DIFFERENTIATION OF THE TUNDRA (CYGNUS COLUMBIANUS COLUMBIANUS) AND TRUMPETER (CYGNUS BUCCINATOR) SWANS AND THEIR HYBRIDS USING MICROSATELLITE REGIONS

by

Lauren Wilson

A Thesis
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Master of Science
Environmental Science and Policy

Committee:

Dr. Patrick M. Gillevet, Thesis Director

Dr. Larry Rockwood, Committee Member

Dr. Ancha Baranova, Committee Member

Dr. Albert Torzilli, Graduate Program Director

Dr. Robert Jonas, Department Chairperson

Dr. Timothy Born, Assistant Dean, Student and Academic Affairs, College of Science

Dr. Peggy Agouris, Interim Dean, College of Science

Date: July 26, 2013

Summer Semester 2013
George Mason University
Fairfax, VA

Differentiation of the Tundra (Cygnus columbianus columbianus) and Trumpeter (Cygnus buccinator) Swans and Their Hybrids Using Microsatellite Regions

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

by

Lauren Wilson
Bachelor of Science
University of Georgia, 2008

Director: Patrick M. Gillevet, Professor
Department of Environmental Science and Policy

Summer Semester 2013
George Mason University
Fairfax, VA

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Pat Gillevet, for helping me through the ups and downs of The Swan Project, and for taking me as a student in the first place. Great thanks to Masi Sikaroodi, for patience and uncountable hours in the lab, and for letting me do more PCRs at one time and fewer gels than she would prefer. Thank you to my committee members, Drs. Rockwood and Baranova, for their input on my proposal and thesis. Much appreciation to Dr. Tom Wood, Director of Environmental Studies on the Piedmont, Dr. Bill Sladen of the Airlie Center, and Jenny Bryant of the U.S. Fish and Wildlife Service in Alaska, for their support of our project. Thank you to the National Science Foundation GK-12 Sunrise Fellowship for supporting my life and studies. Finally, thank you to my roommates, who have been occasional editors, collaborators, and listeners, and constant encouragers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

Swan Research Project ..... SRP
National Wildlife Refuge ..... NWR
Pacific Coast Population ..... PCP
Rocky Mountain Population ..... RMP
Eastern Population ..... EP
Western Population ..... WP
U.S. Fish and Wildlife Service ..... FWS
International Union for the Conservation of Nature ..... IUCN
Deoxyribonucleic Acid ..... DNA
Polymerase Chain Reaction ..... PCR
Chromodomain-helicase DNA ..... CHD

ABSTRACT

DIFFERENTIATION OF THE TUNDRA (CYGNUS COLUMBIANUS COLUMBIANUS) AND TRUMPETER (CYGNUS BUCCINATOR) SWANS AND THEIR HYBRIDS USING MICROSATELLITE REGIONS

Lauren Wilson, M.S.

George Mason University, 2013

Thesis Director: Dr. Patrick M. Gillevet

This thesis describes a molecular method of differentiating two closely related swan species (Trumpeter and Tundra) and their hybrids. The Trumpeter and Tundra Swans are migratory waterfowl which breed in areas of Alaska during summer. They are known to be completely reproductively compatible in captivity, but have been historically allopatric during breeding season due to differing habitat preferences. Changing temperatures have affected the sub-arctic vegetative composition, and the breeding ranges of the two birds now overlap in some areas. The need for identifying these species and their hybrids exists because there is evidence that hybridization is occurring in the wild due to changes in vegetation in the breeding habitat of these species.

We used next-generation sequencing technology to identify and describe seven new polymorphic microsatellite loci. In combination with two previously described markers, these new markers allow differentiation of the Trumpeter Swan, Tundra Swan,

and their captive hybrids. Estimates of differentiation, particularly D, were high and indicate significant divergence between these loci. We then tested this method on unknown wild samples to detect any evidence of genetic introgression from interbreeding. However, genotypes of these individuals adhered to those of either species, not hybrids, and should not be considered of hybrid ancestry.

INTRODUCTION

Overview

The Tundra Swan (formerly the Whistling Swan, *Cygnus columbianus columbianus*) and Trumpeter Swan (*C. buccinator*) are two migratory species with significant breeding populations occupying areas of Alaska. The Bewick's (*C. columbianus bewickii*) and Whooper (*C. cygnus*) Swan are morphologically very similar species which spend winter in areas of Russia and summer in Europe. Some doubt lingers over the appropriate taxonomic status of these birds, with three species being recognized currently: *C. buccinator, C. cygnus,* and *C. columbianus.*

Historically, breeding ranges of the North American species have remained distinct, with the Trumpeter nesting in the boreal forest (taiga) habitat, and the Tundra nesting in tundra habitats. However, there is a recently-developed geographic overlap of breeding areas. Climate change has caused changes in plant ranges in many areas of Alaska, including the expansion of the boreal forest into tundra (Beck *et al.* 2001; Soja *et al.* 2006). Trumpeter Swans have followed this expansion and now may be found in some areas occupied by Tundra Swans during breeding season. Though interspecies breeding is exceptionally common in Family Anatidae, for example between species of genus *Anas,* the production of fertile offspring is less common (Johnsgard 1960). The Trumpeter and Tundra Swans are known to be reproductively compatible, as a population of hybrids is maintained at the Airlie Center's Swan Research Program (SRP), Virginia. Furthermore,

several swan carcasses, exhibiting morphology similar to that observed in hybrids at the Airlie Center, were collected in the wintering grounds in Washington State. This discovery suggests the possibility of hybridization in areas of geographic overlap during breeding season in the wild.

Objectives

This study has two objectives. Its first aim is to develop a method to distinguish the two swan species and their hybrids using genetic markers. Its second aim is to use this novel technique to test for genetic introgression in wild swans taken from the Washington State and the Koyukuk National Wildlife Refuge (NWR) of Alaska, where geographic overlap of Trumpeter and Tundra populations has been reported. I also compared a number of specimens collected from the wintering grounds in Washington state that appear to be hybrids based on bill morphology. Natural interbreeding between these species offers opportunities for the study of the evolutionary process of speciation and the validity of the biological species concept. The successful reproduction between the species suggests extremely close relationships among these "species," and hybridization in the wild challenges the taxonomic validity of these four northern swan populations. Additionally, the breakdown of all reproductive barriers due to a significant biome shift illustrates the evolutionary and ecological implications of climate change.

The overall purposes of this study are to start to develop an understanding of the evolutionary relationships among the four Northern Swan species. The specific goals of this study are to:

1. Isolate microsatellite primers using NextGen sequencing technology and validate loci for distinguishing the two North American swan species and their hybrids.
2. Validate the loci on hybrids from the Swan Research Program at Clifton Farm in Warrenton, VA.
3. Test wild samples from Washington and Alaska for evidence of genetic introgression.

Hypothesis

Based on the above objectives, I propose the following two hypotheses:

H1: Selected microsatellite loci will differentiate Tundra and Trumpeter Swans and known hybrids.

H0: Selected microsatellite loci will not differentiate Tundra and Trumpeter Swans and known hybrids.

H2: Wild swans which exhibit hybrid morphology will exhibit allele ranges and frequencies consistent with the captive hybrid population.

H0: Wild swans which exhibit hybrid morphology will not exhibit allele ranges and frequencies consistent with the captive hybrid population.

BACKGROUND

Taxonomy

The Trumpeter (*Cygnus buccinator*) and Tundra (*C. columbianus columbianus*) Swans of North America, along with the Whooper (*C. cygnus*) and Bewick's (*C. columbianus bewickii*) swans of Europe and Asia, represent a group of closely related species of migratory waterfowl known as the four northern swans. The evolutionary relationships between these species remain unclear. Species of genus *Cygnus*, along with geese and other swans, are members of the tribe Anserini, within subfamily Anserinae of the waterfowl family Anatidae (Johnsgard 1974; Travsky and Beauvais 2004).

The present taxonomic approach treats the northern swans as individual species, despite proposals to consider some or all of these swans as conspecifics due to their evident morphological and behavioral similarities (Delacour and Mayr 1945; Evans and Sladen 1980; Johnsgard 1974) as well as the more recently discovered genetic similarity (Barrett and Vyse 1982). The Tundra and Bewick's swans share a similar habitat preference, vocalization pattern, and nearly-identical external and internal morphology. They are only distinguished by slight variations in the amount of yellow on the bill (Banko 1960; Evans and Sladen 1980; Hansen *et al.* 1971; Sladen *et al.* 2002; Travsky and Beauvais 2004). Accounts indicate that Bewick's Swans interbreed and produce offspring with the Tundra in the rare occasion that a Bewick's Swan migrates to North America (Evans and Sladen 1980). Interbreeding of the Trumpeter Swan with the

Whooper, Bewick's and Tundra has also been confirmed, with fertile offspring resulting from Trumpeter-Tundra hybrids (Sladen *et al.* 2002; Travsky and Beauvais 2004). Additionally, genetic studies performed by Barrett and Vyse (1982) indicate extreme similarity between the Trumpeter, Tundra, and Trumpeter-Tundra hybrids. In fact, Meng and Parkin (1993) determined a high level of similarity between the trumpeter, whooper, and Bewick's swans via DNA fingerprinting.

Proposed taxonomic schemes differ. Delacour and Mayr (1945) suggested joining the Tundra with Bewick's swan as *C. bewickii*, and the Trumpeter with the Whooper swan as *C. buccinator*. They assert that the nearly-identical tundra and Bewick's swans are New and Old World counterparts; the same for the Trumpeter and Whooper swans. Similarly, Palmer (1976) viewed the tundra and Bewick's swans as conspecifics, suggesting the name "Tundra Swan" for both. Johnsgard (1974) proposed the most radical change of lumping together all four swans into one species (*C. cygnus*) and treating each as a subspecies.

Description

All members of tribe Anserini share the traits of large size, elongated necks, downy plumage, and webbed feet. They show no sexual dimorphism, and lack the iridescent speculum common in other waterfowl (Delacour and Mayr 1945; Johnsgard 2010). The northern swans exhibit entirely white adult plumage, which is a morphology observed in all swan species except one, Australia's black swan (*C. atratus*). However, bill coloration distinguishes the four species, and shows a general geographic pattern of more yellow on the bills of the Eurasian species and less or none on that of the North

American species. The Tundra Swan has a black bill with a small yellow or orange-yellow spot in front of the eye, though a completely black bill is occasionally observed. The Trumpeter's bill is entirely black, though rarely a grey or dull yellow spot is visible immediately behind the nostril (Banko 1960; Evans and Sladen 1980). The bill of Bewick's Swan is nearly one-third yellow (Evans and Sladen 1980), and the Whooper has even more yellow (up to half). Therefore, a gradient across the continents exists for bill coloration from least to most yellow; the species may be ordered Trumpeter, Tundra, Bewick's, and Whooper. Internal morphology allows easy identification; the trachea of the Trumpeter Swan is longer with an extra loop (Bank 1960; Wood *et al.* 2002).

The size of the four species exhibits a pattern that is opposite from the geographic cline seen in the bill coloration. As the largest waterfowl species in the world, the Trumpeter is the largest of the four. The Whooper is slightly smaller, followed by the Tundra and Bewick's in size (Banko 1960).

