SYSTEMATICS AND HISTORICAL BIOGEOGRAPHY OF WAGTAILS: DISPERSAL VERSUS VICARIANCE REVISITED

GARY VOELKER[1]

Barrick Museum of Natural History, Box 454012, University of Nevada Las Vegas, Las Vegas, NV 89154

Abstract. Nucleotide sequences of the mitochondrial cytochrome *b* and ND2 genes were used in analyses of phylogenetic relationships of 10 of the 11 currently recognized species of wagtails (Aves: *Motacilla*). All analyses produced similar hypotheses of species relationships. Both *Motacilla citreola* and *M. flava* are paraphyletic, and consist of two and three distinct clades, respectively. *M. lugens* and *M. alba* are paraphyletic with respect to one another. None of the three currently recognized superspecies are monophyletic. A series of alternative tree topologies on which species and superspecies monophyly were enforced were significantly worse estimates of relationships in all cases except that of *lugens* and *alba*. Ancestral area reconstructions suggest that *Motacilla* arose in the eastern Palearctic; molecular-clock dates suggest that this occurred roughly 4.5 million years ago. *Motacilla* colonized Africa around 4.5 million years ago, and accomplished a second ancestral colonization of that continent about 2.9 million years ago. A number of recent dispersals from Eurasia to North Africa and North America (via both Alaska and Greenland) have occurred, and no fewer than 26 dispersals have occurred throughout the evolution of the modern distribution of this genus. This result, similar to results of other recently studied avian lineages, strongly suggests that dispersal has been an important factor in the development of modern-day avian distributions.

Key words: colonization, dispersal, intercontinental biogeography, mitochondrial DNA, molecular clock, Motacilla, *vicariance.*

Sistemática y Biogeografía Histórica de *Motacilla*: Revisión de Dispersión Versus Vicarianza

Resumen. Se usaron secuencias de nucleótidos de los genes mitocondriales citocromo *b* y ND2 en análisis de relaciones filogenéticas en 10 de las 11 especies de aves actualmente reconocidas del género *Motacilla*. Todos los análisis produjeron hipótesis similares sobre las relaciones entre las especies. Tanto *M. citreola* como *M. flava* son parafiléticas, con dos y tres clados distintivos, respectivamente. *M. lugens* y *M. alba* son parafiléticas en relación a una con la otra. Ninguna de las tres super-especies actualmente reconocidas es monofilética. Una serie de árboles topológicos alternativos, en los cuales se forzó la monofilia de las especies y super-especies, brindaron estimaciones de las relaciones significativamente peores para todos los casos, con excepción de aquella para *lugens* y *alba*. Reconstrucciones ancestrales de área sugieren que *Motacilla* surgió en el este Paleártico; las fechas del reloj molecular sugieren que esto ocurrió hace aproximadamente 4.5 millones de años. *Motacilla* colonizó Africa alrededor de 4.5 millones de años atrás, y protagonizó una segunda colonización ancestral de este continente cerca de 2.9 millones de años atrás. Una serie de dispersiones recientes han ocurrido desde Eurasia hacia Africa del Norte y América del Norte (vía Alaska y Groenlandia), y no menos de 26 dispersiones han ocurrido a lo largo de la evolución de la distribución moderna de este género. Este resultado, similar al de otros estudios recientes de linajes de aves, sugiere con firmeza que la dispersión ha sido un factor importante en el desarrollo de las distribuciones modernas de aves.

INTRODUCTION

The avian genus *Motacilla* comprises 11 species, whose combined distribution includes the continents of Africa, Eurasia, and North America, as well as several major islands (Sibley and Monroe 1990). As such, *Motacilla* is one of a relatively small number of songbird genera whose component species are distributed on three or more continents. By having such a distribution, *Motacilla* provides an interesting opportunity to explore questions concerning the relative roles of dispersal and vicariance in the establishment of the genus' modern-day intercontinental distribution.

The obvious key to assessing the roles of dispersal and vicariance is molecular-clock dating of lineage divergences. While these clocks can

be problematic (Martin et al. 1992), their ability to provide rough estimates of the time of lineage splits allows us to seek similarly timed geological and climatic events; this may be especially true when dealing with more recent lineages and geological events. Similarly timed lineage splits and geological events would implicate a vicariant speciation mechanism, whereas a much younger lineage split relative to a plausible geological event would imply that a dispersal had occurred. Such clock calibrations have become increasingly important in testing explicit vicariant and dispersal models (Bermingham et al. 1992, Bowen and Grant 1997, Waters et al. 2000), and in assessing the relative roles of these often competing mechanisms (Voelker 1999a).

The distribution of *Motacilla* also allows us to ask whether geographic species assemblages within the genus are the result of multiple invasions or in situ speciation. The key to assessing the manner in which species assemblages form is simple: one needs only a well-supported phylogenetic hypothesis, from which area cladograms can be constructed. To date, species assemblages in several songbird lineages have been shown to be derived from multiple invasions (Richman 1996, Slikas et al. 1996, Voelker 1999a). These results have important implications for the determination of ancestral areas, especially when ancestral areas are simply deemed to be that area with the largest number of species (i.e., the "center of origin"). These results also render a cautionary note for studies attempting to determine whether the concordance of modern-day species distributions and areas indeed reflect vicariant patterns, or whether these patterns are instead a reflection of distributions modified via past dispersal events.

Finally, the answers to the above two questions can be combined to allow us to determine whether the results are broadly comparable to studies of other avian lineages with similar distributions (Richman 1996, Voelker 1999a; Voelker, in press). By assessing these questions on a cosmopolitan scale, there is almost certainly a better likelihood of correctly identifying dispersal or vicariance mechanisms than there might be at smaller scales (e.g., intracontinental). However, studies largely focused on a cosmopolitan scale have shown the above key elements to be applicable at smaller biogeographic scales as well (Voelker 1999a); thus these elements are indeed worth including in any biogeographic study.

To date, there are no phylogenetic hypotheses of the relationships among *Motacilla* species available, from which dates can be gleaned and biogeographic interpretations made. Although Keith et al. (1992) have postulated relationships based on plumage differences, these relationships have been only broadly expressed through the recognition of superspecies or general groupings, and do not hypothesize sister relationships among *Motacilla* species. Therefore, the current state of our knowledge of *Motacilla* relationships is the recognition of three superspecies. The first of these, the yellow wagtail group, consists of four species: *flava, citreola, capensis*, and *flaviventris*. Both *flava* and *citreola* breed primarily in Eurasia, while *capensis* is found in southern Africa, and *flaviventris* is endemic to Madagascar. The stream wagtail superspecies consists of *clara* and *cinerea*; the former being found in southern Africa, the latter primarily in Eurasia. Lastly, the pied wagtail superspecies consists of *alba, lugens, aguimp, madaraspatensis*, and *grandis*. Both *alba* and *lugens* breed in Eurasia, while *aguimp* is found in southern Africa, *madaraspatensis* on the Indian subcontinent, and *grandis* in Japan (Sibley and Monroe 1990, Keith et al. 1992).

To resolve the relationships among *Motacilla* species, I used 1966 base pairs of mitochondrial DNA sequence data from individuals representing 10 of 11 currently recognized species. I used these data to generate phylogenetic hypotheses of relationships, which in turn provided a framework in which to explore the historical biogeography of the genus.

