11-2019

Analysis of Genetic Variation of Western and Eastern Populations of Socotra Cormorant (Phalacrocorax Nigrogularis) In the UAE

Salama Rashed Al Mansouri

United Arab Emirates University

College of Science

Department of Biology

ANALYSIS OF GENETIC VARIATION OF WESTERN AND EASTERN POPULATIONS OF SOCOTRA CORMORANT *(PHALACROCORAX NIGROGULARIS)* IN THE UAE

Salama Rashed Al Mansouri

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science in Environmental Sciences

Under the Supervision of Dr. Sofyan Alyan

November 2019

Declaration of Original Work

I, Salama Rashed Al Mansouri, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled "*Analysis of Genetic Variation of Western and Eastern Populations of Socotra Cormorant (Phalacrocorax Nigrogularis) in the UAE*", hereby, solemnly declare that this thesis is my own original research work that has been done and prepared by me under the supervision of Dr. Sofyan Alyan, in the College of Science at UAEU. This work has not previously been presented or published, or formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

Student's Signature: _____ Date: 4/12/19

Approval of the Master Thesis

This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Sofyan Alyan
   Title: Associate Professor
   Department of Biology
   College of Science
   Signature _____ Date 14/11/2019

2) Member: Taoufik Ksiksi
   Title: Professor
   Department of Biology
   College of Science
   Signature _____ Date 14/11/2019

3) Member (External Examiner): Mohammad Shahid Nazir
   Title: Associate Professor
   Department of Biotechnology
   Institution: University of Modern Sciences, UAE
   Signature _____ Date 14/11/2019

This Master Thesis is accepted by:

Dean of the College of Science: Professor Ahmed Murad

Signature _____ Date 24/12/2019

Dean of the College of Graduate Studies: Professor Ali Al-Marzouqi

Signature _____ Date 20/1/2020

Copy 6 of 6

Abstract

Understanding the genetic structure of threatened and endangered species is important in management and conservation efforts. Measuring the pattern and scale of genetic variation within and among populations is important for understanding population dynamics. It helps us improve our understanding of the ecological and genetic relation between the populations. The current research looks at the population structure of the endemic seabird of the United Arab Emirates (UAE), the Socotra cormorant (*Phalacrocorax nigrogularis*). The lack of genetic information, increased threats, and small breeding habitats of the Socotra cormorants, makes it important to investigate the genetic population structure in the UAE. The aim of this research is to assess the level of genetic variation and structure within Eastern and Western Socotra Cormorant populations, using descriptive molecular genetic analysis. Using known mtDNA and nDNA primers of avian and cormorant species, we investigated the genetic differentiation and structure of the Socotra cormorants and assessed their genetic diversity. The results revealed that the Western and Eastern populations have low genetic differentiation and high gene flow. Also, they have low genetic diversity across all populations, which might indicate that the UAE population is recovering from a long-term bottleneck or an event of selective sweep. Now we have a closer insight into their genetic diversity, a further study of whole genome sequence (WGS) is required to get a better understanding of the population's genetic history and dynamics. This will enable us to additionally understand the Socotra cormorant global population and their connectivity.

Keywords: Genetic variation, Socotra cormorant, mtDNA and nDNA markers, genetic diversity, genetic differentiation, population genetics, genetic structure.

Title and Abstract (in Arabic)

تحليل التباين الوراثي بين أعداد الغربية والشرقية للغاق السقطري (Phalacrocorax nigrogularis) في دولة الإمارات العربية المتحدة

الملخص

إن فهم التركيب الوراثي لأنواع الكائنات المهددة والمعرضة للانقراض مهم لفهم ادارتها والمحافظة عليها. قياس نمط وحجم التباين الوراثي داخل وبين الأعداد يعد أمر هام ويساعدنا على تحسين معرفتنا للعلاقة الإيكولوجية والجينية بين المجموعات. في بحثنا اننا ننظر في التركيب الجيني للطيور البحرية المستوطنة في دولة الإمارات العربية المتحدة وهو الغاق السقطري (Phalacrocorax nigrogularis). إن قلة المعلومات الوراثية والتهديدات المتزايدة وتقلص مدى السكن الطبيعي للغاق السقطري يحثنا على التحقيق في تركيبتها الوراثية. الهدف من هذا البحث هو تقييم مستوى التباين الوراثي والبنية الجينية داخل مجموعات سرب الغاق في شرق وغرب الدولة ، باستخدام الوصف والتحليل الوراثي الجزيئي. استخدمنا علامات الجينية على mtDNA و nDNA المعروفة لدى الطيور والغاق ودرسنا التمايز والهيكل الوراثي بين مجموعات السرب وقيمنا التنوع البيولوجي. أظهرت النتائج أن الاعداد الغربية والشرقية لديهم تمايز جيني منخفض جدا وتدفق جيني مرتفع. وأن التنوع الجيني منخفض في كلى الأسراب مما قد يشير إلى أن الأعداد تتعافى من حالة bottleneck منذ المدى الطويل أو من حدث اكتساح انتقائي. الآن لدينا نظرة فاحصة على تنوعها الجيني وهناك حاجة إلى مزيد من الدراسات لتسلسل الجينوم الكامل (WGS) للحصول على فهم اوضح للتاريخ الجيني وديناميكية الأعداد. بالإضافة انه سوف يمكننا من فهم تفاعل واتصال أعداد الغاق السقطري ببعضها البعض حول العالم.

مفاهيم البحث الرئيسية: لتباين الوراثي، العاق السقطري، علامات الوراثة mTDNA و nDNA، التنوع الوراثي، التمايز الوراثي، الوراثيات السكانية، التركيب الوراثي.

Acknowledgements

I would first like to thank my thesis advisor Dr. Sofyan Alyan at United Arab Emirates University. The door to Prof. Alyan office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it.

I would also like to thank the experts who were involved in this research project my co-advisors: Professor Saber Bin Muzzafer and Professor Rabah Iratni. Without their valuable participation and input, the establishment of this thesis could not have been successfully conducted.

I would like to thank the chair and all members of the Department Biology at the United Arab Emirates University for assisting me all over my studies and research. My special thanks also extended to the Genomic Laboratory at Biology department at the UAEU and Khalifa Center for Genetic Engineering & Biotechnology for providing me with the needed laboratory assistance and counseling.

I am also grateful to the Scholarship received through the Alain Zoo and the Sandooq Al Watan Fund (Grant 0018), UAE University to undertake my Master.

Special thanks go to my parents, brothers, sisters and husband who helped me along the way. I am sure they suspected it was a never-ending research.

Dedication

To my beloved kids, parents, husband, and family

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ETD	total number of mutation
HD	Haplotype Diversity
K	average number of nucleotide differences
ML	Maximum Likelihood
mtDNA	mitochondrial DNA
nDNA	Nuclear DNA
Nm	gene flow
PCR	Polymerase Chain Reaction
S	Siniya
SNPs	Single Nucleotide Polymorphisms
UQ	Umm Qasr

Chapter 1: Introduction

1.1 Overview

The biological diversity of our planet is rapidly diminishing as a direct and indirect consequence of human behavior. A large number of species are already extinct, and the populations of many others had been reduced to levels where they risk extinction. One of the most critical issues in conservation biology is the decrease of genetic variability and the increased likelihood of extinction caused by habitat fragmentation, hunting, pollution or other anthropogenic pressures. All these threats will directly result in decreasing number of individuals and becoming isolated populations, and which would compromise the species' survivability (Mills, 2013).

Many studies on wild vertebrates suggested that endangered species with very low genetic variation may be less able to adapt to new diseases, climate shifts or other changes, and this can compromise their survival. For example, California sea lions with low heterozygosity were more likely to harbor infectious diseases and parasites. They also took longer to recover compared to others and acted as reservoirs of infectious diseases, and costed more to treat and rehabilitate (Acevedo-Whitehouse *et al.*, 2003). The loss of genetic diversity generally limits evolutionary potential (Johansson *et al*., 2007), and populations are more likely to experience inbreeding depression. The increase of inbreeding and random genetic drifts within the habitats could drive it to extinction. (Keller & Waller, 2002; Willi *et al.*, 2006; Wright *et al*., 2007).

In the past two decades, genetic analyses had been used to determine in details the prospective status of several endangered species. Such approaches have

provided valuable insights that have critically affected management decisions and produced tangible benefits for the species studied (Mills, 2013).

In particular, data from genetic markers with different modes of inheritance, specifically maternally inherited mitochondrial DNA and bi-parentally inherited nuclear DNA, have revealed a complex and often-conflicting patterns of genetic structure. Using biparentally inherited nuclear DNA and maternally inherited mitochondrial DNA (mtDNA) to quantify genetic variation and differentiation among populations at a particular scale can test theories regarding historical and contemporary gene flow. Specifically, measuring the magnitude and pattern of nuclear DNA and mtDNA variation among known breeding ground crossing two or more subspecies, and use cluster analysis to infer the presence of genetic groups across a sampled range (Mills, 2013).

Population genetics chiefly focuses on the processes and mechanisms, which make evolutionary changes; it provides the genetic support for all evolutionary biology. Its goal is to understand the effects of various forces that result in evolutionary changes in species over time. Population genetic studies, especially in combination with other ecological research, can aid our understanding of population history (Hansen & Taylor, 2008) and provide necessary information for the designation of management units (Gonzalez-Suarez *et al*., 2009). Understanding the genetic structure of threatened and endangered species is important in management and conservation efforts. The variation in alleles is crucial for the survival of a species and allows organisms to adapt to changing environments. This makes quantifying the pattern and scale of genetic variation within and among populations is essential for understanding relationships between population dynamics and genetic

structure and improve understanding of the ecological and genetic dynamics of populations (Barlow *et al*., 2011).

The current research focus at the population structure of the endemic Seabird of the UAE, the Socotra cormorant (*Phalacrocorax nigrogularis)*, which will be established using molecular genetic analysis. It will allow us to examine whether the existing populations found in Arabian Gulf and the Arabian Sea are mixing or not by looking into their genetic variation. The lack of information increased threats, and small breeding habitats of the Socotra cormorant, gives urgency to investigating the population in the UAE. If the genetic status of determined, a management strategy can be planned and implemented to ensure the long-term integrity of this population. Therefore, understanding the Socotra cormorant genetic structure will give essential knowledges for future implementations in conservation management, and allow prediction of future population status and condition.

1.2 Objective

The aim of this research is to assess te level of genetic variation within populations of Socotra cormorant, *Phalacrocorax nigrogularis*, using descriptive analysis of mitochondrial control region and nuclear marker sequences. The analyses will allow (1) estimate and comparing the genetic variability of the two populations; (2) describe the genetic differentiation between them; (3) describe the population structuring. Those results will be used to build a baseline for population genetic dataset, to monitor the current, as well predict the future, dynamics of the species.

In breeding season Siniya (eastern) population, move over large area, to Abu Dhabi (western) island, out to Musandam and at intermediate area. It might suggest the

same for Abu Dhabi population and do not expect a lot of genetic difference with all that movement of population mixing. Therefore, it is not expect of them to be genetically distinctive from each other. Therefore, our hypothesis is that the two subpopulations are genetically same and do not represent two distinct subspecies different.

1.3 Literature review

1.3.1 The Socotra cormorant, *Phalacrocorax nigrogularis*, an overview

The cormorants (Figure 1) are medium to large, voracious dark-colored long-necked aquatic birds with an expansive pouch for holding fish. They are characterized by long bills, mostly monochromatic plumage and webbed feet. cormorants distributed worldwide and are important predators of small fish in both marine and freshwater environments (Kennedy & Spencer, 2014).

Figure 1: Adult Socotra cormorant
(Retrieved from: http://www.hawar-islands.com)

Genetically the Socotra cormorant is a monotypic, regionally endemic seabird restricted to the Arabian Gulf, Gulf of Oman and south along the Omani shoreline into the Gulf of Aden and Red Sea (BirdLife International, 2017; Jennings, 2010) and it belongs to the *Phalacrocoracidae* family of the *Phalacrocorax* genus as shown in Table 1.

Table 1: The taxonomy of Socotra cormorant

Kingdom	Phylum	Class	Order	Family
Animalia	Chordata	Aves	Suliformes	Phalacrocoracidae

An adult Socotra cormorant is 76-84 cm long with a wingspan of 102–110 cm. They have almost black plumage overall, with a hint of brown to it. Their wings are lightly brown, and their face, gular region, legs and feet are black. The legs and feet are located at the rear of their body, and they have totipalmate foot; this gives them an advantage for swimming and paddling but not necessary for walking. During the breeding season, its fore-crown has a purplish gloss, and its upperparts have a slaty-green tinge. There are a few white plumes around the eye and neck and a few white streaks at the rump (Jennings, 2010). Similar species is the Great cormorant (*Phalacrocorax carbo*) which is fairly more substantial, with stouter bill, white face, and chin-patch (BirdLife International, 2017).

The cormorants have a long independent history, with >40 million years separating them from their closest relatives darters, also known as Anhinga, a fish-eating bird of warm inland waters (Gibbs et al., 2013). They are known to have a long flexible neck and slender, sharp-pointed bill. Based on recent molecular systematics study, the Phalacrocoracidae splits into seven clades as follows: *Microcarbo* (Micro-cormorant), *Poikiolocarbo* (red-legged shags), *Urile* (North-

Pacific cormorant), *Phalacrocorax* (Old-world cormorant), *Gulosus* (European shags), *Leucocarbo* (blue-eyed shags), *Nannopterum* (American cormorants).

Based on the systematic molecular phylogeny (Figure 2), the Socotra cormorant belongs to the *Phalacrocorax*. It consist of a large morphologically diverse group of species, falling into three subclades (Kennedy & Spencer, 2014) and is a sister clade to *Urile* (North –Pacific cormorants). The estimated time that *Urile* split from its sister clade, *Phalacrocorax*, was around 8.9–10.3 million years ago. Moreover, the Socotra cormorant diverged from the common ancestor of its closest extant relatives, the *Urile*, over four million years ago. This indicates that the Socotra cormorant had a long, independent evolutionary history (Kennedy et al., 2001).

