ORNITOLOGÍA NEOTROPICAL
(2022) 33: 142–153
ORIGINAL ARTICLE

CHARACTERIZATION OF MOLECULAR GENETIC MARKERS IN *SPHENISCUS* BANDED PENGUINS AND THE IDENTIFICATION OF A STRANDED JUVENILE IN CENTRAL AMERICA

Gillian M. Ross[1] · Katelyn I. Schumacher[1] · Gustavo Cañas-Valle[1] · Carlos R. Hasbun[2] · Juan L. Bouzat[1*]

[1] Department of Biological Sciences, Bowling Green State University, Bowling Green, OH 43403, USA.
[2] Fundación Zoológica de El Salvador, El Salvador.
E-mail: Juan L. Bouzat · jbouzat@bgsu.edu

Abstract · In this study, the mitochondrial cytochrome c oxidase 1 gene, three nuclear microsatellite markers, and the MHC class II *DRβ1* exon 2 were assessed for species-specific differences that would allow the diagnostic identification of *Spheniscus mendiculus, S. humboldti,* and *S. magellanicus* specimens. Analyses of reference samples for these species revealed that genetic variation at these markers showed species-specific haplotypes and alleles that can provide positive evidence for species identification. Bayesian cluster analyses demonstrated high probability of assignment (>99%) for individual samples to their corresponding species. The set of nuclear and mitochondrial markers studied proved useful for the identification of a juvenile penguin stranded on the Pacific Ocean shores of El Salvador as a Magellanic Penguin. The negative consequences of accidental captures of Magellanic Penguins by fishermen and the relocation of wildlife through human intervention are discussed.

Resumen · Caracterización de marcadores genético-moleculares en pingüinos *Spheniscus* para la identificación de un espécimen juvenil varado en Centroamérica
En este estudio, el gen mitocondrial citocromo c oxidasa 1, tres marcadores de microsatélites y el exón 2 del gen CMH *DRβ1* de clase II fueron evaluados en busca de diferencias especie-específicas que permitan la identificación de especímenes de *Spheniscus mendiculus, S. humboldti* y *S. magellanicus*. El análisis de muestras de referencia reveló que la variación genética en estos marcadores presenta haplotipos y alelos especie-específicos que pueden proveer evidencia positiva para la identificación de las especies. Análisis bayesianos demostraron altas probabilidades de asignación (>99%) de muestras individuales a sus correspondientes especies. El grupo de marcadores nucleares y mitocondriales estudiados demostró utilidad para la identificación, como pingüino de Magallanes, de un individuo juvenil varado en las costas del océano Pacífico en El Salvador. Las consecuencias negativas de capturas incidentales de pingüinos de Magallanes por la industria pesquera y la relocalización de individuos silvestres mediante intervención humana son discutidas.

Key words: Cytochrome c oxidase 1 gene · MHC class II *DRβ1* exon 2 · Microsatellite markers · Species identification · Vagrant

INTRODUCTION

Spheniscus banded penguins constitute the most tropical penguin genus, containing the most northerly distributed penguin species across the southern hemisphere. Their geographic range encompasses the southern coast of Africa (*S. demersus*), Pacific and Atlantic coastal regions of South America (*S. humboldti* and *S. magellanicus*), and the Galapagos islands (*S. mendiculus*), as well as numerous other islands across their distribution (García Borboroglu & Boersma 2013, Ramos et al. 2018). Three species of *Spheniscus* penguins inhabit coastal regions of the Pacific Ocean, including the Galapagos Penguin, the Humboldt Penguin, and the Magellanic Penguin.

Galapagos Penguins primarily inhabit limited coastal regions of Fernandina and Isabela Islands (Valle 1986, Vargas et al. 2005; Figure 1), where the Cromwell current upwells, creating a pocket of highly productive cold water. The population is estimated to be between 1,500 to 4,700 individuals (Boersma et al. 2013a), and their colonies exhibit low levels of genetic diversity —likely due to founder effects and serial bottlenecks because of El Niño Southern Oscillations (Bollmer et al. 2007, Nims et al. 2008, Arauco-Shapiro et al. 2020). In this species, movements are limited mainly to areas close to their nesting sites, except during El Niño events, when they may move farther to find productive waters (Boersma et al. 2013a).

Humboldt Penguins are endemic to the coasts of Chile and Peru (Figure 1), where the Humboldt current carries productive cold Antarctic waters to the region. The species has an estimated population of 30,000 to 40,000 breeding individuals (De la Puente et al. 2013), and their populations show some genetic structuring among colonies from Peru, northern Chile, and central-southern Chile (Vianna et al. 2014). Humboldt Penguin feeding migration is limited, except during El Niño events, when individuals travel farther south in search of productive waters (Culik et al. 2000).

Submitted 6 October 2021 · First decision 18 November 2021 · Acceptance 8 March 2022 · Online publication 28 October 2022
Communicated by María Andreína Pacheco

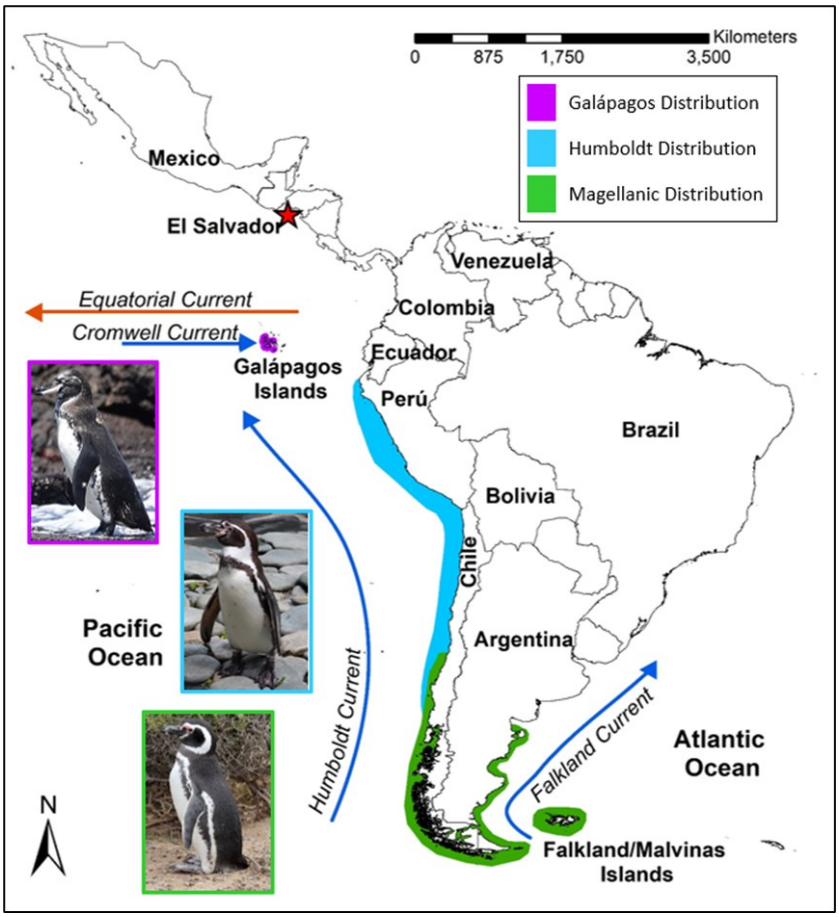

Figure 1. Map of Central and South America showing the distribution ranges of *Spheniscus mendiculus* (Galapagos Penguin), *S. humboldti* (Humboldt Penguin), and *S. magellanicus* (Magellanic Penguin). The stranded penguin was found on the Pacific Ocean shore of El Salvador, indicated by a red star. The warm Equatorial Current is represented by an orange arrow. Cold ocean currents are represented by blue arrows. BirdLife International and Handbook of the Birds of the World (2016) provided spatial data for species distribution ranges; data downloaded May 2021. Penguin pictures by Mike's Birds (2019), Przemek Pietrak (2019), and Dominic Sherony (2014); distributed under Creative Commons public licenses CC BY-SA 2.0 and CC BY-SA 3.0.

Magellanic Penguins inhabit the Atlantic and Pacific coasts of southern South America (Figure 1). The estimated population size is between 1.2–1.6 million breeding pairs (Boersma et al. 2013b), and their primary breeding grounds range from Cape Horn to 40°S and includes the Falkland/Malvinas Islands (Boersma et al. 2013b). Bouzat et al. (2009) found limited genetic structuring among Magellanic Penguin colonies distributed throughout the South Atlantic Ocean, likely a result of population intermixing through natal dispersal and the large effective population size of the species. However, genetic comparisons using both nuclear and mitochondrial markers revealed significant differentiation between populations from the Pacific and Atlantic Oceans (Bouzat et al. 2013). Magellanic Penguins on the Pacific coast have limited feeding migration patterns, remaining close to their colonies (Skewgar et al. 2014). Conversely, those on the Atlantic coast follow the Falkland current, migrating as far north as southern Brazil to feed and overwinter (Stokes et al. 2014).

