

Analyzing the Biodiversity of Biological Vectors from a Forested Environment to a Coastal Environment in Suffolk County, New York William Floyd High School, Cold Spring Harbor Laboratory, & SUNY Stony Brook STEP Program N.Reis. E.Tumbrello. V. D'Ambrosia

Abstract

Despite a decrease in overall biodiversity and ecological stability worldwide since 1780, there has been an increase in biological vector diversity (Shah 2014). Biological vectors are organisms that transmit pathogens but do not directly cause diseases. Vectors are important to people and the environment because they spread diseases by transferring pathogens from one host to another. By analyzing the Cytochrome Oxidase subunit 1 gene in the genetic sequences of vectors collected, the species identity can be found. Scientists can then infer what pathogens may be present based on the vectors present. The purpose of this research was to determine if there is a difference in biodiversity of biological vectors on Long Island between forested and coastal environments. Various vectors have been collected from Manorville and Shirley to barcode their genetic sequences. The data indicates that PBQ-001 is a potential novel sequence. In Shirley, the percentages of vectors collected were: 28.6% Drosophila suzukii, 14.3% Ctenocephalides felis, 14.3% Oryzaephilus surinamensis, 14.3% Mycetophila fungorum, 14.3% Plodia interpuncte, and 14.3% Anopheles maculipennis. In Manorville, the percentages of vectors collected were: 14.3% Trichocera bimacula, 14.3% Savtshenkia sp, 14.3% Ctenocephalides felis, 14.3% Oryzaephilus surinamenis, 14.3% Musca domestica, 14.3% Ixodes laguri, and 14.3% Drosophila melongaster. This indicates that forested environments had a higher biodiversity than coastal environments, based on the organisms collected in this experiment.

Introduction

Biodiversity, or the total variety of life, impacts the stability of an ecosystem. Each organism has a niche, or role, and these organisms depend on each other in order to survive. Biodiversity has continued to decrease around the world for numerous years and many species continue to decrease in population (Shah 2014). Increasing human populations have led to an increase in various vectors' populations due to the vectors using them as hosts (Harvard, 2016). A higher quality diversity of vectors can lead to a diversity of pathogens. Pressures on biodiversity like climate change, habitat loss, and pollution can continue to increase the rate of biodiversity loss across the world (Miller & Levine 2010). In this research project, the biodiversity of biological vectors from a forested environment and a coastal environment in Suffolk County, New York were studied. It was hypothesized that there would be a difference in the biodiversity of biological vectors on Long Island between forested and coastal environments, due to the forested region having more habitat complexity than the coastal region.

The study of biological vectors helps the scientific future because the knowledge gained from analyzing their biodiversity can assist future researchers in discovering what diseases the people of Long Island may be at risk of acquiring in that specific area. Most vectors transmit pathogens through biting and are able to transmit diseases without experiencing the symptoms of the diseases (Becker et al., 2003). Some vectors carry dangerous diseases such as malaria, dengue fever, and yellow fever which are detrimental to human health. Vector-borne diseases account for more than 17% of all infectious diseases and cause over one million deaths each year (World Health Organization, 2016). By analyzing the biodiversity of these vectors, inferences could be made about which diseases the people of Long Island are at most risk from.

DNA barcoding is a method used to identify organisms based on the comparison of specific genes. This project will look specifically at the Cytochrome Oxidase subunit 1 (CO1) gene in the mitochondrial DNA (mDNA). The CO1 gene can serve as the core of a global bioidentification system for animals such as plants. (Herbet et al, 2003). When the CO1 gene is amplified, new sequences are developed to identify the organisms used in the experiment.

Methodology

Twenty vectors were collected from the south shore of Long Island, from Manorville to Shirley. Manorville is associated with pine barrens, while Shirley is coastal with mostly shores. Various samples of vectors were collected with two different types of traps. These vectors were collected by putting a Berlese funnel trap next to a pole by the wooded pine barren region in Manorville and the shores in Shirley. Additionally, vectors were collected through a water bottle trap.

The samples were frozen in order to prevent decomposition of DNA. Using a microscope camera, pictures of the vectors were documented so if the sequence is novel and the sample is entirely used than taxonomists cannot determine the correct species. Following Cold Spring Harbor Laboratory DNA extraction procedure, the vector's DNA was amplified. The CO1 gene was amplified because its variable and can be used to make a distinction between species but not too variable that it is too difficult to compare the sequences accurately. Once the gene is amplified, gel electrophoresis was determined if the gene amplified correctly. The amplified DNA was loaded into the wells of a 2% agarose gel solution and a 130 volt electrical current will be run through the gel for 30 minutes which the DNA was separated into bands based on size and charge. The darker the bands show up, the more success the experiment was. The samples were then sent to Cold Spring Harbor Laboratory's DNA Learning Center and to Genewiz, where the sequences were uploaded onto the bioinformatics tool, DNA Subway. The sequences were then trimmed based on their phred scores. Any nitrogenous bases with phred scores less than 20 on the outer portions of the sequences were taken off. Next, the samples were blasted and the possible species of the sample were determined. Then, a barcode was created comparing the different organisms collected to each other and to other organisms in the database to determine if a novel sequence had been found. Lastly, a cladogram, or phylogenetic tree, was created to compare the ancestry of the organisms collected and to similar organisms in the

database.

