unrelated lineages of bacterial endosymbionts. This work suggests that endosymbionts and lice might evolve together for some period of time but that over long time scales the bacteria possibly get replaced.

Secondly, I have studied a group of Red Colobus monkeys and their parasitic lice from Kibale National Park in Uganda. Here I found that there is gene-flow between the

monkey troops from across the park, suggesting there is contact between the host populations. This movement of the hosts likely affects the rate at which these parasites are able to move around the park.

CHAPTER 1
INTRODUCTION

Overview

The evolutionary history of an organism is strongly affected by its interactions with other species in a shared environment (e.g. predator-prey). Interactions with these organisms may affect the evolutionary trajectory of a lineage. An extreme case of such species interactions is in obligate relationships because one organism is completely dependent on the other. It is important to understand how the evolutionary processes of one partner affects the evolutionary history of the other. Blood-sucking lice form two strong obligate relationships; they are obligate parasites of mammals, and each louse harbors a mutualistic species of bacteria. These relationships make sucking lice ideal for examining the co-evolutionary history of interacting lineages.

Part I: Lice and their Primary Endosymbionts

Primary endosymbiotic bacteria (p-endosymbionts) are thought to be partially responsible for the incredible diversification of insects. P-endosymbionts are found in insects that specialize on nutrient poor diets, where they supplement their insect hosts' diet with nutrients (Buchner 1965). It has been suggested that the acquisition of a bacterial lineage in insect ancestors has played a key role in the species diversification of some insect groups. Endosymbiotic bacteria cannot live outside of the insect hosts and are transmitted from mother to offspring through the eggs. Within insect groups with p-endosymbionts, all individuals share the same lineage of bacteria. For example all aphids have the same genus of bacteria called *Buchnera*. This suggests a long coevolutionary history between the insect-bacterial partners.

The primary-endosymbiont in the human head louse was first discovered over three hundred years ago, (Hooke 1664) but before my research began had not been formally described and named. In Chapter 2 I will describe the endosymbiont in Great Ape lice. In Chapter 3 I examine its rate of molecular evolution in light of Muller's Ratchet and the rate of evolution of other insect/endosymbionts. Then in Chapter 4 I describe endosymbionts in many species of sucking lice to determine a rate of endosymbiont replacement in this group.

Part II: Sucking Lice and their Mammalian Hosts

Organisms involved in obligate relationships are affected on both a long-term species-level time scale and short-term population-level time scale (Clayton and Johnson 2003). For example, the size of a parasite population is strictly tied to the social structure, dispersal ability and population size of its host. In the case of lice and mammals, lice only move between hosts during direct host-to-host contact. Therefore, louse population dynamics are particularly constrained by the social structure of their hosts (e.g. which hosts come into contact and how often they come into contact). The effect of mammalian host population dynamics on louse population dynamics remains poorly understood.

The Red Colobus monkey is parasitized by lice in the genus *Pedicinus*, and these monkeys breed and forage in stable social groups. For these primates, it is thought that only females migrate from troop to troop (Struhsaker 1975). Female migration may therefore provide the mechanism for lice to disperse between monkey troops. To examine the host population structure I have selected six troops of monkeys that reside within Kibale National Park (Uganda). Using molecular markers, in Chapter 5 I calculate

the number of populations of monkeys in the park and the migration rates of the hosts to understand more about louse movement.

CHAPTER 2
EVOLUTIONARY RELATIONSHIPS OF *CANDIDATUS* RIESIA SPP., ENDOSYMBIOTIC *ENTEROBACTERIACEAE* LIVING WITHIN HEMATOPHAGOUS PRIMATE LICE

Introduction

Insects in general are an incredibly successful and diverse group. Part of their success is undoubtedly due to mutualistic primary endosymbiotic bacteria, which have enabled insects to radiate into niches that include nutrient poor diets (Buchner 1965, Douglas 1989, Perotti et al. 2006) such as eating wood (*e.g.*, termites), plant sap (*e.g.*, aphids) and blood (*e.g.*, sucking lice in the Anoplura). Primary endosymbiotic bacteria (as opposed to secondary endosymbionts[S-endosymbionts]) are normally maintained inside specialized host cells called mycetomes (Douglass 1989, Perotti et al. 2006), usually exhibit nucleotide A+T bias greater than 50%, and show elevated sequence evolution with respect to their free-living counterparts (Brynnel et al. 1998, Wernegreen and Moran 1999, Woolfit and Bromham 2003). These endosymbionts are required for host survival and they provide nutrients that are not available in the insect's specialized diet (Buchner 1965). Most primary endosymbionts leave their mycetomes to migrate to the ovaries so that they may be incorporated into developing eggs (transovarial transmission) and thus be passed onto the next host generation (Douglas 1989, Perotti et al. 2006) leading to long-term, shared coevolutionary histories between the insects and their symbionts.

It is estimated that there are 14,000 species of haematophagous insects (Adams 1999, Lehane 2005) but only a few P-endosymbionts have been described from these insects (e.g., *Wigglesworthia*). Because blood is a nutrient poor diet, all sucking lice likely have some form of endosymbiont and many are reported to have obligate primary

endosymbionts based on microscopic observation. The P-endosymbiont in the human head and body louse (*Pediculus humanus*) was first seen over 340 years ago (Hooke 1665) with some of the very first microscopes. It has a complex migration (Buchner 1965) that involves four different mycetomes and two extracellular migrations as it moves to the eggs in the adult female lice (Perotti et al. 2007). This migration was first observed by Ries (Ries 1931) and later shown by scanning and transmission electron microscopy by Eberle and Mclean (1983). *Wolbachia* is the only other bacterium that has been found among human lice (Perotti et al., 2004, Kyei-Poku et al. 2005,). The complex migration associated with transovarial transmission stands as potential evidence of the importance of the relationship between the P-endosymbiont and the lice. If the mycetome is removed from a young female louse, she dies after only a few days and her eggs are deformed (Aschner and Ries 1933). Furthermore, if the bacteria are removed from the eggs directly, the larvae only survive a few days (Aschner 1934). Puchta (1955) demonstrated that lice without P-endosymbionts were able to survive if their diet was supplemented with nicotinic acid, pantothenic acid, and beta-biotin, suggesting a basis for a mutualistic long-term relationship.

Humans have three types of lice, head and body lice (*Pediculus humanus*), which are currently classified as two distinct subspecies (*Pediculus humanus capitis* and *Pediculus humanus humanus*) and pubic lice (*Pthirus pubis*) (Durden and Musser 1994). Body lice are known to transmit three diseases, louse-borne epidemic typhus (LBET), relapsing fever, and trench fever (Buxton 1946). Although head lice *can* transmit LBET in a laboratory setting (Goldberger and Anderson 1912, Robinson et al. 2003), there has never been evidence of LBET transmission by head lice in nature.

Currently, it is not known why one subspecies of *P. humanus* can transmit three deadly bacterial agents while the other subspecies, epidemic in school children, effectively cannot. Although the secondary endosymbiont of Tsetse flies (*Sodalis glossinidius*) has been shown not to affect the ability of its host to transmit *Trypanosoma congolense* (Geiger et al. 2005), it is possible that there are differences in the endosymbiotic bacteria of head and body lice and that these differences may reinforce patterns of disease transmission in human lice.

Primate lice show a history of cospeciation with their hosts, and have been used effectively to infer human evolutionary history (Kittler et al. 2003, Yong et al. 2003, Reed et al. 2004, Leo and Barker 2005). Most of this work has relied upon mtDNA from the lice and to a lesser extent on nuclear markers. If a single lineage of endosymbiont were found among primate lice, the endosymbiont could serve as still another marker of human and primate evolutionary history. In addition, this new three-tiered assemblage of primates, lice, and endosymbiotic bacteria would yield a new system in which to study relative and absolute rates of evolution in three disparate lineages (vertebrates, insects, and bacteria).

In this study, we describe the molecular characterization of the P-endosymbiont of primate lice including head and body lice (*Pediculus humanus*) characterized by Sasaki-Fukatsu et al. (2006), and present new data from chimpanzee lice (*Pediculus schaeffi*) and human pubic lice (*Pthirus pubis*). We use the full-cycle rRNA approach including comparative 16S rRNA gene analysis and the detection of endosymbionts within the host cell by means of fluorescent in situ hybridization using specific 16S rRNA-targeted oligonucleotide probes. By sampling across a phylogenetically diverse assemblage of

louse species, we expect to capture a greater percentage of the diversity in this lineage of P-endosymbiont.

Methods

Specimen Preparation

Specimens of human head lice (*Pediculus humanus capitis*) were collected from patients in West Palm Beach, Florida and from school children in La Rioja, Argentina (Perotti et al. 2004). Specimens of body lice (*Pediculus humanus humanus*) were acquired from the rabbit-adapted strain held at the insect control and research lab in Maryland, USA and in Cambridge, United Kingdom. Chimpanzee lice were collected in Uganda, and samples of *Pthirus pubis* were collected in Utah. All lice were surface-sterilized using either a lysis-sodicum-dodecyl sulfate buffer as described in Reed and Hafner (2002) or with 0.7 % sodium chloride, 0.05 % Triton X-100 for 20 s in an ultrasonic bath, then rinsed with sterile distilled water to remove any surface contamination. The lice were rehydrated in PBTA phosphate buffer with Triton X-100 plus sodium azide (PBTA) in three consecutive steps: 30, 60 and 100% PBTA. Whole mycetomes were dissected manually under a Leica stereoscope with a magnification of X100 using tungsten tips and special carbon steel blades. A maximum dissecting resolution of 10-15 µm was obtained. The dissected bodies were fixed inside 1.5 ml tubes for DNA extraction using the DNA extraction methods described in Reed and Hafner (2002).

Bacterial Diversity in *Pediculus Humanus Capitis* Using Molecular Methods

Universal bacterial 16S rDNA primer 27F (5'- GAG TTT GAT CCT GGC TCA G-3') was used with either 1492R (5'- CAC GGA TCC TAC GGG TAC CTT GTT ACG ACT T -3') or 1525R (5'-AGA AAG GAG GTG ATC CAG CC-3') to amplify 16S rDNA

sequences. PCR was performed on the isolated DNA using standard reaction conditions with 10 ng of template DNA, 300 nM of each primer, 200 µM of each dideoxy nucleotide triphosphate, 2.5 mM $MgCl_2$ and 0.02 U of Taq DNA polymerase per µl of reaction mix. Cycling conditions consisted of an initial denaturation step (95°C, 10 min) followed by 30 cycles of amplification involving denaturation (95°C, 1 min), annealing (50°C, 1 min) and extension (65°C, 1 min), followed by a final extension step at 65°C for 10 min. The PCR product was analyzed by gel electrophoresis. The 1.5-kbp 16S rDNA PCR product was purified with Exo-SAP-IT (USB Corp) - as prescribed by the manufacturer. The PCR product was then cloned into the pTOPO 4.0 vector (Invitrogen) according to the supplier's instructions to generate clone libraries. Recombinant clones were sequenced to completion at the University of Florida sequencing facility and in the Bangor lab, United Kingdom, using vector-specific primers and internal sequencing primers that were designed as sequence information became available. The computer program Sequencher® v. 4.1 (Gene Codes Corporation, Ann Arbor, MI) was used to join contiguous 16S rDNA fragments into a single consensus sequence.

Targeted Sequencing of Endosymbionts from Additional Taxa of Lice

The 1525bp 16S rRNA sequence for the endosymbiont of *P. humanus capitis* was aligned to other 16S sequences in the ARB database (Ludwig et al., 2004). A new PCR primer specific to the endosymbiont sequence was then created with the intention of excluding contaminant sequences coamplified with the endosymbiont target. The single specific primer, in concert with the general Eubacterial primer, preferentially amplifies the endosymbiont from whole-insect preparations. The 1525R primer was paired with

the specific primer 461F (5'-ACA GAA GAA GCA CCG GCT AA-3') to produce a 1,200bp fragment of the 16S rRNA gene. These primers were used to amplify, clone, and sequence the endosymbiont from three additional taxa of body lice (*P. humanus humanus*), chimpanzee lice (*Pediculus schaeffi*), and human pubic lice (*Pthirus pubis*). Genbank accession numbers are (EF110569—EF110574).

Fluorescent In-Situ Hybridization

Specimens were washed as for dissections and selected under the microscope on a slide with a drop of water and immediately fixed with ethanol-glacial acetic acid (3:1) for 2 h. Preparations were kept overnight in a 1:1 mixture of xylene and ethanol (1:1) at 4°C, were then transferred to xylene:ethanol (1:2) for 30 minutes and to ethanol for 30 minutes, washed with -20°C 80% acetone for 20 minutes and dehydrated in ethanol. Specimens were rehydrated using an ethanol/PBTA mixture with ratios of 2:1, 1:1, and 1:2 for 20 minutes each with a final rehydration of only PBTA for 30 minutes, where specimens remained at 4°C until use. Before hybridization PBTA samples were incubated with PBTA-hydration buffer (HB) at a ratio of 1:1 (Hybridization Buffer: TRIS-HCl 0.02 M, sodium chloride 0.09 M, SDS 0.01%, formamide 35%, Denhardt's solution 15%) for 20 minutes followed by only hybridization buffer (around 500µl/tube), sonicated in an ultrasonic bath for 20 seconds and then probes and corresponding helpers added (final concentration of 100 pmol each). Samples were incubated at 47 °C in total darkness for 16 h, then washed for 1 h with the same HB without probe/helpers at 47°C, then changed to HB:PBTA, 1:1, at room temperature and finally to PBTA. Samples were mounted with PBTA:glycerol mounting medium.

In order to ensure that the 16S rRNA sequences that we retrieved were from the same mycetome-bound bacteria described by Ries (1931), we created a species-

specific probe using the ARB database in combination with known probes and helpers of bright fluorescence (Fuchs et al. 1998). The endosymbiont-specific probe 2-Sd-Cy5 (5'-Cy5-GAG ATT GTT GCC TAG GTG-3'), which does not match any published sequence; and the 3 helpers, Helper 1-Sd (5'-ACC TCA CCT ACT AGC TAA TCT C-3'); Helper 2-Sd (5'-GTA TGG GCT CAT CTA AAG-3'); and Helper 3-Sd (5'-TTT AGG TAG ATY CCC ATA T-3') were based on the endosymbiont sequence obtained from human head louse endosymbiont. No-probe and competition suppression controls using excess unlabeled probes were performed. Fluorescent *in situ* hybridization was conducted on whole mount specimens and on 14-μm thick serial sections of whole individuals using the endosymbiont-specific probe. Samples were analysed with a Zeiss LSM510 confocal microscope with Coherent Multiphoton laser.

Phylogenetic Analysis

Nearly complete 16S rDNA sequences were obtained and added to the ARB rRNA sequence database, which contained over 16,000 homologous small-subunit rRNA primary structures. Multiple sequence alignment was achieved using the ARB automated alignment tool (program available at http://www.mikro.biologie.tu-muenchen.de), with modification by eye. Comparative sequence analysis revealed that the 16S rRNA genes of the endosymbionts were novel and showed highest sequence similarities with members of the *Enterobacteriaceae.* Therefore, we downloaded an aligned dataset of additional taxa from the ARB database and GenBank in order to perform more complete phylogenetic analyses (alignment available in TreeBase [www.trebase.org], no. SN3132). We used the computer program ModelTest (Posada and Crandall 1998) as a guide to determine the best-fit maximum likelihood (ML) model

as described by Cunningham et al. (1998). ModelTest examines maximum likelihood models ranging from simple to complex. This method increases the number of parameters in the ML model incrementally until the addition of a new parameter no longer increases significantly the fit between the model and the data. ModelTest calculated likelihood scores for 56 nested ML models and used hierarchical likelihood ratio tests (LRTs) to determine the best-fit model. We incorporated the best-fit model of nucleotide evolution in ML heuristic searches in Paup* (Swofford 2006) and in Bayesian searches in MrBayes (Hulsenbeck and Ronquist 2001) using the maximum likelihood optimality criterion. Multiple outgroups were chosen from phylogenetic studies of *Enterobacteriaceae* especially endosymbiotic taxa, and the phylogenetic tree was rooted on the most divergent outgroup taxon, the alpha proteobacterium *Wolbachia pipientis*. MrBayes was run for the 43-taxon dataset for 10 million generations. Burn-in was achieved within the first 100,000 generations; therefore to be conservative, our posterior probabilities are based on the last 9.8 million generations (phylogenetic trees available in TreeBase, no. M3012).

Results

Phylogenetic Analysis and Taxonomic Position of Louse Endosymbiont

DNA amplification and sequencing with the Eubacterial primers 27f, 1492r and 1525r led to the nearly complete sequencing of the 16S rRNA for two head lice and two body louse (ca. 1,520bp). The design of species-specific primers led to the sequencing of shorter fragments from the 16S rRNA gene ranging from 960bp to 1212bp for the remaining taxa. Pairwise sequence divergences among the four individuals of *P. humanus* ranged from 0.0% to 0.3% (Table 2-1). The two head lice were identical to each other in sequence as were the two body lice sequences. The head lice differed

from the body lice by 0.3% (GTR+I+ Γ model), resulting from five fixed differences between the two types of lice.

BLAST searches of Genbank showed that the *Pediculus humanus* endosymbiont was nearly 100% identical to the sequences of "*Candidatus* Riesia pediculicola" found by Sasaki-Fukatsu et al. (2006). BLAST searches also indicated that our sequences were very similar to those of the S-endosymbiont "*Candidatus* Arsenophonus insecticola", which was supported in the phylogenetic analysis as well (Figure 2-1). Phylogenetic analysis based on the 16S rDNA locus revealed that the symbionts from human head and body lice, chimpanzee lice, and human pubic lice were monophyletic with 100% support from the Bayesian posterior probabilities. The ML analysis produced a topology that agreed fully with the Bayesian analysis presented in Figure 2-1. The endosymbiont sequences obtained from *P. humanus* contain 139 substitutional differences when compared to "*Candidatus* Arsenophonus insecticola" (GenBank accession no. DQ115536), and they share 89% sequence identity over 1,260bp. The sequence from the P-endosymbiont of all lice surveyed were greater than 50% A+T, which is typical of true primary endosymbionts (Woolfit and Bromham 2003; Table 2-1). The endosymbiont sequences from two genera and three species of primate lice form a monophyletic assemblage with an average percent sequence divergences of 4.8% within the clade (Table 2-1).

Fluorescent In-Situ Hybridization

The ARB database was used to generate a species-specific oligonucleotide probe, which was 100% identical to the endosymbiont from *P. humanus* but no other taxa in the database or Genbank. *In situ* hybridization confirmed the specificity of this probe in sectioned and whole lice producing strong signal within the clearly visible stomach disc

of *P. humanus* (Figure 2-2). The *in situ* hybridization confirmed that the bacterium from which we recovered 16S rRNA sequence is indeed the bacterium found in the mycetomes, which have been known and visualized since the 1,600s.

Discussion

Overview

Bayesian phylogenetic analysis and BLAST searches demonstrated that this new lineage of *Enterobacteriaceae* is closely related to a lineage of S-endosymbionts ("*Candidatus* Arsenophonus spp."). "*Candidatus* Arsenophonus" endosymbionts have been found living within Hippoboscid flies (Dale et al. 2006), commonly called louse flies. Louse flies parasitize a wide range of birds and mammals, and are known to physically carry lice as hitchhikers among vertebrate host individuals (called phoresy). The lice, which are particularly bad dispersers having no wings of their own, use the flies as a means of dispersal. While phoresy itself does not involve any humoral interaction between passenger (phoront) and carrier, in one case, a phoretic mite (*Macrocheles subbadius*) has been shown to feed on the hemolymph of its transporting host (*Drosophila nigrospiracula)* (Polak 1996). The close association between the endosymbionts of lice and louse flies might be explained by horizontal transmission, which appears to be more common among microorganisms than previously thought (e.g., Moran and Dunbar 2006, Saito and Bjornson 2006). One might presume that the P-endosymbiont in lice is derived from the S-endosymbiont (i.e., increasing specialization over evolutionary time), however this presents a hypothesis that can be tested directly in future studies. Our phylogenetic analysis shows that the endosymbionts of primate lice are distinctly different from each other and their closest relatives, and yet they represent a strongly supported monophyletic clade.

Nomenclature

These endosymbiotic bacteria have not been completely characterized or grown in pure culture, although they have been known and visualized since the early 1600s (Hooke 1665, Perotti et al. 2007). Our endosymbiont sequences from *Pediculus humanus* were nearly identical to sequences from "*Candidatus* Riesia pediculicola", but the additional primate louse species had closely related but distinctly different endosymbionts. The endosymbionts from *P. humanus* were more than 3% divergent from those from the chimpanzee louse *P. schaeffi*, and endosymbionts from the human pubic louse, *Pthirus pubis*, were more than 10% divergent from both *P. humanus* and *P. schaeffi*. Therefore, following the recommendations for uncultured microorganisms (Murray and Schleifer 1994, Murray and Stackerbrandt 1995), we suggest that endosymbionts of *P. schaeffi* and *Pthirus pubis* be recognized as distinct species within the genus "*Candidatus* Riesia".

The current taxonomic name "*Candidatus* Riesia pediculicola" is ambiguous as to the specific host species with which it is associated. The specific epithet of "*Candidatus* Riesia pediculicola" refers only to the genus of louse host, and yet there are two species of *Pediculus* (*P. schaeffi* and *P. humanus*) which have closely related P-endosymbionts. Therefore in order to reduce further confusion, we propose to use the concatenated scientific names of the louse host for the specific epithet of the endosymbiont. We propose the name "*Candidatus* Riesia pediculischaeffi" for the primary endosymbiont of the chimpanzee louse (*Pediculus schaeffi*). Similarly, we propose the name "*Candidatus* Riesia pthiripubis" for the primary endosymbiont of the human pubic louse (*Pthirus pubis*). By latinizing the full binomial, we can accommodate additional species within a genus, such the endosymbiont from the gorilla louse, *Pthirus gorillae*, when and

if it is characterized. Although the names are verbose, we find this nomenclature appropriate and beneficial.

