Proceedings of The National Conference
On Undergraduate Research (NCUR) 2020
Montana State University, Bozeman MT
March 26-28, 2020

Molecular Diversity of West Nile Virus in Montana

Julia Sciuchetti
Department of Life and Environmental Sciences
Carroll College
1601 N Benton Ave.
Helena, MT 59625 USA

Faculty Advisor: Dr. Jennifer Glowienka

Abstract

West Nile virus (WNV) has been present in the United States since 1999 and like many RNA viruses, it is known to have high mutation rates. At Carroll College, a decade long project continues in which the state of Montana is surveyed for the presence of WNV. This project created and continues to refine a risk assessment map to predict where outbreaks will occur in a given year. The goal of this particular study was to assess and compare genetic variation of West Nile virus sequences found within geographically distinct mosquito populations across the state of Montana using the highly variable *env* gene. Two different primer sets were utilized to perform reverse transcription and DNA amplification on the viral RNA, which was followed by gel electrophoresis to visualize the amplifications. Amplification was achieved in a subset of samples and the products were sequenced, although the resulting sequences were of low quality. Knowledge about the evolution and diversification of the WNV in Montana will contribute to refinement of the risk assessment model, which should help in making better predictions of WNV outbreaks.

Keywords: West Nile Virus, Montana, Evolution

1. Introduction

The West Nile virus that once emerged in North America in New York in 1999 is now spread throughout the world[1]. From 1999 to 2017, there have been 22,999 reported cases of neuroinvasive West Nile virus (WNV) disease in the United States[11]. As of September 4th, 2019, the Centers for Disease Control and Prevention (CDC) reported 326 cases of WNV disease in humans and 45 states have reported infections in birds, mosquitos or humans[4].

West Nile virus is spread through cycles of mosquito vectors and avian hosts and can be passed to humans through bites from infected mosquitos[3]. West Nile virus expresses most severe disease and highest mortality rates in humans and horses[5]. Categorized as a single stranded, positive sense RNA virus that is a member of the Japanese encephalitis serocomplex in the *Flavivirus* genus and Flaviviridae family, WNV disease can be divided into two classifications: neuroinvasive and non-neuroinvasive[2,4]. Neuroinvasive disease manifests as meningitis, acute flaccid paralysis or encephalitis. Non-neuroinvasive manifests as Febrile illness with symptoms such as headache, body aches, joint pains, vomiting, diarrhea or rash with fatigue and weakness[4]. Eighty percent of people infected with the West Nile virus will show no symptoms, but in the remaining twenty percent, the symptoms can be severe[4].

Localized environmental conditions can affect viral persistence and allow for predictive models of outbreaks[3]. Infection risk is not uniform across a state, but the distribution of WNV is dependent on several geographical, biological and climatic factors[8]. These models can be applied to indicate where geographically equine vaccination efforts should be implemented in a particular state[8]. In Helena, MT, these models are currently being constructed and refined in the state of Montana through long term research conducted at Carroll College.

Even with predictive models, outbreaks can still be random and are difficult to control[5]. The Centers for Disease Control and Prevention has surveillance programs throughout the United States that track both the virus and the disease[4]. Equine vaccines have shown great promise, but research continues for a human counterpart[12].

The WNV is rapidly diversifying and evolving[2]. RNA viruses are known to frequently have high mutation rates due to RNA polymerase's inability to proofread to correct mistakes[7]. The West Nile virus has evolved at a rate of about $0.85*10^{-3}$ substitutions per site per year, which is similar to the rate of evolution in the human Influenza B virus at $1.1*10^{-3}$ substitutions per site per year. In contrast, the rate of evolution of the Measles virus is about $0.4*10^{-3}$ substitutions per site per year[2]. The Influenza B virus has a new vaccination every year because of this high mutation rate whereas the Measles virus has a slower mutation rate and needs fewer vaccinations. The high mutation rate on the WNV helps to explain the difficultly in creating an effective vaccination comparable to the influenza B virus.

The first goal of this study is to compare variation geographically of the WNV found within mosquito populations across the state of Montana. This comparison will be achieved through genetic analysis of viral RNA of the WNV envelope gene because it has great variability and can give phylogenic information comparable to whole genome studies[2]. The ENV gene has a sequence length of 1575 nucleotides, which is around 14% of the whole genome that consists of about 11029 base pairs[2]. The gene product is a 53 kDa protein that is the target for virus-neutralizing antibodies in possible vaccination efforts[2].

