UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

FACULTY OF SCIENCE

# INTERSPECIFIC HYBRIDISATION BETWEEN THE MALLARD AND YELLOW-BILLED DUCK IN SOUTH AFRICA

STACEY DE SOUZA
717155

A DISSERTATION SUBMITTED TO THE FACULTY OF SCIENCE, UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG, SOUTH AFRICA, IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

MAY, 2019

# DECLARATION

I, **Stacey de Souza** (**717155**), am a student registered for the degree of Master of Science in the academic year 2017 and 2018.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Dissertation **HAS NOT** been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature:  24th day of May 2019

# STRUCTURE OF THE DISSERTATION

This research dissertation is made up of four chapters. The first chapter introduces the study with a review of the literature that served to inform the study. Chapter two discusses the methodology used in the study. Traditionally, chapter three would describe the results; however, here it is presented in the format of a research paper. The final chapter is a discussion as to the findings, limitations, future work and conclusions drawn from the study. A single reference list can be found at the end of the dissertation. Due to the structure of the dissertation, repetition between chapters is unavoidable.

# RESEARCH OUTPUTS

## Conference outputs

**de Souza S**, Symes C and Mollett J. 2017. Interspecific hybridisation between the Mallard and Yellow-billed Duck in South Africa. University of the Witwatersrand Molecular Biosciences Research Trust (MBRT) Research Day. Wits Club, Braamfontein West Campus. 30 November 2017. *Poster presentation*.

**de Souza S**, Symes C, Reynolds C and Mollett J. 2018. Interspecific hybridisation between the Mallard and Yellow-billed Duck in South Africa. SASBi/SAGS Joint Congress. Golden Gate Resort, Free State Province. 16-18 October 2018. *Oral presentation*.

**de Souza S**, Symes C, Reynolds C and Mollett J. 2018. Interspecific hybridisation between the Mallard and Yellow-billed Duck in South Africa. 9$^{th}$ Cross Faculty Postgraduate Symposium. New Commerce Building, Braamfontein West Campus. 29-30 October 2018. *Poster presentation*.

**de Souza S**, Symes C, Reynolds C and Mollett J. 2018. Interspecific hybridisation between the Mallard and Yellow-billed Duck in South Africa. University of the Witwatersrand Molecular Biosciences Research Trust (MBRT) Research Day. Wits Club, Braamfontein West Campus. 29 November 2018. *Poster presentation*.

## Publication outputs

**de Souza S**, Symes C, Reynolds C, Smit-Robinson H, and Mollett J. 2018. Minimal evidence of interspecific hybridisation between the Mallard and Yellow-billed Duck in central and north-west South Africa. *Submitted to Ostrich: Journal of African Ornithology*. *Manuscript ID: TOST-2019-0028.*

# ABSTRACT

Interspecific hybridisation is the interbreeding of genetically distinct groups and can lead to introgression, the exchange of genetic material between species, thereby leading to a loss of local genetic adaptations and genetic diversity. Hybridisation is a significant threat for dabbling ducks where interbreeding with the closely related invasive Mallard (*Anas platyrhynchos*) is extremely common. Mallards have been introduced into South Africa, and phenotypic evidence now suggests that they are hybridising with the native Yellow-billed Duck (*Anas undulata*). The aim of this research was to estimate the presence and extent of hybridisation between Yellow-billed Ducks, representing populations in central and north-west South Africa, and the introduced Mallard. Two mitochondrial DNA markers, *cyt b* and *COI*, and eleven microsatellite markers were used to assess genetic variation between these species and their putative hybrids. The microsatellite markers showed no significant admixture within the putative hybrids, however, one phenotypic Mallard sample showed significant admixture. The mitochondrial DNA markers separated the Mallards and Yellow-billed Ducks, and grouped the putative hybrids with the latter. Both the microsatellite and mitochondrial DNA markers were found to be successful at identifying admixture between Mallards and Yellow-billed Ducks, and separating these species both inter- and intra-specifically. However, an insufficient number of markers were used to infer backcrossed or late generation hybrids, and sampling was too localised to determine if hybridisation is occurring at different rates across South Africa. Future research should focus on whole-genome sequencing of the Yellow-billed duck to develop more markers, and sampling should occur throughout South Africa to better determine the different rates at which hybridisation may be occurring.

# DEDICATION

*This dissertation is dedicated to my grandfather, Manuel Quintal, for teaching me the most important values in life: hard work, dedication, love, respect and family. Love you always, and you are forever in my heart.*

# ACKNOWLEDGEMENTS

I would like to acknowledge and extend my sincere appreciation to the following individuals and institutions:

My supervisor, Dr Jean Mollett, for the continuous patience and guidance throughout these past two years. Thank you for always encouraging me to keep on going when the times got tough.

My co-supervisor, Prof. Craig Symes, thank you for the continued guidance and advice.

My co-supervisor, Dr Chevonne Reynolds, thank you for all the ecological knowledge you've given me and for the collection of blood samples.

My advisor, Dr Nikki Gentle, and Dr Vanessa Meyer, thank you for all your extra help along the way.

My colleagues in the Gene Marker and Bioinformatics laboratories, for all the advice and support throughout this study.

James Bird, I couldn't thank you enough for all that you have done for me during this period. Thank you for being my constant support.

Most importantly, to my parents, I am so truly blessed for all the opportunities you have given me. I would not be where I am today if it weren't for the two of you. Words could never express how thankful I am.

Lastly, much gratitude to the National Research Foundation as well as Birdlife South Africa for the financial support.

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AEWA | African-Eurasian Migratory Waterbirds Agreement |
| AMOVA | Analysis of Molecular Variance |
| bp | Base Pairs |
| BIC | Bayesian Information Criterion |
| CBD | Convention on Biological Diversity |
| *CHD* | *Chromodomain-Helicase-DNA* |
| *cyt b* | *Cytochrome b* |
| *COI* | *Cytochrome c Oxidase subunit I* |
| DNA | Deoxyribonucleic Acid |
| DNase | Deoxyribonuclease |
| dNTPs | Deoxyribonucleotide Triphosphates |
| ddNTPs | Dideoxynucleotide Triphosphates |
| Dr | Doctor |
| EDTA | Ethylenediaminetetraacetic Acid |
| HWE | Hardy-Weinberg Equilibrium |
| HKY | Hasegawa-Kishino-Yano |
| HCl | Hydrochloric Acid |
| LD | Linkage Disequilibrium |
| MCMC | Markov Chain Monte Carlo |
| mtDNA | Mitochondrial DNA |
| MEGA | Molecular Evolutionary Genetics Analysis |
| NEMBA | National Environmental Management: Biodiversity Act |
| PCR | Polymerase Chain Reaction |
| PopART | Population Analysis with Reticulate Trees |
| PCA | Principal Component Analysis |
| Prof. | Professor |
| rpm | Revolutions Per Minute |

| | |
|---|---|
| RNase | Ribonuclease |
| SNP | Single Nucleotide Polymorphism |
| SDS | Sodium Dodecyl Sulfate |
| STE | Sodium Chloride, Tris-HCl, EDTA |
| SA | South Africa |
| SABAP | Southern African Bird Atlas Project |
| *Taq* | *Thermophilus aquaticus* |
| UV | Ultraviolet |
| USA | United States of America |

# LIST OF SYMBOLS

| | |
|---|---|
| A230/280 | Absorbance ratio at 230 nm and 280 nm |
| A260/280 | Absorbance ratio at 260 nm and 280 nm |
| Ta | Annealing Temperature |
| q | Assignment Probability or Posterior Probability |
| Df | Degrees of Freedom |
| ΔK | Delta K |
| E | East |
| ε | Epsilon |
| = | Equals |
| $H_{exp}$ | Expected Heterozygosity |
| 5′ | Five Prime |
| $F_{ST}$ | Fixation Index |
| > | Greater Than |
| ≥ | Greater Than or Equal To |
| $F_{IS}$ | Inbreeding Coefficient |
| < | Less Than |
| $H_{obs}$ | Observed Heterozygosity |
| % | Percent |
| Q | Phred Quality Score |
| K | Number of Populations |
| S | South |
| $r^2$ | Squared Coefficient of Determination |
| SS | Sum of Squares |
| 3′ | Three Prime |
| Va | Variance |

# LIST OF UNITS OF MEASUREMENT

| | |
|---|---|
| cm | centimetre |
| ° | degrees |
| °C | degrees Celsius |
| g | grams |
| $\mu$l | microlitre |
| $\mu$M | micromolar |
| ' | minutes |
| ng | nanograms |
| ng/ $\mu$l | nanograms per microlitre |
| nm | nanometre |
| " | seconds |
| U | units |
| V | volts |

# LIST OF FIGURES

# LIST OF TABLES

# Chapter 1

# Introduction

## 1.1 Hybridisation

Anthropogenic activity has facilitated the introduction of biological invaders for many years, allowing for sympatry of once allopatric organisms (Chown *et al.* 2015). Biodiversity and ecosystem integrity are threatened by alien invasive species worldwide (Vilà *et al.* 2010). Alien species can have major ecological effects such as habitat alteration through physical and chemical process, disruption of food webs and its processes, increased competition and predation as well as the introduction of novel pathogens and/or parasites (Josefsson and Andersson 2001, Vilà *et al.* 2010). However, invasive species can also have genetic consequences on native species through hybridisation.

Hybridisation is defined as the interbreeding of genetically distinct groups (Rhymer and Simberloff 1996) and can occur naturally or anthropogenically. Hybridisation can lead to introgression which is the exchange of genetic material between interbreeding organisms (Rhymer 2006, Guay *et al.* 2014). Natural hybridisation can be beneficial in that it allows for adaptability and diversification of species, by increasing fitness through means such as increased survival or improved immuno-efficiency, known as heterosis (Mallet 2005, Tompkins *et al.* 2006). On the Chatham Islands, New Zealand, immune function was enhanced in the hybrid population of the Forbes' Parakeet (*Cyanoramphus forbesi*) and the Red-crowned Parakeet (*Cyanoramphus novaezelandiae*), which led to an increased survival rate in the new populations as compared to their "pure" populations (Tompkins *et al.* 2006). Natural hybrids amongst Darwin's finches (*Geospiza scandens* and *Geospiza fortis*) were found to be fitter than parental populations by producing phenotypes that were better adapted to current conditions on Galapagos (Grant *et al.* 2003, Mallet 2005).

Anthropogenic hybridisation may be a consequence of habitat alterations or species introductions due to human activity (Malukiewicz *et al.* 2015). Anthropogenic hybridisation can lead to a loss of local genetic adaptations and diversity, ultimately causing extinction of native and/or introduced species (Josefsson and Andersson 2001). Introduced Green Iguana (*Iguana iguana*) males are larger and are sexually active earlier than endemic Lesser Antillean Iguana (*Iguana delicatissima*) males, therefore allowing them to displace the native males and reproduce with the females, resulting in a substantial decline of the Lesser Antillean Iguana in the Lesser Antilles (Vuillaume *et al.* 2015). Habitat alteration, due to agriculture, allowed for sympatry of Blue-winged Warblers (*Vermivora cyanoptera*) and Golden-winged Warblers (*Vermivora chrysoptera*) in eastern North America, consequently causing introgression of Blue-winged Warbler genes into nearly a third of phenotypically representative Golden-winged Warbler genomes (Vallender *et al.* 2007).

Hybridisation can lead to species decline through outbreeding depression (Rhymer 2006). This is expressed through lowered fitness due to embryonic inviability or sterility of hybrids (Rhymer 2006). Introduced American Mink (*Mustela vison*) males mate earlier with native European Mink (*Mustela lutreola*) females, preventing other males from mating with them (Maran and Henttonen 1995, Rhymer and Simberloff 1996). These females produce inviable embryos and therefore produce no offspring, which with continued hybridisation has been found to result in significant reductions in the European Mink population (Maran and Henttonen 1995, Rhymer and Simberloff 1996). Hybrids of Red Hartebeest (*Alcelaphus buselaphus*) and Blesbok (*Damaliscus dorcas*) are sterile, contributing to the decline of both species within game farms in South Africa (Robinson *et al.* 1991, Rhymer and Simberloff 1996).

In extreme cases hybrid swarms or genetic swamping can also occur, causing genetic extinction of both parental species and only hybrids remain. The introduction of the Italian Tree Frog (*Hyla intermedia*) into the Grangettes natural reserve has allowed for introgressive hybridisation with the native European Tree Frog (*Hyla arborea*), resulting in the extinction of both parental species in this region. Hybrid individuals can also have ecological consequences within an

ecosystem. Plains Leopard Frog (*Rana blairi*) and Southern Leopard Frog (*Rana sphenocephala*) hybrids, in south eastern Missouri, were found to be significantly more infected by a chytrid fungus (*Batrachochytrium dendrobatidis*) compared to the parental species (Parris 2004). Hybrids as such could harbour pathogens and diseases that the parental species never could, potentially harming the ecosystem in which they occur (Parris 2004). Ryan *et al.* (2009) found that the survival rate of two native community members, the Pacific Chorus Frog (*Pseudacris regilla*) and the California Newt (*Taricha torosa*), was drastically reduced due to increased predation by large tiger salamander hybrids in Salinas Valley, California. These large hybrids are a result of interbreeding between the native California Tiger Salamander (*Ambystoma californiense*) and the introduced Barred Tiger Salamander (*Ambystoma tigrinum mavortium*).

In animals, hybridisation is hypothesised to occur due to limited availability of conspecifics, flexible mate choices, errors in mate choice, or forced-extra pair copulations (Grant and Grant 1997, Randler 2002, Willis 2013). Sex-ratio biases may cause limitations on the availability of conspecific mates, thereby instigating interbreeding with heterospecifics (Grant and Grant 1997). In birds, where mate pair-bonds occur, it has been found that individuals may be flexible when choosing a mate once all conspecific mates have already been paired, thus mating with heterospecifics (Randler 2002). Flexibility in mate choice may also occur when heterospecifics have supernormal stimuli, for example brighter appearance, complex songs or social dominance (Randler 2002). Often errors in conspecific recognition may occur, especially when species have similar mating behaviours and phenotypes, leading to interbreeding between species (Randler 2002). In order to maximise the chances of producing offspring, some males mate with as many females as possible, including heterospecifics (Randler 2002, Willis 2013). This is known as forced extra-pair copulations (Randler 2002, Willis 2013).

## 1.2 Hybridisation with the Mallard

Hybridisation is relatively common amongst birds, especially within the order Anseriformes, where over 500 hybrid crosses have been reported in the wild and in captivity (McCarthy 2006). One member of this group, the Mallard (*Anas platyrhynchos*) is especially well-known for hybridising with many species.

The Mallard is a sexually dimorphic dabbling duck that is 37-41 cm in length and weighs about 850-1400 g (**Figure 1.1a & 1.1b**, Sinclair *et al.* 2011). Males have glossy green heads and chestnut breasts, while females have spotted brown plumage, and both have orange feet (Sinclair *et al.* 2011). Domestic Mallards usually have plain white plumage and are often known as the Pekin Duck (*Anas platyrhynchos domesticus*, **Figure 1.1c**, Ashton and Ashton 2001). It is common for the wild and domestic strains to interbreed allowing for a range of phenotypes within Mallards (**Figure 1.1d – 1.1f**, Wood-Gush 2012). The Mallard is the most abundant and well-known waterfowl species that is native to the northern hemisphere (Kulikova *et al.* 2005). They possibly have the largest range of all the dabbling ducks with a near global distribution (Rogers 2001). Mallards have been introduced purposefully into many countries worldwide for game hunting and aesthetic purposes, and have been domesticated (Desforges and Wood-Gush 1975, Rhymer 2006). Both wild and domestic Mallards now highly threaten endemic dabbling ducks within these countries due to competition, displacement and hybridisation (Williams and Basse 2006, Taysom *et al.* 2014). Mallards are generally larger, more aggressive and more dominant over other duck species therefore allowing them to sexually outcompete these species (Williams and Basse 2006). Most cases of hybridisation with Mallards involve sexually monomorphic dabbling ducks, thus, it has also been suggested that the ornate plumage of the male Mallard provides a supernormal stimulus to females of other species, therefore favouring hybridisation (Guay *et al.* 2014). Domesticated Mallards have been shown to have physiological and behavioural changes, which are thought to promote hybridisation (Hepp *et al.* 1988). Such changes include: increased testosterone, longer breeding seasons, increased promiscuity and forced extra-pair copulations (Desforges and Wood-Gush 1976, Hepp *et al.* 1988).

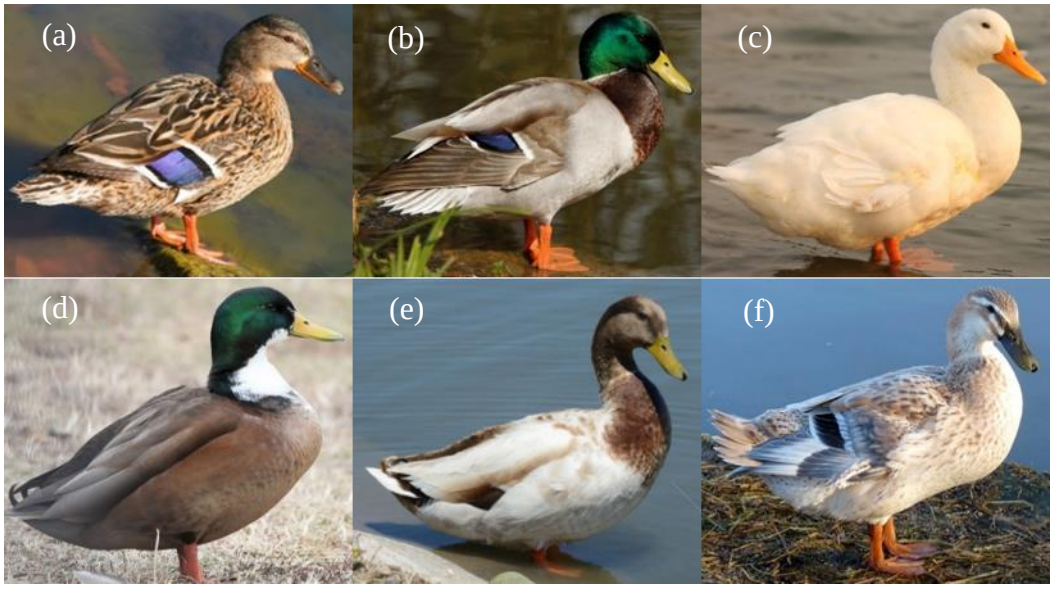


**Figure 1.1**: Photographic representations of pure Mallards (*Anas platyrhynchos*). (a) A female Mallard. Photograph accredited to B Lally. (b) A male Mallard. Photograph accredited to P Brown. (c) A domestic Mallard (Pekin Duck, *Anas platyrhynchos domesticus*). Photograph accredited to Peakpx.com. (d-f) Photographic displays of the range of phenotypes of the Mallard. Photographs accredited to (d) W Alexander, (e) DM Crawford, and (f) M Fuller.

Mallards are known to hybridise with approximately 40 Anatidae species in the wild and a further 20 in captivity (**Appendix Table A1**, McCarthy 2006). Hybridisation with the Mallard is, particularly, a significant threat for dabbling ducks belonging to the Mallard Complex (**Appendix Table A2**), whereby hybridisation events have been shown to cause large genetic effects, population declines and local extinctions in many of these species (Rhymer 2006, Guay *et al.* 2014). The Mallard Complex is made up of 13 closely-related dabbling duck species, which all have similar life history traits and behaviours, making them more vulnerable to interspecific hybridisation with the Mallard (Delacour and Mayr 1945, Lavretsky *et al.* 2014).

Changes in land-use in southwestern America enabled the sympatry between Mallards and Mexican Ducks (*A. diazi*) (Rhymer 2006). This allowed for extensive hybridisation between the species, which led to the genetic swamping of the Mexican Duck in America and northern Mexico (Perez-Arteaga *et al.* 2002, McCarthy 2006, Rhymer 2006). However, pure populations of Mexican Ducks still remain in central Mexico (Perez-Arteaga *et al.* 2002, Rhymer 2006). The release of

farm-reared Mallards in Florida allowed for non-native Mallards to invade Mottled Duck (*A. fulvigula*) breeding grounds (Williams *et al.* 2005, Ford *et al.* 2017). Also, Mottled Ducks were translocated into South Carolina, outside of their native range and into the breeding range of Mallards (Williams *et al.* 2005, Ford *et al.* 2017). This facilitated hybridisation between the species to such an extent that the populations in South Carolina are implicated in a hybrid swarm, while, hybridisation is occurring at a much slower rate in Florida where Mallards are less abundant (Williams *et al.* 2005, Guay *et al.* 2014, Ford *et al.* 2017). In eastern North America habitat alteration due to agriculture allowed for the sympatry between Mallards and American Black Ducks (*A. rubripes*, Rhymer 2006). Extensive introgressive hybridisation between these two species has resulted in a loss of genetic differentiation between them (Mank *et al.* 2004).

The introduction of the Mallard onto the Hawaiian Islands permitted severe hybridisation with the endemic Hawaiian Duck (*A. wyvilliana*), causing the extinction of the Hawaiian Duck from all the islands except Kaua'I (Rhymer 2001, Fowler *et al.* 2009). The critically endangered Laysan Duck (*A. laysanensis*) is also endemic to Hawaii and occurs only on Laysan Island (Rhymer 2001). This species has survived severe bottlenecks and population declines causing them to have a very small population with low genetic diversity, which is now highly threatened due to hybridisation with the introduced Mallard (Rhymer 2001, McCarthy 2006). The endangered Meller's Duck (*A. melleri*), endemic to Madagascar, is also threated by hybridisation with game-farm Mallards released from Lac Alaotra (McCarthy 2006, Rhymer 2006). In the West Pacific Islands of Guam, Tinian and Saipan, Lesser Grey Ducks (*A. superciliosa pelewensis*) and Mallards have both gone locally extinct and only their hybrids can be found (Guay *et al.* 2014). The introduction of Mallards onto Lord Howe Island has led to the local extinction of the native New Zealand Grey Duck (*A. s. superciliosa*), due to extensive hybridisation (Guay *et al.* 2014). Similarly, in New Zealand, extensive hybridisation with the introduced Mallard has caused a massive decline in the New Zealand Grey Duck population, to the extent that they are now endangered and on the verge of extinction (Rhymer 2006, Guay *et al.* 2014). It is now believed that

Australian Black Ducks (*A. s. rogersi*) in Australia will suffer this same fate due to introduced Mallards (Rhymer 2006).

## 1.3 The Mallard and the Yellow-billed Duck

Mallards have been introduced both deliberately and accidentally into South Africa, and have subsequently become established (Owen *et al.* 2006). Both wild and domestic strains from Britain were introduced into the Western Cape Province in the 1940's (Brooke and Siegfried 1991). The wild Mallards were used for game-hunting and ornamental purposes, whereas the domestic Mallards were used for farming. The Department of Environmental Affairs has since declared the Mallard an alien invasive species in South Africa (DEA 2016). The Mallard is listed as category 2 on the national list of invasive bird species. This denotes that a permit is required in order to own, import, trade or translocate this species, the species must be contained and fall under a government management programme (DEA 2016). According to data from the Southern African Bird Atlas Project (SABAP) their population has increased dramatically in the past 28 years, especially within urban and peri-urban areas (**Figure 1.2,** Berruti 1992, Owen *et al.* 2006, SABAP2 2019). However, these data come from public sightings and therefore reporting bias may be present due to the disproportionate reporting in urban areas as compared to more rural areas. Nonetheless, the population is increasing and this may, consequently, displace native waterbirds, through competition for food and breeding sites, and through hybridisation (Banks *et al.* 2008). Eradication programmes have been implemented; however, Mallards are still considered a major problem, especially within the Western Cape Province (Owen *et al.* 2006, Banks *et al.* 2008). There is anecdotal evidence that these Mallard populations hybridise with the native Yellow-billed Duck (*Anas undulata*) and produce fertile offspring (Milstein and Osterhoff 1975, Owen *et al.* 2006, Lavretsky *et al.* 2014).

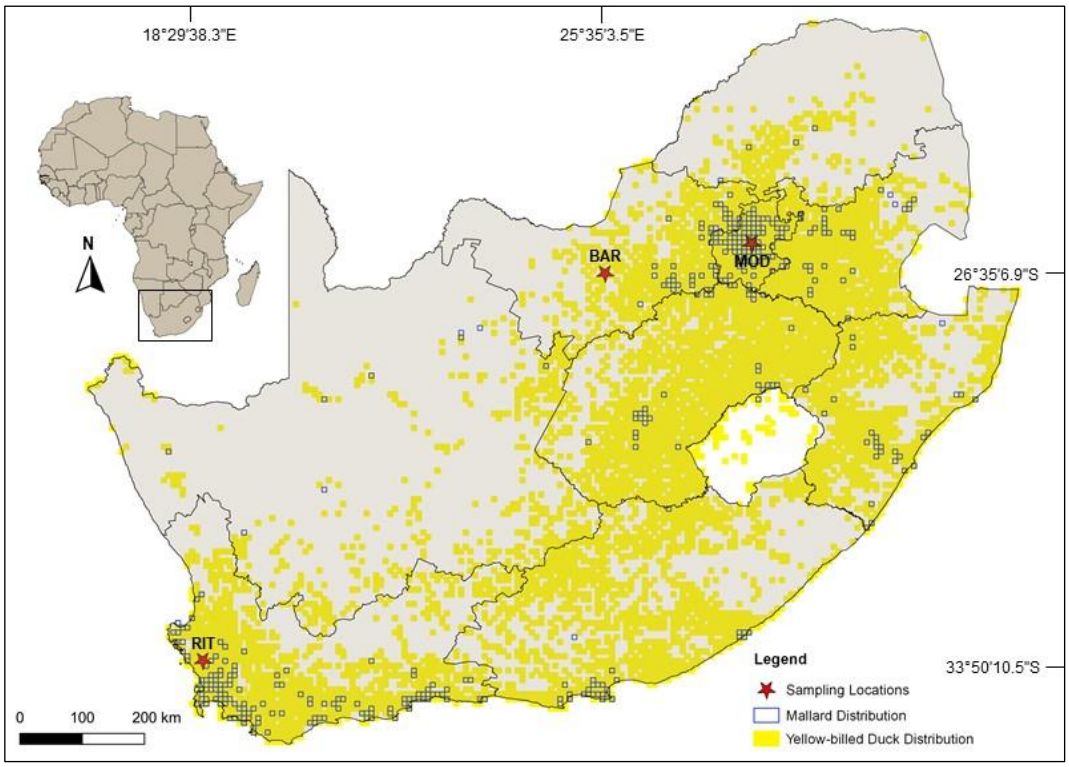


**Figure 1.2**: A map of South Africa showing the distribution of Mallards (*Anas platyrhynchos*) and Yellow-billed Ducks (*Anas undulata*), as well as the sampling localities of the study. Distribution data was obtained from the Southern African Bird Atlas Project 2 (SABAP2 2019). RIT - Rietvlei Wetland Reserve, BAR - Barberspan Nature Reserve and MOD - Modderfontein Nature Reserve.

