Accounting for Phylogenetic Uncertainty in Biogeography: A Bayesian Approach to Dispersal-Vicariance Analysis of the Thrushes (Aves: *Turdus*)

JOHAN A. A. NYLANDER,[1,2] URBAN OLSSON,[3] PER ALSTRÖM,[2,4,6] AND ISABEL SANMARTÍN[5]

[1]*School of Computational Science, Florida State University, Tallahassee, Florida 32306, USA; E-mail: nylander@scs.fsu.edu*
[2]*Department of Systematic Zoology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden*
[3]*Department of Zoology, Göteborg University, Box 463, SE-405 30 Göteborg, Sweden*
[4]*Department of Vertebrate Zoology and Molecular Systematics Laboratory, Swedish Museum of Natural History, P.O. Box 50007, SE-104 05 Stockholm, Sweden*
[5]*Department of Biodiversity and Conservation, Real Jardin Botanico-CSIC, Plaza de Murillo 2; 28014 Madrid, Spain*
[6]*Current Address: Swedish Species Information Centre, Swedish University of Agricultural Sciences, Box 7007, SE-750 07 Uppsala, Sweden*

Abstract.— The phylogeny of the thrushes (Aves: *Turdus*) has been difficult to reconstruct due to short internal branches and lack of node support for certain parts of the tree. Reconstructing the biogeographic history of this group is further complicated by the fact that current implementations of biogeographic methods, such as dispersal-vicariance analysis (DIVA; Ronquist, 1997), require a fully resolved tree. Here, we apply a Bayesian approach to dispersal-vicariance analysis that accounts for phylogenetic uncertainty and allows a more accurate analysis of the biogeographic history of lineages. Specifically, ancestral area reconstructions can be presented as marginal distributions, thus displaying the underlying topological uncertainty. Moreover, if there are multiple optimal solutions for a single node on a certain tree, integrating over the posterior distribution of trees often reveals a preference for a narrower set of solutions. We find that despite the uncertainty in tree topology, ancestral area reconstructions indicate that the *Turdus* clade originated in the eastern Palearctic during the Late Miocene. This was followed by an early dispersal to Africa from where a worldwide radiation took place. The uncertainty in tree topology and short branch lengths seems to indicate that this radiation took place within a limited time span during the Late Pliocene. The results support the role of Africa as a probable source area for intercontinental dispersals as suggested for other passerine groups, including basal diversification within the songbird tree. [Bayesian inference; dispersal-vicariance analysis; historical biogeography; *Turdus*.]

Dispersal-vicariance analysis (Ronquist, 1997; as implemented in the program DIVA, Ronquist, 1996) is a parsimony "event-based" biogeographical method that optimizes ancestral areas onto the internal nodes of a phylogeny by minimizing the number of duplication and extinction events required to explain the terminal distributions (Ronquist, 2003). Although allopatric speciation in response to vicariance is assumed as the null model, DIVA also considers dispersal and extinction as alternative processes in the shaping of current distribution patterns (Ronquist, 1997). In recent years, DIVA has become popular as a method of biogeographic inference because of its ability to reconstruct the geographic history of an individual lineage without relying on any previous knowledge of the history of the areas. Unlike cladistic biogeography, area relationships are not forced to conform to an "area cladogram" but free to unite and split as ancestral areas are optimized onto the phylogeny.

One problem with the current implementation of DIVA is that it can only handle fully bifurcated trees. In cases of hard polytomies or when internal branches are short, this is particularly troublesome, as the tree topology is likely to be unreliable, unstable, and highly sensitive to taxon sampling. In fact, a common problem to all current biogeographic methods is that they ignore the uncertainty in phylogenetic inference: ancestral areas are reconstructed onto a fixed tree topology (usually the most parsimonious tree) assuming that the phylogeny is known without error (e.g., Vigilant et al., 1991; Sanmartín and Ronquist, 2004; Brooks and Ferrao, 2005; Goodson et al., 2006). Because this is seldom the case, and different trees can give different biogeographic histories, bio-geographic conclusions from these analyses are at the most tentative. A second source of uncertainty in DIVA is that associated with ancestral area optimization: multiple equally parsimonious optimal distributions are often suggested at ancestral nodes, particularly if there are many widespread extant distributions (Ronquist, 1997). Furthermore, this tendency increases towards the root, where optimal distributions also tend to become more ambiguous (Ronquist, 1997).

The thrushes (Aves, Passeriformes, Muscicapoidea, Turdinae) is a large and diverse group of medium-sized passerine birds of several genera (Clement and Hathway, 2000; Collar, 2005). Cibois and Cracraft (2004) and Voelker and Spellman (2004) assessed the relationships of these and other genera in the superfamily Muscicapoidea. Klicka et al. (2005) estimated the phylogeny of the "true thrushes" (subfamily Turdinae), and Voelker et al. (2007) did the same for the genus *Turdus*. *Turdus* is by far the largest thrush genus, comprising approximately 65 species, with considerable disagreement between authors regarding the taxonomic status of certain taxa (Clement and Hathway, 2000; Collar, 2005; Dickinson, 2003; Ripley, 1964; Sibley and Monroe, 1990). This is one of the few genera of passerines that has radiated extensively both in Africa, South America, and the Palearctic and Oriental regions (Clement and Hathway, 2000; Collar, 2005). The highest number of species occurs in South and Central America. Eurasia has only two thirds as many species, but many of these are polytypic, and the total number of taxa is of the same magnitude in Eurasia as in South America (Ripley, 1964; Clement and Hathway, 2000; Collar, 2005). Africa is home to only 15% of the taxa (Clement and Hathway, 2000).

In the most extensive phylogenetic treatment of *Turdus* to date, Voelker et al. (2007) analyzed mitochondrial DNA data sampled from 60 of the 65 extant species of *Turdus*, as well as the four species from the genera *Cichlherminia*, *Nesocichla*, and *Platycichla*, all of which have been shown to fall inside *Turdus* (Klicka et al., 2005). They identified four major clades within *Turdus*: a "Largely South American Clade," a "Eurasian Clade," an "African Clade," and a "Central American-Caribbean Clade." However, despite their extensive taxon sampling, only one of their recognized clades received high support (and only by Bayesian inference, not in maximum likelihood bootstrap or parsimony bootstrap analyses), and the relationships among the major clades were to some extent uncertain. The study of Voelker et al. (2007) indicates that the choice of molecular markers, the method of inference, and the included taxa (see also Klicka et al., 2005) have a profound impact on the *Turdus* tree topology. This uncertainty may be a consequence of rapid radiation (e.g., Goodson et al., 2006; Rokas et al., 2005) or simultaneous fragmentation of a wide-ranging ancestral area (e.g., Haffer, 1969), and a well-supported resolved topology may never be possible to reconstruct, which renders ancestral distribution reconstruction problematic.

The *Turdus* clade thus presents all of the challenges to biogeographic analysis outlined above: many internal branches are short, and support is low or lacking both for some of the clades and for their relative phylogenetic position. To overcome the problems with phylogenetic uncertainty, we introduce a Bayesian approach to DIVA analysis, where ancestral reconstructions are averaged over all trees, weighted by the probability that each tree is correct (see also Huelsenbeck et al., 2000; Huelsenbeck and Immenov, 2002; Lutzoni et al., 2001; Pagel et al., 2004; Ronquist, 2004). We base our analyses on nuclear and mitochondrial DNA sequence data (∼3 kb) for ∼95% of the generally recognized species in the genus *Turdus*, and an extensive collection of closely related taxa, as indicated by Voelker and Spellman (2004) and Klicka et al. (2005).

MATERIAL AND METHODS

Phylogenetic Analysis

We used a combined data set of mitochondrial (12S, cytochrome *b*) and nuclear (ornithine decarboxylase (ODC) introns 6 to 7, myoglobin intron 2) DNA sequences (Table 1). A posterior distribution of trees was approximated by Bayesian Markov chain Monte Carlo (MCMC) using MrBayes v. 3.1.2 (Ronquist and Huelsenbeck, 2003). In the MCMC, the data were partitioned and each locus was allowed to have partition-specific model parameters (Ronquist and Huelsenbeck, 2003; Nylander et al., 2004). The choice of nucleotide-substitution models for the data partitions in the MCMC was determined based on the Akaike information criterion (Akaike, 1973) calculated using MrAIC v. 1.4.2 (Nylander, 2004) in conjunction with PhyML v. 2.4.4 (Guindon and Gascuel, 2003). For cytochrome *b*, 12S, and ODC, posterior probabilities were calculated under the general time-reversible model (Lanave et al., 1984; Rodríguez et al., 1990; Tavaré,

1986), assuming rate variation across sites according to a discrete gamma distribution with four rate categories (Γ_4; Yang, 1994) and an estimated proportion of invariable sites (Gu et al., 1995). For myoglobin, the HKY (Hasegawa et al., 1985) with Γ_4 distributed rate variation was the best AIC model. Default settings in MrBayes were used for prior and MCMC proposal distributions. Four Metropolis-coupled Markov chains with incremental heating temperature of 0.2 were run for 20 million generations and sampled every 100th generations. The simulation was repeated three times, starting from random trees. Convergence of the MCMC was graphically checked by monitoring cumulative posterior split probabilities and among-run variability of split frequencies using the online tool AWTY (Nylander et al., 2007). The first 1 million generations, before the chains reached apparent stationarity, were discarded and the remaining samples from the independent runs were pooled to obtain the final approximation of the posterior distribution of trees. To yield a single hypothesis of phylogeny, the posterior distribution was summarized as a majority-rule consensus.

