8-1-2022

IMMUNE VARIATION THROUGH MHC GENE ANALYSIS OF THE MOURNING GECKO (Lepidodactylus lugubris)

Elizabeth Rose Weeresinghe

IMMUNE VARIATION THROUGH MHC GENE ANALYSIS OF THE MOURNING GECKO (*Lepidodactylus lugubris*)

by

ELIZABETH ROSE WEERESINGHE

Presented to the Faculty of the Honors College of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of Biology

HONORS BACHELOR OF SCIENCE IN BIOLOGY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2022

ACKNOWLEDGMENTS

I would like to acknowledge and thank Dr. Matthew K. Fujita and Mr. Joshua Rivera for their assistance with this project. Without them, this project would not have been possible, and I would not have learned as much as I have about MHC genes.

I was fortunate to take Genetics with Dr. Fujita during a quarantining semester, and really enjoyed his teaching style along with the material I learned in that class. When I viewed his seminar presentation, I knew I wanted to do my capstone under his mentorship. Seeing where we have come from where we started in the Fall of 2022 makes me smile immensely. I am so happy that I obtained the opportunity to do research under one of the coolest professors ever.

I met Mr. Rivera during a Microsoft Teams meeting with Dr. Fujita early this semester when we went over next steps regarding this project. Even though we were both a little lost and I was freaking out about how Dr. Fujita had not obtained the lizards needed for our first project proposal, his presence put me at ease. We later had an encounter in Fujita's lab when going over the data obtained through the BLAST test. Mr. Rivera was more than just helpful that day. He gave me so many points that allowed me to complete this project with no issue. I will forever be thankful to him for helping me complete this project.

April 22, 2022

ABSTRACT

IMMUNE VARIATION THROUGH MHC GENE

ANALYSIS OF THE MOURNING GECKO

(*Lepidodactylus lugubris*)

Elizabeth Rose Weeresinghe, B.S. Biology

The University of Texas at Arlington, 2022

Faculty Mentor: Matthew K. Fujita

A Major histocompatibility complex (MHC) is a cell-surface molecule encoded by a large gene family in all vertebrate DNA (deoxyribonucleic acid). The MHC determines susceptibility to autoimmune diseases. These genes are highly polymorphic, meaning that in a non-endogamic population, each organism has a unique set of MHC genes and molecules. Evolution of the MHC through polymorphism ensures a population will not succumb to a new or mutated pathogen, because some individuals will develop an adequate immune response to defeat the pathogen. Because of this, finding MHC genes in transcriptomes is important for finding genetic variability in the immune systems of a species. Looking at the transcriptomes of the mourning gecko (*Lepidodactylus lugubris*) will allow observation and analysis of the genetic variation in this species. Even though this particular gecko is asexual and reproduces through parthenogenesis, environmental

factors may play a part in immune variation resulting in inheritance of different MHC molecules.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

All animals are inhabited by a variety of microorganisms such as bacteria, archaea, fungi, and viruses. These microorganisms influence numerous aspects of host physiology, including digestion and nutrient absorption, organ growth and development, behavior, and the development and function of the immune system. Development and maintenance of these communities and their composition are influenced by several factors: primarily diet, environment, and host genetics.

Mourning geckos (*Lepidodactylus lugubris*) provide a unique opportunity to investigate the relationship between host genetic diversity and physiology due to its presence as a parthenogenetic species. When making comparisons within and between populations, the parthenogens will serve as perfect opportunity to observe genetic variability due to evolution through time as they would commonly exhibit little-to-no genetic variation in comparison to an androgenetic species. Because host physiology is closely connected with the immune system, I am specifically interested in genes that are essential for immune functions. We expect populations with low diversity in immunity genes to have different microbiome compositions than populations with high genetic diversity.

CHAPTER 2

LITERATURE REVIEW

2.1 Genetic Variability

Genetic variation is the presence of differences in the sequence of genes when comparing individual organisms in a species. These differences enable natural selection and immune enhancement. Genes are units of hereditary information, which carry instructions for building proteins. The genes that are encoded within these proteins are what enable our cells to function. Most organisms that reproduce sexually have two copies of each gene provided by each parent cell, because each one donates a single copy of its genes to its offspring. Additionally, genes can exist in slightly different forms, called alleles, which further adds to genetic variation. The combination of alleles of a gene that an individual receives from both parents determines what biologists call the genotype for a particular trait. The genotype that an individual possesses for a trait, in turn, determines the phenotype, which is the observable characteristics.