The Yellow-billed Duck is a sexually monomorphic dabbling duck species native to southern and eastern Africa (Owen *et al.* 2006). There are two subspecies, the Northern Yellow-billed Duck (*A. u. rueppelli*) and the Southern Yellow-billed Duck (*A. u. undulata*, Sinclair *et al.* 2011). This species falls under the African-Eurasian Migratory Waterbirds Agreement (AEWA), whose main aim is to maintain migratory waterbird species in a favourable conservation status or to restore them to such a status (AEWA 2018). The Yellow-billed Duck is usually 52-58 cm in length and weighs about 700-1150 g, displays a spotted brown plumage, yellow bill and black legs, and exhibits a similar voice and behaviour to the Mallard (**Figure 1.3a,** Delacour and Mayr 1945, Sinclair *et al.* 2011). Orange colouration has been found in the usually black feet and legs of Yellow-billed Ducks, leading to the hypothesis that these ducks may be hybridising with Mallards (**Figure 1.3b**, Sinclair *et al.* 2011). In southern Africa, there is little known about hybridisation between native duck species in the wild (Harebottle and Vanderwalt 2014). Yet, there is anecdotal evidence that Yellow-billed Ducks have hybridised with both the

Red-billed Teal (*Anas erythrorhyncha*) and the Southern Pochard (*Netta erythrophthalma*) in South Africa (Milstein 1979, Sinclair *et al.* 2011). This suggests that there could be a potential for these species to also hybridise with the Mallard. The Mallard is also thought to be potentially hybridising with the African Black Duck (*Anas sparsa*) and Cape Shoveler (*Anas smithii*) in South Africa (Sinclair *et al.* 2011). Mallards are often thought to be hybrids due to the range of phenotypes that they may present.

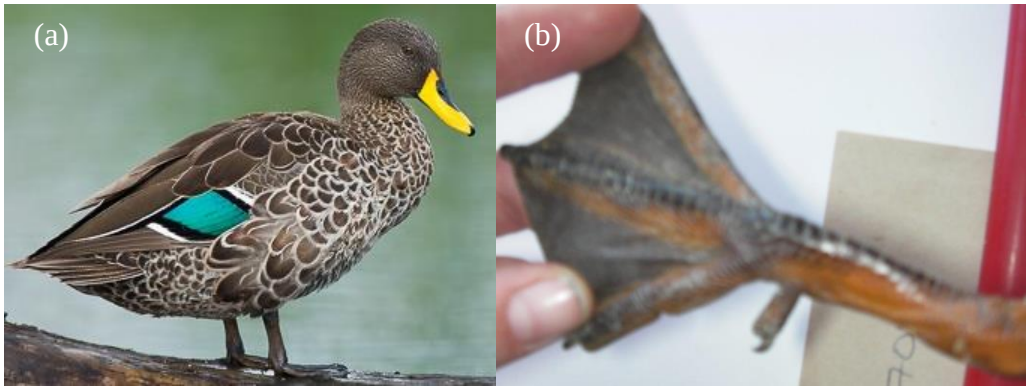


**Figure 1.3**: Photographic representation of (a) the Yellow-billed Duck (*Anas undulata*) and (b) the orange variation found in some of their feet and legs. Photographs accredited to (a) S Davey and (b) C Reynolds.

## 1.4 Managing hybridisation

The scenarios given above exemplify just some of the conservation concerns caused by anthropogenic hybridisation. Managing these hybridisation events can be crucial and often challenging, as hybridisation rates can rapidly increase over time (van Wyk *et al.* 2017). Four management strategies have been suggested by Guay *et al.* (2014) when dealing with hybridisation, however for these to be effective extensive evaluation of the situation should first take place before any management strategy is implemented. Firstly, prevention is usually the most cost-effective strategy, therefore the introduction and establishment of alien species into new environments should be prevented (Guay *et al.* 2014). Secondly, once an alien species has established, eradication of both the alien species and the hybrids should be implemented, particularly when hybridisation is localised (Rhymer 2006, Guay *et al.* 2014). Thirdly, containment may be another strategy to consider. This would be

employed when hybridisation is known to occur throughout sub-populations of native species and movements are manageable between the sub-populations (Allendorf *et al.* 2001, Guay *et al.* 2014). Finally, if hybridisation is widespread, there is little that can be done. In this situation, captive breeding programmes or translocations should be carefully considered (Rhymer 2006, Guay *et al.* 2014). Allendorf *et al.* (2001) also suggested that hybrids be protected in cases where they contain the only remaining genetic information of a taxon. Regular monitoring with molecular markers can provide early detection of hybridisation, allowing for more management options than is available if introgression is already extensive before detection (Schwartz *et al.* 2007).

## 1.5 Detection of hybridisation using molecular techniques

Hybrids were traditionally identified using intermediate phenotypic characteristics (e.g. Braithwaite and Miller 1975). While this method is simple in determining F1 hybrids, a disadvantage is the underestimation of the overall frequency of hybridisation, as after a few generations of backcrossing, hybrids may be indistinguishable from parent species (Braithwaite and Miller 1975, Guay *et al.* 2014). Today, molecular techniques are preferred as they allow for a more accurate estimation of the incidence of hybridisation.

Genetic methods may include the use of nuclear deoxyribose nucleic acids (DNA) and mitochondrial DNA (mtDNA) markers. Brown *et al.* (1979) found that mtDNA has a much higher mutation rate when compared to nuclear DNA. This allows for the accumulation of unique nucleotide variations between closely related species, and therefore can be used as a potential marker for phylogenetic analyses (Brown *et al.* 1979, Hebert *et al.* 2004). It was also found that mtDNA is more conserved within species and highly variable between species (Avise 1987, Hebert *et al.* 2004). Due to this, mtDNA is widely used to infer phylogenetic relationships, especially at the species level (Hebert *et al.* 2004). Mitochondrial DNA barcoding, amplifying and sequencing the 648 bp *cytochrome c oxidase I* (*COI*) mitochondrial gene, has been shown to be extremely successful in identifying and distinguishing between bird species (Hebert *et al.* 2004, Yoo *et al.* 2006, Tavares and Baker 2008).

Another important mitochondrial gene, in phylogenetic studies of species to family levels, is *cytochrome b* (*cyt b*, Kartavtsev and Lee 2006). This gene has also been found to contain phylogenetic information that is able to discern deep avian relationships, such as relationships between the orders Passeriformes, Anseriformes and Galliformes (Johnson 2001). Both *cyt b* and *COI* have proved to be advantageous over many other mitochondrial genes inferring divergencies up to the family level within many animal groups (Kartavtsev and Lee 2006). These genes have also proven to successfully separate species both inter- and intra-specifically (Su *et al.* 1999, Kartavtsev and Lee 2006, Arif *et al.* 2011).

It is thought that mtDNA haplotypes introgress preferentially into one of the hybridising species, as was found between Coyotes (*Canis latrans*) and Gray Wolves (*Canis lupus*) in Northern America (Lehman *et al.* 1991). Haldane's rule predicts that the heterogametic sex will be preferentially affected by hybrid incompatibilities, preventing mitochondrial transmission between species when the heterogametic sex is female (Mallet 2005, Wang *et al.* 2018). This could explain why mitochondrial introgression is uncommon in birds, where the heterogametic sex is the female. Despite this, mtDNA sequencing can be used in hybridisation studies due to the maternal inheritance patterns that allow the identification of the directionality of hybridisation, by identifying the maternal species of a hybrid individual (Watanabe *et al.* 1985). It should be noted that, although studies have shown paternal leakage of mtDNA (Kvist *et al.* 2003, Morgan *et al.* 2013, Luo *et al.* 2018), the characteristic that mtDNA inheritance is only maternal is currently still valid and the aforementioned studies are exceptions (Luo *et al.* 2018). If hybridisation is occurring between Mallard drakes and Yellow-billed Duck hens, then the mitochondrial haplotypes of the hybrids should resemble that of the Yellow-billed Duck. Mitochondrial DNA on its own provides very little information about hybridisation, however, when coupled with highly polymorphic loci, such as microsatellites, these markers together can provide credible evidence of introgression (Mallet 2005).

Microsatellites, a type of nuclear DNA marker, are abundant in genomes, highly reproduceable, easy to use and relatively cheap, making them ideal markers for genetic analyses (Bhargava and Fuentes 2010). These markers are thought to be

the most dependable method for identifying hybridisation. Microsatellite markers have high polymorphism rates, which allows for allele frequency differences to be detected between different species, and ultimately for the identification of hybrids (Väli *et al.* 2010). This method does not rely on phenotype, nor does it only measure hybridisation along the maternal line.

Microsatellites, also known as short tandem repeats or simple sequence repeats, are tandem repeats of 1 - 6 nucleotides, ranging between 5 and 40 repeats in length, and are widespread in both prokaryotes and eukaryotes (Selkoe and Toonen 2006, Bhargava and Fuentes 2010). Tetranucleotide repeats are generally the preferred type of microsatellite loci for genetic analyses as they give a high degree of error free data during allele scoring (Edwards *et al.* 1991). This study will utilise mostly dinucleotide repeat microsatellite loci. This type of marker has been suggested to be useful in molecular genetic studies of birds, as dinucleotide repeats occur frequently within avian genomes and are hypervariable (Ellegren 1992).

The mutation rate of microsatellites varies among loci, alleles and organisms, and depends on intrinsic features such as the repeat number, repeat type, perfection of repeat structure and the flanking sequence (Bhargava and Fuentes 2010). Microsatellite loci that have an increased number of tandem repeats, shorter repeating units, a decrease in repeat type, and a perfect repeat sequence have higher mutation rates. The repeat sequence of a perfect microsatellite is not interrupted by any base not belonging to the motif, whereas an imperfect microsatellite has bases between repeated motifs that do not belong (Bhargava and Fuentes 2010). Imperfect microsatellites have lower mutation rates and therefore decreased variability and often their alleles are more difficult to score, hence perfect microsatellites are preferred for genetic analyses. The flanking region is a single-copy DNA sequence directly upstream and downstream of the microsatellite loci (Goldstein and Schlotterer 1999). These regions are generally conserved within a species and are therefore used as the primer binding site (Selkoe and Toonen 2006). Mutations within these regions may also modify the mutation rate of the microsatellite loci (Bhargava and Fuentes 2010).

Although microsatellites are easy to use, they are not without their errors, as heterozygote individuals can sometimes go undetected. Null alleles are alleles that fail to amplify due to mutations in primer binding sites within flanking regions (Wattier *et al.* 1998, Guichoux *et al.* 2011). Short allele dominance can occur in heterozygotes containing a long and short allele, where only the short allele is detected (Wattier *et al.* 1998). Users should be cautious of these errors during allele scoring, and these errors should also be tested for after allele scoring and before further analyses.

Most hybridisation genetic studies utilise less than ten microsatellite loci, however, a comparative study done by Vähä and Primmer (2006) suggest that twelve loci with pairwise $F_{ST} = 0.21$ is necessary to accurately identify between first generation hybrids and pure species. An increased number of nuclear markers with species-specific allele frequencies are required if species have only recently diverged, if introgressive hybridisation has occurred over a long period, and if backcrosses and later-generation hybrids are to be distinguished (Vähä and Primmer 2006, Väli *et al.* 2010).

## 1.6 Project rationale

Biological invasions through the introduction of alien species is one of the many drivers negatively affecting biodiversity (Walthe *et al.* 2009). With globalisation and anthropogenic activities increasing, alien invasions are becoming ever more significant. It is therefore important that we study these invasive species in order to shed light on their current impacts on our biodiversity. Biodiversity is a vital asset for economics as it is connected to many ecosystem services (CBD 2014). Such services include: various jobs in agriculture; fisheries; forestry and more, supplying food; clean air and water, and supplying goods and services essential for human health (CBD 2014). Biodiversity is also a key element in sustainable development. It is therefore vital that we protect our biodiversity in South Africa (CBD 2014). The first goal, as part of the Convention on Biological Diversity's (CBD) Strategic Plan for Biodiversity 2011-2020, is to address underlying causes of biodiversity loss with biodiversity of inland waters as one of the main areas of interest (CBD

2014). Also, scientific monitoring serves as one of the founding elements in the South African National Biodiversity Institute's approach to protecting biodiversity as well as in the action plan of AEWA (AEWA 2018, SANBI 2019). Genetic monitoring to investigate the occurrence of hybridisation between the Mallard and Yellow-billed Duck in South Africa will serve as part of these goals in protecting our biodiversity and waterbirds.

## 1.7 Aims and objectives

The aim of this study was to estimate the presence and extent of hybridisation between Yellow-billed Ducks (*Anas undulata*) from North West Province, South Africa, and the introduced Mallard (*Anas platyrhynchos*) by assessing genetic variation in the study species using a reference population of each. Based on this aim, the objectives included 1) assessing and validating a set of microsatellite loci that can detect genetic variation between the Mallard and Yellow-billed Duck, 2) microsatellite genotyping to detect hybridisation between these duck species and 3) applying mitochondrial DNA markers to determine the direction of the detected hybridisation.

# Chapter 2

# Materials and methods to identify interspecific hybridisation

## 2.1 Introduction

Molecular methods have provided new tools with which to address many ecological and evolutionary questions. For example, molecular markers and DNA sequencing have led to the discovery of whole new groups of microbes, such as those related to *Acidobacterium capsulatum* (Gordon *et al.* 2000). DNA markers have also been used to determine the host distribution of herbivorous insects and mycorrhizae (Bruns *et al.* 1998, Jackson *et al.* 2002). Additionally, highly polymorphic markers have provided insights to patterns of gene flow and population partitioning in marine organisms such as Green Turtles (*Chelonia mydas*) (Avise 1998). And finally, genetic techniques have provided essential information in conservation biology, by analysing pedigree and mating systems of the Serengeti Cheetah (*Acinonyx jubatus*, Kelly 2000), and the tracing of the source and spread of introduced species, such as Cheatgrass (*Bromus tectorum*) in North America (Novak and Mack 2001).

Ecological questions concerning introgressive hybridisation can also be answered using molecular methods. Most data on hybridisation in wild animals comes from phenotypic evidence (Mallet 2005, Goulet *et al.* 2017). However, identification of hybrids using this technique is not accurate due to phenotypic plasticity and backcrossing (Valentini *et al.* 2009). Molecular genetic techniques have, therefore, become an important method for identifying hybrid species in ecological studies.

Although many techniques can be used, such as multilocus allozymes; cytology; restriction fragment length polymorphism or single nucleotide polymorphisms (SNPs), microsatellite markers are thought to be the most dependable method for identifying hybridisation due to the potential to determine relatedness of individuals (Selkoe and Toonen 2006, Guay *et al.* 2014).

Microsatellites have high mutation rates, generating high levels of allelic diversity (Selkoe and Toonen 2006). Using multiple polymorphic loci provides a genotypic profile of the different alleles that can be used to assign individuals to a species (Selkoe and Toonen 2006, Guay *et al.* 2014). van Wyk *et al.* (2013) used 13 polymorphic microsatellite loci and found a hybridisation incidence of 33%, occurring between the Blesbok (*Damaliscus pygargus phillipsi*) and endangered Bontebok (*Damaliscus pygargus pygargus*), in samples of unknown origin.

Furthermore, mtDNA can be used to determine the directionality of hybridisation (Watanabe *et al.* 1985). Mitochondrial DNA is maternally inherited and can therefore be used to determine which species is the maternal parent of a hybrid individual (Watanabe *et al.* 1985). Yamazaki *et al.* (2008) found directional introgression by hybridisation between Japanese Stream Toad (*Bufo torrenticola*) males and Japanese Common Toad (*Bufo japonicus formosus*) females, using the *cyt b* mitochondrial gene region. Similarly, Zaya *et al.* (2015) used a maternally inherited chloroplast DNA marker and found that introduced Oriental Bittersweet (*Celastrus orbiculatus*) males are hybridising with native American Bittersweet (*Celastrus scandens*) females. However, mtDNA alone provides little information about hybridisation and is therefore often used alongside microsatellite loci (Mallet 2005).

Dalton *et al.* (2014) used both mtDNA and nine microsatellite loci to confirm the presence of a Nyala (*Tragelaphus angasii*) x Greater Kudu (*Tragelaphus strepsiceros*) sterile hybrid within a game farm in the North West Province of South Africa. Dufresnes *et al.* (2015) also used both mitochondrial and microsatellite markers, which identified a hybrid swarm population of European Tree Frog (*Hyla arborea*) x Italian Tree Frog (*Hyla intermedia*) in the Grangettes Natural Reserve in Switzerland. Using both mitochondrial and microsatellite markers, another recent study by Miller *et al.* (2017) on Northern Spotted Owls (*Strix occidentalis caurina*) and California Spotted Owls (*Strix occidentalis occidentalis*) found that 50% of the individuals sampled within the species range overlap were F2 hybrids or backcrosses with bidirectional introgression.

Microsatellite and mtDNA markers have been used to study hybridisation between the Mallard (*Anas platyrhynchos*) and many other dabbling ducks, such as the Hawaiian Duck (*A. wyvilliana*, Fowler *et al.* 2009), Mottled Duck (*A. fulvigula*, Williams *et al.* 2005, Ford *et al.* 2017), American Black Duck (*A. rubripes*, Mank *et al.* 2004), Pacific Black Duck (*A. superciliosa*, Taysom *et al.* 2014), and Eastern Spot-Billed Duck (*A. zonorhyncha*, Kulikova *et al.* 2004).

Based on the success of microsatellite genotyping and mtDNA barcoding in identifying introgressive hybridisation, as illustrated by the aforementioned studies, these techniques will be used to verify the occurrence of introgressive hybridisation between the Mallard and Yellow-billed Duck (*A. undulata*) in South Africa. This chapter will deal solely with the molecular techniques used in this study.

## 2.2 Materials and methods

### 2.2.1 Sample collection

Blood samples (N = 48) were collected from phenotypic Yellow-billed Ducks (n = 20) and Mallards (n = 20, reference populations); and putative Mallard x Yellow-billed Duck hybrids (n = 8, **Appendix Figure A1**), as part of a previous study by Dr Chevonne Reynolds (2016, University of Cape Town Animal Research Ethics Committee, permit number 2011/V17/GC). Both Yellow-billed ducks and putative hybrids were sampled from Barberspan Nature Reserve, North West Province (26°35'6.9"S, 25°35'3.5"E, **Figure 1.2**) during May/June in 2013 and 2014. These live ducks were captured in baited funnel traps or mistnets, sampled and ringed (**Appendix Table A3**). The Mallards were culled from Rietvlei Wetland Reserve, Western Cape Province (33°50'10.5"S, 18°29'38.3"E, **Figure 1.2**) during January/February in 2015. A further eight blood samples were collected, by Prof. Craig Symes and Dr Jean Mollett, from phenotypic Mallards, that were sampled and ringed (ring numbers: 879551-879558) at Modderfontein Nature Reserve, Gauteng Province (26°4'47.2"S, 28°8'23.8"E) in 2015 (**Figure 1.2**, University of the Witwatersrand Animal Ethics Screening Committee, permit number: 2015/04/13/C).

Blood, from all samples, was collected from the brachial vein and stored in Queen's (Tris- ethylenediaminetetraacetic acid (EDTA), Sodium Chloride) lysis buffer at 4 °C. Queen's lysis buffer is a blood lysis buffer used as a DNA preservative. The buffer contains n-lauroylsarcosine, an anionic detergent, which solubilises membrane proteins, as well as EDTA to protect the DNA from degradation by sequestering divalent cations essential for deoxyribonuclease (DNase) activity (Marjorie and Borrelli 1958, Filip *et al.* 1973).

### 2.2.2 Total genomic DNA extraction

Total genomic DNA was extracted from blood samples using the phenol-chloroform methodology adapted from Mathew (1984). Samples were centrifuged (13 000 rpm for 6 minutes) in order to pellet the DNA and proteins. The supernatant was discarded and replaced with STE (Sodium Chloride, Tris-hydrochloric acid (HCl), EDTA) buffer, sodium dodecyl sulfate (SDS), proteinase K and ribonuclease (RNase), vortexed and incubated overnight at 56 °C. STE further lyses the cells, SDS denatures the proteins, proteinase K digests the proteins, and RNase removes ribonucleic acid. The lysate was then treated with phenol:chloroform:isoamylalcohol (25:24:1), to separate the proteins and DNA. The solution was vortexed and centrifuged (13 000 rpm for 8 minutes) to separate the mixture into the lower organic phase and the upper aqueous phase separated by a band of denatured protein (Mathew 1984). The aqueous phase, containing nucleic acids, was then recovered and the phenol:chloroform:isoamylalcohol step repeated in order to maximise DNA yield. The DNA was precipitated out using 95% ethanol and sodium acetate with an overnight incubation at -20 °C. The precipitated DNA was collected by centrifugation (13 000 rpm for 10 minutes) and subsequently washed in 70% ethanol. The DNA was then eluted in TE (Tris-HCl, EDTA) buffer and stored at 4 °C. TE is a common storage buffer containing Tris, a pH buffer, and EDTA to prevent DNA degradation (Marjorie and Borrelli 1958).

### 2.2.3 DNA visualisation and quantification

The DNA products were visualised through gel electrophoresis in order to validate successful extraction. Gel electrophoresis of nucleic acids involves the application of an electric field across a gel matrix through which DNA molecules can be separated based on their size (Yilmaz *et al.* 2012). Smaller molecules migrate faster than larger ones (Yilmaz *et al.* 2012). The extracted DNA was mixed with a loading dye and run on a 1.5% agarose gel at 100 V using TBE (Tris, Boric acid, EDTA) buffer. The loading dye adds colour and density to the sample so that it sinks to the bottom of the well and allows for tracking and visualisation of the sample on the gel (Lee *et al.* 2012). The fragments were viewed by fluorescence using ultraviolet (UV) irradiation on a Bio-Rad GelDoc XR Imager (Bio-Rad Laboratories Incorporated, California, USA). EZ-Vision In-Gel Solution (VWR Life Science AMRESCO, Ohio, USA) was used to stain the agarose gel. This is a dye that intercalates between the DNA bases allowing for fluorescent visualisation of the DNA bands when exposed to UV light (Lee *et al.* 2012).

The concentration and purity of the extracted DNA was quantified using a Nanodrop ND-1000 spectrophotometer (Inqaba Biotechnical Industries, Pretoria, SA) at a wavelength of 260 nm, 280 nm and 230 nm. DNA absorbs light at a wavelength of 260 nm and proteins absorb light at a wavelength of 280 nm (Cler *et al.* 2006). The DNA purity is provided as ratios of A260/280 and A260/230. Pure nucleic acids typically yield A260/280 values of 1.8 and A260/230 values in the range of 1.8-2.2 (Desjardins and Conklin 2010). Significant deviations from these values may indicate contamination by proteins, phenol or other co-purified products (Desjardins and Conklin 2010). The average DNA concentration, A260/280 and A260/230 purity ratios were 240.98 ng/$\mu$l; 1.69 and 1.61, respectively (**Appendix Table A4**).[1]

[1] Unfortunately, the 8 putative Mallards from Modderfontein Nature Reserve had to be omitted from the study, due to problems with amplification possibly stemming from incorrect blood to Queen's lysis buffer ratios.

### 2.2.4 Microsatellite loci analysis

#### 2.2.4.1 Validation

Twenty-seven microsatellite loci (**Table 2.1**), used by Taysom *et al.* (2014), were validated for cross-amplification across all study species. This was performed using polymerase chain reaction (PCR) on a subset of samples. The subset included one putative hybrid, two Yellow-billed Ducks and two Mallards.

PCR allows for large quantities of target DNA to be amplified due to repetitive cycles of template denaturing, primer annealing and elongation (Erlich 1989). PCR requires template DNA; specific forward and reverse primers, which will amplify the region of interest; a supply of deoxyribonucleotide triphosphates (dNTPs), for the extension of the newly synthesized DNA; *Taq* DNA polymerase; magnesium chloride, a cofactor for the *Taq* polymerase; and water. DNA polymerases are enzymes that synthesise new DNA strands (Erlich 1989). *Taq* polymerase, isolated from *Thermophilus aquaticus*, is a heat tolerant variant of DNA polymerase, and is therefore stable during the denaturation steps (Erlich 1989). Amplified target regions can then be genotyped or sequenced for further analyses.

PCRs, with a final volume of 25 $\mu$l, contained 1X Phusion High-Fidelity PCR Master Mix with HF Buffer (New England Biolabs Incorporated, Massachusetts, USA), 0.2 $\mu$M of both the forward and reverse primer, and 1 $\mu$l of genomic DNA. The reactions were heated to 98 °C for 30 seconds, followed by 35 cycles of 98 °C for 10 seconds, primer pair specific annealing temperatures for 30 seconds, 72 °C for 20 seconds, and a final extension step of 72 °C for 5 minutes on a GeneAmp PCR System 9700 (Applied Biosystems Incorporated, California, USA). Successful amplicons were visualised using gel electrophoresis as described in section 2.2.3. In order to obtain better resolution to determine any polymorphism of the loci, 4% agarose gels were used. Eighteen microsatellite loci successfully amplified in both the Mallard and Yellow-billed Duck (**Table 2.1**).