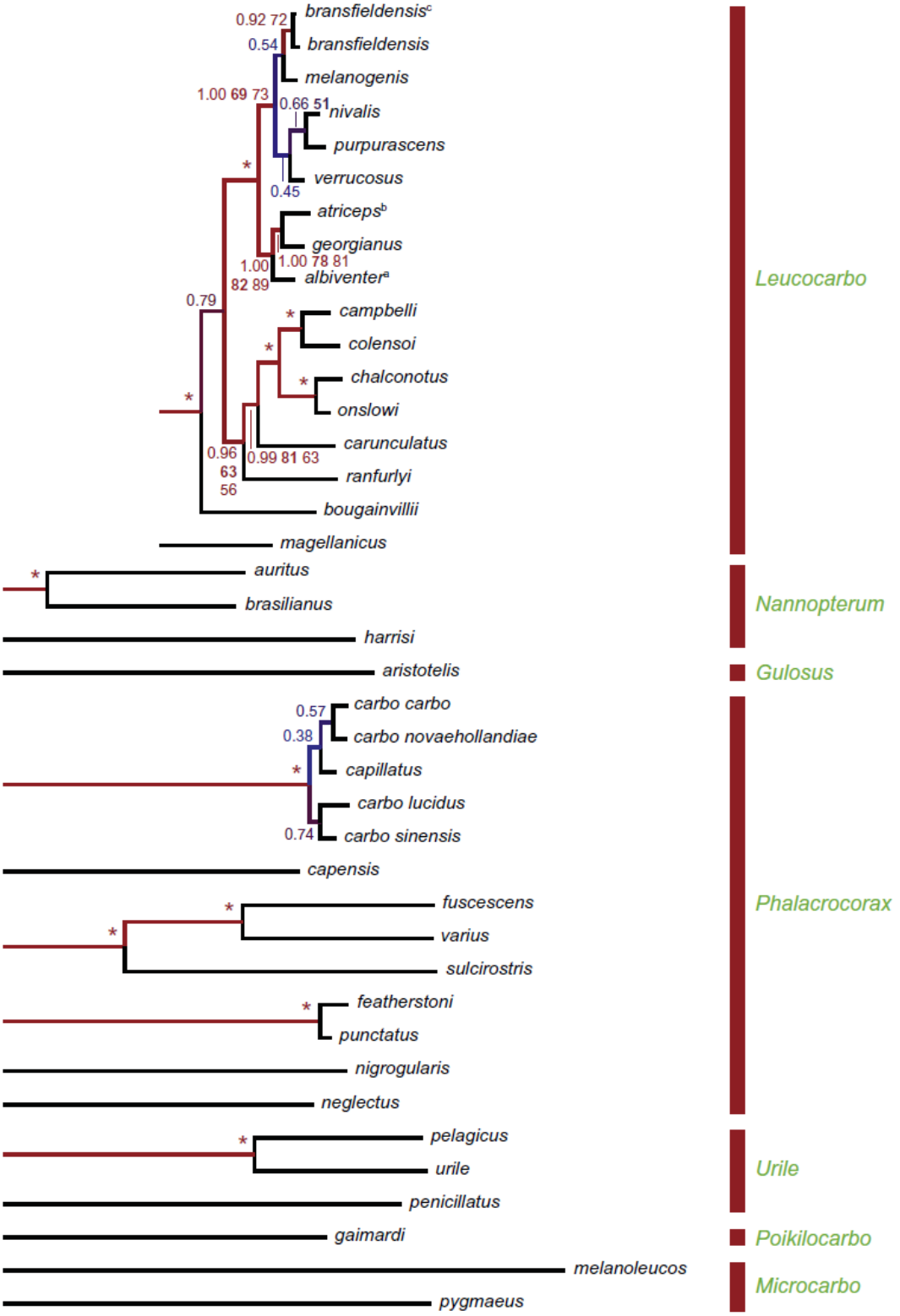


Figure 2: A phylogenetic tree of the world cormorants classification based on recent molecular systematics (Kennedy & Spencer, 2014)

The divergence of the Socotra cormorant possibly evolved within the last 10,000 years as consequence of the forming of the Arabian Gulf. It is likely that the species inhabited the Tigris/Euphrates river valleys and adapted to a marine habitat after the Gulf flooded or that it might have colonized the Gulf coming from another marine area, i.e., from the Indian Ocean (Kennedy *et al.*, 2001).

The Socotra cormorant range is restricted to the Arabian Gulf and the Arabian Sea and sometimes can extend up to the Red Sea (Figure 3). They exist in two subpopulations (Gallagher *et al.,* 1984): a Northern subpopulation that breeds on islands off the Arabian Gulf coasts of United Arab Emirates (UAE), Bahrain, Saudi Arabia, Qatar and possibly Iran (no recorded breeding confirmed since 1972). The Southern subpopulation is seemingly smaller, and breeds on islands off the coast of Oman in the Arabian Sea and in the Gulf of Aden off Yemen. Census of birds is 60,000 total (Aspinall, 1996b; Gallagher *et al.,* 1984; Jennings, 2000; Symens *et al.,* 1993). Approximately 34% of this population breeds in the United Arab Emirates (Aspinall, 1995; Jennings, 2010).

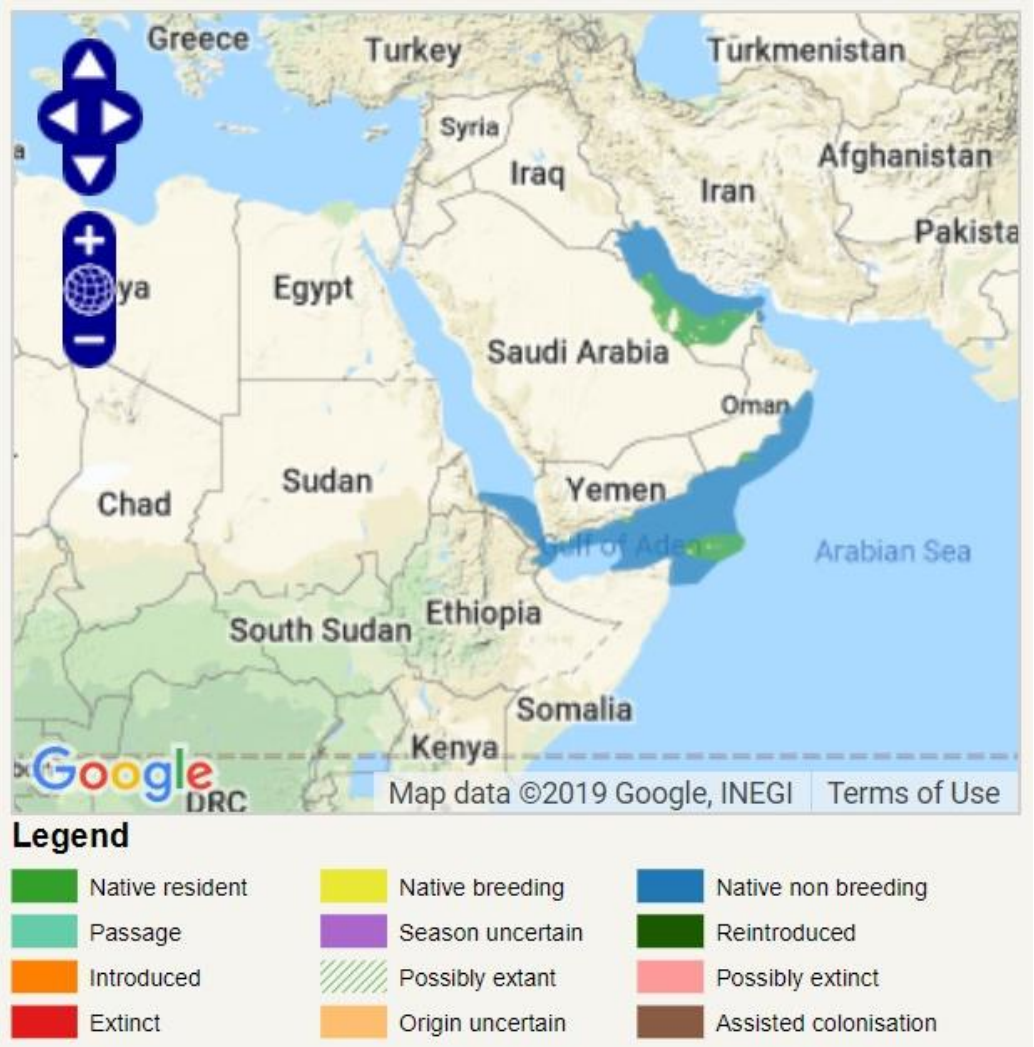


Figure 3: Socotra cormorant geographical range based on IUCN

The Socotra cormorant is a highly social aquatic species and tends to aggregate in large groups throughout the year (Johnsgard, 1993; King, 2004; Nelson, 2005). Socotra cormorants typically nest either in a small group of hundreds of nests or large, dense clusters of thousands of nests (Nelson, 2005). Colonies elsewhere in UAE fall into the small group of hundreds of nests, mainly due to declines caused by a disturbance and habitat destruction (Aspinall, 1995; Jennings, 2010; Symens *et al.*, 1993).

Clearly, they are highly communal and tend to hunt in a mass; their feeding frenzy may last up to 30 minutes offshore in a small or large group. Scrounging for food that occurs offshore in large groups, is thought to be communal hunting rather than cooperative (Gallagher & Woodcock, 1980; Nelson, 2005). They form a v-line or a long single file after hunting 2-50 meter above the wave, which is typical for marine species exhibit such behavior.

They roost in tight packs and occupy smallest footprint ground as possible to maximize shade to the feet on isolated sandbars or rocky cliffs (King, 2004). Socotra cormorants nest on islands with sandy or loose gravel substrate to allow construction of a bowl-shaped nest (Jennings, 2010) of gravel, compact sand and residual debris (Muzaffar, 2013). Many seabirds favor nesting under vegetation cover to avoid severe weather or predation (Pattern *et al.,* 2005; Olsson *et al*., 2001; Yorio *et al.,* 1995). Exposure to high temperature may affect reproductive performance, as a result of an effect on avian incubation behavior (Woinarski, 1973; Conway & Martin, 1999; Yasue & Dearden, 2006). On Siniya Island, nesting occurred on sandy-gravel substrates under either trees or open areas with scattered Haloxylon/Arthorocnemum shrub complex. The Siniya Island habitat consists of mixed desert scrub or gravel and plantations of *Acacia spp.* and *Prosopis juliflora.* Mangroves *Avecennia marina,* occur in patches, and Haloxylon *Arthrocnemum macrostachyum* scrub complex borders the periphery of the island (Muzaffar, 2012).

Breeding season is variable within years, egg-laying recorded in most months, but each colony is internally synchronized (Hoyo *et al.*, 1992). In UAE, it takes place between October to March. Some records state that pairs breed twice per season, producing 2-3 egg that are laid directly on the gravel, ledges or among

boulders (Aspinall, 1995; King, 2004; Jennings, 2010). While on Halne Island in the Arabian Gulf they begin to breed from January to March (Meinertzhagen, 1954). On the islands off Saudi Arabia, breeding can occur twice a year first in Spring, during April and May, and a second time throughout fall in September, October, and November (Bundy *et al.*, 1989). Abundance and availability of food may be a response to cause breeding to occur irregularly (Johnsgard, 1993).

Shallow coastal waters are of high importance for breeding and migrating of cormorants and other seabirds, primarily because of the vast schools of forage fish that migrate along the southern shoreline of the Gulf. Socotra cormorant primarily consume small pelagic shoaling fish and dive from the surface to depths more than 18 m (King, 2004). Adult cormorants feed on a different variety of species, but there is little information of the prey species they feed on (Johnsgard, 1993; Nelson, 2005). Although sardines (*Sardinella spp.),* scads (*Selar crumenophthalmus and Atule mate*), Silverside *Atherinomorphus lacunosus*, Spotted Half-beak *Hemiramphus* far and Streaked Rabbit-fish (*Siganus javus*) are probably among the species taken (King, 2004).

Phytoplankton phenology is often associated with distribution of seabirds globally (Ramirez *et al.,* 2016; McKnight *et al.,* 2013). The Arabian Gulf and the Gulf of Oman has a well characterized pattern of primary productivity (Nezlin *et al.,* 2007; Muzaffar *et al.,* 2017b) whereby high phytoplankton growth begins in October within the Gulf of Salwa, in the coastal areas near Siniya Island, around the Strait of Hormuz, and into the Gulf of Oman. This productivity, characterized by phytoplankton blooms, expands through much of the northern Gulf, bordering the coastlines of Iran (John *et al.,* 1990; Nezlin *et al.,* 2007). The of Socotra cormorants

abundance in and around the Musandam peninsula and Siniya Island, correspond to phytoplankton productivity and coincide with it (Muzaffar *et al.*, 2017b). It is presumed that the cormorants within the Gulf of Salwa also benefit from phytoplankton blooms during that same period, resulting in the large congregation of breeding cormorants in Hawar islands and the Saudi coastal islands (Jennings, 2010).

1.3.2 Conservation genetics of the Socotra cormorant

According to the IUCN Red List, the Socotra cormorant population trend is decreasing and have been listed as Vulnerable, for the last 20 years (Table 2). It is the most persecuted breeding bird in the emirates and is subjected to shooting, harvesting eggs and chicks, and nest destruction (EAD, 2016).

Table 2: IUCN Red list of the Socotra cormorant

Previously published Red List assessments:	2016 – Vulnerable (VU) 2012 – Vulnerable (VU) 2010 – Vulnerable (VU) 2008 – Vulnerable (VU) 2006 – Vulnerable (VU) 2005 – Vulnerable (VU) 2004 – Vulnerable (VU) 2000 – Vulnerable (VU) 1994 – Lower Risk/near threatened (LR/nt) 1988 – Near Threatened (NT)

Most recent count shows that the latest estimates of the total population size of breeding Socotra cormorants in the UAE is 50,000 to 60,000 breeding pairs. This represents about half of the global breeding population (Khan *et al.*, 2018). The Environment Agency Abu Dhabi estimated that there were around 51,812 nesting pairs during 2016–2017 in Abu Dhabi. Highest records were from three colonies: Rufayq, Ghagah and Butinah. However, largest existing colony in UAE is found on Siniya Island in Umm Al Quwain Emirate, with an estimated breeding population of 26,000-41,000 pairs during 2011-2016 (Muzaffar, 2013).

