The Analysis of the Ebony Gene in *Culex tarsalis* Across the Major Drainage Basins of Montana

Life and Environmental Sciences, Carroll College, Helena, MT
BI-499 Senior Thesis

Nathaniel Zebulon Antonioli
Carroll College, Helena MT
April 29, 2022

Table of Contents

Acknowledgements

A huge thank you to Dr. Hokit for his mentorship, assistance with the data analysis, and hiring me to be on the West Nile Virus research team for the past two summers. This has been a very fruitful endeavor thanks to Dr. Hokit. Thank you to Dr. Otto-Hitt for her assistance and guidance with the RT-PCR, other molecular techniques, and her assistance with the many issues that arose during this research. Thank you to my parents for the support throughout my college and research career. I would not be where I am today without their guidance and support. Thank you to all of my professors and especially my advisor Dr. Carpenter for educating me in the fields of Biochemistry and Biology and for being a reader for this thesis. A final thank you to Montana INBRE and all those who have worked on the West Nile Virus research team.

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM103474. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

Many diseases are transmitted by vectors one of which is West Nile Virus (WNV). WNV has been in the United States since 1999 and Montana since 2002. The main vector for WNV in Montana is *Culex tarsalis*. Carroll College monitors the presence and range of WNV every summer since 2009. Better understanding of the vector can help to control the spread and frequency of the disease. Previous studies into the population genetics of *Cx. tarsails* in Montana have not yielded much insight. The identification of possible isolated populations with the use of naturally occurring variations within the species can be used to direct future inquiry. The analysis of expression of pigmentation via the ebony gene was used to compare the Upper Missouri, Yellowstone, Clark Fork, and Headwaters drainage basins as well as individual sites to determine if any were significantly different. This was done by RT-PCR with RNA extractions. The results suggested no significant differences between individual sites, but differences were suggested between the Clark Fork and Headwaters as well as the Yellowstone and Headwaters drainage basins. The addition of the other major drainage basins and biological replicates of sites to expand on both data sets is the goal for future inquiry.

Introduction

Many diseases need a biological intermediate or vector that carries a pathogen from organism to organism and by biting or exchanging of other bodily fluids, infect other individuals. Vectors for disease are always an interest to research especially when the disease can be lethal in humans. One vector that is known to transmit various diseases is the mosquito. Mosquitoes can transmit a variety of diseases such as Yellow Fever, Malaria, Zika Virus, and Dengue Fever according to the CDC just to name a few (CDC, 2016). Due to the possible adverse effects of being infected with any of these diseases, there is much interest in understanding the vector and the disease itself to prevent further loss of life and infection. An additional disease that can be transmitted by mosquito vectors is West Nile Virus (WNV; CDC, 2016).

WNV was first discovered in Uganda in the 1930's. Throughout the 20th century, many outbreaks were studied spanning from Northern Africa, to the Middle East, and even to the Mediterranean. Observations gathered through these outbreaks demonstrated that the virus could infect many animals including birds and other non-human mammals; however, the virus was only able to be isolated from mosquito samples and no other insects. This was evidence that mosquitoes were the primary vector for WNV. West Nile Virus was also shown to cause encephalitis or swelling of the brain which could lead to death in severe cases. WNV continued to infect the aforementioned areas as well as being reported as far north as Romania. WNV was first introduced in the United States in 1999 when a cluster of encephalitis cases occurred in New York in tandem with many avian deaths. From that point on, WNV spread across the US with the most notable year being the summer of 2002 when WNV spread as far westward as the Continental Divide with

over 4,000 cases being reported in the US alone, according to the CDC (Sievar, 2003; CDC, 2021). This was the same year when WNV was detected in Montana (Johnson et al., 2010). There are two species of mosquito that are known to carry WNV in the state of Montana, but the most prolific and common, which is found in all 56 counties of Montana, is *Culex tarsalis*. The second is *Culex pipiens*.

Since this initial detection in Montana, yearly monitoring for the disease within *Cx. tarsalis* has been carried out. WNV was found to be carried in local bird populations and then is carried by a mosquito that has fed on the blood of an infected bird. WNV can infect a variety of mammals, but the most common infections documented are those of humans and horses with a fatality rate in horses around 33% and in humans around 5% (Swinker, 2022; CDC, 2021). In the United States, there have been 2,456 deaths attributed to WNV between 1999 and 2020 with Montana reporting 622 cases in this same time period, according to the CDC (2021). Between 2014 and 2020, Montana has experienced 3 deaths due to WNV according to the Montana Department of Public Health and Human Services (DPHHS). The most recent positives were detected during summer of 2021 which included Eastern Montana and as far west as Gallatin County.

Due to the widespread nature of the disease and its main vector *Cx. tarsalis*, it remains of the upmost importance to gain a better understanding of the ecological factors and how these influence the biological factors such as genetic exchange between populations or selective pressures that keep populations relatively isolated. Understanding these relationships can help to better manage the spread of WNV across the state by controlling the vector *Cx. tarsalis* in various ways.

Since 2009, Carroll College has conducted WNV surveillance. Over the years, many different models have been used to show various ecological relationships as well as some exploration into the genetic relations of the populations of *Cx. tarsalis* such as the analysis of COI sequences by Fiocchi (2016) and microsatellite sites by Stromberg (2019). The genetic models have yielded little information into the distinct population structure of *Cx. tarsalis*. Similar models employed in other parts of the United States have worked such as the model of the Sonra Desert populations using COI sequences (Pfeiler et al., 2013), but Montana seems to be a difficult area to distinguish genetic relatedness and gene flow between populations with the same methods. Populations that are distinct usually have varied features due to either separation by time and distance which allows genetic drift to occur, or by selective pressures that can prevent other individuals from other populations with a different characteristic from adding their genes to the population's gene pool.

One way to help determine distinct populations is by looking at the natural variations that are localized to certain areas such as differences in pigmentation, body size, and alterations in color pattern which are reflected in the genetic code. In this study, pigmentation was used, and more specifically, the analysis of the ebony gene responsible for darker coloration. The reason behind this choice was due to my initial observations that many *Cx. tarsalis* across the state varied greatly between very dark markings and more tan colored markings. The ebony gene has been studied in a closely related mosquito species which allowed for the analysis of the gene's expression (Feng et al., 2021). Measuring the expression of the ebony gene can help to determine possible distinct populations of *Cx. tarsalis*. These populations can then be targeted to elucidate a

more effective model for the gene flow and relatedness of *Cx. tarsalis* across Montana. The hypothesis was that expression of the ebony gene could be used to distinguish between populations of *Cx. tarsalis* in Montana by the observation of significantly different expression of the gene between drainage basins and specific sites.

Methods & Materials

Mosquitoes were trapped from many locations across the state. These sites were either historic sites known to host *Cx. tarsalis* or new sites intended to explore the distribution boundaries of the species. The trapping process used the CDC Light Trap baited with CO_2 as described by Hokit et al. (2013). The traps were set an hour before sunset and were collected the next morning no later than 9:00 a.m. The traps were then placed into a cooler with ice as to keep the samples preserved and prevent degradation during transport to the lab.