In order to get estimates of lineage divergence times, we used a penalized-likelihood approach (Sanderson, 2002) as implemented in the program r8s v.1.71 (Sanderson, 2004). Commands used were Set penalty=log checkgradient=yes smoothing=1 num_time_guesses=25 num_time_guesses=10; Divtime method=pl algorithm=tn. The smoothing value was established using the cross-validation routines implemented in r8s. The r8s analysis was done on the majority-rule consensus from the tree sample in MrBayes, as well as on a random sample (*n* = 500) from the MCMC output. The same smoothing value (1.0) was used for all trees in the MCMC sample. A number of nodes were constrained in the analysis using either a minimum or a maximum age based on information from the fossil record (see Supplementary Material; available at www.systematicbiology.org). To derive absolute ages from the resulting chronograms, all trees were calibrated using a fossil assigned to the "core *Turdus*" clade (see Fig. 1a and Supplementary Material) at a fixed age of 6.8 Ma (Jánossy, 1991). Using a fixed age for a node is clearly a drawback because it underestimates the uncertainty in age estimation (e.g., Ho, 2007), but it is typically a necessary part of the r8s analysis (Sanderson, 2004). Our approach that uses a sample from the MCMC output does, however, partly compensate for this (potentially artificial) overall precision.

In the present analyses, 113 terminal taxa were included, representing nearly 95% of the species, and several taxa currently treated as subspecies, within the genus *Turdus* (Table 1). Species missing from our analyses but included in previous studies (Klicka et al., 2005; Voelker et al., 2007) were *Turdus feae* and *T. rufopalliatus*. We also included all the species allocated to the genera *Cichlherminia*, *Nesocichla*, *Platycichla*, and *Psophocichla*, shown in previous studies to belong to the same clade as *Turdus* (Klicka et al., 2005; Voelker et al., 2007), as well as representatives of the Turdinae clades found by

TABLE 1.   List of taxa and DNA sequence data used, together with locality data and voucher information. A dash (—) indicates missing sequence data. Aligned DNA data are found on TreeBASE (www.treebase.org; accession nos. S1978, M3685).

| Taxon | Locality | GenBank accession no. | | | | Museum no.[a] |
		Cytb	12S	Myo	ODC	
Catharus						
C. dryas	Ecuador	EU154579	EU154479	EU154691	EU154805	ZMUC–120579
C. gracilirostris	Costa Rica	AY049497	EU154480	EU154692	EU154806	AMNH–GFB998
C. occidentalis	Mexico	AY049506	EU154481	EU154693	EU154807	AMNH–GFB1325
C. ustulatus	Ecuador	EU154582	EU154482	EU154694	EU154808	ZMUC–121362
Chlamydochaera						
C. jeffreyi	Malaysia: Sabah	EU154583	EU154483	EU154695	EU154809	LSUMZ–B36487
Cichlherminia						
C. lherminieri dominicensis	Dominica	AY752380	EU154484	EU154696	EU154810	STRI–DOCLH1
C. lherminieri lawrencii	Montserrat	EU154585	EU154485	EU154697	EU154811	STRI–MOCLH1
Cochoa						
C. viridis	Vietnam	EU154587	EU154486	EU154698	EU154812	AMNH–RTC580
Hylocichla						
H. mustelina	USA	AY049504	EU154487	EU154699	EU154813	AMNH–PAC691
Ixoreus						
I. naevia	USA	EU154589	EU154488	EU154700	EU154814	AMNH–GFB3236
Myadestes						
M. ralloides	Bolivia	AF295087	EU154489	EU154701	EU154815	AMNH–CJV141
Nesocichla						
N. eremita	Tristan da Cunha Islands	AY752384	—	EU154702	EU154816	PF–464706 BMNH–1953.55.144
Oenanthe						
O. moesta	Morocco	EU154591	EU154490	EU154703	EU154817	MIUT–2003-103(26)
O. oenanthe	Sweden	EU154592	EU154491	EU154704	EU154818	NRM–20046483
O. pleschanka	Iran	EU154593	EU154492	EU154705	EU154819	MIUT–2003-26(30)
Platycichla						
P. flavipes polionota	Venezuela	EU154594	EU154493	EU154706	EU154820	AMNH–CJW15
P. leucops	Venezuela	EU154595	EU154494	EU154707	EU154821	AMNH–PRS 845
Psophocichla						
P. litsipsirupa litsitsirupa	South African Republic	EU154596	EU154495	EU154708	EU154822	NMBV–6507
P. litsipsirupa simensis	Eritrea	EU154597	—	EU154709	EU154823	BMNH–1953.67.5
Ridgwayia						
R. pinicola	Mexico	AY752371	—	—	EU154824	BMNH–25591
Sialia						
S. sialis	USA	AY049488	EU154496	EU154710	EU154825	AMNH–PAC757
Sturnus						
S. vulgaris	Sweden	AF378103	—	AY228322	EU154804	NRM–966615
Turdus						
T. albicollis paraguayensis	Paraguay	EU154600	EU154498	EU154712	EU154827	NRM– 976739
T. albocinctus	Nepal	EU154601	EU154499	EU154713	EU154828	unvouchered
T. amaurochalinus	Paraguay	EU154602	EU154500	EU154714	EU154829	
T. assimilis dague	Ecuador	EU154603	EU154501	EU154715	EU154830	ZMUC–119943
T. aurantius	Jamaica	EU154604	EU154502	EU154716	EU154831	STRI–JA TAU1
T. bewsheri	Comoro Island	EU154605	EU154503	EU154717	EU154832	MNHN –9-49
T. boulboul	Nepal	EU154606	EU154504	—	EU154833	AMNH–JGG1175
T. cardis	Japan	EU154607	EU154505	EU154718	EU154834	NRM–20076471
T. celaenops	Japan	EU154608	EU154506	EU154719	EU154835	NRM–20076472
T. chiguanco chiguanco	Peru	AY752394	EU154507	EU154720	EU154836	ZMUC–118373
T. chrysolaus	Philippines	EU154610	EU154508	EU154721	EU154837	ZMUC–118 376
T. dissimilis	Captive, Denmark	EU154611				NRM–20076473
	Vietnam		EU154509	EU154722	EU154838	NRM– 20026681
T. falcklandii magellanicus	Argentina	EU154613	EU154511	EU154724	EU154840	AMNH–PRS 1846
T. fulviventris	Ecuador	EU154615	EU154512	EU154726	EU154842	ZMUC–118388
T. fumigatus	Grenada	EU154616	EU154513	EU154727	EU154843	STRI–GR TFU1
T. fuscater fuscater	Bolivia	AY752387	EU154514	EU154728	EU154844	ZMUC–123743
T. grayi	Costa Rica	EU154618	EU154515	EU154729	EU154845	AMNH–GFB1036
T. haplochrous	Bolivia	EU154619	EU154516	EU154730	EU154846	LSUMZ–B7620
T. hauxwelli	Bolivia	EU154620	EU154517	EU154731	EU154847	LSUMZ–B18251
T. hortulorum	China	EU154622	EU154519	EU154733	EU154849	NRM–20076474
T. ignobilis debilis	Ecuador	EU154623	EU154520	EU154734	EU154850	ZMUC–123780
T. iliacus iliacus	Norway	EU154624	EU154521	EU154735	EU154851	NRM–20076475
T. infuscatus	Guatemala	DQ910953	—	EU154736	—	BMNH–1897.10.1.101
T. jamaicensis	Jamaica	EU154625	EU154522	EU154737	EU154852	STRI–JA TJA1
T. kessleri	China	EU154626	—	EU154738	EU154853	NRM–569382
T. lawrencii	Ecuador	EU154627	EU154523	EU154739	EU154854	ZMUC–121372
T. leucomelas leucomelas	Paraguay	EU154628	EU154524	EU154740	EU154855	NRM– 967095
T. libonyanus peripheris	South African Republic	EU154629	EU154525	EU154741	EU154856	UWBM–52923
T. maculirostris	Ecuador	EU154631	EU154526	EU154743	EU154858	ZMUC–118399
T. maranonicus	Peru	EU154634	EU154528	EU154745	EU154860	LSUMZ–B32836
T. merula mandarinus	Vietnam			EU154744	EU154859	NRM– 20046938
	China	EU154632	EU154527			NRM 20076494

