On the importance of geographic and taxonomic sampling in phylogeography: A reevaluation of diversification and species limits in a Neotropical thrush (Aves, Turdidae)

Jorge Enrique Avendaño a,⇑, Enrique Arbeláez-Cortés b,c, Carlos Daniel Cadena a

a Laboratorio de Biología Evolutiva de Vertebrados, Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia
b Colección de Tejidos, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Palmira, Colombia
c Grupo de Estudios en Biodiversidad, Escuela de Biología, Facultad de Ciencias, Universidad Industrial de Santander, Bucaramanga, Colombia

A R T I C L E   I N F O

Article history:
Received 1 November 2016
Revised 9 February 2017
Accepted 23 March 2017
Available online 24 March 2017

Keywords:
Biogeography
Colombia
Phylogenetic accuracy
Taxonomy
Taxon sampling
Turdus ignobilis

A B S T R A C T

Phylogeographic studies seeking to describe biogeographic patterns, infer evolutionary processes, and revise species-level classification should properly characterize the distribution ranges of study species, and thoroughly sample genetic variation across taxa and geography. This is particularly necessary for widely distributed organisms occurring in complex landscapes, such as the Neotropical region. Here, we clarify the geographic range and revisit the phylogeography of the Black-billed Thrush (Turdus ignobilis), a common passerine bird from lowland tropical South America, whose evolutionary relationships and species limits were recently evaluated employing phylogeographic analyses based on partial knowledge of its distribution and incomplete sampling of populations. Our work employing mitochondrial and nuclear DNA sequences sampled all named subspecies and multiple populations across northern South America, and uncovered patterns not apparent in earlier work, including a biogeographic interplay between the Amazon and Orinoco basins and the occurrence of distinct lineages with seemingly different habitat affinities in regional sympatry in the Colombian Amazon. In addition, we found that previous inferences about the affinities and taxonomic status of Andean populations assumed to be allied to populations from the Pantepui region were incorrect, implying that inferred biogeographic and taxonomic scenarios need re-evaluation. We propose a new taxonomic treatment, which recognizes two distinct biological species in the group. Our findings illustrate the importance of sufficient taxon and geographic sampling to reconstruct evolutionary history and to evaluate species limits among Neotropical organisms. Considering the scope of the questions asked, advances in Neotropical phylogeography will often require substantial cross-country scientific collaboration.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Phylogeographic studies examining the spatial distribution of gene lineages (Avise et al., 1987) have revolutionized our knowledge of the origin, diversification, and historical processes responsible for the population structure and geographic distribution of genetic variation within species (Hickerson et al., 2009). This is especially true for birds from Central and South America, which have been subject of considerable attention from phylogeographers working at the interface between populations and species (e.g., Bates et al., 2008; Smith and Klicka, 2010; Smith et al., 2014; Weir, 2009; Weir et al., 2009). Phylogeographic studies on Neotropical birds have been instrumental for describing and understanding geographic patterns of population differentiation and area relationships (Cabanne et al., 2008; Cheviron et al., 2005; Fernandes et al., 2013; Lovette, 2004; Miller et al., 2008), the tempo and mode of speciation (Bonaccorso, 2009; Chaves et al., 2011; Maldonado-Coelho et al., 2013; Ribas et al., 2012; Smith et al., 2012), the role of features of the landscape related to population divergence (Arbeláez-Cortés et al., 2010; Benham and Witt, 2016; Burney and Brumfield, 2009; Rodríguez-Gómez et al., 2013; Weir et al., 2015; Winger and Bates, 2015), and demographic processes associated with physical and climatic changes over Earth history (Cabanne et al., in press; Lougheed et al., 2013; Maldonado-Sánchez et al., 2016; Pulido-Santacruz et al., 2016; Valderrama et al., 2014). In addition, phylogeographic work on Neotropical birds has been crucial to reveal previously unrecognized diversity at the species level (Cadena and Cuervo, 2010; Isler et al., 2012; Milá et al., 2012; Nyári, 2007; Thom and Aleixo, 2015).

