Development and characterization of DNA markers for two avian species

By

Davida F. Kamara

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Animal and Poultry Sciences

Dr. E. J. Smith, Chair
Dr. A. Dickerman
Dr. E. Hallerman

May 22, 2006
Blacksburg, Virginia

Keywords: Budgerigars, Turkeys, Microsatellites, Single nucleotide polymorphism

Development and characterization of DNA markers for two avian species

Davida F. Kamara

ABSTRACT

Central to the application of genomics to animal agriculture are DNA markers, especially microsatellites and single nucleotide polymorphisms. These markers are the resources necessary for constructing genetic maps and for determining how improved and unimproved animal breeds are related. Here, DNA markers were developed for two avian species, the turkey, *Meleagris gallopavo* and the budgerigar (budgie), *Melopsittacus undulatus.* Genomic libraries enriched for simple sequence repeats were used to generate about 70 budgie sequences of a total length of 38 kb. From these sequences, 9 primer pairs were designed and used to screen for informativeness in a panel of DNA samples from unrelated budgie samples. All but one of the nine primers evaluated were polymorphic with the number of alleles ranging from two to four. Comparative analysis involving the use of these budgie primers showed moderate sequence similarity to turkey and chicken. The genomic libraries and the comparative sequences provide useful genomic reagents that could be used to construct a budgie genome map. In the turkey, ten previously described microsatellites and a gene-based single nucleotide polymorphism (SNP) were used to evaluate the relatedness of heritage varieties to a commercial strain. Estimates of Nei's genetic distance (*D*) and genetic differentiation (R_{st}) between populations using microsatellite markers showed that the commercial strain is genetically more closely related to the Bourbon Red and Narragansett and least related to the Royal palm and Spanish Black. Gene flow (Nm) level was highest between the commercial and Bourbon Red populations. The SNP analysis by PCR-RFLP revealed that the commercial strain was more closely related to the Spanish black and Narragansett and least related to the Bourbon red and Blue slate. Though results of the two marker systems, microsatellite and SNP, were inconsistent, they provide insights into using heritage turkeys to genetically improve commercial populations by introgression. The present thesis investigation showed that DNA markers provide a strong opportunity to develop genomic reagents needed to test hypotheses in little-studied agriculturally important and model avian species.

Keywords: DNA markers, Microsatellite, SNP, Genomic Libraries, Genetic Relatedness

ACKNOWLEDGEMENTS

With a great sense of gratitude, I wish to express my sincere thanks to my advisor, Dr. Ed Smith for giving me the opportunity to work on this project. My sincere thanks are also due to my committee members, Dr. Eric Hallerman and Dr. Allan Dickerman for their support and for serving on my committee.

I wish to thank Touyu "Leo" Geng, Choya Amenkhienan, Yan Geng and Stephanie Hartman for their help in data collection. I would also like to thank the rest of my colleagues, Xiaojing Guan, Kwaku Gyenai, and Kuan-Chin Lin, for their constant support. I am very fortunate to have known you, not only as lab mates, but as friends. Thank you all for your help and for creating an enjoyable work environment with all your silly jokes.

With great appreciation for their influence on whom I have become, I would like to thank my parents, David Kamara and Esther Kamara, who taught me the value of hard work by their own example. Thanks for the life lessons. I hope I've made you proud. Special thanks to my aunt, Pat Kamara, who always believed in me and has cheered me on all the way. Thank you for those care packages and the many cards with words of encouragement.

Lastly, I'm grateful to my friends for their constant support and words of encouragement. Thank you for being there always and for cheering me on through out my studies.

TABLE OF CONTENTS

LIST OF TABLES

Tables

LIST OF FIGURES

Figures

CHAPTER I

INTRODUCTION

DNA marker technologies have revolutionized genetic research in model species and animals of economic importance including livestock and poultry. The advances from these technologies have resulted in the development of genomic reagents essential for the construction of genetic maps and for the assignment of genotypes to phenotypes. Using DNA markers, useful genetic linkage maps have been constructed for many animals including the cattle (Barendse et al., 1997), pig (Rohrer et al., 1996), sheep (Maddox et al., 2001), and chicken (Groenen et al., 2000). These DNA marker-based genetic maps have provided the foundation for identifying genes that influence economically important traits or quantitative trait loci (QTL). The genome map of the bovine, for example, has been used to identify QTLs of economic importance such as milk yield, protein and fat content (Coppieters et al., 1998), and to identify the gene responsible for the renal tubular dysplasia disease (Ohba et al., 2000). Map-based QTL analysis has also been used to identify genes for fat deposition in the chicken, a trait of interest to the broiler industry because of consumer desire for healthy poultry meat (Ikeobi et al., 2002). Despite the value of these maps for genetic improvement, little progress has been made in developing them for other birds of agricultural and biomedical importance.

Except for the chicken, for which the draft genome sequence is now available (International Chicken Genome Sequencing Consortium, 2004), genome maps of many economically important and model avian species are either publicly unavailable or have very few markers. Like the chicken, the turkey (*Meleagris gallopavo)* is an economically important animal with significant impact on U.S. total farm receipt. However, efforts to develop genetic information for the turkey, and thus the construction of a useful linkage map, have lagged behind the chicken genome project.

Another example of an avian species for which genetic information including a linkage map could be useful for biomedical studies is the budgerigar, *Melopsittacus undulatus*. Birds, including the budgerigar, are good models for understanding the molecular processes involved in hearing, as they have the ability to regenerate hair-cells and restore their hearing after deafness. Studies have shown that the budgerigar regenerates hair-cells after experiencing hair-cell loss after the administration of kanamycin, an ototoxic drug (Dooling et al., 1997). Further, the

budgie has been used as a model to conduct extensive behavioral studies on the effect of auditory hair-cell regeneration on vocal production. However, due to the lack of genetic information for the budgie, to my knowledge, no studies have reported the identification of the genes involved in their ability to regenerate auditory hair-cells and regain hearing. Understanding the molecular processes for hair-cell regeneration in the budgie may aid in stopping or reversing the progression of hearing loss. This may also provide some insights into the processes by which hearing loss leads to the changes in the quality of vocal production.

The budgerigar has also been proposed as a model for understanding some aging phenomena. Currently, animal models best developed for aging research are rats and mice. However, these animals are poorly adapted for aging studies as their maximum longevity is 3-5 years. According to the free radical theory of aging, reactive oxygen species (ROS), the natural by-products of oxidative energy metabolism, are the cause of aging and age-related diseases such as cancer and Alzheimer's syndrome (Hasty et al., 2003; Harman, 1956). Normally, an increased rate of metabolism will lead to an increased rate of production of ROS, which causes damage to DNA and proteins leading to the development of diseases and decreased maximum life span. Despite their high metabolic rate, some birds are exceptionally long-lived so they have been proposed as models for human longevity (Holmes and Ottinger, 2003). Birds are prone to many of the same diseases of aging that afflict mammals, including infertility, cardiovascular disease, cancer, cataracts, osteoarthritis, and diabetes. The aging rates and progression of these diseases in birds, however, are generally much slower than in similar-sized mammals (Holmes and Ottinger, 2003). Among avian species, the budgerigar is considered an appropriate model for aging because of some unique characteristics. For example, though the rate of metabolism of the budgerigar is very high, their maximum lifespan is five times that of the mouse, a mammal of similar body size (Austad, 1997). This suggests that the budgerigar may have an innate mechanism to successfully manage reactive oxygen species, and that they may be useful in understanding the genetic basis of tolerance or susceptibility to oxidative stress.

A genome map for the budgie may provide insights into the genetic basis of auditory hair-cell regeneration that leads to the reversal of hearing loss and the recovery of vocal production. Hearing loss is a relatively common disorder in humans that has been attributed to both environmental and genetic factors. Approximately 1 in 800 children are born with a serious hearing impairment, and a significant proportion of the population suffers progressive hearing

loss as they age. Whereas single-gene defects probably account for over half of the cases of childhood deafness, the nature of the genetic contribution to progressive hearing loss has not yet been clearly defined (Steel and Kros, 2001). Study of model species offers strong opportunities to change this situation. Some progress has been made in identifying deafness genes in the mouse. However, relatively few of the identified mouse deafness genes have been shown to be involved in human deafness (Steel and Kros, 2001). The limited overlap in deafness genes identified in mice and humans indicates that there are still many deafness genes that have yet to be discovered for which other model species may be useful (Steel and Kros, 2001).

DNA markers can and have also been extensively used in population genetics and genetic relatedness studies (MacHugh et al., 1997; Yang et al., 2003). Their use in these types of research could enhance our knowledge of biomedically and agriculturally important avian species such as the budgerigar and the turkey. Though an animal of economic importance, genetic studies of the turkey continues to be limited. One such limitation is that the genetic relationship between commercial and heritage turkey varieties has never been established. Due to intensive selection for higher meat production, commercial turkey flocks are highly susceptible to diseases that result in significant economic losses. Understanding the genetic relationships between commercial and heritage turkey strains may be useful in breeding programs that could involve introgression of novel genes important for economic traits including disease resistance.

The present thesis project is resource-development driven so that genomic reagents developed could be used to subsequently test different hypotheses. More specifically, however, it will test the hypothesis that novel and previously described microsatellites are useful for the genetic characterization of budgies and turkeys, species that remain little characterized. The overall objective of this thesis research is to develop and characterize genomic reagents for the budgerigar through direct and comparative genome analyses and to use microsatellite-based DNA markers to evaluate the genetic relatedness between commercial and heritage domestic turkey varieties. Specific aims include:

1. To test the informativeness of microsatellites developed from diverse microsatellite-enriched budgie genomic DNA libraries,
2. To conduct comparative genome analysis of the budgerigar, turkey, and the widely-studied avian species, *Gallus gallus*,

3. To conduct molecular genetic analysis of relatedness among commercial and heritage domestic turkeys.

The rationale for this thesis project is that a budgie genome map will be useful in understanding the genetic and molecular basis of diverse vertebrate disorders such as age-related diseases and hearing loss. Also, information from the evaluation of the genetic relationships between commercial and heritage turkeys could be used in breeding programs for the introgression of novel genes important for economic traits such as disease resistance.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Some avian species are not only economically important animals, but they hold strong potential as models for biomedical research. However, efforts to develop genetic information necessary to increase our understanding of the biology of avian species other than the chicken have been limited. For example, the genetic relationship among commercial and non-commercial domestic turkey strains, especially at the molecular level, have never been assessed or reported. Genetic analysis has been conducted for the wild subspecies from which these domestic turkeys derived (Mock et a l., 2002), but similar study has not been conducted to evaluate the genetic relatedness between the domestic turkeys. Another avian species for which genetic material is necessary is the budgerigar. Due to their attractive appearance with colorful feathers and their ability to mimic human speech, the budgerigar, a native of Australia, is a very popular pet in Europe and the United States of America (USA). They are also recognized as models for studying biological abnormalities in vertebrates such as deafness and age-related diseases. However, the use of the budgie as a model for scientific research is limited due to the lack of information on its genome.

DNA markers such as microsatellites are excellent resources that could prove useful in increasing our understanding of the genetics of budgie and turkey and thus increase their value in agriculture and as models for biomedical research. In this review, I will describe the turkey and budgie, and discuss the use of DNA markers in the construction of genetic maps and in comparative and phylogenetic analyses.

2.2 Turkeys

The turkey, domesticated in Mexico between 200 B.C. and 700 A.D., is one of the few domestic animals native to the Americas (Crawford, 1992). There are two known species of turkeys: *Meleagris ocellata* of Central America and *M. gallopavo* of North America. The *Meleagris ocellata* species is also known as the Ocellated turkey, and is found in Belize,

Honduras, and the Yucatan. Ocellated turkeys are typically smaller than North American turkeys (Christman and Hawes, 1999). *M. gallopavo* is the ancestor of the domestic turkey and is divided into six subspecies that are distinguished by geographic range and plumage differences. They are the Eastern (*silvestris*), Rio grande (*intermedia*), Merriam's (*merriami*), Florida (*osceola*), Gould (*mexicana*), and Mexican (*gallopavo*) turkeys.

The Eastern wild turkey is native to eastern United States and is the most numerous and widely distributed. The Rio Grande wild turkey is native to the open lands of the central and southern plains and is the second most numerous. The Merriam's wild turkey is native to the ponderosa pine and juniper wood-lands and mountains of the southwestern United States from Colorado to Mexico. This subspecies almost became extinct earlier in the 20th century but is now the third most numerous. The Florida wild turkey is found on the peninsula of Florida and is similar to the Eastern turkey though smaller and darker. It is the fourth most numerous. The fifth subspecies, the Gould's turkey, is found in the southwestern United States and northern Mexico. The Mexican turkey was historically found in southern Mexico, with a range between Puerta Vallarta and Acapulco on the Pacific coast and Vera Cruz and Tuxpan on the Gulf of Mexico. This taxonomic classification (*M.g. gallopavo*) also includes all varieties of improved domesticated turkeys (Christman and Hawes, 1999; Mock et al., 2002) though with some genetic input from other subspecies.

Turkeys are large and impressive birds, standing as tall as four feet with a wing span of nearly six feet. The male turkeys are called toms, cocks, or gobblers. The females are called hens and young turkeys are called poults (Christman and Hawes, 1999). A distinguishing characteristic of this species is the courting behavior of the male turkey in which he struts and displays his wing and tail feathers (Crawford, 1984). The turkey head and neck are bare of feathers, but covered with wrinkled skin. The skin color is generally red but changes readily from red to dark blue to bluish white. Attached above the beak and hanging down beside it is a fleshy snood or frontal caruncle. Though found in both sexes, the snood is more prominent in the males. Male turkeys, and occasionally females, have a long hair-like appendage on the breast known as a beard which is composed of primitive contour feathers. These feathers are not molted and can attain a length of twelve inches in old males (Christman and Hawes, 1999).

Turkeys eat a variety of foods including seeds, nuts, fruits, insects, and plants and while they thrive in a wide range of environments and climatic conditions, they also have

characteristics that make them well suited for domestication. These characteristics include the fact that they are flock-oriented rather than solitary; they lack a strong pair-bond and can thus be kept with fewer males than females; they may tolerate close proximity to people in order to explore new sources of food; and they can be fed a diet consisting almost entirely of plant material plus insects they find on their own (Christman and Hawes, 1999).

The term breed, in most species of poultry and livestock, describes a group of animals that are distinctive in conformation, color, performance, behavior, and other qualities with each breed having a unique combination of these characteristics. A breed's consistency of appearance reflects an underlying genetic consistency, so when individuals of the same breed are mated together, they reproduce the same distinct type, a true breed. Varieties are sub-breed populations which show variations in color or type or comb and feather type as in the case of chickens, but similar in form and function to other varieties in the breed. The turkey species is classified differently: the American Poultry Association considers the domestic turkey to be a single breed with eight varieties: Black, Bronze, Narragansett (NA), Blue Slate (BS), White Holland, Bourbon Red (BR), Beltsvil le Small White, and Royal Palm (RP, Christman and Hawes, 1999).

The Black variety originated in Europe from Mexican turkeys which had been imported beginning in the 1500s. Turkeys of the Black variety with black legs are usually referred to as Norfolk Black and those with pink legs are referred to as Spanish Black (SB). The Norfolk Blacks, when crossed with wild birds, were a foundation for the Bronze, NA, and BS varieties, as well as being the basis for the Black variety in America (Christman and Hawes, 1999). The Bronze variety has been the most popular turkey variety for most of American history. This variety originated from crosses between the Mexican and Eastern wild turkeys (Crawford, 1992). The NA variety is derived from a cross between wild turkeys and the domestic Norfolk Black and was named for Narragansett Bay in Rhode Island, where it was developed. Its color pattern is made up of black, gray, tan, and white. The BS variety is named for its color, solid to ashy blue over the entire body with or without a few black flecks. It may have been derived from the cross between the Black and the White Holland varieties. The White Holland was the most important white-feathered variety in most of American history but is now one of the rarest and most difficult to authenticate. The name White Holland implies Dutch origin as it is believed that Dutch settlers or other European immigrants came to the New World with the white turkeys. The white feather color is due to a mutation (Christman and Hawes, 1999). This variety crossed

with the Broad-breasted Bronze produces the Broad-breasted White or Large White that dominates commercial production in the turkey industry today. The BR turkeys originated in Bourbon County in Kentucky's Bluegrass region. This variety is a cross between the Buff, Bronze, and White Holland turkeys. The Buff is another turkey variety which declined due to difficulty in breeding birds to fit the Buff color standard and was removed from the APA standard of perfection. The Beltsville Small White was developed by the U.S. Department of Agriculture in Beltsville, Maryland. This variety was developed in an attempt to produce small, white-feathered turkey, early maturing and easily reproduced for home and small-scale production. It was selected from a genetic foundation which included the White Holland, White Austrian (small white turkey from Scotland), NA, Bronze and wild turkey. The Royal Palm is derived from a mix of Black, Bronze, NA, and wild turkeys. This is a strikingly attractive and small-sized turkey variety with a mixture of white and metallic black. This variety has not been purposefully selected for either growth rate or muscling, but is used primarily as an ornamental variety (Christman and Hawes, 1999).

