Identifying Risk Factors of Anaplasma Infection in Plains Zebra of Etosha National Park, Namibia

Madison Stahle
IDENTIFYING RISK FACTORS OF *ANAPLASMA* INFECTION IN PLAINS ZEBRA

OF ETOSHA NATIONAL PARK, NAMIBIA

by

Madison H. Stahle

A Thesis Submitted in Partial Fulfillment
of the Requirements for a Degree with Honors
(Animal Veterinary Science)

The Honors College
University of Maine
May 2021

Advisory Committee:
Pauline Kamath, Assistant Professor of Animal Health, Advisor
Julie DellaMattera, Associate Professor of Early Childhood Development and Education
Sue Ishaq, Assistant Professor of Animal and Veterinary Studies
Melanie Prentice, Postdoctoral Research Associate, School of Food and Agriculture
James Weber, Associate Professor of Animal and Veterinary Sciences
ABSTRACT

Anaplasmosis is a tick-borne illness that is caused by bacteria from the genus *Anaplasma* which infect the blood cells of their hosts. Anaplasmosis can affect the health and life expectancy of zebra, however, little is known about what makes drives the variability in infection within this species. The purpose of this study is to (1) determine the prevalence, and (2) identify risk factors for *Anaplasma* infections in plains zebra (*Equus quagga*) from the Etosha National Park, Namibia, including whether *Anaplasma* infections correlate with other parasite infections (ectoparasites and gastrointestinal parasites). Other possible risk factors for *Anaplasma* infection that are investigated here include age, sex, and major histocompatibility (MHC) genotype. I used polymerase chain reaction to detect the presence of *Anaplasma* species in DNA extracted from blood samples. I used generalized linear models to determine the relationships between possible individual risk factors and infected individuals. Sanger sequencing was used to identify *Anaplasma* strains present in these samples to the species level, and a phylogenetic analysis was performed to characterize the genetic relationships between strains found in zebra with strains from other species, including livestock, humans, and ticks. This study revealed that 35.9% of ENP zebra were infected with *Anaplasma*, and out of the risk factors evaluated, only age was significantly associated with *Anaplasma* infection, showing that as age increased the probability of *Anaplasma* infection decreased. Sequencing data revealed that there were 3 variant strains, and the *Anaplasma* species in this population is 99.7% similar to *A. platys*. These findings will be important in helping researchers better understand the risks of *Anaplasma* infections in plains zebra.
and how these infections correlate with other parasite infections. These results could also be helpful in studying risk factors of other diseases in the population in the future.
ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Pauline Kamath and Melanie Prentice for the extensive time and guidance they have provided throughout this process. This project would not have been possible without the support and patience you both have given me. Thank you for all the time you have taken to help me learn and succeed.
# TABLE OF CONTENTS

Introduction 1

Anaplasmosis Overview 1

Study Area 1

Potential Risk Factors of Infection 2

Research Goals and Purpose 4

Methods 6

Field Sampling and Risk Factor Data Collection 6

*Anaplasma* Prevalence 6

Statistical Analyses 8

Sanger Sequencing 9

Results 10

Initial Results 10

Statistical Analyses Results 11

*Anaplasma* Sequencing Results 12

Discussion 14

Conclusion 20

Literature Cited 21

Appendix 25

Author’s Biography 35
LIST OF TABLES AND FIGURES

Tables

Table 1: Explanatory Variables used in the AIC Model Selection Process 25
Table 2: Explanatory Variables and Associated P-Values 26
Table 3: Phylogenetic Analysis Information 27

Figures

Figure 1: Gel of PCR 4 28
Figure 2: Gel of PCR 12 28
Figure 3: Gel of PCR 13 29
Figure 4: *Anaplasma* Infection Prevalence by Sex 29
Figure 5: *Anaplasma* Infection Prevalence by Species of Tick 30
Figure 6: *Anaplasma* Infection Prevalence in Heterozygotes 30
Figure 7: *Anaplasma* Infection Prevalence by Type of Allele 31
Figure 8: *Anaplasma* Infection Prevalence by Age 31
Figure 9: *Anaplasma* Infection Prevalence by Gastrointestinal Parasite Load 32
Figure 10: Age vs. Probability of *Anaplasma* Infection 32
Figure 11: Gastrointestinal parasite load vs. Probability of *Anaplasma* Inf. 33
Figure 12: *H. truncatum* ticks vs. Probability of *Anaplasma* Infection 33
Figure 13: Phylogenetic Tree 34
INTRODUCTION

Anaplasmosis Overview

Anaplasmosis is a hemoparasitic disease that is caused by intracellular bacterial parasites from the genus *Anaplasma* (Weny et al. 2017) which infect the blood cells of their hosts. There are many different species of *Anaplasma* bacteria, including *A. phagocytophilum, A. marginale, A. platys,* and *A. centrale* (Ybanez & Inokuma 2016). *Anaplasma phagocytophilum* has been found to be the cause of equine granulocytic anaplasmosis in species of the Equidae family. This disease is characterized by fever, swelling of the joints, anemia due to bone marrow hypoplasia, enlarged lymph nodes, and inappetence, in addition to other symptoms (Dziegiel et al. 2021). Other species of *Anaplasma* infect different hosts based on species, location, and environmental factors (Ybanez & Inokuma 2016). *Anaplasma* bacteria are known to be transmitted by Ixodid ticks (Dziegiel et al. 2021), which are a common ectoparasite for plains zebra in the Etosha National Park (ENP) (Kamath et al. 2014). In ENP, high anaplasmosis prevalence in the zebra population is a concern for both the health of the wildlife in the park, and livestock in the surrounding regions.

