The shifting roles of dispersal and vicariance in biogeography

Robert M. Zink1*, Rachelle C. Blackwell-Rago1 and Fredrik Ronquist2

1J. F. Bell Museum of Natural History, University of Minnesota, St Paul, MN 55108, USA
2Department of Systematic Zoology, Evolutionary Biology Centre, Uppsala University Norbyvägen 18D, SE-752 63 Uppsala, Sweden

Dispersal and vicariance are often contrasted as competing processes primarily responsible for spatial and temporal patterns of biotic diversity. Recent methods of biogeographical reconstruction recognize the potential of both processes, and the emerging question is about discovering their relative frequencies. Relatively few empirical studies, especially those employing molecular phylogenies that allow a temporal perspective, have attempted to estimate the relative roles of dispersal and vicariance. In this study, the frequencies of vicariance and dispersal were estimated in six lineages of birds that occur mostly in the aridlands of North America. Phylogenetic trees derived from mitochondrial DNA sequence data were compared for towhees (genus Pipilo), gnatcatchers (genus Polioptila), quail (genus Callipepla), warblers (genus Vermivora) and two groups of thrashers (genus Toxostoma). Different area cladograms were obtained depending on how widespread and missing taxa were coded. Nonetheless, no cladogram was obtained for which all lineages were congruent. Although vicariance was the dominant mode of evolution in these birds, approximately 25% of speciation events could have been derived from dispersal across a pre-existing barrier. An expanded database is now needed to estimate the relative roles of each process. Applying a molecular clock calibration, nearly all speciation events are of the order of a million or more years old, much older than typically presumed.

Keywords: aridland birds; North America; dispersal; vicariance; Pleistocene; molecular clocks

1. INTRODUCTION

Explaining the geographical deployment of biodiversity through time describes the science of historical biogeography. During the past 30 years, this science has endured several paradigm shifts, and concomitant partnership. For example, prior to 1970, most workers assumed that species’ ancestors dispersed across pre-existing barriers, became isolated, and evolved into new species (Udvardy 1969). In the late 1970s, a series of papers appeared in which a new paradigm was advanced. This, the vicariance school (Nelson & Platnick 1981), proposed that the major way in which biodiversity was generated was through the fragmentation of widespread ancestors by vicariant (isolating) events, such as uplifting of mountain ranges. Hence, dispersal occurred prior to the presence of the isolating barrier, not afterwards. Critics of the vicariance school argued that dispersal must play a more prominent role in the evolution of diversity, especially for such highly vagile organisms such as birds, whereas vicariance advocates (Rosen 1978) countered that dispersal should be invoked only after falsification of a vicariance model. Recently, methods of reconstructing biogeographical history have been proposed in which both dispersal and vicariance are allowed (Ronquist 1997). Hence, there is a growing plurality in the theoretical and methodological tools of biogeography. Nonetheless, few empirical studies document the relative roles of vicariance and dispersal (Lynch 1989). In the present study, these roles are estimated for a series of avian lineages that occur predominantly in the aridlands of North America.

*Author for correspondence (rzink@biosci.umn.edu).

Received 5 August 1999 Accepted 20 October 1999

It is widely agreed that the spatial component of speciation in most vertebrates involves allopatry (Chesser & Zink 1994). Both vicariance and dispersalist schools of biogeographical analysis are compatible with allopatric speciation being dominant but they differ in how they construe the interplay between dispersal and allopatry. In the vicariance paradigm, rare but extensive dispersal (range expansion) is followed by a series of allopatric isolation events (Bush 1975; Endler 1977), interrupted by occasional random dispersals. If the isolation events affect many organisms simultaneously, this process will generate congruent tree topologies. Dispersalists consider range expansion to be a more common and regularly occurring phenomenon, and both dispersal and vicariance processes are viewed as possibly resulting in predicted as well as unpredicted (random) events.

Testing hierarchical vicariance patterns requires phylogenetic analyses of multiple lineages, each containing at least three endemic taxa, chosen at random from a single large region, followed by assessment of congruence of tree topologies (Nelson & Platnick 1981). Congruence among phylogenetic trees is explained most parsimoniously by assuming that a single series of vicariant events fragmented widespread ancestors in each lineage in a common sequence, leading to modern-day phylogenetic congruence among geographical equivalents. Conflicting, or incongruent, trees can be explained by differential dispersal across pre-existing barriers. ‘Trees may also appear to conflict’ if they have unequal numbers of terminal taxa, which can result from lack of differentiation in response to a barrier (i.e. widespread species), or because some lineages have experienced extinctions; however, such trees can be compatible with vicariance.