⇑ Corresponding author.
E-mail address: jorgeavec@gmail.com (J.E. Avendaño).
and to identify instances of non-monophyly of species (Cuervo et al., 2014; D’Horta et al., 2013; Dingle et al., 2006; Gutiérrez-Pinto et al., 2012), resulting in substantial changes to classification (Renssen et al., 2016).

The Black-billed Trush *Turdus ignobilis* (Aves, Turdidae) is a common and widely distributed songbird found in the lowlands of the Orinoco and Amazon basins, and in low- to mid-elevations of the Northern Andes including interAndean valleys and the Pantepui region (Collar, 2015). Recently, Cerqueira et al. (2016, hereafter CEA) conducted a phylogeographic study of *T. ignobilis* in which they made biogeographic inferences and concluded that this taxon should be divided in three different species based on mutlitocus DNA sequence data and plumage characters. However, CEA (2016) did not analyze samples of Andean populations (i.e., nominate *T. i. ignobilis* and *T. i. goodfellowi* from Colombia), which were assumed to be closely allied to and conspecific with populations from the Pantepui region based on putative similarities in plumage characters and ecology based on their occurrence in highland environments.

Because the lack of samples from highly diverse and geographically complex areas of northwestern South America may result in incomplete pictures of the history and evolutionary relationships of Neotropical organisms (Arbeláez-Cortés, 2013; Avendaño et al., 2016; Cadena et al., 2007), we were curious as to how Andean populations of *T. ignobilis* fit in with the biogeographic and taxonomic scenarios proposed by CEA (2016). In addition, we noted that CEA (2016) underestimated the geographic range of *T. ignobilis* and did not sample populations from the Colombian Amazon and Orinoco regions, where two different subspecies occur. Accordingly, we here clarify the distribution range of *T. ignobilis* and its constituent subspecies, and we reassess their phylogeography employing sequence data obtained from several populations from the Andes and lowlands east of the Andes of Colombia. Our results based on more complete taxonomic and geographic sampling work alter some of the conclusions reached by CEA (2016). More broadly, our work underscores that crucial to any phylogeographic study seeking to describe biogeographic patterns, to infer evolutionary processes, and to revise species-level classification is to (1) properly understand the distribution ranges of study species and (2) thoroughly sample genetic variation across relevant taxa and geographic regions.

2. Materials and methods

2.1. The geographic and elevational distribution of *Turdus ignobilis*

Differences in plumage characters, and body and bill dimensions allow one to recognize five to six subspecies of *T. ignobilis* (Collar, 2015; Dickinson and Christidis, 2014; Hellmayr, 1934). Andean taxa occur along the Central and Eastern cordilleras and the Magdalena and lower Cauca valleys of Colombia (*T. i. ignobilis*), and in the Western cordillera and arid Pacific and Upper Cauca valley of Colombia (*T. i. goodfellowi*). East of the Andes, *T. i. debilis* ranges from the foothills of the Andes of Colombia and Venezuela in the Maracaibo and Orinoco basins to western Amazonia into Brazil and northern Bolivia, where it inhabits gallery and terra firme forests. A second cis-Andean taxon, *T. i. arthuri*, ranges across Amazonian savannas and white-sand forest in the Orinoco Basin and tepuis of Colombia, Venezuela, Guyana and Surinam (this subspecies includes the synonymized taxon cururuensis from white-sand forest between the Xingu and Tapajós rivers in Brazil; CEA, 2016). Finally, *T. i. murinus* occurs in subtropical elevations in the tepuis of south Venezuela, Guyana and northwestern Brazil (Collar, 2015).

To characterize the distribution of *T. ignobilis*, we produced maps based on (1) locality data available from specimens from the Biomap Database (Darwin Database, 2014) and the Global Biodiversity Information Facility (http://www.gbif.org/), and (2) recordings of vocalizations archived in xeno-canto (http://www.xeno-canto.org/) and field observations documented in e-bird (Sullivan et al., 2009). Although these data are not free of problems (see Discussion), when used judiciously they may provide a useful complement to distribution polygons based on expert knowledge (BirdLife International and NatureServe, 2015) that served as the basis for range characterizations by CEA (2016). In addition, we extracted elevation data from localities to describe the elevational distribution of Andean populations. Maps and plots showing elevational distributions were generated in the R statistical software (R Core Team, 2014).