**Table 2.1**: Twenty-seven microsatellite loci, validated for cross-amplification across all study species, with their characteristics.

| Locus | Primer Sequence (5′-3′) | Motif | Ta (°C) | Size (bp) | Reference |
|---|---|---|---|---|---|
| **APH12** | **F: TTA GTA GCA TGT CAG GTT TAT T** | **$(GAAA)_4A_2(GAAA)_2$** | **52** | **165** | **Maak *et al.* 2003** |
| | **R: GCT TGT AGA CTT CAG AGT TC** | | | | |
| APH13 | F: CAA CGA GTG ACA ATG ATA AAA | $(GA)_{10}$ | 52 | 179 | Maak *et al.* 2003 |
| | R: CAA TGA TCT CAC TCC CAA TAG | | | | |
| APH15 | F: TGA ATA TGC GTG GCT GAA | $(CA)_9$ | 60 | 179 | Maak *et al.* 2003 |
| | R: CAG TGA GGA ATG TGT TTG AGT T | | | | |
| APH16 | F: CCT TCT GAA CCT TCG TAG | $(CA)_7$ | 52 | 146 | Maak *et al.* 2003 |
| | R: AAA TAT AGA CTT TTG TCC TGA A | | | | |
| ***APH17** | **F: GGA CAT TTT CAA CCA TAA ACT C** | **$(CA)_{14}$** | **60** | **222** | **Maak *et al.* 2003** |
| | **R: CAT CCA TGA CAG ACA GAA GA** | | | | |
| ***APH19** | **F: CAT GGA GCA AGC AAT CGT CTG** | **$(CA)_8$** | **54** | **166** | **Maak *et al.* 2003** |
| | **R: ACC ACG TCC ATC CTG AAG AAA** | | | | |
| ***APH20** | **F: ACC AGC CTA GCA AGC ACT GT** | **$(CA)_9$** | **58** | **150** | **Maak *et al.* 2003** |
| | **R: GAG GCT TTA GGA GAG ATT GAA AAA** | | | | |
| ***APH21** | **F: CTT AAA GCA AAG CGC ACG TC** | **$(CA)_8$** | **59** | **137** | **Maak *et al.* 2003** |
| | **R: AGA TGC CCA AAG TCT GTG CT** | | | | |
| **APH23** | **F: TCC TCT GCT CTA GTT GTG ATG G** | **$(CA)_4TA(CA)_{12}$** | **58** | **205** | **Maak *et al.* 2003** |
| | **R: CCT CAG CAG TCT TCC TCA GTG** | | | | |

Ta - Annealing temperature Continued on next page

Bold text represents loci that successfully cross-amplified

* represents loci selected for further analyses

**Table 2.1 – Continued from previous page**

| Locus | Primer Sequence (5′-3′) | Motif | Ta (°C) | Size (bp) | Reference |
|---|---|---|---|---|---|
| **APH24** | **F: TCA ACC AGT GGT CAG AGA AAA A** | **$(CA)_2TA(CA)_9$** | **58** | **147** | **Maak *et al.* 2003** |
| | **R: AGG TCA GCC CCC ATT TTA GT** | | | | |
| ***APH25** | **F: CCG TCA GAC TGT AGG GAA GG** | **$(GA)_9$** | **58** | **167-187** | **Maak *et al.* 2003** |
| | **R: AAA GCT CCA CAG AGG CAA AG** | | | | |
| ***Apl11** | **F: AAC TAC AGG GCA CCT TAT TTC C** | **$(GA)_{25}$** | **60** | **92-139** | **Denk *et al.* 2004** |
| | **R: TTG CAT CAG GGT CTG TAT TTT C** | | | | |
| Bca*μ*4 | F: ACA CAA CCT TCA AAG TCA ATC CAA T | $(CA)_{12}$ | 60 | 173 | Buchholz *et al*. 1998 |
| | R: TCC TGA CGC TCT CGG ACG AGT | | | | |
| ***Bca*μ*6** | **F: TTT AAC CCA GTA GCC TAT CAT GTC A** | **$(CA)_{10}$** | **60** | **151** | **Buchholz *et al.* 1998** |
| | **R: GTC TGA AGA TAA TGC TGC ATG GTT** | | | | |
| ***Bca*μ*11** | **F: TAG AAA AGG CTG AAG GAG TGG C** | **$(CA)_8$** | **55** | **134** | **Buchholz *et al.* 1998** |
| | **R: TGA GGA AGC AAC TGT AAA TAG GAG A** | | | | |
| **Blm5** | **F: GCC ACT TCT TTT GAA GTC ACC** | **$(GAAA)_3GATA(GAAA)_7$** | **55** | **231** | **Guay and Mulder 2005** |
| | **R: GAA GCA TCT TGT ATG GCT TGC** | | | | |
| Blm10 | F: CAA AGT ATA TCT TCT CAG GGA CAC G | $(GAAA)_{14}$ | 55 | 340-376 | Guay and Mulder 2005 |
| | R: TGC ATT GCT GTG AAG AGA CC | | | | |
| ***Blm12** | **F: TTC TGT GGG AGA AGA CAA AGG** | **$(CTTT)_{15}$** | **55** | **221-257** | **Guay and Mulder 2005** |
| | **R: CAC TTG CCT GCT TCA CTC C** | | | | |

Ta - Annealing temperature

Continued on next page

Bold text represents loci that successfully cross-amplified

* represents loci selected for further analyses

**Table 2.1 – Continued from previous page**

| Locus | Primer Sequence (5′-3′) | Motif | Ta (°C) | Size (bp) | Reference |
|---|---|---|---|---|---|
| Cam2 | F: TCC ACA AGG ACA CCA TTA GG | $(GA)_2GAAAAA(GAAA)_{15}$ | 54 | 180-218 | Carew *et al.* 2003 |
| | R: GGT ATT TCT TTT GC | | | | |
| **Cam3** | **F: AAC ATC TAC TTT GGC CTC TCC** | **$AAAAAG\ (GAAA)_2GAA(GAAA)_3$ $GAAG(GAAA)_8GAA(AAG)_2$** | **54** | **181-225** | **Carew *et al.* 2003** |
| | **R: TCT GTG CCC TGT TCT ACT GC** | | | | |
| Cam9 | F: AAT TGC AGC ACT AAT GAG C | $(A)_7(GAAA)_4(GGAAGAAA)_5(GGAA)_3$ | 52 | 260-296 | Carew *et al.* 2003 |
| | R: GCT CAT CAA TCA AAA CAT TCC | $(GAAAGGAA(GAAA)_2GAGAAAAGA$ | | | |
| ***CM09** | **F: GGA TGT TGC CCC ACA TAT TT** | **$(AT)_{16}$** | **55** | **110-140** | **Taysom *et al.* 2014** |
| | **R: TTG CCT TGT TTA TGA GCC AT** | | | | |
| CmAAT16 | F: TCC CAA GGG TAC CAG TGA A | $(AAT)_{10}$ | 55 | 171 | Stai and Hughes 2003 |
| | R: TGT TGG CTC CCT GCT TAA AT | | | | |
| JCC1 | F: GGA TTG GAG ATT TTC AGG AGC | - | 55 | 160 | Taysom *et al.* 2014 |
| | R: AGG GAA CTG ATG CCC CA | | | | |
| **Sfi*μ*2** | **F: ATA AAC GGC TAA TAT GAA GTC T** | **$(A)_{17}(CA)_{12}$** | **56** | **327** | **Fields and Scribner 1997** |
| | **R: AGG CTA GAT ATT GCT CTT ATC CT** | | | | |
| ***Sfi*μ*4** | **F: TGA GGG GGA AGA GAA TAA GAG A** | **$(GA)_{13}$** | **60** | **163** | **Fields and Scribner 1997** |
| | **R: CAG GGC AGT ATT TTC AGG ACA TT** | | | | |
| ***Smo7** | **F: TTT TCA CCC AGT TCA CTT CAG CC** | **$(GT)_{12}$** | **60** | **186-224** | **Paulus and Tiedemann 2003** |
| | **R: GAT TCA AAT TTG CCG CAG GAT TA** | | | | |

Ta - Annealing temperature

Bold text represents loci that successfully cross-amplified

* represents loci selected for further analyses

### 2.2.4.2 Amplification

Of the eighteen microsatellite loci that were found to cross-amplify, twelve were selected for further analyses (**Table 2.1**). The remaining six loci were omitted as they contained imperfect motifs. Of the twelve loci, eleven contained dinucleotide repeats, with Blm12 being the only locus to contain tetranucleotide repeats. The forward primers of each primer pair were synthesised with a 5′-M13 tag (5′-CACGACGTTGTAAAACGAC-3′) for use in the universal dye labelling method (Boutin-Ganache *et al.* 2001). A primary round of amplification was performed to amplify the loci of interest, this was then followed by a secondary round of amplification to fluorescently label the loci. The primary PCRs, with a final volume of 25 $\mu$l, contained 12 $\mu$l 1X Phusion Master Mix (New England Biolabs Incorporated, Massachusetts, USA), 0.2 $\mu$M of both the M13-forward and reverse primer, and 1 $\mu$l of genomic DNA. The secondary PCRs, with a final volume of 25 $\mu$l, contained 1X Phusion Master Mix (New England Biolabs Incorporated, Massachusetts, USA), 0.2 $\mu$M of the fluorescent dye (FAM, NED, PET or VIC; DS-33 Matrix Standard G5 Dye Set; Thermo Fisher Scientific Incorporated, Massachusetts, USA), 0.2 $\mu$M of the reverse primer, and 2 $\mu$l of PCR product from the primary round. PCR thermocycle conditions for both amplification cycles were set as described in the validation section. Successful amplification was determined by gel electrophoresis as mentioned in section 2.2.4.1.

### 2.2.4.3 Genotyping

Microsatellite loci are genotyped by fluorescently labelling the forward PCR primers, so that amplicons can be detected by laser-induced fluorescence on a capillary electrophoresis system (Guichoux *et al.* 2011). In diploid organisms, PCR will therefore produce a pair of fluorescent allelic products that may vary in size according to the number of microsatellite repeat units. Using different fluorescent colours and allele sizes to distinguish between fragments, it is possible to pool multiple markers in a single capillary, usually to a maximum of eight markers at a time (Guichoux *et al.* 2011).

Final PCR products were pooled into two panels based on the allele size range as well as the fluorescent dye used (**Table 2.2**). Pooled products were sent to Inqaba Biotechnical Industries, Pretoria, for fragment analyses. Fragment analysis occurs using an automated genetic analyser. Fluorescently labelled PCR products, mixed with a fluorescently labelled internal size standard, migrate through electrophoresis capillaries past a laser beam, exciting the fluorophores (Flores-rentería and Krohn 2013). As each dye emits light at a different wavelength when excited by the laser, all colours, and therefore many loci, can be detected and distinguished in one capillary injection. Emission spectra, from individual fluorophores, are separated by a diffraction system (Flores-rentería and Krohn 2013). A charge-coupled device camera detects the fluorescence signal and converts this into digital data (peaks) (Flores-rentería and Krohn 2013). These data are then stored in a file format that is compatible with an analysis software. Allele sizes were detected using a LIZ500 (Thermo Fisher Scientific Incorporated, Massachusetts, USA) internal size standard on an ABI 3500XL Genetic Analyzer (Applied Biosystems Incorporated, California, USA).

Alleles were scored manually using GeneMarker v2.7.0 (Liu *et al.* 2011). GeneMarker is a highly accurate, user-friendly genotyping software that allows for the visualisation of peaks to determine allele sizes of multiple samples at once (Liu *et al.* 2011). To reduce any bias or errors, allele scoring was done by two researchers independently. Any discordant alleles were then scored by a third party. Due to the amount of missing data contained within the APH20 locus, this locus was omitted from further analyses. MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004) was used to check for genotyping errors. MICRO-CHECKER can identify genotyping errors caused by null alleles, short allele dominance or stuttering (Van Oosterhout *et al.* 2004).

**Table 2.2**: Pooling of microsatellite loci PCR products into two panels (A and B) based on the allele size and the fluorescent dye. The colours represent the fluorescent dye used for each locus, FAM: blue, NED: yellow, PET: red and VIC: green.

**Panel A**

| Locus | Size (bp) | 90 | 100 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| APH19 | 166 | | | | | | | | | | | | | | | | | |
| Bca$\mu$11 | 134 | | | | | | | | | | | | | | | | | |
| Blm12 | 221-257 | | | | | | | | | | | | | | | | | |
| CM09 | 110 | | | | | | | | | | | | | | | | | |
| APH20 | 150 | | | | | | | | | | | | | | | | | |
| Apl11 | 92-139 | | | | | | | | | | | | | | | | | |
| Sfi$\mu$4 | 163 | | | | | | | | | | | | | | | | | |

**Panel B**

| Locus | Size (bp) | 90 | 100 | 110 | 120 | 130 | 140 | 150 | 160 | 170 | 180 | 190 | 200 | 210 | 220 | 230 | 240 | 250 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| APH25 | 167 - 187 | | | | | | | | | | | | | | | | | |
| APH21 | 137 | | | | | | | | | | | | | | | | | |
| Bca$\mu$6 | 151 | | | | | | | | | | | | | | | | | |
| APH17 | 222 | | | | | | | | | | | | | | | | | |
| Smo7 | 186 - 224 | | | | | | | | | | | | | | | | | |

#### 2.2.4.4 Genetic diversity analyses

Levels of genetic diversity were estimated using Arlequin v3.1 (Excoffier *et al.* 2005). Arlequin determined the mean number of alleles per locus (A), observed and expected heterozygosities, deviations from Hardy-Weinberg Equilibrium (HWE), linkage disequilibrium (LD) and F-statistics ($F_{ST}$ and $F_{IS}$). LD is defined as the non-random association of alleles at different loci (VanLiere and Rosenberg 2008). Typically, population genetic models exhibit less complicated behaviour when there is no LD (VanLiere and Rosenberg 2008). For biallelic markers, one of the most commonly used measures for LD is $r^2$, the squared coefficient of determination (VanLiere and Rosenberg 2008). The $r^2$ statistic ranges from 0–1, where 1 indicates complete LD, and 0 indicates complete linkage equilibrium. The $r^2$ statistic is commonly used to help identify informative markers, and generally has a threshold of 0.8 (VanLiere and Rosenberg 2008). This means that any two markers with an $r^2 \geq 0.8$ is considered to be in LD.

F-statistics describe the partitioning of genetic diversity within and among populations (Holsinger and Weir 2009). F-statistics include both $F_{ST}$ (fixation index), which measures the amount of genetic differentiation among populations and $F_{IS}$, which measures the inbreeding coefficient within a population (Holsinger and Weir 2009). The F-statistics can be estimated using an analysis of molecular variance (AMOVA). Excoffier *et al.* (1992) introduced AMOVA for analysis of haplotype variation, however it can also analyse variation within microsatellite loci and SNP data. The analysis is based on differences in the number of repeats between alleles at each microsatellite locus (Holsinger and Weir 2009). The value of $F_{IS}$ ranges between -1 and +1. Negative $F_{IS}$ values indicate heterozygote excess (outbreeding) and positive values indicate heterozygote deficiency (inbreeding) compared with HWE expectations (Wright 1965). $F_{ST}$ is always positive and ranges between zero (panmixis) and one (complete isolation) (Wright 1965). $F_{ST}$ values between 0 and 0.05 indicate slight population differentiation, values between 0.05 and 0.15 indicate moderate differentiation, values between 0.15 and 0.25 indicate strong differentiation, and values above 0.25 indicate very strong differentiation (Wright 1978).

### 2.2.4.5 Admixture analyses

Identification of any hybrid individuals was done using three different approaches: 1) principal component analysis (PCA), and two Bayesian clustering programmes, 2) STRUCTURE v2.3.4 (Pritchard *et al.* 2000) and 3) NEWHYBRIDS v1.1 (Anderson and Thompson 2002).

Relationships between individuals was inferred using a PCA of a pairwise, individual-by-individual, covariance matrix without spatial components nor a priori population assignment in R v3.5.1 (R Core Team 2018) using the adegenet package (Jombart 2008). The final PCA plot was composed using the ggplot2 package (Wickham 2016) in R. PCA is a multivariate statistical technique often used to find relationships between objects, by reducing the dimensionality of observations whilst retaining most of the variation within a dataset (Wold *et al.* 1987, Abdi and Williams 2010).

STRUCTURE was used to assign individuals to a particular genetic cluster and determine any admixture. In order to do this, STRUCTURE uses a Bayesian clustering approach, which assumes a model with K populations, each characterized by a set of allele frequencies at each locus (Pritchard *et al.* 2000). The programme also assumes that the loci used are not linked and that the populations are in Hardy-Weinberg Equilibrium (Pritchard *et al.* 2000). Assignments were conducted using the ADMIXTURE model with no prior population information. Following the methods by Evanno *et al.* (2005) to determine the optimum number of clusters, STRUCTURE was run for twenty replicates from K=1–10, with a run-length of 50,000 repetitions of Markov chain Monte Carlo (MCMC), following the burn-in period of 10,000 iterations. The twenty values for the estimated ln(Pr(X—K)) were averaged, from which ΔK was calculated (Evanno *et al.* 2005), using StructureSelector (Li and Liu 2018). The K value with the largest ΔK was identified as the optimum number of clusters. The assessment was then repeated ten times using the identified optimum number of clusters, K=2, with a run-length of 500,000 repetitions of MCMC, following the burn-in period of 100,000 iterations. Longer run-lengths and burn-ins were used to ensure more accurate results. CLUMPP

v1.2.2 (Jakobsson and Rosenberg 2007) was used to find the optimal assignment probabilities from the ten repetitions using the FullSearch algorithm. The final structure plot was composed using the ggplot2 package in R. Following Williams *et al.* (2005) and Väli *et al.* (2010) individuals were assigned to the following groups based on their q (assignment probability) value: Mallard, $q < 0.10$; hybrid, $0.15 < q < 0.85$; Yellow-billed Duck, $q > 0.90$.

NEWHYBRIDS can identify hybrid individuals beyond the F1 generation. The programme does so using a Bayesian statistical method that estimates the posterior probability that an individual belongs to each of six classes (Anderson and Thompson 2002). These six classes include each parental species, F1 hybrids, F2 hybrids and back-crosses to each parental species. The program considers prior information on allele frequencies and the amount of admixture. In the absence of this information, both the Jeffreys-like and Uniform priors were used. These priors are used to infer allele frequencies within the populations. The Uniform prior assumes that at least one copy of every allele has been found in both of the populations, therefore down-weighing the influence of private alleles, conversely, the Jeffreys-like prior does not assume this (Anderson 2003). The assessment ran for ten repetitions with a run-length of 100,000 repetitions of MCMC, following the burn-in period of 100,000 iterations. CLUMPP was used to find the optimal posterior probabilities from the ten repetitions using the Greedy algorithm with 1000 random input orders. The final structure plot was composed using the ggplot2 package in R. Individuals were assigned to one of the aforementioned classes based on their q (posterior probability) value with a threshold of 0.5. This means that individuals that presented with at least one q-value above 0.5 were assigned into that particular group, while those with posterior probabilities less than 0.5 remained unidentified (Väli *et al.* 2010).

### 2.2.5 Mitochondrial DNA analysis

#### 2.2.5.1 Amplification and visualisation

Mitochondrial DNA analysis was done using two primer sets (**Table 2.3**) following the methodology used by Hebert *et al.* (2004) to amplify the *cyt b* and *COI* gene regions. PCRs, with a final volume of 25 $\mu$l, contained 1.25 U OneTaq DNA Polymerase (New England Biolabs Incorporated, Massachusetts, USA), 1X OneTaq Standard Reaction Buffer (New England Biolabs Incorporated, Massachusetts, USA), 200 $\mu$M dNTPs, 1 $\mu$M of each forward and reverse primer and 1 $\mu$l of genomic DNA. PCRs were run using a GeneAmp PCR System 9700 (Applied Biosystems Incorporated, California, USA) with the following cycle conditions: a denaturation step at 94 °C for 60 seconds, followed by 5 cycles of denaturation at 94 °C for 60 seconds, annealing at 45 °C for 1.5 minutes and elongation at 72 °C for 1.5 minutes, followed by 30 cycles of 94 °C for 1 minute, 54 °C for 1.5 minutes, 72 °C for 1.5 minutes and a final elongation step of 72 °C for 5 minutes. PCR products were then visualised by gel electrophoresis, as described in section 2.2.3, in order to verify amplification of the correct gene regions.

**Table 2.3**: Mitochondrial gene regions, *cytochrome b* (*cyt b*) and *cytochrome c oxidase subunit I* (*COI*) with their characteristics, used in this study for phylogenetic and haplotype analyses.

| Gene | Primer | Primer Sequence (5′-3′) | Ta (°C) | Size (bp) |
|---|---|---|---|---|
| *cyt b* | L14816 | F: CCATCCAACATCTCAGCATGATGAAA | 54 | 358 |
| | H15173 | R: CCACTCAGAATGATATTTGTCCTCA | | |
| *COI* | BirdF1 | F: TTCTCCAACCACAAAGACATTGGCAC | 54 | 751 |
| | BirdR2 | R: ACTACATGTGAGATGATTCCGAATCCAG | | |

Ta – Annealing temperature

#### 2.2.5.2 Sequencing

The PCR products were sent to Inqaba Biotechnical Industries, Pretoria, for forward, Sanger, sequencing using an ABI 3500XL Genetic Analyzer (Applied Biosystems Incorporated, California, USA).

Sanger sequencing is a chain-termination sequencing method that uses dideoxynucleotide triphosphates (ddNTPs) each fluorescently labelled with a unique colour (Sanger *et al.* 1977, Karger *et al.* 1991). When these ddNTPs are incorporated by DNA polymerase they cease termination, as they lack a 3′-OH group, required to form phosphodiester bonds between successive nucleotides (Sanger *et al.* 1977). Cycle sequencing is performed, using both the standard four dNTPs and ddNTPs, in order to synthesise numerous DNA strands that are complementary to the template (Murphy *et al.* 2005). The DNA polymerase randomly incorporates either of the two types of nucleotides, resulting in all possible product lengths. These DNA molecules are then separated by size using capillary electrophoresis (Karger *et al.* 1991). As the molecules migrate through the capillary gel, they pass a laser beam that excites the fluorophore, emitting a specific wavelength (Karger *et al.* 1991). Computer software is then used to interpret these emission spectra into a DNA sequence trace chromatogram (Karger *et al.* 1991).

Chromatograms, with Phred quality scores, $Q \geq 20$, were produced. A Phred quality score is a measure of the quality of the identification of the nucleobases generated by automated DNA sequencing (Ewing *et al.* 1998). Sequence chromatograms were edited and aligned using Molecular Evolutionary Genetics Analysis (MEGA) v7.0.14 (Kumar *et al.* 2016). Sequences for *cyt b* and *COI* were 251 bp and 634 bp, respectively. All sequences have been submitted to GenBank (Accession numbers: MG654787-MG654810 and MK129186–MK129241).

#### 2.2.5.3 Phylogenetic analysis

Mallard sequences, from Asia, were obtained from GenBank to act as positive controls and to better resolve relationships within the dataset.

MEGA was used for the phylogenetic analysis. Sequences were aligned using MUSCLE (Edgar 2004). MUSCLE is a multiple sequence alignment tool that is often quicker and more accurate than CLUSTALW, as the algorithm includes a distance estimation using k-mer counting, progressive alignment using a log-

expectation score, and refinement using tree-dependent restricted partitioning (Edgar 2004).

The Hasegawa-Kishino-Yano (HKY) evolutionary model proposed by Hasegawa *et al.* (1985) was used based on a model test conducted through MEGA. Models of DNA evolution are used to estimate evolutionary distances between sequences in phylogenetic analyses. Choosing inappropriate evolutionary models can affect the outcome of a phylogenetic analysis by incorrectly estimating tree topology, influencing the estimation of branch length, and biasing statistical support values (Shapiro *et al.* 2006). Many models of DNA sequence evolution have been proposed, each differing in the parameters describing the rate of nucleotide substitution (Shapiro *et al.* 2006). Generally, transition substitutions are more likely in nucleotide evolution than transversions (Hasegawa *et al.* 1985). Therefore, the HKY model distinguishes between the rate of transitions and transversions and allows for unequal base frequencies (Hasegawa *et al.* 1985). The best model (HKY + G) was selected based on the Bayesian Information Criterion (BIC, **Appendix Table A5**).

Methods for model selection in phylogenetics include the hierarchical likelihood ratio tests (Posada and Buckley 2004). When Bayesian methods are used, the result is to choose the model with the highest posterior probability, however, this method requires computationally exhaustive techniques, such as MCMC (Posada and Buckley 2004). The BIC was, therefore, developed as an approximation to the log marginal likelihood of a model, thus, making this method less computationally intensive (Posada and Buckley 2004). Choosing the model with the smallest BIC is equivalent to selecting the model with the maximum posterior probability (Posada and Buckley 2004).

A maximum-likelihood phylogenetic tree was inferred using 1000 bootstrap replicates. Bootstrapping statistics were calculated for the evaluation of the reliability of the inferred clades, where a high bootstrap value (>50) provides more confidence that the branch point is correct (Felsenstein 1985). The Baikal teal (*Anas formosa*) was used as the outgroup, which also belongs to the Anatidae family, but is not part of the species found within the Mallard complex (Lavretsky *et al.* 2014).

A concatenated dataset was used, as datasets for the individual gene regions, *cyt b* and *COI*, produced congruent phylogenetic trees (**Appendix Figure A2 & A3**).