Coevolution and Rates of Substitution

The two 16S sequences from the endosymbionts of human head lice differ from the two endosymbiont sequences from body lice by five fixed differences (sites 218, 492, 877, 950, and 1203) in our aligned matrix (available in TreeBase). Although this appears to suggest distinction between the endosymbionts of head and body lice, recent additional sequence data acquired in the lab of DLR suggests this is an artifact of sampling and is not indicative of real phylogenetic differences between the endosymbionts of head and body lice. The head and body lice of modern humans have been reasonably well sampled, and their mtDNA can be divided into three distantly related haplotypes (Kittler et al. 2003, Leo and Barker 2005, Yong et al. 2003). The first mtDNA haplotype (Type A) is worldwide in distribution, and is the most common in both the head and body louse morphotypes. The second mtDNA haplotype (Type B) has been found in the New World, Europe, and Australia and exists only in head lice. The third mtDNA haplotype (Type C) has only been found among head lice from Nepal and Ethiopia. Unpublished data from our lab suggest that the endosymbionts of *P. humanus* mirror the host mtDNA haplotypes for *P. humanus*. This is perhaps not surprising given that both the endosymbionts and mtDNA are maternally inherited in these lice.

The phylogenetic relationships of the endosymbionts of primate lice (both species of *Pediculus* plus *Pthirus*) are identical in topology to the phylogenies of their hosts (Reed et al. 2004) suggesting a long-term coevolutionary history between primate lice and their endosymbionts. Reed et al. (2004) demonstrated that *P. humanus* and *P. schaeffi* diverged from one another ca. 5.6 million years ago based on mtDNA, which

coincides precisely with the estimated divergence of their human and chimpanzee hosts based on mtDNA. The average percent sequence divergence (model corrected) between the endosymbionts of *P. humanus* and *P. schaeffi* is 3.75% (Table 2-1). We can use the 5.6 million year divergence date as a calibration point to determine the rate of nucleotide substitution in these endosymbionts, which equates to an absolute rate of 0.67% per million years. This rate is 15 to 30 times faster than the rates of 16S rRNA evolution estimated previously for *Buchnera*, the P-endosymbiont in Aphids (1 to 2% per 50 million years) (Moran et al. 1995). A more thorough comparison of the "*Candidatus* Riesia" lineage with *Buchnera* will determine whether our rate calculations are indeed correct. It is conceivable that the same mechanisms that cause more mtDNA substitutions in lice than in aphids (Johnson et al. 2003) may have the same effect on the endosymbionts as well.

Because endosymbionts evolve more quickly than their hosts (Moran et al. 1995), they may record more nucleotide substitutions during recent (or rapid) evolutionary events than do their hosts. Because lice have strictly coevolved with their primate hosts, and provide clear resolution of both recent and rapid evolutionary events (e.g., the population expansion of humans out of Africa is also evident in louse mtDNA) they have been used as a fast evolving marker to examine different events in human evolutionary history (Kittler et al. 2003, Yong et al. 2003, Reed et al. 2004, Leo and Barker 2005). The evidence for coevolution between "*Candidatus* Riesia spp." and their primate louse hosts, along with the estimate of faster evolutionary rates in "*Candidatus* Riesia", leads us to conclude that this new lineage of endosymbiotic bacteria shows promise as another independent evolutionary marker of primate and human

evolutionary history. If faster-evolving markers can be found within the endosymbionts (markers much faster than the 16S rRNA), then very recent events in human evolutionary history (e.g., Peopling of the Americas) might be studied from the perspective of the endosymbiont of a human parasite. This new three-tiered assemblage of host, parasite, and endosymbiont is among the first tripartite assemblages of this type and will permit many new tests relating to their shared coevolutionary history.

Table 2-1. Pairwise differences for p-endosymbionts from human head and body lice *Pediculus humanus*. Abbreviations P. h.h. *Pediculus humanus humanus,* human body lice and P. h. h. *Pediculus humanus capitis*, the human head louse. Sequence divergences were caluclated using a GTR +I+G model of molecular evolution and compared to the most closely related taxon from Genbank ("*Candidatus* Arsenophonus insecticola", [GenBank accession no. DQ115536]).

	P. h. h. (USA)	P. h. h. (Wales)	P. h. c. (USA)	P. h. c. (Wales)	P. schaeffi	Pthirus pubis
P. h. humanus (USA, 50.9%)						
P. h. humanus (UK, 50.9%)	0.0000					
P. h. capitis (USA, 50.7%)	0.0031	0.0033				
P. h. capitis (Arg., 50.8%)	0.0032	0.0033	0.0000			
P. schaeffi (50.4%)	0.0395	0.0378	0.0359	0.0366		
Pthirus pubis (52.5%)	0.1239	0.1127	0.1085	0.1110	0.1055	
Arsenophonus (52.2%)	0.1315	0.1557	0.1533	0.1540	0.1174	0.1342

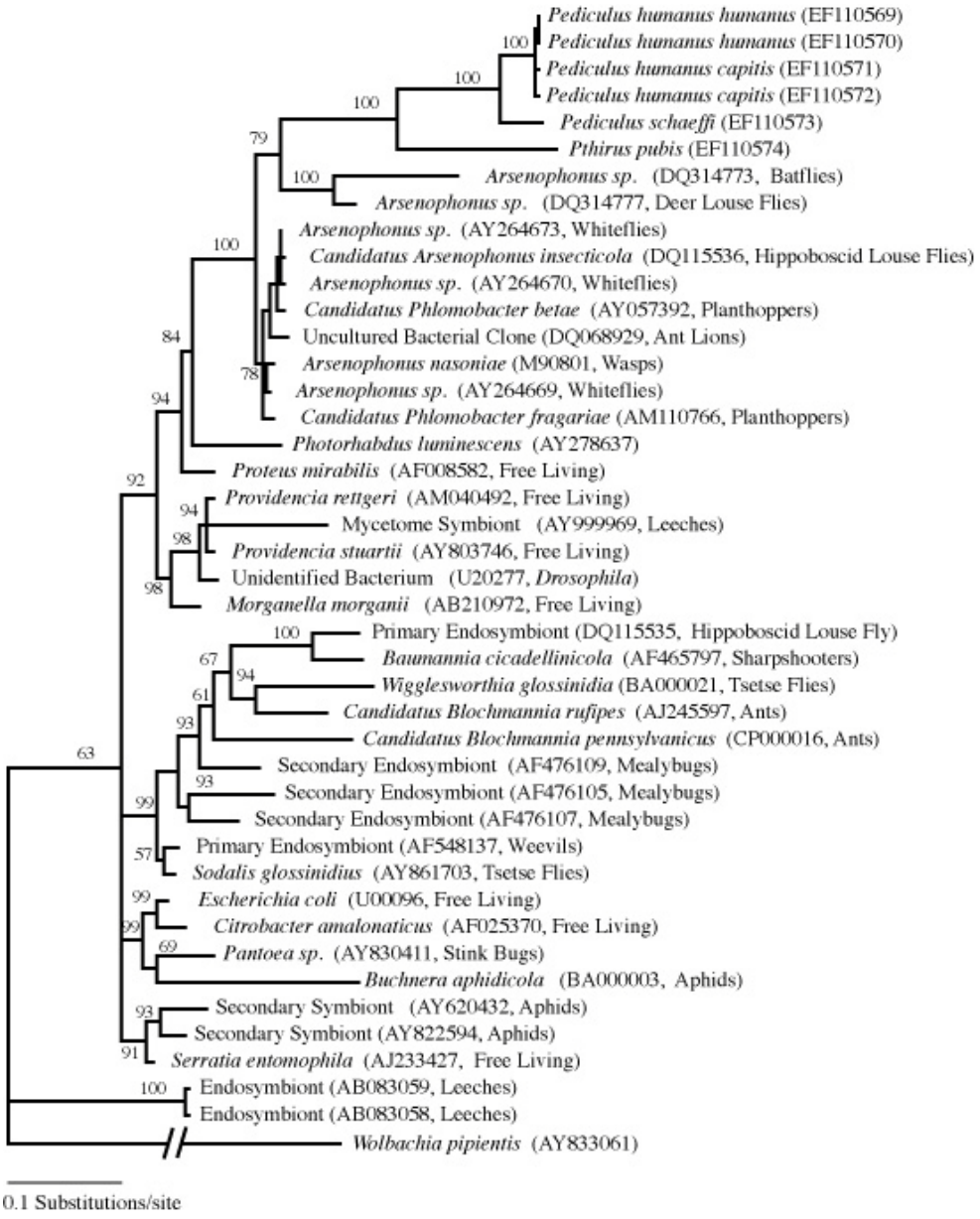

Figure 2-1. Phylogenetic tree of endosymbiont from sucking lice (Anoplura) and other bacteria. Tree is based on Bayesian phylogenetic analysis of 10 million generations using MrBayes. Posterior probabilities above 0.50 are shown above nodes. Endosymbionts of primate lice are monophyletic with a posterior probability of 1.00. GenBank accession numbers are given in parenthesies.

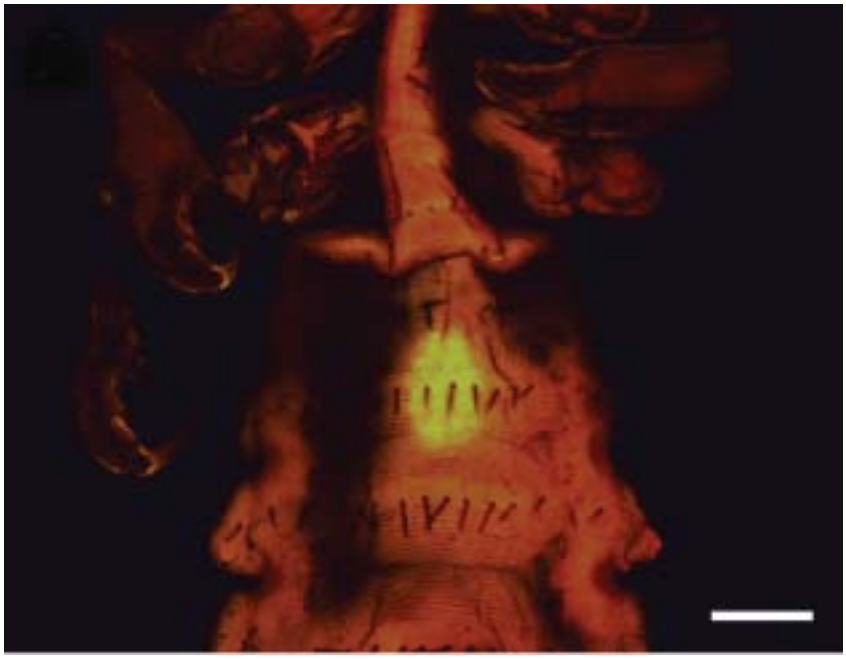

Figure 2-2. Fluorescent in situ hybridization microphotograph of thorax and abdomen (ventral view) of a second instar nymph of a head louse (*Pediculus humanus capitis*). This whole-insect mount was probed with the species-specific probe designed from the 16S rRNA gene endosymbiont sequence obtained from another individual of *P. humanus capitis*. Note the bacteria inside the mycetome in yellow (scale bar = 50 µm).

CHAPTER 3
MUTATIONAL MELTDOWN IN PRIMARY ENDOSYMBIONTS: SELECTION LIMITS MULLER'S RATCHET

Introduction

Primary endosymbiotic bacteria (p-endosymbionts) are thought to have enabled insects to become ecologically diverse by facilitating radiations into niches with nutrient–poor diets such as plant sap, wood, and vertebrate blood. P-endosymbionts live within specialized host organs called mycetomes and are transmitted transovarially (vertically) from mother to offspring (Douglass 1989). Some p-endosymbionts are required for host reproduction (Dedeine et al. 2003, Perotti et al., 2006,) whereas others provide essential services for their hosts such as light emission, or synthesis of amino acids, cofactors, and vitamins that are lacking in the host's specialized diet (Buchner 1965).

Because of their endosymbiotic lifestyle and strict vertical transmission, all p-endosymbionts share many characteristics such as small populations, reduced genomes, and AT bias (Moran 1996, Clark et al., 1999, Lutzoni and Pagel 1997, Gill and Moya 2004). P-endosymbionts also accrue slightly deleterious mutations at a faster rate than free-living bacteria (Moran 1996). This is thought to be due to genetic drift acting on already small populations that go through population bottlenecks at each host generation (O'fallon 2008). Furthermore, because p-endosymbionts are maternally transmitted, it is thought that recombination cannot occur between different strains (Moran 2007). The steady accumulation of these deleterious mutations is a process called Muller's ratchet (Muller 1964, Felsenstein 1974, Moran 1996).

Muller's ratchet states that in small populations, due to genetic drift, there is a chance that individuals with the fewest mutations will fail to reproduce (Muller 1964, Lynch and Gabriel 1990). When this happens, the ratchet clicks (Felsenstein 1974)

irreversibly increasing the overall mutational load of the population. As mutational load increases, the relative fitness decreases through reduced reproductive rate or reduced survivorship (Haldane 1937, Wallace 1987). If deleterious mutations continually get fixed over time, the p-endosymbiont may experience a mutational meltdown ultimately resulting in extinction (Lynch and Gabriel 1990).

It is thought that once a p-endosymbiont has deteriorated to the point of being nonfunctional, it may be replaced by another bacterium (Andersson and Kurland 1998). Evidence for p-endosymbiont replacement, however, is scarce, having only been found in a few species of aphids (Moran and Baumann 1994, Perez-Brocal et al. 2006), weevils (Lefevre et al., 2004), and more recently in sucking lice (Hypsa and Krizek 2007, Perotti et al. 2009). In contrast, some insect/p-endosymbiont assemblages have existed for hundreds of millions of years without evidence of p-endosymbiont replacement, suggesting that Muller's ratchet may slow or stop over time.

Several mechanisms have been proposed to explain how Muller's ratchet might slow or stop over time. These mechanisms include back mutations (Atwood et al., 1951), compensatory responses (Hurst and Mcvean 1996), and selection. The probability of a back mutation "correcting" each slightly deleterious mutation is so minimal compared to the probability of a forward mutation that it has been ignored in models of Muller's ratchet (Atwood et al. 1951). Compensatory responses have been suggested in the case of the GroEL protein. The GroEL protein mediates the folding of polypeptides, and it is found to be highly expressed in *Buchnera* and other p-endosymbionts. Up-regulation of the GroEL protein may reduce the effects of other slightly deleterious mutations that may change the folding of important proteins (Moran

1996). However, little is known about the overall effect of compensatory responses on Muller's ratchet. Selection may be acting to slow or stop Muller's ratchet through long term bottlenecks which cause the variance in fitness to be increased among hosts for selection to act upon (Bergstrom and Pritchard 1998), and through epistatic interactions between slightly deleterious mutations (Charlesworth et al. 1993, Kondraschov 1994) where the effect on fitness of the p-endosymbiont does not increase linearly with each mutation that becomes fixed in the population. Epistatic interactions may make the ratchet slow down but not necessarily stop (Butcher 1995). In this study, we are specifically interested in determining if selection is acting to reduce the number of slightly deleterious mutations that become fixed in the population, thus slowing the process of Muller's ratchet in some insect/p-endosymbiont assemblages.

Early studies of non-synonymous nucleotide substitution rates suggested that selection was weak in p-endosymbionts (Moran 1996, Moran and Baumann 2000). Recent studies of the genomes of the p-endosymbiont of aphids (*Buchnera*), however, show that selection may play a role in slowing genome degradation (i.e., gene loss) and AT bias. For example, Tamas et al. (2002) found long-term genomic stasis in two *Buchnera* genomes that diverged around 50–70 Mya. They concluded that gene loss must have occurred early in the association between *Buchnera* and its host, only to stabilize later due to selective constraints. Early and rapid gene loss was also found in another lineage of *Buchnera* with divergence dates of 80-150 My (van Ham et al. 2003, Klasson and Anderson 2004). Clark et al. (1999) suggested that selection might also reduce AT bias over time in *Buchnera*, and therefore slow the speed of Muller's ratchet.

Selection could decrease the rate at which slightly deleterious mutations become fixed in the population, especially as the p-endosymbiont/insect association ages. As the mutational load increases in p-endosymbionts, selection may act to remove individuals with the highest mutational loads (i.e., the least functional individuals). This would then slow the rate of fixation of slightly deleterious mutations, which would result in a reduction in the overall nucleotide substitution rate.

Early estimates of nucleotide substitution rate in p-endosymbionts consistently averaged 1-2% per 50 My (Ochman et al. 1999). However, recent studies have documented much faster evolving p-endosymbionts (Degnan et al. 2004, Allen et al. 2007). The fastest rate reported to date is 33.5% per 50 My for *Candidatus* Riesia (hereafter *Riesia*), the p-endosymbiont of primate sucking lice (Anoplura: Pediculidae and Pthiridae) (Allen et al. 2007). The p-endosymbionts with the highest nucleotide substitution rates appear to be among the youngest insect/p-endosymbiont associations, which suggest that rates may vary in relation to the age of the association. However, the age of the association of *Riesia* with its host is unknown. Therefore, we first determine the age of the association of *Riesia* with its host and calculate more rigorously the rate of molecular evolution for *Riesia*.

Additionally, we estimate the nucleotide substitution rates from the *16S* ribosomal DNA gene (*16S rDNA*) for a diverse assemblage of p-endosymbionts to test the prediction that selection reduces the effect of Muller's ratchet over time. If selection slows Muller's ratchet over time, then we should observe an inverse relationship between nucleotide substitution rates and the age of the insect/p-endosymbiont assemblage. A decline in substitution rates would be consistent with an increase in

selection over time. The mutational meltdown model predicts that given no opposing force, p-endosymbionts should steadily accrue slightly deleterious mutations until extinction. An increase in selection over time might allow p-endosymbionts to stave off extinction, which would explain the existence of ancient insect/p-endosymbiont associations.

Materials and Methods

Age of the Riesia/Louse Assemblage

The oldest split found within *Riesia* dates to 12.95 Ma and occurs between p-endosymbionts associated with the louse genera *Pediculus* and *Pthirus* (Allen et al. 2007). However, some of the oldest divergences among related lice (those of the genus *Pedicinus*; Anoplura: Pedicinidae) date back to the split between their Anthropoid and Cercopithecoid primate hosts, ca. 25-30 Ma. At present it is not known whether lice of the genus *Pedicinus* carry the *Riesia* lineage of p-endosymbionts. Therefore, we have molecularly characterized the endosymbiont of *Pedicinus* to evaluate the age of the *Riesia*/louse association. If the p-endosymbiont in the louse genus *Pedicinus* does not belong to the *Riesia* lineage, then the *Riesia*/louse association is between 12.95 and 25 My.

Specimen collection and DNA sequencing

To determine the age of the *Riesia*/louse association, specimens of *Pedicinus badii* were collected from Red Colobus monkeys (*Procolobus rufomitratus*) from Kibale National Park in Uganda. Three human head louse specimens (*Pediculus humanus capitis*) and a single body louse specimen (*Pediculus humanus humanus*) were collected from individuals in West Palm Beach, Florida, USA, and the rabbit-adapted strain held at the Insect Control and Research Lab in Maryland, USA, respectively, to

determine the absolute rate of nucleotide substitution in *Riesia*. Whole lice were washed twice with 400µl saline EDTA, 15µl of 20% SDS and 5µl lysozyme to remove any external bacteria. The sample then was crushed and genomic DNA was isolated using the DNeasy Tissue Kit (QIAGEN Inc., Valencia, California). PCR amplification of the endosymbiont *16S rDNA* gene (1.5-kbp) was performed in 25µl reactions with primers 27F (5' – AGA GTT TGA TCC TGG CTC AG – 3') and 1392R (5' – CAC GGA TCC ACG GGC GGT GTG TRC – 3') for the *Pedicinus* endosymbiont, and the *Riesia* specific primer 461F (5' – ACA GAA GAA GCA CCG GCT AA – 3') and general reverse primer 1525R (5' – AGA AAG GAG GTG ATC CAG CC – 3') for *Pediculus* endosymbionts. Each amplification was performed using standard reaction conditions with 10ng of template DNA, 300nM of each primer, 200µM of each dNTP, 2.5mM $MgCl_2$ and 0.02 U of Taq DNA polymerase (Promega, Madison, Wisc.) per µl of reaction mix. Cycling conditions consisted of an initial denaturation step (94°C, 10 min), 30 cycles of amplification involving denaturation (94°C, 1 min), annealing (50-52°C, 1 min) and extension (65°C, 1 min), and a final extension step at 65°C for 10 min. The *16S rDNA* PCR product was purified with ExoSAP-IT (USB Corporation) and then cloned into the pTOPO 4.0 vector (Invitrogen). Recombinant clones were sequenced in both directions at the University of Florida sequencing facility using vector-specific primers and internal sequencing primers as in Reed and Hafner (2002). Sequences were edited using Sequencher Version 4.1 (Gene Codes Corporation, Ann Arbor, Michigan) and deposited in GenBank (Accession numbers: EU827259-EU827263 *Reisia pediculicola* from *Pediculus humanus humanus*, three sequences of *Riesia pediculicola* from *Pediculus humanus capitis* and the primary-endosymbiont from *Pedicinus badii*, respectively).