Genetic variation within a species can result from certain different situations. Mutations, the changes in the sequences of genes in DNA, are one source of genetic variation. Another source is gene flow, or the movement of genes between different groups of organisms. Finally, genetic variation can be a result of sexual reproduction, which leads to the creation of new combinations of genes.

Mutations are changes in the information contained in genetic material. In most instances, this means a change in the sequence of DNA, the hereditary material of life. An

organism's DNA affects its physiology, so a change in an organism's DNA can cause changes in all aspects of its life. (Understanding Evolution, 2022).

Genetic variation in a group of organisms enables some organisms to survive better than others in the environment in which they live. Organisms of even a small population can differ strikingly in terms of how well suited they are for life in their particular habitat. The process by which organisms who are better suited for their environment have a better rate of survival, is the process of natural selection, and it is the main force that drives evolution (Collins, F. S., n.d. & National Geographic Society, 2019).

Specifically, immune variation is the topic of inspection in this project. Finding genes that display immune variation in the mourning gecko will further expand on the effects of environmental stimulus to a parthenogenetic species.

2.2 Transcriptomes

A genome is made up of DNA, a long, winding molecule that contains the instructions needed to build and maintain cells. These instructions are spelled out in the form of base pairs of four different nucleobases: adenine (A), guanine (G), thymine (T), and cytosine (C). These chemicals are organized into genes. For the instructions to be carried out, DNA must be read and then transcribed into RNA (ribonucleic acid). The difference in nucleobase between DNA and RNA is the switch of thymine to uracil (U). These gene readouts are called transcripts, and a transcriptome is a collection of all the gene readouts present in a cell.

There are various kinds of RNA, but the major type, called messenger RNA (mRNA), plays a vital role in making proteins. In this process, mRNA is transcribed from genes, and then the transcripts are delivered to ribosomes, the molecular machines located

in the cell's cytoplasm. The ribosomes translate the sequence of chemical letters in the mRNA and assemble the amino acid building blocks into proteins.

An RNA sequence mirrors the sequence of the DNA from which it was transcribed. By analyzing the entire collection of RNA sequences in a cell (the transcriptome) researchers can determine when and where each gene is turned on or off in the cells and tissues of an organism. Depending on the technique used, it is often possible to count the number of transcripts to determine the amount of gene expression in a certain cell or tissue type.

In many organisms, nearly every cell contains the same genes, but different cells show different patterns of gene expression. These differences are responsible for the many different properties and behaviors of various cells and tissues, both in health and disease.

By collecting and comparing transcriptomes of different types of cells, researchers can gain a deeper understanding of what constitutes a specific cell type, how that type of cell normally functions, and how changes in the normal level of gene activity may reflect or contribute to disease. In addition, transcriptomes may enable researchers to generate a comprehensive, genome-wide picture of what genes are active in which cells (Srivastava, A. et.al., 2019 & NHGRI. 2019).

Analyzing the DNA sample collected from the mourning gecko will allow the display of its transcriptomes to illustrate which genes may be present that are responsible for immune variability in this species. By creating and searching through a transcriptome database, I will obtain a list of all the tissues in which a gene is expressed, providing clues about its possible function in immune variation.

2.3 MHC Genes

A Major histocompatibility complex (MHC) is a cell-surface molecule encoded by a large gene family in all vertebrate DNA. They contain a set of closely linked polymorphic genes that code for cell surface proteins essential for the adaptive immune system called MHC molecules. MHC molecules bind an antigen derived from self-proteins, or from pathogens, and bringing the antigen presentation to the cell surface for recognition by the appropriate T-cells and display a molecular fraction (epitope) and mediate interactions between leukocytes and/or body cells. The MHC determines donor compatibility for organ transplant, as well as one's susceptibility to autoimmune diseases.

