Determinants of tickborne pathogen prevalence

Julia Frederick

2019-12-06

# Summary/Abstract

Tickborne disease is increasingly a public health concern in the United States. With more that 50,000 cases being confirmed annually by the CDC and estimating a 10-fold higher case rate actually occuring, understanding the disease system at play is of the upmost importance. Through tick dragging, almost 2000 individuals were collected, IDed through morphology, and tested using PCR for pathogens in order to determine how the environment determines the risk for disease. This data is use to determine how habitat can influence pathogen prevalence, which readily available factors are contributing to disease risk, and how specifc disease system vary from overall tickborne disease. Through this research it is confirmed that tick population size, and season are two major factors in determining disease risk. When you look specifically at the Lyme disease system using *Ixodes scapularis* and *Borrelia spp.* you see that the season and region are the main contributing factors, and population size falls off. This shows that understanding the overall dynamics of tickborne disease can be useful but that it cannot apply across all individual disease systems.

# Introduction

## General Background Information

Tickborne disease prevalence is on the rise in the United States with 50,000 cases being confirmed by the CDC each year. The number one vector borne disease in the US is Lyme disease with 30,000 cases reported annually, and the Center for Disease Control and Prevention (CDC) estimating over 300,000 cases actually occurring (Vector-Borne Disease, 2019). Understanding how these pathogens are being moved throughout the environment and what influences higher prevalence gives insight into risk prevention efforts and effective targets. Especially when considering that a single species of tick is capable of transmitting multiple disease pathogens (Adelson et al., 2004). There are multiple tick species native to the U.S. that can spread at least one human disease, and the consequences of increased globalization and livestock trade are becoming apparent with invasive species making landfall (Jongejan & Uilenberg, 2004). For example, native species, *Ixodes scapularis* has a range covering the entire eastern half of the U.S. and is capable of transmitting pathogens that cause Lyme disease, anaplasmosis, ehrlichiosis, babesiosis, and Powassan virus (Adelson et al., 2004; Dantas-Torres, Chomel, & Otranto, 2012). Multiple of these pathogens have been found in a single tick during field studies (co-infection) (Moutailler et al., 2016). Current research focusing on the dynamics of these co-infections has shown there to be correlations between specific pathogens occuring in higher frequencies (Abraham et al., 2017). These co-infections can have an impact on human health and diagnostic tests used for detecting the specific pathogens (Krause et al., 2018, 2014). This isn’t only an issue within species, because many tick species are the main vector for specific pathogens and these species can feed on the same source which will mingle the pathogens between species. When investigating multiple tick species, it becomes clear that each species has its own variation in lifecycle and disease prevalence that comes along with it. Knowing what these differences are give researchers the power to create control and prevention efforts for human disease. Current vector research is focusing on understanding the biology of these vectors in order to put future range expansions (Ginsberg et al., 2017), and increased disease threat (Arsnoe, Hickling, Ginsberg, McElreath, & Tsao, 2015; Eisen & Eisen, 2018) into a clearer context. Our research aims to define trends over time in tick species looking into pathogen prevalence, and microbiome shifts. This information will inform future efforts to model disease expansion across the U.S. and into Canada.

## Description of data and data source

Through the Southeastern Coopertative Wildlife Disease Study (SCWDS) ticks were collected from the state of Georgia in the United States over the course of a year. When collected the habitat type and location was recorded for future reference. These ticks were then IDed through morphology, and used for PCR in order to determine the presence of pathogens. The pathogens will not be investigated genetically, instead we will be determining changes in prevalence over time and space. The data used for this analysis includes the pathogens *Rickettsia*, *Ehrlichia ewingii*, *Ehrlichia chaffeensis*, *Anaplasma phagocytophilum*, *Borrelia spp*, and Panola Mountain Ehrlichia (PME). There are multiple tick species that were identified, and samples from each life stage is present. The habitat information includes: season, region, site, and general habitat. Samples were numbered, and the transect they were collected in was accounted for, however, this data will not be used for analysis.

## Questions/Hypotheses to be addressed

**1. How the habitat effects the prevalence of pathogens in ticks?**  
The habitats are broken into a few different locations, some more rural than others and with different tree communities. The effect of habitat on pathogen prevalence is likely due to the different community make-up in each location. Doing this will give us an idea of how the community can sway pathogen prevalence in the select environment.  
**2. What factors have the greatest influence on a tick carrying a pathogen?**  
We will look further into what general variables are influencing pathogen prevalence in Georgia. This data will consider all variables collected that are biologically relevant.  
**3. What influences *Ixodes scapularis* to have *Borrelia spp.*?**  
We apply the concepts used to model for overall tickborne pathogen on a specific disease system to see how similar the outcome is for that population of ticks.