Based on the geographic distribution and migratory patterns of *Spheniscus* species, and given the patterns of oceanographic currents, finding vagrant individuals outside of their natural range is not uncommon. Changes in ocean currents due to cyclic El Niño and La Niña events alter patterns of upwelling and productivity, thereby influencing the movements and population sizes of Galapagos (Boersma 1998, Boersma et al. 2013a) and Humboldt (Hays 1986, Culik et al. 2000) Penguins, which could potentially lead to vagrant penguin sightings. For example, during the 1997–1998 El Niño event, Humboldt individuals satellite-tracked from Pan de Azucar Island (northern Chile) traveled up to 895 km south in search for better conditions and productive waters (Culik et al. 2000). There are some reports of long-distance vagrants, such as Humboldt Penguins displaced to the Northern hemisphere (reviewed by Van Buren & Boersma 2007, Scordino & Akmajian 2012) and Little Penguins (*Eudyptula minor*) have been observed in Chile, more than 10,000 km from their native range in New Zealand (Valverde & Oyarzo 1996, Brito 1999, Wilson et al. 2000). The finding of unidentified penguins (e.g., juveniles) or biological remnants (e.g., tissue remnants or feathers) of vagrant specimens stranded on ocean shores warrants the need for a genetic method for the assignment of species identity. Thus, we assessed a set of genetic markers to determine their usefulness for species identification among Galapagos, Humboldt, and Magellanic Penguins —the three banded penguin species distributed in South America. The mitochondrial cytochrome c oxidase subunit 1 (mtDNA *COI*) gene, three nuclear microsatellite markers, and exon 2 of the Major Histocompatibility Complex (MHC) class II *DRβ1* gene were assessed for species-specific alleles that would allow diagnostic identification of the three *Spheniscus* species.

The use of mtDNA *COI* as a DNA barcoding gene has been remarkably successful in discriminating avian species (Hebert et al. 2004, Kerr et al. 2007, Tavares & Baker 2008), since it is not only highly conserved across taxa (due to its essential

Figure 2. Photographic images of the stranded penguin specimen found on the Pacific Ocean shores of Ahuachapán Department in El Salvador. Measuring tape approximates the body length at 54 cm. Stranded penguin photo credits: Fundación Zoológica de El Salvador (FUNZEL) and SalvaNatura (www.salvanatura.org).

role in metabolism), but also relatively variable between species due to the higher evolutionary rate of mtDNA genes compared to nuclear genes (Hebert et al. 2003a, 2003b). Microsatellites are short, tandem repeats of about one to six nucleotides found at high frequency in non-coding regions of the nuclear genome (Toth et al. 2000, Ellegren 2004, Selkoe & Toonen 2006). The high mutation rate of these markers creates high levels of allelic diversity, including unique alleles, which can be used to detect hybridization events or to determine population origin and species identification of individual specimens (Schlötterer 2000, Jan et al. 2010, Dawson et al. 2013, Hibbets et al. 2020). Finally, the MHC class II *DRβ1* gene encodes antigen-presenting surface proteins, which are crucial to the vertebrate immune system (Benacerraf 1981, Snell 1981, Piertney & Oliver 2006, Ujvari & Belov 2011). Given their role in the immune system, class II loci exhibit extraordinarily high degrees of polymorphism (Doherty & Zinkernagel 1975, Sallaberry-Pincheira et al. 2016), which are associated to disease resistance; therefore, their high levels of variation make these markers valuable for species identification, as well as evolutionary studies on the adaptive value of ecologically relevant genes.

Here, we assessed the genetic variability of mtDNA *COI*, microsatellites, and MHC *DRβ1* markers in reference populations of Galapagos, Humboldt, and Magellanic Penguins, to evaluate the presence of species-specific variation useful for species identification. These genetic markers were then used to perform a Bayesian quantitative assignment of species identity of a juvenile banded penguin stranded on the Pacific shores of Ahuachapán, El Salvador (Jones & Komar 2007). We then compared lorum morphology (specifically the shape of the loreal edge) of this stranded individual with reference pictures of the three species to identify a diagnostic character that would complement our genetic data. Our study provides a set of molecular markers for the identification of *Spheniscus* species from South America and the Galapagos Islands. The identification of a stranded juvenile penguin on the Pacific shores of Central America is discussed in relation to the distribution ranges, migratory patterns, and conservation of these species.

METHODS

Sample collection and DNA extraction. Genetic data from reference populations of Galapagos, Humboldt, and Magellanic Penguins, used to assess the suitability of markers for species identification, were collected from studies performed in our lab (Bouzat et al. 2009, 2013; Arauco-Shapiro et al. 2020, Hibbets et al. 2020). Blood samples from 38 Galapagos Penguins were collected in 1997–1998 from breeding colonies on Fernandina and Isabela islands (with permits from the Galapagos National Park Service, Ecuador, to Dee Boersma). Samples from 23 Humboldt Penguins were obtained during routine exams by veterinary personnel in 2017 from a captive population at the Chicago Zoological Society's Brookfield Zoo (Illinois, USA), which had known ancestry traced back to individual penguins captured in the wild in northern Peru. The Magellanic Penguin reference population included samples from 45 individuals collected in 2004 at three breeding colonies (Sebastiana = 18; Puñihuil = 16; Ahuenco = 11) located in the northern distribution of the species on the Pacific Ocean coast of Chile, along with samples from 20 individuals previously collected by Prof. Dee Boersma (University of Washington) in 1998 from Punta Arenas, on the southern tip of South America. The sampling of Magellanic Penguins was conducted with corresponding permits from the Division of Forestry (CONAF) and Chile's Secretary of Fisheries (Permit # 3523, 2003), and then transported to the United States with importation permits from the USDA -APHIS (USDA permit # 51802 to Juan L. Bouzat, Bowling Green State University).

Whole blood samples were collected by puncture of the brachial vein and stored in Queen's lysis buffer (0.01M Tris, 0.01M NaCl, 0.01M EDTA, and 1% *n*-lauroylsarcosine, pH 7.5;

Seutin et al. 1991) or stored in lithium-heparin tubes. Total DNA was extracted using standard phenol–chloroform extraction protocols, followed by ethanol precipitation (Sambrook et al. 1989) or through Qiagen DNeasy Blood and Tissue DNA Extraction kits (Qiagen Inc., Valencia, CA, USA).

On 7 June 2007 an emaciated juvenile banded penguin was reported stranded on the Pacific shores of Ahuachapán, El Salvador and ultimately died later that day (Jones & Komar 2007). Morphology clearly indicated the stranded penguin was a juvenile banded penguin (Figure 2); however, it was not clear to which particular species this individual belonged because the plumage had not developed the distinctive white bands that are characteristic of adult banded penguins. Tissue samples were collected by Carlos R. Hasbun (Fundación Zoológica de El Salvador) from the liver and muscle of the deceased stranded penguin and stored in 90% ethanol. DNA was extracted using the Qiagen DNeasy Blood and Tissue DNA Extraction kit. Extractions from liver tissue were discarded, as gel electrophoresis revealed the DNA was highly degraded.

DNA sequencing and genotyping. A section of the mtDNA *COI* gene and the contiguous tRNA-Cys (partial) and tRNA-Tyr (complete) genes were amplified by PCR using primer sets EM5287 (5′-CAC ATC AAT GAG CTT GCA ACTC-3′) and COI-R722 (5′-TAA ACT TCA GGG TGA CCA AAA AAT YA-3′). These primers have previously shown to amplify mtDNA of *Spheniscus* penguins (Bouzat et al. 2009, 2013, Arauco-Shapiro et al. 2020, Hibbets et al. 2020). PCR amplifications were performed in 25 µl volumes containing approximately 40 ng of DNA, 1X of GoTaq Flexi Buffer, 1 mM of $MgCl_2$, 0.08 mM of each dNTP, 0.4 µM of each primer, and 0.5 units of GoTaq Flexi DNA polymerase (Promega, Madison, WI, USA). The PCR amplification profiles included an initial denaturing step for three min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 52°C, and 30 s at 72°C, with a final extension step of 5 min at 72°C. PCR products were purified by ethanol precipitation and then sent for direct sequencing (forward and reverse) at the University of Chicago Comprehensive Cancer Center DNA Sequencing and Genotyping Facility.

Three microsatellite markers (G3-6, G2-2, and M1-2) originally developed for *Spheniscus* penguin species (Akst et al. 2002, Bouzat et al. 2013) were used for species identification. These markers have previously shown variation within Galapagos, Humboldt, and Magellanic Penguin species (Akst et al. 2002, Bouzat et al. 2009, 2013; Arauco-Shapiro et al. 2020, Hibbets et al. 2020). PCR amplifications were performed using forward (F) and reverse (R) primers for microsatellite G3-6 (F: 5′-TCT TAA GGT CTT GCA CAC-3′ and R: 5′-CAG CTC AGT AAC TGC AGG CA-3′), microsatellite G2-2 (F: 5′-ATG ACA TAT TGA TTG GC-3′ and R: 5′-CTG CCT GAA CTA AGC TTT GTC-3′), and microsatellite M1-2 (F: 5′-GCT TCC AAG AAG CTT GTG AC-3′ and R: 5′-ACT GAA CTT TGT CTG CGT GC-3′). Amplification reactions were performed by PCR in 25 µl volumes containing approximately 40 ng of DNA, 1X of GoTaq Flexi Buffer, 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.4 µM of each primer, and 0.625 units of GoTaq Flexi DNA polymerase. The PCR amplification profiles included an initial denaturing step for 2 min at 95°C followed by 30 cycles of 45 s at 95°C, 45 s at the respective annealing temperature, and 45 s at 72°C, with a final extension step of 10 min at 72°C. The annealing temperature for microsatellite markers G2-2 and M1-2 was set at 51°C, while the annealing temperature for microsatellite marker G3-6 was set at 53°C. Amplification products were sent for fragment analysis at the University of Chicago.