There are three subgroups in MHC gene family: MHC class I, class II, and class III. Among all the genes present in MHC, there are two types of genes coding for the proteins MHC class I and class II molecules that are directly involved in the antigen presentation. MHC class I molecules are expressed in all nucleated cells and also in platelets (no RBCs). It presents epitopes to killer T cells, also called cytotoxic T lymphocytes (CTLs). MHC class II can be expressed by all cell types, but normally occurs only on professional antigen-presenting cells (APCs): macrophages, B cells, and dendritic cells (DCs). Class III molecules have physiologic roles unlike classes I and II but are encoded between them in the short arm of human chromosome 6.

The MHC genes are highly polymorphic, meaning there are many different alleles in the different individuals inside a population. The polymorphism is so high that in a mixed population (non-endogamic) there are not two individuals with exactly the same set of MHC genes and molecules, except for identical twins. The polymorphic regions in each allele are located in the region for peptide contact, which will be displayed to the

lymphocyte. Because of this, the contact region for each allele of MHC molecule is highly variable. The polymorphic residues of the MHC will create specific clefts where only certain types of residues of the peptide can enter. This imposes a very specific link between the MHC molecule and the peptide. Each MHC variant will be able to bind specifically only to those peptides that are able to properly enter in the cleft of the MHC molecule, which is variable for each allele. In this way, the MHC molecules have a broad specificity, because they can bind many, but not all, types of possible peptides.

MHC is the chaperone for intracellular peptides that are complexed with MHCs and presented to T cell receptors (TCRs) as potential foreign antigens. MHC interacts with TCR and its co-receptors to optimize binding conditions for the TCR-antigen interaction, in terms of antigen binding affinity and specificity, and signal transduction effectiveness. MHC is also the tissue-antigen that allows the immune system, more specifically T cells, to bind to, recognize, and tolerate itself through autorecognition.

Evolution of the MHC through polymorphism ensures that a population will not succumb to a new pathogen or a mutated one, because at least some individuals will be able to develop an adequate immune response to win over the pathogen. The variations in the MHC molecules are the result of the inheritance of different MHC molecules, and they are not induced by recombination, as it is the case for the antigen receptors (Weeresinghe, E. R., & Fujita, M. K., 2022).

These are the genes that will be searched and identified from the transcriptome reading derived from the DNA sample of the mourning gecko. If MHC genes are identified, analyzed, and compared to by those of other reptile species, the resulting will be evidence of the potential presence for immune variation in this species of parthenogenetic gecko.

2.4 The Mourning Gecko

The mourning gecko (*Lepidodactylus lugubris*) is quite prevalent in coastal areas surrounding the Pacific and Indian oceans and have also been introduced to areas in South America and Hawaii. From the Maldives to the Galapagos Islands, mourning geckos are widespread. The average mourning gecko size is usually between three and four inches when fully grown, and upon hatching, juveniles are a mere 1.5 inches long. The average mourning gecko can survive for an approximate ten-year lifespan.

Mourning geckos like to stay out of sight in their natural environments. They are cryptically colored to blend into their surroundings. Most of the body is dark to light tan. Dark brown or black markings adorn its back. This species is also able to modify their color. By firing up or down, meaning that they can change the color and patterns of their skin to lighter or darker shades, they allow themselves to adjust to their surroundings. As a result, the same gecko may appear lighter or darker at different times during the day.

For the most part, mourning geckos are peaceful creatures. They are friendly creatures who do best in large social groups. These geckos are also cathemeral, meaning they only come out of hiding sporadically. They have irregular patterns of behavior throughout the day, but most tend to be active at night.

This species of gecko is parthenogenetic, which means that the females reproduce asexually without males. Because reproduction only requires egg-laying, a vast majority of mourning geckos are females. Males do exist, but they're rare and are usually infertile (Briggs H., 2021 & N. E. H., n.d.).

Since this species is parthenogenetic, immune variability through genetic lineage is a rare occurrence. For this reason, these geckos are perfect subjects for testing variability through other means such as environmental factors and natural selection.

