Mitochondrial Genomic Rearrangements in Songbirds

Staffan Bensch* and Anna Härlid†
*Department of Animal Ecology, Lund University, Lund, Sweden; and †Department of Evolutionary Biology, Uppsala University, Uppsala, Sweden

The organization of the mitochondrial genome is generally very conserved among vertebrates. Because of this, examination of the rare rearrangements which do occur has been suggested as offering a powerful alternative to phylogenetic analyses of mitochondrial DNA sequences. Here, we report on an avian mitochondrial rearrangement in a group of oscine passerines (warblers of the genus Phylloscopus). This rearrangement is identical to the mitochondrial organization recently identified in representatives of four orders of birds, including suboscine Passeriformes. The rearrangement involves the movement of three genes (tRNA^Phe, NADH6, and tRNA^Glu) from their normal position in birds between rRNA^Thr and the control region (CR), to a new location between the CR and a novel, supposedly noncoding (NC), region. Our results suggest that this derived arrangement cannot be used to distinguish between suboscine and oscine passerines, as it has multiple origins both within Passeriformes and within birds as a whole. We found short stretches of DNA with high degrees of similarity between the CR and each NC region, respectively, all of which could be located in the same area of the CR. This suggests that the CR and the NC region are homologous and that the mechanism behind this mitochondrial rearrangement is a tandem duplication followed by multiple deletions. However, the similarities between the control and NC regions of each species were less pronounced than those between the control or NC regions from the different species, supporting the hypothesis of a single basal rearrangement in the Phylloscopus warblers.

Introduction

To date, about 90 complete animal mitochondrial (mt) genome sequences have been determined, representing several phyla (Boore 1999). The gene content of animal mtDNA is highly conserved and, with only a few exceptions, encodes 13 proteins, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNAs). Apart from a noncoding region with regulatory functions (the control region [CR]), animal mitochondrial genomes usually contain very little noncoding DNA. Gene order is highly conserved among most vertebrate mtDNAs; however, there are some exceptions. For example, the gene order in birds differs from that in mammals (Desjardins and Morais 1990), marsupials have a uniquely rearranged tRNA gene cluster (Piábo et al. 1991), and the sea lamprey has two noncoding regions separated by two tRNA genes (Lee and Kocher 1995).

It is commonly believed that rearrangements in mitochondrial genomes represent unique evolutionary events. As such, they have been used to reconstruct phylogenetic relationships, particularly between distantly related taxa (Smith et al. 1993; Boore et al. 1995; Macey et al. 1997). Certain difficulties sometimes associated with phylogenetic analyses of DNA sequences (e.g., different evolutionary rates, nonhomogenous nucleotide composition, ambiguous alignment, and multiple substitutions) can thus be avoided. However, using mitochondrial organization to determine phylogenetic relationships involves two assumptions: that mitochondrial rearrangements are rare, and that a shared organization reflects a common ancestry. A mitochondrial genomic organization which appears to have arisen independently in several lineages was recently discovered in birds (Mindell, Sorenson, and Dimcheff 1998). This finding suggests that certain mitochondrial rearrangements are more likely to occur than others and that convergence must be considered when reconstructing phylogenies using mitochondrial organization.

Most avian orders have the following arrangement between the cytochrome b (Cyt b) and 12S subunit ribosomal RNA (12S rRNA) genes: tRNA^Thr/tRNA^Pro/NADH6 (nicotinamide adenine dinucleotide dehydrogenase subunit 6)/tRNA^Glu/CR/tRNA^Phe. This differs from the typical vertebrate gene order in an apparent single translocation of NADH6 and tRNA^Glu and has been found in 12 bird orders (Desjardins and Morais 1990; Härlid, Janke, and Arnason 1997, 1998; Mindell, Sorenson, and Dimcheff 1998; Härlid and Arnason 1999). A different mitochondrial arrangement (tRNA^Thr/CR/tRNA^Pro/NADH6/tRNA^Glu/tRNA^Phe) was recently reported for the orders Piciformes, Cuculiformes, and Falconiformes and for suboscine passeriforms (Mindell, Sorenson, and Dimcheff 1998). This second avian mitochondrial arrangement requires at least two evolutionary changes in order to have evolved directly from the usual vertebrate arrangement (translocation of the NADH6/tRNA^Glu genes and translocation of the CR). However, it can be derived from the first avian mitochondrial arrangement with only one step. For this reason, it seems likely that the first reported avian gene order represents the ancestral avian mitochondrial arrangement, whereas the second mitochondrial arrangement is derived from the first one. The two avian mitochondrial arrangements are hereafter referred to as the “ancestral” arrangement and the “derived” arrangement, respectively.