# Methods and Results

## Data aquisition

Through the help of SCWDS ticks were collected from the state of Georgia in the United States over the course of a year. When collected the habitat type and location was recorded for future reference. In order to collect the ticks typical drag sampling was done in each transect of the area, with checks being done of the drag every 10m. When collected the habitat type and location was recorded for future reference. These ticks were then IDed through morphology, and used for PCR in order to determine the presence of pathogens. We will be determining changes in prevalence over time and space. The data used for this analysis includes the pathogens *Rickettsia*, *Ehrlichia ewingii*, *Ehrlichia chaffeensis*, *Anaplasma phagocytophilum*, *Borrelia spp*, and Panola Mountain Ehrlichia (PME). The habitat information includes: season, region, site, and general habitat. The raw data can be found in the project file ./data/raw\_data/RawDataTickPathLoc.csv. This data is used for the analysis.

## Data import and cleaning

The data is originally imported from a .csv file which includes 16 variables and 1968 observations. These observations are from either a single nymph, a single adult, or a pool of larvae. The larvae need to be pooled as individually they do not contain enough DNA for pathogen detection. However, larvae could only be infected with pathogens that are transovarially passed to the offspring (like *Rickettsia*) and the two pathogens consistently tested for (*A. phagocytophilum* and *Borrelia spp.*) are not known to be transmitted in this way. In order to clean this data we first will import the data so all empty observations are coded as NA. Then we need to adjust a few levels within our variables because they have extra spaces or a spelled incorrectly. Once these are fixed we calculate the total number of pathogens found within single individuals. This is our main outcome of interest instead of looking at individual pathogens which weren’t consistently tested.

## Univariate analysis

The main variable that can account for pathogen prevalence is tick species. We are interested in the breakdown of ticks that were sampled. As we can see in Table 1, *A. americanum* is the most frequently collected tick, with *I. scapularis* in second, with about a 3 fold decrease in sample number. The other species collected are negligable, but still important to note they exist within in the same environment so we aren’t going to remove them from the whole data set because it is possible they introduce a new pathogen. This data will be further examined by pathogen breakdown.

*Table 1. Frequency of collection per species. Each species and the number of times collected through drag sampling.*

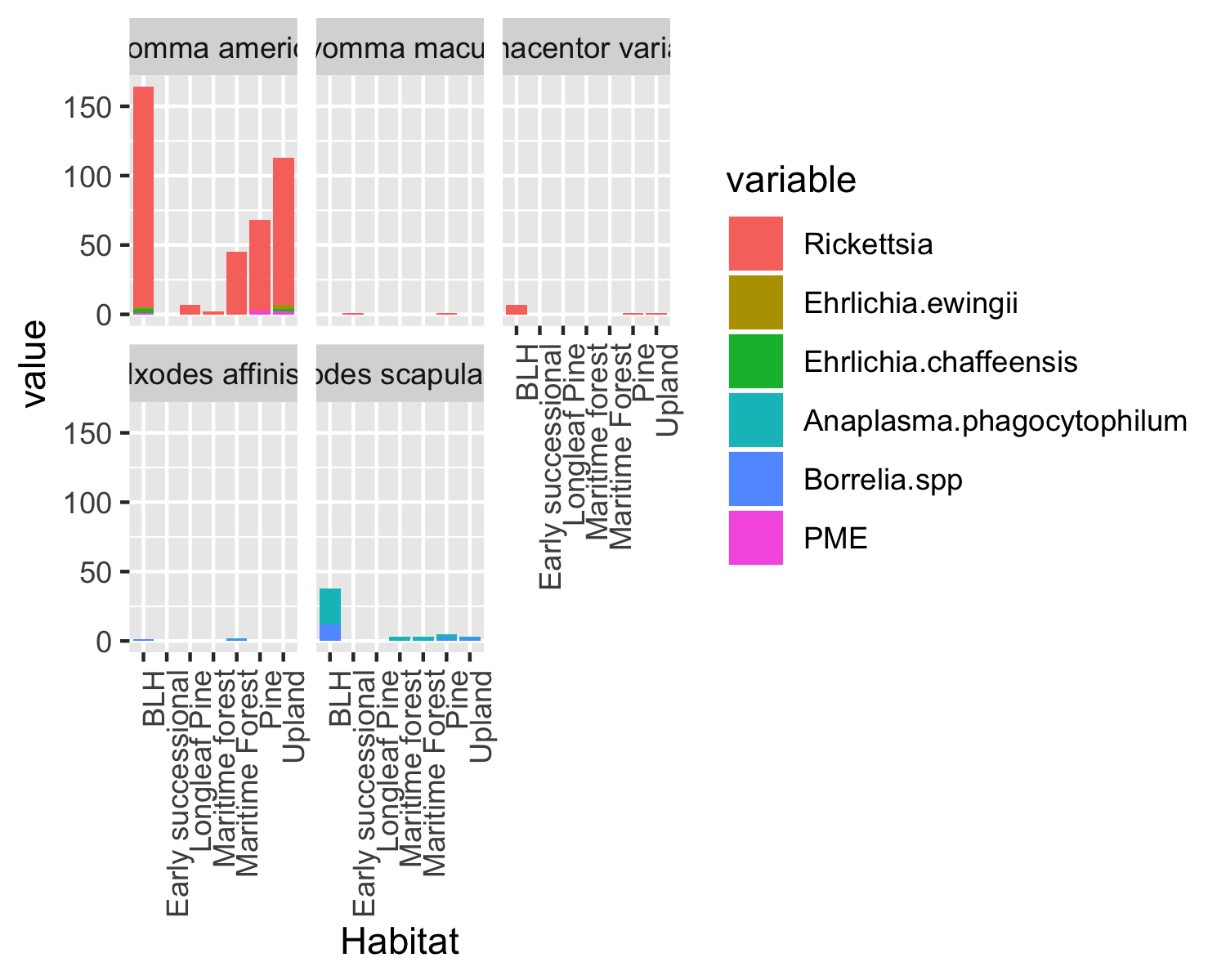
|  |  |  |
| --- | --- | --- |
| Species | frequency | relative\_frequency |
| Amblyomma americanum | 1527 | 77.59 |
| Ixodes scapularis | 414 | 21.04 |
| Ixodes affinis | 16 | 0.81 |
| Dermacentor variabilis | 9 | 0.46 |
| Amblyomma maculatum | 2 | 0.1 |

