Stony Brook University

Ringtailed lemur social networks and their role in pathogen transmission

A Dissertation Presented

by

Gena Sbeglia

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

May 2017

Stony Brook University
The Graduate School

Gena Sbeglia

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Daniel Dykhuizen – Dissertation Advisor
Ecology and Evolution

Liliana Dávalos - Chairperson of Defense
Ecology and Evolution

Patricia Wright
Anthropology

Jesse Hollister
Ecology and Evolution

Susan Perkins – Outside member
Molecular Systematics
Richard Gilder Graduate School

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Ringtailed lemur social networks and their role in pathogen transmission

by

Gena Sbeglia

Doctor of Philosophy

in

Ecology and Evolution

Stony Brook University

2017

Many of the pathogens that cause disease are transmitted through physical contact, which makes patterns of social behavior potential routes of transmission. Questions about socially facilitated transmission are best addressed by combining data on the observed contacts of the host and the haplotype-level genetic differentiation of the pathogen because individuals must harbor the same or related haplotypes of a particular pathogen for transmission to be deduced. However, few studies simultaneously collect data on both the behavior of the host and the genetics of the pathogen and those that do are limited by their use of culture-based methods. Culture-based methods involve growing a sample on a nutrient plate and identifying the genetic variation in each bacterial isolate across multiple loci. These methods are time- and labor-intensive and unrealistic to accurately differentiate the multiple bacterial haplotypes of the same species that can reside within an infected host. One-locus methods do not require a culturing step and allow sequencing of every amplified haplotype in a sample. Such an approach is used in the microbiome literature via the 16S gene, which can differentiate species or genera of bacteria, but

is not appropriate for the haplotype-level differentiation that is necessary to identify incidences of transmission. In this dissertation, I determined patterns of association in wild ringtailed lemurs and developed a one-locus, culture-independent approach for the differentiation of *E. coli* haplotypes that could be used to test hypothesized routes of pathogen movement. Although not usually pathogenic, *E. coli* is valuable as a model "pathogenic" organism for social transmission studies because its ubiquity and within-host diversity in mammals allows for the inference of fine-scale patterns of transmission among *all* individuals, instead of just those infected by an occasional pathogen. Furthermore, its well-known genetics makes it suitable for the development of a novel approach to haplotype differentiation because it is possible to assess the haplotype diversity that is and is not captured by this new method.

I collected over 1000 hours of detailed social behavior data on 29 individually identifiable ringtailed lemurs living in three sympatric social groups in Beza Mahafaly Specieal Reserve in south western Madagascar from March–September 2015. Active and passive affiliation had different temporal patterns with individuals decreasing the overall time in active affiliation and increasing the time in passive affiliation from the pre- to the post-mating season. Further, there was substantial variation across individuals in their network centrality for both affiliative and agonistic interactions, but sex and dominance did not explain this pattern for active or passive affiliation, which are the behavior modes most likely to cause pathogen transmission. I also found that social groups differed in their connectedness and that living in degraded habitat may coincide with properties of the social network that could cause heightened pathogen transmission. Animals living in degraded habitat are often expected to have a higher rate of infection because of the increased exposure to pathogens from humans and livestock, but

the results presented here suggest a possible amplification of these effects by an increase in network connectedness.

To test these hypothesized patterns of transmission, I developed the first one-locus, culture-independent approach for the differentiation of *E. coli* haplotypes. I identified and tested primers at the FumC locus that target a single, highly variable 294-bp region and found it could differentiate 91-172% of the haplotype diversity as compared to standard multi-locus methods. When applying this method to wild-collected feces sampled bi-weekly throughout the observation period, the results demonstrated its potential to capture much more within-host *E. coli* haplotype diversity than previously identified in any study to date. When coupled with detailed data on social contact patterns, this method can revolutionize our ability to determine fine-scale transmission dynamics and assess *E. coli* population genetics within a wild host.

Dedication Page

I dedicate this dissertation to my Grandma, Phyllis Mina, who passed away on April 28th, 2017, two weeks to the day after I defended my dissertation and 11 days before I gave birth to her first great grandchild. She offered unconditional love, support, stability and understanding when it mattered the most. She was a model for how to live your life with grace, kindness, and a balanced checkbook. I am grateful that I got to know her and call her my Grandma.

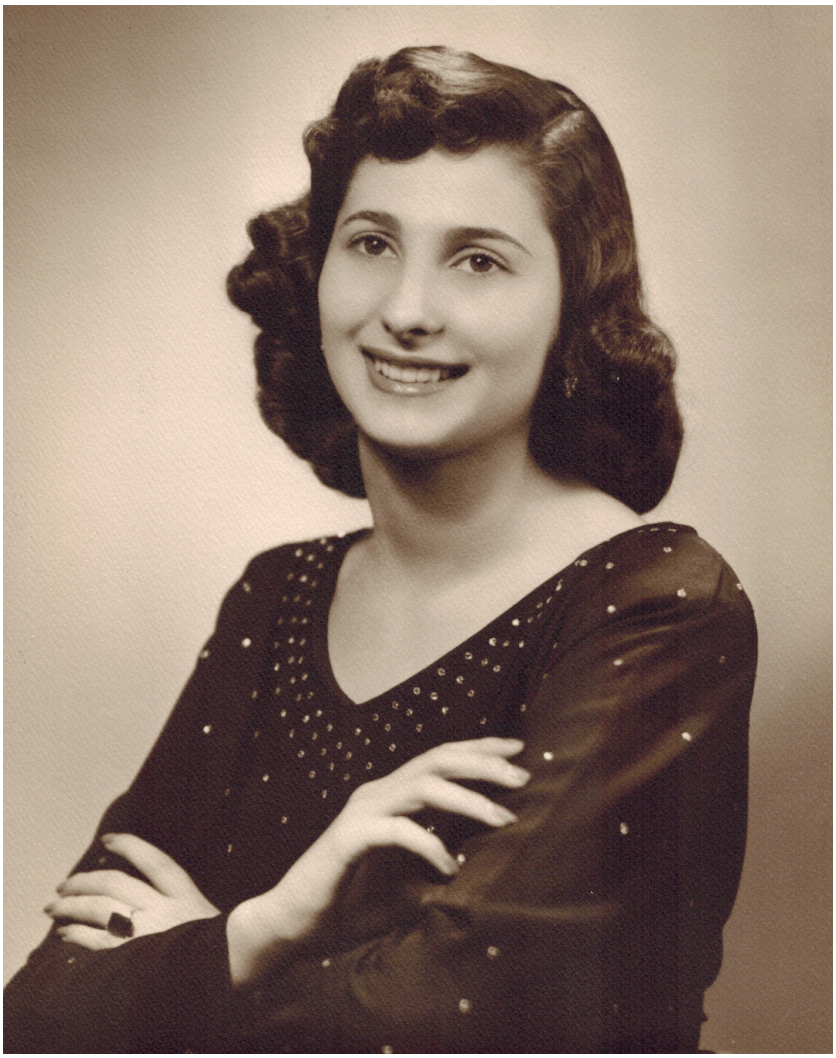


Table of Contents

List of Figures

List of Tables

List of Abbreviations

LI – Low intensity agonism

HI – High intensity agonism

ETs – Electrophoretic types

MLST – Multi-locus sequence typing

CH-typing – FumC-FimH typing

CH-type – FumC-FimH type

Acknowledgments

I thank the staff at Beza Mahafaly Special Reserve for making the field work for this dissertation possible and for making Beza feel like home for a year. I especially want to thank Jean Sibien Mahereza for helping me with everything and anything I needed to live and work at Beza, Ralalasoa Victorine for her incredible cooking and laughter, and Efitroarany, my incredibly talented guide, for his dedication, hard work, and constant willingness to learn new skills. I also thank the team at Centre ValBio, especially Tharcisse Ukizintambara, Desiree Randriarisata and Prisca Oliva Andrianmbintsoa, for making me feel welcome when I visited and for providing access to the lab that stored my painstakingly collected fecal samples. I thank Institut Halieutique et des Sciences Marines (IHSM) and The Fondation Merieux for help storing my samples and Madagascar Institute for the Conservation of Tropical Environments (MICET) for assistance with permits and supplies, especially, Benjamin Andriamihaja and Tiana Razafindrasata.

I thank my wonderful committee. Daniel Dykhuizen, my advisor, listened to my research ideas on the first day we met and told me to run with them even though we both knew they were risky. The knowledge and brilliance he brought to my project, some of which took me months to fully understand, will always amaze me. I am grateful for his willingness to teach me and proud to be his final student before his well-earned retirement. I also thank Patricia Wright, whose story and study abroad program introduced me to Madagascar 13 years ago and inspired me to choose research as a life-long career. Her fearless approach to research, fieldwork, and life gave me the confidence, both professionally and personally, to just go for it. With Dan and Pat, I was privileged to have one of the foremost experts on both *E. coli* population genetics and lemur behavior and conservation on my team. I was also privileged to have Liliana Dávalos as a

committee member and mentor. Both being mentored by her and observing her mentor others was a lesson in best practices for mentorship and student engagement. She has demonstrated over and over again that being a stellar scientist and a supportive mentor are complements to each other, although I doubt that most could do both quite as well as she manages on a daily basis. She is one of the few people that deeply understood the genetic work I was laboring to carry out and offered invaluable advice every step of the way. I also thank Jesse Hollister for his contributions to the bioinformatics planning for this dissertation and Susan Perkins for her incredible knowledge and fresh perspective that has saved this project on several occasions.

I also thank my labmates from the Dyhuizen lab, Fabrizio Spagnolo and Omar Warsi, for being my first teachers of all things lab work. Their patience and willingness to teach me made all the difference in the early stages of this project and will mold how I mentor novice students in the future. I also thank my labmates from the Wright lab and the Dávalos lab, especially Alicia Lamb, Kate Thompson, Katherine Kling, Mariah Donohue, Rachel Jacobs, Caitlin Karanewsky, Fanny Cornejo, Elise Lauterbur, Laurel Yohe, and Katie Martin for their support and feedback on many aspects of this project and for their friendship. All of them deserve a paragraph of thanks each for all the little and big ways they supported me and my work these past years. I am honored to know each of these inspiring women and scientists.

I thank Zuleyma Tang-Martinez and Robert Sussman for their mentorship and encouragement throughout my master's program, which is the primary reason I went on to do a Ph.D. Zuleyma was an incredible advisor who made me a better scientist and taught me to the realities of academia. The mistakes I made while being her student were some of the best lessons I ever learned in science, for which I credit her supportive and honest approach to mentorship. Bob introduced me to ringtailed lemurs as a study system and made it possible to conduct

research at Duke Lemur Center and Beza Mahafaly Special Reserve. His encouragement made me feel like I could actually be a scientist and his excitement made me feel like I could contribute something amazing one day. I also thank Brenda Bradley for sharing her DNA extraction protocol and Adam Leache for sharing the recipe for the feces storage solution used in this study.

I thank Rima Madan and Catherine Sbeglia for their help with laborious and repetitive data entry and lab work, Dean Bobo for help with the bioinformatics, Evgeni Sokurnko and Veronika Tchesnokova for their assistance with the Sanger sequencing and interpretation of data, and Sasha Levey, Ben Weinstein, John Borelli, Kash Bandaralage, Joshua Rest, Bob Thacker and Dianna Padella for their help and advice on many essential components of the project. I also must thank Tinah Martin for all the help she provided during my fieldwork and most of all, for her friendship.

Finally, I must thank my family and friends for their never-ending support for the non-traditional decisions I have (and will probably continue) to make. They trusted that I could reach the goals I set, even though there were times when I wasn't so sure myself. No one demonstrated their faith in me more than my parents, Richard and Therese Sbeglia, my in-laws, Robin and Hal Gold, and my husband, Noah Gold. I give special thanks to Noah for being the most amazing partner and best friend I could ever imagined. Building our adult lives together while both of us worked way too many hours per week was anything but easy. In that time, he embraced my professional goals and did absolutely everything I asked, and didn't ask, to help me reach them. In addition to my partner in life, he was also a constant partner in the completion of this dissertation and would impress a seasoned scientist with his ability to discuss principles of

ecology and evolution. I am excited to work through our future challenges together, especially the one that will be arriving any day now.

I have received permission from Springer publishers (License #: 4100929955576) to use the copyrighted figure from Brockman et al. (2008). The procedures used in this study were approved by the IACUC at Stony Brook University (Protocol # 319914) and complied with all legal requirements of the United States. All fieldwork was conducted with appropriate permissions and in accordance with the laws of Madagascar.

Introduction

Many of the pathogens that cause disease are transmitted through physical contact, which makes patterns of social behavior potential routes of transmission (Rushmore et al. 2013, 2017 - for examples, see Hoogland 1979, Nunn et al. 2000, Whiteman and Parker 2004). However, there are relatively few studies in any animal group that simultaneously collect host infection data (van Hoek et al. 2013, Rushmore et al. 2013) and only four that simultaneously genotype the pathogen (i.e. Blyton et al. 2014, Springer et al. 2016, VanderWaal et al. 2014 a, b). Without data on both host social behavior and pathogen genetics, attempts to answer questions about socially facilitated pathogen transmission have a high possibility of mistakenly scoring independently acquired infections of the same pathogen as social transmission, which limits the ability to accurately test how the type and frequency of interactions contribute to fine-scale transmission dynamics within and between groups (Craft 2015).

The merging of these two types of data has the potential to refine our understanding of the link between social behavior and infection, challenge long-held ideas about the high cost of socially transmissible diseases to social living, and revolutionize our ability to construct predictive models for the future spread of pathogens. However, a major limitation in the assessment of pathogen genetics is the use of culture-based methods to differentiate pathogen haplotypes (Dias et al. 2010). This challenge is particularly relevant in mixed-haplotype samples, which is a common condition of many infections (e.g. Bachmann et al. 2015, Eyre et al. 2013, Taylor et al. 1995). Differentiating many bacterial haplotypes using culture-based methods is time- and labor- intensive and likely to grossly underestimate the true haplotype diversity in a

sample (Bachmann et al. 2015, Dias et al. 2010). Culture-based methods involve growing a sample on a nutrient plate and identifying the genetic variation in each bacterial isolate across multiple loci. Using multiple loci allows for the incorporation of more genetic variation in the differentiation of haplotypes, but it also makes it necessary to grow each isolate in culture so that the relationship among loci in the same cell can be established. Therefore, a one-locus, culture-independent approach to haplotype differentiation, similar to that used for genus-level differentiation in the microbiome literature, would greatly improve the ability distinguish among many bacterial haplotypes. When coupled with data on social contact patterns, routes of bacterial transmission can be identified and demonstrate the types of interactions that are sufficient for transmission to occur (Craft 2015). Unfortunately, there are no one-locus, culture-independent methods currently available to differentiate haplotypes within a species. The 16S gene can be used to differentiate species or genera of bacteria, but it is not appropriate for haplotype-level differentiation, which is the necessary scale to identify incidences of transmission.

In the first two chapters of this dissertation, I examined patterns of association and the traits that help generate them in three groups of wild ringtailed lemurs (*Lemur catta*) using Social Network Analysis. These data are useful to generate hypothetical routs of pathogen movement. In the third chapter, I developed and tested a culture-independent, one-locus approach to haplotype differentiation in *E. coli* that bypasses the culturing step and is appropriate for high throughput sequencing. This approach makes it possible test hypothetical routs of pathogen transmission using the actual movement patterns of a real bacterial organism. I used *E. coli* as a model "pathogen" because its near-ubiquity in mammals and high haplotype diversity within a single host (Hartl and Dykhuizen 1984) allows the assessment each individual's *E. coli* community and contribution to bacterial transmission in a variety of species at all time points.

Furthermore, to develop a one-locus method, I required a bacterium whose diversity is well-known so that I could evaluate how well this method captured that diversity. In the fourth chapter, I report the application of the novel one-locus, culture-independent method developed in Chapter 3 to fecal *E. coli* of wild ringtailed lemurs in a manner appropriate for the analysis of socially facilitated *E. coli* transmission. I also provide preliminary results showing the potential of this method to revolutionize our ability to determine fine-scale transmission dynamics and assess *E. coli* population genetics within a wild host. The one-locus, culture-independent method described and applied in this dissertation is the first to allow the large-scale haplotype-level differentiation of bacteria, setting the stage for one of the most direct investigations into socially facilitated transmission to date.

The magnitude and stability of dominance during the ringtailed lemur mating season using the Elo-rating method

Chapter 1

Abstract: Ringtailed lemurs are seasonal breeders with all females entering estrus within two weeks of one another for less than 24 hours each year. The rapid and frequent shifts in patterns of social behavior and group membership during this period make it difficult to study dominance characteristics with traditional matrix-based methods, which often cannot capture these shifting dynamics. I used the Elo-rating method to investigate the magnitude and rank order stability of the dominance relationships and hierarchies of three groups of ringtailed lemurs at Beza Mahafaly Special Reserve, before, during and after the annual mating season. Elo-rating can estimate a rank order at any point in time that incorporates the sequence of dominance interactions and is independent of demographic changes. Linearity of the dominance hierarchy and stability in the magnitude of dominance scores declined during the mating season. However, these occurrences did not coincide with marked instability of the order of the dominance rank for most groups. The exception was the mating season for one group, which was characterized by reduced rank order stability. The mating season was also the period for this group in which the starkest changes in female rank order and the arrival of new males occurred. Therefore, while instability in the magnitude of dominance could contribute to rank order instability, this is not always the case, suggesting that dominance rank order may be more robust in this species than expected. This study is the first to use the Elo-rating method in a lemur population and may enable analyses of ephemeral dominance patterns, particularly those that change rapidly, in a more nuanced and detailed manner than traditional methods. Ultimately, this method may lead to different characterizations of social behavior than are currently available in the literature.

Introduction:

Dominance relationships result from the interaction patterns of pairs of individuals. These relationships can be characterized as egalitarian (i.e. a lack of dominance relationships), despotic (i.e. one individual dominant to all others), or somewhere in between (van Schaik 1989; Sterck et al. 1997). Dominance hierarchies emerge from the collective dominance relationships of all dyads in a social group and allow individuals to be ranked by their ability to "win" dominance interactions (Bernstein 1981). The determination of dominance ranks is a frequent goal of behavioral work in social animals because an individual's rank is often associated with its fitness. High ranking individuals can receive priority access to food (Isbell et al. 1999, Whitten 1984), mates (Altmann 1962, Koyama 1988), and affiliative interactions (Snyder-Mackler et al. 2016), which can reduce external parasite load, reduce stress, and reinforce collaborations. As a result, high ranking individuals are expected to have higher reproductive success (Majolo et al.

2012, Ostner et al. 2008, Pusey et al. 1997, Taylor 1986). There are many methods to assess the dominance hierarchy of a social group but they come with important challenges. Most of these methods are based on square interaction matrices in which the frequency of "wins" and "losses" for each dyad are tabulated (Albers and de Vries 2000, Neumann et al. 2011). One of the challenges associated with these matrix-based methods includes accounting for empty cells in the matrix, which occur when dyads were not observed in a dominance interaction. Additionally, dynamic social environments can cause rapid or frequent changes in dominance relationships and group membership (Neumann et al. 2011), which become problematic for accurate assessments of dominance patterns. Matrix-based methods to assess dominance cannot readily incorporate the temporal sequence of dominance interactions or the presence or absence of group members at particular times. Therefore, it is difficult to use matrix-based dominance methods to reflect the shifting nature of dominance relationships over time and during periods of rapid social change (Albers and de Vries 2000, Neumann et al. 2011). Separating interaction data into multiple matrices across periods to represent different dynamics can sometimes mitigate this difficulty, but this approach could cause more empty cells in each matrix and reduce the reliability of the resulting conclusions (Neumann et al. 2011).

To deal with the shortcomings of matrix-based methods, a method called Elo-rating has been proposed to assess dominance hierarchies in social animals (Albers and de Vries 2000, Neumann et al. 2011). Elo-rating was developed in the 1960s to assess and rank chess players (Elo 1961). This method does not rely on matrices but rather on the sequence with which dominance interactions occur. Each individual's Elo-rating can be used to estimate a rank order at any point in time (Neumann et al. 2011). Unlike matrix-based methods, Elo-rating can be applied to groups of any size, generate scores independent of demographic changes, and

incorporate the sequence of dominance interactions into the determination of dominance scores, which allows tracking of changing dominance patterns through time (Albers and de Vries 2000, Neumann et al. 2011).

Elo-rating has so far been used sparingly in the social behavior literature and has never been applied to analyses of lemur dominance structures. In this paper, I applied this methodology to dominance behavior in three groups of wild ringtailed lemurs. Its application to dominance patterns in this species is particularly valuable because ringailed lemurs can exhibit flexible dominance hierarchies that have been observed to be linear, transitive, and stable for long periods of time (Norscia and Palagi 2015), but also non-linear and unstable (Sauther et al. 1999). This variation in dominance patterns is particularly evident during the mating season, which occurs over a few consecutive weeks each year with each female being receptive for under 24 hours (Jolly 1966). Though short in duration, these few weeks have been characterized as socially chaotic because the prevailing dominance hierarchy has been observed to temporarily break down (Jolly 1966, Budnitz and Dainis 1975). In addition, adult males temporarily or permanently transfer groups every few years and this migration typically occurs immediately before, during, or immediately after mating (Sussman 1992, Gould 1996). Therefore, shifts in patterns of social behavior and group membership are frequent and rapid during this several-week period and are difficult to study with matrix-based methods. Unlike matrix-based methods, Elo-rating allows the incorporation of unstable grouping and interaction patterns during these narrowly defined periods without compromising the integrity of the hierarchy estimation. In this paper, I investigate the structure and stability of dominance hierarchies in three groups of wild ringtailed lemurs during a six-month time period that spans the highly chaotic mating season. In

doing so, this paper represents the first analysis of ringtailed lemur dominance relationships using the Elo-rating method.

Methods:

Subjects and Data Collection

I collected detailed social behavior data on 29 individually identifiable lemurs in three sympatric social groups (Group 1, Group 2, and Group 3) in or near Parcel 1 of Beza Mahafaly Special Reserve in southwestern Madagascar from March – September 2015. These months reflect a transition between the wet and the dry season and include the period of female mating (May) and gestation (late May to September or October). Parcel 1 consists of 80 hectares of mixed vegetation forest (figure 1). The western portion of Parcel 1 consists of xerophytic/scrub forest and the eastern portion consists of lush gallery forest adjacent to the Sakamena River (Sussman and Rakotozafy 1994). The territory of Group 1 was located in the western scrub forest and the territory of Group 3 was located within the lush gallery forest. The territory of Group 2 was south and southwest of parcel

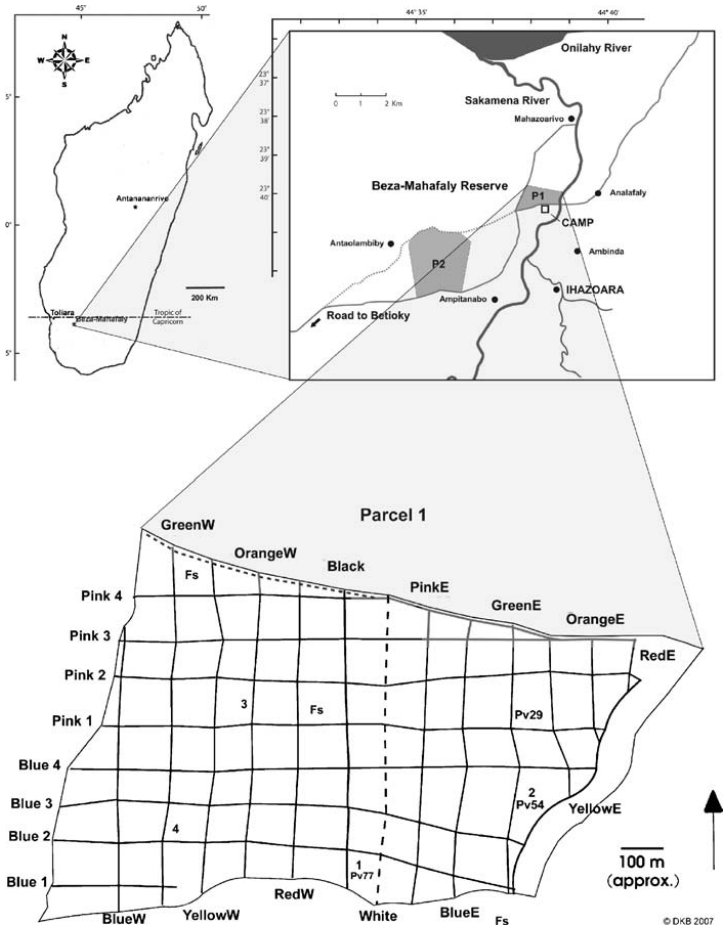


Figure 1. Map of Beza Mahafaly Special Reserve. From Brockman et al. 2008.

1 and almost entirely outside of the reserve. The area outside of parcel 1 is heavily disturbed dry forest with less canopy cover and sparse vegetation due to human logging and livestock (Gemmill and Gould 2008) and a lower availability of preferred food (Gould 1996; Sauther 1992, 1993).

Group 1 consisted of 13 individuals (4 adult females, 1 subadult female, 1 juvenile female, 2 adult males, 1 subadult male, 1 juvenile male, 3 infant males). At its maximum size, Group 2 consisted of 12 individuals (4 adult females, 5 adult males, 1 sub-adult male, 2 infant males). At its maximum size, group 3 consisted of 18 individuals (5 adult females, 1 sub-adult female, 3 juvenile females, 4 adult males, 1 subadult male, 2 juvenile males and 2 infant males). Only the adult and sub-adult individuals were treated as focal animals resulting in 15 female and 14 male focal animals. The details of all interactions of the focal animal were documented and included in relevant analyses.

I conducted full day follows (~7:00-17:00) of each group twice per week. During these observation periods, I conducted continuous all occurrence focal sampling (Altmann 1974) of social behaviors (active affiliation, passive affiliation, submission, low intensity (LI) agonism and high intensity (HI) agonism; see table 1) of each adult and subadult individual for 45 – 60 min, aiming to observe each focal animal in the same group in the same day. The sequence of observations was designed to generate similar observation times in the morning and afternoon for each focal animal. On average, each focal animal was observed for approximately 1.5-2 hours every week for a total of 1,038 hours of observation (Group 1: 344 hours, Group 2: 341 hours; Group 3: 353 hours). For the purposes of this paper, all submissive, LI, and HI agonistic interactions were documented from their time of initiation to the time of completion. Behaviors that ended almost immediately after they begin, were considered to last for a total duration of

one second. These duration times were summed for each dyad for each observation day and used to generate dominance scores, as described below.

High-intensity agonism		**Low-intensity agonism**	
Stink fight[a]	Grab	Plosive bark[b]	Chutter[b]
Lunge	Push	Mouth to face threat	Yip/spat call[b]
Chase	Bite	Squeal[b]	Displace/supplant
Hit	Contact fight	Cackle/deep spat[b]	Chase/lunge threat

Table 1. List of high and low intensity agonistic interactions. [a]Between males only. [b]Terms taken from Macedonia (1993).

Establishing the Mating Period

Ringtailed lemurs live in stable multi-male, multi-female groups and mate polygamously during the singular annual mating season. They are seasonal breeders with all females entering estrus for less than 24 hours each year and within one to three weeks of one another (Sauther et al. 1999). At Beza Mahafaly Special Reserve, the mating period occurs in May and the peak of births occur in October (Sauther et al. 1999). The period of mating for these three groups was determined to begin at the first sign of mating-related behaviors, such as the temporary absence of one or more adult males for long periods of time (observed and assumed to be prospecting other groups in possible preparation for a group transition), or the appearance of sexual swellings and copulatory plugs, the latter of which occur 1-2 days after mating and last one day (Sauther 1991). Additionally, all females were surveyed biweekly for the presence of an infant from early September to mid-November. Ringtailed lemurs at Beza Mahafaly Special Reserve have gestation times between 136 and 145 days (Sauther 1991). I estimated the date of conception in this study as 136 days prior to the first day an infant was observed. The last day of the mating period was determined based on the secession of observed mating-related behaviors or the last estimated date of conception. Because some males were away from their group for extended

periods of time, a guide was tasked with following the males starting early in the morning so that their regularly scheduled observations could continue and so that we could confirm prospecting behavior.

Linearity, Magnitude, and Rank Order Stability of Dominance Hierarchies

I evaluated the dominance hierarchy for each group using all dyadic and decided (clear "winner "and "loser") submissive, LI and HI agonistic interactions. The direction of behavior for all interactions was organized from the dominant to the subordinate individual. I categorized each interaction into one of three time periods for each group; before mating, during mating, and after mating. I calculated the linearity of the dominance hierarchy using statistical methodology developed by de Vries (1995) and available in the R package compete v 0.1 (Curley 2016). This methodology calculates linearity by placing all interactions in a square matrix and calculating how many triads (i.e. groups of three individuals) in a given matrix were circular (i.e. A>B>C>A) as opposed to linear (i.e. A>B>C, A>C) as compared to the maximum number of possible triads (Appleby 1983). The output of this analysis is Landau's index (*h*), which indicates the degree of linearity of a matrix on a scale of 0 to 1 (de Vries 1995). A linearity value of 0 indicates that all individuals are dominant to the same number of group members (i.e. non-linear) and 1 indicates strongly linear interactions. I consider a linearity between 0.80 - 1.0 to be strongly linear, 0.50-0.80 to be moderately linear and 0.50 – 0 to be weakly or non-linear (Bergstrom and Fedigan 2010).

I calculated the rank order of dominance hierarchies using the Elo-rating method via the R package EloRating v. 0.43 (Elo 1978, Neumann and Kulik 2014). Elo-rating works by assigning each individual in a social group a dominance score of 1000 at time zero. Scores are

then updated after each sequential dominance interaction (Group 1: 989 interactions, Group 2: 925 interactions; Group 3: 1220 interactions) based on the outcome of the interaction (i.e. win, lose, or tie) and the pre-interaction probability of that outcome. Interactions with a high probability of occurring alter each individual's score less than interactions with a low probability of occurring (Albers and de Vries 2001). For example, a dominant individual winning over a subordinate individual is an expected outcome and changes dominance scores by only a small amount. Likewise, if the subordinate animal were to win, both scores would change in the opposite direction by a much larger fraction. The actual magnitude of the change in score for each dominance interaction is a function of a pre-set value, k (100 in this study), weighted by the expected probability of the outcome of a particular interaction (Albers and de Vries 2001). I also included data on which individuals were absent each day, which allowed the program to omit them from the final plot for those days. The ability to specify presence and absence of individuals is valuable for constructing hierarchies in which individuals leave or join a social group on a temporary or permanent basis during the observation period.

I evaluated the number of rank changes within and across dyads by determining the difference in score for each observation day for each dyad and adding up the number of times that score changed its sign (i.e. "+" to "-" or "-" to "+"). Days with missing Elo scores were assigned the Elo score from the last available observation day. I did not consider the Elo scores from the first two observation days because there often are not enough interactions at this point to provide reliable scores. Rank changes were considered permanent if rank positions that were observed for at least 10 observation days, switched for at least another 10 observation days and remained until the end of the observation period. Because the two new males that entered the group during the mating season, entered with a pre-determined Elo score of 1000 (as did all focal

individuals during their first observation period), I replaced the first two Elo scores with scores observed from observation day 3 after interactions had time to accumulate.

I measured the magnitude of dominance changes among and within individuals in each time period by calculating the slope and standard deviation of each individual's set of Elo scores. A negative slope for the Elo scores of a particular lemur indicated a decline in dominance and a positive slope indicated an increase in dominance. A large slope indicated large fluctuations in the magnitude of within-individual dominance. Finally, a large standard deviation in the slopes of Elo scores indicated a large variation in the trajectory of dominance across individuals.

I also calculated the degree of hierarchy stability (*S*), which is a measure of the ratio of rank changes per individual over a given time period weighted by the standardized Elo-rating of the highest-ranked individual involved in a rank change. *S* ranges from 0 and 1, with 1 indicating a stable hierarchy with no rank changes (Neumann et al. 2011).

Results:

Mating Behavior and the Mating Period

Mating-related behaviors were directly observed in all groups between late April and early June. In Group 1, the two subordinate males (Dy and Vn; figure 3) displayed confirmed prospecting behavior, usually as a pair, beginning in late April, which continued through late May, at which point both males remained in their original group. No matings were directly observed in this group but all four surviving sexually mature females (Rz, Mm, Cc, Sy) successfully birthed infants (all first observed on October 3rd; figure 2). The date of conception

was estimated to be on or around May 21st, two days after the secession of male prospecting behaviors.

In Group 2, the two dominant males (Ev and Cy; figure 3) displayed confirmed prospecting behavior in early May (again, as a pair) but this behavior was not regularly observed throughout the remainder of the mating season and both males remained in their original group. Two non-native males (Fr and Ha) that eventually joined this group were consistently observed nearby beginning in mid-May. Although Fr was observed mating with a resident female during the mating season and both males became accepted members of the group in the post-mating season, they fell to the lowest rank after mating (figure 3). Three matings were directly observed between one female (Bn) and three males (Ev, Fr, Cy, in that order; figure 3), but all sexually mature females successfully birthed infants over a two-week time period (September 29th – October 16th; figure 2), roughly in the order of the pre-mating dominance rank (Sa, Ma, Bn, Dt). This timing puts the date of conception for all females in this group between May 17th and June 3rd, an 18-day period.

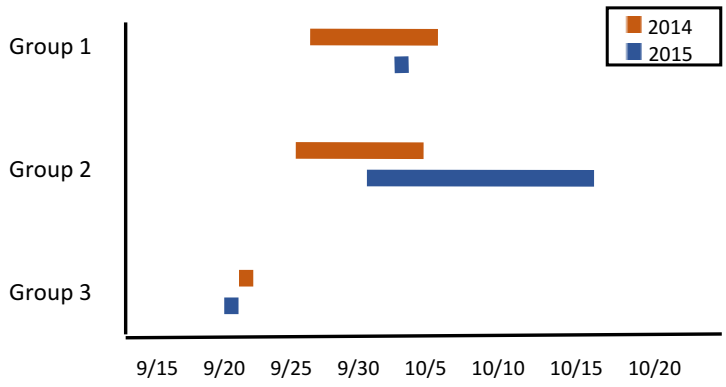


Figure 2: Timing of births in 2014 and 2015

In Group 3, one mating was directly observed (between Mg and Nt) but most surviving sexually mature females (Yl, Rd, Gy, Vi) successfully birthed infants (all infants first observed on September 20th; figure 2), putting the date of conception for all females on or around May 8th. Although female Ht did not birth an infant, she was observed with a sexual swelling around the time of conception of the other females in her group. Ht is an older, subordinate female who had

one of the only two surviving infants in the prior year. It is not known if she mated, got pregnant, or lost an infant before or soon after birth. One of the most subordinate males (Pa; figure 3) was last observed in the group on May 24th and did not return. Given the timing of his absence, it is presumed that he attempted to disperse to a different group.

Observed or inferred mating behaviors mostly overlapped in the three groups and spanned a similar number of days (Group 1: 28 days, April 24th - May 21st; Group 2: 31 days May 4th - June 3rd; Group 3: 24 days, May 1st - May 24th; Figure 2). However, Group 3 was estimated to complete conception approximately two weeks before the females in Groups 1 and 2. As a result, the birthing dates of Group 3 were earlier than is typical for this species at Beza Mahafaly Reserve (Sauther et al. 1999; figure 26). Furthermore, the length of time during which indicators of female receptivity or recent matings could be observed (i.e. sexual swellings, mating events, copulatory plugs) differed across groups (Group 1: not directly observed; Group 2: 24 days; Group 3: nine days). This length of time seems to roughly coincide with the timing and estimated range of conception dates within each group. Specifically, females in Groups 1 and 3 had no or few days of observable estrus and a small range of dates during which conception was estimated to occur. However, females in Group 2 had a much longer period of observed estrus and their conception dates were estimated to span an 18-day period. Interestingly, the timing of birthing dates for each group was similar to what was observed in the prior year, with Group 3 completing the birthing of infants by mid-September, and Groups 1 and 2 beginning the birthing of infants at the end of September and continuing through the beginning of October (figure 2), roughly in the order of the 2015 dominance hierarchy (Group 1: Rz, Bl, Mm; Group 2: Sa, Ma, Bn, Dt).

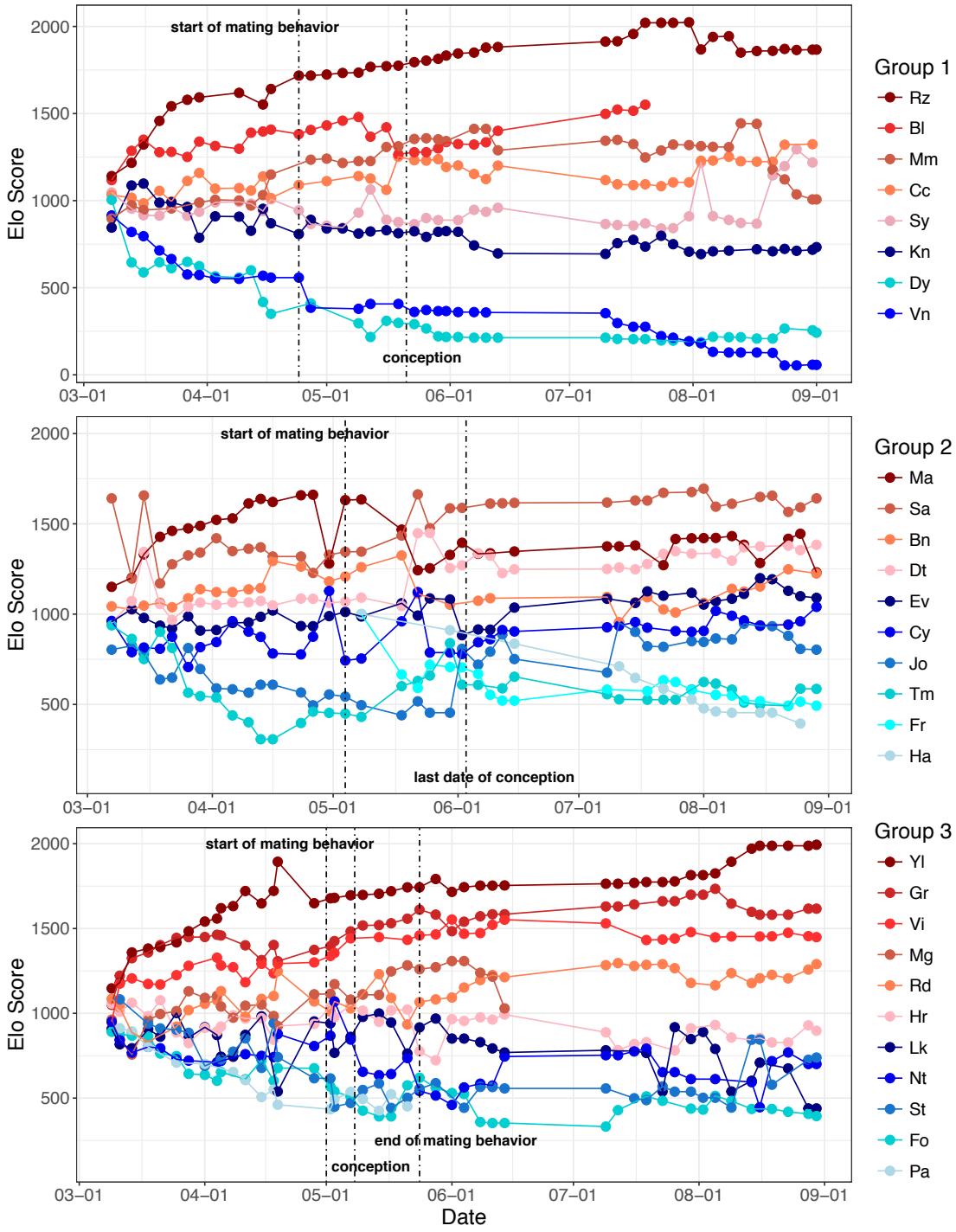


Figure 3. Elo scores through time for Groups 1 (top), 2 (middle), and 3 (bottom). Females are shown in shades of red and males in shades of blue. The dashed vertical lines indicate the first day that mating behaviors were observed or inferred, the estimated or last estimated date of conception, and the last day that mating behaviors were observed or inferred. In the case of Groups 1 and 2, the estimated date of conception was the final indicator of mating.

