Spring 4-2020

Diversity of the Major Histocompatibility Complex in African Penguins (Spheniscus demersus) in situ

Athena Schalk
athena.schalk@otterbein.edu

Diversity of the Major Histocompatibility Complex in African Penguins (*Spheniscus demersus*) *in situ*

Otterbein University
Department of Biology and Earth Science
Westerville, Ohio 43081
Athena Schalk

April 8th 2020

Submitted in partial fulfilment of the requirements for graduation with Honors

Simon Lawrance Ph.D. Project Advisor	_____ Advisor's Signature
Sarah Bouchard Ph.D. Second Reader	_____ Second Reader's Signature
Prof. Erica Van Dop Honors Representative	_____ Honors Rep's Signature

Acknowledgements

I would like to thank Dr. Simon Lawrance for showing me how to run DGGE gels and providing the guidance necessary to complete the research. Since starting this project, I have increased my confidence while working in the lab and have learned a lot about the Major Histocompatibility and about immunology as well as how it can be applied to various fields. I would also like to thank Dr. Mike Cranfield for providing us with the penguin DNA samples from South Africa and purifying the DNA before sending it to Otterbein University. Without those samples, it would be difficult to conduct a wide range DNA analysis on the African penguin populations and communicating with him has showed me the importance that our work has to other organizations as well. I'd also like to thank Dr. Bouchard and Prof. Van Dope for looking over my thesis and helping me refine and edit my paper. Lastly, I'd to thank Erin Ulrich for providing the reagents and buffers necessary to conduct my experiments.

Abstract

The Major Histocompatibility Complex (MHC) is responsible for the immune response in all jawed vertebrates and protects individuals against a variety of pathogens and diseases. Maintaining genetic diversity within the MHC exons is critical to protecting endangered species. African penguins (*Spheniscus demersus)* are in danger of losing their MHC diversity in isolated populations due to reductions in population size associated with environmental changes and human activity. This study analyzes the diversity within the exons in the DNA encoding the MHC by amplifying the exons through polymerase chain reaction and identifying alleles through denaturing gradient gel electrophoresis. Wild populations of African penguins from Dassen Island, Dyer Island, and Robben Island in South Africa were sampled. Four alleles were present in the populations analyzed, but the frequencies differed between each island. Cape001 was present at high frequencies within each island population. The Robben Island population had the highest frequency of cape001 compared to the other island populations, with 78.9% of individuals carrying cape001. Analyzing the variability of the MHC II region contributes to previous studies done on populations of African penguins and provides insight into how to further protect the genetic diversity of this endangered species.

Table of Contents

Introduction

The Major Histocompatibility Complex

The Major Histocompatibility Complex (MHC) is present throughout all vertebrates. The MHC regions are highly conserved and are vital to an organism's ability to fight disease. T cells – white blood cells that hunt down pathogens and targeted molecules during an immune response – are activated by MHC molecules that bind to a peptide structure, such as a protein from a foreign substance (Janeway et al. 1999). There are two classes of MHC – MHC class I and MHC class II. MHC class I is responsible for binding to cytosolic pathogens – like viruses. MHC class II is responsible for binding to extracellular pathogens and toxins (Janeway et al. 1999). DNA contains a polymorphic and extensive region that codes for the complex and different recognition sites. MHC class I and II genes code for cell surface glycoproteins that bind to antigens and present them to effector cells.

T-cells recognize intracellular and intercellular pathogens that may be a threat to the organism. MHC class I molecules present peptides of intracellular origin (such as viral proteins or peptides from cancerous cells). Presence of the loci within the MHC region in DNA indicates that there is innate protection present from the immune system. Since glycoproteins are coded within the DNA, immune cells are coded to recognize foreign particles that enter the body, regardless of previous exposure.

Since its discovery, there has been extensive research on the complex across various fields. From the MHC's role in medicine and bodily processes in humans to its usefulness as a biological marker of trans-species allele distribution between closely related species, the MHC still remains vital to understanding how immune systems defend against pathogens as well as evolutionarily how previous species have defended against said pathogens.

Role in Conservation of Endangered Species

Analysis of the MHC is essential to conservation efforts among various types of species throughout the world, particularly endangered species subjected to bottleneck and genetic drift. Limited MHC variation can also be due to a lack of selection pressure on specific alleles where rare alleles are lost because there is no selection for them, as shown in the low MHC variation in

the endangered Galapagos penguin (Bollmer and et al 593) and sequence diversity of gazelles raised in Sanliurfa of Turkey (Bozkaya and Şükrü 897). Limited variation within MHC genes could be a cause of concern and lead to inadequate protection against novel pathogens, as discussed with black howler monkeys (*Alouatta pigra*) whose low MHC-DRB exon 2 variation is suggested to be a result of a historical founder effect (Arguello 2018).

African Penguins (*Spheniscus demersus)* and Population Trends

African penguins (*Spheniscus demersus)* are flightless birds native to South Africa. It is believed that they are closely related to other penguin species, such as the Humboldt penguin (*Spheniscus humboldt*), Magellanic penguin (*Spheniscus magellanicus*), and Galapagos penguin *(Spheniscus mendiculus)*. Typically, African penguins can be found along their breeding range along the coasts of South Africa, including Dassen Island, Robben Island, Dyer Island. Populations of African penguins in situ are endangered and the species are considered at risk due to a combination of climate change and human interference. As a result, the species has experienced a bottleneck, and genetic diversity has been lost.