Phylogenetic analysis

Phylogenetic analyses were used to determine the placement of the *Pedicinus* p-endosymbiont with respect to other known bacteria, and to estimate the age of the *Riesia*/louse association. The *16S rDNA* sequence of the *Pedicinus* p-endosymbiont obtained above was compared to 32 bacterial *16S rDNA* sequences downloaded from GenBank which included louse p-endosymbionts, insect p-endosymbionts, free-living *Escherichia coli*, and sequences with the highest sequence similarity to the *Pedicinus* and other louse p-endosymbionts obtained form GenBank BLAST searches (Table 3-2). All sequences were aligned using Clustal X (Thompson et al. 1997), then manually adjusted by eye.

Modeltest v. 3.7 (Posada and Crandall 1998) was used to determine a model of nucleotide evolution according to an Akaike Information Criterion (GTR+I+G; (Hulsenbeck and Rannala 1997, Posada and Buckley 2004). This best-fit model was used in Maximum Likelihood (ML) and Bayesian phylogenetic analyses performed in PAUP*4.0b10 and MrBayes 3.12 (Hulsenbeck and Ronquist 2001, Swofford 2002), respectively.

For the ML analyses, full heuristic ML and bootstrap (200 pseudoreplicates) searches were conducted with 10 random addition replicates and tree bisection-reconnection branch swapping using the best-fit model in PAUP* 4.0b10 (Swofford 2002). In the Bayesian analyses, model parameters were treated as unknown variables with uniform priors and were estimated as part of the analysis. Bayesian analyses were initiated with random starting trees, run with four incrementally heated chains (Metropolis-coupled Markov chain Monte Carlo; (Hulsenbeck and Ronquist 2001) for 10 million generations, and sampled at intervals of 1000 generations. Two independent

Bayesian analyses were run to avoid entrapment on local optima. Stationarity was assessed by plotting the log-likelihood scores of sample points against generation, and a conservative burn-in period of 25% was discarded. The retained equilibrium samples were used to generate a 50% majority rule consensus tree with the percentage of samples recovering any particular clade representing that clade's posterior probability (Huelsenbeck and Ronquist 2001).

Alternative phylogenetic hypotheses were compared statistically using the Kishino-Hasegawa (KH) and the Shimodaira-Hasegawa (SH) tests as implemented in PAUP*4.0b10 (MP and ML analyses using RELL optimization and 1,000 bootstrap replicates; (Kishino and Hasegawa 1989, Shimodaira and Hasegawa 1999, Goldman et al. 2000). Suboptimal trees from the Bayesian analyses also were examined to assess alternative phylogenetic hypotheses. The frequency of the Markov chain Monte Carlo trees in agreement with an alternative hypothesis equals the probability of that alternative hypothesis being correct (Ihlen and Ekman 2002). The probability of trees agreeing with alternative subfamily hypotheses was calculated by applying constraint-based filter trees implemented in PAUP*4.0b10 (Ihlen and Ekman 2002, Swofford 2002,).

Absolute rates of nucleotide evolution in *Riesia*

To determine the absolute rate of nucleotide substitution in *Riesia*, the *16S rDNA Riesia* sequences obtained above were aligned with p-endosymbiont sequences of human head lice (*R. pediculicola*; GenBank Accession Numbers AB263105, EF110570, and EF110571), human body lice (*R. pediculicola*; EF110569, EF110572, and AB236101), chimpanzee lice (*R. pediculischaeffi*; EF110573), and human pubic lice (*R. pthiripubis*; EF110574). Sequences were aligned using Clustal X (Thompson et al.

1997) and manually adjusted using MacClade v. 4.06 (Maddison and Maddison 2000). These closely related sequences were easily aligned by eye with no ambiguity as to positional homology. Modeltest v. 3.7 (Posada and Crandall 1998) was used to determine a model of nucleotide evolution (GTR+G) for the *Riesia 16S rDNA* data as described above. A branch and bound ML analysis with a subsequent bootstrap analysis (200 replicates) was conducted using the best-fit model in PAUP* 4.0b10 (Swofford 2002).

Reed et al. (2007) estimated divergence dates in the phylogenetic tree of primate lice, and estimated the split between the genera *Pediculus* and *Pthirus* to be 9.42–17.38 Ma ago. Because this node in the louse tree has a corresponding node of cospeciation in the endosymbiont tree (Allen et al. 2007), we are able to calculate an absolute rate of nucleotide substitution within *Riesia*, using the calibration range of 9.42-17.38 Ma for the split between *Pediculus* and *Pthirus* endosymbionts. Divergence times were estimated using penalized likelihood (TN algorithm) in the program r8s (Sanderson 2003). A smoothing parameter of 0.32 was determined using the cross-validation procedure.

Substitution Rates Among Host/Endosymbiont Lineages

To determine whether the age of the host/endosymbiont association correlates with nucleotide substitution rate, we retrieved data from the literature of insect/p-endosymbiont assemblages having both estimates of the age of the association (through fossil evidence) as well as either rates of p-endosymbiont nucleotide evolution for *16S rDNA* or pairwise sequence divergences. In the absence of pairwise sequence divergences for a particular assemblage, we estimated these values by examining the two most divergent sequences as in Ochman et al. (Ochman et al. 1999). The systems

examined included primate lice and *Riesia* (Allen et al. 2007), aphids and *Buchnera* (Moran et al. 1993), cockroach/termites and *Blattabacterium* (Bandi et al. 1995), whiteflies and *Portiera* (Thao and Baumann 2004) (date from Poinar (1992)), tse-tse flies and *Wigglesworthia* Askoy et al. 1995), Auchenorrhyncha (cicadas, hoppers and spittlebugs) and *Sulcia* (Moran et al. 2005), psyllids and *Carsonella* (Thao et al. 2000), weevils and *Nardonella*, weevils and the S-clade of p-endosymbionts (Lefevre 2006), and ants and *Blochmannia* (Degnan et al. 2004).

Rates of nucleotide substitution were plotted against the age of host/endosymbiont association. Because there is error in both the estimate of the rate of nucleotide evolution and the age of association, a reduced major axis regression was performed on the log-transformed data to better estimate the relationship between the age of the association and rate of nucleotide evolution in these systems.

Results

Age of *Riesia*/Louse Association

The age of the association between the fast-evolving p-endosymbiont *Riesia* and the primate sucking lice in which it lives was previously unknown. In order to estimate this age, we examined the p-endosymbiont from a closely related louse genus, *Pedicinus*. The p-endosymbiont from *Pedicinus badii* (a louse that parasitizes Old World monkeys) does not group with the anthropoid primate louse p-endosymbionts (the *Riesia* lineage) in our Maximum Likelihood or Bayesian (not shown) phylogenetic analyses (Figure 3-1). The Maximum Likelihood analysis groups the p-endosymbiont of *Pedicinus badii* at the base of a clade containing the p-endosymbionts *Wigglesworthia* and *Baumannia* (p-endosymbionts of tse-tse flies and leafhoppers, respectfully), some free-living bacteria, and the p-endosymbionts of distantly related sucking lice of rodents

(Figure 3-1). Bayesian phylogenetic trees were largely identical, and placed the p-endosymbiont of *Pedicinus badii* at the base of the same clade. Analyses constraining the *Pedicinus* p-endosymbiont to group with the *Riesia* lineage produced trees that were significantly worse than the best Maximum Likelihood tree according to the Kishino-Hasegawa ($p = 0.004$) and Shimodaira-Hasegawa ($p = 0.004$) tests. Furthermore, none of the suboptimal trees from the Bayesian analysis were consistent with this topological constraint ($p < 0.001$). We can therefore formally reject the hypothesis that the p-endosymbiont sequences from *Pedicinus badii* are sister to or embedded within the *Riesia* lineage. Because *Pedicinus* is the closest living relative of *Pediculus* and *Pthirus*, this phylogenetic analysis demonstrates that the age of the association between *Riesia* and primate lice has an upper bound at 25 My for the split between *Pedicinus* and *Pediculus* and *Pthirus*. Thus, the age of association between *Riesia* and their louse hosts is between 12.95 and 25 My, making this one of the youngest insect/p-endosymbiont assemblages known.

Absolute Rates

Using the 9.42–17.38 My split between *Pediculus* and *Pthirus* as a calibration date (Reed et al. 2004, Reed et al. 2007), we estimated the divergence time between *Riesia pediculicola* (human head and body louse p-endosymbionts) and *Riesia pediculischaeffi* (chimp louse p-endosymbionts) at 5.42 My, which is very close to the ages estimated for these lice and for their vertebrate hosts (Reed et al. 2004). We further estimate that the p-endosymbionts of the human head lice originated 0.90 My (Figure 3-2), which is similar to the estimate of 1.2 My for the lice (Reed et al. 2004). The pairwise sequence divergence for *Riesia* p-endosymbionts of *Pediculus* and *Pthirus* is 12.90% (GTR+I model), therefore the absolute rate of evolution of *Riesia* p-

endosymbionts is 0.0037-0.0684 substitutions per site per million years which translates to 18.56-34.24% per 50 My (Table 3-1).

Substitution Rates Among Host/Endosymbiont Lineages

When the rates of nucleotide substitution for *Riesia* are compared to other known insect/p-endosymbiont systems, we find that the rate of substitutions in *16S rDNA* decreases with age of association and levels off after 100 My (Figure 3-3). Although the majority of the systems are evolving at a rate similar to what was reported for *Buchnera* (1-2% per 50 MY), the younger systems are evolving much faster (3-34% per 50 MY; Figure 3-3). Reduced major axis regression of the log-transformed data indicates that 78% of the variation in rates of nucleotide evolution can be explained by the age of the association (Figure 3-4) and that the decrease in rates is exponential. The pairwise sequence divergence in *Riesia* calculated here (18.56–34.24% per 50 My) was corrected with a best-fit model of nucleotide substitution. Some previous studies did not use the best-fit evolutionary model to correct for multiple substitutions. Therefore, to test the impact of the substitution model, we also evaluated the same pairwise divergences using the Jukes-Cantor model. The Jukes-Cantor distances still provide a much faster rate of nucleotide substitution in *Riesia* (12.9-23.9% per 50 My), and it is important to note that this more simplistic model of molecular evolution underestimates the substitution rate by 30%.

Discussion

Overview

In this study, we find that there is considerable variation (15 to 30-fold) in the rate of p-endosymbiont nucleotide evolution for the *16S rDNA* gene. The association between anthropoid primate lice and their p-endosymbionts in the genus *Riesia*

represents one of the youngest insect/p-endosymbiont assemblages known to date (between 12.95 and 25 Ma), and compared to other insect/p-endosymbiont assemblages, *Riesia* is experiencing the highest rate of nucleotide substitution yet measured (18.56–34.24% per 50 My; Figures 3-3 and 3-4). Among all insect/p-endosymbiont assemblages examined, we find that 78% of the variation in nucleotide substitution rate can be explained by the age of the association (Figure 3-4). Higher rates of nucleotide substitution are associated with the youngest host/p-endosymbiont assemblages despite correcting for multiple substitutions (Figures 3-3 and 3-4). Nucleotide substitution rates decrease to approximately 1-2% per 50 My when the insect/p-endosymbiont assemblages reach approximately 100 My of age. These findings are consistent with the hypothesis that selection reduces the effect of Muller's ratchet over time.

An alternative explanation is that substitution rate variation is driven by variation in p-endosymbiont population size (i.e., smaller populations evolve faster than larger ones due to genetic drift). Estimates of effective population size are not available for the taxa used in this study, therefore this cannot be tested directly. It is likely that p-endosymbiont effective population size is governed largely by host effective population size, transmission dynamics, and other aspects of the p-endosymbiont/host relationship. Additional research, however is needed to directly test these hypotheses.

As proposed in the mutational meltdown model, selection is likely the force reducing the number of deleterious mutations that become fixed in p-endosymbiont populations. The slow but steady accumulation of deleterious mutations is predicted to impair the p-endosymbiont's ability to function if the process of Muller's ratchet goes

unchecked. Our data suggest that selection may steadily grow stronger in older assemblages and thereby slow the rate of Muller's ratchet by removing individuals with the highest mutational load. An increase in selection over time explains why some p-endosymbionts have ancient associations with their insect hosts and remain functional for hundreds of millions of years. However, younger assemblages do not always have a higher rate, especially among co-endosymbionts. Weevils (Insecta: Coleoptera) have two lineages of p-endosymbionts (termed the R- and S-clades) that are evolving at roughly the same rate even though R-endosymbionts have been associated with their hosts for 75 My longer (Lefevre et al. 2004). The rate differences between weevil p-endosymbionts, however, are minimal and fit well within the limits of other p-endosymbionts (Figure 3-4).

It has been predicted that selection plays a major role in slowing down or stopping Muller's ratchet. As slightly deleterious mutations go to fixation, they reduce the fitness of the host. If there were synergistic epistatic interactions between mutations we would expect an exponential increase in selection over time, which is consistent with our data.

Endosymbiosis

The importance of p-endosymbionts to insects, concerning their radiation into nutrient poor niches, cannot be overstated (Douglas 1989, Moran 2007, Perotti et al. 2009). Yet, very little is known about how bacteria become endosymbionts, although it is thought that they might originate from attenuated pathogens (Corsaro et al. 1999, Moran and Wernegreen 2000, Dale et al. 2001, Braig et al. 2009). Regardless of the mechanism, the basic requirements for becoming an endosymbiont are substantial. The endosymbiont must overcome many host physical, cellular, and molecular barriers for internalization (Ochman and Moran 2001), and a mechanism must develop for

transmission of the bacteria to the insect's offspring (Gil et al. 2004). Within the *Riesia* lineage alone, these bacteria undergo two extra-cellular migrations and are housed in no fewer than four distinct mycetomes (Perotti et al. 2007). From an evolutionary perspective this complex host/p-endosymbiont interaction seems highly specialized and the likelihood of repeated endosymbiont replacement over time is unknown.

If slightly deleterious mutations were to continue unabated in insect/p-endosymbiont associations, then we would expect to see a steady increase in the number of nucleotide substitutions over time, maintaining a high rate of molecular evolution. Instead we see a decline in the substitution rate (Figure 3-3). Our interpretation is that as the host/p-endosymbiont association ages, and the mutational load of p-endosymbionts increases, the role of selection increases and slows the rate of accumulation of slightly deleterious mutations. This is consistent with the studies of Tamas et al. (2002), van Ham et al. (2003) and Clark et al. (1999), who found that the rate of genome degradation and AT bias also decreases over time. Our findings are also consistent with that of Delmotte et al. (Delmotte et al. 2006), who found that the genes lost at the beginning of the association were those that were the least selectively constrained. We propose that there are selective constraints embodied in the process and maintenance of endosymbiosis that could mitigate the effects of Muller's ratchet in late-stage or well-established endosymbionts. Bergstrom and Pritchart (1998) suggested that long-term bottlenecks increase the selection pressure on deleterious mutations by increasing the variance in fitness among hosts. Therefore host-level selection may help to maintain the endosymbiosis over the long-term.

Although our data show that Muller's ratchet slows through time, Muller's ratchet may not cease to act entirely in p-endosymbionts, and three outcomes have been recorded. An endosymbiont may become so degraded that it effectively becomes an organelle such as *Carsonella*, the p-endosymbiont of psyllids (Nakabachi et al. 2006). *Carsonella* has been associated with its host for 100 to 250 My (Thao et al. 2000) and has a low rate of nucleotide evolution (Figure 3-3). It may be possible that *Carsonella* remains functional only because many of its genes have been transferred to the host genome and the products of these genes are shipped back to the symbiont (Moran 2007). Alternatively, the biological functions of an endosymbiont may be so reduced that a second endosymbiont is required. This is the case with the co-primary endosymbionts *Baumannia* and *Sulcia*, which have lost so many metabolic genes that by themselves they would not be viable or functional as endosymbionts (McCutcheon and Moran 2007). *Sulcia*, the more ancient p-endosymbiont, has been associated with its host for 250 My and has a genome size of 245 kb (McCutcheon and Moran 2007) whereas *Baumannia*, the younger p-endosymbiont, has only been associated with sharpshooters for 25-40 My (Takiya et al. 2006) and has a larger genome of ~686 kb (Wu et al. 2006). These co-primary endosymbionts only survive by complementing each other. Finally, an endosymbiont may become so degraded it is eventually replaced, possibly out-competed, by another bacterial lineage. In fact, Anderson and Kurland (1998) suggested that obligate bacteria may replace each other at rates determined by Muller's ratchet. The gradual accumulation of slightly deleterious mutations, slowly degrading the genome over time, may make the endosymbiont unable to compete with relatively benign pathogens that have the ability to participate in the mutualism. These

less-attenuated pathogens could then replace the older degraded p-endosymbiont lineages, which may have been the case with some aphid lineages (Moran and Baumann 1994, Perez-Brocal 2006), weevils (Lefevre et al. 2004), and sucking lice (Hypsa and Krizek 2007). The relationship of the pathogen and host at this point would change to a mutualistic one thereby giving the new bacterial lineage the benefits of this relationship such as potentially escaping host immune defense through provision of various host transported mycetomes that protect the new p-endosymbiont, which has been found in lice for many stages (Perotti et al., 2007). For the new mutualist, however, in some cases this new arrangement might hasten its extinction as Muller's ratchet engages.

Table 3-1. Percent sequence divergence and rate of nucleotide substitution of the *16S rDNA* gene of *Riesia*. *Riesia* is the the primary endosymbiont from anthropoid primate lice, these numbers are calibrated at 9.42 and 17.38 Million years.

	9.42 My	17.38 My
Substitutions / My	0.0137	0.0074
Substitutions / site / My	0.0068	0.0037
Percent / 50 My	34.24%	18.56%

Table 3-2. Bacteria taxa, their hosts, and GenBank accession numbers used in the phylogenetic analysis presented in Figure 3-1. Some bacteria lack a species name because they have not been characterized completely. Secondary rather than primary endosymbionts are indicated with an asterisk. Free-living bacteria have no associated host.

Endosymbiont	Host Species	GenBank
Riesia pediculicola	*Pediculus humanus capitis*	EF110571
Riesia pediculicola	*Pediculus humanus humanus*	EF110569
Riesia pediculischaeffi	*Pediculus schaeffi*	EF110573
Riesia pthiripubis	*Pthirus pubis*	EF110574
---	*Haematomyzus elephantis*	DQ076663
---	*Haematopinus apri*	DQ076665
---	*Haematopinus eurysternus*	DQ076661
---	*Haematopinus suis*	DQ076662
---	*Solenopotes capillatus*	DQ076664
---	*Polyplax spinulosa*	DQ076666
---	*Polyplax serrata*	DQ076667
Blochmannia ulcerosus	*Camponotus* ants	AY334375
Blochmannia sp.	*Camponotus* ants	AY334375
Sodalis glossinidius	*Glossina*	AJ245596
Baumannia cicadellinicola	leafhoppers	AY676882
Arsenophonus sp.	*Triatoma melanosoma*	DQ508172
Buchnera aphidicola	*Acyrthosiphon pisum*	BA000003
Wigglesworthia glossinidia	*Glossina brevipalpis*	BA000021
---*	*Melanococcus albizziae*	AF476106
---*	*Planococcus citri*	AF476107
---*	*Erium globosum*	AF476105
---*	*Aphalaroida inermis*	AF263556
---*	*Glycaspis brimblecombei*	AF263561
---*	*Glossina austeni*	GAU64869
---*	*Planococcus citri*	AF476107
*Serratia symbiotica**	*Aphis craccivora*	AY822594
Tatlockia micdadei	--NA--	AF227162
Legionella adelaidensis	--NA--	Z49716
Photorhabdus luminescens	--NA--	EF592562
Providencia alcalifaciens	--NA--	AY994312
Providencia vermicola	*Steinernema thermophilum*	AM040495
Arsenophonus arthropodicus	*Pseudolynchia canariensis*	DQ115535
Escherichia coli	--NA--	AP009048

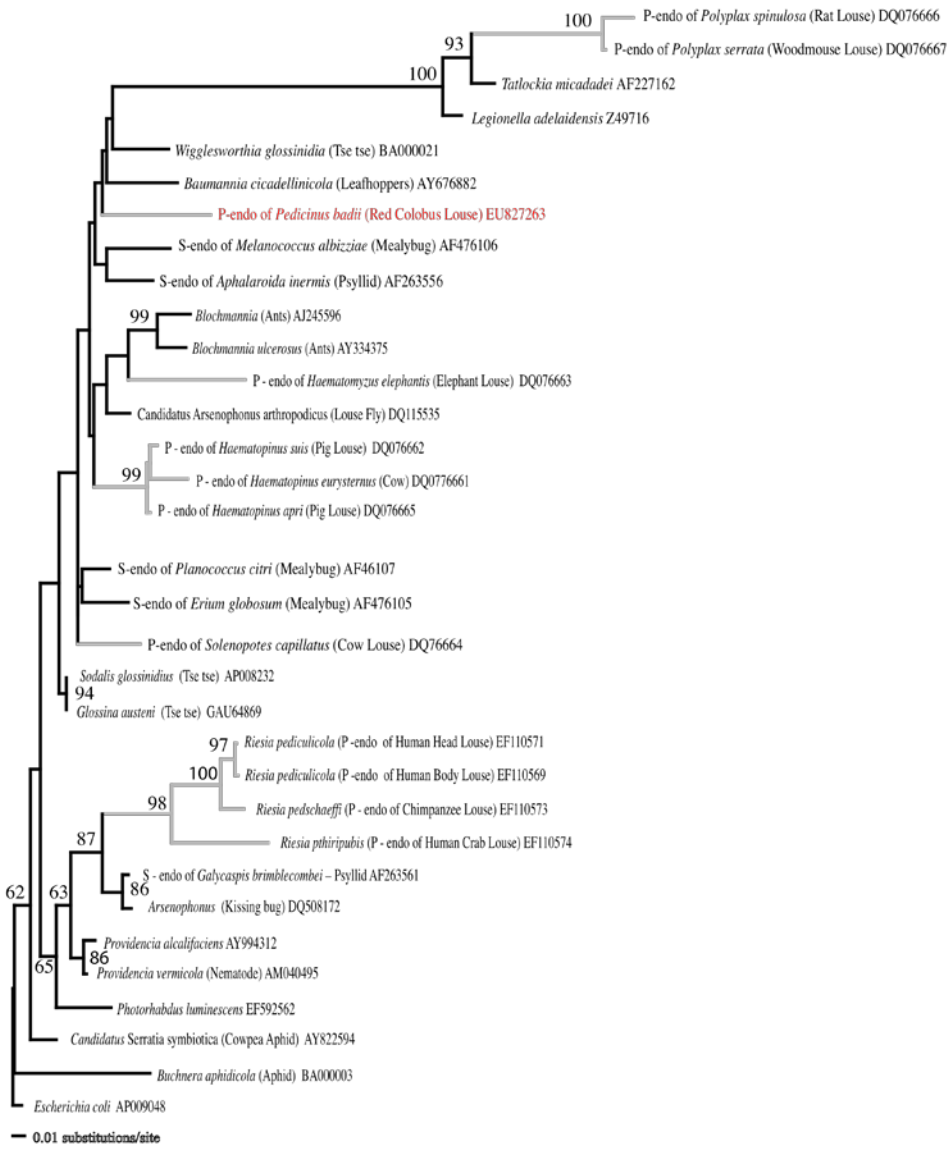

Figure 3-1. Maximum likelihood phylogram representing phylogenetic relationships of alll bacteria used in this study. Common insect p-endosymbionts, and closely related taxa as determined from a BLAST search of each louse endosymbiont sequence. Numbers at nodes indicate maximum likelihood support values greater than 60. Gray lines indicate louse p-endosymbionts. The p-endosymbiont from *Pedicinus badii* (the louse that parasitizes Red Colobus monkeys) is shown in red demonstrating that it does not group with the *Riesia* p-endosymbionts. There are now at least six distinct lineages of p-endosymbionts sampled from sucking lice.

