

Research



Cite this article: Batista R, Olsson U, Andermann T, Aleixo A, Ribas CC, Antonelli A. 2020 Phylogenomics and biogeography of the world's thrushes (*Aves*, *Turdus*): new evidence for a more parsimonious evolutionary history. *Proc. R. Soc. B* **287**: 20192400. <http://dx.doi.org/10.1098/rspb.2019.2400>

Received: 14 October 2019

Accepted: 18 December 2019

Subject Category:

Evolution

Subject Areas:

evolution, genomics

Keywords:

Turdus, dispersal, next-generation sequencing, species radiations

Author for correspondence:

Romina Batista

e-mail: rominassbatista@gmail.com

[†]Joint senior authors.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4800375>.

Phylogenomics and biogeography of the world's thrushes (*Aves*, *Turdus*): new evidence for a more parsimonious evolutionary history

Romina Batista^{1,2,3}, Urban Olsson^{3,4}, Tobias Andermann^{3,4}, Alexandre Aleixo⁵, Camila Cherem Ribas^{6,†} and Alexandre Antonelli^{3,4,7,†}

¹Programa de Pós-Graduação em Genética, Conservação e Biologia Evolutiva, PPG GCBEV – Instituto Nacional de Pesquisas da Amazônia (INPA) Campus II, Av. André Araújo, 2936, Petrópolis, CEP 69067-375 Manaus, Amazonas, Brazil

²Coordenação de Zoologia, Laboratório de Biologia Molecular, Museu Paraense Emílio Goeldi, CEP 66077-830 Belém, Pará, Brazil

³Gothenburg Global Biodiversity Centre, Box 461, 405 30 Gothenburg, Sweden

⁴Department of Biological and Environmental Sciences, University of Gothenburg, SE-413 19 Gothenburg, Sweden

⁵Finnish Museum of Natural History, University of Helsinki, P.O. Box 17, 00014 Helsinki, Finland

⁶Instituto Nacional de Pesquisas da Amazônia, (INPA) Campus II, Av. André Araújo, 2936, CEP 69060-000 Manaus, Amazonas, Brazil

⁷Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AE, UK

id RB, 0000-0002-7030-5227; UO, 0000-0003-1435-5411; TA, 0000-0002-0932-1623; AA, 0000-0002-7816-9725; CCR, 0000-0002-9088-4828; AA, 0000-0003-1842-9297

To elucidate the relationships and spatial range evolution across the world of the bird genus *Turdus* (*Aves*), we produced a large genomic dataset comprising *ca* 2 million nucleotides for *ca* 100 samples representing 53 species, including over 2000 loci. We estimated time-calibrated maximum-likelihood and multispecies coalescent phylogenies and carried out biogeographic analyses. Our results indicate that there have been considerably fewer trans-oceanic dispersals within the genus *Turdus* than previously suggested, such that the Palaearctic clade did not originate in America and the African clade was not involved in the colonization of the Americas. Instead, our findings suggest that dispersal from the Western Palaearctic via the Antilles to the Neotropics might have occurred in a single event, giving rise to the rich Neotropical diversity of *Turdus* observed today, with no reverse dispersals to the Palaearctic or Africa. Our large multilocus dataset, combined with dense species-level sampling and analysed under probabilistic methods, brings important insights into historical biogeography and systematics, even in a scenario of fast and spatially complex diversification.

1. Introduction

The family Turdidae (*Aves*, Passeriformes) comprises 21 genera distributed worldwide, where *Turdus* is the most species-rich genus (86 spp. [1]), accounting for half of the family's species diversity. Species in the genus are generally similar in shape but show a large variation in plumage colouration and ecology, exhibiting a variety of colour patterns and inhabiting almost all biomes in the world—including, savannahs, alpine areas, and both tropical and temperate forests [1]. This genus contains many common and familiar species throughout the world, such as the Blackbird (*Turdus merula*), the American Robin (*Turdus migratorius*), and the Clay-coloured Thrush or yigüirro (*Turdus grayi*).

Previous attempts to resolve phylogenetic relationships within *Turdus* have yielded conflicting phylogenetic hypotheses, and, consequently, different biogeographic interpretations [2–4]. One of their main limitations so far has been the reliance on a very limited number of independent genetic markers, making the analyses vulnerable to stochastic effects. Voelker *et al.* [2], using only mitochondrial DNA, recovered four main clades (African, Central

America-Caribbean, South American, and Eurasian), but their topology lacked nodal support for all but the Eurasian and a South American clade, with the latter also including an African species. Based on their results, the authors suggested including the genera *Platycichla*, *Cichlerminia*, and *Nesocichla* within *Turdus*. In another study, Nylander *et al.* [3] analysed a combined dataset from mitochondrial and nuclear markers and inferred one Eurasian, three American, and three African clades. Their results led those authors to suggest that: (i) the apparent non-monophyly of the African taxa might be illusory, (ii) several independent dispersal events have taken place from an African source to South America, and (iii) that *Platycichla*, *Cichlerminia*, and *Nesocichla* should indeed be placed within *Turdus*. Owing to the partially conflicting results of these studies, the evolutionary relationships and biogeographic history within *Turdus* remain unresolved.

The phylogenetic uncertainties concern not only relationships and biogeographic history among the major clades in *Turdus* but also the taxonomic and phylogeographic implications of recently diverged polytypic species and species complexes. Recent phylogeographic studies of species complexes within *Turdus* revealed additional taxonomic complexity at the level previously regarded as intraspecific, indicating that current taxonomy underestimates the true species diversity within the genus [5–8].

In summary, the controversies surrounding the early diversification and biogeography of *Turdus*, the circumscription and relationships among major clades and species, and our insufficient knowledge of intraspecific divergences and species limits, make *Turdus* a prime candidate for testing the utility of phylogenomic approaches for tackling evolutionary questions among and within bird species.