The nesting population of African penguins *Spheniscus demersus* was about 575,000 adults at the start of the 20th century (Crawford at el. 2010). By late 1970s, the numbers have decreased by 30% and have continued to decrease (Crawford at el. 2010). The present –day population is about 400,000 birds fewer than at the start of the 20th century, resulting in a 70% decrease in the nesting population. Possible reasons for the decrease in populations of African penguins include excessive egg harvests in the 20th century, oil spills which has impacted Dyer Island in particular, scarcity of food, predation, competition for space with seals, and human disturbance (Crawford at el. 2010). Robben Island and Dassen Island populations since the 1970s have slowly increased. From 1991-1994, there were 8,957 recorded adults on Robben Island and 30,045 recorded adults on Dassen Island. However, the number of adults counted at Robben Island decreased by almost 900 between 1993 and 1995 due to oil spills that occurred off the coast (Crawford at el. 1999). Robben Island populations have since stabilized and indicate that the loss to oiling has not halted the growth of the penguin colony (Crawford at el. 1999). From 1989 to 1995, it was estimated that 59-87% of new adults at the colony have arisen from immigration, with many of the new adults immigrating from Dassen Island and Dyer Island (Crawford at el. 1999). Dyer Island populations have continued to decline in the 1990s.

Unfortunately, there has been a decline in South African populations since the mid-2000s, which has resulted in the species national classification of Endangered. Currently, wild populations of African penguins are estimated to be about 52,000 birds (Maryland Zoo 2020) A decline in available food sources, sporadic oiling events, habitat destruction, human disturbance, and predation are contributing to declines in the African penguin populations (Sherley 2020). In 2013, the South African government has put into place a Biodiversity Management Plan to halt the decline of the species. While the plan failed to halt the decline of penguin populations, they have provided a more coordinated approach to penguin conservation and have implemented interventions to preserve the penguin populations currently present and have been revising the plan to deal with the most pressing threats to the species (Sherley 2020).

Figure 1. Map of Cape Town, South Africa. Islands of interest are noted within the map. (Crawford at el. 2010). Penguin DNA samples were taken from Dassen Island, Robben Island, and Dyer Island.

MHC and African Penguins

The purpose of this study was to gather information on the diversity of MHC II genes in wild populations of African penguins, potentially find rare alleles, and further analyze allele frequencies to better understand how they are representative of their current population. Studying rare alleles not present within captive populations will give us insight into how to introduce those alleles to captive populations and increase gene flow.

There are many reasons why there may be a decrease in MHC variability. One of which are neutral forces, such as genetic drift and gene flow. Another proposed reason is reduced selection pressure due to low exposure to pathogens. Endangered species in particular are at risk of losing crucial alleles because of small population sizes or past bottlenecks.

Previous research compared MHC diversity of captive penguins and wild penguin populations and found that captive penguin populations did not carry as much genetic diversity within their MHC sequences compared to wild populations (Willmer 2015). The number of alleles found within captive Mystic Aquarium population in Stonington, Connecticut was half of the number of alleles found within the populations within South Africa. It is essential to maintain diversity within both captive and wild populations, especially with the rapid decline of populations currently present off the coast of Africa.

Maintaining the diversity of the MHC regions within captive populations ensures that vital alleles are not lost and that individual penguins will have adequate protection against pathogens. Pathogens are a major threat to wildlife, in many cases causing population extinction. There are several pathogens that African penguins in particular are currently threatened by. Avian malaria is a mosquito-borne disease caused by parasitic protozoans (*Plasmodium)* are common throughout the world, yet are not associated with mass mortality within bird populations within the wild. However, avian malaria is detrimental to captive penguin populations. A survey completed in zoos show that 12.5% of zoos at some point have diagnosed cases of malaria in their penguin collections of African penguins and Humboldt penguins *Spheniscus humboldti* (Grillo 2016). They have also been recorded in wild penguins as well, however none of the wild penguins afflicted with malaria showed outward signs of the disease and parasiaemia were generally low (Grillo 2016).

Methods

Analysis of the Major Histocompatibility Complex through Molecular Biology Techniques

To analyze polymorphism in DNA samples, the MHC II regions of the DNA are isolated through polymerase chain reaction (PCR) and run through denaturing gradient gel electrophoresis (DGGE) followed by direct sequencing for unambiguous identification of all products.

Polymerase chain reaction (PCR) is a technique that amplifies regions of DNA from a small amount of DNA. DNA template, buffer containing Mg^{2+}, nucleotides, DNA polymerase, and primer sets that determine which regions of DNA will be amplified. To maximize separation of alleles that differ only by one or two nucleotides, a G +C rich sequence (GC-clamp) was attached to the PCR product (Knapp 1997). Since PCR products of the MHC II region are essentially the same size, a chemical gradient will separate out the alleles based on the sequences. When DNA fragments are electrophoresed through a gradient of urea and formamide at 65:35 ratio, alleles dissociate at different rates (Knapp 1997). Together, PCR, DGGE, and direct sequencing provides rapid method for high-resolution identification for alleles.

Sampling and DNA Extraction

Blood samples from African Penguins were obtained by Dr. Michael Cranfield in South Africa in 2008. Penguins were randomly sampled on Dassen Island, Dyer Island, and Robben island. The samples were processed within 12 hours to serum and blood clots and imported to USDA at Plum island to be eradiated. - treated with radiation. The DNA was extracted by adding equal volume of Phenol and TB in a 50 mL blue cap tube. The samples were gently rocked overnight then centrifuged at 3,000 rpm to separate the phases for 5-10 minutes. The top aqueous phase was extracted equal volume of phenol/chloroform/isoamyl alcohol and gently rock overnight then centrifuged at 3,000 rpm to separate the phases for 5-10 minutes. The top aqueous phase was extracted again and equal volume of chloroform was added. The mixture was gently agitated for 1 hour and the top aqueous phase was removed again. 0.1 volume Ammonium acetate and 2x volume of cold ethanol was added to the aqueous solution. The samples were then shipped to Otterbein University. The ethanol was poured off and the samples were rinsed with 70% ethanol to remove the salt. The tubes were spun down at 3,000 rpm and the 70% ethanol

decanted carefully. The pellet was air dried overnight and the DNA pellet resuspended in DNAase/RNAase free water overnight in a fridge.