The voices of the Trumpeter and Tundra Swans are easily distinguishable. The call of the Trumpeter Swan is a melodious yet nasal honk, while the Tundra Swan has a short, higher-pitched, "barking" honk. The loud, robust call of the Trumpeter Swan is likely due to the elongation and loop of the trachea, which the Tundra Swan lacks (Banko 1960). The call of the hybrid, however, is described as dull and raspy – entirely distinct from either species (Wood *et al.* 2002). Vocal characteristics have been used in multiple studies for species identification (Banko 1960; Hansen *et al.* 1971).

History and Distribution

Trumpeter Swan

The pre-historic range of the Trumpeter Swan was vast and spread across most of North America. Alexander Wetmore identified Trumpeter remains from the Pleistocene Era in Oregon, Illinois, and Florida. In Illinois the bones were discovered along with those of the mastodon (*Mastodon americanus*) and giant beaver (*Castoroides ohioensis*), and were believed to be deposited in a post-glacial lake bed (Banko 1960; Wetmore 1935). In Florida the Trumpeter Swan bones were associated with the California Condor (*Gymnogyps californianus*), Whooping Crane (*Grus americana*), and Jabiru Stork (*Jabiru mycteria*), all of which are no longer indigenous to that region of the continent. Surprisingly, a Flamingo (*Phoenicopterus copei*) was found with the Trumpeter in Oregon. The ancestral species of the Trumpeter Swan reflects changes in North American geology, climate, and ecology, as well as changes in habitat preferences and adaptability of the swan species (Banko 1960).

The historic range of the Trumpeter included most of North America, including wintering grounds near the Gulf of Mexico and east to Hudson's Bay (Banko 1960; Barrett and Vyse 1982). Accounts from settlers suggest the presence of Trumpeter Swans in Massachusetts, Connecticut, and New Hampshire. The notes of John Lawson, Surveyor-General of North Carolina, clearly describe Trumpeters not only wintering, but also breeding in lakes along the east coast (Banko 1960).

However, as with many other birds and mammals, hunting during the 19th century drastically depleted populations across the continent (Banko 1960; Barrett and Vyse 1982; Hansen *et al.* 1971; Johnsgard 1978; Shea *et al.* 2002; Travsky and Beauvais

2004). Swan plumage was used for powder puffs and writing quills, and their eggs sought by collectors. From 1853 to 1877, nearly 18,000 swan skins imported from Hudson's Bay Company were sold in London alone. By the mid-1800s, the Trumpeter Swan was quite rare east of the Mississippi River (Johnsgard 1978). In 1918 the Migratory Bird Treaty Act forbid hunting of the Trumpeter Swan, but by this time ornithologists agreed that the species was near extinction. Some hunting continued and lead-poisoning from consumption of lost lead bullets was also a significant cause of mortality (Banko 1960; Johnsgard 1978). During the 1930s, surveys indicate fewer than 100 Trumpeter Swans survived in the wild, mostly in Yellowstone National Park, though this figure does not include Alaska, where Trumpeter Swans were confused as Tundra Swans and the harsh climate and terrain made ventures to the interior of the state rare. The establishment of Montana's Red Rock Lakes Migratory Waterfowl Refuge in 1935 is perhaps the primary reason for the return and success of the species in the continental United States. After this time, the species began to colonize other protected areas, such as the Grand Teton National Park, established 1950 (Hansen *et al.* 1971; Johnsgard 1978; Shea *et al.* 2002).

To contrast the decimation that occurred in the 19th century, the 1900s included significant increases in population sizes, most notably since the 1940s (Banko 1960; Caithamer 2001; Johnsgard 1978; McKelvey *et al.* 1983; Moser 2006; Oyler-McCance *et al.* 2007; Shea *et al.* 2002). Three primary populations are now recognized. Aerial surveys in 1959 confirmed the Pacific Coast Population (PCP), which winters in Oregon, Washington, and British Columbia, and migrates to Alaska and the Yukon Territories to breed. The Rocky Mountain Population (RMP) winters as far west as eastern Oregon and

northern Nevada, and east into Wyoming. The RMP includes a non-migratory "tri-state flock," which resides in Idaho, Montana, and Wyoming, largely in Yellowstone National Park. In the 1960s, the Interior Population was established in central and eastern North America by translocation efforts, which have continued into the 1990s (McKelvey *et al.* 1983; Oyler-McCance *et al.* 2007; Shea *et al.* 2002). By 1999, restoration efforts had been attempted in South Dakota, Minnesota, Missouri, Michigan, Wisconsin, Iowa, Ohio, and Ontario, with 1629 captive bred swans released total (Shea *et al.* 2002). Figures 1 and 2 below depicts the overall range of the Trumpeter Swan, and comparison reflects the expansion of the range from 2000 to 2005 (Caithamer 2001; Moser 2006). As a result of the growth of this species, the trumpeter is now listed as "least concern" with the International Union for the Conservation of Nature (IUCN 2013).

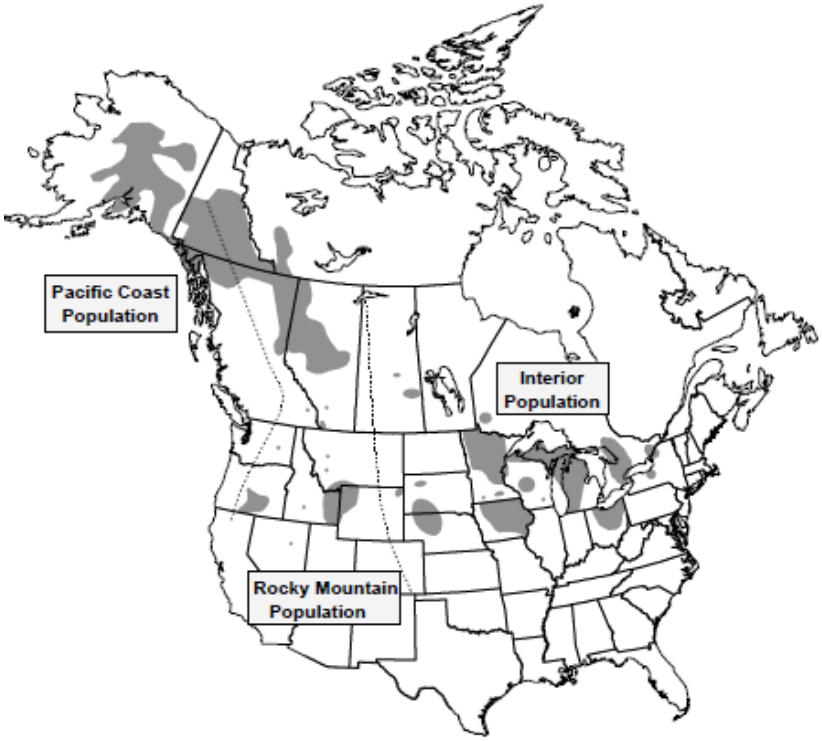


Figure 1. Distribution map of Trumpeter Swans in 2000. Caithamer 2001.

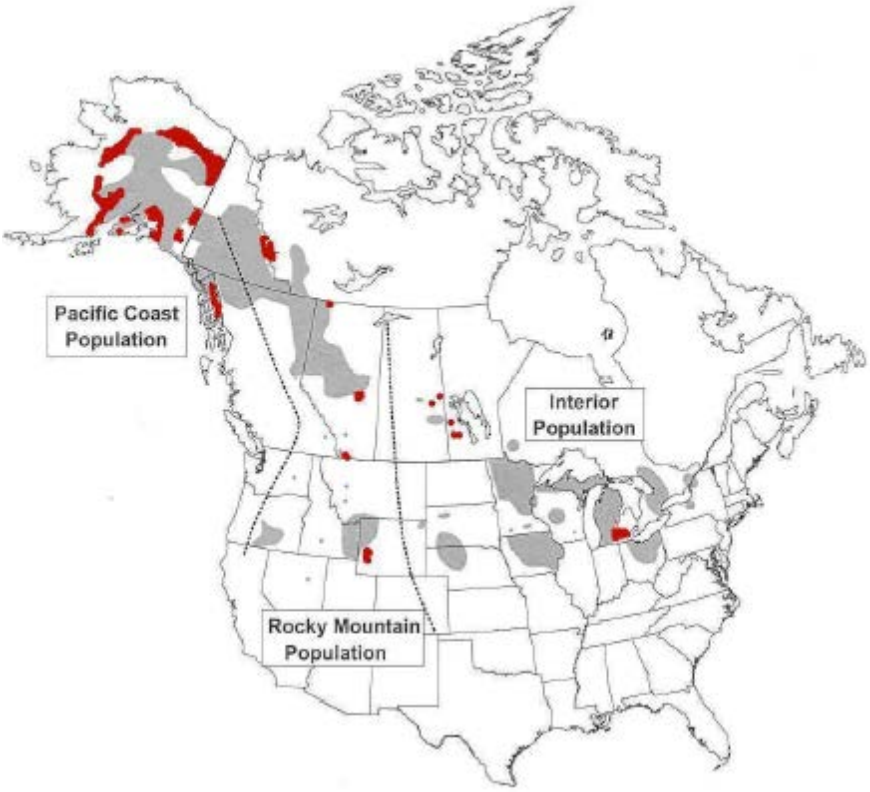


Figure 2. Distribution of Trumpeter Swans in 2005. Red areas indicate expansion as compared to distribution in 2000. Moser 2006.

A study by Oyler-McCance *et al.* (2007) indicates that the Interior Population is not genetically distinct from the other populations and should be treated with the same management strategy. In this study, 16 microsatellite loci characterized by St. John *et al.* (2006) were used, though most loci exhibited low heterozygosity. Additionally, results showed that the Trumpeter Swan exhibits lower genetic diversity in the mitochondrial DNA than other waterfowl species, a discrepancy which indicates that a bottleneck event has occurred in both the PCP and RMP.

Tundra Swan

The Tundra Swan, whose common name was changed from the whistling swan in 1982 by the American Ornithologist's Union, has two recognized populations: East Coast (EP) and West Coast (WP). Hunting significantly reduced population sizes in the 1800s, though not to the degree of the Trumpeter Swan. During the 1950s the estimated size of the WP alone doubled, and rose steadily until the late 1990s. The highest population size for the WP was recorded in 1997, with 122,521 swans, and nearly repeated in 1999. The EP, historically larger than the WP, began to increase by the 1970s, and reached 110,000 in 1992. The Migratory Bird Treaty Act combined with the increasing area of preserved land doubtless had a great impact on the increase in numbers for the Tundra Swan (FWS 2001). However, recently both populations have begun to utilize agriculture fields for foraging, especially during spring migration, raising a concern that quality stop-over habitat has become limited and the species has become dependent on this new resource (Earnst 1994; Petrie *et al*. 2002).