CHAPTER 3

METHODOLOGY

3.1 Specimens

This study uses a specimen of *L. lugubris*. These animals were collected using proper collecting permits and under UTA IACUC protocol #A16.010. Dorsal skin samples from the gecko was collected and stored in RNALater (Qiagen) and the specimens have been deposited in the Amphibian and Reptile Diversity Research Collection at the University of Texas at Arlington.

3.2 Transcriptome Sequencing and Assembly

Total RNA was extracted from the skin sample using the Promega Total SV RNA Extraction kit (Promega). The quality of the RNA was assessed using a bioanalyzer, which compares the relative peaks of two target rRNA molecules to determine whether the sample is adequate for library preparation. Because this value, called the RNA Integrity Number (RIN), was > 7 for the sample, it was approved for sequencing at the North Texas Genome Center (NTGC). The NTGC prepared the library and subsequently sequenced the samples on a NovaSeq6000.

The data was processed to remove low quality reads using Trimmomatic v.32 and the following parameters to trim and remove failed reads: a 4-base sliding window trimming nucleotides with a Q score <5 and discarding reads <25bp long (Bolger et al., 2014; MacManes, 2014). To ensure quality trimming and filtering, the processed reads were run through FASTQC v0.10.1 (Babraham Bioinformatics) to evaluate read quality,

length, and the number of reads retained. De novo assemblies were carried out using Trinity short-read assembler v2.2.1 for each sample (Grabherr et al., 2011). TransDecoder v3.0.1 was used to identify candidate coding genes from our assembled transcript sets with the longest open reading frames (ORFs) (Haas and Papanicolaou, 2015). To maximize the number of ORFs captured and to ensure we did not lose any potential coding genes, TransDecoder's optional homology search was run against the PFAM database (Finn et al., 2016).

3.3 Data Analysis

Using the program BLAST+ (2.10.1), Dr Fujita, Mr. Rivera, and I created a transcriptome database for L. lugubris (Camacho et al., 2009). After creating the database for the sample, the data derived was compared to that of data already collected on different reptile species where MHC gene sequences have already been identified and characterized in tblastx searches (e-value cutoff of 1e-30). The other reptiles used in the comparison were Amblyrhynchus, Ambystoma, Sphenodon, and Anolis. The program tblastx translates the database and query nucleotide sequences into their six different reading frames (three forward and three backward) and then conducts six different searches against the database based on amino acid sequences. Based on this initial blast search 21 genes were identified from the *L. lugubris* transcriptomes that matched genes found in the Amblyrhynchus. This initial collection of genes was then used as a comparison for an additional tblastx search, which was performed using a more stringent e-value cutoff (1e-50). This search identified 114 matched genes.

Using the program MACSE (2.03), Dr. Fujita aligned the genes between the *L. lugubris* specimens and those from the other squamates (Ranwez et al., 2011; Ranwez et

al., 2018). MACSE aligns the nucleotide sequences by comparing their respective amino acid translations. By importing the resulting alignments into the program Geneious Prime 2020.1.2 (Biomatters), a comparison of the aligned genes using phylogenetic trees and distances were made. Genes that were too divergent or too small to be full genes were removed from the alignments. Final phylogenetic trees were made with the *L. lugubris* genes that are homologous to the other squamate MHC genes.

CHAPTER 4

DISCUSSION

Judging by the number of matches displayed in the more stringent comparison, the mourning gecko has a sizeable number of MHC genes present in its DNA. Genes present in the comparison that have been previously classified in the lizard database as MHC genes are as follows: 769, 17616, 1015, 16859, & 16858. Future research can be conducted to identify additional genes which may display MHC potential. Other future studies could be conducted to demonstrate variation of immunity through generational lineage by waiting for a large enough population of *L. lugubris* to form and collecting samples from each present generation.

CHAPTER 5

CONCLUSION

The number of MHC genes that were identified from the database comparisons signify that the mourning gecko is indeed limited regarding immune variation. Since this species is asexual and reproduces through parthenogenesis, genetic variation present in this population is rare. In retrospect, MHC genes present that have the function of immune variation are also uncommon. These findings support our hypothesized statement that the immune variation would not be vast in this species.