Polymerase Chain Reaction (PCR)

Rehydrated DNA samples were diluted tenfold to prepare for polymerase chain reaction (PCR). Pen1 and pen4 primers were used to amplify a 198-bp fragment of exon 2 of a class II MHC *DRB1*-like gene (Tsuda et al.2001; Kikkawa et al. 2005). The pen1 primer sequence is 3'AACGGCACCGAGCGGGTGAGGT5' and the pen4 primer sequence is 3'CTGCCAACACAACTACGGG5'. A GC-clamp was added to the 5' end of pen4 to prevent degradation of DNA as it passes through the denaturing gradient gel. 12.5 μL of Buffer D, 10 μL of dH_2O, 1 μL of pen1/pen4 primer mix, and 0.5 μL of Failsafe enzyme was added to each tube and run through PCR for 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds.

Denaturing Gradient Gel Electrophoresis (DGGE)

Denaturing gradient gel was poured with 65:45 ratio of low and high solution. 100% denaturing solution was created by mixing together 40 mL of Formamide, 20 mL of acrylamide, 2 mL of 50X TAE, and 18 mL of dH_2O. 0% denaturing solution was made by mixing 20 mL of acrylamide, 2 mL of 50X TAE, and 79 mL of dH_2O. A low solution was mixed by adding 9.6 mL of 0% denaturing solution and 6.4 mL of 100% denaturing solution. A high solution was mixed by adding 5.6 mL of 0% denaturing solution and 10.4 mL of 100% denaturing solution. After adding 52.5 μL of 10% ammonium persulfate and 52.5 μL of TEMED (N, N, N', N' – Tetramethylethylene) into each high and low solution, the gels were immediately poured into the gel box set up with the 65% high solution and 45% low solution gradient. After the gels have set, 5 μL of PCR product and 5 μL of 2X loading dye are pipetted into the wells of the gel. The gels were run at 300 V for 3.5 hours at 55°C after the samples were loaded. The gels were stained in ethidium bromide and visualized in an imaging system. Each band on the gels contains a unique DNA sequence or allele.

Direct Sequencing of Alleles

Once alleles of interest were identified, gel fragments containing the band were cut out and soaked in 1 mL of dH_2O overnight. 12.5 μL of Buffer D, 10 μL of dH_2O, 1 μL of pen1 and

pen4 primer mix without the GC clamp, 0.5 μL of Failsafe enzyme, and 1 μL of gel wash were prepared for PCR. The prepared mixtures ran through 25 cycles of 95°C for 15 seconds, 55°C for 15 seconds, and 72°C for 15 seconds. 5 μL of each PCR product, 5 μL of dH_2O, 1 μL of SAP (Shrimp Alkaline Phosphatase), and 0.5 μL of Exonuclease I enzyme underwent cleanup of remaining nucleotides and primers present after initial PCR reaction, with 1 cycle at 37°C for 15 minutes followed by 80°C for 15 minutes. After clean up, 5 μL of cleanup, 5 μL of dH_2O, 5 μL of forward or reverse primer was added to an Eppendorf tube to be sent to Eurofins for sequencing. For each allele sequenced, there should be two tubes – one containing the forward primer and one containing the reverse primer.

After receiving the sequences, the forward and reverse sequences of each sample were aligned using the MacVector program to ensure that each band had unique sequencing and to pinpoint the single nucleotide differences. Each sequence was run through a BLAST search to identify if the sequences have previously been submitted before and to find similar sequences present within other penguin species.

Results

Nineteen samples from Dassen Island (DAI), sixteen penguin samples from Dyer Island (DYI), , and twenty samples from Robben Island (RI) have been analyzed through DGGE. Twelve DAI, ten DYI, and sixteen RI samples have allele #1, which has been identified as cape001. Five DAI, three DYI, and two RI samples have allele #2, which has been identified as cape002. Six DAI, eight DYI, and three RI samples have allele #3, which has been identified as cape003. Eight DAI, three DYI, and one RI samples have allele #4, which has been identified as cape004.

Fifteen of the DAI samples present heterozygosity (78.9% of DAI samples), with four samples carrying cape001 and cape003 (26.6% of heterozygous DAI samples), two samples carrying cape001 and cape002 (13.3% of heterozygous DAI samples), one sample carrying cape002 and cape003 (6.7% of heterozygous DAI samples), two samples carrying cape002 and cape004 (13% of heterozygous DAI samples), one sample carrying cape003 and cape004 (6.7% of heterozygous DAI samples), and five samples carrying cape001 and cape004 (33.3% of heterozygous DAI samples). Two DAI samples were homozygous (10.5% of DAI samples) for cape001 and no samples were homozygous for cape002, cape003, and cape004.

Heterozygosity is present in eleven DYI samples (68.8% of DYI samples), with 5 samples carrying cape001 and cape003 (45.5% of heterozygous DYI samples), three samples carrying cape001 and cape004 (27.3% of heterozygous DYI samples), two samples carrying cape002 and cape003 (18.2% of heterozygous DYI samples), and one sample carrying cape002 and cape004 (9.1% of heterozygous DYI samples). There are four homozygous samples (25% of DYI samples), two samples homozygous for cape001 (50% of homozygous DYI samples), one sample homozygous for cape002 (25% of homozygous DYI samples), and one sample homozygous for cape003 (25% of homozygous DYI samples).

In the RI samples, there are more individuals with homozygous alleles present (85% of RI samples). Fourteen samples are homozygous for cape001 (82.4% of homozygous RI samples). Two samples are homozygous for cape004 (11.8% of homozygous RI samples). One sample is homozygous for cape002 (5.9% of homozygous RI samples). Three samples are heterozygous (15% of RI samples), with two samples heterozygous for cape001 and cape003 (66.7% of heterozygous RI samples) and the other heterozygous for cape002 and cape003 (33.3$ of heterozygous RI samples).

BLAST results indicate that the alleles have been evolutionarily conserved, with strong similarities to closely related penguin species as shown in Figure 2. All alleles sequenced have been reported before as MHC II alleles present in African penguin populations.