Dominance Linearity, Magnitude and Stability

The hierarchies of all groups were significantly linear before the mating season with Groups 1 and 2 being strongly linear (Group 1: $h^2_{t1} = 1$, n = 8, $p < 0.001$; Group 2: $h^2_{t1} = 0.83$, n = 8, $p < 0.01$) and Group 3 being moderately linear (Group 3: $h^2_{t1} = 0.75$, n = 11, $p < 0.001$). The hierarchies of all groups declined in linearity during mating although two groups remained significantly linear (Group 1: $h^2_{t2} = 0.93$, n = 8, $p < 0.001$; Group 3: $h^2_{t2} = 0.71$, n = 11, $p < 0.001$) and one became non-linear (Group 2: $h^2_{t2} = 0.43$, n = 10, $p = 0.18$). All hierarchies were again linear in the months following the mating season (Group 1: $h^2_{t3} = 1$, n = 8, $p < 0.001$; Group 2: $h^2_{t3} = 0.78$, n = 10, $p < 0.001$; Group 3: $h^2_{t3} = 0.91$, n = 11, $p < 0.001$) with Group 1 and Group 2 regaining a linearity score that was similar to their pre-mating score and Group 3 displaying its highest linearity out of all three time periods.

Before the mating season, the slope of Elo scores was near zero or negative for all males ($\overline{X}slope$ = -4.3, SD = 3.82) and near zero or positive for all females ($\overline{X}slope$ = 3.06, SD_{female} = 3.39). Therefore, females tended to increase in dominance and males tended to decline in dominance during this period. During the mating season, however, the patterns of the Elo scores changed dramatically with individuals of both sexes varying widely in the magnitude and direction of their slopes, as indicated by the large standard deviations of the Elo scores (SD_{male} = 6.74, SD_{female} = 7.09). This increase in the standard deviation of the slope of Elo scores from the pre-mating to the mating period indicated a large variation among individuals in the magnitude of within-individual dominance change, with some individuals having a large slope and others having a near-zero slope. After the mating season, the within-individual slope of Elo scores for both sexes regained a similarly small standard deviation to the pre-mating period, but no longer showed a sex-specific directionality. Rather, both males and females had markedly smaller, near-

zero slopes, indicating a pattern of high within-individual consistency of dominance ($\overline{X}slope_{\text{male}}$ = -0.59, SD_{male} = 2.48, $\overline{X}slope_{\text{female}}$ = 0.92, SD_{female} = 3.99).

All adult and subadult females were dominant to all adult and subadult males for the majority of all time periods as indicated by the generally higher Elo scores among females (figure 3; females shown in shades of red and males in shades of blue). Therefore, sex and Elo score were strongly correlated in the three social groups (Spearman: $r = 0.74$, $p < 0.001$). Furthermore, the dominance rank order was generally stable for all three groups in all time periods. In fact, two of the groups displayed consistently higher *S* indices in each consecutive time period (Group 1: $S_{t1} = 0.982$, $S_{t2} = 0.993$, $S_{t3} = 0.999$; Group 3: $S_{t1} = 0.978$, $S_{t2} = 0.987$, $S_{t3} = 0.996$) indicating few rank changes. Group 2, on the other hand, had similarly high *S* indices in the pre- and post-mating period as the other groups but declined in rank stability during mating (Group 2: $S_{t1} = 0$ .982, $S_{t2} = 0.966$, $S_{t3} = 0.995$). Therefore, in all groups, the post-mating period was the most stable of all three time periods, but the mating season showed different stability patterns across groups, which was likely caused by the permanent rank shifts of several females. Specifically, the first (Ma; figure 3) and second ranked female (Sa; figure 3) in the pre-mating period permanently switched rank positions during the mating season. Similarly, the third (Bn; figure 3) and fourth ranked female (Dt; figure 3) in the pre-mating period also permanently switch rank positions during the mating season. As a result of these two rank order shifts, the first (Mn) and fourth (Dt) ranked females (Mn) in the pre-mating period ended up with very similar Elo score to one another at the end of the mating season, a pattern that persisted throughout the post-mating period. Three other permanent rank shifts occurred in this group, all in the post-mating period and all occurring among resident and transfer males or between the two

transfer males (Fr-Ha, Jo-Ha, Tm-Ha; figure 3). Therefore, these post-mating rank shifts occurred because of changes in group composition, not changes in previously established relationships.

No permanent rank shifts occurred between males and females in any group (figure 4). Permanent rank shifts also did not occur in Groups 1 or 3 throughout the entire observation period but there were multiple temporary rank switches in all groups (Group 1, 29% of dyads; Group 2, 40% of dyads; Group 3, 38% of dyads). Temporary rank switches generally lasted for one or two days and were most likely to occur among dyads with similar Elo scores (figure 4). Furthermore, those with more similar scores switched ranks more often than those with more dissimilar scores (figure 4). In fact, no rank switching (permanent or temporary) occurred for dyads with over a 600-point difference in their Elo scores (figure 4). Overall, 64% of dyads never switched rank, 16% switched rank two or fewer times, and 20% switched rank three or more times (figure 4). Of those dyads that switched rank (temporarily or permanently), 49% were male-male dyads, 28% were female-male dyads, and 23% were female-female dyads. Therefore, male-male dyads were more unstable in terms of rank-order than dyads involving females. Finally, most rank switching, 42%, occurred before the mating season, 30% occurred during the mating season, and 28%

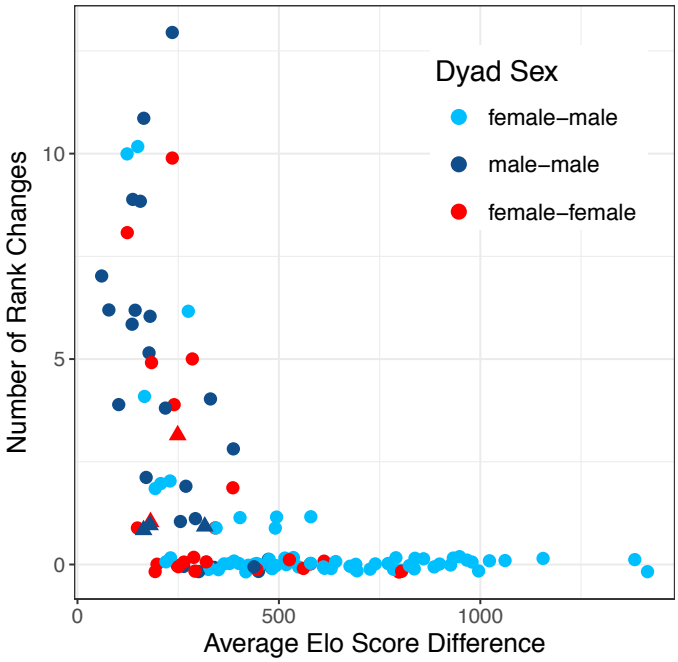


Figure 4. Relationship between average Elo score difference between pairs of individuals and their total number of rank changes. A triangle (▲) denotes a permanent rank change.

occurred after the mating season. The distribution of rank change events across time was similar for dyads of all sex combinations.

Discussion:

Dominance is important for the fitness and social functioning of many species (Altmann 1962, Koyama 1988, Isbell et al. 1999, Snyder-Mackler et al. 2016, Whitten 1984). However, it is traditionally assessed using matrix-based methods, which often cannot capture short-term and rapid shifts in dominance dynamics, a well-known characteristic of the ringtailed lemur mating season (Budnitz and Dainis 1975, Gould 1994, Jolly 1966, Sauther 1991, Sussman 1991). I used the Elo-rating method to investigate the structure and stability of the dominance relationships and hierarchies of ringtailed lemurs before, during, and after the highly chaotic, annual mating season. Ringtailed lemurs, like all lemur species, have a single mating period lasting only a couple of weeks each year. Elo-rating can easily deal with these dynamic patterns and is thus a better method to answer questions about short-term dominance stability. This paper represents the first application of the Elo-rating method to assess the dominance patterns of a lemur species. Thanks to this methodological advance, I found that the magnitude of dominance did fluctuate during the mating season and some dominance ranks did change as a result, but rank order in these three lemur groups was generally stable at all time periods, suggesting that dominance rank order may be more robust in this species than is often described in the literature (e.g. Jolly 1966, Budnitz and Dainis 1975). The idea that mating behavior is associated with social instability and changes in the characteristics of dominance in a social group is well-supported (e.g. Budnitz and Dainis 1975, Gould 1994, Huchard and Cowlishaw 2011, Jolly 1966, McCauley 2010, Sauther 1991, Sussman 1991, Wingfield et al. 1990).

Although observable mating-related behaviors largely overlapped in the three groups in this study, the timing of births differed, with females in Group 3 giving birth almost two weeks before females in the other two groups. The timing of birthing dates for each group was similar to what was observed in the prior year, with Group 3 giving birth earlier than the other groups. The variation among the three groups may be explained by the habitat quality of their territories. The habitat of Group 3 is in the eastern gallery forest and has a much higher and denser canopy than the habitats of Group 1 or 2. Group 2, in particular, lived in the poorest quality territory, which was located primarily outside of the reserve in degraded habitat with less vegetation and canopy cover (Gemmill and Gould 2008), lower food quality (Gould 1996; Sauther 1992, 1993), and high human and livestock presence (personal observation). The timing of ringtailed lemur reproduction is tightly tied to food availability (Sauther et al. 1999), which may explain the earlier reproduction and birthing in Group 3. Furthermore, females in Group 2 gave birth over the longest period of time in 2015 and roughly in the order of dominance, which may reflect differential access to resources or mates in accordance with dominance rank. Females in Group 1 and 2 also gave birth in the order of dominance rank in 2014. Given the tight synchrony of birthing dates in Group 3 for two years in a row, dominance in groups living in high quality habitats may be less important for structuring mating and food priority, and thus birthing dates, than it is for groups living in lower quality habitat. In a previous study at Beza Mahafaly Reserve, high ranking ringtailed lemurs living in disturbed habitat outside of the reserve expressed more pronounced feeding priority than did high ranking females in a group living within the reserve (Gemmill and Gould 2008).

Using a combination of matrix-based methods and Elo-Rating, I generated multiple measures of instability in dominance patterns in the three social groups from the pre-mating to

the post-mating season. I assessed the linearity of dominance using traditional matrix-based methods and found it to be highest in the pre- and post-mating period and decline during the three- to four-week mating season in all groups. Using Elo-rating, I assessed the frequency, permanence, and temporal pattern of rank changes. Male-male dyads had more temporary rank changes than dyads that included females but these rank changes were similarly spaced out among the three time periods. Therefore, males had more unstable dominance relationships in all time periods, possibly because the magnitude of rank differences for male-male dyads was often relatively small, making temporary rank changes more likely. Male-female and female-female temporary rank changes were similarly common as one another. This pattern of temporary rank changes did not differ across groups.

Because these temporary rank changes tended to occur for only one or two days at a time and among individuals with similar Elo scores, most time periods, including mating, had high overall rank stability. The post-mating period had the highest rank stability in all groups. This rank stability persisted even though the mating season displayed a relatively large inconsistency in the fluctuations of within-individual dominance scores, with some individuals changing the magnitude of their dominance score greatly and others remaining stable throughout the mating season. However, the inconsistent fluctuations in the magnitude of dominance across individuals did not coincide with marked instability in the order of the dominance rank for Groups 1 and 3. Group 2, however, did experience a decline in rank order stability and a non-linear dominance rank during the mating season. This shift was the result of rank changes occurring between two previously established female-female dyads.

Females were overwhelmingly dominant to males in all group and across all time periods, as has been observed by other authors (Gould 1994, Kappeler 1990, Taylor 1986, Wright 1999).

Interestingly, most females *increased* their dominance magnitude throughout the pre-mating season, while most males *decreased* their dominance magnitude at this time. However, the change in dominance magnitude did not correspond with a permanent change in dominance rank in either sex, and corresponded to only a slightly higher number of temporary rank shifts than in the mating and post-mating period. During the pre-mating period when the sex-specific directional shifts in dominance magnitude occurred in this study, males are usually preparing for the upcoming mating season, as evidenced in other studies by the gradual enlargement of their testicles (Pereira 1993) and increase in number of erections (Sauther 1991). Female dominance and male submission in lemurs has been hypothesized to be the result of seasonal energy constraints during reproductive periods (Pereira 1999, Wright 1999) and it has been suggested that males may submit readily to females (Pereira 1993), especially in relation to feeding priority near or during the mating season, which can` have the effect of better female nutrition and fertility (Pochron et al. 2003). Furthermore, it has been proposed in other lemur species that females may prefer to mate with males that submit to them (Richard 1992), suggesting that submission to females may be an effective reproductive strategy. Therefore, female choice has been hypothesized to govern reproductive outcomes in ringtailed lemurs (Pereira and Weiss 1991, Sauther et al. 1999). An increase in the magnitude of female dominance in conjunction with a complementary decline in the magnitude of male dominance just prior to mating suggests a role of female dominance in structuring reproductive outcomes in the subsequent mating season. In support of this point is the observation that females in two of the groups in this study gave birth to infants roughly in the order of dominance rank. Infant mortality is high in ringtailed lemurs (Sussman 1991) and being born earlier in the season may be beneficial because the infant is weaned during periods of high food availability or at an older age (Sauther 1991).

Male dominance over other males may also structure reproductive outcomes as dominant males have been observed to have first mating priority, followed by transfer/extra-group males, and then subordinate males (Sauther 1991). The first male to mate with a female often fathers the infant (Pereira and Weiss 1991). However, ringtailed lemurs have also been observed to successfully father infants irrespective of their dominance status (Pereira and Weiss 1991). In fact, several authors have described the mating period in ringtailed lemurs as a time when the male dominance hierarchy temporarily breaks down, facilitating sexual access to all troop females (Budnitz and Dainis 1975, Gould 1994, Jolly 1966, Sauther 1991, Sussman 1991). Therefore, the role of male dominance in paternity and reproductive success remains unclear.

Mating behavior is known to be associated with social instability and changes in the characteristics of dominance in a social group (e.g. Budnitz and Dainis 1975, Gould 1994, Huchard and Cowlishaw 2011, Jolly 1966, McCauley 2010, Sauther 1991, Sussman 1991, Wingfield et al. 1990). However, as the above results suggest, instability in the linearity and magnitude of dominance did not generally coincide with rank order instability (with the exception of Group 2 during the mating season), suggesting that dominance rank order may be more robust in this species than is often described in the literature. The amount of rank instability in a social group has health and fitness implications because it has been reported to be associated with high levels of stress, particularly for dominant animals (Sapolsky 1983, 1992). Furthermore, the declines in linearity and the inconsistency in the magnitude of dominance during mating did not seem to be caused by a breakdown of the *male* hierarchy. In fact, I only found a decline in rank order stability during mating in one group and its cause was likely *female* rank changes.

Elo-rating may offer the ability to analyze dominance patterns, particularly those that change rapidly and temporarily, in a more nuanced and detailed manner than traditional matrix-

based methods and may lead to different characterizations than are currently available in the literature. This paper represents the first application of the Elo-Rating method to assess the dominance patterns of a lemur species. Using this method, I found that dominance may be important for structuring mating and food priority in groups living in poor quality habitats. I also found that changes in the magnitude of dominance did not necessary imply that dominance rank order became unstable, adding nuance to the observation that dominance breaks down during mating in this species. Researchers studying organisms that experience concentrated periods of structural changes in group membership and interaction patterns would benefit from applying Elo-rating methodology to their social behavior data and may identify nuanced patterns that modify or add detail to those described in the literature for even the most well-studied of species.

Literature Cited

Albers, P. C. H. and de Vries, H. 2001. Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour* 61: 489–95.

Altmann, J. 1974. Observational study of behavior: sampling methods. *Behaviour* 49: 227–267.

Altmann, S. A. 1962. A field study of the sociobiology of rhesus monkeys *Macaca mulatta*. *Annals of the New York Academy of Sciences* 102: 338–435.

Appleby, M. C. 1983. The probability of linearity in hierarchies. *Animal Behaviour* 31:600-608.

Bergstrom, M. L. and Fedigan, L. M. 2010. Dominance among female white-faced capuchin monkeys (*Cebus capucinus*): hierarchical linearity, nepotism, strength and stability. *Behaviour* 147: 899-931.

Bernstein, I. S. 1981. Dominance: The baby and the bathwater. *Behavioral and Brain Sciences* 4: 419-457.

Brockman, D. K., Godfrey, L. R., Dollar, L. J. and Ratsirarson, J. 2008. Evidence of invasive *Felis silvestris* predation on *Propithecus verreauxi* at Beza Mahafaly Special Reserve, Madagascar. *International Journal of Primatology* 29: 135–52.

Budnitz, N. and Dainis, K. 1975. *Lemur catta*: Ecology and Behavior. In: Lemur Biology. (Ed. by I. Tattersall and R. W. Sussman), pp. 219-235. New York and London: Plenum Press.

Curley, J. P. 2106. Compete: Analyzing competitive interaction data (R package, version 0.1), https://cran.r-project.org/web/packages/compete/index.html.

Elo, A. E. 1978. The Rating of Chess Players, Past and Present. New York: Arco.

Gemmill, A. and Gould, L. 2008. Microhabitat variation and its effects on dietary composition and intragroup feeding interactions between adult female *Lemur catta* during the dry season at Beza Mahafaly Special Reserve, southwestern Madagascar. *International Journal of Primatology* 29: 1511–33.

Gould L. 1994. Patterns of affiliative behavior in adult male ring-tailed lemurs (*Lemur catta*) at the Beza-Mahafaly Reserve, Madagascar. Ph.D. dissertation. St. Louis: Washington University.

Gould, L. 1996. Male-female affiliative relationships in naturally occurring ringtailed lemurs (*Lemur catta*) at the Beza-Mahafaly Reserve, Madagascar. *American Journal of Primatology* 39: 63–78.

Huchard, E. and Cowlishaw, G. 2011. Female-female aggression around mating: An extra cost of sociality in a multimale primate society. *Behavioral Ecology* 22: 1003–11.

Isbell, L. A., Pruetz, J. D., Lewis, M. and Young, T. P. 1999. Rank differences in ecological behavior: A comparative study of patas monkeys (*Erythrocebus patas*) and vervets (*Cercopithecus aethiops*). *International Journal of Primatology* 20: 257–272.

Jolly, A. 1966. Lemur social behavior and primate intelligence. *Science* 153: 501–506.

Kappeler, P. M. 1990. Female dominance in *Lemur catta*: more than just female feeding priority? *Folia Primatologica* 55: 92–95.

Koyama, N. 1988. Mating behavior of ring-tailed lemurs (*Lemur catta*) at Berenty, Madagascar. *Primates* 29: 163–175.

Majolo, B., Lehmann, J., De Bortoli Vizioli, A. and Schino, G. 2012. Fitness-related benefits of dominance in primates. *American Journal of Physical Anthropology* 147: 652–60.

Macedonia, J. M. 1993. Vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatologica* 61: 186–217.

McCauley, S. J. 2010. Body size and social dominance influence breeding dispersal in male *Pachydiplax longipennis* (*Odonata*). *Ecological Entomology* 35: 377–85.

Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., Widdig, A. and Engelhardt, A. 2011. Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour* 82: 911–921.

Neumann, C. and Kulik, L. 2014. EloRating: Animal dominance hierarchies by Elo rating (R package, version 0.43), https://CRAN.R-project.org/package=EloRating.

Norscia, I. and Palagi, E. 2015. The socio-matrix reloaded: From hierarchy to dominance profile in wild lemurs. *PeerJ* 3: e729. doi:10.7717/peerj.729.

Ostner, J., Nunn, C. L. and Schulke, O. 2008. Female reproductive synchrony predicts skewed paternity across primates. *Behavioral Ecology* 19: 1150–1158.

Pereira, M. E. 1993. Seasonal adjustment of growth rate and adult body weights in ringtailed lemurs. In: Lemur Social Systems and Ecological Basis (Ed. by P. M. Kappeler, J. U. Ganzhorn), pp 205–221. New York: Plenum Press.

Pereira, M. E. and Weiss, M. L. 1991. Female mate choice, male migration, and the threat of infanticide in ringtailed lemurs. *Behavioral Ecology and Sociobiology* 28: 141-152.

Pochron, S. T., Fitzgerald, J., Gilbert, C. C., Lawrence, D. Grgas, M., Rakotonirina, G., Ratsimbazafy, R., Rakotosoa, R. and Wright, C. W. 2003. Patterns of female dominance in *Propithecus diadema edwardsi* of Ranomafana National Park, Madagascar. *American Journal of Primatology* 61: 173–85.

Pusey, A., Williams, J. and Goodall, J. 1997. The influence of dominance rank on the reproductive success of female chimpanzees. *Science* 277: 828–831.

Richard, A. F. 1992. Aggressive competition between males, female-controlled polygyny and sexual monomorphism in a Malagasy primate, *Propithecus verreauxi*. *Journal of Human Evolution* 22: 395–406.

Sapolsky, R. M. 1983. Endocrine aspects of social instability in the olive baboon (*Papio anubis*). *American Journal of Primatology* 5: 365–379.

Sapolsky, R. M. 1992. Cortisol concentrations and the social significance of rank instability among wild baboons. *Psychoneuroendocrinology* 17: 701–702.

Sauther, M. L. 1991. Reproductive behavior of free-ranging *Lemur catta* at Beza Mahafaly Special Reserve, Madagascar. *American Journal of Physical Anthropology* 84: 463–77.

Sauther, M. L. 1992. Effects of reproductive state, social rank, and group size on resource use among free-ranging ringtailed lemurs (*Lemur catta*) of Madagascar. Dissertation, Washington University.

Sauther, M. L. 1993. Resource competition in wild Populations of ringtailed lemurs (*Lemur catta*): Implications for female dominance. In: Lemur Social Systems and their Ecological Basis (Ed. by P. M. Kappeler and J. U Ganzhorn), pp 135-152. New York: Plenum Press.

Sauther, M. L, Sussman, R. W. and Gould, L. 1999. The socioecology of the ringtailed lemur: Thirty-five years of research. *Evolutionary Anthropology* 8: 120–32.

van Schaik, C. P. 1989. The ecology of social relationships amongst female primates. In: Comparative Socioecology: The behavioural ecology of humans and other animals. (Ed. by V. Standen and R. Foley), pp 195-218. Oxford: Blackwell Scientific.

Snyder-Mackler, N., Kohn, J. N., Barreiro, L. B., Johnson, Z. P., Wilson, M. E. and Tung, J. 2016. Social status drives social relationships in groups of unrelated female rhesus macaques. *Animal Behaviour* 111: 307-317.

Sterck, E. H. M., Watts, D. P. and van Schaik, C. P. 1997. The evolution of female social relationships in nonhuman primates. *Behavioral Ecology and Sociobiology* 41: 291-309.

Sussman, R. W. 1991. Demography and social organization of free-ranging *Lemur catta* in the Beza Mahafaly Reserve, Madagascar. *American Journal of Physical Anthropology* 84:43–58.

Sussman, R W. 1992. Male life-history and intergroup mobility among ringtailed lemurs (*Lemur catta*). *International Journal of Primatology* 13: 395–413.

Sussman, R. W. and Rakotozafy, A. 1994. Plant diversity and structural analysis of a tropical dry forest in southwestern Madagascar. *Biotropica* 26: 241–254.

Taylor, L. 1986. Kinship, dominance and social organization in a semi free-ranging group of ringtailed lemurs (*Lemur catta*). Ph.D. dissertation, Washington University, St. Louis, Missouri.

de Vries, H. 1995. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour* 50:1375-1389.

Whitten, P. L. 1984. Competition among female vervet monkeys. In: *Female Primates* (Ed. by M. F. Small), pp 127–140. New York: Liss.

Wingfield J. C., Hegner, R. E., Dufty, R. E. and Ball, G. F. 1990. The challenge hypothesis: theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist* 136: 829–846.

Wright, P. C. 1999. Lemur traits and Madagascar ecology: Coping with an island environment. *Yearbook of Physical Anthropology* 42: 31-72.

A longitudinal analysis of the ringtailed lemur social network and its implications for pathogen transmission

Chapter 2

Abstract: Social relationships are a potential cause of pathogen transmission, but the direct and indirect dynamics of these relationships are often difficult to observe directly. I collected >1000 hours of detailed social behavior data on 29 ringtailed lemurs living in three sympatric social groups in Beza Mahafaly Reserve in Madagascar from March–September 2015, a period that spaned the annual mating season. Using Social Network Analysis, I measured the direct and indirect connectedness of each social group through agonistic and affiliative interactions, the latter of which is the mode of interaction most likely to cause the transmission of pathogens. Active and passive affiliation had different temporal patterns with individuals decreasing the overall time spent in active affiliation and increasing the time spent in passive affiliation from the pre- to the post-mating period. Further, the connectedness of a social group was associated with its microhabitat, such that the group living in degraded habitat displayed the lowest modularity for active affiliation, highest graph density for active and passive affiliation, and the most relationships and highest network-wide centrality overall and through time for active affiliation, all of which could cause heightened pathogen transmission. Animals living in degraded habitat are often expected to have a higher infection rate because of the increased presence of pathogens from humans and livestock, but these results suggest the effects may be amplified by an increase in network connectedness.

Introduction:

Social animals build complex relationships with members of their group. The elements of social relationships are most readily quantified at the level of the dyad but there is additional information to be gained from quantifying them at the level of the individual and the entire group, facilitating the identification of emergent properties of dyadic social interactions that are difficult to observe directly. Assessing these emergent properties provides insight into aspects of these networks that would otherwise remain elusive (Foster et al. 2008) and is an important step towards understanding how underlying social structure influences the transmission of pathogens within a social group (Craft 2015). Pathogen transmission is an important cost of sociality (Alexander 1974, Corner et al. 2003, Freeland 1976, 1979) and social structure has been hypothesized to regulate pathogen sharing through the mediation of contact patterns (Freeland 1976, 1979, Griffin and Nunn 2012, Hess 1996, Romano et al. 2016, Wilson et al. 2003). For example, a simulation of the spread of infection across populations with different levels of inter-individual interaction using data from real primate groups demonstrated that greater modularity

in community network structure reduced the spread of socially transmitted parasites. Therefore, modular groups may have a lower prevalence, abundance and diversity of parasites (Griffin and Nunn 2012). Conversely, highly connected individuals are more likely to become infected with pathogens and to transmit that infection faster and to more individuals than their less-connected group members (Rimbach et al. 2015, Romano et al. 2016, VanderWaal et al. 2014a, b). As a result, pathogens have been hypothesized to cause selection on the evolution of mating systems and social interactions (Møller et al. 2001). However, little is known about the network dynamics of hosts and pathogens in real social groups (Romano et al. 2016).

One approach to assessing the emergent properties in social interaction is through Social Network Analysis, which allows mapping of relationships onto a visual space and measuring of the flow of different modes of contact among individuals to identify underlying social structure. A powerful aspect of this method is that both direct and indirect structure can be identified for individual group members as well as for the social group as a whole (Boyland et al. 2016). These analyses have provided insight into the mechanisms underlying variation in an individual's role in the social network, which is important to advance our understanding of how social groups function (Craft 2015). For example, traits such as age, sex, dominance rank, and family size have been found to influence pathogen load (Ezenwa 2004) and structure networks in several species (Friant et al. 2016, red-capped mangabeys; MacIntosh et al. 2012, Japanese macaques; Rimbach et al. 2015, brown spider monkeys; Rushmore et al. 2013, chimpanzees), but there are few consistent patterns across studies (Rushmore et al. 2017). Furthermore, little is known about whether individuals maintain their roles over time (Sih et al. 2009; Wilson et al. 2003) or how the networks as a whole change over time, particularly during high stakes periods such as seasons of high predation (but see Kelley et al. 2011 for an example) and mating (but see

Hamede et al. 2009 for an example). Seasonality in network patterns can result in temporal variation in pathogen transmission (Altizer et al. 2006, Rushmore et al. 2017) and there is evidence that the period during and near the mating season can have particularly high disease transmission (e.g. van Schaik and Kerth 2017, Zohdy et al. 2012). While several recent studies have examined the dynamics of social networks through time (e.g. Boyland et al. 2016; Blonder and Dornhaus 2011; Hamede et al. 2009, Jeanseon 2012; Kelley et al. 2011, Rushmore et al. 2013), static networks, which dilute temporal variation in contact patterns, are still overwhelmingly used to study host-parasite relationships (Rushmore et al. 2017). Therefore, understanding the implications of temporally dynamic networks for disease transmission routes is still in its infancy.

In this study, I investigated the dynamic structure of ringtailed lemur social networks in three sympatric groups from March-September 2015 at Beza Mahafaly Special Reserve in southwestern Madagascar. This observation period spanned the singular annual mating season and provided the opportunity to determine affiliative and agonistic interaction patterns before, during, and after mating using Social Network Analysis. Although I expect affiliation to be the more effective mode of behavior for the transmission of pathogens, I present network patterns for both affiliation and agonism because the role of agonism in transmission has never been assessed despite the mechanisms of socially facilitated pathogen transmission being poorly understood (Blyton et al. 2014). Furthermore, using multiple measures of social behavior is important to understand the social relationships and underlying social structure of group-living animals (Hinde 1976, Lehman and Ross 2011). Finally, I use the resulting network patterns to develop hypotheses for the transmission of parasites in these groups. Specifically, I ask: 1) How do the individual-level and group-level characteristics of the social network change over time? 2)

How do an individual's characteristics (i.e. group membership, sex, and dominance) influence its role in the network?

Methods:

Study Site, Subjects and Data Collection

I collected detailed social behavior data on 29 individually identifiable ringtailed lemurs living in three sympatric social groups: Group 1(Pink Group), Group 2 (Purple Group), and Group 3 (Red Group) in Parcel 1of Beza Mahafaly Special Reserve in southwestern Madagascar from March –September 2015. Parcel 1 consists of 80 hectares of mixed vegetation forest (figure 1). The western portion of Parcel 1 consists of xerophytic/scrub forest and the eastern portion consists of lush gallery forest adjacent to the Sakamena River (Sussman and Rakotozafy 1994). The territory of Group 1 was located in the western scrub forest and the territory of Group 3 was located within the lush gallery forest. The territory of Group 2 was south and southwest of parcel 1 and was almost entirely outside of the reserve. The area outside of the reserve is heavily disturbed dry forest with low canopy cover and sparse vegetation due to human logging and livestock foraging (Gemmill and Gould 2008) and a lower availability of preferred food (Gould 1996; Sauther 1992, 1993).

Group 1 consisted of 13 individuals (4 adult females, 1 subadult female, 1 juvenile female, 2 adult males, 1 subadult male, 1 juvenile male, 3 infant males). At its maximum size, Group 2 consisted of 12 individuals (4 adult females, 5 adult males, 1 subadult male, 2 infant males). At its maximum size, group 3 consisted of 18 individuals (5 adult females, 1 subadult female, 3 juvenile females, 4 adult males, 1 subadult male, 2 juvenile males and 2 infant males). Only the adult and subadult individuals were treated as focal animals resulting in 15 female and

14 male focal animals. The details of all interactions of the focal animal were documented and included in relevant analyses.

I conducted full day follows (~7:00-17:00) of each group twice per week. During these observation periods, I conducted continuous all occurrence focal sampling (Altmann 1974) of social behaviors (active affiliation, passive affiliation, low intensity (LI), and high intensity (HI) agonism; see table 1) of each adult and subadult individual for 45 – 60 min, aiming to observe each focal animal in the group in the same day. The order of observations was selected before observations began and was designed to generate similar observation times in the morning and afternoon for each focal individual every two weeks. This approach made it possible to treat each two-week time period as comparable bins that could be used to meaningfully assess the social behavior patterns of each individual. On average, each focal animal was observed for approximately 3-4 hours every two weeks for a total of 1,038 hours of observation (Group 1: 344 hours, Group 2: 341 hours; Group 3: 353 hours). To normalize the time spent observing each focal animal, I divided the total observation time for each focal animal in each two-week bin by the average observation time for each group in those 2-weeks.

When a social behavior occurred, I documented the time of initiation and completion, the type of behavior, the identity of relevant individuals, and the direction of the interaction. I documented all social interactions of the focal animal including, active affiliation, passive affiliation, high intensity agonism (HI agonism) and low intensity agonism (LI agonsim) (table 1). Active affiliation predominantly included grooming behaviors, whereas passive affiliation predominantly included resting in contact. HI agonism included behaviors that involved physical contact or the possibility of physical contact such as biting, hitting, lunging and chasing. LI agonism included behaviors that involved submission or threats of aggression. The directionality

of all agonistic behaviors were recorded in the direction of "winner" to "loser". Behaviors that ended almost immediately after they begin, as was common for agonism, were considered to last for a total duration of one second.

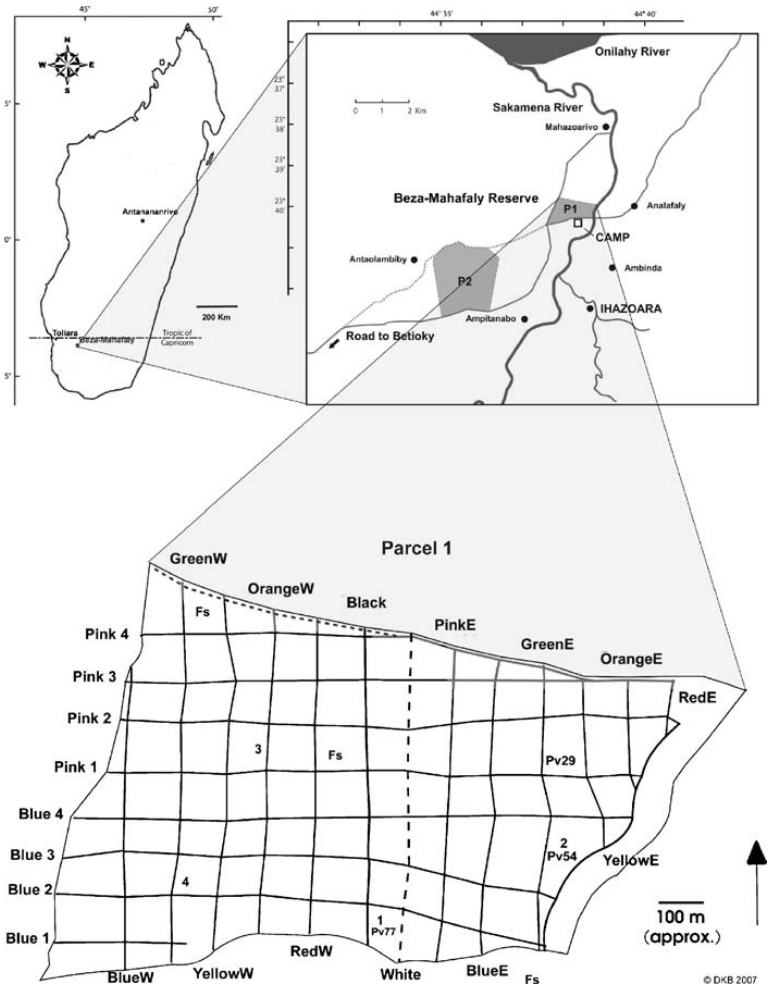


Figure 1. Map of Beza Mahafaly Special Reserve. From Brockman et al. 2008.

Active affiliation		**Passive affiliation**	
Groom Mutual groom Play		Any behavior not listed as "active affiliation" that occurs in contact with another individual	
High-intensity agonism		**Low-intensity agonism**	
Stink fight[a]	Grab	Plosive bark[b]	Chutter[b]
Lunge	Push	Mouth to face threat	Yip/spat call[b]
Chase	Bite	Squeal[b]	Displace/supplant
Hit	Contact fight	Cackle/deep spat[b]	Chase/lunge threat

Table 1. List of social interactions. [a]Between males only. [b]Terms from Macedonia (1993).

Network Analysis

The behavioral data were used to construct social networks for the four modes of social behavior across three time periods: pre-mating (March-April, mating/group transitions (May - June), post-mating (July – August). Ringtailed lemurs live in stable multi-male, multi-female groups and mate polygamously during the singular annual mating season. They are seasonal breeders with all females entering estrus for less than 24 hours each year and within one to three weeks of one another (Sauther et al. 1999). The mating period occurs in May and the peak of births occur in October (Sauther et al. 1999). My guide and I took note of all observed copulations, sexual swellings, copulatory plugs, and births and used these data to establish the timing of mating (Chapter 1).

To construct the networks, I first summed the interaction data in each consecutive two-week bin for each pair of focal animals. However, because each focal animal had slightly different lengths of observation time across and within bins, I normalized the interaction data in each two-week bin by dividing the raw pairwise interaction time among individuals in that bin by the normalized observation time (described above). This step allowed the time spent in social interaction for each individual in the social network to be directly comparable to all other individuals in its network within and across time periods. I then averaged together the bins

within the three time periods. For animals that died or left the group during the study period, I scaled up their interactions as though they were present for the entire time period in which they disappeared. This step is recommended only when assessing the structure of social interactions and is not recommended if interaction networks are to be correlated with patterns of pathogen transmission.

I used the R package igraph v 1.0 (Csardi and Nepusz 2006) to visualize interaction networks from each of the three time periods for each group. Social networks consist of nodes, which represent each focal individual, and edges, which represent interactions between individuals. The shape and hue of the node denotes the sex of the animal. The width of the edges was designed to be proportional to the normalized average time each pair of individuals spent in that mode of social behavior during that time period. For behavior modes in which there was a sender and receiver of each behavior, directed networks were constructed. In these networks, the arrows represent the directionality of each behavior. Passive affiliation is an undirected network and active affiliation and HI and LI agonism are directed networks. In agonism networks, the arrow points from the submissive to the dominant individual for all interactions, including submission. In active affiliation networks, mutual grooming was counted as two interactions with arrows going in both directions for the entire length of time the mutual grooming occurred. The position of the nodes in the network were determined by the Fruchterman-Reingold Algorithm, which is a force-directed layout algorithm in which the force between nodes is considered in determining their final position in network space (Fruchterman and Reingold 1991). The weight of the edges attracts the nodes to one another and the nodes themselves are repelled from one another. This algorithm maximizes these attractive forces and minimizes the repulsive forces (Fruchterman and Reingold 1991).

Next, I calculated multiple measures of connectedness for each network. I calculated two measure of network-level connectedness: graph density and modularity. Graph density is the proportion of potential direct connections in a network that exist as actual connections. It is a measure of how saturated the network is with direct connections (i.e. relationships) among individuals. I calculated modularity using the Spin Glass clustering algorithm, which finds communities within a network (Zhu et al. 2015). A community is a set of nodes with many connections inside the community but few outside the community. I visualized these communities by drawing polygons around nodes of the same community in all networks.

I also calculated six measures of node-level connectedness (i.e. centrality). Centrality refers to the connection to and impact of a node on the entire network. The measures of centrality are indegree, outdegree, in strength, out strength, in closeness, and out closeness. Degree is the number of connections a node has in a network. In a directed network where there is a performer and a receiver of each behavior, the directionality of the interactions can be taken into account. Indegree is the number of connections directed *towards* a node. An individual with high indegree has many groupmates directing behavior directly towards it and is described as prominent. Outdegree is the number of connections directed *away* from a node. An individual with high outdegree directs behavior towards many group mates and is described as influential. Strength is the sum of the weighted edges that connect one node to another. It incorporates the time nodes spent interacting with each other. In strength reflects the total time that behaviors are directed *towards* a node and out strength reflects the total time that behaviors are directed *away* from a node. Closeness is the inverse of the sum of the shortest path (i.e. geodesic) between all pairs of nodes. It measures how indirectly connected an individual is to all other members in its network. A high closeness score indicates high connectivity to the entire network. In closeness measures

the indirect or network-wide connectedness of a node through its *inward*-directed edges and out closeness measures the indirect or network-wide connectedness of a node through its *outward*-directed edges. All centrality measures (except strength) were normalized by the number of individuals in the network and were designed to range from 0-1. Therefore, these scores can be compared across social groups and time periods.