In this study, we address the following questions: (1) can phylogenomic data confidently delimit major clades within *Turdus* and elucidate the relationships among them? (2) How many trans-Atlantic dispersals happened during the diversification of *Turdus*, leading to its current cosmopolitan distribution? (3) Has the African continent functioned as a source for intercontinental dispersals? To address these questions, we infer the phylogeny of the genus *Turdus* based on the novel and extensive genomic data (ultraconserved elements (UCEs), intragenic and untranslated regions, respectively) and multiple analytical approaches (maximum-likelihood, species tree inference, ancestral area reconstructions). To maximize taxonomic sampling, we also compiled and analysed a dataset of sequences of the mitochondrial marker cytochrome *b* (*cytb*), for all currently recognized species within the genus [2,3,5,6,9].

2. Material and methods

A full description of the data and methods can be found in electronic supplementary material, tables S1–S4 and figure S1. The overall sampling included 115 individuals from 53 species (approx. 62%) of 86 currently recognized species within the genus *Turdus* [1]. We used a UCE probe set developed by Faircloth *et al.* [10] and developed additional specific probes for *Turdus* based on 49 of the loci described in Backström *et al.* [11].

(a) Phylogenomic and divergence time estimates

We performed unpartitioned concatenated maximum-likelihood (ML) analyses in RAxML 8.2.9 [12] for two datasets: (1) a matrix including UCEs (two different thresholds of missing data, 75%

and 85%) and (2) a matrix including the additional non-UCE loci (intragenic and untranslated regions). We assessed support for the best ML topology by performing 150 nonparametric bootstrap (BS) replicates using the autoMRE option in RAxML with the GTR GAMMA site-rate substitution model.

To avoid loci that could bias species tree analyses, we calculated the number of parsimony-informative sites (PIS) for each UCE locus and created a subset of UCE loci having a number of PIS in the upper quartile of the range (every locus with equal or more than 65 PIS). For the additional non-UCE loci, we included all markers recovered. We performed nonparametric bootstrapping by sites for each locus using RAxML with the GTR-GAMMA model of site-rate substitution and sorted the bootstrapped gene trees following Seo [13], an approach for estimating phylogeny and calculating bootstrapping using multilocus sequence data in the context of a distance method. To infer the species tree and calculate nodal support from the frequency of splits in trees constructed, we inputted the results to ASTRAL-II 4.7.8 [14]. We set ASTRAL to run 150 BS replicates of species trees. We also performed a species tree analysis using SVDquartets [15] for the UCE loci, including all loci recovered. By analysing fewer loci for the species tree, we were able to compare the results using different approaches.

To estimate divergence times, we used two late Miocene fossils from Hungary, correlated to the Mammal Neogene zone 13 (MN13) (7.2–4.9 Ma). These two fossils were compared to *T. viscivorus*, *T. pilaris*, and *T. torquatus*, and *T. merula* and *T. philomelos*, respectively. We tentatively interpret this as that they represent forms contemporary with the stem node of the main global radiation of thrushes, and use them for setting the minimum age for the clade including all *Turdus* species minus *T. philomelos* and *T. viscivorus* (figure 3). To fix a maximum age for this same node, we used the minimum age for the split between Regulidae and Certhioidea (17.2 Ma), based on fossil data and assigned in [16] (figure 3). As an additional secondary calibration point, we added the split between Muscicapidae and Turdidae, with a confidence interval of 22–13.5 Ma gleaned from [16] (figure 3). A UCE ML tree (the topology recovering highest support values overall) was time-calibrated under penalized likelihood using treePL [17]. The rate smoothing parameter was set to 100. The smoothing value was established using the cross-validation parameter available in treePL, we left the cvstart = 1000 (default) and cvstop = 0.1 (default).

(b) Biogeographic analyses

To infer the spatial range evolution for *Turdus*, we used Biogeography with Bayesian (and likelihood) Evolutionary Analysis in R Scripts—BioGeoBEARS, version 0.2.1 [18]. We used the time-calibrated phylogeny for *Turdus* calculated in treePL and coded each species as present or absent, according to their ranges following del Hoyo [1]. Eight biogeographic areas were defined based on the zoogeographical realms estimated in [19], with adaptations (figure 3): 1. Afrotropical (Afr), 2. Western Palearctic (WP), 3. Eastern Palearctic (EP), 4. Oriental (O), 5. Nearctic (N), 6. Panamanian (C), 7. Antilles (Ant), 8. South America (SA).

3. Results

(a) Phylogenetic relationships

Our results using different thresholds of concatenated UCE datasets (75% and 85%) recovered *T. viscivorus* (historical samples GNM8238 and GNM172, from Sweden) and *T. philomelos* (U4329, from Belarus) as incremental sisters to all other *Turdus* species (figure 1 and electronic supplementary material, figures S2 and S3). In the main clade, all taxa from the Afrotropical region, including *T. pelios*, previously

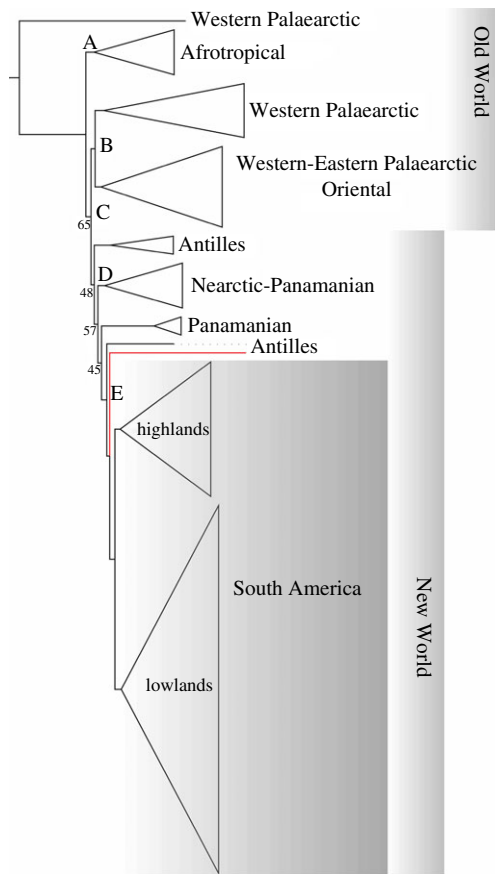


Figure 1. Phylogeny of *Turdus*. Topology with branch lengths, based on concatenated ML analysis of 1931 UCE loci. Numbers on nodes indicate ML BS support. Nodes without numbers indicate ML BS support greater than 75%. Letters A–E depict the main nodes mentioned in the text. Branch in red represents a historical sample (toepads) included in this study. Zoogeographical regions, according to the current distribution of *Turdus* species, are provided to the right of the tree. Shaded boxes in grey show the subdivision of Old/New World areas and two main groups within South America (highland and lowland). (Online version in colour.)