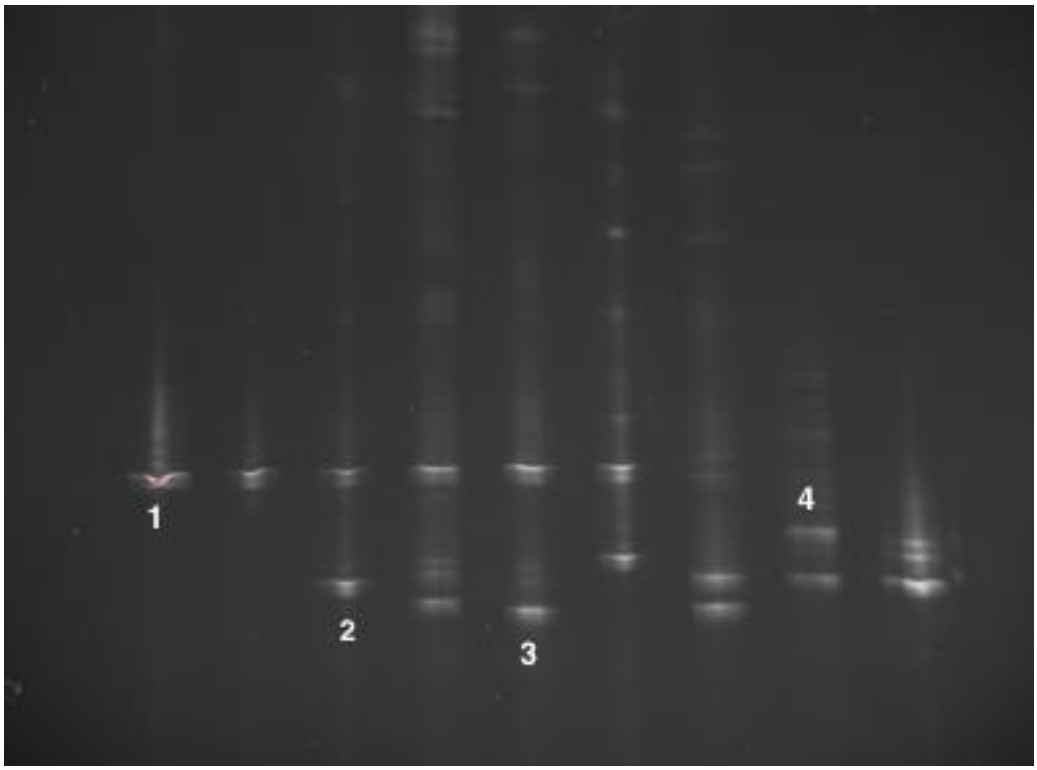
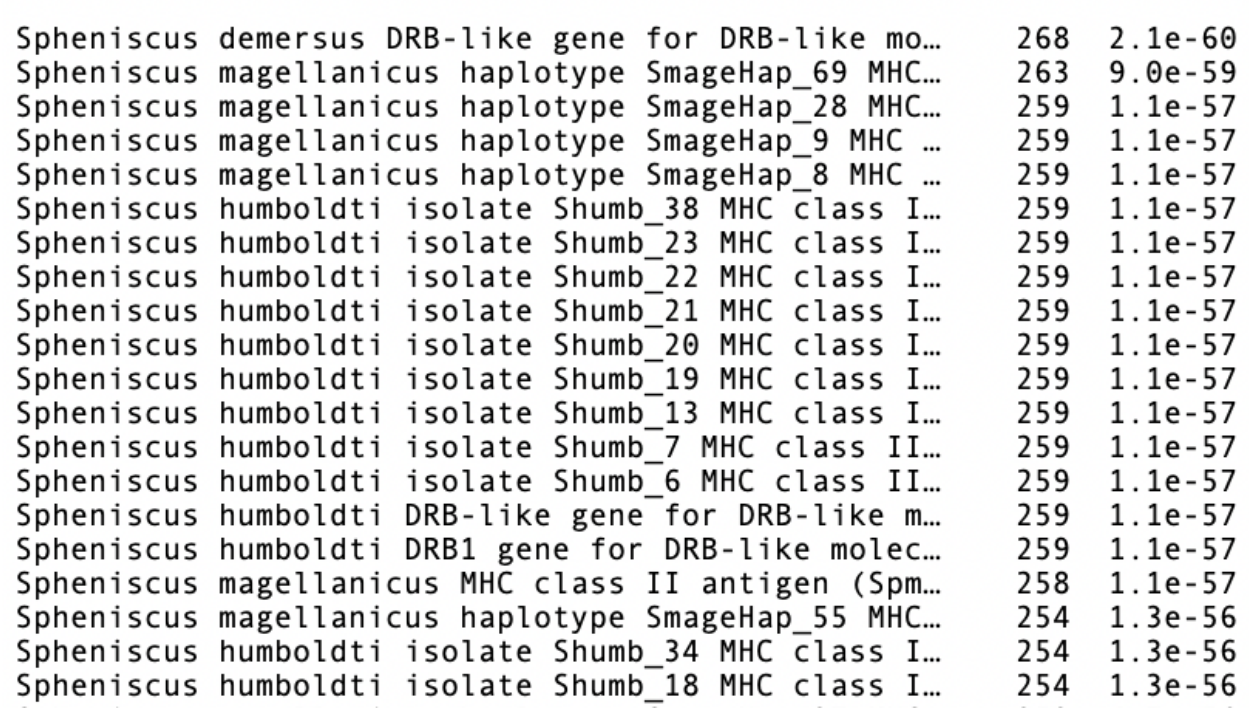


Figure 2. Denaturing gradient gel electrophoresis of 9 Dassen Island penguin PCR products generated from penguin DNA samples. Each lane within the gel contains a PCR product from an individual penguin sample whose MHC gene region has been amplified. Numbers displayed show bands sequenced. Bands 1-4 have been identified though BLAST search as followed: cape001, cape002, cape003, cape004. Lanes with only one band present is identified as homozygous for a particular allele while lanes that contain multiple bands are identified as heterozygous.



