Challenges and prospects of population genetic studies in terns (Charadriiformes, Aves)

Patrícia J. Faria¹², Erika Baus², João S. Morgante¹ and Michael W. Bruford²

¹Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, SP, Brasil.
²Biodiversity and Ecological Processes Group, Cardiff University, Cardiff, UK.

Abstract
Little information is available about the population structure of communally nesting terns (Sternidae) and skimmers (Rynchopidae) throughout the world. In order to fill this gap, a survey of molecular markers was carried out for six species of terns (Anous stolidus, Sterna hirundinacea, S. fuscata, S. supercilialis, Thalasseus maximus and Phaetusa simplex) and one species of skimmer (Rynchops niger). First, we describe the results of the construction of genomic DNA libraries and document problems encountered during this procedure. Secondly, we tested the cross-amplification of 18 microsatellite loci previously described for related species (the number of polymorphic loci ranged from three to seven). Thirdly, we tested the usefulness of mtDNA (control region, ND2, Cytochrome b and ATPase 6/8) for phylogeographic studies in this group of birds. The occurrence of nuclear copies of the mitochondrial control region is reported. Nucleotide divergence in the mtDNA genes analyzed ranged from 0.0 to 0.006. Despite the difficulties associated with the selection of variable markers in this group of seabirds, we were able to select polymorphic markers for each species tested and we anticipate these results will help the development of genetic studies concerning important biological questions in terns.

Key words: molecular markers, microsatellites, mitochondrial DNA, terns.

Introduction
Terns (Sternidae, Charadriiformes) are migratory seabirds widely dispersed throughout the world and include several threatened species (Gochfeld and Burger, 1996). Despite this, there is not enough information about many species to enable effective conservation measures to be taken and several important biological questions regarding their dispersal behavior are not yet understood.

Molecular markers have been useful for addressing many questions in conservation and population biology. From 44 extant species of terns, only three have been genetically studied at the population level (S. fuscata, S. dougalli and S. albifrons) (Avise et al., 2000; Peck and Congdon, 2004; Szczys et al., 2005; Whittier et al., 2006). To carry out such genetic studies, it is important to obtain markers that provide suitable levels of variation for the scope of the proposed questions, which are so far unknown for terns.

Two main types of molecular markers can be used: general (e.g. RAPDs, ISSR and AFLP) or locus-specific markers. Among the locus-specific markers, microsatellites have become widely applied for several types of study, due to their advantages over other markers, such as the relative ease in obtaining markers, their use with non-invasive samples, their high level of variation and the possibility of cross amplification (Bruford et al., 1996; Beaumont and Bruford, 1999).

However, while in many organisms the first and apparently simple step of isolating variable microsatellites is an easy task, in others it has proved to be problematic and can be both time-consuming and expensive (Beaumont and Bruford, 1999). It is known that the abundance of microsatellites across the taxa varies by orders of magnitude (Nève and Meglécz, 2000), and in some groups, such as birds, their frequency is relatively low (Primmer et al., 1997; Neff and Gross, 2001). In addition, major difficulties in the isolation of these markers have been reported for groups such as lepidoptera (Saccheri and Bruford, 1993), birds (Beaumont and Bruford, 1999) and some plants (Squirrell et al., 2003).

Mitochondrial DNA is also extensively used in evolutionary studies because of its unique characteristics such as predominantly maternal inheritance, lack of recombination...
and high evolutionary rate. Conserved primers can be used to amplify DNA in divergent taxa (Kocher et al., 1989) and this enhances the application of mitochondrial DNA. Nevertheless, a high number of studies reported the existence and amplification of nuclear copies of mitochondrial DNA in many organisms (cats: Lopez et al., 1994; gorillas: Thalmann et al., 2005; humans: Mourier et al., 2001; Tourmen et al., 2002 and other eukariotes: Richly and Leister, 2004), including several species of birds (Quinn, 1992; Arctander, 1995; Kidd and Friesen, 1998; Sorenson and Quinn, 1998). When undetected, these nuclear sequences, which are known as nuclear mitochondrial translocated sequences (Numts) (e.g. Zhang and Hewitt, 1996; Sorenson and Quinn, 1998; Richly and Leister, 2004), can lead to misinterpretations of phylogenetic and population genetic data.