PCR amplification of the MHC class II *DRβ1* exon 2 was performed using primers Lpen.hum1F2 (5′-ACT CCT GGC ACA GCC GCG TG-3′) and Lpen.hum2R (5′-ACA CGC TCT CCC CTC CTG TG-3′), which were originally developed by Kikkawa et al. (2005, 2009) and have been shown to be locus-specific for the MHC of *Spheniscus* penguins (Kikkawa et al. 2005, 2009, Knaffler et al. 2012, Arauco-Shapiro et al. 2020, Hibbets et al. 2020). PCR amplification reactions were performed in 25 µl volumes containing approximately 40 ng of DNA, 1X of GoTaq Flexi Buffer, 1.5 mM of $MgCl_2$, 0.2 mM of each dNTP, 0.5 µM of each primer, and 0.5 units of GoTaq Flexi DNA polymerase. The PCR amplification profiles included an initial denaturing step for 2 min at 95°C followed by 27 cycles of 1 min at 94°C, 1 min at 62°C, and 2 min at 72°C, with a final extension step of 15 min at 72°C. PCR products were purified through standard ethanol precipitation and sent to the University of Chicago for direct sequencing to classify individuals as either homozygotes or heterozygotes. Individual alleles were then confirmed through cloning and sequencing of amplification products, following the protocol outlined by Arauco-Shapiro et al. (2020). Cloning and sequencing allowed confirmation of individual MHC alleles for every genotype in all samples analyzed.

Finally, mitochondrial haplotype, microsatellite genotypes, and MHC sequences of the stranded penguin specimen were confirmed through three independent reactions, using independent DNA extractions as templates for DNA amplification of haplotype/genotypes.

Analyses of genetic data from reference species and the stranded penguin. All DNA sequences and genotypes were analyzed using Geneious v. R6.1.8 (Kearse et al. 2012). Final *COI* sequences analyzed included 807 base pairs, with the first 93 bases corresponding to the tRNA-Cys gene (partial) and tRNA-Tyr gene (complete) of the mitochondrial genome. The sequences of each haplotype detected in the reference populations are available on GenBank under accession numbers MN565806, MN565807, MN565811, MN830876–MN830879, MZ852655, and MZ852656. Sequences of MHC alleles included 419 base pairs, with exon 2 of the *DRβ1* gene located at bases 93–362. The sequences of each allele detected in the reference populations are available on GenBank under accession numbers MN565812–MN565815, MN565817, MN565819–MN565823, MN565831, MN565832, MN830880, MN830881, MN830884–MN830888, MN830894, MN830897–MN830904, and MZ852649–MZ852654.

Variation at each marker was assessed to estimate differences in haplotype/allele frequencies between reference samples and to identify potential species-specific differences. A marker was considered informative of species identity if there were species-specific haplotypes/alleles detected in at least one of the three reference species studied. The genetic profile of the stranded penguin was then compared to the reference samples. A match of a haplotype or allele observed in the stranded penguin to a species-specific marker in a reference population was considered positive evidence for species identification.

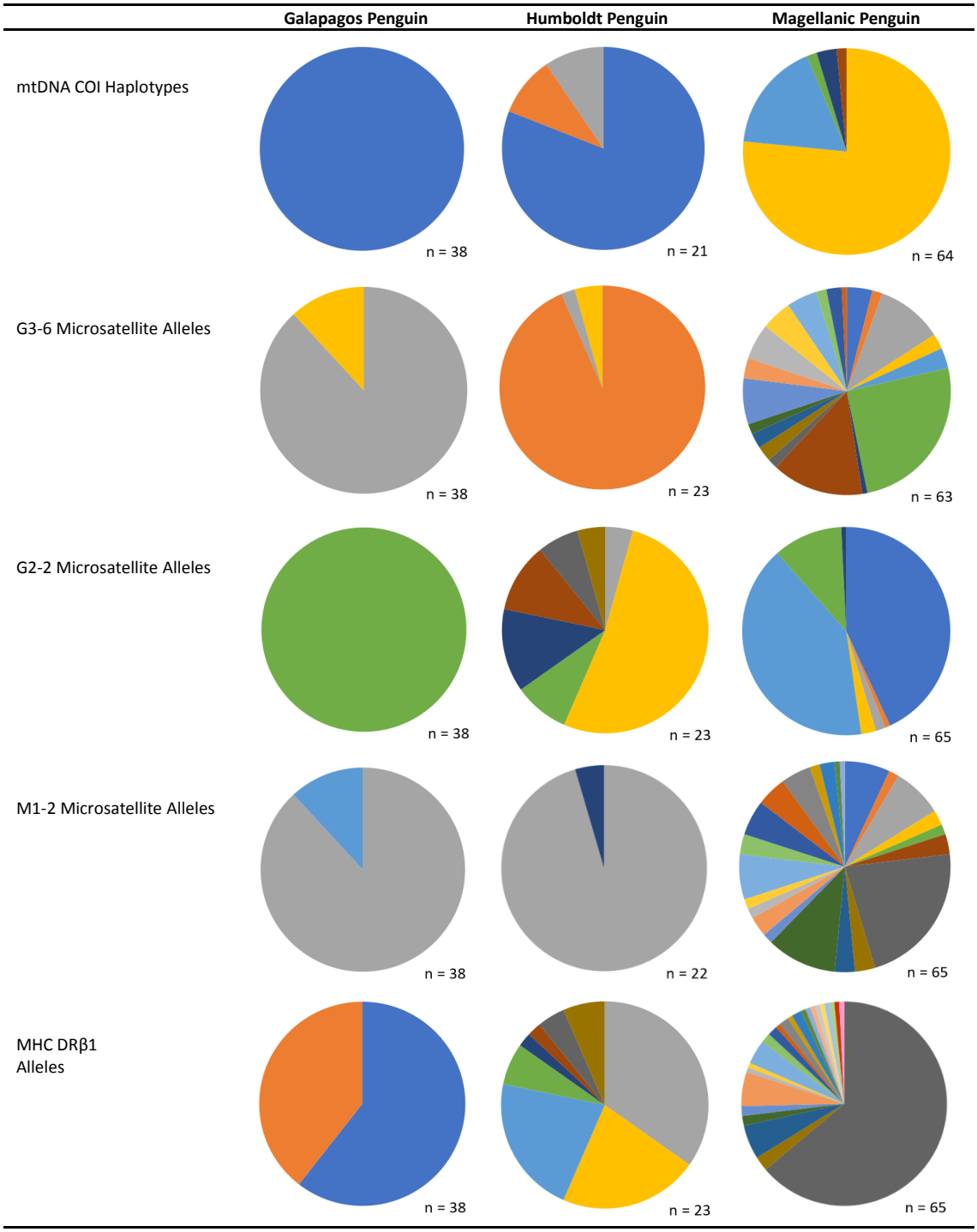

Figure 3. Haplotype/allele frequencies of the mtDNA *COI* gene, microsatellite loci G3-6, G2-2, and M1-2, and the MHC *DRβ1* gene found in Galapagos, Humboldt, and Magellanic Penguin reference populations. For each marker, different colors represent distinct haplotypes/alleles. Identical haplotypes/alleles from different species are represented by the same color. The number of individuals sampled for each marker (n) appears in the bottom right corner of each chart.

Structure analysis and assignment probability of the unidentified stranded penguin. Cluster analyses using Structure v2.3.4 (Pritchard et al. 2000) were performed to characterize the reference samples and estimate assignment probabilities of the stranded specimen to the reference species. Structure uses Bayesian model-based clustering methods and multilocus genotype data to identify distinct genetic clusters and probabilistically assign individuals to said clusters (Pritchard et al. 2000). Structure is traditionally used to discover population structure within species; however, in our analyses, it was used to evaluate the performance of genetic markers in defining reference populations (in this case three species) and then quantify the genetic identity of the stranded penguin. This approach allowed estimates of the relative genetic contribution of each reference species to the genome of the unidentified specimen as the basis for its species assignment.