TABLE 1. List of taxa and DNA sequence data used, together with locality data and voucher information. A dash (—) indicates missing sequence data. Aligned DNA data are found on TreeBASE (www.treebase.org; accession nos. S1978, M3685). *(Continued)*

Taxon	Locality	GenBank accession no.				
		Cytb	12S	Myo	ODC	Museum no.[a]
T. merula maximus	China	EU154635	EU154529	EU154746	EU154861	NHMO–17030
T. merula intermedius	Xinjiang, China	EU154638	EU154531	EU154749	EU154864	NRM–20076476
T. merula merula	Sweden	EU154637	EU154530	EU154748	EU154863	NRM–20076477
T. merula simillimus	India	EU154674	EU154564	EU154787	EU154902	NRM–20076478
T. migratorius migratorius	USA	EU154639	EU154532	EU154750	EU154865	AMNH–PRS1971
T. mupinensis	China	EU154640	EU154533	EU154751	EU154866	NRM–20076479
T. naumanni eunomus	Hongkong	EU154612	EU154510	EU154723	EU154839	NRM–20000927
T. naumanni naumanni	Hongkong	EU154641	EU154534	EU154752	EU154867	NRM–20000926
T. nigrescens	Costa Rica	EU154642	EU154535	EU154753	EU154868	AMNH–GFB990
T. nigriceps	Ecuador	EU154643	EU154536	EU154754	EU154869	ZMUC–118412
T. nudigenis	Grenada	EU154645	EU154538	EU154756	EU154871	STRI–GR TNU1
T. obscurus	China	EU154646	EU154539	EU154757	EU154872	NRM–20076480
T. obsoletus parambanus	Ecuador	EU154647	EU154540	EU154758	EU154873	ZMUC–118413
T. olivaceuofuscus	Sao Tomé and Principe	EU154648	EU154541	EU154759	EU154874	GUDZ–2003356
T. olivaceus abyssinicus	Kenya	EU154599	EU154497	EU154711	EU154826	ZMUC–124138
T. olivaceus helleri	Kenya	EU154621	EU154518	EU154732	EU154848	UG–TT20
T. olivaceus ludoviciae	Somalia	EU154630	—	EU154742	EU154857	NRM–569379
T. olivaceus nyikae	Tanzania	EU154649	EU154542	EU154760	EU154875	ZMUC–131596
T. olivaceus olivaceus	South African Republic	AY251574	—	—	—	
T. olivaceus roehli	Tanzania	EU154668	EU154559	EU154779	EU154894	NRM2007 6484
T. olivaceus smithi	South African Republic	EU154675	EU154565	EU154788	EU154903	NRM 20076481
T. olivater ssp.	Venezuela	EU154650	EU154543	EU154761	EU154876	AMNH–PRS849
T. pallidus	China	EU154651	EU154544	EU154762	EU154877	NRM–19991106
T. pelios saturatus	Nigeria	EU154654	EU154547	EU154765	EU154880	NRM–20076482
T. pelios	Cameroon	EU154653	EU154546	EU154764	EU154879	LSUMZ–B27170
T. pelios centralis or *pelios*	Central African Republic	EU154652	EU154545	EU154763	EU154878	AMNH–ALP094
T. philomelos philomelos	Sweden	DQ008573	EU154548	EU154766	EU154881	NRM–20046801
T. pilaris	Sweden	EU154656	EU154549	EU154767	EU154882	NRM–20076483
T. plebejus	Panama	EU154657	EU154550	EU154768	EU154883	STRI–PA TPJ492
T. plumbeus ardosiaceus	Dominican Republic	EU154658	EU154551	EU154769	EU154884	AMNH–NKK736
T. plumbeus plumbeus	Bahamas	EU154659	EU154552	EU154770	EU154885	AMNH–NKK872
T. poliocephalus hygroscopus	Sulawesi, Indonesia	EU154661	EU154553	EU154772	EU154887	AMNH–RWD24648
T. poliocephalus kulambangrae	Solomon Islands	EU154662	EU154554	EU154773	EU154888	AMNH–PRS2740
T. poliocephalus niveiceps	Taiwan	EU154644	EU154537	EU154755	EU154870	TESRI–1414
T. poliocephalus poliocephalus	Norfolk Island	EU154663	—	EU154774	EU154889	BMNH–1919.7.15.101
T. poliocephalus rennellianus	Solomon Islands	EU154664	EU154555	EU154775	EU154890	AMNH–MKL57
T. poliocephalus thomassoni	Luzon, Philippines	EU154665	EU154556	EU154776	EU154891	ZMUC–118420
T. reevei	Ecuador	EU154667	EU154558	EU154778	EU154893	ZMUC–118422
T. rubrocanus rubrocanus	India	EU154671	—	EU154782	EU154897	BMNH–1949.Whi.1.2492
T. rubrocanus gouldi	China	EU154670	EU154561	EU154781	EU154896	MNHN –9-60
T. ruficollis atrogularis	China	EU154666	EU154557	EU154777	EU154892	NRM–19980930
T. ruficollis ruficollis	Hongkong	EU154669	EU154560	EU154780	EU154895	NRM–20050609
T. rufitorques	Mexico	DQ910991	—	EU154783	EU154898	BMNH–1954.5.59
T. rufiventris	Paraguay	EU154672	EU154562	EU154784	EU154899	NRM– 937276
T. serranus serranus	Bolivia	EU154673	EU154563	EU154786	EU154901	ZMUC–120529
T. swalesi swalesi	Dominican Republic	EU154676	EU154566	EU154789	EU154904	AMNH–NKK1013
T. tephronotus	Kenya	EU154677	—	EU154790	EU154905	ZMUC–135529
T. torqatus torqatus	Sweden	EU154678	EU154567	EU154791	EU154906	NRM–20076485
T. unicolor	Nepal	EU154679	EU154568	EU154792	EU154907	AMNH–JGG1182
T. viscivorus viscivorus	Sweden	EU154680	EU154569	EU154793	EU154908	NRM–20076486
Zoothera						
Z. cameronensis graueri	Zaire	EU154681	EU154570	EU154794	EU154909	ZMUC–122075
Z. citrina innotata	Vietnam	EU154682	EU154571	EU154795	EU154910	,NRM– 20046884
Z. crossleyi	Zaire	EU154683	EU154572	EU154796	EU154911	ZMUC–122069
Z. dauma aurea	China	EU154684	EU154573	EU154797	EU154912	NRM–20076487
Z. dixoni	China	EU154685	EU154574	EU154798	EU154913	NRM–2007 6488
Z. erythronota mendeni	Bangai, Indonesia	EU154686	EU154575	EU154799	EU154914	AMNH–RWD24739
Z. monticola	Vietnam	EU154687	—	EU154800	EU154915	NRM–569384
Z. princei batesi	Central African Republic	EU154688	EU154576	EU154801	EU154916	ZMUC–122087
Z. schistacea	Tanimbar, Indonesia	EU154689	EU154577	EU154802	EU154917	NRM–20076489
Z. sibirica davisoni	Japan	EU154690	EU154578	EU154803	EU154918	NRM–20076490

[a] *Museum abbreviations*. AMNH: American Museum of Natural History, New York; BMNH: The Natural History Museum, Tring, UK; GUDZ: Department of Zoology, Göteborg University, Göteborg, Sweden; MNHN: Muséum National d'Histoire Naturelle, Paris, France; LSUMZ: Louisiana State University Museum of Natural Science, Baton Rouge, LA; MBM: Marjorie Barrick Museum of Natural History, University of Nevada, Las Vegas, NV; MIUT: Museum of Ispahan University of Technology, Isphahan, Iran; NHMO: National Centre for Biosystematics, Natural History Museum, Oslo, Norway; NRM: Swedish Museum of Natural History, Stockholm, Sweden; PF: Percy FitzPatrick Institute of African Ornithology, Cape Town, South Africa; STRI: Smithsonian Tropical Research Institute, Washington, DC; TESRI: Taiwan Endemic Species Research Institute, Chi-chi, Taiwan; UG: University of Ghent, Belgium; ZMUC: Zoologisk Museum, Copenhagen, Denmark.

Voelker and Spellman (2003), Cibois and Cracraft (2004), and Klicka et al. (2005) (Table 1). The species in the latter groups were selected to maximize geographic sampling within each genus. The genera *Monticola* and *Myophonus* were used as outgroups based on the results of Voelker and Spellman (2004) and Klicka et al. (2005). One species of *Sturnus* and three species of *Oenanthe* were used as external outgroups (see Table 1 and Supplementary Material). A fossil of *Oenanthe* has been recorded from the Late Pliocene (2.0 Ma) of France (Clot et al., 1976a, 1976b) and this was used to place a further external constraint in the dated phylogeny (see Supplementary Material).

Biogeographic Analyses

The 10 areas of geographical distribution used in the analysis are defined in Figure 1b. The zoogeographic boundaries were modified from Good (1974), taking the actual breeding distribution patterns of thrushes into account. In a few cases, sections of ranges that extend marginally into a neighboring zoogeographic area have been disregarded; e.g., for *T. mupinensis* and *T. grayi*.