Using the crosses as a base, the turkey production industry has had tremendous economic success due to the increased demand for turkey meat. In 1998, the National Turkey Federation estimated that 270 million turkeys were produced. This growth of the industry has been based on an increase in the year-round market for turkey as it is used in school lunch programs, and processed as deli meat, sausages, hot dogs and other products for retail sale. Today, only about 20% of turkey sales are associated with the holiday season (Christman and Hawes, 1999). Meeting the increased demand is made possible by producing the commercial birds that have been highly selected for increased growth rate. There are two types of commercial birds, the Broad-breasted bronze (a cross between the U.S Bronze variety and the English Cambridgeshire Bronze) and the Large White (a cross between the Broad-breasted bronze and the White Holland). The Large White is not considered a variety by the APA because an application for its recognition has never been made. Today, the Large White accounts for well over 90% of the commercial market, and breeding stock is held primarily by one international company, Nicholas Turkeys. The greatest problem facing the turkey industry today is the narrow genetic foundation of industrial turkeys (Christman and Hawes, 1999). Intensive selection for higher meat production has resulted in a highly efficient and consistent production of meat, but at the same time, this strategy has also led to increase in health problems. Survival characteristics, such as

disease resistance have not been a priority, thus increasing the disease vulnerability in these highly selected turkeys. The uniformity of the commercial turkey population has made it possible for them to be similarly vulnerable to infectious diseases and genetic diseases (Cristman and Hawes, 1999).

2.3 Budgerigars

The free-living budgerigar, *Melopsitaccus undulatus*, is a small xerophilous psittacine species that inhabits arid and semiarid grasslands of inland Australia. This native Australian bird is a nomadic parakeet that usually weights 26 to 29 grams (del Hoyo et al., 1997). The budgerigars are known as budgies in short, and like many Australian birds and animals, its name was derived from that given to it by the Aboriginals who were living on the East coast at the time of the arrival of the Europeans. Throughout the years, the budgie has been given different names. These names include undulated parakeet, undulated grass parakeet, zebra parakeet, zebra grass parakeet, and "betcherrygar", an old aboriginal word from which the now universal name budgerigar is derived (Watmough, 1948; Samour, 2002). This Aboriginal name of the budgie is a reference to its value as a source of food (Scoble , 1981). The budgerigar was first described by the naturalist George Shaw in the classics Naturalists' Miscellany (1790-1813) and the Zoology of New Zealand (1794) (Watmough, 1948; Samour, 2002). In this edition, it was given the scientific name *Psittacus undulatus*, with Psittacus meaning parrot and undulates meaning diversified with waves (Scoble, 1981). John Gould, the famous naturalist later added the generic name Melo, meaning song, thus giving the budgerigar the name by which we know it today, *Melopsitaccus undulatus*. This name refers to the common green bird seen in the wild. Some subspecies, representing only slight variations are: *M. undulatus intermedius*, which inhabits the Northern Territory and has paler colouration on its back and neck; and *M. undulatus pallidiciceps*, which inhabits Western Australia and is also paler, particularly on the head (Scoble, 1981).

The budgies are governed in their movement by the availability of food and water and are widely distributed across Australia. They exist in large heterosexual flocks throughout the year. The cocks are recognized by a rather plump and pale blue cere, and the hens are recognized by a flatter and paler cere that is generally tinged white around the nostrils (Scoble, 1981). They feed on seeds mainly from *Astrebla pectinata, Boerhavia diffusa* and *Atriplex* species in the hot

months and *Iseilema* species in cooler periods. In the northern range, the breeding season is usually from June to September and from August to January in the southern range. However, because breeding is influenced by the rains and the abundance of food, it can occur at any time of the year, allowing in some cases two breeding seasons in the same year (del Hoyo et al., 1997). Budgies reach maturity at six months of age, however it is recommended that birds should be at least nine months old before they are paired for breeding. Hens left until they are twelve months old will give better results. They lay on alternate days with average clutches of four to six white eggs. The hen incubates the eggs. In the early stages of hatching, the cock feeds the hen at the entrance hole and she, in turn, feeds the young by the same regurgitation process (Scoble, 1981).

Due in large part to their ability for vocal mimicry, budgerigars have become the most commonly kept pet birds world wide and the most important birds for shows and exhibitions around the world (Binks, 1974) since they were first introduced in England in 1840 (Gould, 1865). Throughout 160 years of selective breeding, budgie breeders have produced more than 100 different color variations from the original green color and the size of show birds is considerably larger, 48 to 50 grams, than that of their native ancestors who weigh 26 to 29 grams (Binks, 1974).

The need for more information on the budgerigar is ever-increasing as it is now realized that these birds do not only make good pets, but they are good models for biomedical research such as hearing loss in humans. Deafness is a relatively common disorder in humans and can appear at any age and with any degree of severity. Profound congenital hearing loss occurs in an estimated 1 in 1000 births and large proportions of the population suffer progressive hearing loss as they age (Petit et al., 2001; ACMG, 2002) . Approximately 0.3% of the population manifest a hearing loss greater than 65 dB between the ages of 30 and 50 years and 2.3% between the ages of 60 and 70 (Petit et al., 2001). Approximately 50% of cases are thought to be due to environmental factors and the remainder to genetic causes (Morton, 1991; Gorlin et al., 1995; ACMG, 2002). Examples of environmental factors that contribute to hearing loss include acoustic trauma, ototoxic drug exposure, and bacterial or viral infections such as rubella or cytomegalovirus. Approximately 70% of congenital cases associated with genetic factors are classified as nonsyndromic, meaning the deafness is not associated with other clinical findings that define a recognized syndrome. The remaining 30% has one of more than 400 forms of

syndromic deafness that can be diagnosed because of associated clinical findings (Gorlin et al., 1995; Steel and Kros, 2001). Among the many forms of syndromic hearing loss, the auditory pathology varies widely and includes both conductive and sensorineural deficits that may be unilateral or bilateral, symmetrical or asymmetrical, and progressive or stable (Resendes et al., 2001). The auditory pathology of nonsyndromic hearing impairment can also vary, but the deficits are most often sensorineural (ACMG, 2002).

Over the past few years, remarkable progress has been made in identifying new hearing impairment loci and cloning new genes related to deafness. At least 77 loci for nonsyndromic hearing impairment have been mapped to date. These include 40 autosomal dominant, 30 autosomal recessive, and 7 X-linked (Van Camp and Smith, 2001). At least 31 genes for syndromic hearing loss have been identified (ACG, 2002). Although significant advances have been made in identifying genes that influence deafness, more genes and mutations await discovery. Different mutations at the same locus have been found to cause some syndromic and nonsyndromic forms of deafness. Also, the mice used as a model have not proved to be promising in identifying genes that influence deafness in humans. Relatively few of the identified mouse deafness genes have been shown to be involved in human deafness (Steel and Kros, 2001).

Though evolutionarily distant from humans, birds provide the only animal model for studying hearing restoration by renewed sensory-cell input and for examining the effect of such recovery on learned vocalization. Hair-cell regeneration in the inner ear of birds does not only provide a useful model for studying the molecular mechanisms by which hearing is lost and regained, but also provide the opportunity to study the effect of renewed auditory input on auditory perception, vocal production, and vocal learning in vertebrates. The loss of hearing in humans can have a profound effect on the quality of speech, so the ultimate value of the regenerative capacity of auditory hair cell in birds depends on whether it results in functional recovery of auditory and vocal behavior (Dooling et al., 1997). With this connection in mind, only 3 out of 28 orders of birds have considerable vocal learning abilities and would be good models for understanding the molecular processes that lead to hearing loss and also provide insights into the processes by which hearing loss leads to the changes in the quality of vocal production.

The Passeriformes (oscine songbirds), Psittaciformes (parrots), and Apodiformes (hummingbirds) are the only order of birds in which vocal learning occur (Kroodsma, 1982; Kroodsma and Baylis, 1982; Gaunt et al., 1994; Lavenex, 2000). Parrots, including the budgie, appear to be more appropriate models because they exhibit an array of vocal plasticity. They are open-ended learners; therefore, like humans, and in contrast to many songbirds, their vocal learning is not constrained by age or season (Farabaugh et al., 1994). They can mimic vocalizations of other species (Cruickshank et al., 1993; Lavenex, 2000) as well as the sounds of human speech (Pepperberg, 1981; Eda-Fujiwara and Okumura, 1992; Banta, 1998; Banta Lavenex, 1999). Budgies make good models for understanding the molecular processes that influence hearing loss and its effect on vocal production, due to the fact that they exhibit an array of vocal plasticity and have the ability to regenerate auditory hair cells. In these studies, kanamycin was administered to 32 birds for six days. After 6 days of kanamycin injection, virtually all of the hair cells were missing in the basal 40% of the papilla. Hair cells begin to regenerate in the basal 40% during the following 6 days. Hair cell number is almost normal within 4 weeks of kanamycin cessation and by 12 weeks hair-cell number was within normal limits (Dooling et al., 1997).

Budgerigars have also been proposed to be useful models for studies of aging. Recent progress in the science of aging is driven largely by the use of model systems ranging from yeast and nematodes to mice (Hasty et al., 2003). The vertebrate animal models currently best developed for studies of aging are laboratory rats and mice. However, these animals are poorly adapted for aging studies as they are short-lived and rapidly-aging species (Holmes and Ottinger, 2003; Smith et al., 2005). Research findings to date, suggest that birds will be better models for the study of longevity. Avian aging processes might be more similar to ours in some respects than those of the short lived rodents typically used to model basic mammalian aging processes (Holmes and Ottinger, 2003). They are prone to many of the same diseases of aging afflicting mammals, including infertility, cardiovascular disease, cancer, cataracts, osteoarthritis and diabetes.

Reactive oxygen species (ROS) or free radicals, the natural by-products of oxidative energy metabolism, are often considered to be the cause of aging (Hasty et al., 2003; Harman, 1956). They cause damage to macromolecules such as DNA and proteins which leads to aging and diseases such as cancer and Alzheimer's syndrome. Free radicals are normal by-products of

oxidative metabolism, and theoretically should be generated at higher rates by organisms with higher metabolic rates (Holmes and Ottinger, 2003). The long life spans and slow aging rates of birds, however, seem paradoxical to the free radical theory of aging. Compared to mammals of equivalent body size, they have very high metabolic rates (2-2.5 times higher), body temperatures (approximately 3˚c higher), and blood glucose levels (two to four-fold higher) (Holmes et al., 2001). Yet, many birds live up to three times longer than mammals of equivalent body mass. Studies by Barja and Herrero (1998) suggest that mitochondrial free radical production correlates better with the rate of aging and the maximum life-span than the metabolic rate. In their studies, they found that the H_2O_2 production in three bird species, pigeons, canaries, and parakeets is lower than those in rats and mouse inspite of the higher metabolic rate of the birds. Recent studies, like that of Barja and Herrero (1998), acknowledge birds to possess the ability to resist oxidative stress, but the genetic bases of this phenomenon are yet to be explored.

Among avian species with strong potential as an animal model for aging research is the budgerigar, *Melopsitaccus undulatus*. Attributes of the budgerigar that make it a useful aging model include exceptional longevity. They can live up to 20 years (Holmes and Ottinger, 2003), which is five times that of mammals with comparable body mass, including the mouse. They have high relative oxygen consumption, high glucose level and an above average body temperature. Their oxygen consumption is 50% higher than the similar-sized house mouse with an estimated average lifetime metabolic potential of about 2000 kcal/g, almost 10 fold more than mice and more than twice that of humans (Austad, 1998). High oxygen consumption and glucose level should lower the longevity of budgerigars, yet they can live 20 or more years. The longevity of the budgerigar despite their high rate of metabolism indicates that they have an innate ability to resist oxidative stress damage. Understanding the budgie's apparent resistance to oxidative stress is of great value in understanding the molecular processes of aging.

2.4 Genome maps

Genome maps, the ordered arrangement of markers within chromosomes of specific genomes, serve as guides to scientists in their hunt for genes that influence their traits of interest. There are two types of genome maps, genetic, also known as linkage maps, and physical maps (Graves, 1998). Genetic maps are derived from recombination frequencies between markers

during meiosis. This type of map depends entirely on the availability of polymorphisms, the property of having two or more alleles at a locus. Physical maps are constructed from information about the physical location of DNA sequences, including genes, on the chromosome. With physical mapping, genes may be assigned to physical positions within chromosomes or chromosome regions by somatic cell genetics, radiation hybrid mapping, and in situ hybridization (Graves, 1998). The construction and analysis of genome maps for many animals, including humans, is having a tremendous impact on improving human health (Soragna et al., 2003) and the genetics of animals for economically important traits (Coppeiter et al., 1997; Ikeobi et al., 2002).

2.5 Genetic markers

As with any type of map, genetic and physical maps consist of distinctive features known as genetic markers. Genetic markers are defined as specific locations on a chromosome which serve as landmarks for genome analysis. Genetic markers are of two types, morphological and molecular markers (Kumar, 1999).

Morphological markers are those whose inheritance can be monitored visually without specialized biochemical or molecular techniques. They are inherited in a Mendelian manner and their morphological traits that are controlled by a single locus can be used as genetic markers if their expression is reproducible over a range of environments. Epistatic and pleiotropic interactions can also alter the expression of these markers. With dominant markers, it is impossible to distinguish heterozygous individuals from homozygous individuals. The number of morphological markers is very limited, and their alleles often interact in a dominant-recessive manner (Kumar, 1999).

There are two types of molecular markers, biochemical and DNA markers (Kumar,1999). Biochemical markers are molecular markers that reveal polymorphisms generally at the protein level. They are proteins produced as a result of gene expression and can be separated by electrophoresis to identify the alleles. Isozymes, variant forms of the same enzyme, are the most commonly used biochemical markers. Protein markers reveal differences in the gene sequence and generally function as co-dominant markers, which allows for the genotype at any locus to be determined within any breeding scheme (Vodenicharova, 1989; Kumar, 1999). Molecular markers that reveal polymorphism at the DNA level are known as DNA markers (Kumar, 1999;

Vignal, 2002). DNA markers are the most frequently employed markers in the construction of genome maps and their variation can be detected in a number of ways. Basepair differences can be identified as restriction fragment length polymorphisms (RFLPs) if they happen to fall within a recognition site for a restriction enzyme. They may also be detected by hybridizing with allele-specific oligonucleotides, or allele-specific primers for PCR, which only bind if there is a perfect match (Graves, 1998). Sequencing of DNA products could also be used to reveal basepair differences.

2.6 DNA markers

DNA markers can be grouped into three types: Type I, Type II and Type III markers. The Type I markers are coding genes, Type II markers are noncoding DNA segments, and Type III markers are the bi-allelic single-nucleotide polymorphisms (SNPs O'Brien, 1999). DNA markers can also be classified as either clone/sequence-based (CSB) or fingerprint (Dodgson et al., 1997; Weber and May, 1989). The clone/sequence-based markers require the isolation of a cloned DNA fragment and often determination of some, if not all, of its DNA sequence. These markers include microsatellites also called simple sequence repeats (SSR, Weber and May, 1989), restriction fragment length polymorphisms (RFLP), sequence-tagged site (STS) and expressed sequence tags (EST). The fingerprint markers require no prior knowledge of the sequence of the polymorphic region or isolation of a cloned DNA fragment and they include the random amplification of polymorphic DNA (RAPD, Williams et al., 1990), minisatellites, also called variable number of tandem repeats (VNTR, Jeffreys et al., 1985), amplified fragment length polymorphisms (AFLP, Vos et al., 1995), and restriction landmark genome scanning (RLGS, Hirotsune et al., 1994). The fingerprint markers are inherently Type II markers, whereas the CSB markers include both Type I and Type II markers (Dodgson et al., 1997). If we consider molecular genetic DNA markers in terms of the type of information they provide at a single locus, only three main categories can be described: the bi-allelic dominant markers such as RAPDs and AFLPs; the bi-allelic co-dominant markers such as RFLPs and the more recent single nucleotide polymorphism (SNP) marker; and the multi-allelic co-dominant markers such as the minisatellites and microsatellites which are collectively called simple-sequence length polymorphisms (SSLPs, Griffiths et al., 1999; Vignal et al., 2002).

