Molecular phylogeny and evolution of Amazon parrots in the Greater Antilles

by

Sofiia Kolchanova

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER IN SCIENCE

in

Biology

University of Puerto Rico
Mayagüez Campus
2018

Approved by:

_____	_____
Taras Oleksyk, Ph. D. President, Graduate Committee	Date
_____	_____
Audrey Majeske, Ph. D. Member, Graduate Committee	Date
_____	_____
Juan Carlos Martínez Cruzado, Ph.D. Member, Graduate Committee	Date
_____	_____
Ana Vélez Díaz, M.S. Chairperson of the Department	Date
_____	_____
Pablo Cáceres Valencia, Ph.D. Graduate Studies Representative	Date

ABSTRACT

Genus *Amazona* is a unique group of Neotropical parrots distributed across South and Central America, as well as in the Caribbean Islands. It is characterized by vocal learning, high intelligence, social structure, longevity, diverse coloration patterns and habitats. There is a number of island species among them, which represent promising study cases for evolutionary biology. According to the 2016 IUCN Red List of Threatened Animals, 20 of the 35 recognized species of Amazons are listed as either vulnerable, endangered or critically endangered. Despite all notoriety, this is a relatively poorly studied group: substantial parts of its evolutionary history remain unresolved, since the so far published phylogenies suggest contradictory scenarios concerning directions of islands-mainland colonization. In our study we are hoping to shed light on the phylogenetic relationships within the group of Greater Antillean amazons, evolution of their mitochondrial genomes and to provide some useful data for further genomic research as well as for the endangered species recovery programs. We have sequenced, assembled and annotated mitochondrial genomes of 7 species – *A. vittata* (Puerto Rican amazon), *A. ventralis* (Hispaniolan amazon), *A. leucocephala* (Cuban amazon), *A. albifrons* (white-fronted amazon), *A. agilis* (black-billed Jamaican amazon), *A. collaria* (yellow-billed amazon), *Amazona rhodocorytha* (red-browed amazon) and a potential extinct subspecies of *A. vittata* from Vieques island. We have run Bayesian analyses on this data in order to resolve phylogenetic relationships within the group and to assess possible island colonization scenarios.

RESUMEN

El género *Amazona* es un grupo único de cotorras neotropicales distribuidas en América del Sur y Central, así como en las islas del Caribe. Se caracteriza por aprendizaje vocal, alta inteligencia, estructura social, longevidad, diversos patrones de coloración y hábitats. Existe una cantidad de especies insulares entre ellas, que representan casos de estudio prometedores para la biología evolutiva. Según la Lista Roja de Animales Amenazados de la UICN de 2016, 20 de las 35 especies reconocidas de amazonas figuran como vulnerables, en peligro o en peligro crítico. A pesar de toda la notoriedad, este es un grupo relativamente poco estudiado: partes sustanciales de su historia evolutiva siguen sin resolverse, ya que las filogenias hasta ahora publicadas sugieren escenarios contradictorios sobre las direcciones de la colonización de la parte continental y las islas. En nuestro estudio, esperamos arrojar luz sobre las relaciones filogenéticas dentro del grupo de las amazonas de las Antillas Mayores, la evolución de sus genomas mitocondriales y proporcionar algunos datos útiles para futuras investigaciones genómicas, así como para los programas de recuperación de especies en peligro de extinción. Hemos secuenciado, ensamblado y anotado genomas mitocondriales de 7 especies - *A. vittata* (cotorra puertorriqueña), *A. ventralis* (de la Española), *A. leucocephala* (cubana), *A. albifrons* (frentiblanco), *A. agilis* (amazona de pico negro), *A. collaria* (a. de pico amarillo), *A. rhodocorytha* (a. de ceja roja) y una subespecie potencial extinta de *A. vittata* de la isla de Vieques. Hemos realizado análisis Bayesianos sobre estos datos para resolver las relaciones filogenéticas dentro del grupo y evaluar posibles escenarios de colonización de las islas.

ACKNOWLEDGEMENTS

I want to express my infinite gratitude to all those who have made it possible for this project to be carried out and completed and who were there for me during this period of my life.

First of all, I would like to thank to my family and friends, who supported me in the decision to move to Puerto Rico and who motivated and encouraged me throughout the whole time I spent on the island.

Second, I want to express endless gratitude to my advisor, Dr. Taras Oleksyk for his supervision and all kinds of help during these last two years. It's impossible to imagine this project and my whole presence here without him and his wife, Dr. Audrey Majeske, who was also very supportive at all times. During the hurricane me and another student stayed at their house and who knows what we would happen if not for them.

I want to express my gratitude to the Administration of the Biology Department, in particular to Ms. Mary Jimenez, who was always incredibly nice and helpful when I was confused about something related to documentation and formal procedures.

Thanks to Dr. Juan Carlos Martínez-Cruzado for the assistance and counsel he kindly provided throughout the whole time I was working on this project.

I am grateful to the team of the Dobzhansky Center for Genome Bioinformatics in St. Petersburg, in particular to Pavel Dobrynin and Alexey Komissarov for sharing their expertise and consulting me on technical questions. Thanks for Dr. Klaus-Peter Koepfli from the Smithosian Institute, who has also consulted me and ensured sequencing of certain samples.

Thanks to Jafet Velez-Valentin and all the Puerto Rican Parrot Recovery Program team for providing the valuable samples, which were used in this study.

Thanks to my labmates, in particular to Walter Wolfsberger and Stephanie O. Castro for their help and just for being my very good friends.

This study was supported in part by the NSF grant NSF DUE 1432092, and by the Russian Ministry of Science Mega-Grant no. 11.G34.31.0068.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Ever since Darwin's observations about finches, mockingbirds and tortoises on the Galapagos in the Voyage on the Beagle [Darwin, 1845], islands have been an important source of observational as well as experimental evidence for evolution [McArthur and Wilson, 1967; Grant and Grant, 2014] and continue to advance the understanding of the evolutionary process [Sangeet et al., 2017]. Islands still provide valuable model systems for the fundamental studies of migration, diversification, and extinction [Whittaker 2017]. Because many individual islands are young and contain relatively few species, evolutionary adaptations and species proliferation may be obvious and easier to study compared to the continental populations with large population sizes and boundaries that are difficult to define [Losos 2009]. Geographic isolation of many islands allows evolution of resident species to take its own course, free of influence from other areas, resulting in unusual faunas and floras, often unlike those found anywhere else [McArthur and Wilson, 1967, Losos 2009]. This is reflected in the amazing biodiversity that is observed across the islands overall: comprising just 3.5% of Earth's land area, islands contribute 15 to 20% of terrestrial species [Whittaker 2017]. However, the combination of diversity and vulnerability means that the islands also contribute to a larger extent to the contemporary extinction crisis [Diamond 1989].

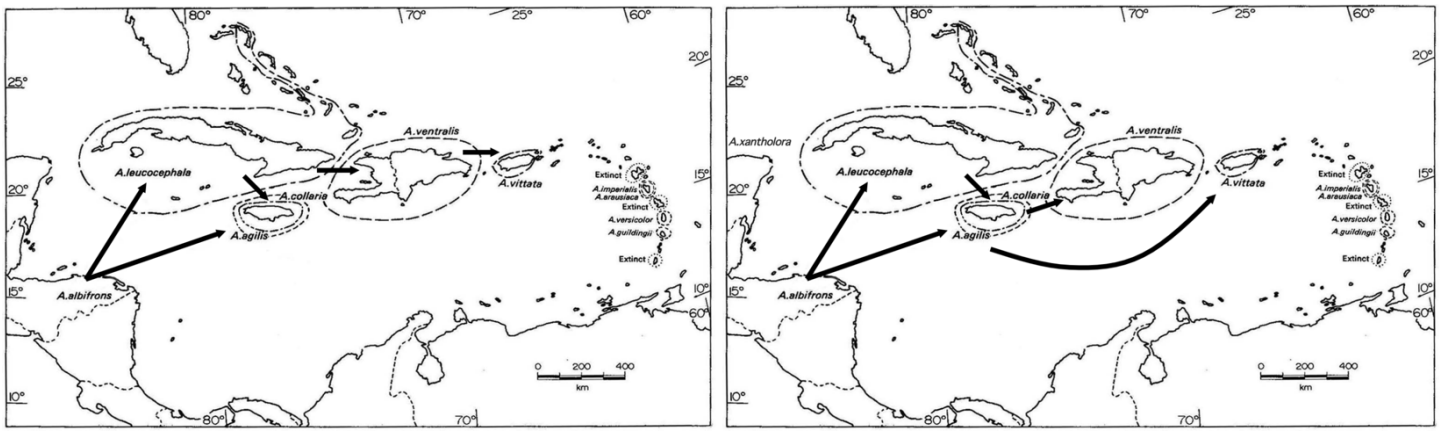
The most important question being asked about evolution is what drives speciation/what are the underlying mechanisms of species formation and adaptive radiation [Losos 2009]. Islands are great natural laboratories, providing innumerable

replicated “experiments” in the factors controlling the distribution, character, and diversity of species. They vary in many ways, and the evolutionary proliferation of life has progressed much further on some islands than on others. The two principal theories that have been used to explain the origins of island species are overwater dispersal and vicariance (fragmentation of habitats). Island vicariance can occur by geologic factors (plate tectonics) or sea level changes that can result in the isolation of ancestral biota [Ottens-Wainright 2004]. Combined with the effect of dispersals from the continent this should result in species on the islands closest to the mainland having the least unique, and the most remote islands - the most unique biodiversity, as these have been separate from the mainland for the longest period of time [McArthur and Wilson 1967]. The varied outcomes of evolution in island settings can indicate a great deal about how evolutionary and spatial processes have built biological diversity through the formation and differentiation of species [Losos 2009].

Amazon parrots (*Amazona sp*.) that currently inhabit the Greater Antillean Islands of Cuba, Jamaica, Hispaniola and Puerto Rico, represent a fascinating model of speciation on islands similar to the Darwin’s finches of the Galapagos [O’Brien 2012]. Taking into consideration significance of the parrots to the island history and ecology, we believe it is very important to understand how these species came to be and how they adapted to specific island environments.

Based on morphological and geological data two alternative migration-speciation scenarios have been proposed for the Greater Antillean *Amazona* by Lack (1976) and

Snyder (1987) – for details see Figure 1. Later, several studies attempted to reconstruct the evolutionary history of speciation in this clade, but these efforts have been limited by the state of the sequencing technology allowing only partial reconstruction of mitochondrial sequences [Rusello & Amato, 2004]. In the last decade the rapid development of NGS (Next-Generation sequencing) technologies has made it possible to identify and interpret differences from closely related genomes in their evolutionary context [Goodwin et al 2016; Allendorf et al 2010]. Through these techniques, a large phylogenetic reference framework has been created thanks to the concerted efforts of sequencing whole genomes for different vertebrates such as Genome 10K [Koepfli et al. 2015]. Therefore, the description of small-scale ("local") variation between closely related species that occupy various environments and display diverse phenotypes with an aim to find functional elements and signatures of adaptation at molecular level would be the next logical step towards better understanding of evolutionary biology. Direct comparisons of genome sequences obtained from organisms with related but different phenotypes can be used to discover genetic variation and identify functions of particular genes. Chromosome-level assemblies hold the potential to reveal answers to all kinds of evolutionary questions (e.g. evolutionary history of genes, gene families and of the species itself, past population dynamics, exact times of speciation events and phylogenetic relationships with other species, understanding of gene functions and evolution of gene function etc.). As this trend continues in the near future, interrogating ever more complete genomic datasets has the potential to give answers to such longstanding questions of evolutionary biology as the role of genetic drift, selection and gene flow in the speciation process [Oleksyk et al., 2010].



A B

Figure 1. The two alternative migration-speciation scenarios proposed for Greater Antillean *Amazona*. A - after Lack 1976, B - after Snyder 1987

While whole genomes hold ultimate answers to many intriguing questions, the computational complexity of most of these problems is still immense, and even the task of proper *de novo* whole genome assembly is still not completely solved [Zhang et al., 2014; Jarvis et al., 2014]. Fortunately, in many cases assembly of every single chromosome is not necessary to address many important questions about evolutionary history. Given today's technology, the mitochondrial genome (thousands of bp) is much easier to assemble than a complete nuclear genome (billions of bp), and it does carry a lot of important information [Chen an Butow, 2005]. Mitochondrial sequences have been used for a long time for phylogenetic and population studies because of its mode of inheritance that does not involve recombination [Cann et al., 1987]. Due to the predominantly maternal transmission of mitochondrial genetic material, it is possible to reconstruct mitochondrial phylogenetic trees and track ancestors of many diverged animal populations as they migrated from their geographical origin [Avise 2000]. These advantages have made mtDNA an important phylogenetic marker, long before the full nuclear genome sequences became available [Boore and Brown 1998; Avise 2000]. At

the same time, mtDNA studies have their limitations, mainly because the entire mitochondrial genomes behave as a single locus, and their phylogeny does not fully describe the evolutionary history of a sexually reproducing species, as half of the lineages are lost in each generation. Phenomena such as incomplete lineage sorting and gene flow following the speciation event interfere with proper interpretation of phylogenetic trees built using mitochondrial genetic data [Avise 2000].

In this study, we take advantage of the new genomic technologies and bioinformatics approaches to ask questions about the evolutionary history of the Amazon parrots in the Caribbean that were previously difficult to answer. As genomic DNA collected in multiple collaborations has grown to represent the entire branch of parrot evolution in the Greater Antillean islands, we describe the evolutionary history of the entire avian branch by harnessing the power of whole genome sequencing, ancient DNA reconstruction, and big data analysis. For the completion of this thesis, we do the following: 1) sequence and assemble the mitochondrial genome sequences of seven species of amazon parrots from the Greater Antilles and the adjacent continental regions of Central and North America (*Amazona vittata, A. ventralis, A. leucocephala, A. agilis, A. collaria, A. xantholora, and A. albifrons*), as well as an extinct potential subspecies of *A. vittata* from the island of Vieques; 2) annotate genes and other functional elements of the mitochondrial genomes; 3) use the new data to re-validate the phylogenetic relationships within the group, 4) test different migration/speciation hypotheses. The results of these analyses are discussed and presented in the light of the knowledge about geological history of the Caribbean, current biogeographical

theory, and recent parrot evolution. Using advanced genomics and bioinformatics tools available, this study contributes new information and expands the context of the ecological and conservation studies of these endemic and often endangered species of birds.

LITERATURE REVIEW

1. GEOLOGICAL HISTORY OF THE CARIBBEAN

The origin of the Caribbean islands is largely the consequence of the movement of the Caribbean plate, a lithospheric plate that consists mainly of anomalously thick, oceanic plateau pushed between the two major continental plates (North and South American, Figure 2). The contact between the thick oceanic crust and continental crust plates causes complex tectonic interactions, leading to rising, submerging and moving of the various small islands [Lidiak 1998]. The origin of the Caribbean plate can be traced to the breakup of Pangaea in mid-Jurassic, when Laurasian supercontinent began to separate from Gondwana [Hedges 1996]. As the Atlantic Ocean slowly widened, the North American and South American plates were pushed westward, separated for some time by oceanic crust [Meschede 2002]. The initial Pacific Ocean floor that formed the gap between the two continents subsequently disappeared through subduction, but the Caribbean plate overrode the ocean floor, and continued to move eastward. Ultimately, the formation of the Isthmus of Panama 3-10 million years ago, cut off the last connection between the Caribbean and the Pacific [Meschede 2002]. As the Caribbean plate continued to move, pushing the North American plate beneath, volcanic islands formed along its northern and eastern margins creating the proto-Antillean island arc. This volcanic arc was the subject of constant tectonism and sea-level fluctuations [Pindell 1994]. Paleogeographic reconstructions in Cuba and Hispaniola suggest that some land areas were emergent throughout the Cenozoic (65-0 Mya), possibly forming

small islands [Hedges 1996]. Most of Puerto Rico and the Virgin Islands, was probably submerged from the late Eocene to the Pliocene. Palynological and other evidence indicates that that some emergent land areas (with uplift up to several kilometers) may have persisted on the Puerto Rican Bank throughout the Cenozoic [Hedges 1996].

Today, the Caribbean islands form a large arch from the Yucatan Peninsula to the coast of Venezuela. The Greater Antilles constitute nearly 90% of the landmass of the entire West Indies and include the islands of Cuba, Hispaniola, Jamaica, Puerto Rico and the Cayman Islands. The largest island is Cuba, with area of 110860 square kilometers. Since these islands have been emergent (at least partially) and separated from the mainland since late Cretaceous [Woods 2001], they have been natural laboratories for the development of diverse and endemic natural biodiversity. Knowledge of area's geologic history is crucial for any biogeographic inference since evolution of any organism is ultimately linked to the evolution of the landmass where it exists.

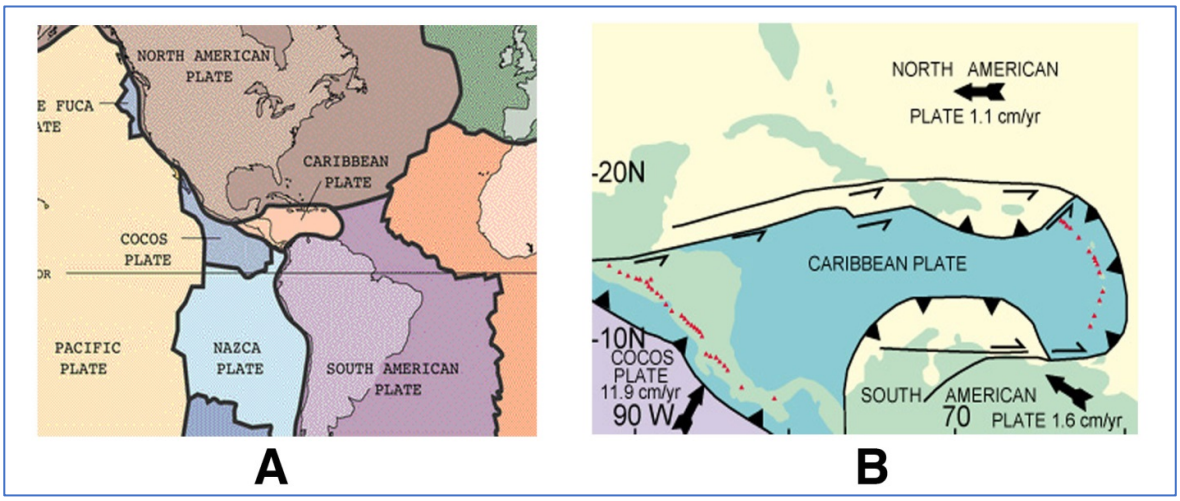


Figure 2. Location of the Caribbean plate (A), its movement and subduction zones (B) (from Kious 1996, Greely 2012)

2. BIOGEOGRAPHIC MECHANISMS

The historical biogeography of the Caribbean has been studied for more than a century and remains an active area of research [Weaver 2016; Matos-Maraví 2014; Grigorev 2017]. The history of the field can be divided into two periods: before and after the general acceptance of continental drift (c.1960s) [Hedges 2006]. Most of the early researchers assumed that there was no change in the position of the islands and their water barriers, thus leaving only overwater mechanism of dispersal [Matthew, 1915; Simpson, 1956]. An alternative viewpoint was that ancient land bridges existed between the continents and the Antilles, permitting dispersal over land [Schuchert, 1935]. However, existence of the continental drift proved that the foundations for both views were incorrect [Hedges 2006].