contentiously placed as part of the South American radiation, form a strongly supported monophyletic radiation (Node A, figure 1 and electronic supplementary material, figure S2), that is sister to a clade composed of all other species. The Palearctic clade contains *T. iliaceus* and *T. merula*, with strong support. In previous studies, these two species have been found in varying phylogenetic positions that suggest different biogeographic histories from other WP species [2,3]. We consider the following cut-offs for BS values: 50–74% weak; 75–84% moderate and greater than 85% as strong support (electronic supplementary material, figure S2 and figure 1).

While support for both the African and the Eurasian clades is strong, support for a monophyletic group [Node D, figure 1] in the Americas is weak. The important first splits within this clade are weakly supported, but the topology suggests a first divergence between Antillean taxa (a strongly supported clade including *T. aurantius* from Jamaica, *T. plumbeus schistaceus* from Cuba, and *T. plumbeus ardosiacus* from Puerto Rico) and remaining taxa (figure 1). The following basal divergences are all between taxa now inhabiting the Nearctic, Panamanian, or Antillean regions [19]. Although the divergence pattern is mostly weakly supported, a biogeographic structure may be discernible. A strongly supported clade comprises mainly species found

in Central America (*T. rufitorques*, *T. nigrescens*, *T. infuscatus*), but also *T. migratorius*, a species widely distributed in North America. An exclusively Panamanian clade (strongly supported) consists of two subspecies of *T. plebejus* (*rafaelenses* from Nicaragua and *plebejus* from Panama). A large and strongly supported monophyletic Neotropical radiation (Node E; figure 1) includes the early (but poorly supported) sequential divergence of two species from the Antilles (*T. swalesi* from Hispaniola and *T. lherminieri* from Dominica) and is further subdivided into two main subclades corresponding to highland and lowland species, respectively.

The topology based on 49 additional loci (electronic supplementary material, figure S4) includes a concatenated matrix of 65 574 bp, using a taxon set with 105 individuals. BS support for a monophyletic *Turdus* is 100% (electronic supplementary material, figure S4). *Turdus viscivorus* is sister to the remainder of the clade with strong support. The next lineage to diverge is *T. philomelos*, which in turn is sister to the remainder of the clade. A basal split separates two clades with low BS support values—one large group includes species from the Palearctic, Africa, Nearctic, Panamanian, as well as the Antillean regions, and is sister to a group which includes species from the Antilles (*T. swalesi* and *T. lherminieri*) and all samples from the continental South America region (electronic supplementary material, figure S4, weakly supported). Support for a monophyletic African radiation is strong, and *T. swalesi* and *T. lherminieri* are recovered (although with weak support) as part of the Neotropical radiation, as for the UCE dataset (Node E, figure 1 and electronic supplementary material, figures S2 and S3).

The analysis of the *cytb* dataset (electronic supplementary material, figure S5) recovers a monophyletic *Turdus* clade (excluding *Otocichla mupinensis* and *Psophocichla litsisirupa*, which were recovered as sisters to the *Turdus* clade and which have both previously been included in *Turdus*), where *T. viscivorus* and *T. philomelos* are again the first lineages to diverge, being sisters to the remainder of the clade with high support. Within the main clade, most early-diverging nodes are unsupported, but otherwise, there are many similarities to previous studies [2,3]. African species are divided into six clades, with *T. pelios* sister to Antillean species and a Eurasian clade with the exception of *T. merula* and *T. iliaceus*. Unlike previous studies [2,3], there is, for the first time, high support for a monophyletic South American clade based on a single mitochondrial locus. A number of unexpectedly deep or shallow divergences of potential taxonomic interest are highlighted in electronic supplementary material, figure S5 and discussed in the electronic supplementary material.

The UCE species tree (electronic supplementary material, figure S6, with gene trees of 483 UCE loci as input in ASTRAL) recovers a similar topology to the concatenated UCEs and *cytb* regarding *T. viscivorus* and *T. philomelos*. The placement of these species was also recovered, although in reverse order compared to concatenated UCEs and *cytb*. Here, *T. viscivorus* and *T. philomelos* are sisters to a strongly supported clade, which consists of two major groups: one including Panamanian, WP, Antillean, African, Eastern Palearctic; and one South American group (Node D, figure 1), with a few Panamanian and Antillean species, plus two Eastern Palearctic species, *T. eunomus* and *T. naumanni*. Support for the diverging order is weak, but

several of the same groups as in figure 1 are recovered with strong support, e.g. *T. plebejus*, the Central American clade (*T. rufitorques*, *T. nigrescens*, *T. infuscatus*, *T. migratorius*), the African clade, and the Eurasian clade. Also, the Neotropical clade is similar to the topology in figure 1, including the division between the highland and lowland clades.

The UCE species tree (electronic supplementary material, figure S7, with 1931 UCE loci as input in SVDquartets) shows *T. lherminieri* from the Antilles as sister species to all other species of the crown group. Again, this analysis recovers most of the same clades as the other analyses performed in this study, while there is no resolution regarding the phylogenetic relationships among these large groups. The similarities between analyses are the recovery of a monophyletic African group, a monophyletic Eurasian group, a monophyletic Panamanian group, and a generally similar Neotropical clade.