The Tundra Swan has a more extensive distribution than the Trumpeter Swan. During summer breeding, their range spans across Canada from the Quebec-Ontario border west to Bristol Bay, Alaska (Wilk 1988). Most of the WP swans (76%) nest near Alaska's Yukon-Kuskokwim Delta near the Bering Sea. Migration routes for the WP include the Central and Pacific Flyways, while the EP utilizes the Atlantic and Mississippi Flyways. A few swans belonging to the WP winter as far north as Alaska's southern peninsula, but the majority winter along coastal California. Both coastal and interior wintering sites exist in between, and as far east as Utah. A small group (300 – 500) winter in southern British Columbia with Trumpeter Swans of the Pacific Coast

Population (FWS 2001). Members of the EP winter along the east coast, primarily in North Carolina, Virginia, and Maryland (FWS 2001; Sladen 1973). A distribution of these populations can be seen in Figure 4.

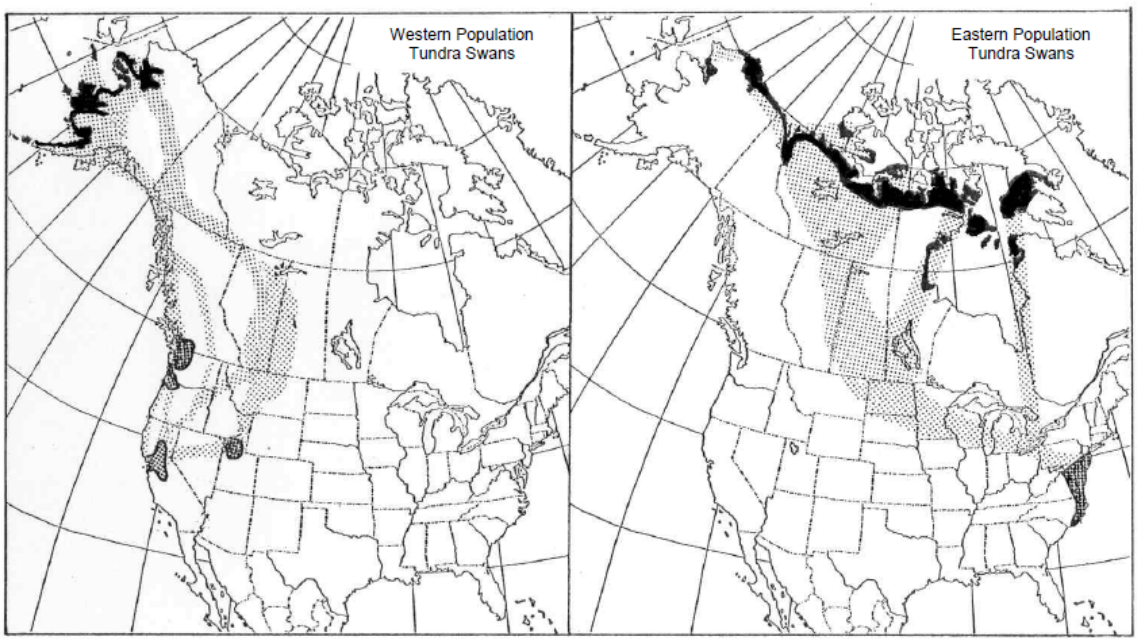


Figure 3. Distribution of Eastern and Western Populations of Tundra Swans. FWS 2001.

Habitat

Trumpeter Swan

The Trumpeter Swan has been recorded nesting in a range of habitat types, from closed boreal forest to short grass prairie. However, the consistently preferred habitat for this species is the open boreal forest type (Banko 1960). The Trumpeter Swan constructs a nest in still, shallow water (< 1 m) on top of natural mounds from small islands or

vegetation (Banko 1960; Hansen *et al.* 1971; Travsky and Beauvais 2004). Basic requirements include a calm body of water with little seasonal fluctuations in depth. Territories are rarely less than 5 acres and may reach 100 acres to ensure sufficient food resources (Johnsgard 1978). The pond or lake may be of variable depth (Hansen *et al.* 1971), though shallow margins are preferred as this increases foraging potential for aquatic plants (Travsky and Beauvais 2004). Hansen *et al.* (1971) surveyed vegetation in water bodies used for trumpeter nesting, and commonly found the following plants: *Potamogeton, Hippuris, Sparganium, Equisetum, Hippuris, Utricularia, Glyceria, Callitriche, Potentilla, Carex, Menyanthes, Calamagrostis,* and moss. The water body must be fairly open with few trees and other obstructions, as 100 m is required for takeoff and landing. Elevation is one restrictive factor; 140 to 154 ice-free days are required for a complete reproductive cycle, resulting in most breeding to occur at 500 feet elevation, and none above 2700 feet. Beaver ponds are frequently used, as the beaver lodge often maintains the water level and provides substrate for nest building (Hansen *et al.* 1970).

During migration, Trumpeter Swans fly in family groups and seek ponds, lakes, rivers, and wetlands for stopover habitat (Travsky and Beauvais 2004). During spring migration, high-quality foraging habitat is critical, because energy and nutrient intake is positively correlated with reproductive output. This is supported by behavioral studies which report that Trumpeter Swan adults spend the majority of time foraging during spring migration (LaMontagne *et al.* 2001). This theory is also supported by pond preference in migrating Trumpeter Swans; migrating groups consistently choose ponds with high rhizome and tuber biomass and few frozen days (LaMontagne *et al.* 2003).

Wintering habitat may be more limited, especially later in the season when weather is more severe. The Trumpeter Swans must find bodies of water which are least likely to freeze and thus provide the best access to aquatic vegetation. Movement of water within streams and large lakes prevents or delays freezing, and often swans are found near geothermal features such as geysers, hot springs, and thermal ponds. Such features are common within Yellowstone National Park, making this area critical for Trumpeter Swans as well as many other wintering waterfowl species. A second factor in habitat selection is the plant community surrounding the body of water. The size of the trumpeter makes take-off impossible on waters surrounded by dense timber, so any wintering habitat must be surrounded by open terrain (Banko 1960; Travsky and Beavais 2004).

Tundra Swan

Unlike the Trumpeter Swan which builds a nest in water in forested areas, the Tundra Swan builds a mound on land near a pond or lake in open tundra habitat. Wilk (1988) confirmed use of many wetland breeding habitats in Alaska below 100 m elevation, and only one pair breeding above this elevation. Most bodies of water above this were open ponds and lakes with insufficient emergent and shoreline vegetation. Most pairs nested in flat coastal areas near lakes with ample vegetation for foraging, cover, and nest building (Wilk 1988). A study in the Arctic National Wildlife Refuge found that 89% of breeding pairs nested less than 1 km from coastal lagoons (Monda *et al.* 1994). This study also found that primary habitat types included aquatic-marsh dominated by sheathed pondweed, mare's tail, and Hoppner sedge; grass-marsh dominated by pendent

grass, tundra grass, tall cotton grass, water sedge, and mare's tail; saline grass-shrub dominated by Hoppner sedge, bear sedge, tundra grass, alkali grass, and round-leaf willow; grass-shrub-water-sedge dominated by water sedge, pendent grass, tundra grass, and willows; partially vegetated with a mixed plant community of less than 15% cover (uncommon); and upland dominated by dryas, polar grass, lyme grass, and willows (uncommon).

Typically all swans forage for aquatic plants in wetland habitats, but Tundra Swans more frequently use agricultural fields for stopover habitat. This shift has occurred relatively recently with the mass conversion of wild lands for agricultural uses (Earnst 1994; Petrie 2002). Quality stopover habitat is especially important, as swans actually increase their lipid and protein reserves during migration. The lowest body mass is reached immediately before spring migration, which must be spent acquiring sufficient nutrition for summer breeding. Successful reproduction is also extremely costly with regard to energy, so autumn migration must also include high quality habitat to replenish nutritional reserves. A study in the eastern basin of Lake Erie indicated that foraging in agricultural fields is significantly more common during spring migration than autumn (Petrie 2002). Sago pondweed produces high-energy tubers which are thought to be the primary food of the Tundra Swan during both spring and autumn migration. Habitat preference during migration is strongly correlated with abundance of sago pondweed in wetlands (Earnst 1994).

It is of note that Hansen *et al*. (1971) found one confirmed record of an active Trumpeter Swan nest on the tundra alongside nesting Tundra Swans. They concluded that

Trumpeter Swans rarely enter Tundra Swan habitat to reproduce. Johnsgard (1974) also claimed that sympatry was almost entirely absent between the species at the time of publication.

EXPERIMENTAL DESIGN

Molecular Methods: Microsatellites

Microsatellites (simple tandem repeats) are DNA sequences containing a series of one to six nucleotide bases repeated. They are highly variable in length due to their high mutation rates, which are significantly higher than substitution mutation rates. This difference in rate results from a relatively high frequency of the gain or loss of a repeat during replication, a process called "DNA slippage" (Schlotterer 2000). Microsatellites are commonly used for population-level genetic studies and paternity testing, and may also be informative in identifying closely related species (Domingo-Roura 2002; Schlotterer 2000; Selkoe and Toonen 2006). Having a multi-locus dataset allows a more statistically powerful comparison of species groups and should also be helpful in identifying hybrid individuals. Additionally, the ease of methods for microsatellites are often more favorable than sequencing. The length of the target region is smaller, enabling successful amplification even with degraded DNA from non-invasive samples such as feathers or eggs. Additionally, these hypervariable loci are usually species-specific, reducing the possibility of cross-contamination. For these reasons, I have chosen to characterize some previously undescribed loci and compare all selected microsatellites to achieve our objective of genetically distinguishing among the two swan species and their hybrids.

However, use of microsatellites as genetic markers does pose some difficulties or disadvantages. For example, the flanking regions of microsatellites, which would include the primer sequence, usually differ among species. Because of that, species-specific primers need to be developed. Comparison of alleles can also be difficult and possibly misleading. While fingerprinting is much cheaper, looking at fragment length alone can result in incorrect inferences because allele sequences of the fragments may be different even if the length is the same. In addition, since the mutation mechanism may be locus-specific, statistical analysis on allele frequencies can be challenging because of the use of mutation models in some analyses (Selkoe and Toonen 2006). Lastly, in a heterozygote, the longer allele may not amplify as well as the shorter and may not be detected well during fingerprinting. This process known as allele dropout can result in the determination of incorrect alleles during the genotyping process. This would result in a homozygote designation for a heterozygote individual. However, results of allele dropout may be avoided by multiple replicates for each loci and each individual.

Sample Collection

Trumpeter and Tundra Swan samples of blood, tissue, feather and egg were collected in four areas of Alaska (Koyukuk, Kaiyuh Flats, Selawik, North Slope) and from swans residing at the Airlie Center in Northern Virginia. For a map of collection sites, see Figure 1. The samples from swans in the Swan Research Program at the Airlie Center are especially useful, as these include known hybrids. Hybrid sequences may be compared to those from wild birds. DNA has already been extracted from some samples

using the Quiagen DNeasy tissue kit, following the specified protocols. Future extractions will be performed in the same way.

Figure 4. Map of Alaska depicting collection sites for wild samples.