The mosquitoes were euthanized in the -80 °C freezer for one hour or the -25 °C freezer for three hours and then were sorted by species by the same protocol described by Hokit et al. (2013). The target species *Cx. tarsalis* was isolated and separated from the rest of the sample and put into collection tubes with a maximum of 50 mosquitoes per collection tube. If a sample had more than 50 mosquitoes, the sample was broken into subsamples. Each sample also had a 5 mm stainless steel bead from Qiagen added as a means of homogenization for RNA extraction. These samples were then kept in the -80 °C freezer until RNA extractions were performed. The samples selected for extraction came from the Upper Missouri, Yellowstone, Clark Fork, and Headwaters drainage basins. The samples had to contain at least 5 mosquitoes. (**Figure 1**, **Tables 1-4**)

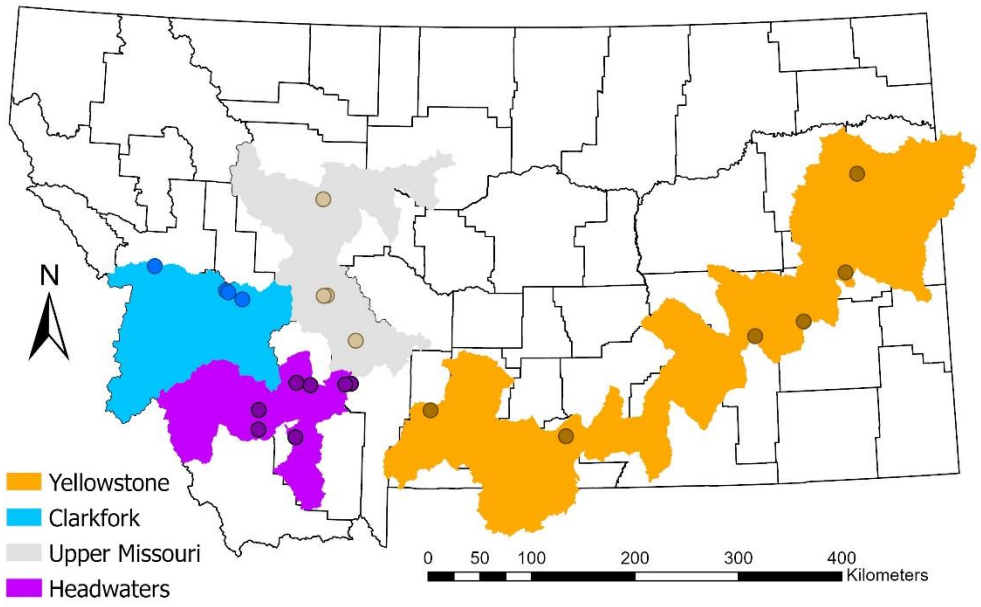

Figure 1: The map above shows the four drainages of the Yellowstone, Clark Fork, Upper Missouri, and Headwaters. The dots represent the individual sites of collection within each drainage basin.

Table 1: The table shows the county, location, date, sample number, and number of *Cx. tarsalis* within the individual samples selected for the Upper Missouri drainage basin.

Drainage	County	Location	Date Sampled	Sample #	# of *C. tarsalis*
Upper Missouri	Broadwater	York Island	8/14/2021	4381	8
	Cascade	West Simms	6/24/2021	4078	14
	Lewis & Clark	Gould	70/9/2021	4183	50
	Lewis & Clark	Gould	7/26/2021	4259A	50
	Lewis & Clark	Gould	7/26/2021	4259B	50
	Lewis & Clark	Norman	7/26/2021	4260A	50
	Lewis & Clark	Norman	7/26/2021	4260B	50
	Lewis & Clark	Norman	8/07/2021	4374	50
	Lewis & Clark	Regulating Reservoir	6/21/2021	4069	19
	Lewis & Clark	Helena Irr District	8/01/2021	4264	16

Table 2: The table shows the county, location, date, sample number, and number of *Cx. tarsalis* within the individual samples selected for the Yellowstone drainage basin.

Drainage	County	Location	Date Sampled	Sample #	# of *C. tarsalis*
Yellowstone	Custer	Spotted Eagle	8/16/2021	4389	27
	Dawson	Richey	8/09/2021	4337	5
	Rosebud	Farwest FA	8/16/2021	4388 A	50
	Rosebud	Farwest FA	8/16/2021	4388 E	50
	Rosebud	Farwest FA	8/16/2021	4388C	50
	Park	Weed Shed	7/14/2021	4174	21
	Prairie	Terry Lagoons	8/16/2021	4390	50
	Carbon	Fulop Drive	7/02/2020	3815	14

Table 3: The table shows the county, location, date, sample number, and number of *Cx. tarsalis* within the individual samples selected for the Clark Fork drainage basin.

Drainage	County	Location	Date Sampled	Sample #	# of *C. tarsalis*
Clark Fork	Granite	Hwy271m1	8/15/2021	4385	20
	Granite	Drummond	7/21/2020	4237	16
	Missoula	Missoula	8/01/2021	4364	50
	Powell	Gold Creek	7/16/2021	4186	43

Table 4: The table shows the county, location, date, sample number, and number of *Cx. tarsalis* within the individual samples selected for the Headwaters drainage basin.

Drainage	County	Location	Date Sampled	Sample #	# of *C. tarsalis*
Headwaters	Beaverhead	Glen	7/12/2021	4142B	50
	Beaverhead	Glen	7/12/2021	4142A	50
	Beaverhead	Glen	7/12/2021	4142C	50
	Silver Bow	Melrose Cemetery	7/19/2021	4191	41
	Gallatin	Leo Lane	8/01/2021	4304	50
	Gallatin	Leo Lane	8/01/2021	4310	50
	Gallatin	Leo Lane	8/09/2021	4401	34
	Gallatin	City	8/01/2021	4343	50
	Gallatin	City	8/01/2021	4344	50
	Gallatin	City	8/01/2021	4345	50
	Gallatin	Clarkston	7/18/2021	4208	50
	Gallatin	Clarkston	8/01/2021	4294	39
	Gallatin	Clarkston	7/11/2021	4154	26
	Jefferson	Antonioli Farm	7/16/2021	4187A	50
	Jefferson	Antonioli Farm	7/16/2021	4187B	50
	Jefferson	Antonioli Farm	8/06/2021	4335	50
	Jefferson	Cassagrands	7/31/2020	3933	39
	Madison	Silver Star Creek	8/13/2020	3980	11

To prepare the samples for RNA extractions, the samples were first combined with the reagents RNAlater from Thermo Fisher Scientific and Buffer RLT from Qiagen. For samples equal or greater than 10 mosquitoes, 1000 µL of RNAlater was added and 500 µL of Buffer RLT. For samples less than 10 mosquitoes, 600 µL of RNAlater was added and 300 µL of Buffer RLT. Samples that had already been extracted for West Nile Virus testing over the summer used the ratio of 2:1 of RNAlater to Buffer RLT to bring the samples back up to an adequate volume for extraction.

The samples were homogenized using a Qiagen TissueLyser LT for one minute at 50 oscillations per second. The samples were then centrifuged at 10,000 rpm for three minutes. For the extractions, 800 µL of supernatant was pipetted into a separate collection tube. At this point, 700 µL of 70% ethanol was added. From the point of homogenization until the addition of ethanol, biosafety level two protocols (BSL2) and personal protective equipment (PPE) were used as to prevent accidental infection by West Nile Virus. The BSL2 protocols included wearing googles, a splashguard apron, an N95 mask, gloves, and the handling of all samples in a laminar flow hood.