Restriction fragment length polymorphisms were the first generation of DNA markers. Before they were replaced by microsatellites, the first large scale effort to produce a human genetic map was performed mainly using RFLP markers (Strachan and Read, 1999; Brown, 2002; Vignal et al., 2002). Restriction enzymes cut DNA molecules at specific recognition sites. This sequence specificity means that treatment of a given DNA molecule with a restriction enzyme should always produce the same set of fragments. However, this is not always the same with genomic DNA molecules. Some restriction sites are polymorphic, existing as two alleles, one allele displaying the recognition sequence for the restriction site and therefore being cut when the DNA is treated with the enzyme, and the second allele having a different sequence so the restriction site is no longer recognized. This sequence polymorphism result in the two adjacent restriction fragments remaining linked together after treatment with the enzyme, leading to a length polymorphism which is the RFLP and its position on a genome map can be worked out by following the inheritance of its alleles, just as when genes are used as markers (Brown, 2002).

RFLPs were initially typed by preparing Southern blots from restriction digests of the test DNA, and hybridizing with radiolabeled probes. However, this technology was highly priced, required plenty of time and DNA and made a whole genome search a heroic undertaking. This is now becoming less of a problem as PCR is now frequently used to type RFLPs. With the PCR method, a sequence including the variable restriction site is amplified, the product is incubated with the appropriate restriction enzyme and then electrophored through a gel to see if it has been cut (Strachan and Read, 1999).

A disadvantage of RFLPs is that they can be labor-intensive in both the development stage and in the typing stage. A more fundamental limitation is their low informativeness. RFLPs are biallelic, so the site is either present or absent. Also, their heterozygosity can be low, a maximum of 0.5. If one allele is relatively uncommon in relation to the other, the proportion of heterozygotes, the crucial individuals useful in mapping will be low. For these reasons, the use of RFLP in animal genetic maps has been limited. Disease mapping using RFLPs is frustrating because more often, a key meiosis in a family will turn out to be uninformative (Dodgson, 1997; Strachan and Read, 1999).

RAPD has a great advantage in terms of their ease of use in the laboratory. RAPDs are produced by using standard oligonucleotides sometimes in combination with restriction enzymes

(Vignal, 2002). RAPDs derive from the fact that short oligonucleotide primers in PCR reactions with lowered annealing condition will generally amplify a spectrum of fragments from almost any template DNA. One or more of these fragments is often polymorphic and this polymorphism can be genetically mapped. Large panels of RAPD primers can be purchased at reasonable price from commercial suppliers thus reducing the cost for RAPD mapping (Levin et al., 1993; Dodgson et al., 1997). Also, they are fairly evenly distributed throughout the genome (Levin et al., 1994; Cheng et al., 1995) and RAPD-generated polymorphic bands can be readily cloned for further analysis (Dodgson et al., 1997). A major limitation of RAPDs is their dependence on the exact PCR conditions employed which can lead to reproducibility problems. This is probably due to the fact that the outcome of the amplification is extremely sensitive to the competition of inexact primer binding sites in the template for primers and polymerase in the critical early cycles (Dodgson et al., 1997). Another drawback of RAPD markers is that they are dominant which makes them less efficient in population genetic studies (Lewis and Snow, 1992; Lynch and Milligan, 1994; Robinson and Harris 1999) but which makes them especially useful for facile mapping of sex chromosomes (Levin et al., 1993; Dodgson et al., 1997).

Amplified fragment length polymorphisms are fragments of DNA that have been amplified using directed primers from restriction digested genomic DNA. Like the RAPD technique, but unlike the RFLPs and SSRs, prior sequence information is not needed (Matthes et al., 1998; Robinson and Harris, 1999). This significantly decreases the investments, time and cost, for AFLP mapping. Another major advantage of the AFLP technique is the large number of polymorphisms it generates. Considering its ability to differentiate individuals in a population, it is a useful technique for paternity analyses (Krauss, 1999), gene-flow experiments, and also for plant variety registration (Law et al., 1998). Other disadvantages of AFLP markers are that like RAPD markers, they are dominant, with polymorphisms detected as either band presence or absence (Robinson and Harris, 1999), and that they are technically difficult to generate.

The SNP marker is a the recently described DNA marker which involves a single base change of either a substitution or a deletion in a DNA sequence. For a base position with sequence alternatives in genomic DNA to be considered a SNP, it is considered that the least frequent allele should have a frequency of 1% or greater (Vignal et al., 2002). Most eukaryotic genome studied to date have been shown to contain vast numbers of SNPs. Some of these give

rise to RFLPs, but many do not because the sequence in which they lie is not recognized by any restriction enzyme. SNPs are usually bi-allelic markers, though in principle, any of the four possible nucleotide bases can be present at each position. As Type I markers (genes), SNPs are very useful in the construction of reference maps for species as they are necessary for both the development of comparative maps and the generation of positional candidate genes (Vignal et al., 2002).

The advantages of SNPs are their abundant numbers and the fact that they can be scored on solid-state arrays without recourse to gel electrophoresis (Wang et al., 1998; Strachan and Read, 1999). Gel electrophoresis has proved difficult to automate so any detection method that uses it will be relatively slow and labor-intensive. The detection of SNP is more rapid because it is based on oligonucleotide hybridization analysis. An oligonucleotide, a short single-stranded DNA molecule, is synthesized in the test tube and if the conditions are just right, then an oligonucleotide will hybridize with another DNA molecule only if the oligonucleotide forms a completely base-paired structure with the second molecule. Hybridization will not occur if there is a single position within the oligonucleotide that does not form a base pair.

Many screening strategies, including DNA chip and solution hybridization techniques, have been devised for the detection of SNPs (Mir and Southern, 2000). A DNA chip is a wafer of glass or silicon, 2.0 cm^2 or less in area, carrying many different oligonucleotides in a high-density array (Brown, 2002). With the DNA chip screening, a fluorescent marker is used to label the DNA to be tested. The labeled DNA is pipetted onto the surface of the chip and hybridization is detected by examining the chip with a fluorescence microscope, the positions at which the fluorescent signal is emitted indicating which oligonucleotides have hybridized with the test DNA. Many SNPs can therefore be scored in a single experiment (Brown, 2002; Wang et al., 1998; Gerhold et al., 1999). Solution hybridization techniques are carried out in the wells of a micro titer tray, each well containing a different oligonucleotide. A detection system is used to discriminate between unhybridized single-stranded DNA and the double-stranded product that results when an oligonucleotide hybridizes to the test DNA (Brown, 2002).

A disadvantage of SNPs include the fact that most of the methods used to detect them are expensive. Since they are bi-allelic markers, they suffer the same drawbacks as RFLPs with regard to human genetic mapping. There are two frequent errors associated with SNP analysis. The first is the non-detection of one of the two alleles, in which case a heterozygote individual is

genotyped as a homozygote, resulting in the SNP not displaying any variability in the family that is being studied. The second is the inverse, which is the false genotyping of a homozygote as a heterozygote (Brown, 2002). There are indeed some instances in which the lack of information due to the bi-allelic nature of SNPs is a limitation, however, there are cases in which they can provide valuable data on associations between specific genes or other DNA structures and phenotypes, or on population and genome dynamics (Vignal et al., 2002).

Although SNPs are now in the scene and have gained high popularity, simple-sequence length polymorphism markers remain to be key players in the construction of genetic maps and animal genetic studies. SSLPs are arrays of repeat sequences that display length variations, different alleles containing different numbers of repeat units (Brown, 2002). Unlike all other DNA markers, these markers are multi-allelic markers. There are two types of SSLP markers, the minisatellites and microsatellites. The minisatellites, also known as variable number of tandem repeats (VNTRs), form clusters up to 20kb in length with repeat units up to 25bp (Brown, 2002). Minisatellites have many alleles and high heterozygosity. Although they are very informative, the technical problems of Southern blotting and radioactive probes are still an obstacle to easy mapping and also they are not evenly spread across the genome (Strachan and Read, 1999).

Due to the fact that microsatellites are conveniently spaced throughout the genome, and their easy use by simple PCR, microsatellites are more popular than minisatellites (Brown, 2002; Strachan and Read, 1999). Microsatellites, also known as simple sequence repeats (SSR) or simple tandem repeats (STR), are tandem repeats of 1 to 6bp which are interspersed throughout the DNA of animal genome (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989; Dodgson et al., 1997). These markers are highly polymorphic, in addition to which their co-dominance and reproducibility make them ideal for genome mapping, as well as for population genetic studies (Dayanandan et al., 1998; Robinson and Harris, 1999). Variations in microsatellites result from differences in the number of repeat units and these differences are thought to be caused by errors in DNA replication (Jarne and Lagoda, 1996; Moxon and Willis, 1999). During DNA replication, the DNA polymerase slips when copying the repeat region, changing the number of repeats by inserting, or less frequently, deleting one or more of the repeat units (Jarne and Lagoda, 1996;). Larger changes in repeat number are thought to be the result of processes such as unequal crossing over (Strand et al., 1993). Microsatellites have not only been used to

construct useful genetic maps (Rohrer et al., 1996; Barendse et al., 1997; Groenen et al., 2000) but they've been employed in diversity studies (Kumar et al., 2006), and paternity analysis (Spence et al., 2006).

The co-dominant nature of microsatellite markers allows for hetrozygotes to be readily identified. This feature also increases the efficiency and accuracy of population genetic measures that are based on these markers compared with other markers, such as AFLPs and RAPDs (Robinson and Harris, 1999). Another advantage of microsatellite markers is their ease of use by simple PCR. Once microsatellite primers have been identified, screening of materials using the technique is fairly inexpensive. One disadvantage of microsatellite markers is that each microsatellite region must initially be cloned and the surrounding sequence determined. This requires plenty of effort and money. Another disadvantage is polymerase slippage. This can be a significant problem when analyzing mono and di nucleotide repeats During amplification, the thermopolymerase can slip, leading to the production of differently sized products that differ by approximately 1-5 repeat units form the expected product. Inaccurate allele identification may also be caused by the tendency of Taq polymerase to add an adenosine nucleotide to the 3' end of the amplified product (Robinson and Harris, 1999). Despite the problems associated with microsatellites, they are one of the most popular DNA markers used with applications in many different fields.

2.7 Comparative genomics

Comparative genomics, the cross-referencing of information on genome organization between species, includes comparisons of every type of gene map at every level from the cytological to the molecular (Graves, 1998). Through the use of genomic techniques, genome comparisons have revealed that the vertebrate genome is highly conserved, therefore, information about location and function of genes can be directly transferred across species. Such conservation should considerably speed up the search for genes that influence disease conditions in domestic mammals and humans as well as genes that specify economically important traits. The human genome is the point of all genetic comparisons as it is the best known and for which there is a well developed map. Advances of the Human Genome Project can be directly applied to increase our knowledge of every mammalian species and even other vertebrates such as birds and fishes of interest (Graves, 1998).

The utility of genome maps for comparative analysis relies on the markers used to construct them. Using DNA sequence comparison and comparative mapping, Type I markers (coding genes) are essential for identifying gene orthologs in distantly related species. However, Type I markers are not efficient in assessments of pedigree or population diversity due to their low rate of polymorphism. Type II markers, like the hypervariable microsatellites, are highly informative in pedigree and population assessment because they are ubiquitous and highly polymorphic. However, these Type II microsatellite markers are not as useful in recognizing orthologs between mammalian species of different orders. This is because particular type II markers and their distinctive flanking DNA sequence, which is required for polymerase chain reaction, arose after the divergence of existing phylogenetic orders. SNP markers are also valuable for pedigree and population assessments, but are uninformative when used for comparative ortholog identification between orders (O'Brien et al., 1999).

An area in which comparative genome analysis is particularly useful is in the construction of genetic maps for different species. The conservation of synteny relationships facilitates the relating of genetic information between taxa and molecular markers produced in one species may be used to accelerate the development of genetic linkage maps in species with less characterized genetic maps (Kumar, 1999). Cross-species amplification of microsatellite loci has been exploited in the construction of genetic maps of many livestock species including the cattle (Kappes et al. 1997), and sheep (deGortari et al., 1996). Because avian species contain fewer microsatellite markers compared to mammals (Primmer et al., 1997), microsatellite markers are more difficult and costly to develop for avian species (Reed et al. 2003). To reduce the considerable costs associated with microsatellite development, cross-species amplification has also been used to develop genetic maps for avian species. One such example is the study by Reed et al. (2000) to estimate the usefulness of primer pairs of chicken microsatellite loci to develop a genetic map of the turkey.

In their study, a total of 520 chicken specific microsatellite primer pairs were screened on DNA from a single commercial female turkey from Nicholas turkey Breeding farms. A total of 280 (54%) of the 520 chicken primer pairs screened produced amplified products. Of the 280 primer sets that produced positive results, 210 (75%) amplified a single product from turkey genomic DNA that was similar, if not identical in size to that expected based on the fragment size of the corresponding chicken locus. The remaining 70 (25%) primer pairs amplified a

product substantially different in size from that expected. These products of different size were sequenced to confirm their homology to the chicken loci. The sequence analysis revealed differences in repeat length and mutation between the species. To assess the usefulness of these chicken specific markers in mapping the turkey genome, allelic polymorphism was determined for 57 of the 280 amplified loci by examining 12 individuals from two distinct populations. Thirty-seven of the 57 loci examined were monomorphic with a single allele present in all the twelve individuals and 20 (35%) were polymorphic. If this value is representative of all the amplified loci (280 of 520, 54%), approximately 20% of chicken microsatellite loci (100 of 520) are expected to be polymorphic in the turkey and will be useful in constructing a genetic map of the turkey, an animal of economic importance.

2.8 Phylogenetics

Phylogenetics, the study of the relationship between groups of organisms as imposed by their evolutionary history, is one of the most exciting uses of comparative data (Graves, 1998; Byrne, 2002). Phylogeny uses genetic data to determine evolutionary relationships at all taxonomic levels such as between species and between populations within species (Byrne, 2002). Phlogenetic analysis, however, is no longer just for systematists, it has become an important tool throughout biology for comparing morphological, behavioral or molecular information about genes, individuals , populations and species. Historical relationships among genes or species are estimated through phylogenetic analysis and these relationships are depicted in a form of a branching diagram, a phylogenetic tree. Traditionally, phylogeny has been used in evolutionary biology. Now, there is an increasing need for phylogenetic analysis in any study in which biological variation is compared across samples. This need arises from the objects of study not being statistically independent, but, rather, connected through historical relationships (Hillis, 1997). Since phylogenetic studies focus on all taxonomic levels, Type II markers and SNPs are highly informative in assessing population diversity (O'Brien et al., 1999).

Phylogenetics can generally be divided into two types of approaches, the phenetic and cladistic approaches. The phenetic approaches are those that propose hypothetical evolutionary organization based on overall similarity. Minimum Evolution is the basic phenetic approach which attempts to minimize the overall observed change between the entities being considered, known as taxa. In practice, heuristic methods such as Neighbor-Joining are employed to infer

phenetic relationships. With phonetic approaches, all pair-wise comparisons are first computed between all taxa, organized into a similarity matrix, and then hierarchically clustered according to the matrix (Sarkar, 2006). In contrast, cladistic methods describe evolutionary history in terms of individual units of evolution and can be broken into two major classes, non-statistical and Statistical. Non-statistical methods include maximum parsimony and assume ambiguous models of evolution. The statistical approach includes the maximum likelihood and Bayesian, and uses the statistical models of evolution to infer phylogenetic relationships. Cladistic methods, compared to phenetic methods, are computationally intense. Using the cladistic, all possible tree topologies are first considered, then they are evaluated according to a specified criterion that satisfies a specific model of evolution to determine the best tree (Sarker, 2006).

Applications of phylogenetics include, reconstructing the ancestral gene sequences from which extant genes are derived, studying the origin and epidemiology of human diseases, inferring the evolution of ecological and behavioral traits through time, estimating historical biogeographic relationships, prioritizing the conservation of endangered populations or species, and reconstructing the historical relationships across all of life (Hillis, 1997). Phylogenetics has been used in various studies, including assessment of the genetic variation across the historical range of the wild turkey (Mock et al., 2002), analyzing the genetic relationships between 10 cattle breeds (Moazami-Goudarzi, 1997), and analyzing the genetic variation and relationships among eighteen Chinese indigenous pig breeds (Yang et al., 2003).