#### 2.2.5.4 Haplotype network analysis

Population Analysis with Reticulate Trees (PopART) v1.7 (Leigh and Bryant 2015) was used to construct a minimum spanning haplotype network, with ε = 0, using the same concatenated sequence dataset used for the phylogenetic analysis. Haplotype networks are commonly used to illustrate relationships between genotypes in a given dataset (Leigh and Bryant 2015). Networks often better represent these relationships, compared to tree-formats, as they do not limit the connections to linear, bifurcating nodes and they show the number of base-pair changes between sequences (Clement *et al.* 2000). A minimum spanning network is one of several methods for inferring haplotype networks. This method connects all given sequence types without inferring any additional (ancestral) nodes (Bandelt *et al.* 1999). Therefore, the total length of the network will be the sum of all the distances between linked sequence types. This uses a distance value, ε, between sequences. An increase in ε adds links between all sequence types from different connected components that are at that specified ε value (Bandelt *et al.* 1999). This will, therefore, create more vectors linking nodes together.

The following diagram (**Figure 2.1)** summarises the methods discussed in this chapter.

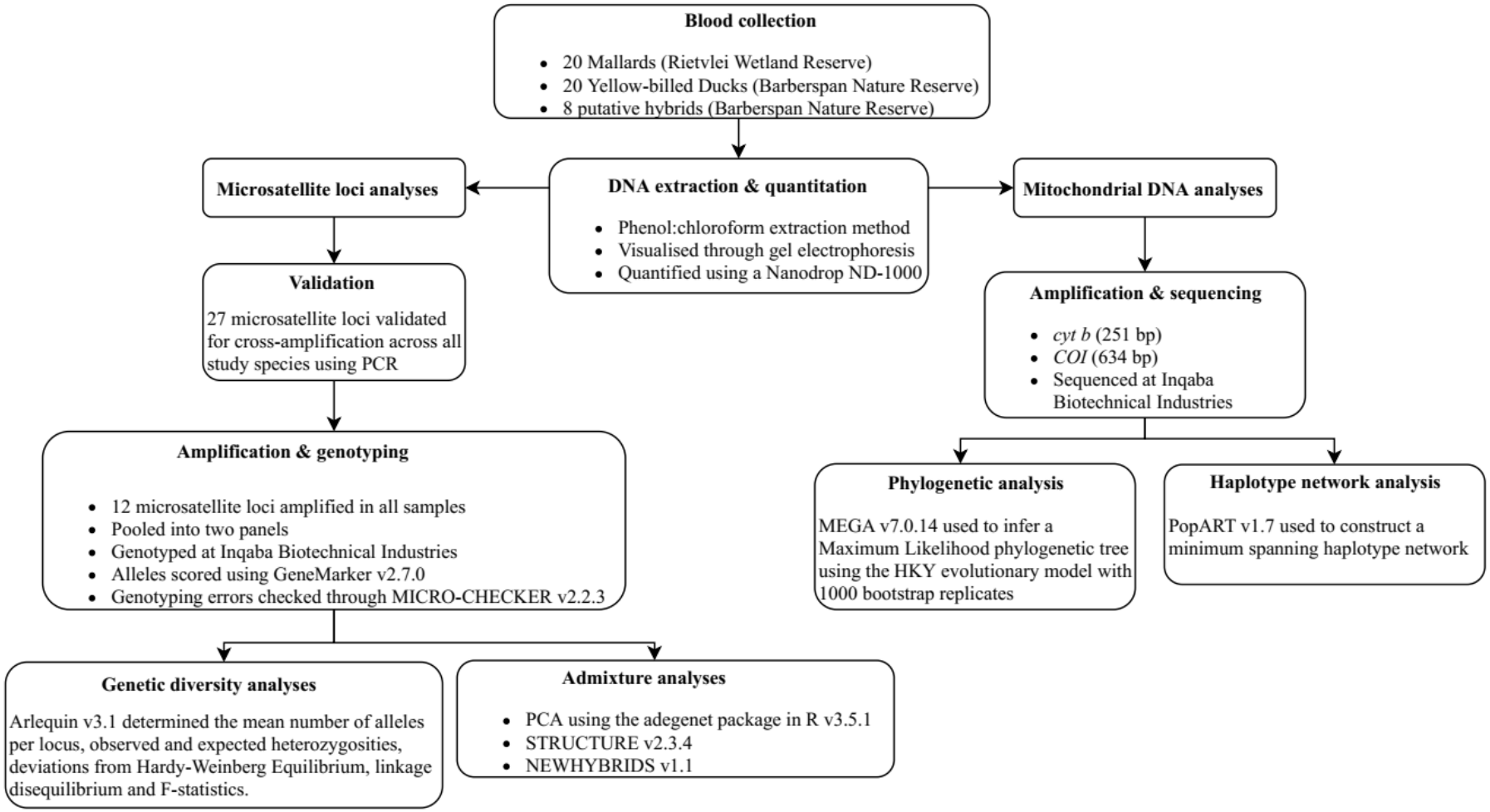


**Figure 2.1**: A summary of the methods and work flow of this study.

# Chapter 3

# Minimal evidence of interspecific hybridisation between the Mallard and Yellow-billed Duck in central and north-west South Africa[2]


## Abstract

Hybridisation is the interbreeding of genetically distinct groups and can lead to introgression - the exchange of genetic material between species. Hybridisation is of conservation concern when alien invasive species are involved, as it can lead to a loss of local genetic adaptations and genetic diversity. Hybridisation is a significant threat for many dabbling ducks where interbreeding with the closely related invasive Mallard (*Anas platyrhynchos*) is extremely common. Phenotypic evidence suggests that Mallard populations in South Africa hybridise with the indigenous Yellow-billed Duck (*Anas undulata*). The aim of this study was to estimate the presence and extent of hybridisation between Yellow-billed Ducks, representing central and north-west South Africa, and introduced Mallards. Genetic variation between Mallards, Yellow-billed Ducks, and their putative hybrids was assessed using mitochondrial and microsatellite DNA markers. All samples inferred to be hybrids based on phenotype were found to have Yellow-billed Duck mitochondrial DNA and show minimal evidence of admixture across the microsatellite markers. Thus, these results do not support the notion that hybridisation between Mallards and Yellow-billed Ducks is prevalent in central and north-west South Africa. However, hybridisation could be occurring within the Western Cape Province where Mallards are found in higher abundance. Therefore, continued monitoring of this potential hybridisation should be performed frequently and throughout South Africa.


[2] The manuscript of this chapter was submitted to Ostrich: Journal of African Ornithology (Manuscript ID: TOST-2019-0009) on the 8th February 2019. A revised manuscript according to reviewers' comments (ID TOST-2019-20028) was resubmitted on the 17th of April 2019.

## 3.1 Introduction

Anthropogenic activity has facilitated biological invasions, causing changes in the abundance and distribution of species and removing barriers that once isolated species (Chown *et al.* 2015). This creates sympatry between species, which can promote hybridisation (Rhymer and Simberloff 1996, Allendorf *et al.* 2001). Hybridisation is the interbreeding of genetically distinct groups and can lead to introgression - the exchange of genetic material between species (Rhymer and Simberloff 1996, Rhymer 2006, Guay *et al.* 2014). Hybridisation can occur naturally or be human induced through species introductions and environmental change (Malukiewicz *et al.* 2015). The latter is known as anthropogenic hybridisation.

Hybridisation is often a natural phenomenon, especially amongst Anseriformes, and usually has positive evolutionary consequences ultimately allowing for adaptability and diversification of species (Mallet 2005, McCarthy 2006). However, anthropogenic hybridisation, depending on the prevalence, can also have negative impacts leading to losses in local genetic adaptations and diversity (Josefsson and Andersson 2001). This may be due to introgression causing defective modifications of gene pools, homogenization, genetic swamping and ultimately the loss of distinctly recognised species (Rhymer and Simberloff 1996, Allendorf *et al.* 2001). Many examples of this have been shown in the literature. For example; the hybrids of Red Hartebeest (*Alcelaphus buselaphus*) and Blesbok (*Damaliscus dorcas*) are sterile, contributing to the decline of both these species (Robinson *et al.* 1991, Rhymer and Simberloff 1996). Additionally, Plains Leopard Frog (*Rana blairi*) and Southern Leopard Frog (*Rana sphenocephala*) hybrids in south eastern Missouri, were found to be significantly more infected by chytrid fungi (*Batrachochytrium dendrobatidis*) compared to the parental species (Parris 2004). Finally, the introduction of the Italian Tree Frog (*Hyla intermedia*) into the Grangettes natural reserve has allowed for introgressive hybridisation with the native European Tree Frog (*Hyla arborea*), resulting in a hybrid swarm (Dufresnes *et al.* 2015).

The Mallard (*Anas platyrhynchos*, **Chapter 1 Figure 1.1a and 1.1b**), is a sexually dimorphic dabbling duck (tribe Anatini) native to the northern hemisphere (Kulikova *et al.* 2005). Presently, this species has a near global distribution as they have been introduced purposefully into countries worldwide for hunting and aesthetic purposes (Desforges and Wood-Gush 1975, Rhymer 2006). These introduced Mallards are now of great threat to many endemic dabbling ducks due to hybridisation (Williams and Basse 2006, Taysom *et al.* 2014). Mallards were recently introduced (in the 1940's) into the Western Cape Province, South Africa, and have subsequently become established, especially within urban and peri-urban areas (**Chapter 1 Figure 1.2**, Brooke and Siegfried 1991, Berruti 1992, Owen *et al.* 2006). Mallards are, on average, larger; more aggressive; and more dominant over other duck species, possibly allowing them to sexually outcompete these species (Williams and Basse 2006). Furthermore, it has been suggested that the ornate plumage of Mallard drakes provide a supernormal stimulus to hens of other species, thereby favouring hybridisation (Guay *et al.* 2014). Domesticated Mallards have been shown to have increased testosterone, longer breeding seasons, increased promiscuity and force extra-pair copulations, therefore possibly promoting hybridisation (Desforges and Wood-Gush 1976, Hepp *et al.* 1988). As such, Mallards are known to hybridise with approximately 60 Anatidae species, causing large genetic effects, population declines and local extinctions in many of them (McCarthy 2006, Guay *et al.* 2014). For example, in eastern North America habitat alteration, due to agriculture, allowed for the sympatry between Mallards and American Black Ducks (*Anas rubripes*, Rhymer 2006). Extensive introgressive hybridisation between these two species has resulted in a loss of genetic differentiation between them. Mank *et al.* (2004) found that the measure of genetic divergence between populations ($G_{ST}$) decreased from 0.1460, for Mallard and American Black Ducks in 1940, to 0.0008 for samples collected in 1998. Additionally, the introduction of Mallards onto Lord Howe Island has resulted in the local extinction of the native New Zealand Grey Duck (*Anas superciliosa superciliosa*), due to extensive hybridisation (Guay *et al.* 2014). Furthermore, on all but one of the Hawaiian Islands (Kaua'i), the endemic Hawaiian Duck (*Anas*

*wyvilliana*) has gone extinct due to hybridisation with the introduced Mallard (Rhymer 2006, Fowler *et al.* 2009).

The Yellow-billed Duck (**Chapter 1 Figure 1.3a**) is a monomorphic dabbling duck species native to southern and eastern Africa and is relatively common throughout South Africa (**Chapter 1 Figure 1.2**, Owen *et al.* 2006). There is anecdotal evidence that Mallards are hybridising with the indigenous Yellow-billed Duck (*Anas undulata*) and producing fertile offspring (Milstein and Osterhoff 1975, Owen *et al.* 2006, Lavretsky *et al.* 2014). For example, Yellow-billed ducks from Barberspan Nature Reserve were found to have orange colouration (speculated to be from the Mallard) in their feet and legs (**Chapter 1 Figure 1.3b**, Reynolds 2016), leading the authors, of this paper, and Sinclair *et al*. (2011) to speculate that these birds were Mallard x Yellow-billed Duck hybrids. To date, no molecular evidence of hybridisation between the Mallard and Yellow-billed Duck has been found. Therefore, the aim of this study was to estimate the presence and extent of hybridisation between Yellow-billed Ducks representing central and north-west South Africa, and introduced Mallards. To do so, we assayed 12 microsatellite loci and two mitochondrial genes across a reference population of each species, and their putative hybrids to determine whether inference of hybrid status based on phenotype (i.e., feet coloration) is correct.

## 3.2 Materials and Methods

### 3.2.1 Study system

Barberspan Nature Reserve is listed as an Important Bird and Biodiversity Area (IBA) of South Africa (BirdLife South Africa 2019). The reserve contains a large permanent alkaline pan located in the North West Province, South Africa (BirdLife South Africa 2019). The pan is important for waterbirds, particularly as a moulting site (BirdLife South Africa 2019). Waterbirds congregate in large numbers here during moult when smaller wetlands in the surrounding central and north-west districts are drying up (BirdLife South Africa 2019). Ring and recovery data for Yellow-billed ducks, that are ringed at Barberspan, show that each year many of

these ducks return to the same residential site that is usually within about a 400 km radius from Barberspan (SAFRING 2019). Therefore, ducks sampled within Barberspan during moult likely represent the collective central and north-west regions of South Africa.

Mallards are 50-65 cm in length and weigh 850-1400 g (Madge and Burn 1988, Sinclair *et al.* 2011). Distinct phenotypes of pure female and male Mallards can be seen in **Chapter 1 Figure 1.1a** and **Figure 1.1b**, respectively. Both males and females have orange feet and legs (Madge and Burn 1988). When females are incubating eggs, males gather in small flocks and migrate to moulting sites where they are flightless for approximately 4 weeks (Madge and Burn 1988). They are omnivorous and opportunistic with a diverse diet, consisting of seeds, plants, invertebrates and sometimes amphibians and fish (Madge and Burn 1988).

Yellow-billed Ducks are 52-58 cm in length and weigh 700-1150 g, and have grey-black feet and legs (**Figure 2b**, Kear 2005, Sinclair *et al.* 2011). Adults undergo a post-nuptial wing moult three or four months after the peak of the breeding season, rendering them flightless for about four weeks (Kear 2005). They are omnivorous with a large diet consisting of fruits, seeds, plants, invertebrates and agricultural grains (Kear 2005).

### 3.2.2 Sample collection and DNA isolation

Blood samples (N = 48) were collected from phenotypic Yellow-billed Ducks (n = 20) and Mallards (n = 20, reference populations); and putative Mallard x Yellow-billed Duck hybrids (n = 8), as part of a previous study by Reynolds (2016) (University of Cape Town Animal Research Ethics Committee and University of the Witwatersrand Animal Ethics Screening Committee, permit numbers 2011/V17/GC and 2015/04/13/C, respectively). Putative hybrids were identified purely based on the orange colouration found within their feet and legs. Reference populations were assigned based on individuals that represented phenotypically pure Yellow-billed Ducks and Mallards. Wild Yellow-billed ducks and putative hybrids were captured, in baited funnel traps or mistnets, sampled, ringed, and

released in Barberspan Nature Reserve, North West Province (26°35'6.9"S, 25°35'3.5"E, **Chapter 1 Figure 1.2**) during May/June in 2013 and 2014. Wild Mallards were culled and sampled from Rietvlei Wetland Reserve, Western Cape Province (33°50'10.5"S, 18°29'38.3"E, **Chapter 1 Figure 1.2**) during January/February in 2015. Using a gauge needle, blood, from all samples, was collected from the brachial vein and stored in Queen's (Tris-EDTA, Sodium Chloride) lysis buffer at 4 °C. Total genomic DNA was extracted from the blood samples using the phenol-chloroform methodology adapted from Mathew (1984).

### 3.2.3 Microsatellite amplification and validation

Twelve microsatellite loci (**Chapter 2 Table 2.1**) cross amplified in Yellow-billed Ducks, Mallards and putative hybrids. Polymerase chain reactions (PCRs), with a final volume of 25 $\mu$l, contained 1X Phusion High-Fidelity PCR Master Mix with HF Buffer (New England Biolabs Incorporated), 0.2 $\mu$M of both the forward and reverse primer, and 1 $\mu$l of genomic DNA. The reactions were heated to 98 ℃ for 30 seconds, followed by 35 cycles of 98 ℃ for 10 seconds, primer pair specific annealing temperatures for 30 seconds, 72 ℃ for 20 seconds, and a final extension step of 72 ℃ for 5 minutes on a GeneAmp PCR System 9700 (Applied Biosystems Incorporated). Successful amplicons were visualised through gel electrophoresis using 4% agarose gels. PCR products were pooled together, into two panels, and run against a LIZ500 (Thermo Fisher Scientific Incorporated) internal size standard on an ABI 3500XL Genetic Analyzer (Applied Biosystems Incorporated) at Inqaba Biotechnical Industries. Alleles were scored manually using GeneMarker v2.7.0 (Liu *et al.* 2011). Due to the amount of missing data contained within APH20, this locus was omitted from further analyses. MICRO-CHECKER v2.2.3 (Van Oosterhout *et al.* 2004) was used to check for genotyping errors and null alleles.

### 3.2.4 Mitochondrial DNA amplification

Mitochondrial DNA (mtDNA) was amplified using two primer sets (**Chapter 2 Table 2.3**) following the methodology used by Hebert *et al.* (2004) to amplify the

*cytochrome b* (*cyt b*) and *cytochrome c oxidase subunit 1* (*CO1*) gene regions. PCRs, with a final volume of 25 $\mu$l, contained 1.25 U OneTaq DNA Polymerase (New England Biolabs Incorporated, Massachusetts, USA), 1X OneTaq Standard Reaction Buffer (New England Biolabs Incorporated, Massachusetts, USA), 200 $\mu$M dNTPs, 1 $\mu$M of each forward and reverse primer and 1 $\mu$l of genomic DNA. PCRs were run using a GeneAmp PCR System 9700 (Applied Biosystems Incorporated, California, USA) with the following cycle conditions: a denaturation step at 94 °C for 60 seconds, followed by 5 cycles of denaturation at 94 °C for 60 seconds, annealing at 45 °C for 1.5 minutes and elongation at 72 °C for 1.5 minutes, followed by 30 cycles of 94 °C for 1 minute, 54 °C for 1.5 minutes, 72 °C for 1.5 minutes and a final elongation step of 72 °C for 5 minutes. The PCR products were sent to Inqaba Biotechnical Industries for forward sequencing, using an ABI 3500XL Genetic Analyzer (Applied Biosystems Incorporated). Sequence chromatograms were edited and subsequently aligned using MEGA v7.0.14 (Kumar *et al.* 2016). All sequences have been submitted to GenBank (Accession numbers: MG654787-MG654810 and MK129186–MK129241).

### 3.2.5 Population structure and diversity estimates

For microsatellite loci, levels of genetic diversity were estimated using Arlequin v3.1 (Excoffier *et al.* 2005). Arlequin v3.1 (Excoffier *et al.* 2005) determined the mean number of alleles per locus (A), observed and expected heterozygosities, and deviations from Hardy-Weinberg Equilibrium for each species. To determine the level of genetic differentiation within and between species, an analysis of molecular variance (AMOVA) was performed to estimate F-statistics ($F_{ST}$ and $F_{IS}$), using Arlequin v3.1 (Excoffier *et al.* 2005). Arlequin v3.1 (Excoffier *et al.* 2005) was also used to evaluate linkage disequilibrium between pairs of microsatellite loci within and between each species. Population structure and identification of putative hybrids was done using three analyses. First, relationships between individuals was inferred using a principal component analysis (PCA) of a pairwise, individual-by-individual, covariance matrix without spatial components nor a priori population assignment in R v3.5.1 (R Core Team 2018) using the adegenet package (Jombart

2008). The final PCA plot was composed using the ggplot2 package (Wickham 2016) in R v3.5.1 (R Core Team 2018). Next, individual assignment probabilities (q) were estimated in the program STRUCTURE v2.3.4 (Pritchard *et al.* 2000), to determine any possible admixture. Assignments were conducted using the ADMIXTURE model with no prior population information. STRUCTURE v2.3.4 (Pritchard *et al.* 2000) was run for twenty replicates from K=1–10, with a run-length of 50,000 repetitions of Markov Chain Monte Carlo (MCMC), following the burn-in period of 10,000 iterations. The twenty values for the estimated ln(Pr(X—K)) were averaged, from which ΔK was calculated (Evanno *et al.* 2005), using StructureSelector (Li and Liu 2018). The K value with the largest ΔK was identified as the optimum number of clusters. The assessment was then repeated ten times using the K value, K=2, with the largest ΔK, with a run-length of 500,000 repetitions of MCMC, following the burn-in period of 100,000 iterations. CLUMPP v1.2.2 (Jakobsson and Rosenberg 2007) was used to find the optimal assignment probabilities from the ten repetitions using the FullSearch algorithm. The final structure plot was composed using the ggplot2 package in R v3.5.1 (R Core Team 2018). Williams *et al.* (2005) used $q > 0.90$ to define Mottled Ducks (*Anas fulvigula*), $0.90 < q < 0.10$ to define hybrids, and $q < 0.10$ to define Mallards. However, Väli *et al.* (2010) defined further generation hybrids and used a 5% gap threshold between categories, treating those individuals who fell within that gap as unidentifiable. Accordingly they used the following, Lesser Spotted Eagle (*Aquila pomarina*) $q < 0.10$; backcross F1 x Lesser Spotted Eagle $0.15 < q < 0.35$; F1 hybrid $0.4 < q < 0.6$; backcross F1 x Greater Spotted Eagle (*Aquila clanga*) $0.65 < q < 0.85$; Greater Spotted Eagle $q > 0.9$ (Väli *et al.* 2010). Based on the methods of Williams *et al.* (2005) and Väli *et al.* (2010) individuals were assigned to the following groups using their q-value: Mallard $q < 0.10$; hybrid $0.15 < q < 0.85$; Yellow-billed Duck $q > 0.90$, and those individuals falling within the gaps were unidentified. Finally, NEWHYBRIDS v1.1 (Anderson and Thompson 2002) was used to assign hybrid status to samples beyond the F1 generation. The programme does so by estimating the posterior probability (q) that an individual belongs to each of six classes (Anderson and Thompson 2002). These six classes include each parental species, F1 hybrids, F2 hybrids and backcrosses to each of the parental

species. F1 hybrids are the first filial generation produced by the interbreeding of two parental species. The interbreeding of two F1 hybrids, results in the second filial generation, or F2 hybrids. Backcrosses arise when F1 or F2 individuals interbreed with a parental generation. Individuals were assigned to one of the aforementioned groups based on their q-value with a threshold of 0.5. This means that individuals that presented with at least one q-value above 0.5 were assigned into that particular group, while those with posterior probabilities less than 0.5 remained unidentified (Väli *et al.* 2010). Assessments used both the Jeffreys-like and Uniform priors, and each ran for ten repetitions with a run-length of 100,000 repetitions of MCMC, following the burn-in period of 100,000 iterations. CLUMPP v1.2.2 (Jakobsson and Rosenberg 2007) was used to find the optimal posterior probabilities from the ten repetitions using the Greedy algorithm with 1000 random input orders. The final plots were composed using the ggplot2 package in R v3.5.1 (R Core Team 2018).

For mitochondrial DNA, we visualized population structure and gene flow using a haplotype network and phylogenetic tree. Mallard sequences, from Asia, were obtained from GenBank to act as positive controls and to better resolve relationships within the dataset. MEGA v7.0.14 (Kumar *et al.* 2016) was used for the phylogenetic analysis. Sequences were aligned using MUSCLE (Edgar 2004). Substitution models were tested and ranked based on the lowest Bayesian Information Criterion (BIC) scores that identified the Hasegawa–Kishino–Yano (HKY) model (Hasegawa *et al.* 1985) as the most appropriate model for the dataset. Models of DNA evolution are used to estimate evolutionary distances between sequences in phylogenetic analyses and can greatly affect the outcome of a phylogenetic analysis (Shapiro *et al.* 2006). The HKY model distinguishes between the rate of transitions and transversions and allows for unequal base frequencies, as generally, transition substitutions are more likely in nucleotide evolution than transversions (Hasegawa *et al.* 1985). A maximum-likelihood phylogenetic tree was inferred using 1000 bootstrap replicates with the Baikal Teal (*Anas Formosa*) as the outgroup. PopART v1.7 (Leigh and Bryant 2015) was used to construct a minimum spanning haplotype network, with ε set to zero.

## 3.3 Results

### 3.3.1 Microsatellite validation and variation

The eleven microsatellite loci that successfully genotyped (**Table 2.1**) were found to be polymorphic across all species. Three loci, Apl11; Smo7; and CM09, contributed to most of the variation between the study species, whereas Bcaμ11; Aph21; and Aph25 contributed the least. Four loci, APH21; APH25; Blm12; and Smo7, showed evidence of null alleles in MICRO-CHECKER. The dataset was not adjusted, as it was found that omitting these loci from further analyses did not have a significant effect on the final genetic diversity or admixture results. No significant linkage disequilibrium ($r^2 < 0.8$) was found between or within populations. The number of alleles per locus ranged between two and nine, with an average of four alleles per locus. Allele frequencies varied between populations. Seven loci yielded private alleles, there were a total of twelve private alleles: ten in the Mallards and two in the Yellow-billed Ducks (**Appendix Table A6**). The results indicate slightly higher diversity in the Mallard samples than in the Yellow-billed Ducks with respect to mean number of alleles per locus (A) as well as observed and expected heterozygosities (**Table 3.1**). The observed heterozygosity within all populations ranged from 0.36 to 0.58. The expected heterozygosity ranged from 0.45 to 0.55. The mean number of alleles per locus (A) was highest in the Mallards (3.64) and lowest (2.91) in the putative hybrids. Populations did not deviate from Hardy-Weinberg Equilibrium across all loci ($p \geq 0.05$, **Table 3.1**). The AMOVA indicated strong genetic differentiation between the populations (global $F_{ST} = 0.25$, **Table 3.2a**). However, it was found that the putative hybrids and Yellow-billed Ducks were strongly genetically differentiated from the Mallard (pairwise $F_{ST} = 0.25$ and 0.30, respectively), whereas there was negligible genetic differentiation between themselves (pairwise $F_{ST} = 0.01$, **Table 3.2b**). Within populations, the $F_{IS}$ values indicate slight heterozygote excess (-0.4) for the Mallards, and slight heterozygote deficiency (0.07 and 0.17) for the putative hybrids and Yellow-billed Ducks, respectively (**Appendix Table A7**).