Study Area

The Etosha National Park is a 22,915 km$^2$ park in northern Namibia. This park is home to more than 114 species, including a population of approximately 15,000 zebra and is characterized by vastly opposing seasons (Zidon et al. 2017). The park has a dry season from April to October in which there are limited watering holes, and animals must compete for resources. In this time, due to limited resources, animals live in larger herds
and are restricted to smaller areas of the park. The park becomes barren and is covered in a white layer of salt from the Etosha pan, a 4,800km² salt basin which covers almost a quarter of the park. During the dry season, it is estimated that about 70% of the ENP zebra population migrates to the eastern side of the park where waterholes are more abundant (Zidon et al. 2017). During the wet season, from October to April, the park becomes green and luscious with many species of plants and many watering holes to support the animals who live on the park, and portions of the zebra population disperse throughout the park, inhabiting both the eastern and western regions (Zidon et al. 2017).

Due to the tourism draw that ENP brings to this area of Namibia, the park is surrounded by many game reserves and farms, which serve as accommodations for guests to the park (Biggestleaf Media, 2021). In addition, northern Namibia is known for farming of grazing cattle which roam the area. Sheep and goats are also common in this area, and commercial farming is a large source of income for the country (BDO Namibia, 2021). Due to the proximity of grazing cattle to the park and surrounding game farms, the spread of disease through parasites between wildlife populations in the park and farm animals could be a concern.

Potential Risk Factors of Infection

Possible risk factors of parasite infections in plains zebra include age, sex, genetics, and coinfection with other parasites, such as ectoparasites and gastrointestinal nematodes. Age could pose as a reasonable risk factor due to the level of fitness associated with both young and old animals (Niccolli, & Partridge 2012). In addition, age could play a part in the level of adaptive immunity an animal has acquired.
Typically, older animals have been exposed to a wider range of pathogens which allows them to build a larger antibody repertoire and increased pathogen-driven immunity (Niccolli, & Partridge 2012). Sex could also be an important risk factor for many diseases, as the differences in the hormones, anatomy, and behavior of males versus females could make the sexes differentially susceptible to a disease (El-Sherbini &Abosdera 2013). In addition, other individual factors, such as genetics or infections with other parasites, could affect an individual’s immune response to infection.

The MHC is a group of genes in vertebrates that is responsible for initiating an immune response in the presence of foreign pathogens. The MHC works by using cell surface receptors to display fragments of pathogens for recognition by helper T cells (Roche & Furuta 2015). Class II MHC genes are responsible for parasite immunity in hosts. It is understood that class II MHC genes undergo rapid evolutionary changes in response to infections from parasites, which leads to great diversity in the MHC genes of a population (Eizaguirre et al. 2021). This may occur through frequency dependent selection, which predicts that an MHC allele which is effective at resisting a particular pathogen will become dominant in a population and become more frequent in that population over time (Hacking et al. 2018). This phenomenon may occur when MHC alleles are associated with resistance; for example, a study done on a wild agamid lizard population showed that individuals possessing MHC class 1 supertype 4 had a lower tick parasite prevalence than individuals without this supertype, providing a survival advantage that was selected for in the population (Hacking et al. 2018). Furthermore, in the ENP zebra population, Kamath et al. (2014) found support for frequency dependent selection at the MHC locus, ELA-DRA Exon 2, with rare alleles at the locus associated
with increased gastrointestinal parasite loads, and common alleles associated with increased tick loads (Kamath et al. 2014). This study also suggests that there is an immunogenetic tradeoff between multiple parasite infections and MHC diversity (Kamath et al. 2014).

Another mechanism by which hosts MHC genes may evolve to resist parasites is through heterozygote advantage. Under the heterozygote advantage hypothesis, a heterozygote individual (i.e. an individual possessing two different alleles at a particular locus) will be better able to resist pathogens than either homozygote parental individuals (i.e. individuals with two of the same alleles at a locus). In addition, heterozygotes should be able to detect and present a wider range of pathogen-driven antigens due to a larger number of different MHC molecules (Sommers, 2005). An example of the heterozygote advantage is that of the connection between individuals who are heterozygous for sickle cell anemia and resistance to malaria. It was found that individuals with incomplete dominance or heterozygosity in the gene associated with sickle cell anemia had enough dysmorphism of the red blood cells to prevent the malaria Plasmodium from surviving in the cells without causing full sickle cell anemia symptoms. Because of this resistance to malaria, heterozygosity in this allele is prevalent and favored in African populations where Malaria is prevalent (Allison 1954).