It is relatively uncommon to find reports documenting difficulties related to the use and development of molecular markers in some groups. In this paper, we describe problems encountered while conducting population genetic studies in terns that breed in Brazil. First, we describe issues associated with isolating microsatellite loci; second, the occurrence of nuclear copies of mtDNA involving the control region and third, low variation of the few molecular markers available. In order to select and test markers, six species of terns (Anous stolidus, Sterna hirundinacea, S. fuscata, S. superciliaris, Thalasseus maximus and Phaeusa simplex) and one species of black-skimmer (Rynchops niger) were evaluated. Our aims were to select polymorphic and useful molecular markers for conducting population genetic studies in different species of terns and skimmers and also to report our results to researchers attempting to select informative markers in ‘recalcitrant’ species. By reporting this information we hope to improve general understanding of the dynamics of these markers in different groups.

Material and Methods

Samples and DNA extraction

The blood samples used in this study were taken during the field work of 2001 to 2004 and are part of a broader study involving population genetics of different species of tern that breed in Brazil. Chicks and adults were captured using hand nets and mist nets at 16 different locations in Brazil. During this survey, 1200 blood samples (200 mL of blood) from 11 species were taken from the brachial vein. In laboratory, these samples were kept in absolute ethanol at 4 °C and the genomic DNA was isolated using standard phenol-chloroform extraction followed by ethanol precipitation (Sambrook et al., 1989).

Microsatellites

Enriched and non-enriched genomic libraries were constructed for two species (Sterna hirundinacea and Rynchops niger), using a well-established protocol described by Hammond et al. (1998), e.g. Williams et al. (2002), Harper et al. (2006). DNA was digested with MboI and selected fragments (300-1000 bp) were isolated from a 2.0% agarose gel by electrowooling into dialysis tubing and purified using Centricon tubes (Amicon). For the enriched library, fragments were enriched for CA and GA repeats using biotin-labelled probes (Hammond et al., 1998). The fragments of both enriched and non-enriched libraries were ligated into a pUC18 vector and the ligation products were cloned into Invitrogen Topo-one shot competent cells (Top10). Colonies were screened following the PCR-based method of Lunt et al. (1999). Positive clones were sequenced directly from the PCR products using Big Dye Terminator Sequencing kit (version 2.0, ABI) following the manufacturer’s instructions. Primers for those sequences containing microsatellites were designed using Primer 3 (http://www.basic.nwu.edu/biotools/Primer3.html) and DSGene 1.0 (Accelrys Inc).

The microsatellites isolated from the genomic libraries, together with eighteen additional microsatellite loci previously described in the literature for other species including the Red-billed Gull (RBG 13, RBG18, RBG20, RBG27, RBG28, RBG29 and RBG39; Given et al., 2002), the Black-legged Kittiwake (K6, K16, K31, K32, K56, K67 and K71; Tirard et al., 2002) and the Roseate Tern (Sdaat20, Sdaat27, Sdaat46 and Scaac20; Szczys et al., 2005), were tested in at least 14 specimens of each studied species.

Amplifications were performed in a Perkin Elmer thermocycler (9700) using a fluorescent primer (TET, HEX or 6-FAM) and the Qiagen Multiplex PCR Kit (Qiagen) according to the manufacturers instructions. PCR products were visualized and sized on an automated DNA sequencer (ABI PRISM 377) using TAMRA-350 size standard (ABI) and GENESCAN version 2.0 (ABI). The software GENOTYPER 1.1 was also used to identify alleles and genotypes.