From our data set that focuses on samples that were tested for *Borrelia spp.*, the only species that is included is *I. scapularis*. This species is the main vector for *Borrelia burgdorferi*, and the other species collected in this study are either confirmed to not be vector species for this pathogen or are assumed through prevalence studies to not carry this pathogen. Knowing which region *I. scapularis* is in is an important metric for disease prevention. In table 2, 50% of *I. scapularis* samples were collected from the Peidmont region. Further studies will be done to know which regions have the highest prevalence of *I. scapularis* ticks testing positive for *Borrelia spp.*

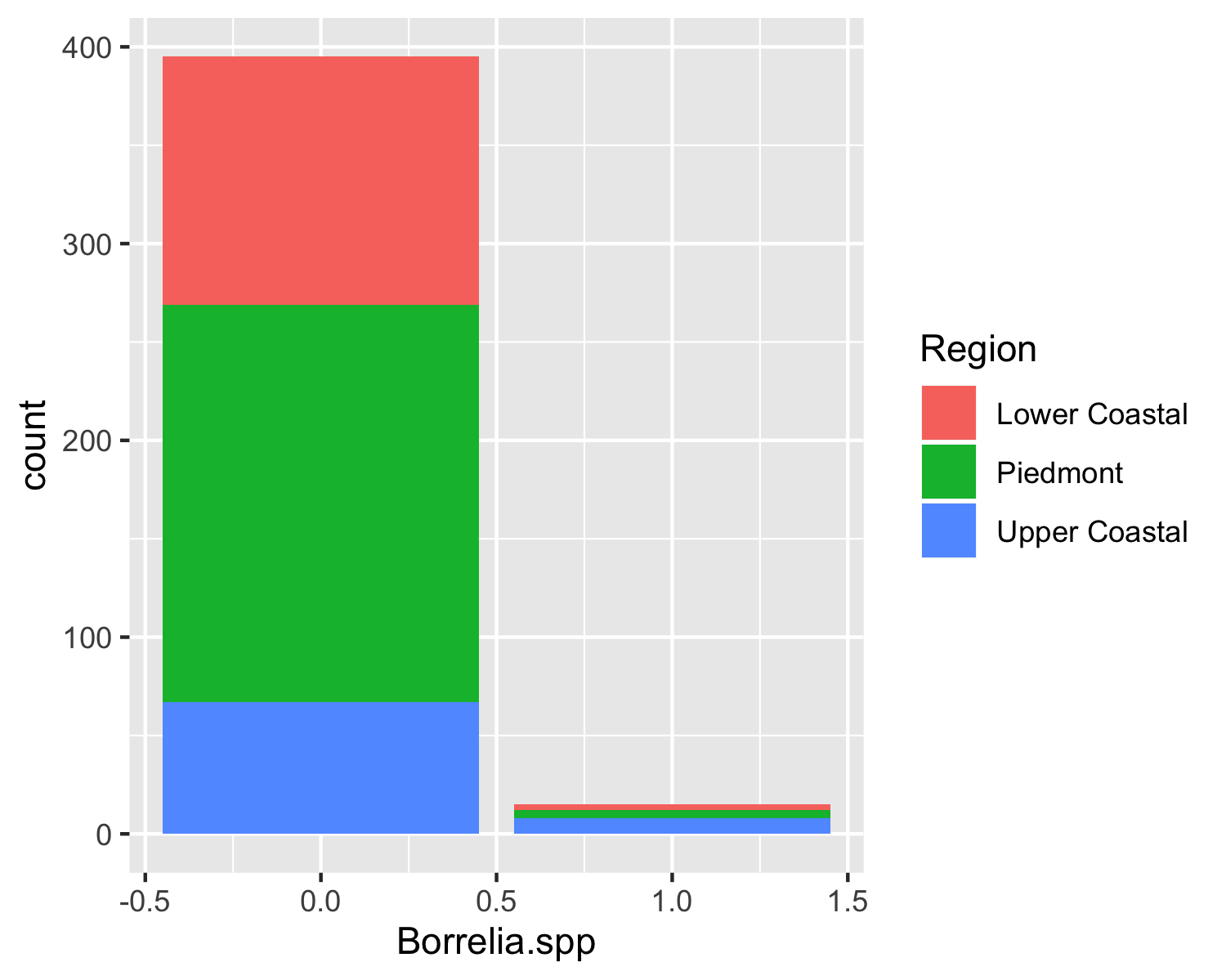
*Table 2.* I. scapularis *collections by region. Shows the frequency of* I. scapularis *samples collected across the different regions.*

|  |  |  |
| --- | --- | --- |
| Region | frequency | relative\_frequency |
| Piedmont | 206 | 50.24 |
| Lower Coastal | 129 | 31.46 |
| Upper Coastal | 75 | 18.29 |

## Bivariate analysis

Many factors influence where a tick resides, and what pathogens are in the population. However, it is important to know which species carry with pathogens. When comparing the prevalence of pathogens in all tick species in different habitats, we can see that certain pathgoens have higher prevalence in particular species. Here we look at how the total number of positive individuals are broken down by species (fig 1). From here we can see that *Rickettsia* is the most common pathogen in *A. americanum*. However, for *I. scapularis* we see that Borrelia species and *Anaplasma* are the two most common. These results trend with what is in current literature. Further analysis of significance is included in the supplemental materials. That analysis shows that the Upland and BHL habitats have significantly more pathogens than other locations.   
*Figure 1. Pathogen prevalence per species. This figure shows the number of individuals per species that tested positive for a particular pathogen.*