**Table 3.1**: Measures of genetic diversity for each population, showing the mean number of alleles per locus (A), observed and expected heterozygosities ($H_{obs}$ and $H_{exp}$), and *p*-values of Hardy-Weinberg proportions.

| **Population** | **A** | **$H_{obs}$** | **$H_{exp}$** | ***p*-value** |
|---|---|---|---|---|
| Putative Hybrid | 2.91 | 0.49 | 0.55 | 0.43 |
| Mallard | 3.64 | 0.58 | 0.54 | 0.09 |
| Yellow-billed Duck | 3.36 | 0.36 | 0.45 | 0.10 |

**Table 3.2**: Analysis of molecular variance (AMOVA) showing the global $F_{ST}$ (a) and pairwise $F_{ST}$ (b) between all populations.

**(a)**

| **Source of variation** | **Df** | **SS** | **Variance** | **Variation (%)** | **$F_{ST}$** |
|---|---|---|---|---|---|
| Among populations | 2 | 63.34 | 0.95444 | 25.18 | |
| Among individuals within populations | 45 | 136.65 | 0.20062 | 5.29 | |
| Within individuals | 48 | 126.5 | 2.63542 | 69.53 | |
| **Total** | **95** | **326.49** | **3.79048** | | **0.25** |

Df - degrees of freedom

SS - sum of squares

**(b)**

| | **Putative Hybrid** | **Mallard** | **Yellow-billed Duck** |
|---|---|---|---|
| Putative Hybrid | - | - | - |
| Mallard | 0.25 | - | - |
| Yellow-billed Duck | 0.01 | 0.30 | - |

### 3.3.2 Population structure and molecular diversity

For microsatellite data, plotting PC1 and PC2 of the PCA separately grouped all Mallard and Yellow-billed Duck samples (**Figure 3.1a**). Moreover, all putative hybrid samples grouped with the Yellow-billed Ducks. One Mallard (Mal024) grouped together with the Yellow-billed Ducks. The most likely number of genetic clusters, as determined by STRUCTURE, was two (**Figure 3.1b**). All putative hybrid samples had assignment probability to the Yellow-billed Duck cluster ($q > 0.90$). Nearly all Mallard and Yellow-billed Duck samples were assigned to their expected clusters, except for three Mallards (Mal001, Mal003 and Mal024). Individual Mal024 is now inferred to be a hybrid ($q = 0.678$). Individuals Mal001 and Mal003 cannot be identified based on the chosen thresholds ($q = 0.138$ and 0.103, respectively). The Jeffreys-like prior, in NEWHYBRIDS, inferred all putative hybrids to be Yellow-billed Ducks and four Mallards (Mal001, Mal003, Mal013 and Mal024) to be F2 hybrids ($q = 0.735$, 0.712, 0.571 and 0.6182, respectively, **Appendix Figure A4a**). However, the Uniform prior did not infer these same results (**Appendix Figure A4b**). Rather, Mal001, Mal003 and Mal013 were inferred to be Mallards ($q = 0.761$, 0.849 and 0.888, respectively) and Mal024 was unidentifiable but highly admixed. Therefore, choice of either the Jeffreys-like prior or the Uniform prior had an effect on the assignment probabilities of individual samples, warranting these results unreliable (Väli *et al.* 2010).

For mitochondrial DNA, 19 (2.15%) variable sites were found between the Mallard, Yellow-billed Duck, and putative hybrid samples used in this study. Of these 19 variants, 14 were transitions, in equal ratios of A↔G and T↔C, as per the HKY evolutionary model. This corresponds to 12 haplotypes between these species, with two major groups; the Mallards (Group A) and the Yellow-billed Ducks (Group B, **Figure 3.2, Appendix Table A8**). The putative hybrid samples grouped together with the Yellow-billed Ducks. An identical pattern was found in the maximum-likelihood phylogenetic tree (**Figure 3.3**). The phylogenetic tree inferred two major clades, the Mallards (67% confidence) and the Yellow-billed Ducks (84% confidence). The putative hybrids grouped together with the Yellow-billed Ducks.

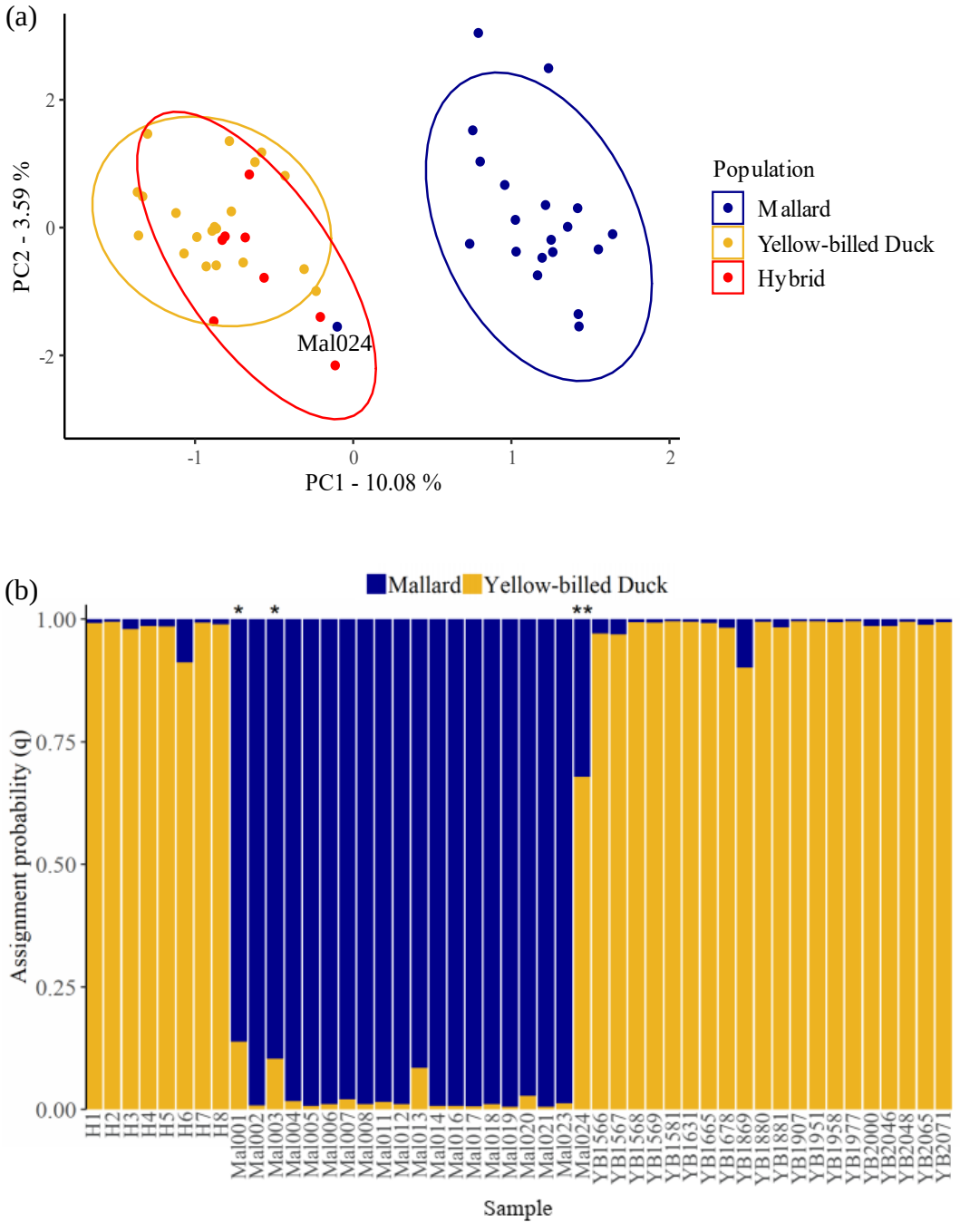


**Figure 3.1**: Results from the microsatellite loci analyses. (a) Principle component analysis (PCA) showing the relationships between Mallards (*Anas platyrhynchos*), Yellow-billed Ducks (*Anas undulata*), and putative hybrids. (b) STRUCTURE analysis (performed with K = 2) of microsatellite genotypes for putative hybrids, Mallards (*Anas platyrhynchos*) and Yellow-billed Ducks (*Anas undulata*). Each individual is represented by a single horizontal line, with lengths proportional to the estimated membership in each cluster. Putative hybrids: samples H1-H8; Mallards: samples M001-M024 and Yellow-billed Ducks: samples YB1566-YB2071.
* represents unidentified individuals ($0.10 < q < 0.15$, $0.85 < q < 0.90$).
** represents hybrid individuals ($0.15 < q < 0.85$).

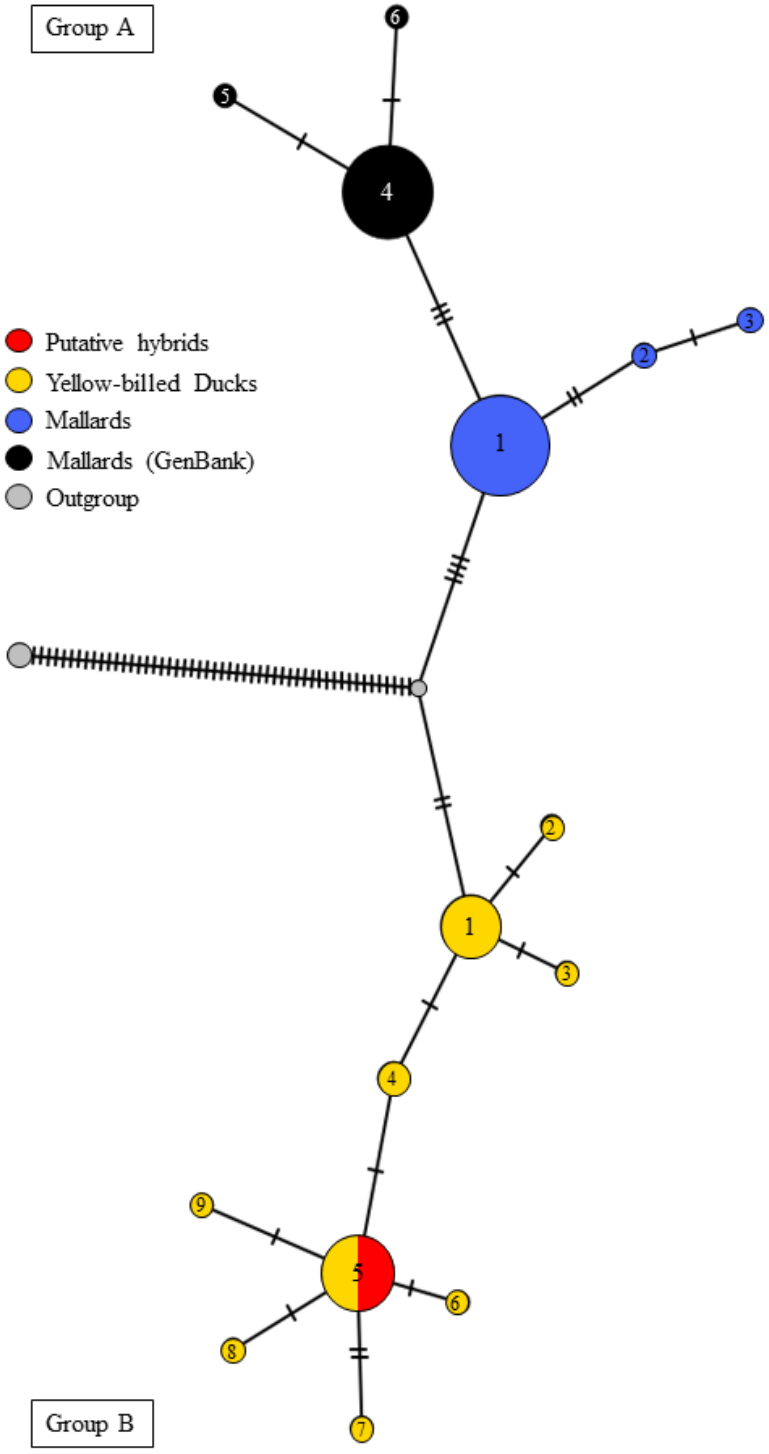


**Figure 3.2**: A rooted minimum spanning haplotype network showing the relationship between mitochondrial DNA, genes *cytochrome b* (*cyt b*) and *cytochrome c oxidase subunit I* (*COI*), haplotypes found within the Mallards (*Anas platyrhynchos*), Yellow-billed Ducks (*Anas undulata*) and putative hybrids, using the Baikal teal (*Anas formosa*) as the outgroup. The size of each circle is proportional to the number of individuals within that haplotype.

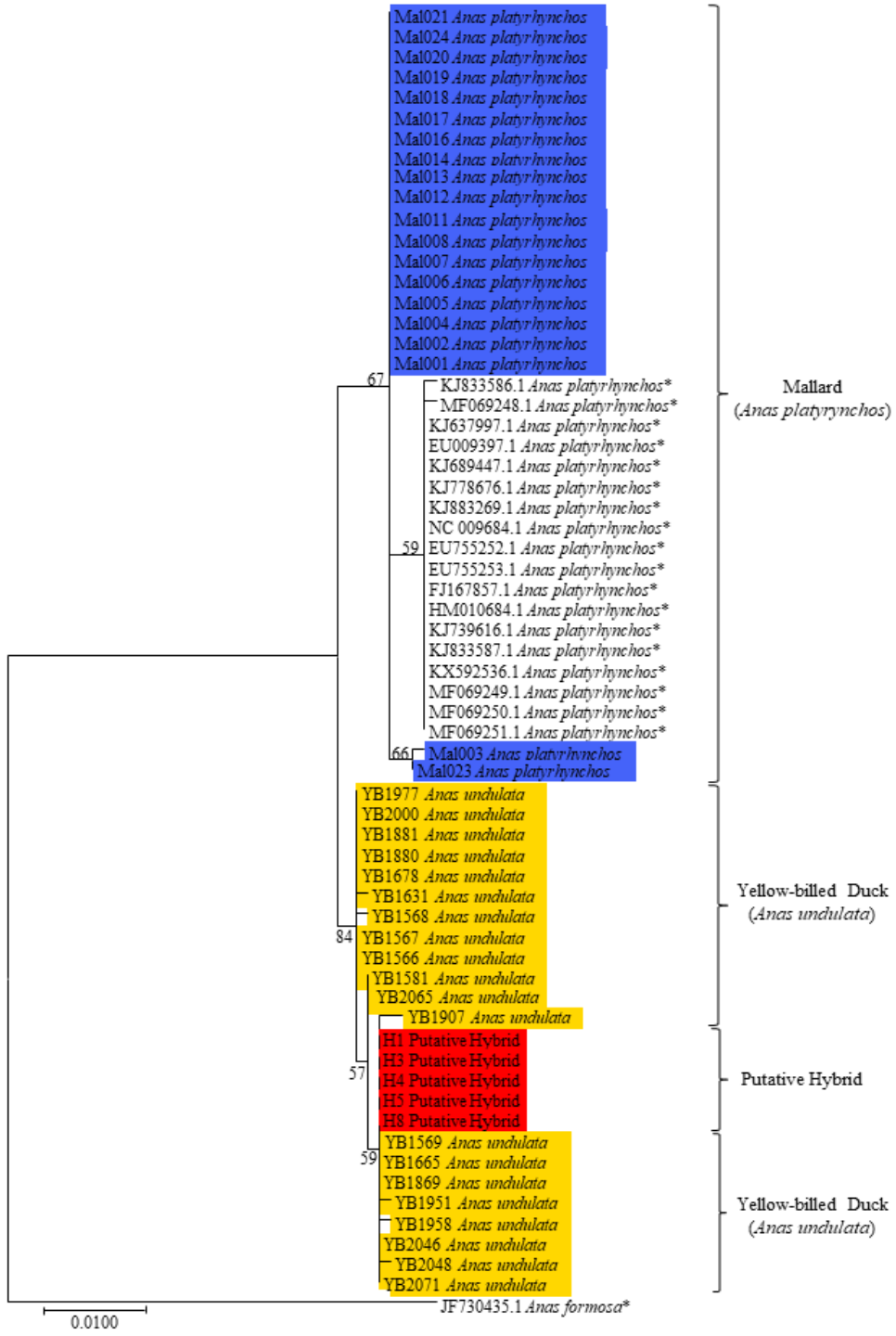


**Figure 3.3**: A maximum-likelihood phylogenetic tree constructed using concatenated *cytochrome b* (*cyt b*) and *cytochrome c oxidase subunit I* (*COI*) mitochondrial DNA sequences (885 bp). The Hasegawa-Kishino-Yano (HKY) evolutionary model was used in a bootstrap test (1000 replicates) using the Baikal teal (*Anas Formosa*) as an outgroup. Sequences highlighted in blue, red and yellow indicate Mallard (*Anas platyrhynchos*), putative hybrid and Yellow-billed Duck (*Anas undulata*) samples used in this study, respectively. * indicates sequences obtained from GenBank.

## 3.4 Discussion

The genetic diversity of many species worldwide is being threatened by introgressive hybridisation due to anthropogenic influences (Kulikova *et al.* 2004, Mank *et al.* 2004, Rhymer 2006, Fowler *et al.* 2009, Dufresnes *et al.* 2015, Ford *et al.* 2017). This may be the case for the Mallard and Yellow-billed Duck, where captive-bred escaped Mallards and the introduction of Mallards for game and ornamental purposes, has potentially allowed for hybridisation between these two species. However, microsatellite genotyping and mitochondrial DNA sequencing showed that this may not be the case in central and north-west South Africa, as there was negligible evidence of admixture within the putative hybrids and Yellow-billed Ducks representing those regions. Yet, hybridisation may be occurring within the Western Cape Province.

### 3.4.1 Molecular diversity estimates

Measures of genetic diversity in the Yellow-billed Duck were slightly lower than that of the Mallard, as was expected. Mallards are expected to have more variability given their large population sizes and near global distribution (Fowler *et al.* 2009). Also, it is hypothesised that a Mallard-like ancestor gave rise to the various monochromatic species within the Mallard complex (Omland 1997, McCracken *et al.* 2001, Wang *et al.* 2018). Based on this it would then be expected that the Mallard would more ancestral and have had time to acquire more mutations, allowing for more variation. There was clear genetic differentiation between the two reference populations as well as the putative hybrids and Mallards (pairwise $F_{ST}$ = 0.30 and 0.25, respectively). This finding is consistent with that of Lavretsky *et al.* (2014), whereby a pairwise $\Phi_{ST}$ of 0.22 between the Old World Mallard and Yellow-billed Duck was found. However, there was no significant genetic differentiation between the putative hybrids and Yellow-billed Ducks (pairwise $F_{ST}$ = 0.01). These findings were further supported by the admixture, phylogenetic and haplotype analyses, which separated the Yellow-billed Ducks and Mallards into distinct populations, but grouped the putative hybrids with the Yellow-billed Ducks.

### 3.4.2 Population structure and admixture

Two populations were inferred and negligible admixture of these two populations was found within the putative hybrid samples according to the q-value thresholds. However, one phenotypic Mallard sample was shown to be admixed (q = 0.678). The Jeffreys-like and Uniform priors showed slightly different results and, therefore, it is suggested that the results be treated with caution (Anderson 2003). In order to further identify more complex hybrid categories and further backcrossing additional markers are required, with significant allele frequency differences between the populations (Anderson and Thompson 2002, van Wyk *et al.* 2017). Vähä and Primmer (2006) found that ≥ 95% of backcrosses were identified when a minimum of 48 microsatellite loci with an average $F_{ST}$ = 0.21 were used. While this study only used 11 microsatellite loci, the NEWHYBRIDS analyses demonstrated minimal admixture within the putative hybrid samples, and considerable admixture within the one putative Mallard (Mal024). This trend was further supported through the PCA, where Mal024 was found to group with the Yellow-billed Duck samples.

The results using these particular markers, within these putative hybrids, give negligible support for hybridisation between the Mallard and Yellow-billed Duck in central and north-west South Africa. Mallards are found in higher abundance in and around urban areas (Berruti 1992), which may explain the lack of hybridisation found within these samples from Barberspan. Furthermore, due to the small sample size of this study, low levels of hybridisation within these regions cannot be completely ruled out. One phenotypic Mallard (Mal024) from Western Cape Province was found to be a hybrid (0.15 < q > 0.85). Two other Mallards (Mal001 and Mal003) were unidentifiable, according to the q value, but showed substantial admixture. This, therefore, suggests the possibility of hybridisation occurring between the Yellow-billed Duck and the Mallard within this province. The mtDNA results would then suggest that hybridisation may be occurring in the reverse direction to what was hypothesised. It was hypothesised that male Mallards were hybridising with female Yellow-billed Ducks, based on the aggressive nature and unique plumage of Mallard drakes. If this were the case, these Mallard individuals would have grouped with the Yellow-billed Ducks, as mtDNA is

maternally inherited. Instead, the results showed that all the Mallards grouped together in their own clade. This suggests that a Yellow-billed Duck drake may have mated with a Mallard hen to produce this hybrid. Fowler *et al.* (2009) found the same trend, where male Hawaiian Ducks were hybridising with female Mallards. Similarly, Kulikova *et al.* (2004) found that Mallard hens were producing hybrids with Eastern Spot-Billed Duck drakes. These findings could be explained by mispairings during pair formation. Mispairings occur due to errors in mate recognition and/or due to scarcity in conspecifics (Randler 2002). Errors in mate recognition may occur when two species have similar body sizes, plumage or colouration and vocalisations (Ford *et al.* 2017). All of these characteristics are similar in the females of Mallards and Yellow-billed Ducks. Also, it is well known that female Mallards display relaxed mate choice and will sometimes mate with heterospecifics when the latter are in abundance (Fowler *et al.* 2009, Wang *et al.* 2018). Mallards were initially introduced into the Western Cape Province and have established populations. Through this, and in part due to urbanisation, Mallards are currently found in abundance in this region (Berruti 1992). This abundance could promote hybridisation. Studies done in North America have shown that hybridisation incidences, between the Mallard and Mottled Duck, vary depending on the region (Williams *et al.* 2005, Ford *et al.* 2017). The hybridisation incidence in western Gulf Coast was found to be about 5%, in Florida ~9% and in South Carolina hybridisation had led to a hybrid swarm (Williams *et al.* 2005, Ford *et al.* 2017). This pattern was attributed to the increased abundance of Mallards from the western Gulf Coast moving north east through to South Carolina.

### 3.4.3 Management implications

The highly admixed Mallard (Mal024, q = 0.678) was found to group with the Yellow-billed Ducks and to consist mainly of Yellow-billed Duck DNA. This individual was identified phenotypically as a Mallard and thus culled as part of the National Environmental Management: Biodiversity Act (NEMBA) regulations. However, due to these findings, there is a possibility that this individual was misidentified. To the untrained eye female Mallards and Yellow-billed Ducks can

look alike, as they have similar plumage and colouration (Sinclair *et al.* 2011). This individual was sexed as being female using the *chromodomain-helicase-DNA* (*CHD*) gene and could therefore have been misidentified. This could have important implications in conservation, as culling authorities could be misidentifying ducks and culling the native species instead. A similar case happened in a study by Ford *et al.* (2017) whereby two putative Mottled Ducks were assigned as Mallards in STRUCTURE. It was hypothesised that these samples were misidentified in the field as the females of the two species resemble each other year-round.

The putative hybrids in this study had phenotypic differences in their feet and legs, i.e. they all had orange colouring, to varying degrees, leading to the hypothesis of hybridisation. Most studies of hybridisation include hybrid samples comprising many phenotypic changes, such as changes in feather shape and colour, bill and leg colour, wing bar sizes and face striping (Gillepsie 1985, Williams *et al.* 2005, Peters *et al.* 2014). This colouration has also been noted by Sinclair *et al.* (2011) who believed it to be due to hybridisation. However, it could be due to phenotypic plasticity within the Yellow-billed Duck population. Phenotypic plasticity is the phenomenon of a genotype producing several different phenotypes in response to different environmental conditions, and can have adaptive, maladaptive or neutral fitness effects (Bradshaw 1965, Ghalambor *et al.* 2007). Carotenoids are red, yellow and orange pigments produced by plants and algae (Goodwin 1984, Grether 2005). These pigments cannot be synthesised *de novo* by animals and are usually obtained through diet which can then be assimilated and deposited in the integument of animals (Goodwin 1984, Grether 2005). Studies done by Hudon and Brush (1989) and Mulvihill *et al.* (1992) found that the yellow tail tip and belly feathers of Cedar Waxwings (*Bombycilla cedrorum*) turn orange when they eat the berries of introduced Honeysuckle (*Lonicera morrowi*) which are rich in rhodoxanthin, a red carotenoid. Yellow-billed Ducks are omnivorous, with a broad diet including fruits, seeds, roots, leaves and stems of plants (Kear 2005). Yellow-billed Ducks are distributed throughout South Africa, and thus the different sub-populations may experience slightly different environments and diets. The Yellow-billed Ducks and putative hybrids of this study were sampled during moult

from Barberspan which is, therefore, not indicative of individual residential or breeding sites. It can be hypothesised that the putative hybrids of this study could come from a habitat supplying a diet rich in orange carotenoids, producing phenotypic plasticity with potential neutral fitness.