The focus of the biogeographic analysis is the *Turdus* clade, for which we have a good taxon sampling. However, as pointed out by Ronquist (1996), ancestral area optimizations in DIVA become less reliable as we approach the root node. This is manifested in DIVA as a tendency for the root distribution to be large and include all the areas analyzed. In part, this is due to the cost assignment rules that favor vicariant speciation (Ronquist, 1997), but also to the fact that global optimal states at each node are dependent on the optimizations of the nodes below, which in the case of the root node is the rest of the tree of life that is not part of the phylogeny. To improve the reliability at the basal node of the *Turdus* clade, we run the DIVA analyses on the complete data set including numerous representatives of taxa shown to be closely related to *Turdus* (Table 1), so that the basal node of *Turdus* is no longer the root node in our analysis (Ronquist, 1996).

Ancestral areas were reconstructed using the program DIVA v. 1.2 (Ronquist, 2001). A preliminary analysis with no restriction in the number of areas forming part of the ancestral distribution ("unconstrained") suggested that the ancestor of the *Turdus* clade was distributed in all areas presently occupied by the terminal taxa. Because it is highly improbable that the ancestor of *Turdus* had a distributional range spanning all continents now separated by ocean barriers (including Australia), we carried out a second analysis constraining the maximum number of areas in ancestral distributions to the maximum size of extant ranges; i.e., two areas (e.g., Fig. 1a, *T. philomelos*: AB) using the maxareas option in DIVA. This is equivalent to assuming that the ancestors of *Turdus* have the same ability to disperse as their extant descendants and therefore ancestral ranges were similar in size to extant ranges. Under this assumption the question raised is "What would be the areas most likely to form part of the group ancestral distribution?" (Sanmartín, 2003).
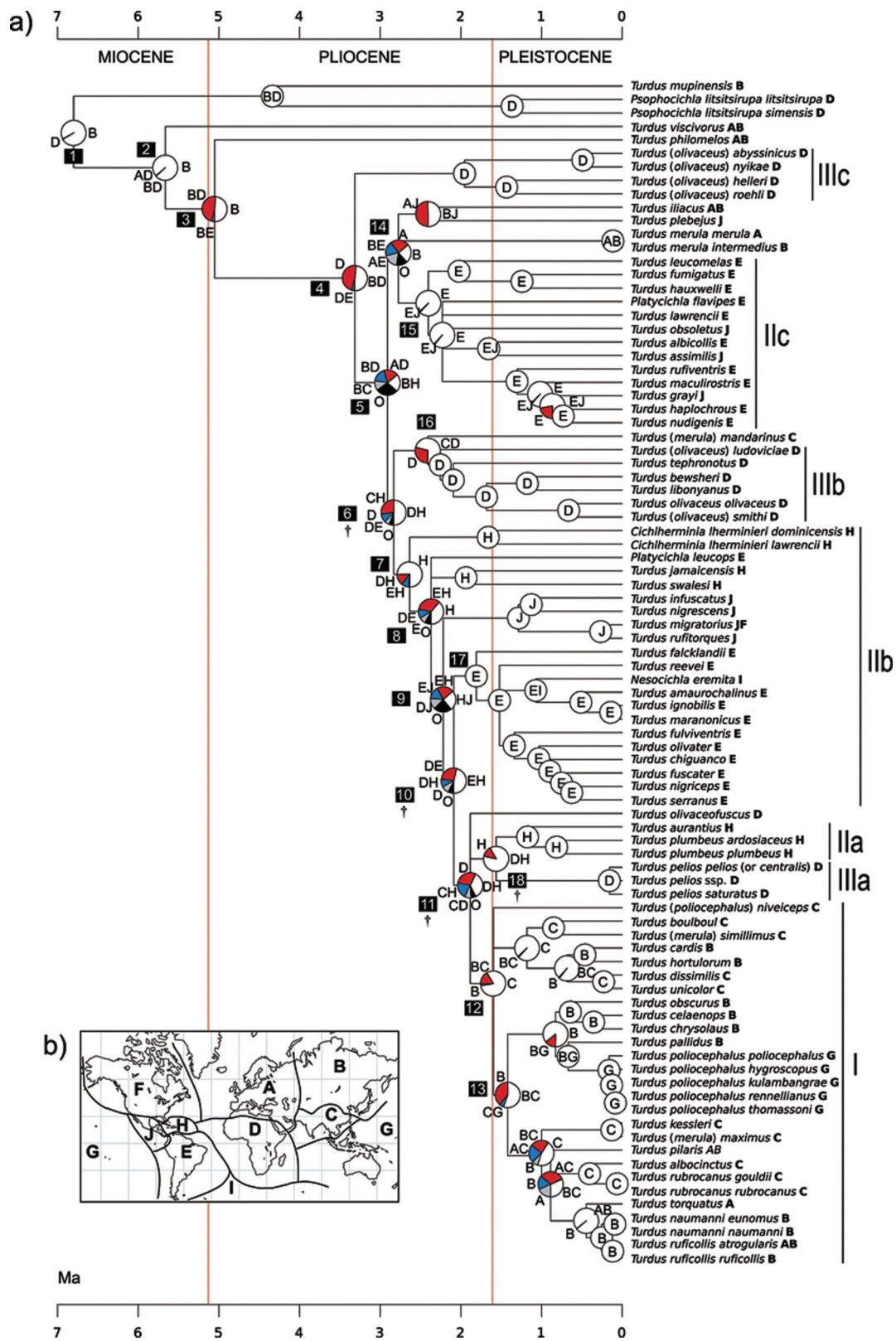
To account for phylogenetic uncertainty in the biogeographic analysis, we randomly sampled 20,000 trees from the MCMC output and ran DIVA analyses on all of them (a "Bayes-DIVA" analysis). The frequency of ancestral areas for clades was then recorded and plotted as marginal distributions on the majority-rule consensus derived from the MCMC. If, in any given tree, multiple ancestral areas were reconstructed for a clade (e.g., "A, B, or AB"), the occurrence of each area was recorded as a fraction (e.g., "A:1/3, B:1/3, AB:1/3"). Therefore, the marginal distributions for alternative ancestral areas at each node in the tree (Fig. 1a) are the product of the phylogenetic uncertainty (clade posterior probability) in the rest of the tree and the uncertainty in the biogeographic reconstruction (multiple equally parsimonious reconstructions) in the node of interest, conditional on that node to occur.

RESULTS

Support for the tree topology was generally high, with >0.95 posterior probability for the majority of clades (Fig. 1a, Supplementary Material). An even more comprehensive phylogenetic analysis, as well as a detailed discussion of the phylogeny in relation to other published work (Klicka et al., 2005; Voelker et al., 2007), is now in preparation. Several geographically restricted subclades or assemblages can be distinguished within the *Turdus* clade: a Eurasian/Australasian clade (I), three American groups (clades IIa, IIc, and a paraphyletic assemblage IIb), and three unrelated African clades (IIIa to IIIc), indicating the importance of local radiation. However, in general, our phylogeny of *Turdus* does not support the monophyly of continental assemblages suggested by the previous study by Voelker et al. (2007). In particular, African and South American *Turdus* seem to be the result of several, independent radiation events. Moreover, the Bayes-DIVA analysis suggests a complex biogeographic history in which intercontinental dispersal has been the primary biogeographic process in the shaping of current distributions in Turdinae. It is possible to deduce at least 16 internal dispersal events from Figure 1a, with the majority of them located at the most basal nodes that form the "backbone" of the tree. The vicariance ratio (average number of vicariant events per area) in the complete phylogeny—including all outgroup taxa (Fig. 1a, Supplementary Material)—is 0.25, whereas dispersal events explain nearly 35% of the speciation events in the tree (ratio dispersal/speciation events).

Bayes-DIVA postulates that the ancestor of *Turdus* originated in the East Palearctic region (B) (>0.99 confidence in ancestral reconstructions at basal nodes 1 to 2) around the Late Miocene; the oldest fossil of *Turdus* is a specimen from Hungary dated between 6.8 and 4.9 Ma (Jánossy, 1991). This was followed by an early dispersal to Africa (+D, the internode leading to node 3 or to node 4) in the Early Pliocene. The ancestral reconstruction for node 4—including all *Turdus* species except *T. philomelos, T. viscivorous*, and *T. mupinensis*—is ambiguous, but it favors either Africa (D) as the sole ancestral area or Africa + Eastern Palearctic. From this point, the ancestral area reconstruction with the highest probability suggests a sequence of trans-Atlantic dispersal events.

