The perils of single gene trees – mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae)

S. M. DEGNAN
Zoology Department, University of Queensland, Queensland 4072, Australia

Abstract

Phylogenetic relationships among animal populations and species commonly have been inferred from patterns of variation observed within a single gene system, most often the mitochondrial genome. Analysis of restriction site variation in the mitochondrial DNA of two species of white-eye (Zosterops lateralis and Z. lutea) in Australia produced a single gene tree that does not accurately represent the organismal tree. In contrast, patterns of variation at two anonymous, single-copy nuclear DNA loci revealed a phylogeography consistent with traditional classification of the species. Discordance between mitochondrial DNA and single-copy nuclear DNA variation is probably the result of past hybridization between Z. lateralis and Z. lutea, evidence of which has been lost from the nuclear genome by recombination. This study provides a clear empirical demonstration that single gene genealogies cannot be assumed to accurately represent the true phylogenies, and emphasizes the need for composite genetic analyses.

Keywords: evolutionary genetics, hybridization, mitochondrial DNA, molecular phylogeny, nuclear DNA, Zosterops

Received 4 January 1993; accepted 23 March 1993

Introduction

Modern comparative biology relies heavily upon the construction of robust phylogenetic hypotheses to provide frameworks for evolutionary interpretations, and the advantages of molecular character sets for evolutionary and phylogenetic studies are now widely acknowledged. In practice, since the adoption of DNA technology, the bulk of molecular analyses of animals have addressed variation only in the mitochondrial genome (e.g. Lamb, Avise & Gibbons 1989; Thomas et al. 1989; Avise, Ankney & Nelson 1990; Miyamoto, Kraus & Ryder 1990; Edwards, Arctander & Wilson 1991; Hedges, Bezy & Maxson 1991; Normark, McCune & Harrison 1991; Shields & Kocher 1991; Zink, Rootes & Dittmann 1991; Dowling et al. 1992; Lanyon 1992; Patton & Smith 1992). Because of the lack of recombination, the entire mitochondrial genome acts as a single genetic locus, providing only a single ‘gene tree’ which might not accurately reflect the ‘organismal tree’ (Pamilo & Nei 1988). This is of vital importance because correct inference of the organismal tree is the goal of most phylogenetic reconstructions. Although several authors acknowledge that lineage bias may be associated with single gene genealogies (e.g. Avise et al. 1983; Avise, Neigel & Arnold 1984; Neigel & Avise 1986; Wilson et al. 1985; Avise et al. 1987; Dittmann & Zink 1991), empirical demonstration of this bias is scarce, and studies of genetic variation relying solely upon mitochondrial DNA (mtDNA) apparently have become widely accepted.

Nuclear DNA provides a source of numerous essentially independent genealogies, under varying degrees of selective constraint, with different mechanisms and rates of evolution, and with a different (biparental) mode of inheritance to that of mtDNA. Allozyme markers have already proved extremely valuable in addressing evolutionary and phylogenetic questions, and DNA markers based on restriction fragment length polymorphisms (RFLPs) are much greater in number and show greater levels of variation than do allozyme markers for closely related species.
groups of organisms (Hartl & Clark 1989). Led by the exemplary studies of Drosophila populations (e.g. Kreitman & Aguade 1986; Thomas & Hunt 1991; and references therein), the enormous potential of direct analysis of DNA variation at single-copy nuclear loci in natural animal populations is gradually being realized (e.g. Quinn & White 1987; Quinn et al. 1987; Nevo, Ben-Shlomo & Maeda 1989; Hall 1990; Gibbs et al. 1990; Yuhki & O'Brien 1990; Bork et al. 1991; Karl & Avise 1992; Karl, Bowen & Avise 1992; Lu et al. 1992).

Here I report an analysis of RFLPs at anonymous single-copy nuclear DNA (scnDNA) loci in two species of white-eye (genus Zosterops), undertaken to complement a previous study of mtDNA variation in the same species (Degnan & Moritz 1992). The white-eye family Zosteropidae is a palaeotropical passerine group represented by two species on the Australian mainland (Blakers, Davies & Reilly 1984). The silvereye Zosterops lateralis is distributed along the eastern, southern and south-western coasts of the continent. The morphologically distinct yellow white-eye Z. lutea is found across the northern and north-western coasts. The ranges of the two species overlap slightly along the west coast, and there is a small outlying population of Z. lutea on the east coast, within the range of the silvereye.