Preliminary Data

Previous work focused on comparing sequenced intron markers, for example the nuclear CHD gene Intron A and the mitochondrial 12S and D-loop (control region). The Gillevet lab has performed sequencing for these markers for both species. All of these segments have been used in previous avian phylogenetic or population genetics studies (Donne-Gousse *et al.* 2002; Kulikova *et al.* 2005; Peters *et al.* 2005; Poyarkova *et al.* 2010; Ruokonen *et al.* 2000; Ruokonen *et al.* 2004; Ruokonen *et al.* 2005; Zink and Barrowclough 2008). However, neighbor joining trees created in PAUP for each locus

provided little or no resolution between the species, which suggests a high degree of genetic similarity. The similarity of the control region in mitochondrial DNA samples of avian species is supported by the Ruokonen (2002) study that found that the control region is genus-specific and in general more variable than cytochrome b. Work performed by Elizabeth Dingess at the Microbiome Analysis Center at George Mason University developed the use of the CHD intron as a method of deciphering species and gender of the Trumpeter and Tundra Swans. Substantially different lengths for Intron A in the Z and W sex chromosomes allowed gender identification. Phylogenetic trees created with CHD sequence data did not provide a consistent branching pattern separating species. Different sequence lengths were observed between the species, but only on the female-specific W chromosome.

The limited variation exhibited in the sequence markers necessitated the use of microsatellites. St John *et al*. (2006) described 16 new microsatellite loci, which were then used in a rangewide population genetic study of the Trumpeter Swan by Oyler-McCance *et al.* in 2007. We selected eleven primer sets for the most allele-heterogeneous loci. In our small sample set, one primer set did not amplify loci in both species, five amplified very poorly, and nine did show sufficient variation between species. Therefore, these loci were not sufficiently informative to provide a method of distinguishing the species and identifying their hybrids. For this reason, we chose to use next-generation sequencing technology to quickly and inexpensively identify, describe, and use new polymorphic microsatellites in the Trumpeter and Tundra Swans. Appendix 6 shows the alleles for these microsatellite loci.

CHARACTERIZATION OF MICROSATELLITE LOCI BY GENOMIC SEQUENCING

Abstract

Microsatellites have become commonly used as markers in studies of population-level genetics studies. Although traditional methods of microsatellite discovery and primer design are costly, here we use genomic "shotgun" sequencing methods to produce reads which were then filtered for repeat sequences under specific sequence criteria using the bioinformatics tool Msatcommander. We then implemented a series of steps to cull undesirable sequences and designed primers for target sequences. After testing these potential loci, we discovered eight polymorphic microsatellite sequences that may be used in future genetics studies for these species. This process has already been used for some non-model organisms, and it represents a much faster and cost-effective method of primer design for species with limited genetic data available.

Introduction

Microsatellites regions, or simple tandem repeats, are widely applicable for genetics studies, particularly at the individual and population levels. These regions have become commonly used in the past 15 years, and are highly useful in questions of gene flow, disease transmission, population *structure*, relatedness and evolutionary history, and other applications, many with important wildlife management implications (Ferreira

da Silva 2012; Schlotterer 2000). Their high level of variability makes them informative for closely related groups, but it also necessitates the design of species-specific primers. This process of designing primers for organisms with limited genetic knowledge was often cost-restrictive with traditional methods, which involved the creation of libraries of repeat-rich sequences using base-specific probes (Abdelkrim *et al.* 2009). St John *et al.* (2006) used such methods, but in my initial tests of these loci as markes, I found insufficient variation to distinguish the species and their hyrbids (see Appendix 6). Therefore, new loci were needed for our study.

High-throughput sequencing technology represents a superior alternative because it produces a sufficiently large number of short sequence reads to have a statistically high chance of containing thousands of microsatellite regions in one run. However, most of these microsatellites will be undesirable as markers, particularly with regard to primer sequence and location relative to the repeat region. Use of a bioinformatics tool to screen for desirable sequences (perfect repeats of specific lengths) and design primers provides a rapid process of preparing a large but manageable set of potential microsatellite markers (Abdelkrim *et al.* 2009). This process of using high-throughput sequencing technology to assemble genomic libraries from which to sample for microsatellite regions is faster and more cost-effective than traditional methods, and has already been employed for some non-model organisms (Abdelkrim *et al.* 2009; Fatemi *et al.* 2013; Saarinen and Austin 2010).

This process was used to discover seven polymorphic microsatellites found in the two native North American swan species, the Trumpeter (*Cygnus buccinator*) and Tundra

(*C. columbianus columbianus*). Both species were seriously impacted by overharvesting before the Migratory Bird Treaty Act of 1918, and the Trumpeter Swan was nearly driven to extinction (Banko 1960; Johnsgard 1978). Though plumage is identical, the Trumpeter is larger than the Tundra and bears an entirely black bill, contrasted to the Tundra Swan's black bill with yellow lores. The species are closely related to the Whooper (*C. cygnus*) and Bewick's (*C. columbianus bewickii*) Swans of Europe and Asia, and taxonomy has been disputed for these four "northern swans" (Banko 1960; Evans and Sladen 1980).

The Trumpeter and Tundra Swans have been chosen for study for two reasons. First, there has been no comprehensive molecular study on the four northern swans or the North American swans. Second, evidence of interbreeding exists. Their breeding ranges have recently begun to overlap in some areas of Alaska (Jenny Bryant, personal communication), a possible result of climate change (Beck *et al.* 2001; Soja *et al.* 2006), and swan carcasses exhibiting hybrid morphology have been found on their Washington wintering grounds. Because the Trumpeter and Tundra are known to be reproductively compatible (Johnsgard 1974), we seek to develop a molecular method of species distinction, in order that this may be applied to wild populations which may exhibit genetic introgression from interbreeding. St. John *et al.* (2006) described 16 microsatellite loci, and primers for 11 markers with the highest heterozygosity were chosen as potential markers. However, we found insufficient variation among these markers to distinguish the two species. Here we describe our methods of discovering new microsatellite regions that will be part of a molecular method of distinguishing the two swan species and their hybrids.

Methods and Results

We have received blood samples from swans at the Airlie Center and from areas of Alaska, taken by the U.S. Fish and Wildlife Service in the Koyukuk National Wildlife Refuge (both species) and the North Slope (Tundra Swan). Two samples, one from each species, were digested with Proteinase K. Methods for DNA extraction from blood samples followed the standard protocol for the Quiagen DNeasy Tissue kit. Extractions were approximately quantified by running 4 uL of product through a 1% agarose gel and visualized on a Kodak EDAS 120. Library preparation followed the protocol described in the Roche Rapid Library Preparation Method Manual (Life Science Corp. 2010), and shotgun sequencing was then performed using a Roche 454 GS-Junior (Basel, Switzerland) at the Microbiome Analysis Center in the Department of Environmental Sciences and Policy, George Mason University. Reads were aligned to concatenate sequences and viewed and analyzed in Applied Biosystems Genemapper v4.1 (Foster City, California).

A total of 91,849 sequence reads in FASTA format were generated for the Trumpeter Swan, and 103,942 reads for the Tundra Swan. We used the program Msatcommander (Faircloth 2008; Rozen and Skaletsky 2000) to find repeat sequences under particular criteria; acceptable reads were those without the same repetitive element elsewhere in the sequences, a repeat sequence of at least 30 bases, and total target sequences less than 500 bases. The resulting 566 sequences were aligned to the chicken (*Gallus gallus*) genome in Roche GS Reference Mapper (Basel, Switzerland) to ensure

species specificity of the loci and to check for repetitive elements in the genome. Primers were aligned to read sequences in Sequencher 4.7 (Gene Codes Corp., Ann Arbor, Michigan). Primers which did not align to their intended sequence were removed. Remaining sequences were sorted by length and those less than 120 bases long were deleted. Next we examined each of the remaining 118 contiguous fragments (contigs) assembled by Sequencher by eye. We removed all sequences and primers which were deemed poor based on the sequence of the primer, microsatellite repeat, and overall sequence (some primers contained small repetitive sequences, some microsatellites were very short relative to the overall target sequence, and some microsatellites were imperfect). Twenty loci were selected as potential markers, because I wanted eight to fifteen microsatellite markers, as is common in the literature, and I anticipated that some loci would not be polymorphic. The primers designed in MSATCOMMANDER were tested on seven swan samples (three Trumpeter, four Tundar). Of those 20 primer sets, twelve worked well with both samples of both species, and FAM-labeled primers were then used for fingerprinting.

Polymerase chain reactions (20 uL) used Applied Biosystems *Taq* Gold DNA polymerase (Foster City, California). Reactions for amplification contained 10 ng DNA and Applied Biosystems (Foster City, California) 25 mM $MgCl_2$, 0.5 mM of each deoxynucleotide triphosphate, 5 M betaine or 0.1% bovine serum albumin, and Invitrogen 10 uM primer (Invitrogen, Carlsbad, California). Reactions were be performed in an MJ Research (now Bio-Rad) PTC-200. Thermal profiles will be 95˚C for 11 min, 35 or 38 cycles at 95˚C for 40 sec, 1 min 30 sec at annealing temperature (Table 1), and

72˚C for 2 min, with a final extension of 10 min at 72˚C. Annealing temperatures and number of cycles may be found in Table 1 below. To address the issue of allele dropout, we performed three replicate PCRs for each loci on each sample. See Table 1 for primer sequence and cycle information

Table 1. Primer sequence and PCR protocol

Locus Name	Annealing Temp	Number of Cycles	Primer	Primer Sequence
TR.MS6	54	35	TR.MS6 Forward	ATCCCTTGGCTGTCATCCTC
			TR.MS6 Reverse	AGCTCCATGAAGCTTGAGATC
TR.MS10	54	38	TR.MS10 Forward	TGGCACTACATCTACCCAGC
			TR.MS10 Reverse	AGTTCAAGCAGGTAGGTGGG
TR.MS11	54	35	TR.MS11 Forward	TGCTCTGTGGATATAGGCCC
			TR.MS11 Reverse	TTCGCTCATTGGAACTGCTG
TU.MS1	54	35	TU.MS1 Forward	GGAGTCTTACGTTTGGGCAC
			TU.MS1 Reverse	AGCCCTGGAGCAAGGATTG
TU.MS2	54	35	TU.MS2 Forward	CTTGGGCTCTGCATCCTCTC
			TU.MS2 Reverse	CATCCTAAGCATGCATCGGG
TU.MS5	54	35	TU.MS5 Forward	AGGATTCAGTAGGTGCCTGC
			TU.MS5 Reverse	GCCGTTCTACAGGTTTGCAG
TU.MS6	54	35	TU.MS6 Forward	GTGAACTCCAGGGACAAGTG
			TU.MS6 Reverse	ATGCAGTATCCTCCCACACC

To quantify product, four microlitres of PCR product were be run with brophenol blue and xylene cyanol dye in 40% sucrose solution in Tris-acetate buffer through a 1.5% agarose gel in a C.B.S Scientific Company PES-2000 Series II (Del Mar, California). Gels were stained in ethidium bromide solution. Products were diluted at a 1/10, 1/15, or 1/20 ratio depending on the relative concentrations of PCR product and primer viewed in the gel image. These dilutions were added to a 1/20 ratio of ILS600 size standard (Promega Corporation, Fitchburg, Wisconsin) to HiDi formamide (Applied Biosystems, Foster City, California) and an Applied Biosystems 3130xl Genetic Analyzer was used for analysis. Fingerprinting results were viewed in Genemapper v4.1.

Fingerprinting results revealed seven polymorphic loci that amplified well with both Trumpeter and Tundra samples. We amplified and fingerprinted 16 additional samples (8 Trumpeter, 8 Tundra) samples to assess heterozygosity. See Table 2 for marker details. Appendices 1 and 2 has additional information, including the original read sequence read name assigned and its correlating locus name, the repeat motif described by Msatcommander, and the full read sequence for the seven chosen markers.