Prior to running Structure on the microsatellite and MHC alleles, microsatellite locus M1-2 was eliminated because previous linkage analysis by Hibbets et al. (2020) revealed significant disequilibrium between loci M1-2 and G3-6 within Humboldt and Magellanic populations. For the Structure analyses, we used the no admixture-independent allele frequencies model. The no admixture ancestry model was selected because the reference populations represented three discrete species. The model selected assumed independent allele frequencies for the three species, as levels of variation at individual loci are driven by population-level processes associated with the independent demographic dynamics of

Table 1. Haplotypes/alleles detected in the unidentified stranded penguin and the Galápagos, Humboldt, and Magellanic Penguin reference populations. Bold font designates species-specific haplotypes/alleles. Green highlighting indicates alleles detected in the stranded penguin sample that were informative of species identity. Pink highlighting indicates alleles detected in the stranded penguin that were uninformative as they were shared between at least two species. Haplotypes/alleles from different species with identical sequences (*Smen01*/*Shum01*, *Spma001*/*Sphu007*, *Spma013*/*Sphu008*) are underlined.

	Stranded penguin	Galapagos Penguin	Humboldt Penguin		Magellanic Penguin		
MtDNA *COI*	**Smag01**	Smen01	Shum01 **Shum02**	**Shum03**	**Smag01** **Smag02**	**Smag03** **Smag04**	**Smag08**
G3-6	266 **298**	266 268	264 266 268		**262** 264 266 268 **270** **272** **274**	**276** **278** **282** **284** **286** **288** **290**	**292** **294** **296** **298** **300** **310**
G2-2	**373** **381**	383	377 379 383 387	**389** **391** **393**	**373** **375** 377 379	**381** 383 387	
M1-2	212 **247**	212 **214**	212 **216**		**209** **211** 212 **213** **215** **217** **219** **221**	**223** **225** **227** **231** **233** **235** **237** **239**	**241** **243** **245** **247** **249** **255** **259**
MHC *DRβ1*	Spma001/ Sphu007 **Spma038**	**Spme001** **Spme002**	**Sphu001** **Sphu002** **Sphu003** **Sphu004** **Sphu005** **Sphu006** Sphu007 Sphu008		Spma001 Spma013 **Spma002** **Spma003** **Spma004** **Spma006** **Spma008** **Spma009**	**Spma010** **Spma011** **Spma012** **Spma015** **Smpa019** **Spma021** **Spma025** **Spma026**	**Spma028** **Spma030** **Spma031** **Spma033** **Spma038** **Spma042** **Spma043** **Spma045**

each species.

An initial Structure analysis including only the reference samples was performed to assess the ability of the markers and the selected model in properly assigning reference individuals to Galapagos, Humboldt, or Magellanic clusters. Then, we added the stranded individual to determine the assignment probability of this individual to the reference populations. For each Structure run, we used a burn-in period of 50,000 iterations and 50,000 iterations after the burn-in to compute posterior probabilities for identifying the number of distinct clusters (*K*) estimated from the data. We performed 10 replicate runs for each *K* (*K* = 1 through 5) to verify the consistency of our estimates between runs. At the inferred *K*, individual assignments (*Q*) were consistent across runs. The online program Structure Harvester (Earl & vonHoldt 2012) was used to calculate mean likelihoods per *K* as well as Δ*K* (Evanno et al. 2005) from the Structure output.

Morphological characterization of the stranded penguin. Photographs of the full body and head of the stranded penguin were used to evaluate potential diagnostic morphological traits. Pictures of the stranded penguin were taken on 7 June 2007, the same day the penguin was found and then died. Images showed a juvenile banded penguin (Figure 2); however, plumage had not yet developed the characteristic black and white bands on the head and throat, which are commonly used for species identification.

Given the heritable nature of beak morphology in birds, the loreal region (lore, or the area located between the eyes and nostrils) was evaluated across Galapagos, Humboldt, and Magellanic Penguin specimens. A qualitative comparative analysis of the loreal region was done using 15 field photographs of each species, obtained from the eBird image database (Sullivan et al. 2009). It was determined that the edge between the base of the bill (rhamphotheca) and the bare skin of the face forms a distinctive, species-specific shape in each of the three *Spheniscus* species evaluated. One of us (GC-V) used the field photographs to create reference sketches of Galapagos, Humboldt, and Magellanic heads with the species-specific loreal edge shapes highlighted. We then compared the loreal edge morphology between the stranded specimen and the three reference species to confirm the species identification that resulted from our genetic analysis.

RESULTS

Characterization of molecular markers for species identification. Overall, eight unique *COI* haplotype sequences were observed among the three species reference populations (Table 1). The Galapagos Penguin exhibited only one haplotype (Smen01), which was identical in sequence to the most common Humboldt Penguin haplotype (Shum01; Figure 3, Supplementary Figure 1). Of the seven species-specific haplotypes, two were observed in the Humboldt Penguins and five in the Magellanic Penguins (Table 1); therefore, the *COI* marker would have the potential to provide positive evidence only for Humboldt and Magellanic species identity. Most of the *COI* base pair sites were non-variable across all three species, with only 20 polymorphic nucleotide positions defining differences between haplotypes (Supplementary Figure 1). Haplotypes observed within a species only differed by one to three base pairs, whereas there were 14 fixed

differences when comparing all Magellanic Penguin haplotypes to all Galapagos/Humboldt Penguin haplotypes (Supplementary Figure 1). By fixed differences, we mean that, at these 14 sites, all Magellanic sequences were identical to one another, whereas all Galapagos/Humboldt sequences were identical to one another, but the nucleotides appearing in the Magellanic Penguins were different from those in the Galapagos/Humboldt Penguins (Supplementary Figure 1).

The three microsatellite markers studied (G3-6, G2-2, and M1-2) exhibited varying levels of diversity. Galapagos and Humboldt Penguins lacked diversity at the G3-6 locus compared to the Magellanic Penguin, with 2, 3, and 20 alleles observed, respectively (Table 1, Figure 3). Of the 20 unique alleles observed among the three species, 17 were specific to the Magellanic Penguin, and the remaining three were shared between two or more species (Table 1). Although no species-specific alleles were observed for the Galapagos and Humboldt reference populations, there were marked differences in the frequencies of shared alleles across the species (Figure 3). Based on the presence of species-specific alleles, the G3-6 marker had the potential to provide positive evidence for Magellanic species identity.

Compared to the Magellanic and Humboldt Penguins, which exhibited seven alleles each at the G2-2 locus, Galapagos Penguins were nonpolymorphic and fixed for an allele that was shared across the three species (Table 1 and Figure 3). Of the 10 unique alleles observed among the three species, three were specific to the Humboldt Penguin, three were specific to the Magellanic Penguin, and the remaining four were shared between two or more species (Table 1). There were marked differences in the frequency of shared alleles between the three species (Figure 3). Based on the presence of species-specific alleles, the G2-2 marker had the potential to provide positive evidence for Humboldt and Magellanic species identity.

Galapagos and Humboldt Penguins lacked diversity at the M1-2 locus compared to the Magellanic Penguin, with 2, 2, and 23 alleles observed, respectively (Table 1, Figure 3). Of the 25 unique alleles observed among the three species, one was shared by all three species, one was specific to the Galapagos Penguin, one was specific to the Humboldt Penguin, and the remaining 22 were specific to the Magellanic Penguin (Table 1). The allele shared among all three species was at similar frequency in the Galapagos (0.88) and Humboldt Penguins (0.96), but at a much lower frequency in the Magellanic Penguin (0.08; Figure 3). Given the presence of at least one species-specific allele in each reference population, the M1-2 marker had the potential to provide positive evidence for Galapagos, Humboldt, and Magellanic species identity.

The analysis of the MHC *DRβ1* sequences revealed 2, 8, and 24 alleles in the Galapagos, Humboldt, and Magellanic Penguins, respectively (Table 1, Figure 3). Of the 32 unique alleles observed among the three species, two were specific to the Galapagos Penguin, six were specific to the Humboldt Penguin, 22 were specific to the Magellanic Penguin, and the remaining two alleles were shared between the Humboldt and Magellanic Penguin (Table 1). The shared allele *Sphu007*/*Spma001* was the most frequently observed allele in the Magellanic population (*Spma001* = 0.64), but it showed a low frequency in the Humboldt population (*Sphu007* = 0.04; Figure 3). The shared allele *Sphu008*/*Spma013* had a low frequency in both the Humboldt (*Sphu008* = 0.07) and Magellanic (*Spma013* = 0.02) reference populations (Figure 3). Given the presence of at least two species-specific alleles in each reference population, the *DRβ1* marker had the potential to provide positive evidence for Galapagos, Humboldt, and Magellanic species identity.

Genetic identification of the stranded penguin. The analysis of the stranded penguin's mtDNA *COI* sequence revealed a match with the Smag01 Magellanic species-specific haplotype, exhibiting the 14 distinctive base pair differences fixed for the species (Table 1, Supplementary Figure 1). The Smag01 haplotype was the most frequently observed haplotype within the Magellanic Penguin reference population (0.77; Figure 3).