CHAPTER 3

Direct and Comparative Genome Analysis of an Avian Model for Aging, *Melopsittacus undulatus*

3.1 ABSTRACT

Though budgies are a valuable resource as model species for many important vertebrate characteristics, their use is limited by the paucity of genetic information. Despite high metabolic rate, the budgie is exceptionally long-lived and is thus considered an excellent model for vertebrate longevity. In this thesis work, genomic reagents were developed that will begin to make the budgie a more useful model species. The genomic reagents developed included microsatellites from genomic libraries enriched for simple sequence repeats. The utility of the budgie-specific microsatellite sequences was tested by genetic analysis of unrelated budgies using 9 primer-pairs, each specific for a microsatellite containing sequence. Eight out of the nine primers evaluated were polymorphic with the number of alleles ranging from two to four. Comparative analysis based on sequences of amplicons produced by budgie primers showed limited sequence similarity to both the turkey and chicken. The genomic libraries and microsatellites as well as the comparative sequences provide tools that can be used to begin to build a budgie genetic map and to use DNA markers to characterize budgie populations.

Keywords: DNA markers, Microsatellite, Comparative genomics

3.2 Introduction

Laboratory rats and mice are currently the animal models used for aging research. These animals, however, are poor models for studies of aging as they are short lived with a maximum longevity of only 3-5 years (Holmes and Ottinger, 2003; Smith et al., 2005). The budgerigar, *Melopsitaccus undulatus* has exceptional longevity, high relative oxygen consumption, high glucose level and an above average body temperature, which makes it a useful animal model for studying some concepts in aging (Austad, 1998). These characteristics would be expected to decrease the maximum life span of the budgie, yet they are able to live for more than 20 years. This suggests that the budgerigar may have an innate ability to successfully manage reactive oxygen species; budgies, therefore, could be useful in understanding the genetic basis of tolerance or susceptibility to oxidative stress. A genomic based approach may facilitate the identification of the genetic factors that influence this characteristic in the budgerigar. It could be assumed that genetic control for aging in the budgerigar, as in other animals, is polygenic. A map-based approach could be used to identify the quantitative trait loci (QTL) that influence longevity in the budgerigar (Risch, 2000). The use of the budgerigar in studying the genetic basis of longevity is, however, limited. There is very little known genetic information about the budgerigar, including DNA markers that are specific for its genome. To address this limitation and to begin to generate genetic information about the budgerigar and thus increase its value as an animal model for aging, a genetic map is necessary.

The objective of this research was to develop and characterize genomic reagents which will form the primary resource needed in the construction of a budgie genome map. Such a map will provide the tools that can be used to identify the gene or quantitative trait loci responsible for the budgies' apparent resistance to oxidative stress damage.

3.3 Materials and Methods

Microsatellite Analysis

Genomic libraries enriched for different microsatellites with Dynabeads. were developed in collaboration with Travis Glenn and Nancy Schable of Savannah River Ecology Laboratory, Aiken, SC. Details of the methods used to develop the enriched budgie genomic libraries, including enrichment in TG, TGG, and GATA sequences, have been described previously (Glenn and Schable, 2005). Briefly, budgie DNA fragments were ligated into PCR 2.1 vectors and cloned using the TA Cloning Kit (Invitrogen). Positive clones containing microsatellite repeats were sequenced using either M13 forward or reverse primers with Big Dye Terminator according to the manufacturer's recommended protocol (Applied Biosystems). The DNA sequences (Table 1) were analyzed using an ABI Prism 377 DNA sequencer (Applied Biosystems).

To test the usefulness (polymorphism) of the microsatellite-containing sequences, 21 sequences (Table 2) from the enriched budgie genomic libraries were randomly selected and used for primer design using the Internet-based computer program PRIMER 3 (Rozen and Skaletsky, 1997). Nine of these sequences were randomly selected and characterized. Genomic DNA was isolated from 50 µl aliquots of blood collected from 21 birds of three different geographical populations according to Smith et al. (1996). Ten of the samples represented a population from Idaho State, 8 of the samples represented a population from North Carolina State and 3 from Oregon State. The blood samples were collected in tubes containing 0.5 M EDTA and the blood mixtures were later transferred into 1.5 ml microcentrifuge tubes in aliquots of 50 µl. Genomic DNA isolated from each sample was air dried and dissolved in sterile water to a concentration of 50 ng/µl. Polymerase chain reaction (PCR) amplifications were carried out in volumes of 10 µl. Each reaction contained 25 ng of genomic DNA, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM Mg^{2+}, 200 µM dNTP, 5 pmol of each primer, and 1 U of Amplitaq DNA polymerase (Eppendorf). The forward primer was labeled with HEX fluorescent dye. Amplifications were carried out in a Mastercycler Gradient Thermocycler (Eppendorf) under the following conditions: an initial denaturation step of 5 min at 95˚C followed by 36 cycles of denaturation for 45 s at 95˚C, annealing for 45 s at temperature T_A (Table 2) and extension for 45

s at 72˚C with final extension for 7 min at 72˚C. Electrophoresis of the amplified products was performed on a 2% agarose gel with SYBR green at 70V for 30 min. Amplicons produced from PCR were analyzed on the ABI 377 sequencer and scanned using Genescan software (Applied Biosystems). For each locus, allele frequencies were computed by direct counting.

Comparative Analysis

To increase the number of genomic reagents developed for the budgie while saving time and cost, comparative genome analysis was conducted using experimental and *in silico* methods . In conducting the experimental comparative analysis, seven budgie specific primers (BGMsat1, BGMsat2A, BGMsat2B, BGMsat8, BGMsat14, BGMsat18, BGMsat20) were used for DNA amplification in the budgie, turkey and chicken. Amplified products from the turkey and chicken were sequenced and compared to the budgie as well as to each other by BLAST 2.

Additionally, 50 microsatellite-containing sequences of the budgie were compared to the recently released chicken genomic DNA sequence using the BLAT (BLAST-Like Alignment Tool) sequence alignment tool which is available at the UCSC Genome Bioinformatics website, (http://genome.ucsc.edu, Kent, 2002). Unlike BLAST, BLAT rapidly scans for relatively short matches (hits) and presents high percent matches as scoring pairs (HSPs). BLAT involving DNA is designed to quickly find sequences of 95% and greater similarity of length 40 bases or more. It may miss more divergent or shorter sequence alignments (Kent, 2002). Where BLAST (Basic Local Alignment Search Tool) builds an index of the query sequence and then scans linearly through the database, BLAT builds an index of the database and then scans linearly through the query sequence.

3.4 Results

A total of 76 microsatellite-containing sequences were isolated and submitted to GenBank and were assigned accession numbers (Table 1). The total length of the sequences is 38,091 bp. The sequence characteristics of the amplicons produced by the 9 primer pairs of sequences tested for usefulness, including locus name, repeat motif, sequence of the PCR primers, the optimal annealing temperature for PCR as well as the expected PCR product size, the alleles detected and the frequency at which they were observed are presented in Tables 2 and 3. Eight of the 9 primer pairs exhibited sequence length polymorphism. A total of 19 alleles were observed from these polymorphic microsatellite markers, which ranged from 2 to 4 with an average of 2.37 per microsatellite locus. The frequencies of the 19 alleles ranged from 0.045-0.958 (Table 3). Eight of the allele frequencies were higher than 0.5. Among the polymorphic markers, the highest and lowest allele frequencies were observed at the BGMsat12 locus.

A total of 7 and 5 budgie primer-pairs amplified DNA from chicken and turkey respectively (Table 4). Overall, the sequence similarity between the budgie reference sequences and the chicken and turkey comparative sequences were higher than expected as most comparative studies in avian species have reported only 10-20% success in cross-species amplification. The *in silico* comparative analysis was carried out using 50 budgie-specific microsatellite containing sequences (Table 5). The percent homology (hits) between most budgie microsatellite sequences and the chicken genome sequence is high, but the size of the exact match is very low (<60%) indicating high sequence divergence in flanking regions. For example, the homology of the sequence of accession number AY568102 with the chicken sequence is 94.0%, but the length of overlap between the two sequences is only 58bp long out of a possible 687bp (Table 5). Out of 50 sequences, only sequences of the accession numbers AY568101, AY568104, AY568113, AY568115, AY568118, AY568119, AY568145, AY568146, and AY568149 showed high overlap.

3.5 Discussion

For the first time, specific genomic reagents have been developed for the budgerigar. The 76 budgie specific genomic DNA sequences developed and submitted to GenBank provide resources that could be used to conduct diverse studies in the budgie including the development of a genome map. All but one of the nine characterized markers were found to be polymorphic, indicating a potential use in the construction of a budgie genome map.

The comparative genomics approach was used to develop additional markers for the budgie while saving time and cost. Comparative genomics has is central in the area of genomics and has been very valuable in the construction of recent genetic maps of many animal species. Cross-species amplification has been exploited in the construction of genetic maps for many livestock species including cattle (Kappes et al., 1997), sheep (de Gortari et al, 1997), and goats (Vaiman et al., 1996). This approach has also been used extensively in developing genomic reagents for less studied avian species such as the turkey (Levin et al., 1995; Reed et al., 199; Reed et al., 2000) and Japanese quail (Pang et al., 1999; Inoue-Murayama et al., 2001). For example, through comparative analysis with the chicken, Reed et al. (2000) found 20 polymorphic microsatellite markers that will be useful for mapping the turkey genome. The addition of these markers to the turkey map will serve to increase marker density which is important for mapping quantitative trait loci.

In this study, comparative analysis between the budgie, chicken and turkey showed promising results. One hundred percent of the budgie specific primers amplified DNA in the chicken and 71% of budgie specific primers amplified DNA in the turkey. The sequences of the chicken and turkey amplicons showed high sequence similarity with the budgie reference sequence. With this result, I infer that additional genomic reagents could be developed for the budgie using the chicken genome map and also the turkey. The chromosomal location of the chicken sequences with high similarity to the budgie sequences could be located and nearby sequences could be investigated to develop useful genomic reagents for the budgie. This approach may be limited with the turkey, since there isn't a comprehensive genetic map of the turkey available at the moment.

BLAT results of the chicken and budgie were not as promising. Only nine out of fifty budgie sequences yielded high sequence match. However, the chicken DNA sequence should

still be considered as a useful resource for developing markers for a budgie genetic map. The nine sequences with high overlap appear to be chicken homologues. Since we know their chromosomal location on the chicken genome, they could be easily mapped on the budgie. Also, markers closely located to these nine markers on the chicken genome could be further investigated to develop additional markers for the construction of a budgie genome map. This approach could show conserved linkage groups in the respective species.

A budgie genome map will be useful in identifying the molecular and genetic basis for the budgies' apparent resistant to oxidative stress. It is quite evident that the budgie has the ability to resist oxidative stress by successfully managing free radicals, however, the mechanism by which this is achieved is still unknown. It is not known if the budgie resists oxidative stress by producing very few radicals or by producing high amounts of free radicals but discard them before they accumulate and cause damage to macromolecules. Understanding this molecular process involved in the budgies' resistance to oxidative stress will aid in understanding the molecular changes that lead to aging in humans and may be useful in the development of drugs for age-related diseases like cancer, Alzheimer's and cardiovascular disease.

A budgie genome map will not only be useful for understanding the aging process in humans, but will also be useful in other biomedical research studies such as identifying genes that are involved in hearing-loss. Currently, the mouse is being used as the animal model for identifying genes that influence hearing loss. While considerable progress has been made in identifying deafness genes in the mouse, only a few of the identified mouse deafness genes are involved in human deafness. This limited overlap in the identification of deafness genes in the mice and humans indicate that more deafness genes await discovery. The budgie genome map may be useful in identifying additional candidate genes that influence hearing loss in humans. The budgerigar is a good model for understanding the molecular processes involved in deafness because like other birds, it has the ability to regenerate auditory hair-cell and regain hearing after being deafened by hair-cell damage. Also, it has the rare ability for vocal production that only occurs in three orders of birds.

Extensive behavioral studies on the effect of auditory hair-cell regeneration on vocal production have been conducted using budgerigars as models. However, no study has focused on identifying the genes that influence the budgie's ability to regenerate auditory hair-cells and regain hearing. This is due to the lack of genetic information for the budgie. A budgie genome

map will allow for the understanding of hair-cell regeneration in the budgie and may aid in stopping or reversing the progression of hearing loss in humans.

CHAPTER 4

DNA marker-based genetic analysis of relatedness among commercial and heritage turkeys (*Meleagris gallopavo)*

4.1 ABSTRACT

The turkey is second only to the chicken in importance as an agriculturally important poultry species. Unlike the chicken, however, genetic studies of the turkey continue to be limited. For example, to date, many genomic investigations have been conducted to characterize genetic relationships among commercial and non-commercial chicken breeds, while the nature of the genetic relatedness among commercial and heritage turkey strains remain unknown. The objective of this research was to conduct molecular genetic analysis of relatedness among commercial and heritage domestic turkeys including Narragansett, Bourbon Red, Blue Slate, Spanish Black, and Royal Palm. Primer-pairs specific for 10 previously described turkey microsatellite markers were used. Additionally, a naturally-occurring *ApaL I* recognition site matching a SNP in cardiac troponin T (*cTnT*) was also used to evaluate the relatedness among the six populations. Results of the microsatellite analysis showed that the Blue Slate, Bourbon Red and Narragansett were genetically more closely related to the commercial strain with Nei distance (*D*) of 0.30, and the Royal Palm and Spanish Black were the least related to the commercial strain with *D*=0.41 and *D*=0.40, respectively. R_{st} values showed higher genetic differentiation between the commercial and the Blue Slate, Spanish Black and Royal Palm and lower genetic differentiation between the commercial and Bourbon Red populations. Gene flow level was highest between the commercial and Bourbon Red populations. The PCR-RFLP analysis revealed that the commercial strain was more closely related to the Spanish Black (*D*=0.00) and Narragansett (*D*=0.02) and least related to the Blue Slate (*D*=0.20) and Bourbon Red (*D*=0.11). Though both marker systems were incongruent, they provide foundation for the basis of using heritage turkeys to genetically improve commercial populations by introgression.

Keywords: Heritage turkey, Microsatellite, PCR-RFLP, SNPs

4.2 Introduction

Heritage turkeys are becoming of interest to the US consumer as documented by Shriver (2003) and Parrish (2002). As more farmers grow them for the US market, the need to increase biological information about them increases. Though the turkey varieties are considered a single breed (American Poultry Association, 2001), evidence is emerging of significant strain differences among the populations (Hartman et al., submitted) and between the heritage turkeys and commercial birds (Gyenai, 2005). In his unpublished thesis work, Gyenai (2005) evaluated phenotypic differences in commercial and heritage turkey populations for their response to toxic levels of furazolidone in order to make inferences about the genetic bases of the incidence and severity of dilated cardiomyopathy (DCM) in turkeys. Variety differences in the turkey's response to diets containing furazolidone were observed, suggesting that an animal's response to furazolidone induced DCM is genetically based. More recently, among heritage varieties, Hartman et al. (submitted) observed significant strain differences for plasma uric acid concentration, a biomarker for many phenotypes in vertebrates, including oxidative stress.

Genetic variation within and among commercial turkey populations were also previously evaluated by Zhu et al. (1996), Ye et al. (1998) and Smith et al. (1996). Though Smith et al. (2005) also analyzed the genetic relatedness of 5 Heritage turkey varieties, the relatedness between the commercial and heritage turkeys have never been investigated. Information regarding the genetic relatedness between commercial and heritage turkey varieties can be used for genetic improvement of the different turkey strains such as the introgression of novel genes important for economic traits including disease resistance.

Commercial turkeys have been highly selected for higher meat production traits. In turn, they have high rate of susceptibility to many diseases because they are produced from a relatively narrow genetic background, a few highly selected strains of the Large White variety which may lack the genetic diversity they need to be able to resist or tolerate diverse disease conditions. Disease resistance has not been a priority in the turkey breeding programs and has therefore led to the increased vulnerability to disease conditions (Christman and Hawes, 1999). The high vulnerability of commercial birds to various disease conditions is of serious concern to the turkey industry. Since there is no comprehensive linkage map available for the turkey,

exploring the turkey genome and thus identifying quantitative trait loci (QTLs) for economic traits such as disease resistance in commercial birds is impossible. The need to evaluate the status of turkey genetic diversity is attracting attention from poultry scientists as it is now realized that the diversity of the non-commercial turkey varieties includes essential genetic resources that will enable breeders to improve their birds' health and vigor or to respond to changing environmental conditions, production systems, or consumer needs (Christman and Hawes, 1999).

The objective of this research was to conduct molecular genetic analysis of relatedness among commercial (CO) and heritage turkey varieties including Bourbon Red (BR), Blue Slate (BS), Narragansett (NA), Royal Palm (RP), and Spanish Black (SB). Understanding the genetic relationship between commercial and heritage turkeys may be useful in breeding programs that could involve the introgression of novel genes important for economic traits including disease resistance.