Table 2. Marker information.

Locus Name	Repeat Motif	Marker Size	No. of Alleles in Trumpeter	No. of Alleles in Tundra	Total No. of Alleles
TR.MS.6	$(CTTT)_{13}$	158-174	3	4	5
TR.MS10	$(TTA)_{12}(GTA)_{8}$	317-329	2	4	5
TR.MS11	$(GTTT)_{8}$	322-330	2	3	3
TU.MS1	$(GGTTT)_{6}$	338-360	4	2	5
TU.MS2	$(AAGG)_{13}$	246-270	6	4	7
TU.MS5	$(AAAAC)_{6}$	202-220	3	4	7
TU.MS6	$(ATTTT)_{23}$	310-326	2	5	6

Conclusion

We investigated the use of previously-described microsatellite loci as markers, but found they were not able to decipher species. We proceeded to use next-generation sequencing technology to discover and describe seven new microsatellite regions in the Trumpeter and Tundra Swans. The allelic richness (number of alleles per locus) for loci described by St. John *et al.* (2006) range from two to seven, with an average of 3.0625. In constrast, the allelic richness of loci described here range from three to seven, with an average of 5.4286. Because these loci exhibit high levels of allelic richness in both species, they represent valuable tools in future population or species-level genetics studies. We plan to incorporate these loci into a molecular method of distinguishing the two species and their hybrids, possibly in combination with previously described loci.

DIFFERENTIATION OF THE TRUMPETER (*CYGNUS BUCCINATOR*) AND TUNDRA (*C. COLUMBIANUS COLUMBIANUS*) SWANS AND THEIR HYBRIDS USING MICROSATELLITE MARKERS

Abstract

The Trumpeter (*Cygnus buccinator*) and Tundra (*C. columbianus columbianus*) Swans are two of the many North American species which experienced severe overharvesting before the protective legislation of the early 1900s. Though now of least concern, they are experiencing shifts in their breeding ecology due to changes in the range of the boreal forest in Alaska, a possible result of climate change in an especially climate-sensitive near-Arctic region. We have developed a molecular method of differentiating these two closely-related species, in an effort to determine whether any evidence of interbreeding exists in these areas of vegetative disturbance, and to elucidate the relationship between the two species and perhaps eventually their Eurasian counterparts.

Introduction

The Trumpeter (*Cygnus buccinator*) and Tundra (*C. columbianus columbianus*) are the two native North American swan species. They are morphologically similar, but the Trumpeter is larger and exhibits yellow lores, while the Tundra Swan typically possesses an entirely black bill (Banko 1960; Evans and Sladen 1980). Additionally they are separated by habitat selection during breeding season. The Trumpeter Swan prefers building nests on emergent vegetation in shallow waterbodies in boreal forest, while

Tundra Swans build nests on the edge of waterbodies in open tunda habitat (Banko 1960; Hansen *et al.* 1971; Travsky and Beauvais 2004).

As the vegetative composition along the ecotone of the boreal forest and tundra in Alaska has changed, the breeding ranges of these species have been affected. Some data suggest that this change is due to the changes in annual soil temperature and weather patterns, a result of climate change (Beck *et al.* 2001; Soja *et al.* 2006). The Trumpeter Swan, which breeds in the boreal forest, has expanded its breeding range as the boreal forest has expanded into the tundra. This range change has caused an unprecedented sympatry between the species during breeding season (Banko 1960; Jenny Bryant, personal communication). Concerns about hybridization have emerged because some swan carcasses with apparent hybrid morphology were found in the wintering grounds in Washington, and because the Trumpeter and Tundra Swans are known to be reproductively viable (Johnsgard 1974). Indeed, a captive hybrid population exists at the Airlie Center at Warrenton, Virginia. The possibility of interbreeding raises questions on the ecological implications of climate change and its role in altering the evolutionary trajectory of two diverged species.

Microsatellites have become commonly used for many questions of involving population genetics. As rapidly evolving loci, they are considered best for closely related groups, such as populations or closely related species. This high level of variation necessitates the development of species-specific primers for each locus. This has been historically very costly and time-consuming (Abdelkrim *et al.* 2009). I tested eleven microsatellite loci described by St. John *et al.* (2006) on our samples, but found

insufficient levels of heterozygosity for our purposes. Dr. Patrick Gillevet and Ms. Masoumeh Sikaroodi used next-generation sequencing technology to create genomic libraries, and I employed the bioinformatics tool MSATCOMMANDER to identify 20 new microsatellite regions. After investigation through molecular fingerprinting, I discovered that seven of these loci were polymorphic and useful in distinguishing the Trumpeter and Tundra Swans. This process is described more fully in Chapter 1. This paper describes how these new loci, in combination with two of the more variable previously described loci, successfully differentiate the Trumpeter and Tundra species and their captive hybrids. Additionally, we used this method to test whether some samples from Washington and Alaska exhibit evidence of genetic introgression.

Methods

Molecular Methods

Blood samples were taken from captive swans from the Airlie Center and from wild Tundra Swans in the North Slope, Alaska. Feather and egg shell samples were collected and sent by the U.S. Fish and Wildlife Service in Alaska, primarily from the Koyukuk National Wildlife Refuge. Heads from carcasses exhibiting hybrid morphology were collected during winter in the state of Washington and also sent. We are using two markers that were previously described by St. John *et al.* (2006), but found that the other loci described here exhibited insufficient heterozygosity to properly distinguish species. Therefore, we used next-generation sequencing technology and the bioinformatics tool MSATCOMMANDER to identify and test several microsatellite regions. This process

and the resulting 7 polymorphic loci discovered are described in Chapter 1. Primer information may be seen in Table 1 (page 28), and marker information may be seen in Table 2 (page 29).

Samples were digested with Proteinase K. Methods for DNA extraction from blood samples followed the standard protocol for the Quiagen DNeasy Tissue kit, except when congealed samples only allowed 50—100 uL of blood to be taken. Feather and egg shell samples will follow methods described by Taberlet and Bouvet (1991). Extractions will be approximately quantified by running 4 uL of product through a 1% agarose gel and visualized on a Kodak EDAS 120. Extractions will then be tested through polymerase chain reaction (PCR) of the 12S mitochondrial region. Samples which show high amounts of DNA after the gel run and which did not properly amplify the 12S region will be diluted at 1/5, 1/10, or 1/100 concentrations and tested again.

Polymerase chain reactions (20 uL) used Applied Biosystems *Taq* Gold DNA polymerase (Foster City, California). Reactions for amplification of the two St. John *et al.* (2006) (TSP_43 and TS_2A) loci contained about 10 ng DNA and Applied Biosystems (Foster City, California) 25 mM $MgCl_2$, 0.5 mM of each deoxynucleotide triphosphate, 5 M betaine, and Invitrogen 10 uM primer (Invitrogen, Carlsbad, California). The remaining seven loci were amplified using 1X bovine serum albumin instead of betaine. Amplification of microsatellites characterized by St. John *et al.* (2006) followed the reaction conditions described by those authors. These cycles were also used for amplification of the microsatellites selected from the Msatcommander sequences. Reactions for TS_2A and TSP_43 were performed in a Bio-Rad iCycler (Hercules,

California). All other reactions were performed in an MJ Research (now Bio-Rad) PTC-200. Thermal profiles were 95˚C for 11 min, 35 or 38 cycles at 95˚C for 40 sec, 1 min 30 sec at annealing temperature (Table 1), and 72˚C for 2 min, with a final extension of 10 min at 72˚C. Annealing temperatures and number of cycles may be found in Table 1. To address the issue of allele dropout, I performed 3 replicate PCRs for each loci on each sample.

Four microlitres of PCR product were run with brophenol blue and xylene cyanol dye in 40% sucrose solution in Tris-acetate buffer through a 1.5% agarose gel in a C.B.S Scientific Company PES-2000 Series II (Del Mar, California). Gels were stained in ethidium bromide solution. Products were diluted at a 1/10, 1/15, or 1/20 ratio depending on the relative concentrations of PCR product and primer viewed in the gel image. These dilutions were added to a 1/20 ratio of ILS600 size standard (Promega Corporation, Fitchburg, Wisconsin) to HiDi formamide (Applied Biosystems, Foster City, California) and were analyzed using an Applied Biosystems 3130xl Genetic Analyzer. Fingerprinting results were viewed in Genemapper v4.1 and allele lengths were determined after examination by eye.

Statistical Analysis Methods

I used three statistical software tools to assess the strength of these loci in differentiating the swan species. *Structure* 2.3.4 was used to test whether these loci successfully differentiated species at all (Falush *et al.* 2003; Pritchard *et al.* 2000). This program uses Bayesian inference to determine population *structure* by testing microsatellite genotype data under multiple population constraints. It determines the

probability of the genotype data given a HWE model under a specified number of populations. I used genotype data from 8 Trumpeter Swans and 8 Tundra Swans. I set a burn-in of 40,000 and run-length of 60,000 steps, a test for one to four populations with admixture, and five iterations for each population number. Admixture analysis was chosen to avoid biasing results with prior information against gene flow. Burn-in was determined by the likelihood dataplot output graphs after the simulation run; likelihood had stabilized by 40,000 iterations. Run-length was tested and better likelihood values were found with 60,000 iterations. These values fall in the middle of those recommended for good estimates of parameter values (P and Q) by the *structure* software manual. This output also provided a value for fixation index (F_{ST}) as one indicator for differentiation.

Additional tools included GenAlEx 6.5 (Peakall and Smouse 2006; Peakall and Smouse 2012) and SMOGD (Crawford 2010). GenAlEx 6.5 was used to produce allele frequency tables and values for expected (H_e) and observed (H_o) heterozygosity. There has been significant debate over the strength of F_{ST} as an indicator of differentiation. This calculation (F_{ST}) seeks to determine genetic distance between populations by comparing heterozygosity within and between populations. Jost (2008) developed an alternative and possibly more accurate calculation, actual differentiation (D), which compares allelic variation between populations and may be calculated easily in SMOGD. Another differentiation estimate, G_{ST_est}, is calculated by SMOGD. This value is similar to F_{ST}, but was generalized for multiple alleles by Nei and Chesser (1983). I have also included the SMOGD calculation of a standardized estimate of G_{ST} (G'_{ST_est}), which divides the G_{ST}

value by the maximum G_{ST} at that locus and is appropriate for microsatellite data (Hendrick 2005).

The five known hybrid samples were added to the input file for *structure*, and re-run with tests for two to five populations. Finally, I incorporated 15 unknown samples from Alaska and Washington to determine whether they exhibited any evidence of genetic introgression. I used the same parameters in *structure*, but tested one to eight populations. See Figure 5 for results for all sample groups.

Results

Trumpeter and Tundra Swans and their hybrids were found to possess distinct microsatellite genotypes. Analysis by *structure* determined that a two-population model was most likely, and members of both detected populations consisted of individuals of one species or another (no mixture of species). See Figure 5 for a barplot output by *structure* containing all genotyped samples, including the Trumpeter and Tundra. For details on the sample names, alleles, and allele frequencies, consult Appendices 1 and 2.