APPENDIX

DATA OUTPUT TABLE

Mourning Gecko MHC Output: 1e-50

AF209117.1.Ambystoma	Gene.16859::spnov37848::g.16859::m.16859	45.312	64	35	0	130
AF209117.1.Ambystoma	Gene.16859::spnov37848::g.16859::m.16859	43.396	53	30	0	466
AF209117.1.Ambystoma	Gene.16859::spnov37848::g.16859::m.16859	50	24	12	0	400
AF209117.1.Ambystoma	Gene.16859::spnov37848::g.16859::m.16859	60	15	6	0	79
AF209117.1.Ambystoma	Gene.16858::spnov37848::g.16858::m.16858	45.312	64	35	0	130
AF209117.1.Ambystoma	Gene.16858::spnov37848::g.16858::m.16858	43.396	53	30	0	466
AF209117.1.Ambystoma	Gene.16858::spnov37848::g.16858::m.16858	50	24	12	0	400
AF209117.1.Ambystoma	Gene.16858::spnov37848::g.16858::m.16858	60	15	6	0	79
AF209117.1.Ambystoma	Gene.769::spnov1749::g.769::m.769	43.077	65	37	0	130
AF209117.1.Ambystoma	Gene.769::spnov1749::g.769::m.769	41.509	53	31	0	466
AF209117.1.Ambystoma	Gene.769::spnov1749::g.769::m.769	50	24	12	0	400
AF209117.1.Ambystoma	Gene.769::spnov1749::g.769::m.769	68.75	16	5	0	79
AF209117.1.Ambystoma	Gene.17616::spnov39675::g.17616::m.17616	42.188	64	37	0	130
AF209117.1.Ambystoma	Gene.17616::spnov39675::g.17616::m.17616	43.396	53	30	0	466
AF209117.1.Ambystoma	Gene.17616::spnov39675::g.17616::m.17616	50	24	12	0	400
AF209117.1.Ambystoma	Gene.17616::spnov39675::g.17616::m.17616	60	15	6	0	79
AF209117.1.Ambystoma	Gene.17615::spnov39675::g.17615::m.17615	42.188	64	37	0	130
AF209117.1.Ambystoma	Gene.17615::spnov39675::g.17615::m.17615	43.396	53	30	0	466
AF209117.1.Ambystoma	Gene.17615::spnov39675::g.17615::m.17615	50	24	12	0	400
AF209117.1.Ambystoma	Gene.17615::spnov39675::g.17615::m.17615	60	15	6	0	79
DQ124233.1.Sphenodon	Gene.16859::spnov37848::g.16859::m.16859	48.438	128	66	0	274
DQ124233.1.Sphenodon	Gene.16859::spnov37848::g.16859::m.16859	46.154	39	21	0	160

DQ124233.1.Sphenodon	Gene.16859::spnov37848::g.16859::m.16859	57.895	19	8	0	97
DQ124233.1.Sphenodon	Gene.16858::spnov37848::g.16858::m.16858	48.438	128	66	0	274
DQ124233.1.Sphenodon	Gene.16858::spnov37848::g.16858::m.16858	46.154	39	21	0	160
DQ124233.1.Sphenodon	Gene.16858::spnov37848::g.16858::m.16858	57.895	19	8	0	97
DQ124233.1.Sphenodon	Gene.769::spnov1749::g.769::m.769	45.736	129	70	0	283
DQ124233.1.Sphenodon	Gene.769::spnov1749::g.769::m.769	43.59	39	22	0	160
DQ124233.1.Sphenodon	Gene.769::spnov1749::g.769::m.769	68.421	19	6	0	97
DQ124233.1.Sphenodon	Gene.17616::spnov39675::g.17616::m.17616	44.715	123	68	0	289
DQ124233.1.Sphenodon	Gene.17616::spnov39675::g.17616::m.17616	51.515	33	16	0	160
DQ124233.1.Sphenodon	Gene.17616::spnov39675::g.17616::m.17616	63.158	19	7	0	97
DQ124233.1.Sphenodon	Gene.17615::spnov39675::g.17615::m.17615	44.715	123	68	0	289
DQ124233.1.Sphenodon	Gene.17615::spnov39675::g.17615::m.17615	51.515	33	16	0	160
DQ124233.1.Sphenodon	Gene.17615::spnov39675::g.17615::m.17615	63.158	19	7	0	97
DQ124233.1.Sphenodon	Gene.22802::spnov51258::g.22802::m.22802	49	100	51	0	358
DQ124233.1.Sphenodon	Gene.22802::spnov51258::g.22802::m.22802	45	40	22	0	160
DQ124233.1.Sphenodon	Gene.22802::spnov51258::g.22802::m.22802	57.143	21	9	0	91
DQ124233.1.Sphenodon	Gene.22801::spnov51258::g.22801::m.22801	49	100	51	0	358
DQ124233.1.Sphenodon	Gene.22801::spnov51258::g.22801::m.22801	45	40	22	0	160
DQ124233.1.Sphenodon	Gene.22801::spnov51258::g.22801::m.22801	57.143	21	9	0	91
DQ124233.1.Sphenodon	Gene.36147::spnov80677::g.36147::m.36147	46.094	128	69	0	283
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	56.923	65	28	0	139
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	56.667	60	26	0	481
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	37.209	43	27	0	358

FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	54.286	35	16	0	642
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	56.25	16	7	0	58
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	57.407	54	23	0	671
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	60	40	16	0	332
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	43.75	48	27	0	483
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	73.684	19	5	0	195
FJ623746.1.Amblyrhynchus	Gene.769::spnov1749::g.769::m.769	57.576	33	14	0	751
FJ623746.1.Amblyrhynchus	Gene.17616::spnov39675::g.17616::m.17616	63.077	65	24	0	139
FJ623746.1.Amblyrhynchus	Gene.17616::spnov39675::g.17616::m.17616	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.17616::spnov39675::g.17616::m.17616	45.652	46	25	0	349
FJ623746.1.Amblyrhynchus	Gene.17616::spnov39675::g.17616::m.17616	56.25	16	7	0	58
FJ623746.1.Amblyrhynchus	Gene.17615::spnov39675::g.17615::m.17615	63.077	65	24	0	139
FJ623746.1.Amblyrhynchus	Gene.17615::spnov39675::g.17615::m.17615	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.17615::spnov39675::g.17615::m.17615	45.652	46	25	0	349
FJ623746.1.Amblyrhynchus	Gene.17615::spnov39675::g.17615::m.17615	56.25	16	7	0	58
FJ623746.1.Amblyrhynchus	Gene.16859::spnov37848::g.16859::m.16859	60	65	26	0	139
FJ623746.1.Amblyrhynchus	Gene.16859::spnov37848::g.16859::m.16859	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.16859::spnov37848::g.16859::m.16859	45.652	46	25	0	349
FJ623746.1.Amblyrhynchus	Gene.16859::spnov37848::g.16859::m.16859	50	16	8	0	58
FJ623746.1.Amblyrhynchus	Gene.16858::spnov37848::g.16858::m.16858	60	65	26	0	139
FJ623746.1.Amblyrhynchus	Gene.16858::spnov37848::g.16858::m.16858	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.16858::spnov37848::g.16858::m.16858	45.652	46	25	0	349
FJ623746.1.Amblyrhynchus	Gene.16858::spnov37848::g.16858::m.16858	50	16	8	0	58