2.2. Taxonomic and geographic sampling

We examined variation in the mitochondrial ND2 gene in 76 individuals of *T. ignobilis*, 57 of which were available from previous studies (Voelker et al., 2007; Cerqueira et al., 2016). The remaining 19 individuals were sequenced by us using samples available in natural history collections in Colombia (Banco de ADN y Tejidos, Museo de Historia Natural ANDES at Universidad de los Andes; Colección de Tejidos, Instituto de Investigación de Recursos Biológicos Alexander von Humboldt). This material represents the two previously unsampled subspecies from Colombia (*T. i. ignobilis*, *T. i. goodfellowii*), as well as previously unsampled populations of subspecies *T. i. debilis* and *T. i. arthuri* from Colombia (Table S1 in Supplementary material). In this expanded data set, the five known subspecies of *T. ignobilis* are represented as follows: *T. i. murinus* (*n* = 19 individuals), *T. i. arthuri* (*n* = 30), *T. i. debilis* (*n* = 21), *T. i. ignobilis* (*n* = 3), and *T. i. goodfellowi* (*n* = 3). These samples correspond to 39 localities distributed over most of the species’ range (Fig. 1). In addition, we considered in analyses a total of 14 published ND2 sequences of eight additional species of *Turdus* selected based on results of previous phylogenetic analyses (Klicka et al., 2005; Nylander et al., 2008; O’Neill et al., 2011; Voelker et al., 2007; Table S2 in Supplementary material).

We also analyzed variation in sequences of the G3PDH intron, one of the two nuclear loci studied by CEA (2016). We sequenced this locus for 12 individuals from Colombia (three from each subspecies) having different ND2 haplotypes and sampled from different populations. When combined with sequences published by CEA (2016), our nuclear/mitochondrial data set consisted of a total of 67 individuals from 39 localities with both G3PDH and ND2 data corresponding to the five known subspecies of *T. ignobilis*: *T. i. murinus* (*n* = 19 individuals), *T. i. arthuri* (*n* = 25), *T. i. debilis* (*n* = 17), *T. i. ignobilis* (*n* = 3), and *T. i. goodfellowi* (*n* = 3). This data set was complemented with 12 G3PDH sequences from seven species of *Turdus* as part of the ingroup and outgroup. The only taxon not having G3PDH sequences was *T. eremita*, but we included it in phylogenetic analyses because of its close relationship with *T. ignobilis* taxa (CEA, 2016).

2.3. Laboratory protocols

We extracted DNA mainly from frozen tissue samples of vouchered specimens (Table S1, for three individuals only blood samples were available) using a Quick-DNATM Universal Kit (Zymo Research) following instructions from the manufacturer. We then PCR-amplified the ND2 mitochondrial gen (1041 pb) and G3PDH intron (382 pb) using primers LS215U and H1065U, and GapdL890 and GapdH950 (Friesen et al., 1997), respectively, in 50 μl reactions following published protocols (Arbeláez-Cortés et al., 2014). PCR products were purified using ExoSAP-IT (USB Corp.) and sequenced at Universidad de los Andes in Bogotá, Colombia.
Sequences were edited and aligned manually using Geneious v.9.1.2 (Kearse et al., 2012).