Currently, it is known that the Greater Antilles once constituted a single geologic unit with North and South America in the late Cretaceous (c. 60-70 million years ago). This raised the possibility that most of the present fauna developed by “vicariance,” meaning that the species formed as they drifted on islands that broke off and continue moving away from the continents [Rosen, 1975]. This mechanism would predict existence of lineages that split before the Cenozoic, but neither fossil nor genetic research has been able to identify more than a few groups that satisfy this model (including *Eleutherodactylus* frogs, Cuban and Hispaniolan solenodons [Roca et al., 2004; Brandt et al., 2014; Grigorev et al., 2017], and the Cuban xantusiid lizard *Cricosaura typica*) [Hedges, 2006]. The virtual absence of the vicariance fauna can also be explained by a catastrophic event that occurred after the islands were separated. It

was proposed that the impact of the Cretaceous-Tertiary (K-T) asteroid (Chicxulub impactor) at 65 Mya probably devastated the ancient West Indian biota due to its very close proximity to that region (only 1-3 crater diameters away). Catastrophic immediate local effects, such as a megatsunami and massive hurricanes, together with draughts and sea level fluctuations occurring over the following years, would almost certainly have caused widespread extinctions of many organisms that might have formed on the Caribbean islands since the Chicxulub asteroid impact [Pindell 1994; Hedges 1996].

Since the early publications concerning Caribbean biogeography, it was noted, that the native fauna is missing many high-level groups, such as salamanders, caecilians, most families of frogs, lizards, snakes and turtles [Matthew 1915, Simpson 1956]. This underrepresented taxonomic composition gives support to the overwater dispersal being the major mechanism of colonization and subsequent speciation in the region [Hedges 2006]. The dispersal hypothesis is also supported by the fossil record, and by the existence of large radiations of species, belonging to the same genera, forming distinct island taxa by filling unoccupied niches on each island [Hedges 2006]. Therefore, most of the modern fauna probably arrived to the Caribbean by flying, swimming, or by hitchhiking on floatsome (e.g. tree logs, rafts of branches etc.). As a consequence, the formation of Caribbean fauna is not as much a result of the geologic factors, but rather depends on air and water currents, as well as the geographic proximity that can influence probability of dispersal in particular directions. For instance, the ocean currents in the Caribbean flow almost unidirectionally from southeast to northwest (Figure 3), making overwater dispersal much more likely from South America, than from Central or North America [Hedges 1996]. Therefore, the species that disperse

on floating vegetation, are most likely to disperse from South America and move along the island arch from southeast to northwest [Lameiro et al. 2012]. On the other hand, the volant species (birds and bats) show greater influence from the west and the north [Hedges 1996], and it has been postulated that, with the exception of the southern Lesser Antilles, the entire West Indian avifauna has arisen by dispersal from North and Central America [Olson 1978; Hedges 1996]. In some avian species, such as Galapagos finches, shiny cowbirds and others, this dispersal is ongoing and often makes it difficult to reconstruct steps of evolutionary history along the island chains [Lamichhaney 2015, Wayne 1984]. The general absence of long-distance migration among modern parrots (order Psittaciformes) and their relatively confined geographic ranges allows for reconstruction of ancient diversification patterns that are less obscured by dispersal than in many other orders of birds [Wright 2008].

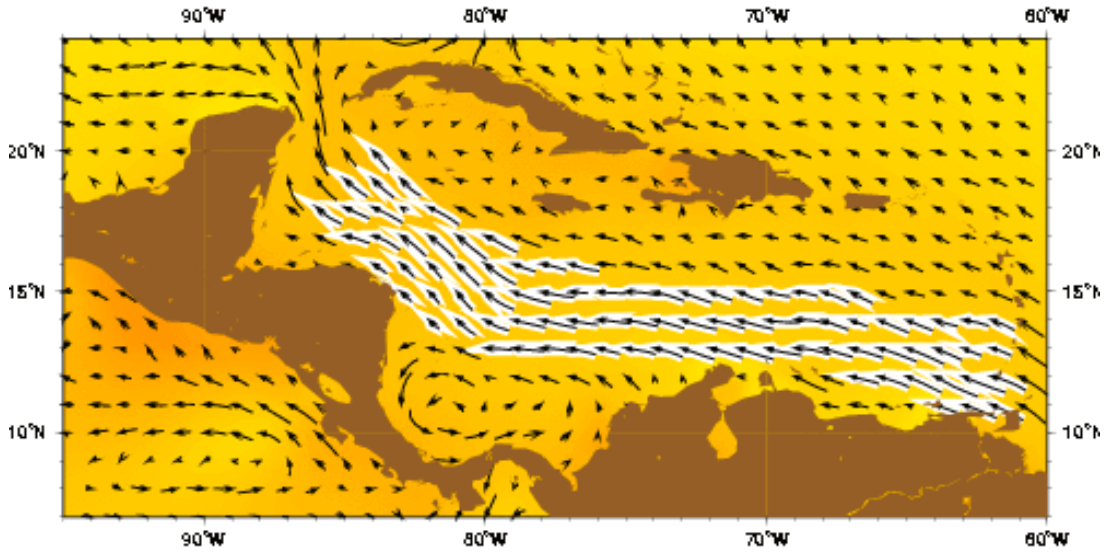


Figure 3. Ocean currents in the Caribbean (from Gyory J. 2005).

Closure of the Isthmus of Panama around 10Mya [Montes et al 2015] has allowed for dispersal of the South American terrestrial fauna into the Central and North America and, subsequently, into the Greater Antilles [Leigh et al 2013]. Periodic overwater dispersal of various vertebrates from the Central and North America to the Greater Antilles and within the archipelago has happened during periods of low sea level (e.g., late Miocene, ca. 10 Mya; latest Miocene, ca. 6 Mya; latest Pleistocene, 10 to 20 kya) when the Nicaraguan Rise, as well as other currently submerged areas in the region, was partially emergent, providing stepping-stone routes for colonization [Wright, Robinson 1993].

3. PARROT EVOLUTION

Based on their present distribution across the southern continents that formerly composed Gondwana, parrots have long been thought to have originated and initially diversified from this giant ancient supercontinent [Cracraft 2001]. Results of ancestral area analyses also support this hypothesis with Australasia being the likeliest region of origin [Wright 2008]. Biogeographic analyses based on phylogenies using multiple loci, extensive taxon sampling, and different analytical approaches, support a hypothesis of an origin and initial diversification during the Cretaceous, soon after the separation of the Africa and India/Madagascar block [Wright 2008].

The order Psittaciformes has long been recognized as a discrete and relatively homogenous group, but considerable uncertainty concerning its relationship to other avian orders has always existed. Various phylogenies produced no consistent placement of outgroups as sister to psittaciforms, reinforcing the idea that they have no

close sister relationship with other existing avian clades [Wright 2008]. The most recent phylogenomic analysis of 48 bird genomes placed parrots in the close proximity with passeriforms, the two orders (Passeriformes and Psittaciformes) together forming a group called “Passerimorphae” and falcons (Falconiformes) being its’ sister clade [Jarvis et al., 2014, Zhang et al. 2014].

Within the order Psittaciformes, the earliest split is between the NZ-restricted family Strigopidae (kakapo, kea and kaka) and all other parrots – 49-42 MYA, followed by divergence between the Australasian cockatoos (*Cacatuidae*) and remaining parrots (*Psittacidae*) – 45-39 MYA [Rheindt 2013], soon after the western half of Gondwana had separated into the continents of Africa and South America in early Cretaceous. The estimated molecular dates are too recent to support the vicariance mode of speciation, but it may be possible, that at the time, dispersal was still possible for some time between the Gondwanan landmasses of South America, Antarctica, and Australia. However, New Zealand, Africa, and Madagascar were slowly drifting away from each other [Teeling 2005].

Caribbean species of the Amazon parrots belong to the family Psittacidae, which is one of three families of true parrots. It consists of about 10 extant species from the Psittacinae subfamily (the Old World or Afrotropical parrots), and 157 of subfamily Arinae (the New World or Neotropical parrots), as well as several species that went extinct in the recent centuries [Joseph L. 2012]. The Arinae (the New World parrots) include, amazons, macaws, conures and parakeets, and have been separated from other groups of parrots for around 35 Mya, or approximately at the time when the major plates of the Gondwana drifted apart (Figure 4). After the emergence of the short-tailed

and the long-tailed New World parrots, which has been placed in the Eocene [Ottens-Wainright 2004], the split between *Amazona* and *Ara* + *Aratinga* clade occurred in late Oligocene 25-23 MYA, followed by the divergence between *Ara* and *Aratinga* in Miocene 10-12 MYA [Rheindt et al 2014].

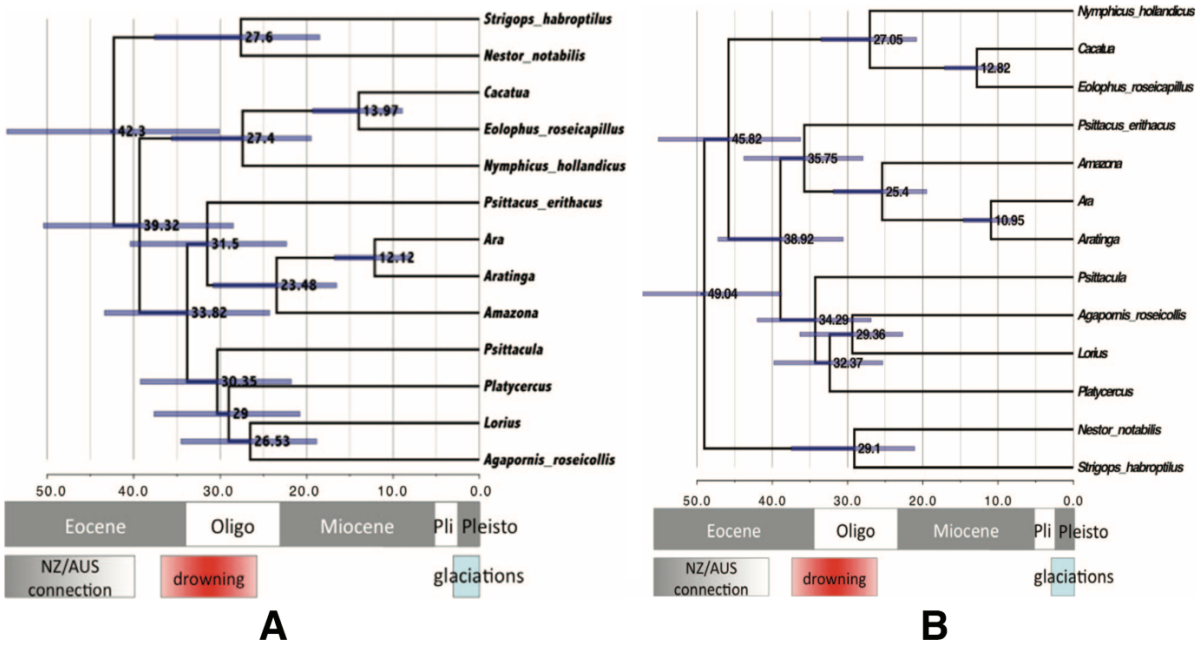


Figure 4. Dating of major psittaciform lineage divergence times using a 7-locus 13-genus sequence dataset and a calibration set based on mitogenomic studies (A) and a nuclear study (B) using independent fossil calibrations [from Rheindt et al 2014].

The neotropical genus *Amazona* includes 27–31 species widely distributed across the Americas continent from Argentina to Mexico, and occupy islands throughout the Caribbean [Forshaw, 1989]. As reported by the phylogenetic analyses described in Rusello & Amato (2004), the genus *Amazona* is not monophyletic: *Amazona xanthops* forms a well-supported sister group relationship with *Graydidascalus.* Unfortunately, the relationships within the short-tailed neotropical parrots are still not fully resolved, with

different analyses suggesting different sister groups to Amazona – either *Graydidascalus* – *A. xanthops* clade, or *Pionus* [Rusello & Amato 2004].

4. PSITTACID EVOLUTION ON THE CARIBBEAN ISLANDS

From 50 to 60 endemic species of psittacids (*Ara, Aratinga, Amazona*) lived on the Caribbean islands in the absence of humans, but only 12 species (3 of *Aratinga*, 9 of *Amazona*) survive today. In the past, each Antillean island once sustained at least one or two indigenous or even endemic species of macaw (*Ara*), and the same was true for the parakeets (*Aratinga*) [Williams 2001]. Today, only one threatened species, *Aratinga euops* can still be found in Cuba, while another one is in declining (*A. chloroptera*) on the Hispaniola and offshore islands. The only species not on the endangered list is *A. nana* in Jamaica. Some of *Amazona* parrots have done a little better, and five native species still survive in the archipelago (Cuban, Hispaniolan, Puerto Rican, yellow-billed and black-billed partots).

The Antillean Amazons have been divided into two main groups based only on plumage characteristics (coloration) and body size: the five medium-sized Greater Antillean species and the seven much larger Lesser Antillean species - 4 extant species and 3 extinct species [Snyder et al., 1987]. Nucleotide divergences analyses and geological data suggest, that *Amazona* colonized the Greater and Lesser Antilles by dispersal during the Pliocene (5.33 to 2.58 MYA) [Ottens-Wainright 2004]. To explain the evolutionary history of parrots in the Caribbean, two main routes were proposed: (1) from the northern South America for many of the Lesser Antillean species, and (2) from

North and Central America for Greater Antillean and some Lesser Antillean species [Hedges 1996, Rusello & Amato 2004].

Phylogenetic analyses show distinct divergence and independent histories of the smaller and mostly green Greater Antillean *Amazona* from the larger, more colorful Lesser Antillean species, and imply that they colonized the West Indies independently [Ottens-Wainright 2004, Rusello & Amato 2004]. The Lesser Antillean *Amazona* are a paraphyletic group, with at least two distinct lineages – *A. arausiaca* (Dominica) + *A. versicolor* (St. Lucia) vs *A. imperialis* (Dominica) and *A. guildingii* (St. Vincent) – both related to South American *Amazona* species [Rusello & Amato 2004]. The five Greater Antillean *Amazona* species belong to a single clade, including parrots endemic to Jamaica (*A. agilis* and *A. collaria*), Cuba (*A. leucocephala*), Hispaniola (*A. ventralis*) and Puerto Rico (*A. vittata*), that form a distinct, monophyletic group with the Central American *A. albifrons* [Rusello & Amato 2004] (Figure 5).

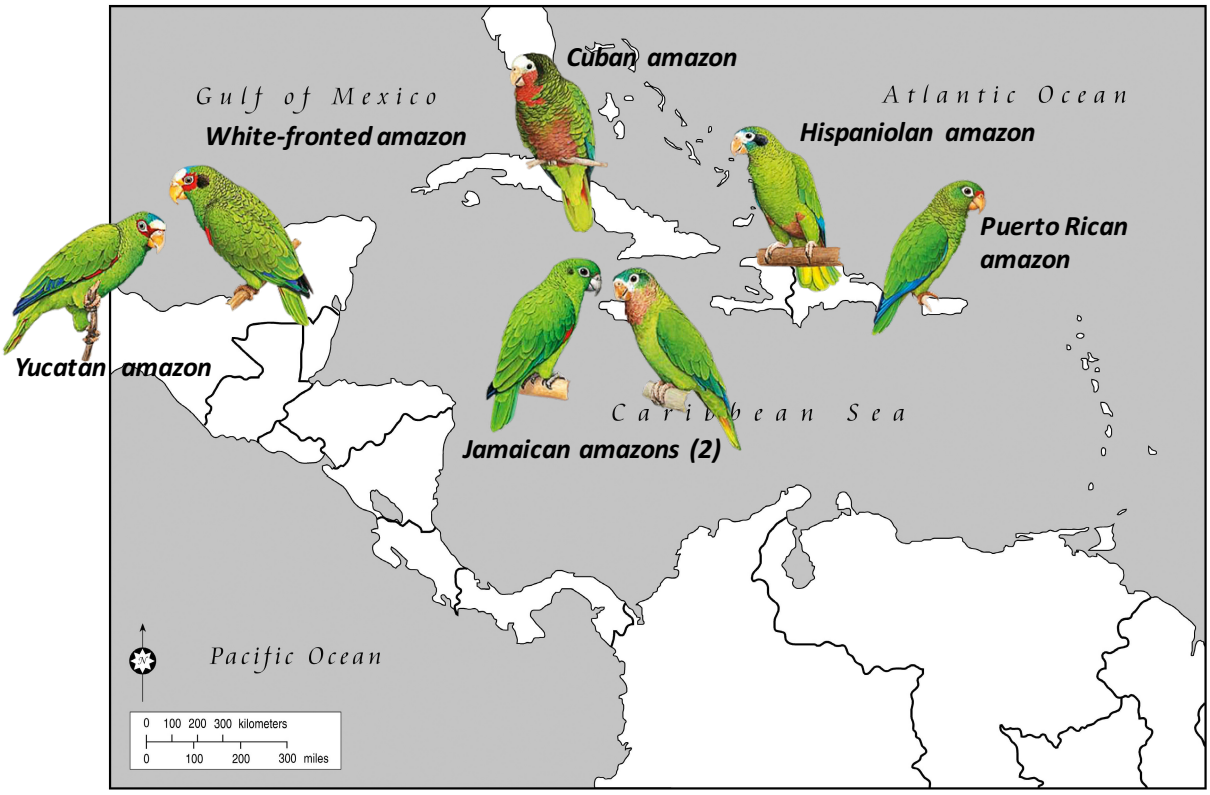


Figure 5. Distribution of surviving Amazon parrot species in the Greater Antilles. Cuban amazon (*A. leucocephala*) can be found between Cuba, Abaco (Bahamas), Isla de Pinos, Grand Cayman and Cayman Brac islands. In Jamaica, there are two species of amazons: the yellow-billed (*A. collaria*) and the black-billed (*A. agilis)*. Hispaniolan amazon (*A. ventralis*) survives on the island of Haiti (Hispaniola) and offshore islands. Finally, Puerto Rico is the home to the critically endangered Puerto Rican amazon (*A. vittata*). The entire clade has been suggested to descend from the ancestor related to the Central American *A. albifrons*. Conservation status of each species can be found in Table 1.