METHODS

TAXA, SEQUENCING, AND PHYLOGENETIC ANALYSIS

I analyzed 960 base pairs (bp) of cytochrome *b*, and 1006 bp of ND2 from 33 individuals of *Motacilla*, and two outgroup species (Table 1). The 33 *Motacilla* taxa represent 10 of the 11 species recognized by Sibley and Monroe (1990). At least two samples were included for all *Motacilla* species except for *madaraspatensis*. Higher numbers of samples for *flava* and *alba* were included as an effort to corroborate findings from ongoing phylogeographic studies of those species (Zink et al., unpubl. data), in the larger framework of the entire genus. The six *alba*

TABLE 1. Species, museum voucher specimen numbers, and collecting localities for *Motacilla* and *Anthus* specimens examined.

Species	Museum voucher[a]	Collection locality
Motacilla		
alba	UWBM 49718	Russia: Murmanskaya Oblast'
	UWBM 59795	Mongolia: Hentiy Aymag
	UWBM 44255	Russia: Avtonomnaya Respublika Yakutia
	UWBM 44591	Russia: Kamchatka
	UWBM 47169	Russia: Khabarovskiy Kray
	UWBM 46304	Russia: Gorno-Altay Republic
lugens	UWBM 44605	Russia: Kamchatka
	UWBM 47267	Russia: Sakhalinskaya Oblast'
madaraspatensis	AMNH23223	Nepal: Betrabati
aguimp	MBM 5840	South Africa: North Cape Province
	UWBM 53198	South Africa: KwaZulu/Natal Province
capensis	MBM 5842	South Africa: North Cape Province
	UWBM 53145	South Africa: KwaZulu/Natal Province
flaviventris	FMNH 352832	Madagascar
	FMNH 352834	Madagascar
citreola	UWBM 49351	Russia: Kurskaya Oblast'
	UWBM 46336	Russia: Avtonomnaya Respublika Buryatia
flava	UWBM 49434	Russia: Kurskaya Oblast'
	UWBM 46234	Kazakhstan: Almaty Oblysy
	UWBM 43953	Russia: Chukotskiy Avtonomnyy Okrug
	UWBM 46360	Kazakhstan: Almaty Oblysy
	UWBM 46381	Kazakhstan: Almaty Oblysy
	UWBM 44273	Russia: Avtonomnaya Respublika Yakutia
	UWBM 47176	Russia: Khabarovskiy Kray
	UWBM 56836	Russia: Astrakhanskaya Oblast'
	UWBM 51665	Russia: Magadanskaya Oblast'
	UWBM 47504	Russia: Sakhalinskaya Oblast'
	UWBM 47505	Russia: Sakhalinskaya Oblast'
cinera	UWBM 46556	Kazakhstan: Alma-Ata Oblys
	UWBM 44455	Russia: Magadanskaya Oblast'
clara	FMNH 355292	Uganda: Kasese District
	UWBM 52913	South Africa: KwaZulu/Natal Province
	UWBM 53178	South Africa: KwaZulu/Natal Province
Anthus		
cinnamomeus	UWBM 52816	South Africa: East Cape Province
brachyurus	UWBM 52901	South Africa: KwaZulu/Natal Province

[a] MBM, Marjorie Barrick Museum, University of Nevada Las Vegas; UWBM, University of Washington Burke Museum; FMNH, Field Museum of Natural History; AMNH, American Museum of Natural History.

samples represent five collecting localities across Russia, and one from Mongolia; the 11 *flava* samples represent seven collecting localities across Russia, and one from Kazakhstan. The outgroup taxa represent two species of the genus *Anthus*, a close relative of *Motacilla* within the family Motacillidae (Voelker and Edwards 1998).

Nineteen tissue samples were prepared via Qiagen extraction (total genomic DNA), according to manufacturer's protocol (Qiagen Inc., Valencia, California); most samples of *flava* and *alba* were isolated at the University of Minne-sota (R. Zink lab, Bell Museum of Natural History, St. Paul). Fifteen frozen tissue samples were prepared via cesium chloride gradient (Dowling et al. 1990), which minimizes the chance of amplifying nuclear mitochondrial sequences. Primers and PCR conditions followed those outlined in Voelker and Edwards (1998) and Voelker (1999b). All sequences were generated on an ABI 377 automated sequencer (PE-Biosystems, Forest City, California). Sequences are deposited in GenBank under the following accession numbers: AF526435–AF526469 (cytochrome *b*) and AF526470–AF526504 (ND2).

Future use of these samples should cite correct museum voucher specimen numbers provided herein, and in GenBank.

All phylogenetic analyses were performed using PAUP* (Swofford 1999). I used MODEL-TEST (version 3.06; Posada and Crandall 1998) to evaluate models of DNA sequence evolution for input into maximum-likelihood (ML) analyses; GTR + I + Γ was identified as the best-fit model to my data. However, I used four different maximum-likelihood (ML) analyses to generate trees, in order to explore the robustness of the data, and tree topologies, to different weighting schemes. First, I generated a neighbor-joining tree, with the HKY85 model of evolution (Hasegawa et al. 1985), to determine overall values of kappa and alpha. These values were then used in a ML analysis. Second, I conducted a ML analysis that estimated values of kappa and alpha while searching for the optimal tree topology. Third, I used a GTR + I + Γ model, in which model values were estimated. Fourth, I conducted a ML analysis that estimated site-specific rate values (Buckley et al. 2001) for each codon partition (three cytochrome *b*, three ND2), under the HKY85 model. All analyses were heuristic searches, using TBR and as-is addition of taxa, except the GTR model, for which I used a neighbor-joining tree as a starting point. All analyses were run until at least 1000 rearrangements (250 for the GTR model) had been performed without a change in likelihood score. I also conducted a heuristic parsimony search, in which I assigned codon-specific transition/transversion ratios that were estimated on a neighbor-joining tree. Cladistic signal was determined for each clade by bootstrapping (500 replicates and 10 random additions). Trees from all analyses were rooted to the two *Anthus* taxa.

I used the parametric bootstrap method (Goldman et al. 2000, Sullivan et al. 2000) to compare likelihood estimates from 100 replicate searches to likelihood scores derived from alternate tree topologies that reflect competing hypotheses as to the relationships among *Motacilla* species; I also used these tests to explore species paraphyly detected in analyses (see Results). The 100 replicate searches were performed on data sets generated in Seq-Gen (version 1.2.5; Rambaut and Grassly 1997), using observed base frequencies and GTR + I + Γ parameters. Because the parametric bootstrap computations are time intensive, I used a pruned data set which includ-ed no more than two individuals from each *Motacilla* species, and both outgroup taxa. I also conducted a new heuristic GTR + I + Γ search on the reduced dataset for the parametric bootstrap comparisons. I used MacClade (Maddison and Maddison 1992) to create alternative tree topologies.

ANCESTRAL AREAS AND MOLECULAR CLOCK

To determine the ancestral area and assess the historical biogeography of *Motacilla*, I converted a ML phylogeny to an area cladogram. There are only three broad-scale areas possible as a *Motacilla* ancestral area: Africa, Madagascar, and Eurasia, and I used these areas in initial analyses. Although North America is included in several species distributions, the distribution of each on North America is restricted to a very small fraction of each species' entire range; therefore North America can reasonably be excluded as an ancestral area.

To attempt to more accurately define an ancestral area, and to assess the number of dispersals necessary to explain modern *Motacilla* distributions, I refined the breeding distribution of each species to more accurately reflect within-continent distributions, and included several other regions (see below). As I have discussed previously (Voelker 1999a), there appear to be few barriers to movement for many avian species within Eurasia. Thus, to make this study generally comparable to other studies, I used the five Eurasian areas that I delimited for the genus *Anthus* (see Fig. 1A in Voelker 1999a): (1) western Palearctic; (2) eastern Palearctic, (3) Mongolia-southeast Asia, (4) the Himalayas, (5) southwestern Asia. Within sub-Saharan Africa, I used the six areas delimited by Crowe and Crowe (1982), with slight modification (see Fig. 1C in Voelker 1999a). Finally, to account for all *Motacilla* distributions, I also include North America, Africa north of the Sahara, and the Canary and Madeira Island groups (Canary Islands hereafter).

