TÍTLE:

Demographic history of the Magnificent Frigatebird (Fregata magnificens) populations off Brazil: an approach using microsatellite genetic markers and approximate Bayesian computation

Historia demográfica de las poblaciones de fragata común (Fregata magnificens) de la costa de Brasil: un enfoque utilizando marcadores genéticos microsatélites y cálculo bayesiano aproximado

AUTHOR:

Lucas Pascual Real
ÍNDICE

Resumen ................................................................. 3
Abstract ................................................................. 4
Introduction ......................................................... 5 – 7
Aims ................................................................. 8
Material and Methods ...................................... 9 – 13
Results and Discusion ..................................... 14 – 17
Conclusion ......................................................... 18
Literature cited .................................................... 19 – 22
La fragata común *Fregata magnificens* es un ave marina distribuida principalmente a lo largo de la costa pacífica y atlántica de América. Un estudio previo utilizando datos genéticos mostró que las poblaciones de Brasil y del Caribe están aisladas, formando linajes independientes, y que la población brasileña tiene una diversidad genética y un tamaño efectivo de la población más pequeño en comparación con el linaje del Caribe. Sin embargo, muchos detalles sobre la historia demográfica del linaje de Brasil permanecen desconocidos. En este estudio, hemos utilizado un conjunto de datos, previamente publicado, de marcadores microsatélites para obtener conocimientos sobre la historia demográfica de las poblaciones brasileñas de fragata común. Más específicamente, se pregunta cuándo se habría producido la reducción de tamaño de la población, y si hay evidencia de efectos de un cuello de botella reciente sobre las colonias de Brasil. El análisis del cálculo bayesiano aproximado encontró que la divergencia entre los linajes de Brasil y del Caribe fue reciente (hace menos de 18.000 años), y que la reducción del tamaño de la población brasileña se remonta a este momento. Por otra parte, no se encontraron pruebas de una mayor reducción de tamaño de la población en las poblaciones brasileñas. Por lo tanto, las diferencias en la diversidad genética pueden ser debidas a um, relativamente, pequeño tamaño de la población efectiva que colonizó Brasil (alrededor de 450 individuos), lo que lleva a la población a una mayor endogamia. La población brasileña de *Fregata magnificens* puede requerir atención especial desde el punto de vista de la conservación con el fin de mantener su diversidad genética y asegurar la conservación de esta especie.
ABSTRACT

The magnificent frigatebird *Fregata magnificens* is a seabird mainly distributed off the Pacific and Atlantic coast of America. A previous study using genetic data showed that Brazilian and Caribbean populations are isolated, representing independent “lineages”, and that the Brazilian population have lower genetic diversity and effective population size compared to populations belonging to the Caribbean lineage. However, several details of the demographic history of the Brazilian lineage remain elusive. In this study, we used a previously published set of microsatellite markers to gain insight on the demographic history of Brazilian populations of *Fregata magnificens*. More specifically, we ask when population size reduction occurred and if there is evidence of recent population bottlenecks in breeding colonies off Brazil. Approximate Bayesian Computation analysis found that the divergence between Brazilian and Caribbean lineages was recent (less than 18,000 years ago), and that population size reduction in the Brazilian lineage dates back to this moment. Moreover, we found no evidence of further population size reductions in the Brazilian populations. Therefore, differences in genetic diversity may be due to a relatively small effective population size that colonized Brazil (around 450 individuals), leading the population to higher inbreeding. The Brazilian population of *Fregata magnificens* may require special attention from the conservation point of view in order to maintain its genetic diversity and ensure the preservation of this species.
I. Introduction

The magnificent frigatebird (*Fregata magnificens*) is a seabird belonging to the Fregatidae family, mainly distributed off the Pacific and Atlantic coasts in the Americas, ranging from California (USA) to Ecuador, including the Galapagos Is., on the Pacific Ocean, and from Florida and the Caribbean to southern Brazil on the Atlantic shore. There is also a relict population on the Cape Verde Islands, in western Africa (Diamond, 2002). In Brazil, frigatebird individuals are found along the whole coast, with breeding colonies distributed in Fernando de Noronha (Pernambuco state), Abrolhos (Bahia state), Ilha do Francês, (Rio de Janeiro state) (Alves, 1993), Ilha dos Alcatrazes (São Paulo state), Ilha do Currais (Paraná state), and Moleques do Sul (Santa Catarina state) (Sick, 1997).