### 3.4.4 Conclusions

Phenotypic changes in the feet and legs of Yellow-billed Ducks lead to the hypothesis of introgressive hybridisation with the Mallard. Minimal admixture was found within these putative hybrids when compared to two "pure" reference populations. The results of this study, therefore, suggest that there may be no loss of genetic diversity of the indigenous Yellow-billed Duck by introgressive hybridisation with the invasive Mallard in central and north-west South Africa. This is promising news for the conservation of this indigenous species in these regions. Although, low levels of hybridisation cannot be discounted and there may be evidence for hybridisation in the Western Cape Province. Therefore, continued monitoring should be undertaken regularly to examine any changes in hybridisation over time, as there is a possibility that hybridisation may be occurring at different rates throughout different regions. It should also be noted that conservation management should be cautious as to the culling of Mallards in the future, due to possibilities of misidentification and the accidental culling of native species.

# Chapter 4

# Concluding remarks

Nuclear and mitochondrial DNA markers were used to genetically assess any hybridisation that may be occurring between the Mallard (*Anas platyrhynchos*) and Yellow-billed Duck (*Anas undulata*) in South Africa, where the introduction of the Mallard has potentially allowed for hybridisation between these two species. No admixture was found within the putative hybrids, rather they were found to group together with the Yellow-billed Ducks.

## 4.1 Findings

Eighteen of the twenty-seven waterfowl microsatellite loci analysed, cross amplified in the Yellow-billed Duck and Mallard. The eleven selected microsatellite markers allowed the assignment of duck samples to either of the two studied species. Two private alleles were present in the Yellow-billed Ducks and ten within the Mallards, within seven loci. Markers Apl11, Smo7, and CM09 contributed to majority of the variation found between the Mallard and Yellow-billed Duck, whereas Bca$\mu$11, Aph21, and Aph25 contributed the least. These are the first microsatellite markers that have been identified and assessed in the Yellow-billed Duck. These markers can be important for other genetic studies of the Yellow-billed Duck, such as genome and/or linkage mapping, pedigree analyses, demographic analyses, population level movement studies and hybridisation studies involving other species (Ellegren 2004, Buschiazzo and Gemmell 2006).

The mitochondrial genes, *cyt b* and *COI*, had few nucleotide variants (~2%) between all the samples. Despite this, these genes successfully separated the samples both interspecifically (Yellow-billed Ducks and Mallards separated) and intraspecifically (Mallards used in this study separated from the Mallards obtained from GenBank) with sufficient bootstrap support. These markers have proven successful in elucidating inter- and intraspecific relationships, with little variation,

in studies on musk deer (genus *Moschus*, Su *et al.* 1999) and bee-eaters (genus *Merops*, Arif *et al.* 2011). This success could also be attributed to concatenation. Urantowka *et al.* (2017) found that single mitochondrial markers incorrectly inferred relationships in parrots (tribe *Arini*), however, concatenation of all the markers used resolved the relationships. Concatenated datasets are thought to increase informativeness of the data; help resolve nodes and basal branching; and ultimately improve phylogenetic accuracy (Fitzpatrick *et al.* 2006). Adding genes together into one dataset increases the absolute number of evolutionary changes on branches and nodes, allowing for the inference of relationships with better accuracy (Gadagkar *et al.* 2005).

Both the microsatellite and mtDNA markers used in this study proved successful in identifying a Yellow-billed Duck x Mallard hybrid and the probable directionality of the mating. One of the Mallards (Mal024) was found to be admixed ($0.15 < q > 0.90$), mainly with Yellow-billed Duck DNA, and grouped with the Yellow-billed Ducks within the PCA. To the untrained eye female Mallards and Yellow-billed Ducks can often look alike, as they have similar plumage and colouration (Sinclair *et al.* 2011). It was therefore hypothesised that this individual, if it was a female, could have been misidentified. In order to test this theory, molecular sexing of this sample was done using the sex-linked *CHD* gene. Female birds are the heterogametic sex (ZW) while males are homogametic (ZZ). Therefore, in most cases, amplification of a sex-linked gene will reveal a double band for females and a single band for males on an agarose gel. The primers 2550F (5′- GTTACTGATTCGTCTACGAGA-3′) and 2718R (5′-ATTGAAATGATC CAGTGCTTG-3′) were used to amplify a different-sized intron of *CHD*-W and *CHD*-Z, following the methodology of Fridolfsson and Ellegren (1999) and Lambert *et al.* (2009). This primer pair amplifies a single gene copy, of different sizes, in both females (400-450 bp) and males (600-650 bp) of the Anatidae family (Fridolfsson and Ellegren 1999). The agarose gel (**Appendix Figure A5**) revealed a single band of ± 450 bp, suggesting that the Mallard (Mal024) was a female. This could have important conservation implications. Mitochondrial DNA analyses grouped this individual (Mal024) together with the Mallards, suggesting that Mallard hens and Yellow-billed Duck drakes may be interbreeding. This individual

(Mal024) was obtained from the Western Cape Province, suggesting the occurrence of hybridisation within that region. However, minimal admixture was found within the putative hybrids or Yellow-billed Ducks sampled from North West Province. Therefore, hybridisation could be occurring only within specific provinces in South Africa and the orange colour found within the legs and feet of the putative hybrids could be due to phenotypic plasticity within the Yellow-billed Duck population.

## 4.2 Limitations

The eight Mallard samples from Modderfontein Nature Reserve had to be omitted from the study due to the continuous failure to amplify the DNA of all eight samples. During DNA extraction, the blood samples were extremely coagulated making it difficult to obtain high yields of intact DNA necessary for downstream analyses. This could be due to two reasons, both involving EDTA. High concentrations of EDTA prevents DNA degradation by sequestering divalent cations essential for DNase activity (Marjorie and Borrelli 1958). EDTA is also an anti-coagulant and therefore prevents clotting of the blood during storage (Marjorie and Borrelli 1958). The problem could have originated during the preparation of the Queen's lysis buffer. Too little EDTA added to the lysis buffer, or a high blood to lysis buffer ratio would result in an inability to prevent blood clotting and DNA degradation. Future studies should be sure of these ratios when making up solutions and collecting samples, or utilise alternative blood collection methods.

The putative hybrids used in this study were sampled from only one area, which was a moulting site, and not representative of breeding or residency sites. The one phenotypic Mallard sample from Rietvlei Wetland Reserve, Western Cape Province, found to be a hybrid indicates the possibility of hybridisation occurring in that region. Mallards are abundant in the Western Cape Province, as this is where they were originally introduced. This abundance could be promoting hybridisation. Studies done in North America have shown that hybridisation incidences, between the Mallard and Mottled Duck (*Anas fulvigula*), vary depending on the region and abundance of Mallards (Williams *et al.* 2005, Ford *et al.* 2017). This study would have benefitted by sampling putative hybrids, Yellow-billed Ducks and Mallards

from various regions within South Africa so as to gain a better understanding of the potential hybridisation events occurring throughout the country.

According to Vähä and Primmer (2006) at least twelve to twenty-four polymorphic markers are needed to accurately identify F1 hybrids between closely related species. Furthermore, approximately forty-eight polymorphic loci are needed to accurately identify backcrossed hybrids and hybrids passed the F1 generation (Vähä and Primmer 2006). Unfortunately, this study could only identify eleven polymorphic, perfect motif microsatellite loci. Although these proved useful in distinguishing between Mallards, Yellow-billed Ducks and admixed individuals, they could not identify later generation and backcrossed hybrids with significant confidence. This highlights the need to identify more informative DNA markers for future work.

## 4.3 Future Work

It is clear that a lot more work needs to be done to accurately identify Yellow-billed Duck x Mallard hybrids. Future studies should focus on sequencing the genome of the Yellow-billed Duck. This will help to identify and develop more genetic markers, such as microsatellite loci or SNPs, with high variability and cross-species applicability, to better identify hybrid individuals (Gramlich *et al.* 2018, Soto-Calderón *et al.* 2018). These studies should also include extensive sampling throughout all provinces in South Africa, with the tracking of samples and the monitoring of behaviours and diets. Once more markers are developed a thorough protocol could be established to frequently monitor hybridisation for conservation efforts.

## 4.4 Conclusions

Based on the genetic markers used in this study, there was minimal evidence of introgressive hybridisation between the species. However, one Mallard (Mal024) hen was found to be a possible hybrid or backcross, and was genetically assigned

as mostly Yellow-billed Duck, indicating the possibility that Yellow-billed Duck drakes could be hybridising with Mallard hens. In conclusion, no loss of local genetic diversity in samples representative of central and north-west South Africa was found. However, continued monitoring should be undertaken regularly to examine any changes in hybridisation over time. In order for accurate and continued monitoring to occur, a protocol needs to be developed that includes many genetic markers with extensive and rigorous sampling throughout the country.

# References


AEWA (African-Eurasian Migratory Waterbirds Agreement). 2018. Agreement text and annexes. Available at: https://www.unep-aewa.org/sites/default/files/basic_page_documents/aewa_agreement_text_2019_2021_final.pdf [Accessed 14 February 2019].

Allendorf FW, Leary RF, Spruell P, Wenburg JK. 2001. The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution* **16**: 613–622.

Anderson EC. 2003. User's guide to the program NEWHYBRIDS version 1.1 beta. Available at: http://ib.berkeley.edu/labs/slatkin/eriq/software/new_hybs_doc1_1Beta3.pdf [Accessed 18 December 2018].

Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217–1229.

Arif IA, Khan H, Shobrak M, Williams J. 2011. *Cytochrome c oxidase subunit I* (*COI*) barcoding of Green Bee-eater (*Merops orientalis*). *Genetics and Molecular Research* **10**: 3992–3998.

Ashton C, Ashton M. 2001. *The Domestic Duck.* Marlborough: The Crowood Press.

Avise JC. 1987. Intraspecific phylogeography. *Annual Review of Ecology, Evolution, and Systematics* **18**: 489–522.

Avise JC. 1998. Conservation genetics in the marine realm. *Journal of Heredity* **89**: 377–382.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**: 37–48.

Banks AN, Wright LJ, Maclean IMD, Hann C, Rehfisch MM. 2008. Review of the status of introduced non-native waterbird species in the area of the African-Eurasian Waterbird Agreement. *AEWA Technical Series*. Available at: http://www.unep-aewa.org/sites/default/files/document/mop4_12_non_native_species_corr1_0.pdf [Accessed 14 May 2017].

Berruti A. 1992. Alien birds. *Birding in SA* 44: 7.

Bhargava A, Fuentes FF. 2010. Mutational dynamics of microsatellites. *Molecular Biotechnology* **44**: 250–266.

BirdLife South Africa. 2019. Barberspan & Leeupan. Available at: https://www.birdlife.org.za/get-involved/join-birdlife-south-africa/item/167-sa026-barberspan-leeupan [Accessed 31 January 2019].

Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF. 2001. M13-tailed primers improve the readability and ysability of microsatellite analyses performed with two different allele-sizing methods. *BioTechniques* **31**: 25–28.

Bradshaw AD. 1965. Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics* **13**: 115–155.

Braithwaite LW, Miller B. 1975. The Mallard, *Anas platyrhynchos*, and Mallard-Black Duck, *Anas superciliosa rogersi*, hybridization. *Australian Wildlife Research* **2**: 47–61.

Brooke RK, Siegfried W.1991. Birds introduced to the fynbos biome of South Africa. In: Groves RH, Di Castri F (eds), *Biogeography of Mediterranean Invasions*. Cambridge: Cambridge University Press. pp 359–364.

Brown WM, George M, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences* **76**: 1967–1971.

Bruns TD, Szaro TM, Gardes M, Cullings K, Pan JJ, Taylor DL, Horton TR, Kretzer A, Garbelotto M, Li Y. 1998. A sequence database for the identification of ectomycorrhizal basidiomycetes by phylogenetic analysis. *Molecular Ecology* **7**: 257–272.

Buchholz WG, Pearce JM, Pierson BJ, Scribner KT. 1998. Dinucleotide repeat polymorphisms in waterfowl (family Anatidae): characterization of a sex-linked (Z-specific) and 14 autosomal loci. *Animal Genetics* **29**: 322–332.

Buschiazzo E, Gemmell NJ. 2006. The rise, fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays* **28**: 1040–1050.

Carew PJ, Adcock GJ, Mulder RA. 2003. Microsatellite loci for paternity assessment in the black swan (*Cygnus atratus*: Aves). *Molecular Ecology Notes* **3**: 1–3.

CBD (Convention on Biological Diversity). 2014. Biodiversity: the foundation for sustainable development. Available at: https://www.cbd.int/development /doc/sdg-feb2014-info-en.pdf [Accessed 12 February 2019].

Chown SL, Hodgins KA, Griffin PC, Oakeshott JG, Byrne M, Hoffmann AA. 2015. Biological invasions, climate change and genomics. *Evolutionary Applications* **8**: 23–46.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1659.

Cler L, Bu D, Lewis C, Euhus D. 2006. A comparison of five methods for extracting DNA from paucicellular clinical samples. *Molecular and Cellular Probes* **20**: 191–196.

Dalton DL, Tordiffe A, Luther I, Duran A, van Wyk AM, Brettschneider H, Oosthuizen A, Modiba C, Kotzé A. 2014. Interspecific hybridization between Greater Kudu and Nyala. *Genetica* **142**: 265–271.

DEA (Department of Environmental Affairs). 2016. National Environmental Management: Biodiversity Act, 2004 (Act No. 10 of 2004) Alien and Invasive Species Lists. *Government Gazette, South Africa*.

Delacour J, Mayr E. 1945. The family Anatidae. *The Willson Bulletin* **57**: 3–55.

Denk AG, Gautschi B, Carter K, Kempenaers B. 2004. Seven polymorphic microsatellite loci for paternity assessment in the Mallard (*Anas platyrhynchos*). *Molecular Ecology Notes* **4**: 506–508.

Desforges MF, Wood-Gush DGM. 1975. A behavioural comparison of domestic and Mallard ducks: spatial relationships in small flocks. *Animal Behaviour* **23**: 698–705.

Desforges MF, Wood-Gush DGM. 1976. Behavioural comparison of Aylesbury and Mallard ducks: sexual behaviour. *Animal Behaviour* **24**: 391–397.

Desjardins P, Conklin D. 2010. NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments* **45**: 1–4.

Dufresnes C, Dubey S, Ghali K, Canestrelli D, Perrin N. 2015. Introgressive hybridization of threatened European Tree Frogs (*Hyla arborea*) by introduced *H. intermedia* in Western Switzerland. *Conservation Genetics* **16**: 1507–1513.

Edgar RC. 2004. MUSCLE : multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Edwards A, Civitello A, Hammond HA, Caskey CT. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49**: 746–756.

Ellegren H. 1992. Polymerase-chain-reaction (PCR) analysis of microsatellites-a new approach to studies of genetic relationships in birds. *The Auk* **109**: 886–895.

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

Erlich H. 1989. Polymerase Chain Reaction. *Journal of Clinical Immunology* **9**: 437–447.

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

Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research* **8**: 175–185.

Excoffier L, Laval G, Schneider S. 2005. Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distance. *Genetics* **131**: 479–491.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Fields RL, Scribner KT. 1997. Isolation and characterization of novel waterfowl microsatellite loci : cross-species comparisons and research. *Molecular Ecology* **6**: 199–202.

Filip C, Fletcher G, Wulff JL, Earhart CF. 1973. Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium lauryl sarcosinate. *Journal of Bacteriology* **115**: 717–722.

Fitzpatrick DA, Logue ME, Stajich JE, Butler G. 2006. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC Evolutionary Biology* **6**: 99–114.

Flores-rentería L, Krohn A. 2013. Scoring microsatellite loci. In: Kantartz SK, (ed.), *Microsatellites: Methods and Protocols, Methods in Molecular Biology*. Berlin: Springer Science+Business Media. pp 319–336.

Ford RJ, Selman W, Taylor SS. 2017. Hybridization between Mottled Ducks (*Anas fulvigula maculosa*) and Mallards (*A. platyrhynchos*) in the western Gulf Coast region. *The Condor* **119**: 683–696.

Fowler AC, Eadie JM, Engilis A. 2009. Identification of endangered Hawaiian Ducks (*Anas wyvilliana*), introduced North American Mallards (*A. platyrhynchos*) and their hybrids using multilocus genotypes. *Conservation Genetics* **10**: 1747–1758.

Fridolfsson AK, Ellegren H. 1999. A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30**: 116–121.

Gadagkar SR, Rosenberg MS, Kumar S. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology Part B Molecular and Developmental Evolution* **304**: 64–74.

Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology* **21**: 394–407.

Gillepsie GD. 1985. Hybridization, introgression, and morphometric differentiation between Mallard (*Anas platyrhynchos*) and Grey Duck (*Anas superciliosa*) in Otago, New Zealand. *The Auk* **102**: 459–469.

Goldstein DB, Schlotterer C. 1999. *Microsatellites: evolution and applications*. Oxford: Oxford University Press.

Goodwin TW. 1984. *The biochemistry of the carotenoids* (2nd edn). New York: Chapman and Hall.

Gordon DA, Priscu J, Giovannoni S. 2000. Origin and phylogeny of microbes living in permanent Antarctic lake ice. *Microbial Ecology* **39**: 197–202.

Goulet BE, Roda F, Hopkins R. 2017. Hybridization in plants: old ideas, new techniques. *Plant Physiology* **173**: 65–78.

Gramlich S, Wagner ND, Hörandl E. 2018. RAD-seq reveals genetic structure of the F2-generation of natural willow hybrids (*Salix* L.) and a great potential for interspecific introgression. *BMC Plant Biology* **18**: 317–328.

Grant PR, Grant BR. 1997. Hybridization, sexual imprinting, and mate choice. *The American Naturalist* **149**: 1–28.

Grant PR, Grant BR, Keller LF, Markert JA, Petren K. 2003. Inbreeding and interbreeding in Darwin's finches. *Evolution* **57**: 2911–2916.

Grether GF. 2005. Environmental change, phenotypic plasticity, and genetic compensation. *The American Naturalist* **166**: 115–123.

Guay PJ, Mulder RA. 2005. Isolation and characterization of microsatellite markers in Musk Duck (*Biziura lobata* : Aves), and their application to other waterfowl species. *Molecular Ecology Notes* **5**: 249–252.

Guay PJ, Taysom A, Robinson R, Tracey JP. 2014. Hybridization between the Mallard and native dabbling ducks: causes, consequences and management. *Pacific Conservation Biology* **20**: 41–47.

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ. 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources* **11**: 591–611.

Harebottle D, Vanderwalt B. 2014. Hybridisation between White-faced and Fulvous Ducks in the wild: further evidence from South Africa. *Ornithological Observations* **5**: 17–21.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* **22**: 160–174.

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

Hepp GR, Novak JM, Scribner KT, Stangel, PW. 1988. Genetic distance and hybridization of Black Ducks and Mallards: a morph of a different color? *The Auk* **105**: 804–807.

Holsinger KE, Weir BS. 2009. Genetics in geographically structured populations: defining, estimating and interpreting $F_{ST}$. *Nature Reviews Genetics* **10**: 639–650.

Hudon J, Brush AH. 1989. Probable dietary basis of a color variant of the Cedar Waxwing. *Journal of Field Ornithology* **60**: 361–368.

Jackson RB, Linder CR, Lynch M, Purugganan M, Somerville S, Thayer SS. 2002. Linking molecular insights and ecological research. *Trends in Ecology & Evolution* **17**: 409–414.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.

Johnson KP. 2001. Taxon sampling and the phylogenetic position of passeriformes: evidence from 916 avian *cytochrome b* sequences. *Systematic Biology* **50**: 128–136.

Jombart T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Josefsson M, Andersson B. 2001. The environmental consequences of alien species in the Swedish lakes Mälaren, Hjälmaren, Vänern and Vättern. *Ambio* **30**: 514–521.

Karger AE, Harris JM, Gesteland RF. 1991. Multiwavelength fluorescence detection for DNA sequencing using capillary electrophoresis. *Nucleic Acids Research* **19**: 4955–4962.

Kartavtsev YP, Lee JS. 2006. Analysis of nucleotide diversity at the *cytochrome b* and *cytochrome oxidase 1* genes at the population, species, and genus levels. *Russian Journal of Genetics* **42**: 341–362.

Kear J. 2005. *Ducks, geese and swans, vol 2, species accounts (Cairina to Mergus)*. New York: Oxford University Press.

Kelly M.J. 2000. Lineage loss in Serengeti cheetahs: consequences of high reproductive variance and heritability of fitness on effective population size. *Conservation Biology* **15**: 137–147.

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

Kulikova IV, Zhuravlev YN, McCracken KG. 2004. Asymmetric hybridization and sex-biased gene flow between Eastern Spot-Billed Ducks (*Anas zonorhyncha*) and Mallards (*A . platyrhynchos*) in the Russian far east. *The Auk* **121**: 930–949.

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.

Kvist L, Martens J, Nazarenko AA, Orell M. 2003. Paternal leakage of mitochondrial DNA in the Great Tit (*Parus major*). *Molecular Biology and Evolution* **20**: 243–247.

Lambert DM, Shepherd LD, Huynen L, Beans-Picón G, Walter GH, Millar CD. 2009. The molecular ecology of the extinct New Zealand Huia. *PLoS ONE* **4**: e8019.

Lavretsky P, McCracken KG, Peters JL. 2014. Phylogenetics of a recent radiation in the Mallards and allies (Aves: Anas): inferences from a genomic transect and the multispecies coalescent. *Molecular Phylogenetics and Evolution* **70**: 402–411.

Lee PY, Costumbrado J, Hsu C, Kim YH. 2012. Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments* **62**: 1–6.

Lehman N, Eisenhawer A, Hansen K, Mech L, Peterson RO, Gogan PJP. 1991. Introgression of Coyote mitochondrial DNA into sympatric North American Gray Wolf populations. *Evolution* **45**: 104–119.

Leigh JW, Bryant D. 2015. PopART: full-feature software for haplotype network construction. *Methods in Ecology and Evolution* **6**: 1110–1116.

Li YL, Liu JX. 2018. StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources* **18**: 176–177.

Liu JC, Hulce D, Li X, Snyder-Leiby T. 2011. GeneMarker genotyping software: tools to increase the statistical power of DNA fragment analysis. *Journal of Biomolecular Techniques* **22**: 35–36.

Luo S, Valencia CA, Zhang J, Lee NC, Slone J, Gui B, Wang X, Li Z, Dell S, Brown J, Chen SM, Chien YH, Hwu WL, Fan PC, Wong LJ, Atwal PS, Huang T. 2018. Biparental inheritance of mitochondrial DNA in humans. *Proceedings of the National Academy of Sciences* **115**: 1–6.

Maak S, Wimmers K, Weigend S, Neumann K. 2003. Isolation and characterization of 18 microsatellites in the Peking duck (*Anas platyrhynchos*) and their application in other waterfowl species. *Molecular Ecology Notes* **3**: 224–227.

Madge S, Burn H. 1988. *Wildfowl*. London: Christopher Helm Publishers.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* **20**: 229–237.

Malukiewicz J, Boere V, Fuzessy LF, Grativol AD, de Oliveira E, Silva I, Pereira LCM, Ruiz-Miranda CR, Valença YM, Stone AC. 2015. Natural and anthropogenic hybridization in two species of eastern Brazilian Marmosets (*Callithrix jacchus* and *C. penicillata*). *PloS ONE* **10**: e0127268.

Mank JE, Carlson JE, Brittingham MC. 2004. A century of hybridization: decreasing genetic distance between American Black Ducks and Mallards. *Conservation Genetics* **5**: 395–403.

Maran T, Henttonen H. 1995. Why is the European mink (*Mustela lutreola*) disappearing? A review of the process and hypotheses. *Annales Zoologici Fennici* **32**: 47–54.

Marjorie B, Borrelli J. 1958. Some effects of divalent cations on the clotting mechanism and the platelets of EDTA blood. *Journal of Applied Physiology* **12**: 453–460.

Mathew CGP. 1984. The isolation of high molecular weight eukaryotic DNA. In: Walker J (ed.), *Nucleic Acids. Methods in Molecular Biology, vol 2*. New York: Humana Press. pp 31–34.

McCarthy EM. 2006. *Handbook of avian hybrids of the world*. Oxford: Oxford University Press.

Miller MP, Mullins TD, Forsman ED, Haig SM. 2017. Genetic differentiation and inferred dynamics of a hybrid zone between Northern Spotted Owls (*Strix occidentalis caurina*) and California Spotted Owls (*S. o. occidentalis*) in northern California. *Ecology and Evolution* **7**: 6871–6883.

Milstein PLS. 1979. The evolutionary significance of wild hybridisation in South Africa highveld ducks. *Ostrich* **13**: 1–48.

Milstein PLS, Osterhoff DR. 1975. Blood-tests identify hybrid wild ducks new to science. *Bokmakierie* 27: 70–71.

Morgan JAT, Macbeth M, Broderick D, Whatmore P, Street R, Welch DJ, Ovenden JR. 2013. Hybridisation, paternal leakage and mitochondrial DNA linearization in three anomalous fish (Scombridae). *Mitochondrion* **13**: 852–861

Mulvihill RS, Parkes KC, Leberman RC, Wood DS. 1992. Evidence supporting a dietary basis for orange-tipped rectrices in the Cedar Waxwing. *Journal of Field Ornithology* **63**: 212–216.

Murphy KM, Berg KD, Eshleman JR. 2005. Sequencing of genomic DNA by combined amplification and cycle sequencing reaction. *Clinical Chemistry* **51**: 35–39.

Novak SJ, Mack RN. 2001. Tracing plant introduction and spread: genetic evidence from *Bromus tectorum* (Cheatgrass). *Bioscience* **51**: 114–122.

Omland KE. 1997. Examining two standard assumptions of ancestral reconstructions: repeated loss of dichromatism in dabbling ducks (Anatini). *Evolution* **51**: 1636–1646.