4.3 Materials and Methods

Samples

Twenty-five birds were used from each of six populations (CO), (NA), (BR), (BS), (SB), and (RP). The birds were obtained from Privett Hatcheries (Portales, NM) and raised at the Virginia Tech Turkey Farm using standard protocols. Blood was collected by brachial venipuncture in tubes containing 0.5M EDTA. Aliquots of 50µl blood were used for genomic DNA isolation according to Smith et al. (1996). Isolated genomic DNA from each sample was air dried and dissolved in sterile water to a concentration of 50ng/µl prior to use in PCR.

Microsatellite DNA analysis

The nucleotide sequences for the ten primer-pairs used in the present work are presented in Table 6. These sequences, some of which were developed in our lab, were previously described by Burt et al. (2003). Polymerase chain reaction (PCR) amplifications were carried out in a final volume of 10 µl. Each reaction contained 50 ng of genomic DNA, 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM Mg^{2+}, 200 µM of each dNTP, 160 ng of each primer, and 1 U of Amplitaq DNA polymerase (Eppendorf). Amplifications were carried out in a Mastercycler Gradient Thermocycler (Eppendorf) under the following conditions: an initial denaturation step of 5 min at 95˚C followed by 38 cycles of denaturation for 45s at 95˚C, annealing for 45s at optimized temperature (Table 6) and extension for 45s at 72˚C with final extension for 7min at 72˚C. Electrophoresis of the amplified products was performed on a 4% metaphor agarose gel with 1% ethidium bromide at 40V for 7hrs.

The microsatellite alleles and allele frequencies were analyzed by direct count within each population. Genetic distance *D* was estimated according to Nei (1972) and these estimates were used to construct a consensus tree with bootstrap values using the Neighbor-Joining method in PHYLIP (Felsenstein, 1989) and visualized using TREEVIEW (Page, 1996). Population differentiation was measured using the RST CALC software (Goodman, 1997). R_{st} is a differentiation measure that assumes a stepwise mutation model (Kimura and Otha,1978). Gene flow (Nm) was also calculated using the RST CALC software (Goodman, 1997).

To test the efficiency of metaphor agarose in genotyping analysis, PCR products of two of the Spanish black samples amplified with the *TUM20* primer and one Spanish black sample

amplified with the RHT0011 primer where randomly selected and genotyped using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and allele sizes were observed using Genescan 3.7.

Troponin T-based phylogenetic analysis

Based on the observation (Gyenai, 2005) that the Heritage turkeys as well as that of commercial turkey vary in their response to toxic levels of furazolidone, which causes DCM, we selected a candidate gene for idiopathic DCM in the human to investigate phylogenetic relationships among N, BR, BS, SB RP, and CO. The protocols used for PCR-RFLP analysis to genotype the turkeys were previously described by Lin (2006). Briefly, oligonucleotide primers specific for the turkey cardiac Troponin T (*cTnT*) described in that thesis were used to amplify the 989bp fragment. The PCR reaction was carried out in a final volume of 25 µl containing 50 ng genomic DNA, 200 µM each of deoxynucleoside triphosphates (dNTPs), 200 ng of each primer, 1.5 mM $MgCl_2$, and 50 units/ml Taq DNA Polymerase in a reaction buffer of pH8.5 (Promega). Amplifications were carried out in a Mastercycler Gradient Thermocycler (Eppendorf). The reaction mixtures were subjected to an initial denaturation at 95˚C for 5 min followed by 35 cycles of denaturation for 45s at 95˚C, annealing for 45s at 62ºC and extension for 45s at 72˚C with final extension for 7 min at 72˚C. Amplification of the 900 bp product of cTnT gene was evaluated by agarose gel electrophoresis in 2% agarose gel stained with SYBR Green. Each of the PCR product was subjected to restriction digestion by *ApaL I* (New England Biolabs) carried out in a final volume of 20 µl. The digestion was carried out at 37ºC for 4hrs and the reaction mixture consisted of 10 U of *ApaL I* and 0.2 µl 100X BSA and 2 µl 10 reaction buffer supplied by the manufacturer. Electrophoresis of products of the enzyme digest was performed on a 2% agarose gel containing SYBR green at 90V for 40 mins. The alleles and allele frequencies were analyzed by direct count within each population. Genetic distance was estimated according to Nei (1972) and these estimates were used to construct a consensus tree with bootstrap values using the Neighbor-Joining method and viewed as described above.

4.4 Results

The metaphor agarose gel was used for genotyping. Based on the results, it proved to be efficient in genotyping analysis. The different alleles were finely separated and easily distinguished from each other (Fig. 1 and 2). Alleles observed using the metaphor agarose gel were consistent with those observed on the ABI PRISM 3100 Genetic Analyzer (Applied Biosystems).

Two alleles were observed for *ADLOO23*, three for the *RHT0024*, *RHT0294*, *RHT0095*, *RHT0009*, *RHT0131*, *RHT0011*, and *TUM16* loci and four for *TUM20* and *RHT0216* (Table 7 and 8). The genetic distance estimates between the six populations are presented in Table 9. The lowest genetic distance was observed between the BS and the NA (0.07) and RP (0.07). The highest genetic distance was observed between the CO and the SB (0.40) and RP (0.41). The extended majority rule consensus tree with bootstrap values (Figure 3) depicts the genetic relationship between the six different populations, clustering the RP with the BS and the BR with the CO population. With the extended majority rule, any set of species that appears in more than 50% of the trees is included. The program then considers the other sets of species in order of the frequency with which they have appeared, adding to the consensus tree any which are compatible with it until the tree is fully resolved (Felsenstein, 1989). The clustering of the turkey populations was not robust as indicated by the low bootstrap values, suggesting that these turkey varieties are very closely related and therefore no group clustering was observed at a significantly high frequency. The R_{st} values (Table 10) ranged from 0.06 (between RP and NA) to 0.31 (between BS and BR). Gene flow levels (Table 11) ranged from 0.53 (between BR and BS) to 3.71 (between NA and RP).

ApaL I recognizes and digests at the 5'…G▼TGCAC…3' and 3'…CACGT▲G…5' site. Figure 4 is a gel representation of observed genotypes in the Troponin *T* gene digestion with *ApaL I* restriction enzyme. Three fragments of 389, 600, and 989bp (Fig. 4) were observed. The two smaller fragments represent the G allele indicating there is a recognition site within that sequence and digestion has occurred. The larger fragment represents the A allele which is the sequence lacking a recognition site and is therefore not digested (cut) by the enzyme. Allele frequencies observed within each population are shown in Table 12. Pairwise genetic distance estimates of *cTnT* allele frequencies are recorded in Table 13. These genetic distance estimates

showed that observed SNPs in the CO population were very similar to that of the SB (0.00) and NA (0.02) populations and least similar to that observed in the BS (0.20) and BR (0.11) populations. The consensus tree (Figure 5) shows this relationship, clustering the BR with the BS and RP and the CO with the NA and SB populations.

4.5 Discussion

Phylogenetic trees showing genetic relationships among taxa are accepted by many biologists in diverse areas of research, as they can provide meaningful insights into taxonomic biology (Soltis and Soltis, 2003). Examples of their use in diverse areas of research include evaluating population history and genetic variation across historical ranges of animals for inferring implications for animal conservation (Mock et al., 2002; Eizirik et al., 2001), and analyzing the genetic relationship between viruses for inferring medical implications (Lindern et al., 2006). Here, the genetic relatedness between commercial and heritage turkeys, were investigated using hypervariable Type II markers and a bi-allelic marker. These markers are valuable for assessing pedigree, family, and population relationships due to their abundance in the genome and their ability to reveal polymorphisms (O'Brien et al., 1999).

As indicated by the genetic distance, R_{st} , and gene flow values and topology of the genetic tree, the microsatellite or Type II DNA marker analysis showed that the BR was genetically more closely related to the CO population while the SB and RP populations were least related. This relationship is consistent with reports (Christman and Hawes, 1999) that the [BR and CO] turkey varieties have a common ancestor, White Holland. The BR is believed to be derived from a cross between the White Holland, Buff, and Bronze varieties while the CO is believed to be derived from the White Holland and Bronze varieties. Though the BS also share the White Holland with the BR and CO, R_{st} and gene flow values and topology of the genetic tree showed that the BS is genetically more closely related to the RP and NA, with whom it also share a common ancestor, the Black variety. The BS is reported to be from the White Holland and the Black varieties. The NA is a cross between the Norfolk Black and wild turkeys. The RP is a mix between the Black, Bronze, Narragansett, and wild turkeys.

The inferred genetic relatedness among the heritage varieties appear to be consistent with that observed by Smith et al. (2005), but with slight difference. When a tree was constructed using only the heritage varieties, the SB clustered with the BR as observed by Smith et al. (2005) However, the tree showed closer relatedness between RP and BS unlike that observed by Smith et al. (2005) which showed closer relatedness between RP and NA. This observation was consistent with the low level of differentiation and high gene flow between the NA and RP. Considering this result and that the NA had the second less genetic distance from the RP, my

results and that of Smith et al.(2005) are actually showing the same relationship that the BS, NA, and RP heritage varieties are genetically more closely related.

The SNP analysis results were not consistent with that of the microsatellite analysis. It showed that the CO population was closely related to the SB and least related to the BS and BR. This difference in results may be due to selection for specific traits such as disease resistance. For example, Gyenai (2005) showed phenotypic differences among the six turkey strains in their response to toxic levels of furazolidone, which induces dilated cardiomyopathy. Dilated cardiomyopathy is a prevalent disease in commercial turkeys that is characterized by enlargement of the heart, thinning of the left ventricular free wall, and impairment of systolic function (Gwathmey et al., 1999; Roberson, 2005). In turkeys, it causes growth retardation, poor blood circulation and death due to heart deformation (Roberson, 2005), costing the industry thousands of dollars each year. This disease also occurs in humans. My genotype-based result of the *cTnT* gene analysis seems to correlate with the phenotypic result of Gyenai (2005). The thesis research reported that the CO, SB and NA were the most susceptible to DCM and the BR and BS were the least susceptible.

Though the results of the SNP analysis was not consistent with the microsatellite analysis, they are both valuable in the management of turkey populations. The current studies provide, provide additional evidence of the distinctiveness of heritage turkeys and for the first time, a foundation on which commercial turkeys could be improved using heritage turkeys. For example, based on these results, I would recommend that the BR be used for improving commercial flocks for DCM. The BR is more resistant to DCM, while the CO is less resistant (Gyenai, 2005). Microsatellite analysis showed that the BR is more closely related to the CO population but PCR-RFLP analysis showed that the Troponin *T* gene, a candidate gene for DCM, of these two populations were least related. The high level of gene flow between the BR and CO populations also makes the BR a good population source for introgressing new genes, like genes that influence DCM resistance in turkeys, into CO population.

CHAPTER 5

SUMMARY OF THESIS

This thesis research is resource development-driven, so that genomic reagents developed could be used to subsequently test different hypotheses related to turkeys, budgies, and other vertebrates. The specific aims of this research were to develop and characterize genomic reagents for the budgie, *Melopsit acus undulatus,* through direct and comparative genome analysis and to evaluate the genetic relatedness between commercial and non-commercial domestic turkeys, *Meleagris gallopavo*. The rationale for this research was that a budgie genome map will be useful in understanding the genetic and molecular basis of diverse vertebrate disorders such as age-related diseases and hearing loss. An additional justification is that understanding the genetic relationship between commercial and non-commercial turkeys would aid in the genetic improvement of turkey flocks including the introgression of genes that influence disease resistance.

Specific findings are:

1. Seventy-six budgie-specific microsatellite-containing sequences totalling 38,091 bp in length were isolated,
2. Twenty-one primer pairs were designed and nine loci characterized. Eight out of the nine markers were polymorphic, indicating a potential for use of all the 76 markers for the construction of a budgie genome map,
3. Comparative analysis of results for the budgie with those for the chicken and turkey suggests that reagents for the chicken and turkey genomes may be limited in their use for development of genomic reagents for the budgie,
4. Genetic distance and differentiation estimates using microsatellite markers showed that BR and NA are genetically more closely related to the commercial strain and SB and RP are least related. Analysis based on SNP in the *cTn T* gene showed that the CO was more closely related to the NA and SB and least related to the BR and BS. This difference in results may be due to selection for specific traits including DCM, a phenotype for which the heritage turkeys were previously shown to differ.

Future work:

More primer pairs will be designed using the remaining six budgie sequences and their usefulness in constructing a budgie genome map evaluated as done here. The chicken and turkey sequences with high similarity with the budgie sequence will be further investigated for syntenic mapping and for development of additional markers for the budgie. In turkeys, crosses will be generated between the CO and heritage turkeys to further characterize the occurrence of DCM in turkeys and to investigate the possibility of marker assisted introgression of genes that influence resistance to DCM into commercial populations.

LITERATURE CITED

ACMG. 2002. Genetics Evaluation Guidelines for the Etiologic Diagnosis of Congenital Hearing Loss. Genetic Evaluation of Congenital Hearing Loss Expert Panel. ACMG statement. Genet Med. 4(3):162-71.

American Poultry Association. 2001. The American Standard of Perfection. Page 492. American Poultry Association, Inc. Troy, New York.

Austad, S. N. 1997. Birds as Models of Aging in Biomedical Research. ILAR J. 38(3):137-141.

Austad, S. N. 1998. Theories of aging: an overview. Aging (Milano) 10(2):146-7.

Banta, P. A. 1998. Neuroethology of Acquired English and Conspecific Vocalizations in the Budgerigar (Melopsittacus undulatus). Ph.D. Dissertation, University of Arizona, Tucson, Arizona.

Banta Lavenex P. 1999. Vocal production mechanisms in the budgerigar (Melopsittacus undulatus): the presence and implications of amplitude modulation. J Acoust Soc Am 106: 491-505.

Barendse, W.; Vaiman, D.; Kemp, S.J.; et al. 1997. A medium-density genetic linkage map of the bovine genome. Mamm. Genome. 8(1):21-28.

Barja, G; Herrero, A. 1998. Localization in a region of complex I and mechanism of the higher free radical production of brain nonsynaptic mitochondria in the short-lived rat than in the longevous pigeon. J. Bioenerg. Biomembr. 30: 235-243.

Binks, G. 1974. Best in Show, Breeding and Exhibiting Budgerigars. London, UK: Ebury Press and Pelham Books.

Brown, T. A. 2002. Mapping Genomes. In: 2nd ed. Genomes. New York: John Wiley & Sons.

Burt, D. W.; Morrice, D. R.; Sewalem, A.; et al. 2003. Preliminary linkage map of the turkey (Meleagris gallopavo) based on microsatellite markers. Anim Genet. 34(6):399-409.

Byrne, M. 2003. Phylogenetics and the conservation of a diverse and ancient flora. C R Biol. 326 Suppl 1:S73-9.

Cheng, H. H.; Levin, I.; Vallejo, R. L. 1995. Development of a genetic map of the chicken with markers of high utility. Poult Sci. 74(11):1855-74.

Crawford, R. D. 1990. Origin and history of poultry species. In: C. RD, ed. Poultry Breeding and Genetics. Amsterdam: Elsevier Science Publishers.

Crawford, R. D. 1992. Introduction to Europe and diffusion of domesticated turkeys from the America. Arch. Zootec. 41 (extra): 307-314.

Christman, C. J.; Hawes, R. O. 1999. Birds of a Feather: Saving Rare Turkeys from extinction. Pittsboro: The American Livestock Breeds Conservancy.

Coppieters, W.; Riquet, J.; Arranz, J. J.; et al. 1998. A QTL with major effect on milk yield and composition maps to bovine chromosome 14. Mamm. Genome. 9(7):540-544.

Cruickshank, A. J.; Gautier, J. P.; Chappuis, C. 1993. Vocal mimicry in wild African Grey parrotsPsittacus erithacus. Ibis 135: 293-299.

Dayanandan, S.; Bawa, K. S.; Kesseli, R. 1997. Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot. 84:1658-1663.

de Gortari, M. J.; Freking, B. A.; Kappes, S. M; et al. 1997. Extensive genomic conservation of cattle microsatellite heterozygosity in sheep. Anim Genet. 28(4):274-90.

del Hoyo, J.; Elliot, A.; Sargatal, J. 1997. Handbook of the Birds of the World. Vol 4. Barcelona, Spain: Lynx Edicions.

Dodgson, J. B.; Cheng, H. H.; Okimoto, R. 1997. DNA marker technology: a revolution in animal genetics. Poult Sci. 76(8):1108-14.