Two main mechanisms causing variation in mitochondrial arrangement have been suggested: tandem duplication of gene regions, followed by subsequent deletion of regions, or inversions (Moritz and Brown 1986). The former can occur either through slipped-
strand mispairing (Levinson and Gutman 1987) or through erroneous initiation of light-strand replication at tRNA stem-and-loop structures (Stanton et al. 1994). A third possible mechanism for mitochondrial rearrangement involves recombination, which normally only occurs between homologous DNA segments but could possibly also involve portions of nonhomologous segments (Thyagarajan, Padua, and Campbell 1996). Here, we suggest that the derived avian mitochondrial arrangement is the result of a tandem duplication of the region rRNA_pro/NADH6/tRNA_Glu/CR followed by the deletion of one of each of the two copies except for part of the duplicated CR (fig. 1). If the derived arrangement did indeed arise in this manner, the NC region and part of the CR are homologous and should show a high degree of similarity. We also investigated whether the derived avian arrangement represents a single early rearrangement or several independent rearrangements in different lineages. If the gene rearrangement occurred in the common ancestor of the group of species examined here, then the NC region of each lineage should be more similar to the NC regions of the other lineages than to the CR of that lineage. On the other hand, if there were several independent rearrangements in different lineages, the CR and NC region of each lineage should be more similar to each other than to those of other lineages.

Here, we report the derived avian mitochondrial arrangement in a group of oscine passeriforms, warblers of the genus *Phylloscopus*. We examined six *Phylloscopus* species representing all of the major branches in the genus (table 1) (Price, Helbig, and Richman 1997).

**Materials and Methods**

We used the polymerase chain reaction (PCR) to amplify mtDNA from total genomic DNA extracted

---

**Table 1**

<table>
<thead>
<tr>
<th>Species</th>
<th>Method</th>
<th>Fragment Length (bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken (Gallus gallus)</td>
<td>S</td>
<td>523</td>
<td>Desjardins and Morais (1990)</td>
</tr>
<tr>
<td>Great reed warbler (Acrocephalus arundinaceus)</td>
<td>S</td>
<td>440</td>
<td>Bensch and Hasselquist (1999)</td>
</tr>
<tr>
<td>Upcher's warbler (Hippolais languida)</td>
<td>A</td>
<td>440</td>
<td>This study</td>
</tr>
<tr>
<td>River warbler (Locustella fluviatilis)</td>
<td>A</td>
<td>520</td>
<td>This study</td>
</tr>
<tr>
<td>Red-faced crombec (Sylvietta whyti)</td>
<td>A</td>
<td>440</td>
<td>This study</td>
</tr>
<tr>
<td>Strong-footed bush warbler (Cettia fortipes)</td>
<td>A</td>
<td>440</td>
<td>This study</td>
</tr>
<tr>
<td>Tawny-flanked prinia (Prinia inornata)</td>
<td>A</td>
<td>440</td>
<td>This study</td>
</tr>
<tr>
<td>Fan-tailed warbler (Cisticola juncidis)</td>
<td>A</td>
<td>440</td>
<td>This study</td>
</tr>
<tr>
<td>Striated canegress warbler ( Megalurus palustris)</td>
<td>A</td>
<td>540</td>
<td>This study</td>
</tr>
<tr>
<td>African yellow white-eye (Zosterops senegalensis)</td>
<td>A</td>
<td>460</td>
<td>This study</td>
</tr>
<tr>
<td>Spotted bush warbler ( Bradypterus thoracicus)</td>
<td>A</td>
<td>480</td>
<td>This study</td>
</tr>
<tr>
<td>Yellow-eyed warbler ( Seicercus barkii)</td>
<td>A (S)</td>
<td>1,310</td>
<td>This study</td>
</tr>
<tr>
<td>Willow warbler ( Phylloscopus trochilus)</td>
<td>A</td>
<td>1,310</td>
<td>This study</td>
</tr>
<tr>
<td>Common chiffchaff ( Phylloscopus collybita)</td>
<td>A (S)</td>
<td>1,340</td>
<td>This study</td>
</tr>
<tr>
<td>Wood warbler ( Phylloscopus sibilatrix)</td>
<td>A (S)</td>
<td>1,310</td>
<td>This study</td>
</tr>
<tr>
<td>Pallas's warbler ( Phylloscopus proregulus)</td>
<td>A</td>
<td>1,240</td>
<td>This study</td>
</tr>
<tr>
<td>Greenish warbler ( Phylloscopus trochiloides)</td>
<td>A (S)</td>
<td>1,360</td>
<td>This study</td>
</tr>
<tr>
<td>Large crowned warbler ( Phylloscopus occipitalis)</td>
<td>A (S)</td>
<td>1,400</td>
<td>This study</td>
</tr>
</tbody>
</table>