Given the results of phylogenetic analyses (see Results), areas for the three *flava* and two *citreola* taxa were somewhat subjective, and necessarily delineated according to samples included here, and in the more inclusive analysis of these species being conducted elsewhere (Zink et al., unpubl. data).
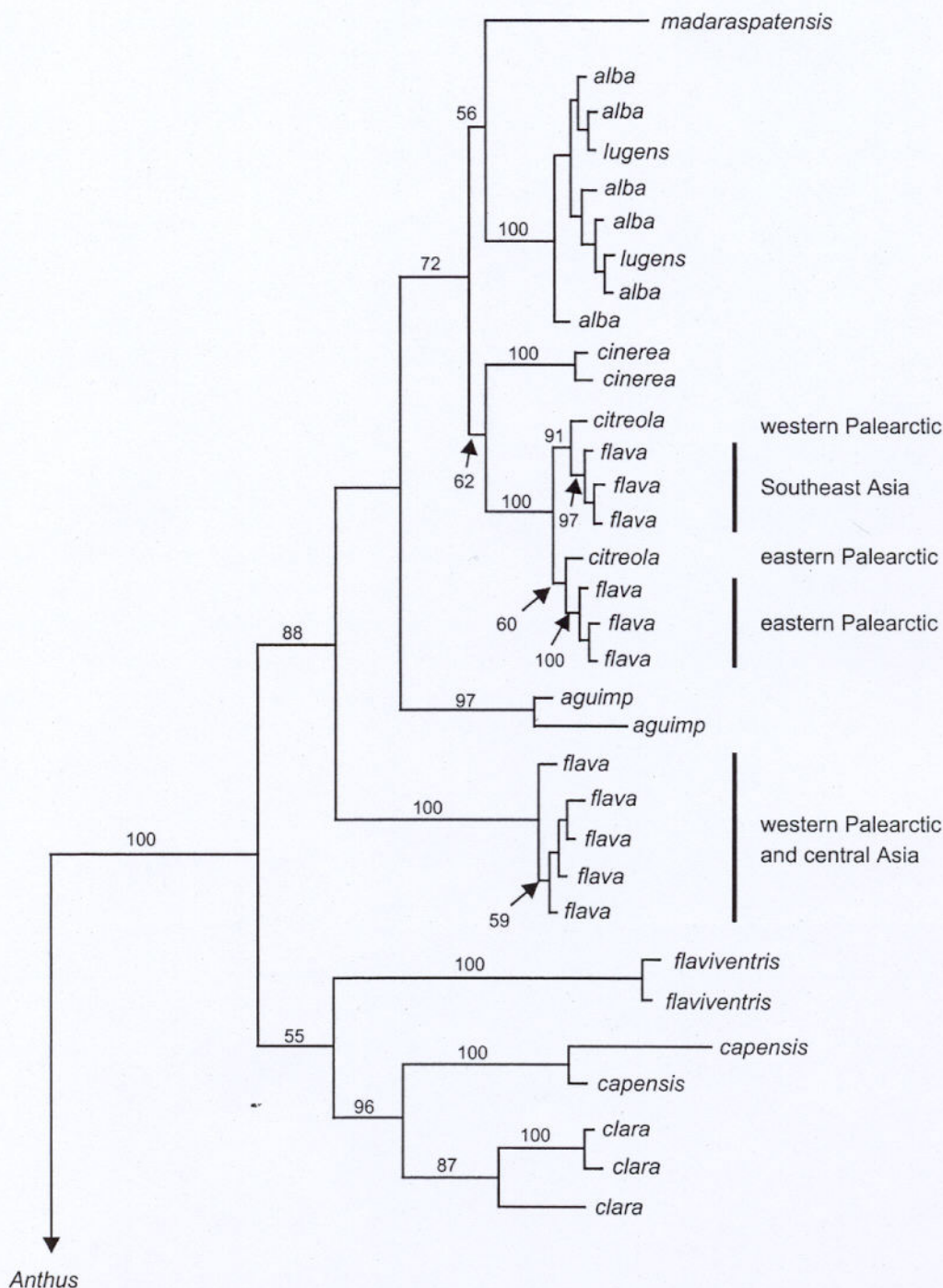
FIGURE 1. Maximum-likelihood tree of *Motacilla* relationships, based on the GTR + I + Γ model of evolution. Numbers at internodes indicate frequencies of occurrence (>50%) for groups in 500 maximum-parsimony bootstrap replicates. Also depicted are broad geographic regions (based on collecting localities) for each subset of those species previously thought to be monophyletic, but which are shown here to be paraphyletic.

Using breeding distributions, I employed dispersal-vicariance analysis (DIVA; Ronquist 1996, 1997) to reconstruct ancestral distributions on the phylogeny (which was pruned to hold just one individual per species or group), and to determine the number of dispersal events. When the ancestral distribution was determined by DIVA to be composed of all, or nearly all, possible areas used to define the distribution of *Motacilla*, I used the maximum areas option to limit the range of ancestral distributions to no more than two areas. This approach determines what the most likely ancestral distribution would be if that distribution were restricted to smaller areas (Ronquist 1997), as historically it must surely have been. In addition to DIVA, I used the ancestral areas method of Bremer (1992). This method is cladistic, relies only on the topological information of the area cladogram, and makes no assumptions as to speciation mechanism. Both methodologies are based upon parsimony criteria.

To test whether lineages within *Motacilla* are evolving in a clocklike fashion, I applied the two-cluster test (Takezaki et al. 1995), with Tajima-Nei distance, to the GTR + I + Γ maximum-likelihood topology. This test assesses whether the two daughter lineages above each node in a given phylogenetic hypothesis are diverging in a clocklike fashion. Here again, I used a pruned tree, including no more than two individuals per species, in order to reduce the difficulty involved in constructing and inputting the ML tree into the test. From species with multiple representatives, individuals from both a basal and a terminal divergence were included; single exemplars were used only if a species cluster showed virtually no divergence (e.g., within some *flava* taxa). If a molecular clock could be inferred at a node linking two daughter lineages, I used the 2% sequence divergence per million years value inferred from several studies of disparate avian lineages (Shields and Wilson 1987, Tarr and Fleischer 1993, Fleischer et al. 1998).

RESULTS

PHYLOGENETIC ANALYSES

I aligned 1966 bp without insertions or deletions. Because (1) the entire cytochrome *b* and ND2 fragments considered here were isolated as a single unit for amplification, (2) there was no difficulty translating sequences into amino acids, and (3) there was a lack of alignment problems and stop codons, the genes amplified were most likely of mitochondrial origin. There were 239 variable sites (165 parsimony informative) in the cytochrome *b* data, and 322 (231 parsimony informative) in the ND2 data. Further details of sequence evolution are available upon request.

All methods of ML analysis resulted in the same tree (Fig. 1), with respect to species relationships. ML trees varied only in the relative position of individuals within the *alba* and *flava* groups. The initial analysis using assigned, overall values of kappa and alpha resulted in a log-likelihood score of -8196, whereas estimation of overall kappa and alpha values during searches gave a log-likelihood score of -7997, GTR + I + Γ resulted in a log-likelihood score of -7987, and site-specific values by codon position resulted in a log-likelihood score of -7764.