The populations zoogeographic range has been declining throughout the years, mostly in response to habitat loss from oil exploitation and disturbances at breeding sites. In the last 5 decades, it had been seen that the breeding ranged has contracted by about 50% (Khan *et al.*, 2018). Many colonies have decreased as a result of disturbance, degradation or straight out destruction of nesting sites because of oil exploitation. 80 % of the global population range within the Arabian Gulf contracted, while the are found in Gulf of Oman. Disturbance from recreational activities such as fishing and camping near the colonies, development such as oil industrialization, change in land use, and introduced predators at breeding colonies, all contributed to reduced nesting or abandonment of breeding colonies and relocation (Symens *et al.*, 1993; Aspinall, 1995; Aspinall, 1996a; Khan *et al.*, 2018; Jennings, 2010; Croxall *et al.*, 2012; Wilson, 2012). The Socotra cormorant have suffered from large-scale egg collection despite the fact that they are legally protected in UAE. In addition, there is evidence of persecution at some breeding and roosting colonies including active displacement, scaring, and killing is still occurring (Aspinall, 1996a; Khan *et al.*, 2018). Oil spills are a constant threat that can have severe effects on the species' numbers (Symens *et al.*, 1993).

Predation has potentially high adverse effects on breeding seabirds (Coulson, 2002). Feral cats introduced to some islands in the western Arabian Gulf had been reported. They are known to be exceptionally adaptive and can survive and expand their populations relatively easily. Hunting by feral cats *Felis catus* and native red foxes *Vulpes vulpes* constitutes a significant threat to Siniya Island colony in Umm Al Quwain (Muzaffar, 2013). A systematic study conducted during 2011/2012 breeding season on Siniya Island examined the predation rate on the breeding birds. Approximately, about 2000 cormorants were hunted by the red fox (*Vulpes vulpes*) during the season (Muzaffar, 2013). Red foxes also hoard eggs (Muzaffar, 2013) further reducing reproductive performance. Another cause of mortality of Socotra cormorant mortality is plastic debris and entanglement on fishing equipment such as nets and fishing lines. Counts of birds dead with fishing line or fishing hooks embedded in their skin or feet has been recorded (Muzaffar, 2013), as well regularly drowning in fishing traps at various depths (King, 2004).

The close proximity of colonial breeding birds helps the transfer of microbes and parasites. The re-use of breeding site year after year may encourage the buildup of parasites, which can act as the vectors for viruses and other microbe (Coulson, 2002). *Ornithodoros muesebecki* is a soft tick that has reported from other seabird colonies in the Arabian Gulf (Hoogstraal, 1970) including Siniya Island (Muzaffar, 2013). High tick abundance can cause breeding failures due to chick mortality, colony abandonment (Duffy, 1983) and disease harboring parasites. The Socotra cormorant is susceptible to ectoparasites; locals associate it with lice because they harbor parasites. However, the relationship of ticks found on the cormorants have not been studied and needs further investigation.

During the outbreeding, season the Socotra cormorant moves to cooler, deeper water in Musandam, during the summer months, where there are many that roosts on the rocky cliffs edge (Wilson, 2012). In a satellite telemetry study, it proves that post-breeding dispersal of Umm Al Quwain population (Siniya island colony) is on North, along with the UAE shoreline out through the Strait of Hormuz and around the Musandam peninsula of Oman (Muzaffer, 2014) also, travelling westwards into Abu Dhabi waters (Muzaffar *et al.*, 2017b). During this time, cormorants apparently roosted on different islands and forage in the general area (Muzaffar *et al.*, 2017b). Whether some of the birds nesting in Abu Dhabi islands subsequently migrate to Siniya Island is not known (Khan *et al.*, 2018). However, this observation does not suggests that the two subpopulations of the cormorants in the UAE are not mixing during breeding season and might considered somewhat different population. areas.

1.3.3 The study site

The United Arab Emirates is located in the Northeastern end of the Arabian Peninsula on the Arabian Gulf, bordering Oman to the east and Saudi Arabia to the south, as well as sharing sea borders with Qatar and Iran. The UAE is a federation of seven emirates; they are Abu Dhabi , Ajman, Dubai, Fujairah, Ras al-Khaimah, Sharjah, and Umm al-Quwain. The UAE occupies an area of 83,600 Km^2 and lies between 22°50′ and 26° north latitude and between 51° and 56°25′ east longitude. The largest emirate, Abu Dhabi, accounts for 87 percent of the UAE's total area (67,340 square kilometers). The smallest emirate, Ajman, covers only 259 Km2. UAE is surrounded by two coasts: The Arabian Gulf that covers 650 km and 9 Km of Gulf of Oman. Between them, is the Strait of Hormuz, the coast mostly consists of salt pans that extend quite far inland. Out in the sea, the country has varied of islands,

coral reef and shoals on its water on both Gulfs. The terrestrial habitat of the UAE consists mainly of flat lands, dunes large desert area and extensive salt flats at the coast. The Hajar Mountains are parallel to the East coast of the United Arab Emirates and continue South-eastward into Oman, along with the UAE border, it rises above the surrounding on about 2000 meters. The climate can be described as bi-seasonal, Mediterranean clime, it is characterized by high temperature and low rainfall with high humidity at the coasts.

The UAE weather is warm and sunny during the winter and hot and humid in summer months. Winter daytime temperature norms at 26°C, while at night it can be relatively cool, temperature between 12-15°C on the coast, and in depth of the desert and into higher mountains temperature can go below 5°C. Local Northwesterly winds often develop during the winter, bringing cooler windy conditions. Summertime temperatures are in the mid-40s but can be higher inland. Humidity averages between 50 and 60 percent, touching over 90 percent in summer and autumn in coastal areas. Inland it is far less humid. The UAE landscape can be categorized as follows: inland sand sheets and dunes; piedmont alluvial and inter-dunal plains; mountains and wadis; coastal sand sheets with dwarf shrub vegetation; and coastal and inland Sabkha. The Arabian Gulf coastline is predominantly flat, with warm, shallow waters. Natural deep-water harbors occur mainly on the eastern coast. Many of the UAE's islands are alongside the coastline, they are at least 200 of them, and most are in Abu Dhabi's Emirates most of which are rather small and flat. However, Zirku Island attains a maximum altitude of 160 meters, even though it is only about 8 square kilometers in size. The largest island in Abu Dhabi is Abu Ab'yad; it is 100 Km west of the capital Abu Dhabi, and 12 Km from the coastline. It

has a sandy, rocky terrain, and a mangrove surrounds its borders. The two island is which the study took place were the following:

Siniya Island (25°36’20.63” N;55°36’28.85” E) in Umm AL Quwain, which is one of the UAE islands. It is 1 Km away from the coastline, and it has a total area of 90 Km^2. The surrounding water contains many fish species such as *Lethrinus Kallpoterus*, Scaled Mullet, Greasy grouper, mackerel, and many more. The Siniya island is one of the biggest island visited by tourist for its greenery that attracts many birdlife such as Pelagic seabirds, flamingo, cormorant also wild gazelles, rabbits and for archaeological sites from pre-Islamic time. The island holds the largest Socotra cormorant colony in the UAE; it is estimated to have a breeding population between 28,000-35,000 pairs.

The second site was Umm Qasr island (24°23.5”N;52°46.6”E) in Abu Dhabi Emirates, it is a scenic, a small flat, sandy and circular desert island. It Cover over 0.5 km at its maximum width, with sandy beach to south and a low 3-4 metres of rocky cliffs margin to the north, west and east side. Other species occupy and nest the island during the cormorant breeding season, including Osprey, *Pandion haliaetus.* Common predators of the Socotra cormorant are the Steppe Gull, Caspian Gull and Heuglin’s Gull nesting on the island too (Wilson, 2012). The latest estimate of the Socotra cormorant during the 2015-2016 report is less than 5000 thousand breeding pairs in Umm Qasr island (EAD, 2016). Umm Qasr island itself had been a historical breeding area for Socotra cormorants, where the breeding numbers grew from 196 pairs in the mid-1990s to 4000 pairs in 2006, after which the island gradually became inactive due to high levels of disturbance with no breeding recorded in 2016 (Khan *et al.*, 2018).

1.3.4 Genetic diversity analyses of mtDNA and nuclear marker

Determining the species identity and distribution can be done through DNA analysis, such analysis uses two genetic markers to identify species. Researchers can identify regions that are variable among species but conserved within species; they can provide diagnostic markers, such as microsatellite's, that can give more than individual specific genotype; it can display Mendelian codominant expression, and identify alleles inherited by the individual from parents. From that, it can calculate allele's frequencies, heterozygotes' and other genetic variation measures. In return, these measures can estimate population structures, connectivity among the population, loss of genetic variation that can lead to inbreeding depression and drastic population reduction in past and present by detecting bottleneck (Cristescu *et al*., 2010)

1.3.4.1 SNP as nuclear DNA markers

Single nucleotide polymorphism known as SNP's is most common form of genetic difference among individuals and because they are distributed approximately uniformly along the chromosomes (Hartl & Jones, 2006) SNP is a single base pair at a particular nucleotide site may differ from one individual to other. Each SNP represents a difference in a single DNA building block, called a nucleotide. For example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA. They occur normally throughout an individual's DNA; they occur almost once every 1,000 nucleotides on an average. There can be somewhere between 4 to 5 million SNP in a human genome. These variation found in the DNA can act as biological markers and assist scientist locate genetic differences (NIH, 2019).

1.3.4.2 Mitochondrial DNA markers

Biologists determine and then compare the mtDNA sequences from different individuals or species. Data from the comparisons are used to create a network of relationships among the sequences, which on the other hand provides an estimate of the relationships among the individuals or species from which the mtDNAs taken. Mitochondrial sequencing has been used to illustrate the genetic diversity and molecular evolution in many species. The mitochondrial control region is the main non-coding area of the mitochondrial genome and contains the mitochondrial origin of replication and transcription. It does not encode protein sequence, but it is highly polymorphic in the mtDNA genome (Stoneking e*t al.,* 1991). For small or degraded samples, mtDNA markers are used for species identification because it has multiple copies in each cell, typically 100-1000 or more copies, compared to Nuclear DNA (Tomar, 2010).

With the information on the genetic application of DNA markers of wildlife population biology, estimating abundance, connectivity, quantifying genetic variation, percentages and determine gene that may affect fitness (Mills, 2013). Mitochondrial DNA sequencing that had been used to explain the origins of many modern species. The studies on the structure and function of mtDNA become highlights in the research area of molecular evolution, classification, population genetic analysis, relative identification and quantitative traits loci. Mitochondrial DNA contains 37 genes, all of which are essential for normal mitochondrial function. In most vertebrate species, mitochondria appears to be inherited through the maternal lineage. The new organism has only the mitochondria which came from the egg of its mother. Therefore, mitochondrial DNA does not get shuffled (go through

recombination) every generation; unlike nuclear DNA, it is presumed to change at a slower rate, which is useful for the study of organism evolution (El-Mahdy, 2012).

The mtDNA is a powerful tool for tracking ancestry through maternal lineage and had been used in this role to trace the ancestry of many species back hundreds of generations. The rapid mutation rate and low effective population size in animals make mtDNA useful for assessing genetic relationships of specific individual or groups within a species and also for identifying and quantifying the phylogeny among different species, provided they are not too distantly related (Brown, George & Wilson, 1979).

Chapter 2: Materials and Methods

2.1 Sample collection

A total of 80 Socotra cormorant samples were collected from Abu Dhabi and Umm Al Quwain, during mid breeding season of 2015 and 2016. The two islands of which the samples were collected from are Siniya island in Umm Al Quwain, a total of 41 samples were collected and 39 collected from Um Qasr island, Abu Dhabi. But only 40 samples were analyzed (20 from Abu Dhabi and 20 from Umm AL Quwain). The gathered samples were obtained mostly from dead chicks and adult Socotra cormorant; they consisted of different types of tissues such as skin, feather, muscle and liver. The samples were bagged, tagged with a number and initials of the location, date and coordinates. Then, stored in deep freezers, 18°C, to reduce tissue degeneration, whereas the feathers kept at room temperature.

2.2 DNA extraction process

First step of analyzing genomic DNA is to isolate DNA from tissues or cell using the combination of physical and chemical methods. All of the DNA extraction processes were performed according to the manufacturers protocol, using the QIAamp Tissue Kit (QiagenTM). In this protocol, the silica-based DNA purification was used in convenient spin-column and 96-well-plate formats. The DNA was purified from tissue or blood and it allowed for more efficient method of DNA extraction from nucleated cells. The basic principle of the silica matrix can be described in four stages. First step is to lyse the sample with reagent like buffer ALT with proteinase K, the buffer contains Sodium Dodecyl Sulfate (SDS) detergent that disrupts cell membrane and dissociates protein DNA complex. Meanwhile, the Proteinase K solution degrades protein including DNA scaffolding proteins and other

protein derbies. The second step is to isolate DNA from cells by adding the Lysis buffer and Ethanol. The AL buffer is employed to lyse the cells and inactivate nuclease. The AL contains Guanidium Hydrochloride Chaotropic salt which removes water from the hydrate molecules in the solution and renders DNA that is susceptible to bind to silica matrix in the spin column / plate. DNA is not soluble in alcohol; therefore, ethanol is used to precipitate DNA out of the fluid suspension. Therefore, it increases DNA affinity to bind to silica spin column for further elution of DNA through spin column. In the third stages, the aim was to produce clean purified DNA bound to a silica matrix; using two wash buffers AW1 and AW2, each having two different concentration of ethanol, where AW1 has highest concentration followed by AW2. This step ensures that the DNA remains bound to silica matrix column. The AW buffer contains Tris[tris (hydroxymethyl) amonimethane $(HOCH_2)_3CNH_2$] and Ethylenediaminetetraccetic acid (EDTA). The EDTA removes debris by chelation metal ions, attaching molecules to itself by means of ionic charge. The fourth and final stage of DNA extraction is elution. The nuclease free water was used to change the binding condition through rehydrating the DNA and removing hydrogen bonds and eluting the DNA from column on 60°c and at a PH of 8.3 for uttermost efficiency.

2.3 Quantification process

DNA quantification is an essential measure for molecular analysis which is later followed by amplification. There are several methods used to quantitate the concentration of DNA in a solution, the most common method of DNA quantification used is spectrophotometric quantification. We used a tabletop spec called a NanoDrop that requires only 1 µl of a sample for quantification.