FJ623746.1.Amblyrhynchus	Gene.22802::spnov51258::g.22802::m.22802	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.22802::spnov51258::g.22802::m.22802	44.681	47	26	0	346
FJ623746.1.Amblyrhynchus	Gene.22802::spnov51258::g.22802::m.22802	38.889	54	33	0	139
FJ623746.1.Amblyrhynchus	Gene.22802::spnov51258::g.22802::m.22802	50	16	8	0	58
FJ623746.1.Amblyrhynchus	Gene.22802::spnov51258::g.22802::m.22802	52.632	19	9	0	297
FJ623746.1.Amblyrhynchus	Gene.22801::spnov51258::g.22801::m.22801	57.407	54	23	0	481
FJ623746.1.Amblyrhynchus	Gene.22801::spnov51258::g.22801::m.22801	44.681	47	26	0	346
FJ623746.1.Amblyrhynchus	Gene.22801::spnov51258::g.22801::m.22801	38.889	54	33	0	139
FJ623746.1.Amblyrhynchus	Gene.22801::spnov51258::g.22801::m.22801	50	16	8	0	58
FJ623746.1.Amblyrhynchus	Gene.22801::spnov51258::g.22801::m.22801	52.632	19	9	0	297
FJ623746.1.Amblyrhynchus	Gene.36147::spnov80677::g.36147::m.36147	60.417	48	19	0	508
FJ623746.1.Amblyrhynchus	Gene.36147::spnov80677::g.36147::m.36147	39.13	46	28	0	349
FJ623746.1.Amblyrhynchus	Gene.36147::spnov80677::g.36147::m.36147	44.444	27	15	0	253
FJ623746.1.Amblyrhynchus	Gene.36147::spnov80677::g.36147::m.36147	40	35	21	0	642
XM_008121064.1.Anolis	Gene.769::spnov1749::g.769::m.769	50.877	57	28	0	499
XM_008121064.1.Anolis	Gene.769::spnov1749::g.769::m.769	53.448	58	27	0	151
XM_008121064.1.Anolis	Gene.769::spnov1749::g.769::m.769	41.379	29	17	0	418
XM_008121064.1.Anolis	Gene.769::spnov1749::g.769::m.769	57.143	21	9	0	82
XM_008121064.1.Anolis	Gene.769::spnov1749::g.769::m.769	64.286	28	10	0	677
XM_008121064.1.Anolis	Gene.22802::spnov51258::g.22802::m.22802	44.118	68	38	0	151
XM_008121064.1.Anolis	Gene.22802::spnov51258::g.22802::m.22802	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.22802::spnov51258::g.22802::m.22802	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.22802::spnov51258::g.22802::m.22802	57.143	21	9	0	82

XM_008121064.1.Anolis	Gene.22801::spnov51258::g.22801::m.22801	44.118	68	38	0	151
XM_008121064.1.Anolis	Gene.22801::spnov51258::g.22801::m.22801	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.22801::spnov51258::g.22801::m.22801	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.22801::spnov51258::g.22801::m.22801	57.143	21	9	0	82
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	41.379	58	34	0	151
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	41.379	58	34	0	151
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	61.905	21	8	0	82
XM_008121064.1.Anolis	Gene.16859::spnov37848::g.16859::m.16859	66.667	9	3	0	328
XM_008121064.1.Anolis	Gene.16858::spnov37848::g.16858::m.16858	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.16858::spnov37848::g.16858::m.16858	41.379	58	34	0	151
XM_008121064.1.Anolis	Gene.16858::spnov37848::g.16858::m.16858	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.16858::spnov37848::g.16858::m.16858	61.905	21	8	0	82
XM_008121064.1.Anolis	Gene.16858::spnov37848::g.16858::m.16858	66.667	9	3	0	328
XM_008121064.1.Anolis	Gene.17616::spnov39675::g.17616::m.17616	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.17616::spnov39675::g.17616::m.17616	43.103	58	33	0	151
XM_008121064.1.Anolis	Gene.17616::spnov39675::g.17616::m.17616	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.17616::spnov39675::g.17616::m.17616	52.381	21	10	0	82
XM_008121064.1.Anolis	Gene.17615::spnov39675::g.17615::m.17615	49.057	53	27	0	499
XM_008121064.1.Anolis	Gene.17615::spnov39675::g.17615::m.17615	43.103	58	33	0	151
XM_008121064.1.Anolis	Gene.17615::spnov39675::g.17615::m.17615	51.724	29	14	0	418
XM_008121064.1.Anolis	Gene.17615::spnov39675::g.17615::m.17615	52.381	21	10	0	82

AF209117.1.Ambystoma	Gene.16859::spnov37848::g.16859::m.16859	45.312	64	35	0	130

The data displayed in this table is the result of the additional tblastx search which was performed using a more stringent e-value cutoff of 1e-50.