The magnificent frigatebird is the largest of the genus, with a length of 89-114 centimeters long and 217-244 centimeters of wingspan. It has long narrow wings and a clearly forked tail. That species is sexually dimorphic: an entirely black body, with a white chest, white and brownish marking on its wings and a blue eye-ring characterizes females. Their legs are flesh tone, and they do not have a gular sac. For its part, males are completely black with brown inner secondaries on the upper wing, and the presence, during the reproductive stage, of a red inflatable throat sac known as gular sac. They produce a purple iridescence on the head and green on the scapulars, neck, and upper wing. Their legs and feet appear black or grey. In addition, females are 15% larger than males. (Audubon, 1950; Diamond and Schreiber, 2002; Orr, 1992).

*Fregata magnificens* do not land on the water but seize prey, such as crustaceans, flying fish, cephalopods, and jellyfish, from the ocean surface thanks to their long beak (O’Brien, 1990). Thus, they are associated both with large marine predators that corner their prey near the sea surface (Weimerskirch, 2010), and with fishing vessels to take unwanted fish that is thrown on the sea. (Branco, 2004). They practice kleptoparasitism by stealing food from other seabirds (Gilardi, 1994).

The reproductive period begins between June and August, with eggs hatched during November and December. Magnificent frigatebirds form monogamous pairs each breeding season, but they generally change the mating pair every season (Diamond, 1973). They usually breed on oceanic islands with high vegetation hedgerow and
arboreal, where they construct their nests. Both sexes share the task of incubation of the eggs, and the first two or three weeks care of the newborn chick. Males leaves the nest after ~20-110 days. (Osorno 1999, Diamond e Schreiber 2002).

Their habitat coincides with places with great availability of food, and with the trade winds, where they can ride warm updrafts under cumulus clouds and fly for ours without flapping its wings. (Orta, 2015). This, added to the fact that the frigatebird has the largest wing area relative to body weight of any bird, allows them to spend days covering distances of up to 200 km without landing. (Weimerskirch, 2003).

Keeping in mind their flight power, it would be expected that *F. magnificens* populations show high gene flow across large geographic distances, but this do not happen (see below), highlighting that are wind and hydrographic characteristics (temperature, seasonality, salinity and nutrient concentration) determine gene flow among populations, as in other seabirds (Schreiber et al., 2002; Baumgartner et al., 2001).

Nuss *et al.* (2016) studied the evolutionary relationship among the populations occurring off Brazil with those in the Caribbean. This study showed that Brazilian and Caribbean populations are effectively isolated, since there are no mitochondrial (mtDNA) haplotype sharing between them. These authors also suggested that the Brazilian population represent a third evolutionary lineage in this species, together with the Caribbean and the Galapagos population. Finally, the Grand Connétable Is., in French Guyana, would represent a genetic “intermediate” between Brazilian and Caribbean populations due to the mtDNA haplotype sharing between Grand Connétable and Brazil, as well as between Grand Connétable and the Caribbean, with whom this population shared most of its genetic diversity, being considered part of the “Caribbean lineage”. In the same study, these authors found large differences between the level of genetic variation in the Caribbean and in Brazil for both mtDNA and microsatellites (short tandem repeats – STR). The discrepancies in the levels of genetic variation between different populations ultimately translates into differences in historical effective population size, suggesting that the populations of Barbuda and Grand Connétable were even three times higher than the populations occurring off Brazil.

Based on these results, and bearing in mind the degree of difference of the mtDNA between the Caribbean and the Galapagos have been used to estimate a
divergence of about 250,000 years between these lineages (Hailer et al., 2010), Nuss et al. (2016) concluded that the population of Brazil would have arisen from a more recent settling from the Caribbean. Thus, the population of the Caribbean would represents an important historical refuge and a reservoir of genetic variation for the species. Nevertheless, Nuss et al. (2016) did not test the possibility that the low genetic diversity showed by the Brazilian populations is the result of population bottlenecks occurring after the separation from the Caribbean. Discriminating among alternative bottleneck scenarios can be important to understand why population size reductions occurred and which factors may affect the demographic history of this species.