The NanodropND-100 is a spectrophotometer that uses a patent sample retention system, which holds 1µl of sample without the need of traditional system devices. It uses fiber optic technology and surface tensions; the 1µl of sample is placed between two optical surfaces that set the path length in vertical orientation. The direct joining of the sample to the two optics of spectrometer removes interference caused by incident light and transmitted light passing through containment wall of traditional cuvette, microcell cuvettes and capillaries. Nucleic acids absorb light at a wavelength of 260 nm. If a 260 nm light source shines on a sample, the amount of light that passes through the sample can be measured, and the amount of light absorbed by the sample can be inferred. For double-stranded DNA, an Optical Density (OD) of 1 at 260 nm correlates to a DNA concentration of 50 ng/µl, so DNA concentration can be easily calculated from OD measurements.

2.4 Amplification process

2.4.1 DNA amplification

To study or detect an individual gene or specific DNA region, it is often required to obtain a large quantity of nucleic acid for study. Instead of isolating single copy of target DNA from large number of cells, it is more useful to create multiple copies of a target from a single molecule of DNA or mRNA, via an in vitro amplification method. Polymerase Chain Reaction (PCR) is the most common DNA amplification method in molecular biology and it is both a thermodynamic and an enzymatic process. PCR process is a process of amplifying low, small quantity or quality of DNA samples. The steps can be simplified in three steps. It starts first with extracting DNA from the sample and preparing the DNA for the PCR reaction. To the isolated DNA following are added,

a. Primers which are again DNA fragments of about 20 bp, that attach to each strand of the DNA at specific location outside the locus to be amplified.

b. Synthetic nucleotides (bases), the building block on DNA strands.

c. Taq polymerase, an enzyme that attaches the nucleotides to synthetic DNA strand.

The reaction requires following ingredients in certain amount to produce a amplified final product, Table 3 shows the necessary components for PCR and their quantities per reaction:

Table 3: Polymerase chain reaction components and quantity per reaction

PCR product	**Quantity (µl)**
DNA	3
Taq PCR Master Mix	12.5
Primer F (10pM)	1
Primer R (10pM)	1
Water	7.5
Total	25

After preparation of the PCR reaction, the next step is placing the samples in PCR thermal cycler (also called thermocycler). It employs precise temperature control and rapid temperature changes to conduct the polymerase chain reaction (PCR) and create millions of copies of a DNA sequence for use in a wide array of downstream procedures including cloning, sequencing, expression analysis, and genotyping.

For each PCR cycle, the thermal cycler machine first heats the DNA up to about 94°C to denature the two strands of DNA, and then cools down at about 55°C -

65°C to anneal the primers to their target sequences on each strand. Next the temperature increases slightly to about 72°C to extend the primer into double-stranded DNA, with the Taq polymerase attaching the nucleotides to their respective complementary bases on the other strand. One PCR cycle doubles the number of DNA strand for each targeted region, this cycle is repeated about 30 to 40 times, nearly doubling each cycle, ending with millions of copies of each original DNA strand. Thermal protocol as shown in Table 4.

Table 4: General thermal cycling parameters used in the study

Cycling conditions for amplifying PCR products			
Step	**Temperature (°C)**	**Time (min)**	**Number of cycles**
Initial denaturation step	95	5	1
Denaturation	95	0.5	25~40
Annealing	52	0.5	
Extension	72	1	
	72	7	1
Hold	4	Indefinite	none

In this study, one Nuclear and three mitochondrial regions was targeted to examine the differentiation and diversity of the two populations, Table 5 represents the primers used to study the amplified regions.

Table 5: Primers employed in this study

Gene	Primer name	Primer sequence (5"-3")	Source
PARK7	PARK.2F	GCAGGCCTRRCTGGAAAAGARCC	Kimball, 2009
	PARK.3R	TTCTGAGCTCCWAGRTTACC	Kimball, 2009
COI	AvMiF1	CCCCCGACATAGCATTCC	Lijtmaer et al., 2012
	AvMiR1	ACTGAAGCTCCGGCATGGGC	Lijtmaer et al., 2012
12s	L1836	GCTTAAAACTCTAAGGAC	Paterson et al., 1995
	H2024	AGAAAATGTAGCCCATTTC	Paterson et al., 1995
ATPase 6	L9265	ATGAACC TAAGCTTCTTCGACCAATT	Desjardins & Morais, 1990
	H9922	TAGGAGTGTGCTTGGTGTGCCATT	Desjardins & Morais, 1990

2.4.2 Run agarose gel of PCR product

PCR final products were checked on 1.8 % agarose 1×TBE electrophoresis, the 1.8 % is good resolution for small DNA of 0.2–1kb fragments. After separation, the DNA molecules can be visualized under UV light after staining with an appropriate dye. Most common dye used to make DNA or RNA bands visible for agarose gel electrophoresis is ethidium bromide, usually abbreviated as EtBr. It fluoresces under UV light when intercalated into DNA (or RNA). By running DNA through an EtBr-treated gel and visualizing it with UV light, any band containing more than ~20 ng DNA becomes distinctly visible

2.5 Sequencing preparation

2.5.1 Clean up of PCR products

Exonuclease I (Exo) - Shrimp Alkaline Phosphatase (SAP) is a one-step enzymatic treatment of PCR products to eliminate unincorporated primers and dNTPs so that they cannot interfere with downstream sequencing reactions. The exonuclease I removes leftover primers, while the Shrimp Alkaline Phosphatase removes the dNTPs that were not consumed in the reaction. It starts with adding 2 µl of the enzyme (ExoSAP) to a 5 µl PCR sample and incubate at 37^{o}C for 15 minutes to activate the enzyme and degrade leftover primers and nucleotides, then incubate at 80^{o}C for 15 minutes to inactivate the ExoSAP enzyme. The final product of 7 µl (ExoSAP + PCR product) ready for downstream sequencing reactions without any additional manipulation.

Cycle sequencing (chain termination with fluorescence dye labeled nucleotides DNA cycle sequencing is a core technique in molecular biology allowing analysis of fmol quantities DNA template. The enzymatic dideoxy chain termination method of Sanger relies on the linear amplification of a single-stranded template DNA using a single primer and thermostable polymerase. The synthesis of the complementary DNA strand starts at the specific priming site and ends with the incorporation of a chain-terminating dideoxynucleotide triphosphate (ddNTP). This generates a multitude of fragments terminated within the desired length of the sequence. By using the four different ddNTPs in four separate reaction vials, a set of extended primer strands terminated at each A, C, G, and T are obtained. The nucleotides are tagged with fluorescence dye and Table 6 corresponds the color to its

nucleotide. When these fragments are separated on a suitable gel matrix the sequence information can be read from the order of the bands.

Table 6: Fluorescence painting of nucleotide

A➔ Green
G➔Black
C➔Blue
T➔Red

The procedure starts with preparing reaction mix of the following ingredient in Table 7 and mixed properly.

Table 7: PCR reaction mix

Reaction Mix	**Quantity (µl)**
5X seq. buffer	2
Ready reaction mix	4
Exosap PCR amplicon	3
Deionized water	8
Primer (we use either Forward or Reverse primers, 10 pM	3
Total	**20**

Then placed in thermal cycler in the following conditions as shown in Table 8 and the final product must be stored in the dark.

Table 8: PCR thermal cycling profile

TO THERMAL CYCLING (PCR)	
PCR Cycles conditions	
96°C	1 min
96°C – 10 sec 50°C – 5 sec 60°C – 4 min	X 25 cycles
4°C	α

2.5.2 Clean up step to remove excess stain

The sample must be cleaned from excess stain using Ethanol-EDTA precipitation, to get a clean sample of DNA. It start with adding 60 µl absolute Ethanol and 5 µl,125 mM EDTA (Ethylenediaminetetraacetic acid) , and mixing it properly, then added 20 µl of the cycle-sequenced sample and incubate in the dark for 15 minutes. Next, transfer the mixture into 1.5 ml microcentrifuge tube and centrifuge at 14,000 RPM for 30 minutes at 4^{o}C. Once done with centrifuging, discard the supernatant from the sample paying attention not removing the pellet, which is not visible but there at the base of the microcentrifuge tube. Making sure there is no Ethanol-EDTA precipitation residue, then added 120 µl of 70% Ethanol and centrifuge at 14,000 RPM for another 15 minutes at 4^{o}C. Again, carefully removed ethanol without disrupting DNA, allowed excess ethanol to evaporate by

placing it in a heat block at about ~40^{o}C for 5 minutes. Then added 10 µl of HiDi Formamide and mix properly, to prevent the double strands formation, because sequencing requires DNA as single stranded. The samples were loaded in a sequencing rack in an orderly manner, then closed the rack top and centrifuge briefly. Once done the rack was placed in thermal cycler to denature at 95^{o}C for 5 minutes, then at 4^{o}C for 2 minutes. Once done the rack is placed in a sequencer, Applied Biosystems™ 3500 Series Genetic Analyzer for sequencing the samples.

2.6 Data analyses

All sequenced data were checked and aligned using the program Sequencing Analysis Software v6.0 from Applied Biosystems ™ while the analyses and edit of the sequence data was done by using MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets (Kumar, Stecher, & Tamura 2015). Representative sequences of COI (KM066529), PARK7 (KM066326), ATPase 8&6 (AF410794), and 12S (AF410797), of *Phalacrocorax nigrogularis* were deposited in the NCBI GenBank.

2.6.1 Population structure

To construct haplotype network PopArt (Population Analysis with Reticulate Trees) version 4.8.4 (available from http://popart.otago.ac.nz/index.shtml) program was used. It is a comprehensive analysis and visual of the relationships among DNA sequences within a population or species. Median joining [MJ] method was used for constructing the haplotype network networks. (Bandelt, Forster, & Röhl, 1999)

The evolutionary history was inferred by using Maximum Likelihood method based on the Tamura-Nei model (Tamura & Nei, 1993) analysis was conducted in MEGA7 (Kumar, Stecher , & Tamura, 2016).

Principle component analysis (PCA) was obtained from the program PAST Version 3.25 (PAlaeontological Statistics) includes several methods for multivariate data analysis specific to paleontology and biology. Principal components analysis (PCA) is a procedure for finding hypothetical variables (components) which account for as much of the variance in the multidimensional data as possible (Davis, 1986; Harper, 1999). These new variables are linear combinations of the original variables. PCA has several applications, one of them are simple reduction of the data set to only two variables, for plotting and clustering purposes. A different representation of genetic distances between individuals or populations and ideally it should Summarize briefly the phylogenetic tree.

2.6.2 Estimating of genetic diversity

To examine the genetic variation several genetic parameters was computed for each, including haplotype, nucleotide diversity and neutrality, Tajima's D (Tajima, 1989a) using DnaSP 5.0 (Librado & Rozas, 2009). Haplotype diversity refers to the frequency and number of haplotypes in the population and was analyzed using the default settings in DnaSP 5.0. Nucleotide diversity estimates the average pairwise differences among sequences, based on all sites. Tajima's D test is based on the differences between the numbers of segregating sites and the average number of nucleotide differences to test for changes in population size.

2.6.3 Population differentiation, structure and gene flow

To determine the level of genetic differentiation among populations, gene flow and genetic differentiation tests were performed with the program DnaSP 5.0. It computes some measures of the extent of DNA divergence among populations, and from these measures it computes the average level of gene flow.

Overall values of genetic differentiation between each pair of populations were quantified using pairwise F-statistics, the population genetic differentiation (F_{ST}) and gene flow between pairwise populations. The levels of genetic differentiation can be categorized as F_{ST} >0.25 (great differentiation), 0.15 to 0.25 (moderate differentiation), and F_{ST} <0.05 (negligible differentiation) (Wright, 1978).

Chapter 3: Results

3.1 Mitochondrial DNA analysis

3.1.1 Genetic diversity analysis

Genetic diversity indices of Um Qasr (UQ) and Siniya (S) Socotras cormorant population based on mtDNA sequence data

Table 9: Genetic diversity measures based on the mtDNA

Population	n	S	Eta	h	*hd*	π	θ	*D*
UQ	20	4	4	5	0.55789	0.00066	0.00107	-1.11111 Statistical significance: Not significant, P > 0.10
S	20	2	2	3	0.57368	0.00060	0.00054	0.26026 Statistical significance: Not significant, P > 0.10
Combined	40	4	4	5	0.5436	0.00065	0.00090	-0.66563 Statistical significance: Not significant, P > 0.10

Note: n, number of samples,S, number of polymorphic sites; Eta, total number of mutation; h, number of haplotype; hd, haplotype diversity; π, nucleotide diversity; k, Average number of nucleotide differences; θ, theta (per site) from Eta; D, Tajima's D

Eight parameters including number of polymorphic sites (S), total number of mutations (Eta), number of haplotype (h), haplotype diversity (hd), nucleotide diversity (π), and (θ) theta (per site) from Eta, and (k) average number of nucleotide differences were used to measure the DNA haplotype and genetic diversity in Table 9. The DNA polymorphism was based on the 1050 bp of mtDNA sequence data. The haplotype diversity (hd) in Siniya(hd=0.57368) was higher than Um Qasr (hd=0.55789), but the difference is insignificant. Nucleotide diversities and theta (per site) from Eta in Um qasr were (π=0.00066 and θ=0.00107), respectively. Nucleotide diversities and theta (per site) from Eta in Siniya were (π=0.00060 and θ= 0.00054).

Um Qasr theta (per site) from eta (θ= 0.00107) is half as large as nucleotide diversity (π=0.00066) and nearly two times as large than Siniya theta (per site) from eta (θ= 0.00054) and nucleotide diversity (π=0.00060). The haplotype diversity (hd), nucleotide diversity (π), and theta (per site) from eta (θ) of the pooled data were 0.5436, 0.00065 and 0.00090, in the order given.

3.1.2 Neutrality tests

Tajima's D test used to execute neutrality test (Table 9) and the results revealed that the positive D value of Siniya population might signifies low levels of both low and high frequency polymorphisms, indicating selection maintaining variation or recent population contraction. Um Qasr Tajima's D value is negative it might signify an excess of low frequency polymorphisms, indicating population size expansion (e.g., after a bottleneck) or selection or moving variation (e.g. selective sweep).