The predicted outcome is that different geographic regions will have significantly different variants of WNV. I hypothesize that the vector mosquito populations should be geographically isolated from each other thus, making the West Nile viruses of the different areas isolated allowing the virus to accumulate different mutations in the variable envelope region.

The second goal of this study is to make a significant contribution to the ongoing project at Carroll College of refining the developed risk assessment map for predicting West Nile virus outbreaks in the state of Montana (Hokit et al. 2013). Knowing more about the specifics of the West Nile virus including its evolution and diversification could help improve current predictive models by indicating the degree of panmixia in WNV. By knowing the WNV sequences from mosquito populations from different geographical locations across Montana, comparison could give insight to how the mosquito and avian populations are interacting. WNV sequences would be expected to be similar between populations geographically close together compared to populations that are isolated either by physical distance or geographic barrier. Tracking population interactions could be informative to overall predictions by understanding how the virus is spreading.

2. Materials and Methods

2.1. Mosquito Collections

Mosquitos were trapped around the state of Montana from June to August of 2019 at trapping sites shown in Figure 1. Figure 1 is a risk assessment map showing areas of low, moderate, high and highest risk for the occurrence of WNV across the state of Montana. The map was generated by applying niche modeling tools to presence-only data using historic and ongoing mosquito surveillance for WNV in the state of Montana. Data on climate, bird-host distribution and land cover were used to build a habitat suitability model for the mosquito vector *Culex tarsalis* and then a WNV risk model for the state (Hokit, unpublished data). Trapped mosquitos were baited using carbon dioxide gas tanks or dry ice over a fan apparatus that effectively sucks the mosquitos into a collection net. The traps were left out over night for approximately 12 hours and then put immediately on ice after collection. Previously trapped mosquitos were also shipped to Carroll College by partner trappers around the state with live mosquitos mailed in coolers over ice. In the lab, the live mosquitos were euthanized in a -80°C freezer for approximately 2 hours to keep the virus intact with mosquito death. The mosquitos were then sorted under microscopes by species over ice packs to identify *Culex tarsalis* and *Culex pipiens* individuals based on morphological traits. After sorting, mosquitos were held in the -20°C freezer placed in pools of ≤50 mosquitos in labeled falcon tubes with their assigned number that corresponded to their trapping location.

2.2. Homogenization

Associated pools of mosquitos where placed in 1.5 mL vials labeled with their identification numbers, and those tubes containing less than 10 mosquitos were also specially marked since they required a different concentration of the reagents described below. A ceramic bead was then placed in all tubes (Qiagen Inc, Valancia, CA, USA). The tubes

with greater than 10 were treated with of 1000 μL of RNAlater and 500 μL of Buffer RLT (Qiagen Inc, Valancia, CA, USA) while the tubes with less than 10 mosquitos were treated with 600 μL of RNAlater and 300 μL of Buffer RLT. Samples were then placed in a FastPrep PF120 (ThermoSavent, Massachusetts) at top speed for 30 seconds. The homogenized samples were then moved to a laminar flow hood, where 300 μL was transferred into appropriately labeled microcentrifuge tubes to be delivered to the Laboratory Services at the State of Montana Department of Health and Human Services (DPHHS). Another 300 μL was pipetted out into appropriately labeled microcentrifuge tubes for RNA extraction and PCR by Carroll College. The remaining homogenate was stored in the -80°C freezer at Carroll College. The homogenates for the Montana State lab were also stored in the -80°C until they were ready to be hand delivered over ice packs in a cooler to the state lab.

2.3. RNA Extraction

Following the protocol outlined by Lanciotti et al. (2000), the labeled microcentrifuge tubes were placed into a QIACUBE (Qiagen Inc. Valencia, CA, USA)[9]. The QIACUBE miniprep protocol automated the extraction process and RNA extraction proceeded according to QIACUBE protocol. The RNA extraction then was stored at -80°C to maintain viral integrity.