The species tree built from the 49 additional loci (electronic supplementary material, figure S8; with gene trees of each additional locus as input) recovers a tree topology where the species *T. viscivorus* and *T. philomelos* are again the first to diverge, followed by two main groups (electronic supplementary material, figure S8): a moderately supported Afro-tropical clade sister to the remainder of the thrushes from all non-African regions. In this tree, the main topology is insufficiently supported, but major clades such as the Eurasian clade including *T. iliacus* and *T. merula* are recovered, although support is weak. The American radiation is recovered as monophyletic, but again with weak support. Although there are differences in detail, major patterns in the Panamanian and Neotropical clades proposed by previous analyses shown here are corroborated.

(b) Divergence time estimates

The crown age of *Turdus* is estimated to *ca* 9.3 million years (Myr) (figure 3 and electronic supplementary material, figure S9), in the late Miocene. The first event of the major global radiation starts *ca* 7.2 million years ago (Ma), corresponding to the divergence of the African clade, rapidly followed by the divergence at *ca* 5.7 Ma between the Palearctic and Oriental clades from the clade comprising species with distributions in the New World. Around 5.3 Ma, the New World radiation begins. Panamanian and South American clades first begin to diverge at *ca* 4.4 Ma and 4.9 Ma, respectively.

(c) Geographical ancestral range reconstruction

Including the 'jump dispersal' (j) parameter in the DEC model (figure 3), implemented in BioGeoBEARS, suggests a similar scenario as the DEC reconstruction (not shown), but causes the reconstructions at the nodes to become less noisy. Independent dispersals originating in the WP and moving towards the Eastern Palearctic continuing to the Oriental region are suggested. The reconstruction of the New World radiation (including N, C, GA, and SA, figure 3 for abbreviations) suggests that it originated from a single dispersal event from the WP to the Antilles, although weak clade support makes this conclusion tentative (figure 1). From the Antilles, one or two dispersals to the Central America region are inferred, but the interpretation is complicated. From the Panamanian region, one lineage (*T. migratorius*) was recovered to have dispersed to North

America. The radiation in South America is inferred to have originated directly from an ancestor dispersing from the Great Antilles. Separate dispersals back to the Central America region from South America are implied for the ancestors of *T. assimilis* and *T. grayi* (figure 3).

4. Discussion

(a) Data considerations

The different evolutionary scenarios that have been suggested by studies of *Turdus* [2,3] propose conflicting scenarios of how they spread across the world. Most likely, the main reasons for the discrepancies are the differences in the amount and source of genetic data, combined with the stochasticity introduced by the short internodes caused by the high rate of speciation during the radiation. The UCE topology (ML) with 1931 concatenated loci (75% completeness of taxa) is the best currently available overview of the *Turdus* radiation, due to the highest number of loci sampled and the better resolution recovered to date (figure 1 and electronic supplementary material, figure S2). The concatenated UCE dataset shows reduced incongruence and improved resolution of the topology, as already documented for this kind of approach [20]. The more limited genetic sampling used by Voelker *et al.* [2] and Nylander *et al.* [3] makes these analyses more vulnerable to disagreement between the genealogy of the loci analysed and the species tree, as well as to the effects of various evolutionary processes, such as incomplete lineage sorting, gene flow, and hybridization. By exploring the different kinds of datasets obtained by genomic sampling, we were able to estimate a novel phylogenetic hypothesis based on 1931 UCEs (figure 1 and electronic supplementary material, figures S2 and S3). These data suggest a much simpler and parsimonious dispersal scenario compared to previous studies.

(b) Divergence time estimates

The age of the split between *Oenanthe* and *Turdidae* was estimated to be 22 Ma, corresponding to the upper extent of the confidence limit given for this split in figure 2 in Oliveros *et al.* [16]. This places the radiation of *Turdidae* within a time frame compatible with several recent studies [16,21–24]. Furthermore, the age of divergences between species is also compatible with the standard 2.1% per million years of divergence rate in the *cytb* gene [25].

(c) Phylogenetic implications and the historical biogeography of *Turdus*

Our time-calibrated phylogeny based on 1931 UCEs (figure 3 and electronic supplementary material, figure S9) sheds new light on the relationships among the majority of species in the genus *Turdus*. Previous studies suggested a series of trans-Atlantic dispersal events, both from Africa to the New World and from the Caribbean Islands to the Old World. Our findings challenge these hypotheses and that the African thrushes acted as a source for dispersal to the New World [3,26] and suggest a different scenario for the evolution of the genus *Turdus*. Although we cannot completely rule out the possibility of more than one trans-Atlantic dispersal (see *Phylogenetic incongruences leading to alternative phylogenetic*

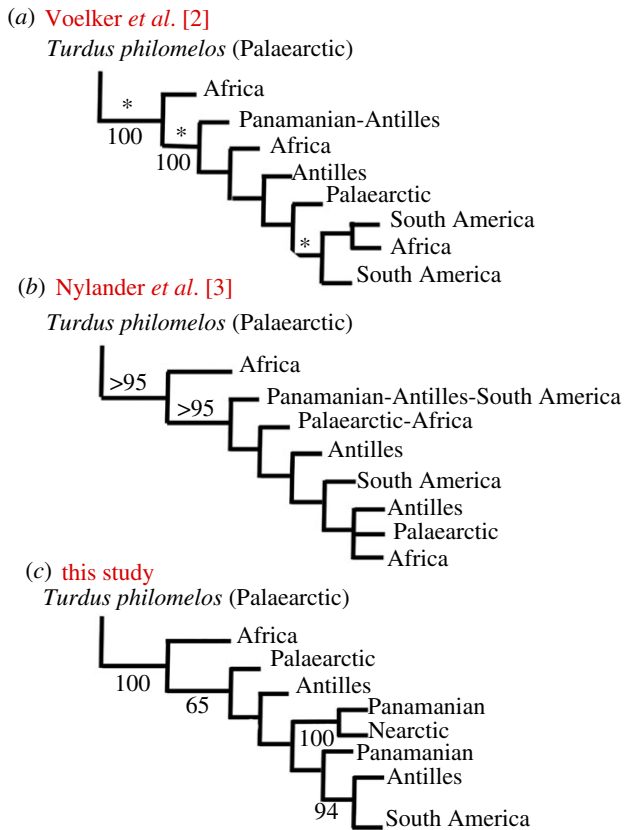


Figure 2. Biogeographic hypotheses for *Turdus*. Summarized from (a) Voelker *et al.* [2] (two mitochondrial genes), asterisks above nodes indicate Bayesian posterior probability greater than or equal to 0.95, numbers below nodes indicate ML BS support (greater than 50%); (b) Nylander *et al.* [3] (two mitochondrial and two nuclear genes), numbers above nodes indicate posterior probability (greater than 0.95); (c) this study (genomic sampling of 1931 UCEs), numbers below nodes indicate maximum-likelihood BS values. Only major clades are shown. (Online version in colour.)