Dooling, R. J.; Ryals, B. M.; Manabe, K. 1997. Recovery of hearing and vocal behavior after hair-cell regeneration. Proc Natl Acad Sci U S A. 94(25):14206-10.

Eda-Fujiwara, H.; Okumura, H. 1992. The temporal pattern of vocalizations in the budgerigar. J Yamashina Inst Ornithol 24:19-31.

Eizirik, E.; Kim, J. H.; Menotti-Raymond, M.; et al. 2001. Phylogeography, population history and conservation genetics of jaguars (Panthera onca, Mammalia, Felidae). Mol Ecol. 10(1):65-79.

Farabaugh, S. M.; Linzenbold, A.; Dooling, R. J. 1994. Vocal plasticity in budgerigars (Melopsittacus undulatus): evidence for social factors in the learning of contact calls. J Comp Psychol 108: 81-92.

Felsenstein, J. 1989. PHYLIP-Phylogeny inference package (version 3.6b). Cladistics 5:164-166.

Gaunt, S. L.; Baptista, L. F.; Sanchez, J. E. 1994. Song learning as evidenced from sharing in two hummingbird species (Colibri coruscans and C. thalassinus). Auk 111: 87-103.

Gerhold, D.; Rushmore, T.; Caskey, C. T. 1999. DNA chips: promising toys have become powerful tools. Trends Biochem. Sci., 24, 168-173.

Glenn, T. C.; Schable, N. A. 2005. Isolating microsatellite DNA loci. Methods Enzymol. 395:202-222.

Goodman, S. J. 1997. "A collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance" Mol Ecol 6: 881-885.

Gorlin, R. J.; Toriello, H. V.; Cohen, M. M. 1995. Hereditary hearing loss and its syndromes. New York: Oxford University Press.

Gould J. 1865. Handbook to the Birds of Australia. Vol 2. London, England.

Graves, J. A. 1998. Background and Overview of Comparative Genomics. ILAR J. 39(2-3):48-65.

Griffiths, A. J.; Gelbart, W. M.; Miller, J. H.; et al. 1999. Structural Genomics. In: Modern Genetic Analysis. New York: W. H. Freeman & Co.

Groenen, M. A.; Cheng, H. H.; Bumstead, N.; et al. 2000. A consensus linkage map of the chicken genome. Genome Res. 10:137-147.

Gwathmey, J. K.; Kim, C. S.; Hajjar, R. J.; et al. 1999. Cellular and molecular remodeling in a heart failure model treated with the beta-blocker carteolol. Am J Physiol. 276(5 Pt 2):H1678-90.

Gyenai, K. B. 2005. Genetic analysis of toxin-induced dilated cardiomyopathy in the turkey (Meleagris gallopavo). MS Thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Harman, D. 1956. Aging: a theory based on free radical and radition chemistry. J Gerontol 11 298-300.

Hartman, S. 2005. Comparison of Plasma Uric Acid Levels in Five Strains of the Domestic turkey, *Meleagris gallopavo.* Submitted for publication.

Hasty, P.; Campisi , J.; Hoeijmakers, J.; et al. 2003. Aging and genome maintenance: lessons from the mouse? Sci. 299 (5611): 1355-9.

Hillis, D. 1997. Phylogenetic analysis. 7 (3) Curr Biol.

Hirotsune, S.; Hirose, K.; Kataoka, H.; et al. 1994. Spot mapping on the standard profile of restriction landmark genomic scanning (RLGS) of sorted chromosome 20 using methylation-insensitive enzyme. Genomics. 24(3):593-6.

Holmes, D .J.; Ottinger, M. A. 2003. Birds as long-lived animal models for the study of aging. Exp Gerontol. 38(11-12):1365-75.

Holmes, D. J.; Fluckiger, R.; Austad, S. N. 2001. Comparative biology of aging in birds: an update. Exp Gerontol. 36(4-6):869-83.

Ikeobi, C. O.; Woolliams, J. A.; Morrice, D. R.; Law, A.; Windsor, D.; Burt, D. W; Hocking, P. M. 2002. Quantitative trait loci affecting fatness in the chicken. Anim. Genet. 33(6):428-35.

Inoue-Murayama, M.; Kayang, B. B.; Kimura, K.; et al. 2001. Chicken microsatellite primers are not efficient markers for Japanese quail. Anim Genet. 32(1):7-11.

International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 432(7018):695-716.

Jarne, P.; Lagoda, P. J. L. 1996. Microsatellites, from molecules to populations and back. Trends Ecol Evol 11: 424-429

Jeffreys, A. J.; Wilson, V.; Thein, S. L. 1985. Individual-specific 'fingerprints' of human DNA. Nature. 316(6023):76-9.

Kappes, S. M.; Keele, J. W; Stone, R. T; et al. 1997. A second-generation linkage map of the bovine genome. Genome Res. 7(3):235-49.

Kent, W.J. 2002. BLAT—The BLAST-Like Alignment Tool. Genome Res. 12(4):656-64.

Kimura, M.; Ohta, T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc. Natl. Acad. Sci. USA 75:2868–2872.

Kroodsma, D. E. 1982. Learning and the ontogeny of sound signals in birds. In: Kroodsma DE , Miller EH , editors. Acoustic communication in birds, vol 2. New York: Academic Press, 1-23.

Kroodsma, D. E.; Baylis, J. R. 1982. A world survey of evidence for vocal learning. In: Kroodsma DE , Miller EH , editors. Acoustic communication in birds, Vol. 2. New York: Academic Press, 311-337.

Kumar, L. S. 1999. DNA markers in plant improvement: an overview. Biotechnol Adv. 17(2-3):143-82.

Kumar, S.; Gupta, J.; Kumar, N.; et al. 2006. Genetic variation and relationships among eight Indian riverine buffalo breeds. Mol Ecol.15(3):593 -600.

Lavenex, P. B. 2000. Lesions in the budgerigar vocal control nucleus NLc affect production, but not memory, of english words and natural vocalizations. J Comp Neurol. 421(4):437-60.

Law, J. R; Donini, P.; Koebner, R. M.; et al. 1998. DNA profiling and plant variety registration III: The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. Euphytica 102: 335-342.

Levin, I.; Crittenden, L. B.; Dodgson, J. B. 1993. Genetic map of the chicken Z chromosome using random amplified polymorphic DNA (RAPD) markers. Genomics. 16(1):224-30.

Levin, I.; Crittenden, L. B.; Dodgson, J. B. 1994. Mapping DNA polymorphisms using PCR primers derived from the sequence of an avian CR1 element. J Hered. 85(2):73-8.

Levin, I.; Cheng, H. H.; Baxter-Jones, C.; et al. 1995 Turkey microsatellite DNA loci amplified by chicken-specific primers. Anim Genet. 26:107-10.

Lewis, P. O.; Snow, A. A. 1992. Deterministic paternity exclusion using RAPD markers. Mol Ecol 1: 155-160.

Lin, K. 2006. Candidate Gene Expression and SNP Analysis of Toxin-Induced Dilated Cardiomyopathy in the Turkey (*Meleagris gallopavo*). MS Thesis, Virginia Plytechnic Institute and State University, Blacksburg, VA.

Litt, M.; Luty, J. A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet. 44(3):397-401.

Lynch, M.; Milligan, B. G. 1994. Analysis of population genetic structure with RAPD markers. Mol Ecol 3: 91-99.

MacHugh, D. E.; Shriver, M. D.; Loftus, R. T.; et al. 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (Bos taurus and Bos indicus). Genetics. 146(3):1071-86.

Maddox, J. F.; Davies, K. F.; Crawford, A. M.; et al. 2001. An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Res. 11:1275-1289.

Matthes, M. C.; Daly, A.; Edwards, K. J. 1998. Amplified fragment length polymorphism (AFLP). In: Karp A, Isaac PG, Ingram DS (eds.). Molecular Tools for Screening Biodiversity. Chapman and Hall, London, 183-190.

Mir, K. U.; Southern, E. M. 2000. Sequence variation in genes and genomic DNA: methods for large-scale analysis. Annu Rev Genomics Hum Genet. 1:329-60.

Moazami-Goudarzi, K.; Laloe, D.; Furet, J. P.; et al. 1997. Analysis of genetic relationships between 10 cattle breeds with 17 microsatellites. Anim Genet. 28(5):338-45.

Mock, K. E.; Theimer, T. C.; Rhodes, O. E . Jr.; et al. 2002. Genetic variation across the historical range of the wild turkey (Meleagris gallopavo). Mol Ecol. 11 (4):643-57.

Morton, N. E. Genetic epidemiology of hearing impairment. 1991. Ann N Y Acad Sci. 630: 16–31.

Moxon, E. R; Wills, C. 1999. DNA microsatellites: agents of evolution? Sci Am. 280(1): 94-9.

Nei, M. 1972. Genetic distance between populations. Am Nat. 106:283-292.

O'Brien, S. J.; Menotti-Raymond, M.; Murphy, W. J.; et al. 1999. The Promise of Comparative Genomics in Mammals. Sci. 286 (5439): 458 – 481.

Ohba, Y.; Kitagawa, H; Kitoh, K; et al. 2000. Homozygosity mapping of the locus responsible for renal tubular dysplasia of cattle on bovine chromosome 1. Mamm. Genome. 11(4):316-9.

Page, D. M. R. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci. 12:357-358.

Pang, S.W.Y.; Ritland, C.; Carlson, J. E. et al. 1999. Japanese quail microsatellite loci amplified with chicken-specific primers. Anim Genet. 30:195-9.

Parrish, M. 2002. Heritage turkey breeds may be making a comeback http://www.post-gazette.com/food/20021121turkey1121fnp2.asp. Accessed 2/14/2006.

Pepperberg, I. M. 1981. Functional vocalizations by an African Grey parrot (Psittacus erithacus). Z Tierpsychol 55: 139-160.

Petit, C.; Levilliers, J.; Hardelin, J. 2001. Molecular genetics of hearing loss. Annual Review of Genetics. 35:589-645.

Primmer, C. R.; Moller, A. P.; Ellegren, H. A. 1996. Wide-range survey of cross-species microsatellite amplification in birds. Mol Ecol. 5(3):365-78.

Reed, K. M.; Mendoza, K. M.; Beattie, C. W. 2000. Comparative analysis of microsatellite loci in chicken and turkey. Genome. 43(5):796-802.

Resendes, B. L.; Williamson, R. E.; Morton, C .C. 2001. At the speed of sound: gene discovery in the auditory nervous system. Am J Hum Genet 69: 923–935.

Risch, N. J. 2000. Searching for genetic determinants in the new millennium. Nature. 405(6788):847-56.

Rohrer, G. A.; Alexander, L. J.; Hu, T. P. L.; et al. 1996. A comprehensive map of the porcine genome. Genome Res. 6:371-391.

Reed, K. M.; Mendoza, K. M.; Beattie, C. W. 1999. Utility of chicken-specific microsatellite primers for mapping the turkey genome. Anim Biotechnol. 10(3):137-41.

Roberson, K. D. 2005. Induction of Round Heart Disease in Turkey Poults by Feed Texture. Int J Poult Sci. 4 (4): 177-181

Robinson, J. P.; Harris. 1999. Amplified fragment length polymorphisms and microsatellites: a phylogenetic perspective. In: Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, Optimization, and Validation of Molecular
Tools for Assessment of Biodiversity in Forest Trees in the European
Union DGXII Biotechnology FW IV Research Programme Molecular
Tools for Biodiversity (ed. Gillet EM). Available online (http://
webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm).

Samour, J. H. 2002. The Reproductive Biology of the Budgerigar (Melopsittacus undulatus): Semen Preservation Techniques and Artificial Insemination Procedures. J Avian Med Surg. 16 (1): 39-49.

Sarkar, I. N. 2006. Phylogenetics in the modern era . J Biomed Inform. 39(1):3-5.

Scoble, J. 1981. Complete book of budgerigars. Sydney: Lansdowne Press.

Shriver, J. 2003. 'Heritage' turkeys bring that old taste home. http://www.usatoday.com/life/2003-11-05-turkey_x.htm. Accessed 2/14/2006.

Smith, E. J.; Jones, C. P.; Bartlett, J.; et al. 1996. Use of randomly amplified polymorphic DNA markers for the genetic analysis of relatedness and diversity in chickens and turkeys. Poult Sci. 75(5):579-84.

Smith, E. J.; Kamara, D.; Pimentel,G; et al. 2005. Avian Genomes: Important Resources for Understanding Vertebrate Biology . Curr Genomics. 6 (2):75-80

Smith, E. J.; Geng, T.; Long, E.; et al. 2005. Molecular analysis of the relatedness of five domesticated turkey strains. Biochem Genet. 43(1-2):35-47.

Soltis, D. E.; Soltis, P. S. 2003. The role of phylogenetics in comparative genetics. Plant Physiol. 132(4):1790-800.

Soragna, D.; Vettori, A.; Carraro, G.; et al. 2003. A locus for migraine without aura maps on chromosome 14q21.2-q22.3. Am J Hum Genet. 72(1):161-7.

Spence, R.; Jordan, W. C.; Smith, C. 2006. Genetic analysis of male reproductive success in relation to density in the zebrafish, Danio rerio. Front Zool. 3(1):5.

Steel, K. P.; Kros, C. J. 2001. genetic approach to understanding auditory function. Nat Genet. Feb;27(2):143-9.

Strachan, T.; Read, A. P. 1999. Genetic Markers. In: 2nd ed. Human Molecular Genetics 2. London: Taylor & Francis.

Strand, M.; Prolla, T. A.; Liskay, R. M.; et al. 1993. Destabilisation of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365: 274-276.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 17(16):6463-71.

Van Camp, G.; Smith, R. J. 2001. Hereditary hearing loss. http://dnalab-www.uia.ac.be/dnalab/hhh. Accessed 2/14/06.

Vignal, A.; Milan, D.; SanCristobal, M.; et al. 2002. A review on SNP and other types of molecular markers and their use in animal genetics.Genet Sel Evol. 34(3):275-305.

Vodenicharova, M. 1989. Use of proteins as molecular-genetic markers in plants. Genet Sel 22**:** 269–277.

von Lindern, J. J.; Aroner, S.; Barrett, N. D.; et al. 2006. Genome analysis and phylogenetic relationships between east, central and west African isolates of Yellow fever virus. J Gen Virol. 87(Pt 4):895-907.

Vos, P.; Hogers, R.; Bleeker, M.; et al. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. Nov 11;23(21):4407-14.

Watmough, W. 1948. The Cult of the Budgerigar. London, England: The Chapel River Press.

Wang, D.G.; Fan, J.B.; Siao, C.J.; et al. 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Sci. 280:1077-1082.

Weber, J. L. 1989. May PEAbundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet. 44(3):388-96

Williams, J. G.; Kubelik, A. R; Livak, K. J.; et al. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18(22):6531-5

Yang, S.; Wang, Z.; Liu, B.; et al. 2003. Genetic variation and relationships of eighteen Chinese indigenous pig breeds. Genet. Sel. Evol. 35: 657-671.

Ye, X.; Zhu, J.; Velleman, S. G.; et al. 1998. Genetic diversity of commercial turkey primary breeding lines as estimated by DNA fingerprinting. Poult Sci. 77(6):802-807.

Zhu, J.; Nestor, K. E.; Patterson, R. A.; et al. 1996. Measurement of genetic parameters within and between turkey lines using DNA fingerprinting. Poult Sci. 75:439-446.

Table 1. GenBank accession numbers (Acc. no) and size of each microsatellite-repeat containing sequence from the budgerigar genomic libraries.

Acc. no	Size (bp)	Acc. no	Size (bp)
AY568101	496	AY568139	572
AY568102	687	AY568140	264
AY568103	661	AY568141	335
AY568104	596	AY568142	321
AY568105	634	AY568143	341
AY568106	359	AY568144	368
AY568107	371	AY568145	457
AY568108	661	AY568146	426
AY568109	537	AY568147	320
AY568110	458	AY568148	597
AY568111	666	AY568149	650
AY568112	664	AY568150	677
AY568113	575	AY568151	584
AY568114	542	AY568152	664
AY568115	632	AY568153	527
AY568116	600	AY568154	510
AY568117	601	AY568155	317
AY568118	678	AY568156	287
AY568119	373	AY568157	632
AY568120	459	AY568158	700
AY568121	293	AY568159	399
AY568122	308	AY568160	315
AY568123	232	AY568161	323
AY568124	364	AY568162	596
AY568125	548	AY568163	460
AY568126	667	AY568164	421
AY568127	475	AY568165	377
AY568128	459	AY568166	292
AY568129	503	AY558876	467
AY568130	703	AY558877	660
AY568131	749	AY558878	713
AY568132	300	AY558879	638
AY568133	557	AY558880	585
AY568134	500	AY558881	600
AY568135	450	AY558882	350
AY568136	431	AY558883	660
AY568137	503	AY558884	442
AY568138	286	AY558885	666

Table 2. Sequence and repeat motif of designed primer pairs and annealing temperature of primers of the budgerigar loci tested for polymorphism.