Research Goals and Purpose

There is limited information available pertaining to the risk and prevalence of Anaplasma infections in wild plains zebra populations. Therefore, this study aims to determine the prevalence and risk factors of Anaplasma infections in the plains zebra
population in Etosha National Park, Namibia. In addition, this study aims to better understand how different types of parasite infections correlate with each other and affect a host’s ability to resist other parasites. By analyzing how gastrointestinal and ectoparasite loads affects the risk of an individual contracting an infection caused by the genus *Anaplasma*, conclusions could be drawn about the risk of multiple parasite infections in this species. By determining the risk factors of *Anaplasma* infections and the prevalence of these infections in ENP plains zebra, surrounding farms could be better able to protect against the spread of anaplasmosis in their herds, and researchers will further understand the risk of this disease in other wild ungulate populations.
METHODS

Field Sampling and Risk Factor Data Collection

In previous work done by Kamath et al. (2014), whole blood, ectoparasites, and fecal samples were collected from 70 adult zebra from Etosha National Park between 2008 and 2010. At the time of sampling, the age of each zebra was estimated by analyzing dental wear on permanent incisors (Smuts, 1974). A modified McMaster fecal floatation approach was used to measure fecal egg counts as an index of gastrointestinal parasite loads (Gordon & Whitlock 1939; Turner & Getz 2010). Ectoparasites were identified to the species level and used to determine tick loads (Horak et al. 1992) by counting the number of ticks of the species *Hyalomma rufipes*, *Hyalomma truncatum*, and *Rhipicephalus evertsi evertsi*, collected from each individual.

Whole genomic DNA was extracted from whole blood samples using Qiagen DNeasy kits (Qiagen inc., Germantown, MD). From these DNA samples, individuals were genotyped at the equid MHC locus, ELA-DRA exon 2, using polymerase chain reaction (PCR) and direct Sanger sequencing (Kamath & Getz 2011; Kamath et al, 2014). In addition, the frequency of each allele at the ELA-DRA exon 2 genetic locus in the ENP zebra population was determined, as in Kamath et. al. (2014). According to their frequency in the population, these alleles were categorized as rare (<5%), mid-frequency (5-10%), and common alleles (>15%) (Kamath et al. 2014).

**Anaplasma Prevalence**

Following the procedure of Barlough et al. (1996), a nested PCR protocol targeting a 928 bp partial sequence of the 16S rRNA locus was amplified to determine
the presence of *Anaplasma* spp. in a subset of individuals (*n* = 64), for which there was sufficient genomic DNA still available. For the first PCR, a master mix was made using 5x GoTaq Buffer with 7.5nM magnesium chloride, 10mM dNTPs (New England Biolabs, Ipswich, MA), 10mM EE1 primer, 10mM EE2 primer (Barlough et al. 1996), 5u/μL GoTaq polymerase (Promega, Madison, WI), and water. Samples concentrations were not standardized before testing, as the initial concentrations were low, so 2μL containing 2 - 13ng of each DNA sample were added to a labelled test tube, in addition to 23μL of master mix for a total volume of 25μL. The master mix for PCR 2 was made with the same reagents as were used in PCR 1 with the exception of the use of 10mM EE3 primer and 10mM EE4 primer (Barlough et al. 1996); also, instead of DNA, two microliters of PCR product from PCR 1 were used in combination with 23 uL of PCR 2 master mix for a total volume of 25μL. A positive and negative control were included in each PCR reaction to prove that the results were reliable. A negative control was made using the same volume of master mix, plus 2 microliters of water. In addition, EQ 118 was tested in many test PCRs with known positive moose samples and was proven to be positive for the presence of an *Anaplasma* spp.; therefore, this sample was used as a positive control in each PCR.

Following the completion of the nested PCR protocols, gel electrophoresis was used to visualize the PCR products. In this process, a standard gel electrophoresis protocol was used in which 2.2g of agarose was added to 110mL of 0.5x TBE solution. A DNA ladder was prepared using 3uL of nuclease free water, 2uL of 100 base pair DNA ladder, and 2 μL of loading dye. Five microliters of PCR product from PCR 2 of the
nested reaction were mixed with 2μL of loading dye and added to each corresponding well of the gel.

Each risk factor was described by the number of individuals possessing this factor, the number of *Anaplasma* positive and negative individuals with this factor, and the proportion of individuals with this factor who tested positive for an *Anaplasma* infection. Epitools, by Ausvet, (https://epitools.ausvet.com.au/), was used to perform a Chi-squared test for a contingency analysis to estimate the prevalence of *Anaplasma* infection with relation to each explanatory risk factor, and to determine whether *Anaplasma* infection prevalence differed significantly by risk factor.

**Statistical Analyses**

I used generalized linear models to evaluate the relationship between individual variables and *Anaplasma* status and identify risk factors of infection in plains zebra. An Akaike Information Criteria (AIC) model selection approach was used to determine which putative risk factors (Table 1) help explain infection of an individual with *Anaplasma*. Generalized linear models of single explanatory variables as predictor values were fit using *Anaplasma* infection status as a binary response variable (0 = not infected and 1 = infected). Multicollinearity was assessed between the explanatory variables which performed better than the null in the AIC model selection process, and these variables were combined in a multivariate generalized linear model to assess significance of the top-ranking risk factors. Factors included in the multivariate final model with a *p*-value below 0.05 were considered significant. All statistical analyses were conducted in R Studio v1.4.1106 (R Core Team 2021).
**Sanger Sequencing**