5. AMAZON PARROTS IN THE GREATER ANTILLES

All the Caribbean islands, except for Barbados and St. Croix were once inhabited by members of the genus *Amazona* [Williams and Steadman, 2001]. Most of the earlier attempts to figure out the phylogenetic relationships within the clade were based on morphology. Ottens-Wainright et al (2004) data supported Lack's (1976) view that *A. ventralis* and *A. vittata* are sister lineages that share a common ancestry with *A. leucocephala* (see Figure 6). According to this study, the most recent common ancestor gave rise to the two lineages that resulted in *A. collaria* and the clade formed by

A.leucocephala, A. ventralis, and *A. vittata*; *Amazona agilis* is basal to these Greater Antillean species. Diversification of *A. leucocephala* into its subspecies (found on Bahamas and Cayman islands) was roughly estimated to be sometime in the middle to late Pleistocene. However, Ottens-Wainright et al. (2004) study used only one molecular marker: partial cytochrome b gene sequence. A study from the same year by Rusello & Amato (2004), used a combined phylogenetic analysis of DNA sequence data from six partitions including mitochondrial (COI, 12S, and 16S), and nuclear (b-fibint7, RP40, and TROP) regions to investigate evolutionary history of the genus *Amazona* and established the relationships between most of the *Amazona* based on the molecular data. Unfortunately, the limitation of data prevented the resolution of a polytomy within the Greater Antillean clade, and it was suggested that further sampling of rapidly evolving regions within the mitochondrial genome, such as control region or ND4, may help to resolve the relationships within the *A. leucocephala–A. ventralis–A. vittata* radiation [Rusello & Amato, 2004]. The short internodes characterizing relationships among the major lineages of Amazon parrots may reflect a rapid rate of diversification during the evolutionary history of the group rather than merely a deficiency in character sampling [Rusello & Amato 2004].-The exact phylogenetic relationships within the group remain unknown (see Figure 7), probably because the radiation is relatively young, and speciation events were not well separated in time and/or because of some remaining gene flow. Resolving the phylogenetic tree of the Greater Antillean amazons remains a scientific question to be answered.

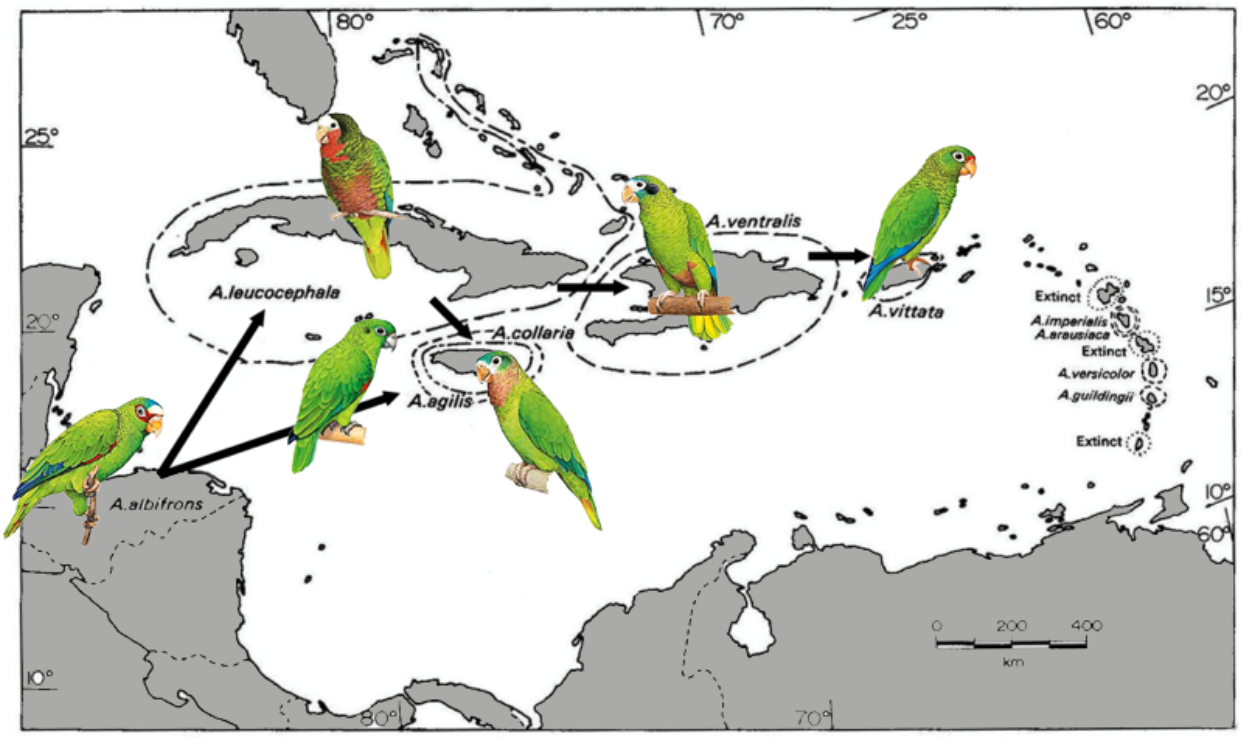


Figure 6. A hypothesized route of colonization of the Greater Antilles by the Amazon parrots (after Lack D. 1986)

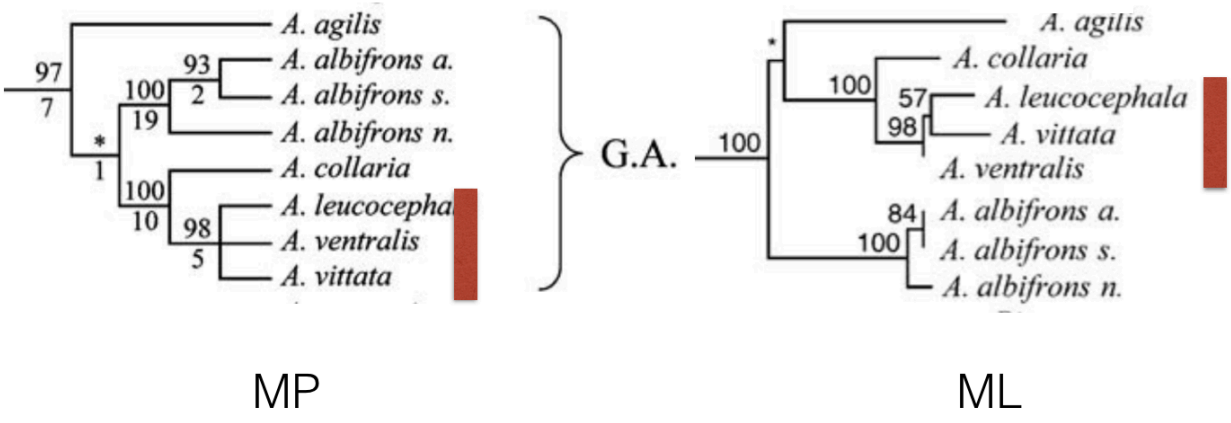


Figure 7. Partially represented MP and ML trees from the 2004 study of Rusello & Amato. Note the polytomy in the *A. vittata - A. ventralis - A. leucocephala* group and uncertain relative positions of the two Amazons from Jamaica: *A. agilis* and *A. collaria*

6. THE CRITICALLY ENDANGERED PUERTO RICAN AMAZON

Many of the *Amazona* species in focus of this study are endangered, threatened, or vulnerable (Table 1). Study of their full genomes, nuclear as well mitochondrial, can empower subsequent studies to estimate sequence variation, population diversity, inbreeding depression, demographic vital rates and viability. Studying the genomes and evolution of the species of interest can provide us with valuable information that can be essential for such management [Allendorf et al., 2010]. This is especially important in the case of the critically endangered Puerto Rican Amazon, the rarest parrot surviving in the wild [O'Brien 2012].

The critically endangered *A. vittata* is the only surviving endemic parrot species anywhere in the U.S. [Brinkley 2009]. Once abundant throughout the island of Puerto Rico, its drastic population decline followed the decimation of the old-growth forest [Snyder et al., 1987]. Recent demographic history inadvertently shaped the genome of *A. vittata* that should be highly invariable due to a recent population bottleneck [Brock and White, 1992]. Recently the 1.58Gb genome of *A. vittata* was sequenced and the data was made publically available, which is useful for genome studies that can focus on conservation and comparative genomics [Oleksyk et al., 2012a, Oleksyk et al., 2012b, Afanador et al., 2014].

The cornerstone of the recovery plan for the critically endangered Puerto Rican parrot (*Amazona vitatta*) is an actively managed, long-term captive breeding and

reintroduction program [Earnhardt 2014]. Before the Hurricane Maria, the species was numbered at less than 500 individuals in its total population, and have been in the focus of an intensive conservation effort since the 1970's. Puerto Rican Parrot Recovery Program (PRPRP) that conducts the management of the species, has included both breeding in captivity and releasing captive-bred individuals into the wild. Currently, managers maintain four distinct sub-populations of this species, two in captivity and two in the wild. The first captive population was created in 1973 within the El Yunque National Forest (EYNF), at that time there were no birds left in the wild. This breeding facility was populated by harvesting all the individuals from the only extant population of Puerto Rican amazons, a relic population located 9 km away in a remote region of the same forest. Some parrots were released to the wild in El Yunque in the 1980's. However, many of them died from Hurricane Hugo, a category 4 hurricane that passed by the eastern tip of the island in 1989, and twenty years after the first captive population was created, a subset of captive parrots was strategically transferred to a second captive-breeding facility located about 100 km away within the Rio Abajo State Forest (RASF). Finally, in 2006, parrots were released from this second breeding facility into the surrounding forested area resulting in the creation of a second population of wild parrots. To date, these four populations represent the entirety of the known Puerto Rican amazon, except for a handful of documented small dispersal events occurring largely in the vicinity of the RASF [White 2014].

Captive breeding programs, especially in case of very low number of founder individuals (such as in this case - only 11 Puerto Rican parrots existed back in 1975,

when they were captured and the breeding program was launched [Jafet Velez, PR Recovery Avicuturist, personal communication]), require careful and scientifically justified management. They must consider not only demographic factors: population growth, generation time and age/sex structure, but also the genetic diversity and genetic structure within the captive population, to avoid inbreeding depression, and to decide which individuals to release to successfully form new wild populations.

Currently the Puerto Rican parrot is confined to small areas on the main island of Puerto Rico, with most of the population surviving in the captive program ran by the FWS [Snyder 1987]. However, historically this species occupied the entire island of Puerto Rico, and formerly its populations also existed on the nearby Culebra, Mona and, possibly, Vieques islands [Forshaw 2017]. Two subspecies of *Amazona vittata* are currently recognized:

1) *A. v. vittata* (Boddaert) – the nominate subspecies, still exists in Puerto Rico and may have formerly (until early 20th century) occurred on the nearby Vieques Island [Forshaw 2017]
2) *A. v. gracipiles* (Ridgway) – this poorly differentiated subspecies formerly existed in Culebra (last record in 1899), but is now extinct.

It is not known, whether the extinct population of Vieques island was isolated and constituted a separate subspecies of its own, or if the birds migrated seasonally from Puerto Rico [Sorrié 1975, Forshaw 2017].

Despite some of the early efforts of DNA fingerprinting [Brock and White, 1992],

the genetic consequences of the severe population bottleneck as well as the population expansion associated with the recent recovery have not been fully evaluated, and a more comprehensive genetic analysis at the population and species level is needed. This has recently become possible, since the samples of almost the entire surviving population of the species along with their detailed pedigrees rising to a dozen founders in 1972 are obtainable through the collaboration with Department of Natural Resources (*DNER*), *Compania de Parques Nacionales*. A detailed study of *A. vittata* population variation can provide an excellent model for conservation genomics to study inbreeding depression, mutation, and adaptation to captivity.

Table 1. Conservation status of the surviving amazon parrot species in the Greater Antilles and Yucatan peninsula.

Species	**Distribution**	**Conservation (IUCN) status**	**Reference**
Amazona vittata (Puerto Rican amazon)	Archipelago of Puerto Rico	Critically Endangered	White 2016
Amazona ventralis (Hispaniolan amazon)	Hispaniola	Vulnerable	Kirwan 2016
Amazona leucocephala (Cuban amazon)	Cuba, Bahamas, Cayman islands	Near Threatened	Kirkconnell 2017
Amazona albifrons (white-fronted amazon)	Central America and Mexico	Least Concern	Butchart 2016
Amazona xantholora (Yucatan amazon)	Belize, Honduras, and Mexico	Least Concern	Butchart 2016
Amazona collaria (yellow-billed amazon)	Jamaica	Vulnerable	Koenig 2016
Amazona agilis (black-billed amazon)	Jamaica	Vulnerable	Koenig 2016

7. REVIEW OF THE CURRENT STATUS IN PARROT GENOMICS RESEACH AND DATA AVAILABILITY

I. AVAILABLITY OF THE WHOLE GENOME DATA

In the last decade, a large effort has been made to sequence and publish genome data for vertebrate species, especially within the Genome 10K project [Koepfli et al, 2015], and the recent 48 avian genomes papers [Jarvis et al., 2015]. However, the quality of genome assembly and therefore the usefulness of genomes for comparative studies still vary to a great degree. The most well-studied avian genomes are those of chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*) and zebrafinch (*Taeniopygia guttata*). The estimated time since divergence for members of Psittaciformes and the chicken (*G. gallus*; Galliformes) is approximately 122–125 MYA, whereas Psittaciformes and Passeriformes (Zebra finch; *T. guttata*) diverged more recently - 78–119 MYA [Seabury 2013]. More recently, some research efforts started to provide genomic insight into the Psittaciforms with the sequencing of the scarlet macaw (*Ara macao*) [Seabury 2013], budgerigar (*Melopsittacus undulatus*) [Ganapathy 2014] and New Zealand parrot Kea (*Nestor notabilis*) [Zhang 2014]. The budgerigar assemblies are comparable to or better than the chicken and zebra finch genome assemblies built from traditional Sanger sequencing reads [Ganapathy 2014]. Data from these studies can be used for the comparative genomics analyses of other representatives of this unique but underserved biological group and understanding the basis of its key features, such as intelligence and longevity [Seabury 2013]. The genome of the Puerto Rican parrot was previously sequenced and a draft assembly with 76% coverage and approximately 26.89x average

coverage depth was achieved [Oleksyk et al 2012a, 2012b]. The assembled sequences provide a starting point towards completing and annotating a draft genome sequence.

Since the Greater Antillean clade species are very closely related, their genomes can be studied in direct genome comparisons. Learning about their demographic history, finding which gene families have expanded or contracted while the new species formed and adapted to their environment, which genes were under selective pressure will allow us to understand better the underlying mechanisms of species formation and adaptive radiation.

For the purposes of population studies and monitoring the recovery of the critically endangered Puerto Rican parrot a set of 18 highly polymorphic microsatellite markers was developed using the genomic sequence of *A. vittata* and verified in a sample of the population [Afanador 2014]. These markers cross-amplify in the closely related species from the nearby islands of Cuba and Hispaniola (*A. leucocephala* and *A. ventralis*) and can be also used for populations assays on those species [Afanador et al., 2014].

II. AVAILABILITY OF THE mtDNA DATA

The mitochondrial genomes of the endangered yellow-shouldered amazon from the Lesser Antilles (*Amazona barbadensis*) [Urantowka 2013], the Hispaniolan amazon (*Amazona ventralis*) [Urantowka 2016], and *Amazona ochrocephala* [Eberhard & Wright 2016] were sequenced. *Amazona barbadensis* was confirmed to have the same mitochondrial gene arrangement as previously reported for other members of the genus - *A. farinosa, A. oratrix* and *A. auropalliata* [Urantowka 2013, Eberhard 2001].

Mitochondrial genome rearrangements that result in control region duplication have been described for a variety of vertebrates, including numerous birds. Control region duplications have arisen independently at least six times across the order Psittaciformes [Eberhard 2016] (see Figure 7). In all cases of such rearrangement, two copies of the control region are present and are highly similar at the sequence level, whereas duplicates of the flanking genes also included in the rearrangement either show signs of degeneration or are lost completely [Eberhard 2016].

Eberhard and Wright (2016) suggest that duplicated control regions are maintained in a functional state in many parrots because they serve to increase the number of initiation zones for mitochondrial genome replication and thus counteract an otherwise unusually slow rate of mitochondrial DNA replication in this avian group (Figure 8). This duplication has been creating difficulties in amplification and sequencing of the mtDNA for the population studies (Oleksyk, personal communication).