Primer ID	Forward (5'-3')	Reverse (5'-3')	Repeats	$^{\S}T_m$ (°C)
BGMsat1‡	agcaaccatatgagtgtcttgtg	tgcagagtttcctctgatttga	(cmpd)*	62
BGMsat2A‡	tgctgtcagagatttgctgc	tttctggaaccctcaggcta	$(TG)_5$	
BGMsat2B	catgtagtagcctgagggttcc	tttttgctcatttttccaataca	$(GAAA)_3$	62
BGMsat3	ccatccctaggagtcacagc	caaacggatctgttgtggtg	$(AC)_9$	
BGMsat4	ctttgggaagcaaaaatcca	gcccgtcctctcatccata	(cmpd)*	
BGMsat5	aagatctgtctctccctcaccc	accgcaatgtccatcttttagt	(cmpd)*	
BGMsat6	tggagctgtgacaaatgacttc	gccagtgtgcggatattagact	$(AC)_{13}$	
BGMsat7	aatcatggtcatgctgtttctg	gagagttagaggaaggggggaag	(cmpd)*	
BGMsat8‡	gtccacctatgagctcttgacc	acctcttgctgaaagcagaaac	$(CACAGA)_5$	62
BGMsat9	catgcaggtctgtccatcac	tgtccccttagcctagaccc	(cmpd)*	
BGMsat10	gaggttgcagtgagctgtca	gcccccgaaaaatgaatct	$(gaaa)_{8*}$	
BGMsat11‡	gtggttattcagaaggcaaagg	tcctgaagtggttttccttgtt	(cmpd)*	62
BGMsat12‡	tgcttcccaggatagcatct	aaggattcttcccctctcca	(cmpd)*	59
BGMsat13	aagatgtggagaggggaagaa	gcaccggctactaggaagaa	(cmpd)*	
BGMsat14‡	gagccgttatcactggctaga	acatctgtggtgctcttaca	(cmpd)*	59
BGMsat15	gccacagcaaattcctccta	aaggcctagctagcagaatcc	(cmpd)*	
BGMsat16	ggcctagctagcagaatcaca	gaaaccgtggtggaagattg	(cmpd)*	
BGMsat17‡	tctcttcaggaacagagaatcaaa	cccaaaagagacaccgaagt	(cmpd)*	59
BGMsat18‡	gctgctgaccagtttttggt	gaatcttttcctccagctcaga	(cmpd)*	59
BGMsat19	aactgcccgaagaaaaatca	tatcggcagaattcggtttt	(cmpd)*	
BGMsat20‡	cccagttcccccagtctatt	cgctcaactctgtgtgagga	(cmpd)*	59

* Compound microsatellite repeat.
‡ Microsatellite loci that were tested for informativeness
$^{\S}T_m$ represents the optimized annealing temperature at which a single amplicon of the expected size was obtained.

Table 3. Expected PCR product size, and allele frequency of each allele detected at nine microsatellite loci in the budgerigar.

Primer ID	Size (bp) [§]	Alleles(bp) [‡]	Frequency *
BGMsat1	159	157	0.867
		173	0.133
BGMsat2B	135	132	0.333
		134	0.670
BGMsat8	153	147	0.476
		153	0.524
BGMsat11	180	178	1.000
BGMsat12	197	188	0.045
		198	0.958
BGMsat14	187	187	0.632
		191	0.368
BGMsat17	198	194	0.125
		196	0.750
		198	0.125
BGMsat18	197	196	0.056
		200	0.611
		204	0.306
		208	0.028
BGMsat20	177	171	0.200
		179	0.800

[§] Estimated size based on primer binding site in the sequences described in Table 2.
[‡] Alleles were estimated using ABI Prism 377.
***** Frequency was estimated from genotypes of 21 unrelated budgies obtained from North Carolina, Idaho and Oregon.

Table 4. Budgerigar primer-based comparative analysis of chicken and turkey genomes.

Budgerigar Sequence ID	Chicken		Turkey	
	% Similarity	Overlap	% Similarity	Overlap
BGMsat1	100	22/22	100	23/23
BGMsat2A	88	138/156	-	-
BGMsat2B	91	128/140	-	-
BGMsat8	89	140/156	89	137/153
BGMsat14	75	143/189	88	167/189
BGMsat18	83	146/175	82	140/169
BGMsat20	73	84/115	72	86/118

Table 5. BLAT-based comparative analysis of budgerigar DNA sequences against the recently released draft of the chicken genomic DNA sequence.

Accesssion	bp	% Similarity	Overlap (bp)	Chromosome
AY568101	496	90.0	120†	8
AY568102	687	94.2	58	1
AY568103	661	89.9	86	12
AY568104	596	85.5	132†	1
AY568105	634	100	36	2
AY568106	359	87.5	77	5
AY568107	371	96.3	48	7
AY568108	661	69.0	27	2
AY568109	537	90.0	56	1
AY568110	458	70.0	28	2
AY568111	666	88.3	30	4
AY568112	664	96.8	29	1
AY568113	575	83.7	206†	3
AY568114	542	92.8	47	Un*
AY568115	632	91.5	107†	1
AY568116	600	51.8	27	5
AY568117	601	66.7	39	17
AY568118	678	88.4	160†	27
AY568119	373	98.2	363†	13
AY568120	459	60.0	24	Un*
AY568121	293	89.8	69	2
AY568122	308	95.1	56	1
AY568123	232	91.4	52	3
AY568124	364	96.9	29	Un*
AY568125	548	100	23	1
AY568126	667	100	23	Un*
AY568127	475	93.3	51	1
AY568128	459	91.7	30	7
AY568129	503	97.0	30	11
AY568130	703	95	37	4
AY568131	749	66.7	22	Un*
AY568132	300	100	27	3
AY568133	557	95.3	38	18
AY568134	500	100	21	1
AY568135	450	81.9	47	Un*
AY568136	431	77.8	30	1
AY568137	503	97.0	31	11
AY568138	286	90.6	44	1
AY568139	572	97.2	66	4
AY568140	264	82.8	26	1
AY568141	33	96.7	28	2
AY568142	321	100	21	6
AY568143	341	97.6	35	7
AY568144	368	96.5	26	3
AY568145	457	88.1	256†	3
AY568146	426	92.6	194†	4
AY568147	320	95.1	56	1
AY568148	597	97.3	34	1
AY568149	650	81.3	105†	3
AY568150	677	91.0	62	11

*Unknown chromosomal location.

†These sequences appear to have, based on the length and high percentage of the matched sequence, a chicken homologue.

Table 6. Repeat motif, nucleotide sequences, 5' to 3', of primers used for polymerase chain reaction, expected size range and optimized annealing temperature for each primer pair.

Primer ID	Repeat motif	Forward primer ‡	Reverse primer ‡	Exp. range (bp)	§T_m °C
RHT0024	$(AC)_{24}$	TCTCTGGGGAAGAAGGGG	AGGCCACACTGGATAACAGC	132–141	56
RHT0294	$(AC)_{12}$	CTGAAAAAGAGTCCTTGAGTGC	AACAACCAGTTTTATAGCCACC	131–138	56
RHT0095	$(AC)_{12}$	AGGAATCCATTGTGGTGGAG	GTTGGAATTTGTGGTGGGAC	242–256	56
TUM20	$(CT)_{24}CA$	TCAGTCCTGGCAGTTTAC	GTTCCTGTGTAGGACAAT	111–156	56
RHT0009	$(AC)_{15}$	TCTCTCACATCCCTGTTCTGC	TTGCCAGATGCTTCTTTGTG	153–173	56
RHT0011	$(AC)_{19}$ G $(AC)_2$ G $(AC)_3$ G $(AC)_7$	GTGTTTCCTCCTAAGCCATACG	GTTCATCTATCTGCTCGCTGC	137–150	56
RHT0216	$(AC)_{17}$	GTAGCTTTGCCCTCAAGACG	CAGTGCCAATCACCCTCC	139–148	56
ADL0023	$(CA)_5(CG)_4(CA)_9$	CTTCTATCCTGGGCTTCTGA	CCTGGCTGTGTATGTGTTGC	158–178	56
RHT0131	$(AC)_7$	CAGCATTCTCGAACTTCTGG	GACTGATTTGGGATGGATGC	151–166	56
TUM16	$(AC)_{16}$	CCCTTGTTGGTAACACTT	TCTCACCATCCTCACCTT	151–188	56

§T_m represents the optimized annealing temperature at which a single amplicon of the expected size was obtained.
‡ Primer pairs previously described by Burt et al. (2003).

Table 7. Allelic frequencies of five turkey microsatellite loci.

	RHT0024‡ (132-141bp) †			RHT0294‡ (131-138bp) †			RHT0095‡ (242-256bp) †			TUM20‡ (111-156bp) †				RHT0009‡ (153-173bp) †		
	122±	130	140	132	136	142	240	246	252	100	120	140	160	144	150	178
NA§	0.250	0.160	0..590	0.740	0.100	0.160	0.360	0.550	0.090	0.020	0.000	0.130	0.850	0.000	0.583	0.417
BR	0.040	0..260	0.700	0.900	0.000	0.100	0.650	0.350	0.000	0.000	0.000	0.330	0.670	0.040	0.340	0.620
BS	0.075	0.375	0.550	0.760	0.040	0.200	0.140	0.610	0.250	0.125	0.000	0.150	0.725	0.146	0.792	0.062
SB	0.190	0..310	0.500	0.940	0.040	0.020	0.630	0.350	0.020	0.040	0.120	0.260	0.580	0.200	0.620	0.180
RP	0.400	0.380	0.230	0.660	0.000	0.340	0.200	0.660	0.140	0.000	0.000	0.348	0.652	0.217	0.566	0.217
CO	0.100	0.000	0.900	0.810	0.070	0.120	0.190	0.560	0.250	0.000	0.000	0.409	0.591	0.022	0.217	0.761

‡Microsatellite loci.

† Range within which alleles for each locus fall.

± Approximate size of alleles observed at each locus.

§ Where NA ,BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain, respectively.

Table 8. Allelic frequencies of additional five turkey microsatellite loci.

	RHT0011‡ **(137-150bp)** †			***TUM16***‡ **(151-188bp)** †			***RHT0216***‡ **(139–148bp)** †				***ADLOO23***‡ **(158-178bp)** †		***RHT0131***‡ **(151-166bp)** †		
	140±	146	150	158	162	170	120	130	150	156	160	170	154	160	168
NA§	0.479	0.313	0.208	0.000	0.100	0.000	0.000	0.326	0.523	0.152	0.040	0.960	0.250	0.330	0.420
BR	1.000	0.000	0.000	0.000	0.952	0.048	0.000	0.022	0.978	0.000	0.000	1.000	0.304	0.413	0.283
BS	0.763	0.000	0.236	0.000	0.813	0.187	0.000	0.417	0.417	0.166	0.150	0.850	0.227	0.545	0.227
SB	0.454	0.318	0.227	0.320	0.620	0.060	0.000	0.140	0.640	0.220	0.555	0.455	0.500	0.260	0.240
RP	0.913	0.021	0.065	0.000	0.909	0.091	0.109	0.630	0.174	0.087	0.105	0.895	0.547	0.310	0.143
CO	0.500	0.500	0.000	0.000	0.000	1.000	0.000	0.435	0.565	0.000	0.087	0.913	0.000	0.977	0.023

‡Microsatellite loci.

† Range within which alleles for each locus fall.

± Approximate size of alleles observed at each locus.

§ Where NA ,BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain, respectively.

Table 9. Pairwise genetic distance (Nei) between six turkey populations based on frequencies determined using microsatellite-based alleles.

	NA †	BR	BS	SB	RP	CO
NA †	0.000000					
BR	0.089772	0.000000				
BS	0.068903	0.139911	0.000000			
SB	0.131444	0.153629	0.152738	0.000000		
RP	0.108381	0.180905	0.069508	0.211093	0.000000	
CO	0.306311	0.299395	0.297193	0.401194	0.411111	0.000000

† Where NA, BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain respectively.

Table 10. Measure of genetic differentiation (R_{st}) between six turkey populations as calculated by RST CALC based on 1000 permutations.

	NA[†]	BR	BS	SB	RP	CO
NA[†]	0.0000					
BR	0.0742[§]	0.0000				
BS	0.1822[§]	0.3187[§]	0.0000			
SB	0.0979[§]	0.1001[§]	0.1670[§]	0.0000		
RP	0.0631[§]	0.2362[§]	0.1256[§]	0.1059[§]	0.0000	
CO	0.1077[§]	0.0956[§]	0.2552[§]	0.1798[§]	0.2079[§]	0.0000

[§] Genetic differentiation between populations were significant at $p<0.05$.
[†] Where NA, BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain respectively.

Table 11. Pairwise estimates of gene flow (Nm) between six turkey populations as calculated by RST CALC.

	NA[†]	BR	BS	SB	RP	CO
NA[†]	0.0000					
BR	3.1174	0.0000				
BS	1.1223	0.5343	0.0000			
SB	2.3043	2.2465	1.2471	0.0000		
RP	3.7144	0.8084	1.7399	2.1112	0.0000	
CO	2.0715	2.3656	0.7297	1.1407	0.9527	0.0000

[†] Where NA, BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain respectively.

Table 12. Frequencies of *cTnT*-based alleles detected in SNP analysis of six turkey population.

Population[†]	Allele [‡]	
	G	A
NA	0.76	0.24
BR	0.40	0.60
BS	0.32	0.68
SB	0.62	0.38
RP	0.44	0.56
CO	0.64	0.36

[†] Where NA, BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain, respectively.
[‡]The allele is based on the G/A SNPs corresponding to the *ApaL I* recognition site in cardiac troponin *T*.

Table 13. Pairwise genetic distance (Nei) between six turkey populations based on allelic frequencies determined using the *cTnT*-based SNP alleles.

	NA[†]	BR	BS	SB	RP	Co
NA[†]	0.000000					
BR	0.249091	0.000000				
BS	0.387868	0.011016	0.000000			
SB	0.030062	0.096802	0.179295	0.000000		
RP	0.191246	0.003042	0.025788	0.064373	0.000000	
CO	0.021477	0.114979	0.204921	0.000702	0.079065	0.000000

[†] Where NA, BR, BS, SB, RP, and CO represent Narragansett, Bourbon Red, Blue State, Spanish Black, Royal Palm and commercial strain, respectively.

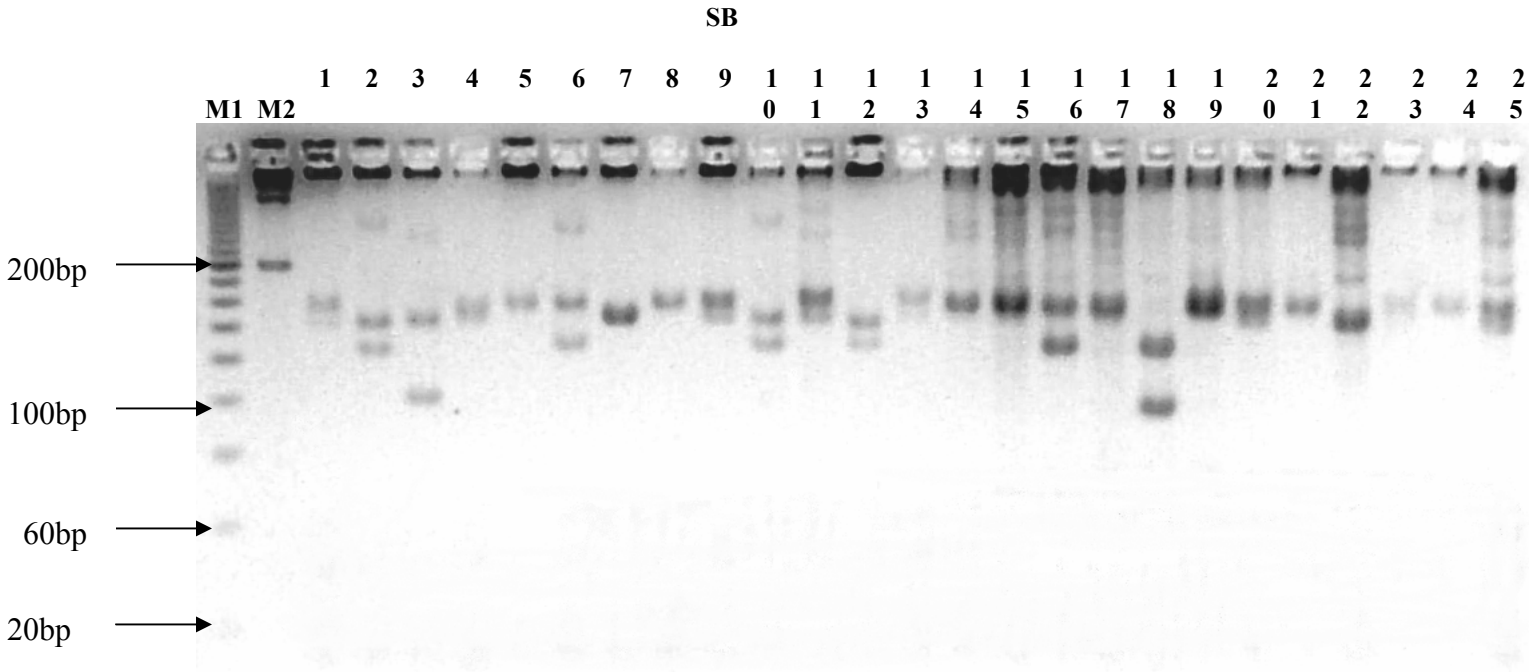

Fig. 1. Metaphor agarose gel electrophoretic patterns of *TUM20* amplicons produced using template from Spanish Black (SB) heritage turkeys. Alleles ranged from ~ 110 to 170bp. M1 and M2 represent DNA ladder I and II, respectively (Gene Choice).