Statistical Analysis

I determined the percent of time individuals spent in each mode of social behavior by dividing the sum of the time in each behavior mode in each time period by the total observation time (with "out of site" time removed) for that period. To determine the influence of sex and time period on the percent of time spent in different modes of social interaction, I fit the behavioral data to a multi-level regression model using the R package lmer v. 1.1-12 (Bates et al. 2016) and tested for significance using a post hoc Tukey test and the R packages lmerTest v. 2.0-33 (Kuznetsova et al. 2016), and multcomp v. 1.4-6 (Hothorn et al. 2016). The regression model treated percent of time in each mode of social behavior as the dependent variable, and sex, group, and time period as fixed factors. Because I collected data on each focal individual at multiple time points, the model also treated the focal animal as a correlated random factor, which generated separate intercepts for each individual. All fixed factors were treated as additive.

To determine the role of time period, sex, dominance, and group membership on an individual's role in the social network, I fit the centrality data from each mode of social behavior to two alternative multi-level regression models. All models treated one of the six centrality scores as the dependent variable, time period and group as fixed factors, and focal animal as a correlated random factor. In addition, the models also included either sex (model A) or

dominance (model B). Dominance was calculated using a non-matrix-based method called Elo-rating (Neumann et al. 2011) via the R package EloRating v. 0.43 (Neumann and Kulik 2014). Using this method, Elo scores were determined for each day of observation for each individual based on the outcome (i.e. win, lose, or tie) of that day's agonistic interactions (HI and LI agonism) and on the pre-interaction probability of that outcome (Albers and de Vries 2001). The results of this analysis are reported in Chapter 1 of this dissertation. Elo score and sex were not included in the same model because they were strongly correlated in these groups (Chapter 1), as is typical for ringtailed lemurs (Kappeler 1990). Therefore, these two variables would explain similar pockets of variation in centrality, making it inappropriate to include them in the same model (Rushmore et al. 2013). Because the influence of both sex/dominance and group on centrality could differentially change with time period in the three groups, I incorporated an interaction effect between sex/dominance and time period (sex*time, dominance*time), and group and time period(group*time). I compared models A and B using AIC and selected the best one for each analysis. Because there were three-six dependent variables for each mode of social behavior, I used a Bonferroni-corrected critical p-value of 0.0083 for all analyses that used centrality as a dependent variable. For all analyses, I used qqplots to check the normality of the residuals of each model and if non-normality was discovered, I converted the dependent variable with a logit transformation and re-ran the model.

Results:

Modes of Social Behavior by Sex, Time and Group

Sex: Overall, affiliation was much more frequent than agonism (figure 2), the latter of which constituted a very small fraction (> 1%) of each individual's time. Males spent significantly less time in active (β= -1.45, df = 29, p < 0.001) and passive affiliation (β= -6.05, df = 27, p < 0.001) than females but there was no significant difference in sex for either mode of agonism.

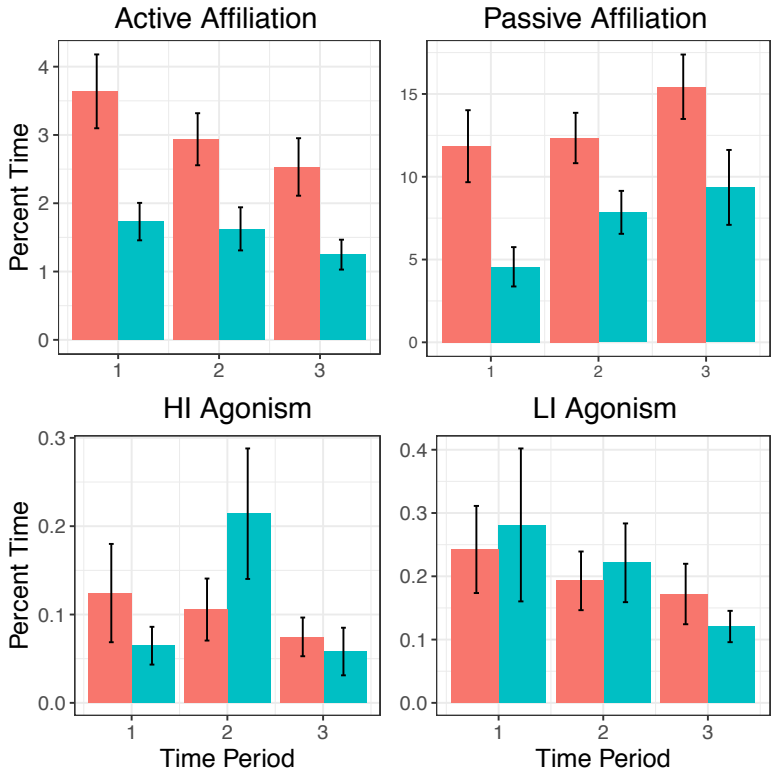


Figure 2. Percent of time in each mode of social behavior by sex and time period (male ; female).

Time Period: There was a gradual *decline* in active affiliation through time with significantly less occurring after the mating season than before the mating season (β= -0.73, df = 57, p < 0.05) (figure 2). Conversely, there was a gradual *increase* in passive affiliation over time with significantly more occurring after the mating season than before the mating season (β= 4.31, df = 56, p < 0.05) (figure 2). The percent of time spent in LI and HI agonism did not change significantly across time periods.

Group Membership: Individuals in Group 3 spent a significantly higher proportion of their overall time in active affiliation than Groups 1 (β = 1.07, df = 29, p < 0.05) and 2 (β = 1.3, df = 29, p < 0.001) but there was no significant difference for passive affiliation across groups

(figure 3). Individuals in Group 1 spent significantly less time in LI agonism than those in Group 2 ($\beta = 0.17$, $df = 28$, $p < 0.05$) and 3 ($\beta = 0.28$, $df = 28$, $p < 0.001$) but there was no significant difference in HI agonism across the groups (figure 3).

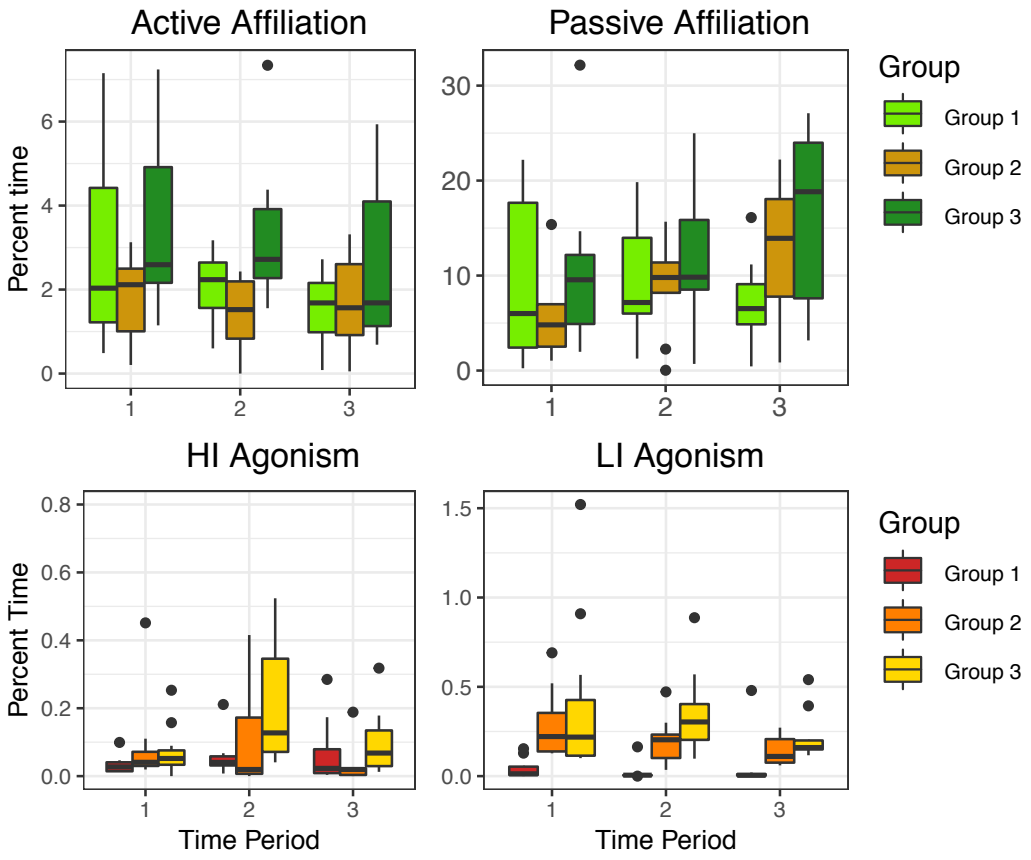


Figure 3. Percent of time in affiliation and agonism by group membership.

Network-Level Connectedness

Graph Density: For each group, the graph density of affiliative interactions was higher than for agonistic interactions, indicating that more dyads in each network were involved in affiliation than agonism (figure 4). The graph density of HI agonism remained relatively stable across time in Group 2 but increased slightly in Groups 1 and 3 during the mating season (figure

4). The graph density of LI agonism declined across time in all groups (figure 4). The graph density for active affiliation declined across time for Groups 1 and 3 but increased for Group 2. Passive affiliation showed variable patterns in the three groups. Overall, Group 2 had the highest graph density for active and passive affiliation, particularly during and after the mating season. Group 2 experienced new transfer males during the mating season that remained poorly integrated for several months. These males were not included in the above analysis. When including the new transfer males, the graph density for passive affiliation in Group 2 declined substantially from time 1 to time 2.

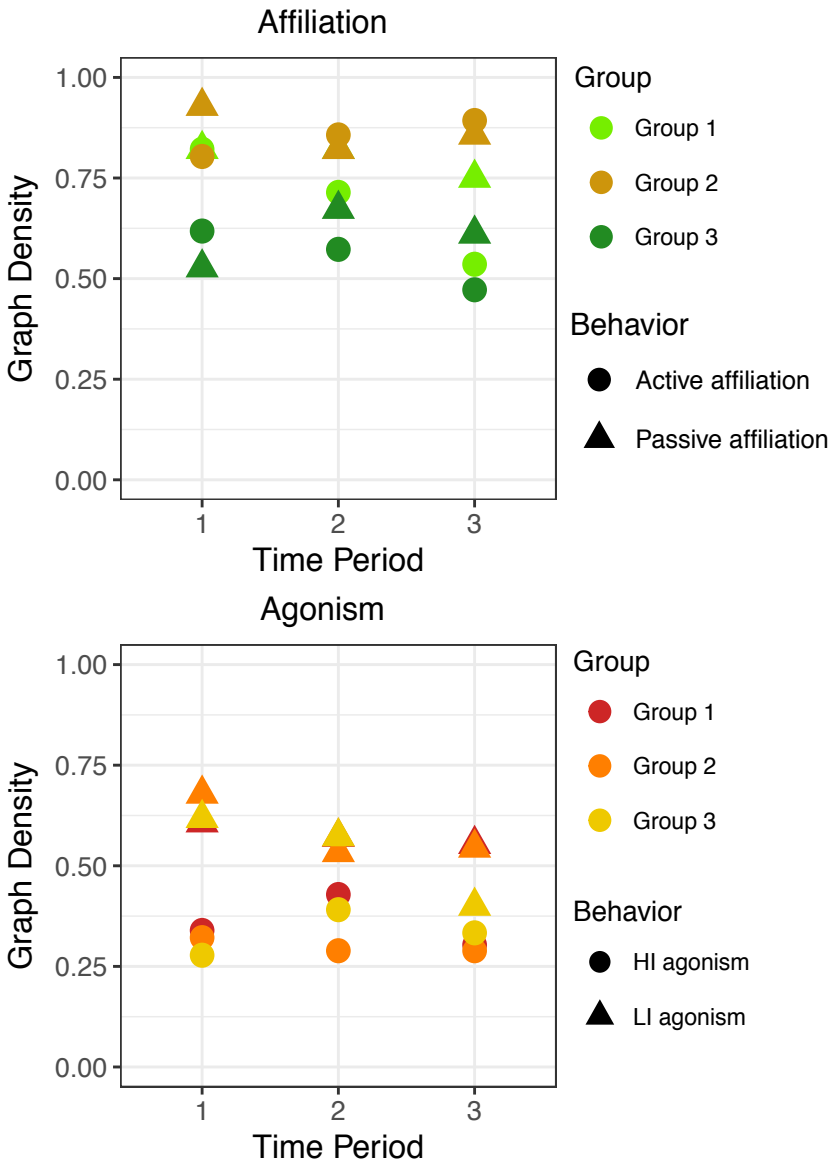


Figure 4. Graph density for each mode of interaction in each group through time.

Modularity: Agonism displayed similar patterns of modularity through time in each of the three groups. Specifically, HI agonism was consistently modular through time in all groups

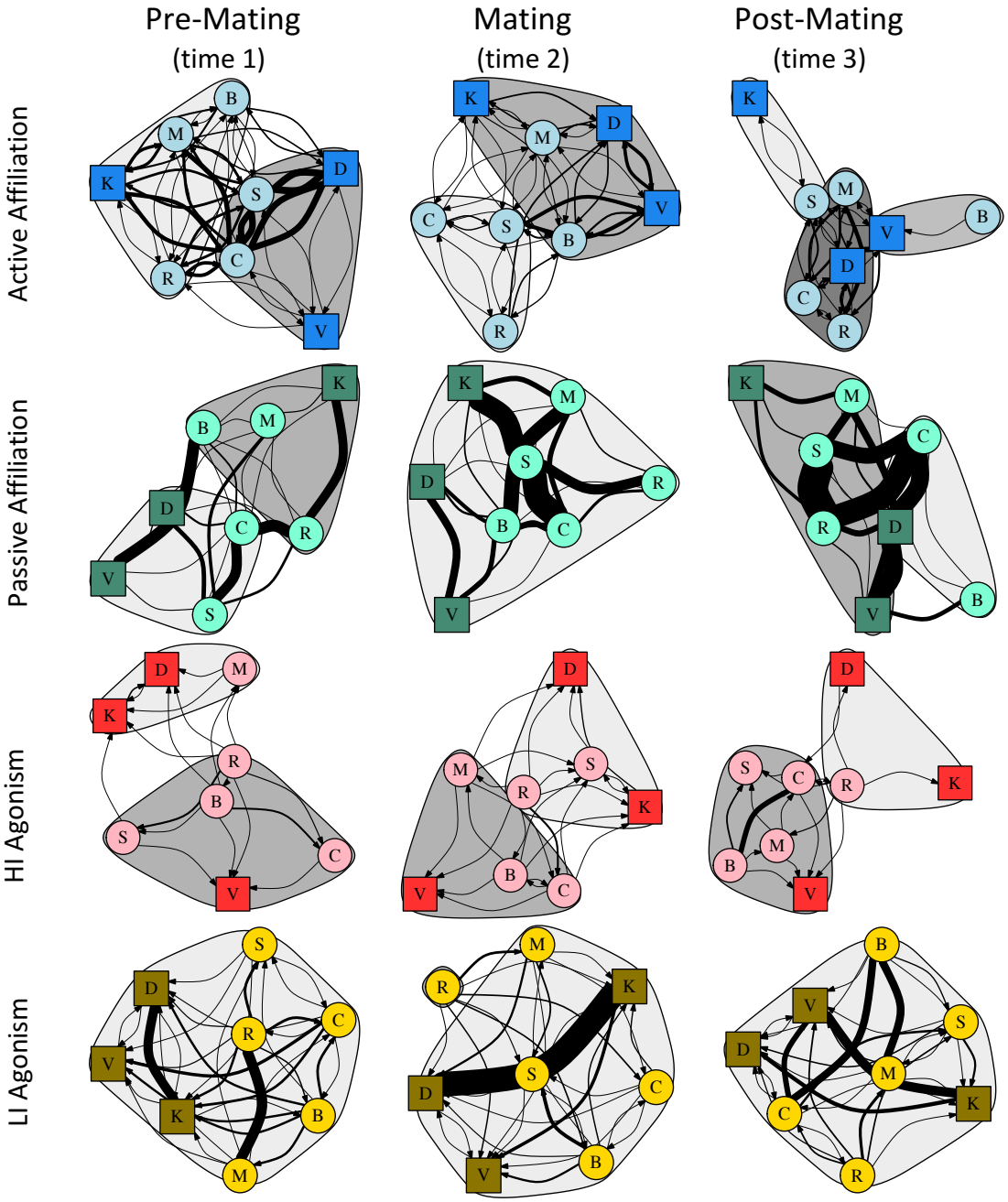
whereas LI agonism was non-modular in the pre-mating period and increased in modularity in either the mating or post-mating period for all groups (figure 5a-c). Conversely, affiliation displayed different modularity patterns in the three groups, with Group 2 being the least modular . Specifically, Group 2 was non-modular in all three time periods for active affiliation, whereas Groups 1 and 3 were always modular (figure 5a-c).

Module membership was fluid as individuals did not remain in the same module through time. However, a consistent pattern that can be observed is that nearly every module contained individuals of both sexes (figure 5a-c). Importantly, in the active affiliation modules of Groups 1 and 3, the dominant male(s) (K in Group 1, L and N in Group 3, figure 5a-c; see data for individuals Kn and Lk Chapter 1) were in a module with the lowest male-female sex ratio and were surrounded by females of the highest dominance status. Group 2 only had had single module for this time period.

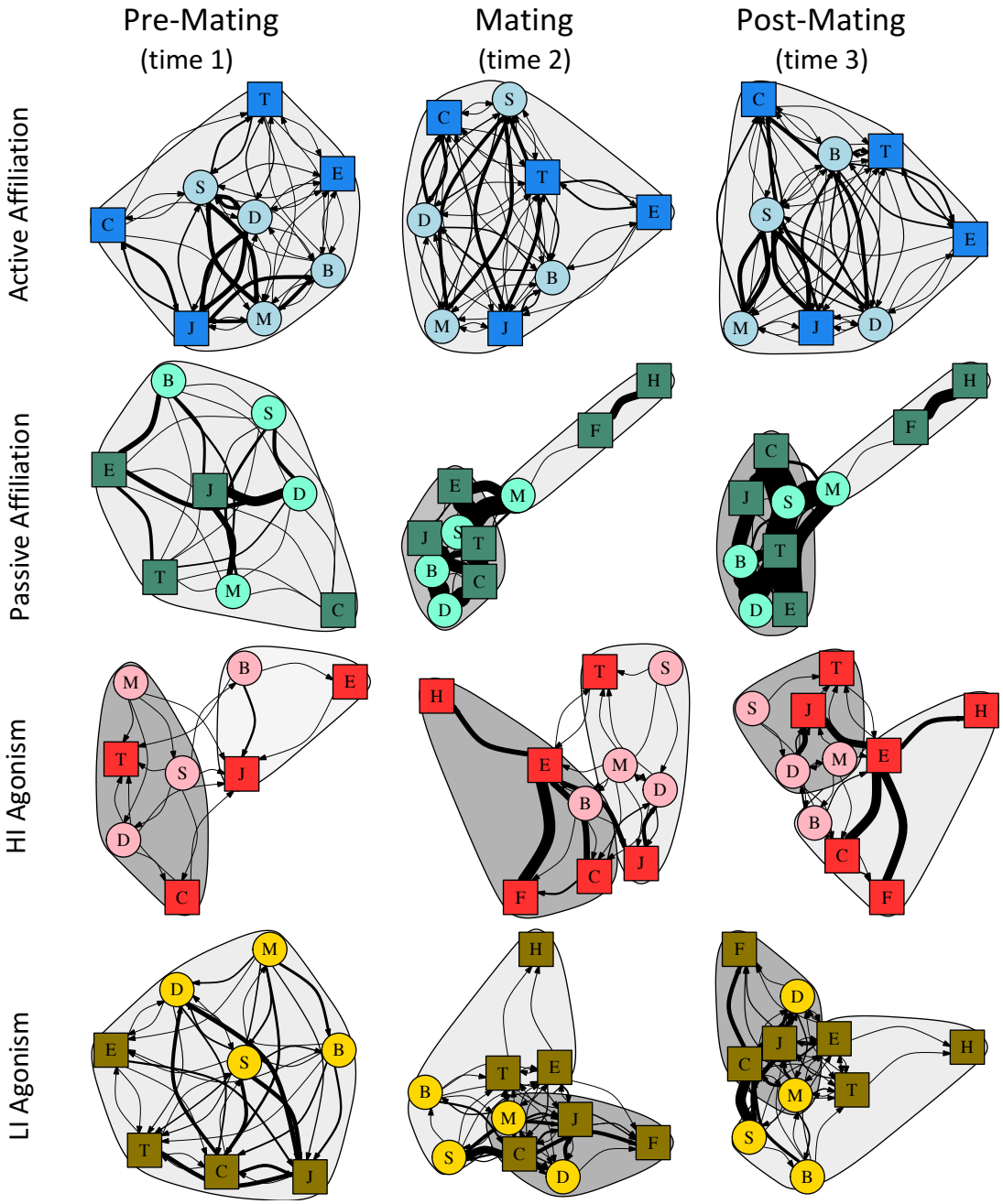
Node-level Centrality

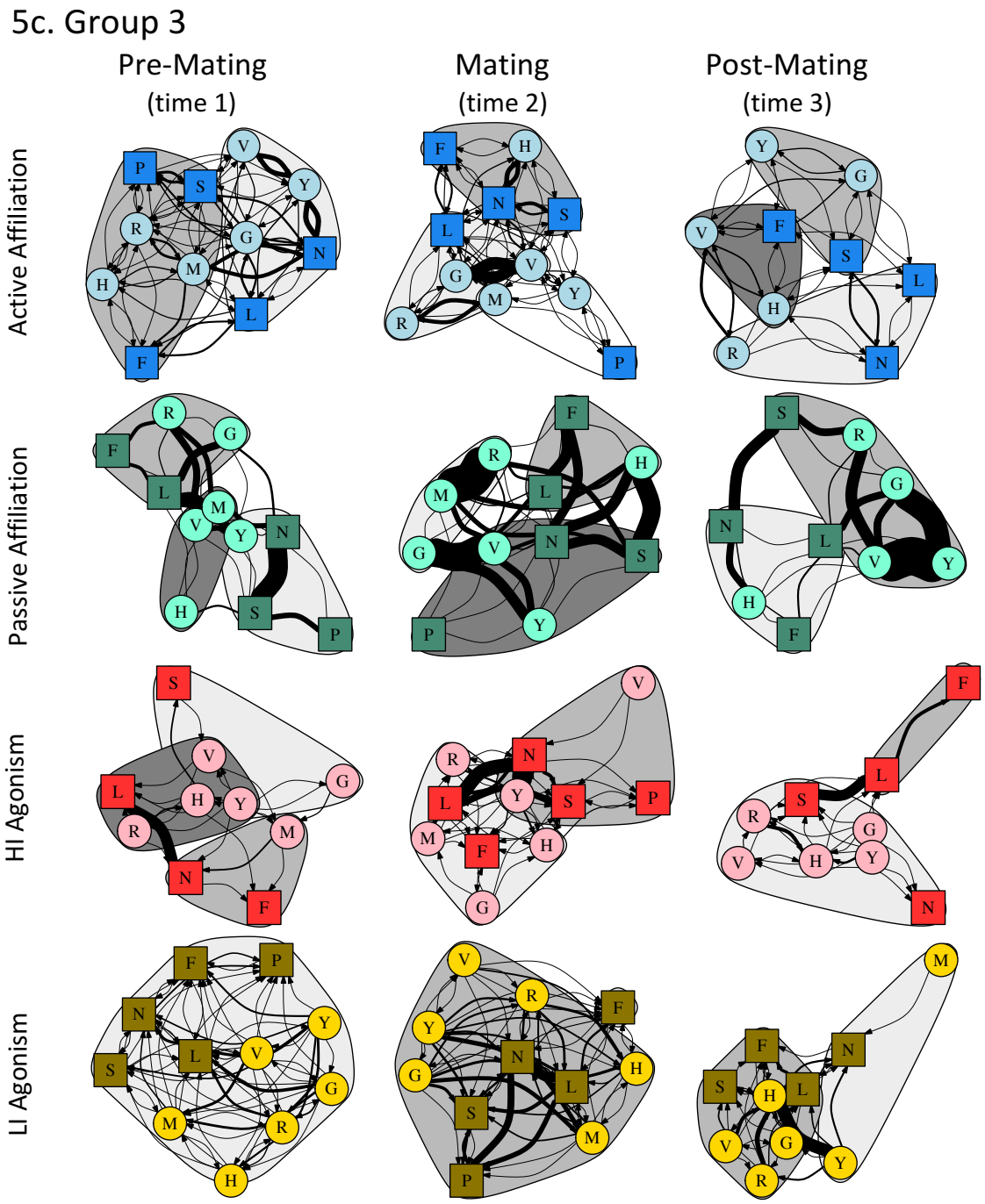
Model Selection: I used models (A and B) to assess the role of sex/dominance, group, and time period in affiliative and agonistic centrality. For affiliation, the model containing sex (model A) had the lowest AIC scores. For HI agonism, the model containing Elo scores (model B) had the lowest AIC score for most centrality measures. Exceptions were indegree and in closeness for LI agonism, which were best explained by the model containing sex (model B). The results for the best models are reported below. Because, sex, group membership, and time period were included in all models, when the outcome for one variables is reported, the other variables were included as covariates.

5a. Group 1


Pre-Mating
(time 1)
Mating
(time 2)
Post-Mating
(time 3)
Active Affiliation
Passive Affiliation
HI Agonism
LI Agonism

5b. Group 2


Pre-Mating
(time 1)
Mating
(time 2)
Post-Mating
(time 3)
Active Affiliation
Passive Affiliation
HI Agonism
LI Agonism



Figures 5a-c. Normalized networks for Groups 1 (5a), Group 2 (5b), and Group 3 (5c). The color of the node indicates the mode of behavior, the shape and shade indicates the sex of the individual (light circles = females; dark squares = males). The width of the edges is proportional to the average normalized time in social interaction in each 2-week bin within each of the time periods. The width of the edges is comparable within each behavior mode across time, and across modes of affiliation and modes of agonism. For maximum visibility, the edges of the affiliation networks were amplified 2x and the edges of the agonism networks were amplified 40x. The polygons around groups of nodes indicates modules calculated using the Spin Glass clustering algorithm.

Sex and Dominance: For affiliation, sex was not a significant predictor of direct centrality (i.e. degree and strength) and some forms of network-wide centrality (i.e. in closeness). Specifically, using a rigorous critical p-value of 0.0083, males and females had a similar number and strength of relationships for passive affiliation (degree: $\beta = -0.14$, $df = 23$, $p = 0.41$; strength: $\beta = -4.66$, $df = 22$, $p = 0.018$) and active affiliation (in degree: $\beta = -0.58$, $df = 23$, $p = 0.07$; outdegree: $\beta = -0.65$, $df = 23$, $p = 0.015$; in strength: $\beta = -0.78$, $df = 23$, $p = 0.14$; out strength: $\beta = -0.77$, $df = 23$, $p = 0.17$) (figures 6 and 7). They also had a similar network-wide centrality for

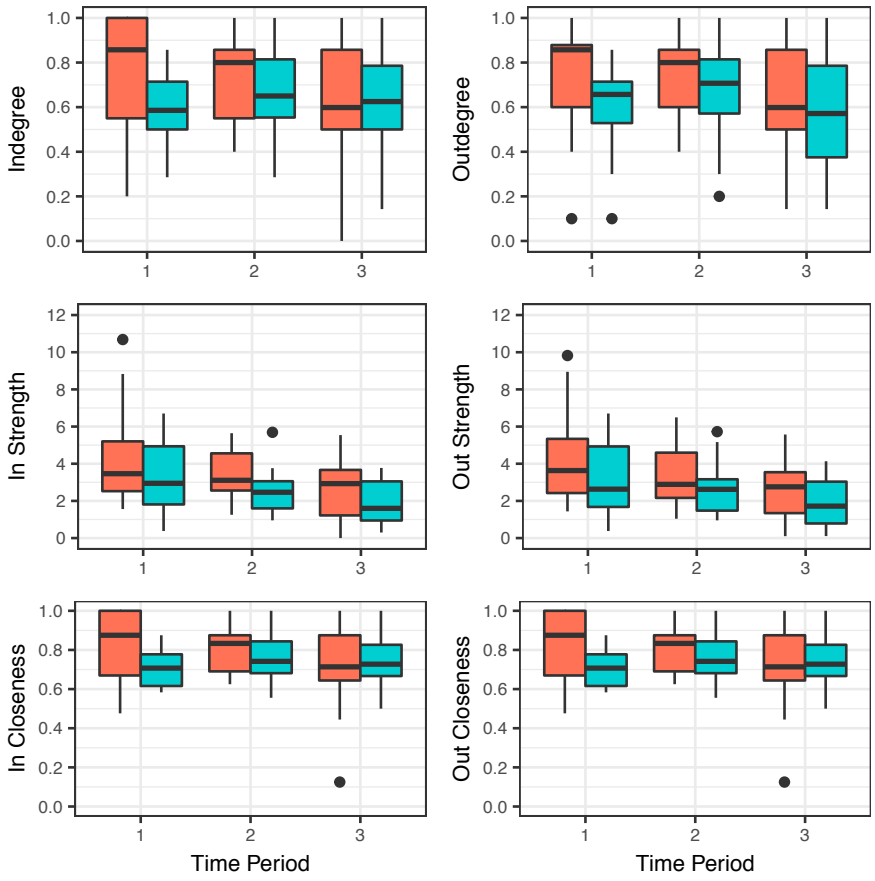


Figure 6. Relationship between sex (female; male) and centrality in the active affiliation network.

passive affiliation (closeness: β = -0.07, df = 23, p = 0.06) and for *received* active affiliation (in closeness: β = -0.48, df = 23, p = 0.042) (figures 6 and 7). Conversely, sex was a significant predictor of network-wide centrality for *performed* active affiliation, with males having lower network-wide centrality than females (out closeness: β = -0.07, df = 23, p < 0.01) (figure 6).

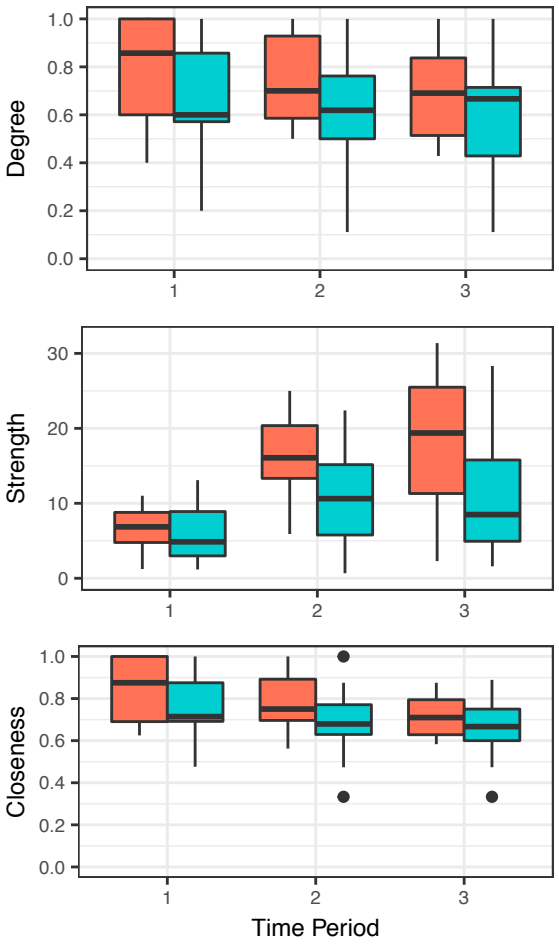


Figure 7. Relationship between sex (● female; ● male) and centrality in the passive affiliation network.

For LI agonism, sex was a significant and strong predictor of the number and network-wide centrality of *inward* relationships, with males receiving LI agonism from more group members than females (indegree: β = 0.40, df = 24, p < 0.001) and being more connected to the entire network through those interactions (in closeness: β = 0.29, df = 25, p < 0.001) (figure 8). Sex was also a significant predictor of the number of *outward* relationships with males performing LI agonism towards fewer group members than females (outdegree: β = -0.40, df = 24, p < 0.001). Dominance was a significant predictor of the network-wide centrality of *outward* relationships with dominant individuals being more connected to the entire network through LI agonistic interactions (out closeness: β = 0.29, df = 34, p < 0.01) (figure 8). Dominance was also a significant predictor of the strength of relationships, with dominant individuals having weaker relationships than subordinate individuals through *inward* LI agonistic interactions (in strength:

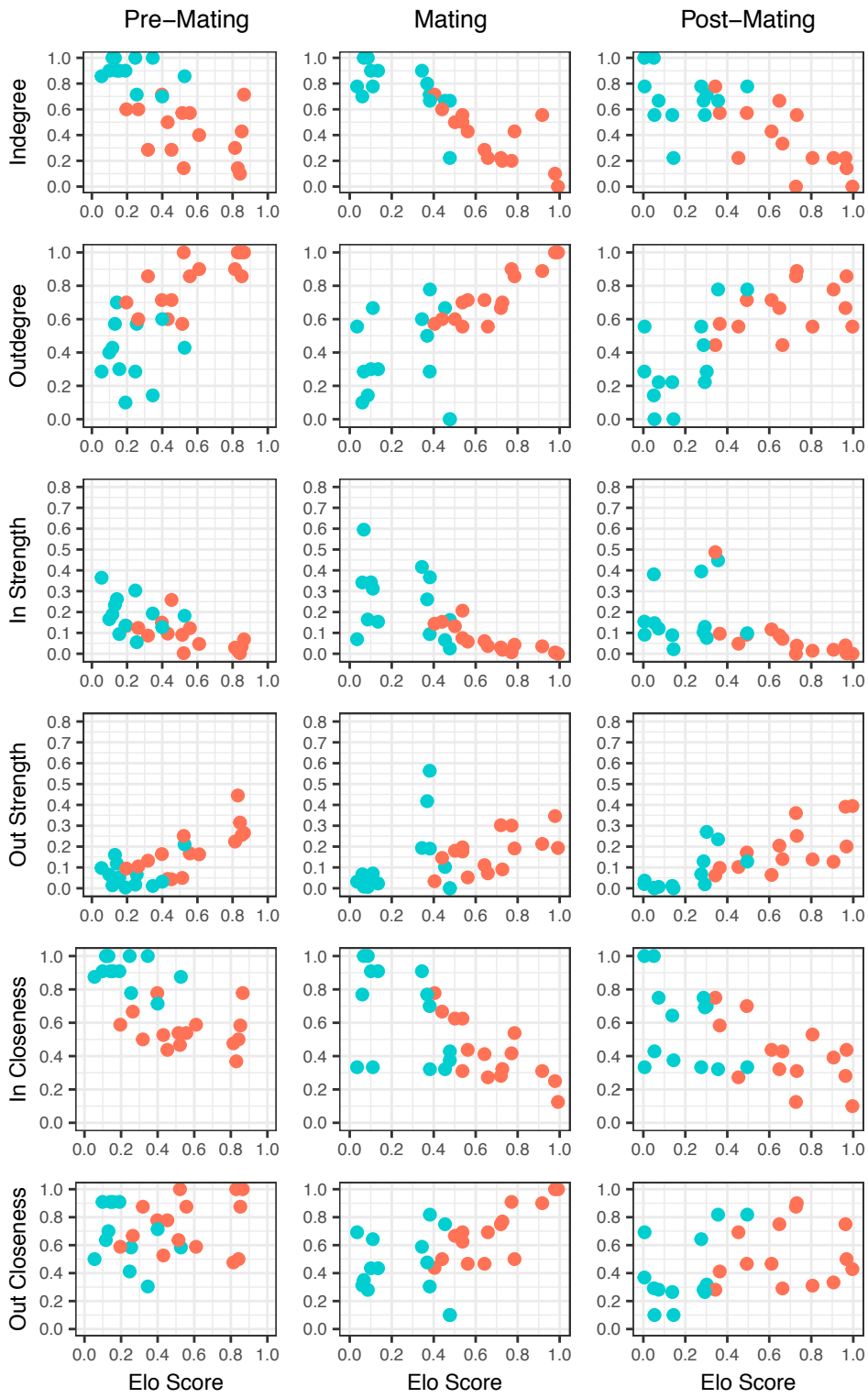


Figure 8. Relationship among sex (● female; ● male), dominance (Elo Scores) and centrality in the LI agonism network.

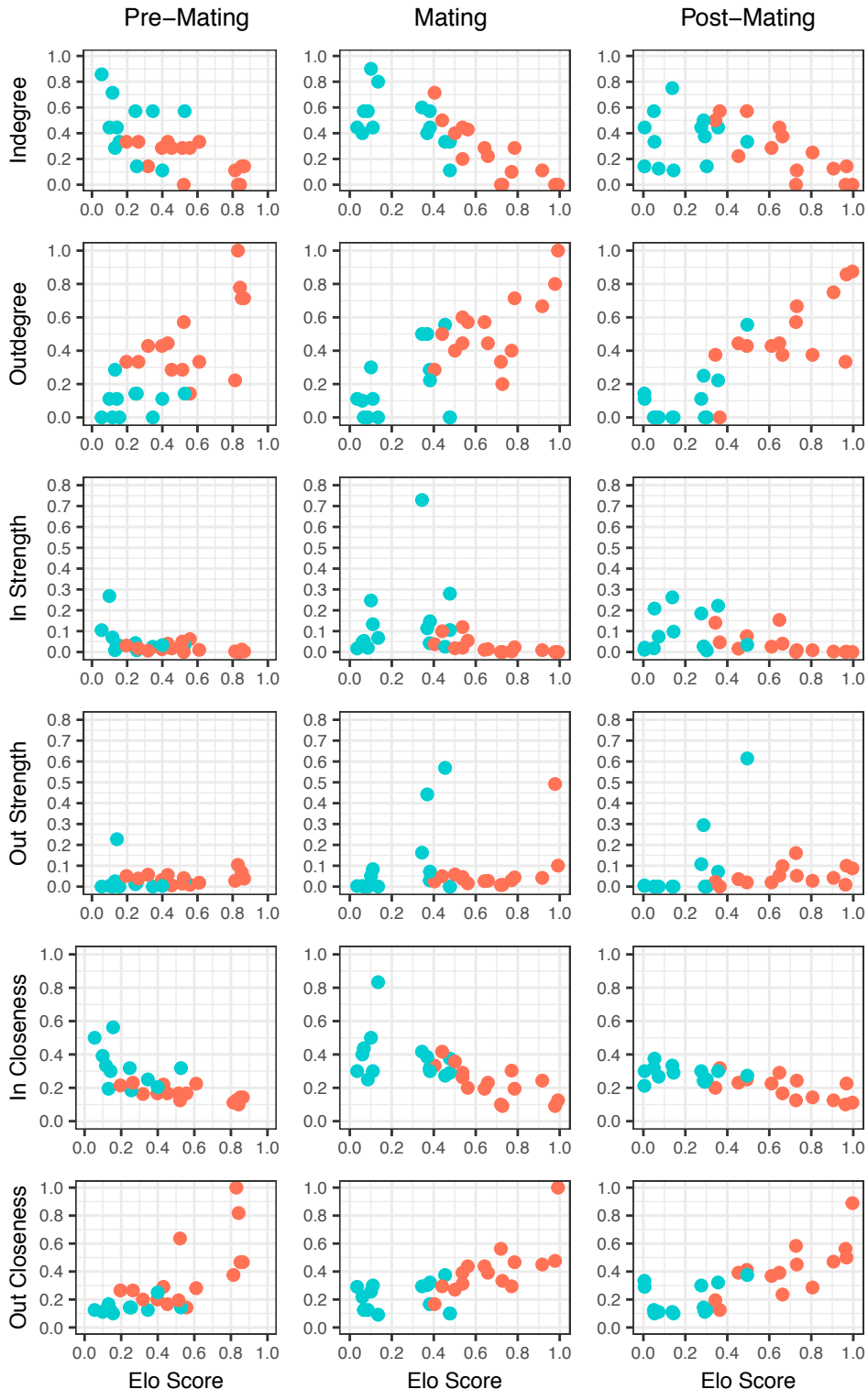


Figure 9. Relationship among sex (● female; ● male), dominance (Elo Scores) and centrality in the HI agonism network.