Mitochondrial DNA

In order to select mitochondrial regions suitable for genetic studies on terns, eighteen previously described (and one newly designed) primer pairs were tested. The first (5') domain of control region was amplified using the primers CH16746L (ACCCCAAGGACTACGGCTTGAA) (We-nink et al., 1994) and T5437R (GGGTGTGATTTCAC GTGA) (Avise et al., 2000). The second domain was amplified with L438 (TCACGTGAAATCAGCAG) (We-nink et al., 1993) and HL13 (GGGTTGCTGATTTCAC GCAA) and the third (3') domain was amplified using H1248 (CATTCTATGCCTACTTCA) and L699 (ATA AACCCTCCAGTGAC) (Crochet and Desmarais, 2000). In order to avoid nuclear copies, long PCR's were also performed and the combinations of the following primers were tested: L14993 (CCATCCACATCTCACG YTGAAYTT), H 16065 (GGAGCTTCTAGTTTTTG GTTACAAGAC), L15569 (ATCCCATCCACCCCTAC
TACTC), H1816 (GCACCGCAAGTCCTTAGAGTT), UUL16076 (AAAACATTGGTCTTGTAAACC) (Helbig and Seibold, 1999), H519 (GGCCTGACATAGGAACAGAGG) and H419 (GGGTGCTGATTTCACGTGA) (Liebers et al., 2001). Figure 1 shows the flanking position of these primers in the control region.

A portion of the cytochrome *b* gene (320 bp) was amplified using L15008 (AACTTCGGATCTCTACTAGG) and H15326 (GAATAAGTTGGTGATGACTG) (Desjardins and Morais, 1990). Partial sequences of the ND2 gene (500 bp) were obtained using the primers MetL (AAGCTATCGGGCCCATACCG) and H5766 (GATGAGAAGGCTAGGATTTTKCG) (Sorenson et al., 1999), while for the amplification of a portion of the ATPase 6/8, Lys L (CAGCACTAGCCTTTTAAGCT) (Sorenson et al., 1999), and ATP6 intH (TGGGATTAGATGTTTTCTTG) (designed for this study) were used.

Amplifications were carried out in a MJ Research thermocycler (PTC-100 and PTC-200) using: an initial denaturation step of 95 °C, followed by 35 cycles of 95 °C for 30 s, 55-60 °C for 30 s and 72 °C for 30 s, with a final elongation step of 72 °C for 10 min. Long PCRs were carried out using the TripleMaster System (Eppendorf) following the manufacturer’s instructions.

PCR products were purified using shrimp alkaline phosphatase/exonuclease I enzymes (Amersham Biosciences) and sequenced using the Big Dye Terminator Sequencing kit (version 3.1, ABI) following the manufacturer’s instructions.

Results and Discussion

Microsatellites

Isolation: the enriched libraries yielded a total of 1162 recombinant clones and from this total only 451 positives (white colonies) were recovered: 288 clones for *S. hirundinacea* and 163 for *R. niger*. The positives were screened for the presence of microsatellites, however due to the low number of clones identified as potentially containing microsatellites (less than 5%), all clones with inserts larger than 250 bp were sequenced: a total of 171 clones for *S. hirundinacea* and 115 for *R. niger*. From these, 47 were microsatellite sequences in *S. hirundinacea* while only 13 in *R. niger*.

The cloning efficiency of our enriched libraries (percentage of positives clones obtained that contains microsatellite sequences from the total of recombinants) was 5.16% (in total for both species). This value is similar to those described in the literature for other bird species, such as the crested ibis *Nipponia nippon* (Ji et al., 2004; 4%), but smaller when compared with other taxa, such as fish and mammals (Zane et al., 2002). Comparisons of these values are problematic due to the different isolation techniques used and their expected efficiency (see review in Zane et al., 2002 and Gregory and Quinn, 2006). Nevertheless, low values of cloning efficiency in enriched protocols have been attributed to the low number of microsatellites present in the taxon (as for example in lepidoptera, where cloning efficiency is around 2.5%, Ji et al., 2003), which has led to several hypotheses about the origin of the reduced abundance of microsatellites in these genomes (Nève and Meglécz, 2000).

The non-enriched protocol yielded 466 recombinant clones for *S. hirundinacea* and 247 for *R. niger*. Here, only the positive clones for the PCR screening were sequenced (16 for *S. hirundinacea* and 8 for *R. niger*) and none of them were found to be true microsatellite sequences. The percentage of positive clones using traditional methods of microsatellite isolation (non-enriched libraries) has been found to be lower in birds (0.46%), compared to other taxa, such as mammals (1.67%) and fish (3.1%; Primmer et al., 1997; Zane et al., 2002). Even considering that in our study a relatively low number of clones were screened (about 200), our results suggested that the frequency of microsatellites in *S. hirundinacea* and *R. niger* may be than for other species of birds studied to date.