Spheniscus demersus DRB-like gene for DRB-like mo…	268	2.1e-60
Spheniscus magellanicus haplotype SmageHap_69 MHC…	263	9.0e-59
Spheniscus magellanicus haplotype SmageHap_28 MHC…	259	1.1e-57
Spheniscus magellanicus haplotype SmageHap_9 MHC …	259	1.1e-57
Spheniscus magellanicus haplotype SmageHap_8 MHC …	259	1.1e-57
Spheniscus humboldti isolate Shumb_38 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_23 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_22 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_21 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_20 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_19 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_13 MHC class I…	259	1.1e-57
Spheniscus humboldti isolate Shumb_7 MHC class II…	259	1.1e-57
Spheniscus humboldti isolate Shumb_6 MHC class II…	259	1.1e-57
Spheniscus humboldti DRB-like gene for DRB-like m…	259	1.1e-57
Spheniscus humboldti DRB1 gene for DRB-like molec…	259	1.1e-57
Spheniscus magellanicus MHC class II antigen (Spm…	258	1.1e-57
Spheniscus magellanicus haplotype SmageHap_55 MHC…	254	1.3e-56
Spheniscus humboldti isolate Shumb_34 MHC class I…	254	1.3e-56
Spheniscus humboldti isolate Shumb_18 MHC class I…	254	1.3e-56

Figure 3. BLAST result from consensus sequence of band 3 (cape003) listed in Figure 1. E-values for band 3 (cape003) are significantly small, which means that the sequences are a close match for the BLAST hits presented. While the first match for the *Spheniscus demersus* DRB-like gene is not zero (a perfect match), Resulting sequences relating to other *Spheniscus* species also display small E-values, which suggests that genetic sequences of the MHC regions of other *Spheniscus* species – specifically *Spheniscus magellanicus* and *Spheniscus humboldti* – are evolutionarily tied to a shared ancestor.

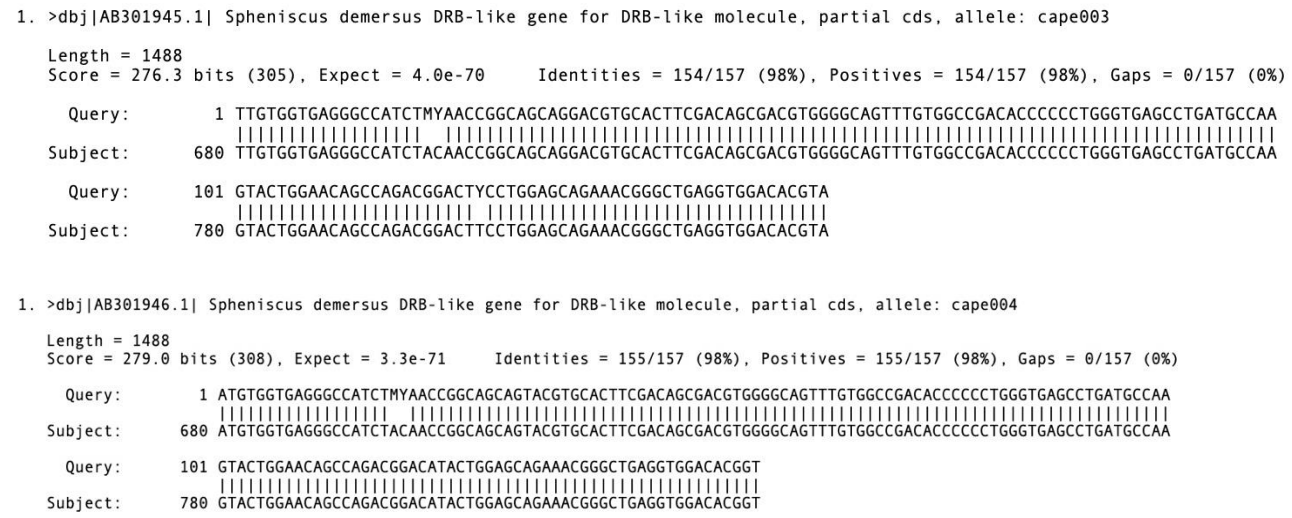


Figure 4: BLAST query search for allele #3 and allele#4. Allele #3 has been previously reported as cape003. Allele #4 has been previously reported as cape004.