$\beta = -0.26$, $df = 37$, $p < 0.001$) (figures 8) and stronger relationships through *outward* LI agonistic interactions (out strength: $\beta = 0.23$, $df = 41$, $p < 0.001$) (figure 8).

For HI agonism, dominance was a significant and strong predictor of the number (but not strength) of relationships and of network-wide centrality (figure 9). Specifically, dominant individuals had fewer relationships and a lower network-wide through *inward* agonistic interactions (indegree: $\beta = -0.54$, 70, $p < 0.001$; in closeness: $\beta = -1.19$, $df = 70$, $p < 0.001$). Dominant individuals also had more relationships and a higher network-wide centrality through *outward* interactions (out degree: $\beta = 0.71$, $df = 40$, $p < 0.001$; out closeness: $\beta = 0.48$, $df = 37$, $p < 0.001$) (figure 9). The strength of HI agonistic relationships was not significantly influenced by sex or dominance.

Time Period and Group Membership: The three groups displayed different patterns of direct and network-wide centrality overall and through time for active affiliation. Group 2 had the most (but not strongest) relationships and highest network-wide centrality overall and Group 3 had the fewest relationships and generally lowest network-wide centrality overall (Indegree: $\beta_{group1\text{-}2} = 1.43$, $df = 22$, $p < 0.001$; $\beta_{group1\text{-}3} = -0.9$, $df = 22$ $p < 0.05$; $\beta_{group2\text{-}3} = -2.32$, $df = 23$, $p < 0.001$; outdegree: $\beta_{group1\text{-}2} = 1.07$, $df = 22$, $p < 0.01$; $\beta_{group1\text{-}3} = -0.89$, $df = 22$, $p < 0.01$; $\beta_{group2\text{-}3} = -1.95$, $df = 23$, $p < 0.001$; in closeness: $\beta_{group1\text{-}2} = 1.2$, $df = 22$, $p < 0.001$; $\beta_{group1\text{-}3} = -0.67$, $df = 22$, $p < 0.05$; $\beta_{group2\text{-}3} = -1.88$, $df = 23$, $p < 0.001$; out closeness: $\beta_{group1\text{-}2} = 0.19$, $df = 21$, $p < 0.001$; $\beta_{group2\text{-}3} = -0.22$, $df = 22$, $p < 0.001$) (figure 10). However, there was no significant difference between group 1 and 3 in their *outward* network-wide centrality ($\beta_{grp1\text{-}3} = -0.03$, $df = 22$, $p = 0.57$) (figure 10). There was also an interaction effect through time such that Group 2's strength and number of relationships, and network-wide centrality increased linearly while that of Group 1 declined (indegree:

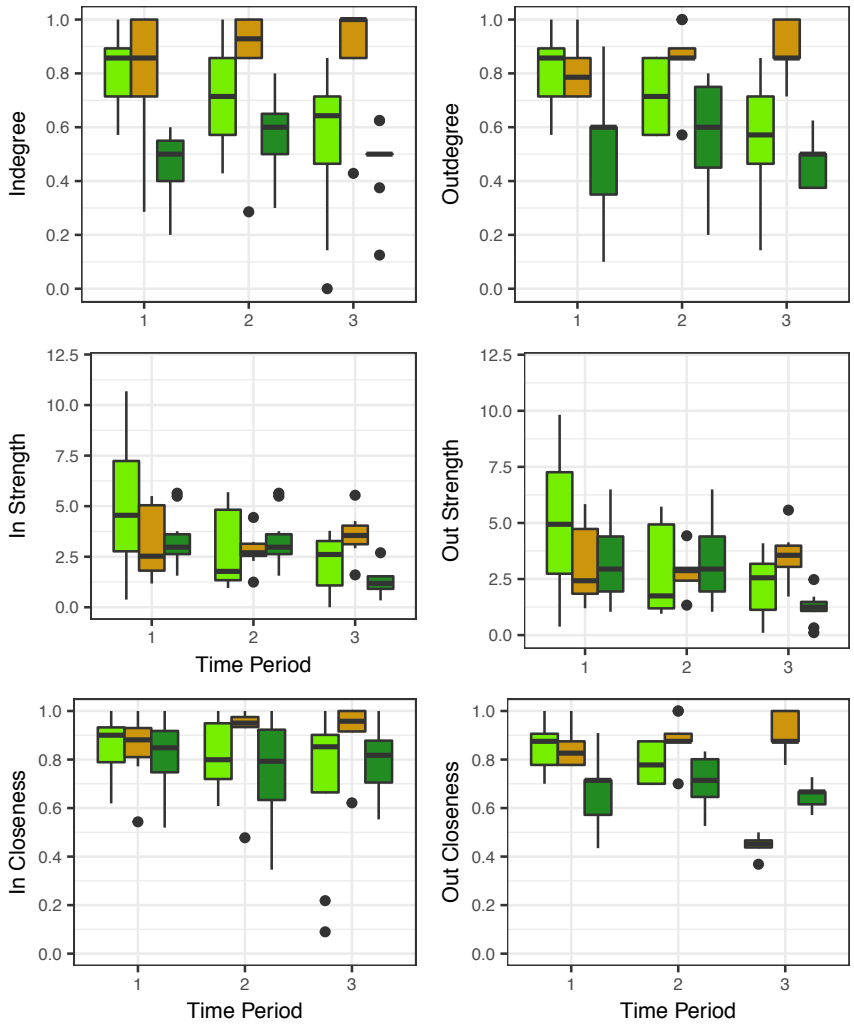


Figure 10. Average centrality scores for each group through time for active affiliation (Group 1 ; Group 2 ; Group 3).

βtime:group1-2 = -1.76, df = 44, p < 0.01; outdegree: βtime:group1-2 = -1.71, df = 44, p < 0.001; in closeness: βtime:group1-2 = -1.3, df = 44, p < 0.01; out closeness: βtime:group1-2 = -0.34, df = 43, p < 0.001; in strength: βtime:group1-2 = -2.36, df = 44, p < 0.001; out strength: βtime:group1-2 = -2.35, df = 44, p < 0.01). Similarly, Group 2's strength increased linearly while that of Group 3 declined (in strength: βtime:group2-3 = -1.75, df = 45, p < 0.001; out strength: βtime:group2-3 = -1.73, df = 45, p < 0.01) (figure 10).

Conversely, for passive affiliation, there was no difference among groups in their number or strength of relationships, or in their network-wide centrality. However, through time, the overall number of relationships (degree: $\beta = -0.22$, $df = 45$, $p < 0.001$) and network-wide centrality of nodes (closeness: $\beta = -0.07$, $df = 23$, $p < 0.001$) declined linearly and the strength of those interactions increased linearly (strength: $\beta = 10$, $df = 48$, $p < 0.001$) (figure 11).

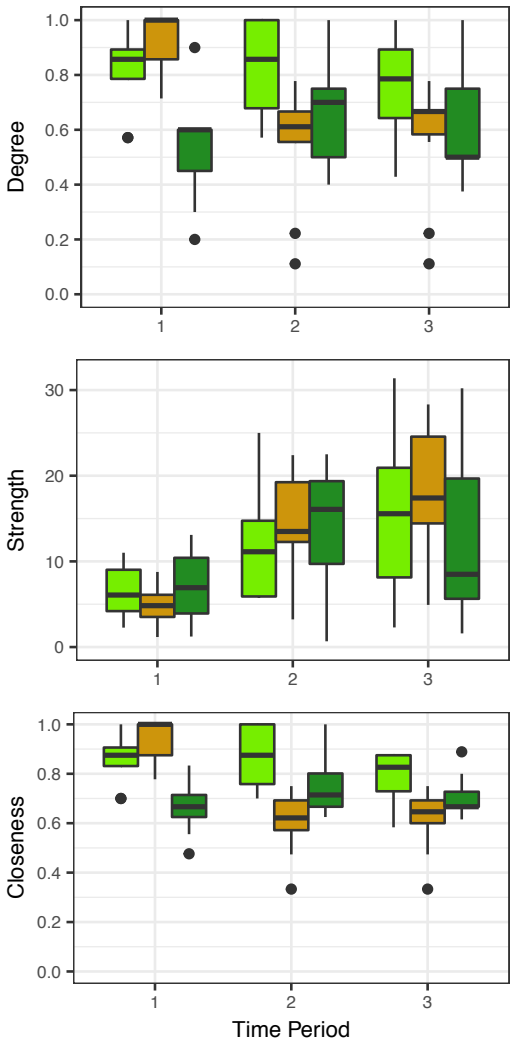


Figure 11. Average centrality scores for each group through time for passive affiliation (Group 1 ; Group 2 ; Group 3).

For HI and LI agonism, there was no significant difference overall among the three groups in their number and strength of relationships, and in their network-wide centrality (figures 12 and 13). HI agonism did not have interaction effects between group membership and time period (figure 12). Conversely LI agonism had several interaction effects between these two variables. Specifically, Group 2 declined in its number (but not strength) of relationships while Group 3 increased (degree: βtime:group2-3 = 0.24, $df = 46$, $p < 0.001$). Additionally, Group 2 declined in its network-wide centrality while Group 1 (βtime:group1-2 = 0.15, $df = 45$, $p < 0.001$) and 3 increased (βtime:grp3-3 = 0.23, $df = 46$, $p < 0.001$) (figure 13).

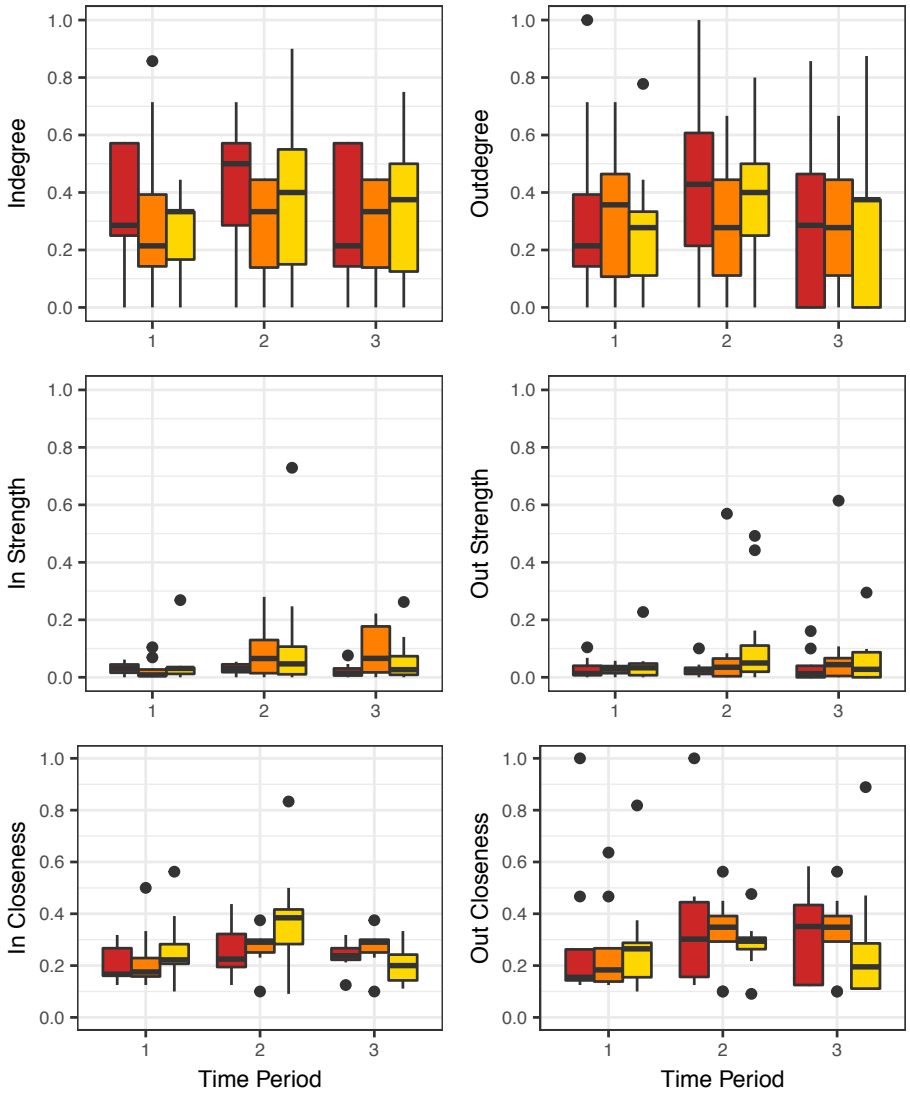


Figure 12. Average centrality scores for each group through time for HI agonism (Group 1 ; Group 2 ; Group 3).

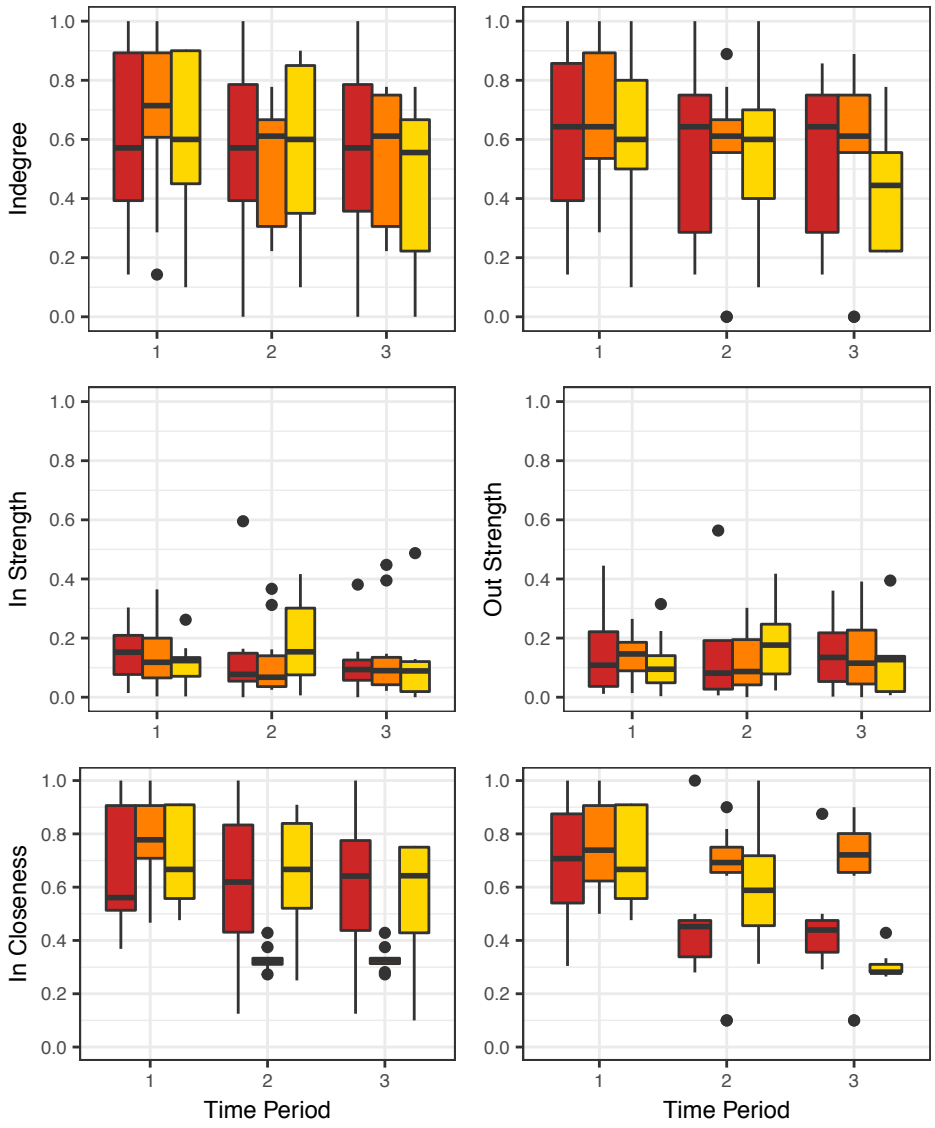


Figure 13. Average centrality scores for each group through time for LI agonism (Group 1 ; Group 2 ; Group 3).

Discussion:

In this study, I measured the network-level connectedness and node-level centrality of three wild ringtailed lemur groups over a six-month period. Although much is known about the social behavior of this species, there is no published work describing this behavior in terms of its emergent network properties, which is invaluable for establishing potential routes of pathogen

transmission. The connectedness of the networks and centrality of individual nodes differed across modes of interaction, group membership, dominance of individuals, and time period, each of which will be discussed below.

Modes of Interaction Through Time

Affiliation was much more common than agonism (as in Sbeglia et al. 2010) and individuals had more and stronger affiliative relationships than agonistic relationships. However, active and passive affiliation had different temporal patterns. Individuals in all groups increased in their strength of passive affiliative relationships and the overall time spent in passive affiliation through time but decreased in their number of relationships. Therefore, individuals invested more time in their relationships after mating even though they had fewer of them. For active affiliation, the overall percent of time in this mode of affiliation decreased through time but the groups behaved differently in terms of centrality, with Group 1 increasing in the number and strength of relationships, and in the network-wide centrality of nodes while the other groups declined or remained constant.

There is corroboration in the literature for the observation that affiliative behavior patterns change seasonally, and more specifically, that active affiliation was most common before and during the mating season and passive affiliation was most common after the mating season. For example, ringtailed lemur males and females displayed a gradual decline in grooming behavior from the mating season through lactation (Rasamimanana et al. 2006). Furthermore, studies in other primate species have shown heighted grooming behaviors occurring in the period before the annual mating season. For example, Taiwanese Macaques (*Macaca cyclopis*), which mate seasonally and polygamously, increased grooming behaviors just

prior to mating (Birky and Su 2005). Furthermore, relatively high male-male affiliation (resting in contact + grooming) has been observed for ringtailed lemurs in the post-mating period (either during gestation or lactation depending on the field site) (Gould 1996, Gabriel et al. 2014), which is consistent with my finding that passive affiliation was most common *after* mating.

Active and passive affiliation are often considered to be the most likely modes of interaction to cause pathogen transmission (e.g. MacIntosh et al. 2012). However, different modes of interaction may have different implications for pathogen movement depending on the characteristics of the pathogen and the nature of its transmission (Craft 2015). Specifically, active affiliation, which mostly includes grooming is likely to cause the transmission of oral-oral or fecal-oral transmitted pathogens such as *E. coli*, tuberculosis, and *Shigella*. Because ringtailed lemurs groom exclusively with their mouth, the role of grooming in fecal-oral pathogen transmission seems likely. Whereas passive affiliation, which mostly includes resting in contact, is likely to cause the transmission of pathogens that require long periods of physical contact or that reside on the skin or fur such as ectoparasites. However, the role of these modes of affiliation on transmission has never been formally tested. In fact, there are only four studies that correlate behavioral observations with genetically-confirmed pathogen sharing (Blyton et al. 2014, Springer et al. 2016, VanderWaal et al. 2014a, b) but they only report the role of proximity or generalized contact.

Sex and Dominance

I also assessed how an individual's traits influenced network patterns. The variation in the magnitude of centrality for agonism was partially explained by sex or dominance status. Specifically, females preformed more, and males received more LI agonism. Furthermore,

dominant individuals, who also tended to be females, had weaker *inward* LI agonistic relationships and were less central to the whole network through those interactions than subordinate individuals, who also tended to be males. Dominant individuals also had stronger *outward* LI agonistic relationships than subordinate individuals. HI agonism displayed similar patterns in terms of the number of agonistic relationships, but the strength of relationships was not influenced by dominance or sex. Therefore, dominant individuals had more but not stronger outward relationships. These findings suggest that HI and LI agonism functioned differently from each other such that individuals demonstrated dominance by performing frequent threats and receiving frequent submissive interactions (most common forms of LI agonism) from the same few individuals and performing relatively little outright aggression (i.e. HI agonism) but to a wider range of individuals. These findings also suggest both the presence of female dominance and the presence of transitivity (a.k.a. linearity) of dominance relationships among individuals. In a linear hierarchy, individual A is dominant to B, B is dominant to C, and A is also dominant to C (de Vries 1995). Importantly, these results are consistent with the complete female dominance and significant linearity reported in previous analyses that focus on these groups (Chapter 1) and that are expected for this species (Kappaler 1990).

Unlike the patterns for HI and LI agonism, sex and dominance did not significantly explain the variation in most measures of centrality for affiliation when applying the conservative critical p-value used in this study. Rather, males and females were generally equally central within the active and passive affiliation networks. The finding that sex was not a pervasive factor structuring affiliation networks is surprising because males spent significantly spent less time in total affiliation than females. However, it appears that they were still able to maintain relationships with comparable levels of centrality through their more limited

interactions. While it is possible that a larger sample size may result in the identification of significant differences between males and females, it is also possible that the lack of significance is a meaningful reflection of ringtailed lemur social structure. For example, successful group membership for ringtailed lemur males has been hypothesized to rely on strong relationships with females, which may be responsible for their maintenance of centrality in the core of the social group (Gould 1996). Furthermore, the role of dominance or sex in network centrality is inconsistent across studies (Rushmore et al. 2017) with some finding a clear relationship (e.g. MacIntosh et al. 2012, Japanese macaques; Rimbach et al, 2015; brown spider monkeys; Romano et al. 2016, Japanese macaque subspecies; Rushmore et al. 2013, chimpanzees) and some finding a weak or non-existent relationship (e.g. Carne et al. 2003, chimpanzees; Romano et al. 2016, Japanese macaque subspecies).

Sex also appeared to play a role in the structuring of membership in particular modules within the network. Although the membership in modules was fluid and inconsistent through time, individuals of both sexes were spread out across most modules, with the resulting pattern that nearly every module contained both males and females. Furthermore, in the pre-mating period, the one or two most dominant males were in a module with the lowest male to female sex ratio and were with females with the highest dominance scores. Females in this species are known to form partner preferences during and near the mating season and high ranking males often have first mating priority (Koyama 1988, Sauther 1991).

Group Membership and Microhabitat

Another finding in this study is that node-level centrality and network-level connectedness were associated with microhabitat. Specifically, the group living in degraded

habitat (Group 2) had the lowest modularity for active affiliation, highest graph density for active and passive affiliation, and the most relationships and highest network-wide centrality overall and through time for active affiliation. These network characteristics all functioned to increase the connectedness of individuals and could therefore cause heightened pathogen transmission along these routes of affiliation. The territory of Group 2 was almost completely outside of the reserve in degraded dry forest, which has less canopy cover, sparse vegetation (Gemmill and Gould 2008), and a lower availability of preferred food items (Gould 1996; Sauther 1992, 1993). There is evidence in the literature that food availability (e.g. Karczmarski et al. 2005, Boesch 1991) and predation risk (Boesch 1991, Cluttonbrock and Harvey 1977, Crook and Gartlan 1966, Jarman 1974, Van Schaik 1983) can influence the characteristics of social organization, specifically group size, group membership, and the pattern of fission-fusion events. However, these analyses often do not assess the influence of these habitat conditions on more complex elements of social dynamics (Kelley et al. 2011), such as the direct and indirect centrality of group members and the modularity and density of their networks. Recent work has demonstrated that there is a link between habitat conditions and an animal's level of spatial (e.g. Jacoby et al. 2012) and social connectedness (e.g. Kelley et al. 2011, Ansmann et al. 2012). For example, guppies in high predation-risk populations had more and stronger associations and more stable social groupings through time than those from low-risk populations (Kelley et al. 2011). Furthermore, there is evidence in humans that high social connectedness is associated with positive health effects such as reduced anxiety (Van Zilk and Van Zilk 2015) and risk of cardiac malfunction (Kim et al. 2016). These findings are consistent with a large body of existing work that shows a relationship between affiliation and certain beneficial neurological and hormonal effects (Bartz and Hollander 2006). These studies are relevant because animals living in

disturbed habitat have higher levels of stress hormones than those living in pristine habitat (Jaimez et al. 2011, Martínez-Mota et al. 2007), which could negatively impact health, fertility, and survival (Martínez-Mota et al. 2007). In fact, Group 2 has the lowest rank instability of the three groups (Chapter 1), which has been found to be associated with stress (Sapolsky 1983, 1992). Therefore, it is possible that high network connectedness for animals living in disturbed habitats functions to lower predation risk and decrease levels of stress often associated with these environments. However, the level of habitat disturbance (Wright et al. 2009), fragmentation (Gillespie and Chapman 2006) and local biodiversity (Keesing et al. 2010) can influence the risk of disease and it is known that lemurs living in degraded habitats have heightened endoparasite richness and infection intensities (Bublitz et al. 2015, Goldberg et al. 2007, Rwego et al. 2008, Wright et al. 2009). The results in this study suggest that this heightened risk of disease, which is often expected due to increased exposure to humans and livestock (e.g. Bublitz et al. 2015, Goldberg et al. 2007, Rwego et al. 2008, Wright et al. 2009), may be amplified by an increase in network connectedness. Therefore, it is possible that these findings represent a tradeoff between the modulation of predation risk and stress through an increase in affiliative connectedness, and the infection risk it may facilitate.

Conclusion

I measured the direct and indirect connectedness of each social group through agonistic and affiliative interactions using Social Network Analysis. Active and passive affiliation had different temporal patterns with individuals decreasing the time spent in active affiliation and increasing the time spent in passive affiliation from the pre- to the post-mating period, suggesting that these time periods may be differentially important for the movement of specific

pathogens depending on its characteristics and the nature of its transmission (Craft 2015). Further, the connectedness of a social group was associated with its microhabitat, such that the group living in degraded habitat displayed the highest network connectedness and node centrality overall and through time for active affiliation, which could cause heightened pathogen transmission. Animals living in degraded habitat are often expected to have a higher infection rate because of the increased presence of pathogens from humans and livestock (e.g. Bublitz et al. 2015, Goldberg et al. 2007, Rwego et al. 2008, Wright et al. 2009), but the results of the present study suggest these effects may be amplified by an increase in network connectedness. Because this study was conducted on only three social groups of one species, more behavioral work on a diversity of social species is necessary to further test these hypotheses. In addition, the simultaneous collection of pathogens and analysis of pathogen genetics, though rarely done (Archie et al. 2009), would greatly improve the ability to test the movement of infectious organisms along social routes (Craft 2015). Madagascar faces extreme habitat disruption and degradation (Bublitz et al. 2015) and it is estimated that 90% of the native forest has either been cleared or fragmented (Harper et al. 2007). As a result, many of Madagascar's abundant endemic flora and fauna are now endangered (IUCN 2017) and living in or near degraded habitat. The potential impact of habitat degradation on the connectedness of individuals within a network may undermine the functioning of social networks that have evolved under certain ecological conditions that no longer exist and in doing so, may amplify the pathogen exposure faced by social animals above and beyond what we have previously anticipated.

Literature Cited

Albers, P. C. H. and de Vries. H. 2001. Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour* 61: 489–95.

Alexander, R. D. 1974. The evolution of social behaviour. *Annual Review of Ecological Systems* 5: 325–383.

Altmann, J. 2009. Observational study of behavior: Sampling methods. *Behaviour* 49: 227–67.

Altizer, S. Dobson, A., Hosseini, P., Hudson, P. Pascual, M. and Rohani, P. 2006. Seasonality and the dynamics of infectious diseases. *Ecology Letters* 9: 467–84.

Ansmann, I. C., Parra, G. J. Chilvers, B. L. and Lanyon, J. M. 2012. Dolphins restructure social system after reduction of commercial fisheries. *Animal Behaviour* 84: 575–81.

Archie, E. A. and Theis, K. R. 2011. Animal behaviour meets microbial ecology. *Animal Behavior* 82: 425–436.

Bartz, J. A. and Hollander, E. 2006. The neuroscience of affiliation: Forging links between basic and clinical research on neuropeptides and social behavior. *Hormones & Behavior* 50: 518–528.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B. Grothendieck, G. and Green, P. 2016. Linear mixed-effects models using 'Eigen' and S4. V. 1.1–12.

Birky, W. A. and Su, H. H. 2005. Temporal variation in adult affiliation patterns in wild Taiwanese macaques (*Macaca cyclopis*). *International Journal of Primatology* 26: 1251–66.

Blonder, B. and Dornhaus, A. 2011. Time-ordered networks reveal limitations to information flow in ant colonies. *PLoS ONE* 6: 1–9.

Blyton, M. D. J., Banks, S. C., Peakall, R., Lindenmayer, D. B. and Gordo, D. M. 2014. Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecology Letters* 17: 970–978.

Boesch, C. 1991. The effects of leopard predation on grouping patterns in forest chimpanzees. *Behaviour* 117: 220–41.

Boyland, N. K., Mlynski, D. T., Richard, J., Brent, J. N. and Croft, D. P. 2016. The social network structure of a dynamic group of dairy cows: From individual to group level patterns. *Applied Animal Behaviour Science* 174: 1–10.

Brockman, D. K., Godfrey, L. R., Dollar, L. J., and Ratsirarson, J. 2008. Evidence of invasive *Felis silvestris* predation on *Propithecus verreauxi* at Beza Mahafaly Special Reserve, Madagascar. *International Journal of Primatology* 29: 135–52.

Bublitz, B. C., Wright, P. C. Rasambainarivo, F. T., Arrigo-Nelson, S. J., Bodager, J. R., Gillespie, T. R. 2015. Pathogenic enterobacteria in lemurs associated with anthropogenic disturbance. *American Journal of Primatology* 77: 330–37.

Carne, C., Semple, S., Morrogh-Bernard, H., Zuberbühler, K. and Lehmann, J. 2013. Predicting the vulnerability of great apes to disease: the role of superspreaders and their potential vaccination. *PLoS One* 8: e84642.

Cluttonbrock, T. H. and Harvey, P. H. 1977. Primate ecology and social organization. *Journal of Zoology* 183: 1–39.

Corner, L. A. L., Pfeiffer, D. U. and Morris, R. S. 2003. Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (*Trichosurus vulpecula*). *Preventative Veterinary Medicine* 59: 147–167.

Craft, M. E. 2015. Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370: 1–12.

Crook, J. H. and Gartlan, J. S. 1966. Evolution of primate societies. *Nature* 210: 1200–1203.

Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695. http://igraph.sf.net.

Ezenwa, V. O. 2004. Host Social Behavior and Parasitic Infection: A Multifactorial Approach. *Behavioral Ecology* 15: 446–54.

Foster, J. A., Krone, S. M., Forney, L. J. 2008. Application of ecological network theory to the human microbiome. *Interdisciplinary Perspectives on Infectious Diseases* 2008: 1–6.

Freeland, W. J. 1976. Pathogens and the evolution of primate sociality. *Biotropica* 8: 12–24.

Freeland, W. J. 1979. Primate social-groups as biological islands. *Ecology* 60: 719–728.

Friant, S., Ziegler, T. E. and Goldberg, T. L. 2016. Primate reinfection with gastrointestinal Parasites: Behavioural and physiological predictors of parasite acquisition. *Animal Behaviour* 117: 105–13.

Fruchterman, T. M. J. and Reingold, E. M. 1991. Graph drawing by force-directed placement software. *Practice and Experience* 21: 1129–1164.

Gabriel, D. N., Gould, L. and Kelley, E. A. 2014. Seasonal Patterns of Male Affiliation in Ring-Tailed Lemurs (*Lemur catta*) in Diverse Habitats across Southern Madagascar. *Behaviour* 151: 935–61.

Gemmill, A. and Gould, L. 2008. Microhabitat variation and its effects on dietary composition and intragroup feeding interactions between adult female *Lemur catta* during the dry season at Beza Mahafaly Special Reserve, Southwestern Madagascar. *International Journal of Primatology* 29: 1511–1533.

Gillespie, T. R. and Chapman, C. A. 2006. Prediction of parasite infection dynamics in primate metapopulations based on attributes of forest fragmentation. *Conservation Biology* 20: 441–448.

Goldberg, T. L., Gillespie, T. R., Rewgo, I. B., Wheeler, E., Estoff, E. L., and Chapman, C. A. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation* 135: 511–517.

Gould, L. 1996. Male-female affiliative relationships in naturally occurring ringtailed lemurs (*Lemur catta*) at the Beza-Mahafaly Reserve, Madagascar. *American Journal of Primatology* 39: 63–78.

Griffin, R. H. and Nunn, C. L. 2012. Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology* 26: 779–800.

Hamede, R. K., Bashford, J., McCallum, H. and Jones, M. 2009. Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: Using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters* 12: 1147–57.

Harper, G. J, Steininger, M. K., Tucker, C. J., Juhn, D. and Hawkins, F. 2007. Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation* 34: 325–333.

Hess, G. 1996. Disease in metapopulation models: Implications for conservation. *Ecology* 77: 1617–1632.

Hinde, R. A. 1976. Interactions, relationships and social structure. Man 11: 1–17.

Hothorn, T., Bretz, F. Westfall, P. Heiberger. R. M. Schuetzenmeister, A. and Scheibe, S. 2016. Simultaneous inference in general parametric models (R package, version 1.4–6), http://multcomp.R-forge.R-project.org.

IUCN 2016. The IUCN Red List of Threatened Species. Version 2016-3. http://www.iucnredlist.org. Downloaded on 22 March 2017.

Jacoby, D. M. P., Brooks, E. J., Croft, D. P. and Sims, D. W. 2012. Developing a deeper understanding of animal movements and spatial dynamics through novel application of network analyses. *Methods in Ecology and Evolution* 3: 574–583.

Jaimez, N. A., Bribiescas, R. G., Aronsen, G. P., Anestis, S. A. and Watts, D. P. 2012. Urinary cortisol levels of gray-cheeked mangabeys are higher in disturbed compared to undisturbed forest areas in Kibale National Park, Uganda. *Animal Conservation* 15: 242–247.

Jarman, P. J. 1974. Social organization of antelope in relation to their ecology. *Behaviour* 48: 215–267.

Jeanson, R. 2012. Long-term dynamics in proximity networks in ants. *Animal Behaviour* 83: 915–923.

Kappeler, P. M. 1990. Female dominance in *Lemur catta*: more than just female feeding priority? *Folia Primatologica* 55: 92–95.

Karczmarski, L., Wursig, B., Gailey, G., Larson, K. W., Vanderlip, C. 2005. Spinner dolphins in a remote Hawaiian atoll: Social grouping and population structure. *Behavioral Ecology* 16: 675–685.

Kelley, J. L., Morrell, L. J., Inskip, C., Krause, J. and Croft, D. P. 2011. Predation risk shapes social networks in fission-fusion populations. *PLoS ONE*: e24280. doi:10.1371/journal.pone.0024280.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T., Ostfeld, R. S. 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647-652.

Kim, D. A., Benjamin, E. J., Fowler, J. H. and Christakies, N. A. 2016. Social connectedness is associated with fibrinogen level in a human social network. *Proceedings of the Royal Society B* 83: 1–7.

Koyama, N. 1988. Mating behavior of ring-tailed lemurs (*Lemur catta*) at Berenty, Madagascar. *Primates* 29: 163–175.

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. 2016. lmerTest: Tests in linear mixed effects models (R package, version 2.0–33), https://cran.r-project.org/web/packages/lmerTest/index.html.

Lehmann, J. and Ross, C. 2011. Baboon (*Papio anubis*) social complexity–A network approach. *American Journal of Primatology* 73: 775–89.

MacIntosh, A. J. J., Jacobs, A., Garcia, C. Shimizu, K., Mouri, K. Huffman, M. A. and Hernandez, A. D. 2012. Monkeys in the middle: Parasite transmission through the social network of a wild primate. *PLoS ONE* 7: 15–21.

Martínez-Mota, R., Valdespino, C., Sánchez-Ramos, M. A. and Serio-Silva, J. C. 2007. Effects of forest fragmentation on the physiological stress response of black howler monkeys. *Animal Conservation* 10: 374–79.

Macedonia, J. M. 1993. Vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatologica* 61: 186–217.

Møller, A. P., Merino, S., Brown, C. R. and Robertson, R. J. 2001. Immune defense and host sociality: A comparative study of swallows and martins. *American Naturalist* 158: 136–45.

Neumann, C. J. Duboscq, C. Dubuc, A. Ginting, A. M. Irwan, M. Agil, A. Widdig, and A. Engelhardt. 2011. Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour* 82: 911–921.

Neumann C. and L. Kulik 2014. EloRating: Animal dominance hierarchies by Elo rating (R package, version 0.43), https://CRAN.R-project.org/package=EloRating.

Rasamimanana, H., Andrianome, V. N., Rambeloarivony, H. and Pasquet, P. 2006. Male and female ringtailed lemurs' energetic strategy does not explain female dominance. In: Ringtailed Lemur Biology (Ed. by A. Jolly, R. W. Sussman, N. Koyama, H. Rasamimanana), pp. 271–295. Chicago: Springer.

Rewgo, I. B., Isabirye-Basuta, G., Gillespie, T. R., Goldberg, T. 2008. Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600–1607.

Rimbach, R., Bisanzio, B., Galvis, N. Link, A, Di Fiore, A. and Gillespie T. R. 2015. Brown spider monkeys (*Ateles hybridus*): a model for differentiating the role of social networks and physical contact on parasite transmission dynamics. *Philosophical Transactions of the Royal Society B* 370: 20140110.

Romano, V., Duboscq, J., Sarabian, C., Thomas, E., Sueur, C. and MacIntosh, A, J. J. 2016. Modeling infection transmission in primate networks to predict centrality-based risk. *American Journal of Primatology* 78: 767–79.

Rushmore, J., Bisanzio, D. and Gillespie, T. R. 2017. Making new connections: Insights from primate–parasite networks. *Trends in Parasitology* 1615: 1–14.

Rushmore, J., Caillaud, D., Matamba, L., Stumpf, R. M., Borgatti, S. P. and Altizer, A. 2013. Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. *Journal of Animal Ecology* 82: 976–86.

Sapolsky, R. M. 1983. Endocrine aspects of social instability in the olive baboon (*Papio anubis*). *American Journal of Primatology* 5: 365–379.

Sapolsky, R. M. 1992. Cortisol concentrations and the social significance of rank instability among wild baboons. *Psychoneuroendocrinology* 17: 701–702.

Sauther, M. L. 1991. Reproductive behavior of free-ranging *Lemur catta* at Beza Mahafaly Special Reserve, Madagascar. *American Journal of Physical Anthropology* 84: 463–477.

Sauther, M. L. 1992. Effect of reproductive state, social rank and group size on resource use among free-ranging ringtailed lemurs (*Lemur catta*) of Madagascar. Ph.D. dissertation, Washington University, St. Louis, MO.

Sauther, M. L. 1993. Resource competition in wild populations of ring-tailed lemurs (*Lemur catta*): implications for female dominance. In: Lemur Social Systems and their Ecological Basis (Ed. by P. M. Kappeler and J. U. Ganzhorn), pp. 135–152. New York: Plenum.

Sbeglia, G. C., Tang-Martinez, Z. and Sussman, R. W. 2010. Effects of food, proximity, and kinship on social behavior in ringtailed lemurs. *American Journal of Primatology* 72: 981–91.

van Schaik, C. P. 1983. Why are diurnal primates living in groups? *Behaviour* 87: 120–144.

van Schaik, J. and Kerth, G. 2017. Host social organization and mating system shape parasite transmission opportunities in three European bat species. *Parasitology Research*. 116: 589–599.

Sih, A., Hanser, S. F. and McHugh, K. A. 2009. Social network theory: new insights and issues for behavioral ecologists. *Behavioral Ecology and Sociobiology* 63: 975– 988.

Springer, A., Mellmann, A., Fichtel, C. and Kappeler, P. M. 2016. Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecology* 16: 1-6.

Sussman, R. W. and Rakotozafy, A. 1994. Plant diversity and structural analysis of a tropical dry forest in southwestern Madagascar. *Biotropica* 26: 241–254.