Taking a closer look at *I. scapularis* and *Borrelia spp.* we can see that many more samples tested negative than positive, but there were some positive tests in each location (fig 2). In Georgia the prevalence of Lyme disease is significantly lower than in the northeast United States; however, 96% of Lyme disease cases are concentrated in the northeast and central US, so we don’t expect a high rate of positive ticks in Georgia. From table 3 we can see that the Upper Coastal region has the largest amount of positive samples, but the smallest number of samples total.

  
*Figure 2.* Borrelia.spp *Prevalence in Different Regions for* Ixodes scapularis. *This figure shows the different Regions have similar sampling efforts and also that* Borrelia.spp *positive samples are significantly less than negative samples.*

*Table 3. Summary of* Ixodes scapularis *samples tested for* Borrelia spp. *In this table 0 is a negative test, and 1 is a positive test by PCR.*

|  |  |  |  |
| --- | --- | --- | --- |
| Borrelia.spp | Lower Coastal | Piedmont | Upper Coastal |
| 0 | 126 | 202 | 67 |
| 1 | 3 | 4 | 8 |

## Full analysis - Random Forest Model

Pathogen prevalence is determined by many factors in the environment. The factors that we measured that are included in this analysis are season, region, site, habitat, species, sex, life stage, and number of ticks found. The other factors that were recorded but not included in this analysis did not have any biological significance or were already being accounted for in other variables listed. This model is being used to determine the outcome of total pathogen. When seeing which variables influence total pathogens in an individual we get a decision tree that covers a few of the main key indicators in an environment (figure 4). This tree is showing that tick density in a region is the largest predictor and then is further defined by species, season, and habitat.

