Examining Tickborne Disease Prevalence and the Effects Habitat and the Microbial Community Play

Julia Frederick

2019-09-13

# Summary/Abstract

*Write a summary of your project.*

Note that this cited reference will show up at the end of the document, the reference formatting is determined by the CSL file specified in the YAML header. Many more style files for almost any journal [are available](https://www.zotero.org/styles). You also specify the location of your bibtex reference file in the YAML. You can call your reference file anything you like, I just used the generic word references.bib but giving it a more descriptive name is probably better.

# Introduction (required for part 1)

## General Background Information

*Provide enough background on your topic that others can understand the why and how of your analysis*  
Tickborne disease prevalence is on the rise in the United States with 50,000 cases being confirmed by the CDC each year. Lyme disease alone is the number one vector borne disease reported in the U.S. each year with over 30,000 cases, and the Center for Disease Control and Prevention (CDC) estimating over 300,000 cases actually occurring [REF]. A single species of tick is capable of transmitting multiple disease pathogens [REFS]. For example, 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 [REF]. 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. When investing 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 [REF], and increased disease threat [REF] 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

*Describe what the data is, what it contains, where it is from, etc.*  
We are currently working to finishing compiling this dataset. This project was begun by prior Master’s student, Madelyn Watson, in Dr. Travis Glenn’s lab. The data was collected but the project did not get completed before Madelyn’s time at UGA. In collaboration with Dr. Michael Yabsley and the Southeastern Cooperative Wildlife Disease Study (SCWDS) ticks were brought in from the eastern region of 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, and used for 16S and PCR in order to determine the presence of pathogens and the microbial community within each specimen. The pathogen and microbial community will not be investigated genetically, instead we will be determining changes in prevalence over time and space. We are currently working to compile all the previous spreadsheets from the different labs into a single location to move forward with analysis. There is the possibility that some variables are in this dataset that had not been discussed in detail with me or we’re forgotten by the current researchers at this time.

## Questions/Hypotheses to be addressed

*State the research questions you plan to answer with this analysis*  
**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. 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. IN order to answer this question we will use an ANOVA, looking at habitat type and specific pathogen.  
**2. How the microbial composition changes with the prevalence of pathogens?**  
This data will give insight into how the microbial community and pathogen prevalence shift with or against each other over a period of time. It will show future intereactions that can be on interest within the microbiome of a tick. We will be using Chi squared tests to look at these changes and the interactions.  
**3. How pathogen prevalence shifts over time within different habitats?**  
This will show how changes in habitat, likely due to temperature and humidity changes, effects the pathogen prevalence in a single area. We will be comparing habitats to each other through ANOVA tests.

# Methods and Results

*In most research papers, results and methods are separate. You can combine them here if you find it easier. You are also welcome to structure things such that those are separate sections.*

## Data aquisition

*As applicable, explain where and how you got the data. If you directly import the data from an online source, you can combine this section with the next.*

## Data import and cleaning

*Write code that reads in the file and cleans it so it’s ready for analysis. Since this will be fairly long code for most datasets, it might be a good idea to have it in one or several R scripts. If that is the case, explain here briefly what each file does. The files themselves should be commented well so everyone can follow along.*

## Univariate analysis

*Use a combination of text/tables/figures to explore and describe your data. You should produce plots or tables or other summary quantities for most of your variables. You definitely need to do it for the important variables, i.e. if you have main exposure or outcome variables, those need to be explored. Depending on the total number of variables in your dataset, explore all or some of the others.*

## Bivariate analysis

*Create plots or tables and compute simple statistics (e.g. t-tests, simple regression model with 1 predictor, etc.) to look for associations between your outcome(s) and each individual predictor variable*

## Full analysis

*Use one or several suitable statistical/machine learning methods to analyze your data and to produce meaningful figures, tables, etc. This might again be code that is best placed in one or several separate R scripts that need to be well documented. You can then load the results produced by this code*

# Discussion

## Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## Conclusions

*What are the main take-home messages?*

*Include citations in your Rmd file using bibtex, the list of references will automatically be placed at the end*

# References