The dispersalist paradigm is more difficult to falsify, and epistemologically it provides a complicated null
hypothesis. One can invoke dispersal to explain nearly any pattern of endemism. Typically a dispersalist argument assumes the unidirectional and idiosyncratic movement of a small number of founders into a new region followed by speciation, which must occur simultaneously in many lineages (to explain congruence among lineages). Hence, the dispersal paradigm requires several independent and rare events. Therefore, methods that assume a null hypothesis of vicariance are appropriate (Ronquist 1997). The strongest statements about dispersal events can be made when dispersal is rare and mixed with vicariance between areas of endemism. Under such conditions, there will be strong phylogenetic constraints on distribution patterns.

Molecular phylogenetic approaches enrich biogeographical analyses because molecular comparisons provide an indication of the time dimension over which speciation events have occurred (Riddle 1996). Assuming that, for example, mitochondrial DNA (mtDNA) divergence accrues at a roughly uniform and time-dependent rate, at least for the same gene within closely related taxonomic groups (Klicka & Zink 1997), sister species in different lineages separated by the same biogeographical barrier ought to exhibit similar mtDNA distances.

In an influential paper about biogeography of aridland birds of North America, Hubbard (1973) described evolutionary patterns in several lineages, including thrashers (Toxostoma), brown towhees (Pipilo), scaled quail (Callipepla) and black-tailed gnatchatchers (Pipilo). These genera contain species that are currently distributed in a highly consistent fashion among several areas of endemism in the aridlands of North America. Additionally, some species of warblers in the genus Vermivora (luciae group) have distributions relatively similar to the above genera. Hubbard (1973) proposed that cycles of Late Pleistocene glacial advance and retreat, allowed for isolation of ancestors in refugia and concomitant evolution of species in these lineages. That is, Hubbard outlined an explicitly vicariance explanation, implicating the two most recent glacial cycles (Wisconsinian and Illinoian) as isolating barriers responsible for speciation. However, Hubbard (1973) lacked explicit phylogenetic hypotheses or independent estimates of the timing of speciation for these groups.

Recent molecular studies provide these data (Zink & Blackwell 1998a,b; Zink et al. 1998, 1999; N. Klein and F. Gill, personal communication). In this study we compare phylogenetic patterns and ascertain degree of congruence, thereby testing whether a single series of vicariance events can explain current distributions. In addition application of a molecular clock allows testing of ideas about the ages of isolating barriers presumed to have influenced these groups. We provide one of the first assessments of the relative roles of dispersal and vicariance in shaping the biodiversity of a continental avifauna.

2. MATERIAL AND METHODS

We acknowledge from the outset that these lineages were not chosen randomly but rather because of the remarkable congruence in distribution noted by Hubbard (1973). The
aridlands were partitioned by Hubbard (1973) into several putative areas of endemism, of which we concentrated on five (figure 1): California plus Baja California (area A), Sonoran desert (area B), Chihuahuan desert (area C), Sinaloan shrubland (area D), and the highlands of southern Mexico (area E). Two other areas (F, G) contain single species. We recognize the difficulty in postulating areas of endemism (Ronquist 1994, 1997; Platnick 1991) but consider these areas to be useful for first-order testing.