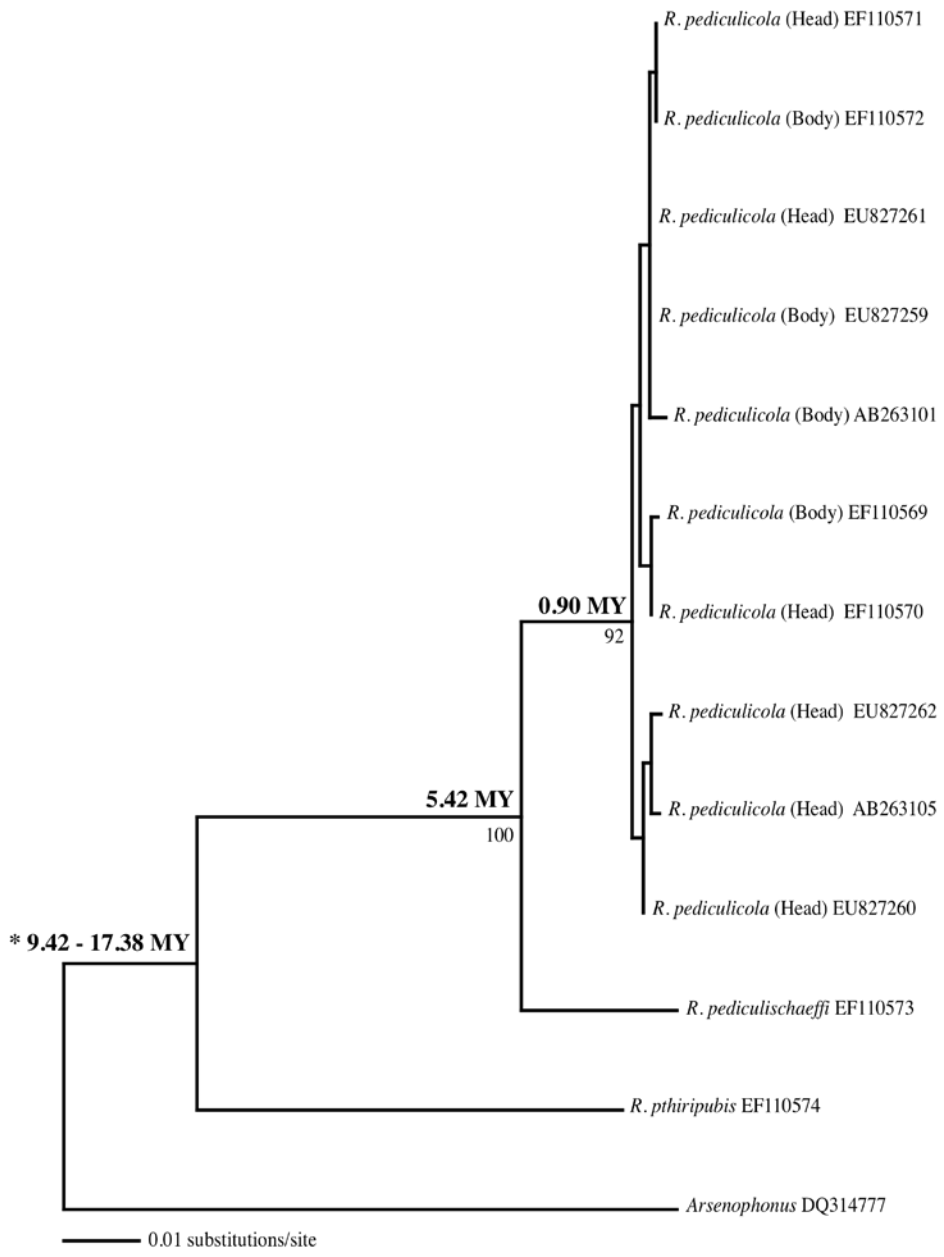

Figure 3-2. Maximum likelihood phylogram representing phylogenetic relationships of *Riesia* p-endosymbionts. Numbers above the node represent divergence dates (in millions of years) whereas numbers below the nodes are bootstrap support values (only numbers greater then 60 are shown). The divergence date calibration point of 9.42 – 17.38 My is indicated with an asterisk.

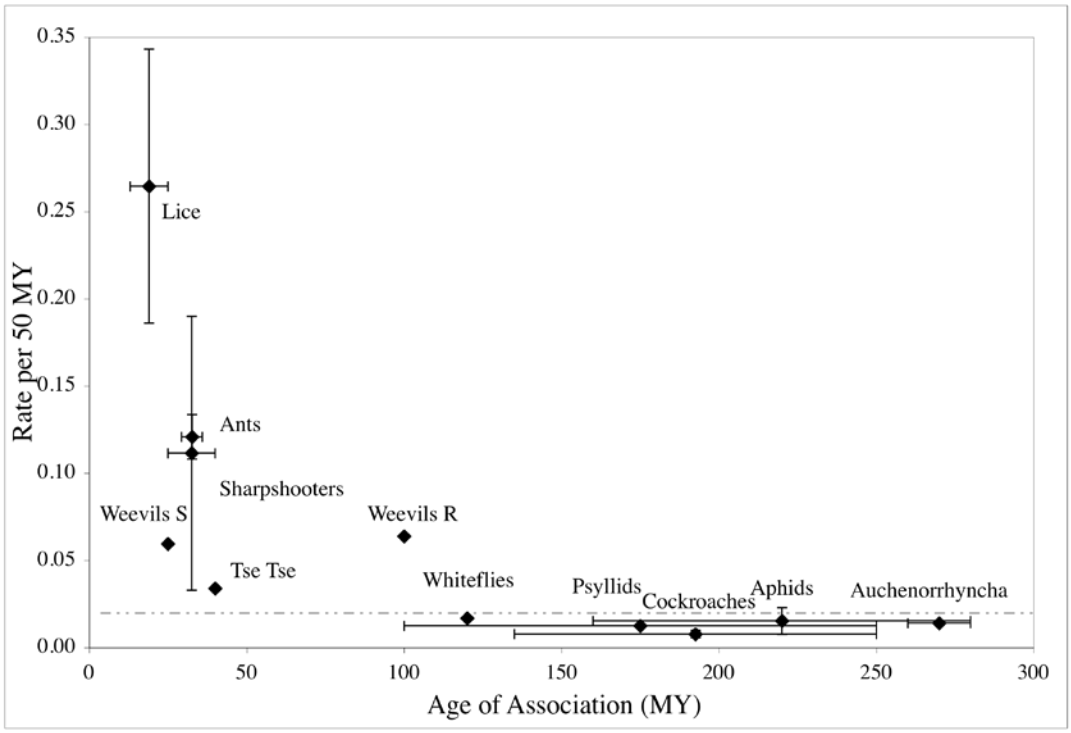

Figure 3-3. Primary endosymbiont nucleotide substitution rates as a function of the age of insect/p-endosymbiont association. Ages and rates were retrieved from the literature or were calculated from pairwise sequence divergences and are shown as percent per 50 My. Point estimates (diamonds) represent the median value between the upper and lower estimates (bars show the full range of dates). The dotted line indicates a rate of 2% per 50 My. It has been suggested that the tse-tse fly and aphid p-endosymbionts (*Wigglesworthia* and *Buchnera*, respectively) are closely related to each other and are the result of a more ancient endosymbiosis event than represented here (Lerat et al. 2003, Canback et al. 2004). Using an older date to calculate rates, however, does not change the results (data available upon request).

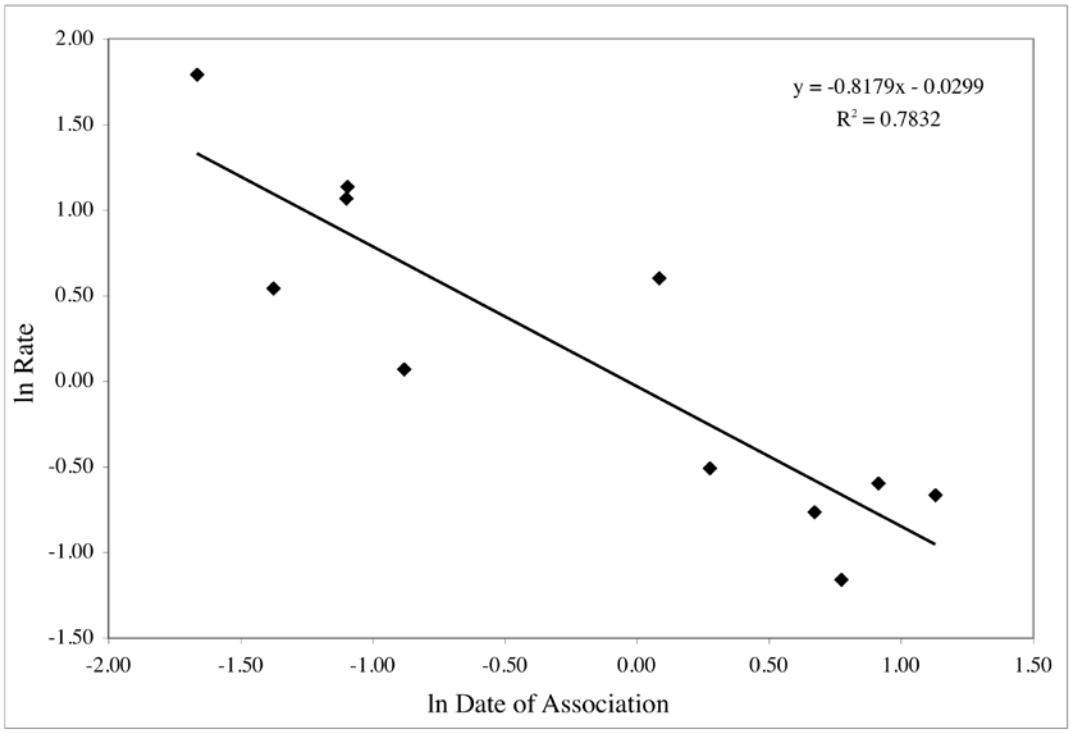

Figure 3-4. Reduced major axis regression of the log transformed data from Figure 3-3. The age of the association is on the x-axis and the rate of nucleotide evolution is on the y-axis. 78% of the variation in rate of nucleotide evolution can be explained by the age of the association suggesting that the rate of nucleotide evolution does decrease over time in p-endosymbionts of insects.

CHAPTER 4
MULTIPLE LINEAGES OF BACTERIA IN ANOPLURA INDICATE A HIGH RATE OF BACTERIAL REPLACEMENT ON A SHORT EVOLUTIONARY TIME-SCALE

Introduction

Bacteria are one of the most common forms of life on the planet; they have evolved into almost every habitat and have every type of interaction with other organisms, from pathogenic to commensal to mutualistic. Humans harbor a community of bacteria that is thought to be critically important in human health and disease prevention (Eckberg et al. 2005). Cows and other fore- and hind-gut fermenters require bacteria to break down and provide nutrients lacking in their diet (Stewart el al. 1997). Many insects also have endosymbiotic bacteria that reside in specialized cells and provide nutrients to the insect (Buchner 1965). Primary endosymbionts have been implicated in helping insects become one of the most diverse animal groups by enabling them to exploit food sources that are lacking in important nutrients, such as plant sap, wood and blood (Buchner 1965), as is the case with blood sucking lice (Pthiraptera: Anoplura), which parasitize mammals.

It is thought that the radiation of many of these insect groups occurred after the insect ancestor acquired the bacteria and was able to exploit a new food source (Buchner 1965). In these insects, the bacteria are transmitted vertically from mother to offspring through the eggs to get incorporated into the next generation. As this process plays out over longer and longer time scales we see that as the insect diversifies, all species of the insect have the same lineage of bacteria. In fact, comparison of the phylogenetic tree of the insect host with that of the bacteria shows that the bacteria have radiated and cospeciated along with their hosts (Moran and Baumann 1994). Although most insect endosymbiont lineages have cospeciated with their hosts, a few

cases have been found where the phylogenetic tree of the bacteria does not match that of its host (Moran and Baumann 1994, Perez-Brocal et al. 2006, Lefevre et al. 2004, Allen et al. 2009). In these cases, the endosymbiont has been replaced with a different endosymbiont lineage. Interestingly, for most of these cases, such replacements have been found only one or two times over hundreds of millions of years of shared evolutionary history between the bacteria and their hosts. These results suggest that bacterial replacement is a rare event.

In Anoplura (sucking lice) however, a different pattern is found. Among the few louse species examined, six different bacterial lineages have been found to date (Hypsa and Kirizek 2007, Allen et al. 2009, Fukatsu et al. 2009). Interestingly, sucking lice are only 60-80 million years old (Light et al. *In Review*), which suggests that many bacterial replacements have happened in a short amount of time compared to other insect/endosymbiont systems (Moran et al. 1993, Bandi et al. 1995). If this trend continues as more lice are sampled, this will likely prove to be the group with the highest known rate of bacterial turnover. With over 540 species of Anoplura represented across 15 families, a much wider sampling of this group is needed to be able to calculate the number of bacterial replacements that have occurred over the last 60 million years.

To determine the rate of bacterial replacement in this group, I sequenced the 16S rDNA gene of primary endosymbiont from many species of Anoplura and analyzed them phylogenetically with other bacteria from Gammaproteobacteria. To understand the coevolutionary history of Anoplura and their primary endosymbionts, the results from the

bacterial phylogeny were mapped onto and compared with the evolutionary history of their louse hosts.

Methods

Lice Sampling

Overall, 27 specimens of lice were collected from museums and mammal collectors, representing 8 families and 21 species for molecular analysis.

To remove external bacteria, lice were washed 3 times in 500ul of 5% bleach, and washed 2 times with sterile water (Meyer 2007). Lice were then completely crushed and extracted using a Qiagen micro kit (Cat No. 56304) according to the manufacturer's protocol with the following modifications. Lice were placed in 80ul of Proteinase K (Quiagen) and incubated overnight on a heating block at 55°C. The DNA was eluted in 50ul of sterile water heated to 55 °C. Water was also extracted as a negative control to ensure no bacterial contamination was introduced during the extraction.

16S rDNA was amplified with general bacterial primers 27 forward (GAG TTT GAT CCT GGC TCA G) and either 1525 (AGA AAG GAG GTG ATC CAG CC) or 1329 (GAC GGA TCCACG GGC GGT GTG TRC) reverse primers. Primers were brought to a final concentration of 0.7uM using Stratagene Hi-Fidelity Master Mix (Cat No. 600650-51). Total PCR volume was 50ul. Cycling conditions included an initial denaturation step at 95°C for 2 min, then 40 cycles of denaturation at 95°C for 40 sec, annealing at 50°C for 30 sec. and extension at 72°C for 2 min. A final extension step at 72°C for 30 min was included to ensure that the majority of the products were polyadenylated for cloning. To separate the different bacterial sequences, products were then cloned using the Invitrogen Cloning Kit (Cat No. 45-0030), and 96 colonies were picked and

sequenced in the forward direction at the ICBR sequencing facility at the University of Florida. Complete sequences for the 16S rDNA region were generated for those products with unique matches when queried against the GenBank non-redundant nucleotide database using BLAST. Due to base pair composition, endosymbionts are likely to match against other endosymbionts in this database, so these matches provided a pool of putative primary endosymbiont sequences. Finally, the primer sequences were removed from all sequences.

Bacterial Sampling Alignment and Tree Building

We downloaded all Gammaproteobacteria 16S rDNA sequences from the Ribosomal Database Project (RDB, ~72,000) and removed all identical sequences (leaving ~65,000 sequences). Next, we removed sequences that were shorter than 750 base pairs (roughly half the gene) to mitigate alignment issues and to retain individuals where the majority of information about the gene was available. The resulting 42,626 taxa were then split into 20 groups of approximately 2,000 taxa each, and one or two of the endosymbiont sequences from each louse were added to each group. Alignments were generated for each group using MUSCLE (Edgar 2004), and checked by eye. Profile alignments (two alignments were maintained and aligned together by adding in gaps) were then created, and again checked by eye. Overall, there were 20 individual MUSCLE alignments and 19 profile alignments. Finally, ambiguous sections of the alignment were removed.

To investigate how taxon sampling would affect the number of endosymbiont lineages identified in Anoplura, four datasets were created. First, all pairwise distances were calculated, and sequences that were at least 80% similar were combined into clusters. Datasets for phylogenetic analysis were generated by selecting one sequence

from each cluster, and adding all endosymbiont sequences to yield an alignment of 76 taxa. This process was then repeated with progressively more stringent sequence similarity thresholds to yield datasets of 215 taxa (85% similarity), 865 taxa (90% similarity) and 4,275 taxa (95% similarity). Each dataset was nested within subsequent datasets, i.e., all 76 taxa from the 80% dataset were included in the 85% dataset, and all taxa in the 85% dataset were included in the 90% dataset etc.

Phylogenetic trees were created for each of these datasets using a maximum likelihood framework with a GTR + G model of molecular evolution in the program RAxML (Stamatakis 2006), and bootstrap values were calculated from 200 replicates. The number of endosymbiont lineages found on each tree was then calculated by counting all of the nodes that grouped sucking lice endosymbionts together to the exclusion of all other bacteria with > 50% bootstrap support. Finally because endosymbionts are thought to have greater than 50% AT bias (Lambert and Moran 1998) in their 16S rDNA, AT% was calculated for all of the putative endosymbionts, this is another indicator that the bacterial sequences came from an endosymbiont. Free-living bacteria have < 50% AT composition at the 16S rDNA.

Results

Endosymbionts

Overall, 27 louse species were examined and 21 putative endosymbiont sequences found (Table 4-1); in one instance two distinct endosymbiont sequences were found within a single louse (*Ancistroplax crocidurae*). For six specimens, I could not recover any sequence that met either the sequence similarity or the base composition percentage criterion for endosymbionts. Most of these came from one louse family, Hoplopleruidae, and five of the six lice species harbored another known

louse pathogen, *Bartonella* (Proteobacteria: alphaproteobacteria). Interestingly, in one specimen, I found both a putative endosymbiont sequence and a *Bartonella* sequence (Table 4-1).

A total of 30 sequences of putative endosymbionts were recovered: 22 generated from this study and 8 downloaded from GenBank. Together, these 30 sequences cover 8 families of Anoplura (Table 4-1). All 30 sequences showed sequence similarity to known endosymbionts in a BLAST search, and 27 of those sequences had AT% of 49 or above. Three sequences from distantly related lice, including the second endosymbiont sequence from *Ancistroplax crocidurae* had AT% of 45, suggesting these may not be a primary endosymbiont based on the 50% rule. The endosymbionts I sequenced from the genus *Pedicinus* had an average pairwise sequence divergence of 3.7% from another known *Pedicinus* endosymbiont, *P. obtusus* (Fukatsu et al. 2009), where the fluorescent in-situ hybridization (FISH) has established that the bacterial sequence is indeed found within the mycetome.

Phylogenetic Analysis

Interestingly, most of the endosymbionts grouped fairly close together within the large bacterial trees (Figure 4-1) and were nested within the genus *Arsenophonus*. Other insect endosymbionts are also in this part of the tree, like *Buchnera* (Aphids) and *Wigglesworthia* (Tsetse flies). Only one group of endosymbionts had a distinctly different placement in the tree, *Polyplax sp*. These endosymbionts were classified to the genus *Legionella* in the Ribosomal Database Project.