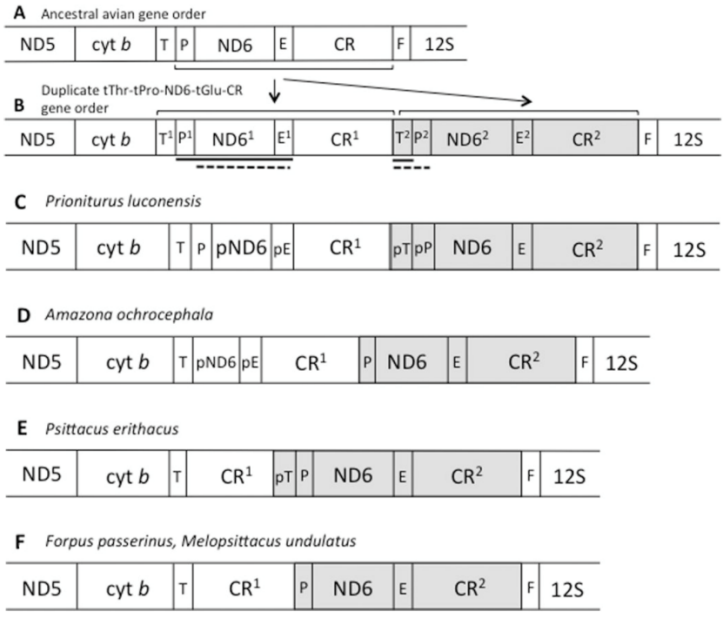


Figure 8. Mitochondrial gene orders for different parrot species. **A**: ancestral gene order. **B**: Gene order resulting from a tandem duplication of the Threonine tRNA gene to the CR segment of the mitochondrial genome, and thought to be ancestral to the gene orders shown below. **C–F**: Gene orders flanking the control region (CR) for five parrot species with duplicated control regions, each representing an independent origin of the duplicate control region state. Transfer RNAs are indicated by their single-letter abbreviations, E (Glutamyl), T (Threonine), P (Proline) and F (Phenylalanine). Pseudogenes are indicated with a "p" preceding the gene's abbreviation. The gene order shown in A is that resulting from a tandem duplication and which is the hypothesized precursor to the gene orders shown in B–E. Shading indicates hypothesized homology between highlighted regions. (Modified from Eberhard & Wright [2016])

8. OBJECTIVES

We set out to sequence and annotate all the mitochondrial genome assemblies for the Greater Antillean *Amazona* parrots. Given the new data, we tried to resolve standing issues that exist in the phylogenetic relationships within the group. At the same time, we set out to evaluate the migration hypotheses about speciation across the Greater Antillean islands. Finally, we contribute new mitochondrial markers that can be used in the future to determine identity, population genetics, and drive ongoing conservation efforts in this highly charismatic endemic species of birds.

METHODOLOGY

1. Sample collection and DNA extraction

This study has been reviewed and approved by the Institutional Animal Care and Use Committee of the University of Puerto Rico. All the required collection and export permits issued by the US government under the Endangered Species Act (ESA), Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), by the Animal and Plant Health Inspection Service (APHIS) had been obtained before any work was started. Origins of the contemporary DNA samples used in this study are listed in Table 2. DNA was extracted from tissue samples listed in the Table 2 using DNEasy kits (Qiagen Inc., Valencia, CA) following the manufacturer's protocol.

Table 2. Origins of the DNA samples in this study. See distribution of parrot species in Figure 1.

Species	Sample origin	Sample type	Sequencing methods
A. vittata	Rio Abajo PR	blood	Illumina PE, MP + PacBio
A. vittata	Vieques PR (1860s)	feather	Illumina PE
A. ventralis	Rio Abajo PR	blood	Illumina PE, MP
A. leucocephala	San Diego Zoo	skin graft	Illumina PE, MP
A. rhodocorytha	from collaboration with aviculturist society (pet)	feather	Illumina PE
A. albifrons	from collaboration with aviculturist society (pet)	feather	Illumina PE
A. agilis	Hope Gardens & Zoo, Kingston Jamaica	blood	Illumina PE
A. collaria	Hope Gardens & Zoo, Kingston Jamaica	blood	Illumina PE

In addition, we have obtained DNA from a private collection in Germany. This sample of *Amazona vittata* originates from the island of Vieques and was received by the private collector in Kassel (Germany) in 1875. As transporting the museum sample from an endangered species skin would be difficult under the CITES convention, DNA has been isolated using the protocol for ancient DNA extraction from bones and teeth [Rohland and Hofreiter, 2007] and sequenced by Illumina PE at the laboratory Dr. Michael Hofreiter at the Institute for Biochemistry and Biology in Potsdam (Germany).

2. DNA sequencing

- **Illumina sequencing**

In the Illumina© sequencing process DNA templates are immobilized on a flow cell surface designed specifically to present them in a way that increases accessibility for enzymes while providing stability of the surface-bound template and low rates of non-specific binding of fluorescently labelled nucleotides. Solid-phase amplification creates up to a thousand identical copies of every single template molecule nearby (one micron or less). Fluorescently-labelled nucleotides are used to sequence in parallel all the clusters on the flow cell surface in the sequencing by synthesis (SBS) technology. Every sequencing cycle, one labelled dNTP (deoxynucleoside triphosphate) is added to the growing nucleic acid chain. The nucleotide label performs a function of a terminator for the polymerization process, so after incorporation of each dNTP, in order to identify the base the fluorescent label is imaged and then, to allow incorporation of the next nucleotide, it is enzymatically cleaved. The labelled desoxynucleotides are all present

as separate molecules, so that incorporation bias is minimized because of competition. During each cycle base calls are made directly from signal intensity measurements, which, compared to other technologies, significantly reduces raw error rates. The highly accurate base-by-base sequencing gets rid of sequence-context specific errors, enabling reliable base calling even in highly repetitive regions and within homopolymers [Bentley 2008].

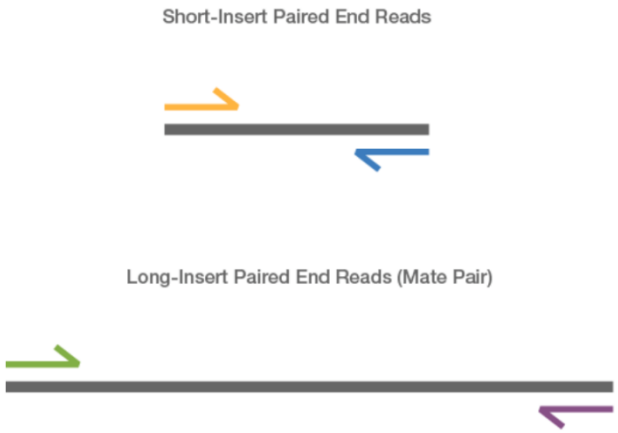


Figure 9. Short-Insert Paired End Reads and long-insert paired end reads (Mate Pair). When sequencing, DNA is chopped into small fragments, and then adaptors are ligated. Then, for single end (SE) only one end of a DNA fragment is sequenced. For PE (and MP as well) both ends of the same fragment are sequenced. Mate pair allows to have read pairs much further apart, which can be more informative than the standard paired-end protocol.

Highest coverage of the genome for *de novo* assembly can be achieved by using a combination of short and long insert sizes with paired-end sequencing (Figure 9). Since larger inserts can pair reads over greater distances, they enable a higher ability to read through regions of large structural rearrangements as well as highly repetitive sequences. Smaller inserts sequenced at higher depths enable to fill in gaps that were missed by longer inserts sequenced at lower depths. Thus, a mixed library of short and

long insert sizes results in higher quality de novo assembly with fewer gaps, larger contigs, and more accurate resulting consensus sequences [Bentley 2008].

Genomes of *A. vittata*, *A. ventralis*, *A. leucocephala, A. albifrons* and *A. rhodocorytha* were sequenced using Illumina Paired End and Mate Pair libraries with different insert sizes. Parameters of the sequencing output are presented in Table 3. For the two Jamaican species – *A. agilis* and *A. collaria* – Paired End Illumina reads were prepared using the 10X genomics protocol.

Table 3. Sequencing outputs for each species

Species	Insert size	Original		Filtered	
		Total Reads	Total Bases	Total Reads	Total Bases
A. vittata	200 bp	267,262,276	26 Mb	201,766,246	20 Mb
A. vittata	3 kbp	153,326,830	15 Mb	83,201,110	8 Mb
A. vittata	5 kbp	51,108,944	5 Mb	28,090,686	2 Mb
A. vittata	8 kbp	385,065,868	38 Mb	283,916,716	28 Mb
A. leucocephala	300 bp	387,336,118	39 Mb	323,253,532	33 Mb
A. leucocephala	3 kbp	88,772,420	8 Mb	49,510,820	4 Mb
A. ventralis	300 bp	408,883,170	50 Mb	378,789,678	32 Mb
A. ventralis	3 kbp	100,678,456	13 Mb	87,567,232	9 Mb
A. rhodocorytha	400 bp	149,097,260	15 Mb	115,678,341	11 Mb
A. albifrons	400 bp	154,323,442	16 Mb	120,391,034	12 Mb

- **PacBio Sequencing**

PacBio© (Pacific Biosciences) uses proprietary Single Molecule RealTime (SMRT) technology, detects nucleotide incorporation events in real time while the elongation of the replicated strand from a single-stranded template is taking place (see Figure 10). The nucleotides that are used in the SMRT technology are labelled not on the base, but on the phosphate chain. The label (fluorophore) is released and dissipated when the phosphate bond is cleaved during the DNA synthesis and therefore the incorporated labelled nucleotides can be identified [Rhoads 2015].

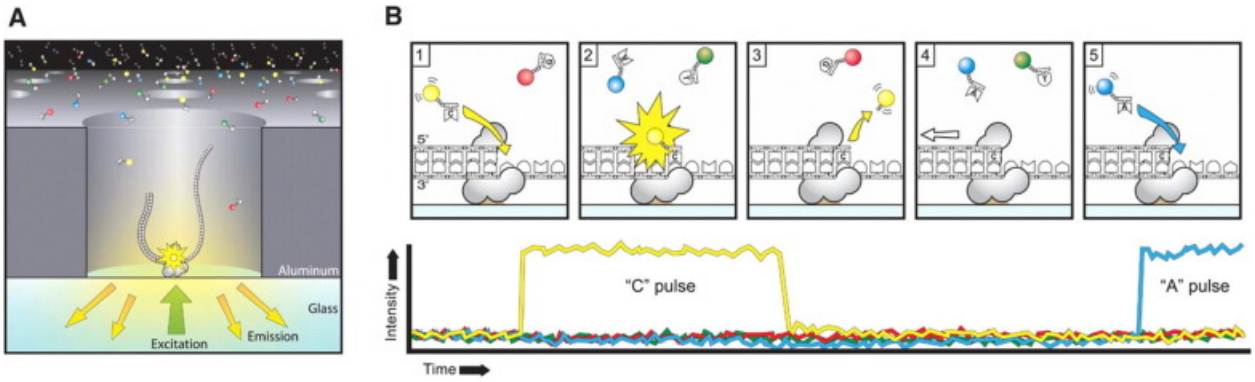


Figure 10. Sequencing via light pulses. **A**. A SMRTbell (gray) diffuses into a zero-mode waveguide, and the adaptor binds to an immobilized polymerase enzyme. **B**. Each of the nucleotides is labeled with its own fluorescent dye (G, C, T, and A - red, yellow, green, and blue) n order for them to have different emission spectrums. While nucleotides are kept inside of the detection volume by the polymerase, a light pulse, which identifies the nucleotide, is produced. (1) A fluorescently-labeled nucleotide joins with the template inside of the polymerase active site. (2) Fluorescence of the color identifying the incorporated nucleotide (yellow for C in the picture here) is emitted. (3) The dye-linker-pyrophosphate product is cut from the base and it diffuses out of the zero-mode waveguide, finishing the fluorescent pulse. (4) The enzyme moves to the adjacent position. (5) The next base joins with the template in the active site of the polymerase, the next fluorescence pulse is therefore initiated (here - base A). The figure is from Rhoads 2015.

PacBio technology provides us with very long reads (average > 10,000 bp, some reads > 60,000 bp) which can be used to close gaps in current reference assemblies

and characterize structural variation (SV) in genomes, enabling highly-contiguous de novo assemblies (Rhoads 2015). A DNA sample extracted from blood of Puerto Rican parrot (*A. vittata*) was sequenced using the PacBio technology yielding around 4Gb of long reads with 2x coverage. This sequence has proven crucial providing support for resolving the repeat structure in the mitochondrial genome of *A. vittata*.

3. Analysis

- **PacBio Methods**

Raw PacBio® reads were corrected with LoRDEC 0.5.1 [Salmela 2014] with k=17 and k=23 and other parameters set as default. Corrected bases were shown in upper case, to estimate correctness rate we counted the number of upper case nucleotides and divided by total number of nucleotides. For k=17 the correctness rate is 89.41% and for k=23 the correctness rate is 21.8%. For the current analysis, we chose corrected reads with better correction rate. PacBio coverage was computed as a total number of nucleotides divided by the expected genome size (1.58 Gb) and this is equal to approximately 2x.

To extract raw mtDNA sequences from the corrected PacBio reads we used Cookiecuter software [Starostina 2015] with 23-mers generated from draft mtDNA assembly. Cookiecutter extracted 11 reads and 7 of them were successfully mapped to the draft assembly and covered it in full with 1x coverage.

- **Mitochondrial Genome Assembly and Annotation**

a) **mtDNA Assembly**

We have found that mitochondrial genome of *Amazona vittata* has a very GC-rich simple sequence region, which was impossible to assemble with usual approaches due to low coverage of Illumina reads and the fact, that it has a nearly exactly duplicated control region, according to length of PCR products. To overcome this complexity of parrot mtDNA control region we developed a naïve approach based on supervised reference-assisted de Bruijn graph walkthrough, implemented in Python language. We used Jellyfish 2 [Marçais 2011] to compute 23-mers frequencies in raw Illumina reads, k-mers were mapped to the closest available high-quality avian genome (*Melopsittacus undulatus*) with kmer_cov_for_fasta.py script. The first 23-mer with exact match to 12S rRNA genes in *M. undulatus* mtDNA genome was used as a starting point for extension in both directions. Python scripts *kmer_continue_left.py* and *kmer_continue_right.py* extend a given k-mer with respect to the difference in genomic and mtDNA 23-mer coverages up until bifurcations, which occur due to repeats in mtDNA. To resolve the fork paths human intervention was required: we had to check every path extension from the fork and choose the correct one, which does not disagree with PCR sizes, Illumina reads and PacBio reads. We found that this is a feasible task due to the small size of mtDNA and a relatively small number of repeats. Assembled mtDNA was verified: 1) by consistency with PacBio reads mapped with Blasr tool [Chaisson 2012] to the assembled mtDNA; 2) length consistency with PCR products; 3) consistency of Illumina reads coverage.

Mitochondrial genome assemblies for other parrot species (*A. ventralis, A. leucocephala, A. albifrons, A. rhodocorytha*) were also performed using the described above method, only without the PacBio dataset, since it was only available for the Puerto Rican parrot. The resulting assemblies do not contain the duplicated control region due to very high similarity between the two sequences.

In order to assemble mitochondrial genomes of Jamaican amazons – *A. agilis* and *A. collaria* – a blast db was created using the NCBI-BLAST-2.3.0+ [Altschul 1990] *makeblastdb* command from Supernova [Weisenfeld 2017] assembly of paired-end Illumina reads prepared using the 10X genomics protocol. *Amazona vittata* mtDNA sequence was queried against the blast database using blastn [Chen 2015], and sequences with hits were extracted from the assembly.

b) "Ancient" mtDNA Assembly

In order to assemble the low-coverage mitochondrial genome of the extinct *A. vittata* population from Vieques island, the following steps were taken. First, the Illumina reads (1.62 Gb, 200 bp insert size PE) were trimmed with Cookiecutter [Starostina 2015]: Illumina adapters were removed, very short sequences were dropped, as a result of which 76% of all sequences were discarded. Remaining reads were mapped with bwa mem [Li 2009 b] to the assembled reference mtDNA of *A. vittata*. Samtools [Li 2009 a] *mpileup* was employed to generate a consensus sequence of all the mapped reads. Each position was then processed semi-manually in accordance with the following rules: 1) if the nucleotide is different from the reference and coverage is 1 – change to

N; 2) if there's no coverage – change to N; 3) if the nucleotide differs from the reference and coverage is > 1 – check it manually and change for an alternative allele, if it's a valid variant. As a result, we have acquired an assembly of the complete mitochondrial genome of this extinct population, with exception of the duplicated control region. Thirteen nucleotide positions, in which the assembled coding regions in mtDNA of the parrot from Vieques island differs from the "core" Puerto Rican parrot, were found.

c) Mitochondrial Genome Annotation

In order to annotate protein-coding genes (PCGs) we first identified all the open reading frames (ORFs) with a custom script implemented in Python. Afterwards we searched for all possible ORFs in the previously assembled nucleotide sequences against the NCBI nucleotide database [National Center for Biotechnology Information 2018]. Non-coding tRNA genes were annotated with tRNAscan-SE [Lowe 2016].

- **Phylogenetic Methods**

BEAST 2 is a cross-platform software for Bayesian phylogenetic analysis of nucleotide and amino acid sequences [Bouckaert 2014]. It is used to estimate time-measured, rooted phylogenies using either strict or relaxed molecular clock models. The software implements relaxed clocks, non-parametric coalescent analysis, multispecies coalescent inference, phylogeography and more. It can be employed as a method of reconstructing phylogenies but at the same time it is a framework for testing evolutionary hypotheses without conditioning on a single tree topology [Bouckaert 2014]. BEAST 2 uses Markov Chain Monte Carlo (MCMC) algorithm to average over

tree space, so that every tree is weighted proportionally to its posterior probability [Bouckaert 2014]. The software package includes a graphical user-interface (GUI) for setting up custom standard analyses (BEAUti 2) and a suit of post-processing tools – LogCombiner, TreeAnnotator and DensiTree.

In order to figure out the phylogenetic relationships within the group of Greater Antillean Amazona we are using mtDNA sequences from the following species: *A. vittata, A. ventralis, A, leucocephala, A. albifrons, A. agilis, A. collaria, A. rhodocorytha, A. aestiva, A. ochrocephala, A. barbadensis, Ara macao, Ara militaris, Ara ararauna, Ara severus, Aratinga mitrata, Aratinga rubritorquis.* Partitionfinder 2 [Lanfear 2017] was used to select best-fit partitioning schemes and models of evolution for phylogenetic analyses with BEAST 2. We selected the following 4 partitions of the mtDNA to assess phylogeny in *Amazona* parrots:

1) 13 protein-coding genes (PCGs) 1st+2nd position;
2) 13 PCGs 3rd position;
3) 12S rRNA + 16S rRNA;
4) 20 complete tRNAs.