*a* This region covers the 3’ end of the control region and the 5’ end of 12S rRNA.

*b* Length was determined either by DNA sequencing (S) or by sizing the fragment on agarose gels (A). A (S) indicates that most of the fragment has been sequenced. For all species, PCR amplifications produced one distinct fragment which was of sufficient quality for direct sequencing.

This fragment could not be amplified from the chicken using the DLL3 and 12SH2 primers due to several substitutions in the chicken sequence in the region corresponding to the 3’ end of DLL3.
from the blood of 16 bird species (see table 1). Because avian erythrocytes are nucleated and lack mitochondria, PCR might possibly favor the amplification of nuclear copies of mitochondrial genes (nun'ts) when blood is used as the source of template DNA (Sorenson and Quinn 1998). To confirm that the sequences obtained here were indeed of mitochondrial origin and not nun'ts, we extracted and purified mitochondrial DNA from the fresh muscle and liver of a willow warbler *Phylloscopus trochilus* using the procedures of Arnason, Gullberg, and Widegren (1991). We used this enriched mitochondrial DNA as a template for PCR, then sequenced various stretches of the resulting PCR products (700 nt in total, from the region between Cyt b and 12S rRNA). These sequences were identical to those obtained from PCR products for which total DNA extracted from willow warbler blood was used as the PCR template.

We also examined the tRNA gene sequences (Pro, Glu, and Phe) from the six *Phylloscopus* species investigated here (see table 1), and they all appeared functional, as they formed stable secondary structures. These observations suggest that the sequences obtained here have a mitochondrial rather than a nuclear origin.

For the six *Phylloscopus* warblers (see table 1), the fragment from Cyt b to 12S rRNA was amplified using either the primer CYTEND (5'-CGAACAACAT-5') or the primer CYTBLON (5'-TTCA- CCTATCTCCATTAT-3'), in combination with the primer 12SH2 (5'-AGCAACACAAACGTAAG-3'). For the warblers of other genera, we used only the primer DLL3 (5'-TGATGCACTTTGACCCCATTCATGG-3') in combination with 12SH2 to amplify a fragment from the 3' end of the control region to 12S rRNA. PCR reactions were set up in 25 μl total volumes and included 25 ng of template DNA, 0.125 mM of each primer, and 0.5 U Taq DNA polymerase. The PCR amplifications started with 3 min at 94°C followed by 35 cycles of 30 s at 94°C, 30 s at 52–57°C, and 120 s at 72°C. Each reaction was terminated by a 10-min step at 72°C. We evaluated 2.5 μl of each finished reaction on a 2% agarose gel using 0.5 × TBE buffer. The remaining 22.5 μl was precipitated by adding 11 μl of 8 M NH4Ac and 32 μl ethanol. Following centrifugation and air-drying, the resulting DNA pellet was dissolved in 20 μl of water. Two to four microliters was then used for sequencing.