Table 3. Summary statistics for the two swan species, showing locus, allele size range in base pairs, number of alleles (Na), number of private alleles, observed heterozygosity (Ho), expected heterozygosity (He), and unbiased expected heterozygosity (uHe).

	C. buccinator						C. columbianus columbianus					
Locus	**Size range**	**N_a**	**No. Private Alleles**	**H_o**	**H_e**	**uH_e**	**Size range**	**N_a**	**No. Private Alleles**	**H_o**	**H_e**	**uH_e**
TR_MS6	158-170	3	1	0.71	0.5	0.54	162-174	4	2	0.75	0.65	0.69
TR_MS10	326-329	2	1	0.13	0.12	0.13	317-329	4	3	0.5	0.74	0.79
TR_MS11	326-330	2	0	0.5	0.38	0.4	322-330	3	2	0.13	0.40	0.43
TU_MS1	344-364	4	3	0.86	0.66	0.71	338-344	2	1	0.13	0.12	0.13
TU_MS2	246-270	6	3	0.75	0.70	0.74	250-262	4	1	0.75	0.68	0.73
TU_MS5	202-217	3	3	0.625	0.54	0.58	210-225	4	4	0.63	0.56	0.59
TU_MS6	316-321	2	1	0.13	0.43	0.46	310-326	5	4	0.75	0.68	0.73
TSP_43	199-203	3	2	0.57	0.57	0.62	199	1	0	0	0	0
TS_2A	227-230	2	2	0.13	0.49	0.53	218	1	1	0	0	0

The *structure* F_{ST} value of 0.2909 and 0.2976 for the Trumpeter and Tundra, respectively, indicate significant levels of divergence between the species at these loci (many papers cite a P-value of 0.001 as reaching "significance"). However, F_{ST} and G_{ST} can be a poor reflection of actual differentation with some genetic *structure*s (Jost 2008). Jost's (2008) estimated D may be more appropriate, and D values calculated at each locus may be seen in Table 4. As with F_{ST}, these values approach one with increased differentiation between groups. Except TR_MS6, all estimates indicate that all loci are highly differentiated between the two species.

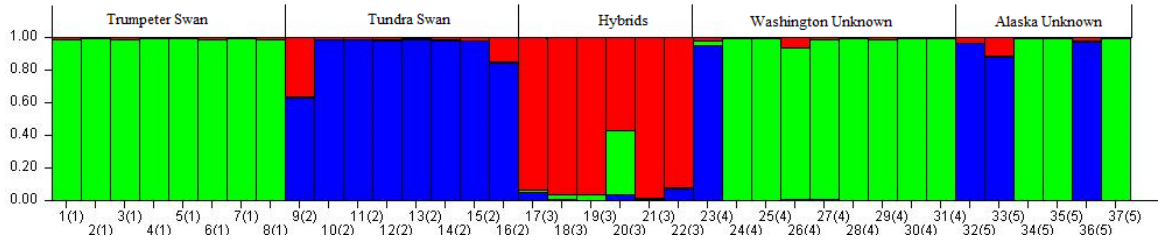
Table 4. SMOGD estimated differentiation for swan species.

Locus	$G_{ST\ est}$	$G'_{ST\ est}$	D_{est}
TR_MS6	-0.006779	-0.0284758	0.0215508
TR_MS10	0.2852792	0.7680594	0.6754808
TR_MS11	0.2391931	0.5750812	0.4414894
TU_MS1	0.4064688	0.9908909	0.9846527
TU_MS2	0.1201375	0.7808936	0.7509766
TU_MS5	0.2631579	1	1
TU_MS6	0.2378397	0.9270893	0.9043367
TSP_43	0.4735605	0.8916001	0.7940887
TS_2A	0.5841584	1	1

When all samples were analyzed (Trumpeter, Tundra, captive hybrids, unknown Washington swans, unknown Alaskan swans) in *structure*, admixture analysis determined that HWE was most likely with three populations. The likelihood table and variance for each run may be seen in Appendix 5. The resulting barplot output is shown in Figure 5 below. Columns represent individuals, which are sorted by prior population information. Colors represent the three populations detected by *structure*. The y-axis gives the Q-value, the membership coefficient estimated for each individual. This value represents the posterior probability that an individual is placed in a particular cluster based on the prior information of alleles of that individual in the model. Individuals with more than one color are not consistently placed in the same cluster, because they exhibit genotypes representative of more than one population. Nearly all individuals exhibit high membership coefficient values for the Trumpeter, Tundra, or hybrid cluster. However, some individuals did exhibit estimated membership in more than one cluster. One Tundra Swan (individual #9) was clustered with the hybrid group and significant proportion of

the time. These were collected from birds on the North Slope of Alaska, an area not inhabited by the Trumpeter Swan. Hybrids should exhibit alleles of both species, so this ancestry assignment to the hybrid population may simply be a result of hybrids and Tundra Swans sharing some alleles. However, these results may also reflect some mixed ancestry in this individual. One hybrid individual (#20) clusters more strongly with the Trumpeter group than the other hybrids. This is not concerning, since the individual is a known hybrid and may simply be a second generation hybrid whose hybrid parent mated with a Trumpeter Swan. One unknown Washington swan appears to be a Tundra Swan, while the remaining eight exhibit primarily Trumpeter genotypes. Unknown samples from Alaska are split - three Trumpeter and three Tundra Swans. This agrees with the tentative species identification for each individual provided with the Alaska samples. Individual #33 from Alaska does exhibit some hybrid genotypes. This could be a result of a small sample size for hybrids or the overlap of alleles in hybrids, as discussed for individual #9.

Figure 5. Structure barplot output with all samples, with three populations.
Each column represents an individual swan, each color represents a cluster detected by *structure*. The extent of color in each column represents the estimated membership coefficient for that individual in that cluster.



Conclusion

Here we present a novel molecular technique of differentiating the Trumpeter and Tundra Swans and their hybrids by using newly identified polymorphic microsatellite regions. This technique was also tested on wild samples of unknown lineage, showing that all individuals primarily exhibited allele frequencies of one or the other species, not hybrids. One wild individual did exhibit some hybrid alleles, but this may be due to a small hybrid sample size rather than genetic introgression. Therefore, sample sizes from all groups (species and hybrids) should be expanded before further testing on wild samples.

Additionally, this method should be expanded to include the Whooper (*C. cygnus*) and Bewick's (*C. columbianus bewickii*) Swans of Europe and Asia. These are two closely related species, but the four swans of genus *Cygnus* have never been compared in a molecular study. Their morphology is very similar - the Whooper is close to the Trumpeter in size but has a bill with yellow extending from the lores halfway down the bill, and the Bewick's Swan is similar to the Tundra Swan in size and has yellow extending from the lores one-third down the bill (Evans *et al.* 1980). These microsatellite loci could prove highly informative in determining relatedness within this genus.

SUMMARY

The development of sympatry in some of the range of the Trumpeter and Tundra Swans during breeding season introduces a possibility of interbreeding in the wild. Investigating this recent phenomenon required the development of a molecular method of species differentiation in order to detect genetic introgression. The inability of mitochondrial markers and previously described microsatellite regions to distinguish species and hybrids necessitated the development of novel microsatellite markers.

The microsatellite markers we described here have proved successful at distinguishing the two species and their hybrids, and may be used in future genetic studies at the population or species level. These loci exhibit higher levels of allelic richness than those described previously by St. John *et al.* 2006 (5.4286 and 3.0625, respectively), and they provide high estimates of differentiation between the two species, making them highly informative as markers. This further proves the effectiveness of using next generation sequencing technology to discover informative microsatellite regions in genomic reads.

To fully investigate the relationships between these two species, the sampling should be expanded to include more swans from across the range of both species. Including the Eurasian species would also be critical, since the Bewick's Swan is more closely related to the Tundra Swan than is the Trumpeter Swan. This comparison of the

four species with these markers could prove highly informative. Additionally, this method may be used to test any wild swans, if an interbreeding population is identified.

APPENDICES

Appendix 1. Newly-described microsatellite library read name, assigned locus name, and motif and position in read as indicated by MSATCOMMANDER.

Read Name	Locus Name	Motif	Start	End
Trump_HNJQXLY01BV8LE	TR_MS6	AAAG	297	349
Trump_HNJQXLY01ARE4I	TR_MS10	AAT	159	192
Trump_HNJQXLY01AVJHI	TR_MS11	AAAC	260	292
Tundra_HNOX2GF01A5ZCX	TU_MS1	AAACC	379	409
Tundra_HNOX2GF01ALTG0	TU_MS2	AAGG	345	397
Tundra_HNOX2GF01BE1FJ	TU_MS5	AAAAC	132	162
Tundra_HNOX2GF01BE1Y0	TU_MS6	AAAAT	210	250

Appendix 2. Sequence reads.
Sequence reads containing the seven selected markers.

Read Name	Read Length	Sequence
Trump_HNJQXLY01BV8LE	537	ATTCAACAAAAAAATAAGAAGCAGGGGA GAAAAAAAAATGTACCCCTTACAAATAAA ATAAATAAATAAATAAATAAATAAAACAA TAAAAAAGTTCCATGATAATAAAGTAAAA ATTGCCAGAGCAGACTGAACAACCAATAA TCAGCCTGTTAGTTGCCTCACAGTGATCCT CTCTACATCCCTTAGAGGAAGGGAAGACT GTAGATTAAGGAATCTACTGTGTGCTTCTC CTGAACTCTGGACAACTTCAAGGGTTAGA TCCCTTGGCTGTCATCCTCTTGAGGTTTTT AACCTCTTTCTTTCTTTCTTTCTTTCTTTCTT TCTTTCTTTCTTTCTTTCTTTCTTTCTATAGT TATGGGTATTCTATGTCATCTTAGGAATCA TCAGGGTAATTTGAAGAGAAGGCATTGAT CTCAAGCTTCATGGAGCTTGAAGTGTGTTT TCTTCTGCAAGTGCCATCCTTAAGTTCTCT GACAAGGAAATACACAAGTATGGAAGAC AGTGGTTTAACACATGAGTTGTGCAACAG TCCAN
Trump_HNJQXLY01ARE4I	449	GACGTTGCAGTTGGCTTCCCACGCTGCTGC CTATGTGAGAATGGGCTTGAACTGTGGGA AGCCTGCATTACCTCTTCTGCAATGCTCAT ATCCAGTTGCTCCTTGGCACTACATCTACC CAGCACACTGTTAACATAAACAGTTACTC CAATAACATTTATTATTATTATTATTATTA TTATTATTATTATTAGTAGTAGTAGTAGTA GTAGTAGTAATAGCAGTAGTAGCAGTAGT AGTAGTAGTAGTAGCTGCCAGTCTCAATG GTACATACCAAAACAAACAAACAGACAA AACCAATAAAATACCTGTCCATAGGGTTT TGGAGACTTAGGAAATCAGATTATAAACT GCAAAATACTGCACCTGTGTTTTGCAACAT TCACGTTTGGTTGCAAATCCTTTCCTCACA GATCAGCCCACCTACCTGCTTGAACTTGCC AGCAAGT