One of the descendants stayed in Africa where it gave rise to clade IIIc. The second lineage (node 5) dispersed to the rest of Eurasia (+A, +C) and to the Caribbean region (+H), with either Africa or the Eastern Palearctic as the source areas. However, the ancestral reconstructions at node 5 and its descending node 14 are ambiguous and the sequence of dispersal events is not clear. Divergence times place the start of the first South American radiation (clade IIc) in the Late Pliocene, with some speciation events extending into the Early Pleistocene. At node 6, the most favored reconstruction includes Africa (D) and the Caribbean (H), indicating a second event of trans-Atlantic dispersal during the Late Pliocene, but the ancestral reconstructions are ambiguous and the node is not well-supported (Fig. 1a, Supplementary Material). The African descendant (node 16) dispersed to South Asia (+C) and later underwent local radiation in Africa (clade IIIb). The other descendant (node 7) stayed in the Caribbean, where it gave rise to the *Cichlerminia* clade, and later dispersed to South America (+E, node 8 or node 10) and Central America (+J, node 9), although the sequence of dispersal events is not clear. These dispersal events seem to have taken place during a very short time span within the Late Pliocene (Fig. 1a, Supplementary Material). This is followed by radiation in Central America (the clade *T. infuscatus*–*T. rufitorques*) and in South America (node 17), with later dispersal to the Antarctic region (+I, *Nesocichla eremita*) and range expansion to North America (+F) in *T. migratorius*. The most favored reconstructions at nodes 11 and 18 indicate two consecutive trans-Atlantic dispersal events from the Caribbean to Africa (DH), giving rise to two African lineages: *T. olivaceousfuscus* and the *T. pelios* clade (clade IIIa). An alternative, slightly less likely, reconstruction is dispersal from South America to Africa (DE, node 10) followed by local speciation in Africa (D, node 11) and a second dispersal event from Africa to the Caribbean (DH, node 18). However, phylogenetic support for nodes 10, 11, and 18 is low (posterior probability 0.50 to 0.66; see Supplementary Material). Thus, it is possible that *T. olivaceousfuscus* and the *T. pelios* clade form a monophyletic lineage, in which case there was only one trans-Atlantic dispersal event from the Caribbean region to Africa. A major radiation starting at node 12 resulted in diversification in South Asia (C) and dispersals to the East Palearctic (B) and to the West Palearctic (A) on at least four separate occasions (node 12 and subsequent nodes). There was also a dispersal event to the Malaysia and Australasia

region (G), resulting in local diversification. Divergence time estimates place these speciation events within the Pleistocene, ranging from the Early Pleistocene (1.6 Ma) to the Pleistocene-Holocene boundary (0.06 Ma).

DISCUSSION

Biogeographic History of the Genus Turdus

Although ancestral area reconstructions are ambiguous at some basal nodes (Fig. 1a), several biogeographic conclusions can be reached from the analysis. The *Turdus* clade seems to have originated in the Eastern Palearctic (B; >0.99 confidence) around the Late Miocene (7 to 6 Ma). Dispersal to Africa took place very early in the *Turdus* history (nodes 3, 4). The ancestor of the main *Turdus* radiation (i.e., the clade comprising all of the descendants of the ancestor in node 4) was probably originally present in Africa or widespread in Africa and the Eastern Palearctic. A similar pattern of eastern Palearctic origin and early dispersal to Africa has been suggested to explain basal diversification within other passerine bird genera (Voelker, 1999, 2002). These Asia to Africa movements have been attributed to the cooling and drying trend at the end of the Miocene (9 to 5 Ma) that led to the development of open, grassy habitats in southwestern Asia and eastern Africa (Vrba, 1993), similar to the habitats occupied by these genera in eastern Asia (Voelker, 1999, 2000a). However, most *Turdus* species are forest dwellers, occupying mainly temperate deciduous forests, but also extending their range to high-altitude shrub land, taiga forest, or tropical/subtropical moist forests. Adaptation to the open-land habitat (e.g., *T. tephronotus*, *T.* (*olivaceus*) *smithi* appears to be a derived feature. Also, grasslands did not develop into a major component of the African landscape until the Late Pliocene, although they were already present in Africa during the Miocene (Hernández-Fernández and Vrba, 2006). Thus, although *Psophocichla* may have colonized Africa as a result of expansion of open-land habitats, it appears more likely that the geographic expansion of *Turdus* was related to the contraction of the evergreen tropical forests and subsequent expansion of deciduous-type (open) woodland forests in Africa, following the drying and cooling trend at the end of the Miocene (Cox and Moore, 2005; Hernández-Fernández and Vrba, 2006).

In addition, the Late Miocene collision of Arabia with Eurasia and the final closing of the Red Sea to the Mediterranean during the Early Pliocene (5 Ma) helped

FIGURE 1. (a) A summary of the Bayesian dispersal-vicariance analysis for the thrush genus *Turdus* and closest relatives. The tree is a chronogram based on a 50% majority-rule consensus tree of a Bayesian Markov chain Monte Carlo (MCMC) analysis of a combined data set of mitochondrial (12S, cytochrome *b*) and nuclear (ODC, myoglobin) DNA sequences. Breeding region for each taxon, as delimited in (b), is given after each taxon name. Pie charts at internal nodes represent the marginal probabilities for each alternative ancestral area derived by using dispersal-vicariance analysis (DIVA) while integrating over tree topologies using MCMC. These probabilities (marginal distributions) are a product of the phylogenetic uncertainty (clade posterior probability) in the rest of the tree and the biogeographic uncertainty (multiple equally parsimonious reconstructions) at each node, conditional on this node to occur. In the pie charts, the first four areas with highest probability are colored according to relative probability in the following order: white > red > blue > gray, and any remaining areas are collectively given with black color. A † indicate nodes discussed in the text with posterior probability < 0.70. Numbers in black boxes and the roman numerals at the right edge of the tree are used in the discussion of the tree (see text for details). (b) Biogeographical regions (modified after Good, 1974). A: West Palearctic; B: East Palearctic; C: South Asia; D: Africa; E: South and Central America; F: North America; G: Malaysia and Australasia; H: Caribbean; I: Antarctic region.

create a new dispersal route between southwestern Asia and Africa via the Arabian Plate and the Levante Region (Thompson, 2000; Fernandes et al., 2006). At least for mammals, this was followed by extensive faunal exchange between Eurasia and Africa (Vrba, 1993; Cox and Moore, 2005). Probably, both the development of new terrestrial routes and the expansion of deciduous woodland biomes in Africa contributed to the geographic expansion of *Turdus* early in its evolutionary history. An alternative possibility is that dispersal from Asia to Africa took place via Europe and North Africa. The earliest fossils of *Turdus* are found in Europe (Janossy, 1991), and desert climate conditions in the Saharan belt were not established until the Early Pliocene (5 to 6 Ma), with increasing desertification during the Late Pliocene and Pleistocene glaciations. Nevertheless, the present sub-Saharan distribution of all extant African species of *Turdus* suggests that the southwestern Asia-Arabian connection is more likely than the Mediterranean-Saharan route, in agreement with other African passerine studies (Voelker, 1999).

After dispersal to Africa, a rapid succession of intercontinental movements followed. Although the exact sequence of events related particularly to nodes 9 to 11, and 18 remains obscure, Africa appears to have functioned as a source area for some of these dispersals, both towards the west and the east. Bayes-DIVA reconstructions suggest that African ancestors could have colonized Eurasia (node 5), the Caribbean (node 6 or node 11), and Southeast Asia (node 16). It also indicates that these dispersals took place within a relatively restricted time span during the Late Pliocene. In some cases, as in the New World, colonization took place at different times and as independent events, so that the present continental assemblages are not monophyletic. Thus, contrary to Voelker et al. (2007), our analysis suggests two independent colonization events of South America. The first event (Late Pliocene) apparently involved a Eurasian/Holarctic route. Two Eurasian (AB) species, *Turdus iliacus* and *T. m. merula/T. m. intermedius*, are placed in the same clade (node 14) as the South American radiation, and reconstructions at nodes 5 and 14, although ambiguous, indicate the possibility of dispersal from Africa to Eurasia and from Eurasia to South America, respectively. Furthermore, although no extant species of *Turdus* are endemic to North America, there is a Late Pliocene (3.0 to 1.8 Ma) fossil from California (Chandler, 1990), indicating the presence of the genus in this continent early in *Turdus* history. This first South American radiation (node 15, Late Pliocene) seems to have comprised a high proportion of low- to mid-elevation species inhabiting subtropical or tropical moist forests, although some species extend their ranges to montane habitats, at least locally (e.g., *T. albicollis*, *T. rufiventris*) The second radiation (node 17) includes a higher proportion of montane species (e.g., *T. fulviventris*, *T. serranus*), or even high-altitude species (e.g., *T. chiguanco*, *T. fuscater*). Colonization of the Andean highlands by lowland species appears to be the common trend in South American birds (e.g., Bates and Zink, 1994; Fjeldså,

1992; Brumfield and Edwards, 2007), although the opposite pattern has also been described (Voelker, 1999). This has usually been attributed to the final uplift of the Northern Andes 3 to 2.5 Ma ago (Gregory-Wodzicky, 2000), which created new ecological habitats and favored speciation by allopatry. However, this event is too early for the second South American radiation, which is dated at the Plio-Pleistocene boundary, 2.26 to 1.47 Ma ago. Instead, our Bayes-DIVA analysis suggests that the second colonization of South America was from the Caribbean and/or Central America. Extant species in these regions are mainly montane species (the *Cichlherminia* clade) or highland species (e.g., *T. rufitorques*, *T. nigrescens*). Thus, it is possible that colonization of the Andean slopes by the second South American clade (node 17) was favored by the preadaptation to montane, high-altitude habitats by ancestors from the Central American/Caribbean highlands where extant species are found today. The ensuing radiation was probably related to the dry-humid cycles during Pleistocene glaciations that repeatedly disrupted the vegetational belt division—the sub-Andean, montane forest was probably the most affected (Hooghiemstra and Van der Hammen, 2004)—and which allowed periodical connections between highland and lowland habitats via the northern end of the Andes (Brumfield and Edwards, 2007). Detailed biogeographic studies with a finer subdivision of the South American region are necessary to further test these hypotheses.