A bottleneck is a drastic reduction in a population size that leads to low genetic variation and increases the power of genetic drift, which will speed up the process of loss of genetic diversity and may result in the fixation of mildly deleterious alleles. When populations become small due to a bottleneck, they become more inbred and less demographically stable, further reducing population size and increasing inbreeding (Frankham et al., 2002). Inbreeding, on average, reduces birth rates and increases death rates and may distort sex-ratios. All this interacts with the basic parameters determining population viability. This feedback between reduced population sizes, loss of genetic diversity and inbreeding is referred to as the extinction vortex, and is illustrated by a direct relationship between population size and persistence (e.g., Berger, 1990).

Genetic variation allows populations to respond to a wide range of environmental challenges, such as climate change or new pathogens (Hoffmann & Parsons, 1997). Small populations lose genetic diversity faster at each generation. Thus, selection response should be reduced compared to large populations. So, from a conservationist standpoint, it is important to know if a bottleneck occurred, since the loss of genetic variation in the population may limit its long-term persistence and its potential for future adaptive evolutionary change (Guerrant, 1996), since the population may lack the necessary characters to adapt to new selective pressures.
II. Aims

The primary goal of this study is to characterize the demographic history of the Brazilian populations of *Fregata magnificens*.

As specific aims, we would like to:

a) Test, for each breeding colony of Brazil, that this population underwent a recent population bottleneck.

b) Based on the results found for the bottleneck test, compare alternative scenarios to determine when the bottleneck happened and estimating its strength.

Our hypothesis is that Brazilian populations suffered a recent bottleneck that is the cause of its low genetic diversity.
III. Material and Methods

Genetic data

In this study, we used the database generated and published by Nuss et al. (2016), which consists of 156 individuals from eight populations: one in the Caribbean (Barbuda), one in French Guyana (Grand Connétable), and six on the east coast of Brazil (Abrolhos, Cabo Frio, Cagarras, Alcatrazes, Currais, and Moleques do Sul). Nuss et al. (2016) considered that both Barnuda and Grand Connétable belong to the Caribbean lineage, even though the latter population may have had recent gene flow with Brazilian populations. Samples were genotyped for eight STR loci, and had a portion of their mtDNA sequenced, though in this study we used only the STR data. The populations of Cabo Frio and Cagarras, both occurring off Rio de Janeiro state were merged in a single population due to their small sample size and to their geographic proximity.

Figure 1. Map of sampled populations: 1-Barbuda, 2-Grand Connétable, 3-Abrolhos, 4-Cabo Frio, 5-Cagarras, 6-Alcatrazes, 7-Currais, 8-Moleques do Sul.
Data analysis

To test the hypothesis that the Brazilian populations underwent a recent population bottleneck we used the method developed by Cornuet and Luikart (1996) and Luikart et al. (1997) and implemented in the software Bottleneck Program. The method is based on the principle that populations that have experienced a demographic bottleneck will show a correlative reduction of both allelic diversity and heterozygosity, but allelic diversity is reduced faster than heterozygosity, resulting in a transient “heterozygosity excess”. This can be illustrated by imagining a locus having 10 alleles at frequency 0.1. In this case, expected heterozygosity is $H_e = 1 - \sum p_i^2 = 0.9$, while for a population losing half of its alleles (5 alleles at 0.2 frequency each), is only slightly reduced ($H_e = 0.8$). To determine whether a population exhibits a significant number of loci with heterozygosity excess, the program has three tests: 'Sign Test', 'Standardized Differences Test', and a "Wilcoxon Sign-Rank Test". However, Sign and Standardized Differences tests have low statistical power for low number of loci (less than 30), so we chose the Wilcoxon-test. The program computes for each population sample and for each locus the distribution of expected heterozygosity conditioned on the number of alleles, given the sample size, under the assumption of mutation-drift equilibrium. This distribution is obtained through simulating the coalescent process of $n$ genes under three possible mutation models, the IAM (Infinite allele Model), SMM (Stepwise Mutation Model) and the two phase model (TPM, which allows multiple-step mutations).