As the number of taxa in the phylogenetic analysis increased, the number of endosymbiont lineages increased as well, the total number of endosymbiont lineages therefore was 13 with a higher number of taxa represented (Table 4-2). A few

endosymbiont groups had phylogenetic relationships that mirrored those of their hosts. These included *Candidatus* Riesia found in Great Ape lice, *Pedicinus* (Old World Monkey Lice), *Haematopinus* (Pig Lice), *Polyplax* (Rodent lice), and *Ancistroplax* (Rodent Lice) (Figure 4-2). Only one group was found that did not match their host's evolutionary history, and it was found consistently in all of the trees with high bootstrap support. This group consists of the endosymbiont from a seal louse (*Proechinophthirus fluctus*) in the family Echinopthiridae, the second putative endosymbiont sequence from *Ancistroplax crocidurae* family Hoplopleuridae, and the endosymbiont sequence from *Sathrax durus* family Polyplacidae (Figure 4-2). These endosymbiont sequences had only 45% AT bias, which suggests they are either not a primary endosymbiont or a very young endosymbiont and have not had time to accumulate AT bias.

Discussion

Endosymbionts

Examination of the results of the base pair composition analysis reveals some interesting findings. Although most endosymbiont sequences had greater than 50% AT bias, a few had less. Two endosymbionts from lice in the genus *Ancistroplax* had 49% and 50% respectively, and they grouped together. Three endosymbionts had 45% AT bias, much less than the 50% cut off for primary endosymbionts, and these three endosymbionts grouped together with high bootstrap support in all of the analyses. They are also the only group that does not mirror host evolutionary history. These endosymbionts are from three species of Anoplura that are very distantly related on the louse tree. One possible explanation is that they are grouping together due to base pair composition and not phylogenetic history. However, when we calculate average pairwise distances in this group we get 0.0074, which is much smaller than the

endosymbionts from *Pedicinus* (0.0373), Haematopinus (0.0329) and *Candidatus* Riesia (0.0579), which all group together with high bootstrap support with the same branching pattern as their hosts.

One explanation for this surprising finding is that this bacterium may actually be a secondary endosymbiont, or a pathogen that is found in many lice, but only sampled in this study a few times. One louse (*Ancistroplax sp*.) had both this sequence and another endosymbiont sequence with 49% AT bias, which is more similar to other primary endosymbionts, so there is some evidence of this lineage co-occurring with other lineages that more closely fit the criteria of a primary endosymbiont. FISH analysis to determine where in the louse this sequence is found would illuminate where it is located. Interestingly, it is thought that the transition to the mutualistic lifestyle speeds up the rate of molecular evolution, if this lineage of bacteria is not in the mycetome, and not the primary-endosymbiont, its rate of molecular evolution may not have increased, which may explain why such a small amount of divergence is found between these bacteria from distantly related hosts.

Taxonomy

In 2007 we proposed to name the endosymbiont of the gorilla louse (*Pthirius gorillae)*, *Riesia pthirigorillae (*Allen et al. 2007*)*; however, recently it has been suggested that *Riesia,* along with other insect endosymbionts, should be placed in the genus *Arsenophonus* (Moran et al. 2009). Here, we find similar results with most of our Anoplura endosymbionts grouping within the genus *Arsenophonus*. In light of this, it is likely that the genus *Riesia* will be subsumed within *Arsenophonus*. However, for louse endosymbionts, we propose that the naming system continues by describing the host name as the species name. Ultimately, however, before names can be formally

described the FISH work needs to be completed to demonstrate that the sequences do come from the mycetome.

Bartonella as an Endosymbiont

Bartonella is a known louse pathogen from a different group of bacteria (Alphaproteobacteria). For many louse taxa, we found this pathogen, and for a few lice I did not sequence a primary endosymbiont. One explanation would be that *Bartonella* has become the primary endosymbiont in this group. To determine if this is the case, I downloaded all *Bartonella* sequences from RDB, removed the sequences having less than half of the gene, aligned them using MUSCLE, and checked them by eye. Using a GTR + G model of molecular evolution, I created a best Maximum Likelihood tree in RAxML with 100 bootstrap replicates. Here, I found that only the *Bartonella* sequences from this group of Hoplopleuran lice group together with high bootstrap support (Figure 4-3). The other *Bartonella* sequences were found outside of this clade and throughout the *Bartonella* tree (data not shown). This pattern suggests that in this louse group, this *Bartonella* lineage is different. One possible explanation is that it has become the primary endosymbiont in this group.

Number of lineages

We expect that as we add taxa to a phylogenetic analysis, some clades will break up, and therefore the number of independent endosymbiont lineages (those that group together to the exclusion of all other taxa) would increase. Also, because there are a finite number of endosymbiont lineages sampled, we expect that the number of endosymbiont lineages will remain the same once a sufficient number of taxa have been included in the analysis. Here, we find that the number of distinct lineages of endosymbionts increased from 10 to 13 (Table 2) at 865 taxa, and remained at 13 at

with 4,275 taxa represented. This suggests that sampling 865 taxa from clusters of sequences that are 90% similar from Gammaproteobacteria, is adequate to calculate the number of independent lineages of endosymbionts in this group, and may provide clues as to the appropriate number of taxa needed to describe phylogenetic relationships among Gammaproteobacteria with 16S rDNA.

We found 13 distinct lineages of bacteria in Anoplura; however, it is likely that there are more because only 8 of the 15 families have been examined. If we suggest that *Bartonella* is the endosymbiont in that clade of Hoplopleuran lice, then we have 14 replacements over 80 million years, which gives us a rate of 0.175 changes/million years, or one turnover every 6 million years. We know that the age of the *Riesia*/Louse association is between 8 and 25 million years old, which supports the finding that these bacteria get replaced roughly every six million years.

Evolution of Sucking Lice and their Endosymbionts

There are five groups of endosymbionts that grouped together with high bootstrap support those of the louse genera *Pediculus* and *Pthirus*, *Pedicinus*, *Haematopinus*, *Ancistroplax*, and *Farenholzia* (Figure 4-2). Within these groups, they have the same branching pattern as their hosts, which suggests that, after a bacterial replacement, the new endosymbiont then coevolved and cospeciated with its louse host for some time.

It has been suggested that endosymbionts are replaced at a rate corresponding to their rate of genome degradation (Andersson and Kurland 1998). As endosymbionts accumulate slightly deleterious mutations over time, they are unable to compete with invading bacteria and may be replaced. However, there have been a number of papers suggesting that selection may act to reduce the rate of genome degradation in this group (Clark et al. 1999, Tamas et al. 2002 and van Ham et al. 2003, Allen et al. 2009),

and may prevent the endosymbionts from being replaced. Anoplura as a whole is only around 80 million years old, and many of these endosymbiont groups are likely much younger than 50 My.(Light et al In Review) In this group though, they are being replaced at a much higher rate.

Blood Feeding as a Source of Endosymbionts

Many insects whose endosymbionts grouped within the *Arsenophonus* genus also feed on mammalian blood. For example, the endosymbiont from Great Ape lice, *Candidatus* Riesia (Great Ape Lice endosymbionts), grouped together with endosymbionts from Hippoposcid flies (many genera), deer keds (Liptoptena) and a dog tick (Dermacentor) (Data not shown). One way to interpret this is that there may be some common *Arsenophonus* lineages in mammalian blood that have the tendency to become primary endosymbionts in insects. Therefore, when lice switch to a new mammalian host, they may come into contact with a different lineage of *Arsenophonus* that is then able to outcompete the myectomic bacteria and become the primary endosymbiont. In lice however, this bacterium is likely to have a short evolutionary history with its host as it will likely get replaced sometime later.

Table 4-1. Louse taxa used in this study (arranged by family), host associations, and GenBank accession numbers. Museum acronyms for host taxa installed in Natural History Museums are as follows: Louisiana State University Museum of Natural Science (LSUMZ), Moore Laboratory of Zoology, Occidental College (MLZ), New Mexico Museum of Natural History (NMMNH), and University of Alaska Museum of the North (UAM). √ indicates endosymbionts sequenced, X indicates no endosymbiont sequenced, other bacteria sequenced as indicated.

Louse Family and Species (Locality)	Taxon Label	Host (Order: Family; Museum Voucher)	Endosymbiont	%AT
Echinophthiriidae				
Proechinophthirus fluctus (USA: AK)	Echin3.17.09.2	*Callorhinus ursinus* (Carnivora: Otariidae)	Yes	45%
Haematopinidae				
Haematopinus eurysternus	--	*Bos taurus* (Artiodactyla: Bovidae)	DQ076661	52%
Haematopinus suis (Florida, USA)	Hpsu7.14.09.4	*Sus scrofa* (Artiodactyla: Suidae)	Yes	52%
Haematopinus suis	--	*Sus scrofa* (Artiodactyla: Suidae)	DQ076662	52%
Haematopinus apri	--	*Sus scrofa* (Artiodactyla: Suidae)	DQ076665	52%
Hoplopleuridae				
Ancistroplax crocidurae 1 (Vietnam)	Axcro4.26.09.1	*Crocidura* sp. (Soricomorpha: Soricidae)	Yes	50%
Ancistroplax crocidurae 2 (China)	Axsp7.14.09.5	*Crocidura attenuata* (Soricomorpha: Soricidae)	Yes(2)	49%, 45%
Hoplopleura ferrisi 2 (MX: Puebla)	Hofer7.14.09.8	*Peromyscus difficilis* (Rodentia: Cricetidae; LSUMZ 36247)	No-*Bartonella*	
Hoplopleura hirsuta (USA: TX)	Hosp4.17.09.7	*Sigmodon hispidus* (Rodentia: Cricetidae; LSUMZ 36377)	No-*Bartonella*	
Hoplopleura onychomydis (USA: AZ)	Hoony8.27.08.6	*Onychomys torridus* (Rodentia: Cricetidae; NMMNH 4394)	No-*Bartonella*	
Hoplopleura reithrodontomydis 2 (USA: AZ)	Hosp7.14.09.6	*Reithrodontomys* sp. (Rodentia: Cricetidae; NMMNH 4411)	No	
Hoplopleura sicata (China)	Hosic7.14.09.9	*Niviventer fulvescens* (Rodentia: Muridae)	No-*Bartonella*	
Linognathidae				
Linognathus spicatus (Zimbabwe)	Linog6.22.09.1	*Connochaetes taurinus* (Artiodactyla: Bovidae)	Yes	52%
Solenopotes capillatus	--	*Bos tarus* (Artiodactyla:Bovidae)	DQ076664	50%
Pedicinidae				
Pedicinus badii (Uganda)	Qnbad7.24.06.8	*Piliocolobus tephrosceles* (Primates: Cercopithecidae)	Yes - EU827263	53%
Pedicinus pictus 1 (Ivory Coast)	Qnpic3.31.08.1	*Piliocolobus badius* (Primates: Cercopithecidae)	Yes	54%
Pedicinus pictus 2 (Ivory Coast)	Qnpic6.30.09.2	*Colobus polykomos* (Primates: Cercopithecidae)	Yes	53%
Pedicinus pictus 3 (Ivory Coast)	Qnsp3.31.08.3	*Colobus polykomos* (Primates: Cercopithecidae)	Yes	54%
Pedicinus obtusus (Nagano, Japan)	--	Macaca fuscata (Primates: Cercopithecidae)	AB478979	54%
Pediculidae				
Pediculus humanus capitis	Pdcap1.19.05.1	*Homo sapiens* (Primates: Hominidae)	Yes - EF110571	51%

Table 4-1. Continued.

Louse Family and Species (Locality)	Taxon Label	Host (Order: Family; Museum Voucher)	Endosymbiont	%AT
Pediculus humanus capitis (USA: FL)	Pdcap9.20.05.2 NW	*Homo sapiens* (Primates: Hominidae)	Yes	51%
Pediculus humanus humanus (USA:MD)	Pdhum5.19.05.2	*Homo sapiens* (Primates: Hominidae)	Yes	51%
Pediculus schaeffi (Uganda)	Pdsch4.30.03.8	*Pan troglodytes* (Primates: Hominidae)	Yes - EF110573	51%
Polyplacidae				
Fahrenholzia ehrlichi 1 (USA: TX)	Fzehr8.20.08.1	*Liomys irroratus* (Rodentia: Heteromyidae; LSUMZ 36395)	Yes - Bartonella	52%
Fahrenholzia ehrlichi 2 (MX: Puebla)	Fzehr6.30.09.4	*Liomys irroratus* (Rodentia: Heteromyidae; LSUMZ 36299)	Yes	51%
Lemurpediculus verruculosus 1 (Madagascar)	Lesp4.26.09.2	*Microcebus rufus* (Primates: Cheirogaleidae)	Yes	53%
Linognathoides marmotae 1 (USA: CO)	Lnlae6.30.09.3	*Marmota flaviventris* (Rodentia: Sciuridae)	Yes	54%
Neohaematopinus neotomae (USA: CA)	Neneo8.20.08.2	*Neotoma lepida* (Rodentia: Cricetidae; MLZ 1921)	No - *Bartonella*	
Neohaematopinus sciuropteri (USA: OR)	Nescp6.30.09.5	*Glaucomys sabrinus* (Rodentia: Sciuridae)	Yes - *Bartonella*	53%
Sathrax durus (Vietnam)	Sathrax4.26.09.3	*Tupaia belangeri* (Scandetia: Tupaiidae)	Yes	45%
Polyplax serrata	--	Apodemus sylvaitcus (Rodentia:Muroidea)	DQ076667	52%
Polyplax spinulosa	--	Rattus norvegicus (Rodentia:Muridae)	DQ076666	53%
Pthiridae				
Pthirus gorillae (Uganda)	Ptgor9.14.08.1	*Gorilla gorilla* (Primates: Hominidae)	Yes	53%
Pthirus pubis (Scotland)	Ptpub8.14.06.2	*Homo sapiens* (Primates: Hominidae)	Yes - EF110571	53%
Rhynchophthirina				
Haematomyzus elephantis	059_*Haem_elephantis*	*Elephas maximus* (Proboscidea: Elephantidae)	DQ076663	55%

Table 4-2. Taxon sampling, and endosymbiont determination. Here we built phylogenetic trees with different numbers of taxa, found by clustering all of the gammaproteobacterial sequences according to % similarity and selecting a representative sequence from each cluster. As the number of taxa examined increased, there was an increase in the number of distinctly different endosymbiont lineages found, ranging from 10 to 13. However, going from 865 to 4275 taxa did not change the endosymbiont number, suggesting that 865 taxa is adequate sampling to identify the number of endosymbiont lineages.

Cluster	80%	85%	90%	95%
No. Taxa	76	217	865	4275
No. Endosymbionts	11	10	13	13

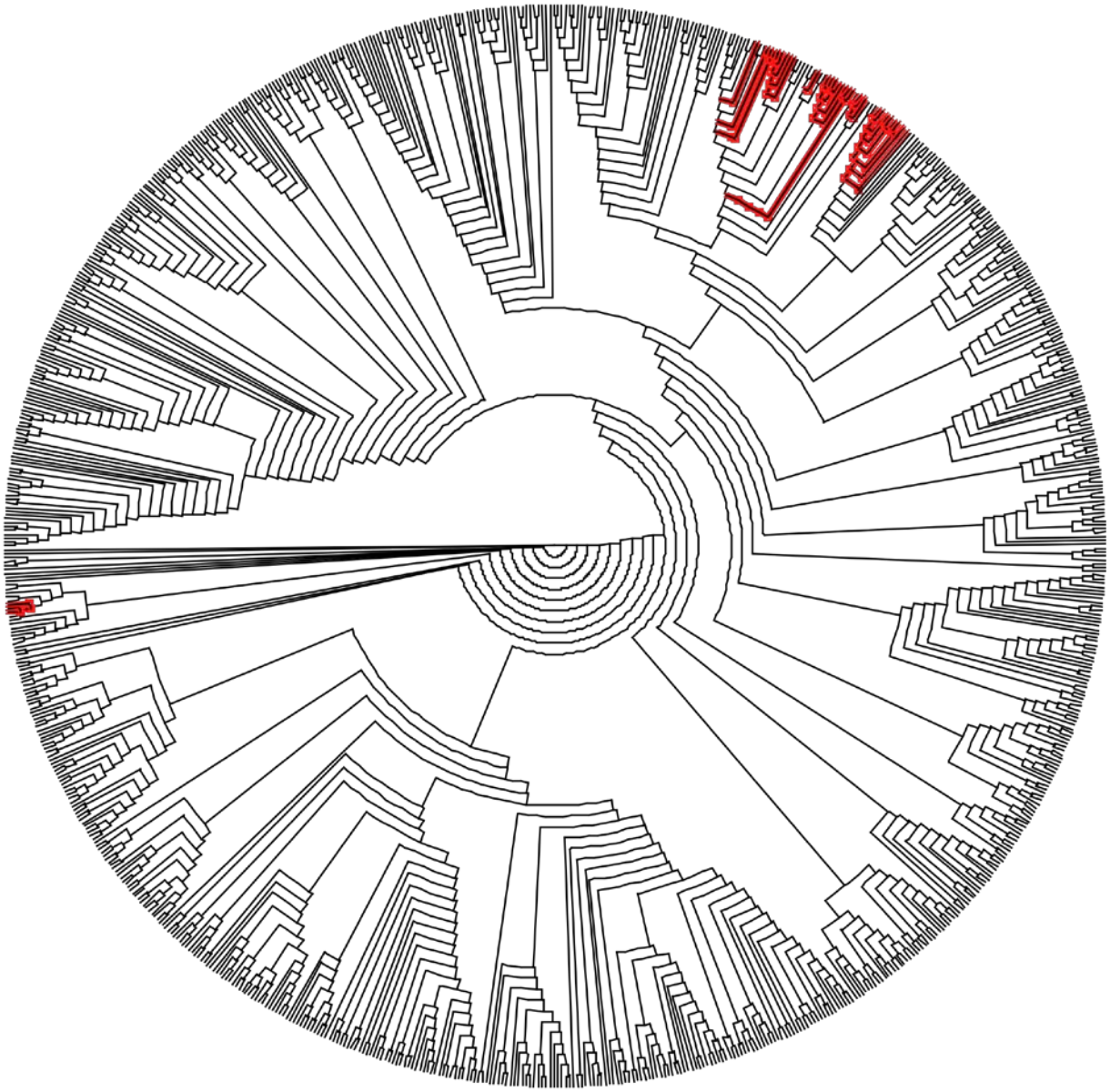

Figure 4-1. Bacterial tree showing placement of Anopluran endosymbionts. Maximum likelihood tree of Gammaproteobacteria from the 90% cluster (865 taxa). Red lines indicate the placement of the endosymbionts from sucking lice (Anoplura). These lice are mostly nested within the Genus *Arsenophonus*, with the exception of the primary endosymbionts from *Polyplax sp*. Identified to be in the genus Legionella from the Ribosomal Database Project.

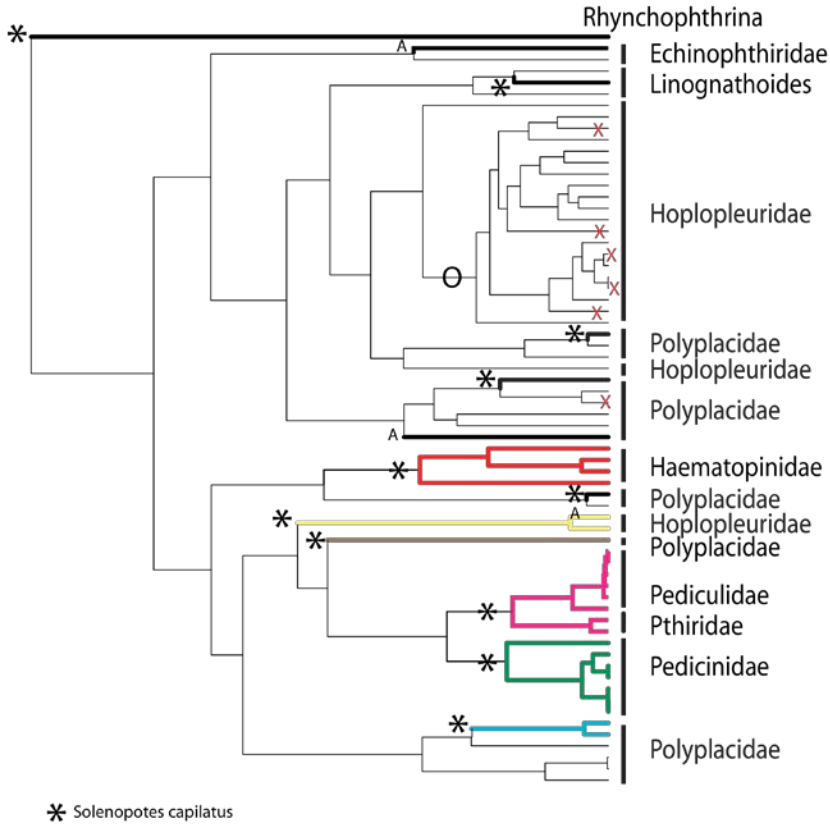

Figure 4-2. Phylogenetic tree of Louse taxa from Light et al. (In Review). Names on the right are the Host Family names. Thicker branches indicate ones with a putative endosymbiont sequence. Colors represent endosymbionts that group together phylogenetically and also show relationships the same as their hosts, suggesting cospeciation of the bacteria and their hosts in these groups. Asterisks represent independent lineages of bacteria, *Solenopotes capilatus* is not represented on the tree because we did not have the molecular data, but an endosymbiont sequence was on GenBank. Subscript A represents the three putative endosymbionts that always group together. Zero indicates a potential loss of the endosymbiont or a switch to a different type of endosymbiont like *Bartonella*.