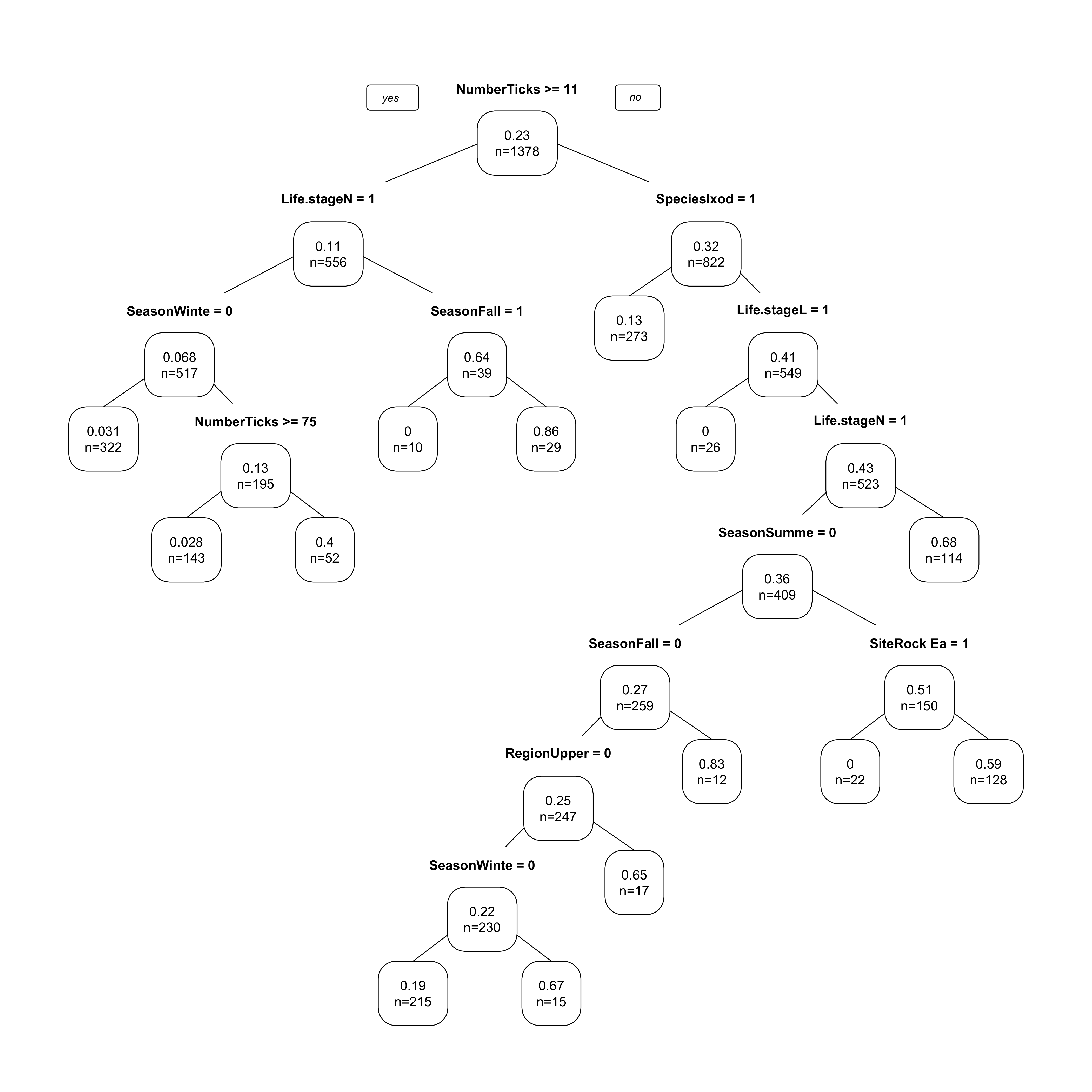


Figure 4. Decision Tree for Predictors of Total Pathogen. Bolded words are test, 0 is FALSE/Negative, and 1 is TRUE/Positive.

To take a closer look at how a specific species and pathogen dynamic can change this model, we look further into *I. scapularis* and *Borrelia spp.* The factors that are included in this analysis are season, region, site, habitat, sex, life stage, and number of ticks found. All other factors were removed as they are not biologically relevant, or are accounted for in a different variable. We used a random forest model, training it on a subst of data, in order to produce the decision tree present in figure 5.