REFERENCES

Briggs, H. (2021, August 17). Mourning Gecko 101. Reptile Direct. https://www.reptiledirect.com/mourning-gecko/

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and T. L. Madden. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Collins, F. S. (n.d.). Genetic Variation. National Human Genome Research Institute. https://www.genome.gov/genetics-glossary/Genetic-Variation

Currie, K. L., & Fujita, M. K. (n.d.). Identifying MHC class II genes in whiptail lizards (Aspidoscelis) file:///C:/Users/12146/AppData/Local/Temp/Temp3_Lepidodactylus_MHC.zip/JMIH 2019_FinalPoster.pdf

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and A. Bateman. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research 44:D279–D285. https://doi.org/10.1093/nar/gkv1344

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29:644–652. https://doi.org/10.1038/nbt.1883

Janeway C. A. J., Travers, P., & Walport, M. (2001). The major histocompatibility complex and its functions. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/books/NBK27156/Llc

N. E. H. (n.d.). NEHERP - Mourning Gecko Care. © Copyright New England Herpetoculture LLC. All Rights Reserved. https://www.neherpetoculture.com/mourninggeckocaresheet

Lumen. (n.d.). The Major Histocompatibility Complex (MHC) | Boundless Microbiology. Lumen Candela. https://courses.lumenlearning.com/boundless-microbiology/chapter/the-major-histocompatibility-complex-mhc/#:%7E:text=Key%20Points%201%20Diversity%20of%20antigen%20presentation%2C%20mediated,different%20individuals%20inside%20a%20population.%20More%20items. . .%20

McDaniels, C., & Fujita, M. K. (2020, July). Opsins in the Skin of Two Color Forms of Common Lesser Earless Lizards (Holbrookia maculata). The University of Texas at Arlington.

National Geographic Society. (2019, June 6). Genetic Variation. https://www.nationalgeographic.org/encyclopedia/genetic-variation/

NHGRI. (2019, March 9). Transcriptome Fact Sheet. National Human Genome Research Institute. https://www.genome.gov/about-genomics/fact-sheets/Transcriptome-Fact-Sheet

Ranwez, V., Harispe, S., Delsuc, F., and E. J. P. Douzery. 2011. MACSE: multiple alignment of coding sequences accounting for frameshifts and stop codons. PLOS ONE 6:e22594. https://doi.org/10.1371/journal.pone.0022594

Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N., and F. Delsuc. 2018. MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Molecular Biology and Evolution 35:2582-2584. https://doi.org/10.1093/molbev/msy159

Srivastava, A., George, J., & Karuturi, R. K. M. (2019). Transcriptome Analysis. Encyclopedia of Bioinformatics and Computational Biology. https://doi.org/10.1016/B978-0-12-809633-8.20161-1.

Understanding Evolution. (2022, February 12). Mutations - Understanding Evolution. Berkeley.Edu. https://evolution.berkeley.edu/evolution-101/mechanisms-the-processes-of-evolution/mutations/

Weeresinghe, E. R., & Fujita, M. K. (2022, February). MHC Gene Summary. M.K. Fujita. file:///C:/Users/12146/Downloads/MHC%20Gene%20Summary%20-%20Elizabeth%20Rose%20Weeresinghe%20-%20BIOL%204393%20007.pdf

BIOGRAPHICAL INFORMATION

Elizabeth Rose Weeresinghe is a senior at the University of Texas at Arlington striving for a BS in Biology and a Minor in Korean. During her freshman year, Elizabeth joined The Honors College, Minority Association of Pre-Medical Students at UT Arlington, Korean Cultural Association, and the Leadership Honors Program and continued those memberships until the last semester of senior year. Junior year, Elizabeth joined The Big Event as a Site Leader, and continued through her senior year as the Programs Director. Elizabeth aspires to be a pediatric cardiologist in the future. Children have always been Elizabeth's main interest, and the cardiovascular system has peaked interest recently. By combining the two interests, Elizabeth strives towards the goal of becoming a specialized physician. Doing research with Dr. Fujita has given Elizabeth a new outlook towards research. As a freshman, the concept of research did not interest Elizabeth. Now that the research project with the Honors College is complete, Elizabeth hopes to do research during her time at medical school. Elizabeth with always be thankful for the relationships made here at UTA and will continue to keep the time spent here close.