Microsatellites G3-6, G2-2, and M1-2 also provided positive evidence for species identification of the stranded penguin, which was heterozygous for all three microsatellite markers (Table 1). For microsatellite G3-6, the stranded penguin revealed alleles 266 and 298. Although allele 266 was uninformative because it was found in all three reference species, allele 298 was informative as it was specific to the Magellanic Penguin. For microsatellite G2-2, the stranded penguin exhibited alleles 373 and 381, both unique to the Magellanic Penguin species. For microsatellite M1-2, the stranded penguin exhibited alleles 212 and 247. However, allele 212 was uninformative because it was found in all three reference species, whereas allele 247 was specific to the Magellanic Penguin. In summary, four of the six microsatellite alleles detected in the stranded penguin were alleles that were species-specific to the Magellanic Penguin (Table 1), whereas the remaining two alleles were found in all three reference species (Table 1).

The stranded penguin had a heterozygote MHC *DRβ1* genotype with alleles *Spma001*/*Sphu007* and *Spma038*. Although the shared allele (*Spma001*/*Sphu007*) was found in both Humboldt and Magellanic Penguins, it was the most frequently observed in the Magellanic reference population (*Spma001* = 0.64), and it was found at a low frequency in the Humboldt reference sample (*Sphu007* = 0.04; Figure 3). In contrast, *Spma038* was found to be species-specific to the Magellanic Penguin (Table 1).

Structure analyses. The Structure analyses confirmed the use of microsatellites and MHC as informative markers for species identification and provided quantification of the assignment probability of the unidentified stranded sample to one of the reference species. The initial Structure analysis of reference samples under the no-admixture model with independent allele frequencies, revealed the highest posterior probabilities for K = 3 (-1110.98, SD = 0.43; see Supplementary Figure 2). Analysis of ΔK revealed that K = 3 was the optimal number of clusters estimated from the data (ΔK at 3 = 399.97), representing the three reference populations corresponding to the Galapagos, Humboldt, and Magellanic Penguins (Supplementary Figure 2). Assignment probabilities for individual samples revealed that all individuals were unambiguously assigned to their respective species (mean Q for Galapagos = 100.00%, Humboldt = 99.99%, Magellanic = 99.96%).

The Structure analysis with the inclusion of the stranded individual also resulted in higher posterior probabilities for K

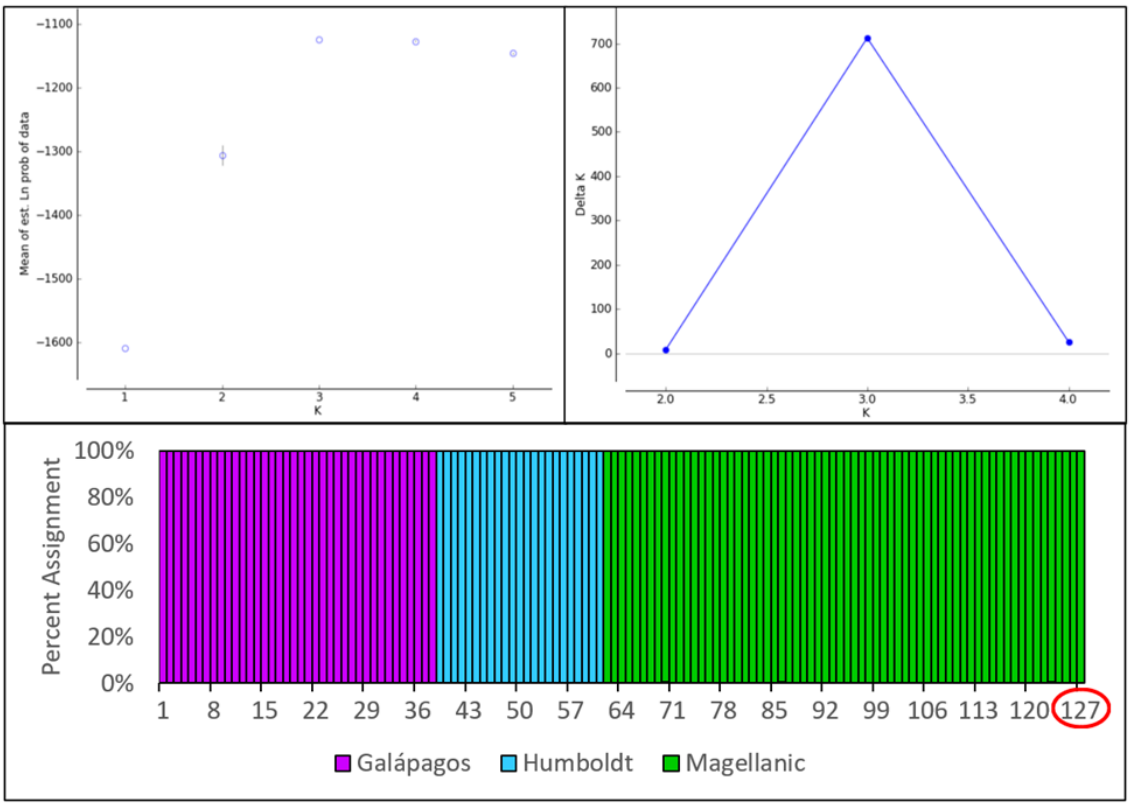

Figure 4. Structure analysis of the stranded penguin with reference samples from Galapagos, Humboldt, and Magellanic Penguins. Top left: Mean log probabilities of data (posterior probabilities) for each tested *K,* indicating highest probability and lowest variability at *K* = 3. Plot created using Structure Harvester (Earl & vonHoldt 2012). Top right: Plot of Δ*K* values created using Structure Harvester (Earl & vonHoldt 2012), showing a definitive peak in Δ*K* at *K* = 3. Bottom: Assignment probabilities (*Q*) for all individuals at *K* = 3. Each vertical bar represents an individual penguin. The three genetic clusters, each represented with a different color, corresponded exactly to the three reference species and all reference individuals were strongly assigned (>99%) to the proper species group. The stranded penguin (bar 127, highlighted by a red circle) was 100% assigned to the Magellanic reference species.

= 3 (–1123.65, SD = 0.26), consistently clustering the stranded penguin with the Magellanic Penguin group (Figure 4). Analysis of Δ*K* revealed that *K* = 3 was the optimal number of clusters estimated from the data (Δ*K* at 3 = 712.26; Figure 4 and Supplementary Figure 3). This analysis resulted in 100% probability of assignment of the stranded penguin to the Magellanic Penguin cluster (Figure 4).

Morphological characterization of the stranded penguin. Photographs of the body and head of the stranded penguin indicated that the specimen represented a juvenile banded penguin with a body length of approximately 54 cm (Figure 2). The qualitative analysis of Galapagos, Humboldt, and Magellanic Penguin images revealed that the shape of the loreal edge, between the base of the bill (rhampotheca) and the bare skin of the face, forms a species-specific shape in each of the three *Spheniscus* species evaluated. The curvature of the loreal edge appeared to resemble the “W” shape characteristic of Galapagos Penguins, an “S” shape found in Humboldt Penguins, and a single convex curve in Magellanic Penguins (Figure 5). These patterns were confirmed across 15 independent pictures for each species. Lastly, images of the stranded penguin were compared to each reference species, revealing that the loreal edge of the stranded penguin was a match to that of the Magellanic Penguin (Figure 5).

DISCUSSION

Our study provides a set of molecular markers that can be used for species identification of individuals or byproducts (e.g., skins or feathers) from the three *Spheniscus* species analyzed. These markers can also aid in identification of Humboldt and Magellanic individuals in areas of sympatry, where mixed species colonies are common and there is documentation of hybridization (Simeone et al. 2009, Hibbets et al. 2020). We identified species-specific haplotypes/alleles for at least one of the three species at each of the five markers evaluated; each marker type (mtDNA *COI*, microsatellites, and MHC *DRβ1*) revealed, however, certain limitations as well as advantages for species identification.