2.4. Viral Detection

Extracted RNA samples were then analyzed for the presence of WNV by using a Taqman RT-PCR assay that made use of the same primers (Envelop and 3'), solution concentrations and protocol as described by Lanciotti et al. (2000)[9]. Control measures were taken through duplication of four total wells of each sample, two ENV and two 3' primers, in hopes of preventing false positives. All RT-PCR experiments included positive control reactions that used RNA samples that were positively detected for WNV in previous years and all negative control reactions used water as template instead of RNA. All RT-PCR reactions were assembled in a sterile laminar flow hood in order to prevent contamination. The samples were then covered and tested in an IQ5™ (Bio-Rad, Hercules, CA, USA).

2.5. RT-PCR

Using extracted RNA with positive detection of WNV, the same primers and the protocol for reverse transcription and polymerase chain reaction were used following Bertolotti et al. (2007) as shown in Table 1[2]. This included amplification of two overlapping regions of the envelop gene using the QIAGEN OneStep RT-PCR kit with 2.5 μL RNA template. The amplification was done in order to have sufficient WNV RNA to reverse transcribe into DNA and amplify for sequencing of the two regions of the *env* gene for comparison. Reactions were performed in 25 μL or 50 μL volumes using a Bio-Rad Thermocycler (Bio-Rad, Hercules, CA, USA). The resulting amplicons were then analyzed using gel electrophoresis with 1% agarose gels that contained 2 μL SybrSafe (Invitrogen) per 50mL of agarose. Each lane had 4 μL of sample with 2 μL of purple dye. The first lane of each gel had a 1000 bp ladder. The gels electrophoresed at 110V for 30 minutes, and then was visualized under Trans UV light. After visualization, 15 μL of ExoSap-It PCR Cleanup Reagent (ThermoFisher) were added to amplified samples and reacted in the BioRad Thermocycler.

After isolating and purifying RT-PCR product amplifications, 100 μL of 10μM primer for the 1700 and 871 primer set and 50μL of 10 μM primer for the 1595 and 2666 primer set were sent to Macrogen, USA for sequencing.

Table 1**:** Sequences of primers used for WNV gene RT-PCR and PCR amplification following Bertolotti et al. (2007)[2].

	Primer Name	Primer Sequence (5'-3')	Direction
1st Primer Set	WNV-871	CTGGTGGCAGCCGTCATTGGTTGG	Forward
	WNV-1700c	CACACGCCACGAAGCAGTCTG	Reverse
2nd Primer Set	WNV-1595	CGTTCTTGGTCCATCGTGAGTG	Forward
	WNV-2666c	AAATGTGGGAAGCAGTGAAGGACG	Reverse

3. Results

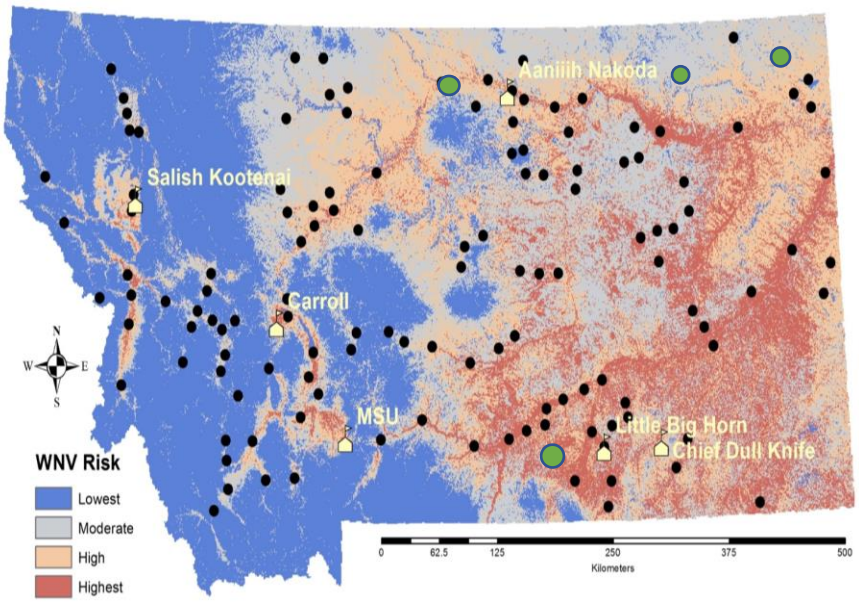

Figure 1: Map of Montana indicating where risk of WNV is highest (Hokit, unpublished data). The black dots represent places where WNV has been found since surveillance began in 2009. The green dots represent where WNV was found the summer of 2019: Yellowstone, Sheridan, Blaine and Valley Counties.