However, IAM assumes that any new mutation generates a new allele, and SMM assumes that every mutation can adds or removes only a single “motif” repeat, which is not realistic for STR data (Eriksson et al., 2012). On the other hand, TPM assumes that some mutations create alleles differing by a single repeat from its parental allele, but that the other part create alleles which differ from its ancestral allele by more than a single repeat according to a geometric distribution. Thus, we used the TPM model assuming mutation parameters which seem to be reasonable for most microsatellites (Cornuet and Luikart, 1996): A SMM ratio of 0 and a variance of the geometric distribution of TPM = 0.36, following the program’s manual.

For comparing alternative bottleneck scenarios, we used a strategy based on a approximate Bayesian computation (ABC) (Bertorelle et al., 2010; Csilléry et al., 2010). All ABC analyses were performed using the package DIY-ABC (Cornuet, 2014). In short, the ABC approach compares “observed” summary statistics estimated from the
real dataset with simulated summary statistics estimated from datasets simulated based on previously specified demographic scenarios defined by a set of parameters whose distribution is given a priori. Simulations resulting in summary statistics close to those from the real (observed) dataset are considered more likely, and, therefore, the “best” (usually the top 100 or top 1,000) are retained to estimate the posterior distribution for each parameter and for each scenario.

Based on the results found for Bottleneck test (see “Results”), we built four different evolutionary sets scenarios (Figure 2) aimed to understand how, or when, the Brazilian lineage underwent a reduction in population size. All scenarios assume two isolated populations: Brazil and the Caribbean (Barbuda). Grand Connétable population is not included in the analysis because previous results suggest it represents a sort of genetic intermediate between the Caribbean and Brazilian lineages (Nuss et al., 2016), and because DIY-ABC simulations cannot handle migration. In scenarios DIV1 and DIV2, the Caribbean population never changed in size, and at a given time, the Brazilian population was separated maintain a constant (thought smaller) size ever since. Thus, in these scenarios, population size reduction in Brazil results directly from population divergence. Scenarios SEC1 and SEC2 are similar to the first, but allows for a more recent bottleneck in the Brazilian lineage after its divergence from the Caribbean.

For the relevant demographic parameters, we used uniform prior distributions for population sizes: $N_{\text{ancCAR}}$ for the Caribbean between 1 and 3,000, $N_{\text{ancBRA}}$ for Brazil between 1 and 3,000, for $N_{\text{botBRA}}$ between 1 and 2,000. These values were chosen based on maximum likelihood estimates of effective population sizes presented by Nuss et al. (2016). The divergence time between the Caribbean and Brazil in scenarios DIV1 and SEC1 were set by parameter T2, which had an uniform distribution between 1 and 10,000 generations, while in the two “SEC” scenarios, there was also a parameter for a more recent bottleneck, T1, which had an uniform distribution between 1 and 10,000 generations. This range is equivalent to 15-150,000 years and to 15-22,500 years, respectively, based on a generation time of 15 years coming from estimates of survival and sexual maturity for frigatebirds (BirdLife International, 2014). Scenarios DIV2 and SEC2 are different only in the time of separation of Brazilian and Caribbean populations, replacing T2 by T3, between 5,000-10,000 generations. These scenarios were evaluated because Nuss et al. found some evidence for an old divergence between
Brazil and Caribbean (unpublished results). Therefore, we wanted to be sure that we were able to select best scenario even if we forced divergences times to be older.

**Figure 2.** Scheme of four possible scenarios. Time is not to scale. Please see the text for details about the demographic parameters.

Overall, we computed 12 summary statistics in DIY-ABC, which included mean number of alleles, heterozygosity, mean size variance, mean Garza-Williamson’s M, pairwise $R_{ST}$, and pairwise $\delta \mu^2$ distances. The best model was selected based on the direct counting and on logistic regression (Fagundes et al., 2007). Each scenario was simulated 1,000,000 times, and the 5,000 simulations closer to the observed data were retained for computation the posterior density of the parameters. Importantly, our main goal with this analysis is selecting the best evolutionary scenario taking in account the
uncertainty in specific parameters. In other words, even though our data cannot estimate all parameters in these scenarios with precision, they are treated as nuisance parameters during the model choice procedure. Finally, we used DIY-ABC to evaluate model misspecification, based on how close the model can reproduce the observed summary statics based on the range of prior values. We also compute bias in model choice and in parameter estimation using “fake” datasets obtained from prior and posterior distributions. In ABC estimation, this step is crucial because, knowing beforehand the “true” scenario and the “true” parameter values, it is possible to evaluate how well the whole estimation procedure is doing. In addition, because models can be “nested” (for example, in model SEC1, if $N_{\text{ancBRA}}$ and $N_{\text{botBRA}}$ have similar posterior values, the model converges to model DIV1, which has a single population size for this lineage), model checking based on posteriors is more effective to evaluate power in finding the best model, while model checking based on priors is more effective to evaluate if there is any bias for choosing a specific model.
IV. Results and Discussion