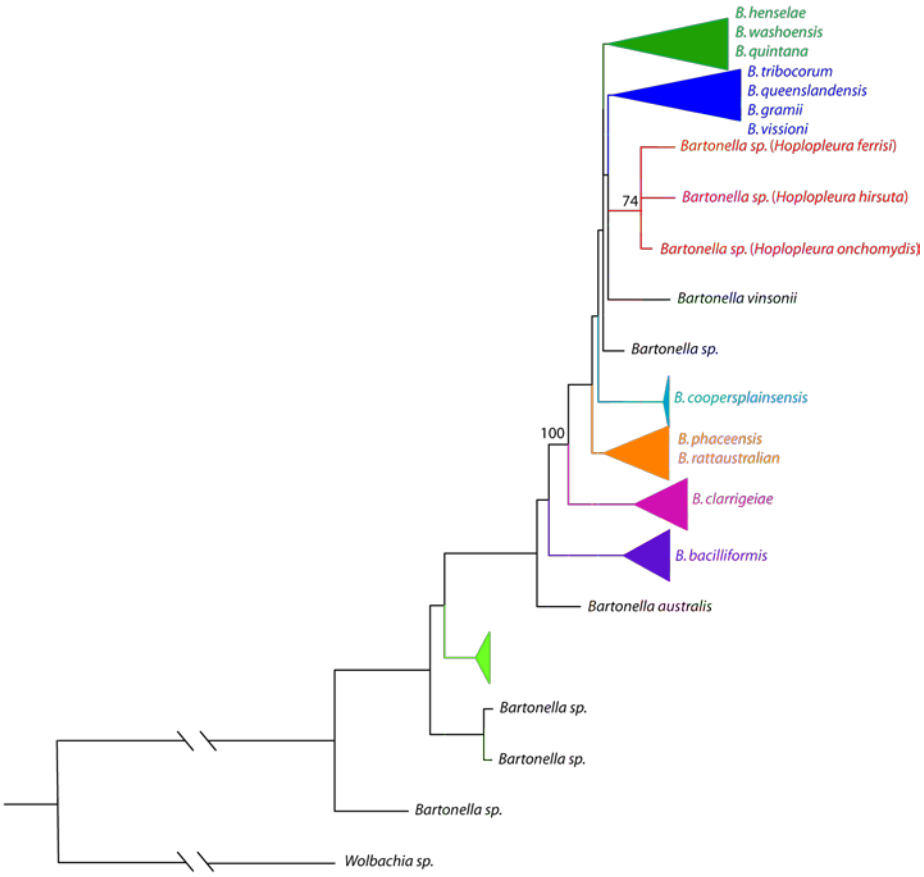

Figure 4-3. Maximum Likelihood tree of 279 *Bartonella* sequences. *Bartonella* from Sucking lice are shown in red, and they group together with 74% bootstrap support, suggesting that in this group of Hoplopleuran lice *Bartonella* may in fact be the endosymbiont. The other *Bartonella* endosymbionts sequenced from Anoplura are nested within different groups on this tree (not shown).

CHAPTER 5
POPULATION GENETICS OF HABITAT SENSITIVE RED COLOBUS SUGGEST LONG-TERM STABILITY OF KIBALE NATIONAL PARK

Introduction

Red Colobus monkeys (Cercopithecidae: Colobus: *Piliocolobus*) are considered to be one of the most endangered primates in the world (IUCN 2010). These monkeys are patchily distributed across equatorial, all in the genus *Piliocolobus*. The radiation of Colobines began in the Pliocene around 7.5 (±1.2) Ma and the Red Colobus (*Procolobus* (*Piliocolobus*)) diverged from the Olive Colobus (*Procolobus* (*Procolobus*)) around 6.4 (±1.1) Ma (Ting, 2008). Recent molecular work suggests that the extant forms of *Pilocolobus* radiated only around 3 Ma (Ting 2008). The taxonomy of this group has been debated for some time, with estimates of the number of species ranging from one to 16 (Ting, 2009). Mittermeier et al. (2009) recognized nine members of this genus, and ranked three within the top 25 most endangered primates. The IUCN (2010) lists six members of the genus *Piliocolobus,* and five of those are considered endangered or critically endangered.

Red Colobus are critically endangered in part because they are very sensitive to habitat change. They are large, folivorous monkeys (Struhsaker, 1975, Oates 1996b) that live in social groups ranging from 25 to 150 individuals (Struhsaker 1975, Chapman pers. com.), necessitating large tracts of land for foraging (Oates, 1987), Due to this most populations of Red Colobus today are threatened by habitat destruction (Struhsaker and Leland 2008). Another issue facing Red Colobus is hunting pressure. Because of their large size, they are often hunted by humans and other predators (Oates and Davies 1994); it has been documented that one population of Red Colobus

in west Africa was recently eradicated due to human hunting within the last 30 years (Oates et al., 2000). Currently, it is thought that there is only one area where Red Colobus are not threatened, and that is Kibale National Park, Uganda (Struhsaker 1997).

Kibale National Park (KNP) (795 km^2) is located in western Uganda near the foothills of the Ruwenzori Mountains (Struhsaker, 1975; 1997). Currently, KNP harbors 13 different species of primates including the highest numbers of Red Colobus in the world; census estimates range from 30,000 – 80,000 monkeys (Chapman pers. com.). These red colobus are thought to have diverged from a sister population less than 600,000 years ago (Ting 2008). They are also the most extensively studied population where data have been collected on behavior (Struhsaker 1975), dietary preferences (Chapman et al. 2005), and even viral load (Goldberg et al. *in press*). However, the reasons these Red Colobus are so abundant in KNP while other populations are on the verge of local extinction remain unclear. Because they are so sensitive to habitat change, the protection the monkeys have received inside the park likely plays a major role. There are good records on human impacts (i.e. logging) in the park that go back at least 30 years, and evidence of disturbances (clearings) within the last 2,000 years (Chapman et al. 2000). Interpreting how anthropogenic disturbances have affected KNP monkey populations in the recent past can provide insight into the stability of this species and inform conservation measures in the future. A critical first step is to develop a basic understanding of the population genetic structure of these monkeys.

To understand the population history of Red Colobus in KNP, we used genetic data to first determine the number of populations. Based on those results, we calculated

the effective population size, coalescent point, and looked for evidence of a population bottleneck. Together with the information about disturbances in the park, these analyses illuminate how the population size of the Red Colobus has been affected by habitat change.

Methods

DNA Collection and Storage

Blood and fecal samples were collected from 6 groups throughout KNP from 2007 – 2009: Small Camp (SC), Large Mikana (LM), K-30, Dura, Sebatolli, Mainaro (Figure 5-1). The first two groups have been part of an ongoing behavior study, and blood samples were obtained when individuals were fitted with collars. Later, to increase the number of individuals sampled for this study, non-collared monkeys were followed to obtain fecal material. For the remainder of the groups, at least five individuals were sampled to look for evidence of gene flow throughout the park (Figure 5-1).

Blood samples were stored on a Classic FTA Card (Whatman WB-12-0205). Fecal samples were treated with the two-step procedure described by Nsubuga et al. (2004): samples were placed in 95% ethanol for 24-36 hours, then moved to silica beads (Sigma S7625) for longer-term storage.

DNA Extraction and Quantification

DNA was extracted from blood samples using the FTA protocol described by the manufacturer. Fecal samples were extracted using a QIAmp®Stool kit (Qiagen) according to the manufacturer's protocol. Feces only contain a small amount of DNA from the cells lining the rectum, which can make genotyping unreliable (Goosens et al 2000). In order to obtain adequate DNA for genotyping analyses, DNA was extracted

from 2-4 fecal samples from each individual (Taberlet et al. 1996, Goosens et al. 2000). Overall, this resulted in 230 extracts. We quantified the DNA in a sub-sample of these extracts (N=143), in order to determine the number of PCR amplifications necessary to obtain an accurate genotype (after Harris et al. 2009; Arandjelovic et al. 2009). Extracts had DNA concentrations of 0.48 to 50pg per reaction, consistent with other fecal studies (Harris et al., 2009, Arandjelovic et al. 2009). A majority (>70%) of samples had greater than 5 pg of DNA per extract, which previous studies have determined requires 2-3 amplifications from independent extracts per individual to determine accurate genotypes (Harris et al., 2009, Arandjelovic et al. 2009).

Genotype analysis

Eleven loci, originally identified from the human genome, (Table 5-1) and found to be variable in Red Colobus (Difore pers com.), were amplified using a three-primer nested PCR as in Schuelke (2000), with one primer a universal M13 fluorescently-labeled forward primer. PCR reactions were conducted in a total volume of 25 µl: 12.5 µl of the Type-it Microsatellite PCR Kit (Qiagen), final concentrations of 0.04 µM for each of the primers (non-labeled forward primer, labeled universal M13 primer and reverse primer), and 8.5 µl water, and .48 – 50pg of DNA per reaction. Cycling conditions included an initial activation step of 95° for 5 min, a first round of cycling consisting of 15 cycles of 95° for 30 seconds, 52° for 1 min 30 seconds and 72° for 30 seconds, followed by a second round of cycling consisting of 28 cycles of 95° for 30 seconds, 48° for 1 min 30 seconds and 72° for 30 seconds, and a final extension step of 72° for 30 min. Final products had a fluorescently-labeled universal M13 primer (-21) at the 5’ end, and were run on a AB3730xl 96 capillary automated sequencer and scored

using the program GeneMarker® (Softgenetics). Only one person (JMA) did all of the scoring to reduce error. Each PCR included a positive control (Monkey 13) to ensure proper scoring and a negative control (water) to check for contamination.

Each locus was amplified from at least two different extracts. If the two reactions produced different genotypes (i.e. missing one allele), another reaction was conducted until two reactions produced consistent genotypes, with up to four reactions conducted for each locus. To decrease the probability of allelic dropout, homozygotes were verified by at least 3 successful reactions. If four PCR reactions failed to yield two consistent genotypes, the sample was scored as missing data. Finally, individuals that had fewer than 9 loci typed were excluded. The final dataset included 85 individuals (SC=38, LM=26, K30=5, Dura=5, Sebatoli=6, Mainaro=5).

Number of Populations

The number of populations of Red Colobus was calculated three ways in order to verify our results, because all subsequent analyses relied on these results. First, using RSTCALC (Goodman 1997), which estimates R_{ST} (an estimation of F_{ST} from microsatellite data) and *Nm* (number of migrants per generation) and secondly, using STRUCTURE 2.3.3 (Pritchard et al., 2000), which assigns individuals to populations where the number of populations is defined *a priori* by the user. STRUCTURE analysis was performed using an admixture model and correlated allele frequencies among populations. Because six troops were sampled throughout the park, the *a priori* number of populations (k) was tested from 1 to 6. Each test was allowed to run for 1×10^6 generations, after excluding a burn in of 1×10^5 generations. Each k was tested ten times to examine the variation between runs. The likelihood scores for each run were examined visually for stationarity, and the maximum likelihood from each run plotted

against k (Figure 5-2). Finally, MIGRATE 3.14 (Beerli 2006, 2008), which uses Bayesian inference and coalescent theory to calculate the likelihood of a one population model and a two population model clustering the data into the two populations based on the Rst results above (Beerli 2010). For the one population model, MIGRATE was run for 100 X 10^6, and a 10% burn-in period was excluded. Multiple tests of the two-population model were run, using different ranges for the prior distribution of M (number of migrants per allele per generation per mutation rate) from 0-1, 0-5, 0-10, 0-15, 0-25, and 0-500. Each model was run for 100 X 10^6 generations, with 10% burn-in, and repeated twice to examine variability between the runs. The results were uploaded into TRACER, and likelihood scores examined visually for stationarity. Marginal likelihoods of the models were then compared using bayes factors (Beerli 2010).

Analysis

To look for evidence of genotyping errors (scoring errors, null alleles and allelic dropout), the loci were examined in Micro-Checker 2.2.3 (van Oosterhout et al., 2004). Evidence of non-random associations between alleles was examined in FSTAT 2.9.3.2 using 550,000 randomizations among the 11 loci (Gouded, 2002), and deviations from Hardy-Weinberg equilibrium were calculated in POPGENE 1.31 using a likelihood ratio (G^2) test (Yeh et al, 1997).

Estimates of Theta, Population Size and Bottlenecks

The results from the analyses above strongly suggested that KNP Red Colobus make up a single population. Therefore, for the remainder of the analysis, all samples were included as a single population. We calculated effective population size (*Ne*) by calculating θ from MIGRATE 3.14 with a one-population model. For diploid biparental loci, θ is equal to 4*Ne,* scaled by mutation rate (μ) (Watterson 1975; Tajima 1983;

Hudson 1990; Hein *et al.* 2004). The human mutation rate of 5 X 10^{-4} was used as a proxy for Red Colobus microsatellite mutation rates (Whittaker et al. 2003). Finally, to look for evidence of a bottleneck, the data were examined in M_p_val.exe (Garza and Williamson 2001) and BOTTLENECK 1.2.02 (Cornuet and Luikart 1996).

Results

Genotyping Errors, Linkage and Hardy Weinberg Equilibrium

Evidence of null alleles was found in only two loci, D1S207 and D17S1290, and there was also evidence of scoring errors in D17S1290 due to an excess of homozygotes. No evidence for allelic dropout was found. Out of 55 tests, only one locus pair (D2S1399 and C2A) was found to be significantly non-randomly associated at the Bonferroni corrected alpha of <0.001 (Rice 1989). Estimates of homozygosity ranged from 0.3 to 0.9 and heterozygosity from 0.5 to 0.85 (Table 5-1), and only one locus (D20S206) significantly deviated from Hardy-Weinberg equilibrium (HWeq) using the Bonferroni corrected p-value in POPGENE. To determine if this was consistent across different analyses, FSTAT was used to calculate Weir and Cockerham's (1984) Fis value with 11,000 randomizations and permutation tests. Here, D20S206 was not found to deviate from HWeq; interestingly, however, D17S1290 was found to deviate from Hardy-Weinberg in the FSTAT analysis.

Taken together, we found minimal evidence for null alleles and genetic disequilibrium (only one locus pair out of 55 tests). The most problematic locus, D17S1290, had high homozygosity levels and deviated from Hardy-Weinberg in only one analysis. Micro-Checker suggested that the heterozygote deficiency at this locus is due either to null alleles or to possible scoring errors, and is therefore due to

methodological, not biological, reasons. Because it was not consistently found to be out of HWeq, and to increase power of the analysis, this locus was retained in the analysis.

Number of Populations

RSTCALC gave an overall mean RHO of 0.036, which was not found to be significantly different from zero (p=0.142; Table 5-2). A mean Nm of 11.48 was calculated with confidence intervals that include negative numbers, which indicate that the number of migrants is so high that it cannot be quantified (Goodman 1997). Not all pairwise Rst (RHO) values overlapped zero (7 of 15), although none were found to be significantly different from zero (p>0.05 for all tests; Table 5-2). The results from STRUCTURE indicate strong support for one population (Figure 5-2). The one-population model had the highest likelihood for all 10 runs and the least amount of variation between them.

Because the 95% confidence intervals of all RHO values did not overlap zero between all comparisons, the genetic distances ($\Delta\mu^2$) from RSTCALC for each population were analyzed with a Neighbor Joining tree (data not shown). Here two clusters were apparent, one cluster included populations SC, Sebatoli and Mainaro (Cluster1), and a second cluster included LM, K-30, and Dura (Cluster2), showing some support for these two clusters as populations. Furthermore, the average $\Delta\mu^2$ is less for each group (Cluster1 = 0.62 and Cluster2 =1.3) relative to average $\Delta\mu^2$ across the groups (1.46). Therefore, these two populations were then used in a two-population model in MIGRATE.

Interestingly, for low values of M, the likelihood score for the one-population model is better than that of a two-population model (Bayes Factor >5) (Figure 5-3). As the prior distribution of M was increased, our likelihood score increased above that of

the one-population model, suggesting that a two-population model is more appropriate for the data. Because M is the number of migrants per generation scaled by mutation rate, if we assume the mutation rate is constant, then, as the prior distribution of M is increased, the number of migrants per generation (*Nm*) is increasing. Our likelihood scores increase above the one-population model when the values of *Nm* increase from 23.8 to 43.5 (Figure 5-3). The highest likelihood score was for the prior distribution of M going from 0-100, with an *Nm* of 171.47 migrants per generation. Because this level of migration suggests a single population even when the two-population model produces a higher likelihood score, all calculations of theta (θ), coalescence time and bottleneck estimates were done considering all of the dataset as one population.

Effective population size, Coalescent Point and Bottleneck

Our overall theta (θ) value estimated from MIGRATE, integrating across all loci, was 7.03. We used a microsatellite mutation rate of 5×10^{-4} (Whittaker et al. 2003) to calculate an effective population size (*Ne*) of 3,905 individuals. A ratio of 1:5 effective population size to census size has been suggested (Tempelton, 1998), and if we take that into account, our estimate of the census size is 19,527 Red Colobus in Kibale National Park. Using the one-population mode in MIGRATE we get a coalescence time of 21,708 generations translating to 108,542 years using a generation time of 5 years (Thomas Struhsaker pers com.). Finally no evidence of a bottleneck was found in the program M_P_Val (Garza and Williamson 2001) or BOTTLENECK (Cornuet and Luikart 1996).

Discussion

Since Red Colobus are one of the most endangered primates in the world (IUCN 2010) and the only area where they are not considered to be threatened is Kibale National Park, Uganda where census estimates range from 30,000 – 80,000 monkeys in the park (Chapman pers com.). Here we examine the number of populations in the park, and based on those results calculate the coalescent point of the population and look for evidence of population bottlenecks, to determine if the Red Colobus in Kibale have gone through a population reduction.

Number and size of populations

When we used MIGRATE to compare a one-population model against a two-population model we found that, at low migration rates, a one-population model best describes the data, and, at higher migration rates, a two-population model is superior. Our likelihood scores for the two-population model increase above those for the one-population model when the values of *Nm* (number of migrants between the populations) are high enough to keep two populations genetically similar (Mills and Alendorf 1996). If this trend were to continue, the highest likelihood score would be that of a two-population model with infinite numbers of migrants between the two populations essentially one panmictic population. Therefore, based on the results from STRUCTURE, RST and MIGRATE, we propose that all of the Red Colobus in Kibale constitute one large population.

Here, we find the effective population size (*Ne*) of the Red Colobus is around 4,000 monkeys translating to a census estimate of 20,000 monkeys and correlates to a density of 23 to 100 monkeys per kilometer. Interestingly, two previous census estimates of Red Colobus in Kibale range from 30,312 to 80,675; however, it is thought

that the true number lies on the lower end of those two estimates (Chapman pers. com.). Our 95% confidence intervals put our estimate of Red Colobus between 6,204 and 39,400, within and below range of the census. We know that this population of Red Colobus is female-biased and male–bonded, so the contribution to the next generation is not equal between the sexes, which would make our estimates of *Ne* smaller compared to the census population sizes, and may explain why our numbers are on the lower end of the census estimates. With this high density of primates, it is therefore not surprising that there is no genetic structure across the park.

History of Red Colobus in Kibale National Park

Although Red Colobus are sensitive to habitat change (Struhsaker 1997), they have an extremely large population in Kibale, which may suggest that there has not been a lot of disturbance over the long term. Here we did not find any evidence of a population bottleneck and preliminary runs in BEAST also suggest that this population has been stable over its history (data not shown). Based on our MIGRATE runs, we get a coalescence point of 108,542 years, and this lineage of Red Colobus is thought to have diverged from its sister taxon around 600,000 years ago (Ting 2008). Taken together, these results suggest that there has been a stable population of Red Colobus in Kibale for over 100 thousand years, with no evidence of a population decline.

Interestingly, there is evidence that KNP has a history of some disturbance. For example, logging has recently been an issue (Chapman et al. 2000), and evidence from the recruitment rate of trees in Kibale suggests that there was a large-scale disturbance sometime in the more distant past (100-3,000 years ago, Taylor et al. 1999, Chapman et al. In Prep). Some possible explanations are drought (Bessems et al 2008), elephant activity, or human clearing. However, based on our results, these disturbances have not

been at the scale where they affected the Red Colobus, likely due to a large tract of land being available to them. Because they are so sensitive to habitat change these results suggest that Kibale National Park may have been a refugium for the Red Colobus over the last 100,000 years and therefore has likely been a refugium for other species of animals during that time as well.

Finally, with a small sampling of COI from lice from the two large troops I find a high level of population structure, even though all of these lice have been identified to the same species. Overall, average p-distances range from 0-0.15 (ave = 0.049, N=6) on a single host, in the SC 0-0.16 (ave= 0.81, N=6) and LM 0-0.004 (ave = 0.0013). Up to 15% difference on a single host suggests quite a bit of structure exists in the parasite populations, which have all been identified as the same species.

Taken together we find that there is no population structure in the Red Colobus in KNP and find some preliminary evidence of population structure in their lice. Because the lice only move between host individuals during direct host-to-host contact these results suggest that perhaps the lice are not moving around as much as their hosts. Alternatively, the lice may maintain population structure in the absence of host population structure. The next step of this work will be to examine the microsatellite dataset of the lice to look for gene flow between lice to determine the level of population structure in the parasites.

Table 5-1. Microsatellite loci genotyped for Red Colobus (*Piliocolobus tephrosceles*) from Kibale National Park. N – number of individuals, Na – Number of alleles, Ne – effective number of alleles, Ho – observed heterozygosity, He – expected heterozygosity calculated in POPGENE 1.3.1. Fis values were calculated in FSTAT 2.9.3.2. * indicates significance at the bonferroni corrected p-values (0.0045). ** indicates overall FIS value for all loci calculated in GDA with 95% confidence intervals which overlap zero suggesting overall no Homozygote or Heterzygote excess when all loci are accounted for. No locus was found to deviate from Hardy-Weinberg equilibrium using both the the G-test and the Fis permutation test suggesting these loci are suitable for other population genetic analysis.