3.1. West Nile Virus Detection

A total of 423 sample pools of mosquitos from across the state of Montana were tested in 2019 using RT-PCR and TaqMan assays, and seven had isolated West Nile Virus representing four counties. As seen in Table 2, three samples from Valley County (sample number:m827, m849, and m850), one sample from Blaine County (sample number:3731A), two samples from Sheridan County (sample number:3718A and 3718D), and one sample from Yellowstone County (sample number:3728G) had positively detected West Nile virus in 2019.

Table 2: Positive West Nile Virus Detection 2019

County	Sample Number
Valley County	m827, m849, m850
Blaine County	3731A
Sheridan County	3718A, 3718D
Yellowstone County	3728G

3.2. Sequence Amplification

Initial experiments yielded no RT-PCR amplified products with no banding patterns on agarose gels, including lack of ladder florescence. After eleven trials, it was discovered the viewing camera was nonfunctional. New gels were then run again and viewed using TransUV illumination with a different camera. As shown in Figure 2, there was RT-PCR product amplification of 871 and 1700 base pair primer sets. The desired base pair length of the product can be seen at 829 base pairs by using the 100 base pair ladder in lane one as a reference (Figure 2). Lanes 2-4 have amplification of the desired length (Figure 2). There were many smaller undesired fragments also amplified in lanes 2-6 (Figure 2). No initial amplifications were seen for the 2666 and 1595 base pair primer sets as seen in Figure 3. After repeating RT-PCR and gel electrophoresis, amplification did occur for the 2666 and 1595 base pair primers as seen in Figure 4. The desired amplification length was 1071 base pairs. Figure 4 shows lane 2 had desired amplification and lanes 2-6 have undesired amplification of smaller fragments. For the RT-PCR step, the gels in Figures 2, 3 and 4 utilized 25 μL reaction volumes. RT-PCR and gel electrophoresis were repeated four more times with 50μL reaction

volumes until there was amplification for desired products of the correct lengths that were visualized as seen in Figure 5.

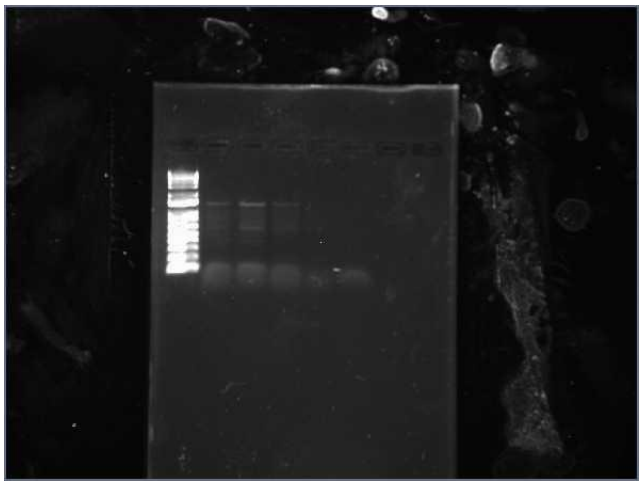

Figure 2: Agarose gel visualizing RT-PCR products using 871 and 1700 base pair primer sets. A 100BP ladder is in well 1. Well identities starting in lane 2 are 3718D, 3731A,2718A, 3728G, and Negative Control.