The extraction of the RNA from the supernatants was achieved by following the protocol described by Hokit et al. (2013) and the use of the Qiagen RNAeasy kit. The volume of 700 µL of the supernatant and ethanol solution was added to the spin column and centrifuged for 30 seconds at 10,000 rpm. The flow through was discarded. This was repeated with the rest of the sample. Next, 350 µL of Buffer RW1 was added to the spin column and was centrifuged at 10,000 rpm for 30 seconds. The flow through was discarded and the step was repeated. Then, 500 µL of Buffer RPE was added to the spin column and centrifuged for 30 seconds at 10,000 rpm. The flow through was discarded

and the step was repeated, but the time centrifuged was increased to two minutes. The spin column was placed in a microcentrifuge tube and 40 µLs of water was added and centrifuged for one minute. This was repeated once more. The RNA was collected in the microcentrifuge tube. If the RNA concentration was too low, the same procedure was followed, but 40 µL of water were added and allowed to stand for one minute before centrifugation and the sample was only centrifuged once with water as to maximize the amount of RNA collected. Once the final RNA extracts were collected, they were frozen in the -80 °C freezer as to preserve the RNA.

The RNA concentration and purity for each sample was determined by the use of a Thermo Scientific NanoDrop Lite Spectrophotometer. The concentration was measured in ng/µL and the purity was measured in A280/A260 and pure RNA is measured at 2.00 A280/A260. The dilutions of the RNA were found so that 3 µL of RNA solution had 50 ng of RNA. In other words, the concentration of the RNA solution after dilution was 16.667 ng/µL. This was done by the use of **Equation 1** and the amount of RNA extraction was found and the use of **Equation 2** found the amount of H_2O that was added with the RNA extraction volume to have a working RNA solution of 25 µL at a concentration of 16.67 ng/µL. The volumes used in the dilution as well as the stock RNA concentrations are seen on **Tables (5-12)**. The dilution volumes were rounded to the nearest hundredths place to provide values that were able to be accurately pipetted.

Equation 1

$$\frac{\frac{\mathbf{50}\ (\boldsymbol{ng})}{\mathbf{3}\ (\mu\boldsymbol{L})} * \mathbf{25}\ (\mu\boldsymbol{L})}{[\boldsymbol{RNA}](\frac{\boldsymbol{ng}}{\mu\boldsymbol{L}})} = \boldsymbol{RNA}\ (\mu\boldsymbol{L})$$

Equation 2

$$\mathbf{25}\ (\mu\boldsymbol{L}) - \boldsymbol{RNA}\ (\mu\boldsymbol{L}) = \boldsymbol{H_2O}\ (\mu\boldsymbol{L})$$

Table 5: The table shows the location, sample number, RNA concentration in ng/µL, and the RNA purity in the ratio A280/A260 for the Upper Missouri drainage.

Drainage	Location	Sample #	RNA Concentration (ng/µL)	A280/A260
Upper Missouri	York Island	4381	22.1	2.00
	West Simms	4078	45.6	2.00
	Gould	4183	97.6	2.13
	Gould	4259A	65.0	1.99
	Gould	4259B	49.1	2.07
	Norman	4260A	40.7	1.93
	Norman	4260B	41.2	1.99
	Norman	4374	27.0	2.08
	Regulating Reservoir	4069	76.6	2.15
	Helena Irr District	4264	29.8	1.97

Table 6: The table shows the location, sample number, and the RNA and water volumes in µL that were mixed to create the working RNA solutions for the RT-PCR at the RNA concentration 16.667 ng/µL for the Upper Missouri drainage.

Drainage	Location	Sample #	RNA µL	Water µL
Upper Missouri	York Island	4381	18.85	6.15
	West Simms	4078	9.14	15.86
	Gould	4183	4.27	20.73
	Gould	4259A	6.41	18.59
	Gould	4259B	8.49	16.51
	Norman	4260A	10.24	14.76
	Norman	4260B	10.11	14.89
	Norman	4374	15.43	9.57
	Regulating Reservoir	4069	5.44	19.56
	Helena Irr District	4264	13.98	11.02

Table 7: The table shows the location, sample number, RNA concentration in ng/µL, and the RNA purity in the ratio A280/A260 for the Yellowstone drainage.

Drainage	Location	Sample #	RNA Concentration (ng/µL)	A280/A260
Yellowstone	Spotted Eagle	4389	29.8	1.99
	Richey	4337	35.5	1.87
	Farwest FA	4388 A	26.4	2.06
	Farwest FA	4388 E	52.5	2.07
	Farwest FA	4388C	38.6	2.13
	Weed Shed	4174	70.3	2.08
	Terry Lagoons	4390	58.5	2.03
	Fulop Drive	3815	27.2	1.88

Table 8: The table shows the location, sample number, and the RNA and water volumes in µL that were mixed to create the working RNA solutions for the RT-PCR at the RNA concentration 16.667 ng/µL for the Yellowstone drainage.

Drainage	Location	Sample #	RNA µL	Water µL
Yellowstone	Spotted Eagle	4389	13.98	11.02
	Richey	4337	11.74	13.26
	Farwest FA	4388 A	15.78	9.22
	Farwest FA	4388 E	7.94	17.06
	Farwest FA	4388C	10.79	14.21
	Weed Shed	4174	5.93	19.07
	Terry Lagoons	4390	7.12	17.88
	Fulop Drive	3815	15.32	9.68

Table 9: The table shows the location, sample number, RNA concentration in ng/µL, and the RNA purity in the ratio A280/A260 for the Clark Fork drainage.

Drainage	Location	Sample #	RNA Concentration (ng/µL)	A280/A260
Clark Fork	Hwy271m1	4385	27.8	1.97
	Drummond	4237	33.1	2.09
	Missoula	4364	52.0	1.95
	Gold Creek	4186	104.8	2.03

Table 10: The table shows the location, sample number, and the RNA and water volumes in µL that were mixed to create the working RNA solutions for the RT-PCR at the RNA concentration 16.667 ng/µL for the Clark Fork drainage.

Drainage	Location	Sample #	RNA µL	Water µL
Clark Fork	Hwy271m1	4385	14.99	10.01
	Drummond	4237	12.59	12.41
	Missoula	4364	8.01	16.99
	Gold Creek	4186	3.98	21.02

Table 11: The table shows the location, sample number, RNA concentration in ng/µL, and the RNA purity in the ratio A280/A260 for the Headwaters drainage.

Drainage	Location	Sample #	RNA Concentration (ng/µL)	A280/A260
Headwaters	Glen	4142B	53.2	1.95
	Glen	4142A	26.8	2.05
	Glen	4142C	26.8	2.14
	Melrose Cemetery	4191	88.3	1.96
	Leo Lane	4304	37.6	1.98
	Leo Lane	4310	50.0	2.03
	Leo Lane	4401	80.6	1.95
	City	4343	64.9	2.00
	City	4344	41.4	1.95
	City	4345	23.1	1.97
	Clarkston	4208	38.2	1.99
	Clarkston	4294	37.3	1.90
	Clarkston	4154	34.4	1.98
	Antonioli Farm	4187A	89.9	2.02
	Antonioli Farm	4187B	43.2	2.01
	Antonioli Farm	4335	27.2	2.07
	Cassagrands	3933	19.6	1.87
	Silver Star Creek	3980	38.1	1.89

Table 12: The table shows the location, sample number, and the RNA and water volumes in µL that were mixed to create the working RNA solutions for the RT-PCR at the RNA concentration 16.667 ng/µL for the Headwaters drainage.