2.4. Phylogenetic analyses

We first reconstructed the phylogeny of *T. ignobilis* and allies based on sequences of the ND2 gene; previous studies on thrushes have shown that phylogenies based on this gene are similar to those obtained using other mitochondrial and nuclear markers (Klicka et al., 2005; Nylander et al., 2008; O’Neill et al., 2011; Voelker et al., 2007), and this was one of the markers chosen by CEA (2016) for their phylogeographic analysis. We used Bayesian inference (BI) and maximum-likelihood (ML) methods to reconstruct phylogenetic relationships based on the ND2 data set, consisting of 1041 base pairs. TrN + I (for 1st and 2nd codon positions) and HKY + I (2nd codon position) were selected as the best-fit models of molecular substitution for analyses employing a partitioned scheme according to the Bayesian Information Criterion in PartitionFinder v.1.1.0 (Lanfear et al., 2012). BI was implemented using MrBayes 3.2.2 (Huelsenbeck and Ronquist, 2001) with four parallel runs consisting of one cold and three heated chains with a temperature of 0.175 for 30 million generations, sampling every 1000 generations. The first 25% of sampled trees was discarded as burn-in, and a majority rule consensus tree was calculated among the trees retained. We used Tracer version 1.5 (Rambaut and Drummond, 2007) to assess convergence across independent runs and to examine effective sample size values for all parameters of Bayesian analyses; in all cases these were >200. We also used the RAxML v.8 algorithm (Stamatakis, 2014) to search for the maximum-likelihood (ML) tree and evaluated nodal support using 1000 bootstrap replicates. BI and ML analyses were run on XSEDE v.3.2.6 and XSEDE v.8.2.8 from the CIPRES portal, respectively (Miller et al., 2010; http://www.phylo.org/).

To analyze the G3PDH and the mitochondrial/nuclear concatenated data sets, we applied the same methods of phylogenetic inference described above. However, due to heterozygous sites in the nuclear sequences, we reconstructed haplotype phases using the PHASE algorithm in DNAsp V.5.10 (Librado and Rozas, 2009), ran for 10,000 iterations, with a thinning interval of 10 and a burn-in of 1000 iterations. Then, the obtained nuclear haplotypes were concatenated with their respective ND2 sequences. The best-fit models of molecular substitution according to the Bayesian Information Criterion in PartitionFinder v.1.1.0 (Lanfear et al., 2012) were K80 + I + G for all positions in the nuclear region, and HKY + G (for 1st codon position), HKY + I (2nd codon position) and TrN + I (for 3rd codon position) for the mitochondrial region. We also analyzed relationships among individuals using a median-joining haplotype network in the software PopART (http://popart.otago.ac.nz); for this analysis we excluded sites with missing data, which resulted in the alignments being reduced to 1002 bp (ND2) and 295 pb (G3PDH). Finally, we calculated mean uncorrected-p distances on the ND2 data set using a partial deletion option within and between selected clades (or taxa) of *T. ignobilis* and allies recovered in phylogenetic analyses in MEGA6 (Tamura et al., 2013).

2.5. Coalescent-based species delimitation

Following CEA (2016), we implemented a coalescent species delimitation method using the software BP&P v. 3.1, which uses a Bayesian modeling approach to generate speciation probabilities of closely related taxa from multilocus sequence data (Yang and Rannala, 2010). We conducted a species delimitation analysis (A10 model) for the concatenated data set implementing algorithms 0 and 1 and different fine-tuning parameters (ε = algorithm 0, and x and m for algorithm 1) following the software’s recommendations. We ran the analyses for 100,000 generations, sampling every five generations, and discarding the first 20,000 generations as burn-in. Because BP&P does not allow specifying an unresolved topology as the tree guide, we performed analyses using different trees based on our BI concatenated hypothesis and the multilocus tree of CEA (2016). We followed previous studies (Smith et al., 2013; CEA, 2016) by implementing three combinations of priors that represent different population sizes and ages for taxa/lineages as follows: (1) large Ne and deep divergence: θ [theta] and τ [tau] gamma priors G(1,10) and G(1,10); (2) small Ne and shallow divergence: θ and τ gamma priors G(2,2000) and G(2,2000); and (3) large Ne and shallow divergence: θ and τ
gamma priors G(1,10) and G(2,2000). Each analysis was performed twice from different starting trees to confirm appropriate mixing of MCMC chains. Speciation probabilities were estimated from the sum of the probabilities of all models for speciation events at each node of the tree guide. Daughter lineages (terminal taxa) from nodes that with speciation probabilities >0.95 under all three prior scenarios have been treated as genetically divergent species in previous studies (Smith et al., 2013; CEA, 2016).