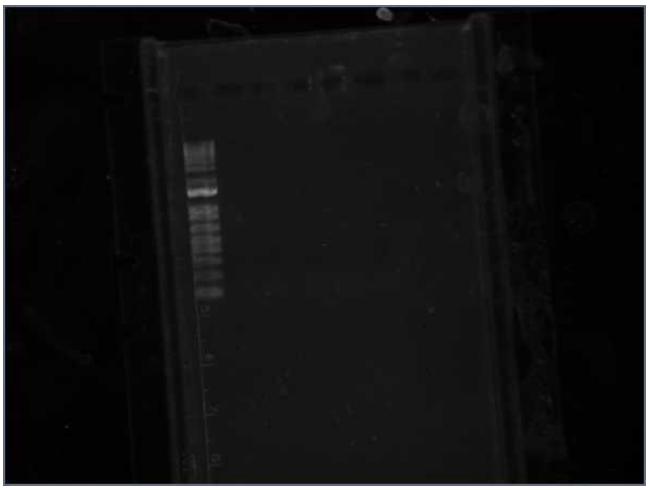

Figure 3: Agarose gel visualizing RT-PCR products using 2666 and 1595 base pair primer sets. A 100BP ladder is in well 1. Well identities starting in lane 2 are 3718D, 3731A,2718A, 3728G, and Negative Control.

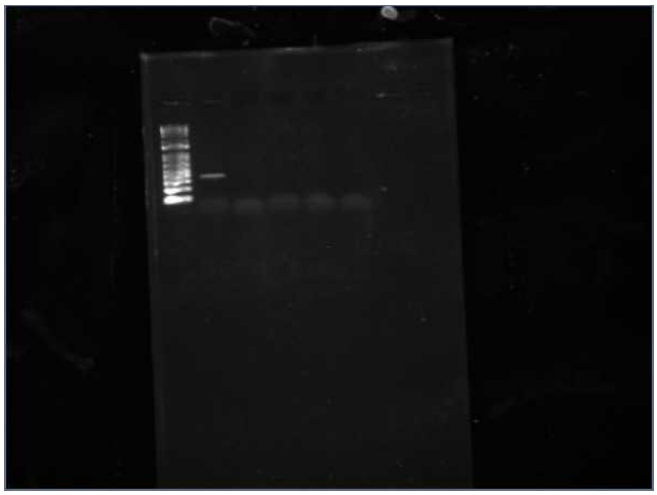

Figure 4: Agarose gel visualizing RT-PCR products using 2666 and 1595 base pair primer sets. A 100BP ladder is in well 1. Well identities starting in lane 2 are 3718D, 3731A,2718A, 3728G, and Negative Control.

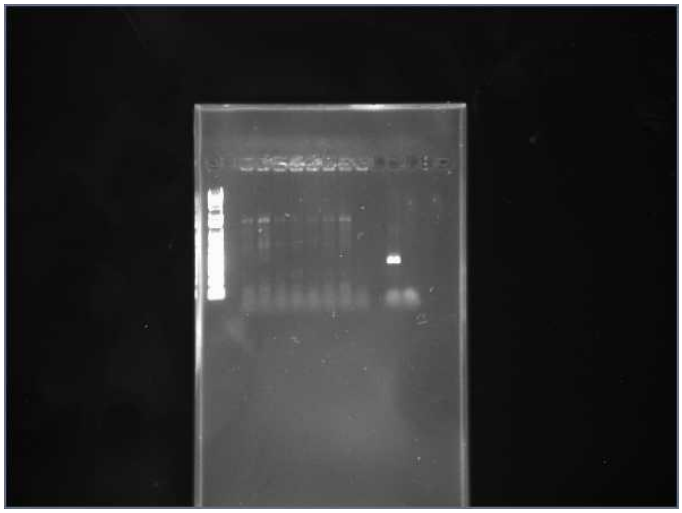

Figure 5**:** Agarose gel visualizing RT-PCR products using 871 and 1700 in wells 2-8 and 2666 and 1595 base pair primer sets in well 10. A 100BP ladder is in well 1. Well identities starting in lane 2 are 3718D, 3731A,2718A, 3728G, M827, M849, M850, Negative Control, 3718D and Negative Control.

3.3. Sequencing Comparison

The sequences returned were of low quality and were unable to be analyzed or compared.

4. Discussion

This study had two aims: 1) to compare sequence diversity of positively detected West Nile Virus samples from 2019 across the state of Montana and 2) to contribute to the refining of the previously developed risk assessment map (Hokit, unpublished data) for predicting West Nile virus outbreaks in the state of Montana. Conclusions about sequence diversity are unable to be drawn as sequenced samples were unable to be compared. The hypothesis that different geographic locations will have significantly different WNV sequences can neither be supported nor rejected until further tests are preformed, and the protocol can be further refined. This study did support the predictive power

of the previously developed risk assessment map as four 2019 WNV pools tested positive and were all in areas of high and highest model predicted risk (Figure 1).