Positive PCR products were purified using an enzymatic clean up protocol, ExoSap-IT Express (Applied Biosystems, Foster City, CA), which involves the addition of exonuclease I shrimp alkaline phosphatase. Cleaned PCR products were then sent to GENEWIZ, an outside sequencing lab headquartered in New Jersey, for Sanger Sequencing of the 16S ribosomal RNA gene locus. Forward and reverse sequences of this amplified region were generated, using the EE3 and EE4 primers from PCR 2. Geneious Prime v2021.1.1 was then used to align and trim sequences which were then manually checked to confirm all polymorphisms. After editing this data, the sequences were BLASTed to the National Center for Biotechnology Information (NCBI) nucleotide database (http://blast.ncbi.nlm.nih.gov) to find the most closely related strains from the *Anaplasma* genus that matched the sequences found in the ENP plains zebra population, and identified the *Anaplasma* spp. present. Haplotypes (i.e., sequence variants) were then identified that represented the mutant strains present in the population. These haplotypes were aligned with additional *Anaplasma* strain sequences obtained from GenBank (NCBI) using the Geneious alignment tool, including *Anaplasma* strains which have been found in domestic animals from surrounding regions in Africa, and global *Anaplasma* strains previously found in livestock, humans, and arthropod vectors (Table 3). Finally, a phylogenetic analysis was conducted using the MrBayes plugin in Geneious, and visualized by a tree showing the evolutionary relationships among all *Anaplasma* strains.
RESULTS

*Anaplasma* Infection Prevalence in Zebra

Of the 64 zebras screened for *Anaplasma* bacterial infection in this study, I found 23 *Anaplasma*-positive individuals. Given these results, I estimated that the ENP zebra population has an *Anaplasma* infection prevalence of 35.9%. Figures 1-3 show the gels which were used to determine the positive or negative status of each sample following the PCRs.

I found that 22.2% of male (95% CI = 0.06 - 0.55) and 38.2% of female (95% CIs = 0.27 – 0.51) zebras in this population were *Anaplasma*-positive, and a contingency analysis showed no significant differences in infection between sexes ($\chi^2 = 0.30, p = 0.58$; Figure 4). In addition, 45.5% of ENP zebras that carried *H. rufipes* species ticks were *Anaplasma* positive (95% CI = 0.21 - 0.72), while 50% of zebras that carried *H. truncatum* species ticks were positive (Figure 5; 95% CI = 0.29). The third species of tick, *R. e. memiticus* ticks, were not included in the analysis because only 1 individual out of the 64 tested possessed this species of tick. The contingency analysis showed that there was no significant difference in *Anaplasma* infection between individuals with or without *H. rufipes* ($\chi^2 = 0.09, p = 0.76$) as well as between individuals with or without *H. truncatum* ($\chi^2 = 0, p = 1.0$) species ticks. 41.4% of heterozygotes (95% CI = 0.26 - 0.59; Figure 6) and 29% of homozygotes (95% CI = 0.16 – 0.45) in this population tested positive for *Anaplasma* infection. The contingency analysis showed that there was no significant difference in heterozygosity and *Anaplasma* infection ($\chi^2 = 0.66, p = 0.42$). Individuals who possessed rare, mid, and common alleles had an *Anaplasma* prevalence of 38.5%, 26.3%, and 40.8%, respectively (Figure 7). The contingency analysis showed
that there was no significant difference in rare (95% CI = 0.18 – 0.64), mid-frequency (95% CI = 0.12 – 0.49), and common alleles (95% CI = 0.28 – 0.55) at the MHC ELA-DRA exon 2 locus with respect to *Anaplasma* infection ($\chi^2 = 1.25, p = 0.54$). Age was broken into two groups which were zebras ages one to six years old (age group 1), and zebras ages 7 to 12 years old (age group 2). The zebras in age group 1 had an *Anaplasma* infection prevalence of 52.6% (95% CI =0.37 - 0.68) while age group 2 had a prevalence of 11.5% (95% CI = 0.04 - 0.29; Figure 8). There was a significant difference in the risk of *Anaplasma* infection by age group according to the contingency analysis ($\chi^2 = 9.61, p = 0.0019$). Lastly, the gastrointestinal parasite load was described in terms of eggs per gram (EPG), and individuals were broken into two groups. EPG group 1 was characterized by having less than 4000 EPG, while EPG group 2 was made up of individuals containing greater than 4000 EPG. I found that 25.8% of EPG group 1 individuals were positive (95% CI = 0.14 - 0.43) while 47.6% of EPG group 2 individuals were positive for *Anaplasma* infection (95% CI = 0.28 - 0.68; Figure 9). However, the contingency analysis for these factors showed no significant difference in *Anaplasma* infection between gastrointestinal parasite load groups 1 and 2 ($\chi^2 = 1.76, p = 0.19$)

**Statistical Analyses Results**

When AIC model selection was used, four out of the twenty-one explanatory variables performed better than the null (19%). These variables were age, gastrointestinal parasite load (EPG), *H. truncatum* species tick count, and the possession of ELA-DRA exon 2 allele 1. In the multivariate generalized linear model the only significant variable
was age ($p = 0.020$; Table 2). The estimate provided for age ($-4.428e^{-1}$) indicates that when the age increases, the prevalence of *Anaplasma* infections decreases in the population, which means that older zebras are less likely to contract an *Anaplasma* infection (Figure 10). Gastrointestinal parasite load, *H. truncatum* species tick count, and possession of the ELA-DRA exon 2 allele 1 had $p$-values of 0.16, 0.50, and 0.99, respectively. The estimates for gastrointestinal parasite load ($2.804e^{-4}$) and the number of *H. truncatum* species ticks ($1.998e^{-1}$) indicate that these variables have a positive relationship with *Anaplasma* infection, meaning that as these factors increase, so does the probability of *Anaplasma* infections (Figures 11-12). The estimate for ELA-DRA exon 2 allele 1 ($-1.737e^{1}$) indicates a negative relationship between possession of this allele and *Anaplasma* infection, meaning that when an individual possesses this allele, probability of *Anaplasma* infection decreases. Models are shown relating the probability of an individual having an *Anaplasma* infection with age (Figure 10), gastrointestinal parasite load (Figure 11), and *H. truncatum* species ticks (Figure 12).