As is the case in some other genera of passerine birds (Voelker, 1999, 2002), dispersal has played a major role in determining the modern intercontinental distribution of the genus *Turdus*. Unlike these other genera, however, where many intercontinental movements are restricted to terminal events (Voelker, 1999, 2002; Outlaw et al., 2003), intercontinental dispersal started early in *Turdus* and has been constant throughout its history. At least 16 intercontinental movements (i.e., successful colonizations) are inferred to explain the extant ranges, most of them at basal nodes. This is not surprising since thrushes are strong flyers. For example, *Turdus migratorius* regularly strays across the Atlantic from North America to Europe (Clement and Hathway, 2000). In the reverse direction, both *T. pilaris* and *T. iliacus* have attempted to colonize Greenland from Europe in recent times and have been recorded in North America (Clement and Hathway, 2000). Although our analysis also indicates considerable duplication (i.e., speciation within the area) as explanation for geographic species assemblages within the genus (e.g., the South American clades IIa to IIc), many of these radiations are more likely the result of allopatric speciation within the biogeographic boundaries defined here. At a finer geographical scale it is likely that even more dispersal events would have been suggested.

Within-continent analyses would also probably reveal a finer pattern of vicariant speciation. There are a number of examples where members of recently radiated clades occupy parapatric ranges that may be the result of fragmentation of a previously continuous distribution (Fig. 1a). For example, *T. celaenops*, *T. chrysolaus*,

T. obscurus, and, according to Voelker et al. (2007), *T. feae,* form a clade with allopatrically distributed species in eastern Siberia, northeastern China, and Japan (Clement and Hathway, 2000; Collar, 2005). A similar case may explain the origin of the clade *T. cardis–T. unicolor,* with an allopatric distribution in the Himalayas, eastern Russia, China, and Japan. Thus, for the *Turdus* clade, as for some other passerine birds (Voelker, 1999, 2002), the eastern Palearctic appears to have acted as a "cradle for diversification." The higher overall diversity of this region compared to other Holarctic infraregions (Latham and Ricklefs, 1993) has traditionally been explained by a lower extinction rate during the Pleistocene glaciations (the "refugium hypothesis") but recent studies (Qian and Ricklefs, 1999; Sanmartín et al., 2001) suggest that a higher speciation rate is a more likely explanation, probably related to extensive orogenic activity during the Tertiary. In the case of *Turdus,* divergence time estimates suggest that radiation events within these Eurasian clades were more probably related to the fragmentation of ancestral ranges (i.e., forest contraction) during the Late Pliocene-Pleistocene climatic changes (cf. Voelker, 1999) (e.g., the *T. torquatus–T. ruficollis* clade). Repeated climatic shifts, involving cycles of cold-dry and humid-warm weather, have occurred in southwestern Asia and the Himalayas region over at least the past 4 Myr (cf. Voelker, 1999), which could explain the recent radiation in these Eurasian (Himalayan) clades. An alternative explanation is tectonic/vicariance events driven by the continuous orogenic activity that has taken place in the Himalayan region since the Miocene to the present: the Higher Himalaya is still uplifting at more than 4.4 mm/year (Sakai et al., 2006).

The biogeographic history of *Turdus* reconstructed here differs from that of some other passerine birds (Voelker, 1999, 2002) in placing the source area for many intercontinental dispersal events in Africa instead of Asia. The Bayes-DIVA basal node reconstructions suggest that the African ancestors of *Turdus* could have colonized independently Eurasia, the Caribbean, and Southeast Asia. The main exception to this "out-of-Africa" pattern is the dispersal event from the Caribbean/South America (H/E) to Africa (D) that led to *T. olivaceousfuscus* and the *T. pelios*-clade (nodes 10, 11). The striking morphological similarity between these two African species and the other African *Turdus* (clades IIIb, IIIc) would thus be the result of parallel evolution or convergence. However, the short branches, ambiguous reconstructions, and low support at nodes 10 and 11 make it difficult to exclude alternative explanations. Given that Africa (D) appears as part of some ancestral distributions in all nodes from 6 to 11, another possibility is that the entire "backbone" of the tree is African and that the extant African *Turdus* clades IIIa to IIIc are all descendants of a common ancestor, having evolved within Africa independently in successive radiations. The non-African sister clades would then be the result of several independent waves of dispersal out of Africa. This explanation is not the most likely in our Bayes-DIVA reconstructions.

However, in a recent analysis, Beresford et al. (2005) argued for a similar pattern of radiation in and subsequent dispersal out of Africa to explain the current diversity of Old World warblers. Moreover, the inference of several African lineages at the deepest, basalmost positions in the songbird (Passeriformes) phylogeny suggests that the African continent has played a key role in the early divergence history of songbirds (Beresford et al., 2005; Fuchs et al., 2006). Our biogeographic analysis shows that, irrespective of the origin of the *T. olivaceousfuscus-T. pelios* clade, Africa does appear to have played a role in the early divergence of genus *Turdus* as well.

Advantages of the Bayes-DIVA Approach

DIVA analyses are usually based on a single, fully resolved tree (e.g., Beier et al., 2004; Voelker, 2002; Sanmartín, 2003; Oberprieler, 2005). The ancestral area reconstruction is then conditional on this tree being true and without error. Inferred trees are, however, usually subject to both systematic and stochastic error, leading to uncertainty in results (Huelsenbeck et al., 2000). The uncertainty in tree topology can be accommodated in a DIVA analysis by using a set of weighted trees. This set could be generated using, e.g., the bootstrap, and the results could be summarized in a similar way as was done in Figure 1a (see also Ronquist, 2003). The exact interpretation of the variation generated by the bootstrap is, however, more problematic. Importantly, bootstrap frequencies do not easily lend themselves to a probabilistic interpretation (see discussion in, e.g., Felsenstein and Kishino, 1993; Newton, 1996; Holmes, 2003). It seems preferable, then, to use the collection of trees from a Bayesian MCMC because they are sampled in direct proportion to their (posterior) probability, and any frequencies derived from such a sample can potentially be treated as probabilities (see below).

Our approach is a straightforward application of Bayesian reasoning where the uncertainty is accommodated by utilizing the posterior tree distribution. The advantages of using this approach are numerous. For example, the ancestral area reconstructions can be displayed as marginal distributions, allowing for an intuitive interpretation of the underlying uncertainty. Inferences of ancestral ranges are thus robust to all alternative topological solutions. Furthermore, if there are multiple optimal solutions for a single node on a particular tree, integrating over the posterior distribution of trees often reveals a preference for a single, or a more restricted, set of solutions. Hence, if needed, the Bayesian analysis can help us "choose" between equally parsimonious alternatives.

A similar approach for accommodating topological uncertainty in historical biogeography was adopted by Huelsenbeck and Immenov (2002). They inferred the geographic distribution for the most recent common ancestor to Hominids by integrating over trees drawn from the posterior distribution of a Bayesian MCMC. They did not, however, use DIVA for the area reconstruction but relied on simple parsimony (Fitch) optimization. In Fitch parsimony, ancestors are assumed to be monomorphic

(i.e., occurring in single areas), whereas polymorphism (i.e., widespread distributions) is restricted to terminal taxa. For biogeographic analysis, this implies that only dispersal (change of area) and duplication (speciation within the area) events can be inferred from Fitch optimization; vicariance and extinction events are simply not allowed (Ronquist, 2003; Sanmartín, 2006). In contrast, allopatric speciation of widespread ancestors by vicariance is the key of the DIVA analysis. Because several extant species of *Turdus* are currently distributed in two areas while (apparently) maintaining genetic cohesion (e.g., *T. migratorius*), restricting ancestors to single areas in the analysis does not seem realistic. The DIVA approach adopted here, which allows combination of dispersal and vicariance events, seems therefore more appropriate.