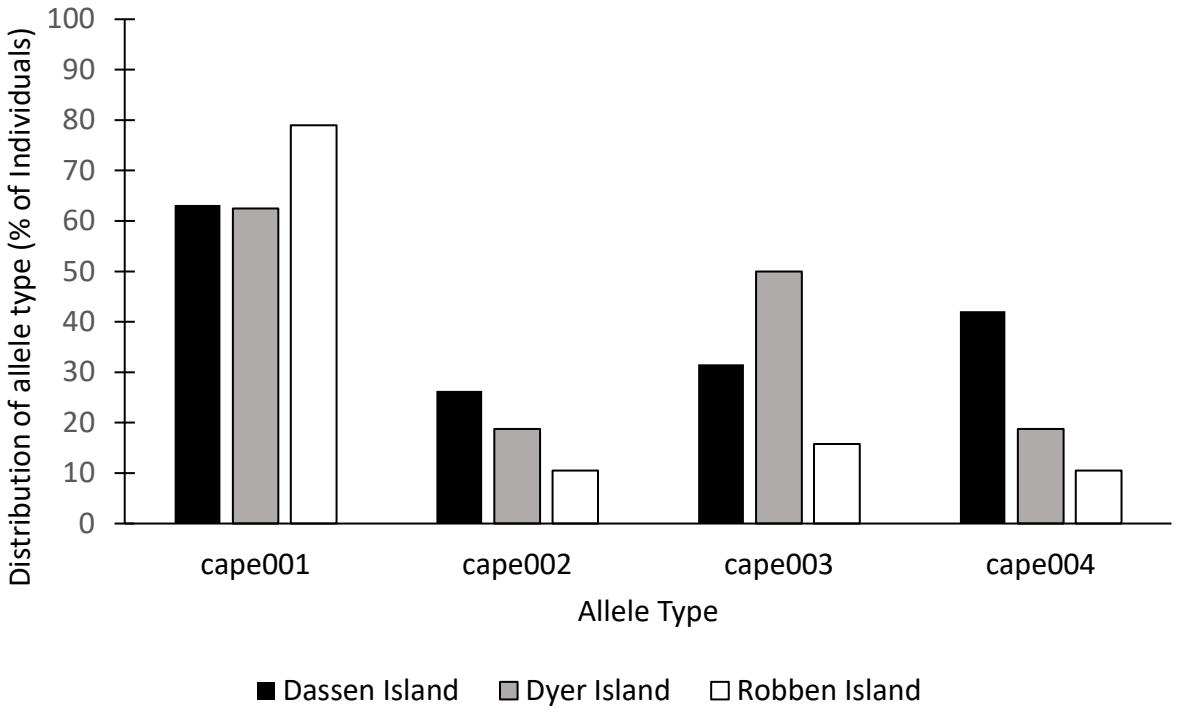


Figure 5: Distribution of allele type frequencies between Dassen Island, Dyer Island, and Robben Island samples from South Africa in 2008. N = 19 for Dassen Island. N = 16 for Dyer Island. N = 19 for Robben Island. Frequencies of allele appearance compared to the number of individuals per island population. N = number of individuals sampled per island. The frequencies of allele appearance within each population differs between each island, but cape001 is the most prevalent allele that appears in all three populations, with 63% of DAI individuals, 62.5% of DYI individuals, and 78.9% of RI individuals carrying cape001.

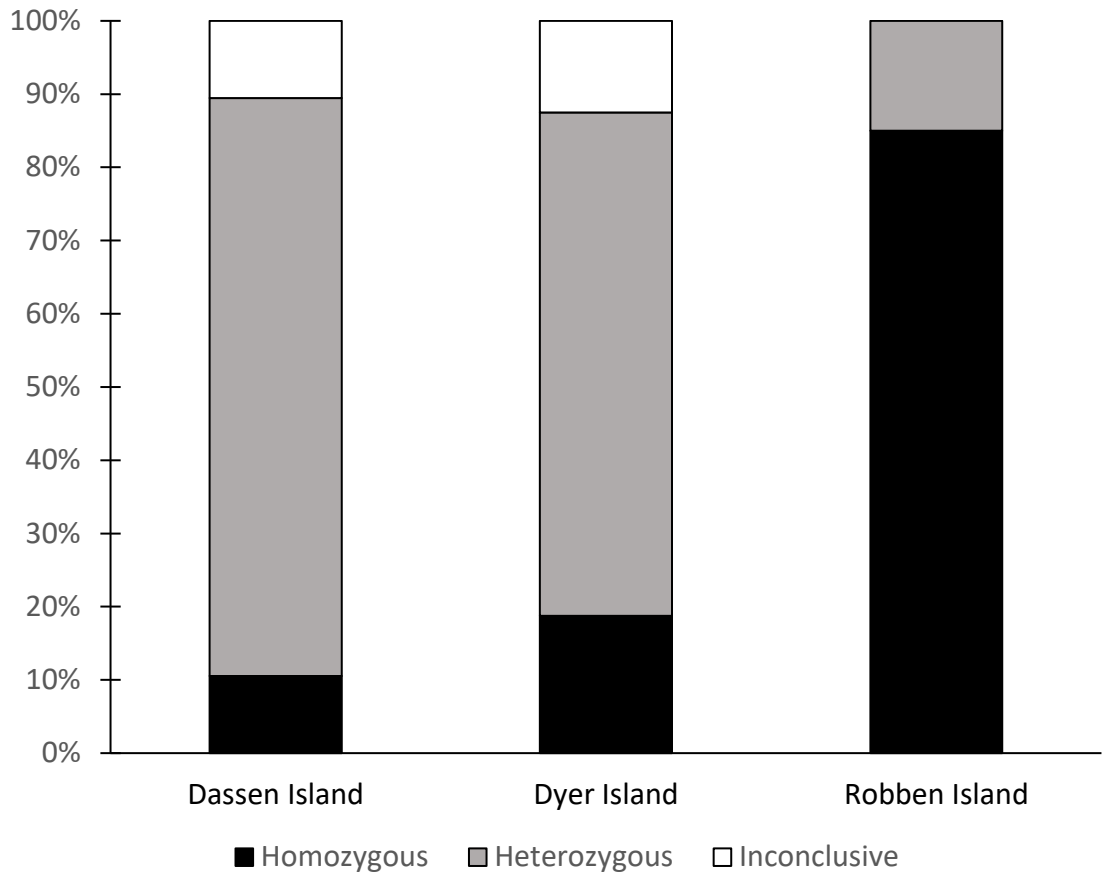


Figure 6: Percentage of African penguin individuals who displayed homozygous and heterozygous bands on the DGGE gel per population. The inconclusive samples either didn't display bands because of the absence of DNA within the PCR product or did not display distinct bands that are identifiable.

Table 7: Seventeen samples analyzed from Dassen Island. Two penguin samples displayed one singular band on the gel, which indicates that the samples are homozygous for the tested allele. The only alleles displaying homozygosity is allele 1, which is also known as cape001. *PCR product for DAI25 displayed no bands, which suggests that no DNA was present to be amplified through PCR. It is possible that the DNA pellet was swept out by 70% ethanol when the solution was discarded. Homozygous expression is shown by having only one type of allele present within the gel (1/1, 2/2, 3/3, 4/4). Heterozygous expression is shown by having two different types of alleles present within the gel (1/2, 1/4, 3/4, etc).