Base substitution models for each partition are presented in Table 4.

Table 4. Best partitioning scheme and base substitution models for parrot mitochondrial genomes acquired by Partitionfinder.

Partition Names	# sites	Best Model
PCGs 1st and 2nd position	7258	TRN + G
PCGs 3rd position	3629	GTR + I
16S, 12S rRNA	2560	TIM + G
tRNAs	1331	TVM + I

Due to scarcity of fossil record in the Psittacidae clade it was not possible to find any primary calibration priors that could be used for this phylogenetic analysis, therefore we had to employ secondary calibration priors. Secondary calibration priors for the BEAST 2 analysis were taken from the Rheindt et. al 2014 paper – 12.12 Mya split between *Ara* and *Aratinga* and 23.48 Mya split between *Amazona* and *Ara* + *Aratinga* (see Figure 3).

The evolutionary history was also inferred from all coding sequences of mtDNA using the Maximum Likelihood method based on the Tamura-Nei model [Tamura 1993]. Initial tree for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.2013)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 0.0010% sites). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 16 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 14544 positions in the final dataset. 500 bootstrap replications were performed [Felsenstein1985]. This evolutionary analysis was conducted in MEGA7 [Kumar 2016].

Additionally, evolutionary history was inferred from cytochrome B (CytB) sequences using the Maximum Parsimony method and the Maximum Likelihood

method based on the Tamura-Nei model [Tamura 1993]. These evolutionary analyses were conducted in MEGA7 [Kumar 2016]. The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (pg. 126 in ref. [Nei 2000]) with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 17 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 946 positions in the final dataset. Five hundred bootstrap replications were performed [Felsenstein1985].

In the Maximum Likelihood method, initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 5.9736)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 34.5137% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 17 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 946 positions in the final dataset. Five hundred bootstrap replications were performed [Felsenstein1985].

RESULTS

1. Assembly and annotation

The combined approach, which incorporated both short Illumina and long PacBio reads allowed us to assemble the complete mitochondrial genome of *A. vittata* including the nearly exactly duplicated control region, which was impossible to assemble using traditional methods. We have annotated 23 tRNAs and tRNA pseudogenes, 14 protein coding genes and pseudogenes, the duplicated control region, repetitive sequences and rRNAs. The resulting map of *A. vittata* mitochondrial DNA, including genes and other features annotation, coverage by Illumina and PacBio reads and PCR products is presented in Figure 13. The illustration was created using Circos [Naquin 2014].

Mitochondrial genomes of *A. ventralis*, *A. leucocephala*, *A. albifrons, A. collaria, A. agilis* and *A. rhodocorytha* were assembled with the exception of the duplicated control regions due to their high level of sequence similarity and the lack of PacBio sequences from all species but *A. vittata*. Alignment of all assembled sequences is presented in Figure 12.

Mitochondrial genome of the extinct Vieques population of *A. vittata* was assembled, excluding the duplicated control region. After aligning and comparing the nucleotide sequences of the two possible subspecies (see Figure 11), we found that mtDNA of the parrot from Vieques island differs from the “core” Puerto Rican parrot in at least 13 positions within the coding regions, that are listed in Table 5.

Table 5. Distinct positions in coding mtDNA of the two *A. vittata* specimens from different islands.

Position	„Core“ PR parrot	Vieques parrot	Gene
74	A	C	12S_rRNA
458	G	A	12S_rRNA
892	A	G	12S_rRNA
1462	A	G	16S_rRNA
2566	T	C	16S_rRNA
7999	A	T	ATP6_Prot
9941	T	C	ND4L_Prot
10666	C	T	ND4_Prot
13810	A	G	CYTB_Prot
15190	T	C	CR_control
15239	A	G	CR_control
15255	C	T	CR_control
15688	G	A	CR_control

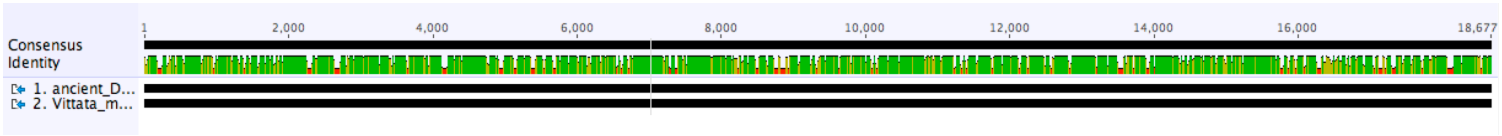


Figure 11. Alignment of the mitochondrial genomes of “core” *A. vittata* and ancient DNA samples obtained from Vieques island (“ancient DNA”), excluding the duplicated control region.

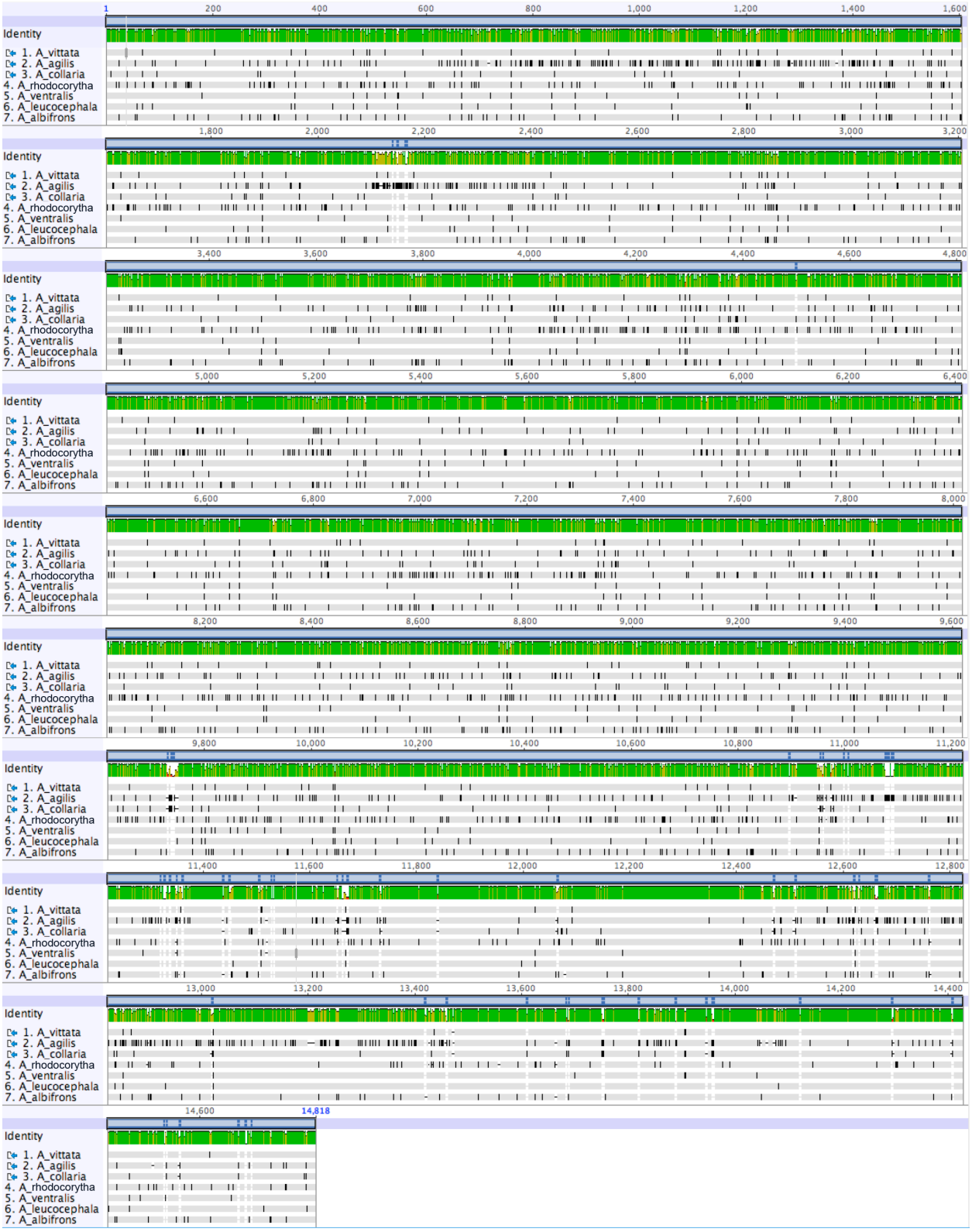


Figure 12. Alignment of all assembled parrot mitogenomes, excluding the duplicated control region

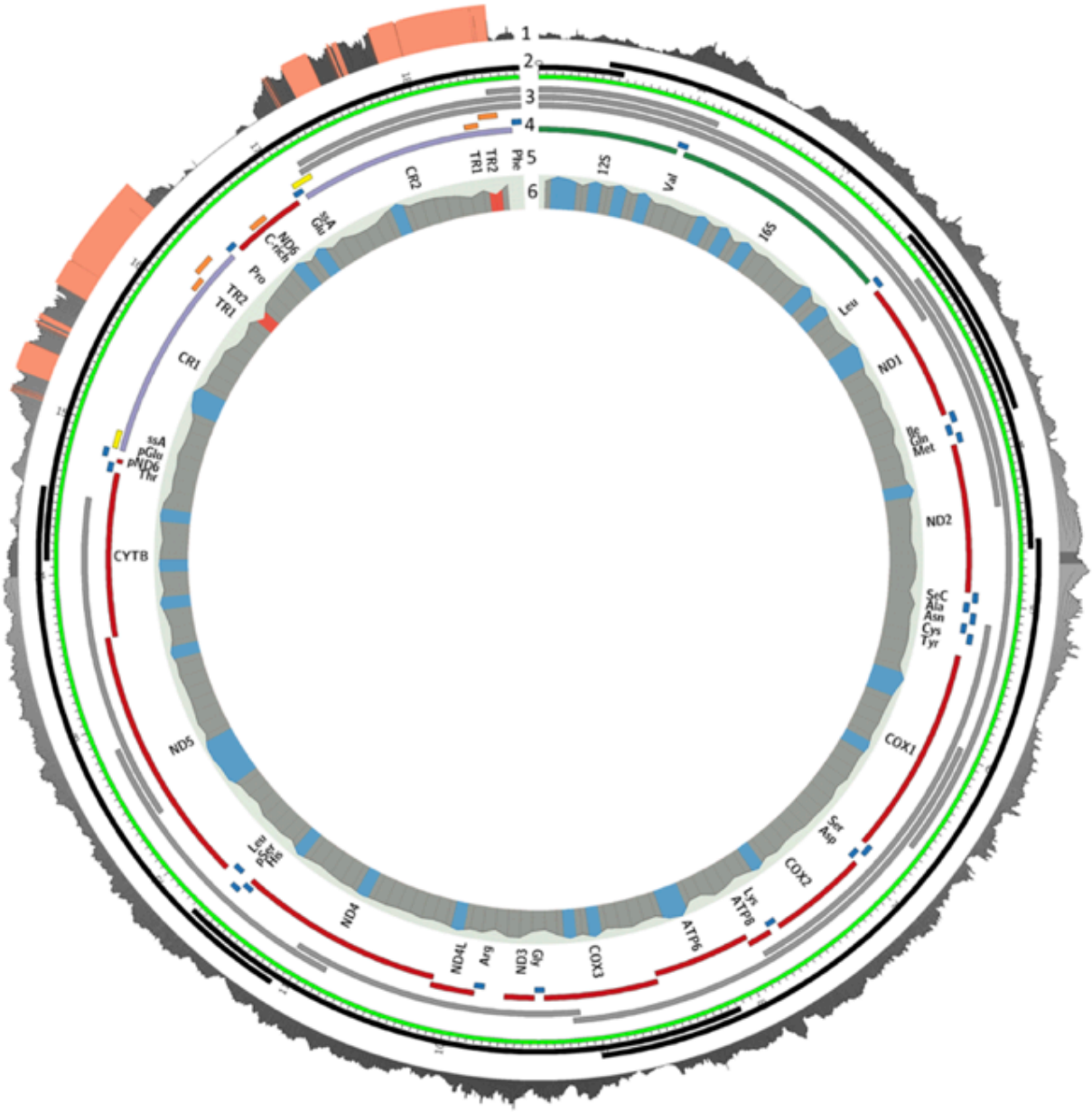


Figure 13. *A. vittata* mitochondrial genome assembly map. (1) Coverage by Illumina reads. Red if coverage > 140. (2) Mapped PacBio reads. (3) Mapped PCR products used for assembly validation. (4) Gene and features annotation. Red – protein coding genes, Green – rRNA, Blue – tRNAs, Purple – control regions, Orange – tandem repeats, Yellow – ssA. (5) Gene names. 6) GC content. Red if GC < 30% and Blue if > 50%.

2. Phylogenetics results

Bayesian phylogeny of all coding mitochondrial loci

Bayesian phylogenetic analysis resulted in topologies that were consistent with the prior knowledge about phylogenetic relationships and node ages. This study is the first one to have resolved phylogenetic relationships between *A. vittata, A. ventralis* and *A. leucocephala* - our Bayesian analysis that included all coding mitochondrial loci has resolved the node with posterior probability of 0.9983 (see Fig. 14), confirming *A. leucocephala* to be the sister group of *A. ventralis* + *A. vittata.* We have also confirmed that *A. collaria* is the closest living relative of all three and that *A. agilis*, in its' turn, is not as closely related to any of them as *A. albifrons*.

According to our estimates, the split between South American and Central American + Caribbean amazon parrots has occurred 5.5 – 10.5MYA, that is, only after complete Central American Seaway closure and formation of the Isthmus of Panama 10-11MYA [Montes et al 2015]. First round of island colonization has occurred 4.1 – 8.1 MYA by ancestors of *A. agilis*, who have migrated either directly to Jamaica, which at that time was closer to the mainland than it is now [Boschman 2014, Hedges 1996] and has recently re-emerged from underwater (around 10MYA) [Hedges 1996, Wadge 1984], or initially to Cuba, where they would have gone extinct afterwards, maybe because of competition with those who arrived later. Cuba's present island shape has formed by late Miocene or early Pliocene [Donovan 1994, Hedges 1996] – around 5.5MYA – and this is approximately when the ancestors of *A. leucocephala, A. ventralis* and *A. vittata* have divided from the mainland species *A. albifrons*, that is, colonization of Cuba by the

ancestral population has occurred. This ancestral population has given rise to *A. collaria* by spreading to Jamaica around 2MYA and to *A. ventralis* and *A. vittata* via stepping-stone migration 0.7 - 1MYA.

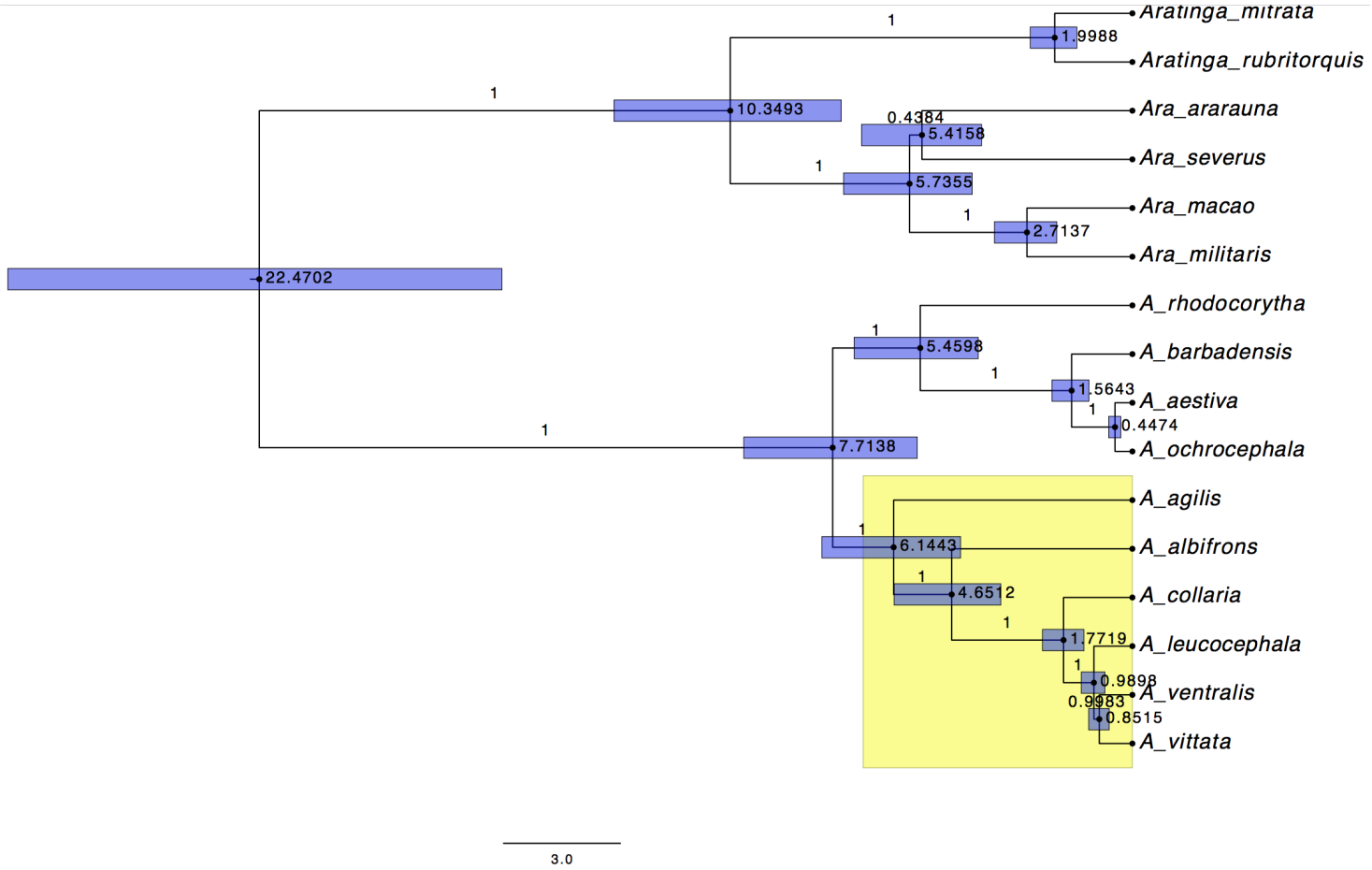


Figure 14. BEAST2 tree using all coding partitions from mitochondrial DNA of 16 parrot species and 2 secondary calibration priors obtained from the Rheindt et. al 2014 paper – 12.12 Mya split between *Ara* and *Aratinga* and 23.48 Mya split between *Amazona* and *Ara* + *Aratinga*. Greater Antillean clade is highlighted in yellow. Numbers at the nodes represent estimated node ages in millions of years. 95% confidence intervals are represented with the blue bars. Numbers on branches represent their posterior probabilities. The time scale bar is in millions of years.