We did not find evidence, in any breeding colony, for significant heterozygosity excess under the TPM mutation model, even though all Brazilian colonies showed a trend towards higher probability of heterozygosity excess ($P_{Hex}$): Abrolhos ($P_{Hex} = 0.840$), Cabo Frio + Cagarras ($P_{Hex} = 0.875$), Alcatrazes ($P_{Hex} = 0.680$), Currais ($P_{Hex} = 0.711$) and Moleques do Sul ($P_{Hex} = 0.902$). On the other hand, Barbuda exhibited a tendency towards heterozygosity deficiency ($P_{Hdef}=0.727$) and Grand Connétable present no trend towards of heterozygosity excess or deficiency ($P_{Hex} = 0.527$; $P_{Hdef}=0.473$).

Concerning the approximate Bayesian computation analysis, as shown in Figure 3, all four scenarios can result in summary statistics similar to those observed directly in the dataset, and thus, none of the four models seem to be obviously misspecified.

![Figure 3. Distribution of summary statistics based on the observed dataset (large yellow circle) and based on simulated datasets for the four models (see legend in the Figure).](image_url)

The scenario with the highest posterior probability ($PP$), determined by logistic regression calculated from 4,000,000 simulated data sets, was DIV1 ($PP = 0.805$; 95% credible interval (CI) = 0.794-0.815), suggesting that population size reduction in the Brazilian lineage was a direct outcome of population divergence with the Caribbean. The second best model was SEC1 ($PP = 0.186$, 95% CI = 0.176-0.196), followed by SEC2 ($PP = 0.005$, 95% CI = 0.003-0.008) and DIV2 ($PP = 0.004$, 95% CI = 0.002-0.007). These results also suggest that divergence between the Caribbean and Brazil was
recent (see below), provided that the two scenarios in which divergence time was enforced to be older (DIV2, SEC2) had the worst performances. DIV1 was also the best scenario based on direct counting how many times each scenario was represented in the 500 simulations closest to observed data. Based on this method, DIV1 had $PP = 0.760$, followed by SEC1 ($PP = 0.226$), DIV2 ($PP = 0.008$), and SEC2 ($PP = 0.006$).

For evaluating the confidence in scenario choice, we simulated 400 datasets sampling from the posterior distribution of the parameters. Results show that when DIV1 was the true scenario, we were able to recover it in 95% of times, suggesting that our dataset have enough statistical power to discriminate among alternative scenarios, with a low rate of false negatives. However, when scenario DIV1 is not the true scenario, our analysis select DIV1 as the best scenario 58% of the times, indicating a high rate of false positives, especially when the true scenario was SEC1, when this rate rises to 64%. However, because these simulations were based on the posterior distribution of the parameters, the high rate of false positive might be explained by similar values of $N_{ancBRA}$ and $N_{botBRA}$ on scenario SEC1, which in makes SEC1 “converge” into DIV1. This means that a very subtle bottleneck would go undetected in our analysis, which would suggest DIV1 as the preferred scenario. This explanation is corroborated by the analysis based on the prior distribution of parameters, which shows that when scenario DIV1 is true, it will be recovered as the best scenario only 26% of the time, while when SEC1 is the true scenario, DIV1 is selected as the preferred scenario only 23% of the time. These results clearly show that there is no a priori bias for choosing DIV1 as the best scenario.