The Bayes-DIVA method presented here is not a strict Bayesian implementation of ancestral area reconstruction. DIVA is a parsimony-based method and has the ability to give multiple optimal solutions if there is ambiguity in the data. The method does not, however, truly account for the uncertainty associated with the reconstruction of ancestral states onto a phylogeny ("mapping uncertainty"; Ronquist, 2004), because only the most parsimonious (minimum change) reconstructions are considered, even though alternative reconstructions could be almost as likely. For our method to incorporate the uncertainty in ancestral area reconstruction within a strict Bayesian framework, we would have needed an explicit model of biogeographic evolution specifying the transition probabilities between alternative ancestral areas and the relative rates of different biogeographic events such as dispersal and extinction (see Ree and Smith, 2008). We would also have needed priors on these rates in order to calculate the posterior probability distribution of all alternative ancestral areas—not only the most parsimonious—at each node in the tree (Pagel et al., 2004). The rate parameters of the model themselves could be estimated from the data (or treated as nuisance parameters) and integrated over all possible tree topologies using MCMC.

The lack of a parametric model for the DIVA reconstruction also somewhat compromises the straightforward interpretation of the frequencies as probabilities. In order for a strictly probabilistic interpretation we need to make the assumption that the parsimony solution is also the maximum likelihood solution, or at least a good approximation to it. The method is perhaps best viewed as a (nonparametric) empirical Bayes method (NPEB; see, e.g., Casella, 1985; Carlin and Louis, 2000), where some parameters are point estimates instead of products of integration over a prior. In this setting, the DIVA reconstruction is treated as a random variable, but with an unspecified prior (hence the nonparametric EB). Again, we need to assume that the parsimony solution is the ML solution, an assumption that is left to be proven since a stochastic model for a DIVA reconstruction is yet to be formulated (but see Ree and Smith, 2008). It is interesting to note, however, that comparisons between DIVA and a recently developed maximum likelihood approach

to dispersal-vicariance analysis (Ree et al., 2005; see below) suggest that for most biogeographic scenarios the most-parsimonious reconstruction(s) in DIVA also correspond to the maximum likelihood scenario (Ree et al., 2005; Moore et al., 2008).

Recently, Ree et al. (2005) proposed a likelihood-based alternative to dispersal-vicariance analysis that integrates over the uncertainty associated with reconstructing ancestral areas onto a phylogeny. Unlike parsimony-based approaches where all branches are considered equal in terms of number of events, maximum likelihood approaches make use of an explicit model of biogeographic evolution and branch length information to estimate the probability of change between ancestral areas along a given branch. Because all alternative reconstructions are considered in estimating the relative probability of ancestral states, the analysis incorporates the uncertainty in ancestral state reconstruction. However, ancestral areas are still reconstructed over a single topology, i.e., the maximum likelihood tree (Ree et al., 2005), so this approach does not incorporate the uncertainty in phylogenetic inference. We do, however, take full advantage of the power of the Bayesian approach by integrating the ancestral area reconstructions over the posterior distribution of all trees from a Bayesian phylogenetic analysis. Thus, the Bayes-DIVA method described here can be said to provide the most parsimonious reconstruction of ancestral areas onto a phylogeny while at the same time accounting for the uncertainty in phylogenetic reconstruction.

A known drawback of the DIVA approach is the unrealistic treatment of extinction and this is also reflected in the Bayes-DIVA analysis and our reconstruction of the *Turdus* history. Extinction events will never appear in dispersal-vicariance optimizations unless a geological model that places explicit geographic constraints in the original cost matrix is used (Ronquist, 1996; Ree et al., 2005; Sanmartín, 2006). For example, movement from one single area to another single area could be constrained to go through an intermediate widespread state followed by extinction (e.g., Ree and Smith, 2008). Instead, extinction events in DIVA are usually inferred ad hoc after the analysis in order to explain widespread ancestral distributions among areas that are not geographically adjacent (Sanmartín, 2003). A common error among DIVA studies is to interpret the results from a DIVA analysis by inferring extinction in cases where one descendant is widespread and the other has a more restricted distribution (e.g., Voelker, 1999; Yoo et al., 2005). For example, in the clade *T. ruficollis atrogularis (AB)–T. ruficollis ruficollis (B)*, DIVA will explain the unequal range size by secondary (post-speciation) dispersal in the most widespread descendant; i.e., dispersal of *T. r. atrogularis* to A (Fig. 1a). The alternative explanation of extinction of the restricted descendant (*T. r. ruficollis*) in part of the ancestral range will never be inferred by DIVA, because the method assumes that widespread distributions cannot be maintained across speciation (Ronquist, 1997). In practice, this means that duplication (speciation within the area) in widespread ancestors is not allowed—

it only occurs within single areas—whereas widespread ancestral distributions are allopatrically divided at each speciation node. The approach of Ree et al. (2005) does not permit inheritance of widespread distributions either, because it will imply the simultaneous origin of both daughter species in each area in the range, a scenario considered very unrealistic (Ree et al., 2005). However, their method, unlike DIVA, accepts peripatric speciation as an alternative explanation; one of the descendants could inherit the widespread ancestral range, whereas the other "buds off" within one of the areas. This reduces the tendency of DIVA to overestimate terminal dispersal events when there are many widespread terminals. In groups like birds, where widespread extant taxa are a common feature, this is likely to become a major problem. Peripatric speciation could be easily introduced in Bayes-DIVA by modifying the original DIVA optimization rule that forces widespread ancestral distributions to be divided by vicariance at each speciation node into "two mutually exclusive sets of areas" (Ronquist, 1997). This clearly would be something to improve in the future. On the other hand, a possible strength of the Bayes-DIVA approach compared to Ree et al.'s maximum likelihood method is its independence from a geological model. Ancestral areas can be reconstructed without any prior knowledge of the geological history of the areas studied (i.e., the timing of geographical barriers and connection routes) or even of lineage divergence times. We agree, however, that when this information is available, it seems reasonable to make use of it as Ree et al. (2005) do. Finally, we conclude that any method for historical biogeography would benefit from adopting a strategy for accommodating phylogenetic uncertainty as the one presented here.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle. Pages 267–281 *in* Second International Symposium on Information Theory (B. N. Petrov and F. Csaki, eds.). Akademiai Kiado, Budapest.

Bates, J. M., and R. Zink. 1994. Evolution into the Andes: Molecular evidence for species relationships in the genus *Leptopogon*. Auk 111:507–515.

Beier, B.-A., J. A. A. Nylander, M. W. Chase, and M. Thulin. 2004. Phylogenetic relationships and biogeography of the desert plant genus *Fagonia* (Zygophyllaceae), inferred by parsimony and Bayesian model averaging. Mol. Phylogenet. Evol. 33:91–108.

Beresford, P., F. K. Barker, P. G. Ryan, and T. M. Crowe. 2005. African endemics span the tree of songbirds (Passeri): Molecular systematics of several evolutionary "enigmas." Proc. R. Soc. Lond. B. Biol. Sci. 272:849–858.

Brooks, D. R., and A. L. Ferrao. 2005. The historical biogeography of co-evolution: Emerging infectious diseases are evolutionary accidents waiting to happen. J. Biogeogr. 32:1291–1299.

Brumfield, R. T., and S. V. Edwards. 2007. Evolution into and out of the Andes: A Bayesian analysis of historical diversification in *Thamnophilus* Antshrikes. Evolution 61:346–367.

Carlin, B. P., and T. A. Louis. 2000. Bayes and empirical Bayes methods for data analysis, 2nd edition. Chapman & Hall, New York.

Casella, G. 1985. An introduction to empirical Bayes data analysis. Am. Stat. 39:83–87.

Chandler, R. M. 1990. Fossil birds of the San Diego Formation, Late Pliocene, Blancan, San Diego County, California. Ornithol. Monogr. 44:73–161.

Cibois, A., and J. Cracraft. 2004. Assessing the passerine "tapestry": Phylogenetic relationships of the Muscicapoidea inferred from nuclear DNA sequences. Mol. Phylogenet. Evol. 32:264–273.

Clement, P., and R. Hathway. 2000. Thrushes. Christopher Helm. A & C Black. London.

Clot, A., J. Chaline, E. Heintz, D. Jammot, C. Mourer-Chauvire, and J.-C. Rage. 1976a. Montoussé 5 (Hautes Pyrénées), un nouveau remplissage de fissure à faune de vertébrés du Pléistocène inférieur. Geobios 9:511–514.

Clot, A., J. Chaline, D. Jammot, C. Mourer-Chauvire, and J.-C. Rage. 1976b. Les poches ossifères du Pléistocène moyen et inférieur de Montoussé (Hautes Pyrénées). Bull. Soc. Hist. Nat. Toulouse. 112:146–161.

Collar, N. 2005. Family Turdidae (thrushes). Pages 514–807 *in* Handbook of birds of the world, volume 10 (J. del Hoyo, A. Elliott, and D. A. Christie, eds.). Lynx Edicions, Barcelona.

Cox, B. C., and P. D. Moore. 2005. Biogeography: An ecological and evolutionary approach, 7th edition. Blackwell Publishing, London, UK.