Drainage	Location	Sample #	RNA µL	Water µL
Headwaters	Glen	4142B	7.83	17.17
	Glen	4142A	15.55	9.45
	Glen	4142C	15.55	9.45
	Melrose Cemetery	4191	4.72	20.28
	Leo Lane	4304	11.08	13.92
	Leo Lane	4310	8.33	16.67
	Leo Lane	4401	5.17	19.83
	City	4343	6.42	18.58
	City	4344	10.06	14.94
	City	4345	18.04	6.96
	Clarkston	4208	10.91	14.09
	Clarkston	4294	11.17	13.83
	Clarkston	4154	12.11	12.89
	Antonioli Farm	4187A	4.63	20.37
	Antonioli Farm	4187B	9.65	15.35
	Antonioli Farm	4335	15.32	9.68
	Cassagrands	3933	21.26	3.74
	Silver Star Creek	3980	10.94	14.06

For the quantitative PCR analysis, primers were designed to target the ebony gene as described by Feng et at. (2021). The primers used were specific to *Culex quinquefasciatus*, but after preliminary testing, the primers were found to amplify the *Cx. tarsalis* homolog. The forward primer was 5'-GAAACGATCCTGAACCGTCTGCAG -3' and the reverse primer was 5'- ATGCTGATGGCGTGACCCTTTTGAC -3'. To normalize the expression of the ebony gene, primers were designed to target the actin gene of *Cx. tarsalis*; the forward primer sequence was 5' – GTCGGTATGGGCAACAAGGA – 3' and the reverse primer sequence was 5' – CGTGCTCGATCGGGTACTTC – 3'. The primers came in 100 mM solutions. This was diluted 10 mM by adding 1.00 µL of the stock primer solution to 9.00 µL of water. This dilution was scaled accordingly to obtain enough working primer solution.

The RT-PCR was executed with the use of the Bio-Rad iTaq Universal SYBR Green One-Step Kit. The 20 µL reaction set up was used to determine the amounts of the reagents as summarized in **Table (13)**. The one note of the reaction set up is that the primer, RNA, and water volumes were the only variable quantities and were determined with the help of Dr. Otto-Hitt. The RNA volume was 3 µL which added the desired 50 ng of RNA to the reaction. The primer volume considered the 10mM concentration and was set to 1.5 µL of both forward and reverse primer.

Table 13: The table shows the volumes of the reagents used in the RT-PCR protocol for the iTaq Universal SYBR Green One-Step Kit 20 µL reaction volume.

Reagent	Volume (µL)
SYBR	10.00
Water	3.75
Reverse Transcriptase	0.25
RNA	3.00 (50 ng RNA or 16.667 ng/µL RNA Concentration)
Primers	1.50 (Forward, 10 mM) & 1.50 (Reverse, 10 mM)

For the RT-PCR protocol, there were two that were used. Both protocols can be seen on **Table (14)**. For the Clark Fork and Yellowstone drainages, **Method #1** was used and for the Headwater and Upper Missouri Drainages, **Method #2** was used. The use of the two separate protocols was possible due to the use of the actin gene as a reference gene to determine expression. The expression should remain the same when based off of the actin gene assuming both genes underwent the same protocol. The completion of each run was followed with the exportation of the data files associated with the PCR including the Cq values which were used in the data analysis.

Table 14: The table shows the two methods used for the RT-PCR runs.

Method	Step 1	Step 2	Step 3	Step 4	Repeats
#1	50 °C:10 min	95 °C:1 min	95 °C:10 s	62 °C:30 s	Steps 3&4:39 Repeats
#2	50 °C:10 min	95 °C:1 min	95 °C:10 s	61 °C:30 s	Steps 3&4:44 Repeats

To test the replicability of the protocols used, the standard deviation was calculated for the Cq values of the actin gene of the technical replicates (**Table 15**) which are defined as three replicate subsamples of the same mosquito pool. Using the criteria described in Brown (2006), the acceptable value of standard deviation (STD) was $\leq$0.25 which indicated consistent sample processing and machine function. The only value to exceed an STD of 0.25 was the Richey site in the Yellowstone drainage. The conversion of Cq to RQ values for the technical replicates was completed by finding the mean of both of the genes. The mean Cq values were then used to find the RQ values as described by the Bio-Rad Real-Time PCR Applications Guide (2006) by the delta Ct or Cq method.

Table 15: The table shows the location, sample number, the actin Cq values for the three technical replicates performed. The Standard Deviation (STD) between the technical replicates is also shown for each location. The STD above 0.25 was bolded and marked with an asterisk.

County	Location	Sample #	Actin 1 Cq	Actin 2 Cq	Actin 3 Cq	Standard Deviation
Upper Missouri	York Island	4381	20.026	20.164	20.082	0.057
	West Simms	4078	20.511	20.471	20.434	0.031
	Regulating Reservoir	4069	19.957	20.138	20.072	0.075
	Helena Irr District	4264	19.750	19.697	19.817	0.049
Yellowstone	Spotted Eagle	4389	17.461	17.306	17.519	0.090
	Richey	4337	16.465	16.488	18.602	**1.002***
	Weed Shed	4174	19.615	19.491	19.560	0.051
	Terry Lagoons	4390	16.827	16.817	16.850	0.014
	Fulop Drive	3815	17.047	16.994	17.052	0.027
Clark Fork	Hwy271m1	4385	17.660	17.597	17.644	0.027
	Drummond	4237	18.720	18.803	18.944	0.093
	Missoula	4364	18.109	18.179	18.108	0.033
	Gold Creek	4186	18.944	18.792	18.678	0.109
Headwaters	Melrose Cemetery	4191	21.454	21.398	21.592	0.082
	Cassagrands	3933	19.670	19.889	19.697	0.097
	Silver Star Creek	3980	19.084	19.083	19.121	0.018

Table 16: The table shows the locations and sample numbers of the technical replicates. These are shown with their associated actin and ebony Cq values.

County	Location	Sample #	Actin 1 Cq	Actin 2 Cq	Actin 3 Cq	Ebony 1 Cq	Ebony 2 Cq	Ebony 3 Cq
Upper Missouri	York Island	4381	20.026	20.164	20.082	37.232	37.726	38.143
	West Simms	4078	20.511	20.471	20.434	36.049	37.201	38.482
	Regulating Reservoir	4069	19.957	20.138	20.072	39.518	39.967	37.002
	Helena Irr District	4264	19.750	19.697	19.817	38.108	38.569	37.755
Yellowstone	Spotted Eagle	4389	17.461	17.306	17.519	32.558	34.623	32.014
	Richey	4337	16.465	16.488	18.602	34.034	35.353	36.208
	Weed Shed	4174	19.615	19.491	19.560	37.497	33.186	37.380
	Terry Lagoons	4390	16.827	16.817	16.850	32.949	38.535	36.265
	Fulop Drive	3815	17.047	16.994	17.052	34.281	31.527	33.659
Clark Fork	Hwy271m1	4385	17.660	17.597	17.644	36.079	37.42	34.785
	Drummond	4237	18.720	18.803	18.944	34.885	35.637	38.020
	Missoula	4364	18.109	18.179	18.108	36.776	38.809	39.229
	Gold Creek	4186	18.944	18.792	18.678	34.943	36.704	39.68
Headwaters	Melrose Cemetery	4191	21.454	21.398	21.592	36.297	35.302	36.622
	Cassagrands	3933	19.670	19.889	19.697	35.682	36.680	35.184
	Silver Star Creek	3980	19.084	19.083	19.121	30.142	35.743	33.633

Biological replicates were processed from independent mosquito pools collected from the same site. The biological replicate RQ values were calculated using the same delta Cq method, but independent samples rather than the average as in the technical replicates (**Table 16**). RQ values were organized by drainage including the Yellowstone, Clark Fork, Upper Missouri, and Headwaters drainage basins (**Tables 17-21**). The biological replicate RQ values were then organized in **Table (22)**. To test the hypothesis that the ebony gene varied between sites and drainage basins, the analysis of variance (AVOVA) test was used.