histories in the electronic supplementary material), the topologies allowing for more than one trans-Atlantic dispersal are unsupported in that respect, and in some details most likely spurious. Our statistically most well-supported results based on 1931 UCEs differ most importantly from previous hypotheses in that (i) the African radiation is monophyletic, (ii) Eurasian taxa did not originate from a west to east trans-Atlantic dispersal, and (iii) that the American taxa originate from a single dispersal from the Old World (WP) to the Americas, a result recovered also with low support in the analysis of 49 independent loci.

Ancestral lineages are here inferred to having occurred in the WP, and from there, dispersals to the Afrotropical, Eastern Palearctic, and South American regions took place. The Afrotropical radiation is inferred to have been entirely *in situ*, with no indication that it had any role in the colonization of South America (figure 3 and electronic supplementary material, figure S9). The radiation in the Old World seems to have originated in the WP, with a subsequent spread towards the Eastern Palearctic and Oriental regions on at least two separate occasions (figure 3 and electronic supplementary material, figure S9). We note, however, that genomic sampling of Palearctic species is incomplete, making these conclusions tentative. Contrary to previous hypotheses, *T. pelios* is recovered in the African clade and is no longer suggested to be more closely related to species from the Caribbean (figures 2 and 3), thus contradicting

both a close relationship with other taxa from South America [2], as well as a reversed trans-Atlantic dispersal event, as inferred by both Voelker *et al.* [2] and Nylander *et al.* [3]. *T. philomelos* and *T. viscivorus* represent ancient lineages that seem to have been unaffected by the processes driving the global radiation from the Palearctic region (figures 1 and 3 and electronic supplementary material, figures S2–S9). *T. iliacus* is inferred to be sister to *T. merula* and both are part of the Palearctic clade, with strong support. By contrast, both Voelker *et al.* [2] and Nylander *et al.* [3] recovered *T. iliacus* within the New World clade, as sister to *T. plebejus*, a Panamanian species.

We find evidence for the Antilles as the location of the first colonization event (figure 3). In the ancestral range reconstruction shown in figure 3, an Antillean distribution receives the highest probability, but Central America is also a possibility. *Turdus* thus seems to be an example of a long-distance dispersal from the Palearctic where the birds first settled in the Caribbean region and then proceeded to radiate and colonize the mainland. These new data receive higher phylogenetic support than previous studies and generate a both more parsimonious explanation and a more logical scenario for the colonization of the Americas by *Turdus*.

The initial events in the colonization of the Americas seem to be centred around the Caribbean region. The most likely scenario according to the ancestral range reconstruction (figure 3) is a dispersal from a founder population in the Greater Antilles to Central America, resulting in one limited radiation including the ancestors of the *T. rufitorques* clade and in the *T. plebejus* lineage, followed by a reverse dispersal back to the Antilles from the Panamanian region by the ancestors of *T. swalesi* and *T. lherminieri*. In an alternative scenario, including two dispersals from the Greater Antilles to the Panamanian region, the ancestor of the clade including *T. rufitorques* would have had to disperse first, followed by the ancestor of *T. plebejus*, while the ancestors of *T. swalesi* and *T. lherminieri* would emanate from dispersal events within the Antilles.

Thrushes probably entered the Panamanian region by the time a land bridge was already fully or partly established [27–29]. There is an ongoing debate on the emergence of the Panamanian land bridge and the timing and processes underlying the Great American Biotic Interchange [29], which is partly due to difficulties in the interpretation of the exposure of land and shallow waters [28]. The closure of the Panamanian land bridge has had an important impact on the diversification of a variety of New World birds, as shown by Smith and Klicka [30]. Weir *et al.* [31], Töpel *et al.* [32], Antonelli *et al.* [33], Oliveros *et al.* [34], among others, have shown a marked increase in the rates of interchange of South and Central American birds, after or in connection with land bridge completion. These studies suggest that the route of most dispersal events was primarily south to north.

While a common pattern for birds, thrushes are likely to only have been marginally affected by the emergence of the Panamanian uplift. The colonization of South America is instead inferred to be a dispersal event directly from the Antilles, with no evidence of any connection to the Panamanian land bridge. This is not surprising, since these birds evidently dispersed to and between the Antillean islands, and could thus have dispersed between the Panamanian or Antillean regions and continental South America