It is key to note that, in our study, the mtDNA *COI* was only a reliable barcoding gene for the Magellanic Penguin. Kerr et al. (2007) tested 643 species of North American birds and found that 6% of the species possessed *COI* sequences that were shared with other species, likely explained by recent divergence of sister taxa or hybridization events. Since diversification amongst extant *Spheniscus* species occurred within the last four million years and Galapagos and Humboldt Penguins are considered sister species (Baker et al. 2006, Gavryushkina et al. 2017), it was not entirely surprising that Galapagos and Humboldt penguins shared a common mitochondrial haplotype. However, Tavares & Baker (2008) evaluated *COI* for 60 avian sister species pairs and found that even closely related pairs had unique DNA sequences with 5-64 fixed nucleotide differences that could be used for species identification. It seems that the intense effects of genetic drift due to founder effects and subsequent bottlenecks in Galapagos Penguins, combined with their recent divergence from Humboldt Penguins, may explain the lack of diversity and interspecific sharing of a haplotype between these sister

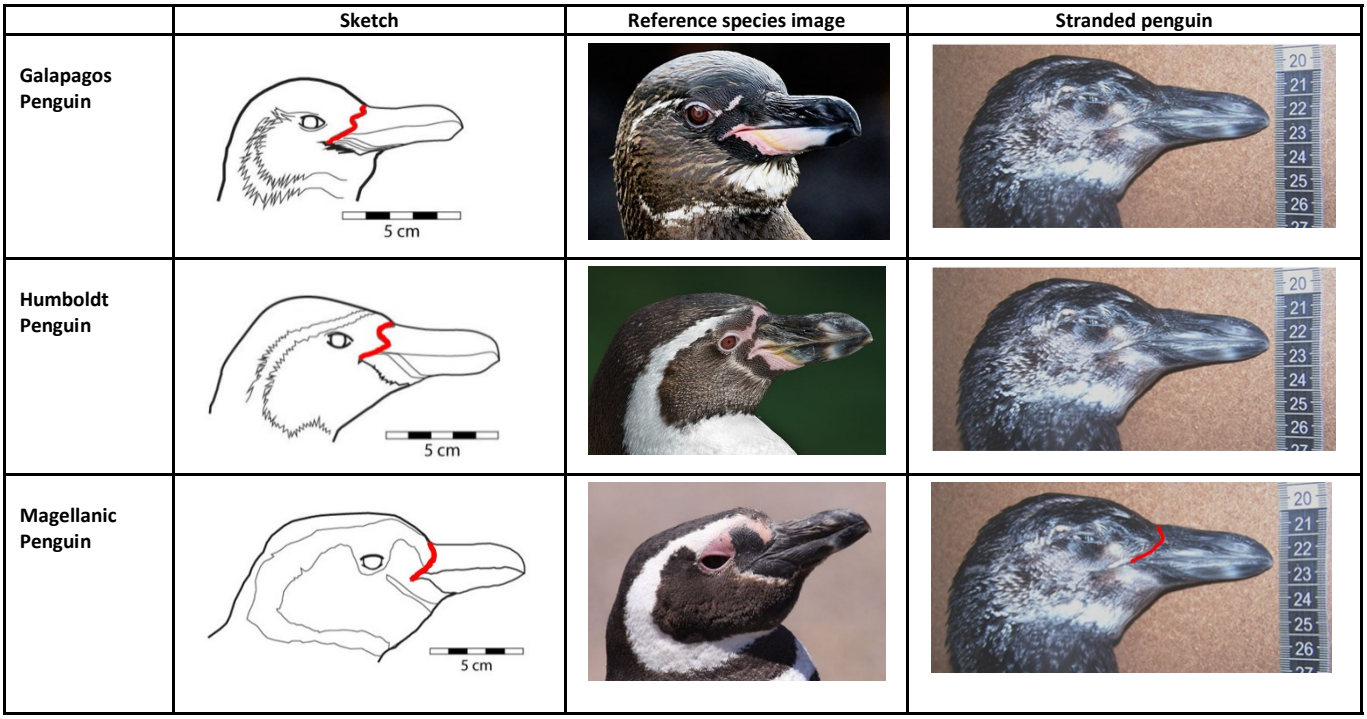

Figure 5. Morphological characterization of the loreal edge in reference samples of Galapagos, Humboldt, and Magellanic Penguins, and the stranded specimen found in El Salvador. Composite sketches of the head of each reference species are shown for qualitative comparative analysis. The loreal edge on the sketches and that of the stranded penguin are highlighted in red. The pictures show that the loreal edge of the stranded penguin specimen matches that of the Magellanic Penguin. Sketches created by Gustavo Cañas-Valle. Reference photographs taken by Pedro Szekely (2018), Dori (2008), and Dfaulder (2010); distributed under Creative Commons public licenses CC BY-SA 2.0 and CC BY-SA 3.0.

species. Indeed, Galapagos and Humboldt Penguin reference mitogenomes revealed that the two species shared identical sequences for six of the thirteen mtDNA genes, including *COI* (Ramos et al. 2018).

Approximately 84% of the microsatellite alleles observed in the Magellanic Penguin were species-specific. Due to the low proportion of species-specific alleles in the Galapagos (20%) and Humboldt Penguins (33%), the microsatellite loci studied are not ideal for identifying these species based solely on allele identity. However, alleles that are at high frequency in one species but rare in another can potentially provide supportive evidence of species identity, despite shared ancestral polymorphisms. For example, there were no Galapagos- or Humboldt-specific alleles observed at the G3-6 locus, and yet allele frequencies were quite divergent among the three species (see Figure 3), which would be informative during cluster analyses. Such was the case in our study, as the nuclear molecular markers were sufficient for Structure to resolve the three reference populations with high assignment probabilities (> 99%). These observations highlight the importance of not relying solely on the identification of species-specific alleles in an unidentified sample, but on utilizing highly polymorphic markers to produce quantitative estimates of species identity based on distributions of allele frequencies.

In contrast to the *COI* gene sequences, the MHC *DRβ1* exon 2 sequences revealed no fixed differences between species, despite showing approximately five times as many polymorphic sites. This finding is not unexpected as Kikkawa et al. (2009) showed that *DRβ1* exon 2 sequences from *Spheniscus* species do not form into distinct species clades, as it was observed for the three *Pygoscelis* penguin species. Although we found that the two MHC alleles exhibited by the Galapagos Penguin were species-specific, the Humboldt and Magellanic Penguins shared two distinct MHC alleles. Identical *DRβ1* exon 2 sequences in Magellanic and Humboldt Penguins have been previously reported (Kikkawa et al. 2009, Sallaberry-Pincheira et al. 2016, Hibbets et al. 2020). The sharing of ancestral polymorphisms in the MHC between *Spheniscus* species has been attributed to recent and rapid speciation, maintenance by balancing selection, or homoplasia due to similar environmental pressures (Kikkawa et al. 2009, Sallaberry-Pincheira et al. 2016). Despite the potential limitations described above, the Structure analyses confirmed the combined use of microsatellites and MHC as informative markers for species identification.

We tested the efficacy of the studied markers through the species assignment of a juvenile banded penguin stranded in El Salvador. The genetic analysis of the unidentified stranded penguin provided unambiguous evidence of species identity, revealing that the specimen was indeed a Magellanic Penguin (Tables 1 and 2, Figure 4). The mitochondrial haplotype of the stranded specimen showed positive evidence for Magellanic species identity and negative evidence for Galapagos or Humboldt Penguin species (Table 2). Additionally, microsatellites provided confirmatory evidence that the stranded penguin specimen was a Magellanic Penguin. Four of six microsatellite alleles detected in the stranded specimen revealed positive evidence for Magellanic Penguin identity and negative evidence for Galapagos or Humboldt identity (Table 2), and the remaining two microsatellite alleles were shared between all reference species, thereby making them uninformative.

Furthermore, the MHC *DRβ1* genotyping of the stranded penguin revealed that one MHC allele was unique to the Magellanic species, indicating positive evidence for Magellanic Penguin identity (Table 2). The other MHC allele sequence was shared between Humboldt and Magellanic spe-

Table 2. Genetic evidence for species identification of the stranded penguin found in El Salvador. The table summarizes results from each genetic marker (mtDNA, microsatellites, and MHC) used to characterize the unidentified penguin specimen. Positive evidence for species identification (i.e., presence of species-specific haplotypes/alleles) is designated by a positive symbol ("+"), and negative evidence (i.e., presence of haplotypes/alleles that are not found in the corresponding species) is designated by a negative symbol ("-").

	Galapagos	Humboldt	Magellanic
mtDNA *COI*	–	–	+
Microsatellites	– – – –	– – – –	+ + + +
MHC *DRβ1*	– –	–	+

cies and was therefore uninformative beyond providing negative evidence of Galapagos species identity. Structure analysis based on nuclear markers confirmed the species assignment of the stranded penguin by consistently clustering the individual with 100% assignment to the Magellanic Penguin reference group (Figure 4), thus suggesting that our conclusions from the genetic characterization of the stranded penguin are consistent with the initial species identification reported by Jones & Komar (2007). The qualitative analysis of images of Galapagos, Humboldt and Magellanic Penguins revealed that the shape of the loreal edge may also be used as a trait that is informative of *Spheniscus* species identity, as it showed species-specific differences. The comparison of the loreal edge in the three reference species with that of the stranded penguin specimen revealed a clear match to the Magellanic Penguin, making it consistent with the species assignment based on the genetic analysis.

Given the geographic distribution and migratory patterns of *Spheniscus* species, as well as the patterns of major Pacific Ocean currents off the shores of South America, we had an *a priori* expectation that the banded penguin stranded on the shores of El Salvador was either a Galapagos or a Humboldt Penguin. Changes in ocean currents due to cyclic El Niño and La Niña events alter patterns of upwelling and productivity, thereby influencing the movements and population sizes of Galapagos (Boersma 1998, Boersma et al. 2013a) and Humboldt (Hays 1986, Culik et al. 2000) Penguins, which could potentially lead to vagrant penguins reaching Central America. Our genetic analysis revealed, however, that the stranded specimen was a juvenile Magellanic Penguin. In fact, this individual represents the first record of a Magellanic Penguin being found north of the Equator (Jones & Komar 2007), over 6,400 km away from Chiloé Island, at the northern range of the distribution of Magellanic Penguin breeding colonies in southern Chile.