Table 17: The table shows the locations and sample numbers of the biological replicates. These are shown with their associated actin and ebony Cq values.

Drainage	Location	Sample #	Actin Cq	Ebony Cq
Upper Missouri	Gould	4183	22.172	37.094
	Gould	4259A	21.794	35.632
	Gould	4259B	21.402	37.331
	Norman	4260A	23.053	38.611
	Norman	4260B	22.154	36.971
	Norman	4374	22.027	37.417
Yellowstone	Farwest FA	4388 A	16.683	37.194
	Farwest FA	4388 E	16.743	34.297
	Farwest FA	4388C	17.093	32.995
Headwaters	Glen	4142B	20.198	35.832
	Glen	4142A	19.696	37.842
	Glen	4142C	20.776	34.910
	Leo Lane	4304	20.173	35.610
	Leo Lane	4310	21.166	36.173
	Leo Lane	4401	22.293	35.708
	City	4343	20.610	39.860
	City	4344	19.552	36.573
	City	4345	19.452	37.236
	Clarkston	4208	20.774	32.633
	Clarkston	4294	19.130	33.447
	Clarkston	4154	20.634	37.273
	Antonioli Farm	4187A	22.615	37.773
	Antonioli Farm	4187B	20.812	36.577
	Antonioli Farm	4335	20.434	30.457

Table 18: The table shows the average actin and ebony Cq values for the technical replicates (**Table 16**) and the actin and ebony Cq values of the biological replicates (**Table 17**). The difference between the value was taken to find the Delta Cq. The RQ value was determined with the Delta Cq method. The RQ values underwent a Log transformation and the values are shown above. These are the values for the Upper Missouri drainage.

Drainage	Location	Actin Cq	Ebony Cq	ΔCq	RQ	Log (RQ)
Upper Missouri	Gould	21.40	37.33	15.93	1.60E-05	-4.80
	Gould	22.17	37.09	14.92	3.22E-05	-4.49
	Gould	21.79	35.63	13.84	6.83E-05	-4.17
	Helena Irr District	19.75	38.14	18.39	2.91E-06	-5.54
	Norman	23.05	38.61	15.56	2.07E-05	-4.68
	Norman	22.03	37.42	15.39	2.33E-05	-4.63
	Norman	22.15	36.97	14.82	3.46E-05	-4.46
	Regulating Reservoir	20.06	38.83	18.77	2.23E-06	-5.65
	West Simms	20.47	37.24	16.77	8.94E-06	-5.05
	York Island	20.09	37.70	17.61	5.00E-06	-5.30

Table 19: The table shows the average actin and ebony Cq values for the technical replicates (**Table 16**) and the actin and ebony Cq values of the biological replicates (**Table 17**). The difference between the value was taken to find the Delta Cq. The RQ value was determined with the Delta Cq method. The RQ values underwent a Log transformation and the values are shown above. These are the values for the Yellowstone drainage.

Drainage	Location	Actin Cq	Ebony Cq	ΔCq	RQ	Log (RQ)
Yellowstone	Farwest FA	16.68	37.19	20.51	6.69E-07	-6.17
	Farwest FA	16.74	34.30	17.55	5.20E-06	-5.28
	Farwest FA	17.09	32.99	15.90	1.63E-05	-4.79
	Fulop Drive	17.03	33.16	16.12	1.40E-05	-4.85
	Richey	17.19	35.20	18.01	3.78E-06	-5.42
	Spotted Eagle	17.43	33.07	15.64	1.96E-05	-4.71
	Terry Lagoons	16.83	35.92	19.09	1.80E-06	-5.75
	Weed Shed	19.56	36.02	16.47	1.10E-05	-4.96

Table 20: The table shows the average actin and ebony Cq values for the technical replicates (**Table 16**) and the actin and ebony Cq values of the biological replicates (**Table 17**). The difference between the value was taken to find the Delta Cq. The RQ value was determined with the Delta Cq method. The RQ values underwent a Log transformation and the values are shown above. These are the values for the Clark Fork drainage.

Drainage	Location	Actin Cq	Ebony Cq	ΔCq	RQ	Log (RQ)
Clark Fork	Drummond	18.82	36.18	17.36	5.95E-06	-5.23
	Gold Creek	18.80	37.11	18.30	3.09E-06	-5.51
	Hwy271m1	17.63	36.09	18.46	2.77E-06	-5.56
	Missoula	18.13	38.27	20.14	8.66E-07	-6.06

Table 21: The table shows the average actin and ebony Cq values for the technical replicates (**Table 16**) and the actin and ebony Cq values of the biological replicates (**Table 17**). The difference between the value was taken to find the Delta Cq. The RQ value was determined with the Delta Cq method. The RQ values underwent a Log transformation and the values are shown above. These are the values for the Headwaters drainage.

Drainage	Location	Actin Cq	Ebony Cq	ΔCq	RQ	Log (RQ)
Headwaters	Antonioli Farm	20.81	36.58	15.77	1.80E-05	-4.75
	Antonioli Farm	22.61	37.77	15.16	2.73E-05	-4.56
	Antonioli Farm	20.43	30.46	10.02	9.62E-04	-3.02
	Cassagrands	19.75	35.85	16.10	1.43E-05	-4.85
	City	20.61	39.86	19.25	1.60E-06	-5.79
	City	19.45	37.24	17.78	4.43E-06	-5.35
	City	19.55	36.57	17.02	7.51E-06	-5.12
	Clarkston	20.63	37.27	16.64	9.80E-06	-5.01
	Clarkston	19.13	33.45	14.32	4.90E-05	-4.31
	Clarkston	20.77	32.63	11.86	2.69E-04	-3.57
	Glen	19.70	37.84	18.15	3.45E-06	-5.46
	Glen	20.20	35.83	15.63	1.97E-05	-4.71
	Glen	20.78	34.91	14.13	5.56E-05	-4.25
	Leo Lane	20.17	35.61	15.44	2.25E-05	-4.65
	Leo Lane	21.17	36.17	15.01	3.04E-05	-4.52
	Leo Lane	22.29	35.71	13.42	9.15E-05	-4.04
	Melrose Cemetery	21.48	36.07	14.59	4.05E-05	-4.39
	Silver Star Creek	19.10	33.17	14.08	5.79E-05	-4.24

Table 22: The table shows the location, actin and ebony Cq, Delta Cq, RQ, and Log transformed RQ values for only the biological replicates. These values are also found within **Tables 18-21**.