The PCR fragments were sequenced from both ends and with internal primers using the AmpliCycle sequencing kit (Perkin Elmer) and an ABI PRISM 310 (Perkin Elmer). The sequences of internal primers can be obtained from the authors on request. The sequences were aligned using Pileup (GGC program package), and the alignments were edited after visual inspection in Gene code (Deverueux, Haberli, and Smithies 1984; Nicholas, Nicholas, and Deerfield 1997). The latter program was also used to estimate the degree of similarity between the different DNA sequences. The degree of similarity between the different CRs and NC regions was identified with the program Bestfit (GGC) using gap weight of 5.0 and a length weight of 0.3. The sequences have been deposited at the DDBJ/EMBL/GenBank International Nucleotide Sequence Database with the accession numbers AJ237645–AJ237657.

**Results**

The derived avian mitochondrial gene order was detected in willow warblers during development of a system for population studies on mitochondrial CR variation (Bensch, Andersson, and Åkesson 1999). The primer pair DLL3 and 12SH2, designed to amplify 440 nt from the 3' end of the control region plus tRNA^Phe^ (unpublished data), did indeed amplify a fragment of the expected length from most examined species of the superfamly Sylvioidae (table 1). However, a fragment of approximately 1,300 nt was obtained from the *Phylloscopus* warblers (table 1). This suggested the presence of an insertion or a duplication in the 3' end of the CR in *Phylloscopus* warblers. This longer fragment was also detected in the yellow-eyed warbler *Seicercus burkii*, which is probably more closely related to the *Phylloscopus* species than to any of the other warbler taxa examined here (table 1).

In this study, we sequenced the complete region between the 3' end of Cyt b and the 5' end of 12S rRNA in the willow warbler, along with the complete control and NC regions, including their respective flanking tRNA genes, from the other five investigated species of *Phylloscopus*.

In the willow warbler, the fragment between Cyt b and 12S rRNA showed a mitochondrial arrangement which differed from the ancestral avian mitochondrial arrangement (fig. 1a and b). The sequences from the other five *Phylloscopus* species indicated that they had the same mitochondrial arrangement as the willow warbler. Remarkably, the willow warbler arrangement was identical to the derived arrangement recently described in Piciformes, Cuculiformes, Falconiformes, and suboscine passerines (Mindell, Sorenson, and Dimcheff 1998), with three genes (tRNA^Pro^/NADH6/tRNA^Glu^) having moved from their normal position between tRNA^Thr^ and the CR to a location between the CR and a novel, supposedly noncoding, region. This derived avian mitochondrial arrangement could be the result of a tandem duplication of the region tRNA^Pro^/NADH6/tRNA^Glu^/CR, followed by subsequent deletions of the original copies of the tRNA^Pro^, NADH6, and tRNA^Glu^ genes and part of the new CR (fig. 1c). In order to investigate whether this was indeed the mechanism behind the rearrangement, we examined the degree of similarity between the NC regions and the CRs in the six *Phylloscopus* species. We also examined the 5' end of the CR for remnants of tRNA^Pro^, NADH6, and tRNA^Glu^.

**Control Region**

The CRs are about 1,100 nt long in all of the *Phylloscopus* species examined here (table 2). The conserved sequence block 1 (CSB-1), previously described for the chicken and the ostrich, was identified in all six species. CSB-1 is located in the 5' end of the CR, and the position of this sequence in the *Phylloscopus* species corresponds to that in the chicken and the ostrich. As in
birds in general, all Phylloscopus warbler CRs have C-rich sequences in the 5’ ends which have been interpreted as being involved in forming hairpin-like structures (Quinn and Wilson 1993). CSB-2 and CSB-3, which have been identified in mammalian CRs, were not conclusively identified.

The Phylloscopus CR sequences showed ≈63% similarity between species (table 3). The CRs were 89% identical in the most closely related species pair, the willow warbler and the common chiffchaff (Phylloscopus collybita), and 78% identical between the large crowned warbler (Phylloscopus occipitalis) and the greenish warbler (Phylloscopus trochiloides). Comparisons between the other species pairs showed lower levels of sequence similarity. This pattern of sequence similarity agrees with phylogenetic relationships reconstructed by analyses of the Cyt b gene (Price, Helbig, and Richman 1997).

Noncoding Region

The NC region varies slightly in length between the six Phylloscopus species (table 2). The longest NC region (308 nt) was found in the large crowned warbler, and the shortest (171 nt) was found in Pallas’s warbler.