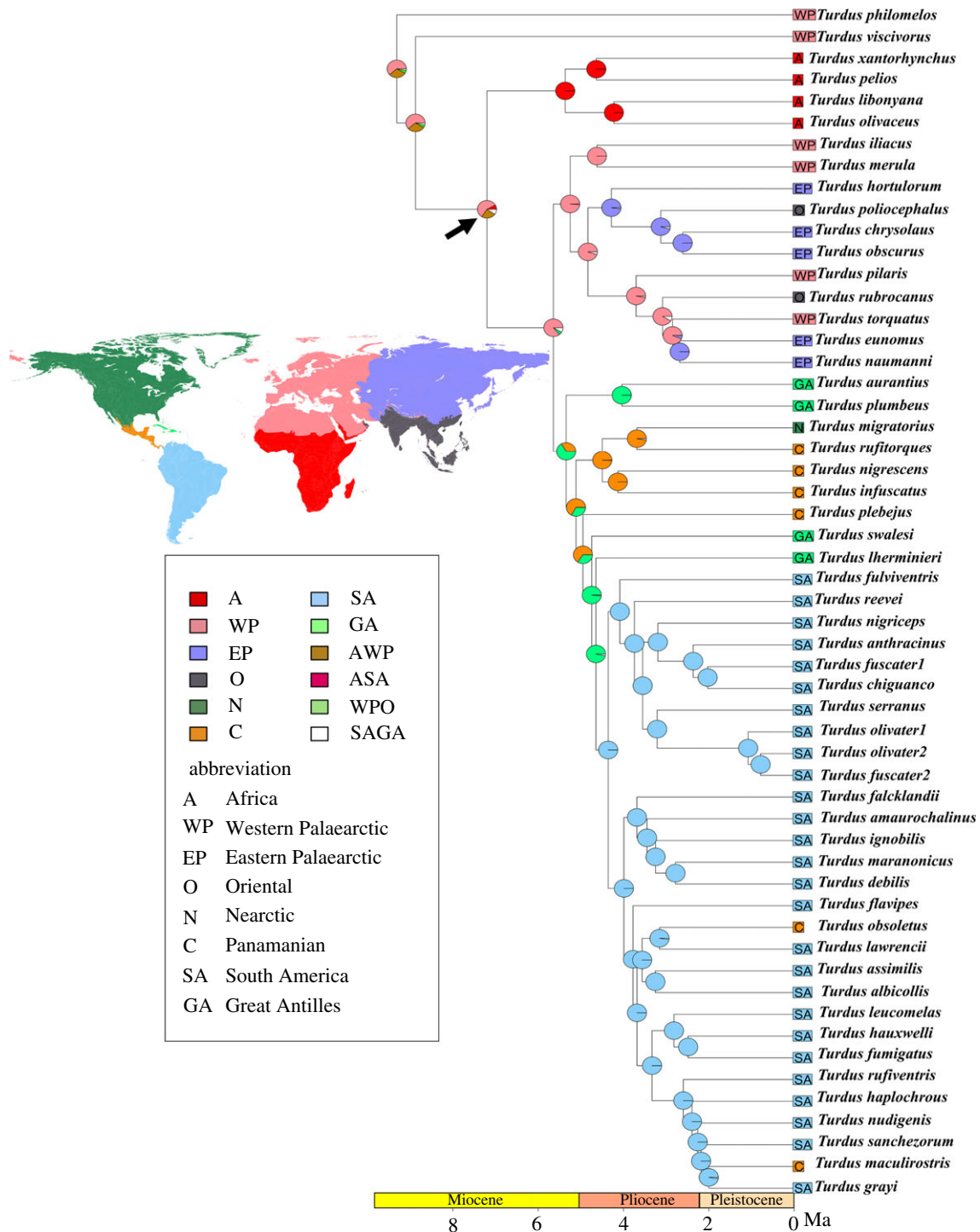


Figure 3. Ancestral range reconstruction. *Turdus* time-calibrated phylogeny, based on 1931 UCEs, calibrated with one fossil calibration (black arrow), and with range evolution estimated under the DEC + j model. Pie charts show the probability of each range. All nodes including secondary colours represent the uncertainty to recover a unique ancestral range for certain nodes within the phylogeny. The map shows the biogeographic regions used for the analyses (adapted from Holt *et al.* [19]). (Online version in colour.)

independently of a land bridge. This result shows similarities to findings from studies of other taxa. For instance, Sturge *et al.* [35] showed, based on mitochondrial DNA (cytb and ND2), that the colonization of South America by New World orioles (*Icterus*) is likely to have involved ancestors from the Caribbean, and a recent biogeographic reconstruction recovered reverse colonization by short-faced bats from the islands (e.g. Antilles) to South America [36]. Other examples include the genus *Megascops* (Aves, Strigidae) endemic to the New World with a likely origin in the Caribbean Islands [37].

By improving the taxonomic and geographical sampling within South America, we recovered for the first time a single, monophyletic South American clade within *Turdus*. By contrast, Nylander *et al.* [3] inferred three South American clades, and thus suggested independent colonization events. Our results regarding the South American radiation are similar to those of Voelker *et al.* [2] in this respect, apart from the African species (*T. pelios*) that was part of the American clade in their analysis but is part of the African clade in this study (figures 1 and 2 and electronic supplementary material, figure S2).

According to our phylogenetic hypothesis, the South American clade is divided into two main groups. One comprises mainly highland species (mostly Andean and Tepuian, but one Atlantic Forest species) and corresponds to an Andean clade already identified in Voelker *et al.* [2]. The second main clade includes species from the lowlands in Amazonia, Chaco, Cerrado, and Pampas, as well as Central America (*T. grayi* and *T. assimilis*). Paraphyly of *T. olivater* was detected, with *T. olivater kemptoni* (U4322, Venezuela) being recovered as sister to *T. fuscater fuscater* (U4285, Bolivia) and the Tepuian *T. olivater roraimae* sister to these two. This pattern has been documented in a previous study using a multilocus dataset [37]. Similarly, *T. fuscater* appears as paraphyletic, with *T. fuscater gigantoides* (U4287, Peru) being sister to *T. chiguanco conradi* (U4279, Peru) [38]. From continental South America, a number of independent reversals—dispersal events to both the Antilles and the Panamanian region—are inferred, which in the latter case could have been aided by the land bridge.

Future studies could focus on clarifying the geographical range evolution of *Turdus* within South America and along altitudinal zones, as well as investigating the influence of landscape and climatic changes on their past and current distribution and genetic diversity.

5. Conclusion

By employing broad genomic data with dense taxonomic sampling (see electronic supplementary material, figure S10 for an overview of all approaches used in this study), we were able to not only provide an improved phylogenomic framework for continued taxonomy and classification revisions, but also a starting point for more detailed studies of the morphological, behavioural, and ecological evolution in thrushes. In addition, our results provide strong evidence to reconsider ongoing controversies on the early evolution and range evolution of *Turdus*, including the logical but novel inference of a monophyletic *in situ* radiation in Africa, no American origin of the major Palearctic radiation, and a single rather than multiple trans-Atlantic dispersals giving rise to the American radiation. In summary, phylogenomic analyses are not only useful at higher taxonomic levels [39] but also hold the potential to uncover new insights into the diversification and biogeographic history of rapid species radiations.