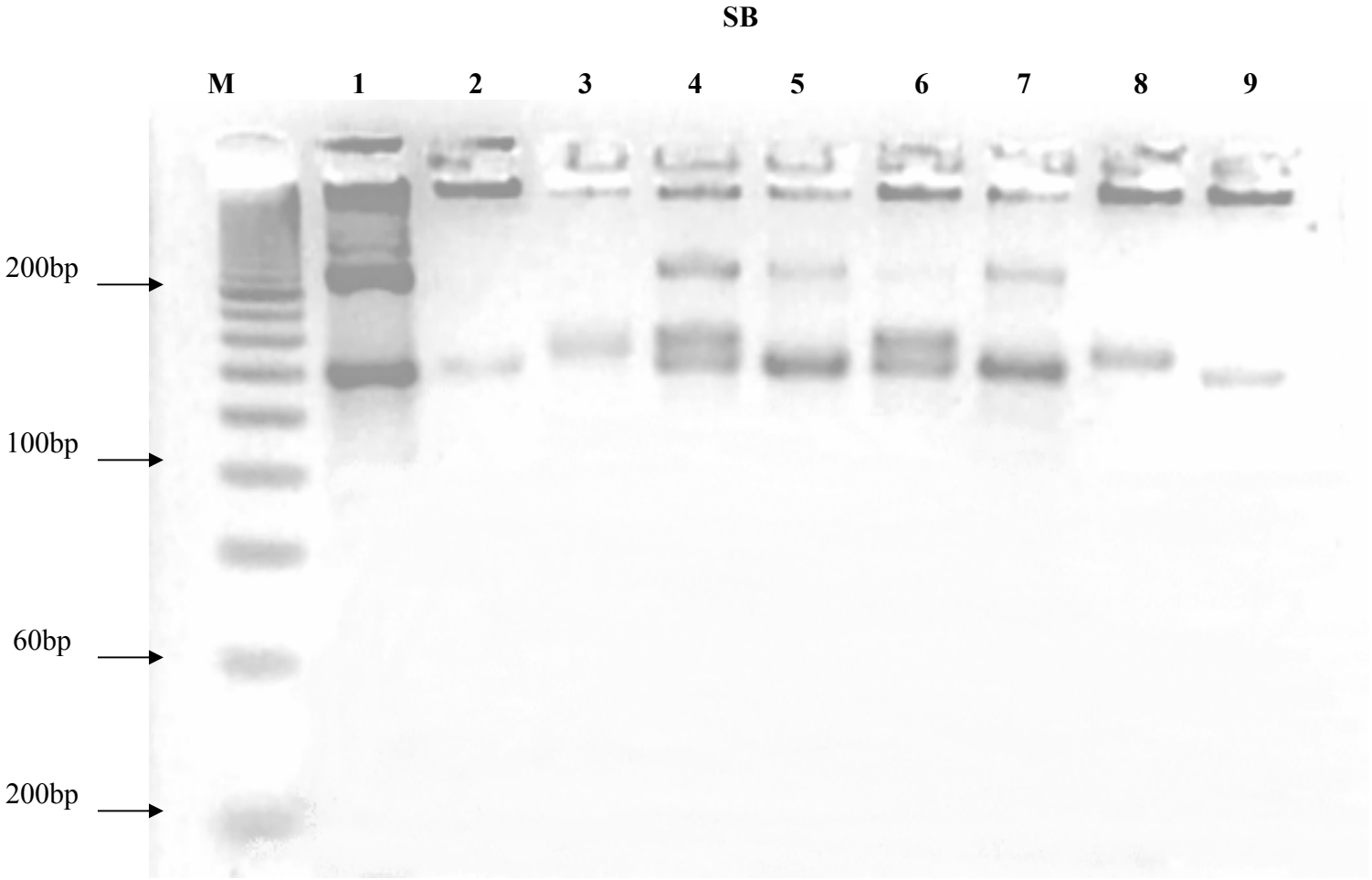

Fig. 2. Metaphor agarose gel electrophoretic pattern of *RHT0011*-based amplicons using template from Spanish Black (SB) heritage turkeys. Alleles ranged from ~ 138 to 150bp.

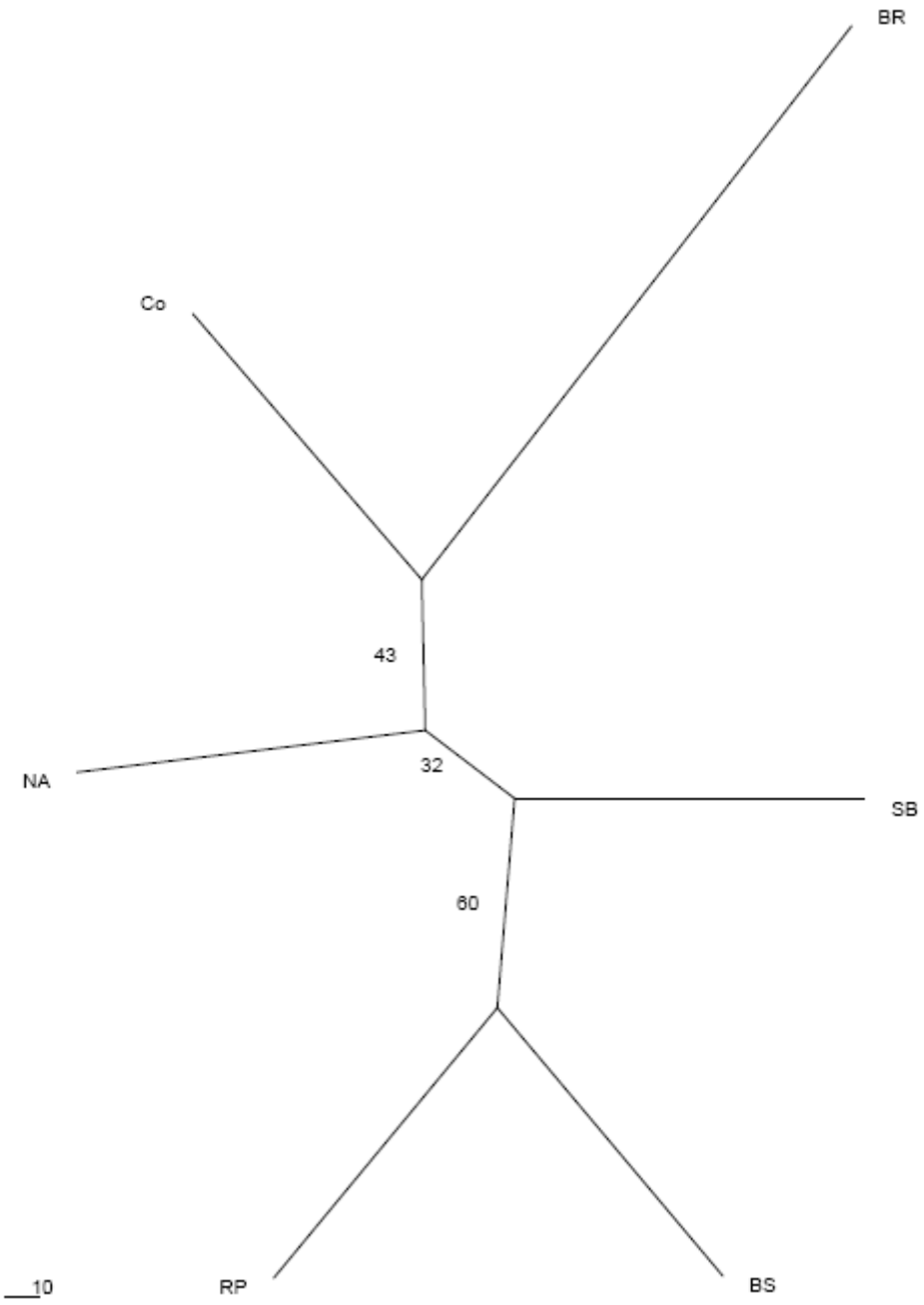

Fig. 3. A Neighbor-joining tree using Nei's distance *D* for six turkey populations sampled at 10 microsatellite loci. Bootstrap values indicated at the nodes were calculated from 100 resamplings.

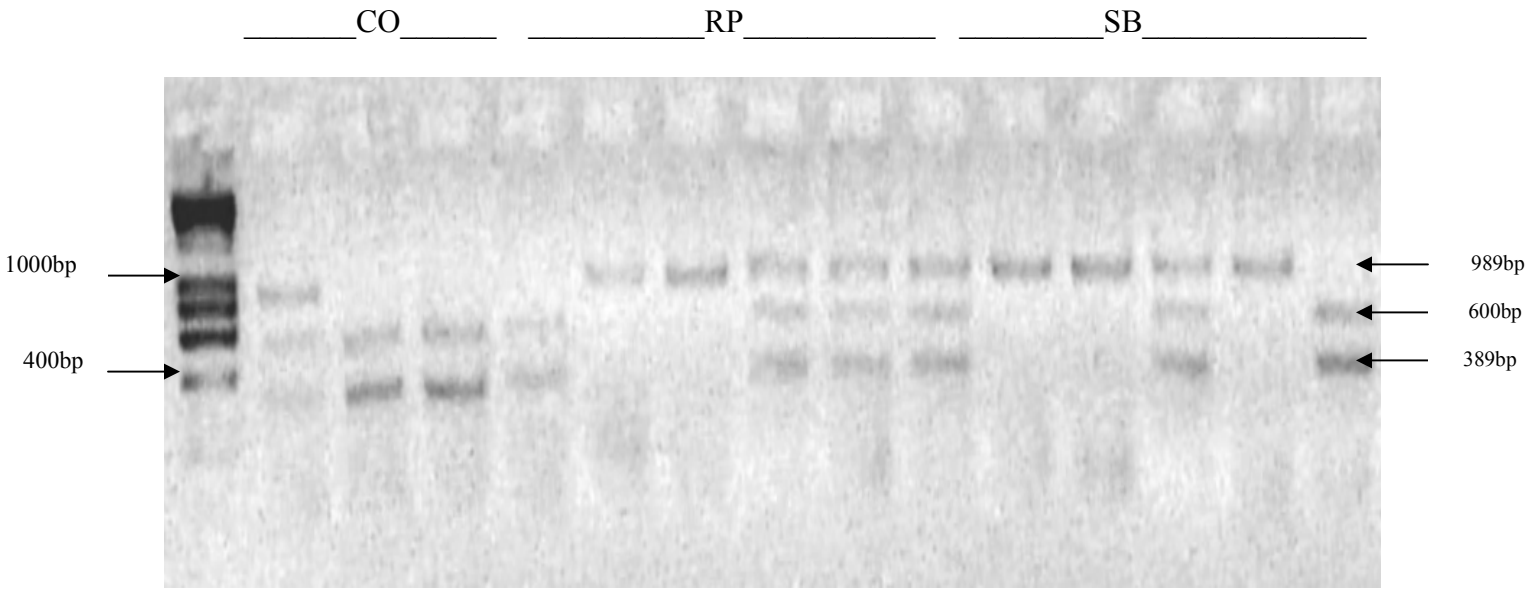

Fig. 4. Gel pattern of *ApaL I* digested cTnT-based PCR products for commercial (CO), Royal Palm (RP), and Spanich Black (SB) turkeys. Lanes with three and two fragments represent amplicons from birds heterozygous and homozygous respectively for the G/A SNP in cardiac *TnT* gene. Single fragments represent amplicons without the G allele that is the site of digestion by *ApaL I*.

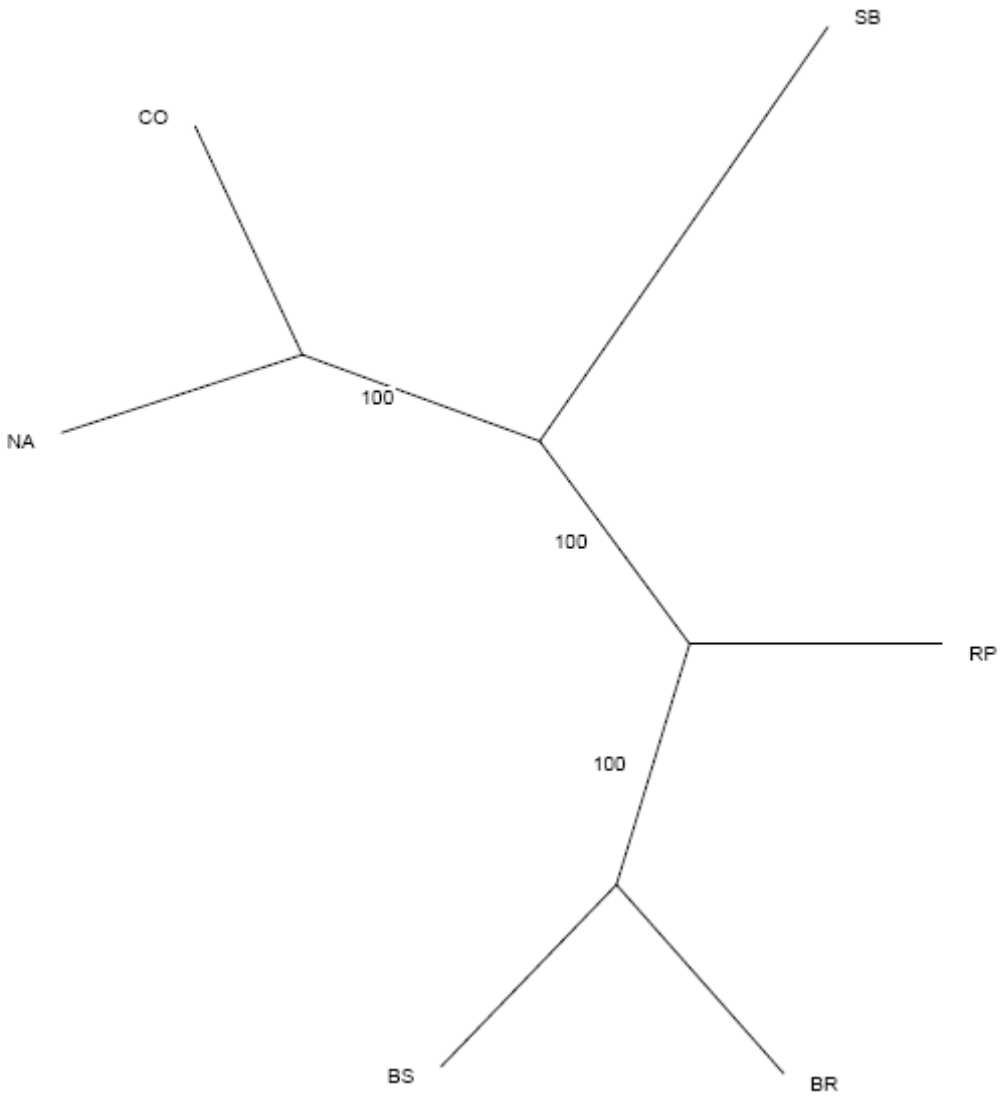

Fig. 5 Neighbor-Joining tree using Nei's distance *D* for six turkey populations sampled at the *cTnT* locus. Bootstrap values indicated at the nodes were calculated from 100 resamplings.