**Anaplasma Sequencing Results**

After editing the *Anaplasma* 16s rRNA sequences, I determined that there were two mutations present, in the 27 samples that were sequenced, which resulted in three haplotypes of this 16s gene in the ENP zebra population. Haplotype 1 had a frequency of 71.9%, haplotype 2 had a frequency of 7.8%, and haplotype 3 had a frequency of 20.3%. The consensus *Anaplasma* sequence found in zebra was 99.7% similar to an *A.platys* strain (Accession number: KU586031) previously found in a mosquito (Guo et al. 2016), but that is capable of infecting diverse host species. The phylogenetic analysis showed
that the strains most closely related to the three haplotypes extracted from this population were KU586031, a strain of *A. platys* found in mosquitoes in China, and MN401148, a strain of *A. platys* found in dogs in Angola.
DISCUSSION

As there is little known about the prevalence of and risk factors for *Anaplasma* infections in wildlife, this study aimed to fill this gap in knowledge. It also provided data that may be valuable for further study of the effects *Anaplasma* infection has on the ENP plains zebra population as well as neighboring domestic animal populations. The prevalence of *Anaplasma* infection in ENP plains zebra was estimated to be 35.9%. and age was a significant risk factor in for *Anaplasma* infections. The consensus *Anaplasma* strain found in zebra was 99.7% similar to *A. platys*, with three haplotypes of the 16S rRNA gene locus. In addition, gastrointestinal parasite load, presence of *H. truncatum* species ticks, and possession of the ELA-DRA exon 2 allele 1, although not significant, were included in the final model, and were found to explain some of the variation in *Anaplasma* infection among individuals of the ENP zebra population.

The finding that there was an inverse relationship between *Anaplasma* infection and age indicates that younger zebras may be more susceptible to *Anaplasma* infections than older zebras. While it is not known why this relationship exists, it has been shown with infections of other mammalian species as well. I hypothesize that this relationship could be explained by the lack of time juveniles have been exposed to a wide range of pathogens, which is necessary for developing adaptive immunity and the ability to resist future pathogen infections. In African buffalo, the median age of first infection of tick-borne pathogens, including *Anaplasma*, were found to be the lowest in comparison to other pathogens examined (Combrink et al. 2020), and therefore individuals may be more likely to acquire these pathogens early in life and then develop immunity with continued exposure.
The prevalence of *Anaplasma* infection in this population was estimated to be 35.9%. A study done on roaming cattle and goats from villages in Uganda found *Anaplasma* infection rates of 1.6% (cattle) and 6.0% (goats) (Weny et al. 2017). In addition, a study done in the Czech Republic found an *Anaplasma* infection prevalence of 85.7% in wild red deer (Zeman & Pecha 2008), thus the prevalence of *Anaplasma* infection has been found to vary greatly depending on location and species.

The prevalence of *Anaplasma* infection is provided here as an estimate because the ENP plains zebra population includes more than 15,000 individuals, of which this study tested only a small subset (N=64) sampled in the central region of the park. In order to obtain a more accurate infection prevalence in this population, this study would need to be repeated on a larger group of individuals from across the park. This is true for the risk factors analysis as well. To obtain dependable results, a large sample size which is representative of the population would need to be tested. For example, it cannot be proven that sex is not a risk factor from this study because the distribution of sexes was not representative of the population (9 males and 55 females), even though sex performed much lower than the null model in the AIC model selection process and was insignificant in the risk factor analysis.

Other risk factors that were included in the top multivariate linear model included gastrointestinal parasite load and coinfection with *H. truncatum* species tick. Because gastrointestinal parasite load and number of a certain species of tick were factors that performed better than the null, it is possible that these relationships could indicate an additive negative effect of co-infections, such that increased parasite loads could result in an increased susceptibility to other pathogens, including *Anaplasma*. This could be due
to a depressed immune response resulting from multiple parasite infections, or an existing infection decreasing physiological condition, leading to increased susceptibility to other parasites. For example, a review analyzing tradeoffs in multiple parasite infections showed that competition for resources of parasites within a host could affect the virulence of the parasite or the susceptibility of the host to new parasite infections (Cressler et al. 2017). In addition, a study done on immune tradeoffs associated with malaria showed that individuals infected with this disease had a decreased immune response to helminth infections (Mabbott, 2018). While it was not significant, here, I showed that as the gastrointestinal parasite load and number of H. truncatum ticks in/on an individual increase, the probability of Anaplasma infection increased as well. Future research with larger sample sizes and more detailed parasite information should be done to test the relationship between gastrointestinal parasites, ectoparasites, and blood parasites to see if there is an immune tradeoff present.