References

- 1. Fillinger, U., Knols, B. G. J. and Becker, N. (2003), Efficacy and efficiency of new *Bacillus thuringiensis var. israelensis* and *Bacillus sphaericus* formulations against Afrotropical anophelines in Western Kenya. Tropical Medicine & International Health, 8: 37–47. doi:10.1046/j.1365-3156.2003.00979.
- 2. Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003, January 8). Biological identifications through DNA barcodes. The Royal Society, 1, 313-320. Retrieved September 30, 2016.
- 3. Institute of Medicine (US) Committee on Lyme Disease and Other Tick-Borne Diseases:
- The State of the Science. Critical Needs and Gaps in Understanding Prevention, Amelioration, and Resolution of Lyme and Other Tick-Borne Diseases 4. (US); 2011. 4
- 5. Emerging Infections, Tick Biology, and Host–Vector Interactions.
- Miller, K. R., & Levine, J. S. (2012). *Biology*. Upper Saddle River, New Jersey: Pearson. 6.
- 7. Statistics & Maps. Retrieved March 09, 2017, from https://www.cdc.gov/

Results:

	С	1	2	3	4	5	6	7	8	9	10	11	12	13	14
С		80.35	79.80	82.37	84.67	80.93	83.11	84.94	86.96	89.37	87.55	88.95	89.95	92.29	91.79
1	80.35	•	100.00	74.20	76.09	75.55	76.71	78.69	76.02	77.06	75.25	78.04	75.95	75.69	74.62
2	79.80	100.00		74.20	76.09	75.55	76.71	78.69	76.02	77.00	75.25	77.98	75.95	75.69	74.62
3	82.37	74.20	74.20		76.12	75.16	75.80	75.16	78.21	77.40	76.92	78.04	78.85	78.85	80.93
4	84.67	76.09	76.09	76.12	•	78.57	78.57	81.11	81.86	80.45	79.18	81.86	79.60	81.34	80.24
5	80.93	75.55	75.55	75.16	78.57		89.37	78.53	80.35	78.75	75.55	79.91	77.78	81.08	79.18
6	83.11	76.71	76.71	75.80	78.57	89.37	•	78.68	80.79	78.31	76.71	80.49	79.91	82.10	81.31
7	84.94	78.69	78.69	75.16	8 <mark>1.1</mark> 1	78.53	78.68	•	82.24	83.38	80.29	82.10	80.52	82.21	81.61
8	86.96	76.02	76.02	78.21	81.86	80.35	80.79	82.24		82.61	80.48	82.47	81.28	83.70	83.74
9	89.37	77.06	77.00	77.40	80.45	78.75	78.31	83.38	82.61	•	88.12	85.45	84.47	85.74	85.71
10	87.55	75.25	75.25	76.92	79.18	75.55	76.71	80.29	80.48	88.12		83.88	85.24	85.59	85.87
11	88.95	78.04	77.98	78.04	81.86	79.91	80.49	82.10	82.47	85.45	83.88	•	89.35	87.19	86.17
12	89.95	75.95	75.95	78.85	79.60	77.78	79.91	80.52	81.28	84.47	85.24	89.35		86.61	86.91







Coastal Region



Figure 2a: This pie graph demonstrates the percentages of biological vectors present in the forested region based on barcode identification. This data indicates that there are more organisms identified in the forested environment than the coastal region,, therefore showing that the forested environment was more biodiverse than the coastal environment

Figure 2b: The pie graph demonstrates the percentage of biological vectors present in the coastal region based on barcode identification. This indicates that there are less organisms found on coastal regions than forested regions. This contributes to that there's not as many areas in coastal regions for biological vectors to live and produce compared to forested regions.

Discussion & Conclusion

The following research questions were posed: Is there a difference in the biodiversity of biological vectors in forested and coastal regions on Long Island? and Are there any novel sequences of biological vectors present in forested and coastal regions on Long Island? The pie graphs shown above shows that there is a larger biodiversity of biological vectors within the forested received region than the coastal region. Although the shoreline region has a larger percentage, the forested region has many more vectors, therefore, Manorville has more biodiversity and variation of biological vectors than Shirley. The data supports the hypothesis presented. This could be due to the environment the forested region has. Unlike Shirley, Manorville has a three dimensional canopy which allows for a higher probability that vectors will find a suitable habitat to live and reproduce in. Shirley is a coastal region which means there was no third dimensional canopy which may have led to less biodiversity in that region.

The data shown indicates that there might be novel sequences present. It was hypothesized that there is novel sequences found due to rapid environmental changes, which could lead to adaptation. After the results were finished, it is shown that there is one novel sequence (PBQ-001). When barcoding the organisms, it was shown in the data table that the sequence was not presented in bold, which was labeled unsequenced.

Acknowledgements

We would like to acknowledge Dr. Pepenella, Dr. Nash, and Dr. Marco-Fernandez from Cold Spring Harbor Laboratory DNA Learning Center. This research would not have been feasible without their help. We would also like to acknowledge Stony Brook's Science and Technology Entry Program, the Department of Environmental Conservation, Brookhaven National Laboratory, William Floyd High School's Research Program, and the Pine Barrens Commission or their support. Finally, we have to thank our research teacher Ms. D'Ambrosia. Without her help, none of this project would have existed.



Figure 1: This figure demonstrates the barcodes of the samples analyzed and other reference species. The colored bars indicate differences in nucleotides and the grey sections indicate similarities in nucleotides. The percentages indicate similarities between samples and reference insects with percentages greater than 95% indicating that the two organisms are the same species.

Percentages Of Biological Vectors Present In



Figure 3: This figure demonstrates the evolutionary relationships between the tested samples and reference organisms in the form of a cladogram. The data indicates that the novel sequence, PBQ-001 has a common ancestor which is the American Cockroach. The bootstrap scores indicates how many times out of 100 the organisms were paired together to have a common ancestor.