Limitations of this study included difficulty with the Trans-UV viewing camera to visualize fluorescing DNA segments in agarose gels. Another limitation was that actual RT-PCR protocol itself did not always foster amplification of the correct length, and for one of the data sets multiple bands were present. Undesired amplification caused difficulties in sequencing, making sequence comparison impossible. In addition, the RT-PCR was originally preformed on pools of 50 mosquitos, so this could cause limitations in sequencing in that it may be difficult to analyze and separate 50 separate individual sequences. Agarose gel extraction of the correct size PCR product and reamplification and/or cloning could help achieve sequencing the desired product.

Because there was amplification of the desired length for both the 871 and 1700 base pair primer set and the 2666 and 1595 base pair primer set (829 base pairs and 1071 base pairs respectively), there is potential for comparisons to be informative once sequence data is acquired. Knowing how sequences compare within the state of Montana could be useful for understanding the distribution and evolution of the West Nile virus in Montana. This information could support the larger project goal of refining the risk assessment map for Montana that can be used to predict when and where there will be outbreaks of the WNV. There were differences in the geographic location of the positive virus carrying populations. While all were located East of the Rockies, one location was much further South (Yellowstone County) and the other locations were far North (Valley, Blaine and Sheridan Counties). The expectation is that the population in the South would be significantly different from those in the North because the mosquito and avian populations would be isolated and not interacting with each other. However, in the North, it is expected that populations would be more similar in sequence because those populations have the potential to interact.

Once sequence comparison data has been acquired, it would be useful to compare the state of Montana to other states. Bertolotti et al. (2007) found that geographic sub-division had little effect on the variability of sequences from the state of Illinois[2]. They believe this is due to the relatively high, steady mutation rate of the virus at $0.85*10^{-3}$ substitutions per site per year. Ebel et al. (2004) also found viral populations to be homogenous throughout the state of New York, but they believe a new genotype was introduced in 2002 that displaced many previous West Nile viruses in the state that had been around since 1999 because of its more rapid and facilitated transmission[6]. New York had more viral diversity in the years previous to the introduction to the new genotype[6]. The difference in the genotype was located in the envelop gene, which Ebel et al. (2004) believe was under immune selection because of its antigenic importance[6]. The appearance of this new genotype also coincided with more cases of human West Nile virus disease[6]. Both of these studies demonstrate that the WNV can rapidly disperse and evolve in different areas. Through research on the evolution and variability of the WNV, this study could have further application in future possible human vaccination development efforts.

Understanding the transmission and evolution of WNV will help to predict where outbreaks will occur in the state of Montana. Better prediction will allow for improved public health measures to protect both human and equine populations that are most vulnerable. The locations of mosquito populations that tested positive for WNV as shown in Table 2, support and contribute to the current risk assessment map (Figure 1). More data will continue to help refine the map and allow for better and more accurate predictions. Future research for this project, beyond current sequencing and comparison methods, would be to compare positive WNV populations East and West of the Rocky Mountains as the mountains are a much larger geographic barrier than exists in Eastern Montana. Such a study could demonstrate more diversity in WNV in the state of Montana. In addition, other genes or whole genome sequencing efforts could be tested and used to measure diversity. Finally, more replicates of samples and samples from future and past years could be used to determine how the WNV is diversifying and evolving over time.

5. Conclusions

The first goal of this study was to compare variation geographically of the WNV found within mosquito populations across the state of Montana. This goal has inconclusive results as sequenced samples were unable to be analyzed. The hypothesis that different geographic locations will have significantly different WNV sequences is currently unable to be supported until further tests are preformed, and the protocol can be further refined.

The second goal of this study was to make a significant contribution to the ongoing project at Carroll College of refining the developed risk assessment map for predicting West Nile virus outbreaks in the state of Montana. The locations of mosquito populations that tested positive for WNV used in this study support and contribute to the current risk assessment map project. More data will continue to help refine the map and allow for better and more accurate predictions of WNV outbreaks in the state of Montana.