The ELA-DRA exon 2 allele 1 also performed better than the null model in this study, while not being significant. The single model comparing possession of the ELA-DRA exon 2 allele 1 and probability of Anaplasma infection suggested that possession of this allele decreased the probability of an Anaplasma infection. This allele was characterized as a mid-frequency allele (between 5-10%) in the ENP plains zebra population. Similarly, a study done of half-smooth tongue sole (Cynoglossus semilaevis), showed that there were 4 MHC class II alleles associated with resistance to V. anguillarum, a pathogen responsible for causing the disease vibriosis in fish (Du et al. 2011). In addition, a study done on domestic sheep (Suffolk, Corridale, and crossbred) found that the OLA-DRB *n23 allele was associated with increased resistance to bovine
leukemia virus-induced ovine lymphoma (Nagaoka et al. 1999). It is important to consider that DRA is a class II MHC gene, which is known to respond to extracellular parasites. This gene could play a role in responding to Anaplasma when it is outside of the host cell, however, Anaplasma is an obligate intracellular parasite that needs to be inside the host cell to replicate. Therefore, it is not expected to find many relationships between this gene and Anaplasma infection. The results of this study generally agree with this, with the exception of the ELA-DRA exon 2 allele 1, which was included in the top multivariate generalized linear model. It is possible that the inclusion of this variable in the top model could be due to this allele having an indirect effect on infection through another unmeasured variable.

The 16s rRNA sequencing data revealed that the Anaplasma strain present in this population was highly similar (99.7% sequence identity) to A. platys, a bacteria which is known for causing anaplasmosis infections in canine species (Carvalho et al. 2016). The A. platys strain (GenBank Accession No. KU586031) closely clustered with the three haplotypes in the phylogenetic tree was originally found in mosquitos in China (Guo et al. 2017). However, this bacterium has also been found to infect livestock and other ungulates. A study in Thailand found that a population of Thai water buffalo (Bubalis bubalis) had a high A. platys infection prevalence (Nguyen et al. 2020). There have also been cases of A. platys in cattle in Algeria (Dahmani et al. 2015). In addition, the BLAST search revealed the strain was also very similar (99.5% similarity) to an Anaplasma sequence (GenBank Accession No. MN401148) recently found in domestic ruminants and dogs in Angola (Sili 2020). Because A. platys is commonly found in canines, the prevalence of A. platys infections in ENP zebra could be due to wild canids that reside in
the fenced park with the zebra (Biigestleaf Media, 2021). In particular, ENP has a large population of foxes and jackals that could be reservoirs for *A. platys* infections. If a tick were to drop from one of these canines and attach to a zebra in the park, it could transmit this *Anaplasma* infection from canine to zebra. Currently, it is unknown whether either foxes or jackals are reservoirs for *A. platys* in ENP. Thus, further research should seek to assess infection in these species to determine the source of *Anaplasma* in ENP.

Possible errors of this study include varying concentrations of DNA in the PCR samples which could have affected the ability of the PCR to provide accurate results, as a sample with a very low concentration of DNA may not be amplified enough to display a positive result. In addition, a contamination issue occurred in the lab, which required me to run many different PCRs with repeat samples. Samples were declared positive if they appeared positive in any PCR that did not display evidence of contamination. Lastly, there is a chance that gene sequences could have minor errors from editing, but these are unlikely to change the results of the phylogenetic analysis.

This study identified a key risk factor, age, that is associated with *Anaplasma* infections in plains zebra of ENP. The identification of risk factors will help researchers better understand the drivers of parasite infections in this population for future research and management. While I examined the correlations between gastrointestinal parasites, ectoparasites, and blood parasites, further work is needed to better understand the effect of multiple co-infecting parasites on host health. While the gastrointestinal parasites and ectoparasites did not have significant effects on *Anaplasma* infection in this study, their inclusion in the top model suggests there could be a connection between the these factors in an individual’s ability to resist infection. Further study is recommended to confirm
this result using larger sample sizes. Characterizing the prevalence of parasite infections and understanding the risk factors associated with infection and the role of multi-parasite co-infection is an important goal in better understanding wild zebra populations and managing the potential spread of *Anaplasma* infections to other species.
CONCLUSION

The ENP plains zebra population has an estimated *Anaplasma* infection prevalence of 35.9%, which I identified as *A. platys*. The probability of an individual contracting *Anaplasma* infection decreased with age, possibly suggesting acquired immunity following exposure to the parasite. While gastrointestinal parasites and ectoparasites were not found to be significant risk factors for this infection, it is possible multiple parasite infections interact, with infection by one parasite suppressing an individual’s immune response to infection by another parasite. Lastly, while also not a significant risk factor in this study, the ELA-DRA exon 2 allele 1 may confer resistance to *Anaplasma* infections in zebra. The information gathered in this study could be used in future studies to better understand risk factors for disease in ENP plains zebra and how the prevalence of *Anaplasma* infections relates to other diseases in the ENP plains zebra population.


Sili G. 2017. Species composition of ticks and tick-borne pathogens in domestic ruminants and dogs in Tchicala-Tchoaloanga Huambo Province Angola. *University of Pretoria.* https://repository.up.ac.za/handle/2263/67906


**Table 1**: Putative risk factors for *Anaplasma* infection in plains zebra (*Equus Quagga*) of Etosha National Park used as explanatory variables in the AIC model selection process. Continued on Page 26.