Figure 1 - Schematic diagram showing the locations of the primers used to amplify the mtDNA control region.
Given et al. (2002) reported that for the development of polymorphic microsatellites for Larus novaehollandiae scopulinus, a gull species, it was necessary to carry out screening of 150,000 clones (using both procedures: enriched and non-enriched) of which only 39 contained microsatellites and seven were polymorphic. Until now, only four microsatellite loci have been isolated for the genus Sterna (Szczyzys et al., 2005). However, to obtain these markers, the authors reported that it was necessary to screen approximately 250 million base pairs of the Roseate tern (Sterna dougallii) genome, which produced 166 positive clones and only four polymorphic markers. In this study it was concluded that genomes in the genus Sterna are particularly poor in microsatellites (Szczyzys et al., 2005).

Our data could also be explained due to the low efficiency of the protocols used in this study or even by potential problems during the isolation of microsatellites. However, this seems unlikely for three main reasons. First, several studies using the same protocol (Hammond et al., 1998) have been successfully done (examples can be found in: Aransay et al., 2001; Bailey et al., 2005; Berendonk and Evans, 2004; Berendonk and Dobson, 2006; Burland et al., 2001; Horning et al., 2003; Liu et al., 2005; Liu et al., 2006; Pai et al., 2003; Tero and Schlötterer, 2005; Harper et al., 2006); Second, our laboratory is experienced in the development of microsatellites in different group of organisms, including birds (Jeffery et al., 2001), reptiles (Ciofi and Bruford, 1998); mammals (Radespiel et al., 2001), insects (Wilcock et al., 2001; Seabra et al., 2002; Williams et al., 2002) and earthworms (Harper et al., 2006); and third, the few reports on microsatellite isolation in this group of birds (only one record for Sterna; Szczyzys et al., 2005 and two for seagulls; Given et al., 2002; Tirard et al., 2002) and several personal communications from researchers that were not successful in the development of polymorphic loci in tern species.

From sixty positive clones containing true microsatellites obtained in our study (47 for S. hirundinacea and 13 for R. niger), using the enriched protocol, eighteen did not possess suitable flanking sequences for primer design, seven possessed very small number of repeat units (<4) and twenty-seven were replicate sequences present in other clones. The libraries constructed in our study generated only eight primers pairs, six of which produced an amplification profile comprising a large number of bands, one of which presented a single product, but not of the expected size and one produced a monomorphic product in six different species of terns. Similar results of non-specific or repetitive DNA amplification have also been reported in lepidoptera. Recent studies of this group have shown the occurrence of microsatellite DNA families, with loci containing similar or almost identical flanking regions, that can complicate the selection of single locus markers (Nève and Meglécz, 2000; Zhang, 2004). These findings have been related to the genesis and spread of microsatellites in lepidoptera (Nève and Meglécz, 2000; Zhang, 2004). It has been suggested that mobile elements could be involved in the evolutionary dynamics of microsatellites, including generation of the sequence followed by dispersal and diversification. Accordingly, in organisms that have low microsatellite frequencies and high redundancy, microsatellites would still be in the early stage of evolution and their flanking regions would not have accumulated many mutations (Ellegren, 2004; Zhang, 2004).

In conclusion, further effort is needed to isolate microsatellites specific for terns and skimmers before general conclusions can be drawn. However, it seems to date that the abundance of microsatellites in this group is lower than in other species and the level of redundancy is high. Further investigation into the evolutionary trajectory of microsatellites within this group is needed, as well as the development of new methods to increase the efficiency of microsatellite isolation.