Parsimony analysis resulted in a single best tree, generally similar to the ML trees, with moderate or better bootstrap support at most species divergences (Fig. 1). Note that even were the moderately supported nodes excluded, the paraphyly detected in several currently recognized taxa (discussed below) would still be highly supported at other nodes.

Four species (*flava, citreola, alba,* and *lugens*) were paraphyletic. The *flava* samples included here fall into three distinct clades. One of these clades contains individuals from the western Palearctic and central Asia, a second contains individuals from southeastern Asia and is most closely related to a western Palearctic *citreola*, and the third clade contains individuals from the eastern Palearctic, and is most closely related to an eastern Palearctic *citreola* (Fig. 1). Parametric bootstrap tests of two trees on which all *flava* taxa were constrained to be closest relatives resulted in significantly worse estimates of the data, when compared to the best ML tree (Fig. 2). One of these alternative trees placed all *flava* together at the western Palearctic *flava* position on the best tree (Fig. 1; denoted as "basal" in Fig. 2), the other placed all *flava* together as sister to *citreola* (Fig. 1; denoted as "tip" in Fig. 2). A parametric bootstrap test also found significantly worse a tree forcing *citreola* to be monophyletic (Fig. 2).

Exemplars of *lugens* and *alba* were not each other's closest relatives, with each *lugens* being more closely related to different *alba* lineages
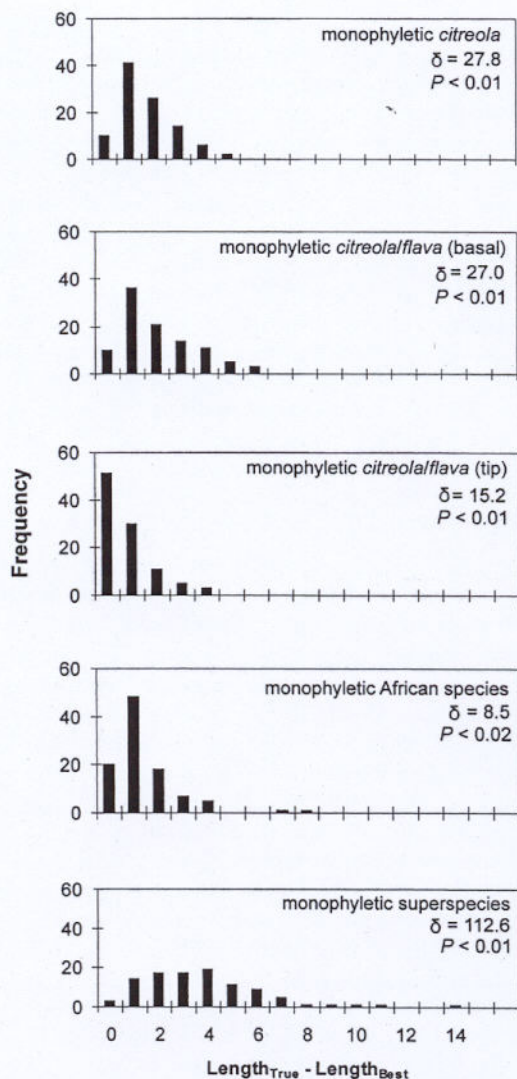
FIGURE 2. Parametric bootstrap tests of alternative topologies. One hundred replicate data sets were simulated for each alternative topology, and searches were conducted under maximum-likelihood criteria. Plots show the distribution of log-likelihood differences obtained for the 100 replicate data sets ($Length_{best}$), as compared to each respective alternative topology ($Length_{true}$). Scores in the top right of each box indicate the log-likelihood difference between each alternative topology and the best estimate of relationships (i.e., Fig. 1, pruned as outlined in Methods), and the associated P-value, based on the actual data set. In each instance, alternative topologies can be rejected as significantly worse estimates of the data.

(Fig. 1). All ML analyses supported paraphyly of these two species, but I did not perform a parametric bootstrap test of this question, as the branch lengths among individuals is extremely short, leaving the possibility of monophyly unlikely to be rejected.

Results suggested that *madaraspatensis* is closely related to *alba* (Fig. 1). The African species were clearly not part of a single radiation; constraining a tree to place *aguimp* with the other African taxa (*capensis* and *clara*) produced a significantly worse estimate of relationships when compared to the ML tree (Fig. 2). Constraining a tree to reflect traditional superspecies relationships also proved a significantly worse estimate of the data (Fig. 2).

MOLECULAR CLOCK AND DATING NODES

Within *Motacilla*, 9 of 11 nodes at which species or group divergences occur are evolving in a manner consistent with a molecular clock (Fig. 3). Only the *clara/capensis* and *madaraspatensis/alba/lugens* lineages are not diverging in a clocklike fashion (Fig. 3).

Dating of nodes (other than intraspecific nodes; Fig. 3) suggests that *Motacilla* arose about 4.5 million years ago, and that five divergences occurred in the late Pliocene, and four divergences occurred since the start of the Pleistocene, about 1.7 million years ago (Fig. 3). Although the date for the *flaviventris* clade is slightly older than the date for the origin of *Motacilla* as a whole (Fig. 3), the associated standard errors overlap considerably, which suggests a rapid radiation.

ANCESTRAL AREAS AND INTERCONTINENTAL MOVEMENTS

Results of dispersal-vicariance analysis suggested that either the eastern Palearctic or southwestern Asia, in combination with either Madagascar, South Africa, or the Southern Savannah Zone of Africa, constitutes the most plausible ancestral area for *Motacilla* (Fig. 4).

Bremer's (1992) ancestral area method provided a more restricted view of a *Motacilla* ancestral area, by suggesting two things. First, that *Motacilla* probably arose in Eurasia (Table 2A); and second, that when multiple areas within Eurasia are considered, the most probable area of origin for *Motacilla* is the eastern Palearctic (Table 2B).
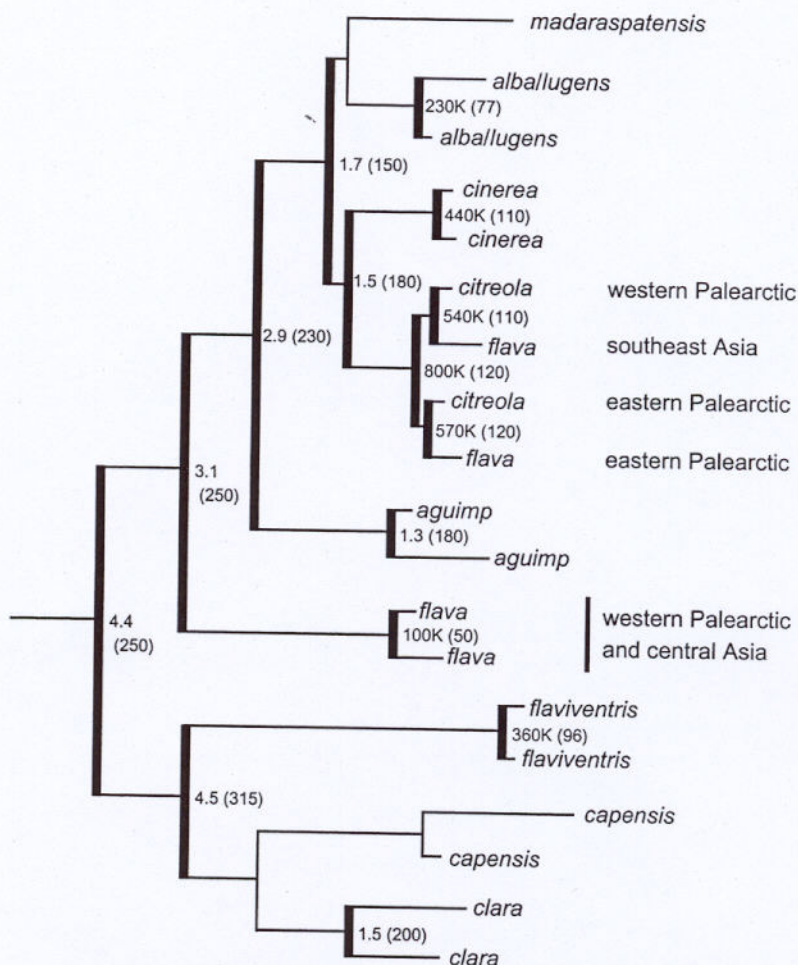
FIGURE 3. Maximum-likelihood topology showing nodes above which daughter lineages are (thick bars) or are not (narrow bars) consistent with a molecular clock. Numbers at nodes represent time in millions of years, or where denoted (with K), thousands of years; numbers in parentheses are standard error estimates, in thousands of years. Both time and consistency with a molecular clock were determined with the two-cluster test.