3.1.3 Genetic differentiation between populations

The genetic differentiation within the population was significantly low among populations (*Fst* 0.04960), and the gene flow value (Nm) among the population is 4.79.There is negligible differentiation between the populations and an extensive gene flow.

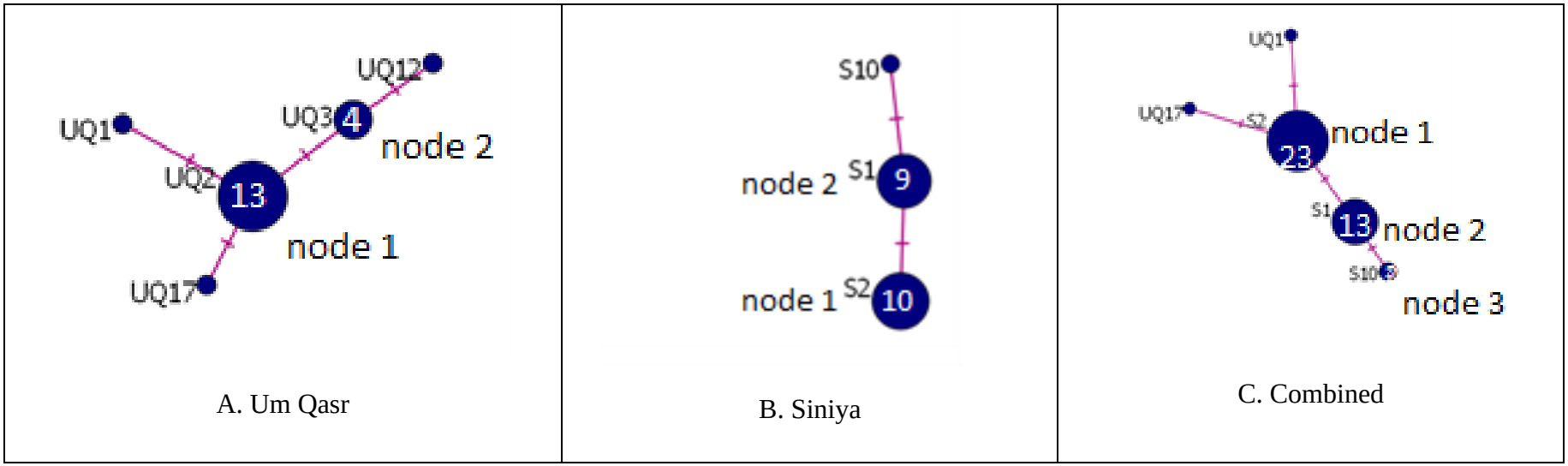
3.1.4 Population structure

The haplotype network (Figure 4) reflects the results of the diversity indices presented above. A total 5 haplotypes were identified from 20 samples in Um Qasr population,4 variable site and 1 parsimony informative site (Figure 4A). Node 1 includes 13 individuals with similar haplotypes, and 3 branches extending out, 2 of

the 3 branches are haplotype with singular nucleotide change. Third branch out into two nodes one with 4 individuals sharing the same one mutation step and a single haplotype stretches out of it with another mutation step. The Siniya population shares three haplotypes, 2 variable site and 1 parsimony informative site (Figure 4B) clustered into main two nodes. Both population shares the two-haplotype nodes with an exception of one haplotype branching out from node 2. Combining 40 haplotypes there is 4 variable site and 2 parsimony informative sites (Figure 4C), it shows that they share three common haplotype nodes with two haplotypes bifurcates from node 1 and they are from Um Qasr population. Third node is the smallest with only two haplotypes shared from both populations.

3.1.5 Principal component analysis

The two dimensional dispersion of the Um Qasr individuals (Figure 5A) support the idea that the population is alike, little variation is observed, with exception of two individuals. The Siniya population (Figure 5B) shows less variation and they are closely related. Combining the two population shows that the populations are clustered together, and lack any variation that might set the two populations from each other.



A. Um Qasr

B. Siniya

C. Combined

Figure 4: Median joining Haplotype network of mtDNA data sequence

Note: Number in the circle are number of haplotypes copies observed individuals. Major circles represent predominant haplotypes. A branch represents a single nucleotide change

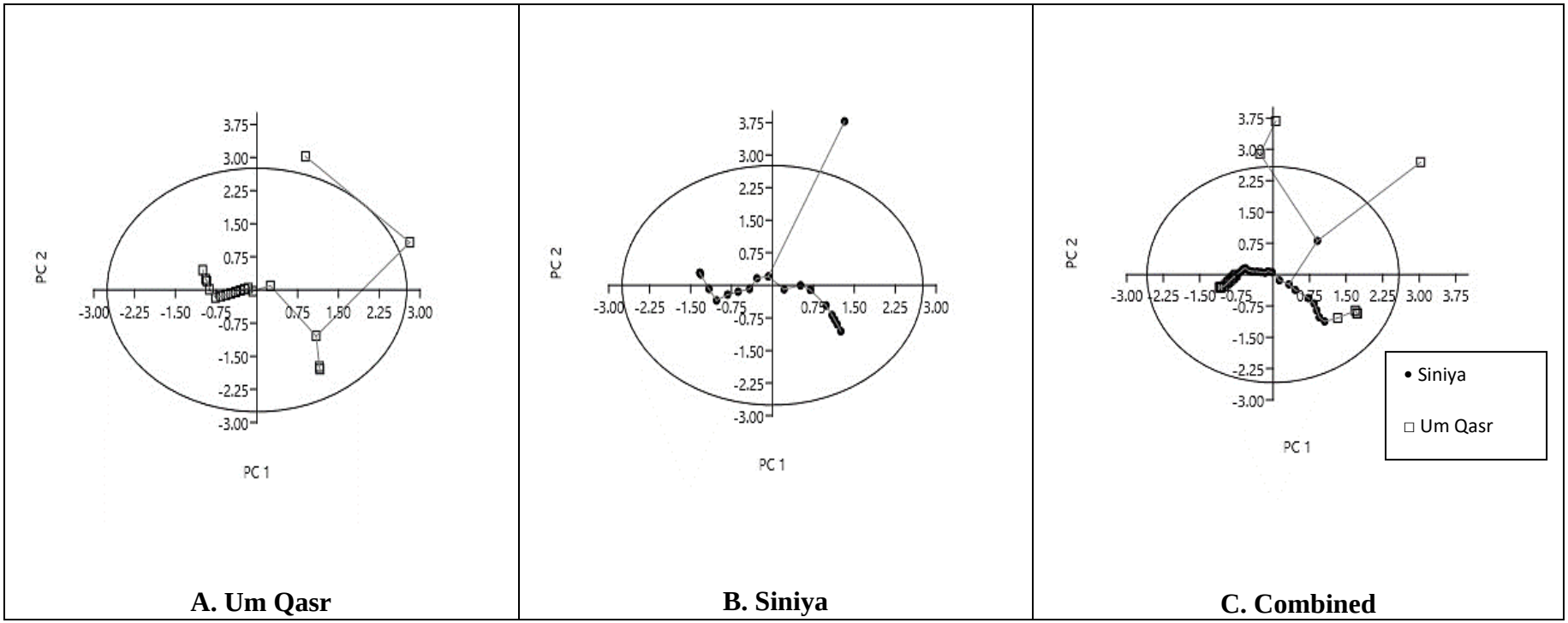


Figure 5: Principle component analysis of the mtDNA sequence data, two-dimensional scatter plots of the first two principal components

3.2 Nuclear DNA analysis

3.2.1 Genetic diversity analysis

Genetic diversity indices of Um Qasr (UQ) and Siniya (S) Socotras cormorant population based on nuclear marker sequence data (Table 10).

Table 10: Genetic diversity measures based on the nDNA

Population	**n**	**S**	**Eta**	**H**	***hd***	***π***	***θ***	***D***
UQ	18	3	3	3	0.21569	0.00054	0.00140	-1.71304 Statistical significance: Not significant, 0.10 > P > 0.05
S	20	1	1	2	0.10000	0.00016	0.00045	-1.16439 Statistical significance: Not significant, P > 0.10
Combined	38	4	4	4	0.15363	0.00034	0.00153	-1.88258 Statistical significance: *, P < 0.05

Note: n, number of samples,S, number of polymorphic sites; Eta, total number of mutation; h, number of haplotype; hd, haplotype diversity; π, nucleotide diversity; k, Average number of nucleotide differences; θ, theta (per site) from Eta; D, Tajima's D

The genetic diversity indices were based on 622 bp of nuclear DNA sequenced data. The highest DNA polymorphism showed in Um Qasr while the lowest polymorphism was found in Siniya. The haplotype diversity (hd), nucleotide diversity (π), and theta (per site) from the total number of mutation (θ) in Um Qasr were 0.21569, 0.00054, and 0.00140, respectively, and in Siniya the haplotype diversity (hd), nucleotide diversity (π), and theta (per site) from the total number of mutation (θ) were 0.10000, 0.00016 and 0.00045. The Um qasr population theta (per site) from eta (θ= 0.00140) three times as large as nucleotide diversity (π=0.00054) and three times larger than Siniya theta (per site) from eta (θ=0.00045) and

nucleotide diversity (π=0.00016). When combining set of 38 samples a total of 4 haplotypes were identified: three of the four are shared among the populations, and one is unique to Um Qasr. Haplotype diversity (hd), nucleotide diversity (π), and theta (per site) from eta (θ) of the population combined were 0.15363, 0.00034 and 0.00153.

3.2.2 Neutrality tests

Tajima's *D* test used to carry out neutrality test (Table 10) and the results of all values were negative which might indicate an excess of low frequency polymorphisms, suggesting that there is a population size expansion or selection or moving variation.

3.2.3 Genetic differentiation between populations

The genetic differentiation within the population was zero (*Fst*= 0.0000); neither genetic differentiation or gene flow (Nm) were detected, suggesting that the populations is one.

3.2.4 Population structure

In Um Qasr haplotype network (Figure 6A) there are 3 haplotype, 3 variable sites and 0 parsimony informative site, and 16 individuals share the same haplotype with two single haplotype with unique mutation. In the Siniya haplotype network (Figure 6B) there are 2 haplotype group and 1 variable site and also, 0 parsimony informative site, and 19 individual share same haplotype apart from 1 haplotype with one mutation. In Figure 6C, 38 samples are combined, and there are 4 haplotypes,3 variable sites and 0 parsimony informative site., 35 individuals share a common haplotype, with three unique haplotype with one or two mutation steps,

these single haplotypes can be an indication of rapid population expansion. The size of the circle is proportional to the frequency of that haplotype in the sample.

3.2.5 Principal component analysis

To further verify the clustering observed in the haplotype network, a PCA was conducted using the same samples and SNP set (Figure 7). The Um qasr (Figure 7A) shows that the population is combined into one with some mutation variations. In the Siniya (Figure 7B) there are more closely related and little variation in the population.

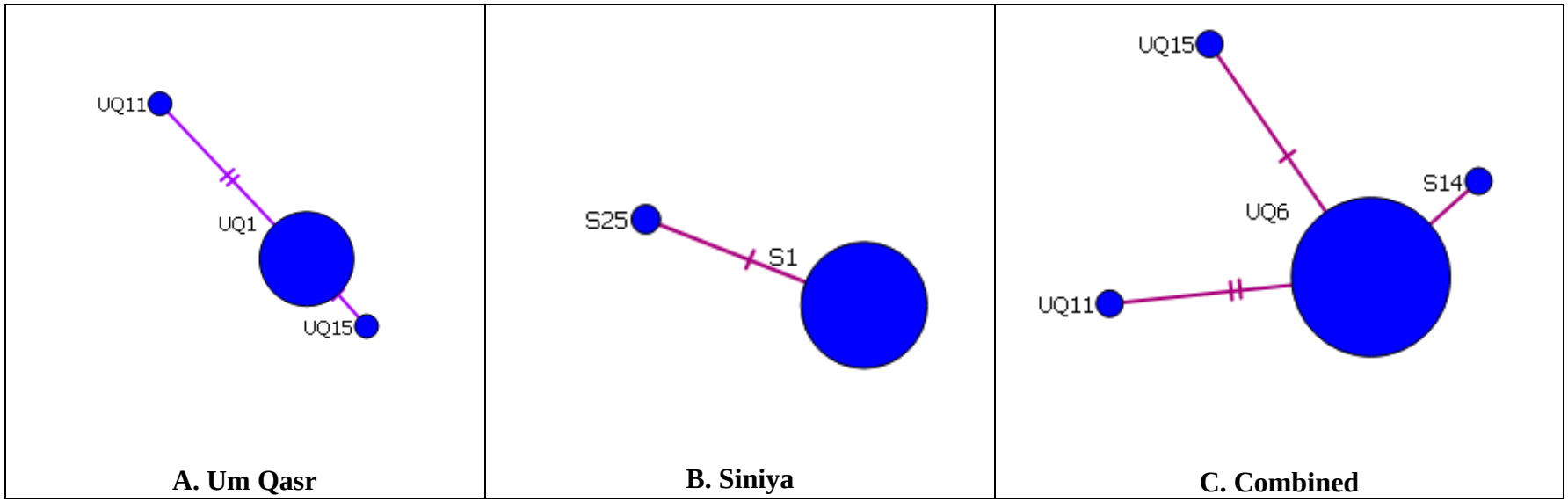


Figure 6: Median joining Haplotype network of nDNA data sequence

Note: Number in the circle are number of haplotypes copies observed individuals. Major circles represent predominant haplotypes. A branch represents a single nucleotide change