Estimated divergence times are in accordance with the prior knowledge – estimated node ages for split between *Ara+Aratinga* and *Amazona* as well as between *Ara* and *Aratinga* [Rheindt et al 2014] fall nicely within the 95% confidence interval for

the corresponding node ages in our Bayesian tree (see Fig. 13). Estimated divergence times for Caribbean *Amazona* also coincide with previously proposed hypotheses concerning their diversification and its timing [Ottens-Wainright 2004] – the split between *A. collaria* and *A. leucocephala* + *A. ventralis* + *A. vittata* according to our data has occurred 2.2 – 1.2MYA, followed by the separation of *A. leucocephala* and *A. ventralis* + *A. vittata* (1MYA), and then of *A. ventralis* and *A. vittata* (0.85MYA).

The previously proposed migration-speciation scenarios [Lack 1976, Snyder 1987] agree on that the two Jamaican species – the Black-billed (*A. agilis*) and the Yellow-billed (*A. collaria*) parrots - are probably not the closest living relatives of each other (see Fig 13), both hypotheses suggesting close relationship between *A. collaria* and *A. leucocephala*; *A. agilis* and *A. albifrons*. However, they disagree on whether *A. ventralis* and *A. vittata* have descended from *A. leucocephala* via “stepping-stone” colonization of the islands [Lack 1976] or from *A. collaria* and *A. agilis*, respectively [Snyder, 1987]. Our results support the hypothesis of *Amazona* speciation in the Greater Antillean islands formulated in Lack et al. (1976), and further supported by Russello and Amato (2004) by confirming out that *A. vittata* is more related to *A. leucocephala* in Cuba than to *A. agilis* in Jamaica.

Another phylogenetic tree inferred from all coding loci in mtDNA of these parrot species was constructed using Maximum Likelihood method based on the Tamura-Nei model [Tamura 1993] (see Figure 15). This phylogenetic tree has the same topology as observed in the discussed above Bayesian tree, with the exception of relative positions

of *A. agilis* and *A. albifrons* – here the white-fronted amazon is basal to all the other parrots from the region, however support of this node is extremely low – 33.8% (bootstrap). Another unresolved branch in this analysis remain *A. ventralis* + *A. vittata*, with bootstrap support of 75.8%. This method does not allow for adequate estimation of node ages in this case as well.

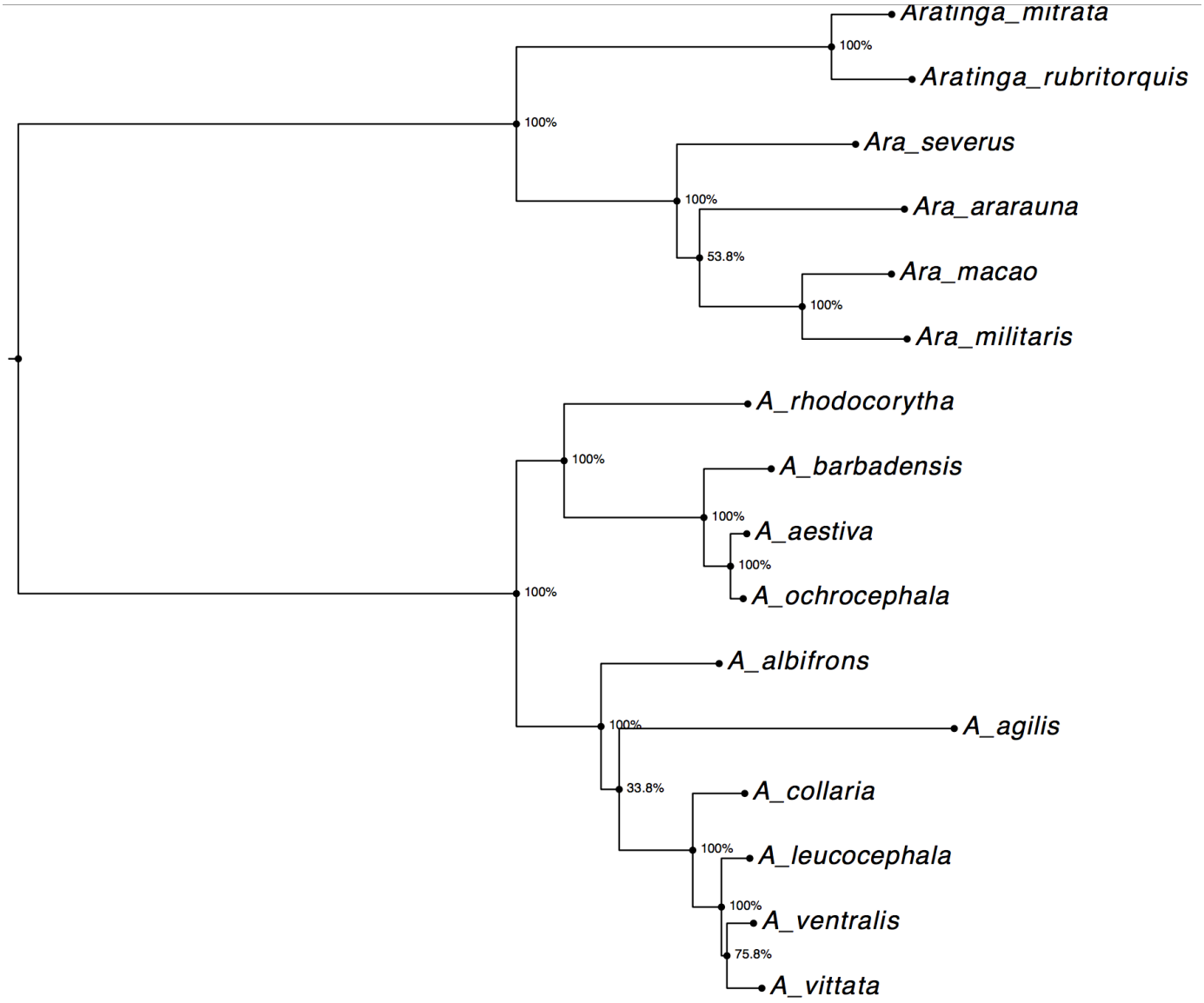


Figure 15. Molecular Phylogenetic analysis of all mitochondrial coding loci by Maximum Likelihood method. The tree with the highest log likelihood (-53968.3801) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches.

Cytochrome B phylogeny

In addition we have reconstructed the phylogeny using only sequences of cytochrome B gene – similarly to how it was previously done in the literature [Ottens-Wainright 2004]. We have used two methods – Maximum Likelihood and Maximum Parsimony. This allowed to include the sequence from the ancient DNA sample of *A. vittata* from Vieques into this dataset, since the Cytochrome B gene in it was found to be nearly complete. The resulting trees (presented in Figures 16 and 17) do not resolve some of the nodes, which confirms the need for use of multiple loci in the analysis.

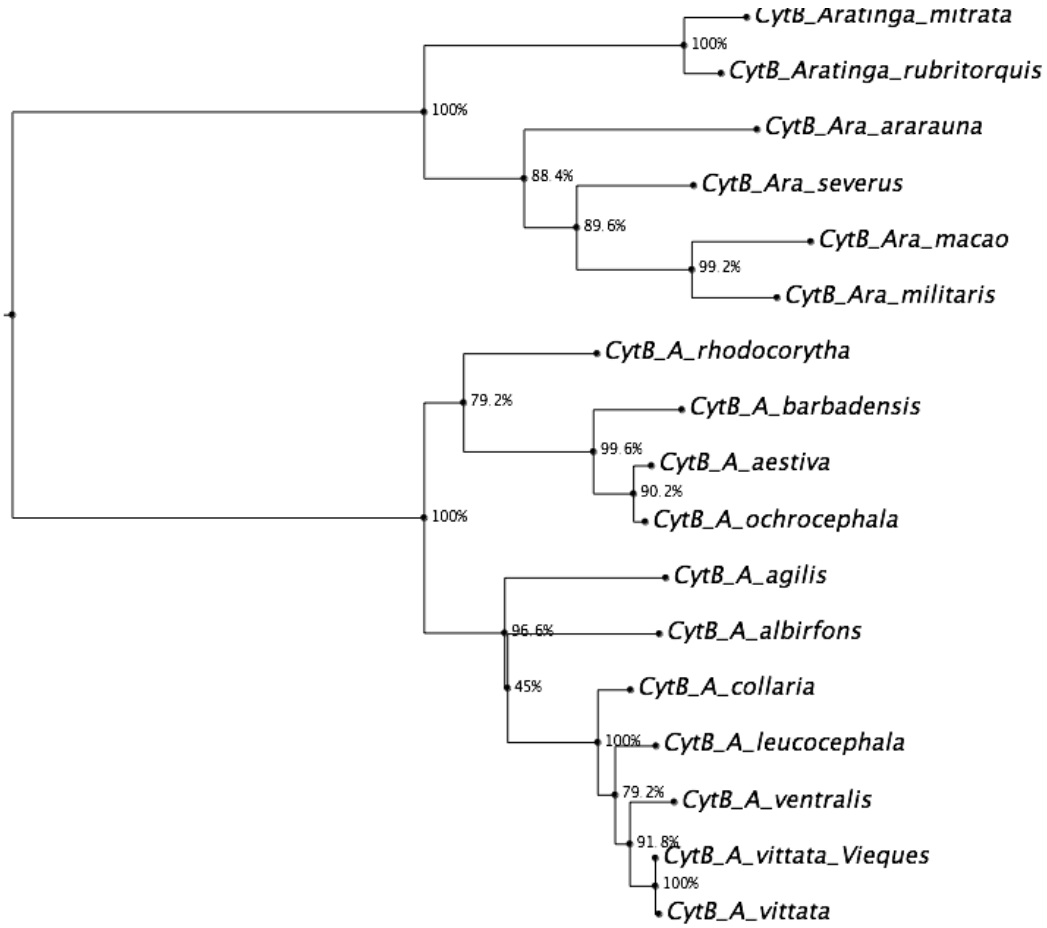


Figure 16. Molecular Phylogenetic analysis by Maximum Likelihood method based on the Tamura-Nei model [Tamura 1993]. The tree with the highest log likelihood (-3754.0729) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches.

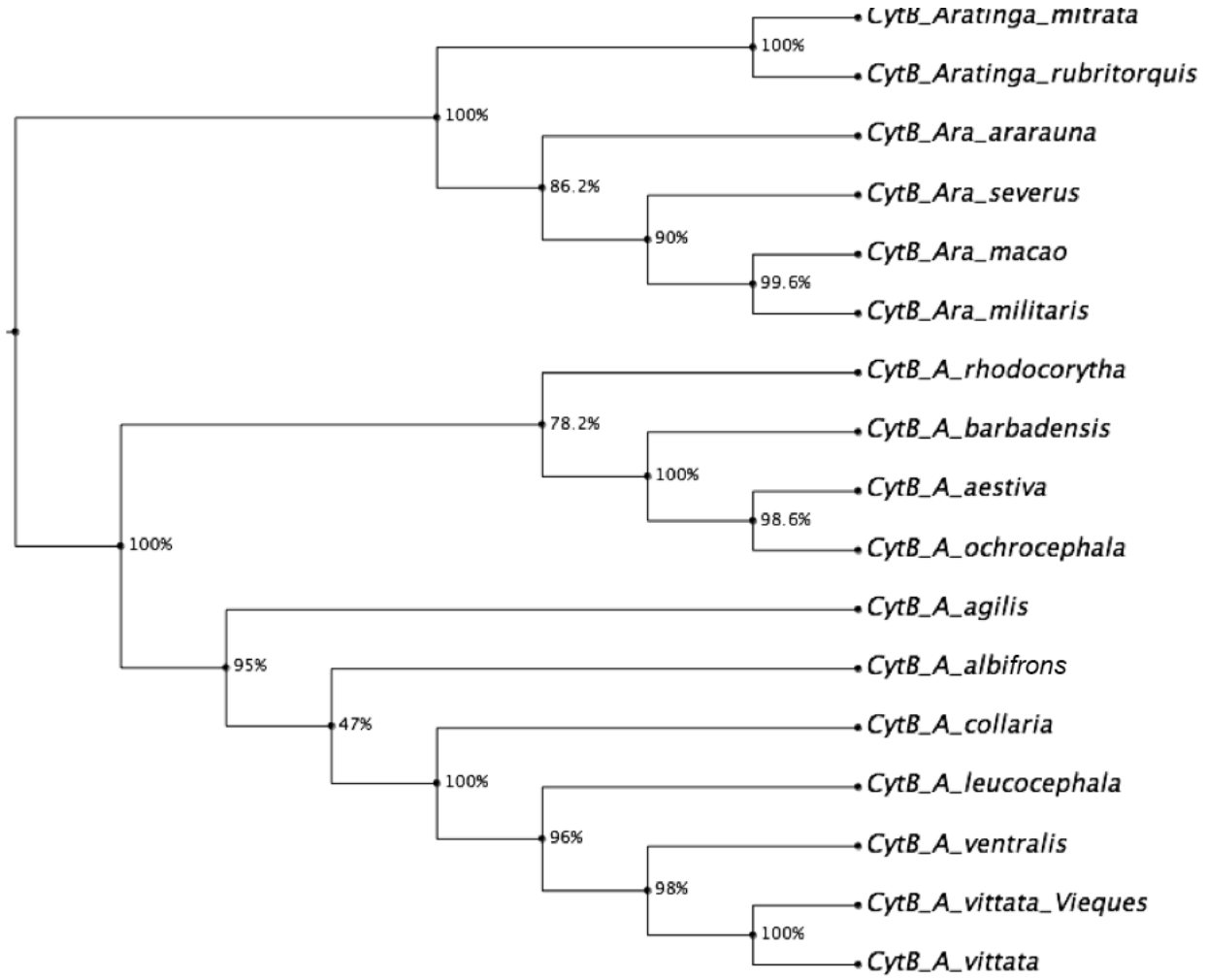


Figure 17. Maximum Parsimony analysis of taxa. Tree #1 out of 2 most parsimonious trees (length = 546) is shown. The consistency index is (0.558091), the retention index is (0.738007), and the composite index is 0.450103 (0.411875) for all sites and parsimony-informative sites (in parentheses). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [Felsenstein 1985].

DISCUSSION

1. Overview

In this study we sequenced and assembled mitochondrial genomes of seven representatives of neotropical parrots of genus *Amazona*, six of which inhabit Greater Antillean islands and Yucatan peninsula. Our approach enabled us to assess previously unresolved phylogenetic relationships within the monophyletic clade of Greater Antillean Amazons, estimate divergence times between them and with other genera of family Psittacidae and test different proposed models of island colonization by these avian species. The study will provide valuable insights into natural history of the archipelago, and adds information that can be incorporated into the phylogenetic tree of birds.

2. Mitochondrial genome assembly and annotation

In our study we have confirmed, that mitochondrial genome of all representatives of the Greater Antillean amazons contain the nearly exactly duplicated control region (CR) and gene arrangement around it characteristic for all the other so far sequenced *Amazona* (also see Fig. 7 D):

-ND5-cytb-T-pND6-pE-CR1-P-ND6-E-CR2-F-12S-

The duplicated control region interferes with regular assembly methods, because it is impossible to map short reads properly onto such repetitive sequences (~1600 bp each). Using long read data, such as PacBio or Oxford Nanopore gives us an

opportunity to overcome such complexities by covering the whole region of duplication and therefore resolving it, like it was performed in our study.

3. Extinct parrot from Vieques

It is not known, whether the extinct population from Vieques island was isolated and constituted a separate subspecies of its own, or if the birds migrated seasonally from Puerto Rico [Sorrié 1975, Forshaw 2017]. Mitochondrial genome of a specimen from the extinct population of *A. vittata* from Vieques island was shown to be different from mitochondrial genome of “core” Puerto Rican parrot in at least 13 positions, which are within coding regions, but the mutations are silent. This is a relatively large difference on the population level, since the previous studies have seen very few polymorphic loci in the DNA of the Puerto Rican parrot [Brock and white, 1992; Afanador et al., 2014]. However, the captive population was likely founded by second degree relatives [Brock and white, 1992] and the existing variation does not represent the historic variability in the species. Therefore, this ancient DNA sample shows the possible extent of the genetic diversity within *Amazona vittata* species for the first time.

In contrast, there are 164 differences between Puerto Rican and Cuban parrot, and 176 differences between Puerto Rican and Hispaniolan parrot in the coding regions of mitochondrial DNA in our study. While the difference between Vieques and core Puerto Rican parrot is substantial, it is insignificant on the species level, and there is not enough evidence to suggest that there was a separate subspecies on the Vieques island, any different from that in Puerto Rico. In the future studies, the diversity of the

mitochondrial (and genomic) DNA sequences between and within populations of both islands needs to be addressed in order to understand the historic diversity within and between these islands and to asses properly the status of the extinct Vieques population.