DIVA inferred 26 dispersal events to account for the modern-day distribution of *Motacilla* species. This number of dispersals is based on the assumption that any or all 14 defined distributional areas can form the ancestral area; under this assumption, all 14 are defined as the ancestral area. Restricting the number of possible ancestral areas to two (the minimum allowed in DIVA), under the assumption that ancestral areas surely did not consist of entire modern distributions, results in 35 dispersals being necessary to explain the modern distributions of *Motacilla* species.

There have been several Eurasia to Africa interchanges during the pre-recent history of *Mo-*

tacilla (Fig. 5). *Motacilla* arrived on Africa very soon after its origin in the eastern Palearctic roughly 4.5 million years ago, and colonized Madagascar shortly thereafter (Fig. 5). A second colonization of Africa occurred 2.9 million years ago (Fig. 5). Seven additional intercontinental and continent to island colonizations have occurred recently (Fig. 5).

DISCUSSION

TAXONOMIC IMPLICATIONS

All methods of analysis resulted in the same tree of *Motacilla* species relationships, with paraphyly in *flava*, *citreola*, *lugens*, and *alba*. The paraphyly in *flava* is not surprising, given that
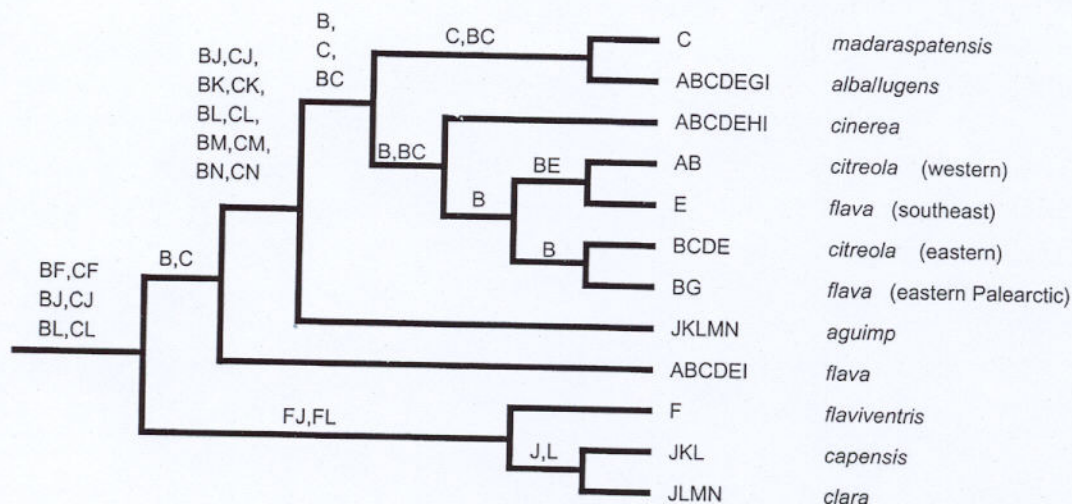
A = Western Palearctic
B = Eastern Palearctic
C = Southwest Asia
D = Himalayas
E = Mongolia-Southeast Asia

F = Madagascar
G = North America
H = Canary Islands
I = Africa, north of the Sahara
J = South Africa

K = Southwest Arid Zone
L = Southern Savannah Zone
M = Forest Zone
N = Northeast Arid Zone

FIGURE 4.   Ancestral area reconstruction using dispersal-vicariance analysis, for all *Motacilla* taxa included in this study, with the maximum number of possible ancestral areas limited to two. Letters K–N refer to regions within sub-Saharan Africa (see Methods).

TABLE 2.   Ancestral area reconstructions based on Bremer's (1992) method. Ancestral area scores represent probabilities that an area is part of the ancestral area for the group. In (A) only the three most probable major landmasses were used to identify a general ancestral area for *Motacilla*. Based on this result, in (B), Eurasian areas were included to refine the likely *Motacilla* ancestral area within Eurasia.

Area	Gains	Losses	Gains/ losses	Ancestral area
(A)				
Africa	2	3	0.67	0.67
Madagascar	1	2	0.5	0.5
Eurasia	2	2	1	1
(B)				
Africa	2	3	0.66	0.53
Madagascar	1	2	0.50	0.40
Western Palearctic	4	5	0.80	0.64
Eastern Palearctic	5	4	1.25	1
Southwest Asia	4	4	1	0.80
Himalayas	4	4	1	0.80
Mongolia-Southeast Asia	5	5	1	0.80

current taxonomy recognizes 14 distinct races; three races of *citreola* are also recognized (Cramp 1988). Results of this study show that *flava*, as currently recognized, comprises three clades, which are not sisters, and that *citreola* comprises two. These results corroborate results found elsewhere (Zink et al., unpubl. data), and suggest that the species limits of these two species are in need of further research.

Separation of *lugens* from *alba* was based on evidence of limited hybridization in two regions of Russia (as cited in AOU 1998). Although I present no well-supported evidence to reject the notion that each is indeed a distinct species, all tree topologies found *lugens* and *alba* to be paraphyletic with respect to one another. Given the detected paraphyly, I suggest that *lugens* may not be a valid species, and should instead be considered conspecific with *alba*, or part of the *alba* superspecies complex, as it has been in the past (Dement'ev and Gladkov 1954, Sibley and Monroe 1990), until more complete analyses of their relationships can be performed.
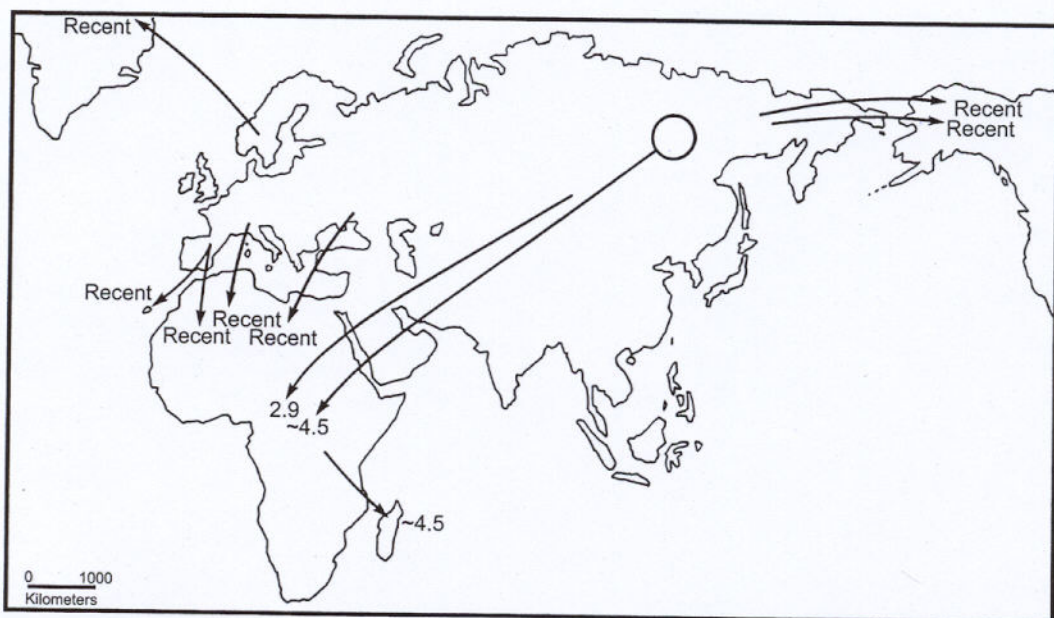
FIGURE 5. Major intercontinental and continent to island colonizations by *Motacilla*. The postulated ancestral area for *Motacilla* is the eastern Palearctic (large circle). The approximate time of each colonization is given either in millions of years (pre-recent dispersal), or is defined as recent, where an extant species has dispersed onto another landmass.