Trump_HNJQXLY01AVJHI	492	CCATCACTATGAATTCCTGTAGACGTTGGC ACTTCCATCTCACACTTTGCATCACTGAAT ATCACAATGCTTACAATGCTTGGTCTCCTT GGTGAAGGAGATTTCATCAAACCTCTTCA TTCTGATTCTTTTTGCCAGAAAAATAGTTG TGTACCAAAATTGCTCTGTGGATATAGGC CCTATCTGTATTCCTACATTGCTGTTTGAG TGACAATGTATCTTGAAATTTTCATATTTT GTTTAAACTGAAAACTGATGTTGTTTGTTT GTTTGTTTGTTTGTTTGTTTGTTTTTAACAT CATAGCTGGGTGCCCTAGAGAGTTTTTTTT CCTGTTTTTTTAACTTTTTTTTTCCTAATTT ATAATCTAACTCAAGAAAACTGTGTGTGT GGCTTTGTCATTAAAAAAACATTGTAGAG AGATGGAAGGTATGCATTAGTACTTGGTT GACAACTGATCCCAGCAGTTCCAATGAGC GAAAGGCTAGTATATC
Tundra_HNOX2GF01A5ZCX	492	AGTGTGTGGTAGTAACTATAGGTGTTATTC TGTAGCTATCTGCCTTTTCTTCTTCTATGCT TCTGAAGCAGAGCTACACTCATCCTCATGT CAGCTACTAAAGCTCTCTGAAGGCAGAGC CAGTGCAGTGAGACCTGTTGCTGAGGAGT CTTACGTTTGGGCACCTGAGTTCTGCAAGG TAATTTGTCTGTGGGAACATAGAATTAAT ATTACTATTCATTCTACTTACTTTTATTAA GAAGATGAACCAACCTAAAAAGAAACACT AGTTGAGAAAGAAAGTTTTGGGGCAATGA GTCCTTATAGAACTTGTTTCCCTTAATTCA GCATTTTGAGACAAATAATATCTGTTTTGT TGTTGTTGTTTGTTTGTGTGTTTGGTTTGGT TTGGTTTGGTTTGGTTTGGTTTTTTTTGAGG GTATGAAATTGGTACAGAAGTAGTAACTG CAAAGATAGCTTGTGCAATCCTTGCTCCA GGGCTGTTATCAAAGG
Tundra_HNOX2GF01ALTG0	502	ACCAATACTTCTCTTACTCTTCAGCTCCTA AGCAGAAAGAAATGAACAGACTAATGGC AAGCAGGCATGGCCCACATCCCAAGATCT TAATTACCCTGTGCAGCAGCATGTACAGC AGAAGTTGTCTAACTTTTGCACAAAGGAG AAAGATCACAAGGTCAGGAGCTACACTAA CACTCCTTCAGACTTGAGCTTGGGCTCTGC ATCCTCTCCATGGCTTAAAGAATACAGCTT TCCTTCTTGGTATATGGTGTTAGTCTTGAA GCTCGGTTCAATATGCAGTATGCAAGGTT CAGAGCTCCAAATTCTGCCAATATTTCAAT GGTGAGGGTTACTCTGGCTGGGAAGGAAG GAAGGAAGGAAGGAAGGAAGGAAGGAAG GAAGGAAGGAAGGAAGGAGGACTGGCTT ACTGTTATTTAGGTTACAAAGGCTCCCCCG ATGCATGCTTAGGATGGTGTTGACATTAAT GTCAGATCTGTAATAAAGTCAGTAGCTCA TGCTC

Tundra_HNOX2GF01BE1FJ	459	ATCTCCACTCCACTCTGGCCCACTTGTTTT CCTCCATCTTCTCTCCAAGCTCCTTTCCCC ACAAGGATTCAGTAGGTGCCTGCCTGGTC TGGCCTCTACCTGGAGCAGGGAAAGCCAA GCACTGCTGGGAAAAAAACAAAACAAAA CAAAACAAAACAAAACAAACAAAAAAAA CACCCAAGCTCTGTAATAGACTTTTCATGA CATTTACTAAATCTTTTAGAAGTTTTCCTT CTTTCTGCACTAAATATCGGTTGCTTAAGA TAATCTGCAAACCTGTAGAACGGCCATTA GGAGCTCCCTGGTGACTAGGTTAATTCAC AGTAAATTACAGTCACTGCCAAGGAGATT GCAGGGAGGTTAGGAGGAGATTGGAAGC GGCAGGCAGTAAAGCACTGTAGCATTTAC CAAAAAAAAAAAAACGAAAAAAAAACGAA AAGAAAATAAGAACAATAAATAC
Tundra_HNOX2GF01BE1Y0	484	AACTAAACAGTCTGTTTGAGTAAATATTTC TTTATGGAAGTGTTAAAGTAATTCCTTCCT AACACAACAGTGAACTCCAGGGACAAGTG GAGCTTGATGATTTACAGGATAAAGATTT TCTTAGTAAACAAAAGACACTGCAGTGTT GACCAGGGACTTTTATTTATTCTATTTTAT TTTATTTTTATTTTATTTATTTTATTTTACTT TATTTTATTTTATTTTATTTTATTTTATTTT ATTTTATTTTATTCTATTTTATTTATTTTAT TTTATTTTATTTTATTTATTTTTAGTAAATG CAAGGTATTCATAGCTTAAAGAGAAAAGC AGAAGGGCAGATATATTGTATGAGATGTG TAGTATTAGTGGTGTGGGAGGATACTGCA TAAGCCATCTTTTTGGCATTTCTGTCACTA TGTATTTAATATTCTTAATAAACATCCTGC TTACATTTTTGACCATACCAAGAGAGGTG AATATC

Appendix 3. Sample descriptions.

Species or Origin	Sample Number	Sample Name
Trumpeter (known captive, Airlie)	1	R07
	2	R34
	3	R37
	4	R43
	5	Z580
	6	Z582
	7	Z587
	8	Z614
Tundra (known wild, North Slope)	9	T192
	10	T231
	11	T233
	12	T234
	13	T235
	14	T236
	15	T237
	16	T248
Hybrid (known captive, Airlie)	17	Z219
	18	Z363
	19	Z507
	20	Z530
	21	Z531
	22	Z595
Washington (unknown wild)	23	SRP13
	24	SRP14
	25	SRP15
	26	SRP16
	27	SRP21
	28	SRP22
	29	SRP23
	30	SRP24
	31	SRP25
Alaska (unknown wild, Koyukuk)	32	707
	33	708
	34	717
	35	719
	36	727
	37	728

Appendix 4. Allele lengths and frequencies by locus and species.

Locus	Allele (bp)	Trumpeter	Tundra	Hybrid
TR_6	**N**	7	8	6
	158	0.071	0.000	0.000
	162	0.000	0.188	0.417
	166	0.286	0.250	0.333
	170	0.643	0.500	0.167
	174	0.000	0.063	0.083
TR_10	**N**	8	8	6
	317	0.000	0.250	0.250
	320	0.000	0.313	0.000
	323	0.000	0.250	0.500
	326	0.063	0.000	0.000
	329	0.938	0.188	0.250
TR_11	**N**	8	8	6
	322	0.000	0.063	0.000
	326	0.250	0.750	0.583
	330	0.750	0.188	0.417
TU_1	**N**	7	8	6
	338	0.000	0.938	0.167
	344	0.143	0.063	0.000
	354	0.143	0.000	0.000
	360	0.500	0.000	0.833
	364	0.214	0.000	0.000
TU_2	**N**	8	8	6
	246	0.500	0.000	0.000
	250	0.063	0.188	0.500
	254	0.125	0.063	0.000
	258	0.125	0.375	0.250
	262	0.000	0.375	0.000
	266	0.125	0.000	0.000
	270	0.063	0.000	0.250
TU_5	**N**	8	8	6
	202	0.375	0.000	0.000
	210	0.000	0.188	0.000
	212	0.063	0.000	0.083
	215	0.000	0.625	0.167
	217	0.563	0.000	0.583
	220	0.000	0.125	0.167
	225	0.000	0.063	0.000
TU_6	**N**	8	8	6
	310	0.000	0.063	0.000
	315	0.000	0.500	0.000
	316	0.688	0.000	0.167
	320	0.000	0.125	0.000

	321	0.313	0.125	0.417
	326	0.000	0.188	0.083
	330	0.000	0.000	0.333
TSP_43	**N**	7	8	6
	199	0.143	1.000	0.750
	203	0.571	0.000	0.167
	205	0.286	0.000	0.083
TS_2A	**N**	8	8	6
	218	0.000	1.000	0.833
	227	0.563	0.000	0.000
	230	0.438	0.000	0.167

Appendix 5. Likelihood and variance for each structure run.
Chosen run highlighted.

Set	K	Ln P(D)	Var[LnP(D)]
SwanParam2_run_30	1	-905.7	19.8
SwanParam2_run_29	1	-904.8	18.5
SwanParam2_run_28	1	-905.4	19.5
SwanParam2_run_27	1	-904.6	17.2
SwanParam2_run_26	1	-904.9	18.6
SwanParam2_run_5	2	-705.9	50.5
SwanParam2_run_4	2	-705.7	50.4
SwanParam2_run_3	2	-706	50.9
SwanParam2_run_35	2	-706.1	50.7
SwanParam2_run_34	2	-705.9	50.7
SwanParam2_run_33	2	-705.7	50.2
SwanParam2_run_32	2	-705.9	50.3
SwanParam2_run_31	2	-705.7	50.2
SwanParam2_run_2	2	-705.2	49.6
SwanParam2_run_1	2	-706.5	51.2
SwanParam2_run_9	3	-684.2	75.8
SwanParam2_run_8	3	-683.3	74.1
SwanParam2_run_7	3	-682.1	72
SwanParam2_run_6	3	-703.1	77
SwanParam2_run_40	3	-682.8	73.2
SwanParam2_run_39	3	-684.8	76.7
SwanParam2_run_38	3	-686	78.9
SwanParam2_run_37	3	-699.5	102.4
SwanParam2_run_36	3	-684.5	76.6

SwanParam2_run_10	3	-682.7	73.1
SwanParam2_run_45	4	-710	145.6
SwanParam2_run_44	4	-713.7	153.6
SwanParam2_run_43	4	-704.3	118.6
SwanParam2_run_42	4	-703.1	154.9
SwanParam2_run_41	4	-722.6	165.6
SwanParam2_run_15	4	-715.7	151.8
SwanParam2_run_14	4	-691	89.3
SwanParam2_run_13	4	-714.3	141.6
SwanParam2_run_12	4	-714.8	152.3
SwanParam2_run_11	4	-710.8	145.8
SwanParam2_run_50	5	-723.7	173.6
SwanParam2_run_49	5	-750.2	229.9
SwanParam2_run_48	5	-714.3	150.1
SwanParam2_run_47	5	-717.5	163.7
SwanParam2_run_46	5	-722.1	167.9
SwanParam2_run_20	5	-725.5	161.5
SwanParam2_run_19	5	-767	267.4
SwanParam2_run_18	5	-740.1	197.7
SwanParam2_run_17	5	-725.2	174.7
SwanParam2_run_16	5	-724.2	174.2
SwanParam2_run_25	6	-741.9	207.6
SwanParam2_run_24	6	-732.1	195.1
SwanParam2_run_23	6	-722.4	165.4
SwanParam2_run_22	6	-724.3	169.3
SwanParam2_run_21	6	-763.6	263.7

Appendix 6. Alleles of the St. John et al. microsatellites for Trumpeter, hybrid, and Tundra swans.