Location	Actin Cq	Ebony Cq	ΔCq	RQ	Log (RQ)
Antonioli Farm	20.81	36.58	15.77	1.80E-05	-4.75
Antonioli Farm	22.61	37.77	15.16	2.73E-05	-4.56
Antonioli Farm	20.43	30.46	10.02	9.62E-04	-3.02
City	20.61	39.86	19.25	1.60E-06	-5.79
City	19.45	37.24	17.78	4.43E-06	-5.35
City	19.55	36.57	17.02	7.51E-06	-5.12
Clarkston	20.63	37.27	16.64	9.80E-06	-5.01
Clarkston	19.13	33.45	14.32	4.90E-05	-4.31
Clarkston	20.77	32.63	11.86	2.69E-04	-3.57
Glen	19.70	37.84	18.15	3.45E-06	-5.46
Glen	20.20	35.83	15.63	1.97E-05	-4.71
Glen	20.78	34.91	14.13	5.56E-05	-4.25
Leo Lane	20.17	35.61	15.44	2.25E-05	-4.65
Leo Lane	21.17	36.17	15.01	3.04E-05	-4.52
Leo Lane	22.29	35.71	13.42	9.15E-05	-4.04
Gould	21.40	37.33	15.93	1.60E-05	-4.80
Gould	22.17	37.09	14.92	3.22E-05	-4.49
Gould	21.79	35.63	13.84	6.83E-05	-4.17
Norman	23.05	38.61	15.56	2.07E-05	-4.68
Norman	22.03	37.42	15.39	2.33E-05	-4.63
Norman	22.15	36.97	14.82	3.46E-05	-4.46
Farwest FA	16.68	37.19	20.51	6.69E-07	-6.17
Farwest FA	16.74	34.30	17.55	5.20E-06	-5.28
Farwest FA	17.09	32.99	15.90	1.63E-05	-4.79

Results

An initial ANOVA test showed no significant differences of the RQ values based on drainage basin; however, a Shapiro-Wilk normality test and standard tests of skewness suggested that the raw RQ values were significantly skewed (**Table 25)**. The use of the non-parametric Kruskal-Wallis test suggested a significant difference between basins (Chi= 11.26, df=3, p-value=0.010). The use of a post-hoc Dunn Test suggested that the Headwaters drainage was significantly different from the Clark fork and Yellowstone drainages (**Table 23, Figure 2**).

Table 23: The pairwise non-parametric comparison between the drainage RQ values performed by the post-hoc Dunn Test is shown in the table. The Z score, unadjusted and adjusted p-values are shown for the respective comparisons. The adjusted p-values of significance are bolded and marked with an asterisk.

Drainage	Drainage Comparison	Z Score	p-Value Unadjusted	p-Value Adjusted
Clark Fork	Headwaters	-2.828	0.00468	**0.0281***
Clark Fork	Upper Missouri	-2.024	0.04295	0.0859
Headwaters	Upper Missouri	0.928	0.35352	0.4242
Clark Fork	Yellowstone	-0.890	0.37320	0.3732
Headwaters	Yellowstone	2.396	0.01657	**0.0497***
Upper Missouri	Yellowstone	1.375	0.16912	0.2537

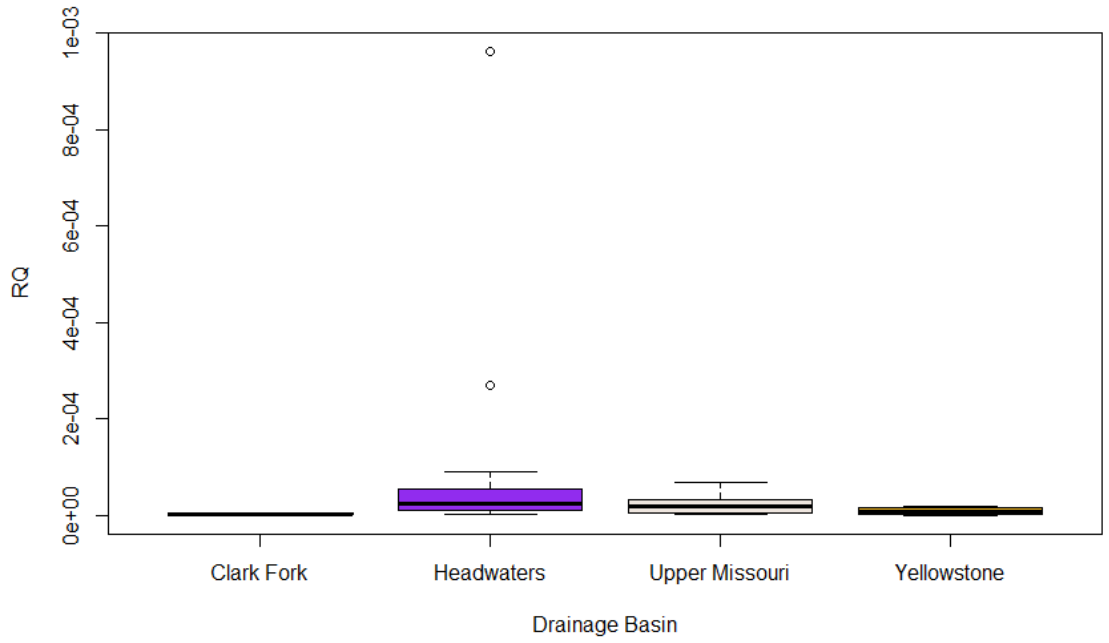

Figure 2: The boxplot above displays the raw RQ values for each drainage in comparison to one another (**Tables 18-21**).

The non-parametric results were supported by repeating the ANOVA analysis after a log transformation of the RQ values to correct for the skewed distribution. The Shapiro-Wilk normality test suggested no significant skewing of the log transformed RQ values (**Table 25**). Replication of the ANOVA on the log transformed RQ values suggested a significant difference between drainage basins (F=4.62, df=3, p-value=0.008) and a post-hoc Tukey HSD test suggested the same significant difference between the

Headwaters and Yellowstone as well as the Headwaters and the Clark Fork drainages (**Table 24, Figure 3**).

Table 24: The table displays the results of the parametric pairwise comparison of the drainages via a Tukey HSD test. The p-values accompany the drainage comparisons and the significant p-values are bolded and marked with an asterisk.

Drainage	Drainage Comparison	p-Value Adjusted
Clark Fork	Headwaters	**0.0166***
Clark Fork	Upper Missouri	0.1742
Headwaters	Upper Missouri	0.5863
Clark Fork	Yellowstone	0.7561
Headwaters	Yellowstone	**0.0527***
Upper Missouri	Yellowstone	0.5476

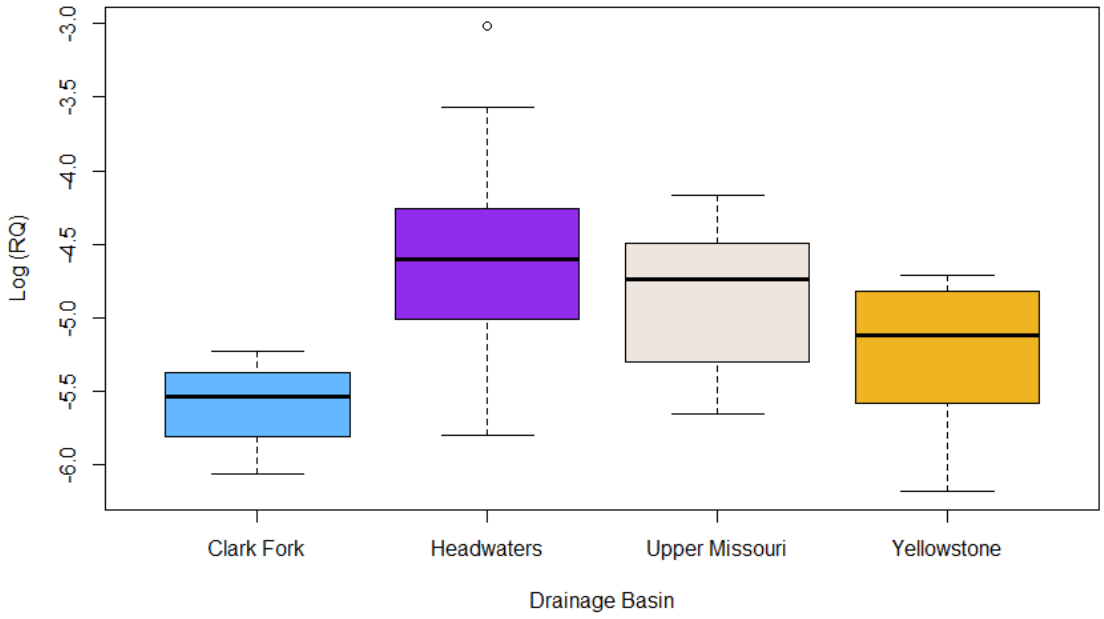

Figure 3: The boxplot above shows the Log transformed RQ values for each drainage (**Tables 18-21**).