Samples	**Alleles**	**Homozygous**
DAI-1	1/3	No
DAI-2	1/2	No
DAI-3	1/1	Yes
DAI-5	2/3	No
DAI-6	2/4	No
DAI-12	1/4	No
DAI13	1/4	No
DAI-14	3/4	No
DAI-15	1/4	No
DAI-16	N/A	N/A
DAI-17	1/4	No
DAI-18	1/3	No
DAI23	1/1	Yes
DAI24	1/4	No
DAI25*	N/A	N/A
DAI26	1/3	No
DAI28	2/4	No
DAI-94	1/3	No
DAI-95	1/2	No

Table 8: Sixteen samples analyzed from Dyer Island. PCR product for DYI-7 displayed no bands, which suggests that no DNA was present to be amplified through PCR. It is possible that the DNA pellet was swept out by 70% ethanol when the solution was discarded. The bands for PCR product DYI 51 were inconclusive and requires more testing to identify alleles within the sample. Homozygous expression is shown by having only one type of allele present within the gel (1/1, 2/2, 3/3, 4/4). Heterozygous expression is shown by having two different types of alleles present within the gel (1/2, 1/4, 3/4, etc).

Samples	**Alleles**	**Homozygous**
DYI-1	1/3	No
DYI-2	1/3	No
DYI-3	3/3	Yes
DYI-4	1/1	Yes
DYI-5	2/2	Yes
DYI-6	1/4	No
DYI-7	N/A	N/A
DYI-8	1/3	No
DYI 36	1/3	No
DYI43	1/4	No
DYI47	1/4	No
DYI48	1/3	No
DYI49	2/3	No
DYI50	2/3	No
DYI51	N/A	N/A
DYI52	1/1	Yes

Table 9: Samples from Robben Island analyzed through DGGE. All PCR products displayed bands when run onto DGGE gels. Homozygous expression is shown by having only one type of allele present within the gel (1/1, 2/2, 3/3, 4/4). Heterozygous expression is shown by having two different types of alleles present within the gel (1/2, 1/4, 3/4, etc).

Samples	Alleles	Homozygous
RI6	1/1	Yes
RI7	1/1	Yes
RI8	1/1	Yes
RI9	1/1	Yes
RI10	2/2	Yes
RI12	4/4	Yes
RI13	1/1	Yes
RI14	1/1	Yes
RI15	1/1	Yes
RI16	1/3	No
RI17	1/1	Yes
RI18	1/1	Yes
RI19	1/1	Yes
RI20	2/3	No
RI21	4/4	Yes
RI22	1/1	Yes
RI24	1/3	No
RI25	1/1	Yes
RI26	1/1	Yes
RI27	1/1	Yes

Discussion

There is variation within the MHC II gene region in African penguin populations. Heterozygosity is present in many of the penguins sampled, 78.9% of DAI samples, with 68.8% of DYI samples, and 85% of RI samples testing heterozygous. However, the majority of the penguin populations for Dassen, Dyer, and Robben Island express cape001, which suggests that there is selective pressure on the populations that promote the presence of cape001. For example, the threat of parasites within South Africa, such as malaria, may give the populations that carry cape001 an advantage and allow them to reproduce and survive at higher rates. As a result, the frequency of all the alleles could potentially be a result of natural selection and environmental pressures that select for that particular allele.

Variability in allele frequency is also apparent between the populations present at Robben Island, Dassen Island, and Dyer Island. Robben Island has higher incidences of homozygous gene expression (with 85% of RI samples) compared to Dassen Island (10.5% of DAI samples) and Dyer Island (18.8% of DYI samples). Robben Island has also seen drastic decreases in population in recent years due to oil spills, habitat destruction, and food scarcity. Increase in homozygous expression may be a result of bottleneck and as a result the allele frequency of less common alleles have decreased. This is particularly concerning because rapid decreases in populations can impact allele frequencies in a short amount of time. The reduction in the MHC variability could potentially put populations on Robben Island at risk by decreasing the population's ability to fight off a wider range of infections and parasites, specifically novel pathogens.

In general, preserving the diversity of the MHC II region in captive African penguins are of upmost importance. With the captive penguin population at the Mystic Aquarium, Spde001 (or cape001 as I have identified it) is the most predominant allele present and heterozygous expression showing similar patterns to Dassen and Dyer Island (Wilmer 2015). However, cape004 is not present within the captive population, as well as two new alleles that have been found within wild populations in Wilmer's research (Wilmer 2015). There is more diversity present within wild populations compared to captive populations, with cape004 present in Dassen, Dyer, and Robben Island. Unfortunately, the two novel alleles found in previous research has not been identified within the tested individuals. However, additional testing of individuals and populations may reveal the presence of these novel alleles. Parasitic diseases are

prevalent in many zoos throughout the world and put many captive populations at risk. By preserving the diversity of the MHC II region, their immune systems would be capable of defending more individuals from different diseases, such as malaria. Captive populations have limited gene flow, which would also impact the frequencies of the alleles and make it substantially different from wild populations. Understanding the allele frequencies of wild populations and how the alleles impact the immune system is critical to protecting captive populations as well. For example, breeding programs at zoos can introduce a rare allele or two that are mostly found in wild populations to increase the genetic diversity of the MHC region in captive penguins and improve their immune systems.