Locus	N	Na	Ne	Ho	He	Fis
D14S306	84	6	4.25	0.76	0.77	0.01
D3S1766	82	9	6.31	0.76	0.85	0.11
D2S1399	84	10	6.57	0.91	0.85	-0.06
D7S1817	78	9	4.88	0.69	0.80	0.14
D20S206*	79	7	5.08	0.79	0.81	0.03
D8S260	82	4	1.91	0.54	0.48	-0.12
D8S165	85	3	2.07	0.55	0.52	-0.63
D1S207	85	14	6.12	0.74	0.84	0.12
D17S1290	80	4	1.66	0.30	0.40	0.25*
C2A	85	7	2.39	0.55	0.59	0.06
D5S1457	82	8	3.62	0.67	0.73	0.08
Average	82	7.36	4.08	0.66	0.69	0.047**
						(-0.0024 0.0987)

Table 5-2. Results from RSTCALC. Each of the 6 populations were compared and RHO estimates calculated. Nm - number of migrants per generation included many negative numbers which suggest that the number of migrants is so high that it cannot be quantified (Goodman 1997). * indicate where the 95% confidence intervals did not overlap zero for RHO and include negative numbers for Nm.

Group 1	Group 2	RHO	*Nm*	P value
SM	LM	0.000	-516.56	0.47
SM	K-30	-0.014	-17.89	0.55
SM	DURA	0.045*	5.28*	0.07
SM	SEB	0.013	19.01	0.23
SM	MAIN	0.036*	6.70*	0.13
LM	K-30	-0.040	-5.94	0.97
*LM	DURA	0.067*	3.48*	0.06
LM	SEB	0.013	19.12	0.33
LM	MAIN	0.063*	3.74*	0.10
K-30	DURA	0.052*	4.55*	0.09
*K-30	SEB	0.005	48.09	0.31
K-30	MAIN	0.062	3.76	0.23
*DURA	SEB	0.082*	2.81*	0.10
DURA	MAIN	0.082*	2.83*	0.21
SEB	MAIN	0.049	4.90	0.32
TOTAL	- PT–	-0.015	-17.42	0.42
TOTAL	95% BS	(-0.008 - 0.1117)	(-80.48 - 94.77)	

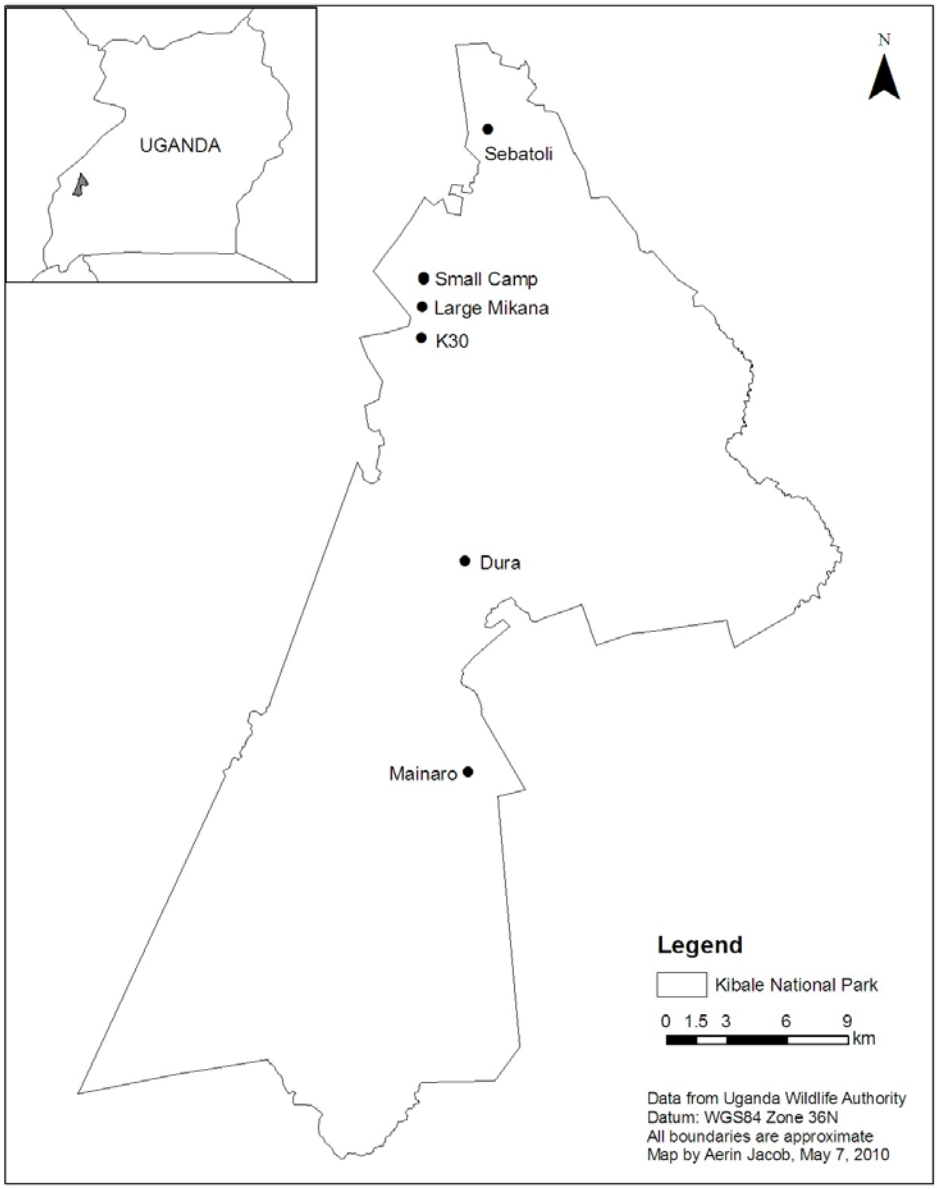

Figure 5-1. Map of Kibale National Park (KNP). Map shows orientation in Uganda and black points indicate populations of Red Colobus (*Piliocolobus*) that were sampled in this study.

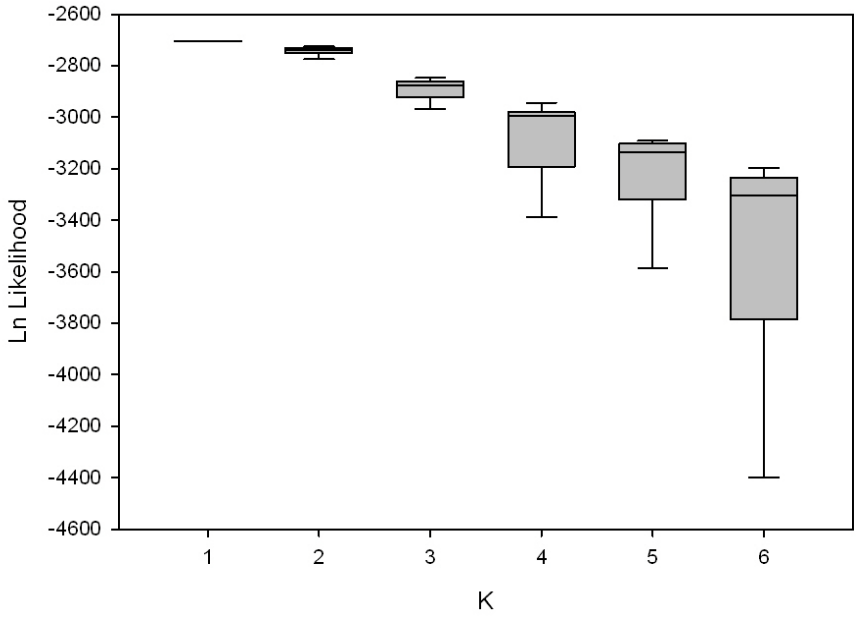

Figure 5-2. Log likelihood scores for output from Structure. It was run using an admixture model and correlated allele frequencies among populations. The number of populations (k) was tested from 1 to 6 for 1 X 10^6 generations after a burnin of 1 X 10^5 generations. Each k was tested ten times to examine the variation between runs. The likelihood scores for each run were examined visually for stationarity.

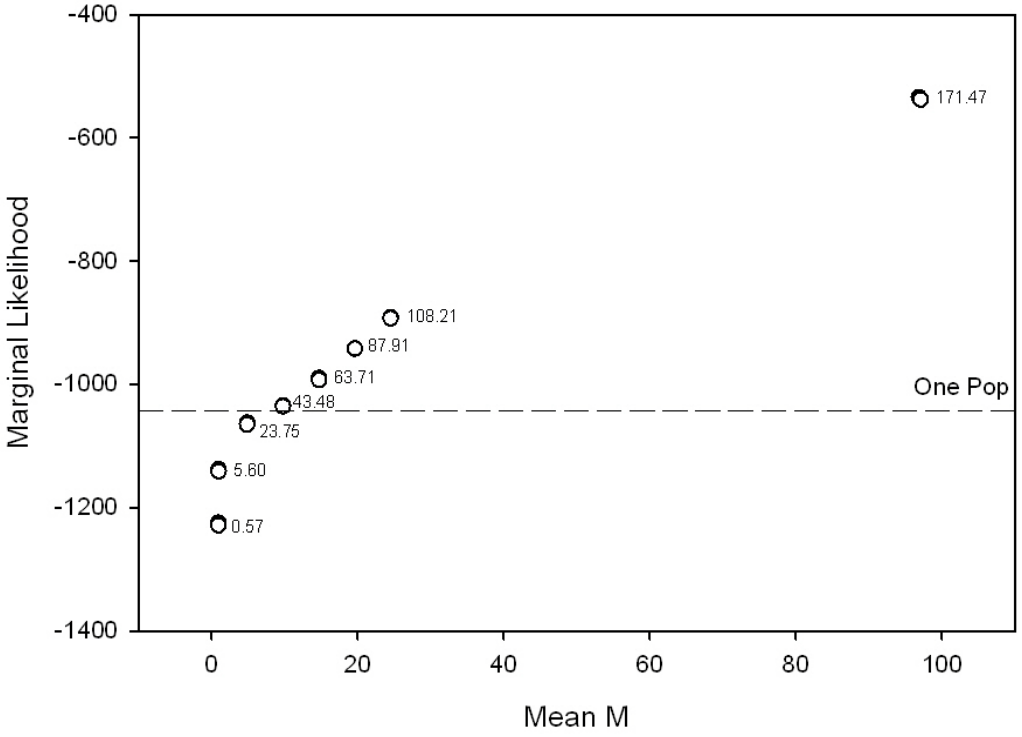

Figure 5-3. Log likelihood scores for the program Migrate. Dotted line indicates likelihood score for one population run for 10 X 10^6 generations with a burnin of 10%. Two populations were tested using different prior distributions of M (m/μ, migrants by mutation rate) between the two populations from 0-1, 0-5, 0-10, 0-15, 0-25, and 0-100 each was run for 10 X 10^6 generations with 10% burnin and repeated twice to examine variability between runs. Numbers by nodes are number of migrants per generation (*Nm*). This graph indicates that for low values of M the one-population model has a higher likelihood score, but for higher values of M the two-population model has a better likelihood score. As M increases the number of migrants between the two populations is increasing effectively making the two populations panmictic, and therefore we conclude that here is one population.

CHAPTER 6
CONCLUSIONS

As stated at the beginning of this discussion, the evolutionary history of an organism is strongly affected by its interactions with other species in a shared environment (e.g. predator-prey interactions), and these interactions may affect the entire evolutionary trajectory of a lineage. I have examined two associates of sucking lice, their primary endosymbionts and their mammalian hosts to determine the phylogenetic as well as population genetic patterns of these two associates.

Part I: Lice and their Primary Endosymbionts

Primary endosymbiotic bacteria (p-endosymbionts) are thought to be partially responsible for the incredible diversification of insects. P-endosymbionts are found in insects that specialize on nutrient-poor diets, where they supplement their insect hosts' diet with nutrients (Buchner 1965). It has been suggested that the acquisition of a bacterial lineage in insect ancestors has played a key role in the species diversification of some insect groups. Endosymbiotic bacteria cannot live outside of the insect hosts and are transmitted from mother to offspring through the eggs. Some insect groups have had the same lineage of endosymbiont for millions of years, such as aphids and their endosymbionts in the genus *Buchnera*. These long-term associations suggest a long coevolutionary history between the insect-bacteria partners.

The primary-endosymbiont in the human head louse was first discovered over three hundred years ago (Hooke 1664), but was only recently described and formally named (Allen et al., 2007, Fukatsu et al. 2007). This was the first p-endosymbiont to be identified from sucking lice. It is thought that all sucking lice have a p-endosymbiont that synthesizes vitamins lacking in the louse's specialized diet of mammalian blood

(Buchner, 1965). I hypothesized that, like aphids, all sucking lice would have the same genus of endosymbiont. I found, however, that there have been as many as 14 different genera of endosymbionts in sucking lice (Allen et. al, 2009, Allen et al. In Prep.). While multiple independent origins of endosymbiotic bacteria is one possible explanation for this pattern, because lice cannot survive without endosymbionts (Buchner 1965), it is more likely that the ancestor of all lice acquired an endosymbiont, and among some louse groups, the endosymbiont has been out-competed and replaced by new bacterial lineages over time.

Bacterial endosymbionts are weakened over time by a high rate of irreversible and harmful mutations due to their small population sizes and lack of recombination (Moran, 1996). It is thought that frequent mutations will eventually degrade the p-endosymbiont genome and diminish the reproductive health of the bacteria (Moran, 1996) enabling other bacteria to outcompete the p-endosymbionts and replace them. Interestingly, only a few groups of insects have shown p-endosymbiont replacement similar to that found in sucking lice (Allen et. al, 2009). Further work has verified that the rate of molecular evolution for the endosymbiont in human sucking lice is an order of magnitude faster than most primary endosymbionts, suggesting that they are accumulating harmful mutations at a faster rate than other p-endosymbionts and likely being outcompeted and replaced faster as well (Allen et al., 2009). The cause of this unusual mutation rate is unknown, although it is interesting that the mutation rate of the louse itself is high relative to other insects (Yoshizawa and Johnson 2003).

Part II: Sucking Lice and their Mammalian Hosts

Organisms involved in obligate relationships are affected on both a long-term species-level time scale and short-term population-level time scale (Clayton and

Johnson 2003). For example, the size of a parasite population is strictly tied to the social structure, dispersal ability and population size of its host. In the case of lice and mammals, lice primarily move between hosts during direct host-to-host contact. Therefore, louse population dynamics are particularly constrained by the social structure of their hosts (e.g. which hosts come into contact and how often they come into contact). The effect of mammalian host population dynamics on louse population dynamics remains poorly understood.

A louse host, the Red Colobus monkey, breeds and forages in stable social groups called troops. I have determined that there is a high level of migration and gene flow across Kibale National Park, the number of host individuals is large, in particular for primates, and is around 20,000 – 40,000 individuals, providing a large host population for their sucking ice, as well as opportunities for movement across the park.

Interestingly, with a small sampling of COI from lice from the two large troops I find a high level of population structure, even though all of these lice have been identified to the same species. Overall, average p-distances range from 0-0.15 (ave = 0.049, N=6) on a single host, in the SC 0-0.16 (ave= 0.81, N=6) and LM 0-0.004 (ave = 0.0013). Up to 15% difference on a single host suggests quite a bit of structure exists in the parasite populations, which have all been identified as the same species. The next step of this work will be to examine the microsatellite dataset of the lice to look for gene flow between lice on different hosts and determine how many populations of parasites exist on one large population of hosts.

This work will provide insight into long-term changes in louse population structure, and we may in fact find that there is genetic structure in the parasite populations in the

absence of host population structure. This would then suggest that the lice are not able to move as easily between host populations, and may be constrained by the amount of physical interactions between host troops. My future goal is to examine obligate relationships on both evolutionary and population level time scales and examine how population level processes affect long-term evolutionary patterns.

LIST OF REFERENCES

Adams, T.S. 1999. Hematophagy and hormore release. Annals of the Entomological Socitey of America **92**:1-13.

Allen, J.M., D.L. Reed, M.A. Perotti, and H.R. Braig. 2007. Evolutionary relationships of "*Candidatus* Riesia spp.," endosymbiotic Enterobacteriaceae living within hematophagous primate lice. Applied and Environmental Microbiology **73**:1659-1664.

Andersson, S.G.E., and C.G. Kurland. 1998. Reductive evolution of resident genomes.Trends in Microbiology **6**:263-268.

Arandjelovic, M., K. Guschanski, and G. Schubert. 2009. Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. Molecular Ecology Resources **9**:28–36.

Atwood. K.C., L.K. Schneider, and F.J. Ryan. 1951. Selective mechanisms in bacteria. Cold Spring Harbor Symposia on Quantitative Biology **16**:345-355.

Aschner, M. 1934. Studies on the symbiosis of the body louse. I. Elimination of the symbionts by centrifugation of the eggs. Parasitology **26**:309-314.

Aschner, M., and E. Ries. 1933. Das verhalten der kleiderlaus beim ausschaltender Zoomorphology **26**:529.

Askoy, S., A.A. Pourhosseini, and A. Chow. 1995. Mycetome endosymbionts of tsetse flies constitute a distinct lineage related to *Enterobacteriaceae*. Insect Molecular Biology **4**:15-22.

Bandi, C., M. Sironi, G. Damiani, L. Magrassi, and C.A. Nalepa. 1995. The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. Proceedings of the Royal Society of London Series B **259**:293-299.

Beerli, P. 2006. Comparison of Bayesian and maximum likelihood inference on population genetic parameters. Bioinformatics **22**:341-345.

Beerli, P. 2008. MIGRATE version 3.0 – a maximum likelihood and Bayesian estimator of gene flow using the coalescent.

Beerli, P., M. Palczewski 2010. Unified framework to evaluate panmixia and migration direction amond multiple sampling locations. Genetics10.1534/genetics.109.112532.

Bergstrom, C.T. and J. Pritchard. 1998. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. Genetics **149**:2135-2146.

Bessems, I., D. Verschuren, J.M. Russell, J. Has, F. Mees, B.F. Cumming. 2008. Paleolimnological evidence for widespread late 18th century drought across equatorial East Africa. Palaeogeography, Palaeoclimatology, Palaeoecology **259**:107-120.

Beuning, K.R., K. Kelts, J. Russell, and B.B. Wolfe. 2002. Reassessment of Lake Victoria-Upper Nile River paleohydrology from oxygen isotope records of lake sediment cellulose. Geology. **30**:559-562.

Boutin-Ganache, I., M. Raposo, M. Raymond, C.F. Deschepper. 2001. M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. Biotechniques **31**:24-28.

Braig, H.R., B.D. Turner, and M.A. Perotti. 2009. Symbiotic Rickettsia. In: Bourtzis K,Miller TA, editors. Insect Symbiosis 3. Boca Raton: Taylor and Francis. pp. 221-252.

Brynnel, E. U., C. G. Kurland, N. A. Moran, and S. G. Andersson. 1998. Evolutionary rates for tuf genes in endosymbionts of aphids. Molecular Biology and Evolution **15**:574-582.

Buchner, P. 1965. Endosymbiosis of animals with plant micro-organisms. Interscience, New York.

Busse, 1977. Chimpanzee predation as a possible factor in the evolution of red colobus monkey social organization. Evolution **31**:907-911.

Butcher, D. 1995. Mullers Ratchet, Epistasis and Mutation Effects. Genetics **141**:431-437.

Buxton, P. A. 1946. The Louse, 1st ed. Williams & Wilkins Co., Baltimore, MD.

Chapman, C.A., L.J. Chapman, R.W. Wrangham, G. Isabirye-Basuta, K. Ben-David. 1997. Spatial and temporal variability in the structure of a tropical forest. African Journal of Ecology **35**:287-302.

Chapman, C.A., S.R. Balcomb, T. Gillespie, J. Skorupa, T.T. Struhsaker T.T. 2000. Long-term effects of logging on African primate communities: A 28 year comparison from Kibale National Park, Uganda. Conservation Biology **14**:207-217.

Canback, B., I. Tamas, and S.G.E. Andersson. 2004. A phylogenomic study of endosymbiotic bacteria. Molecular Biology and Evolution **21**:1110-1122.

Chapman, C.A. T. Webb, R. Fronstin, M.D. Wasserman, and A.M. Santamaria. 2005. Assessing dietary protein of colobus monkeys through feacal sample analysis: A tool to evaluate habitat quality. African Journal of Ecology 43:276-278.

Charlesworth, D., M.T. Morgan, and B. Charlesworth. 1993. Mutation accumulationin finite outbreeding and inbreeding populations. Genetic Research **61**:39-56.

Clark, M.A., N.A. Moran, and P. Baumann. 1999. Sequence evolution in bacterial endosymbionts having extreme base compositions. Molecular Biology and Evolution **16**:1586-1598.

Cornuet, J.M., and G. Luikart 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001-2014.

Corsaro, D., D. Venditti, M. Padula, and M. Valassina. 1999. Intracellular life. Critical Reviews in Microbiology **25**:39-79.

Cunningham, C. W., H. Zhu, and D. M. Hillis. 1998. Best-fit maximum likelihoodmodels for phylogenetic inference: empirical tests with known phylogenies. Evolution **52**:978-987.

Dale, C., S.A. Young, D.T. Haydon, S.C. Welburn. 2001. The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. Proceedings of the National Academy of Sciences USA **98**:1883-1888.

Dale, C., M. Beeton, C. Harbison, T. Jones, and M. Pontes. 2006. Isolation, pureculture, and characterization of "*Candidatus* Arsenophonus arthropodicus," anintracellular secondary endosymbiont from the hippoboscid louse fly Pseudolynchia canariensis. Applied and Environmental Microbiology **72**:2997-3004.

Dedeine, F. C., M. Bandi, M. Boulétreau, and L.H. Kramer. 2003. Insights into *Wolbachia* obligatory symbiosis. In: Bourtzis K, Miller T, editors. Insect Symbiosis CRC Press: LLC. 267-282.

Delmotte, F., C. Rispe, J. Schaber, F.J. Silva, and A. Moya. 2006. Tempo and mode of early gene loss in endosymbiotic bacteria from insects. BMC Biology. **6**:56.