To test for differences between sites, the ANOVA procedure was replicated and tested for differences between sites both within and across drainage basins; however, neither the non-parametric Kruskal Wallis test or the ANOVA on the log transformed data suggested significance between site specific RQ values (**Table 25, Figures 4-5**).

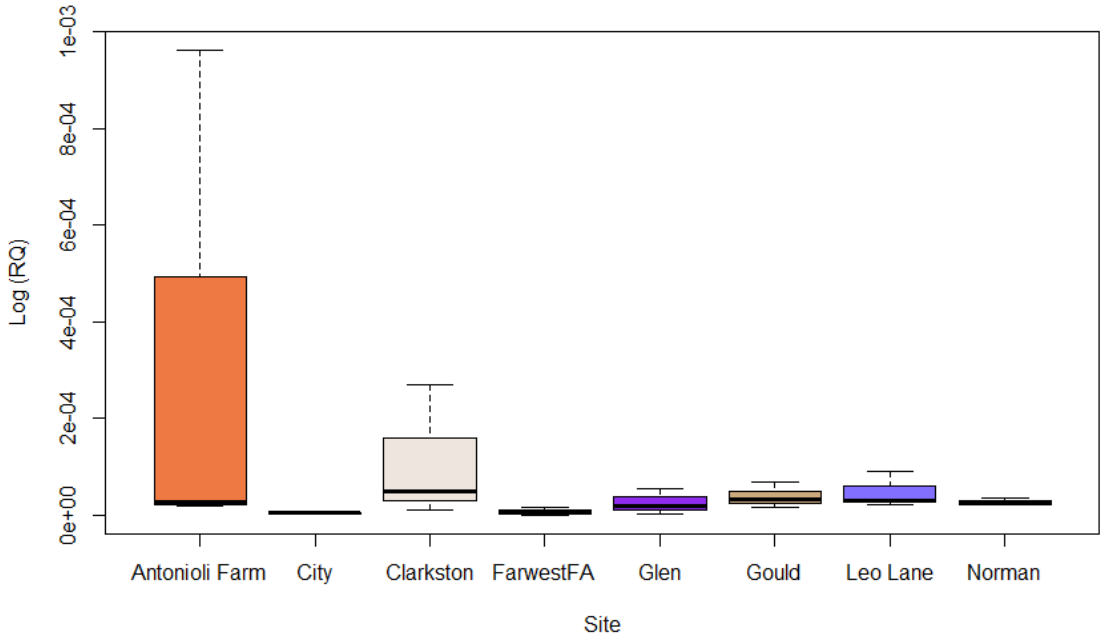

Figure 4: The boxplot shows the raw RQ values for each site with sufficient biological replicates (**Table 22**)

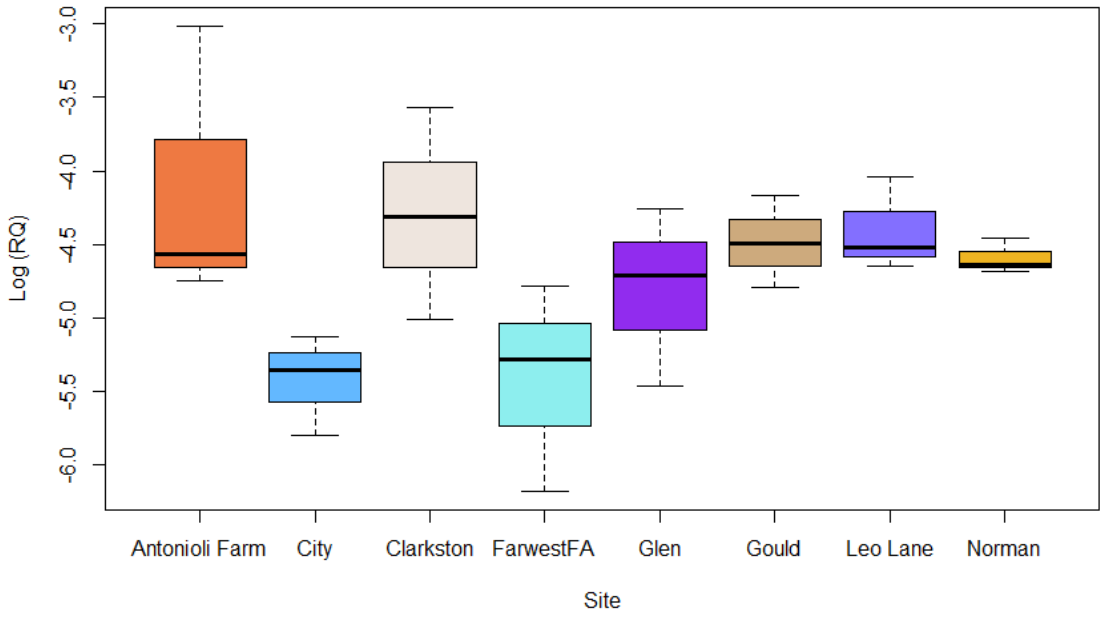

Figure 5: The boxplot shows the Log transformed RQ values for each site with sufficient biological replicates (**Table 22**)

In summary, with the use of non-parametric testing on the raw data, the log transformation of the data, and the use of parametric testing on the normalized data, it was found that there was significant Drainage Basin effect in the case of the Headwaters and the Yellowstone as well as the Headwaters and the Clark Fork.

Table 25: The table shows the summary of all normality, non- parametric, and parametric tests that were performed on the data sets to determine normality, skewness, and significance. The significant values are bolded and marked with an asterisk.

Test Performed	Data Used	p-Value
ANOVA	Raw RQ- Drainage	0.454
Shapiro-Wilk	Raw RQ- Drainage	**1.30E-11***
Kruskal-Wallis	Raw RQ- Drainage	**0.010***
Shapiro-Wilk	Log (RQ)- Drainage	0.595
ANOVA	Log (RQ)- Drainage	**0.008***
Shapiro-Wilk	Raw RQ- Site	**3.07E-06***
ANOVA	Raw RQ- Site	0.513
Kruskal-Wallis	Raw RQ- Site	0.117
Shapiro-Wilk	Log (RQ)- Site	0.778
ANOVA	Log (RQ)- Site	0.088

Discussion

The results of the data analysis offer a preliminary look into the drainage basin effect for select Motnana drainages. The original goal of this research was to obtain a comprehensive analysis of the drainage basin effect across the state of MT, but due to budget constraints and contamination error, the four drainages of the Yellowstone, Clark Fork, Upper Missouri, and the Headwaters were used in order to begin the process of obtaining a state level view of differences in ebony gene expression. In the case of the Headwaters and Yellowstone as well, as the Headwaters and Clark Fork, the headwaters had a significantly higher expression of the ebony gene when compared to these two drainages. While the Clark Fork had a p-value well below 0.05 indicating a strong certainty in this significance, the Yellowstone and Headwaters had 0.0497 for the post-hoc Dunn Test of the raw RQ values and 0.0527 for the post-hoc Tukey HSD test of the

log transformed RQ data. This suggests that the expression levels of the ebony gene are significant, but not as significant as the difference between the Clark Fork and the Headwaters. This means that, in general terms, the *Cx. tarsalis* found in the Headwaters should be significantly darker than those found in the Yellowstone and the Clark Fork as this gene is associated with dark pigmentation (Seivar, 2013). The standard deviations of the Cq values of the actin gene, with the exception of the Richey site in the Yellowstone Drainage, showed that the precision and accuracy of the techniques was high. The low standard deviation can be used to show the reliability of the data overall.