As the NC regions from the different species were highly divergent, the alignment of these sequences was complicated (fig. 2). We were, however, able to align most of the NC regions from the two most closely related species (the willow warbler and the chiffchaff), which had a similarity of 78%. There was also a substantial sequence similarity (65%) between the NC region of the large crowned warbler and that of the greenish warbler. For the more distantly related species, pairwise comparisons of the NC regions gave very low degrees of similarity. When comparing all six species, we were only able to find two conserved parts in the NC region. The first was between positions 152 and 207 in the alignment while the other was between positions 246 and 325 (fig. 2). Apart from these two regions, the alignments were no better than random.

Comparisons Between the CR and the NC Region

Only parts of the NC region of each species could be aligned with the CR of that species. However, each species’ NC region had short stretches which showed high similarity to a portion of the CR from that species, located between positions 867 and 1136 (fig. 3). These stretches of the NC region were not the same as the conserved regions identified across species (fig. 2). It is noteworthy that the similarities between the stretches of CR and NC region sequences could all be mapped to the same area in the CR. The detected similarity between the CR and the NC region supports the hypothesis that the control and NC regions are homologous and that the derived gene order arose through a tandem duplication followed by deletions. We were not, however, able to detect any significant similarities between the tRNAPro, NADH6, and tRNAGlu genes and any part of the CR.

Discussion

The derived avian mitochondrial arrangement described here has now been identified in representatives from the avian orders Piciformes, Cuculiformes, and Falconiformes (Mindell, Sorenson, and Dimcheff 1998) and from both of the major passeriform groups; suboscines (Mindell, Sorenson, and Dimcheff 1998) and oscines (the present study). It seems most probable that this rearrangement occurred independently several times. The hypothesis of a single evolutionary event in a common ancestor leading to this arrangement requires several unlikely phylogenetic assumptions. For example, a nonmonophyletic origin of oscine passeriforms with the Phylloscopus warblers and several suboscine groups forming a monophyletic group would have to be assumed. In addition, the two passeriform groups (suboscines and Phylloscopus) would form a monophyletic grouping with the orders Piciformes, Cuculiformes, and Falconiformes. These relationships strongly disagree with current hypotheses of avian relationships (Cracraft 1988; Sibley and Ahlquist 1990; Mindell et al. 1999), and we therefore conclude that the derived arrangement has multiple origins. Thus, the value of the derived arrangement in reconstructing phylogenetic relationships among avian groups remains at best uncertain.

Macey et al. (1997) recently examined the known gene rearrangements in vertebrate mtDNA and proposed

<table>
<thead>
<tr>
<th></th>
<th>P. collybita</th>
<th>P. trochiloides</th>
<th>P. occipitalis</th>
<th>P. proregulus</th>
<th>P. sibilatrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. trochilus</td>
<td>89</td>
<td>65</td>
<td>64</td>
<td>73</td>
<td>71</td>
</tr>
<tr>
<td>P. collybita</td>
<td>66</td>
<td>74</td>
<td>73</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>P. trochiloides</td>
<td>78</td>
<td>65</td>
<td>65</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>P. occipitalis</td>
<td>63</td>
<td>73</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. proregulus</td>
<td>73</td>
<td>73</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* For common names of these species, see table 1.
a mechanism which could account for these rearrangements. They suggested that a mutational event disrupts the origin of replication of the light strand (OL) and that alternative stem-and-loop structures (e.g., of tRNA genes) subsequently take over as the start and end points of replication. Because the initiation and termination of DNA replication can now take place at a number of different sites, the occurrence of tandem duplications is facilitated (Fig. 1c). The duplicated regions are then the target of multiple deletions, which may lead to a new mitochondrial organization. However, with this mechanism, the independent occurrence of identical rearrangements in different lineages is highly unlikely because of the many possible combinations of duplication and deletion (Macey et al. 1997). Thus, the occurrence of repeated and identical movements of the tRNAPro, NADH6, and tRNAGlu genes is unlikely to be the result of the above mechanism alone.