Degnan, P.H., A.B. Lazarus, C.D. Brock, and J.J. Wernegreen. 2004. Host-symbiont stability and fast evolutionary rates in an ant-bacterium association: Cospeciation of *Camponotus* species and their endosymbionts, *Candidatus* Blochmannia. Systematic Biology **53**:95-110.

Douglas, A. E. 1989. Mycetocyte symbiosis in insects. Biological Reviews, Cambrige Philosophical Society **64**:409-34.

Durden, L. A., and G. G. Musser. 1994. The sucking lice (Insecta: Anoplura) of theworld: A taxonomic checklist with records of mammalian hosts and geographical distributions. Bulletin of the American Museum of Natural History **218**:90.

Eberle, M. W., and D. L. McLean. 1983. Observation of symbiote migration in Human body lice with scanning and transmission electron microscopy. Canadian Journal of Microbiol. **29**:755-762.

Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gil, K.E. Nelson and D.A. Relman. 2005. Diversity of the human intestinal microbial flora. Science **308**:1635-1638.

Felsenstein, J. 1974. Evolutionary Advantage of Recombination. Genetics **78**:737-756.

Fuchs, B. M., G. Wallner, W. Beisker, I. Schwippl, W. Ludwig, and R. Amann.1998. Flow cytometric analysis of the in situ accessibility of *Escherichia coli* 16SrRNA for fluorescently labeled oligonucleotide probes. Applied and Environmental Microbiology **64**:4973-4982.

Fukatsu T., T. Hosokawa, R. Koga, N. Nikoh, T. Kato, S. Hayama, H. Takefushi and I. Tanaka. 2009. Intestinal endocelluar symbiotic bacterium of the macaque louse Pedicinus obtusus: Distinct endosymbiont origins in anthropoid primate lice and the old world mokey louse. Applied and Environmental Microbiology **75**:3796-3799.

Garza, J.C. and E.G. Williamson 2001. Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10:305-318.

Geiger, A., S. Ravel, R. Frutos, and G. Cuny. 2005. *Sodalis glossinidius* (Enterobacteriaceae) and vectorial competence of *Glossina palpalis gambiensis* and *Glossina morsitans morsitans* for *Trypanosoma congolense* savannah type. Current Microbiology **51**:35.

Gil, R., A. Latorre, and A. Moya. 2004. Bacterial endosymbionts of insects: insights from comparative genomics. Environmental Microbiology **6**:1109-1122.

Goldberg,T.L., D.M. Sintasath, C.A. Chapman, K.M. Cameron, W.B. Karesh, S. Tang, N.D. Wolfe, I.B. Rwego, N. Ting, and W.M. Switzer. Co-infection of Ugandan red colobus (*Procolobus [Piliocolobus] rufomitratus tephrosceles*) with novel, divergent delta-,lenti- and spuma- retroviruses. Journal of Virology (In Press).

Goldberger, J., and J. F. Anderson. 1912. The transmissions of typhus fever, with special reference to transmission by the head louse (*Pediculus capitis*). Public Health Reports **27**:297-308.

Goldman, N., J.P. Anderson, A.G. Rodrigo. 2000. Likelihood-based tests oftopologies in phylogenetics. Systematic Biology **49**:652-670.

Goodman, S.J. 1997. RST CALC2.2: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. Molecular Ecology **6**:881-885.

Goossens, B., L. Chikhi, S.S. Utami, J. Ruiter, and M.W. Bruford. 2000. A multi-samples, multi-extracts approach for microsatellite analysis of faecal samples in an arboreal ape. Conservation Genetics 1:157-162.

Goudet, J. 2002. FSTAT Software, v. 2.9.3.2. http://www2.unil.ch/popgen/softwares/fstat.htm

Haldane, J.B.S. 1937. The effect of variation on fitness. American Naturalist **71**:337-349.

Harris, T.R., D. Caillaud, C.A. Chapman, and L. Vigiland, 2009. Neither genetic nor observational data alone are sufficient for understanding sex-biased dispersal in a social group-living species. Molecular Ecology **18**:1777-1790.

Hein, J., M. Schierup, C. Wiuf. 2004. Sequence Variation, Genealogies and Evolution. Oxford University Press, New York, New York.

Hooke, R. 1664. Micrographia or some physiological description of minute bodies made by magnifying glasses with observations and iquiries thereupon. Council of the Royal Society of London for Improving Natural Knowledge, London, United Kingdom.

Hypsa, V., and J. Krizek. 2007. Molecular evidence for polyphyletic origin of the primary symbionts of sucking lice (Phthiraptera, Anoplura). Microbial Ecolology **54**:242-251.

Huelsenbeck, J.P., and B. Rannala. 1997. Phylogenetic methods come of age:Testing hypotheses in an evolutionary context. Science **276**:227-232.

Huelsenbeck, J. P., and F. Ronquist. 2001. MR-BAYES: Bayesian inference of Phylogeny. Bioinformatics **17**:754-755.

Hudson, R.R. 1990. Gene genealogies and the coalescent process. In: Oxford Surveys in Evolutionary Biology. Oxford University Press United Kingdom

Hurst, L., and G. Mcvean. 1996. Evolutionary genetics. and scandalous symbionts.Nature **381**:742-742.

IUCN 2010. IUCN Red List of Threatened Species. Version 2010.1.

Ihlen, P.G., and S. Ekman. 2002. Outline of phylogeny and character evolution inRhizocarpon (Rhizocarpaceae, lichenized Ascomycota) based on nuclear ITS and mitochondrial SSU ribosomal DNA sequences. Biological Journal of the Linnean Society **77**:535-546.

Johnson, K. P., R. H. Cruickshank, R. J. Adams, V. S. Smith, R. D. M. Page, and D. H. Clayton. 2003. Dramatically elevated rate of mitochondrial substitution in lice (Insecta: Phthiraptera). Molecular Phylogenetics and Evolution **26**:231-242.

Khuner, M. 2008. Coalescent Geneology samplers: windows into population history. Trends in Ecology and Evolution **24**:86-93.

Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA-sequence data, and the branching rrder In Hominoidea. Molecular Ecology **29**:170-179.

Kittler, R., M. Kayser, and M. Stoneking. 2003. Molecular evolution of *Pediculus humanus* and the origin of clothing. Current Biology **13**:1414-1417.

Klasson, L., S.G.E. Andersson. 2004. Evolution of minimal-gene-sets in hostdependent bacteria. Trends in Microbiology **12**:37-43.

Kondrashov, A.S. 1994. Mullers ratchet under epistatic selection. Genetics **136**:1469-1473.

Kyei-Poku, G. K., D. D. Colwell, P. Coghlin, B. Benkel, and K. D. Floate. 2005. On the ubiquity and phylogeny of *Wolbachia* in lice. Molecular Ecology **14**:285-294.

Lefévre, C., H. Charles, A. Vallier, B. Delobel and B. Farrell B. 2004. Endosymbiont phylogenesis in the *Dryophthoridae* weevils: Evidence for bacterial replacement. Molecular Biology and Evolution **21**:965-973.

Lehane, M.J. 2005. The biology of blood feeding insects, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.

Leo, N. P., and S. C. Barker. 2005. Unravelling the evolution of the head lice andbody lice of humans. Parasitological Research **98**:44-47.

Lerat E, V. Daubin, and N.A. Moran. 2003. From gene trees to organismal phylogenyin prokaryotes: The case of the gamma-proteobacteria. PLoS Biology **1**: 101-109.

Light, J.E., V.S. Smith, J.M Allen, L.A. Durden, and D.L. Reed. 2010. Evolutionary history of mammalian sucking lice (Pthiraptera: Anoplura). *In Review*: BMC Biology.

Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W.Ginhart, O. Gross, S.Grumann, S. Hermann, R. Jost, A. Konig, T. Liss, R.Lussmann, M. May, B. Nonhoff, B.Reichel, R. Strehlow, A. Stamatakis, N.Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K. H. Schleifer. 2004. ARB: a software environment for sequence data. Nucleic Acids Research **32**:1363-1371.

Lutzoni, F., and M., Pagel. 1997. Accelerated evolution as a consequence oftransitions to mutualism. Proceedings of the National Academy of Sciences, USA **94**:11422-11427.

Lynch, M., and W. Gabriel. 1990. Mutation load and the survival of small populations. Evolution **44**:1725-1737.

Maddison, W.P., and D.R. Maddison. 2000. MacClade: Analysis of phylogeny and character evolution. Sunderland, Massachusetts: Sinauer Associates.

Mittermeier, R. A., J. Ratsimbazafy, A.B. Rylands, L. Williamson, J.F. Oates, D. Mbora, J.U. Ganzhorn, E. Rodríguez-Luna, E. Palacios, E.W. Heymann, M.C.M. Kierulff, L. Yongcheng, J. Supriatna, C. Roos, S. Walker, and J.M. Aguiar. 2007. Primates in Peril: The World's 25 Most Endangered Primates, 2006–2008. Primate Conservation **22**:1-40.

Mills, S.,L., and F.W. Allendorf. 1996. The one-migrant-per-generation rule in conservation and management. Conservation Biology **10**:1509-1518.

McCutcheon, J.P., and N.A. Moran. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. Proceedings of the National Academy of Sciences, USA **104**:19392-19397.

Moran, N.A., M.A. Munson, P. Baumann, and H. Ishikawa. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proceedings of the Royal Society of London Series B **253**:167-171.

Moran, N., and P. Baumann. 1994. Phylogenetics of cytoplasmically inherited microorganisms of arthropods. Trends in Ecology and Evolution **9**:15-20.

Moran, N. A., C. D. von Dohlen, and P. Baumann. 1995. Faster evolutionary rates in endosymbiotic bacteria than in cospeciating insect hosts. Journal of Molecular Ecology **41**:727-731.

Moran, N.A. 1996. Accelerated evolution and Muller's rachet in endosymbioticbacteria. Proceedings of the National Academy of Sciences, USA **93**:2873-2878.

Moran, N.A., and P. Baumann. 2000. Bacterial endosymbionts in animals. Current Opinions in Microbiology **3**:270-275.

Moran, N.A., and J.J. Wernegreen. 2000. Lifestyle evolution in symbiotic bacteria: Insights from genomics. Trends Ecolology and Evolution **15**:321-326.

Moran, N.A., P. Tran, and N.M. Gerardo. 2005. Symbiosis and insect diversification: An ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. Applied and Environmental Microbiology **71**: 8802-8810.

Moran, N. A., and H. E. Dunbar. 2006. Sexual acquisition of beneficial symbionts in aphids. Proceedings of the National Academy of Sciences, USA **103**:12803-12806.

Moran, N.A. 2007. Symbiosis as an adaptive process and a source of phenotypic complexity. Proceedings of the National Academy of Sciences, USA **104**:8627-8633.

Muller, H.J. 1964. The relation of recombination to mutational advance. Mutation Research **1**:2-9.

Murray, R. G. E., and K. H. Schleifer. 1994. Taxonomic Notes - a proposal for recording the properties of Pputative taxa of prokaryotes. International Journal of Systematic Bacteriology **44**:174-176.

Murray, R. G. E., and E. Stackebrandt. 1995. Taxonomic Note - Implementation of the provisional status *Candidatus* for incompletely described prokaryotes. International Journal of Systematic Bacteriology **45**:186-187.

Nakabachi, A., A. Yamashita, H. Toh, H. Ishikawa, and H.E. Dunbar. 2006. The 160 kilobase genome of the bacterial endosymbiont *Carsonella*. Science **314**:267-267.

Nsubuga, A.M., M.M. Robbins, A.D. Roeder. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Molecular Ecology **13:**2089–2094

Oates, J.F. 1987. Food distribution and foraging behavior. In Primate societies, ed. B.B. Smuts, D.L. Cheney, R.M. Seyfarth, R.W. Wrangam and T.T. Struhsaker, 169-209. Chicago: University of Chicago Press.

Oates, J.F. and A.G. Davies. 1994. Conclusions: Past, present and future of the colobines. In Colobine monkeys: Their ecology, behavior and evolution, ed. A.G. Davies and Oates, J.F. 45-73. Cambridge: Cambridge University Press.

Oates, J.F. 1996b. Habitat alteration, hunting and the conservation of folivorous primates in African forests. Australian Journal of Ecology. **21**:1-9.

Oates, J.F., M. Abedi-Larty, S.W. McGraw, T.T. Struhsaker, and G.H. Whitesides. 2000. Conservation Biology **14**:1526-1532.

Ochman, H., S. Elwyn, and N.A. Moran. 1999. Calibrating bacterial evolution. Proceedings of the National Acadademy of Sciences. USA **96**:12638-12643.

Ochman, H., and N.A. Moran. 2001. Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. Science **292**:1096-1098.

O'Fallon, B. 2008. Population structure, levels of selection, and the evolution of intracellular symbionts. Evolution **62**:361-373.

Perez-Brocal, V., R., Gil, S. Ramos, A. Lamelas, and M. Postigo. 2006. A smallmicrobial genome: The end of a long symbiotic relationship? Science **314**:312-313.

Perotti, M. A., S. S. Catala, A. D. Ormeno, M. Zelazowska, S. M. Bilinski, and H.R. Braig. 2004. The sex ratio distortion in the human head louse is conserved overtime. BMC Genetics **5**:10.

Perotti, M.A., H.K. Clarke, B.D. Turner and H.R. Braig. 2006. Rickettsia as obligate and mycetomic bacteria Federation of American Societies for Experimental Biology **20**:2372-2374.

Perotti, M. A., J. M. Allen, D. L. Reed, and H. R. Braig. 2007. Host-symbiont interactions of the primary endosymbiont of human head and body lice. Federation of American Societies for Experimental Biology **21**:1058-1066.

Perotti, M.A., D.L. Reed and H.R. Braig. 2009. Endosymbionts of lice. In: Bourtzis K, Miller TA, editors. Insect Endosymbionts 3. Boca Raton: Taylor andFrancis. pp. 205-220.

Poinar, G.O. 1992. Life in Amber: Stan. Univ. Press.

Polak, M. 1996. Ectoparasitic effects on host survival and reproduction: The Drosophila Macrocheles Association. Ecology **77:**1379-1389.

Posada, D., and K. A. Crandall. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics **14**:817-818.

Posada, D., and T.R. Buckley. 2004. Model selection and model averaging in phylogenetics: Advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Systematic Biology **53**:793-808.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics **155**:945-959.

Puchta, O. 1955. Experimentelle untersuchungen uber die symbiose der kleiderlaus *Pediculus vestimenti* Burmeister. Z Parasitenkd **17**:1.

Reed, D. L., and M. S. Hafner. 2002. Phylogenetic analysis of bacterial communities associated with ectoparasitic chewing lice of pocket gophers: A culture-independent approach. Microbial Ecology **44**:78-93.

Reed, D. L., V. S. Smith, S. L. Hammond, A. R. Rogers, and D. H. Clayton. 2004. Genetic analysis of lice supports direct contact between modern and archaichumans. PLoS Biology **2**: 1972-1983.

Reed DL, J.E. Light, J.M. Allen and J.J. Kirchman. 2007. Pair of lice lost or Parasites regained: the evolutionary history of anthropoid primate lice. BMC Biology **5**:1-11.

Ries, E. 1931. Die symbiose der lause und federlinge. Zeitschrift fuer Morphologie und Oekologie der Tiere **20**:233-367.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution **43**:223-225.

Robinson, D., N. Leo, P. Prociv, and S. C. Barker. 2003. Potential role of headlice, *Pediculus humanus capitis*, as vectors of *Rickettsia prowazekii*. Parasitological Research **90**:209.

Saito, T., and S. Bjornson. 2006. Horizontal transmission of a microsporidium fromThe convergent lady beetle, Hippodamia convergens Guerin-Meneville (Coleoptera: Coccinellidae), to three coccinellid species of Nova Scotia. Biological Control **39**:427.

Sanderson, M.J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics **19**:301-302.

Sasaki-Fukatsu, K., R. Koga, N. Nikoh, K. Yoshizawa, S. Kasai, M. Mihara, M.Kobayashi, T. Tomita, and T. Fukatsu. 2006. Symbiotic bacteria associated with Stomach Discs of Human Lice. Applied and Environmental Microbiology **72**:7349-7352.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology **18**:233-234.

Shimodaira. H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution **16**:1114-1116.

Southworth, J., H. Hartter, M.W. Binford, A. Goldman, C.A. Chapman, L.J. Chapman, P. Omeja, and E. Binford 2010. Parks, people and pixels: evaluating landscape effects of and East Afrian national park on its surroundings. Tropical Conservation Science **3**:122-142.

Stewart, C.S., H.J. Fligt, M.P. Bryant. 1997. The rumen bacteria In: The rumen microbial ecosystem. P.N Hobson and C.S. Stewart (eds.) Blackie Academic and Professional.

Struhsaker, T.T. 1975. The red colobus monkey. University of Chicago Press.

Struhsaker, T.T. and M. Leakey. 1990. Prey selectivity by crowned hawk-eagles on monkeys in Kibale Forest, Uganda. Behavioral Ecology and Sociobiology **26**:435-443.

Struhsaer, T.T. 1997. Ecology of an African rain forest: logging in Kibale and the conflict between conservation and exploitation. University of Florida Press, Gainesville.

Swofford, D. L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and othermethods). 4.0 b10 ed. Sinauer, Sunderland, MA.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics **105**:437-460.

Tamas, I., L. Klasson, B. Canback, A.K. Naslund, A.S. Eriksson. 2002. 50 millionYears of genomic stasis in endosymbiotic bacteria. Science **296**:2376-2379.

Takiya, D.M., P.L. Tran, C.H. Dietrich, and N.A. Moran. 2006. Co-cladogenesis spanning three phyla: leafhoppers (Insecta: Hemiptera: Cicadellidae) and their dual bacterial symbionts. Molecular Ecology **15**:4175-4191.

Taylor, D., R.A. Marchant, P. Robertshaw. 1999. A sediment-based history of medium altitude forest in central Africa: A record from Kabata Swamp, Ndale volcanic field, Uganda. Journal of Ecology **87**:303-315.

Tempelton, A. 1998. Human Races: A genetic and evolutionary perspective. American Anthropologist **100**:632-650.

Ting, N. 2008. Mitochondrial relationships and divergence dates of the African colobines: evidence of Miocene origins for the living colobus monkeys. Journal of Human Evolution **55**:312-325.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins.1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research **25**:4876-4882.

Thao, M.L., N.A. Moran, P. Abbot, E.B. Brennan, and D.H. Burckhardt. 2000. Cospeciation of psyllids and their primary prokaryotic endosymbionts. Applied and Environmental Microbiology **66**:2898-2905.

Thao, M.L., and P. Baumann. 2004. Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. Applied and Environmental Microbiology **70**:3401-3406.

Hypsa, V, and J. Kirizek. 2007. Molecular evidence for the polyphyletic origin of the primary symbionts of sucking Lice (Pthiraptera: Anoplura). Microbial Ecology **54**:242-251.

van Ham, R., J. Kamerbeek, C. Palacios, C. Rausell, and F. Abascal. 2003. Reductive genome evolution in *Buchnera aphidicola*. Proceedings of the National Academy of Sciences, USA **100**:581-586.

van Oosterhout, C., W.F. Hutchinson, D.P. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes **4**:535-538.

Wallace, B. 1987. 50 Years Of Genetic Load. Jounal of Heredety **78**:134-142.

Watterson, G.A. 1975. On the number of segregating sites in genetical models without recombination. Theoretical Population Biology **7**:256-276.

Weir, B.S. and C.C. Cockerham. 1984. Estimating F-statistics for the analysis of poulation structure. Evolution **38**:1358-1370.

Wernegreen, J. J., and N. A. Moran. 1999. Evidence for genetic drift in endosymbionts (*Buchnera*): analyses of protein-coding genes. Molecular Biology and Evolution **16**:83-97.

Whittaker, J.C., R.M. Harbord, N. Boxall, I. Mackay, G. Dawson, and R.M. Sibly. 2003. Likelihood-based estimation of microsatellite mutation rates. Genetics **164**:781-787.

Wing, L.D., I.O. Buss. 1970. Elephants and forests. Wildlife monographs 19.

Woolfit, M., and L. Bromham. 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Molecular Biology and Evolution **20**:1545-1555.

Wu, D., S.V. Daugherty, S.E. Van Aken, G.H. Pai, and K.L. Watkins. 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biology **4**:e188.

Yeh, F.C., and T.J.B. Boyle. 1997. Population genetic analysis of co-dominat markers and quantitative traits. Belgian Journal of Botany **129**:157.

Yong, Z., P. E. Fournier, E. Rydkina, and D. Raoult. 2003. The grographical segregation of human lice preceded that of *Pediculus humanus capitis* and *Pediculus humanus humanus*. Canadian Research on Biology **326**:565-574.

Yoshizawa, K. and K.P. Johnson. 2003. Phylogenetic position of Pthiraptera (Insecta: Paraneoptera) and elevated rate of evolution in mitochondrial 12S and 16S rDNA. Molecular Phylogenetics and Evolution **29**:102-114.

BIOGRAPHICAL SKETCH

Julia M. Allen was born in Provo, Utah in 1978. She grew up in Sandy, Utah and attended Skyline High School where she graduated in 1997. She then attended Southern Utah University for a few years before moving to the University of Utah. Here, she worked with Dr. Jim Ehlringer and Dr. Joan Coltrain studying the stale isotope chemistry of the large mammals of the LaBrea Tar Pits. She graduated from the University of Utah in 2003 with a B.Sc. in Biology and a chemistry minor. She worked for a year with Dr. Denise Dearing studying hantavirus transmission of desert rodents in central Utah. She then started graduate school at the University of Florida in the Zoology Department in 2004. Here she studied evoluionary biology of sucking lice and their obligate partners.