Cross amplification of microsatellites. In birds, a wide survey of the literature has reported a success rate of cross amplification of 84% in several passerine species (Galbusera et al., 2000), while the level of polymorphism ranged from 39% (Galbusera et al., 2000) to 46% (Primmer et al., 1996). The probability of success seems to be higher in species that belong to the same genus or recently separated genera (Zane et al., 2002)

The results of the cross amplification in our data are shown in Table 1. Some loci did not amplify in some species in T. maximus (K67 and K16), S. hirundinacea (K67), S. fuscata (K16) and A. stolidus (Sdaat20) (Table 1). This result could be a consequence of the occurrence of null alleles in those species, therefore those primers were not used further.

The proportion of polymorphic markers in each species ranged from 16.6% in A. stolidus to 38% in S. fuscata (Table 2). These results are within the range obtained for other species of seabirds. In Rissa tridactyla, only 14.28% of the loci were polymorphic when used in Stercorarius parasiticus, a species belonging to a closely related family (Tirard et al., 2002). A recent survey of seventeen microsatellite loci developed for whiskered auklet (Aethia pygmaea) in 42 species of seabirds, showed that as the genetic distance increased, the level of polymorphism decreased more rapidly than amplification success (Dawson et al., 2005). The average level of polymorphism within the family was 52%, while outside the family it dropped to 5% (Dawson et al., 2005). About 90% of the loci developed for Larus novaehollandiae scopulinus were polymorphic within the same genus (Given et al., 2002).

The use of primers developed for other species has thus proven to be a viable alternative, mainly when difficulties are associated with the isolation of specific markers. It is also useful when the study involves several related species and it is time-consuming and potentially inefficient to obtain specific markers for each one.
Mitochondrial DNA

Considering the difficulties associated with the isolation of polymorphic microsatellite loci, it was decided to select mitochondrial DNA sequences that could also be highly informative for addressing questions concerning the population genetics of terns.

Control region. The most suitable region of the mtDNA for population studies is the control region (D-loop) because it evolves rapidly. This segment of the mtDNA has been used in several phylogeographic studies of seabirds (Kidd and Friesen, 1998; Burg and Croxall, 2001; Liebers et al., 2001; Abbot and Double, 2003). However, in some avian species, the control region is not more variable than coding genes (Randi and Lucchini, 1998; Crochet and Desmarais, 2000). In gulls, the third domain of the D-loop showed the same variation as the second domain and cytochrome b gene. This slower rate of evolution could be partly explained by the existence of secondary structures in the third domain of the D-loop of these species (Crochet and Desmarais, 2000).

Table 1 - Number of alleles found in each species of seabird for each primer combination. The number of individuals tested is shown in brackets.