Finding a Magellanic Penguin stranded on a beach in El Salvador raises questions about how the bird journeyed so far away from its natural range. Magellanic Penguins exhibit long-distance migration and are the only *Spheniscus* species where migration during the non-breeding season is part of the annual cycle (García Borboroglu & Boersma 2013). There have been reports of vagrant penguins traveling thousands of kilometers within the southern hemisphere (see Van Buren & Boersma 2007, García Borboroglu & Boersma 2013). The distance traveled by this stranded Magellanic Penguin is similar to that of other vagrant penguins; however, the marine environments are significantly different. Vagrant penguins in the southern hemisphere can exploit cold, productive ocean currents, such as the Antarctic Circumpolar current, and travel thousands of kilometers without exceeding water temperatures of 10–15°C (Van Buren & Boersma 2007). However, a penguin traveling from southern Chile or Argentina —where Magellanic colonies are commonly found— to El Salvador would have to cross warm, unproductive waters near the Equator. The subcutaneous fat deposits and dense overlapping feathers that enable penguins to thrive in cold climates would likely cause death in warm climates (Van Buren & Boersma 2007), and therefore Magellanic Penguins' physiology would unlikely tolerate swimming distances from southern South America to El Salvador. Instead, it is more likely that the stranded Magellanic Penguin was transported to El Salvador by, for example, a commercial fishing ship or a sailing boat. Van Buren & Boersma (2007) reviewed several cases of human transport of penguins, and it is therefore possible that the Magellanic Penguin found stranded on the Pacific shores of El Salvador could have been transported over long distances by a fishing vessel or a sailing boat before being released close to Central America. Moreover, Van Buren & Boersma (2007) stated that fishermen transporting penguins may release birds from bycatch before entering international ports to avoid prosecution under health or wildlife laws.

Conservation implications. This study highlights the potential dangers of human intervention on wildlife. The transportation of wildlife outside of its natural range may have fatal consequences —as was the case for the stranded Magellanic Penguin. Previous reports suggest that transportation of penguins through human intervention may not be rare, as many cases may be undetected (Crawford et al. 2017). *Spheniscus* penguins constitute the most tropical penguin genus, containing the most northerly distributed penguin species across the southern hemisphere, and are consequently more exposed to human activities; therefore, they might be more likely to be affected by fisheries, oil tankers, accidental capture, global change, and other anthropogenic factors than their relatives in the Southern Ocean.

On the southern coasts of Argentina and Chile, the accidental capture of Magellanic Penguins by fishing vessels is a large conservation concern. For example, Crawford et al. (2017) reviewed the ample record of Magellanic Penguins appearing as bycatch in fisheries, and subsequently described the species as being at high risk of substantial impacts on local populations of the species. Cardoso et al. (2011) examined the impact of accidental capture of Magellanic Penguins by fishing nets in the Atlantic Ocean, which is home to populations of this species that are known to migrate thousands of kilometers north to southern Brazil to feed and overwinter (Stokes et al. 2014). During a single, 16-day fishing trip in 2011, 56 Magellanic Penguins were captured and killed by drift nets, and 12 additional Magellanic Penguins were killed by bottom gillnets (Cardoso et al. 2011).

Additionally, Magellanic Penguins in the Pacific Ocean migrate up to 800 km north to overwintering grounds in the Gulf of Arauco, Chile, which hosts a large concentration of industrial fisheries because it is one of the most productive

areas along the Chilean coast (Skewgar et al. 2014). Although there are few specific reports of Magellanic Penguin bycatch from fisheries off Chile (Crawford et al. 2017), Schlatter et al. (2009) reported entanglement in gillnets as cause of injuries and death in the mass stranding of 1380 Magellanic Penguins in southern Chile. Future studies should explore the dynamics between *Spheniscus* penguin migration and potential capture by fisheries, especially during overwintering migrations and El Niño years, when banded penguins travel farther in search of food.

ACKNOWLEDGMENTS

We would like to thank Manuel R. Paredes Oyarzún and Hernán Rivera Meléndez (CONAF, Chile) for helping with CONAF permits and logistic support during the collection of Magellanic Penguin samples at Chiloé. Tim Snyder, Curator of Birds at the Chicago Zoological Society's Brookfield Zoo, coordinated the sampling and transferring of Humboldt Penguin samples from the Brookfield Zoo. Prof. Dee Boersma, from the University of Washington at Seattle, coordinated collection of samples from Galapagos Penguins and Magellanic Penguins from Punta Arenas. We would also like to thank Eunice Echeverria, from the Natural History Museum of El Salvador and Fundación Zoológica de El Salvador (FUNZEL), who coordinated the collection of samples from the stranded penguin specimen, which is currently deposited at the Museo de Historia Natural de El Salvador. Additionally, we thank Susan J. Chivers (NMFS/NOAA-La Jolla) for assistance to CRH during the initial phases of molecular identification and to Oliver Komar for his support on initial morphological identification. Finally, we would like to thank the Editor and two anonymous reviewers who provided helpful comments on the manuscript. This study was part of Gillian M. Ross' thesis submitted to Bowling Green State University in partial fulfillment of the requirements for the degree of Master of Science.

REFERENCES

Akst, EP, D Boersma & RC Fleischer (2002) A comparison of genetic diversity between the Galápagos Penguin and the Magellanic Penguin. *Conservation Genetics* 3: 375–383.

Arauco-Shapiro, G, KI Schumacher, D Boersma & JL Bouzat (2020) The role of demographic history and selection in shaping genetic diversity of the Galápagos Penguin (*Spheniscus mendiculus*). *PLoS One*: e0226439.

Baker, AJ, SL Pereira, OP Haddrath & K-A Edge (2006) Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proceedings of the Royal Society B: Biological Sciences* 273: 11–17.

Benacerraf, B (1981) Role of MHC gene products in immune regulation. *Science* 212: 1229–1238.

BirdLife International and Handbook of the Birds of the World (2016) *The IUCN Red List of Threatened Species. Version 2021-1.* Available at https://www.iucnredlist.org/. [Accessed 03 May 2021]

Boersma, PD (1998) Population trends of the Galápagos Penguin: impacts of El Niño and La Niña. *The Condor* 100: 245–253.

Boersma, PD, A Steinfurth, G Merlen, G Jiménez-Uzcátegui, FH Vargas & PG Parker (2013a) Galápagos Penguin. Pp. 285–302 *in* García Borboroglu, PG & PD Boersma (eds). *Penguins: natural history and conservation*. University of Washington Press, Seattle, Washington, USA.

Boersma, PD, E Frere, O Kane, LM Pozzi, K Pütz, A Raya Rey, GA Rebstock, A Simeone, J Smith, A Van Buren, P Yorio & P García Borboroglu (2013b) Magellanic Penguin. Pp. 233–363 *in* García Borboroglu, PG & PD Boersma (eds). *Penguins: natural history and conservation*. University of Washington Press, Seattle, Washington, USA.

Bollmer, JL, FH Vargas & PG Parker (2007) Low MHC variation in the endangered Galápagos Penguin (*Spheniscus mendiculus*). *Immunogenetics* 59: 593–602.

Bouzat, JL, AC Lyons, GJ Knafler & PD Boersma (2013) Environmental determinants of genetic structure in Magellanic Penguin breeding colonies of the Atlantic and Pacific Oceans. *VIII International penguin Conference*. September 2–6, 2013. Bristol, UK.

Bouzat, JL, BG Walker, & PD Boersma (2009) Regional genetic structure in the Magellanic Penguin (*Spheniscus magellanicus*) suggests metapopulation dynamics. *The Auk* 126: 326–334.

Brito, JL (1999) Segundo registro para Chile del pingüino azul *Eudyptula minor* (Spheniscidae) en la costa de Santo Domingo. *Boletín Chileno de Ornithología* 6: 45–46.

Cardoso, LG, L Bugoni, PL Mancini & M Haimovici (2011) Gillnet fisheries as a major mortality factor of Magellanic Penguins in wintering areas. *Marine Pollution Bulletin* 62: 840–844.

Crawford, R, U Ellenberg, E Frere, C Hagen, K Baird, P Brewin, S Crofts, et al. (2017) Tangled or drowned: a global review of penguin bycatch in fisheries. *Endangered Species Research* 34: 373–396.

Culik, B, J Hennicke & T Martin (2000) Humboldt Penguins outmaneuvering El Niño. *Journal of Experimental Biology* 203: 2311–2322.

Dawson, AD, AD Ball, LG Spurgin, D Martín-Gálvez, IRK Stewart, GJ Horsburgh, J Potter, et al (2013) High-utility conserved avian microsatellite markers enable parentage and population studies across a wide range of species. *BMC Genomics* 14: 176–199.

De la Puente, S, A Bussalleu, M Cardeña, A Valdés-Velásquez, P Majluf & A Simeone (2013) Humboldt Penguin. Pp. 265–283 *in* García Borboroglu, PG & PD Boersma (eds). *Penguins: natural history and conservation*. University of Washington Press, Seattle, Washington, USA.