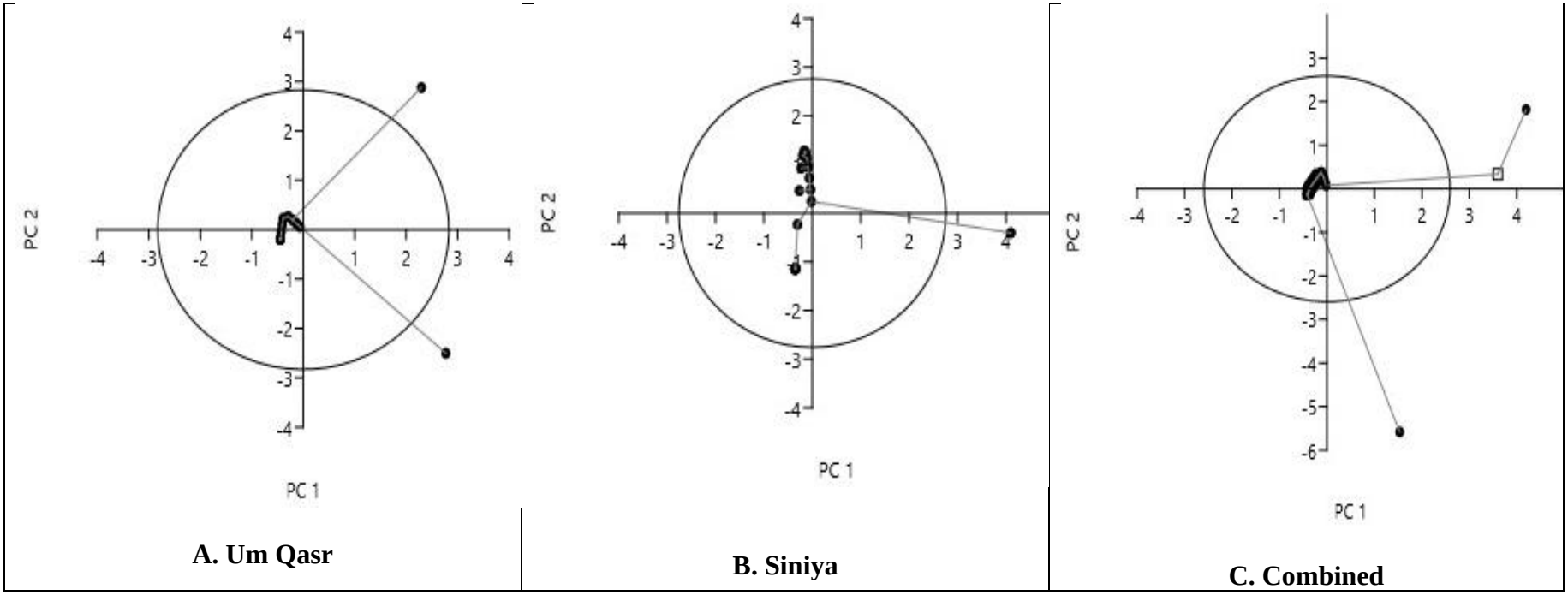


Figure 7: Principle component analysis of the nDNA sequence data

Note:Two-dimensional scatter plots of the first two principal components. The squares stand for the Um Qasr; the dots are for Siniya.

3.3 Phylogenetic tree analysis of both markers sequenced data

3.3.1 mtDNA phylogenic tree analysis

A maximum likelihood (ML) phylogenetic tree, based on mtDNA sequence of all populations, was constructed for all haplotypes (Figure 8). The populations are clustered together into 2 clades, and the haplotypes are distributed randomly between clades. Clade I contains 23 haplotypes and two subclade deviates from it. Clade II includes 13 haplotypes and a sub clade of two haplotypes.

3.3.2 Nuclear DNA SNP phylogenic tree analysis

A (ML) phylogenic tree consisting of 38 sequence was constructed based on 622 bp of nuclear DNA (Figure 9). The tested individuals cluster into two clades, almost all group into Clade I, it consists of 35 individuals and two subclades that bifurcate. Clade II consist of one sample separated from the rest by two mutation steps.

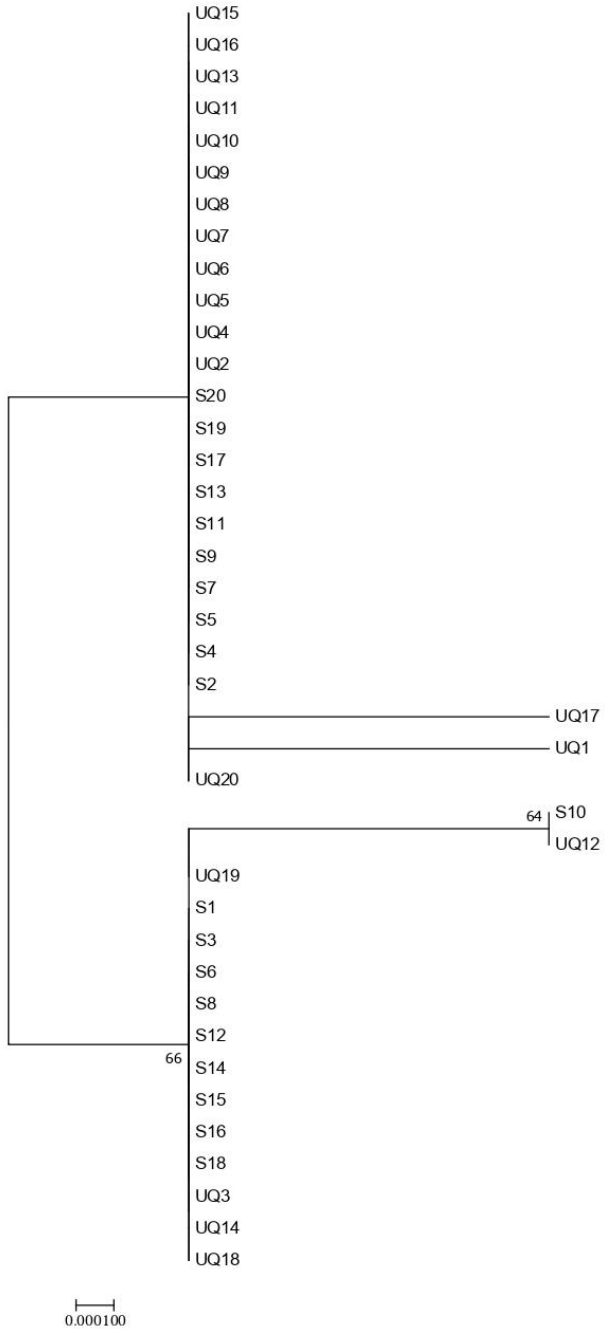


Figure 8: mtDNA Phylogenetic analysis by maximum likelihood method.
The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-1431.84) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 40 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1050 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

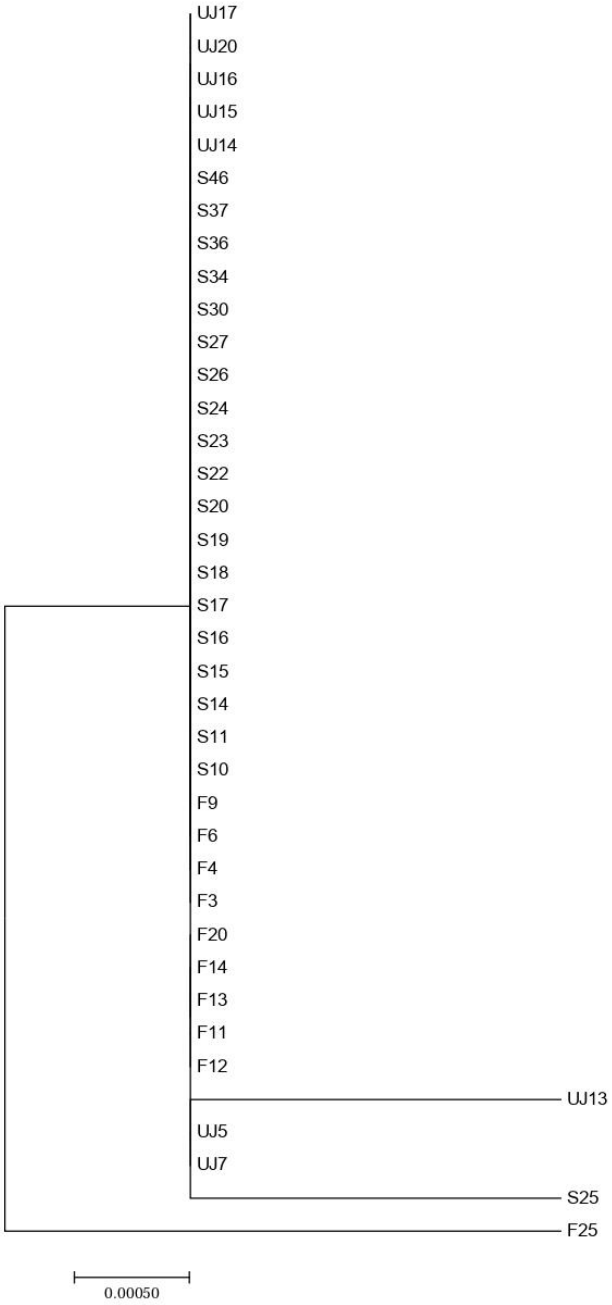


Figure 9: nDNA Phylogenetic analysis by maximum likelihood method.
The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-884.10) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 38 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 621 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

Chapter 4: Discussion

The study, explored the genetic variation of the *Phalacrocorax nigrogularis* using variety of approaches based on different nuclear and mitochondrial dataset. Throughout all analyses, it can clearly deduce that there is no differentiation between the two subpopulations of Um Qasr and Siniya, which indicates gene flow between populations. Therefore, it concludes that the two subpopulations can be treated as one population.

Examining the genetic structure within these populations, we consistently find a relation between the two populations. This close relationship between Um Qasr and Siniya becomes apparent in all phylogenetic trees (Figures 8 and 9). It is further supported by interrelation analyses in PopArt haplotype network (Figures 5 and 6), as well as by Principle component analysis PAST (Figures 5 and 7). Figure 4C clearly reveals that the Siniya and Um Qasr are closely related and overlapping. Also, there are two haplotype variables (Hap 2 and 3) emerging from the main haplotype (node 1) in Figure 4C, and they differ from each other by one mutational step, those might be diverged from the main population as newly emerging subpopulations. The haplotype network of nDNA data set (Figure 6C) shows that populations are sharing one major haplotype with three individual deviating out and this demonstration of a star like pattern is a common haplotype network pattern for recently expanding population.

The phylogenetic tree was characterized by shallow divergence between haplotypes (Figures 8 and 9) and general lack of affiliation with the distance between the two colonies. There is no east-west division between the individuals of Socotra cormorant, as suggested by phylogenetic tree and the PCA (Figures 5C and 7C).

This study presents evidence for the lack of population diversity. The results suggest a low level of genetic diversity throughout the population and seemingly high level of gene flow between populations. This is in contrast to findings by (Duffie *et al.,* 2009) where the measured genetic differentiation by the two populations was significant, leading the authors to conclude that the two *Phalacrocorax harrisi* populations are genetically distinct, and should be managed separately. Similar findings report for other cormorants populations (Winney *et al,* 2001).

The negative value of pooled Tajimas' D (D= -0.66563 mtDNA and -1.88258 nDNA) suggests that the population is expanding and singleton mutation found in (Figures 4C and 6C), may be a result of expansion population growth i.e., a pattern expected at neutrally evolving genetic markers, if population is not in mutation drift equilibrium. The star like haplotype pattern (Figure 6C) and somewhat star like in (Figure 4C) along with the singleton haplotypes might suggest that they are recently diverged haplotypes and at low frequency in populations. It might imply a bottleneck event followed by expansion and low decline in haplotype diversity is expected. A demographic expansion details are outside of the scope of this study, and dating back these expansions accurately are difficult but would be beneficial for further research.

Our mtDNA and nDNA data set provided the first view of Socotra cormorant genetic structure. All analyses strongly supported a genetic indifference between UmQasr and Siniya. The nDNA data set showed that the genetic differentiation within the population was zero (*Fst*= 0.), i.e. no genetic differentiation is detected, suggesting that the populations are one mixed population. Meanwhile, our mtDNA data set also reveals that genetic differentiation within the populations was

significantly low (*Fst* =0.04960). The gene flow value among the populations is (Nm = 4.79), where a value of Nm>4 indicates that the populations are generally mixing. This further supports out conclusion that there is negligible differentiation between the two populations and an extensive gene flow.

The variation estimates of genetic diversity based on the mtDNA sequence observed in Siniya population were low in comparison, but not significantly different to Um Qasr. The general low haplotype diversity (hd), nucleotide diversity (π) and pair wise comparison (θ) overall mtDna and nDNA marker (Table 9 and 10) imply the possibility that the populations are recovering from a long term of bottleneck event. It might suggest that the population are genetically suffering and yet to slowly recover. The observed mtDNA diversity in Sinya is feature of a population that has undergone a reduction in population size (π= 0.00060 and θ=0.00054). Such a population is considered likely to show rapid increase in the proportion of identical or closely identical haplotypes. In addition, a small or bottlenecked population would be lacking in genetic diversity. The nDNA analysis of the Siniya genetic parameter (Table 10) suggests that the population is in expansion (π= 0.00016 and θ=0.00045). However, Siniya mtDNA Tajima's D value is positive (Table 9) and may imply that the population is contracting. We assume that the Siniya population underwent a selection event, affecting the maternally inherited genes, the nucleotide and haplotype diversity is low with positive value of Tajimas D all attribute to it.

Tajima's D value is negative for the pooled data of both DNA data sets (Table 9 and 10), which signifies an excess of low frequency polymorphisms. This indicates total population size expansion (e.g., after a bottleneck), selection, or moving variation (e.g., selective sweep). Also, negative and close to zero value of

Tajimas D (Table 9) might indicate that the population in general is neutrally evolving population (Tajima, 1989a).

The amount of DNA polymorphism can be measured by the average number of nucleotide differences per site (π), the proportion of segregating (polymorphic) sites (s), and the minimum number of mutations per site (s*). Since the latter two quantities depend on the sample size, θ is often used as a measure of the amount of DNA polymorphism, where $\theta = 4Nm\mu$, N_e is the effective population size, and μ is the neutral mutation rate per site per generation (Misawa & Tajima, 1997).

The nucleotide diversity of Um qasr nDNA ($\theta = 0.00140$) is three times larger than Siniya ($\theta = 0.00045$), and the mtDNA nucleotide diversity $\theta = 0.00107$ is nearly two times as large than Siniya ($\theta = 0.00054$). We can assume that Um qasr has a θ that is 3 times higher than the Saniya θ. Therefore, Um Qasr has higher N (effective population size) than Siniya. The idea coincides with the Tajima's D value of the mtDNA of Siniya ($D = 0.26026$) where it is a positive value, which might indicate recent population contraction (Tajima, 1989a). In contrast, to findings by (Butkauskas *et al.,* 2016), where the measure of mtDNA genetic diversity of the Great cormorant (Phalacrocorax carbo sinensis) was significant leading the authors to conclude that the genetic population diversity in different breeding colonies are highly variable and well-adapted populations and possibly participating in the process of colonization of new breeding areas in the Baltic region.