References

- del Hoyo J, Elliott A, Sargatal J, Christie DA, Kirwan G. 2019 *Handbook of the birds of the world alive*. Barcelona, Spain: Lynx Edicions.
- Voelker G, Rohwer S, Bowie RCK, Outlaw DC. 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. (doi:10.1016/j.ympev.2006.07.016)
- Nylander JA, Olsson U, Alstrom P, Sanmartin I. 2008 Accounting for phylogenetic uncertainty in biogeography: a Bayesian approach to dispersal–vicariance analysis of the thrushes (*Aves: Turdus*). *Syst. Biol.* **57**, 257–268. (doi:10.1080/10635150802044003)
- Nagy J, Végvári Z, Varga Z. 2019 Phylogeny, migration and life history: filling the gaps in the origin and biogeography of the *Turdus* thrushes. *J. Ornithol.* **160**, 529–543. (doi:10.1007/s10336-019-01632-3)
- Melo M, Bowie RCK, Voelker G, Dallimer M, Collar NJ, Jones PJ. 2010 Multiple lines of evidence support the recognition of a very rare bird species: the Principe thrush. *J. Zool.* **282**, 120–129. (doi:10.1111/j.1469-7998.2010.00720.x)
- Núñez-Zapata J, Peterson AT. 2016 Pleistocene diversification and speciation of white-throated thrush (*Turdus assimilis*; *Aves: Turdidae*). *J. Ornithol.* **157**, 1073–1085. (doi:10.1007/s10336-016-1350-6)
- Cerqueira PV, Santos MPD, Aleixo A. 2016 Phylogeography, inter-specific limits and diversification of *Turdus ignobilis* (*Aves: Turdidae*). *Mol. Phylogenet. Evol.* **97**, 177–186. (doi:10.1016/j.ympev.2016.01.005)

Data accessibility. All DNA sequence data produced in this project are made freely available at NCBI Sequence Read Archive (SRA) under the BioProject accession PRJNA574741 (electronic supplementary material, table S1). The project's documentation, comprising the origin of all data, download information, and basic information, is freely available on GitHub: <https://github.com/RominaSSBatista/Turdus-Project>.

Authors' contributions. R.B., C.C.R., A.A.I., A.A.n., and U.O. conceived, designed, and coordinated the study and helped draft the manuscript; R.B. and U.O. carried out the molecular laboratory work; R.B., T.A., and U.O. participated in data analysis; R.B. led the writing with contributions from all authors. All authors gave final approval for publication.

Competing interests. The authors have no competing interests to declare.

Funding. This research was supported by the Knut and Alice Wallenberg Foundation through a Wallenberg Academy Fellowship, the Swedish Research Council, and the Swedish Foundation for Strategic research (grants to A.A.n.); the US National Science Foundation (NSF DEB 1241056), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2012/50260-6) and CNPq (fellowships 306843/2016-1 to A.A.I. and 308927/2016-8 to C.C.R.). R.B. received a Doctoral Fellowship and a Fellowship for Internship abroad from Coordination for the Improvement of Higher Education Personnel [CAPES, Processo 99999.000566/2015-02]. Genomic work was further supported by Science for Life Laboratory and the National Genomics Infrastructure funded by the Swedish Research Council.

Acknowledgements. We thank the curator and curatorial assistants of Burke Museum (UWBM); American Museum of Natural History (AMNH); Louisiana State University Museum of Natural Science (LSUMZ); South Western Forestry University (SWFU); Zoological Museum, University of Copenhagen (ZMUC); Naturhistoriska riksmuseet (NRM); Department of Biology and Environmental Science, University of Gothenburg (DZUG); Göteborg Naturhistoriska Museum (GNM and GNM-Av.ex); Martim Melo collection; Instituto Nacional de Pesquisas da Amazônia (INPA); Laboratório de Genética e Evolução Molecular de Aves (LGEMA); Divisão de Aves do Museu de Zoologia da Universidade Estadual de Feira de Santana (DAM-ZUFS) and Museu Paraense Emílio Goeldi (MPEG), for loaning tissue samples (electronic supplementary material, table S1). We thank the Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure. Bioinformatics analyses were run on the Albiorix computer cluster (<http://albiorix.bioenv.gu.se/>) at the University of Gothenburg. The authors also acknowledge the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer, which have contributed to the research results reported within this paper. We also thank Anna Ansebo for her support in the entire process of the wet laboratory; Mats Töpel, for his support with bioinformatics and constructive advice that greatly improved the data process. The first author also thanks Chrysoula Gubili, Erico Polo, and Mateus Ferreira for helpful comments and guidance toward improving the manuscript.