4. Speciation of The Greater Antillean Amazons and Migration Models

Previously, attempts have been made to address the issue of the *A. vittata, A. ventralis* and *A. leucocephala* polytomy [Ottens-Wainright et al., 2004; Rusello and Amato, 2004]. Unfortunately, these otherwise comprehensive studies either did not have sufficient sequence length or lacked representation of all the species forming the Greater Antillean clade. Specifically, Ottens-Wainright et al. (2004) relied on short fragments that may have been insufficient to resolve deep nodes with sufficient support (<75%). The Russello and Amato (2004) study had better node support overall, but disagreed on the positioning of Jamaican amazons (*A. collaria* and *A. agilis*) and the *A. abifrons* in the Yucatan. That study could not resolve the *A. vittata, A. ventralis* and *A. leucocephala* polytomy. In our study, we have attempted to resolve the phylogenetic relationships between parrots in the West Indies. Specifically, we wanted to understand the order and timing of the colonization of the Greater Antilles by the *Amazona* species, and resolve the polytomy between *A. vittata, A. ventralis* and *A. leucocephala.*

We have attempted to reconstruct the phylogeny of this group, like it was done previously, with only a limited amount of molecular data – cytochrome B gene sequences. This didn't allow us to resolve the questionable nodes. We have then tried

to infer this phylogeny from a larger dataset – using all coding mitochondrial loci – with Maximum Likelihood method based on Tamura Nei model. This approach did not result in the desired degree of resolution and neither was it capable of adequate divergence times estimation.

This study used all coding mitochondrial loci in the Bayesian analysis. The resulting tree supported the order of speciation of the Jamaican amazons as previously described in Russello and Amato (2004), but it has also resolved the *A. vittata, A. ventralis* and *A. leucocephala* polytomy with very high posterior probability of 0.9983 (see Figure 13). We have also attempted to obtain speciation dates for the clade. Unfortunately, a very limited amount of parrot fossils is currently available [Boles 1993, Waterhouse et al. 2008]. For this reason, calibrations of parrot phylogenetic trees previously relied on fossils from outside the Psittaciformes. In this study, we used dates indirectly inferred from non-parrot fossils for split between *Ara*+*Aratinga* and *Amazona* as well as between *Ara* and *Aratinga,* recently published by Rheindt et al (2014). Our results support the hypothesis of *Amazona* speciation in the Greater Antillean islands formulated by Lack et al. (1976) and supported by Russello and Amato (2004) by pointing out that *A. vittata* is more related to *A. leucocephala* than to *A. agilis.*

One of the Jamaican species, *A. agilis*, is more basal than *A. albifrons,* that still inhabits Central America (Figure 14). This node has been reported previously by Ottens-Wainright et al (2004), however the bootstrap support for it was less than 50%. This order of branching has been interpreted earlier to suggest that an early branch of

Amazons could have arrived to Jamaica directly, and not from Cuba like the rest of the species in the clade [Snyder 1976; Ottens-Wainright et al. 2004]. Our study suggests the timing of the speciation events.

Our estimated age of the first split leading to the Greater Antillean (GA) amazons (between South American and Central American species) at 7.7 Mya may indicate south-to-north dispersal across the isthmus of Panama. Since at the time the isthmus was covered with savannah-like vegetation, the first species to cross had to be non-arboreal. Molecular evidence from birds indicates that the south-to-north transfer happened around the Panama land bridge completion, with flycatchers crossing around 12.5 Mya and tanagers at 8.5 Mya, while the antbirds and woodcreepers, which prefer arboreal habitats, had to "wait" until a much later date around 3 Mya [Weir et al., 2009]. Today's *A. albifrons* natural habits include a variety of habitats from savannas to palm groves, scrub forest and rainforest. Therefore it is reasonable to suggest that the ancestor of this species would have no problem dispersing north in the absence of rainforest as soon as the Central American Seaway had grown narrow enough circa 10 Mya [Leigh et al. 2013, Montes et al 2015]. This would also imply that the common ancestor of *A. agilis* and *A. albifrons* reached Nicaragua before 6-7 Mya, arriving there from South America.

The inferred dates on the tree (Figure 13) suggest that the first split with *A. agilis* has occurred around 6.14 Mya. At this time over-water dispersal directly to the Caribbean islands seems unlikely. Jamaica has been submerged below the sea level

from 42 until 10 Mya [Ricklefs & Bermingham, 2008; Montes et al., 2015]. Today, amazon parrots have been known to fly 20-25 miles over water, specifically, the flights between mainland Puerto Rico and Vieques have been reported in the XIX century [Snyder, Wiley and Kepler 1987, Sorrié 1975, Forshaw 2017], but the parrots have never been reported in the Virgin islands only 50 miles away - parrots in general are known for absence of long-distance migrations [Wright 2008]. When the sea levels were similar or higher than present, the distance between Jamaica and the mainland of Central America would have been more than 500 km, which wouldn't permit the parrots to reach the island. Periodic overwater dispersal of various vertebrates from Central or North America to Jamaica has happened during periods of low sea level (e.g., late Miocene, ca. 10 Mya; latest Miocene, ca. 6 Mya; latest Pleistocene, 10 to 20 kya) when the Nicaraguan Rise was partially emergent [Wright, Robinson 1993, p418]. Ancestors of *A. agilis* could have migrated to Jamaica around 6MYA, during one of these periods, which is around our estimated data. Alternatively *A. agilis could have* reached Jamaica much later than 6.14 Mya, possibly as late as 3 Mya when the sea level has dropped more significantly and for longer time in the Nicaraguan Rise and provided a sure stepping-stone path for the parrots to cross over from the continent to the island (Figures 18 and 19). Interestingly, the modern range of *A. albifrons* does not extend to Nicaragua. This could imply that the split between *A. agilis* and *A. albifrons* has occurred on the continent before the dispersal event, and afterwards the founding population of *A. agilis* has gone extinct.

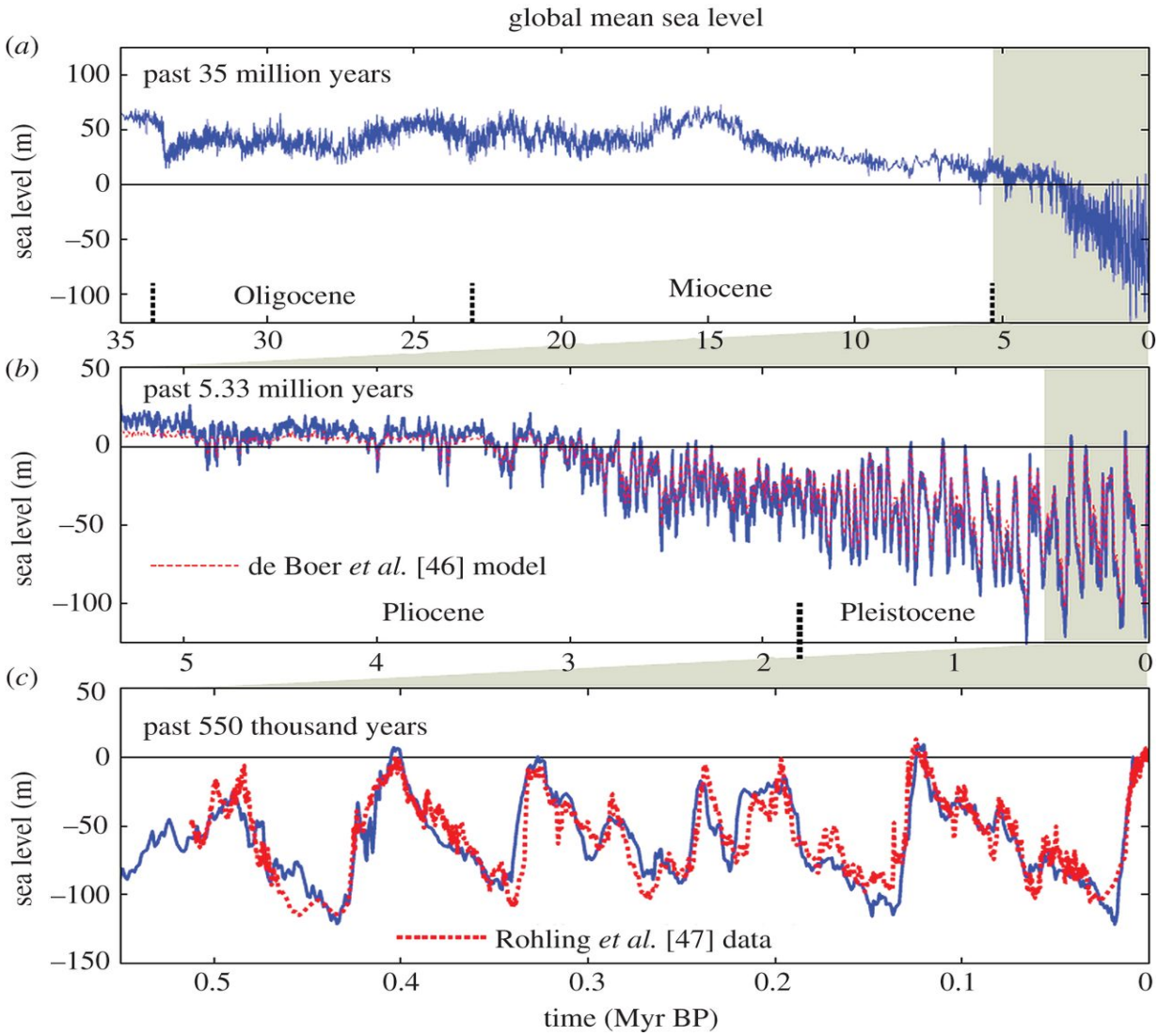


Figure 18. Global sea levels in the Cenozoic. The recent fluctuations since the end of Pliocene have often lowered the sea level by 50 meters or more (From Hansen et al. 2013).

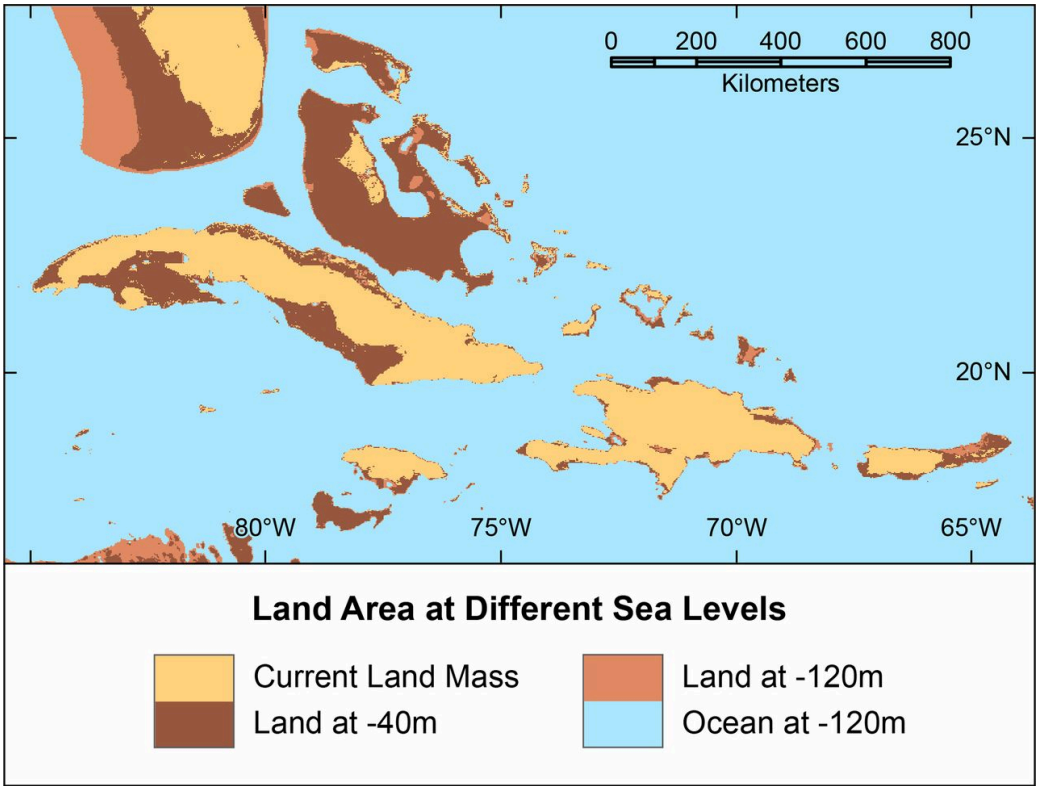


Figure 19. Changing sea-level and extent of Caribbean islands through time. Dropping of sea level by 40 meters would open up a stepping-stone path across the Nicaraguan Rise that could have been used by birds to colonize Jamaica from the south (from David W. Steadman and Janet Franklin 2017).

The fluctuating sea levels occurring in the late Pliocene could also have made it possible for the parrots to cross from Yucatan peninsula to Cuba. Cuba's present island shape has formed by late Miocene or early Pliocene around 5.5 Mya [Donovan 1994, Hedges 1996]. The node separating *A. abifrons* with the rest of the species on the island is dated around 4.65 Mya, and is within the confidence interval [6.14 – 3.16 Mya].

Despite of the shorter distance between Cuba and Hispaniola, the next split in the node that occurred approximately 1.78 Mya leads to the branch of *A. collaria* that cohabits the island of Jamaica with *A. agilis*. Today the two species do not hybridize and occur sympatrically only in few locations such as in the Cockpit Country, however

black-billed parrots occur at greater frequency than yellow-billed parrots in edge habitats [Koenig 2001].

The distances between the islands of Cuba, Hispaniola and Puerto Rico are approximately 100 km each, and the dispersal between the island could have happened directly, or indirectly through the Bahamas as well as Turks and Caicos also during one of the more recent sea level minimums in the late Miocene and Pleistocene (Figures 18 and 19). Our data shows with certainty that Cuban species *A. lucocephala* is the most basal in this trio, suggesting origin of *A. vittata* and *A. ventralis* is consistent with Lack (1976) hypothesis.

5. Recommendations For The Future Work

We recommend that in the nearest future more data is added to this research. First, mitochondrial sequences from another species, *A. xantholora*, which inhabits Yucatan peninsula alongside with *A. albifrons* and is thought to be it's close relative both by morphological and molecular data (though the latter is very insufficient) [Silva 2017], should be incorporated into the study. Addition of subspecies of *A. leucocephala* from Bahamas and Caiman Islands should also improve our understanding of the past parrot migration events in the region. Second, whole-genome sequencing and analysis is needed in order to obtain the best understanding of evolutionary processes which shaped these species. Analysis based solely on mitochondrial DNA is insufficient in

many cases due to phenomena such as incomplete lineage sorting and gene flow following the speciation event, which interfere with proper interpretation of phylogenetic trees. In addition, population diversity in the present-day populations of the Puerto Rican parrot needs to be assessed in order to estimate how much of the original genetic diversity has been lost. This will be an important next step in conservation genetics of this endangered species.

LITERATURE CITED

1. Afanador Y., J. Velez-Valentín, R. Valentín de la Rosa, J.-C. Martínez-Cruzado, B. vonHoldt, T. K. Oleksyk. Isolation and characterization of microsatellite loci in the critically endangered Puerto Rican parrot (*Amazona vittata*). Conservation Genetics Resources. December 2014, Volume 6, Issue 4, pp 885–889.

2. Allendorf, F. W., Hohenlohe, P. A. & Luikart, G. Genomics and the future of conservation genetics. *Nature Rev. Genet.* **11**, 697–709 (2010).

3. Altschul, S.F. et al., 1990. Basic local alignment search tool. Journal of molecular biology, 215(3), pp.403–410.

4. Avise, John. Phylogeography: The History and Formation of Species. Harvard University Press, 2000

5. Bankevich A, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, Sergey I. Nikolenko, Son Pham, Andrey D. Prjibelski, Alexey V. Pyshkin, Alexander V. Sirotkin, Nikolay Vyahhi, Glenn Tesler, Max A. Alekseyev, Pavel A. Pevzner. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 2012 May; 19(5): 455–477. doi:10.1089/cmb.2012.0021

6. Bentley D.R. et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53-9. doi: 10.1038/nature07517.

7. Boetzer M., Christiaan V. Henkel, Hans J. Jansen, Derek Butler, Walter Pirovano. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics, Volume 27, Issue 4, 15 February 2011, Pages 578–579, https://doi.org/10.1093/bioinformatics/btq683

8. Boles, W. E. 1993b. A new cockatoo (Psittaciformes: Cacatuidae) from the Tertiary of Riversleigh, northwestern Queensland, and an evaluation of rostral characters in the systematics of parrots. Ibis, 135: 8–18

9. Bolger Anthony M., Marc Lohse, Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1; 30(15): 2114–2120. Published online 2014 Apr 1. doi: 10.1093/bioinformatics/btu170.

10. Boore, J.L. and Brown, W.M. (1998). Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr. Opin. Genet. Dev. 8: 668-674

11. Boschman LM, van Hinsbergen DJJ, Torsvik TH, Spakman W, Pindell JL (2014) Kinematic reconstruction of the Caribbean region since the Early Jurassic. Earth-Science Reviews 138: 102–136

12. Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C-H., Xie, D., Suchard, MA., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology, 10(4), e1003537. doi:10.1371/journal.pcbi.1003537

13. Brinkley D: The wilderness warrior : Theodore Roosevelt and the crusade for America, 1st edn. New York: HarperCollins; 2009.

14. Brock MK, White BN: Application of DNA fingerprinting to the recovery program of the endangered Puerto Rican parrot. *Proceedings of the National Academy of Sciences* 1992, 89(23):11121-11125

15. Butchart, S. & Symes, A. BirdLife International. 2016. *Amazona albifrons*. The IUCN Red List of Threatened Species 2016: e.T22686222A93103351. http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22686222A93103351.en. Downloaded on 26 December 2017

16. Butchart, S. & Symes, A. BirdLife International. 2016. *Amazona xantholora.* The IUCN Red List of Threatened Species 2016: e.T22686226A93103597. http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22686226A93103597.en. Downloaded on 26 December 2017.