Madaraspatensis is alternatively considered part of the *alba* superspecies (Sibley and Monroe 1990), or as a distinct species (Ali and Ripley 1973). The results of this study imply that *madaraspatensis* may warrant species status. Bootstrap support for the node containing *madaraspatensis* and *alba* as sisters is poor, lending support to the idea that *madaraspatensis* may not be part of the *alba* complex, and may be more closely related to another lineage. Furthermore, the age of the *madaraspatensis/alba* divergence can be estimated via percent sequence divergence to be roughly 1.5 million years, implying the two are distinct lineages. Another line of evidence supporting species status is molt pattern (molt is under genetic control, Berthold and Querner 1982): *madaraspatensis* molts just once per year (Ali and Ripley 1973), whereas *alba* molts twice per year (Cramp 1988).

In no instance did all of the species thought to be part of any one of the three described superspecies (yellow, stream, and pied wagtails) cluster as closest relatives, and these superspecies relationships were clearly rejected by parametric bootstrap tests. This obviously renders superspecies designations in this genus invalid.

ANCESTRAL AREAS

DIVA suggested that Eurasian areas, with African areas or Madagascar are part of a *Motacilla* ancestral area. It seems unlikely that Eurasia and Madagascar (which has been isolated from other landmasses for at least 90 million years) could be part of the same ancestral area, especially when considering that an ancestral population must have moved through either South Africa or the Southern Savannah Zone (assuming no subsequent extinction, and barring overwater dispersal from Asia). It is therefore reasonable to exclude Eurasian + Madagascar combinations as ancestral areas, leaving only Eurasian + southern African areas as possibilities.

Bremer's (1992) ancestral area method identified the eastern Palearctic as having a high probability of being the *Motacilla* ancestral area, and this area was also identified by DIVA. Based on the combination of these two methodologies, then, I exclude Africa as being part of the *Motacilla* ancestral area, and conclude that *Motacilla* arose in the eastern Palearctic; this result holds when wintering, rather than breeding, distributions are used (not shown).

The conclusion of an eastern Palearctic origin seems reasonable when considering the relationships of *Motacilla* to other motacillid genera, and their ancestral areas. The monotypic genus *Dendronanthus* is endemic to eastern Asia, which has also been inferred as the ancestral area of the globally distributed genus *Anthus* (Voelker 1999a). These three genera form the basal radiations within the Motacillidae (Voelker and Edwards 1998).

PRE-RECENT INTERCONTINENTAL MOVEMENTS

Motacilla colonized Africa very soon after its origin in the eastern Palearctic roughly 4.5 million years ago; other widespread genera have shown a similar pattern of establishing an intercontinental presence very soon after their inferred time of origin (Voelker 1999a; Voelker, in press). A second colonization from Eurasia occurred around 2.9 million years ago. Although *Motacilla* occupies forested habitats, most species are associated with more open habitats near water; the above dates suggest that the Eurasia to Africa movements may have been during or near periods when open habitats dominated northeastern Africa. From 9 to 5 million years ago, climate and habitat changes generated an increase in open, grassy habitats across both eastern Africa and Pakistan (Vrba 1985, 1993). Following a re-expansion of tropical forests between 5–3 million years ago (isolating the first ancestral movement into Africa), a second major shift to grasslands began in northeast Africa about 3 million years ago (allowing the second movement into Africa). Evidence of this second shift is reflected by the increase of micromammal fossils (Wesselman 1985), pollen changes (Bonnefille 1983), distributions of open-habitat francolins (*Francolinus*; Crowe et al. 1992), and Eurasia to Africa invasions inferred to have occurred during the same time periods by *Anthus* species, which also occupy open habitats in Africa (Voelker 1999a).

Neither the Mediterranean Sea, which has been flooded for over 5 million years, or the Red Sea, which flooded between 4–3.5 million years ago (Hsü et al. 1977) and was implicated as a possible barrier to ancestral intercontinental movements in *Anthus* (Voelker 1999a), appear to have played a role as a barrier to the second *Motacilla* invasion of Africa. East African or southwest Asian mountains have not played a

role in speciation, as none (including recent volcanic mountains) form a continuous barrier (Visser 1984, Alavi 1994, Burgess et al. 1996, Young 1996).

Very shortly after colonizing Africa 4.5 million years ago, *Motacilla* also colonized Madagascar. This colonization is clearly the result of dispersal, as Madagascar has been isolated from Africa for roughly 130 million years (Brown and Gibson 1983).

RECENT INTERCONTINENTAL MOVEMENTS

Seven intercontinental and continent-to-island colonizations have occurred recently. These recent colonizations have been achieved by *alba* (Africa, and North America via both Alaska and Greenland), eastern Palearctic *flava* (North America), *cinerea* (Africa and Canary Islands), and western *flava* (Africa). In all instances, I assumed that these recent colonizations were initiated from Eurasia, given that the amount of area occupied by these species outside of Eurasia is comparatively tiny. It is not impossible however, that movement occurred *into* Eurasia from another area, the most likely being northern Africa. For example, if Eurasian species, like North American species, were affected by glacial activity (e.g., Frenzel 1968), then the toehold of the above species in northern Africa may instead reflect the oldest extant lineages within each species, and the area from which the recolonization of Eurasia occurred. This concept could be tested were North African breeding samples available for comparison to those from Eurasia. However, given that regions around northern shores of the Mediterranean, as well as much of central Asia, were not glaciated (Frenzel 1968), it seems unlikely that northern Africa could have served as the sole refugium for these *Motacilla* species.

Further evidence suggests that the toehold of these species in northern Africa is a recent occurrence. Much of northern Africa has been affected by repeated cooling–drying cycles, which have greatly affected climate and vegetation (Kutzbach and Liu 1997). The most recent of these cycles was a return to dry (desert) conditions about 6000 years ago. Because the three species with distributions in north Africa do not breed in, or south of, the Sahara, it is unlikely that they were in a position to take advantage of grassland corridors opened across the desert by

the climatic cooling prior to the latest return to dry conditions.

Because Greenland was covered by glaciers during the Quaternary (Brown and Gibson 1983), the presence of *alba* there must reflect recent dispersal, and thus Greenland can readily be excluded as a source area for Eurasian colonization. North America can likewise be excluded from this scenario, because, although portions of western Alaska were ice free during the last glaciation (Porter et al. 1983), widespread available *Motacilla* habitat in Alaska and Canada has not since been occupied, as one would expect from long occupation on the continent.