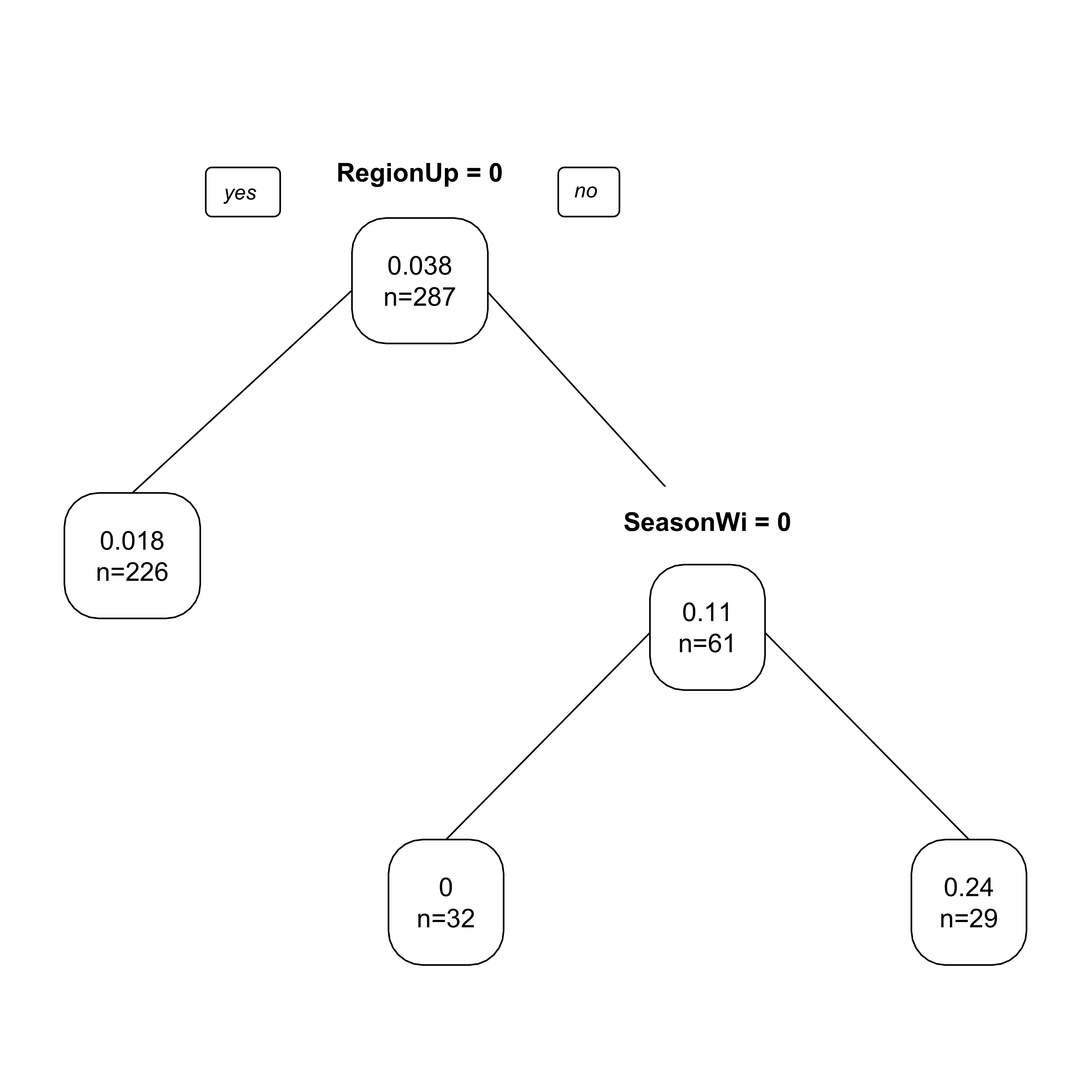


Figure 5. Decision Tree for Predictors of Borrelia.spp in Ixodes scapularis. Bolded words are test, 0 is FALSE/Negative, and 1 is TRUE/Positive.