A survey of restriction endonuclease site variation among mtDNAs from geographically dispersed individuals of the two species (Degnan & Moritz 1992) revealed high levels of intraspecific divergence (Fig. 1a); regions of current overlap between the two species were not sampled. Unexpectedly, the distribution of mtDNA haplotypes of Z. lutea was paraphyletic; eastern samples were more closely related to eastern Z. lateralis than to western Z. lutea. The discordance between morphological species boundaries and the distribution of mtDNA haplotypes, coupled with a large sequence divergence between east and west Z. lutea lineages, supported a hypothesis of interspecific hybridization in the north-eastern region of the continent (Degnan & Moritz 1992). This unexpected result prompted the present analysis of additional (single copy nuclear) loci to construct a more robust organismal tree, and, in doing so, to assess the degree of lineage bias associated with a single gene genealogy. In view of the mtDNA result, an additional aim of this study was to search for evidence of interspecific hybridization in the nuclear genome.

Materials and methods

Two scnDNA clones were selected from a partial genomic library constructed in the plasmid vector Bluescript SK+ (Stratagene Cloning Systems) using DNA extracted from Z. lateralis. Details of library construction, and of selection and further analysis of single-copy clones will be described elsewhere (S. Degnan et al., in preparation). Clone pDZL46 (p = plasmid, D = DNA source, ZL = Zosterops lateralis, 46 = 4.6 kb insert size), chosen because it showed high levels of polymorphism with several endonucleases in a preliminary study, was screened in combination with the restriction endonucleases EcoRI, TaqI and BgIII. Clone pDZL18, chosen because it showed potential as a species marker in a preliminary study, was screened in combination with the endonucleases EcoRI and TaqI.

RFLPs at the two selected loci were detected by Southern blot hybridization. Nuclear DNA was purified by a standard phenol/chloroform method (Sambrook, Fritsch & Maniatis 1989) from 22 Z. lateralis and six Z. lutea collected from across their range of distribution. The individuals analysed were the same as those used in the mtDNA study; details are presented in Appendix I of Degnan & Moritz (1992). Three to five micrograms of endonuclease-digested DNA was electrophoresed in 0.8% agarose/TBE, and transferred onto Hybond-N' (Amer sham) using an LKB Vacugene apparatus. DNA was depurinated (3 min), denatured (10 min), neutralized (5 min) and transferred in 20 x SSC (40 to 50 min) at a pres-
sure of 40 cmH₂O. Membranes were fixed in 0.4-N NaOH. Probe DNA (25 to 50 μg) was [³²P]-random-prime labelled (Feinberg & Vogelstein, 1984) using an Amersham Multi-Prime kit. Hybridizations were performed at high stringency (65°C), with 1–3 × 10⁶ c.p.m. of labelled probe per ml of phosphate-buffered hybridization solution (Church & Gilbert 1984). Filters were washed to a high stringency of 0.2 × SSC at 65°C for 30 min, and exposed to Kodak XAR film for 2–10 days, with two intensifying screens.

Results

Alleles for each probe–enzyme combination were determined by the presence of homozygotes. This was especially straightforward when the combination generated only a single polymorphic restriction site (e.g. pDZL46–BglII; pDZL18–EcoRI; Fig. 2), but was more complex when numerous rare alleles were generated by several polymorphic sites (e.g. pDZL46–TaqI). Allelic designations of RFLPs for which homozygotes were not observed generally could be derived from their occurrence with previously determined alleles. Only for pDZL46–TaqI was the allelic status of some observed restriction fragments ambiguous; for the purpose of the present qualitative analysis, ambiguous fragments have been designated as unique alleles. The distribution of observed alleles across the continent was summarized using the eastern and western geographic subdivisions identified by the mtDNA analysis.

The pDZL18 probe identified fixed allelic differences between Z. lateralis and Z. lutea with EcoRI, and frequency differences with TaqI (Table 1; Fig. 2). For EcoRI, both species were monomorphic, with a 1.9-kb fragment exclusive to Z. lutea and a 1.8-kb fragment exclusive to Z. lateralis. TaqI digests also revealed no variation at this locus within Z. lutea, but identified variable sites within Z. lateralis, especially among east coast individuals. The pDZL18 probe thus proved to be a species diagnostic marker. Genotypic diversity was relatively low at this locus, with few unique genotypes observed among 23 Z. lateralis and only one among six Z. lutea.