<table>
<thead>
<tr>
<th>Species</th>
<th>SH</th>
<th>SF</th>
<th>SS</th>
<th>TM</th>
<th>PS</th>
<th>AS</th>
<th>RN</th>
</tr>
</thead>
<tbody>
<tr>
<td>K6</td>
<td>3  (12)</td>
<td>5 (14)</td>
<td>3 (3)</td>
<td>1 (10)</td>
<td>4 (8)</td>
<td>6 (14)</td>
<td>3 (10)</td>
</tr>
<tr>
<td>K16</td>
<td>4 (12)</td>
<td>2 (2)*</td>
<td>2 (4)</td>
<td>2 (10)</td>
<td>1 (8)</td>
<td>1 (14)</td>
<td>3 (10)</td>
</tr>
<tr>
<td>K31</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>K32</td>
<td>1 (12)</td>
<td>6 (14)</td>
<td>2 (4)</td>
<td>1 (10)</td>
<td>4 (8)</td>
<td>1 (14)</td>
<td>1 (12)</td>
</tr>
<tr>
<td>K56</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>K67</td>
<td>2 (8)*</td>
<td>2 (14)</td>
<td>2 (4)</td>
<td>1 (6)*</td>
<td>2 (8)</td>
<td>1 (14)</td>
<td>1 (12)</td>
</tr>
<tr>
<td>K71</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>1(14)</td>
</tr>
<tr>
<td>RBG13</td>
<td>2 (17)</td>
<td>6 (14)</td>
<td>2 (2)</td>
<td>1 (10)</td>
<td>3 (8)</td>
<td>1 (14)</td>
<td>1 (11)</td>
</tr>
<tr>
<td>RBG18</td>
<td>3 (12)</td>
<td>1 (14)</td>
<td>2 (5)</td>
<td>2 (10)</td>
<td>2 (8)</td>
<td>1 (14)</td>
<td>5 (12)</td>
</tr>
<tr>
<td>RBG20</td>
<td>1 (12)</td>
<td>2 (14)</td>
<td>3 (4)</td>
<td>2 (8)*</td>
<td>1 (8)</td>
<td>2 (75)</td>
<td>1 (11)</td>
</tr>
<tr>
<td>RBG27</td>
<td>7 (12)</td>
<td>1 (14)</td>
<td>3 (4)</td>
<td>2 (10)</td>
<td>1 (8)</td>
<td>1 (14)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>RBG28</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>1 (14)</td>
</tr>
<tr>
<td>RBG29</td>
<td>4 (12)</td>
<td>8 (14)</td>
<td>1 (5)</td>
<td>1 (10)</td>
<td>2 (8)</td>
<td>1 (14)</td>
<td>1 (12)</td>
</tr>
<tr>
<td>RBG39</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Sdaat20</td>
<td>7 (11)</td>
<td>4 (14)</td>
<td>?</td>
<td>1 (10)</td>
<td>4 (8)</td>
<td>2 (7)*</td>
<td>4(5)</td>
</tr>
<tr>
<td>Sdaat27</td>
<td>1 (53)</td>
<td>2 (14)</td>
<td>?</td>
<td>2 (10)</td>
<td>1 (8)</td>
<td>2 (15)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Scaac20</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Scaac46</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Scuat27</td>
<td>8 (11)</td>
<td>3 (14)</td>
<td>2 (5)</td>
<td>1 (10)</td>
<td>4 (6)*</td>
<td>1 (13)</td>
<td>?</td>
</tr>
</tbody>
</table>

Abbreviations: SH (S. hirundinacea), TM (T. maximus), SF (S. fuscata), SS (S. superciliaris), PS (Phaetusa simplex), AS (Anous stolidus), and RN (Rynchops niger). NS = non-specific amplification; (-) non-amplification; *lack of amplification in some specimens; ? not tested.

Table 2 - Summary of the monomorphic and polymorphic microsatellite loci and its proportion in each species.

<table>
<thead>
<tr>
<th>Species</th>
<th>Monomorphic</th>
<th>Polymorphic</th>
<th>Percentage of polymorphic markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. hirundinacea</td>
<td>K32, K67, RBG13, RBG20, and Sdaat27</td>
<td>K6, K16, RBG18, RBG27, RBG29, Scac20 and Sdaat20</td>
<td>58.33</td>
</tr>
<tr>
<td>S. fuscata</td>
<td>K16, K67, RBG18, RBG20 and RBG27</td>
<td>K6, K32, RBG13, RBG29, Scac20, Sdaat27 and Sdaat20</td>
<td>58.33</td>
</tr>
<tr>
<td>S. superciliaris</td>
<td>K16, K67, RBG13, RBG18, RBG28, RBG29 and Scaac20</td>
<td>K6, K32, RBG20 and RBG27</td>
<td>36.36</td>
</tr>
<tr>
<td>T. maximus</td>
<td>K6, K16, K32, K67, RBG13, RBG29, Scaac20 and Sdaat20</td>
<td>RBG18, RBG27 and Sdaat27</td>
<td>27.27</td>
</tr>
<tr>
<td>P. simplex</td>
<td>K16, RBG18, RBG20, RBG27, RBG29 and Sdaat27</td>
<td>K6, K32, K67, RBG13, Scac20 and Sdaat20</td>
<td>50.00</td>
</tr>
<tr>
<td>Anous stolidus</td>
<td>K16, K32, K67, K71, RBG13, RBG18, RBG28, RBG29, RBG27, RBG28, RBG29 and Scaac20</td>
<td>K6, RBG20 and Sdaat27</td>
<td>21.42</td>
</tr>
<tr>
<td>Rynchops niger</td>
<td>K32, K67, RBG13, RBG20 and RBG29</td>
<td>K6, K16, RBG18, RBG27, RBG28 and Sdaat20</td>
<td>54.54</td>
</tr>
</tbody>
</table>
they used the primers CH1674 GL and T5437R, previously described for other species of birds (Wenink et al., 1994). In the present study, several trials were performed using this set of primers in the six species tested. For all species analyzed in our study, two fragments were obtained, sized 450 bp (expected size) and 600 bp. Peck and Congdon (2004) also related difficulties associated to the use of these primers on the same species (Sterna fuscata) analyzed by Avise et al., (2000). After changes to the PCR conditions, a single fragment of 600 bp was isolated. This fragment was sequenced and the resulting sequences were of good quality and without ambiguities. These sequences were submitted and aligned using BLASTn with those in the GenBank database (Accession number EF028082) and were matched to nuclear copies described for Cephus (Alcidae) (Kidd and Friesen, 1998).