17. Butow Ronald A. 2005. The organization and inheritance of the mitochondrial genome. Nature Reviews Genetics 6, 815–825

18. Cann Rebecca L., Mark Stoneking & Allan C. Wilson. 1987. Mitochondrial DNA and human evolution. Nature 325, 31–36

19. Chaisson MJ, Tesler G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. BMC Bioinformatics 2012 13:238 https://doi.org/10.1186/1471-2105-13-238

20. Chen Xin Jie & Ronald A. Butow 2005. The organization and inheritance of the mitochondrial genome. Nature Reviews Genetics 6, 815–825

21. Chen Y., Weicai Ye Yongdong, Zhang Yuesheng Xu. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic Acids Research, Volume 43, Issue 16, 18 September 2015, Pages 7762–7768, https://doi.org/10.1093/nar/gkv784

22. Cracraft J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. Proc R Soc Lond B Biol Sci. 268:459–469.

23. Darwin, Charles (1845). Journal of researches into the natural history and geology of the countries visited during the voyage of H.M.S. Beagle round the world, under the Command of Capt. Fitz Roy, R.N. (Second ed.), London: John Murray (The Voyage of the Beagle)

24. Diamond J. M. 1989. The present, past and future of human-caused extinctions. Phil. Trans. R. Soc. Lond. B 1989 325 469-477

25. Dobin A., Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson and Thomas R. Gingeras. STAR: ultrafast universal RNA-seq alignerBioinformatics. 2013 Jan; 29(1): 15–21. Published online 2012 Oct 25. doi: 10.1093/bioinformatics/bts635

26. Donovan SK, Jackson TA, ed. 1994. Caribbean Geology: An Introduction. Kingston, Jamaica: Univ. West Indies Publishers' Assoc.

27. Eberhard J.R., Wright TF. 2016. Rearrangement and evolution of mitochondrial genomes in parrots. Mol Phylogenet Evol. 94:34–46.
28. Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.

29. Forshaw J, Knight F. Vanished and Vanishing Parrots: Profiling Extinct and Endangered Species. Csiro Publishing, 2017

30. Ganapathy G., Howard J.T., James M Ward, Jianwen Li, Bo Li, Yingrui Li, Yingqi Xiong, Yong Zhang, Shiguo Zhou, David C Schwartz, Michael Schatz, Robert Aboukhalil, Olivier Fedrigo, Lisa Bukovnik, Ty Wang, Greg Wray, Isabelle Rasolonjatovo, Roger Winer, James R Knight, Sergey Koren, Wesley C Warren, Guojie Zhang, Adam M Phillippy, Erich D Jarvis. De novo high-coverage sequencing and annotated assemblies of the Budgerigar genome. GigaScience 2014 3:11. https://doi.org/10.1186/2047-217X-3-11

31. Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci USA. 2011 Jan 25;108(4):1513-8. doi:10.1073/pnas.1017351108.

32. Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**, 333–351 (2016).

33. Grant PR, and Grant BR. 2014. 40 Years of evolution. Darwin's Finches on Daphne Major Island. Princeton, NJ: Princeton University Press.

34. Greely T., Geologic History of Lesser Antilles. 2012, Joides Resolution

35. Grigorev K., Sergey Kliver, Pavel Dobrynin, Aleksey Komissarov, Walter Wolfsberger, Ksenia Krasheninnikova, Yashira M. Afanador-Hernández, Liz A. Paulino, Rosanna Carreras, Luis E. Rodríguez, Adrell Núñez, Filipe Silva, David Hernández-Martich, Audrey J. Majeske, Agostinho Antunes, Alfred L. Roca, Stephen J. O'Brien, Juan Carlos Martinez-Cruzado, Taras K. Oleksyk. Innovative assembly strategy contributes to the understanding of evolution and conservation genetics of the critically endangered *Solenodon paradoxus* from the island of Hispaniola. 2018 doi: https://doi.org/10.1101/164574

36. Gyory J., Arthur J. Mariano, Edward H. Ryan. "The Caribbean Current." Ocean Surface Currents. (2005). http://oceancurrents.rsmas.miami.edu/caribbean/caribbean.html.

37. Haas BJ, Alexie Papanicolaou, Moran Yassour, Manfred Grabherr, Philip D Blood, Joshua Bowden, Matthew Brian Couger, David Eccles, Bo Li, Matthias Lieber, Matthew D MacManes, Michael Ott, Joshua Orvis, Nathalie Pochet, Francesco Strozzi, Nathan Weeks, Rick Westerman, Thomas William, Colin N Dewey, Robert Henschel, Richard D LeDuc, Nir Friedman & Aviv Regev. De novo transcript sequence reconstruction from

RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8, 1494–1512 (2013) doi:10.1038/nprot.2013.084

38. Hansen J, Sato M, Russell G, Kharecha P (2013) Climate sensitivity, sea level, and atmospheric CO2. Philos Trans R Soc A 371: 20120294, 2013.

39. Jarvis E.D. et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. Science VOL 346 ISSUE 6215 (2014): 1320-1331

40. Joseph L., Alicia Toon, Erin E. Schirtzinger, Timothy F. Wright & Richard Schodde. (2012) A revised nomenclature and classification for family-group taxa of parrots (Psittaciformes). Zootaxa 3205: 26–40

41. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T, "Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads". Genome Res. 2014 Aug;24(8):1384-95. doi: 10.1101/gr.170720.113
42. Kious W.J., Tilling R.I. This Dynamic Earth: The Story of Plate Tectonics. U.S. Geological Survey 1996

43. Kirkconnell, A., Mitchell, A., Cañizares, M. & Cant-Woodside, S. BirdLife International. 2017. *Amazona leucocephala*. (amended version published in 2016) The IUCN Red List of Threatened Species 2017: e.T22686201A110628050. http://dx.doi.org/10.2305/IUCN.UK.2017-1.RLTS.T22686201A110628050.

44. Kirwan, G., Woolmer, G. & White, T. BirdLife International. 2016. *Amazona ventralis*. The IUCN Red List of Threatened Species 2016: e.T22686215A93103000. http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22686215A93103000

45. Koenig, S. BirdLife International. 2016. *Amazona agilis.* The IUCN Red List of Threatened Species 2016: e.T22686231A93103793. http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22686231A93103793.en. Downloaded on 26 December 2017.

46. Koenig, S. BirdLife International. 2016. *Amazona collaria*. The IUCN Red List of Threatened Species 2016: e.T22686209A93102664. http://dx.doi.org/10.2305/IUCN.UK.2016-3.RLTS.T22686209A93102664.en. Downloaded on 26 December 2017.

47. KOENIG, SUSAN E. (2001). The breeding biology of Black-billed Parrot *Amazona agilis* and Yellow billed Parrot *Amazona collaria* in Cockpit Country, Jamaica Bird Conservation International / Volume 11 / Issue 03 / September 2001, pp 205 – 225

48. Koepfli K-P, Paten B, Genome 10 K Community of Scientists, O'Brien SJ (2015) The Genome 10 K Project: a way forward. Annu Rev Anim Biosci 3:57–111.

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

50. Lack, D. (1976) *Island biology illustrated by the land birds of Jamaica.* Oxford: Blackwell (Studies in Ecology 3).

51. Lamichhaney S., Jonas Berglund, Markus Sällman Almén, Khurram Maqbool, Manfred Grabherr, Alvaro Martinez-Barrio, Marta Promerová, Carl-Johan Rubin, Chao Wang, Neda Zamani, B. Rosemary Grant, Peter R. Grant, Matthew T. Webster & Leif Andersson. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518 (2015), 371–375

52. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol. 2017 Mar 1;34(3):772-773. doi: 10.1093/molbev/msw260.

53. Leigh E, O'Dea A, Vermeij G. (2013) Historical biogeography of the Isthmus of Panama. Biol. Rev. 89:148–72.

54. Li H., Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009 Jul 15;25(14):1754-60. doi: 10.1093/bioinformatics/btp324.

55. Li H., Durbin R. Inference of human population history from individual whole-genome sequences. Nature 475, 493–496 (28 July 2011) doi:10.1038/nature10231

56. Li H., Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9. doi: 10.1093/bioinformatics/btp352

57. Lidiak Edward G., D. K. Larue. Tectonics and Geochemistry of the Northeastern Caribbean. Geological Society of America, Jan 1, 1998 - Science - 215 pages

58. Lowe TM., Chan PP. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res. 2016 Jul 8; 44(Web Server issue): W54–W57.Published online 2016 May 12. doi: 10.1093/nar/gkw413

59. MacArthur, Robert H.; Edward O. Wilson (1967). The Theory of Island Biogeography. Princeton, N.J.: Princeton University Press. pp. 203

60. Marçais G., Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics, Volume 27, Issue 6, 15 March 2011, Pages 764–770, https://doi.org/10.1093/bioinformatics/btr011

61. Matos-Maraví P., Rayner Núñez Águila, Carlos Peña, Jacqueline Y Miller, Andrei Sourakov and Niklas Wahlberg. Causes of endemic radiation in the Caribbean: evidence from the historical biogeography and diversification of the butterfly genus *Calisto* (Nymphalidae: Satyrinae: Satyrini). BMC Evolutionary Biology 2014 14:199

62. Matthew, W.D. (1915): Climate and evolution. Ann. NY Acad. Sci. **24**: 171-213.

63. Meschede, M.; Frisch W. (2002). "The evolution of the Caribbean Plate and its relation to global motion vectors: geometric constraints for an inter-american origin". In Jackson T. A. Caribbean geology: into the third millennium : transactions of the Fifteenth Caribbean Geological Conference. University of West Indies Press. p. 279. ISBN 978-976-640-100-9.

64. Montes C, et al. (2015) Middle Miocene closure of the Central American Seaway. Science 348(6231):226–229

65. Naquin D., Yves d'Aubenton-Carafa, Thermes C., Silvain M. CIRCUS: a package for Circos display of structural genome variations from paired-end and mate-pair sequencing data. BMC Bioinformatics 2014 15:198. https://doi.org/10.1186/1471-2105-15-198

66. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2018 Apr 20]. Available from: https://www.ncbi.nlm.nih.gov/

67. Nei M. and Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

68. O'Brien S.J. Genome empowerment for the Puerto Rican parrot – Amazona vittata. Gigascience. 2012; 1: 13

69. Oleksyk et al.: A locally funded Puerto Rican parrot (Amazona vittata) genome sequencing project increases avian data and advances young researcher education. GigaScience 2012 1:14.

70. Oleksyk T. K., Smith, M. W. & O'Brien, S. J. Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc. Lond. B***365**, 185–205 (2010).

71. Oleksyk, T, K; Guiblet, W; Martinez-Cruzado, J, C; Pombert, J, F; Valentin, R (2012): Genomic data of the Puerto Rican Parrot (*Amazona vittata*) from a locally funded project. GigaScience. http://dx.doi.org/10.5524/100039

72. Olson SL. 1978. A paleontological perspective of West Indian birds and mammals. Proc. Acad. Nat. Sci. Philadelphia, Spec. Publ. 13:99-117

73. Ottens-Wainright, P., K. M. Halanych, J. R.Eberhard, R. I. Burke, J. W. Wiley, R. S.Gnam, X. G. Aquilera. 2004. Independent geographic origin of the genus Amazona in the West Indies. Journal of Caribbean Ornithology: 23-49.

74. Page RDM, Lydeard C. 1994. Towards a cladistic biogeography of the Caribbean. *Cladistics 10:21-41*

75. Parra G, Bradnam K, Korf I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics. 2007 May 1;23(9):1061-7

76. Pindell, J.L., 1994. Evolution of the Gulf of Mexico and theCaribbean. In: Donovan, S.K., Jackson, T.A. (Eds.), Caribbean Geology: An Introduction, U.W.I. Publ. Ass. Kingston, pp. 13–39

77. Rheindt F.E., Les Christidis, Sylvia Kuhn, Siwo de Kloet, Janette A. Norman, Andrew Fidler. 2014. The timing of diversification within the most divergent parrot clade. Journal of Avian Biology 45 (2): 140-148.

78. Rhoads A., Kin Fai Au. PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics, 13 (5) (2015), pp. 278-289

79. Ricklefs, R.E. & Bermingham, E. (2008) The West Indies as a laboratory of biogeography and evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 2393–2413.

80. Rohland, Nadin & Hofreiter, Michael. (2007). Ancient DNA extraction from bones and teeth. Nature protocols. 2. 1756-62. 10.1038/nprot.2007.247.

81. Rosen DE. 1975. A vicariance model of Caribbean biogeography. *Syst. Zool.* 24:431-64

82. Russello, M. A. and G. Amato. 2004. A molecular phylogeny of Amazona: implications for Neotropical parrot biogeography, taxonomy, and conservation. Mol. Phylogenet. Evol 30:421–437

83. Salmela L., Rivals E. LoRDEC: accurate and efficient long read error correction. Bioinformatics. 2014 Dec 15;30(24):3506-14. doi: 10.1093/bioinformatics/btu538

84. Salmela, L., and E. Rivals. 'Lordec: Accurate and Efficient Long Read Error Correction'. Bioinformatics 30.24 (2014): 3506-3514. Web. 13 Sept. 2015.

85. Savage JM. 1982. The enigma of the Central American herpetofauna: dispersals or vicariance? Ann. Mo. Bot. Gar- den 69:464-547

86. Schuchert, C. (1935): Historical Geology of the Antillean-Caribbean region. John Wiley and Sons, New York

87. Seabury CM, Dowd SE, Seabury PM, Raudsepp T, Brightsmith DJ, Liboriussen P, Halley Y, Fisher CA, Owens E, Viswanathan G, Tizard IR: A Multi-Platform Draft de novo Genome Assembly and Comparative Analysis for the Scarlet Macaw (Ara macao). PLoS One. 2013, 8: e62415-10.1371/journal.pone.0062415.

88. Silva T, Guzmán A, Urantówka AD, Mackiewicz P. (2017) A new parrot taxon from the Yucatán Peninsula, Mexico—its position within genus Amazona based on morphology and molecular phylogeny. PeerJ 5:e3475

89. Simpson, G.G. (1956): Zoogeography of West Indian land mammals. Amer. Mus. Novit. **1759**: 1-28.

90. Slater G., Birney E.. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 2005 6:31 https://doi.org/10.1186/1471-2105-6-31

91. Snyder, N., Wiley, J.W., Kepler, C.B., 1987. The Parrots of Luquillo: Natural History and Conservation of the Puerto Rican Parrot. Western Foundation of Vertebrate Zoology, Los Angeles.

92. Sorrié B. A. Observations on the birds of Vieques Island, Puerto Rico. Carib. J. Sci., 15 (l-2). June 1975.

93. Stanke M., Morgenstern B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 2005 Jul 1; 33(Web Server issue): W465–W467. Published online 2005 Jun 27. doi: 10.1093/nar/gki458

94. Starostina E. et al. 'Cookiecutter: A Tool for Kmer-Based Read Filtering And Extraction'. (2015): n. pag. Web. 13 Sept. 2015.

95. Steadman DW, Franklin J (2017) Origin, paleoecology, and extirpation of bluebirds and crossbills in the Bahamas across the last glacial–interglacial transition. Proc Natl Acad Sci USA 114:9924–9929

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

97. Teeling E. C., Springer M. S., Madsen O., Bates P., O'Brien S. J. & Murphy W. J. (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307:580–84.

98. Urantowka A.D., Aleksandra Kroczak & Pawel Mackiewicz. Complete mitochondrial genome of the greater Antillean parrot *Amazona ventralis* (Hispaniolan amazon). Mitochondrial DNA Part B Resources Volume 1, 2016 - Issue 1. 864-866

99. Urantowka A.D., Hajduk K, Kosowska B (2013) Complete mitochondrial genome of endangered Yellow-shouldered Amazon (*Amazona barbadensis*): two control region copies in parrot species of the Amazona genus. Mitochondrial DNA 24(4): 411–3.

100. Wadge G, Dixon TH. A Geological Interpretation of Seasat-Sar Imagery of Jamaica. The Journal of Geology 1984 92:5, 561-581

101. Waterhouse, D. M., B. E. K. Lindow, N. V. Zelenkov, and G. J. Dyke. 2008. Two new parrots (Psittaciformes) from the lower Eocene Fur Formation of Denmark. Palaeontology 51:575–582

102. Wayne J. Arendt and Tomás A. Vargas Mora. Range Expansion of the Shiny Cowbird in the Dominican Republic. Journal of Field Ornithology Vol. 55, No. 1 (Winter, 1984), pp. 104-107

103. Weaver P.F, Alexander Cruz, Steven Johnson, Julia Dupin, Kathleen F. Weaver. Colonizing the Caribbean: biogeography and evolution of livebearing fishes of the genus Limia (Poeciliidae). Journal of Biogeography 43, 9 (2016): 1808-1819

104. Weir,J.T.,Bermingham,E.& Schluter, D. (2009). The Great American Biotic Interchange in birds. Proceedings of the National Academy of Sciences of the United States of America 106, 21737–21742

105. Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. Genome Res. 2017; 27:757-767.

106. White, Jr, Thomas & A. Collazo, Jaime & J. Dinsmore, Stephen & Llerandi-Roman, Ivan. (2014). Niche Restriction and Conservatism in a Neotropical Psittacine: the Case of the Puerto Rican Parrot.. 1-83.

107. White, T., Engeman, R. & Anadon, V.. BirdLife International. 2016. *Amazona vittata*. The IUCN Red List of Threatened Species 2016: e.T22686239A93104104. http://dx.doi.org/10.2305/IUCN.UK.2016

108. Whittaker R. J. et al. 2017. Island biogeography: taking the long view of nature's laboratories. – Science 357: eaam8326.

109. Woods Charles A., Florence E. Sergile. Biogeography of the West Indies: Patterns and Perspectives, Second Edition. CRC Press, Jun 27, 2001 - Nature - 608 pages

110. Wright et. 2008. *A Multilocus Molecular Phylogeny of the Parrots (Psittaciformes): Support for a Gondwanan Origin during the Cretaceous.* Mol. Biol. Evol. 25(10):2141–2156. 2008

111. Wright RM, Robinson E. Biostratigraphy of Jamaica. Geological Society of America, 1993

112. Xue W, Li JT, Zhu YP, Hou GY, Kong XF, Kuang YY, Sun XW. L_RNA_scaffolder: scaffolding genomes with transcripts. BMC Genomics. 2013 Sep 8;14:604. doi: 10.1186/1471-2164-14-604

113. Zhang G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 2014 Dec 12;346(6215):1311-20.

114. Zhang G., B. Li, C. Li, M. T. P. Gilbert, E. D. Jarvis, J. Wang, The Avian Genome Consortium, Genomic data of the Avian Phylogenomics Project I. GigaScience 3, 26 (2014) doi:10.1186/2047-217X-3-26