<table>
<thead>
<tr>
<th>Putative Risk Factor</th>
<th>Description of Risk Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Estimated age determined from tooth wear</td>
</tr>
<tr>
<td>Gastrointestinal Parasite Load</td>
<td>Number of eggs per gram</td>
</tr>
<tr>
<td><em>H. truncaturn</em> species tick</td>
<td>Number of ticks belonging to this species, found on the individual at the time of sampling</td>
</tr>
<tr>
<td><em>H. rufipes</em> species tick</td>
<td>Number of ticks belonging to this species found on the individual at the time of sampling</td>
</tr>
<tr>
<td><em>H. sulcatus</em> species tick</td>
<td>Number of ticks belonging to this species found on the individual at the time of sampling</td>
</tr>
<tr>
<td><em>R.e. mimeticus</em> species tick</td>
<td>Number of ticks belonging to this species found on the individual at the time of sampling</td>
</tr>
<tr>
<td>Total ticks</td>
<td>Total number of ectoparasites found on the individual at the time of sampling</td>
</tr>
<tr>
<td>Sex</td>
<td>Male or female</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 1</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 3</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 4</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 5</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 7</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
</tbody>
</table>
Table 1: Continuation

<table>
<thead>
<tr>
<th>Putative Risk Factor</th>
<th>Description of Risk Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELA-DRA Exon 2 allele 9</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 10</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>ELA-DRA Exon 2 allele 11</td>
<td>A binary variable describing individuals possessing this allele on the ELA-DRA Class II MHC locus (1) and those without (0)</td>
</tr>
<tr>
<td>DRA SNPS</td>
<td>Number of Single Nucleotide Polymorphisms at the ELA-DRA locus</td>
</tr>
<tr>
<td>Possession of a Common Allele</td>
<td>A binary variable describing individuals possessing an allele with a frequency greater than 15% in the population (1) and those that don’t (0)</td>
</tr>
<tr>
<td>Possession of a mid-frequency allele</td>
<td>A binary variable describing individuals possessing an allele with a frequency between 5-10% in the population (1) and those that don’t (0)</td>
</tr>
<tr>
<td>Possession of a rare allele</td>
<td>A binary variable describing individuals possessing an allele with a frequency less than 5% in the population (1) and those that don’t (0)</td>
</tr>
<tr>
<td>Heterozygote individual</td>
<td>A binary variable describing individuals heterozygous at the ELA-DRA exon 2 locus (1) versus individuals that are homozygous at the locus (0)</td>
</tr>
</tbody>
</table>

Table 2: Putative risk factors used as explanatory variables in a multivariate generalized linear model predicting *Anaplasma* infection in plains zebra (*Equus Quagga*) of Etosha National Park. Included in this table are variables which performed better than the null model and their associated P-values and estimates. Significant variables are indicated in bold.