However, the pigmentation difference may have broader implications. The data also suggests that due to environmental factors currently unknown, there may be a selective pressure at work in the Headwaters drainage that is selecting for darker mosquitoes. Alternatively, in this case, there may be a lack of gene flow to this region. In the Yellowstone and Clark Fork drainages, there may be a selective pressure for lighter coloration of the mosquitoes. Future research could use these drainages to analyze the population genetics of *Cx. tarsalis* as there is a probability that these drainages, being significant from one another, lack abundant genetic flow between them. Another area of future research is the analysis of environmental factors that could be causing the selective pressure in these areas for the observed gene expression. As these examples are conjecture, another possibility for the differences found could be some passive effect in the development or possibly even the diet of the *Cx. tarsalis* in these regions. Overall, the reason for the differences described between the previously mentioned drainages is unknown, but could be the focus or basis of additional inquiry.

The site-specific analysis suggested no significant differences. This could be due to multiple factors, but the most likely is low sample size. The number of sites analyzed was hindered by the lack of biological replicates. The addition of biological replicates for the current sites as well as the addition of new sites could aid in the determination of a site-specific effect and a higher resolution of analysis.

The future direction of this analysis could follow many different directions. The main goal would to obtain a comprehensive analysis of all the drainages in the state as well as a comprehensive site analysis across the state. The site-specific analysis would allow for a higher resolution of analysis including being able to compare sites within the same drainage and sites across the state in a very high-resolution map of the ebony gene expression across Montana. In order for the analysis of the drainages to be completed, the addition of the Flathead, Lower Missouri, Marias, Milk River, and Musselshell would be needed. This would allow for the comprehensive analysis of the drainage basin effect across MT. This could also allow for further analysis of population genetics across the state. This could be done by focusing efforts on determining a protocol or method that that reflects the analysis of the ebony gene in terms of differing populations, sites, and drainages. This again could allow for a state wide comprehensive analysis of possible environmental factors that may be responsible for the selective pressure or may be passively causing the differences in pigmentation.

As mentioned prior, there were many issues with contamination of the PCR runs. This was the contamination of the actin and ebony negative controls that would show a Cq value rather than remaining without a value at the end of the PCR protocol. Due to the budget constraints, the contamination caused the inability to finish the state wide

analysis. Looking to the future, there are several ways around this issue. The first is to continue the PCR protocol once more funding is obtained. The second is to look for alternative methods to analyze the gene. One mode of analysis could be the use of SAGE (serial analysis of gene expression). In this case, the RNA samples would have to be sent to anther institute or company in order for the analysis to be done, but it would also allow for the analysis of novel mRNA sequences as well as analysis of the expression of the ebony gene.

In the case of the PCR being the route followed for the future analysis of the ebony gene, other caveats would have to be considered. The technical replication of biological replicates that are outliers would have to be completed to rule out contamination. The addition of reruns such as the Richey technical replicate due to the standard deviation being > 0.25 as well as the addition of other technical replicates to ensure precision and accuracy in subsequent trials.

In either of these cases, the comprehensive analysis of the ebony gene across the state of MT would be completed and the investigation of the drainage basin and site-specific effects would be more well understood across the state of MT.

Conclusion

With the combination of both non-parametric KW testing of the RQ values as well as the ANOVA testing of the RQ values after a log transformation, it was found that the Headwaters drainage ebony gene expression was significantly greater than both the Yellowstone and the Clark Fork drainages. The exact reason is currently unknown, but these results indicated that the Headwaters and/or the Yellowstone and the Clark Fork have selective pressure(s) at work selecting for significantly darker mosquitoes in the

Headwaters when compared to the other two drainages or significantly lighter colored mosquitoes in the case of both the Yellowstone and Clark Fork. Further inquiry is needed to better understand the environmental factors at work as well as the continuation of the current research to obtain a comprehensive analysis of all MT drainages and the addition of biological replicates to further explore and determine if there is a site-specific effect on the expression of the ebony gene and a comprehensive view of all drainage basin effects across the state of MT.

References

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, New York: Guilford Press.

CDC. Final Cumulative Maps and Data for 1999-2020. (2021, December 17). Retrieved April 29, 2022, from https://www.cdc.gov/westnile/statsmaps/cumMapsData.html#five

CDC. Other Mosquito-Borne Diseases. (2016, March 21). Retrieved April 29, 2022, from https://www.cdc.gov/niosh/topics/outdoor/mosquito-borne/other.html

Feng, X., Kambic, L., Nishimoto, J., Reed, F. A., Denton, J. A., Sutton, J. T., & Gantz, V. M. (2021). Evaluation of Gene Knockouts by CRISPR as Potential Targets for the Genetic Engineering of the Mosquito *Culex quinquefasciatus*. *The CRISPR Journal*, 4(4), 595–608. https://doi.org/10.1089/crispr.2021.0028

Fiocchi, J. (2016). An evaluation of population differentiation in *Culex tarsalis* in Montana using COI sequences. Helena, Montana]: Carroll College.

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

ITaq Universal SYBR Green One-Step Kit. (n.d.). Bio-Rad Laboratories. 1000 Alfred Nobel Drive Hercules, CA.

Johnson G, Nemeth N, Hale K, Lindsey N, Panella N, Komar N. 2010. Surveillance for West Nile Virus in American White Pelicans, Montana, USA. 2006-2007. *Emerging Infectious Diseases* 16(3):406-411.

Montana Department of Health and Human Services. West Nile Virus. Retrieved April 29, 2022, from https://dphhs.mt.gov/publichealth/cdepi/diseases/westnilevirus

Pfeiler, E., Flores-López, C., Mada-Vélez, J., Escalante-Verdugo, J., & Markow, T. (2013). Genetic Diversity and Population Genetics of Mosquitoes (Diptera: Culicidae: Culex spp.) from the Sonoran Desert of North America. *TheScientificWorld*, 2013, 724609-11.

Real-Time PCR Applications Guide. Bulletin 5279 (2006). Bio-Rad Laboratories. 1000 Alfred Nobel Drive Hercules, CA.

Sejvar, J. (2003). West nile virus: An historical overview. *The Ochsner Journal*, 5(3), 6-10.

Stromberg, K. (2019). Microsatellite analysis of *Culex tarsalis* populations in Montana. Helena, Montana]: Carroll College.

Swinker, A. M. (2022, April 12). West Nile Hirus - A Threat to Horses. Retrieved April 29, 2022, from https://extension.psu.edu/west-nile-virus-a-threat-to-horses#:~:text=Horses%20and%20humans%20are%20considered,WNV%20infection%20is%20approximately%2033%25.