	Locus											
	TS_32		TSP_9		TS_54		TS_57		TS_51		TSP_2B	
Trumpeter	155	157	157	171	221	221	186	188	155	161	-	-
	-	-	157	171	221	221	188	188	-	-	-	-
	155	155	157	157	221	221	188	188	-	-	-	-
	155	157	157	157	221	221	188	188	-	-	-	-
	157	157	-	-	-	-	188	188	159	159	203	203
Hybrid	155	157	157	171	221	221	188	188	-	-	-	-
	157	157	-	-	221	221	188	188	159	161	203	205
	157	157	-	-	221	221	188	188	159	161	203	203
	157	157	-	-	221	221	188	188	159	161	205	205
Tundra	155	155	-	-	221	221	186	188	159	161	203	203
	155	155	173	175	221	221	188	188	-	-	-	-
	155	155	169	171	221	221	188	188	-	-	-	-
	155	155	171	171	221	221	188	188	-	-	-	-
	155	155	-	-	221	221	188	188	155	159	203	203
Number of Alleles	2		4		1		2		3		2	

	TSP_16		TS_25		TS_30		TSP_43		TS_2A	
Trumpeter	143	145	149	149	176	178	205	205	230	230
	-	-	-	-	-	-	205	205	230	230
	-	-	-	-	-	-	203	205	230	230
	-	-	-	-	-	-	203	205	230	230
	139	143	149	149	176	176	205	205	-	-
Hybrid	-	-	-	-	-	-	199	205	218	218
	143	143	149	149	176	178	199	203	218	218
	143	143	149	149	-	-	203	203	218	218
	143	143	149	149	176	176	199	199	218	218
Tundra	143	143	149	149	174	176	199	199	218	218
	-	-	-	-	-	-	199	199	218	218
	-	-	-	-	-	-	199	199	218	218
	-	-	-	-	-	-	199	199	-	-
	143	143	149	149	174	176	199	199	218	218
Number of Alleles	3		1		3		3		2	

LITERATURE CITED

Abdelkrim J, Robertson BC, Stanton, JL, Gemmell N (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. BioTechniques, 46, 185–192.

Banko, WE (1960) The Trumpeter Swan. North American Fauna 63. U.S. Fish and Wildlife Service, Washington, D.C.

Barrett VA, Vyse ER (1982) Comparative genetics of three Trumpeter Swan populations. The Auk, 99, 103–108.

Beck PSA, Juday GP, Alix C, Barber VA, Winslow SE, Sousa EE, Heiser P, Herriges JD, Goetz SJ. (2011) Changes in forest productivity across Alaska consistent with biome shift. Ecology Letters, 14(4), 373–379.

Caithamer DF (2001) Trumpeter Swan population status 2000. U.S. Fish and Wildlife Service. Laurel, MD. Available http://library.fws.gov/Bird_Publications/trumpeterswan_pop00.pdf. Accessed January 23, 2013.

Crawford, NG (2010) SMOGD: software for the measurement of genetic diversity. Molecular Ecology Resources, 10, 556–557.

Delacour J, Mayr E (1945) The family Anatidae. The Wilson Bulletin 57(1), 3–55.

Domingo-Roura X (2002) Genetic distinction of marten species by fixation of a microsatellite region. Journal of Mammalogy, 83(3), 907–912.

Donne-Gousse C, Laudet V, Hanni C (2002) A molecular phylogeny of anseriformes based on mitochondrial DNA analysis. Molecular Phylogenetics and Evolution, 23,339-356.

Earnst, SL (1994) Tundra Swan habitat preferences during migration in North Dakota. Journal of Wildlife Management, 58(3), 546–551.

Evans, ME, William, Sladen WJL (1980) A comparative analysis of the bill markings of whistling and Bewick's swans and out-of-range occurrences of the two taxa. The Auk, 97, 697–703.

Faircloth, BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources, 8, 92–94.

Falush D, Stephens M, Pritchard JK (2003) Inference of population *structure*: Extensions to linked loci and correlated allele frequencies. Genetics, 164, 1567–1587.

Ferreira da Silva MJ, Minhos T, Sa RM, Bruford MW (2012) Using genetics as a tool in primate conservation. Natural Education Knowledge, 3(10), 89.

Hansen HA, Shepherd PEK, King JG, Troyer WA (1971) The Trumpeter Swan in Alaska. Wildlife Monographs, 26, 3–83.

Hendrick PW (2005) A standardized genetic differentiation measure. Evolution, 59(8), 1633–1638.

Kulikova IV, Drovetski SV, Gibson DD, Harrigan RJ, Rohwer S, Sorenson MD, Winker K, Zhuravlev YN, McCracken KG (2005) Phylogeography of the Mallard (*Anas platyrhynchos*): hybridization, dispersal, and lineage sorting contribute to complex geographic structure. The Auk, 122(3), 949–965.

IUCN (2013) International Union for the Conservation of Nature Red List. Available at http://www.iucnredlist.org/. Accessed January 23, 2013.

Johnsgard PA (1974) The taxonomy of the northern swans. Wildfowl, 25, 155–161.

Johnsgard PA (1978) Revised 2010. University of Nebraska Press. Lincoln, NE.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. Molecular Ecology, 17(18), 4015-4026.

LaMontagne JM, Barclay RMR, Jackson LJ (2001) Trumpeter Swan behavior at spring-migration stopover areas in southern Alberta. Canadian Journal of Zoology, 79, 2036–2042.

LaMontagne JM (2003)Characteristics of ponds used by Trumpeter Swans in a spring migration stopover area. Canadian Journal of Zoology, 81, 1791–1798.

McKelvey RW, Dennington MC, Mossop D (1983) The status and distribution of Trumpeter Swans (*Cygnus buccinator*) in the Yukon. Arctic, 36(1), 76–81.

Meng A, Parkin DT (1993) Genetic differentiation in natural populations of swans revealed by DNA fingerprinting. Acta Zoologica Sinica, 39, 209–216.

Moser TJ (2006) The 2005 North American Trumpeter Swan Survey. U.S. Fish and Wildlife Service. Denver, Colorado. Available http://library.fws.gov/Bird_Publications/trumpeterswan_survey05.pdf. Accessed January 23, 2013.

Monda MJ, Ratti JT, McCabe TR (1994) Reproductive ecology of Tundra Swans on the Arctic National Wildlife Refuge, Alaska. Journal of Wildlife Management, 58(4), 757–773.

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Annals of Human Genetics, 47(3), 253–259.

Oyler-McCance SJ, Ransler FA, Berkman LK, Quinn TW (2007) Conservation Genetics, 8, 1339–1353.

Pacific Flyway Council (2001) Pacific Flyway management plan for the western population of tundra swans. Pacific Flyway Study Comm., Subcomm. on Tundra Swans. Unpubl. Rept. [c/o USFWS], Portland, OR. 28 pp.+ appendices. Available at http://pacificflyway.gov/Documents/Wts_plan.pdf. Accessed January 23, 2013.

Peakall R, PE Smouse (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288–295.

Peakall R, PE Smouse (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics, 28, 2537–2539.

Peters JL, McCracken KG, Zhuravlev YN, Lu Y, Wilson RE, Johnson KP, Omland KE (2005) Phylogenetics of wigeons and allies (Anatidae: *Anas*): the importance of sampling multiple loci and multiple individuals. Molecular Phylogenetics and Evolution, 35, 209–224.

Petrie SA, Badzinski SS, Wilcox KL (2002) Population trends and habitat use of Tundra Swans staging at Long Point, Lake Erie. Waterbirds, 25(Special Publication 1), 143–149.

Poyarkova N, Klenova A, Kholodova M (2010) Genetic diversity of Swan Goose (*Anser cygnoides L.*) in Russia: analysis of the mitochondrial DNA control region polymorphism. Russian Journal of Genetics, 46(4), 493–496.

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

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (eds Krawetz S, Misener S). Humana Press, Totowa, NJ.

Ruokonen M, Kvist L, Lumme J (2000) Close relatedness between mitochondrial DNA from seven Anser goose species. Journal of Evolutionary Biology, 13, 532–540.

Ruokonen M, Kvist L, Aarvak T, Markkola J, Morozov VV, Oien IJ, Syroechkovsky EE, Tolvanen P, Lumme J (2004) Population genetic structure and conservation of the lesser white-fronted goose (*Anser erythropus*). Conservation Genetics, 5, 501–512.

Ruokonen M, Aarvak T, Madsen J (2005) Colonization history of the high-arctic pink-footed goose *Anser brachyrhynchus*. Molecular Ecology, 14, 171–178.

Schlotterer C (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma, 109(6), 365–371.

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

Shea RE, Nelson HK, Gillette LN, King JG, Weaver DK (2002) Restoration of Trumpeter Swans in North America: a century of progress and challenges. Waterbirds, 25 (Special Publication 1), 296–300.

Soja AJ, Tchebakova NM, French NHF, Flannigan MD, Shugart HH, Stocks BJ, Sukhinin AI, Parfenova EI, Chapin III FS, Stackhouse Jr. PW (2007) Climate-induced boreal forest change: Predictions versus current observations. Global and Planetary Change, 56(3-4), 274-296.

St. John J, Ransler FA, Quinn TW, Oyler-McCance SJ (2006) Characterization of microsatellite loci isolated in Trumpeter Swan (*Cygnus buccinator*). Molecular Ecology Notes, 6, 1083–1085.

Taberlet P, Bouvet J (1991) A single feather as a source of DNA for bird genetic studies. The Auk, 108(4) 959–960.

Travsky A, Beauvais GP (2004) Species assessment for the Trumpeter Swan (*Cygnus buccinator*) in Wyoming. BLM, Cheyenne, WY.

Wetmore A (1935) A record of the Trumpeter Swan from the late Pleistocene of Illinois. Wilson Bulletin, 47(3), 237.

Wilk, RJ (1988) Distribution, abundance, population *structure* and productivity of Tundra Swans in Bristol Bay, Alaska. Arctic, 41(4), 288–292.

Wood TC, Brooks TO, Sladen WJL (2002) Vocal characteristics of trumpeter and Tundra Swans and their hybrid offspring. Waterbirds, 25 (Special Publication 1), 360–362.

Xia C, Hu T, Yang T, Wang L, Xu G, Lin C (2005) cDNA cloning, genomic *structure* and expression analysis of the goose (*Anser cygnoides*) MHC class I gene. Veterinary Immunology and Immunopathology, 107, 291–302.

Zinc RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology, 17(9), 2107–2121.

BIOGRAPHY

Lauren Wilson graduated from Westminster Preparatory School in Augusta, Georgia, in 2005. She received her Bachelor of Science in Wildlife Management from the Warnell School of Forestry and Natural Resources, University of Georgia, in 2008. She was employed as a teacher in Edgefield County, South Carolina, for one year and then worked as the Assistant Education Coordinator for Reed Creek Wetlands Park and Interpretive Center in Augusta, Georgia.