# Discussion

## Summary and Interpretation

Understanding the distribution of pathogens across different species is important for prevention efforts. Knowing which species are commonly found with a particular pathogen creates a knowledge base for a targeted approach to risk prevention efforts. From this data we see that *A. americanum* mainly has *Rickettsia* species, not all *Rickettsia* species are pathogenic but a majority are. *Amblyomma americanum* is known in literature for being highly competent for *Rickettsia spp.* and this pathogen is transmitted transovarially (from mother to offspring) which causes for an accumulation of this pathogen in the tick population. *Ixodes scapularis* is the main vector for *Borrelia burgdorferi*, the causitive agent of Lyme disease. It is highly competent for this pathogen, with some studies showing 97% survivial of the pathogen transtadially (through molt). However, *Borrelia spp.* are not transmitted transovarially, which can account for why it appears in a lower porportion than *Rickettsia* in *A. americanum*. The other species collected and sampled are common disease vectors throughout the United States; however, their main pathogen is not commonly found in Georgia, where these samples originated from.

Seeing which habitats have the majority of ticks in them, and at what lifestage helps to tease out how the environment could be playing a role in the aggregation of ticks into a select few environments. We expect to see this because ticks are sensitive to desiccation and prefer to be in specific humidity and temperature levels. *Ixodes scapularis* is a species that mainly feeds on small rodents so it is found in the highest amount in the BHL environment which has a large population of rodents. When you compare this to *A. americanum* which is a more generalist species that feeds on a variety of mammals, this species is found across a wide range of habitat types. This data can also speak some to the seasonality and general abundance of tick species. *Ablyomma americanum* is one of the most populous tick species in Georgia, and is typically active in the afternoon during the Summer. This led to a large amount of *A. americanum* being collected over the course of the study. While *I. scapularis* is a species that is spread across the entire eastern half of the United States, and needs a period during the Winter where temperatures drop below 4C in order to complete its lifecycle. This difference in lifecycle compared to other species causes for *I. scapularis* to be active later into Winter and for them to be less active during the afternoon. This species peak time of activity tends to occur closer to sunrise and sunset which isn’t when a majority of drag sampling takes place. Knowing about the phenology of a tick species and seeing it occur in the environment is solidfying what researchers assumed would occur. This gives public health efforts more concrete information on the phenology of their tick speices of interest to move forward with disease prevention efforts and prevention of geographic spread of these populations.

Focusing in on *I. scapularis* a further look into how *Borrelia spp.* distribution occured across regions can correlate to Lyme disease risk in the region. Georgia has a low case rate for Lyme disease but there are established populations of *I. scapularis* throughout, just like in other regions of the United States that have a high case rate. Seeing that a majority of *I. scapularis* ticks samples tested negative for *Borrelia spp.* supports this phenomenon. In similar studies in the Northeast United States, over 30% of *I. scapularis* ticks tested positive for *Borrelia burgdorferi*. However, the rate of *Borrelia* in the different regions in Georgia appear to be similar, showing that *I. scapularis* presence is a major determining factor.

#### Modeling

Being able to determine the variables that have the largest influence on total pathogen outcome and on *Borrelia spp.* outcome is important for future prevention efforts. Within the total pathogen model the most important perdictor is the number of ticks found in a region. Having a large population of ticks increases the total pathogen prevalence. In an environment with a large number of vectors in close contact typically increases the circulation of pathogens, especially through the mammalian hosts. If there is a large number of ticks and the next important factor is the life stage being nymphal. Nymphs are the typically the life stage that has the highest transmission rate, they are the most aggressive and have the highest proportion of individuals that are infected and seeking a bloodmeal. Some species of ticks do not feed as adult males, cutting the adult population that feeds in half. With each life stage the number of surviving individuals also decreases typically 50%, which can also decrease the amount of pathogens circulating in an environment. Another important factor is season, which is highly related to life stage and species that is active. Typically the Summer has the largest proportion of *A. americanum* and *I. scapularis* nymphs active, and their activity decreases as Winter gets closer. The activity of ticks is highly correlated to pathogen prevalence because if they are not active getting samples is more difficult. The variables that the tree identified are all biologically relevant to this system, and can aid in defining where and when humans would be at the most at risk for obtaining a pathogen from a tick. In future efforts this data can also be extrapolated from to determine regions that are not currently at risk for tick borne pathogens but can be as the climate changes, and the species continue to shift their range. This total pathogen tree shows the overall threat to human health for tick borne pathogens; however, more specialized models are likely more informative for that particular disease system.

Through the specific example of *I. scapularis* and the prevalence of *Borrelia spp.*, we can see how the total pathogen model can be parsed down for a particular disease system. *Ixodes scapularis* is a tick species that is more active in the Winter than other species, it’s nymphs are active through the Fall into the Winter. Their shifted phenology compared to other species that were included in the total pathogen model causes season to be a main decider on if *Borrelia spp.* is prevelant in the population. Specifically, whether or not it is Winter is a major deciding factor because the main mammalian host is small rodent, which are not active during the Winter, which is why there is a very low prevalence rate during that season.