Phylogenetic analyses are described in the publications cited above (see §1). In several of these studies, a single tree was not obtained, and here we evaluated some competing topologies not shown in figure 1. For the *Toxostoma curvirostre* complex, we included the taxa that Hubbard (1973) and previous taxonomists considered as members, even though it was not supported by mtDNA sequences (Zink et al. 1999); the tree favoured by mtDNA sequence data requires more dispersal events than that used here, but a large number of trees could not be eliminated with the data at hand). Many methods exist for estimating relationships among different areas (Morrone & Crisci 1995). For example, information from each lineage can be pooled by coding each node in each cladogram that is informative about area relationships, and then using maximum parsimony to derive a phylogenetic tree that interrelates areas. Nodes potentially informative about area relationships are labelled with a ‘T’ in figure 1. For example, *Pipilo aberti* and *Pipilo crissalis* and *Pipilo arbei* are sister species and occur in areas A and B, respectively. Thus character T1 constitutes a synapomorphy for areas A and B. Differences of opinion exist over how to code redundant, missing and widespread taxa. For example, in Brook’s parsimony analysis (Wiley 1988) widespread terminals and missing (absent) taxa. The data in table 1 yielded two trees (figure 2) of length 29 (CI 79, RI 60), the consensus of which is shown in figure 2c. Bootstrap support was 80% for areas A and B and 63% for areas C and D. Excluding characters 3, 8, 12 and 16 resulted in three trees (length 22, CI 86, RI 73), the consensus of which has the same topology as figure 2c. Bootstrap support was 92% for areas A and B and 32% for C and D; hence, bootstrap analysis of this reduced data set also supported the topology of figure 2c.

### Table 1. Data derived from informative nodes labelled with ‘T’ in figure 1

<table>
<thead>
<tr>
<th>Area</th>
<th>Node Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>101110001110110101111</td>
</tr>
<tr>
<td>B</td>
<td>111111110111111111111111</td>
</tr>
<tr>
<td>C</td>
<td>01110111011011111101111</td>
</tr>
<tr>
<td>D</td>
<td>?????011111????0010111??</td>
</tr>
<tr>
<td>E</td>
<td>0001011011???????????????</td>
</tr>
</tbody>
</table>

To estimate the number of dispersal events required to explain the current distribution of taxa beyond a strict hierarchical model, we used dispersal–vicariance analysis (Ronquist 1997), as implemented in the computer program DIVA (Ronquist 1996). We conducted two analyses in which the number of ancestral areas was unconstrained (favouring vicariance) or set to two (favouring dispersal). The latter analysis can help pinpoint the geographical location of the ancestral area, in a manner analogous to that proposed by Bremer (1992).

Because of the controversial nature of biogeographical analysis, we also used the program Component (Page 1993) to construct an area cladogram from the individual trees, and then to map the phylogeny for each species onto the area cladogram. However, the results of these analyses (e.g. area cladogram, and the fit of individual trees to it) did not differ qualitatively from those derived above. Also, this method does not *a priori* incorporate dispersal (Page 1994), and is not considered further here.

To provide a molecular perspective on speciation events, we computed Kimura (1980) two-parameter distances among taxa in each lineage, with an alpha value set to 0.75 (an empirical value derived for thrashers). We used 433 bp of cytochrome *b* data, except for *Vermivora* (N. Klein and F. Gill, personal communication) for which 1140 bp were used (analyses of *Vermivora* restricted to the same 433 bp fragment did not alter results). The distances were equated to time using the rate of 1.6%/per million years determined by Fleischer et al. (1998). This rate, for Kimura (1980) gamma-corrected two-parameter distances for Hawaiian honeycreepers, is relevant to our songbird taxa. If their calibration were correct, a distance of 3.2% or less would indicate that lineage-sundering originated in the Pleistocene (see Klicka & Zink 1999). We calibrated the galliform sequence divergence using a rate of 3.5% per million years (Zink & Klicka 1999). We plotted phylogenetic hypotheses that were consistent with our data if they were also suggested by Hubbard (1973); otherwise, we used those from our own papers (e.g. for the *Pipilo fuscus* group, where the phylogeny proposed by Hubbard is incorrect). In any event, the dates of divergences would not be any earlier with alternative tree topologies.

### 3. RESULTS

The topology of the area cladogram depended on the coding of widespread terminals and missing (absent) taxa. The data in table 1 yielded two trees (figure 2a,b) of length 29 (CI 79, RI 60), the consensus of which is shown in figure 2c. Bootstrap support was 80% for areas A and B and 63% for areas C and D. Excluding characters 3, 8, 12 and 16 resulted in three trees (length 22, CI 86, RI 73), the consensus of which has the same topology as figure 2c. Bootstrap support was 92% for areas A and B and 32% for C and D; hence, bootstrap analysis of this reduced data set also supported the topology of figure 2c. Assigning missing taxa (and the hypothetical ancestor) the code of ‘0’, and either including or excluding characters 3, 8, 12 and 16, resulted in the tree (length 31 or 24, CI 74 or 79, RI 71 or 62) shown in figure 2d. In the various analyses, the western region (Baja California plus California—aera A) was consistently the sister of the Sonoran desert (area B). Low bootstrap values for other nodes and conflicting area cladograms prevented identification of an unambiguous area cladogram.