Dfaulder (2010) Magellanic Penguin in Peninsula Valdes, Argentina. (Photo) Wikimedia Commons. Available at https://commons.wikimedia.org/w/index.php?curid=34650422 [Accessed 13 May 2021] .

Doherty, PC & RM Zinkernagel (1975) Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature* 256: 50–52.

Dori (2008) A Humboldt Penguin (*Spheniscus humboldti*), taken at the Santa Barbara Zoo, California. Edited by Richard Bartz. (Photo) Wikimedia Commons. Available at https://commons.wikimedia.org/w/index.php?curid=3810182 [Accessed 13 May 2021].

Earl, DA & BM vonHoldt (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.

Ellegren, H (2004) Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics* 5: 435–445.

Evanno, G, S Regnaut & J Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.

García Borboroglu, P & PD Boersma (eds) (2013) *Penguins: natural history and conservation*. University of Washington Press, Seattle, Washington, USA.

Gavryushkina, A, TA Heath, DT Ksepka, T Stadler, D Welch & AJ Drummond (2017) Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology* 66: 57–73.

Hays, C (1986) Effects of the 1982-83 El Niño on Humboldt Penguin colonies in Perú. *Biological Conservation* 36: 169–180.

Hebert, PDN, A Cywinska, SL Ball & JR deWaard (2003a) Biological identifications through DNA barcodes. *Proceedings of the Royal*

Society B: Biological Sciences 270: 313–321.

Hebert, PDN, S Ratnasingham & JR de Waard (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences* 270: S96–S99.

Hebert, PDN, MY Stoeckle, TS Zemlak & CM Francis (2004) Identification of birds through DNA barcodes. *PLoS Biology* 2: e312.

Hibbets, EM, KI Schumacher, HB Scheppler, PD Boersma & JL Bouzat (2020) Genetic evidence of hybridization between Magellanic (*Spheniscus magellanicus*) and Humboldt (*Spheniscus humboldti*) penguins in the wild. *Genetica* 148: 215–228.

Jan, CMI, K Frith, AM Glover, RK Butlin, CD Scott, F Greenaway, M Ruedi, AC Frantz, DA Dawson & JD Altringham (2010) *Myotis alcathoe* confirmed in the UK from mitochondrial and microsatellite DNA. *Acta Chiropterologica* 12: 471–483.

Jones, HL & O Komar (2007) Central America. *North American Birds* 61: 648–651.

Kearse, M, R Moir, A Wilson, S Stones-Havas, M Cheung, S Sturrock, S Buxton, et al. (2012) Geneious basic, an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.

Kerr, KCR, MY Stoeckle, CJ Dove, LA Weigt, CM Francis & PDN Hebert (2007) Comprehensive DNA barcode coverage of North American birds. *Molecular Ecology Notes* 7: 535–543.

Kikkawa, EF, TT Tsuda, TK Naruse, D Sumiyama, M Fukuda, M Kurita, K Murata, et al. (2005) Analysis of the sequence variations in the *Mhc DRB1*-like gene of the endangered Humboldt Penguin (*Spheniscus humboldti*). *Immunogenetics* 57: 99–107.

Kikkawa, EF, TT Tsuda, TK Naruse, D Sumiyama, M Fukuda, M Kurita, RP Wilson, et al. (2009) Trans-species polymorphism of the *Mhc* class II *DRB*-like gene in banded penguins (genus *Spheniscus*). *Immunogenetics* 61: 341–352.

Knaffler, GJ, JA Clark, PD Boersma & JL Bouzat (2012) MHC diversity and mate choice in the Magellanic Penguin, *Spheniscus magellanicus*. *Journal of Heredity* 103: 759–768.

Mike's Birds (2019) Galapagos penguin (Photo). Wikimedia Commons. Available at https://commons.wikimedia.org/wiki/File:Galapagos_penguin_(47725915092).jpg [Accessed 14 July 2021].

Nims, BD, FH Vargas, J Merkel & PG Parker (2008) Low genetic diversity and lack of population structure in the endangered Galápagos Penguin (*Spheniscus mendiculus*). *Conservation Genetics* 9: 1413–1420.

Piertney, SB & MK Oliver (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity* 96: 7–21.

Pietrak, P (2019) *Spheniscus humboldti* in Prague Zoo (1) (Photo). Wikimedia Commons. Available at https://commons.wikimedia.org/wiki/File:Spheniscus_humboldti_in_Prague_Zoo_(1).jpg [Accessed 14 July 2021]

Pritchard, JK, M Stephens & P Donnelly (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Ramos, B, D Gonzalez-Acuna, DE Loyola, WE Johnson, PG Parker, M Massaro, GPM Dantas, MD Miranda & JA Vianna (2018) Landscape genomics: natural selection drives the evolution of mitogenome in penguins. *BMC Genomics* 19: 53–71.

Sallaberry-Pincheira, N, D González-Acuña, P Padilla, GPM Dantas, G Luna-Jorquera, E Frere, A Valdés-Velásquez & JA Vianna (2016) Contrasting patterns of selection between MHC I and II across populations of Humboldt and Magellanic Penguins. *Ecology and Evolution* 6: 7498–7510.

Sambrook, J, EF Fritschi & T Maniatis (1989) *Molecular cloning: a laboratory manual* (2nd ed). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Schlatter, RP, E Paredes, J Ulloa, J Harris, A Romero, J Vasquez, A Lizama, C Hernández & A Simeone. (2009) Mortandad de pingüino de Magallanes (*Spheniscus magellanicus*) en Queule, Región de la Araucanía, Chile. *Boletín Chileno de Ornithología* 15: 78–86.

Schlötterer, C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365–371.

Scordino, JJ & AM Akmajian (2012) No it's not a murre, it's a penguin. *Northwestern Naturalist* 93: 232–235.

Selkoe, KA & RJ Toonen (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellites markers. *Ecology Letters* 9: 615–629.

Seutin, G, BN White & PT Boag (1991) Preservation of avian blood and tissue samples for DNA analysis. *Canadian Journal of Zoology* 69: 82–90.

Sherony, D (2014) Magellanic Penguin (*Spheniscus magellanicus*) (15953854011) (Photo). Wikimedia Commons. Available at https://commons.wikimedia.org/wiki/File:Magellanic_penguin_(Spheniscus_magellanicus)_(15953854011).jpg [Accessed 14 July 2021].

Simeone, A, L Hiriart-Bertrand, R Reyes-Arriagada, M Halpern, J Dubach, R Wallace, K Pütz & B Lüthi (2009) Heterospecific pairing and hybridization between wild Humboldt and Magellanic Penguins in southern Chile. *The Condor* 111: 544–550.

Skewgar, E, PD Boersma & A Simeone (2014) Winter migration of Magellanic Penguins (*Spheniscus magellanicus*) along the southeastern Pacific. *Waterbirds* 37: 203–209.

Snell, GD (1981) Studies in histocompatibility. *Science* 213: 172–178.

Stokes, DL, PD Boersma, JL de Casenave & P García-Borboroglu (2014) Conservation of migratory Magellanic Penguins requires marine zoning. *Biological Conservation* 170: 151–161.

Sullivan, BL, CL Wood, MJ Iliff, RE Bonney, D Fink & S Kelling (2009) eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* 142: 2282–2292.

Szekely, P (2018) Galapagos penguin (Photo) Wikimedia Commons. Available at https://commons.wikimedia.org/w/index.php?curid=82961043 [Accessed 13 May 2021].

Tavares, ES & AJ Baker (2008) Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evolutionary Biology* 8: 81.

Toth, G, Z Gaspari & J Jurka (2000) Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* 10: 967–981.

Ujvari, B & K Belov (2011) Major Histocompatibility Complex (MHC) markers in conservation biology. *International Journal of Molecular Sciences* 12: 5168–5186.

Valle, CA (1986) Status of the Galápagos Penguin and flightless cormorant populations in 1985. *Noticias de Galápagos* 43:16–17.

Valverde, V & H Oyarzo (1996) Registros de *Eudyptula minor* (Spheniscidae) en la costa de la Región de Atacama, Chile. *Boletín Chileno de Ornitología* 3: 42–43.

Van Buren, AN & PD Boersma (2007) Humboldt Penguins (*Spheniscus humboldti*) in the Northern Hemisphere. *Wilson Journal of Ornithology* 119: 284–288.

Vargas, H, C Lougheed & H Snell (2005) Population size and trends of the Galápagos Penguin *Spheniscus mendiculus*. *Ibis* 147: 367–374.

Vianna, JA, M Cortes, B Ramos, D Sallaberry-Pincheira, D Gonzalez-Acuna, GPM Dantas, J Morgante, A Simeone & G Luna-Jorquera (2014) Changes in abundance and distribution of Humboldt Penguin *Spheniscus humboldti*. *Marine Ornithology* 42: 153–159.

Wilson, RP, A Simeone & P McGill (2000) Nota complementaria a la observacíon de un pingüino azul *Eudyptula minor* en la costa de Santo Domingo. *Boletín Chileno de Ornitología* 7: 30.