6. Acknowledgements

The author wishes to express appreciation to the National Conference on Undergraduate Research as well as would like to thank Dr. Jennifer Glowienka for her continuous help and support over the last four years. Thank you for being a mentor and helping to think of solutions to all the different problems that were faced. The author would like to thank both Dr. Jennifer Glowienka and Dr. Grant Hokit for their guidance over the last two summers while working together on this project. The author would like to thank Dr. Grant Hokit as well as Dr. Katherine Greiner for reviewing this research and giving helpful feedback. The author would also like to thank Dr. Gerald Shields for sharing the lab and continually giving a good laugh. In addition, the author would like to thank her peers for their help over the past two summers, including Brandon Brown, Ember Drivdahl, Kendall Fredrics, Patrick Oster, Kaitlin Stromberg and Maddie Nye.

The author would also like to thank the MSU extension offices, Aaniiih Nakoda College, Chief Dull Knife College, Little Big Horn College, and Salish Kootenai College, and all who helped contribute to the trapping and collection of mosquitoes. This research is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM103474.

7. References

1. Beasley D.W.C., Davis, C.T., Guzman, H., Vanlandingham, D.L., Travassos da Rosa, A.P.A., Parsons, R.E., Higgs, S., Tesh, R.B., & Barrett, A.D.T. (2003). Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States, *Virology*, 309(2), 190-195, ISSN 0042-6822.

2. Bertolotti, L., Kitron, U., & Goldberg, T.L. (2007). Diversity and evolution of West Nile virus in Illinois and the United States 2002-2005, *Virology*, 360(1), 143-149.

3. Bertolotti, L., Kitron, U.D., Walker, E.D., Ruiz, M.O., Brawn, J.D., Loss, S.R., Hamer, G.L. & Goldberg, T.L., (2008) Fine-scale genetic variation and evolution of West Nile Virus in a transmission "hot spot" in suburban Chicago, USA, *Virology*, 374(2), 381-389.

4. CDC, Preliminary Maps & Data for 2019 | West Nile Virus | (2019, September 17). Retrieved from https://www.cdc.gov/westnile/statsmaps/preliminarymapsdata2019/index.html

5. Dauphin, G., Zientara, S., Zeller, H., & Murgue, B. (2004). West Nile: worldwide current situation in animals and humans, *Comparative Immunology, Microbiology and Infectious Diseases*, 27, (5), 343-355, ISSN 0147-9571

6. Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A., & Kramer, L.D. (2004). Genetic and Phenotypic Variation of West Nile Virus in New York, 2000-2003. *The American Journal of Tropical Medicine and Hygiene*, 71(4), 493-500.

7. Ebel, G. D., Dupuis, A. P., 2nd, Ngo, K., Nicholas, D., Kauffman, E., Jones, S. A., & Kramer, L. D. (2001). Partial genetic characterization of West Nile virus strains, New York State, 2000. *Emerging infectious diseases*, *7*(4), 650–653. doi:10.3201/eid0704.010408

8. Hokit, D.G., Alvey, S., Geiger, J.M.O., Johnson, G.D., Rolston, M.G., Kinsey, D.T. & Bear, N.T. (2013). Using Undergraduate Researchers to Build Vector and West Nile Virus Surveillance Capacity. *International Journal of Environmental Research and Public Health*, 10, 3192,3202.

9. Lanciotti, R.S., Kerst, A.J., Nasci, R.S., et al. Rapid detection of west nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. J Clin Microbiol. 2000;38(11):4066–4071.

10. Pesko, K. N., & Ebel, G. D. (2012). West Nile virus population genetics and evolution. *Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases*, *12*(2), 181-190. doi:10.1016/j.meegid.2011.11.014.

11. Petersen, L. R. (2019). "Epidemiology of West Nile Virus in the United States: Implications for Arbovirology and Public Health." Journal of Medical Entomology, vol. 56, no. 6, 24 Sept. 2019, pp. 1456–1462., doi:10.1093/jme/tjz085.

12. Tesh, R. B., Arroyo, J., Travassos Da Rosa, A. P., Guzman, H., Xiao, S. Y., & Monath, T. P. (2002). Efficacy of killed virus vaccine, live attenuated chimeric virus vaccine, and passive immunization for prevention of West Nile virus encephalitis in hamster model. *Emerging infectious diseases*, *8*(12), 1392–1397.doi:10.3201/eid0812.020229