Van Zalk, N. and Van Zalk, M. 2015. The importance of perceived care and connectedness with friends and parents for adolescent social anxiety. *Journal of Personality* 83: 346–60.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014a. Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *Journal of Animal Ecology* 83: 406–414.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014b. Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation* 169: 113–146.

de Vries H. 1995. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour* 50: 1375-1389.

Wilson, K., Knell, R., Boots, M. and Koch-Osborne, J. 2003. Group living and investment in immune defense: an interspecific analysis. *Journal of Animal Ecology* 72: 133–143.

Wright, P. C., Arrigo-Nelson, S. J., Hogg, K. L. Bannon, B., Morelli, T. L., Wyatt, J. Harivelo, A. L. and Ratelolahy, F. 2009. Habitat disturbance and seasonal fluctuations of lemur parasites in the rain forest of Ranomafana National Park, Madagascar. In: Primate Parasite Ecology: The Dynamics and Study of Host-Parasite Relationships (Ed. by M.A. Huffman and C.A. Chapman), pp. 311-330. Cambridge: Cambridge University Press.

Zhu, Z. Ochoa, A. J. and Katzgraber, H. G. 2015. Efficient cluster algorithm for spin glasses in any space dimension. *Physical Review Letters* 115: 077201.

Zohdy, S., Kemp, A. D., Durden, L. A., Wright, P. C. and Jernvall, J. 2012. Mapping the social network: Tracking lice in a wild primate (*Microcebus rufus*) population to infer social contacts and vector potential. *BMC Ecology* 12: 4.

One-locus, culture-independent differentiation of *E. coli* haplotype

Chapter 3

Abstract: The differentiation of bacterial haplotypes traditionally involves culturing bacteria on a nutrient plate and sequencing the colonies at multiple loci. However, there are billions of *E. coli* cells comprising many haplotypes within an individual, making this approach time consuming and imprecise. I present a one-locus approach to haplotype differentiation in *E. coli* that will bypass the culturing step and allow for high throughput sequencing of an entire fecal sample at one locus. This novel method is conceptually similar to that used to differentiate bacterial genera in the microbiome and is the first to differentiate haplotypes of *E. coli* (and perhaps any bacteria) using one locus. To test the viability of the one-locus method, I compared the differentiation ability of a single highly-variable locus with that of CH-typing, a well-established multi-locus method. I collected fecal samples from healthy adult ringtailed lemurs living at the Duke Lemur Center in Durham, NC and Sanger sequenced the cultured *E. coli* colonies at the two CH loci, FimH and FumC. I then designed and evaluated new primers within these highly variable loci to identify a region that differentiated a sufficient diversity of *E. coli* haplotypes. Taking sequencing error into account, one primer differentiated 91-172% of the haplotype diversity identified with CH-typing and was thus highly effective at differentiating known *E. coli* haplotype diversity. Although not usually pathogenic, *E. coli* is valuable as a model "pathogenic" organism for social transmission studies because its ubiquity in mammals and high haplotype diversity within a single host allows the inference of fine-scale patterns of transmission among *all* individuals. This approach can refine our understanding of the movement of bacteria within and among populations and can be a powerful method to answer questions about disease ecology in a faster and more fine-grained manner than is currently possible.

Introduction:

Infections by pathogenic organisms frequently contain multiple haplotypes of the same species (Bachmann et al. 2015, Eyre et al. 2013, Taylor *et al.* 1995), which greatly influences the approach and outcome of treatment (Cohen *et al.* 2012). However, assessing the diversity of mixed-haplotype infections within a host is challenging, especially when using traditional culture-based approaches to haplotype identification (Bachmann et al. 2015, Dias et al. 2010). Culture-based methods differentiate haplotypes by growing pathogens on culture plates and identifying the genetic similarities and differences among isolates across multiple loci. This multi-locus approach makes it necessary to know how the variation at each locus is linked to the variation at the other loci by virtue of originating in the same isolate. Therefore, culture-based methods are limited by cell culturing, which is time-, labor-, and resource-intensive (Bachmann

et al. 2015). As a result, the identification of haplotype diversity within a mixed-haplotype sample, particularly of rare or un-culturable variants, is challenging.

In the late 1990s and early 2000s, scientists reported the first uses of a culture-independent approach to the differentiation of bacterial organisms (e.g. Heilig et al. 2002, Suau et al. 1999). This innovation was a remarkable step in the assessment of bacterial diversity because it allowed the identification of rare and difficult to culture strains and species that would have been unlikely to be discovered with traditional culture-based approaches (Rajilić-Stojanović et al. 2007, Suau et al. 1999). Furthermore, the time- and resource-limiting step of growing samples on culture could be completely bypassed, allowing a more accurate identification of the billions of bacterial cells and unknown numbers of strains that often exist within a sample (Bachmann et al. 2015, Dias et al. 2010). These early studies overwhelmingly utilized the variation within the small RNA subunit of the 16S gene for differentiation (Hill et al. 2002). Since then, high throughput sequencing at this particular locus has revolutionized our ability to rapidly and accurately identify bacterial diversity (e.g. Cho et al. 2012, Ezenwa et al. 2012, Odell and Flavell 2016, Sampson et al. 2015). However, because 16S methodology can only differentiate bacteria at the level of the species or genus, questions that require differentiation at the level of the haplotype, such as those related to pathogen transmission disease treatment, cannot be readily answered using this approach (Heilig et al. 2002). Questions about pathogen transmission in particular are most accurately addressed by haplotype-level differentiation because individuals must share the same or a related haplotype of a particular organism for transmission to be deduced (Eyre et al. 2013). Therefore, although one-locus culture-independent methods such as 16S are powerful at detecting diversity and efficient in terms of time and resources, they cannot currently be used to *directly* assess incidences of transmission. Authors

that have used this approach for transmission studies could identify only that individuals shared or did not share the same species or genera of a pathogen or bacterium, but could only speculate that the sharing was due to transmission as opposed to independent infections (e.g. Tung et al. 2015). Recently, authors have proposed full genome culture-independent methods to differentiate mixed-haplotype infections (e.g. Bachmann et al. 2015, *Chlamydia pecorum*; Eyre et al. 2013, *Clostridium difficile*). Although whole genome high-throughput sequencing is powerful for detecting haplotype diversity within samples, they require high variant density within a genome (Eyre et al. 2013) and extremely high coverage (Bachmann et al. 2015) to accurately reconstruct haplotypes, which reduces the number of samples that can be sequenced at one time and limits the ability to re-construct minor haplotypes within a sample (e.g. Bachmann et al. 2015).

At present, studies aiming to assess bacterial transmission through the detection of shared or closely-related haplotypes have primarily used traditional culture-based methods (e.g. Blyton et al. 2014, Caugant et al. 1984, Dubois et al. 2010, Georghiou et al. 1994, Goldberg et al. 2007, Michalak et al. 1998, Montoro et al. 1998, Rewgo et al. 2008, Springer et al. 2016, VanderWaal 2014a, b), although some studies have begun to use a whole genome sequencing approach (e.g. Eyre et al. 2013). While many of these studies convincingly indicate the occurrence of transmission, the haplotype diversity identified either required an unsustainable amount of time and resources or was too limited to allow inferences as to how the characteristics of environments and social interactions contribute to fine-scale dynamics of transmission. Therefore, a culture-independent, one-locus approach to species-specific bacterial identification would be useful (i.e. a species-specific version of 16S).

In this paper, I describe a culture-independent, one-locus approach to haplotype differentiation in *E. coli* that bypasses the culturing step and is appropriate for high throughput

sequencing. *E. coli* is highly valuable as a model "pathogenic" organism for social transmission studies. Although not usually pathogenic, its ubiquity in mammals and high haplotype diversity within a single host (Hartl and Dykhuizen 1984) allows the inference of fine-scale patterns of transmission among *all* individuals, instead of just those infected by an occasional parasite. *E. coli* is transmitted by oral ingestion, either through a fecal-oral route or through the consumption of contaminated water, food or soil. Further, because individuals are known to carry one or two major haplotypes of *E. coli* that is at high frequency for long periods of time and is distinct from the haplotype(s) carried by other individuals (Caugant et al. 1984, Hartl and Dykhuizen 1984), these major haplotypes can potentially act as bacterial identifiers for individuals, facilitating the identification of the directionality of transmission events. Most other haplotypes are minor or transient and tend to be present for only a short period of time and a low frequency (Hartl and Dykhuizen 1984). Furthermore, *E. coli* diversity is well-known, which allows the evaluation of the effectiveness of this new method in terms of its ability differentiate known haplotypes.

Scientists currently differentiate *E. coli* haplotypes through multi-locus, culture-based methods such as multi-locus sequence typing (MLST), which uses 5-7 loci, and CH-typing, which uses two loci (Weissman et al. 2012). These multi-locus approaches require knowledge of the linkage relationship of each variant across each target locus so that the variation can be confirmed to originate within a single cell. An implication of this approach is that each *E. coli* cell must be cultured before sequencing, thus limiting the diversity that can be identified (Bachmann et al. 2015, Dias et al. 2010, Eyre et al. 2013). As far as I know, the methodology described here is the first to use a one-locus, culture-independent, high-throughput approach to differentiate bacterial haplotypes. In this paper, I report the specifics of this novel method as well as evidence that it captures substantial *E. coli* diversity.

Methods

Collecting and Sanger Sequencing Isolates

I collected two fecal samples from each of 16 adult ringtailed lemurs in three semi-free ranging social groups at Duke Lemur Center in August 2012. I stored the samples in glycerol and salts and froze them at -80°C until analysis. I cultured bacteria from each fecal sample on EMB agar and mailed the resulting isolates to the lab of Dr. Evgeni Sokurenko in the Department of Microbiology at the University of Washington Sanger. Dr. Sokurenko's lab sequenced 414 isolates with the physical characteristics of *E. coli* (19-29 isolates per lemur) at the CH-loci, FumC and FimH, using standard primers (Weissman et al. 2012). Although CH-typing uses fewer loci than MLST, it has actually been found to differentiate more haplotypes (Weissman et al. 2012). FumC is a housekeeping gene that encodes for fumerate hyratase and is used in aerobic metabolism. A specific 475 bp-region of this gene is designated as the official FumC typing region for MLST and CH-typing (MLST database at the University of Warwick). This region has the greatest discriminatory power out of all seven loci used for MLST (Weissman et al. 2012). FumC also has high phylogenetic congruence (Wallace index=0.986) (Weissman et al. 2012), which indicates that a haplotype differentiated at both the FumC locus and by its phylogenetic group will be assigned to the same genotype 99% of the time. Therefore, haplotypes with the same FumC profiles are likely closely related to one another.

FimH is not used for MLST, but is one of the two loci used for CH-typing. FimH encodes for type 1 fimbrial adhesion. It is under positive selection for functional mutations related to adhesion of the cell to a substrate, which is relevant for pathogenicity (Weissman et al. 2006). In a sample of global, mostly clinical, *E. coli* isolates, FimH has higher discriminatory power than any single MLST locus. However, FimH is phylogenetically dispersed (Wallace index=0.504)

such that a haplotype differentiated at the FimH locus and by its phylogenetic group will be assigned to the same genotype 50% of the time (Weissman et al. 2012).

Assessing Haplotype Diversity and Designing New Primers

After Sanger sequencing all 414 isolates, I trimmed the resulting sequences at a 5% error probability limit and mapped them to the FumC (GenBank accession #: X04065) and FimH reference genomes (GenBank accession #: JQ658994) using Geneious (version 8.1.7) alignment. I removed isolates with low coverage or that appeared to be a species other than *E. coli*. I then determined the CH-type of each isolate and evaluated how many could be differentiated using only one CH locus (i.e. only FumC and only FimH). I selected the locus that better differentiated *E. coli* haplotypes and used Geneious (version 8.1.7) to design new primers that targeted a small enough region of DNA to be appropriate for high throughput Illumina sequencing. Microbiome researchers use between 200-460 base pairs of the 16S locus to differentiate genera of bacteria. The size of the target region depends on the platform used and the desired overlap in paired-end sequencing. I designed multiple primer pairs that targeted 150-300 bp in the most variable region of this locus, which would allow for little or no overlap in paired-end sequencing when using an Illumina MiSeq platform and a large overlap in paired-end sequencing when using an Illumina HiSeq Rapid platform. I determined the most variable region of this locus by aligning 101 FumC records in GenBank. As a result of designing primers that were appropriate for high throughput sequencing, I targeted a smaller region of the gene than the typical CH primers. I then narrowed down our primers to those that successfully and consistently amplified from a whole fecal sample (see methods below). Because *E. coli* constitutes only about 0.1% of the gut flora (Hartl and Dykhuizen 1984) and because feces contains PCR inhibitors, it is possible that some primers

will amplify a pure culture but will not amplify directly from a fecal sample. The 150-300 bp target region of each successful newly-designed primer was then bioinformatically "extracted" from each Sanger sequence to be further analyzed using analytical methods commonly applied in 16S high throughput studies. I chose to use this bioanalytical approach because I aimed to design a method that is appropriate for multiplex high throughput sequencing. Any isolate that did not have sequence coverage for at least 75% of the target primer region, or was not sequenced at both CH loci (unless it was a novel haplotype in our study at the one successful locus) was removed from the analysis. I then used the program Mothur v.1.37.0 (Schloss et al. 2009) to cluster sequences with the same or similar variants. The resulting unique sequences will be referred to as "Mothur sequences" throughout this paper. Differences in sequencing length did not cause sequences to be categorized as unique. I assumed sequence error rates of 0%, ~0.5%, and 1%, which translated to between 1 - 3 errors per sequence depending on the length of the target region of each primer. Error rates for Sanger sequencing range from 0.0001% to over 1% (Ewing and Green 1998, Hoff 2009). I assessed the number of haplotypes differentiated by the newly-designed primers and compared this differentiation ability to CH-typing at each of these error rates.

After evaluating the number of haplotypes that were differentiated by the one-locus versus the multi-locus approach, I next compared the identities of these haplotypes and used the igraph package in R (Csardi and Nepusz 2006) to visualize the network of haplotype relationships. There are three options for how the haplotype identities resulting from each approach could compare to one another: *Consistent assignment*: The newly-designed primer(s) e differentiated the same haplotypes as CH-typing; *Over collapse*: The newly-designed primer(s) incorrectly combined what should have been multiple CH types; *Differentiating*: The newly-

designed primer(s) identified novel haplotypes. *Differentiating* applies only to primer(s) that encompassed a region of the FumC gene that was completely or partially outside of the CH-typing region. Any evidence of differentiated haplotypes that resulted from primers that were completely within the FumC typing region was due to sequencing error and will instead be termed *Over splitting*. It is also possible that the relationship between the haplotype differentiation of CH-typing and the newly-designed primers was ambiguous. These cases were visualized in the networks but removed from the analysis. To assess the haplotype relationships between the two approaches, I compared how each *E. coli* sequence was categorized using CH-typing versus each successful primer assuming an error rate of 0%.

Testing Primer Functionality in Wild-Collected Fecal Samples

All newly-designed primers were evaluated for their melting temperatures, primer-dimer formation, and ability to amplify the target region out of a fecal sample. To test for functionality of the new primers, I amplified DNA extracts of whole fecal samples collected from wild ringtailed lemurs at Beza Mahafaly Special Reserve in Madagascar in 2015. Samples were collected within 5 minutes of an observed defecation using metal tweezers cleaned with sterile ethanol. I sampled three different areas of the feces and stored the samples in sterile plastic cryotubes containing the following storage solution: 40 ml 0.5 M EDTA, 25 ml 1M Sodium Citrate, 700 gm Ammonium Sulfate and 935 ml of sterile distilled water, pH adjusted to 5.2 using 1M H_2SO_4. Samples were labeled with the identity of the focal individual, and the date and time of collection. I attempted to sample internal portions of the feces to avoid contamination of bacteria from the environment. The samples were kept at ambient temperature for 24 hours and then stored at -10°C for up to three months, at which point they were transferred to a -80°C

freezer unil transported to the lab at Stony Brook University and again stored at -80°C until further analysis. To extract DNA, I thawed and mixed the fecal sample while still in the sample tube using a flamed and cooled metal spatula. I vortexed the sample for 30 seconds and used a pipet to draw 200ul of liquid and place it in a fresh tube. I extracted DNA from fecal samples using a QIAamp Stool Mini kit with a protocol developed by Brenda Bradley at George Washington University. The protocol included a 24-48-hour lysis phase and the addition of DTT and carrier RNA. I quantified the DNA in each sample using a Nanodrop (Thermo Scientific) and only used those samples with a DNA concentration above 10ng/ul and a 360/380 ratio above 1.8 for further analysis. I then PCR amplified the samples via a complete nested PCR using Illustra PuRe Taq Ready-To-Go Beads for the outer nest and Kapa HiFi HotStart taq for the inner nest. The specifics of the PCR protocol are described in the results section. I used a nested PCR because it allowed for the consistent amplification of the target region from an entire fecal sample and because it increased the probability that the PCR targeted the correct species of bacteria.

Results

CH-typing

Of the 414 colonies isolated, 272 were successfully Sanger sequenced at the FimH locus (Appendix 1, supplemental figure 1) and 276 were successfully Sanger sequenced at the FumC locus (Appendix 1, supplemental figure 2) and 253 were successfully sequenced at both loci, resulting in 34 FimH haplotypes, 42 FumC haplotypes, and 57 unique combinations of one or both loci (table 1). Five FumC haplotypes were deemed new additions to the MLST database; sequences 527, 528, 529, 534, and 543. FumC and FimH alone perfectly differentiated 74%

(42/57) and 60% (34/57) of CH-types, respectively. In addition, FumC identified four isolates as non-*E. coli* that FimH identified as one of the known *E. coli* haplotypes. Because FumC differentiated more haplotypes and better distinguished *E. coli* from non-*E. coli*, I selected this locus for development of the one-locus method.

CH Types (FumC-FimH)	Observed Frequency	CH Types (FumC-FimH)	Observed Frequency
100-96	1	4-86	2
103-397	2	41-86	2
11-54	3	45-97	8
187-ns	1	50-299	10
19-31	7	52-14	4
19-86	6	52-63	8
19-87	1	527-270	1
210-218	4	528-218	12
218-172	2	529-123	2
218-47	1	53-370	3
23-31	3	530-124	1
23-38	32	534-123	3
231-58	5	543-124	1
24-339	9	58-30	8
24-9	18	6-32	1
26-65	3	65-26	7
271-394	9	65-32	2
284-395	4	65-38	5
29-38	3	67-222	2
331-31	1	93-124	1
332-30	3	93-331	1
35-49	4	95-32	2
36-48	10	95-39	1
360-38	1	Novel*(35)-218	3
4-24	1	**Novel**(332)-ns**	1
4-31	2	Novel****(251)-123	3
4-38	18	Novel*****(251)-123	3
4-39	1	**Novel*******(224)-ns**	1
4-396	2		

Table 1. List of 57 CH-types (FumC-FimH) identified in this study and their frequencies. The bolded CH-types were only successfully sequenced at one of the two loci (ns = no sequence) but were included in this list because they represented unique haplotypes.

Newly-designed Primers

I designed two forward and five reverse primers that spanned up to 300 bp of the most variable region of the FumC locus. I then PCR-amplified the newly designed primers for this locus from whole fecal samples using a nested PCR protocol. The outer primers in the nested PCR were the standard FumC primers used for MLST and CH-typing (Forward: 5'-GTACGCAGCGAAAAAGATTC-3'; Reverse: 5'-TCACAGGTCGCCAGCGCTTC-3') (Wirth et al. 2006). The inner primers were the newly designed primer pairs, two of which successfully isolated their target regions (figure 1), as confirmed through BLASTing their Sanger sequences. A description of the successful primers is shown in table 2. The ingredients and protocol for each PCR was as follows:

Outer PCR:
Ingredients: Water: 19 ul, MgCl2, 1ul, primers: 1ul of each
Touch down PCR Protocol: 1) 95 degrees for 3 minutes, 2) 94°C for 45 seconds, 3) 65 degrees for 1 minute, 4) 72°C for 1.5 minutes, 5) go to step 2, 9 times, decreasing annealing temperature 1°C each cycle, 6) 94°C for 45 seconds, 7) 55°C for 1 minute, 8) 72°C for 1.5 minute, 9) go to step 6 19 times, 10) 72°C for 10 minutes, 11) hold at 4°C

Inner PCR:
Ingredients: 12.5 ul PCR reactions. Water: 13.35ul, 5x kappa HIFI buffer with MgCl2: 5ul, dNTPs: 0.75ul, BSA (20mg/mL): 0.4ul, primers: 1 ul of each, DNA: 1.5 ul
PCR Protocol: 1) 95°C for 3 minutes, 2) 98°C for 20 seconds, 3) 65°C for 30 sec, 4) 72°C for 15 seconds, 5) go to step 2 24 times, 6) 72°C 5 minutes, 7) hold at 4°C

Primer name	Sequence	GC content	Melting temperature
Primer pair one	Forward 5'-GGCTGTATTCRATATGTTTGAGATT-3' Reverse 5'-AGCCAAAGTTCYAAYGATGTCTTT-3'	33.3 - 41.7%	56.8 - 61.5°C
Primer pair two	Forward 5'-TCAAATTTGTTCGGYGCGGT-3' Reverse 5'-CTGCGCAAGCA ACTCATTCC-3'	45 - 55%	59 - 61.8°C

Table 2. Description of successful primers.

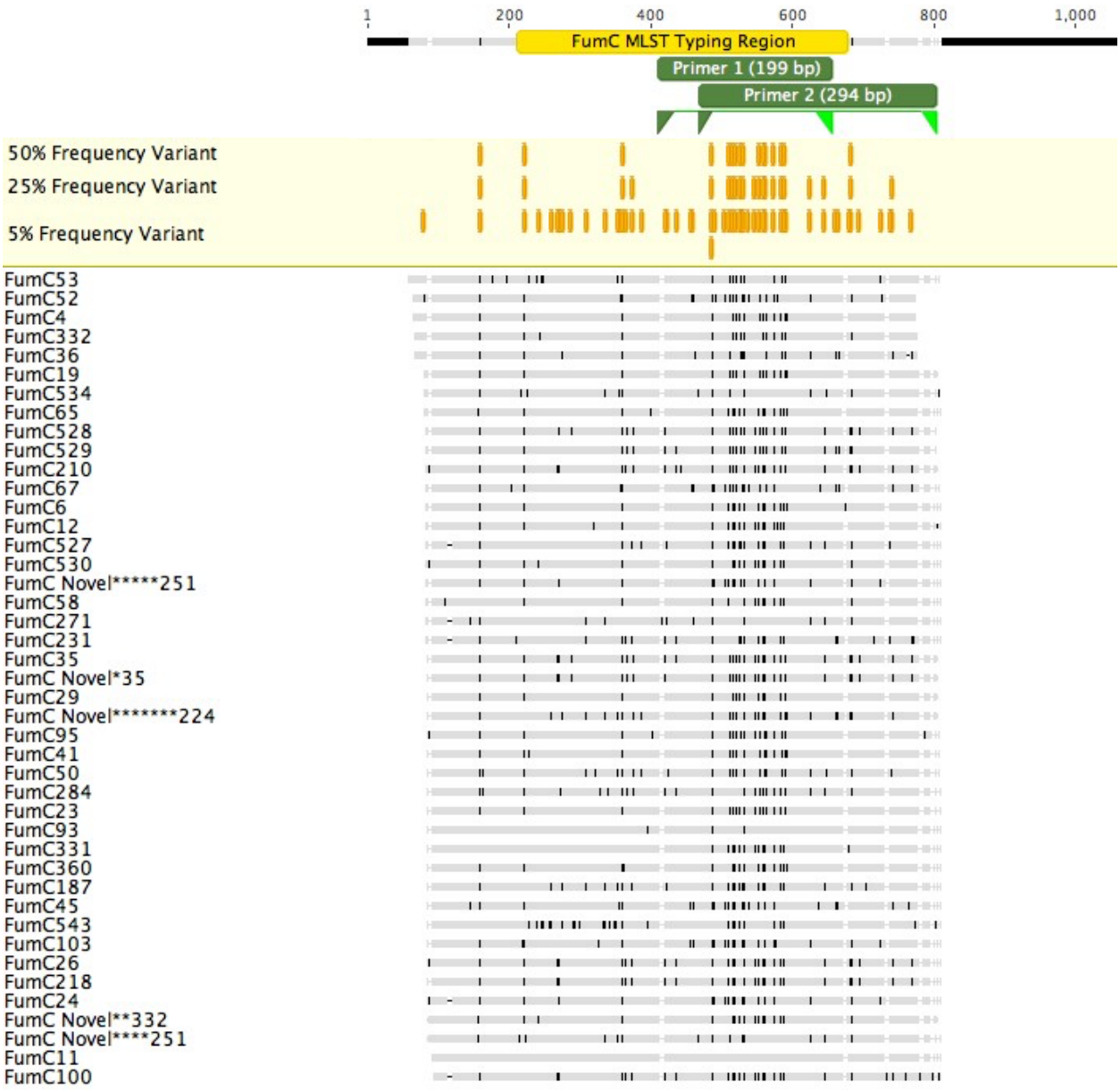


Figure 1. Alignment of reference genome (accession number: NZ_AAJU00000000) with the 43 FumC haplotypes from ringtailed lemurs identified in this study. The haplotypes were determined based on the standard 475 bp-FumC typing region for MLST (Wirth et al. 2006) (indicated in yellow). The names and identities of the FumC haplotypes are consistent with those in the MLST database at the University of Warwick (Wirth et al. 2006). The green brackets indicate regions targeted by the two newly-designed primers.

Both primers encompassed the majority of high- (i.e. 25-50% prevalence) and low- (i.e. 5% prevalence) frequency polymorphisms that were identified in the *E. coli* sequences derived from ringtailed lemurs in this study (figure 1). Primer pair one resides completely within the 475 bp-FumC typing region for MLST and CH-typing. It targets a 199 bp region that encompassed 12 of 18 high frequency polymorphisms and 17 of 37 overall polymorphisms (figure 1). Primer pair 2 resides partially within the 475 bp-FumC typing region. This primer targets a 294 bp region that encompassed 14 of 18 high frequency and 22 of 37 overall polymorphisms (figure 1)

There is evidence that the haplotype diversity captured by these primers reflects known *E. coli* population genetics. Specifically, the observation that a chromatogram of the Sanger sequence from a whole fecal sample resulted in mostly clear, single peaks, despite multiple haplotypes present in the fecal sample (figure 2), supports the presence of a major haplotype, as expected from known *E. coli* population genetics. Furthermore, the observation that nucleotide positions with multiple peaks in the chromatogram were often aligned with and shared the same nucleotide variants with known polymorphisms in the six primary reference genomes indicates the presence of multiple simultaneous haplotypes in the sample (peaks labeled "a", figure 2). However, there is also evidence of variation in the fecal chromatograms that is not present in the six reference genomes (peaks labeled "b", figure 2).

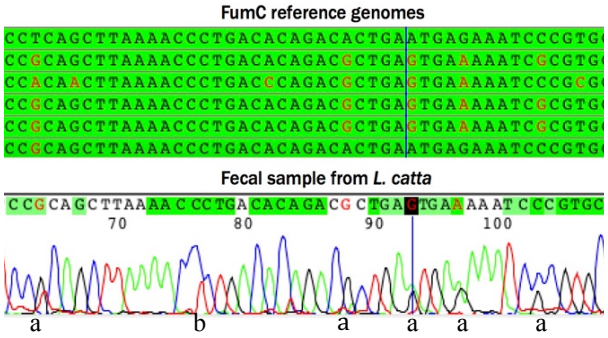


Figure 2. Alignment of FumC reference genomes with a chromatogram of a Sanger sequences from the newly-designed primer target region. Sites where polymorphisms in the reference genome and dual peaks in the chromatogram co-occur is designated by "a". Sites where genetic variation not present in the reference genome occurred is designated by "b".

Differentiation Capability of Newly-Designed Primers Using Sanger Sequencing

Because the newly-designed primers target a smaller region of the FumC gene than the MLST and CH-typing FumC primers, I calculated the number and identity of haplotypes that could be differentiated by these smaller regions alone using standard 16S high throughput sequencing bioanalytics to account for sequencing error in all 276 FumC isolates. I found that Primer pair one differentiated 36 (0 errors), 28 (1 error, 0.5% error rate), and 19 haplotypes (2 errors, 1% error rate) and Primer pair two differentiated 101 (0 errors), 52 (2 errors, 0.68% error rate), and 43 (3 errors, 1% error rate) haplotypes (table 3). When compared to the differentiation ability of CH-typing, which differentiated 57 distinct CH-types, primer pair one identified 33% (36/57) - 53% (36/57) of haplotypes and primer pair two identified 75% (43/57) - 177% (101/57) of haplotypes (table 3). Therefore, primer pair two differentiated more haplotypes than primer pair one at all error levels. Furthermore, at low error levels, primer pair two actually differentiated more haplotypes than CH-typing, which suggests that variation present in the region of primer pair two that was outside of the CH-typing region, resulted in the splitting of haplotypes that would otherwise have been categorized as a single CH-haplotype. Therefore, I recommend primer pair two for one-locus differentiation of *E. coli* at the FumC gene.

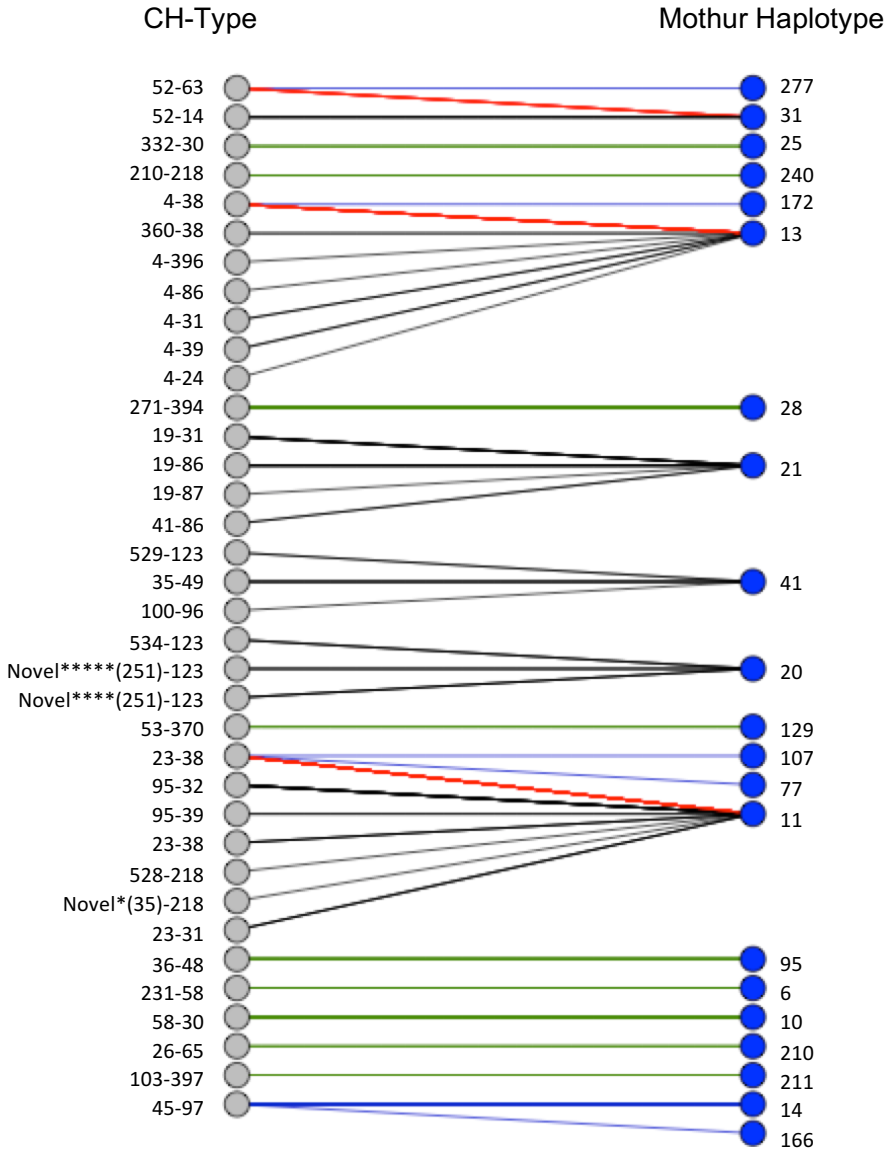
CH-typing	Error	Primer pair 1	Primer pair 2
57	0	36	101
	1 (~0.5%)	28	52

Table 3. Number of haplotypes identified using CH-typing, Primer 1 and Primer 2.

I also compared the identities of the haplotypes resulting from the one-locus and multi-locus differentiation approaches. Of the 36 Mothur sequences differentiated by Primer pair 1, 18 correlated with a unique CH-type (*consistent assignment*, green lines in figure 3) and 11 incorrectly combined multiple CH-types (*over collapse*, black lines in figure 3). Seven of the

Mothur sequences (blue circles with at least one blue line in figure 3) corresponded to five CH-types (grey circles with multiple lines, at least one of which is blue, in figure 3). However, because primer pair one lay completely within the CH-typing region, this apparent increase in diversity found with CH-typing was *over splitting* due to error. Therefore, Mothur haplotypes 14 and 166 and 107 and 77 were incorrectly split (figure 4), reducing the maximum number of haplotypes identified with Primer pair 1 to 34 and the percentage of haplotypes identified to 60%.

Of the 101 Mothur sequences generated by Primer pair 2, 12 correlated with a unique CH-type (*consistent assignment*, green lines in figure 4) and 11 incorrectly combined multiple CH-types (*over collapse*, black lines in figure 4). Because three additional Mothur haplotypes consisted only of ambiguously assigned CH-types (Mothur haplotypes 197, 198, and 222, figure 4), I removed these haplotypes from the total, reducing the maximum number of haplotypes identified with primer 2 to 98 and the percentage of haplotypes identified to 172%. Importantly, 75 Mothur sequences (blue circles with at least one blue line in figure 4, 77% (75/98) of total Mothur haplotypes) corresponded to 25 CH-types (grey circles with multiple lines, at least one of which is blue, in figure 4), which indicates that this primer identified more sequence variation than CH-typing (*differentiating*). Because primer pair two lay partially outside of the CH-typing target region, the increased diversity identified by this primer is likely real variation and not entirely due to error.


CH-Type
Mothur Haplotype
52-63
52-14
332-30
210-218
4-38
360-38
4-396
4-86
4-31
4-39
4-24
271-394
19-31
19-86
19-87
41-86
529-123
35-49
100-96
534-123
Novel*****(251)-123
Novel****(251)-123
53-370
23-38
95-32
95-39
23-38
528-218
Novel*(35)-218
23-31
36-48
231-58
58-30
26-65
103-397
45-97
277
31
25
240
172
13
28
21
41
20
129
107
77
11
95
6
10
210
211
14
166

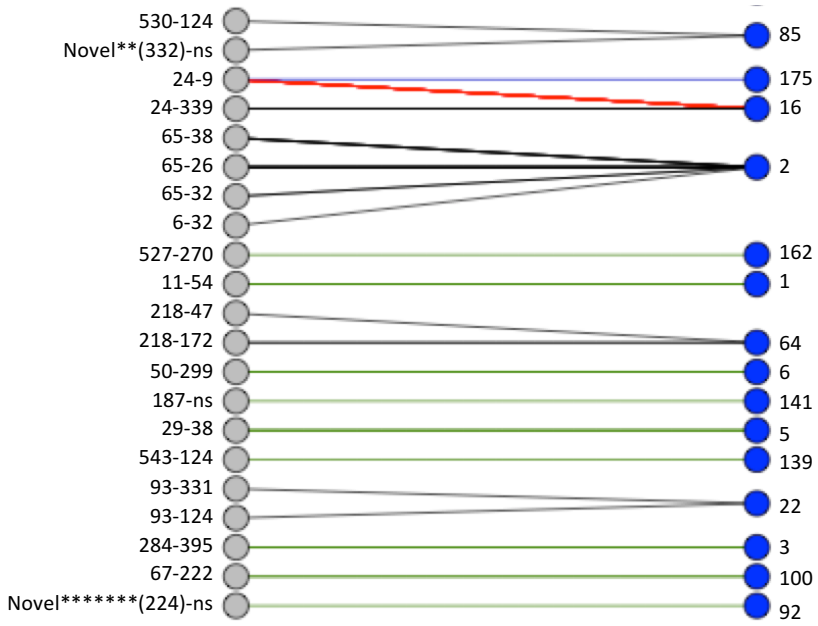
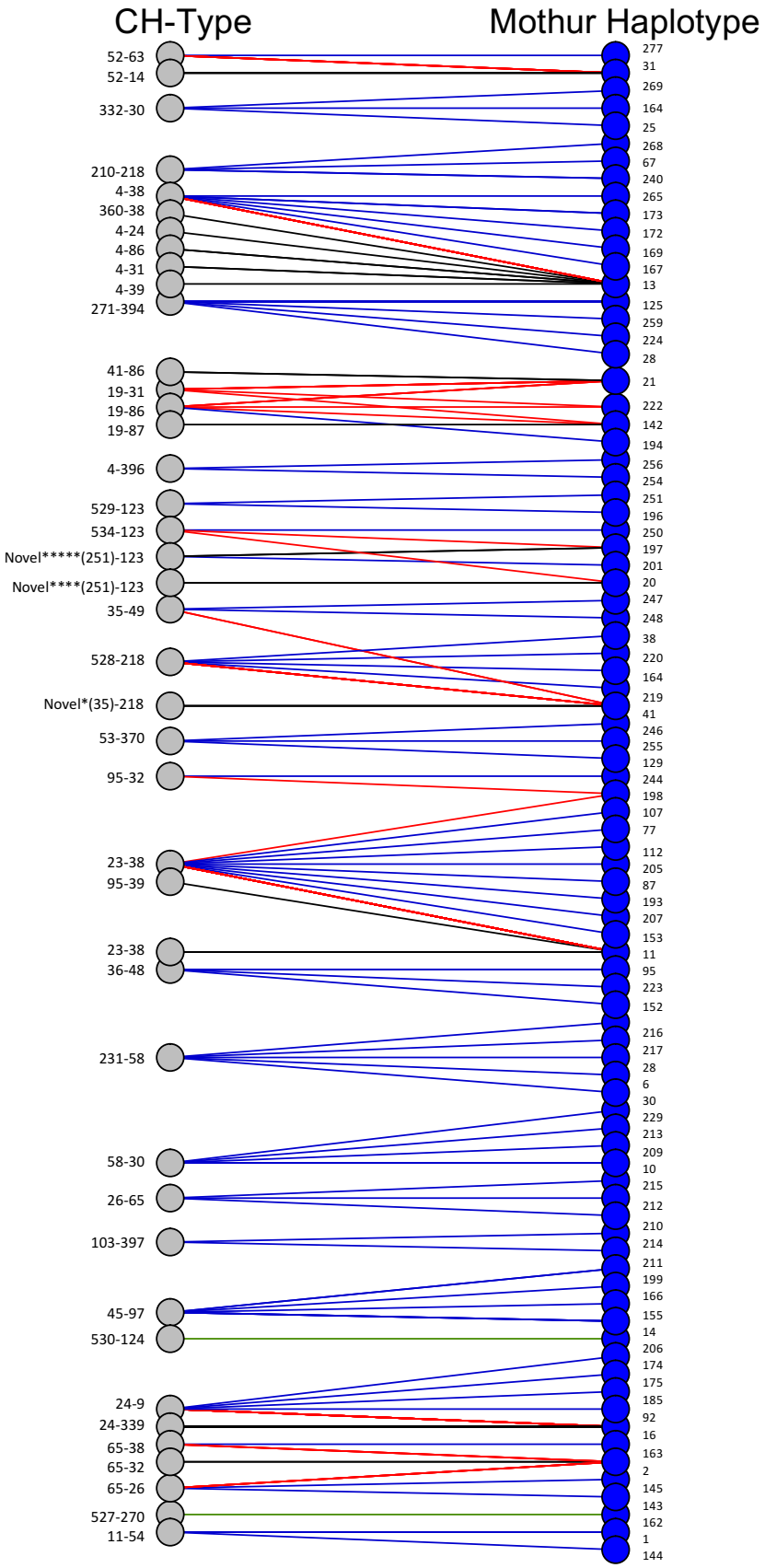


Figure 3. Haplotype network for primer 1. The CH-types are shown on the left (grey circles) and the Mothur Haplotypes (0% error) are shown on the right (blue circles). The green lines indicate a perfect correspondence between CH-types and Mothur haplotypes (*consistent assignment*). The black lines indicate Mothur sequences that correspond to more than one CH-type (*over collapse*). The blue lines indicate CH-types that correspond to more than one Mothur sequence (*over splitting*). The red lines indicate instances where the relationship between CH-types and Mothur sequences was ambiguous.