Parasitic sequences, such as malaria, should be analyzed for further research to determine the role of the MHC II diversity in the immunity of wild African penguins, specifically parasitic proteins that T-cells would recognize inside the body. Since MHC molecules are designed to activate T-cells and have them recognize the presence of foreign proteins, identifying sequences for parasitic proteins could potentially allow us to identify which MHC molecule and allele activates those T-cells and provides the immune protection. There are many different types of proteins that could be released by pathogens. Toxins and external proteins in parasitic pathogens would most likely be identified by MHC II molecules. Understanding the resistance of parasites, such as malaria, and its connection to the MHC will allow conservation programs to target specific alleles that increase resistance in breeding programs. By doing so, captive populations of African penguins will have increased resistance to pathogens such as malaria. Zoo programs throughout the world – such as the Baltimore Zoo – could benefit from understanding which MHC alleles most protect from the pathogens that present the most threats to captive penguin populations. Understanding which alleles are critical to protecting against parasitic molecules could provide conservation programs with information necessary to protect wild populations from diseases native to their particular area. Breeding programs for wild populations of African penguins could focus on maintaining genetic diversity and preventing the loss of less common alleles that could protect the species from a wider range of parasitic molecules.

The close resemblance to sequences to other species suggest that evolutionarily the species are closely related and share a common ancestor. Based on how similar the sequences are, we could potentially construct a cladogram that can trace the ancestry of the African penguins and other species since the MHC regions are heavily conserved and potentially trace

the evolutionary branches of the closely related species. It would be beneficial to analyze the amino acid sequence of the DNA samples and determine how each allele differs from each other as well as provide information on the types of molecules that the immune system would protect against. As always, sequencing more samples would increase the sample size of the population and would provide more accurate data on the MHC diversity of wild African penguins on each island. Dr. Mike Cranfield has also sent over penguin samples from other islands that have not been analyzed within this study. Looking into the MHC diversity of penguins on Betty's Bay and Boulders Island would also provide insight into wild population diversity and the populations could potentially hold rare alleles that are not present with Dassen, Dyers, and Robben Island. Willmer's research had identified alleles that haven't been reported before in Boulders Beach and Dyers island (Wilmer 2015). It would be beneficial to sequence more bands present within the gel as a way of verifying that those are the alleles present within the individual samples.

Bibliography

Aleksandra Biedrzycka, Emily O'Connor, Alvaro Sebastian, et al. "Extreme MHC Class I Diversity in the Sedge Warbler (Acrocephalus Schoenobaenus); Selection Patterns and Allelic Divergence Suggest That Different Genes Have Different Functions." *BMC Evolutionary Biology, Vol 17, Iss 1, Pp 1-12 (2017)*, no. 1, 2017, p. 1.

Arguello, Sánchez, L. E., Arguello, J. R., García–Feria, L. M., et al. "MHC Class II DRB Variability in Wild Black Howler Monkeys (Alouatta Pigra), an Endangered New World Primate." *Animal Biodiversity and Conservation, Vol 41, Iss 2, Pp 389-404 (2018)*

Bollmer, Jennifer L., Vargas, F. Hernán, Farker, Patricia G. "Low MHC Variation in the Endangered Galápagos Penguin (Spheniscus Mendiculus)." *Immunogenetics*, vol. 59, no. 7, July 2007, pp. 593–602.

Bozkaya, Faruk, and Şükrü Gürler. "Sequence Diversity of MHC Class-II DRB Gene in Gazelles (Gazella Subgutturosa) Raised in Sanliurfa of Turkey." *Journal of Genetics*, vol. 97, no. 4, Sept. 2018, pp. 897–903.

Crawford, R.J.M., Shannon L.J., Whittington P.A. 1999. "Population Dynamics of the African Penguin *Spheniscus demersus* at Robben Island, South Africa" . *Marine Ornithology* 27: 139-147.

Crawford, R.J.M. , Williams A.J , J. Hofmeyr H., Klages N.T.W., Randall R.M., Cooper J., Dyer B. M. & Chesselet Y. "Trends of African penguin Spheniscus demersus populations in the 20th century". *South African Journal of Marine Science*, 16:1, 101-118, DOI: 10.2989/025776195784156403 (1995)

Grilo, M.L, Vanstreels, R, Wallace, R., et al. "Malaria in penguins – current perceptions". *Avian Pathology*. Vol. 45, no. 4, 2016. pp. 393-407.

Janeway, Charles A., Travers, Paul, Walport, Mark. "Immunobiology: The Immune System in Health and Disease". New York: Garland Publishing. 1999. Pp.116-118. Print

Kikkawa EF, Tsuda T, Naruse TK, Sumiyama D, Fukuda M, Kurita M, Murata K, Wilson RP, LeMaho Y, Tsuda M, et al. 2005. Analysis of the sequence variations in the mhc DRB1-like gene of the endangered humboldt penguin (spheniscus humboldti). Immunogenetics 57(1-2):99.

Kikkawa EF, Tsuda TT, Sumiyama D, Naruse TK, Fukuda Michio, Kurita M, Wilson RP, LeMaho Y, Miller GD, Tsuda M, et al. May 2009. Trans-species polymorphisms of the mhc class II DRB-like gene in banded penguins (genus spheniscus). Immunogenetics 61(5):341-52.

Knapp, L.A., Lehmann. E., Hennes, L., et al. "High-resolution *HLA-DRB* typing using denaturing gradient gel electrophoresis and direct sequencing." *Tissue Antigens vol.* 50. 1997. pp. 170-177

Maryland Zoo. "African Penguin." *The Maryland Zoo,* 2020. www.marylandzoo.org/animal/african-penguin/.

Sallaberry-Pincheira, Nicole, González-Acuña, Daniel, Padilla, Pamela, et al. "Contrasting Patterns of Selection between MHC I and II across Populations of Humboldt and Magellanic Penguins." *Ecology & Evolution (20457758)*, vol. 6, no. 20, Oct. 2016, pp. 7498–7510.

Sherley, Richard B., Crawford, Robert J.M, Blocq, Andrew, The conservation status and population decline of the African penguin deconstructed in space and time. 2020

Willmer, Emily, "A Comparison of the Genetic Diversity of the Class II Major Histocompatibility Complex in ex situ and in situ Populations of African Penguins (Spheniscus demersus)" (2015). *Honors Thesis Projects*. 5.