Considering DIV1 as the best scenario, it is also important to look at the demographic parameters. Table 1 shows posterior estimates for demographic parameters, and its mean relative bias. Posterior density curves are shown in Figure 4. According to our estimates, the strength of the population size reduction is 0.4. That is, the Brazilian lineage of $F. magnificens$ had an effective population size of ~40% compared to the Caribbean. Moreover, the divergence time between populations is very recent. Parameters for which we used narrow and informative priors had larger bias and a confidence interval of 95% wider. This occurs because our prior range was based on the effective population size estimates reported by Nuss et al., (2016), who used the same STR dataset as us. Thus, there is not statistical power to improve much upon the
prior range. On the other hand, divergence time, whose prior values were given a wide range, had a low bias, reinforcing the late divergence between these lineages.

**Table 1** Posterior estimates for demographic parameters and mean relative for scenario DIV1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior range</th>
<th>Posterior mode</th>
<th>Posterior 95% CI</th>
<th>Relative mean bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{\text{ancCAR}}$</td>
<td>[1-3,000]</td>
<td>1,110</td>
<td>512 – 2,200</td>
<td>0.218</td>
</tr>
<tr>
<td>$N_{\text{ancBRA}}$</td>
<td>[1-2,000]</td>
<td>443</td>
<td>116 – 1,660</td>
<td>0.170</td>
</tr>
<tr>
<td>$T2^*$</td>
<td>[1-10,000]</td>
<td>146</td>
<td>41 – 1,170</td>
<td>0.062</td>
</tr>
</tbody>
</table>

*Time is given in generations.

So far, it was known that the Brazilian population of *Fregata magnificens* had lower genetic diversity than populations of the Caribbean and French Guiana. However, the reasons for this difference were unknown. Our study suggests that the most likely evolutionary scenario comprehends a recent divergence between the Brazilian and Caribbean populations about 146 generations ago (2,190 years ago). Even though the credible interval for this parameter is relatively large, this divergence is more recent than 18,000 years ago, and, therefore, occurred only after the last glacial maximum (Clark *et al.*, 2009). During divergence, the population that originally settled off Brazil, and which has not undergone further significant size changes, had an effective population size of around 443 individuals, or ~40% of the effective population size in the Caribbean. Thus, population divergence is the key to the low genetic diversity in Brazil, rather than a recent population bottleneck.

In small populations, inbreeding and loss of genetic diversity are inevitable. Inbreeding is of profound importance in conservation biology as it reduces reproduction and survival in the short term and diminish the capacity of populations to evolve in response to environmental change in the long term. Therefore, loss of genetic diversity increases the susceptibility of populations to extinction (Frankham *et al.*, 2002). Although there are highly inbred bird populations with no apparent signal of inbreeding depression, like the Mauritius kestrel (Groombridge *et al.* 2000), this may be the exception rather than the rule (Frankham *et al.*, 2002). It is generally considered that the minimum viable population (MVP) for maintaining evolutionary potential is 500 effective individuals (Jamieson and Allendorf, 2012; Franklin *et al.*, 2014), but some
authors propose to increase this number to 1,000 (Frankham et al., 2014). Irrespective of the precise values, the effective size of the Brazilian lineage of *F. magnificens* seems to be below the threshold for long-term survival. It may be necessary to monitor this population to ensure that inbreeding depression is not undermining its adaptation. On a positive note, considering that frigatebirds may cover wide distances, dispersal of individuals from Caribbean populations may be effective to restore genetic diversity (and eventually improve reproductive fitness) in the Brazilian populations in the mid-term (see Westemeier et al. (1998) for an example on grater prairie chickens).

![Graphs showing prior and posterior distribution of parameters for scenario DIV1.](image)

**Figure 4.** Prior and Posterior distribution of parameters for scenario DIV1.
VIII. Conclusion

We can conclude that although our hypothesis was not correct, we have achieved the goal of this study: to know the causes of low genetic diversity of the Brazilian population. Our analyzes indicate that the most likely evolutionary scenario suggests a divergence of Brazilian and Caribbean population more recent than 18,000 years ago, beeing the Brazilian population size 40% of the Caribbean population size. This size difference explains why the Brazilian population has less genetic diversity than the Caribbean, and highlights the need to monitor the population and possible interventions to reverse the situation and prevent the Brazilian population of magnificent frigatebird threatened their persistence over time.
IX. Literature cited