8. Avendano JE, Arbelaez-Cortes E, Cadena CD. 2017 On the importance of geographic and taxonomic sampling in phylogeography: a reevaluation of diversification and species limits in a neotropical thrush (Aves, Turdidae). *Mol. Phylogenet. Evol.* **111**, 87–97. (doi:10.1016/j.ympev.2017.03.020)
9. Bowie RCK, Bloomer P, Clancey PA, Crowe TM. 2003 The Karoo Thrush (*Turdus smithi* Bonaparte 1850), a southern African endemic. *Ostrich* **74**, 1–7. (doi:10.2989/00306520309485364)
10. Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. 2012 Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst. Biol.* **61**, 717–726. (doi:10.1093/sysbio/sys004)
11. Backström N, Fagerberg S, Ellegren H. 2008 Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Mol. Ecol.* **17**, 964–980. (doi:10.1111/j.1365-294X.2007.03551.x)
12. Stamatakis A. 2014 RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
13. Seo TK. 2008 Calculating bootstrap probabilities of phylogeny using multilocus sequence data. *Mol. Biol. Evol.* **25**, 960–971. (doi:10.1093/molbev/msn043)
14. Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. 2014 ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* **30**, i541–i548. (doi:10.1093/bioinformatics/btu462)
15. Chifman J, Kubatko L. 2014 Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**, 3317–3324. (doi:10.1093/bioinformatics/btu530)
16. Oliveros CH *et al.* 2019 Earth history and the passerine superradiation. *Proc. Natl Acad. Sci. USA* **116**, 7916–7925. (doi:10.1073/pnas.1813206116)
17. Smith SA, O'Meara BC. 2012 treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**, 2689–2690. (doi:10.1093/bioinformatics/bts492)
18. Matzke NJ. 2014 Model selection in historical biogeography reveals that founder-event speciation is a crucial process in Island Clades. *Syst. Biol.* **63**, 951–970. (doi:10.1093/sysbio/syu056)
19. Holt BG *et al.* 2013 An update of Wallace's zoogeographic regions of the world. *Science* **339**, 74–78. (doi:10.1126/science.1228282)
20. Salichos L, Rokas A. 2013 Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature* **497**, 327–331. (doi:10.1038/nature12130)
21. Claramunt S, Cracraft J. 2015 A new time tree reveals Earth history's imprint on the evolution of modern birds. *Sci. Adv.* **1**, e1501005. (doi:10.1126/sciadv.1501005)
22. Moyle RG, Oliveros CH, Andersen MJ, Hosner PA, Benz BW, Manthey JD, Travers SL, Brown RM, Faircloth BC. 2016 Tectonic collision and uplift of Wallacea triggered the global songbird radiation. *Nat. Commun.* **7**, 12709. (doi:10.1038/ncomms12709)
23. Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. 2015 A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569–573. (doi:10.1038/nature15697)
24. Selvatti AP, Gonzaga LP, Russo CA. 2015 A Paleogene origin for crown passerines and the diversification of the Oscines in the New World. *Mol. Phylogenet. Evol.* **88**, 1–15. (doi:10.1016/j.ympev.2015.03.018)
25. Weir JT, Schluter D. 2008 Calibrating the avian molecular clock. *Mol. Ecol.* **17**, 2321–2328. (doi:10.1111/j.1365-294X.2008.03742.x)
26. Voelker G, Rohwer S, Outlaw DC, Bowie RCK. 2009 Repeated trans-Atlantic dispersal catalysed a global songbird radiation. *Glob. Ecol. Biogeogr.* **18**, 41–49. (doi:10.1111/j.1466-8238.2008.00423.x)
27. Montes C *et al.* 2012 Evidence for middle Eocene and younger land emergence in central Panama: implications for Isthmus closure. *Geol. Soc. Am. Bull.* **124**, 780–799. (doi:10.1130/b30528.1)
28. Bacon CD, Silvestro D, Jaramillo C, Smith BT, Chakrabarty P, Antonelli A. 2015 Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl Acad. Sci. USA* **112**, 6110–6115. (doi:10.1073/pnas.1423853112)
29. Jaramillo C, Montes C, Cardona A, Silvestro D, Antonelli A, Bacon CD. 2017 Comment (1) on 'Formation of the Isthmus of Panama' by O'Dea *et al.* *Sci. Adv.* **3**, e1602321. (doi:10.1126/sciadv.1602321)
30. Smith BT, Klicka J. 2010 The profound influence of the Late Pliocene Panamanian uplift on the exchange, diversification, and distribution of New World birds. *Ecography* **33**, 333–342. (doi:10.1111/j.1600-0587.2009.06335.x)
31. Weir JT, Bermingham E, Schluter D. 2009 The Great American Biotic Interchange in birds. *Proc. Natl Acad. Sci. USA* **106**, 21 737–21 742. (doi:10.1073/pnas.0903811106)
32. Töpel M, Zizka A, Calio MF, Scharn R, Silvestro D, Antonelli A. 2017 SpeciesGeoCoder: fast categorization of species occurrences for analyses of biodiversity, biogeography, ecology, and evolution. *Syst. Biol.* **66**, 145–151. (doi:10.1093/sysbio/syw064)
33. Antonelli A, Zizka A, Carvalho FA, Scharn R, Bacon CD, Silvestro D, Condamine FL. 2018 Amazonia is the primary source of neotropical biodiversity. *Proc. Natl Acad. Sci. USA* **115**, 6034–6039. (doi:10.1073/pnas.1713819115)
34. Oliveros CH, Andersen MJ, Hosner PA, Mauck WM, Sheldon FH, Cracraft J, Moyle RG. 2019 Rapid Laurasian diversification of a pantropical bird family during the Oligocene-Miocene transition. *Ibis* **162**, 137–152. (doi:10.1111/ibi.12707)
35. Sturge RJ, Jacobsen F, Rosensteel BB, Neale RJ, Omland KE. 2009 Colonization of South America from Caribbean Islands confirmed by molecular phylogeny with increased taxon sampling. *The Condor* **111**, 575–579. (doi:10.1525/cond.2009.080048)
36. Tavares VD, Warsi OM, Balseiro F, Mancina CA, Dávalos LM. 2018 Out of the Antilles: fossil phylogenies support reverse colonization of bats to South America. *J. Biogeogr.* **45**, 859–873. (doi:10.1111/jbi.13175)
37. Dantas SM, Weckstein JD, Bates JM, Krabbe NK, Cadena CD, Robbins MB, Valderrama E, Aleixo A. 2016 Molecular systematics of the new world screech-owls (Megascops: Aves, Strigidae): biogeographic and taxonomic implications. *Mol. Phylogenet. Evol.* **94**, 626–634. (doi:10.1016/j.ympev.2015.09.025)
38. Valderrama E, Pérez-Emán JL, Brumfield RT, Cuervo AM, Cadena CD, Patten M. 2014 The influence of the complex topography and dynamic history of the montane neotropics on the evolutionary differentiation of a cloud forest bird (*Premnoplex brunnescens*, Furnariidae). *J. Biogeogr.* **41**, 1533–1546. (doi:10.1111/jbi.12317)
39. Zhang G *et al.* 2014 Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320. (doi:10.1126/science.1251385)