CH-Type
Mothur Haplotype
52-63
52-14
332-30
210-218
4-38
360-38
4-24
4-86
4-31
4-39
271-394
41-86
19-31
19-86
19-87
4-396
529-123
534-123
Novel*****(251)-123
Novel****(251)-123
35-49
528-218
Novel*(35)-218
53-370
95-32
23-38
95-39
23-38
36-48
231-58
58-30
26-65
103-397
45-97
530-124
24-9
24-339
65-38
65-32
65-26
527-270
11-54
277
31
269
164
25
268
67
240
265
173
172
169
167
13
125
259
224
28
21
222
142
194
256
254
251
196
250
197
201
20
247
248
38
220
164
219
41
246
255
129
244
198
107
77
112
205
87
193
207
153
11
95
223
152
216
217
28
6
30
229
213
209
10
215
212
210
214
211
199
166
155
14
206
174
175
185
92
16
163
2
145
143
162
1
144

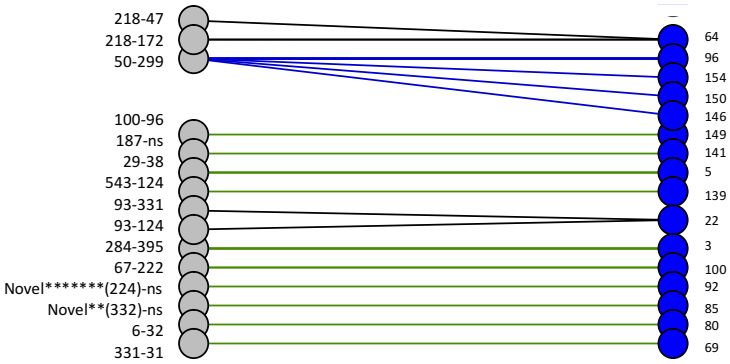


Figure 4. Haplotype network for primer 2. The CH-types are shown on the left (grey circles) and the Mothur Haplotypes (0% error) are shown on the right (blue circles). The green lines indicate a perfect correspondence between CH-types and Mothur haplotypes (*consistent assignment*). The black lines indicate Mothur sequences that correspond to more than one CH-type (*over collapse*). The blue lines indicate CH-types that correspond to more than one Mothur sequence (*differentiating*). The red lines indicate instances where the relationship between CH-types and Mothur sequences was ambiguous.

Discussion:

Traditionally, the genetic variation at multiple loci is collectively used to differentiate haplotypes of *E. coli* (e.g. Blyton et al. 2014, Caugant et al. 1984, Goldberg et al. 2007, Rewgo et al. 2008, Sears et al. 1956, Springer et al. 2016, VanderWaal 2014a, b). However, these multi-locus differentiation methods require a culture-based approach so that the variation across loci can be linked. Such approaches are time- and labor-intensive and likely to under-represent *E. coli* diversity within a sample. A single locus approach to haplotype differentiation is necessary to bypass the cell culturing step. In this study, I quantified the *E. coli* haplotype diversity that could be identified from a short region of a single locus and compared it to the diversity identified from an established, multi-locus method of haplotype differentiation, CH-typing. However, because a one-locus approach necessarily incorporates less genetic variation for differentiation than a multi-locus approach, I determined the extent to which this novel method altered the ability to distinguish among *E. coli* haplotypes as compared to established multi-locus methods.

I found that the CH-typing region of the FumC locus alone differentiated more haplotypes than the CH-typing FimH locus alone, making this locus a reasonable target for the development of a one-locus method to haplotype differentiation. I also found that the most variable of the newly-designed FumC primers, primer pair 2, identified between 75% (at an error rate of 0%) and 172% (at an error rate of 1% with ambiguous assignments removed from the analysis) of the haplotypes that could be identified using CH-typing. Sanger sequencing is considered to have very low error rates, with an error occurring every 10,000 – 100,000 nt (0.00001% - 0.0001%) (Ewing and Green 1998). Therefore, rates of 0.5% and 1%, as used in this study, were likely substantial overestimates. As a result, the true error rate was likely somewhere between 0% and 0.5%, putting the number of identified haplotypes for primer pair 2 between 52 and 98, which is 91%-172% of the diversity identified with CH typing. Therefore, at low error rates, this primer differentiated more haplotypes than CH-typing through the novel differentiation of existing CH-types. In fact, the haplotype network for primer pair two showed that 77% of its haplotypes were the result of the novel differentiation of established CH-types. Because primer pair 2 extended beyond the FumC CH-typing region, it is possible for novel variation to be identified by this primer. In fact, primer pair 1, which lay completely within the FumC typing region and thus could not identify novel variation, had very little over splitting and likely very little error. If, as expected, both target regions had similar error rates, primer pair two probably has a low error rate as well.

Overall, most CH-types were correctly identified as unique using primer pair 2 either by being assigned to a single Mothur haplotype or by being assigned to multiple novel Mothur haplotypes. However, 11 out of 98 Mothur haplotypes over collapsed multiple CH-types. FumC has high phylogenetic congruence (Wallace index=0.986) (Weissman et al. 2012), which

indicates that a haplotype differentiated at the FumC locus and by phylogenetic group will generally be assigned to the same genotype. Therefore, different haplotypes that were incorrectly identified as having the same FumC profiles at the primer 2 target region were likely close relatives. This phylogenetic congruence suggests that distinct haplotypes incorrectly grouped together may still indicate useful patterns for some research questions, such as the social sharing of bacteria where *E. coli* populations are evolving in the gut of the host before being shared.

Although the FumC locus differentiated more haplotypes than the FimH locus in this study, Weissman et al. (2012) actually reported the opposite pattern. Interestingly, while the differentiation ability of FimH was very similar in my study and in the Weissman et al. (2012) paper, the differentiation ability of FumC was much higher in my study. This divergent pattern can likely be explained by the nature of the isolates collected in the two studies. Weissman et al. (2012) used isolates from many species and geographic locations whereas our isolates originated from a comparatively closed system that may have distinct *E. coli* population dynamics. Additionally, the majority (51%) of Weissman et al.'s (2012) isolates were pathogenic, which are likely to experience different selective regimes than non-pathogenic haplotypes in natural systems (Weissman et al. 2012, supplemental material, table S1). Therefore, the ability for a haplotype differentiation approach to distinguish among haplotypes may differ based on the pathogenic vs. non-pathogenic nature of the samples. This difference may be especially relevant for the FimH locus, which is horizontally transmitted across *E. coli* cells and even among bacterial species and is essential for encoding adhesion to surfaces within the host, a major factor determining pathogenicity (Sokurenko et al. 1998). By way of evidence, 60% (12/20) of the 20 CH-types used by Weissman et al. (2012) that originated in wild animals could be differentiated with FumC alone, which is more similar to the FumC differentiation observed in our study.

Therefore, in a natural system, FimH may be less meaningful for haplotype differentiation but in studies where pathogenic haplotypes are the predominant haplotypes being differentiated, FimH may be more appropriate.

In this paper, I presented a one-locus, culture-independent method to differentiate *E. coli* haplotypes that is compatible with high throughput sequencing and captures much of the genetic diversity as an established multi-locus method. This one-locus method renders cell culturing unnecessary and offers the potential to differentiate nearly every one of the billions of *E. coli* cells in a fecal sample. However, there is a tradeoff between a near-perfect differentiation of a limited number of cells at multiple loci and the less-perfect differentiation of nearly every cell in a sample at one-locus. Yet, as demonstrated by the method detailed here, that tradeoff appears minimal or non-existent at typical Sanger sequencing error levels and acceptable at high Sanger sequencing error levels. Furthermore, because nearly all mammals harbor *E. coli* in their guts, this bacterium can be used as a model "pathogen" to study the movement of bacteria within and among populations of humans (e.g. Caugant et al. 1984) and wild mammals (e.g. Goldberg et al. 2006, VanderWaal et al. 2014a). This one-locus method is the first to differentiate bacterial haplotypes of a single species and can be a powerful approach to answer questions about disease ecology in a faster and more fine-grained manner than was previously possible.

Literature Cited:

Bachmann, N. L., Sullivan, M. J., Jelocnik, M., Myers, G. S. A., Timms, P. and Polkinghorne, A. 2015. Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by *Chlamydia pecorum*. *Journal of Clinical Microbiology* 53: 1573–1581.

Blyton, M. D. J., Banks, S. C., Peakall, R., Lindenmayer, D. B. and Gordo, D. M. 2014. Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecology Letters* 17: 970–978.

Caugant, D. A., Levin, B. R. and Selander, R. K. 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. *Journal of Hygiene-Cambridge* 92: 377-384.

Cho, I. and Blaser, M. J. 2012. The human microbiome: at the interface of health and disease. *Nature Review Genetics* 13: 260–270.

Cohen, T., van Helden, P. D., Wilson, D., Colijn, C., McLaughlin, M. M., Abubakar, I. and Warren, R. M. 2012. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clinical Microbiology Reviews* 25: 708–719.

Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695. http://igraph.sf.net.

Dethlefsen, L., Huse, S., Sogin, M. L. and Relman, D. A. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology* 6: e280.

Dias, R. C. S., Moreira, B. M. and Riley, L. W. 2010. Use of FimH single-nucleotide polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. *Journal of Clinical Microbiology* 48: 483-488.

Dubois, V., De Barbeyrac, B., Rogues, A. M., Arpin, C., Coulange, L., Andre, C., M'Zali, F., Megraud, F. and Quentin, C. 2010. CTX-M-producing *Escherichia coli* in a maternity ward: a likely community importation and evidence of mother-to-neonate transmission. *Journal of Antimicrobial Chemotherapy* 65: 1368–1371.

Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* 8: 186–194.

Eyre, D. W., Cule, M. L., Griffiths, D., Crook, D. W., Peto, T. E. A., Walker, S. A. and Wilson, D. J. 2013. Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. *PLoS Computational Biology* 9: e1003059. doi:10.1371/journal.pcbi.1003059.

Ezenwa, V. O., Gerardo, N. M., Inouye, D. W., Medina, M., and Xavier, J. B. 2012. Animal behavior and the microbiome. *Science* 338: 198–199.

Georghiou, P. R., Doggett, A. M., Kielofner, M. A., Stout, J. E., Watson, D. A., Lupski, J. R. and Hamill, R. J. 1994. Molecular fingerprinting of *Legionella* species by repetitive element PCR. *Journal of Clinical Microbiology* 32: 2989–2994.

Goldberg, T. L., Gillespie, T. R., Rewgo, I. B., Wheeler, E., Estoff, E. L. and Chapman, C. A. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation* 135: 511-517.

Hartl, D. L. and Dykhuizen, D. E. 1984. The population genetics of *Escherichia coli*. *Annual Review of Genetics* 18: 31–68.

Heilig, H. G., Zoetendal, E. G., Vaughan, E. E., Marteau, P., Akkermans, A. D. and de Vos, W. M. 2002. Molecular diversity of *Lactobacillus spp*. and other lactic acid bacteria in the human intestine as determined by specific amplification of 16S ribosomal DNA. *Applied and Environmental Microbiology* 68: 114–23.

Hill, J. E, Seipp, R. P., Betts, M., Van Kessel, A. G., Crosby, W. L. Hemmingsen, S. M. and Hawkins, L. 2002. Extensive profiling of a complex microbial community by high-throughput sequencing. *Applied Environmental Microbiology* 68: 3055–3066.

Hoff, J. K. 2009. The effect of sequencing errors on metagenomics gene prediction. *BMC Genomics* 10: 520. doi:10.1186/1471-2164-10-520.

Michalak, K., Austin, C., Diesel, S., Bacon, J. M., Zimmerman, P. and Maslow, J. N. 1998. *Mycobacterium tuberculosis* infection as a zoonotic disease: Transmission between humans and elephants. *Emerging Infectious Diseases* 4: 283–287.

Montoro, E., Valdivia, J. and Leao, S. C. 1998. Molecular fingerprinting *of Mycobacterium tuberculosis* isolates obtained in Havana, Cuba, by IS6110 restriction fragment length polymorphism analysis and by the double repetitive-element PCR method. *Journal of Clinical Microbiology* 36: 3099–3102.

Odell, I. D. and Flavell, R. A. 2016. Microbiome: Ecology of eczema. *Nature Microbiology* 1: 16135

Rajilić-Stojanović, M., Smidt, H. and De Vos, W. M. 2007. Diversity of the human gastrointestinal tract microbiota revisited. *Environmental Microbiology* 9: 2125–36.

Rewgo, I. B., Isabirye-Basuta, G., Gillespie, T. R. and Goldberg, T. 2008. Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600–1607.

Sampson, T. R. and Mazmanian, S. K. 2015. Control of brain development, function, and behavior by the microbiome. *Cell Host and Microbe* 17: 565–576.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., Weber, C. F. 2009. Introducing Mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied Environmental Microbiology* 75: 7537–7541.

Sears, H. J., Janes, H., Saloum, R., Brownlee, I. and Lamoreaux, L. F. 1956. Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *Journal of Bacteriology* 71: 370–372.

Sokurenko, E. V., Chesnokova, V. Dykhuizen, D. E., Ofek, I., Wu, X., Krogfelt, K. A., Struve, C., Schembri, M. A. and Hasty, D. L. 1998. Pathogenic adaption of *Escherichia coli* by natural variation of the FimH adhesin. *PNAS* 95: 8922–8926.

Springer, A., Mellmann, A., Fichtel, C. and Kappeler, P. M. 2016. Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecology* 16: 1–6.

Suau, A., Bonnet, R. Sutren, M., Gordon, J., Gibson, G. R., Collins, M. D., Doré, J. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Applied Environmental Microbiology* 65: 4799–4807.

Taylor, N. S., Fox, J. G., Akopyants, N. S., Berg, D. E., Thompson, N., Shames, B., Yan, L., Fontham, E., Janney, F., Hunter, F. M., et al. 1995. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA-fingerprinting. *Journal of Clinical Microbiology* 33: 918–923.

Tung, J., Barreiro, L. B., Burns, M. B., Grenier, J. C., Lynch, J., Grieneisen, L. E., Altmann, J., Alberts, S. C., Blekhman, R. and Archie, E. A. 2015. Social networks predict gut microbiome composition in wild baboons. *eLife* 4: e05224. http://dx.doi.org/10.7554/eLife.05224.

VanderWaal, K. L., Atwill, E. R., Isbell, L. A. and McCowan, B. 2014a. Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *Journal of Animal Ecology* 83: 406–414.

VanderWaal, K. L., Atwill, E. R., Isbell, L. A., McCowan, B. 2014b. Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation* 169: 136–146.

Weissman, S. J., Chattopadhyay, S., Aprikian, P., Obata-Yasuoka, M., Yarova-Yarovaya, Y., Stapleton, A., Ba-Thein, W., Dykhuizen, D., Johnson, J.R. and Sokurenko, E. V. 2006.

Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*. *Molecular Microbiology* 59: 975–988.

Weissman, S. J., Johnson, J. R., Tchesnokova, V., Billing, M., Dykhuizen, D., Riddell, K., Rogers, P., Qin, X., Butler-Wu, S., Cookson, B. T., Fang, F. C., Scholes, D., Chattopadhyay, S. and Sokurenko, E. 2012. High-resolution two-locus clonal typing of Extraintestinal Pathogenic *Escherichia coli*. *Applied and Environmental Microbiology* 78: 1353–1360.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves, P. R., Maiden, M. C., Ochman, H. and Achtman, M. 2006. Sex and virulence in *Escherichia coli*: An evolutionary perspective. *Molecular Microbiology* 60: 1136–1151.

Socially facilitate *E. coli* transmission in ringtailed lemurs using a novel one-locus approach to haplotype differentiation

Chapter 4

Abstract: Questions about socially facilitated transmission are best addressed by combining data on the observed contacts of the host and the haplotype-level genetics of the pathogen because individuals with known social contact must share the same or related haplotypes of an infectious organism for transmission to be deduced. However, only three studies have used these two types of data simultaneously to assess patterns of transmission in social animals. Further, they have only done so using culture-based methods for the differentiation of bacterial haplotypes. These methods are time- and resource-intensive and inappropriate for identifying the true diversity of haplotypes within a single host. I collected detailed behavioral data and biweekly fecal samples from adult ringtailed lemurs in three social groups over a 6-month period. In this paper, I describe the application of this data set to a novel, culture independent, one-locus method of *E. coli* haplotype differentiation. Although not usually pathogenic, *E. coli* is valuable as a model "pathogenic" organism for social transmission studies because its ubiquity and high within-host haplotype diversity in mammals allows the inference of fine-scale patterns of transmission among *all* individuals. Furthermore, its well-known genetics makes it possible to assess the haplotype diversity that is and is not captured by this approach. I report preliminary results showing the effectiveness of this method on wild-collected fecal samples and its potential to capture much more within-host *E. coli* haplotype diversity than has previously been identified. This method can revolutionize our ability to determine fine-scale transmission dynamics and assess *E. coli* population genetics within a wild host.

Introduction:

The close proximity and frequent social interaction involved in group life provides increased opportunity for the horizontal transmission of microorganisms. Therefore, enhanced disease transmission is a potential cost of group living (Alexander 1974). As a result of this cost to group living, it has been hypothesized that pathogens cause selection on the evolution of mating systems and social interactions (Møller et al. 2001), resulting in behavioral and physiological mechanisms that reduce the risk of disease transmission (Freeland 1976). For example, sexually promiscuous primates have higher leukocyte counts than less promiscuous species (Nunn et al. 2000) and social species of wasp have higher antimicrobial activity than solitary species (Hoggard et al. 2011). Furthermore, it has also been suggested that the formation of modular groups with limited inter-group interaction may act to quarantine pathogens (Freeland 1976, 1979, Griffin and Nunn 2012, Hess 1996, Wilson et al. 2003). Additionally, females may reduce the risk of disease transmission by using secondary sexual characteristics,

such as color intensity, to evaluate the parasite load and genetic superiority of potential mates (Hamilton and Zuk 1982, birds; Ressel and Schall 1989, reptiles; Simmons 1990, crickets; Tinsley 1990, amphibians).

Questions about socially facilitated transmission are best addressed by combining data on the observed contacts of the host and the genetics of the pathogen (Eyre et al. 2013). More specifically, it is essential to genetically differentiate the pathogen to the level of the haplotype because individuals must harbor the same or a related haplotype of a particular pathogen for transmission to be deduced (Craft 2015, Eames et al. 2014). There are four approaches to the study of socially facilitated transmission that achieve elements of these best practices to varying extents (table 1): *Approach 1)* Correlations of shared pathogens species with inferred host affiliations using group size (e.g. Hoogland 1979, Loudon 2009, McGrew et al. 1989, Rubenstein and Hohmann1989), population density (e.g. Nunn et al. 2003, Whiteman and Parker 2004) or self-reports (e.g. Gardy et al. 2011, Rothenberg et al. 1995). Because neither the pathogen's genetics or the host's contacts are directly identified, this approach allows only the speculation of both host contact and pathogen transmission. For example, the risk of contracting sexually transmitted infections was found to increase with the number of self-reported mating partners in humans (Rothenberg et al. 1995) and the prevalence of parasitic nematodes increased with group

	Differentiate pathogen species	**Differentiate pathogen haplotypes**
Inferred affiliations	**Approach 1** Speculate transmission Speculate contact	**Approach 3** Confirm transmission Speculate contact
Observed affiliations	**Approach 2** Speculate transmission Confirm contact	**Approach 4** Confirm transmission Confirm contact

Table 1. Approaches to the study of socially facilitated pathogen transmission

size in olive baboons (*Papio anubis*) (McGrew et al. 1989). *Approach 2)* Correlation of shared pathogen species with observed host affiliations (e.g. Loudon 2009, Nunn et al. 2000). This approach allows the confirmation of host contact but speculates transmission because pathogen genetics are not assessed. For example, ectoparasite intensity in ringtailed lemurs increased seasonally as the frequency of dyadic grooming increased (Loudon 2009). *Approach 3)* Correlation of genetically differeniated pathogen haplotypes with inferred host affiliations using shared space or overlapping territories (e.g. Caugant et al. 1984, Dubois et al. 2010, Eyre et al. 2013, Georghiou et al. 1994, Goldberg et al. 2007, Michalak et al. 1998, Montoro et al. 1998, Rewgo et al. 2008). This approach allows the confirmation of transmission but only speculation of host contact, making it possible that individuals may share pathogens due to shared habitats. For example, the fecal *E. coli* of gorillas (Goldberg et al. 2007) and chimpanzees (Rewgo et al. 2008) were genetically similar to those humans with whom their habitat overlapped. Additionally, cohabitating humans shared over twice the *E. coli* strains as unassociated humans (Caugant et al. 1984). Finally, humans in the same hospital shared clusters of *Mycobacterium turberculosis* isolates (Montoro et al. 1998) and pathogenic *E. coli* (Dubois et al. 2010). *Approach 4)* Correlation of genetically differentiated haplotypes with observed host affiliations (i.e. Blyton et al. 2014, Springer et al. 2016, VanderWaal et al. 2014a, b). This approach allows both the confirmation of host contact and transmission and is the strongest approach for the identification of socially facilitated transmission. Notably, there are only four studies that use this approach (i.e. Blyton et al. 2014, Springer et al. 2016, VanderWaal et al. 2014a, b). For example, the sharing of *E. coli* haplotypes among opossums was better explained by host contacts than spatial proximity (Blyton et al. 2014). Furthermore, individual giraffes that shared

E. coli haplotypes were more likely to be strongly linked in the social network (VanderWaal et al. 2014a).

All of these studies are limited by their use of culture-based method, which is a common way to differentiate bacterial pathogens and requires growing a sample in culture and identifying the genetic similarities and differences among isolates across multiple loci. Culture-based methods are limited by cell culturing, which is time-, labor-, and resource-intensive (Bachmann et al. 2015). Infections by pathogenic organisms frequently contain multiple haplotypes of the same species (Bachmann et al. 2015, Eyre et al. 2013, Taylor et al. 1995), which greatly influences the approach and outcome of treatment (Cohen et al. 2012). Therefore, the identification of haplotype diversity within a mixed-haplotype sample, particularly of rare or un-culturable variants, is challenging with culture-based methods. Culture-independent methods using one highly variable locus allow researchers to bypass the culturing step and identify more diversity within mixed-haplotype samples than is possible with culture-based methods. The most powerful existing approach of this kind involves high throughput sequencing of the small RNA subunit of the 16S gene for (e.g. Cho et al. 2012, Ezenwa et al. 2012, Odell and Flavell 2016, Sampson et al. 2015). However, because the 16S gene can only differentiate bacteria to the level of species or genera, questions that require differentiation at the level of the haplotype, such those related to pathogen transmission and disease treatment, cannot be readily answered using this approach (Heilig et al. 2002). Questions about pathogen transmission in particular are most accurately addressed by haplotype-level differentiation because individuals must share the same or a related haplotype of a particular organism for transmission to be deduced (Eyre et al. 2013). Therefore, although one-locus culture-independent methods such as 16S are powerful at

detecting genus-level diversity and efficient in terms of time and resources, they cannot currently be used to assess confirmed incidences of transmission.

I collected detailed data on the social behavior of three groups of ringtailed lemurs and collected fecal samples from each focal individual every two weeks over a 6-month period. In this paper, I describe the application of this data set to a novel, culture independent, one-locus method of *E. coli* haplotype differentiation (Chapter 3). Here I report the fecal collection and laboratory methods used to generate bacterial genetic data from these samples. I also provide preliminary results demonstrating the potential of this method to revolutionize our ability to determine fine-scale transmission dynamics and assess *E. coli* population genetics within a wild host. *E. coli* is highly valuable as a model "parasitic" organism for social transmission studies. Although not usually pathogenic, its ubiquity in mammals and high haplotype diversity within a single host (Hartl and Dykhuizen 1984) allows the inference of fine-scale patterns of transmission among *all* individuals, instead of just those infected by an occasional pathogen. *E. coli* is transmitted by oral ingestion, either through a fecal-oral route or through the consumption of contaminated water, food or soil. Further, because individuals are known to carry one or two major haplotypes of *E. coli* that is at high frequency for long periods of time and is distinct from other hosts, these haplotypes can potentially act as bacterial identifiers for individuals, facilitating the identification of the directionality of transmission events. Finally, there is evidence that hosts differentially share haplotypes of *E. coli* as a function of their interactions (e.g. Caugant et al. 1984, Blyton et al. 2014, Springer et al. 2016, VanderWaal et al. 2014a).

Methods:

Sample Collection and Storage

I collected feces non-invasively from 29 adult and subadult ringtailed lemurs in three social groups at Beza Mahafaly Special Reserve in south western Madagascar. Feces were collected from all focal individuals in the same groups in the same day and from all groups within 2-3 days of one another. Samples were collected within 5 minutes of an observed defecation using metal tweezers cleaned with sterile ethanol. I sampled three different areas of the feces and stored the samples in sterile plastic cryotubes containing the following storage solution: 40 ml 0.5 M EDTA, 25 ml 1M sodium citrate, 700 gm ammonium sulfate and 935 ml of sterile distilled water, pH adjusted to 5.2 using 1M H_2SO_4. Samples were labeled with the identity of the focal individual, and the date and time of collection. To the estent possible, I sampled internal portions of the feces to avoid contamination of bacteria from the environment. The samples were kept at ambient temperature for 24 hours to allow the storage solution to permiate the bacterial cells and then stored at -10°C for up to three months, at which point they were transferred to a -80°C freezer until transported to the lab at Stony Brook University and again stored at -80°C until further analysis.

Sample Processing

I processed samples according to the protocol I developed and described in Chapter 3 of this dissertation. Specifically, I thawed and mixed the fecal sample while it was still in the sample tube using a flamed and cooled metal spatula. I vortexed the sample for 30 seconds and used a pipet to draw out 200ul of liquid and place it in a fresh tube. I then extracted DNA using a QIAamp DNA Stool Mini Kit with a protocol developed by Brenda Bradley at George

Washington University. The protocol included a 24-48-hour lysis phase and the addition of DTT and carrier RNA. I quantified the DNA in each sample using a Nanodrop (Thermo Scientific) and used those samples with a DNA concentration above 10ng/ul and a 360/380 ratio above 1.8. Any samples that did not pass these requirements, were extracted a second time. I then PCR amplified the samples via a complete nested PCR, which involved five replicates of the outer nest using Illustra PuRe Taq Ready-To-Go Beads that were then mixed together and the mixture was then PCR amplified for three replicates of the inner nest using Kapa HiFi HotStart Polymerase. The replicates functioned to reduce the impact of PCR bias. Equal concentrations of all PCR replicates were combined and then cleaned up with a 1x volume of Agencourt AMPure XP beads and run on a gel to confirm amplification.

The PCR protocols are described in detail in Chapter 3. The primers for the outer nest were the same as those described in Chapter 3. The primers for the inner nest had the same locus-specific region as described in Chapter 3 but with several additional features to make it consistent with the Illumina TrueSeq Nested PCR construct and appropriate for high throughput sequencing. Specifically, these inner primers contained the Illumina adapter region, followed by eight random nucleotides (i.e. denoted by Ns) to increase complexity of the libraries, one of two sets of internal indices (set 1: *Forward:* 5'-**CGATGT**-3', *Reverse:* 5'-**TATATACGC**-3'; set 2 = *Forward:* 5'-**ACAGTG**-3', *Reverse:* 5'-**ACTAGCAGA**-3'), and the locus-specific region. The internal indices were incorporated into the primer to extend the 194 indices that had been developed by the University of Wisconsin so that all 273 samples could be sequenced in a single lane. The primers were constructed as follows: *Forward:* 5'-ACACTCTTTCCCTACACGAC GCTCTTCCGATCTNNNNNNNN**CGATGT**CTGCGCAAGCAACTCA TTCC-3'; *Reverse:* 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNNNNN**TATATACGC**TCAA

ATTTGTTCGGYGCGGT-‘3. I stored the amplified samples at -80°C until they were ready to be mailed on dry ice to the Biotechnology Lab at the University of Wisconsin where a final 8-cycle PCR was conducted to add the remaining stem of the Illumina adapter with the following primers: *Forward:* 5’-AATGATACG GCGACCACCGAGATCTACAC-3’; *Reverse:* 5’-CAAGCAGAAGACGGCATACGAGAT-3’. Libraries were then constructed and quantified with a Quant-iT PicoGreen dsDNA Assay Kit using the published protocol. Several samples were put through an Agilent Bioanalyzer to confirm the size of the amplified region. After normalizing and pooling all libraries, the pool was re-quantified and the quality was checked on an Agilent Bioanalyzer. Finally, the samples were sequenced in one lane of an Illumina Rapid HiSeq sequencer.

Bioinformatics

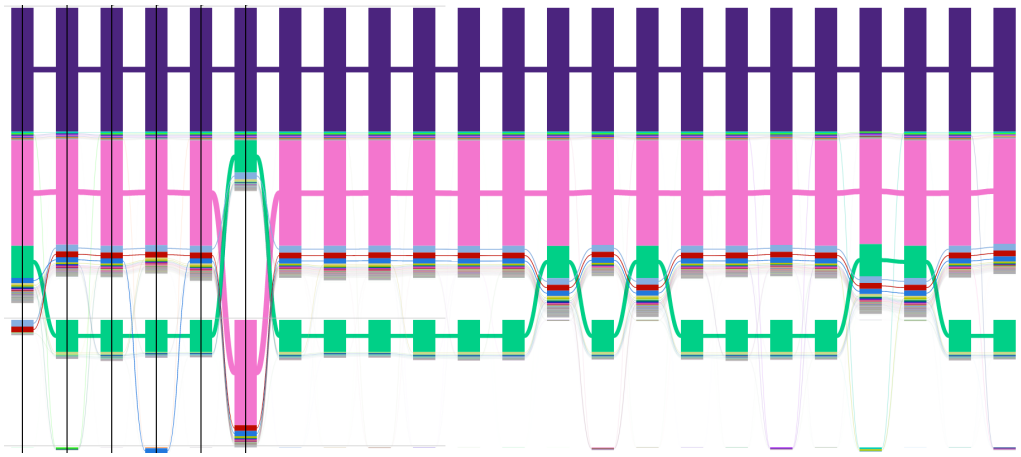
The sequenced samples were de-multiplexed using the two sets of indices. The adapters where then trimmed and the overlapping paired end reads merged. The design of the primers resulted in a >100bp overlap of the paired end reads and the merging of these reads improved coverage at the 3’ end of each read, which typically has a higher error rate. The reads were then mapped to the reference genome (*E. coli* K-12 Sub-strain MG1655, GenBank accession number: GCA_000005845.2) using BWA and sorted with Samtools. Indel realignment was conducted with GATK. Several samples were pulled out to assess their coverage and visualize their nucleotide variation using the program HapFlow version 1.1.2 (Sullivan et al. 2015). HapFlow identifies the variant profiles for all reads in a mixed-haplotype sample and then creates a flow across the 294-bp locus for each unique variant. Each variant was required to be present in at least 10% of the reads in order for it to be considered a real variant as opposed to sequence error.

Results:

Over 98% of extracted fecal samples were considered high quality and used in the remaining steps of this analysis. I confirmed the successful amplification and clean-up of 273 fecal samples (8-10 per lemur over 6 months) and submitted them for sequencing. Of the submitted samples, 256 resulted in promising libraries. The Agilent Bioanalyzer confirmed that sequences of the anticipated size were successfully amplified. After sequencing, nearly all reads aligned to the reference genome.

The following results are preliminary and intended to demonstrate the success of the sequencing and potential usefulness of this method. The average coverage for nine randomly selected samples was 297,649.44X. Furthermore, the variation in two samples (pk247 and pp121) was visualized via the program HapFlow (figure 1). Several observations can be made from figure 1. First, there were 12 and 23 polymorphic sites identified for sample pp121 (figure 1B) and pk247 (figure 1A), respectively (assuming a per-base minimum variant frequency of 10%). Second, the large number of colored bands indicates that there were hundreds of distinct haplotypes within each of these samples. Third, the pattern of haplotype frequencies, indicated by the width of the colored bands, suggests that one or two haplotypes were present at high frequency within a single fecal sample and most were present at very low frequency.

A.



B.

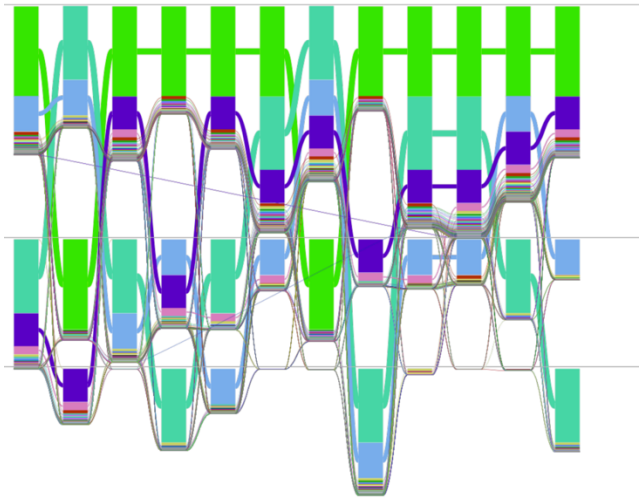


Figure 1. Haplotype variation within a fecal sample, pk247 (A) and pp121 (B). The program HapFlow was used to visualize and identify the variant profiles for all reads in a mixed-haplotype sample and then creates a flow across the 294-bp locus for each unique variant. Each variant was required to be present in at least 10% of the reads in order for it to be considered a real variant as opposed to sequence error. Each column of bars represents a nucleotide position that has variation. The colors represent unique haplotypes. The width of the colored bars represents the number of reads for each haplotype. Flows containing only the major allele group in the top row, flows containing the second most common group in the second row, flows containing the least common allele group at the bottom. mixed flows switch among levels (Sullivan et al. 2015).

Discussion:

The one-locus, culture-independent method described in Chapter 3 and applied to ringtailed lemur fecal samples in this paper, appears to have successfully targeted a region of high variation in the *E. coli* genome. Most fecal samples were successfully extracted, amplified, and sequenced. A randomly-selected subset of these samples have demonstrated exceedingly high coverage, which offers power to identify the nucleotide diversity in even rare haplotypes with high confidence. Attaining high coverage for rare haplotypes is a major challenge in whole genome sequencing (e.g. Bachmann et al. 2015), which is the only existing methodology to assess the nucleotide diversity of mixed-haplotype samples. The ability to determine the sequences of most haplotypes in a fecal sample with high confidence allows a much more fine-grained analysis of transmission dynamics than has previously been possible.

The within-sample haplotype variation identified in the two samples analyzed reflects several aspects of known population genetics of *E. coli*. First, the number of polymorphic sites was on the order of what was anticipated based on the variation in this region of the *E. coli* genome detected in the 43 FumC haplotypes reported in Chapter 3 (22 polymorphic sites). Furthermore, I observed that both samples had one or two major haplotypes at high frequency and many minor or transient haplotypes at low frequency. This major/minor haplotype dynamic is well-established in *E. coli* populations within the mammalian gut. Specifically, one haplotype tends to be present at high frequency and may remain for several months or years (i.e. the major haplotype) (Caugant et al. 1981, 1984). This haplotype is typically unique among individuals and is not influenced by genetic relatedness. Because novel isolates do not easily colonize the gut, individuals tend to carry many transient haplotypes that turn over after a few days or weeks, leading to a pattern of many low-frequency haplotypes (Caugant et al. 1981, Sears et al. 1950,

1956, Sears and Brownlee 1952). For example, in an individual sampled twice a month for 11 months, the number of transient strains varied from 0-11 per sample and the rate of turnover ranged between 2-4 weeks (Caugant et al. 1981). Furthermore, transient haplotypes were not genetically similar to those inhabiting the gut in the previous month, indicating that new haplotypes were the result of successive invasions from the environment or other individuals and not mutation (Caugant et al. 1981). Because the major haplotype acts as a "bacterial identifier" of an individual, this major/minor haplotype dynamic can possibly be used to identify the directionality of transmission. Specifically, directionality of transmission may be evident when an individual's major haplotype is found as a minor haplotype in an individual with whom it interacts.

Additionally, and most importantly, the results reported in this paper show a very high haplotype diversity (>100 haplotypes) within each of the two fecal samples analyzed. In fact, this within-sample diversity is far larger than has ever been reported within a single individual in any study that I know of. For example, Caugant et al. (1981) reported that out of 550 *E. coli* isolates sampled at multiple time points from a single individual, the most electrophoretic types (ETs) identified at a single point in time was 13. Further, Caugant et al. (1984) reported that out of 650 *E. coli* isolates collected from 28 individuals, the most ETs identified within a single sample was 11 with a mean of 2.3. More recently, Springer et al. (2016) reported that out of 83 *E. coli* isolates collected from 39 individuals, the maximum number of haplotypes identified within a single sample was 4. These studies all used culture-based methods (i.e. MLST and electrophoretic typing). Although there is much more work to do to confirm the functionality of the one-locus, culture independent method described in chapter 3, the hundreds of haplotypes that so far appear to be differentiated within each fecal sample in this study required a fraction of

the time and resources to generate. When this analysis is expanded to all focal individuals, it can provide astronomically more transmission information than traditional culture-based methods. Furthermore, the increase in the haplotype diversity identified should allow the determination of fine-scale patterns of transmission that have previously remained elusive. The few studies that have used *E. coli* as a model organism for the social transmission of bacteria (e.g. Caugant et al. 1984, Springer et al. 2016, VanderWaal et al. 2014a) have found evidence of transmission along social routes. However, low genetic variation of the pathogen, which will often be the case for *E. coli* when only small numbers of haplotypes can be identified, limits discriminatory power and hinders the ability to identify fine-scale and nuanced patterns of transmission (Craft 2015) such as the modes and frequencies of social behaviors that act to promote and prevent the spread of bacteria. This low discriminatory power does not appear to be a problem with the novel one-locus method described in Chapter 3 because it successfully differentiates hundreds of haplotypes.

The novel method for *E. coli* differentiation described in Chapter 3 and applied in this paper represents the first approach to differentiate a single species of bacteria at one-locus without the need to culture isolates and that is appropriate for high throughput sequencing. I report here that this method appears to be successful at differentiating a large diversity of haplotypes within a single fecal sample with very high coverage, potentially allowing the identification of even rare variants that have previously remained elusive. When coupled with detailed data on social contact patterns, routes of *E. coli* transmission can be identified and can demonstrate the types of social interactions that are sufficient for transmission to occur.

Literature Cited:

Alexander, R. D. 1974. The evolution of social behaviour. *Annual Review of Ecological Systems* 5: 325–383.

Bachmann, N. L., Sullivan, M. J., Jelocnik, M., Myers, G. S. A., Timms, P. and Polkinghorne, A. 2015. Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by *Chlamydia pecorum*. *Journal of Clinical Microbiology* 53: 1573–1581.

Blyton, M. D. J., Banks, S. C., Peakall, R., Lindenmayer, D. B. and Gordo, D. M. 2014. Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecology Letters* 17: 970–978.

Caugant, D. A., Levin, B. R. and Selander, R. K. 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* 98: 487-490.

Caugant, D. A., Levin, B. R. and Selander, R. K. 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. *Journal of Hygiene-Cambridge* 92: 377-384.

Cho, I. and Blaser, M. J. 2012. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics* 13: 260–270.

Cohen, T., van Helden, P. D., Wilson, D., Colijn, C., McLaughlin, M. M., Abubakar, I. and Warren, R. M. 2012. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clinical Microbiology Reviews* 25: 708–719.