Low genetic diversities of endemic species and isolated population of widespread species tend to be low and this may be an important factor contributing to their extinction risk. Based on this and the known range of genetic diversity for Socotra cormorant, we predicted that island endemic seabird with broad geographic

range would have low haplotype (hd<0.5) and nucleotide diversity (π<0.005). The population (hd) values were low (0.56 for pooled mtDNA and 0.15363 nDNA) and nucleotide diversity (π=0.00065 pooled mtDNA and 0.00034 nDNA) this support our prediction. Low nucleotide variation could result either from low mutation rate or low effective population size, possibly reflecting past census size changes in history. Low levels genetic diversity might be due to result of recent founder effect or bottleneck followed by population expansion or rapid growth from small size after bottleneck or recent contraction.

In satellite telemetry study done by Muzaffer (2014), post breeding dispersal of Siniya colony is taking place with Western region colony but no observation proves any breeding mixing. However, we can see that there is gene flow going on (Nm>4) which indicates that the populations generally mixing. Our data support these findings based on both nuclear and mitochondrial markers. No genetic differentiation was observed between population colony.

Based on our findings the mtDNA and nDNA haplotypes are common to all population at both localities, and some other haplotypes were present within Um Qasr population. Low frequency of haplotype and nucleotide diversity in general breeding and migrating contribute to gene flow and have great effect on the overall genetic differentiation in the Socotra cormorant populations. Geographically closer populations tend to be less different from geographically distant populations.

Generally, seabirds exhibit greater dispersal tendencies than inland birds, which can be a reason for little genetic difference (Schreiber & Burger, 2002).

Geographic separation might not limit dispersal and suggest cormorants are demographically unseparated. In addition, there is no evident morphological or behavioral differences between the populations. However, decline of the zoogeographic range caused by human disturbances and activities (Aspinall, 1995; Jennings, 2010; Symens *et al.*, 1993; Khan *et al.*, 2018) may get them to gather and share similar range. This might cause to be the limited degree of differentiation of the population or might be associated to the recent population expansion, past contributive to dispersal, and absence of zoogeographic boundaries to movement that might obscure any possible relation between phylogeny and geography.

The amount of genetic variation in a population is determined by many factors, such as the population size, mutation rate and natural selection. The amount of genetic variation at the DNA level (DNA polymorphism) can be estimated from the number of segregating sites and nucleotide diversity in a sample of DNA sequences. The number of segregating sites is one of the most important indicators which quantify the amount of DNA polymorphism (Watterson, 1975). This number as well as nucleotide diversity is known to be influenced by change in population size (Tajima, 1989b). The distribution of pairwise nucleotide differences among a sample of DNA sequences is also known to be influenced by change in population size (Slatkin & Hudson, 1991). Combination of moderate values of haplotype diversity coupled with very low values of nucleotide diversity at the pooled mtDNA and nDNA data set was detected in Socotra cormorant population, can be considered consistent with the scenario of population expansion after a period of low (Ne) effective population size or long term of bottleneck event. Those negative events are represented by sharp reduction in size of population would give a rise to reduced variation in gene pool, and smaller genetic diversity, this situation remains to pass on

for generations and are hard to recover from. Robustness of population and its ability to survive selecting environmental changes mainly based on their genetic diversity. It can only increase with gene flow from other populations or slowly increasing in time with random mutation. The lack of diversity can be attributed to the absences of effective law protection, exploitation or other environmental factors adding to its vulnerability. If human disturbance, egg collecting and colonies persecution persist, there is a great chance that the genetic recovery of the populations hindered.

Chapter 5: Conclusion

The analyses suggested the presence of strong/high gene flow within the UAE Socotra cormorant populations; the western cormorant population extends in the eastern cormorant population range and vice versa. No genetically distinct subpopulations can be detected between the Eastern or Western populations.

Additional sampling and whole genome sequencing (WGS) analysis would be necessary to confirm this assumption and represent accurately any potentially unique contemporary population. We note that while locations are evidently connected by sufficient gene flow to prevent significant genetic divergence, the Eastern and Western regions, or portions of the breeding region therein, may still warrant differential consideration from populations in the Northern and Southern regions to possible demographic separation, habitat differences, and documented declines at some colonies. Additional information on regional population demographics, breeding site fidelity, and prolonged movements throughout the annual cycle, would further clarify the demographic connectivity of these regional populations.

References

Acevedo-Whitehouse, K., Gulland, F., Greig, D., & Amos, W. (2003). Disease susceptibility in California sea lions. Nature, 42(27), 35-49.

Aspinall, S. (1995). Why the Socotra Cormorant Phalacrocorax nigrogularis should be formally protected. Tribulus, 5, 10–12.

Aspinall, S. (1996a). International researchon Socotra cormorants. Arabian Wildlife, 2(3), 11-22.

Aspinall, S. (1996b). Status and conservation of the breeding birds of the United Arab Emirates. Dubai, UAE: Hobby Publications.

Bagi, Z., Dimopoulos, E., Loukovitis, D., Eraud, C., & Kusza, S. (2018). MtDNA genetic diversity and structure of Eurasian Collared Dove (Streptopelia decaocto). Plos One, 13(3), 25-38. doi: 10.1371/journal.pone.0193935

Bandelt, H., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution, 16(1), 37–48.

Barahona, S., Oré-Chávez, D., & Bazán, R. Q. (2017). High genetic connectivity and Population Expansion of Scomber japonicus in the Northern Humboldt Current System revealed by mitochondrial control region sequences. Revista peruana de biología, 24(2), 163-174.

Barlow, E., Daunt, F., & Wanless, S. (2011). Weak large-scale population genetic structure in a philopatric seabird, the European Shag Phalacrocorax aristotelis. The International Journal of Avian Science, 153, 768–778.

Bensch, S., Irwin, D., Irwin, J., Kvist, L., & Åkesson, S. (2005). Conflicting patterns of mitochondrial and nuclear DNA diversity in Phylloscopus warblers. Molecular Ecology, 15(1), 161-171.

Berlin, S., Quintela, M., & Höglund, J. (2008). A multilocus assay reveals high nucleotide diversity and limited differentiation among Scandinavian willow grouse (Lagopus lagopus). BMC Genetics, 9(1), 55-67.

BirdLife International (2017). Phalacrocorax nigrogularis.The IUCN Red List of Threatened Species. Retrieved 5 May, 2018 from https://www.iucnredlist.org/species/22696802/132594449

Brooke, M. L. (2002). Seabird Systematics and Distribution:A Review of Current Knowledge. In E.A.Schreiber, & J. Burger, Biology of Marine Birds (pp. 57-85). Florida: CRC Press.

Brown, W., George, M., & Wilson, A. (1979). Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences, 76(4), 1967-1971. doi: 10.1073/pnas.76.4.1967

Bundy, G., Conner, R.J. and Harrison, C.J.O. (1989). Birds of the Eastern Province of Saudi Arabia, Witherby, London: Academic Press.

Butkauskas, D., Chaika, K., Švažas, S., Grishanov, G., Paulauskas, A., & Pūtys, Ž. (2016). Genetic diversity of Great Cormorants Phalacrocorax carbo in the eastern Baltic region. Biologija, 62(3), 64-79.

Conway, C. J. & Martin, K. (1999). Effects of ambient temperature on avian incubation behavior. Behavioral Ecology, 11(2), 178–188.

Coulson, J. C. (2002). Colonial Breeding in Seabirds. In E. A. Schreiber, & J. Burger, Biology of Marine Birds (pp. 87-113). Florida: CRC Press.

Cristescu, S., Sherwin, W., Handasyde, K., Cahill, V., & Cooper, D. (2010). Detecting bottlenecks using Bottleneck 1.2.02 inn wild populations: importance of microsatellite structure. Conservation Genetics, 11, 22-35.

Croxall, J. P., Butchart, S. H., Lascelles, B. E. N., Stattersfield, A. J., Sullivan, B. E. N., Symes, A., & Taylor, P. H. I. L. (2012). Seabird conservation status, threats and priority actions: a global assessment. Bird Conservation International, 22(1), 1-34.

Dantas, G., Meyer, D., Godinho, R., Ferrand, N., & Morgante, J. (2012). Genetic variability in mitochondrial and nuclear genes of Larus dominicanus (Charadriiformes, Laridae) from the Brazilian coast. Genetics And Molecular Biology, 35(4), 847-885. doi: 10.1590/s1415-47572012005000065

Davis, J. C. (1986). Statistics and Data Analysis in Geology. New York: John Wiley & Sons.

Desjardins, P., & Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. Journal of molecular biology, 212(4), 599-634.

Duffie, C. V., Glenn, T. C., Vargas, H., & Parker, P. G. (2009). Genetic structure within an between island population of the flightless cormorant (Phalacrocorax harris). Molecular Ecology, 12, 2103-2111.

Duffy, D. C. (1983). The Ecology of Tick Parasitism on Densely Nesting Peruvian Seabirds. Ecology, 64, 110-119.

EAD (2016). Biodiversity annual report 2016-Staus of birds in Abu Dhabi. Retrieved 3 March, 2019 from Enviromental Agency Abu Dhabi: https://www.ead.ae/Publications/Status%20of%20Birds%20in%20Abu%20Dhabi%202016/Status%20of%20Birds%20in%20Abu%20Dhabi%202016.pdf

El-Mahdy, O. (2012). Mitochondrial DNA as a Marker for Genetic Diversity and Evolution. Advances In Genetic Engineering & Biotechnology, 11, 23-34. doi: 10.4172/2324-9021.1000e102

Encyclopedia of Life (2018). Phalacrocorax nigrogularis. Retrieved 14 March, 2018 from http://eol.org/pages/1049418/details

Eo, S. H., Doyle, J. M., & DeWoody, J. A. (2010). Genetic diversity in birds is associated with body mass and habitat type. Journal of Zoology, 283(3), 220–226.

Fenderson, L. E., Kovach, A. I., Litvaitis, J. A., & Litvaitis, M. K. (2011). Population genetic structure and history of fragmented remnant population of New England cotton tail (sylvilagus transitionalis). Conservation genetic, 12, 44-56.

Gallagher, M. D., Scott, D. A. Ormond, R. F. G.,Connor, R. J. and Jennings, M. C. (1984). The distribution and conservation of seabirds breeding on the coasts and islands of Iran and Arabia. Cambridge, UK: International Council for Bird Preservation.

Gasca-Pineda, J., Cassaigne, I., Alonso, R. A., & Eguiarte, L. E. (2013). Effective Population Size, Genetic Variation, and Their Relevance for Conservation: The Bighorn Sheep in Tiburon Island and Comparisons with Managed Artiodactyls. PLoS ONE, 8(10), 56-69.

Gibbs, G. C., Kennedy, M., & Penny, D. (2013). Beyond phylogeny: pelecaniform and ciconiiform birds, and long-term niche stability. Molecular Phylogenetics and Evolution, 24, 229-238.

Gonzalez-Suarez, M., Flatz, R., Aurioles-Gamboa, D., Hedrick, P. W., & Gerber, L. R. (2009). Isolation by distance among California sea lion populations in Mexico: redefining management stocks. Molecular Ecology, 18(6), 1088–1099. https://doi.org/10.1111/j.1365-294X.2009.04093.x

Hall, B. (2013). Building Phylogenetic Trees from Molecular Data with MEGA. Molecular Biology And Evolution, 30(5), 1229-1235.

Hansen, B. D., & Taylor, A. C. (2008). Isolated remnant or recent introduction? Estimating the provenance of Yellingbo Leadbeater's possums by genetic analysis and bottleneck simulation. Molecular ecology, 17(18), 4039-4052.

Harper, D. (1999). Numerical Palaeobiology. New York: John Wiley & Sons.

Hartl, D. L. & Jones, H. (2006). Essential genetics: a genomics perspective. Burlington, MA: Jones & Bartlett Learning.

Hoogstraal, H. O. (1970). Larva, nymph and life cycle of Ornithodoros (Alectorobius) muesebecki (Ixodoidea: Argasidae), a virus-infected parasite of birds and petroleum industry employees in the Arabian Gulf. Annals of the Entomological Society of America, 63, 1-8.

Hoyo, J. D., Elliott, A., & Sargatal, J. (1992). Handbook of the birds of the world. Barcelona: Lynx Edicions.

Jennings, M. (2000). Atlas of the breeding birds of Arabia (ABBA). Fauna of Arabia, 74, 86-95.

Jennings, M. C. (2010). Atlas of the breeding birds of Arabia. Fauna of Arabia, 25, 216–221.

Johansson, M., Primmer, C. R., & Merilä, J. (2007). Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (Rana temporaria). Molecular Ecology, 16(13), 2693-2700.

Johngard, P. A. (1993). Cormorants, Darters and Pelicans of the World. Washington, USA, Smithsonian Institution Scholarly Press.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. Trends in ecology & evolution, 17(5), 230-241.

Kennedy, M., & Spencer, H. G. (2014). Classification of the cormorants of the world. Molecular Phylogenetics and Evolution, 13, 249–257.

Kennedy, M., Aspinall, S., & Rodric, D. M. (2001). The phylogentic position of the Socotra cormorant Phalacrocorax Nigrogularis a near endemic of the Arabian gulf. Tribulus, 11(2), 5–13.

Khan, S. B., Javed, S., Ahmed, S., Al Hammadi, E. A., Al Hammadi, A. A., & Al Dhaheri, S. (2018). Does a recent surge in Socotra Cormorant Phalacrocorax nigrogularis nesting population and establishment of new breeding colonies ensure long term conservation? Pragmatic assessment of recent augmentation in Abu Dhabi Emirate, UAE. Bird Conservation International, 29(3), 361-369.

Kimball, R. T., Braun, E. L., Barker, F. K., Bowie, R. C., Braun, M. J., Chojnowski, J. L., & Holznagel, W. (2009). A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution.

King, H. (2004). Communal behaviour of Socotra cormorant, Bahrain. Phoenix, 20, 25-28.