Avise et al. (2000), isolated mtDNA from tissues such as liver and muscle through CsCl-ethidium bromide density gradient centrifugation. In the present study, only blood samples were used. Sorenson and Quinn (1998) pointed out that avian blood samples are particularly likely to yield Numts and should be avoided for mtDNA sequencing. Indeed, red blood cells in birds are nucleated and relatively depauperate in mtDNA. Instead of total genomic extractions, the use of purified mtDNA from mtDNA-rich tissues is preferred.

The occurrence of Numts can be detected when: 1) the PCR amplification produces more than one band; 2) there are ambiguities in the sequence or poor quality; 3) there are stop codons, frameshifts, deletions or insertions in gene sequences; and 4) when the analysis produces contradictory results (Zhang and Hewitt, 1996). Unfortunately, it is not always possible to identify Numts at the first three steps (as happened in our study) and sometimes the detection occurs when all the laborious work has been done and misleading results have been reported as a consequence - including in Genbank (Clifford et al., 2004; Anthony et al., 2007).

Identification of nuclear copies of mtDNA is important because unrecognized inclusion of Numts in an analysis can lead to incorrect conclusions and can invalidate the analyses. However, nuclear integrations provide a real opportunity to study the relative rate of evolution of nuclear and mitochondrial sequences and in calibration of sequence divergence (Zhang and Hewitt, 1996; Sorenson and Quinn, 1998; Bensasson et al., 2001).

Other experiments were carried out in an attempt to sequence the second and third domain of D-loop and the whole control region using the primer sets described in Material and Methods. Long PCR was also used in order to avoid nuclear copies and identify the true mtDNA sequence. Non-specific fragments were produced on amplification of the second and third domain and the whole control region.

Mitochondrial genes. Considering the problems related to sequencing of the control region, it was decided to evaluate the utility of mitochondrial genes in this study. Cytochrome b, ND2 and ATPase 6/8 were chosen as candidates. Cytochrome b is one of the most commonly utilized mtDNA genes in phylogenetics, however, it does not always seem to be the optimal choice. According to Sorenson (2003), all mitochondrial protein genes have similar rates of substitution at the third codon position, but vary in the rate of amino acid substitution. Certain proteins, such as the Cytochrome b are subject to more biochemical constraints than others and, as a result, the levels of homoplasy are often higher than in comparatively unconstrained regions (Johnson and Sorenson, 1998).

In terms of variation, ATPase 8, ND6 and ND2 are considered the most variable genes in birds. However, due to the small size of ATPase 8 and the unusual base composition and location of ND6, they often provide little information (Sorenson, 2003). Sequences of ND2 and ATPase 6/8 have been used in several phylogenetic and phylogeographic studies (Garcia-Moreno et al., 2004; Kennedy and Spencer, 2004; Overton and Rhoads, 2004; Given et al., 2005; Zink, 2005).

Partial sequences of cyt b, ND2 and ATPase 6/8 genes were obtained for each of the six species studied (S. hirundinacea, S. fuscata, S. superciliaris, T. maximus, P. simplex and R. niger). The values of nucleotide divergence ranged from 0.0 to 0.006 within species (Table 3). Consequently, their usefulness for genetic studies is limited and depends on the species analyzed and the questions addressed. Unfortunately, it is impossible to know in advance which gene will be informative in which species for conducting studies at population level.