Craft, M. E. 2015. Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370: 1–12.

Dubois, V., De Barbeyrac, B., Rogues, A. M., Arpin, C., Coulange, L., Andre, C., M'Zali, F., Megraud, F. and Quentin, C. 2010. CTX-M-producing *Escherichia coli* in a maternity ward: A likely community importation and evidence of mother-to-neonate transmission. *Journal of Antimicrobial Chemotherapy* 65: 1368–1371.

Eames, K., Bansal, S., Frost, S. and Riley, S. 2014. Six challenges in measuring contact networks for use in modelling. *Epidemics* doi:10.1016/j.epidem.2014. 08.006.

Eyre, D. W., Cule, M. L., Griffiths, D., Crook, D. W., Peto, T. E. A., Walker, S. A. and Wilson, D. J. 2013. Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. *PLoS Computational Biology* 9: e1003059. doi:10.1371/journal.pcbi.1003059.

Ezenwa, V. O., Gerardo, N. M., Inouye, D. W., Medina, M. and Xavier, J. B. 2012. Animal behavior and the microbiome. *Science* 338: 198-199.

Freeland, W. J. 1976. Pathogens and the evolution of primate sociality. *Biotropica* 8: 12–24.

Freeland, W. J. 1979. Primate social-groups as biological islands. *Ecology* 60: 719–728.

Gardy, J. L., Johnston, J. C., Sui, S. J. H., Cook, V. J., Shah, L., Brodkin, E., Rempel, S., Moore, R. Zhao, Y., Holt, R. et al. 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *New England Journal of Medicine* 364: 730–739.

Georghiou, P. R., Doggett, A. M., Kielofner, M. A., Stout, J. E., Watson, D. A., Lupski, J. R. and Hamill, R. J. 1994. Molecular fingerprinting of *Legionella* species by repetitive element PCR. *Journal of Clinical Microbiology* 32: 2989–2994.

Goldberg, T. L., Gillespie, T. R., Rwego, I. B., Wheeler, E., Estoff, E. L. and Chapman, C. A. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation* 135: 511–517.

Griffin, R. H. and Nunn, C. L. 2012. Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology* 26: 779–800.

Hamilton, W. D. and Zuk, M. 1982. Heritable true fitness in bright birds: a role for parasites? *Science* 218: 384-386.

Hartl, D. L. and Dykhuizen, D. E. 1984. The population genetics of *Escherichia coli*. *Annual Review of Genetics* 18: 31-68.

Heilig, H. G., Zoetendal, E. G., Vaughan, E. E., Marteau, P., Akkermans, A. D. and de Vos, W. M. 2002. Molecular diversity of *Lactobacillus spp*. and other lactic acid bacteria in the human intestine as determined by specific amplification of 16S ribosomal DNA. *Applied and Environmental Microbiology* 68: 114–23.

Hess, G. 1996. Disease in metapopulation models: implications for conservation. *Ecology* 77: 1617-1632.

Hoggard, S. J., Wilson, P. D., Beattie, A. J. and Stow, A. J. 2011. Social complexity and nesting habits are factors in the evolution of antimicrobial defenses in wasps. *PLoS One* 6: e21763. doi:10.1371/journal.pone.0021763.

Loudon, J. E. 2009. The parasite ecology and socioecology of ring-tailed lemurs (*Lemur catta)* and Verreaux's sifaka (*Propithecus verreauxi*) inhabiting the Beza Mahafaly Special Reserve. Ph.D. dissertation University of Colorado-Boulder, Boulder, CO.

McGrew, W. C., Tutin, C. E. G., Collins, D. A. and File, S. K. 1989. Intestinal parasites of sympatric *Pan troglodytes* and *Papio* spp. at two sites: Gombe (Tanzania) and Mt. Assirik (Senegal). *American Journal Primatology* 17: 147–155.

Michalak, K., Austin, C., Diesel, S., Bacon, J. M., Zimmerman, P. and Maslow, J. N. 1998. *Mycobacterium tuberculosis* infection as a zoonotic disease: Transmission between humans and elephants. *Emerging Infectious Diseases* 4: 283–287.

Møller, A. P., Merino, S., Brown, C. R. and Robertson, R. J. 2001. Immune defense and host sociality: A comparative study of swallows and martins. *American Naturalist* 158: 136–45.

Montoro, E., Valdivia, J., and Leao, S. C. 1998. Molecular fingerprinting *of Mycobacterium tuberculosis* isolates obtained in Havana, Cuba, by IS6110 restriction fragment length polymorphism analysis and by the double repetitive-element PCR method. *Journal of Clinical Microbiology* 36: 3099–3102.

Nunn, C. L., Altizer, S., Jones, K. E. and Schrest, W. 2003 Comparative tests of parasite species richness in primates. *American Naturalist* 162: 597–614.

Nunn, C. L., Gittleman, J. L. and Antonovics, J. 2000. Promiscuity and the primate immune system. *Science* 290: 1168–1170.

Odell, I. D. and Flavell, R. A. 2016. Microbiome: Ecology of eczema. *Nature Microbiology* 1: 16135.

Ressel, S. and Schall, J. J. 1989. Parasites and showy males: Malarial infection and color variation in fence lizards. *Oecologia* 78: 158–164.

Rewgo, I. B., Isabirye-Basuta, G., Gillespie, T. R. and Goldberg, T. 2008. Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600–1607.

Rothenberg, R. B., Woodhouse, D. E., Potterat, J. J., Muth, S. Q., Darrow, W. W. and Klovdahl, A. S. 1995. Social networks in disease transmission: The Colorado Springs study. In: National Institute on Drug Abuse: Research Monograph Series (Ed. by R. H. Needle, S. L. Coyle, S. G. Genser, R. T. Trotter), pp. 3–19. National Institutes of Health.

Rubenstein, D. I. and Hohmann, M. E. 1989. Parasites and social behavior of island feral horses. *Oikos* 55: 312–320.

Sampson, T. R. and Mazmanian, S. K. 2015. Control of brain development, function, and behavior by the microbiome. *Cell Host and Microbe* 17: 565–576.

Sears, H. J. and Brownlee, I. 1952. Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of Bacteriology* 63: 47–57.

Sears, H. J., Brownlee, I. and Uchlyama, J. K. 1950. Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of Bacteriology*. 59: 293–201.

Sears, H. J., Janes, H., Saloum, R., Brownlee, I. and Lamoreaux, L. F. 1956. Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *Journal of Bacteriology* 71: 370–372.

Simmons, L. W. 1990. Post-copulatory guarding, mate choice and the levels of gregarine infections in the field cricket, *Gryllus bimaculatus*. *Behavioral Ecology and Sociobiology* 26: 403–407.

Springer, A., Mellmann, A., Fichtel, C. and Kappeler, P. M. 2016. Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecology* 16: 1–6.

Sullivan, M. J., Bachmann, N. L., Timms, P., and Polkinghorne, A. 2016. HapFlow: Visualizing haplotypes in sequencing data. *Bioinformatics* 32: 441–443.

Taylor, N. S., Fox, J. G., Akopyants, N. S., Berg, D. E., Thompson, N., Shames, B., Yan, L., Fontham, E., Janney, F., Hunter, F. M., et al. 1995. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA-fingerprinting. *Journal of Clinical Microbiology* 33: 918–923.

Tinsley, R. C. 1990. The influence of parasite infection on mating success in spadefoot toads, *Scaphiopus couchii. American Zoologist* 30: 313–324.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014a. Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *Journal of Animal Ecology* 83: 406–414.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014b. Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation* 169: 113–146.

Wilson, K., Knell, R., Boots, M. and Koch-Osborne, J. 2003. Group living and investment in immune defense: An interspecific analysis. *Journal of Animal Ecology* 72: 133–143.

Whitman, N. K. and Parker, P. G. 2004. Effects of host sociality on ectoparasite population biology. *Journal of Parasitology* 90: 939–947.

Conclusion

Understanding the health and disease of vulnerable species is an essential component of conserving wild populations (Daszak et al. 2001, Dobson and May 1986). Merging data on the social behavior of the host and the genetics of the pathogen generates the most precise dataset for the identification of the dynamics of socially facilitated transmission (Craft 2015). The addition of a one-locus, culture-independent approach to the differentiation of multi-haplotype infections substantially improves the resolution with which these dynamics can be observed and allows for the discovery of more nuanced, fine-grained patterns of pathogen movement than is possible with traditional culture-based methods. The development of such a method for *E. coli*, a typically non-pathogenic bacterium that is ubiquitous in mammals, offers the ability to establish the nature of bacterial movement among all individuals in nearly any mammalian species at all times, regardless of the host's health status. The *E. coli* differentiation method I developed and tested in this dissertation is the first one-locus, culture-independent approach to haplotype-level differentiation of a single bacterial species to date. Using this novel approach, it becomes unnecessary to treat the advancement of our understanding of pathogen transmission as a primarily reactionary endeavor to incidences disease outbreaks. Rather, we can determine the types and frequencies of interactions that are sufficient to facilitate the transmission of pathogens within and among species using real bacterial transmission dynamics regardless of the infection status of the host (Rushmore et al. 2017). In doing so, these data can potentially inform predictive models for the future spread of pathogens and identify vulnerable social groups or species before the occurrence of transmission events that could severely threaten species'

survival. Predictive disease modeling is considered to be a critical tool in the conservation of vulnerable species (Epstein 2009, Junge and Sauther 2006) because understanding how diseases move is essential for reducing their impact (Patz et al. 2004). Much of the world's biota suffer from habitat loss, fragmentation and degradation, which increases the risk of exposure to pathogens (Bublitz et al. 2015, Daszak et al. 2001, Gillespie and Chapman 2006, Smith et al. 2008, Wright et al. 2009), including novel pathogens to which species may not be able to adapt at a fast enough rate. As a result, emerging infectious diseases are a global problem (Daszak et al. 2001) and contribute to species extinctions (Smith et al. 2006). This dissertation presents genetic and analytical methodology that can contribute to a growing understanding of pathogen transmission dynamics in real social groups.

Complete Literature Cited

Albers, P. C. H. and de Vries, H. 2001. Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour* 61: 489–95.

Alexander, R. D. 1974. The evolution of social behaviour. *Annual Review of Ecological Systems* 5: 325–383.

Altmann, J. 1974. Observational study of behavior: Sampling methods. *Behaviour* 49: 227–267.

Altmann, S. A. 1962. A field study of the sociobiology of rhesus monkeys *Macaca mulatta*. *Annals of the New York Academy of Sciences* 102: 338–435.

Altizer, S. Dobson, A., Hosseini, P., Hudson, P. Pascual, M. and Rohani, P. 2006. Seasonality and the dynamics of infectious diseases. *Ecology Letters* 9: 467–84.

Ansmann, I. C., Parra, G. J. Chilvers, B. L. and Lanyon, J. M. 2012. Dolphins restructure social system after reduction of commercial fisheries. *Animal Behaviour* 84: 575–81

Appleby, M. C. 1983. The probability of linearity in hierarchies. *Animal Behaviour* 31: 600-608.

Archie, E. A. and Theis, K. R. 2011. Animal behaviour meets microbial ecology. *Animal Behavior* 82: 425–436.

Bachmann, N. L., Sullivan, M. J., Jelocnik, M., Myers, G. S. A., Timms, P. and Polkinghorne, A. 2015. Culture-independent genome sequencing of clinical samples reveals an unexpected heterogeneity of infections by *Chlamydia pecorum*. *Journal of Clinical Microbiology* 53: 1573–1581.

Bartz, J. A. and Hollander, E. 2006. The neuroscience of affiliation: Forging links between basic and clinical research on neuropeptides and social behavior. *Hormones & Behavior* 50: 518–528.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B. Grothendieck, G. and Green, P. 2016. Linear mixed-effects models using 'Eigen' and S4. V. 1.1–12.

Bergstrom, M. L., and Fedigan, L. M. 2010. Dominance among female white-faced capuchin monkeys (*Cebus capucinus*): hierarchical linearity, nepotism, strength and stability. *Behaviour* 147: 899-931.

Bernstein, I. S. 1981. Dominance: The baby and the bathwater. *Behavioral and Brain Sciences* 4: 419-457.

Birky, W. A. and Su, H. H. 2005. Temporal variation in adult affiliation patterns in wild Taiwanese macaques (*Macaca cyclopis*). *International Journal of Primatology* 26: 1251–66.

Blonder, B. and Dornhaus, A. 2011. Time-ordered networks reveal limitations to information flow in ant colonies. *PLoS ONE* 6: 1–9.

Blyton, M. D. J., Banks, S. C., Peakall, R., Lindenmayer, D. B. and Gordo, D. M. 2014. Not all types of host contacts are equal when it comes to *E. coli* transmission. *Ecology Letters* 17: 970–978.

Boesch, C. 1991. The effects of leopard predation on grouping patterns in forest chimpanzees. *Behaviour* 117: 220–41.

Boyland, N. K., Mlynski, D. T., Richard, J., Brent, J. N. and Croft, D. P. 2016. The social network structure of a dynamic group of dairy cows: From individual to group level patterns. *Applied Animal Behaviour Science* 174: 1–10.

Brockman, D. K., Godfrey, L. R., Dollar, L. J., and Ratsirarson, J. 2008. Evidence of invasive *Felis silvestris* predation on *Propithecus verreauxi* at Beza Mahafaly Special Reserve, Madagascar. *International Journal of Primatology* 29: 135–52.

Bublitz, B. C., Wright, P. C. Rasambainarivo, F. T., Arrigo-Nelson, S. J., Bodager, J. R. and Gillespie, T. R. 2015. Pathogenic enterobacteria in lemurs associated with anthropogenic disturbance. *American Journal of Primatology* 77: 330–37.

Budnitz, N. and Dainis, K. 1975. *Lemur catta*: Ecology and Behavior. In: Lemur Biology (Ed. by I. Tattersall and R. W. Sussman), pp. 219-235. New York and London: Plenum Press.

Carne, C., Semple, S., Morrogh-Bernard, H., Zuberbühler, K. and Lehmann, J. 2013. Predicting the vulnerability of great apes to disease: the role of superspreaders and their potential vaccination. *PLoS One* 8: e84642.

Caugant, D. A., Levin, B. R. and Selander, R. K. 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* 98: 487-490.

Caugant, D. A., Levin, B. R. and Selander, R. K. 1984. Distribution of multilocus genotypes of *Escherichia coli* within and between host families. *Journal of Hygiene-Cambridge* 92: 377-384.

Cho, I. and Blaser, M. J. 2012. The human microbiome: at the interface of health and disease. *Nature Review Genetics* 13: 260–270.

Cluttonbrock, T. H. and Harvey, P. H. 1977. Primate ecology and social organization. *Journal of Zoology* 183: 1–39.

Cohen, T., van Helden, P. D., Wilson, D., Colijn, C., McLaughlin, M. M., Abubakar, I. and Warren, R. M. 2012. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clinical Microbiology Reviews* 25: 708–719.

Corner, L. A. L., Pfeiffer, D. U. and Morris, R. S. 2003. Social-network analysis of Mycobacterium bovis transmission among captive brushtail possums (*Trichosurus vulpecula*). *Preventative Veterinary Medicine* 59: 147–167.

Craft, M. E. 2015. Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 370: 1–12.

Crook, J. H. and Gartlan, J. S. 1966. Evolution of primate societies. *Nature* 210: 1200–1203.

Csardi, G. and Nepusz, T. 2006. The igraph software package for complex network research. InterJournal, Complex Systems 1695. http://igraph.sf.net.

Curley, J. P. 2106. Compete: Analyzing competitive interaction data (R package, version 0.1), https://cran.r-project.org/web/packages/compete/index.html.

Daszak, P., Cunningham, A. A. and Hyatt, A. D. 2001. Anthropogenic environmental change and the emergene of infectious diseases in wildlife. *Acta Tropica 78*: 103–116.

Dethlefsen, L., Huse, S., Sogin, M. L., Relman, D. A. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology* 6: e280.

Dias, R. C. S., Moreira, B. M., and Riley, L. W. 2010. Use of fimH single-nucleotide polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. *Journal of Clinical Microbiology* 48: 483-488.

Dobson, A. P. and May, R. M. 1986. Disease and conservation. In: Conservation Biology - The Science of Scarcity and Diversity (Ed. by M. Soule) pp. 345-365. Sunderland, MA: Sinauer Associates.

Dubois, V., De Barbeyrac, B., Rogues, A. M., Arpin, C., Coulange, L., Andre, C., M'Zali, F., Megraud, F. and Quentin, C. 2010. CTX-M-producing *Escherichia coli* in a maternity ward: a likely community importation and evidence of mother-to-neonate transmission. *Journal of Antimicrobial Chemotherapy* 65: 1368–1371.

Eames, K., Bansal, S., Frost, S. and Riley, S. 2014. Six challenges in measuring contact networks for use in modelling. *Epidemics* doi:10.1016/j.epidem.2014.08.006.

Elo, A. E. 1978. The Rating of Chess Players, Past and Present. New York: Arco.

Epstein, J. M. 2009. Modelling to contain pandemics. *Nature* 460: 687–689.

Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Research* 8: 186–194.

Eyre, D. W., Cule, M. L., Griffiths, D., Crook, D. W., Peto, T. E. A., Walker, S. A. and Wilson, D. J. 2013. Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. *PLoS Computational Biology* 9: e1003059. doi:10.1371/journal.pcbi.1003059.

Ezenwa, V. O. 2004. Host social behavior and parasitic infection: A multifactorial approach. *Behavioral Ecology* 15: 446–54.

Foster, J. A., Krone, S. M. and Forney, L. J. 2008. Application of ecological network theory to the human microbiome. *Interdisciplinary Perspectives on Infectious Diseases* 2008: 1–6.

Freeland, W. J. 1976. Pathogens and the evolution of primate sociality. *Biotropica* 8: 12–24.

Freeland, W. J. 1979. Primate social-groups as biological islands. *Ecology* 60: 719–728.

Friant, S., Ziegler, T. E. and Goldberg, T. L. 2016. Primate reinfection with gastrointestinal Parasites: Behavioural and physiological predictors of parasite acquisition. *Animal Behaviour* 117: 105–13.

Fruchterman, T. M. J. and Reingold, E. M. 1991. Graph drawing by force-directed placement software. *Practice and Experience* 21: 1129–1164.

Gabriel, D. N., Gould, L. and Kelley, E. A. 2014. Seasonal Patterns of Male Affiliation in Ring-Tailed Lemurs (*Lemur catta*) in Diverse Habitats across Southern Madagascar. *Behaviour* 151: 935–61.

Gardy, J. L., Johnston, J. C., Sui, S. J. H., Cook, V. J., Shah, L., Brodkin, E., Rempel, S., Moore, R. Zhao, Y., Holt, R. et al. 2011. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *New England Journal of Medicine* 364: 730–739.

Gemmill, A. and Gould, L. 2008. Microhabitat variation and its effects on dietary composition and intragroup feeding interactions between adult female *Lemur catta* during the dry season at Beza Mahafaly Special Reserve, southwestern Madagascar. *International Journal of Primatology* 29: 1511–33.

Georghiou, P. R., Doggett, A. M., Kielofner, M. A., Stout, J. E., Watson, D. A., Lupski, J. R. and Hamill, R. J. 1994. Molecular fingerprinting of *Legionella* species by repetitive element PCR. *Journal of Clinical Microbiology* 32: 2989–2994.

Gillespie, T. R. and Chapman, C. A. 2006. Prediction of parasite infection dynamics in primate metapopulations based on attributes of forest fragmentation. *Conservation Biology* 20: 441–448.

Goldberg, T. L., Gillespie, T. R., Rewgo, I. B., Wheeler, E., Estoff, E. L., and Chapman, C. A. 2007. Patterns of gastrointestinal bacterial exchange between chimpanzees and humans involved in research and tourism in western Uganda. *Biological Conservation* 135: 511–517.

Gould L. 1994. Patterns of affiliative behavior in adult male ring-tailed lemurs (*Lemur catta*) at the Beza-Mahafaly Reserve, Madagascar. Ph.D. dissertation. St. Louis: Washington University.

Gould, L. 1996. Male-female affiliative relationships in naturally occurring ringtailed lemurs (*Lemur catta*) at the Beza-Mahafaly Reserve, Madagascar. *American Journal of Primatology* 39: 63–78.

Griffin, R. H. and Nunn, C. L. 2012. Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology* 26: 779–800.

Hamede, R. K., Bashford, J., McCallum, H. and Jones, M. 2009. Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: Using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters* 12: 1147–57.

Hamilton, W. D. and Zuk, M. 1982. Heritable true fitness in bright birds: a role for parasites? *Science* 218: 384-386.

Harper, G. J, Steininger, M. K., Tucker, C. J., Juhn, D. and Hawkins, F. 2007. Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation* 34: 325–333.

Hartl, D. L. and Dykhuizen, D. E. 1984. The population genetics of *Escherichia coli*. *Annual Review of Genetics* 18: 31–68.

Heilig, H. G., Zoetendal, E. G., Vaughan, E. E., Marteau, P., Akkermans, A. D. and de Vos, W. M. 2002. Molecular diversity of *Lactobacillus spp*. and other lactic acid bacteria in the human intestine as determined by specific amplification of 16S ribosomal DNA. *Applied and Environmental Microbiology* 68: 114–23.

Hess, G. 1996. Disease in metapopulation models: Implications for conservation. *Ecology* 77: 1617–1632.

Hill, J. E, Seipp, R. P., Betts, M., Van Kessel, A. G., Crosby, W. L. Hemmingsen, S. M. and Hawkins, L. 2002. Extensive profiling of a complex microbial community by high-throughput sequencing. *Applied Environmental Microbiology* 68: 3055–3066.

Hinde, R. A. 1976. Interactions, relationships and social structure. Man 11: 1–17.

Hoff, J. K. 2009. The effect of sequencing errors on metagenomics gene prediction. *BMC Genomics* 10: 520. doi:10.1186/1471-2164-10-520.

Hoggard, S. J., Wilson, P. D., Beattie, A. J. and Stow, A. J. 2011. Social complexity and nesting habits are factors in the evolution of antimicrobial defenses in wasps. *PLoS One* 6: e21763. doi:10.1371/journal.pone.0021763.

Hoogland, J. L. 1979. Aggression, ectoparasitism, and other possible costs of prairie dog (*Sciuridae*, *Cynomys* spp.) coloniality. *Behaviour* 69: 1–35.

Hothorn, T., Bretz, F. Westfall, P. Heiberger. R. M. Schuetzenmeister, A. and Scheibe, S. 2016. Simultaneous inference in general parametric models (R package, version 1.4–6), http://multcomp.R-forge.R-project.org.

Huchard, E. and Cowlishaw, G. 2011. Female-female aggression around mating: An extra cost of sociality in a multimale primate society. *Behavioral Ecology* 22: 1003–11.

IUCN 2016. The IUCN Red List of Threatened Species. Version 2016-3. http://www.iucnredlist.org. Downloaded on 22 March 2017.

Isbell, L. A., Pruetz, J. D., Lewis, M., Young, T. P. 1999. Rank differences in ecological behavior: A comparative study of patas monkeys (*Erythrocebus patas*) and vervets (*Cercopithecus aethiops*). *International Journal of Primatology* 20: 257–272.

Jacoby, D. M. P., Brooks, E. J., Croft, D. P. and Sims, D. W. 2012. Developing a deeper understanding of animal movements and spatial dynamics through novel application of network analyses. *Methods in Ecology and Evolution* 3: 574–583.

Jaimez, N. A., Bribiescas, R. G., Aronsen, G. P., Anestis, S. A. and Watts, D. P. 2012. Urinary cortisol levels of gray-cheeked mangabeys are higher in disturbed compared to undisturbed forest areas in Kibale National Park, Uganda. *Animal Conservation* 15: 242–247.

Jarman, P. J. 1974. Social organization of antelope in relation to their ecology. *Behaviour* 48: 215–267.

Jeanson, R. 2012. Long-term dynamics in proximity networks in ants. *Animal Behaviour* 83: 915–923.

Jolly, A. 1966. Lemur social behavior and primate intelligence. *Science* 153: 501–506.

Junge, R. E. and Sauther, M. L. 2006. Overview on the health and disease ecology of wild lemurs: Conservation implications. In: Lemurs: Ecology and Adaptation (Ed. by L. Gould, M. L. Sauther) pp 425-442. New York: Springer.

Kappeler, P. M. 1990. Female dominance in *Lemur catta*: more than just female feeding priority? *Folia Primatologica* 55: 92–95.

Karczmarski, L., Wursig, B., Gailey, G., Larson, K. W., Vanderlip, C. 2005. Spinner dolphins in a remote Hawaiian atoll: Social grouping and population structure. *Behavioral Ecology* 16: 675–685.

Kelley, J. L., Morrell, L. J., Inskip, C., Krause, J. and Croft, D. P. 2011. Predation risk shapes social networks in fission-fusion populations. *PLoS ONE*: e24280. doi:10.1371/journal.pone.0024280.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T. and Ostfeld, R. S. 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* 468: 647-652.

Kim, D. A., Benjamin, E. J., Fowler, J. H. and Christakies, N. A. 2016. Social connectedness is associated with fibrinogen level in a human social network**.** *Proceedings of the Royal Society B* 83: 1–7.

Koyama, N. 1988. Mating behavior of ring-tailed lemurs (*Lemur catta*) at Berenty, Madagascar. *Primates* 29: 163–175.

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. 2016. lmerTest: Tests in linear mixed effects models (R package, version 2.0–33), https://cran.r-project.org/web/packages/lmerTest/index.html.

Loudon, J. E. 2009. The parasite ecology and socioecology of ring-tailed lemurs (*Lemur catta)* and Verreaux's sifaka (*Propithecus verreauxi*) inhabiting the Beza Mahafaly Special Reserve. Ph.D. dissertation University of Colorado-Boulder, Boulder, CO.

MacIntosh, A. J. J., Jacobs, A., Garcia, C. Shimizu, K., Mouri, K. Huffman, M. A. and Hernandez, A. D. 2012. Monkeys in the middle: Parasite transmission through the social network of a wild primate. *PLoS ONE* 7: 15–21.

Majolo, B., Lehmann, J., De Bortoli Vizioli, A. and Schino, G. 2012. Fitness-related benefits of dominance in primates. *American Journal of Physical Anthropology* 147: 652–60.

Martínez-Mota, R., Valdespino, C., Sánchez-Ramos, M. A. and Serio-Silva, J. C. 2007. Effects of forest fragmentation on the physiological stress response of black howler monkeys. *Animal Conservation* 10: 374–79.

Macedonia, J. M. 1993. Vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatologica* 61: 186–217.

McCauley, S. J. 2010. Body size and social dominance influence breeding dispersal in male *Pachydiplax longipennis* (*Odonata*). *Ecological Entomology* 35: 377–85.

McGrew, W. C., Tutin, C. E. G., Collins, D. A., and File, S. K. 1989. Intestinal parasites of sympatric *Pan troglodytes* and *Papio* spp. at two sites: Gombe (Tanzania) and Mt. Assirik (Senegal). *American Journal Primatology* 17: 147–155.

Michalak, K., Austin, C., Diesel, S., Bacon, J. M., Zimmerman, P., Maslow, J. N. 1998. *Mycobacterium tuberculosis* infection as a zoonotic disease: Transmission between humans and elephants. *Emerging Infectious Diseases* 4: 283–287.

Møller, A. P., Merino, S., Brown, C. R. and Robertson, R. J. 2001. Immune defense and host sociality: A comparative study of swallows and martins. *American Naturalist* 158: 136–45.

Montoro, E., Valdivia, J., and Leao, S. C. 1998. Molecular fingerprinting *of Mycobacterium tuberculosis* isolates obtained in Havana, Cuba, by IS6110 restriction fragment length polymorphism analysis and by the double repetitive-element PCR method. *Journal of Clinical Microbiology* 36: 3099–3102.

Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., Widdig, A. and Engelhardt, A. 2011. Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour* 82: 911–921.

Neumann, C. and Kulik, L. 2014. EloRating: Animal dominance hierarchies by Elo rating (R package, version 0.43), https://CRAN.R-project.org/package=EloRating.

Norscia, I. and Palagi, E. 2015. The socio-matrix reloaded: From hierarchy to dominance profile in wild lemurs. *PeerJ* 3: e729. doi:10.7717/peerj.729.

Nunn, C. L., Altizer, S., Jones, K. E., and Schrest, W. 2003 Comparative tests of parasite species richness in primates. *American Naturalist* 162: 597–614.

Nunn, C. L., Gittleman, J. L. and Antonovics, J. 2000. Promiscuity and the primate immune system. *Science* 290: 1168–1170.

Odell, I. D. and Flavell, R. A. 2016. Microbiome: Ecology of eczema. *Nature Microbiology* 1: 16135.

Ostner, J., Nunn, C. L. and Schulke, O. 2008. Female reproductive synchrony predicts skewed paternity across primates. *Behavioral Ecology* 19: 1150–1158.

Patz, J. A., Daszak, P., Tabor, G. M., Aguirre, A. A., Pearl, M., Epstein, J., Wolfe, N. D., Kilpatrick, A. M., Foufopoulos, J., Molyneux, D., and Bradley, D. J. 2004. Unhealthy Landscapes: Policy recommendations on land use change and infectious disease emergence. *Environmental Health Perspectives* 112: 1092–1098.

Pereira, M. E. 1993. Seasonal adjustment of growth rate and adult body weights in ringtailed lemurs. In: Lemur Social Systems and Ecological Basis (Ed. by P. M. Kappeler, J. U. Ganzhorn), pp 205–221. New York: Plenum Press.

Pereira, M. E. and Weiss, M. L. 1991. Female mate choice, male migration, and the threat of infanticide in ringtailed lemurs. *Behavioral Ecology and Sociobiology* 28: 141–152.

Pochron, S. T., Fitzgerald, J., Gilbert, C. C., Lawrence, D. Grgas, M., Rakotonirina, G., Ratsimbazafy, R., Rakotosoa, R. and Wright, C. W. 2003. Patterns of female dominance in *Propithecus diadema edwardsi* of Ranomafana National Park, Madagascar. *American Journal of Primatology* 61: 173–85.

Pusey, A., Williams, J. and Goodall, J. 1997. The influence of dominance rank on the reproductive success of female chimpanzees. *Science* 277: 828–831.

Rajilić-Stojanović, M., Smidt, H. and De Vos, W. M. 2007. Diversity of the human gastrointestinal tract microbiota revisited. *Environmental Microbiology* 9: 2125–2136.

Rasamimanana, H., Andrianome, V. N., Rambeloarivony, H. and Pasquet, P. 2006. Male and female ringtailed lemurs' energetic strategy does not explain female dominance. In: Ringtailed Lemur Biology (Ed. by A. Jolly, R. W. Sussman, N. Koyama, H. Rasamimanana), pp. 271–295. Chicago: Springer.

Ressel, S. and Schall, J. J. 1989. Parasites and showy males: Malarial infection and color variation in fence lizards. *Oecologia* 78: 158–164.

Rewgo, I. B., Isabirye-Basuta, G., Gillespie, T. R., Goldberg, T. 2008. Gastrointestinal bacterial transmission among humans, mountain gorillas, and livestock in Bwindi Impenetrable National Park, Uganda. *Conservation Biology* 22: 1600–1607.

Richard, A. F. 1992. Aggressive competition between males, female-controlled polygyny and sexual monomorphism in a Malagasy primate, *Propithecus verreauxi. Journal of Human Evolution* 22: 395–406.

Rimbach, R., Bisanzio, B., Galvis, N. Link, A, Di Fiore, A. and Gillespie T. R. 2015. Brown spider monkeys (*Ateles hybridus*): a model for differentiating the role of social networks and physical contact on parasite transmission dynamics. *Philosophical Transactions of the Royal Society B* 370: 20140110.

Romano, V., Duboscq, J., Sarabian, C., Thomas, E., Sueur, C. and MacIntosh, A, J. J. 2016. Modeling infection transmission in primate networks to predict centrality-based risk. *American Journal of Primatology* 78: 767–79.

Rothenberg, R. B., Woodhouse, D. E., Potterat, J. J., Muth, S. Q., Darrow, W. W. and Klovdahl, A. S. 1995. Social networks in disease transmission: The Colorado Springs study. In: National Institute on Drug Abuse: Research Monograph Series (Ed. by R. H. Needle, S. L. Coyle, S. G. Genser, R. T. Trotter), pp 3–19. National Institutes of Health.

Rubenstein, D. I. and Hohmann, M. E. 1989. Parasites and social behavior of island feral horses. *Oikos* 55: 312–320.

Rushmore, J., Bisanzio, D. and Gillespie, T. R. 2017. Making new connections: Insights from primate–parasite networks. *Trends in Parasitology* 1615: 1–14.

Rushmore, J., Caillaud, D., Matamba, L., Stumpf, R. M., Borgatti, S. P. and Altizer, A. 2013. Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. *Journal of Animal Ecology* 82: 976–86.

Sampson, T. R. and Mazmanian, S. K. 2015. Control of brain development, function, and behavior by the microbiome. *Cell Host and Microbe* 17: 565–576.

Sapolsky, R. M. 1983. Endocrine aspects of social instability in the olive baboon (*Papio anubis*). *American Journal of Primatology* 5: 365–379.

Sapolsky, R. M. 1992. Cortisol concentrations and the social significance of rank instability among wild baboons. *Psychoneuroendocrinology* 17: 701–702.

Sauther, M. L. 1991. Reproductive behavior of free-ranging *Lemur catta* at Beza Mahafaly Special Reserve, Madagascar. *American Journal of Physical Anthropology* 84: 463–77.

Sauther, M. L. 1992. Effects of reproductive state, social rank, and group size on resource use among free-ranging ringtailed lemurs (*Lemur catta*) of Madagascar. Dissertation, Washington University.

Sauther, M. L. 1993. Resource competition in wild Populations of ringtailed lemurs (*Lemur catta*): Implications for female dominance. In: Lemur Social Systems and their Ecological Basis (Ed. by P. M. Kappeler and J. U Ganzhorn), pp 135-152. New York: Plenum Press.

Sauther, M. L, Sussman, R. W. and Gould, L. 1999. The socioecology of the ringtailed lemur: Thirty-five years of research. *Evolutionary Anthropology* 8: 120–32.

Sbeglia, G. C., Tang-Martinez, Z. and Sussman, R. W. 2010. Effects of food, proximity, and kinship on social behavior in ringtailed lemurs. *American Journal of Primatology* 72: 981–91.

van Schaik, C. P. 1983. Why are diurnal primates living in groups? *Behaviour* 87: 120–144.

van Schaik, C. P. 1989. The ecology of social relationships amongst female primates. In: Comparative Socioecology: The behavioural ecology of humans and other animals. (Ed. by V. Standen and R. Foley), pp 195-218. Oxford: Blackwell Scientific.

van Schaik, J. and Kerth, G. 2017. Host social organization and mating system shape parasite transmission opportunities in three European bat species. *Parasitology Research*. 116: 589–599.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., Weber, C. F. 2009. Introducing Mothur: Open-

source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied Environmental Microbiology* 75: 7537–7541.

Sears, H. J. and Brownlee, I. 1952. Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of Bacteriology* 63: 47–57.

Sears, H. J., Brownlee, I. and Uchlyama, J. K. 1950. Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *Journal of Bacteriology*. 59: 293–201.

Sears, H. J., Janes, H., Saloum, R., Brownlee, I. and Lamoreaux, L. F. 1956. Persistence of individual strains of *Escherichia coli* in man and dog under varying conditions. *Journal of Bacteriology* 71: 370–372.

Sih, A., Hanser, S. F. and McHugh, K. A. 2009. Social network theory: new insights and issues for behavioral ecologists. *Behavioral Ecology and Sociobiology* 63: 975– 988.

Simmons, L. W. 1990. Post-copulatory guarding, mate choice and the levels of gregarine infections in the field cricket, *Gryllus bimaculatus*. *Behavioral Ecology and Sociobiology* 26: 403–407.

Smith, K. F., Acevedo-Whitehouse, K. and Pedersen, A. B. 2009. The role of infectious diseases in biological conservation. *Animal Conservation* 12: 1–12.

Smith, K. F., Sax, D. F. and Lafferty, K. D. 2006. Evidence for the role of infectious disease in species extinction and endangerment. *Conservation Biology* 20: 1349–1357.

Snyder-Mackler, N., Kohn, J. N., Barreiro, L. B., Johnson, Z. P., Wilson, M. E. and Tung, J. 2016. Social status drives social relationships in groups of unrelated female rhesus macaques. *Animal Behaviour* 111: 307-317.

Sokurenko, E. V., Chesnokova, V. Dykhuizen, D. E., Ofek, I., Wu, X., Krogfelt, K. A., Struve, C., Schembri, M. A. and Hasty, D. L. 1998. Pathogenic adaption of *Escherichia coli* by natural variation of the FimH adhesin. *PNAS* 95: 8922–8926.

Springer, A., Mellmann, A., Fichtel, C. and Kappeler, P. M. 2016. Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. *BMC Ecology* 16: 1-6.

Sterck, E. H. M., Watts, D. P. and van Schaik, C. P. 1997. The evolution of female social relationships in nonhuman primates. *Behavioral Ecology and Sociobiology* 41: 291-309.

Suau, A., Bonnet, R. Sutren, M., Gordon, J., Gibson, G. R., Collins, M. D., Doré, J. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Applied Environmental Microbiology* 65: 4799–4807.

Sullivan, M. J., Bachmann, N. L., Timms, P., Polkinghorne, A. 2016. HapFlow: Visualizing haplotypes in sequencing data. *Bioinformatics* 32: 441–443.

Sussman, R. W. 1991. Demography and social organization of free-ranging *Lemur catta* in the Beza Mahafaly Reserve, Madagascar. *American Journal of Physical Anthropology* 84:43–58.

Sussman, R. W. 1992. Male life-history and intergroup mobility among ringtailed lemurs (*Lemur catta*). *International Journal of Primatology* 13: 395–413.

Sussman, R. W. and Rakotozafy, A. 1994. Plant diversity and structural analysis of a tropical dry forest in southwestern Madagascar. *Biotropica* 26: 241–254.

Taylor, L. 1986. Kinship, dominance and social organization in a semi free-ranging group of ringtailed lemurs (*Lemur catta*). Ph.D. dissertation, Washington University, St. Louis, Missouri.

Taylor, N. S., Fox, J. G., Akopyants, N. S., Berg, D. E., Thompson, N., Shames, B., Yan, L., Fontham, E., Janney, F., Hunter, F. M., et al. 1995. Long-term colonization with single and multiple strains of *Helicobacter pylori* assessed by DNA-fingerprinting. *Journal of Clinical Microbiology* 33: 918–923.

Tinsley, R. C. 1990. The influence of parasite infection on mating success in spadefoot toads, *Scaphiopus couchii. American Zoologist* 30: 313–324.

Tung, J., Barreiro, L. B., Burns, M. B., Grenier, J. C., Lynch, J., Grieneisen, L. E., Altmann, J., Alberts, S. C., Blekhman, R. and Archie, E. A. 2015. Social networks predict gut microbiome composition in wild baboons. *eLife* 4: e05224. http://dx.doi.org/10.7554/eLife.05224.

Van Zalk, N. and Van Zalk, M. 2015. The importance of perceived care and connectedness with friends and parents for adolescent social anxiety. *Journal of Personality* 83: 346–60.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014a. Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). *Journal of Animal Ecology* 83: 406–414.

VanderWaal K. L., Atwill, E. R., Isbell, L, A. and McCowan, B. 2014b. Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation* 169: 113–146.

de Vries, H. 1995. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour* 50: 1375-1389.

Weissman, S. J., Chattopadhyay, S., Aprikian, P., Obata-Yasuoka, M., Yarova-Yarovaya, Y., Stapleton, A., Ba-Thein, W., Dykhuizen, D., Johnson, J.R., Sokurenko, E. V. 2006.

Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*. *Molecular Microbiology* 59: 975–988.