DIVA identified the following number of dispersal events, depending on whether the maximum number of ancestral areas was unconstrained or set to two: *Calipepla* (0, 2), *Poliolota* (0, 2), *Pipilo* (1, 2), *Toxostoma lecontei* complex (2, 2), *T. curvirostre* complex (2, 4) and *Vermivora* (1, 2). Thus, at least six dispersal events (e.g. T2, T6, T7, T13, T14, T23) occurred over all lineages. Out of these six
dispersals, four or five occurred on internal branches (lineages) that were later split by allopatric speciation events. Because there are 16 internal nodes in the data set, this means that at least 25% of the speciation events in these birds were preceded by range expansion. Dispersal–vicariance analysis did not consistently identify a common ancestral area when the maximum number of areas allowed in ancestral distributions was limited to two.

Molecular results show that speciation events span a considerable period (figure 3). All speciation events in *Vermivora* are Pleistocene, whereas many of the more xeric-adapted species originated much earlier. Two species, *P. fusca* and *T. curvirostre*, show marked phyleogeographical differentiation (indicated by the dashed lines), as yet unrecognized taxonomically (but coded in figure 1 and table 1).

4. DISCUSSION

(a) Historical patterns and timing

Different area cladograms emerged from different treatments of the data (figure 2). Zink et al. (1998) suggested that some taxa in areas C, D and E, might have been involved in trichotomous (contemporaneous) splits, which is consistent with figure 2c. However, individual cladograms do conflict. For example, the two endemic taxa in area D have different sister-taxon (area) relationships. Also, the sister-taxon relationships of the two endemic taxa in area E are unresolved, but in the studies discussing their relationships it can be inferred that they probably conflict. Thus, a strong signal about area relationships (except for A and B) was not obtained from the individual trees.

Comparing tree topologies for the different lineages (figure 1) does support two consistent biogeographical patterns, namely that species in California and Baja California (area A) are sister species to those found in the Sonoran desert (area B), and that taxa currently considered conspecific (and whether differentiated or not) are widespread in the Sonoran and Chihuahuan deserts.

The sister-taxon pattern between California–Baja California and the Sonoran desert is of interest because of its generality. An expectation of speciation events caused by the same isolating barrier is that taxa should differ by approximately the same level of genetic distance. For example, speciation events considered by Hubbard (1973) to be contemporaneous are coded with identical boxes in figure 3. Sequence divergence values for the sister species in areas A and B range from 1.0 to 6.0%. It is unclear if this degree of variation is what one would expect from sister taxa isolated contemporaneously, given variation around mean sequence divergences. Some taxa might initially have been more differentiated, and stochastic sampling of haplotypes can yield trees of differing depths even if the initial average divergence values are the same (Neigel & Avise 1986). Evolutionary rates of sequence divergence might differ. However, such explanations seem insufficient to explain two species pairs that stand out as being relatively recent and old, namely *Vermivora luciae–Vermivora ryficappilla* and *Polioptila californica–Polioptila melanura*, respectively.

A potential reason for lack of congruence in timing of splits between areas A and B could be that different habitats were isolated at different times (Riddle 1996). The *Vermivora* warblers have different habitat requirements than some other desert taxa considered here, and the barrier between areas A and B might not have become completed for warblers until more recently. The split between *Polioptila melanura* and *Polioptila californica* seems to pre-date other sister taxa with the same biogeographical
Thus, these two regions contain well-differentiated sister species as well as widespread species, which are either differentiated or not. Molecular comparisons (figure 3) reveal that taxa in these areas represent at least two tiers of evolutionary diversification, namely an older tier between sister species in areas A and B and a more recent episode within widespread biological species in areas B and C (T. curvirostre, P. fuscus). Thus, area B has played a role in diversification at two different times in avian evolution.