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>P-Value</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.020</td>
<td>-4.428e⁻¹</td>
</tr>
<tr>
<td>Gastrointestinal parasite load (Eggs per Gram)</td>
<td>0.160</td>
<td>2.804e⁻⁴</td>
</tr>
<tr>
<td><em>H. truncatum</em> species ticks</td>
<td>0.496</td>
<td>1.998e⁻¹</td>
</tr>
<tr>
<td>Possession of ELA-DRA Exon 2 allele 1</td>
<td>0.994</td>
<td>-1.737e⁻¹</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Taxa ID</th>
<th>Host Species</th>
<th>Origin</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. bovis</em></td>
<td><em>Bos taurus</em></td>
<td>India</td>
<td>MH244925</td>
</tr>
<tr>
<td><em>A. bovis</em></td>
<td><em>Lepus sylvaticus</em></td>
<td>Massachusetts, USA</td>
<td>AY144729</td>
</tr>
<tr>
<td><em>A. centrale</em></td>
<td><em>Bos taurus</em></td>
<td>Southern Italy</td>
<td>EF520690</td>
</tr>
<tr>
<td><em>A. marginale</em></td>
<td>NA</td>
<td>Florida, USA</td>
<td>AF309867</td>
</tr>
<tr>
<td><em>A. ovis</em></td>
<td><em>Ovis aries</em></td>
<td>China</td>
<td>AY262124</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (1)</td>
<td><em>Alces alces alces</em></td>
<td>Norway</td>
<td>KT070819</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (2)</td>
<td><em>Alces alces alces</em></td>
<td>Norway</td>
<td>KT070822</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (1)</td>
<td><em>Alces alces alces</em></td>
<td>Sweden</td>
<td>KC800983</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (2)</td>
<td><em>Alces alces alces</em></td>
<td>Sweden</td>
<td>KC800985</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td><em>Dermacentor albipictus</em></td>
<td>Maine, USA</td>
<td>MW899038</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td><em>Homo sapiens</em></td>
<td>Connecticut, USA</td>
<td>KT454992</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td><em>Ixodes ricinus</em></td>
<td>Warsaw, Poland</td>
<td>MH122891</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td><em>Ixodes scapularis</em></td>
<td>Connecticut, USA</td>
<td>EF123258</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (1)</td>
<td><em>Ixodes scapularis</em></td>
<td>Maine, USA</td>
<td>MW899039</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em> (2)</td>
<td><em>Ixodes scapularis</em></td>
<td>Maine, USA</td>
<td>MW899040</td>
</tr>
<tr>
<td><em>A. phagocytophilum</em></td>
<td><em>Ixodes scapularis</em></td>
<td>Saskatchewan, Canada</td>
<td>HG916767</td>
</tr>
<tr>
<td><em>A. platys</em></td>
<td><em>Canis familiaris</em></td>
<td>Southern Italy</td>
<td>EU439943</td>
</tr>
<tr>
<td><em>A. platys</em></td>
<td><em>Canis familiaris</em></td>
<td>India</td>
<td>KT982643</td>
</tr>
<tr>
<td><em>A. platys</em></td>
<td><em>Culicidae</em></td>
<td>China</td>
<td>KU586031</td>
</tr>
<tr>
<td><em>A. platys</em></td>
<td><em>Bos taurus</em></td>
<td>South Africa</td>
<td>MK814421</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em> (1)</td>
<td><em>Alces alces americana</em></td>
<td>Maine, USA</td>
<td>MW899041</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em> (2)</td>
<td><em>Alces alces americana</em></td>
<td>Maine, USA</td>
<td>MW899042</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em></td>
<td><em>Odocoileus hemionus</em></td>
<td>British Columbia, Canada</td>
<td>JN673772</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em></td>
<td><em>Odocoileus virginianus</em></td>
<td>British Columbia, Canada</td>
<td>JN673768</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em> Dedessa</td>
<td><em>Bos taurus</em></td>
<td>Illubabor zone, Ethiopia</td>
<td>KY924886</td>
</tr>
<tr>
<td><em>Anaplasma sp.</em> Saso</td>
<td><em>Bos taurus</em></td>
<td>Illubabor zone, Ethiopia</td>
<td>KY924885</td>
</tr>
<tr>
<td><em>Ehrlichia chaffeensis</em></td>
<td>NA</td>
<td>NA</td>
<td>M73222</td>
</tr>
<tr>
<td><em>Ehrlichia ewingii</em></td>
<td>NA</td>
<td>NA</td>
<td>U96436</td>
</tr>
<tr>
<td><em>Neorickettsia sennetsu</em></td>
<td>NA</td>
<td>NA</td>
<td>M73225</td>
</tr>
<tr>
<td><em>Rickettsia rickettsii</em></td>
<td>NA</td>
<td>NA</td>
<td>L36217</td>
</tr>
</tbody>
</table>
**Figure 1:** Agarose gel image of PCR 4 with samples positives for *Anaplasma* infection shown as strong black bands in wells 4, 6, 7, 8, 10 and 15.

**Figure 2:** Agarose gel image of PCR 12 showing positive *Anaplasma* infection results as a strong white band in wells 2, 5, 6, 7, 10, 11, 14, 16, 17, 18, 21, 27, 28, and 30.
Figure 3: Agarose gel image of PCR 13 showing positive *Anaplasma* infection results as a strong white band in wells 7, 9, 11, 13, 19, 20, 21, 26, 29, 30, and 31.

Figure 4: Contingency analysis output showing the prevalence estimates for male versus female ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.
Figure 5: Contingency analysis output showing the prevalence estimates for *H. rufipes* and *H. truncatum* species ticks in ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.

Figure 6: Contingency analysis output showing the prevalence estimates for heterozygote versus homozygote ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.
Figure 7: Contingency analysis output showing the prevalence estimates for rare, mid-frequency, and common alleles in ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.

Figure 8: Contingency analysis output showing the prevalence estimates for age group 1 (1-6 years old) versus age group 2 (7-12 years old) ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.
Figure 9: Contingency analysis output showing the prevalence estimates for EPG group 1 (<4000 EPG) versus EPG group 2 (>4000 EPG) ENP plains zebra (*Equus Quagga*) with 95% CI shown by the bars.

Figure 10: Plot of *Anaplasma* infection probability as a function of age.
Figure 11: Plot of *Anaplasma* infection probability as a function of the gastrointestinal parasite load, shown as eggs per gram.

Figure 12: Plot of *Anaplasma* infection probability as a function of the presence of *H. truncatum* species ticks.
Figure 13: *Anaplasma* phylogenetic tree showing the relationship between the sequences found in plains zebra and publicly available sequences from humans, ectoparasites, and wildlife (Table 3). Highlighted in red are the three 16s rRNA haplotypes from sanger sequencing results and in blue are the most closely related strains to these haplotypes.
AUTHOR’S BIOGRAPHY

Madison Stahle was born on August 20, 1998 in Atlanta, Georgia. At the age of seventeen, she moved to Boothbay Harbor, Maine where she completed her final year of high school. She then attended the University of Maine, where she is now a fourth-year undergraduate student, majoring in Animal Veterinary Science, with a concentration in Pre-Veterinary Science. Additionally, she is a student of the University of Maine’s Honors College. During her time at the University of Maine, Madison participated in the UMaine Icelandic Sheep Club, the Wildlife Disease Genetics Lab, and assisted Dr. James Weber in his porcine orthopedics research project. Following her undergraduate career, she plans to attend the Auburn University College of Veterinary Medicine to earn her Doctor of Veterinary Medicine degree.