## Strengths and Limitations

The strengths and limitations of this analysis begins with data collection. We extensively sampled different areas of Georgia over the course of one year. However, when sampling measures like amibient temperature, humidity, and time of day was not recorded. Not having this data could have further informed these models because these are known to be factors in how active different tick species are. The assay used for pathogen testing is well used PCR assays, but some of them are non-specific so we are only getting an idea of what is possibly in the environment. For example, *Rickettsia* is a broad family of bacteria but only a subset is actually pathogenic. In the analysis, we needed to discount the majority of individual pathogens being tested and combine them into a total pathogen value because many of the samples weren’t tested for the full panel. In the future we would like to test all of the samples for the full panel. We channeled our reasources into testing a majority for *Anaplasma phagocytophilum* and *Borrelia spp.* since they would be the most common in the Georgia area.

## Conclusions

Overall, this paper shows how sampling data, and diagnositic assays can be used to inform predictions of disease risk in different locations. We are able to tell the most at risk season, habitat, and region for the state of Georgia. We are also able to provide information for individual species if the model is broken down further. Overall, this type of research can be used in the future, using exisiting data, to inform public health officials on current disease risk. If these models take in more data they are also be used to predict future risk.

# References

Abraham, N. M., Liu, L., Jutras, B. L., Yadav, A. K., Narasimhan, S., Gopalakrishnan, V., … others. (2017). Pathogen-mediated manipulation of arthropod microbiota to promote infection. *Proceedings of the National Academy of Sciences*, *114*(5), E781–E790.

Adelson, M. E., Rao, R. V., Tilton, R. C., Cabets, K., Eskow, E., Fein, L., … Mordechai, E. (2004). Prevalence of borrelia burgdorferi, bartonella spp., babesia microti, and anaplasma phagocytophila in ixodes scapularis ticks collected in northern new jersey [Journal Article]. *J Clin Microbiol*, *42*(6), 2799–2801. <https://doi.org/10.1128/JCM.42.6.2799-2801.2004>

Arsnoe, I. M., Hickling, G. J., Ginsberg, H. S., McElreath, R., & Tsao, J. I. (2015). Different populations of blacklegged tick nymphs exhibit differences in questing behavior that have implications for human lyme disease risk [Journal Article]. *PLoS One*, *10*(5), e0127450. <https://doi.org/10.1371/journal.pone.0127450>

Dantas-Torres, F., Chomel, B. B., & Otranto, D. (2012). Ticks and tick-borne diseases: A one health perspective [Journal Article]. *Trends Parasitol*, *28*(10), 437–446. <https://doi.org/10.1016/j.pt.2012.07.003>

Eisen, R. J., & Eisen, L. (2018). The blacklegged tick, ixodes scapularis: An increasing public health concern [Journal Article]. *Trends Parasitol*, *34*(4), 295–309. <https://doi.org/10.1016/j.pt.2017.12.006>

Ginsberg, H. S., Albert, M., Acevedo, L., Dyer, M. C., Arsnoe, I. M., Tsao, J. I., … LeBrun, R. A. (2017). Environmental factors affecting survival of immature ixodes scapularis and implications for geographical distribution of lyme disease: The climate/behavior hypothesis [Journal Article]. *PLoS One*, *12*(1), e0168723. <https://doi.org/10.1371/journal.pone.0168723>

Jongejan, F., & Uilenberg, G. (2004). The global importance of ticks [Journal Article]. *Parasitology*, *129 Suppl*, S3–14. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/15938502>

Krause, P. J., Carroll, M., Fedorova, N., Brancato, J., Dumouchel, C., Akosa, F., … Lane, R. S. (2018). Human borrelia miyamotoi infection in california: Serodiagnosis is complicated by multiple endemic borrelia species. *PloS One*, *13*(2), e0191725.

Krause, P. J., Narasimhan, S., Wormser, G. P., Barbour, A. G., Platonov, A. E., Brancato, J., … others. (2014). Borrelia miyamotoi sensu lato seroreactivity and seroprevalence in the northeastern united states. *Emerging Infectious Diseases*, *20*(7), 1183.

Moutailler, S., Moro, C. V., Vaumourin, E., Michelet, L., Tran, F. H., Devillers, E., … others. (2016). Co-infection of ticks: The rule rather than the exception. *PLoS Neglected Tropical Diseases*, *10*(3), e0004539.

Vector-Borne Disease, D. of. (2019). [Web Page]. Retrieved from <https://www.cdc.gov/lyme/datasurveillance/index.html>