If an area contains taxa of demonstrably different ages, construction of area cladograms could be affected. For example, taxa currently in some areas, such as B, differentiated at different times (figure 3). However, nodes in individual trees (figure 1) are coded equivalently in the construction of area cladograms, irrespective of the timing of the speciation events. Clearly some nodes conflict, such as those grouping A and B, and B and C. If one realizes that speciation patterns involving these areas occurred at two different points in time, estimation of vicariance and dispersal is made more complicated. It would seem logical in our case that more recent speciation events involving these areas would be attributable to dispersal. Coding of taxa in such areas deserves further analysis.

Given plausible calibrations of molecular divergences, the hypothesis by Hubbard (1973) that speciation in these aridland birds was initiated by the glacial cycles of the Illinoian (250 000 years before present) and Wisconsinian (100 000 years before present) periods is probably incorrect. Instead, divergence dates appear to be much older (Klicka & Zink 1997, 1998, 1999). More work on calibrations, and especially variances around single-locus clocks, is needed (S. Edwards, unpublished data).

(b) Dispersal–vicariance analysis

Given uncertainty over the topology of the area cladogram, estimation of dispersal from it is difficult. Results of dispersal–vicariance analysis (using DIVA), which is not based on an area cladogram, revealed that a strict mode of hierarchical vicariance can be ruled out for at least three (P. fuscus, T. curvirostre, T. lecontei) and probably four (V. ruficapilla) of six lineages. Hence, the question is not dispersal ‘or’ vicariance, but rather the relative importance of each. In our analysis, at least 25% of speciation events appear to have resulted from dispersal, although if molecular data are considered, perhaps one or two of the apparent vicariant events can be attributed to dispersal (see § 4(a)). Thus, we consider 25% to be a lower bound. We think it significant that even in these largely sedentary birds, dispersal is demonstrably important.

We are aware that many alternative methods exist for coding ranges and cladograms. For example different coding schemes influenced the topology of the area cladogram. However, we suggest that given our data, a coding scheme that altogether ruled out either vicariance or dispersal would probably be considered biologically unrealistic by many biogeographers.

The role of vicariance in this avifauna is relevant to an ongoing discussion between historical biogeographers and palaeoecologists (Riddle 1998). Brown (1995, p.191) stated ‘The geographic ranges of the species in many groups have probably experienced so many shifts since the speciation events that it would be difficult or
impossible to reconstruct the spatial pattern of the history of lineages. Specifically, in the North American fauna, one might wonder whether Late Quaternary range shifts caused by glacial cycles have erased the biogeographical (speciation) record present in organismal phylogenies. If species ranges are indeed ephemeral, and widespread ancestral biotas are temporally not coherent assemblages of species, then one would indeed not expect to recover phylogenetic patterns consistent with vicariance (Cracraft 1994). The relatively large role ascribed to vicariance in this study suggests that ranges are sufficiently stable to preserve the correlation between organismal phylogenies and the areas in which they have evolved; a similar result was reported by Cracraft (1991). The apparent depths of gene trees (figure 3) suggest range stability over at least a million years. Riddle (1998) reached a similar conclusion in his study of mammals in North America. Hence, the vestiges of history are apparent in organismal phylogenies for considerable periods, despite the known fluidity of ranges and the faunistic displacements caused by Late Pleistocene glacial cycles.

(c) Conclusions

The power of our study is limited by the fact that some areas of endemism are represented in only two of six lineages (e.g. areas D and E). The lack of strong congruence suggests that these lineages, chosen a priori because of their striking patterns of distributional congruence, have been influenced by both vicariance and dispersal. Whether this is typical of most continental avifaunas remains to be shown (Cracraft 1988). It will be of interest to estimate the relative roles of dispersal and vicariance in more faunas.

Many people who assisted with the studies synthesized here are thanked in the component papers. Much credit is due to John Hubbard for his exceptional paper on aridland biogeography: We thank N. Klein and F. Gill for their Vernivora data. The research was supported by the American Museum of Natural History (Chapman Fellowship to R.M.Z.) and by the National Science Foundation (DEB9317945). J. Klicka, D. Wake, J. Bates, E. Birney, G. Voelker, B. Riddle, A. Simons, F. Gill, N. Klein and A. Kessen made useful comments on the manuscript.

REFERENCES


As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.