All of the preceding scenarios imply fairly recent arrivals (via dispersal) in North America and Africa. These scenarios are not, however, intended to imply that modern distributions reflect the first accomplishment of these colonizing feats. Indeed, given the cyclical nature of glacial activity in North America, and climatic shifts across northern Africa, it is entirely likely that the above species may have undergone repeated cycles of dispersal, establishment, and extinction (or at the very least severe range contraction), before now.

ROLES OF DISPERSAL AND VICARIANCE

Dispersal has obviously played a considerable role in shaping the multicontinent distribution of *Motacilla*. Neither Beringia nor the North Atlantic have proven resistant to dispersal by *Motacilla* species. Neither have these water barriers proven impermeable to the myriad Eurasian avian species recorded each year on the Aleutians, to North American species recorded in Europe, nor to the species from several genera that have permanently colonized Greenland from Eurasia (AOU 1998). Dispersal across the Mediterranean with subsequent colonization also appears to be nonproblematic, given that three *Motacilla* species have accomplished it, as have a number of species in other genera (see Cramp 1988). This is not particularly surprising, given that the distance across the Straits of Gibraltar is less than 20 km. Dispersal is clearly implicated in the establishment by *cinerea* of a population in the Canary Islands, which appear never to have had a mainland connection (Schmincke 1976), and the presence of *flaviventris* on Madagascar.

Vicariance, in the form of repeated climatic shifts and associated vegetation changes, most likely drove speciation between African and Eurasian forms. And, while vicariance may also be driving intracontinental speciation, the widespread overlap of species within each continent makes it difficult to define specific events or barriers.

One possible driving force behind *Motacilla* speciation within Eurasia could be glacial activity. Elsewhere (Voelker 1999a; Voelker, in press), I have hypothesized that glacial activity and associated environmental changes may explain speciation in a number of avian sister species whose ranges abut or overlap in the Ural Mountain-Yenisey River region of the north-central Palearctic. At least 44 glacial-to-interglacial shifts in this region are recorded from the past 2.5 million years in the Loess Plateau of China (Kukla and An 1989, An et al. 1991), and, were clock dating fine enough, we could in theory determine whether glaciers did play a role in generating species diversity. Note that several sister species within *Motacilla* show an eastern-versus western-Palearctic or Asian distribution pattern (Fig. 3, 4) that would be generally consistent with this concept. Indeed, three of these lineage splits occur at 800, 540, and 500 thousand years ago, during a glacial period between two of the three most strongly recorded interglacial periods at 1.2 and 0.5 million years ago.

Motacilla shows evidence of a buildup of species communities based on multiple intercontinental invasions by more distantly related congeners, rather than by in situ speciation of all community members within a given region. Two invasions of Africa, followed by vicariant climatic events, resulted in the establishment of the three-species *Motacilla* community in southern Africa. With respect to the relatively recently established species community in northern Africa, multiple invasions via dispersal can be conclusively implicated as having played an exclusive role. Without clock dates and reference to the timing of geological events, it would be less credible to argue that the North African assemblage was the product of multiple dispersal events, rather than the result of a single vicariant event. The pattern of community buildup through repeated invasions, rather than through within-region speciation, echoes the results of other studies (e.g., Richman 1996, Voelker 1999a), and has a negative impact on the notion of "centers of origin", which rely on relative species densities to determine ancestral areas.

There has been a tendency over the past few decades to reject dispersal, in favor of vicariance, as a primary mechanism in the development of concordant distribution patterns (e.g., Ball 1975, Cracraft 1991), despite the fact that both are well-established natural processes (Morrone and Crisci 1995). However, as Hedges et al. (1994) have pointed out, concordant dispersals can explain area congruence just as easily as vicariance, especially when the areas considered are adjacent. Evidence presented here of the multiple dispersals into northern Africa and North America from adjacent areas by *Motacilla* species (and by other avian species) supports this contention. Indeed, when one considers the four basic models of biotic interaction between areas, dispersal can be implicated as the driving force in three (Hunn and Upchurch 2001). These models include dispersal from area Y to area Z, dispersal from area Z to area Y, and dispersal from a third area, X, to both Y and Z. Only the fourth model, where areas Y and Z become isolated from each other after a combined Y + Z area isolation from area X, involves vicariance (Hunn and Upchurch 2001). Because we can show dispersals within lineages to be recent through dating, through habitat occupancy arguments, and through the lack of potential vicariant events, this raises the tantalizing possibility that dispersal may have played a far larger role in the historical development of modern species distributions than we currently allow. The most critical of these approaches is certainly our ability, however imperfect, to date lineage divergences, and to tie these divergences to events in geological history (i.e., a "chronobiogeographical" approach; Hunn and Upchurch 2001). Where over the past few decades dispersal could be ignored as an unscientific means by which to explain distributions (Rosen 1978, Nelson and Platnick 1981), the timing of geological events and the application of molecular clocks may now render the effects of time less efficient at wiping out the signature of dispersal events.

ACKNOWLEDGMENTS

LITERATURE CITED

ALAVI, M. 1994. Tectonics of the Zagros orogenic belt of Iran: new data and interpretations. Tectonophysics 229:211–238.

ALI, S., AND S. D. RIPLEY. 1973. Handbook of the birds of India and Pakistan. Oxford University Press, London.

AMERICAN ORNITHOLOGISTS' UNION. 1998. Check-list of North American birds. 7th ed. American Ornithologists' Union, Washington, DC.

AN, Z., G. J. KUKLA, S. C. PORTER, AND J. XIAO. 1991. Magnetic susceptibility evidence of monsoon variation on the Loess Plateau of central China during the last 130,000 years. Quaternary Research 36:29–36.

BALL, I. R. 1975. Nature and formulation of biogeographic hypotheses. Systematic Zoology 24:407–430.

BERMINGHAM, E., S. ROHWER, S. FREEMAN, AND C. WOOD. 1992. Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: a test of Mengel's model. Proceedings of the National Academy of Sciences 89: 6624–6628.

BERTHOLD, P., AND U. QUERNER. 1982. Genetic basis of moult, wing length, and body weight in a migratory bird species, *Sylvia atricapilla*. Experientia 38:801–802.

BONNEFILLE, R. 1983. Evidence for a cooler and drier climate in the Ethiopian uplands towards 2.5 million years ago. Nature 303:487–491.

BOWEN, B. W., AND W. S. GRANT. 1997. Phylogeography of the sardines (*Sardinops* spp.): assessing biogeographic models and population histories in temperate upwelling zones. Evolution 51:1601–1610.

BREMER, K. 1992. Ancestral areas: a cladistic reinterpretation of the center of origin concept. Systematic Biology 41:436–445.

BROWN, J. H., AND A. C. GIBSON. 1983. Biogeography. C. V. Mosby Co., St. Louis, MO.

BUCKLEY, T. R., C. SIMON, AND G. K. CHAMBERS. 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Systematic Biology 50:67–86.

BURGESS, N., C. FITZGIBBON, AND P. CLARKE. 1996. Coastal forests, p. 329–359. *In* T. R. McClanahan and T. P. Young [EDS.], East African ecosystems and their conservation. Oxford University Press, Oxford, UK.

CRACRAFT, J. 1991. Patterns of diversification within continental biotas: hierarchical congruence among the areas of endemism of Australian vertebrates. Australian Systematic Botany 4:211–227.