The variation revealed by the pDZL46 probe was not species-specific, and showed only weak geographic and taxonomic structure (Table 1; Fig. 2). For EcoRI, the allele E1 was common throughout the range of Z. lateralis. Less common alleles showed a greater tendency for frequency differences between western and eastern individuals; in particular, the E2 allele was moderately common along the eastern coast of Australia, but was not observed among western birds. Five of the six Z. lutea individuals were homozygous for the E2 allele, which distinguished them from all Z. lateralis individuals. A single eastern Z. lutea was heterozygous for the E1 and E2 alleles, as were four eastern Z. lateralis. pDZL46–TaqI showed similar patterns to EcoRI, except that a greater number of rare alleles were observed, and the distribution of these showed no obvious geographic structure (Table 1). The T1 allele, common across the range of Z. lateralis, was not observed in Z. lutea. All of the TaqI alleles observed in Z. lutea individuals were rare or absent in Z. lateralis. For pDZL46–BglIII, 22 of the 24 Z. lutealis were homozygous for the B1 allele, as were all six Z. lutea. There was no suggestion of taxonomic or geographic structure revealed by this endonuclease–probe combination. Genotypic diversity at the pDZL46 locus was relatively high, with 15 unique genotypes observed among 24 Z. lateralis, and five among six Z. lutea. For the pDZL46 and pDZL18 loci combined, a total of 21 genotypes was observed among the 24 Z. lateralis, and five among the six Z. lutea; most genotypes within a species differed by only one or two restriction fragments.

Allele frequencies were compared between eastern and western groups of individuals that were of special interest because of the significant east–west differentiation re-
By the mtDNA analysis. Pairwise genetic distances were estimated from allele frequencies (Table 1) according to Nei (1978), using the program PHYLIP version 3.4 (J. Felsenstein 1991, University of Washington, Seattle). Genetic distance data analysed by UPGMA clustering, again according to Nei (1978), using the program PHYLIP version 3.4 for the same individuals (Fig. 1a, b). Two characteristics of the scnDNA dendrogram particularly are in contrast to the mtDNA dendrogram. First, eastern Z. lutea clustered with western Z. lutea, rather than with eastern Z. lateralis as in the mtDNA analysis. Secondly, the degree of genetic differentiation between east and west groups of individuals was almost negligible (0.0091 for Z. lateralis and 0.0097 for Z. lutea) in the scnDNA analysis, but was very large in the mtDNA analysis. For each species, east–west allele distributions were compared using chi-square contingency tests (Weir 1990) performed on actual numbers of alleles, derived from the frequencies presented in Table 1. Because of small sample sizes, counts for rare alleles observed at frequencies of less than or equal to 0.17 were pooled within each probe–enzyme category (e.g. for Z. lateralis – pDZL46: E3 + E4; T2 + T3... T12; pDZL18: T2 + T3 + T4). For both Z. lateralis and Z. lutea, chi-square tests confirmed that there was no significant east–west differentiation at either the pDZL18 or the pDZL46 locus (P > 0.01 in all cases; data from each endonuclease were treated separately; counts for rare alleles were pooled, see text). Chi-square values were non–significant in all cases (P > 0.01).

### Table 1 Geographic distribution of allele frequencies for two scnDNA loci (pDZL46 and pDZL18) in Zosterops lateralis and Z. lutea. Distributions are separated into east and west according to geographic divisions defined by mtDNA variation (Degnan & Moritz 1992). Alleles are listed according to endonuclease: E, EcoRI; B, BglII; T, TaqI. For each endonuclease, different alleles are identified by numbers 1 to 12, and their sizes are given. Numbers in parentheses above each column represent the total number of chromosomes scored. For Z. lutea, pDZL46-TaqI frequencies total less than 1.0 because two eastern individuals could not be scored. East–west allele distributions were compared using chi-square contingency tests performed on actual numbers of alleles for each endonuclease separately; counts for rare alleles were pooled (see text). Chi-square values were non–significant in all cases (P > 0.01).
nonetheless predominantly carried _Z. lutea_ scnDNA RFLPs. Only for pDZL46–EcoRI was one eastern _Z. lutea_ carrying a genotype (E1E2) found in eastern _Z. lateralis_, despite the species-specific genotype (E2E2) seen in all other _Z. lutea_. The contrasting patterns of variation between the mitochondrial and nuclear loci confirm that the mtDNA paraphyly is the result of hybridization between the two widely diverged (as demonstrated by the high level of mtDNA sequence divergence) species, as hypothesized previously (Degnan & Moritz 1992). In addition, results of the scnDNA analysis strongly suggest that the hybridization was historical rather than recent or current, especially in view of the fixed species differences at the pDZL18 locus. In the time since the two species hybridized, evidence of the hybridization has been virtually lost from the nuclear genome, presumably by recombination, but has been retained in the mitochondrial genome by virtue of its clonal maternal inheritance.