It seems that the selection of suitable (variable) molecular markers to address questions of population and evolutionary genetics for terns is not a straightforward task and significant effort is needed for each species. However, despite all difficulties encountered during the course of this study, polymorphic markers were selected for each species and can potentially be used in further genetic studies of this group of seabirds. Reports on potential problems encountered during the selection of molecular markers are also im-

Table 3 - Number of haplotypes and values of nucleotide divergence in five species of terns and one species of black skimmer for each mtDNA gene sequences. The number of individuals tested is shown in brackets.

<table>
<thead>
<tr>
<th>Species</th>
<th>Cytochrome b</th>
<th>ND2</th>
<th>ATPase 6/8</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. hirundinacea</td>
<td>2/0.006 (21)</td>
<td>2/0.002 (17)</td>
<td>2/0.002 (11)</td>
</tr>
<tr>
<td>S. fuscata</td>
<td>?</td>
<td>3/0.006 (16)</td>
<td>3/0.006 (13)</td>
</tr>
<tr>
<td>S. superciliaris</td>
<td>1/0 (4)</td>
<td>?</td>
<td>1/0 (5)</td>
</tr>
<tr>
<td>T. maximus</td>
<td>1/0 (10)</td>
<td>1/0 (9)</td>
<td>1/0 (10)</td>
</tr>
<tr>
<td>P. simplex</td>
<td>1/0 (8)</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Anous stolidus</td>
<td>1/0 (9)</td>
<td>2/0.004 (14)</td>
<td>2/0.005 (14)</td>
</tr>
<tr>
<td>Rynchops niger</td>
<td>1/0 (8)</td>
<td>1/0 (15)</td>
<td>2/0.004 (12)</td>
</tr>
</tbody>
</table>

?: not tested.
portant to improve general understanding of their dynamics in different groups.

For future studies involving population genetics of terns, the utilization of other kinds of molecular markers that have proved to be suitable to conduct genetic studies in other species of birds might be considered, such as nuclear introns (Congdon et al., 2000; Pacheco et al., 2002; Friesen et al., 2005; Whittier et al., 2006), SNPs (Saetre et al., 2001; Bensch et al., 2002) and AFLP (Wang et al., 2003; Boulet et al., 2005; Helbig et al., 2005).

In a recent study, cytochrome b and nuclear introns (332bp of α-enolase and 289 bp of glyceraldehydes-3-phosphate dehydrogenase) were used to determine the genetic variation among subspecies of Least tern. Low genetic variability was found using all markers: only three cyt b haplotypes in 50 individuals, a single allele using nuclear introns (332bp of α-enolase and 289 bp of glyceraldehydes-3-phosphate dehydrogenase) were used to determine the genetic variation among subspecies of Least tern. Low genetic variability was found using all markers: only three cyt b haplotypes in 50 individuals, a single allele using Enol and three alleles using Gadp (Whittier et al., 2006). Our findings in addition to the published record suggest that this group of birds presents low levels genetic variation despite the choice of the molecular markers. Further studies need to be done in order to confirm these results, nevertheless if they prove to be general the causes and consequences of this lack of genetic variation needs to be further investigated.

Acknowledgements

We thank Fausto P. Campos, Sérgio L. Pereira, Silvia R. Blanco, Simon P. Shaylor, Sofia Seabra and Trinidad Pérez for suggestions and support in lab work. Patricia Szczys for provided Roseate tern microsatellites primers before their publication. Kate Devlin, Ann Lashko, Gernot Segelbacher, Veronica Neves and Dorit Liebers for sharing information about genetics of seabirds. This research was funded by Cardiff University, Wales, UK and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Process n. 01/01868-7 and 00/13213-2). Blood samples were taken under IBAMA permissions (228/2001, 008/2002, 022/2003, 033/2004; DEUCs: 039/2004, 060/2003, 136/2001 and 057/2001).

References


Molecular markers in terns


 Associate Editor: Fábio de Melo Sene