CRAMP, S. [ED.]. 1988. Handbook of birds of Europe, the Middle East, and North Africa. Vol. 5. Oxford University Press, Oxford, UK.

CROWE, T. M., AND A. A. CROWE. 1982. Patterns of distribution, diversity and endemism in Afrotropical birds. Journal of Zoology, London 198:417–442.

CROWE, T. M., E. H. HARLEY, M. B. JAKUTOWICZ, J. KOMEN, AND A. A. CROWE. 1992. Phylogenetic, taxonomic and biogeographical implications of genetic, morphological, and behavioral variation in francolins (Phasianidae: *Francolinus*). Auk 109:24–42.

DEMENT'EV, G. P., AND N. A. GLADKOV [EDS.]. 1954. The birds of the Soviet Union. Vol. 5. Translated by E. D. Gordon, 1970. Israel Program for Scientific Translations, Jerusalem.

DOWLING, T. E., C. MORITZ, AND J. D. PALMER. 1990. Nucleic acids II: restriction site analysis, p. 250–317. *In* D. M. Hillis and C. Moritz [EDS.], Molecular systematics. Sinauer Associates, Sunderland, MA.

FLEISCHER, R. C., C. E. MCINTOSH, AND C. L. TARR. 1998. Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Molecular Ecology 7:533–545.

FRENZEL, B. 1968. The Pleistocene vegetation of northern Eurasia. Science 161:637–649.

GOLDMAN, N., J. P. ANDERSON, AND A. G. RODRIGO. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49:652–670.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22:160–174.

HEDGES, S. B., C. A. HASS, AND L. R. MAXSON. 1994. Reply: towards a biogeography of the Caribbean. Cladistics 10:43–55.

HSÜ, K., J. L. MONTADERT, D. BERNOULLI, M. B. CITA, A. ERIKSON, R. E. GARRISON, R. B. KIDD, F. MÉLIERÉS, C. MÜLLER, AND R. WRIGHT. 1977. History of the Mediterranean salinity crisis. Nature 267:399–403.

HUNN, C. A., AND P. UPCHURCH. 2001. The importance of time/space in diagnosing the causality of phylogenetic events: towards a "chronobiogeographical" paradigm? Systematic Biology 50:391–407.

KEITH, S., E. K. URBAN, AND C. H. FRY. 1992. The birds of Africa. Vol. 4. Academic Press, London.

KUKLA, G. C., AND Z. AN. 1989. Loess stratigraphy in central China. Palaeogeography, Palaeoclimatology, Palaeoecology 72:203–225.

KUTZBACH, J. E., AND Z. LIU. 1997. Response of the African monsoon to orbital forcing and ocean feedbacks in the Middle Holocene. Science 278:440–443.

MADDISON, W. P., AND D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.04. Sinauer Associates, Sunderland, MA.

MARTIN, A. P., G. J. P. NAYLOR, AND S. R. PALUMBI. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature 357:153–155.

MORRONE, J. J., AND J. V. CRISCI. 1995. Historical biogeography: introduction to methods. Annual Review of Ecology and Systematics 26:373–401.

NELSON, G., AND N. PLATNICK. 1981. Systematics and biogeography: cladistics and vicariance. Columbia University Press, New York.

POSADA, D., AND K. A. CRANDALL. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818.

PORTER, S. C., K. L. PIERCE, AND T. D. HAMILTON. 1983. Late Wisconsin mountain glaciation in the western United States, p. 74–114. *In* S. C. Porter [ED.], Late Quaternary environments of the United States. Vol. 1, the Late Pleistocene. University of Minnesota Press, Minneapolis, MN.

RAMBAUT, A., AND N. C. GRASSLY. 1997. Seq-Gen: an application for the monte carlo simulation of DNA sequence evolution along phylogenetic trees. Computional Applications in Bioscience 13:235–238.

RICHMAN, A. D. 1996. Ecological diversification and community structure in the Old World leaf warblers (genus *Phylloscopus*): a phylogenetic perspective. Evolution 50:2461–2470.

RONQUIST, F. [ONLINE]. 1996. DIVA version 1.1. ⟨www.ebc.uu.se⟩ (2 August 2002).

RONQUIST, F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Systematic Biology 46:195–203.

ROSEN, D. E. 1978. Vicariant patterns and historical explanation in biogeography. Systematic Zoology 27:159–188.

SCHMINKE, H.-U. 1976. The geology of the Canary Islands, p. 67–184. *In* G. Kunkel [ED.], Biogeography and ecology in the Canary Islands. Junk, The Hague.

SHIELDS, G. F., AND A. C. WILSON. 1987. Calibration of mitochondrial DNA evolution in geese. Journal of Molecular Evolution 24:212–217.

SIBLEY, C. G., AND B. L. MONROE JR. 1990. Distribution and taxonomy of birds of the world. Yale University Press, New Haven, CT.

SLIKAS, B., F. H. SHELDON, AND F. B. GILL. 1996. Phylogeny of titmice (Paridae): I. Estimate of relationships among subgenera based on DNA-DNA hybridization. Journal of Avian Biology 27:70–82.

SULLIVAN, J., E. ARELLANO, AND D. S. ROGERS. 2000. Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic events. American Naturalist 155:755–768.

SWOFFORD, D. L. 1999. PAUP*: phylogenetic analysis using parsimony (* and other methods), version 4. Sinauer Associates, Sunderland, MA.

TAKEZAKI, N., A. RZHETSKY, AND M. NEI. 1995. Phylogenetic test of the molecular clock and linearized trees. Molecular Biology and Evolution 12:823–833.

TARR, C. L., AND R. C. FLEISCHER. 1993. Mitochondrial DNA variation and evolutionary relationships in the Amakihi complex. Auk 110:825–831.

VISSER, J. N. J. 1984. A review of the Stormberg Group and Drakensberg volcanics in southern Africa. Palaeontology of Africa 25:5–25.

VOELKER, G. 1999a. Dispersal, vicariance and clocks: historical biogeography and speciation in a cosmopolitan passerine genus (Anthus: Motacillidae). Evolution 53:1536–1552.

VOELKER, G. 1999b. Molecular evolutionary relationships in the avian genus Anthus (Pipits: Motacillidae). Molecular Phylogenetics and Evolution 11:84–94.

VOELKER, G. In press. Molecular phylogenetics and the historical biogeography of dippers (Cinclus). Ibis.

VOELKER, G., AND S. V. EDWARDS. 1998. Can weighting improve bushy trees? Models of cytochrome b evolution and the molecular systematics of pipits and wagtails (Aves: Motacillidae). Systematic Biology 47:589–603.

VRBA, E. S. 1985. African Bovidae: evolutionary events since the Miocene. South African Journal of Science 81:263–266.

VRBA, E. S. 1993. Mammal evolution in the African Neogene and a new look at the Great American Interchange, p. 393–434. In P. Goldblatt [ED.], Biological relationships between Africa and South America. Yale University Press, New Haven, CT.

WATERS, J. M., J. A. LOPEZ, AND G. P. WALLIS. 2000. Molecular phylogenetics and biogeography of galaxiid fishes (Osteichthyes: Galaxiidae): dispersal, vicariance, and the position of Lepidogalaxias salamandroides. Systematic Biology 49:777–795.

WESSELMAN, H. B. 1985. Fossil micromammals as indicators of climate change about 2.4 million years ago in the Omo Valley, Ethiopia. South African Journal of Science 81:260–261.

YOUNG, T. P. 1996. High montane forest and alpine ecosystems, p. 401–424. In T. R. McClanahan and T. P. Young [EDS.], East African ecosystems and their conservation. Oxford University Press, Oxford, UK.