Korstian, J., Hale, A., & Williams, D. (2015). Genetic diversity, historic population size, and population structure in 2 North American tree bats. Journal Of Mammalogy, 96(5), 972-980. doi: 10.1093/jmammal/gyv101

Kumar, S., Stecher , G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, 33, 1870-1874.

Librado , P., & Rozas , J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451–1452.

Lijtmaer, D. A., Kerr, K. C. R., Stoeckle, M. Y., & Tubaro, P. L. (2012). DNA Barcoding Birds: From Field Collection to Data Analysis. DNA Barcodes Methods in Molecular Biolog, 25, 127–152.

Liu, X., Zhang, P., Zhang, G., Li, S., Zhang, L., Xu, Z., Li, D. (2017). A Comparison of Genetic Diversity of COX-III Gene in Lowland Chickens and Tibetan Chickens. BioMed Research International, 20, 1–13.

Low, V., Adler, P., Takaoka, H., Ya'cob, Z., Lim, P., & Tan, T. *et al.* (2014). Mitochondrial DNA Markers Reveal High Genetic Diversity but Low Genetic Differentiation in the Black Fly Simulium tani Takaoka & Davies along an Elevational Gradient in Malaysia. Plos ONE, 9(6), 12-24.

McKnight, A., Allyn, A. J., Duffy, D. C., & Irons, D. B. (2013). 'Stepping stone'pattern in Pacific Arctic tern migration reveals the importance of upwelling areas. Marine Ecology Progress Series, 491, 253-264.

Meinertzhagen, R. (1954). Birds of Arabia. Oliver & Boyd, Edinburgh: CRC Press.

Mellanby, R. j., Ogden, R., Clements, D. N., French, A. T., Gow, A. G., Powell, R. B., *et al.* (2013). Population structure and genetic heterogeneity in popular Dog breeds in the UK. The veterinary Journal, 23, 92-97.

Mills, L. S. (2013). Conservation of wildlife populations: demography, genetics, and management. Hoboken, NJ: Wiley-Blackwell.

Misawa, K., & Tajima, F. (1997). Estimation of the amount of DNA polymorphism when the neutral mutation rate varies among sites. Genetics, 147(4), 59-64.

Muzaffar, S. B. (2012). Reproductive performance of the Socotra Cormorant (Phalacrocorax nigrogularis) on Siniya Island, United Arab Emirates: Planted trees increase hatching success. Waterbirds, 35, 626-630.

Muzaffar, S. B. (2013). The impact of fox and feral cat predation on breeding Socotra Cormorants. Marine Ornithology, 41, 171-177.

Muzaffar, S. B., Clarke, C., Whelan, R., Gubiani, R., & Cook, T. R. (2017a). Short distance directional migration in the threatened Socotra cormorant: link to primary productivity and implications for conservation. Marine Ecology Progress Series, 575, 181-194.

Muzaffar, S. B., Whelan, R., Clarke, C., Gubiani, R. and Benjamin, S. (2017b) Breeding population biology in Socotra Cormorants (Phalacrocorax nigrogularis) in the United Arab Emirates. Waterbirds, 40, 1–10.

Muzaffer, S. B. (2014). Ecology and conservation of the Socotra Cormorant (Phalacoroax Nigrogularis) in the eastern Arabian Gulf. unpublished.

Nabholz, B., Glémin , S., & Galtier , N. (2009). The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. BMC Evolutionary Biology, 9(54), 23-37.

Nelson, J. B. (2005). Pelicans, cormorants and their relatives, The Pelecaniforms. Oxford, U.K.: Oxford University Press.

Nezlin, N. P., Polikarpov, I. G., & Al-Yamani, F. (2007). Satellite-measured chlorophyll distribution in the Arabian Gulf: Spatial, seasonal and inter-annual variability. International Journal of Oceans and Oceanography, 2(1), 139-156.

NIH (2019). What are single nucleotide polymorphisms (SNPs)?. Retrieved from https://ghr.nlm.nih.gov/primer/genomicresearch/snp

Nuss, A., Carlos, C. J., Moreno, I. B., & Fagundes, N. J. (2016). Fagundes Population Genetic Structure of the Magnificent Frigatebird Fregata magnificens (Aves, Suliformes) Breeding Colonies in the Western Atlantic Ocean. Plos One, 11(2), 45-57.

Ogilvie-Grant, W. R., & Forbes, H. O. (1899). The expedition to Socotra. 1. Description of new Birds. Bulletin Liverpool Museum, 2, 2-3.

Olsson, O., Wiktander, U., Malmqvist, A., & Nilsson, S. G. (2001). Variability of patch type preferences in relation to resource availability and breeding success in a bird. Oecologia, 127(3), 435-443.

Paradis, E. (2018). Analysis of haplotype networks: The randomized minimum spanning tree method. Methods In Ecology And Evolution, 9(5), 1308-1317.

Paterson, A. M., Wallis, G. P., & Gray, R. D. (1995). Penguins, petrels, and parsimony: Does cladistic analysis of behavior reflect seabird phylogeny?. Evolution, 49(5), 974-989.

Patten, M. A., Wolfe, D. H., Shochat, E., & Sherrod, S. K. (2005). Habitat fragmentation, rapid evolution and population persistence. Evolutionary Ecology Research, 7(2), 235-249.

Peterman, W. E., Feist, S. M., Semlitsch, R. D., & Eggert, L. S. (2013). Conservation and management of perpherial popuations: Spatial and temporal influences on genetic structure of wood frog (Rana sylvatica) populations. Biological Conservation, 12, 351-358.

Pham, L. D., Do, D. N., Binh, N. T., Van Ba, N., Thuy, T. T. T., Hoan, T. X., ... & Kadarmideen, H. N. (2013). Assessment of genetic diversity and population structure of Vietnamese indigenous cattle populations by microsatellites. Livestock science, 155(1), 17-22.

Ramírez, F., Afán, I., Tavecchia, G., Catalán, I. A., Oro, D., & Sanz-Aguilar, A. (2016). Oceanographic drivers and mistiming processes shape breeding success in a seabird. Proceedings of the Royal Society B: Biological Sciences, 28(26), 20-32.

Schreiber, E. A., & Burger, J. (2002). Biology of Marine Birds. CRC Press, Boca Raton, Florida.

Sharma, M., Fomda, B., Mazta, S., Sehgal, R., Bagicha Singh, B., & Malla, N. (2013). Genetic Diversity and Population Genetic Structure Analysis of Echinococcus granulosus sensu stricto Complex Based on Mitochondrial DNA Signature. Plos one, 8(12), 65-78.

Slatkin, M., & Hudson, R. R. (1991). Pairwise comparison of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics , 129, 555–562.

Sonsthagen, S. A., Wilson, R. E., & Underwood, J. G. (2017). Genetic implications of bottleneck effects of differing severities on genetic diversity in naturally recovering populations: An example from Hawaiian coot and Hawaiian gallinule. Ecology and Evolution, 7(23), 9925–9934.

Soo, H., Doyle, J. M., & Dewoody, J. A. (2010). Genetic diversity in birds is associated with body mass and habitat type. Journal of Zoology , 283(3), 23-34.

Stoneking, M., Hedgecock, D., Higuchi, R. G., Vigilant, L., & Erlich, H. A. (1991). Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. American Journal of Human Genetics, 48(2), 370-379.

Subramanian, S. (2016). The effects of sample size on population genomic analyses – implications for the tests of neutrality. BMC Genomics, 17(1), 33-45.

Symens, P., Kinzelbach, R., Suhaibani, A., & Werner, M. (1993). A review of the status, distribution and conservation of the Socotra Cormorant,Phalacrocorax nigrogularis. Zoology in the Middle East, 8(1), 17–30.

Tajima, F. (1989a). Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics, 123, 585-595.

Tajima, F. (1989b). The effect of change in population size on DNA polymorphism. Genetics, 123, 597–601.

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution, 10, 512-526.

Tan, K., Qu, Y., Wang, Z., Liu, Z., & Engel, M. (2015). Haplotype diversity and genetic similarity among populations of the Eastern honey bee from Himalaya-Southwest China and Nepal (Hymenoptera: Apidae). Apidologie, 47(2), 23-34.

Thomson, C., & Brooke, J. M. (2014). Divergence between Avian Sister Species is Linked to Generation Length and Body Mass. Plos One, 23, 56-68.

Tomar, R. S. (2010). Molecular markers & plant biotechnology. New Delhi: New India Pub. Agency.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. Theoretical population biology, 7(2), 256-276.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. Annu. Rev. Ecol. Evol. Syst., 37, 433-458.

Wilson, K. D. (2012). Breeding of Socotra cormorant at Umm Qasr Island. The Phoenix, 28, 26-28.

Winney, B., Litton, C., Parkin, D., & Feare, C. (2001). The subspecific origin of the inland breeding colonies of the cormorant Phalacrocorax carbo in Britain. Heredity, 86(1), 45-53.

Wright, L., Tregenza, T., & Hosken, D. (2007). Inbreeding, inbreeding depression and extinction. Conservation Genetics, 9(4), 833-843.

Wright, S. (1978). Evolution and the Genetics of Populations. Variability within and among Natural Populations. Chicago, Illinois: University of Chicago Press.

Yasue, M., & Dearden, P. (2006). Simultaneous biparental incubation of two nests by a pair of Malaysian plover Charadrius peronii. Wader Study Group Bull, 109, 121-132.

Yorio, P., Bertellotti, M., & Quintana, F. (1995). Preference for covered nest sites and breeding success in Kelp Gulls Larus dominicanus. Marine Ornithology, 23, 121-128.

Appendix

Nuclear markers

Park7

>P2_UJ20 {length 743, GC 43%}

GGCAGTTGGCCAGATGTCTTCATTTGTCCTGATGCCAGTCTTGAAGATGCCAGAAAAGAGGTTGGT
CATCACCCTAGTCTGATATAATGGCCACCATAGCCTTTCTCACAGTTTCAGCAGTGAATCTCGAAAG
CTTATGCATTACCTTTAGAAATATGTTATGTGGTATGTTAGCTCTGCGCAGTAGTGAATCACAGAAC
TGCACGTCCTGCAGTCCCCCTCTTCAGTGTTGTCCTGCTTCATAGCATGCTAGGATCAGACAGATTTC
TTAACTCTGCATGGGTATTTGTTTTTGCAAAGCATCAAACATAGAGCTTTATTTTGCATGTTAGTCAG
CCCAACAAAACAGCGTGATTTGGAAAAACTGAAAAGTTCTGATGGGAAGCTAAGCAATGCTGTAGT
TAGACCTGTACAGTAAGGTTGGAGCAAAACCACTTGTCTTTACAGTGCAAGAGCTGCCATTTGCTTG
CTGAAGCTCTTAAATGCCATGGGTTAAAGTTAGTTCTGTGACAGACTTTATTGACAATAGGCTCTTC
CTGGAATAGGAAACAGTATTTTGTTGCACAGGTGGAACTGGTTCTCTGGTCATACAGATCCCTTTGC
TTCTCGCTTGACTCAGCTGTAGTGGCCTTAAAACATTGTAACATATCTACTGGCCTCTGTTCAACAGA
ACAGTTTCGGTTTTTGTTTTTTTAGGGGCCTTACGATGTCGTGGTCCTGCCTGGGAGTATAAAGCTTT
GAAA

MtDNA markers

12S

>P8_UJ20 {length 195, CG 51%}

GCACCACGTAGAGGAGCCTGTTCTATAATCGATAACCCACGATTCACCCGACCGCCCCTTGCCAAAA
CAGCCTACATACCGCCGTCGTCAGCCCACCTTCATGAAAGAACAACAGTGAGCGCAATAGCCTCTCC
CGCTAACAAGACAGGTCAAGGTATAGCTCATGAGGCGGAAGAAATGGGCTACATTTTCTTA

ATPs6

>P5_UJ20 {length 675, GC 46%}

GGTCATGTGTATAGGCTTAATAGTAGTACGAATACGTAGGCTTGGATTATGGCTACTGCTACTTCTA
GGATTGTTAATAGTAGTAGGATTAGTGTAGTTAGGATTGAGATTGTTGGCATGATGGGGAGGAGT
ACAGTGGTAGCTGTGGAGATTAGTTGGATGAGGAGGTGGCCTGCTGTGAGGTTTGCTGTGAGTCG
AACACCAAGTGCAAGAGGGCGGATGAGTAGGCTGGTAGTTTCGATTATGATTAGGGCTGGGATTA
GGGGAGTGGGGGTTCCTTCAGGCAGGAGGTGTCCTAGTGAGATAGTTGGTTGGTTTCGTAGGCCA
GTGAGTAAGGTTGCAAGTCAGAGTGGGAATGCCAGTGCTATATTTATTGATAGTTGGGTGGTTGGG
GTGAATGTATATGGGAGTAGCCCTAGTAAGTTGGTTATGAGTAGGAGGGCTATTAATGATGTTAGG
ATTAGGGCCCATTTATGCCCCTTCTTGTTTAGGGGCATTATTAGTTGTTTTGTGATAAGGTGGAGTA
GTCATGATTGTAGGGTGGAGAGTCGGTTGGTGACCCATCGGTTATCTGGTGATGGGAGTAGTAGG
GCAGGGAGTAGTATAGAGAGTAGGATTAGAGGAATGCCTAGGAGGCATGGGCTTGTGAATTGGTC
GAAGAAGCTTAAGGTTCAT

COI

>P11_UJ20 {length 379, CG 49%}

TCGTATTTTCGGTGCCTGAGCCGGCATAGTTGGAACTGCCCTCAGCCTACTTTATCCGTGCAGAACT
AGGCCAACCAGGAACCCTCCTAGGAGATGACCAAATCTACAACGTAATTCGTCACTGCCCATGCTTT
TGTAATAATCTTCTTCATAGTGATGCCCATCATAATCGGGGGATTCGGAAACTGACTGGTTCCCCTC
ATAATCGGTGCCCCCGACATAGCATTCCCACGTATAAATAACATAAGCTTCTGACTCCTTCCACCATC
ATTCCTCCTCCTATTAGCCTCTTCTACAGTAGAAGCAGGCGCAGGTACAGGATGAACTGTATATCCC
CCTCTAGCCGGAAACCTCGCCCATGCCGAAGCCTTCAGT