Secondly, within _Z. lateralis_, the scnDNA variation showed no significant geographic structure across the Australian continent. This was in contrast to the strong phylogeographic structure observed in the mtDNA variation, in which there was evidence of an historical separation of eastern and western populations. If the spatial separation of mtDNA haplotypes was due to long-term extrinsic barriers to gene flow, then the single-copy nDNA should reveal a similar geographic pattern: in fact, there is no clear evidence of a long-term east-west separation in patterns of scnDNA variation. Population subdivision is not expected to have the same quantitative effect on the geographic distribution of variability in nuclear and mitochondrial genomes, because of the different effective population size of the two genomes. The effective size for nuclear genes will be four times that for mitochondrial genes (if the sex ratio is 1:1, as it is expected to be in the monogamous silvereye), so that the mean time to fixation or loss of new mutations will be approximately two times longer for the nuclear genome (Birky, Maruyama & Fuerst 1983). The data for silvereyes are consistent with this theoretical proposal that the smaller effective size and faster mutation rate of mtDNA makes it more sensitive than scnDNA to population subdivision. It would appear that the biogeographical barriers that have produced the mtDNA differentiation between eastern and western silvereyes either have permitted sufficient gene flow to maintain relative panmixia in nuclear DNA, or have existed for insufficient time to produce differentiation at scnDNA loci.

It is now becoming clear that there is variation in the degree of concordance in the distribution of different molecular markers across taxonomic, and even biogeographic, boundaries. This is not unexpected, because each gene genealogy traces only a portion of the organismal pedigree linking individuals to their ancestors, and thus may vary from gene to gene. Indeed, simulations show little association between gene genealogies in the absence of long-term reproductive and/or geographic isolation between subpopulations (Fu & Arnold 1991). Accordingly, some authors (e.g. Wilson et al. 1985; Neigel & Avise 1986; Avise et al. 1987; Nei 1987) make a distinction between genealogies created by the analysis of particular genes or gene products, and the organismal pedigree through which those genes descend. The current study of variation at three independent loci in _Z. lateralis_ and _Z. lutea_ illustrates this discordance, and hence the need for data sets from more than one genetic system to infer organismal phylogenetics. This study provides clear empirical evidence that a single gene analysis can provide incomplete, and even misleading, phylogenies, and emphasizes the need for population genetic analyses to employ a composite of genetic markers under different modes of inheritance and different modes and rates of evolution.

Acknowledgements

L. Christidis (Museum of Victoria) and R. Schodde (Commonwealth Scientific and Industrial Research Organization) generously provided _Zosterops_ tissues from their collections. I am indebted to R. Zink for his valuable ideas and enthusiastic discussion regarding this work, to C. Moritz for providing a stimulating laboratory environment, and to J. Kikkawa for introducing me to _Zosterops_. Comments on the manuscript by B. Degnan, C. Moritz and R. Zink were greatly appreciated. This work was funded by the Australian Research Council and by a University of Queensland Postgraduate Research Award.

References


Nei M (1978) Estimation of average heterozygosity and genetic distance from a smaller number of individuals. Genetics, 89, 583.


After being introduced to Zosterops through Professor Jiro Kikkawa's long-term study of Z. lateralis on Heron Island, Australia, the author initiated work on evolutionary genetics of that species as a University of Queensland PhD student in the laboratory of Dr Craig Moritz. Although the Heron Island and related island populations formed its central focus, the molecular study was approached in a hierarchical manner, the macrogeographic tier of which is described by the work presented and referred to here.