Dickinson, E. C. (ed.), 2003. The Howard and Moore complete checklist of the birds of the world. Christopher Helm, London.

Felsenstein, J., and H. Kishino. 1993. Is there something wrong with the bootstrap? A reply to Hillis and Bull. Syst. Biol. 42:193–200.

Fernandes, C. A., E. J. Rohling, and M. Siddall. 2006. Absence of post-Miocene Red Sea land bridges: Biogeographic implications. J. Biogeogr. 33:961–966.

Fjeldså, J. 1992. Biogeographic patterns and evolution of the avifauna of relict high-altitude woodlands of the Andes. Steenstrupia 18:9–62.

Fuchs, J., J. Fjeldså, R. C. K. Bowie, G. Voelker, E. Pasquet. 2006. The African warbler genus *Hyliota* as a lost lineage in the Oscine songbird tree: Molecular support for an African origin of the Passerida. Mol. Phylogenet. Evol. 39:186–197.

Good, R. 1974. The geography of flowering plants, 4th edition. Longmans, London.

Goodson, B. E., A. Santos-Guerra, and R. K. Jansen. 2006. Molecular systematics of *Descurainia* (Brassicaceae) in the Canary Islands: Biogeographical and taxonomic implications. Taxon 55:671–682.

Gregory-Wodzicky, K. M. 2000. Uplift history of the Central and Northern Andes: A review. GSA Bull. 112:1091–1105.

Gu, X., Y.-X., Fu, and W.-H. Li. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12:546–557.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Haffer, J. 1969. Speciation in Amazonian forest birds. Science 165:131–137.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21:160–174.

Hernández Fernández, M., and E. S. Vrba. 2006. Plio-Pleistocene climatic change in the Turkana Basin (East Africa): Evidence from large mammal faunas. J. Human Evol. 50:595–626.

Ho, S. Y. W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J. Avian Biol. 38:409–414.

Holmes, S. 2003. Bootstrapping phylogenetic trees: Theory and methods. Stat. Sci. 18:241–255.

Hooghiemstra, H., and T. van der Hammen. 2004. Quaternary Ice-Age dynamics in the Colombian Andes: Developing an understanding of our legacy. Phil. Trans. R. Soc. B. 1442:173–181.

Huelsenbeck, J. P., and N. S. Immenov. 2002. Geographic origin of human mitochondrial DNA: Accommodating phylogenetic uncertainty and model comparison. Syst. Biol. 51:155–165.

Huelsenbeck, J. P., B. Rannala, and J. P. Masly. 2000. Accommodating phylogenetic uncertainty in evolutionary studies. Science 288:2349–2350.

Jánossy, D. 1979. Plio-pleistocene bird remains from the Carpathian Basin. IV. Anseriformes, Gruiformes, Charadriiformes, Passeriformes. Aquila 85:11–39.

Jánossy, D. 1991. Late Miocene bird remains from Polgárdi (W-Hungary). Aquila 98:13–35.

Klicka, J., G. Voelker, and G. M. Spellman. 2005. A molecular phylogenetic analysis of the "true thrushes" (Aves: Turdinae). Mol. Phylogenet. Evol. 34:486–500.

Lanave, C., C. Preparata, C. Saccone, and G. Serio. 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20:86–93.

Latham, R. E., and R. E. Ricklefs. 1993. Continental comparisons of temperate-zone tree species diversity. Pages 294–314 in Species diversity in ecological communities (R. E. Ricklefs, D. Schluter, eds.). The University of Chicago Press, Chicago and London.

Lutzoni, F., M. Pagel, and V. Reeb. 2001. Major fungal lineages derived from lichen-symbiotic ancestors. Nature 411:937–940.

Moore, B. R., S. A. Smith, R. H. Ree, and M. J. Donoghue. 2008. Incorporating fossil data in biogeographic inference: A likelihood approach. Evolution In press.

Newton, M. 1996. Bootstrapping phylogenies: Large deviations and dispersion effects. Biometrika 83:315–328.

Nylander, J. A. A. 2004. MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. Available at http://www.ebc.uu.se/systzoo/staff/nylander.html.

Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53:47–67.

Nylander, J. A. A., J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford. 2007. AWTY (Are We There Yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics, Advance Online Access, August 30. Available at http://ceb.scs.fsu.edu/awty.

Oberprieler, C. 2005. Temporal and spatial diversification of the circum-Mediterranean Compositae-Anthemidae. Taxon 54:951–966.

Outlaw, D. C., G. Voelker, B. Mila, and D. J. Girman. 2003. Evolution of long-distance migration in and historical biogeography of Catharus thrushes: A molecular phylogenetic approach. Auk 120:299–310.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53:673–684.

Qian, H., and R. E. Ricklefs. 1999. A comparison of the taxonomic richness of vascular plants in China and the United States. Am. Nat. 154:160–181.

Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution 59:2299–2311.

Ree, R. H., and S. A. Smith. 2008. Maximum-likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57:4–14.

Ripley, S. D. 1964. Family Muscicapidae, subfamily Turdinae. Pages 13–227 in Checklist of birds of the world, a continuation of the work of James L. Peters, volume X. (E. Mayr and R. A. Paynter, Jr., eds.). Museum of Comparative Zoology, Cambridge, Massachusetts.

Rodríguez, J., L. Oliver, A. Marín, and R. Medina. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142:485–501.

Rokas, A., D. Kruger, and S. B. Carroll. 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310:1933–1938.

Ronquist, F. 1996. DIVA 1.1. User's Manual. Evolutionary Biology Centre, Uppsala University. Available at http://www.ebc.uu.se/systzoo/research/diva/diva.html.

Ronquist, F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. Syst. Biol. 46:195–203.

Ronquist, F. 2001. DIVA version 1.2. Computer program for MacOS and Win32. Evolutionary Biology Centre, Uppsala University. Available at http://www.ebc.uu.se/systzoo/research/diva/diva.html.

Ronquist, F. 2003. Parsimony analysis of coevolving species associations. Pages 22–64 in Tangled trees: Phylogeny, cospeciation, and co-evolution (R. D. M. Page, ed.). University of Chicago Press, Chicago, Illinois.

Ronquist, F. 2004. Bayesian inference of character evolution. Trends Ecol. Evol. 19:475–481.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Sakai, H., H. Sakai, W. Yahagi, R. Fujii, T. Hayashi, and B. N. Upreti. 2006. Pleistocene rapid uplift of the Himalayan frontal ranges recorded in the Kathmandu and Siwalik basins. Palaeogeogr. Palaeoclimateol. Palaeoecol. 241:16–27.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

Sanderson, M. J. 2004. R8S, program and documentation. Available at http://loco.biosci.arizona.edu/r8s/.

Sanmartín, I. 2003. Dispersal vs. vicariance in the Mediterranean: Historical biogeography of the Palearctic Pachydeminae (Coleoptera, Scarabaeoidea). J. Biogeogr. 30:1883–1897.

Sanmartín, I. 2006. Event-based biogeography: Integrating patterns, processes, and time. Pages 135–159 in Biogeography in a changing world (M. C. Ebach and R. S. Tangney, eds.). The Systematics Association Special Volume Series 70. CRC Press, Taylor & Francis, London.

Sanmartín, I., H. Enghoff, and F. Ronquist. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linn. Soc. 73:345–390.

Sanmartín, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: Plant versus animal patterns. Syst. Biol. 53:216–243.

Sibley, C. G., and B. L. Monroe Jr. 1990. Distribution and taxonomy of birds of the world. Yale University Press, New Haven, Connecticut.

Tavaré, S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lec. Math. Life Sci. 17:57–86.

Thompson, A. 2000. Origins of Arabia. Oriental Press, Dubai.

Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes, and A. C. Wilson. 1991. African populations and the evolution of modern mitochondrial DNA. Science 253:1503–1507.

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

Voelker, G. 2002. Systematics and historical biogeography of wagtails (Aves: Motacilla): Dispersal versus vicariance revisited. Condor 104:725–739.

Voelker, G., S. Rohwer, R. C. K. Bowie, and D. C. Outlaw. 2007. Molecular systematics of a speciose, cosmopolitan songbird genus: Defining the limits of, and relationships among, the Turdus thrushes. Mol. Phylogenet. Evol. 42:422–434.

Voelker, G., and G. M. Spellman. 2004. Nuclear and mitochondrial DNA evidence of polyphyly in the avian superfamily Muscicapoidea. Mol. Phylogenet. Evol. 30:386–394.

Vrba, E. S. 1993. Mammal evolution in the African Neogene and a new look at the great American interchange. Pages 393–434 in Biological relationships between Africa and South America (P. Goldblatt, ed.). Yale University Press, New Haven, Connecticut.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39:306–314.

Yoo, M.-J., C. D. Bell, P. S. Soltis, and D. E. Soltis. 2005. Divergence times and historical biogeography of Nymphaeales. Syst. Bot. 30:693–704.