Owen M, Callaghan D, Kirby J. 2006. Guidelines on avoidance of introductions of non-native waterbird species. *AEWA Technical Series*. Available at: http://www.unep-aewa.org/sites/default/files/publication/ts12_guidelines_non-native-species_complete_0.pdf [Accessed 14 May 2017].

Parris MJ. 2004. Hybrid response to pathogen infection in interspecific crosses between two amphibian species (Anura: Ranidae). *Evolutionary Ecology Research* **6**: 457–471.

Paulus KB, Tiedemann R. 2003. Ten polymorphic autosomal microsatellite loci for the Eider Duck, *Somateria mollissima*, and their cross-species applicability among waterfowl species (Anatidae). *Molecular Ecology Notes* **3**: 250–252.

Perez-Arteaga A, Gaston K, Kershaw M. 2002. Population trends and priority conservation sites for Mexican Duck, *Anas diazi*. *Bird Conservation International* **12**: 35–52.

Peters JL, Sonsthagen SA, Lavretsky P, Rezsutek M, Johnson WP, McCracken KG. 2014. Interspecific hybridization contributes to high genetic diversity and apparent effective population size in an endemic population of Mottled Ducks (*Anas fulvigula maculosa*). *Conservation Genetics* **15**: 509–520.

Posada D, Buckley T. 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**: 793–808.

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

R Core Team. 2018. R: a language and environment for statistical computing. *R Foundation for Statistical Computing*. Available at: https://www.r-project.org/ [Accessed 15 August 2018]

Randler C. 2002. Avian hybridization, mixed pairing and female choice. *Animal Behaviour* **63**: 103–119.

Reynolds C. 2016. The role of waterbirds in the dispersal of aquatic organisms in southern Africa. PhD Thesis, University of Cape Town, South Africa.

Rhymer JM. 2001. Evolutionary relationships and conservation of the hawaiian anatids. *Studies in Avian Biology* **22**: 61–67.

Rhymer JM. 2006. Extinction by hybridization and introgression in anatine ducks. *Acta Zoologica Sinica* **52**: 583–585.

Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* **27**: 83–109.

Robinson TJ, Morris DJ, Fairall N. 1991. Interspecific hybridization in the bovidae: Sterility of *Alcelaphus buselaphus* × *Damaliscus dorcas* F1 progeny. *Biological Conservation* **58**: 345–356.

Rogers D. 2001. *Anas platyrhynchos*. *Animal Diversity Web*. Available at: http://animaldiversity.org/accounts/Anas_platyrhynchos/ [Accessed 30 March 2017].

Ryan ME, Johnson JR, Fitzpatrick BM. 2009. Invasive hybrid tiger salamander genotypes impact native amphibians. *Proceedings of the National Academy of Sciences* **106**: 11166–11171.

SABAP2 (Southern African Bird Atlas Project 2). 2019. Southern African Bird Atlas Project 2. Available at: http://sabap2.adu.org.za/species_info.php?spp=1016&section=5#menu_left [Accessed 21 January 2019].

SAFRING. 2019. Yellow-billed Duck. Available at http://safring.adu.org.za/Species_summary.php [Accessed 07 February 2019].

SANBI (South African National Biodiversity Institute). 2019. Biodiversity. Available at: https://www.sanbi.org/biodiversity/ [Accessed 12 February 2019].

Sanger F, Nicklen S, Coulson R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**: 5463–5467.

Schwartz MK, Luikart G, Waples R. 2007. Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution* **22**: 25–33.

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

Shapiro B, Rambaut A, Drummond AJ. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Molecular Biology and Evolution* **23**: 7–9.

Sinclair I, Hockey P, Tarboton W, Ryan P. 2011. *Sasol birds of southern Africa*. (4th edn). Cape Town: Struik Nature.

Soto-Calderón ID, Acevedo-Garcés YA, Restrepo-Agudelo T, Llinás-Guerrero JC, Rivillas-Puello Y. López JB. 2018. Phenotypic, genetic, and cytogenetic evidence of hybridization between species of Trans-Andean Tamarins (Genus *Saguinus*). *International Journal of Primatology* **39**: 1022–1038.

Stai SM, Hughes CR. 2003. Characterization of microsatellite loci in wild and domestic Muscovy Ducks *(Cairina moschata*). *Animal Genetics* **34**: 387–389.

Su B, Wang YX, Lan H, Wang W, Zhang Y. 1999. Phylogenetic study of complete *cytochrome b* genes in Musk Deer (genus Moschus) using museum samples. *Molecular Phylogenetics and Evolution* **12**: 241–249.

Tavares ES, Baker AJ. 2008. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evolutionary Biology* **8**: 81–94.

Taysom A, Johnson J, Guay PJ. 2014. Establishing a genetic system to distinguish between domestic Mallards, Pacific Black Ducks and their hybrids. *Conservation Genetics Resources* **6**: 197–199.

Tompkins DM, Mitchell RA, Bryant DM. 2006. Hybridization increases measures of innate and cell-mediated immunity in an endangered bird species. *Journal of Animal Ecology* **75**: 559–564.

Urantowka AD, Kroczak A, Mackiewicz P. 2017. The influence of molecular markers and methods on inferring the phylogenetic relationships between the representatives of the Arini (parrots, Psittaciformes), determined on the basis of their complete mitochondrial genomes. *BMC Evolutionary Biology* **17**: 166–191.

Vähä JP, Primmer CR. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**: 63–72.

Valentini A, Pompanon F, Taberlet P. 2009. DNA barcoding for ecologists. *Trends in Ecology and Evolution* **24**: 110–117.

Väli L, Saag P, Dombrovski V, Meyburg BU, Maciorowski G, Mizera T, Treinys R, Fagerberg S. 2010. Microsatellites and single nucleotide polymorphisms in avian hybrid identification: a comparative case study. *Journal of Avian Biology* **41**: 34–49.

Vallender R, Robertson RJ, Friesen VL, Lovette IJ. 2007. Complex hybridization dynamics between Golden-winged and Blue-winged Warblers (*Vermivora chrysoptera* and *Vermivora pinus*) revealed by AFLP, microsatellite, intron and mtDNA markers. *Molecular Ecology* **16**: 2017–2029.

VanLiere JM, Rosenberg NA. 2008. Mathematical properties of the $r^2$ measure of linkage disequilibrium. *Theoretical Population Biology* **74**: 130–137.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**: 535–538.

van Wyk AM, Dalton DL, Hoban S, Bruford MW, Russo IRM, Birss C, Grobler P, van Vuuren BJ, Kotzé A. 2017. Quantitative evaluation of hybridization and the impact on biodiversity conservation. *Ecology and Evolution* **7**: 320–330.

van Wyk AM, Kotzé A, Randi E, Dalton DL. 2013. A hybrid dilemma: a molecular investigation of South African Bontebok (*Damaliscus pygargus pygargus*) and Blesbok (*Damaliscus pygargus phillipsi*). *Conservation Genetics* **14**: 589–599.

Vilà M, Basnou C, Pyšek P, Josefsson M, Genovesi P, Gollasch S, Nentwig W, Olenin S, Roques A, Roy D, Hulme PE. 2010. How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Frontiers in Ecology and the Environment* **8**: 135–144.

Vuillaume B, Valette V, Lepais O, Grandjean F, Breuil M. 2015. Genetic evidence of hybridization between the endangered native species *Iguana delicatissima* and the invasive *Iguana iguana* (Reptilia, Iguanidae) in the Lesser Antilles: management implications. *PLoS ONE* **10**: e0127575.

Walthe GR, Roques A, Hulme PE, Sykes MT, Pysek P, Kuhn I, Zobel M. 2009. Alien species in a warmer world: risks and opportunities. *Trends in Ecology & Evolution* **24**: 686–693.

Wang W, Wang Y, Lei F, Liu Y, Wang H, Chen J. 2018. Incomplete lineage sorting and introgression in the diversification of Chinese Spot-billed Ducks and Mallards. *Current Zoology* zoy074.

Watanabe T, Mizutani M, Wakana S, Tomita T. 1985. Demonstration of the maternal inheritance of avian mitochondrial DNA in chicken-quail hybrids. *Journal of Experimental Zoology* **236**: 245–247.

Wattier R, Engel CR, Saumitou-Laprade P, Valero M. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Molecular Ecology* **7**: 1569–1573.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis.

Williams CL, Brust RC, Fendley TT, Tiller GR, Rhodes OE. 2005. A Comparison of hybridization between Mottled Ducks (*Anas fulvigula*) and Mallards (*A. platyrhynchos*) in Florida and South Carolina using microsatellite DNA analysis. *Conservation Genetics* **6**: 445–453.

Williams M, Basse B. 2006. Indigenous Gray Ducks, *Anas superciliosa*, and introduced Mallards, *A. platyrhynchos*, in New Zealand: processes and outcome of a deliberate encounter. *Acta Zoologica Sinica* **52**: 579–582.

Willis PM. 2013. Why do animals hybridize? *Acta Ethologica* **16**: 127–134.

Wold S, Esbensen KIM, Geladi P. 1987. Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems* **2**: 37–52.

Wood-Gush D. 2012. *Elements of Ethology: A textbook for agricultural and veterinary students*. Berlin: Springer Science & Business Media.

Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**: 395–420.

Wright S. 1978. *Evolution and the genetics of population variability within and among natural populations, vol 4*. Chicago: University of Chicago Press.

Yamazaki Y, Kouketsu S, Fukuda T, Araki Y, Nambu H. 2008. Natural hybridization and directional introgression of two species of Japanese toads *Bufo japonicus formosus* and *Bufo torrenticola* (Anura : Bufonidae) resulting from changes in their spawning habitat. *Journal of Herpetology* **42**: 427–436.

Yilmaz M, Ozic C, Gok İ. 2012. Principles of nucleic acid separation by agarose gel electrophoresis. In: *Gel Electrophoresis - Principles and Basics*. Rijeka: IntechOpen. pp 33–40.

Yoo HS, Eah JY, Kim JS, Kim YJ, Min MS, Paek WK, Lee H, Kim CB. 2006. DNA barcoding Korean birds. *Molecules and Cells* **22**: 323–327.

Zaya DN, Leicht-Young SA, Pavlovic NB, Feldheim KA, Ashley MV. 2015. Genetic characterization of hybridization between native and invasive Bittersweet Vines (*Celastrus* spp.). *Biological Invasions* **17**: 2975–2988.

# Appendices

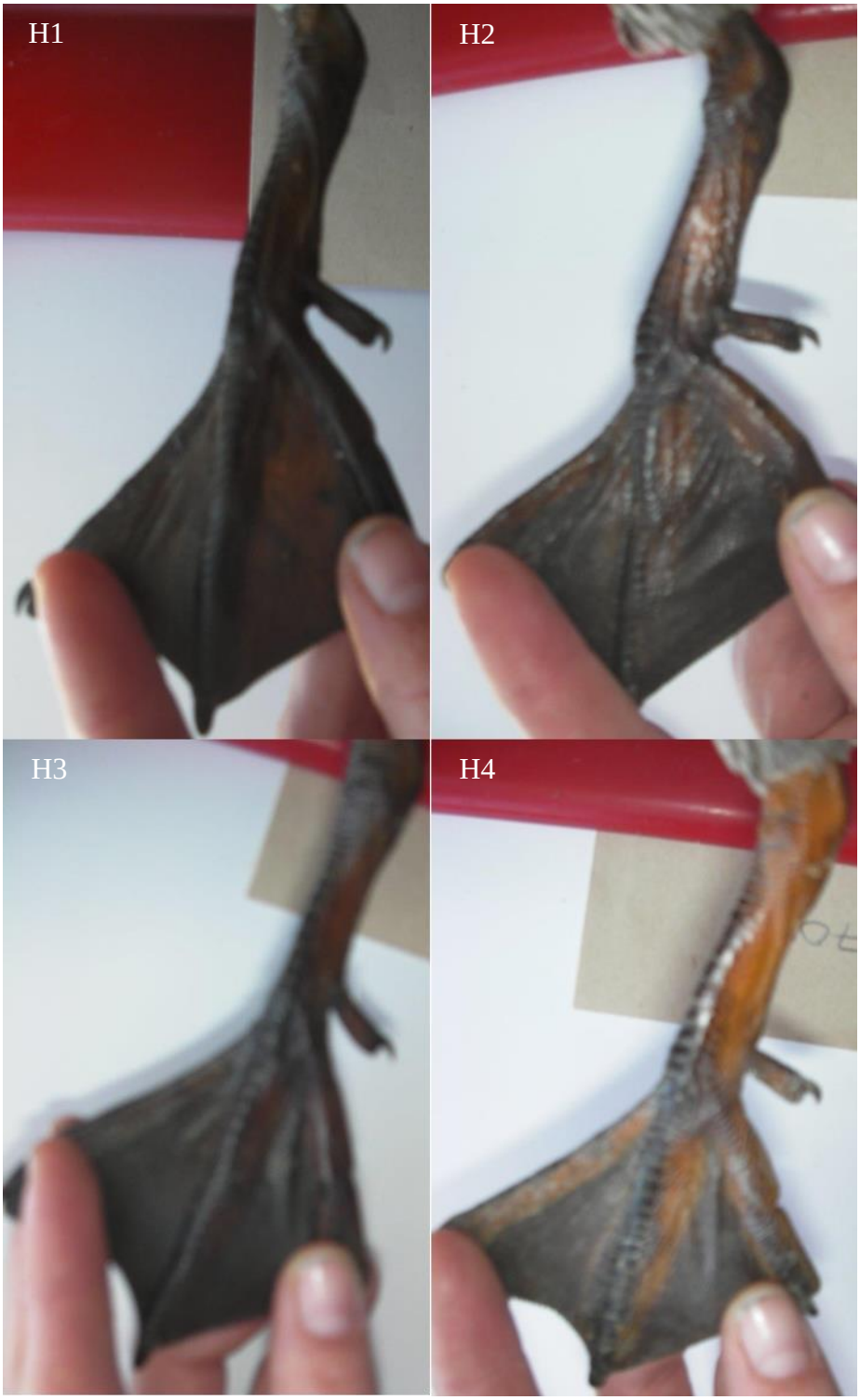


**Figure A1**: Images of the feet of all putative hybrid samples (H1-H8). Photographs accredited to C Reynolds.

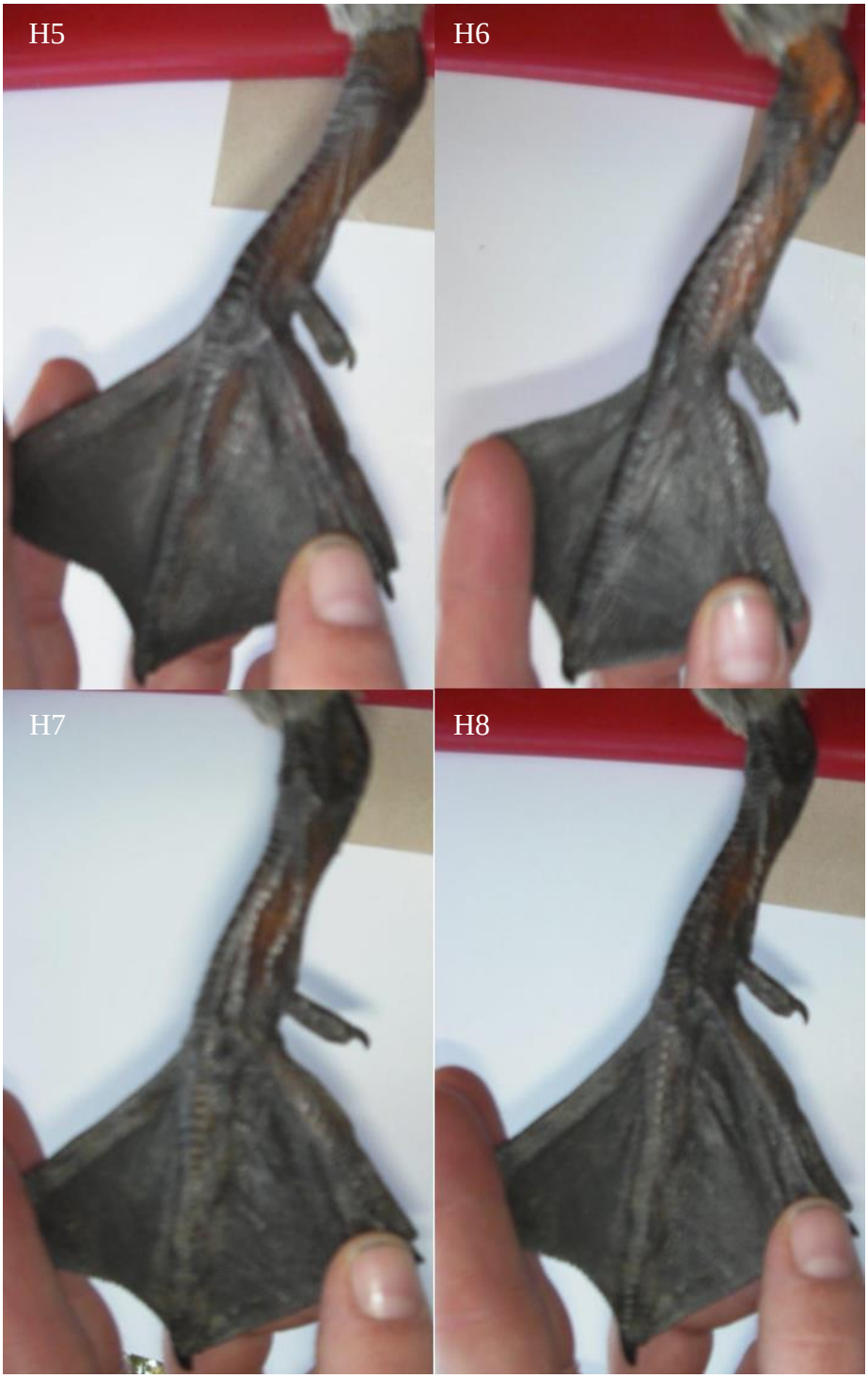


**Figure A1**: Images of the feet of all putative hybrid samples (H1-H8). Photographs accredited to C Reynolds.

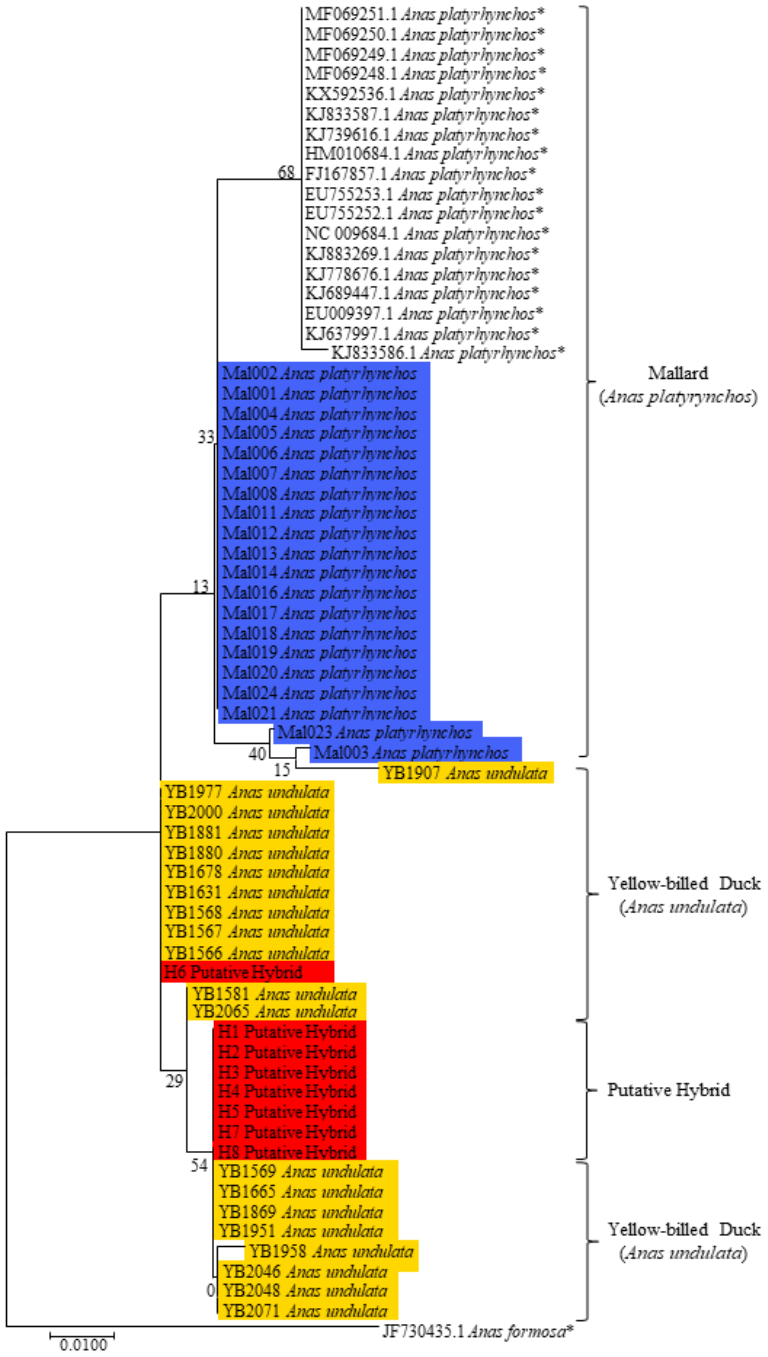


**Figure A2**: A maximum-likelihood phylogenetic tree constructed using *cytochrome b* (*cyt b*) mitochondrial DNA sequences (251 bp). The Hasegawa-Kishino-Yano (HKY) evolutionary model was used in a bootstrap test (1000 replicates) using the Baikal teal (*Anas Formosa*) as an outgroup. Sequences highlighted in blue, red and yellow indicate Mallard (*Anas platyrhynchos*), putative hybrid and Yellow-billed Duck (*Anas undulata*) samples used in this study, respectively. * indicates sequences obtained from GenBank.

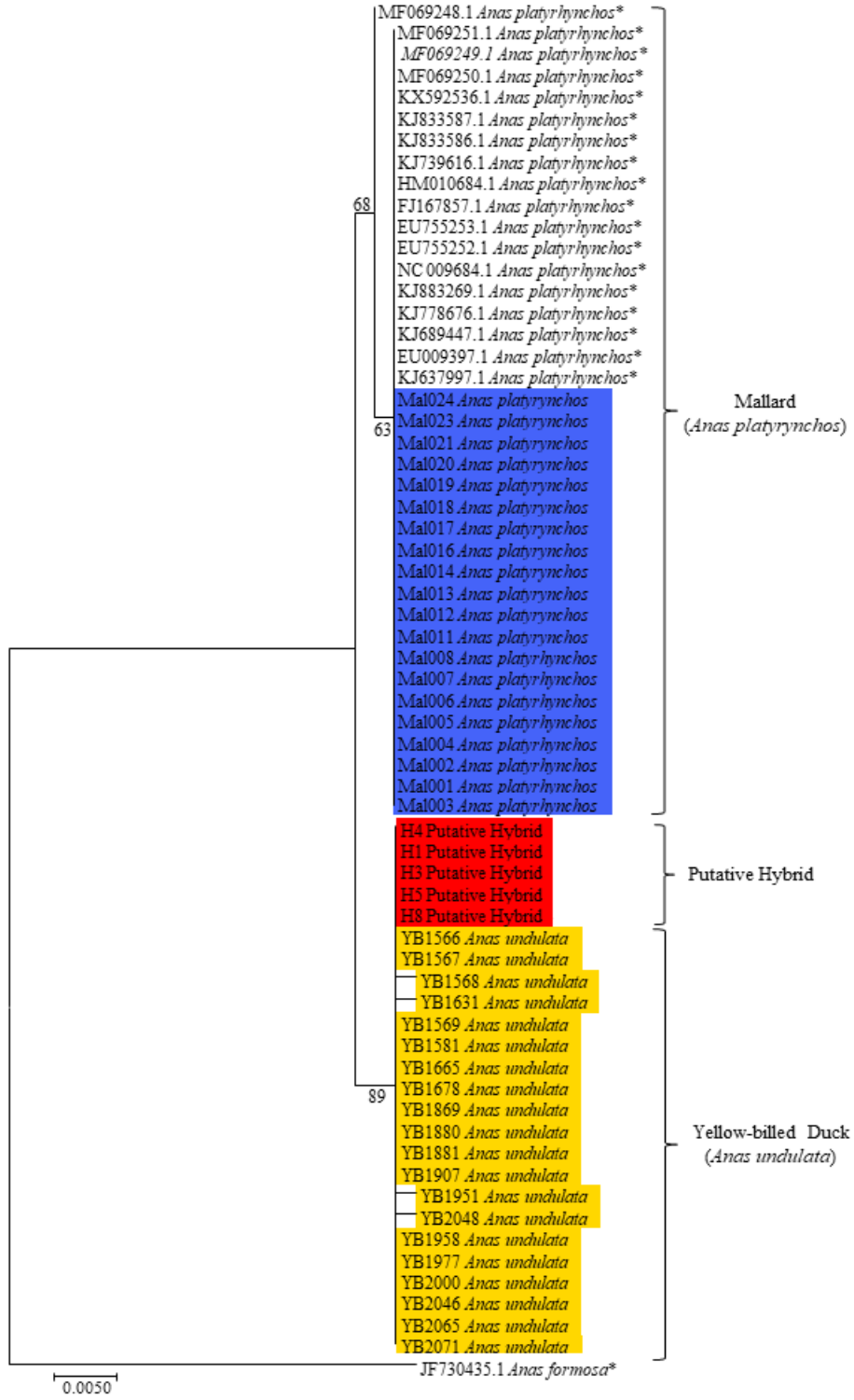


**Figure A3**: A maximum-likelihood phylogenetic tree constructed using *cytochrome c oxidase subunit I* (*COI*) mitochondrial DNA sequences (634 bp). The Hasegawa-Kishino-Yano (HKY) evolutionary model was used in a bootstrap test (1000 replicates) using the Baikal teal (*Anas Formosa*) as an outgroup. Sequences highlighted in blue, red and yellow indicate Mallard (*Anas platyrhynchos*), putative hybrid and Yellow-billed Duck (*Anas undulata*) samples used in this study, respectively. * indicates sequences obtained from GenBank.