The observation of identical mitochondrial rearrangements in bird groups which are not genetically regarded as closely related raises the following questions: Why are the three genes tRNAPro, NADH6, and tRNAGlu (1) prone to move together and (2) prone to moving to the same site in different lineages? One possible explanation is that most mitochondrial rearrangements must be deleterious. Even if duplications and deletions occur relatively frequently, only a few gene combinations might be viable and thus reach fixation. Hence, the ob-

---

**Fig. 2.**—Alignment of the NC regions of six species of *Phylloscopus*: *P. trochilus* (willow warbler), *P. collybita* (common chiffchaff), *P. sibilatrix* (wood warbler), *P. proregulus* (Pallas’s warbler), *P. trochiloides* (greenish warbler), and *P. occipitalis* (large crowned warbler). Dark gray boxes indicate conserved positions, while light gray boxes indicate positions at which only one of the taxa differs.
Fig. 3.—Alignment of the partial control region (positions 867–1136) and the noncoding region of the willow warbler (Phylloscopus trochilus). Conserved sites are indicated by gray shadow.

In the NC region, we were able to identify 80 nt (positions 246–325) with appreciable sequence similarity between the different species. In addition, in each species, we could find parts of the NC region which were similar to parts of the CR from the same species. However, the similarities between the CR and the NC region of each species were less pronounced than those between the different NC regions when we compared closely related species (chiffchaff/willow warbler and greenish/large crowned warblers). Thus, the data presented here support the hypothesis of a single basal rearrangement in the Phylloscopus warbler lineage. In a recent study (Mindell, Sorenson, and Dimcheff 1998), the derived avian mitochondrial arrangement found here to exist in a group of oscine passerines was not identified among any other oscine passeriform despite broad taxon sampling. Future work will focus on whether the derived avian mitochondrial arrangement occurs in other oscine passeriforms, with particular emphasis on the following questions: (1) Is the genus Phylloscopus the only oscine passeriform group to have the derived arrangement? (2) Which is the closest lineage to the Phylloscopus warblers to have the ancestral mitochondrial arrangement? If these questions can be answered successfully, then it will be possible to attempt to date this rearrangement using molecular methods.

Acknowledgments

We wish to thank P. Alström, D. Irwin, U. Olsson, and T. Price for providing DNA samples from several of the species investigated here. A. Janke, K. Slack, and two anonymous reviewers gave many constructive comments on the manuscript. This study was supported by grants from the Swedish Natural Science Research Council and from the Nilsson-Ehle foundation.

LITERATURE CITED


served positions of the three genes (tRNA^Pro/NADH6/tRNA^Glu) either downstream or upstream of the control region in birds might be two of very few functional locations for these genes. In addition, the presence of an NC region approximately 200–950 nt long between tRNA^Glu and tRNA^Phe in all taxa with the derived avian gene order suggests that this region might be important, e.g., for separating these two tRNAs, which in the ancestral arrangement are separated by the CR.

The high degree of similarity between positions 867 and 1136 of the CR and the NC region from each Phylloscopus species, respectively, strongly suggests that the CR and the NC region of each species are homologous and that the mechanism which caused this rearrangement was a tandem duplication followed by multiple deletions. There are only two conserved stretches in the NC region of the Phylloscopus warblers (fig. 2), and whether the NC region has a functional role other than a possible spacing function remains uncertain.

As the derived avian mitochondrial arrangement seems to have originated several times within birds, we cannot rule out the possibility that it has also occurred several times within the genus Phylloscopus. If we assume that the NC region is a partially deleted and degraded copy of the CR and that the rearrangement occurred in the common ancestor of Phylloscopus, we would expect the NC region of each species to be more similar to the NC regions of other species than to the CR of that species. Furthermore, the NC regions of closely related species (e.g., between the willow warbler and the chiffchaff) should be more similar than the NC regions of more distantly related species. Alternatively, if we assume that the rearrangement has occurred independently in all six species, the similarity between the NC region and the CR within each species would be greater than that between NC regions from different species. This pattern could also result from concerted evolution. However, a recent study of snakes carrying two control regions within their mtDNA genomes (Kumazawa et al. 1996, 1998) found that these two copies were more similar to each other than to the single CRs of related species, and it was suggested that this was the result of concerted evolution.


NARUYA SAITOU, reviewing editor

Accepted September 21, 1999