Weissman, S. J., Johnson, J. R., Tchesnokova, V., Billing, M., Dykhuizen, D., Riddell, K., Rogers, P., Qin, X., Butler-Wu, S., Cookson, B. T., Fang, F. C., Scholes, D., Chattopadhyay, S., Sokurenko, E. 2012. High-resolution two-locus clonal typing of Extraintestinal Pathogenic *Escherichia coli*. *Applied and Environmental Microbiology* 78: 1353–1360.

Whitman, N. K. and Parker, P. G. 2004. Effects of host sociality on ectoparasite population biology. *Journal of Parasitology* 90: 939–947.

Whitten, P. L. 1984. Competition among female vervet monkeys. In: *Female Primates* (Ed. by M. F. Small), pp 127–140. New York: Liss.

Wilson, K., Knell, R., Boots, M. and Koch-Osborne, J. 2003. Group living and investment in immune defense: an interspecific analysis. *Journal of Animal Ecology* 72: 133–143.

Wingfield J. C., Hegner, R. E., Dufty, R. E. and Ball, G. F. 1990. The challenge hypothesis: theoretical implications for patterns of testosterone secretion, mating systems, and breeding strategies. *American Naturalist* 136: 829–846.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves, P. R., Maiden, M. C., Ochman, H., Achtman, M. 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Molecular Microbiology* 60: 1136–1151.

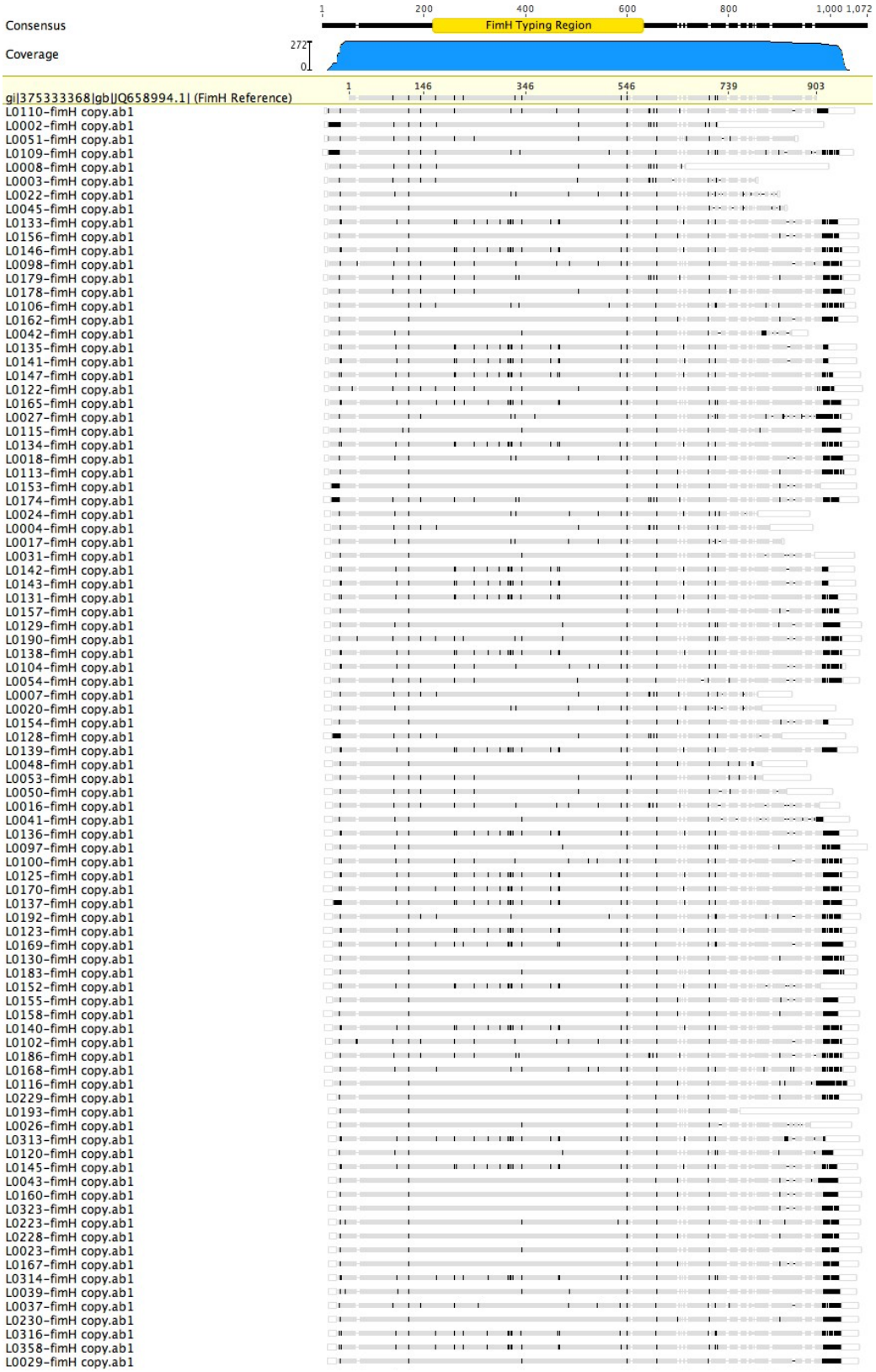
Wright, P. C. 1999. Lemur traits and Madagascar ecology: Coping with an island environment. *Yearbook of Physical Anthropology* 42: 31-72.

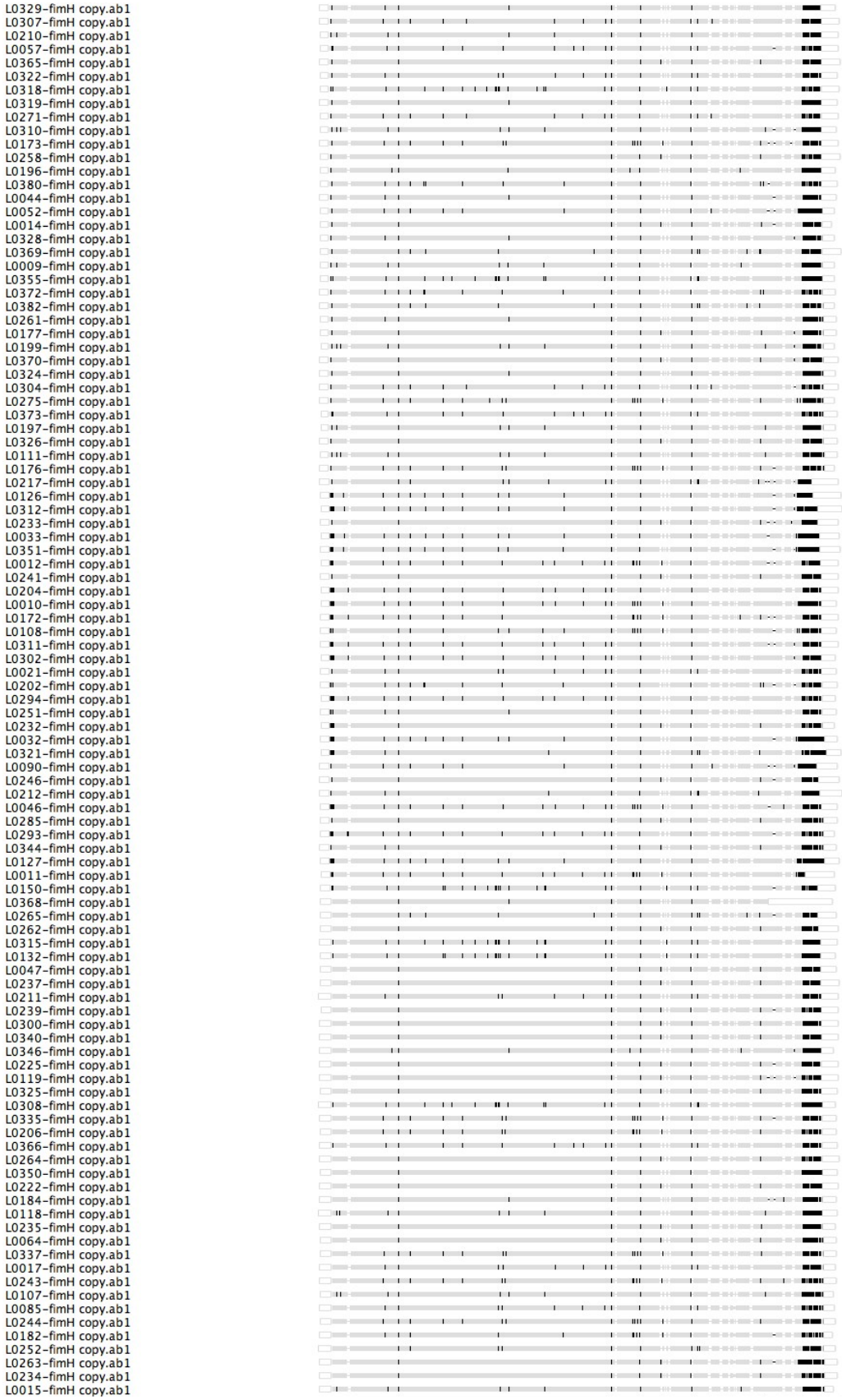
Wright, P. C., Arrigo-Nelson, S. J., Hogg, K. L. Bannon, B., Morelli, T. L., Wyatt, J. Harivelo, A. L. and Ratelolahy, F. 2009. Habitat disturbance and seasonal fluctuations of lemur parasites in the rain forest of Ranomafana National Park, Madagascar. In: Primate Parasite Ecology: The Dynamics and Study of Host-Parasite Relationships (Ed. by M.A. Huffman and C.A. Chapman), pp. 311-330. Cambridge: Cambridge University Press.

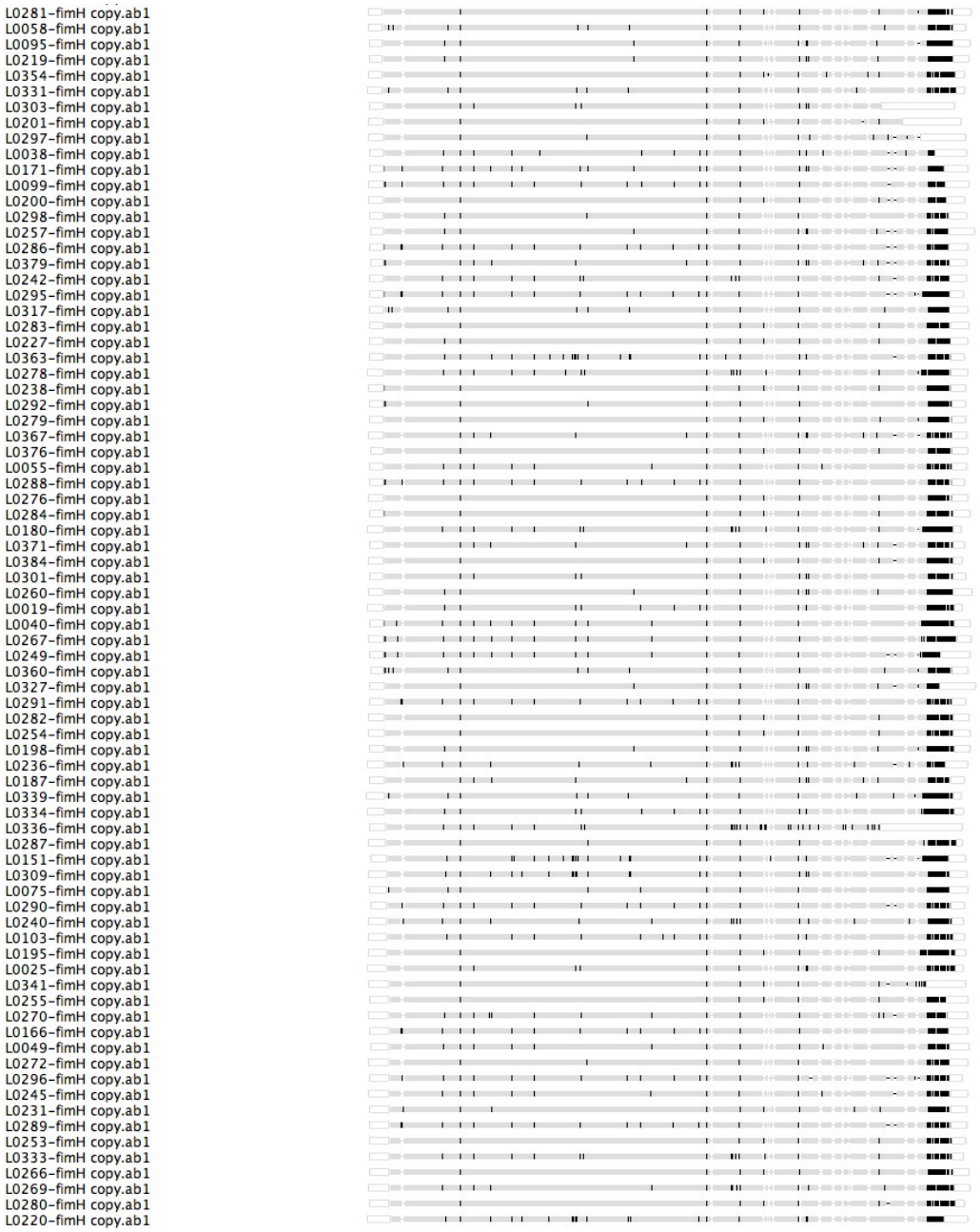
Zhu, Z. Ochoa, A. J. and Katzgraber, H. G. 2015. Efficient cluster algorithm for spin glasses in any space dimension. *Physical Review Letters* 115: 077201.

Zohdy, S., Kemp, A. D., Durden, L. A., Wright, P. C. and Jernvall, J. 2012. Mapping the social network: Tracking lice in a wild primate (*Microcebus rufus*) population to infer social contacts and vector potential. *BMC Ecology* 12: 4.

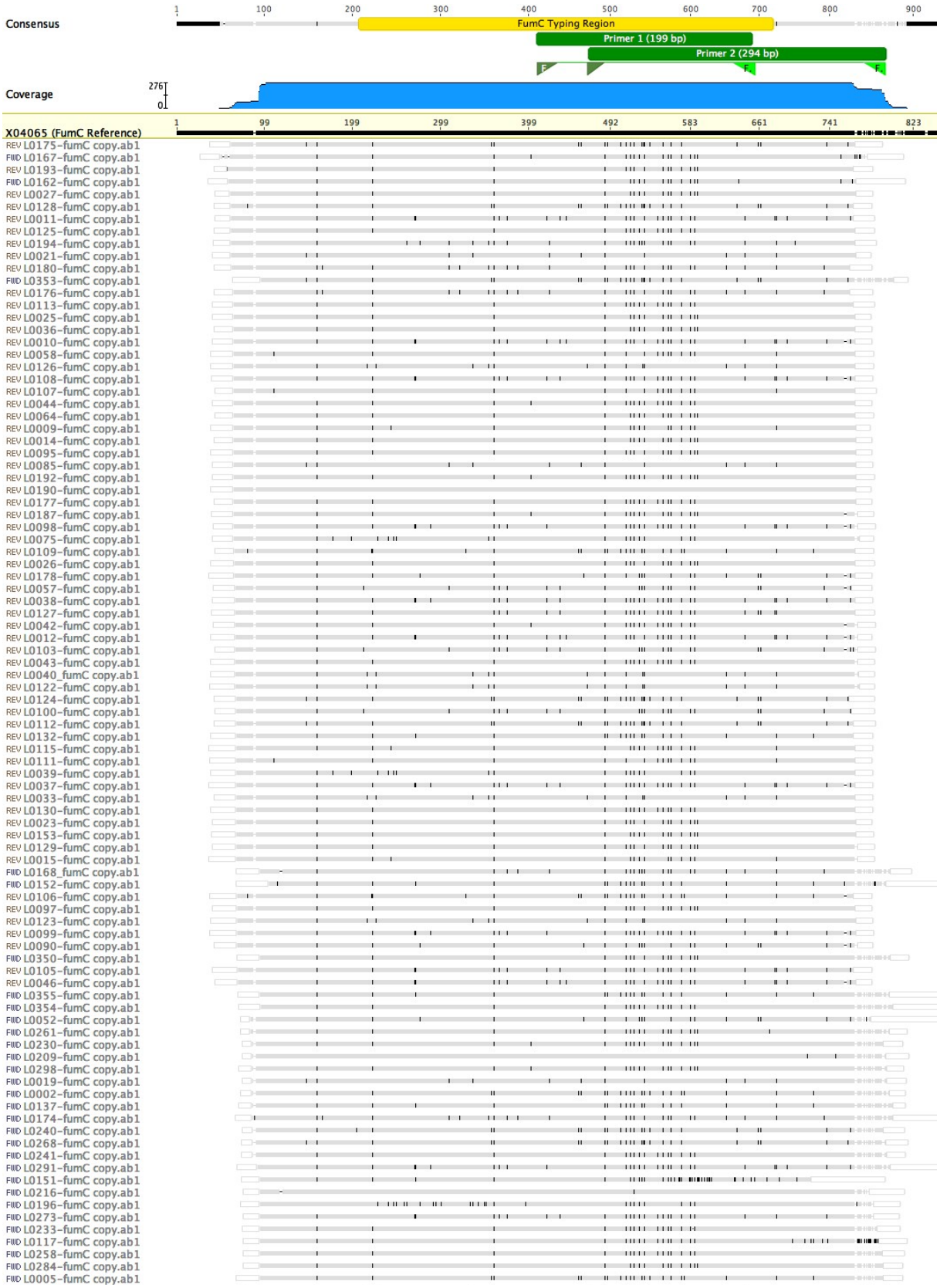
Appendix 1

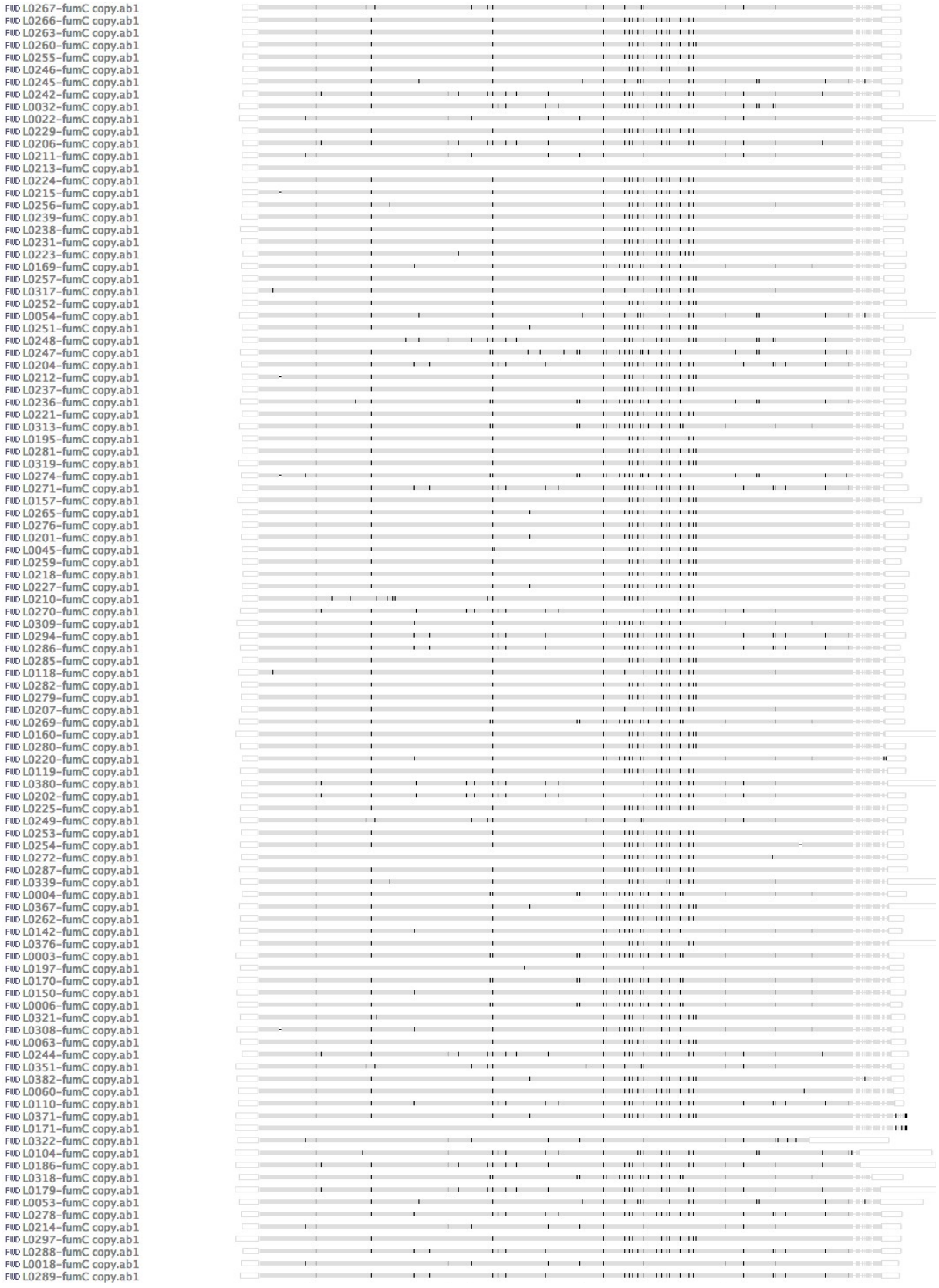

Consensus
1
200
400
600
800
1,000 1,072
FimH Typing Region
Coverage
272
0
gi|375333368|gb|JQ658994.1| (FimH Reference)
1
146
346
546
739
903
L0110-fimH copy.ab1
L0002-fimH copy.ab1
L0051-fimH copy.ab1
L0109-fimH copy.ab1
L0008-fimH copy.ab1
L0003-fimH copy.ab1
L0022-fimH copy.ab1
L0045-fimH copy.ab1
L0133-fimH copy.ab1
L0156-fimH copy.ab1
L0146-fimH copy.ab1
L0098-fimH copy.ab1
L0179-fimH copy.ab1
L0178-fimH copy.ab1
L0106-fimH copy.ab1
L0162-fimH copy.ab1
L0042-fimH copy.ab1
L0135-fimH copy.ab1
L0141-fimH copy.ab1
L0147-fimH copy.ab1
L0122-fimH copy.ab1
L0165-fimH copy.ab1
L0027-fimH copy.ab1
L0115-fimH copy.ab1
L0134-fimH copy.ab1
L0018-fimH copy.ab1
L0113-fimH copy.ab1
L0153-fimH copy.ab1
L0174-fimH copy.ab1
L0024-fimH copy.ab1
L0004-fimH copy.ab1
L0017-fimH copy.ab1
L0031-fimH copy.ab1
L0142-fimH copy.ab1
L0143-fimH copy.ab1
L0131-fimH copy.ab1
L0157-fimH copy.ab1
L0129-fimH copy.ab1
L0190-fimH copy.ab1
L0138-fimH copy.ab1
L0104-fimH copy.ab1
L0054-fimH copy.ab1
L0007-fimH copy.ab1
L0020-fimH copy.ab1
L0154-fimH copy.ab1
L0128-fimH copy.ab1
L0139-fimH copy.ab1
L0048-fimH copy.ab1
L0053-fimH copy.ab1
L0050-fimH copy.ab1
L0016-fimH copy.ab1
L0041-fimH copy.ab1
L0136-fimH copy.ab1
L0097-fimH copy.ab1
L0100-fimH copy.ab1
L0125-fimH copy.ab1
L0170-fimH copy.ab1
L0137-fimH copy.ab1
L0192-fimH copy.ab1
L0123-fimH copy.ab1
L0169-fimH copy.ab1
L0130-fimH copy.ab1
L0183-fimH copy.ab1
L0152-fimH copy.ab1
L0155-fimH copy.ab1
L0158-fimH copy.ab1
L0140-fimH copy.ab1
L0102-fimH copy.ab1
L0186-fimH copy.ab1
L0168-fimH copy.ab1
L0116-fimH copy.ab1
L0229-fimH copy.ab1
L0193-fimH copy.ab1
L0026-fimH copy.ab1
L0313-fimH copy.ab1
L0120-fimH copy.ab1
L0145-fimH copy.ab1
L0043-fimH copy.ab1
L0160-fimH copy.ab1
L0323-fimH copy.ab1
L0223-fimH copy.ab1
L0228-fimH copy.ab1
L0023-fimH copy.ab1
L0167-fimH copy.ab1
L0314-fimH copy.ab1
L0039-fimH copy.ab1
L0037-fimH copy.ab1
L0230-fimH copy.ab1
L0316-fimH copy.ab1
L0358-fimH copy.ab1
L0029-fimH copy.ab1

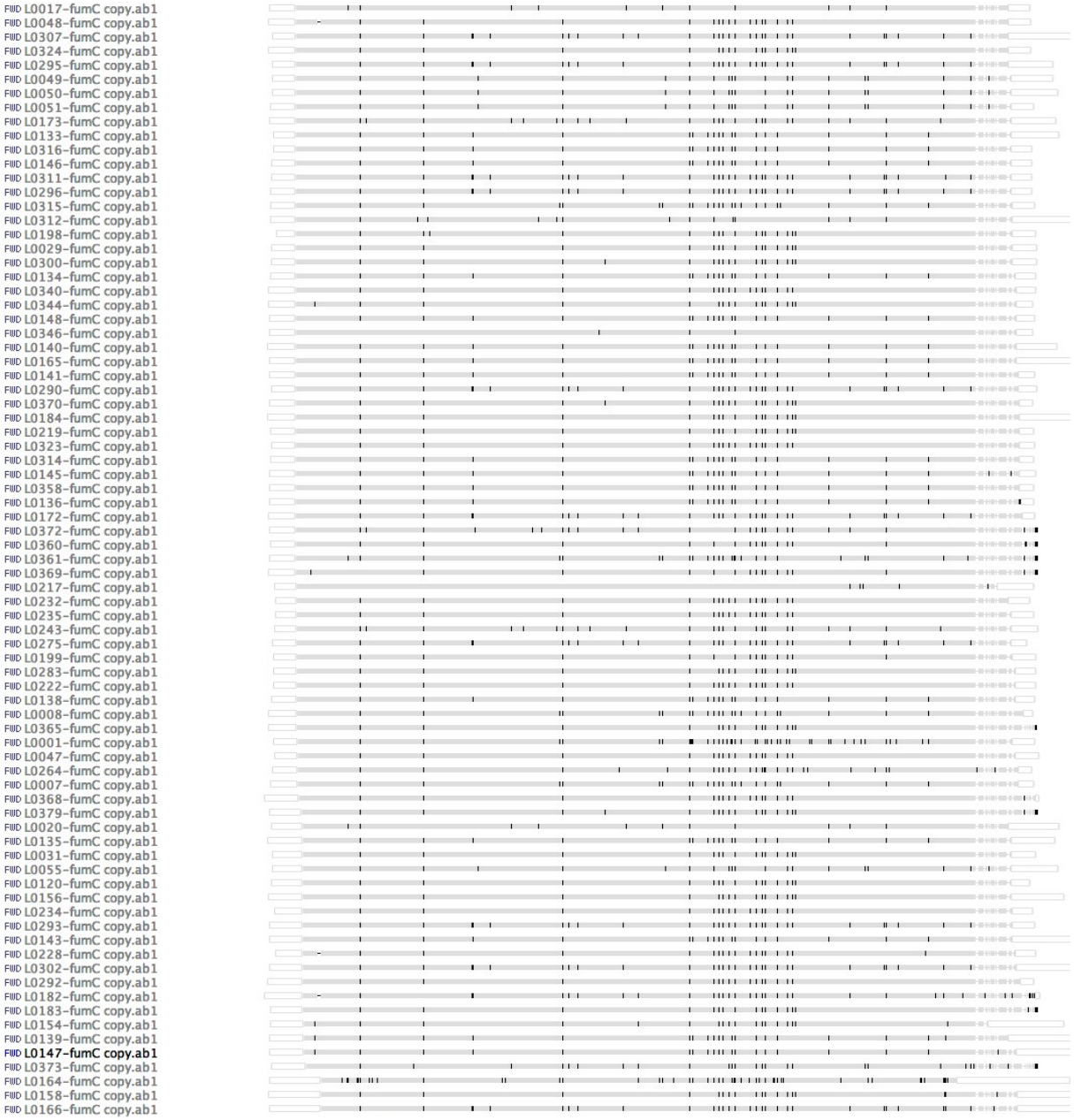

L0329-fimH copy.ab1
L0307-fimH copy.ab1
L0210-fimH copy.ab1
L0057-fimH copy.ab1
L0365-fimH copy.ab1
L0322-fimH copy.ab1
L0318-fimH copy.ab1
L0319-fimH copy.ab1
L0271-fimH copy.ab1
L0310-fimH copy.ab1
L0173-fimH copy.ab1
L0258-fimH copy.ab1
L0196-fimH copy.ab1
L0380-fimH copy.ab1
L0044-fimH copy.ab1
L0052-fimH copy.ab1
L0014-fimH copy.ab1
L0328-fimH copy.ab1
L0369-fimH copy.ab1
L0009-fimH copy.ab1
L0355-fimH copy.ab1
L0372-fimH copy.ab1
L0382-fimH copy.ab1
L0261-fimH copy.ab1
L0177-fimH copy.ab1
L0199-fimH copy.ab1
L0370-fimH copy.ab1
L0324-fimH copy.ab1
L0304-fimH copy.ab1
L0275-fimH copy.ab1
L0373-fimH copy.ab1
L0197-fimH copy.ab1
L0326-fimH copy.ab1
L0111-fimH copy.ab1
L0176-fimH copy.ab1
L0217-fimH copy.ab1
L0126-fimH copy.ab1
L0312-fimH copy.ab1
L0233-fimH copy.ab1
L0033-fimH copy.ab1
L0351-fimH copy.ab1
L0012-fimH copy.ab1
L0241-fimH copy.ab1
L0204-fimH copy.ab1
L0010-fimH copy.ab1
L0172-fimH copy.ab1
L0108-fimH copy.ab1
L0311-fimH copy.ab1
L0302-fimH copy.ab1
L0021-fimH copy.ab1
L0202-fimH copy.ab1
L0294-fimH copy.ab1
L0251-fimH copy.ab1
L0232-fimH copy.ab1
L0032-fimH copy.ab1
L0321-fimH copy.ab1
L0090-fimH copy.ab1
L0246-fimH copy.ab1
L0212-fimH copy.ab1
L0046-fimH copy.ab1
L0285-fimH copy.ab1
L0293-fimH copy.ab1
L0344-fimH copy.ab1
L0127-fimH copy.ab1
L0011-fimH copy.ab1
L0150-fimH copy.ab1
L0368-fimH copy.ab1
L0265-fimH copy.ab1
L0262-fimH copy.ab1
L0315-fimH copy.ab1
L0132-fimH copy.ab1
L0047-fimH copy.ab1
L0237-fimH copy.ab1
L0211-fimH copy.ab1
L0239-fimH copy.ab1
L0300-fimH copy.ab1
L0340-fimH copy.ab1
L0346-fimH copy.ab1
L0225-fimH copy.ab1
L0119-fimH copy.ab1
L0325-fimH copy.ab1
L0308-fimH copy.ab1
L0335-fimH copy.ab1
L0206-fimH copy.ab1
L0366-fimH copy.ab1
L0264-fimH copy.ab1
L0350-fimH copy.ab1
L0222-fimH copy.ab1
L0184-fimH copy.ab1
L0118-fimH copy.ab1
L0235-fimH copy.ab1
L0064-fimH copy.ab1
L0337-fimH copy.ab1
L0017-fimH copy.ab1
L0243-fimH copy.ab1
L0107-fimH copy.ab1
L0085-fimH copy.ab1
L0244-fimH copy.ab1
L0182-fimH copy.ab1
L0252-fimH copy.ab1
L0263-fimH copy.ab1
L0234-fimH copy.ab1
L0015-fimH copy.ab1



Supplemental Figure 1. Alignment of 272 FimH isolates.


Consensus
1
100
200
300
400
500
600
700
800
900
FumC Typing Region
Primer 1 (199 bp)
Primer 2 (294 bp)
F
F
F
Coverage
276
0
X04065 (FumC Reference)
1
99
199
299
399
492
583
661
741
823
REV L0175-fumC copy.ab1
FWD L0167-fumC copy.ab1
REV L0193-fumC copy.ab1
FWD L0162-fumC copy.ab1
REV L0027-fumC copy.ab1
REV L0128-fumC copy.ab1
REV L0011-fumC copy.ab1
REV L0125-fumC copy.ab1
REV L0194-fumC copy.ab1
REV L0021-fumC copy.ab1
REV L0180-fumC copy.ab1
FWD L0353-fumC copy.ab1
REV L0176-fumC copy.ab1
REV L0113-fumC copy.ab1
REV L0025-fumC copy.ab1
REV L0036-fumC copy.ab1
REV L0010-fumC copy.ab1
REV L0058-fumC copy.ab1
REV L0126-fumC copy.ab1
REV L0108-fumC copy.ab1
REV L0107-fumC copy.ab1
REV L0044-fumC copy.ab1
REV L0064-fumC copy.ab1
REV L0009-fumC copy.ab1
REV L0014-fumC copy.ab1
REV L0095-fumC copy.ab1
REV L0085-fumC copy.ab1
REV L0192-fumC copy.ab1
REV L0190-fumC copy.ab1
REV L0177-fumC copy.ab1
REV L0187-fumC copy.ab1
REV L0098-fumC copy.ab1
REV L0075-fumC copy.ab1
REV L0109-fumC copy.ab1
REV L0026-fumC copy.ab1
REV L0178-fumC copy.ab1
REV L0057-fumC copy.ab1
REV L0038-fumC copy.ab1
REV L0127-fumC copy.ab1
REV L0042-fumC copy.ab1
REV L0012-fumC copy.ab1
REV L0103-fumC copy.ab1
REV L0043-fumC copy.ab1
REV L0040_fumC copy.ab1
REV L0122-fumC copy.ab1
REV L0124-fumC copy.ab1
REV L0100-fumC copy.ab1
REV L0112-fumC copy.ab1
REV L0132-fumC copy.ab1
REV L0115-fumC copy.ab1
REV L0111-fumC copy.ab1
REV L0039-fumC copy.ab1
REV L0037-fumC copy.ab1
REV L0033-fumC copy.ab1
REV L0130-fumC copy.ab1
REV L0023-fumC copy.ab1
REV L0153-fumC copy.ab1
REV L0129-fumC copy.ab1
REV L0015-fumC copy.ab1
FWD L0168_fumC copy.ab1
FWD L0152-fumC copy.ab1
REV L0106-fumC copy.ab1
REV L0097-fumC copy.ab1
REV L0123-fumC copy.ab1
REV L0099-fumC copy.ab1
REV L0090-fumC copy.ab1
FWD L0350-fumC copy.ab1
REV L0105-fumC copy.ab1
REV L0046-fumC copy.ab1
FWD L0355-fumC copy.ab1
FWD L0354-fumC copy.ab1
FWD L0052-fumC copy.ab1
FWD L0261-fumC copy.ab1
FWD L0230-fumC copy.ab1
FWD L0209-fumC copy.ab1
FWD L0298-fumC copy.ab1
FWD L0019-fumC copy.ab1
FWD L0002-fumC copy.ab1
FWD L0137-fumC copy.ab1
FWD L0174-fumC copy.ab1
FWD L0240-fumC copy.ab1
FWD L0268-fumC copy.ab1
FWD L0241-fumC copy.ab1
FWD L0291-fumC copy.ab1
FWD L0151-fumC copy.ab1
FWD L0216-fumC copy.ab1
FWD L0196-fumC copy.ab1
FWD L0273-fumC copy.ab1
FWD L0233-fumC copy.ab1
FWD L0117-fumC copy.ab1
FWD L0258-fumC copy.ab1
FWD L0284-fumC copy.ab1
FWD L0005-fumC copy.ab1


FWD L0267-fumC copy.ab1
FWD L0266-fumC copy.ab1
FWD L0263-fumC copy.ab1
FWD L0260-fumC copy.ab1
FWD L0255-fumC copy.ab1
FWD L0246-fumC copy.ab1
FWD L0245-fumC copy.ab1
FWD L0242-fumC copy.ab1
FWD L0032-fumC copy.ab1
FWD L0022-fumC copy.ab1
FWD L0229-fumC copy.ab1
FWD L0206-fumC copy.ab1
FWD L0211-fumC copy.ab1
FWD L0213-fumC copy.ab1
FWD L0224-fumC copy.ab1
FWD L0215-fumC copy.ab1
FWD L0256-fumC copy.ab1
FWD L0239-fumC copy.ab1
FWD L0238-fumC copy.ab1
FWD L0231-fumC copy.ab1
FWD L0223-fumC copy.ab1
FWD L0169-fumC copy.ab1
FWD L0257-fumC copy.ab1
FWD L0317-fumC copy.ab1
FWD L0252-fumC copy.ab1
FWD L0054-fumC copy.ab1
FWD L0251-fumC copy.ab1
FWD L0248-fumC copy.ab1
FWD L0247-fumC copy.ab1
FWD L0204-fumC copy.ab1
FWD L0212-fumC copy.ab1
FWD L0237-fumC copy.ab1
FWD L0236-fumC copy.ab1
FWD L0221-fumC copy.ab1
FWD L0313-fumC copy.ab1
FWD L0195-fumC copy.ab1
FWD L0281-fumC copy.ab1
FWD L0319-fumC copy.ab1
FWD L0274-fumC copy.ab1
FWD L0271-fumC copy.ab1
FWD L0157-fumC copy.ab1
FWD L0265-fumC copy.ab1
FWD L0276-fumC copy.ab1
FWD L0201-fumC copy.ab1
FWD L0045-fumC copy.ab1
FWD L0259-fumC copy.ab1
FWD L0218-fumC copy.ab1
FWD L0227-fumC copy.ab1
FWD L0210-fumC copy.ab1
FWD L0270-fumC copy.ab1
FWD L0309-fumC copy.ab1
FWD L0294-fumC copy.ab1
FWD L0286-fumC copy.ab1
FWD L0285-fumC copy.ab1
FWD L0118-fumC copy.ab1
FWD L0282-fumC copy.ab1
FWD L0279-fumC copy.ab1
FWD L0207-fumC copy.ab1
FWD L0269-fumC copy.ab1
FWD L0160-fumC copy.ab1
FWD L0280-fumC copy.ab1
FWD L0220-fumC copy.ab1
FWD L0119-fumC copy.ab1
FWD L0380-fumC copy.ab1
FWD L0202-fumC copy.ab1
FWD L0225-fumC copy.ab1
FWD L0249-fumC copy.ab1
FWD L0253-fumC copy.ab1
FWD L0254-fumC copy.ab1
FWD L0272-fumC copy.ab1
FWD L0287-fumC copy.ab1
FWD L0339-fumC copy.ab1
FWD L0004-fumC copy.ab1
FWD L0367-fumC copy.ab1
FWD L0262-fumC copy.ab1
FWD L0142-fumC copy.ab1
FWD L0376-fumC copy.ab1
FWD L0003-fumC copy.ab1
FWD L0197-fumC copy.ab1
FWD L0170-fumC copy.ab1
FWD L0150-fumC copy.ab1
FWD L0006-fumC copy.ab1
FWD L0321-fumC copy.ab1
FWD L0308-fumC copy.ab1
FWD L0063-fumC copy.ab1
FWD L0244-fumC copy.ab1
FWD L0351-fumC copy.ab1
FWD L0382-fumC copy.ab1
FWD L0060-fumC copy.ab1
FWD L0110-fumC copy.ab1
FWD L0371-fumC copy.ab1
FWD L0171-fumC copy.ab1
FWD L0322-fumC copy.ab1
FWD L0104-fumC copy.ab1
FWD L0186-fumC copy.ab1
FWD L0318-fumC copy.ab1
FWD L0179-fumC copy.ab1
FWD L0053-fumC copy.ab1
FWD L0278-fumC copy.ab1
FWD L0214-fumC copy.ab1
FWD L0297-fumC copy.ab1
FWD L0288-fumC copy.ab1
FWD L0018-fumC copy.ab1
FWD L0289-fumC copy.ab1



Supplemental Figure 2. Alignment of 276 FumC isolates