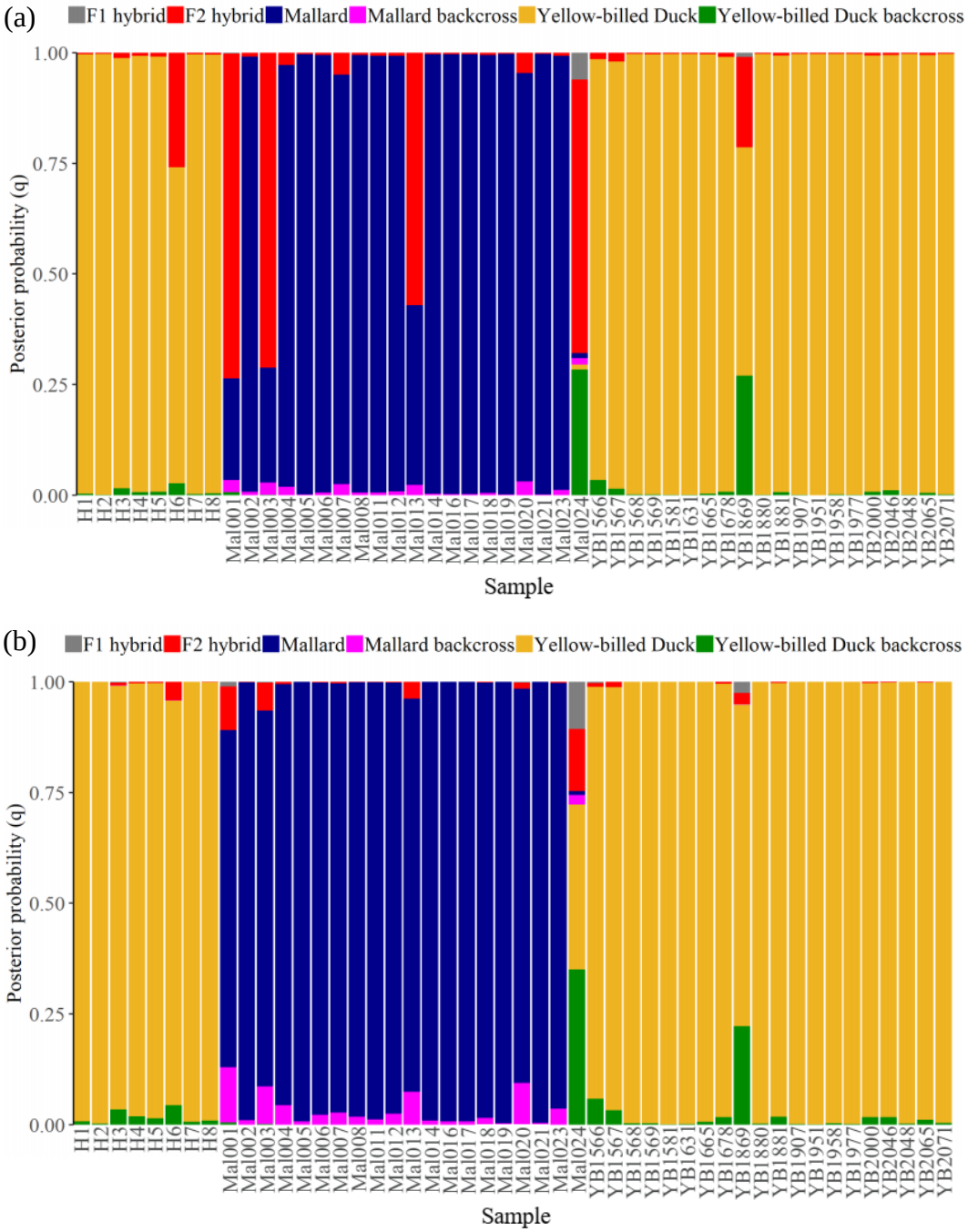


**Figure A4**: NEWHYBRIDS analyses of microsatellite genotypes for putative hybrids, Mallards (*Anas platyrhynchos*) and Yellow-billed Ducks (*Anas undulata*). Each individual is represented by a single horizontal line, with lengths proportional to the estimated membership for each sub-structure class. Putative hybrids: samples H1-H8; Mallards: samples Mal001-Mal024 and Yellow-billed Ducks: samples YB1566-YB2071. (a) Jeffreys-like prior. (b) Uniform prior.

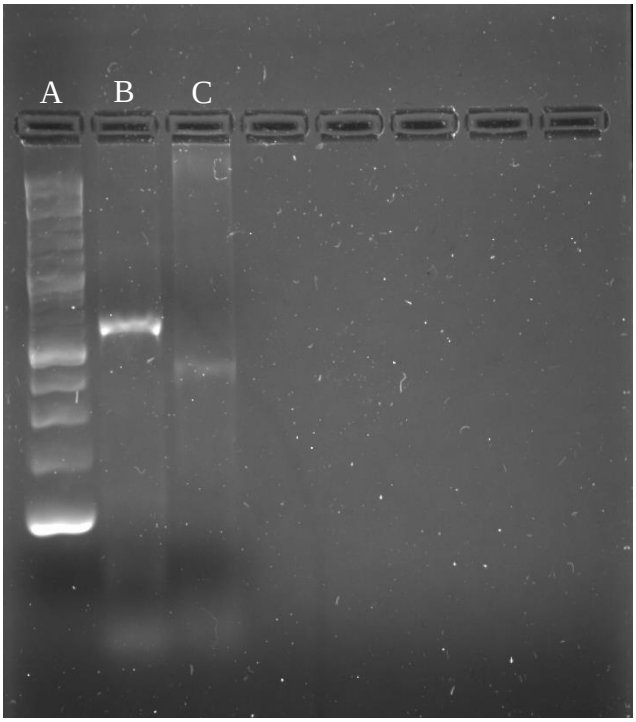


**Figure A5**: A 1.5% agarose gel displaying amplification of the sex-linked *chromodomain-helicase-DNA* (*CHD*) gene. A: 100 bp DNA ladder. B: Male used as a positive control. C: Phenotypic Mallard (*Anas platyrhynchos*) sample Mal024.

**Table A1**: A list of all the Anatidae species that have been identified to hybridise with the Mallard (*Anas platyrhynchos*).

| **Common Name** | **Scientific Name** |
|---|---|
| Mandarin Duck | *Aix galericulata* |
| Wood Duck | *Aix sponsa* |
| Egyptian Goose | *Alopochen aegyptiacus* |
| Common Pintail | *Anas acuta* |
| American Wigeon | *Anas americana* |
| Flightless Teal | *Anas aucklandica* |
| Bahama Pintail | *Anas bahamensis* |
| Chestnut-breasted Teal | *Anas castanea* |
| Brown Teal | *Anas chlorotis* |
| Northern Shoveler | *Anas clypeata* |
| Common Teal | *Anas crecca* |
| Mexican Duck | *Anas diazi* |
| Blue-winged Teal | *Anas discors* |
| Falcated Teal | *Anas falcata* |
| Speckled Teal | *Anas flavirostris* |
| Mottled Duck | *Anas fulvigula* |
| Yellow-billed Pintail | *Anas georgica* |
| Sunda Teal | *Anas gibberifrons* |
| Grey Teal | *Anas gracilis* |
| Laysan Duck | *Anas laysanensis* |
| Philippine Duck | *Anas luzonica* |
| Meller's Duck | *Anas melleri* |
| Eurasian Wigeon | *Anas penelope* |
| Spot-billed Duck | *Anas poecilorhyncha* |
| Garganey | *Anas querquedula* |
| American Black Duck | *Anas rubripes* |
| Chiloe Wigeon | *Anas sibilatrix* |
| Cape Shoveler | *Anas smithii* |
| African Black Duck | *Anas sparsa* |
| Spectacled Duck | *Anas specularis* |

**Table A1 - continued from previous page**

| Common Name | Scientific Name |
|---|---|
| Gadwall | *Anas strepera* |
| Pacific Black Duck | *Anas superciliosa* |
| Yellow-billed Duck | *Anas undulata* |
| Versicolor Teal | *Anas versicolor* |
| Hawaiian Duck | *Anas wyvilliana* |
| Greylag Goose | *Anser anser* |
| Swan Goose | *Anser cygnoides* |
| American Redhead | *Aythya americana* |
| Hardhead | *Aythya australis* |
| Ring-necked Duck | *Aythya collaris* |
| European Pochard | *Aythya ferina* |
| Tufted Duck | *Aythya fuligula* |
| Ferruginous Pochard | *Aythya nyroca* |
| Canvasback | *Aythya valisineria* |
| Canada Goose | *Branta canadensis* |
| Barnacle Goose | *Branta leucopsis* |
| Goldeneye | *Bucephala clangula* |
| Muscovy Duck | *Cairina moschata* |
| White-winged Wood Duck | *Cairina scutulata* |
| Goosander | *Mergus merganser* |
| Red-breasted Merganser | *Mergus serrator* |
| Rosybill | *Netta peposacea* |
| Red-crested Pochard | *Netta rufina* |
| Steller's Eider | *Polysticta stelleri* |
| Knob-billed Duck | *Sarkidiornis melanotos* |
| Common Eider | *Somateria mollissima* |
| Ruddy Shelduck | *Tadorna ferruginea* |
| Common Shelduck | *Tadorna tadorna* |
| Australian Shelduck | *Tadorna tadornoides* |
| New Zealand Shelduck | *Tadorna variegata* |

**Table A2**: A list of all the dabbling ducks belonging to the Mallard Complex.

| **Common Name** | **Scientific Name** |
|---|---|
| Mexican Duck | *Anas diazi* |
| Mottled Duck | *Anas fulvigula* |
| Florida Mottled Duck | *A. f. fulvigula* |
| Gulf Coast Mottled Duck | *A. f. maculosa* |
| Laysan Duck | *Anas laysanensis* |
| Philippine Duck | *Anas luzonica* |
| Meller's Duck | *Anas melleri* |
| Indian Spot-billed Duck | *Anas poecilorhyncha* |
| Mallard | *Anas platyrhynchos* |
| American Black Duck | *Anas rubripes* |
| African Black Duck | *Anas sparsa* |
| Pacific Black Duck | *Anas superciliosa* |
| Lesser Grey Duck | *A. s. pelewensis* |
| Australian Black Duck | *A. s. rogersi* |
| New Zealand Grey Duck | *A. s. superciliosa* |
| Yellow-billed Duck | *Anas undulata* |
| Northern Yellow-billed Duck | *A. u. rueppelli* |
| Southern Yellow-billed Duck | *A. u. undulata* |
| Hawaiian Duck | *Anas wyvilliana* |
| Chinese Spot-billed Duck | *Anas zonorhyncha* |

**Table A3**: Ring numbers for each of the putative hybrid and Yellow-billed Duck (*Anas undulata*) samples used in this study.

| Sample | Population | Ring Number |
|---|---|---|
| H1 | Putative Hybrid | K40839 |
| H2 | Putative Hybrid | K40813 |
| H3 | Putative Hybrid | K40814 |
| H4 | Putative Hybrid | K40815 |
| H5 | Putative Hybrid | K40816 |
| H6 | Putative Hybrid | K40817 |
| H7 | Putative Hybrid | K40818 |
| H8 | Putative Hybrid | K40819 |
| YB1566 | Yellow-billed Duck | K40768 |
| YB1567 | Yellow-billed Duck | K40769 |
| YB1568 | Yellow-billed Duck | K40770 |
| YB1569 | Yellow-billed Duck | K40801 |
| YB1581 | Yellow-billed Duck | K38502 |
| YB1631 | Yellow-billed Duck | K40773 |
| YB1665 | Yellow-billed Duck | K40785 |
| YB1678 | Yellow-billed Duck | K41142 |
| YB1869 | Yellow-billed Duck | K41221 |
| YB1880 | Yellow-billed Duck | K41238 |
| YB1881 | Yellow-billed Duck | K41222 |
| YB1907 | Yellow-billed Duck | K41230 |
| YB1951 | Yellow-billed Duck | K41249 |
| YB1958 | Yellow-billed Duck | K47503 |
| YB1977 | Yellow-billed Duck | K47519 |
| YB2000 | Yellow-billed Duck | K47525 |
| YB2046 | Yellow-billed Duck | K47540 |
| YB2048 | Yellow-billed Duck | K35543 |
| YB2065 | Yellow-billed Duck | K41172 |
| YB2071 | Yellow-billed Duck | K41177 |

**Table A4**: The concentration and purity ratios of DNA extracted from all samples used in this study.

| Sample | Population | Concentration (ng/$\mu$l) | A260/280 | A260/230 |
|---|---|---|---|---|
| H1 | Putative Hybrid | 94.7 | 1.72 | 2.02 |
| H2 | Putative Hybrid | 538.5 | 1.73 | 1.9 |
| H3 | Putative Hybrid | 318.6 | 1.7 | 1.76 |
| H4 | Putative Hybrid | 1131.2 | 1.79 | 1.88 |
| H5 | Putative Hybrid | 385.9 | 1.77 | 2.08 |
| H6 | Putative Hybrid | 90.6 | 1.71 | 1.85 |
| H7 | Putative Hybrid | 416.4 | 1.71 | 1.66 |
| H8 | Putative Hybrid | 57.1 | 1.57 | 2.09 |
| Mal001 | Mallard | 323.7 | 1.86 | 2.03 |
| Mal002 | Mallard | 155.7 | 1.79 | 0.45 |
| Mal003 | Mallard | 74.6 | 1.86 | 1.63 |
| Mal004 | Mallard | 136.1 | 1.89 | 1.82 |
| Mal005 | Mallard | 163.9 | 1.86 | 1.75 |
| Mal006 | Mallard | 100.9 | 1.59 | 1.45 |
| Mal007 | Mallard | 28.3 | 1.71 | 1.56 |
| Mal008 | Mallard | 108.3 | 1.83 | 1.71 |
| Mal011 | Mallard | 52.7 | 1.6 | 1.61 |
| Mal012 | Mallard | 64.4 | 1.52 | 1.61 |
| Mal013 | Mallard | 221.4 | 1.94 | 1.75 |
| Mal014 | Mallard | 76.7 | 1.52 | 1.71 |
| Mal016 | Mallard | 341.5 | 1.86 | 2.04 |
| Mal017 | Mallard | 47 | 1.7 | 1.01 |
| Mal018 | Mallard | 15.9 | 1.66 | 1.54 |
| Mal019 | Mallard | 64.9 | 1.7 | 1.97 |
| Mal020 | Mallard | 56.6 | 1.57 | 1.77 |
| Mal021 | Mallard | 15.6 | 1.41 | 0.92 |
| Mal023 | Mallard | 71.9 | 1.75 | 1.6 |
| Mal024 | Mallard | 3181 | 1.96 | 1.08 |

Continued on next page

**Table A4 - continued from previous page**

| Sample | Population | Concentration (ng/$\mu$l) | A260/280 | A260/230 |
|---|---|---|---|---|
| YB1566 | Yellow-billed Duck | 138.3 | 1.71 | 1.93 |
| YB1567 | Yellow-billed Duck | 184.5 | 1.75 | 1.96 |
| YB1568 | Yellow-billed Duck | 45.4 | 1.62 | 1.67 |
| YB1569 | Yellow-billed Duck | 145.4 | 1.59 | 1.93 |
| YB1581 | Yellow-billed Duck | 142.4 | 1.69 | 2.17 |
| YB1631 | Yellow-billed Duck | 102.9 | 1.55 | 1.26 |
| YB1665 | Yellow-billed Duck | 67.7 | 1.5 | 1.39 |
| YB1678 | Yellow-billed Duck | 199.2 | 1.68 | 1.99 |
| YB1869 | Yellow-billed Duck | 347.1 | 1.81 | 1.78 |
| YB1880 | Yellow-billed Duck | 96.3 | 1.56 | 1.49 |
| YB1881 | Yellow-billed Duck | 267.6 | 1.8 | 1.92 |
| YB1907 | Yellow-billed Duck | 224.7 | 1.6 | 0.74 |
| YB1951 | Yellow-billed Duck | 135.9 | 1.6 | 1.63 |
| YB1958 | Yellow-billed Duck | 135.6 | 1.59 | 0.95 |
| YB1977 | Yellow-billed Duck | 109.2 | 1.65 | 1.47 |
| YB2000 | Yellow-billed Duck | 128.5 | 1.67 | 1.19 |
| YB2046 | Yellow-billed Duck | 208.8 | 1.51 | 0.92 |
| YB2048 | Yellow-billed Duck | 108.8 | 1.61 | 1.41 |
| YB2065 | Yellow-billed Duck | 118.8 | 1.54 | 0.61 |
| YB2071 | Yellow-billed Duck | 325.8 | 1.8 | 2.43 |
| | **Average** | **240.98** | **1.69** | **1.61** |

**Table A5**: Results of the model test conducted in MEGA, showing the BIC value for each of the evolutionary models.

| Model | BIC |
|---|---|
| HKY+G | 4250.66 |
| HKY | 4250.77 |
| HKY+I | 4255.79 |
| TN93+G | 4259.39 |
| TN93 | 4259.45 |
| TN93+I | 4264.52 |
| HKY+G+I | 4265.25 |
| TN93+G+I | 4273.4 |
| K2 | 4275.54 |
| K2+G | 4276.05 |
| K2+I | 4280.71 |
| T92 | 4284.24 |
| T92+G | 4284.66 |
| GTR+G | 4285.3 |
| GTR | 4286.55 |
| T92+I | 4289.39 |
| K2+G+I | 4291.68 |
| GTR+I | 4292.09 |
| T92+G+I | 4300.42 |
| GTR+G+I | 4301.07 |
| JC | 4393.96 |
| JC+G | 4396.19 |
| JC+I | 4399.63 |
| JC+G+I | 4406.09 |

**Table A6**: Frequency of private and shared alleles per locus within the Yellow-billed Ducks (*Anas undulata*), Mallards (*Anas platyrhynchos*), and putative hybrids.

| | | Frequency of alleles | | |
|---|---|---|---|---|
| **Locus** | **Allele** | **Yellow-billed Ducks** | **Mallards** | **Putative Hybrids** |
| APH17 | **234** | 0 | 0.025 | 0 |
| | 236 | 0.7 | 0.125 | 0.4375 |
| | 238 | 0.175 | 0.025 | 0.25 |
| | 240 | 0.125 | 0.825 | 0.3125 |
| APH19 | 178 | 0.225 | 0.225 | 0.5 |
| | 180 | 0.175 | 0.125 | 0.1875 |
| | 182 | 0.2 | 0.275 | 0.25 |
| | 184 | 0.4 | 0 | 0.0625 |
| | **186** | 0 | 0.05 | 0 |
| | **188** | 0 | 0.325 | 0 |
| APH21 | 152 | 0.15 | 0.5 | 0 |
| | 154 | 0.075 | 0.075 | 0.25 |
| | 156 | 0.775 | 0.425 | 0.75 |
| APH25 | 165 | 0.825 | 0.45 | 0.75 |
| | 167 | 0.175 | 0.55 | 0.25 |
| Apl11 | **109** | 0 | 0.05 | 0 |
| | 111 | 0.075 | 0 | 0.1875 |
| | 115 | 0.025 | 0 | 0.0625 |
| | 117 | 0.05 | 0.075 | 0.125 |
| | 119 | 0.425 | 0.225 | 0.3125 |
| | 121 | 0.425 | 0.175 | 0.3125 |
| | **125** | 0 | 0.075 | 0 |
| | **127** | 0 | 0.25 | 0 |
| | **129** | 0 | 0.15 | 0 |
| Bcau6 | 165 | 0.85 | 0.425 | 0.875 |
| | 167 | 0.15 | 0.575 | 0.125 |
| Bcau11 | 152 | 0.325 | 0.55 | 0.5 |
| | 156 | 0.675 | 0.45 | 0.5 |
| Blm12 | 257 | 0 | 0.125 | 0.75 |
| | 261 | 0.025 | 0.775 | 0.125 |
| | 265 | 0.725 | 0.1 | 0 |
| | **269** | 0.2 | 0 | 0 |

Bold text represents private alleles

Continued on next page

**Table A6 - continued from previous page**

| Locus | Allele | Frequency of alleles | | |
|---|---|---|---|---|
| | | **Yellow-billed Ducks** | **Mallards** | **Putative Hybrids** |
| CM09 | 114 | 0.175 | 0.05 | 0.125 |
| | **116** | 0 | 0.35 | 0 |
| | 120 | 0.25 | 0 | 0.375 |
| | 122 | 0.25 | 0 | 0.125 |
| | 124 | 0.325 | 0.6 | 0.375 |
| Sfiu4 | 161 | 0.025 | 0.45 | 0 |
| | 163 | 0.875 | 0.175 | 1 |
| | 167 | 0.05 | 0.375 | 0 |
| | **169** | 0.05 | 0 | 0 |
| Smo7 | **200** | 0 | 0.175 | 0 |
| | **202** | 0 | 0.6 | 0 |
| | 204 | 0.525 | 0.075 | 0.3125 |
| | 206 | 0.05 | 0.05 | 0.0625 |
| | 208 | 0.275 | 0.05 | 0.4375 |
| | 210 | 0.025 | 0.05 | 0.125 |
| | 212 | 0.125 | 0 | 0.0625 |

Bold text represents private alleles

**Table A7**: Analysis of molecular variance (AMOVA) within each population.

| Species | Source of variation | Df | SS | Va | Variation (%) | $F_{IS}$ |
|---|---|---|---|---|---|---|
| Yellow-billed Duck | Among individuals within populations | 19 | 54.78 | 0.42895 | 17.48 | |
| | Within individuals | 20 | 40.5 | 2.025 | 82.52 | |
| | **Total** | **39** | **95.28** | **2.45395** | | **0.17** |
| Mallard | Among individuals within populations | 19 | 56.53 | -0.1125 | -3.64 | |
| | Within individuals | 20 | 64 | 3.2 | 103.64 | |
| | **Total** | **39** | **120.53** | **3.0875** | | **-0.04** |
| Putative Hybrid | Among individuals within populations | 7 | 19.63 | 0.18304 | 6.98 | |
| | Within individuals | 8 | 19.5 | 2.4375 | 93.02 | |
| | **Total** | **15** | **39.13** | **2.62054** | | **0.07** |

Df - degrees of freedom

SS - sum of squares

Va - variance

**Table A8**: The clustering of each individual into each haplotype group.

| Haplotype Group | Sample |
|---|---|
| Group A1 | Mal001 *Anas platyrhynchos* |
| | Mal002 *Anas platyrhynchos* |
| | Mal004 *Anas platyrhynchos* |
| | Mal005 *Anas platyrhynchos* |
| | Mal006 *Anas platyrhynchos* |
| | Mal007 *Anas platyrhynchos* |
| | Mal008 *Anas platyrhynchos* |
| | Mal011 *Anas platyrhynchos* |
| | Mal012 *Anas platyrhynchos* |
| | Mal013 *Anas platyrhynchos* |
| | Mal014 *Anas platyrhynchos* |
| | Mal016 *Anas platyrhynchos* |
| | Mal017 *Anas platyrhynchos* |
| | Mal018 *Anas platyrhynchos* |
| | Mal019 *Anas platyrhynchos* |
| | Mal020 *Anas platyrhynchos* |
| | Mal021 *Anas platyrhynchos* |
| | Mal024 *Anas platyrhynchos* |
| 2 | Mal023 *Anas platyrhynchos* |
| 3 | Mal003 *Anas platyrhynchos* |
| 4 | KJ637997.1 *Anas platyrhynchos* |
| | EU009397.1 *Anas platyrhynchos* |
| | KJ689447.1 *Anas platyrhynchos* |
| | KJ778676.1 *Anas platyrhynchos* |
| | KJ883269.1 *Anas platyrhynchos* |
| | NC 009684.1 *Anas platyrhynchos* |
| | EU755252.1 *Anas platyrhynchos* |
| | EU755253.1 *Anas platyrhynchos* |
| | FJ167857.1 *Anas platyrhynchos* |
| | HM010684.1 *Anas platyrhynchos* |
| | KJ739616.1 *Anas platyrhynchos* |
| | KJ833587.1 *Anas platyrhynchos* |
| | KX592536.1 *Anas platyrhynchos* |
| | MF069249.1 *Anas platyrhynchos* |
| | MF069250.1 *Anas platyrhynchos* |
| | MF069251.1 *Anas platyrhynchos* |
| 5 | KJ833586.1 *Anas platyrhynchos* |
| 6 | MF069248.1 *Anas platyrhynchos* |

Continued on next page

**Table A8 - continued from previous page**

| Haplotype Group | Sample |
|---|---|
| Group B1 | YB1566 A*nas undulata* |
| | YB1567 *Anas undulata* |
| | YB1678 *Anas undulata* |
| | YB1880 *Anas undulata* |
| | YB1881 *Anas undulata* |
| | YB1977 *Anas undulata* |
| | YB2000 *Anas undulata* |
| 2 | YB1568 *Anas undulata* |
| 3 | YB1631 *Anas undulata* |
| 4 | YB1581 *Anas undulata* |
| | YB2065 *Anas undulata* |
| 5 | H1 Putative Hybrid |
| | H3 Putative Hybrid |
| | H4 Putative Hybrid |
| | H5 Putative Hybrid |
| | H8 Putative Hybrid |
| | YB1569 *Anas undulata* |
| | YB1665 *Anas undulata* |
| | YB1869 *Anas undulata* |
| | YB2046 *Anas undulata* |
| | YB2071 *Anas undulata* |
| 6 | YB1958 *Anas undulata* |
| 7 | YB1907 *Anas undulata* |
| 8 | YB1951 *Anas undulata* |
| 9 | YB2048 *Anas undulata* |