3. Results

3.1. The geographic and elevational distribution of Turdus ignobilis

The maps we obtained for voucher specimens and observation records of T. ignobilis (Fig. 1) revealed that the distribution of this species is more extensive than described by CEA (2016). In particular, T. ignobilis is considerably more widely distributed in the Colombian foothills and lowlands East of the Andes because the species occurs widely in the Orinoco and Amazon basins. Although CEA (2016) showed T. i. debilis to occur in parts of the southern Colombian Amazon, they did not note that this taxon extends broadly to the east and north in Amazonia or that it ranges widely into the Orinoco and Maracaibo basins and the Andean foothills. In addition, CEA (2016) did not note that T. i. arthuri occurs in Colombia, where it has been recorded in the Orinoco Basin close to the Venezuelan border and in several localities in the Colombian Amazon; in the latter region it occurs in close geographic proximity with T. i. debilis in departments Guainía (arthuri: specimen catalogue no. ICN 3315; debilis: IAvH-A 2740, 2742), Vaupés (arthuri: ICN 3799, 3825, 3891; debilis: IAvH-A 1224, 1225), and Cauquetá (arthuri: ICN 31972, 31982, 32882; debilis: ICN 2511, 2512, 2541, 2542, 2543; Fig. 1).

We found that populations of T. ignobilis from the Andean region have a broad elevational distribution, with 90% of the records occurring from 250 to 2370 m (Fig. 2). The median elevation of records was 1390 m, indicating that these populations do concentrate in montane areas as indicated by CEA (2016). However, more than one quarter (27%) of the records we analyzed were from elevations lower than 1000 m. This indicates that these taxa are not restricted to highlands and that they frequently occur in foothill to lowland areas. In addition, the elevational (but not geographic) distributions of these populations overlap considerably with that of T. i. debilis (median = 365 m, range of 90% of records 190–990 m; Fig. 2).

3.2. Phylogenetic relationships and genetic distances

ML and BI phylogenetic reconstructions based on ND2 sequences recovered similar topologies (Fig. 3). Both analyses found a well-supported clade (posterior probability = 1.0; ML bootstrap = 97%) that grouped all subspecies of T. ignobilis with four other species of Turdus (T. lawrencii, T. eremita, T. amaurochalimus, T. maranonicus). For the most part, subspecies of T. ignobilis formed well-supported clades. The only exception was T. i. debilis; this taxon seemed to form two separate clades not clearly sister to each other, one of which was not strongly supported.

Consistent with the results of CEA (2016), we found that T. ignobilis is not monophyletic. Although T. i. murinus is a deeply divergent lineage of uncertain affinities (it was recovered as sister to T. lawrencii but with low support), the remaining subspecies of T. ignobilis from the Andes and the Amazon and Orinoco basins form a well-supported clade (1.0/99%) with the Marañón endemic T. maranonicus. This ignobilis/maranonicus clade grouped with T. amaurochalimus and T. eremita also with strong support (1.0/100%). The previously unsampled subspecies T. i. goodfellowi and nominate T. i. ignobilis formed a polytomy with T. i. arthuri, T. i. debilis, and T. maranonicus.

The ND2 haplotype network confirmed the shallow genetic differentiation among populations within the T. ignobilis/T. maranonicus clade, where most taxa differed by one to four mutational steps (Fig. 4). The network also illustrates that one haplotype of T. i. debilis was shared among samples from the Orinoco and Amazon basins. In contrast with results of the phylogenetic analyses, the two samples of T. maranonicus did not group together, with one grouping with haplotypes of T. i. ignobilis and the other with T. i. debilis (Fig. 4).

Genetic divergence (uncorrected-p distance) ranged from 0.4% to 8.6% between taxa and from 0.10% to 0.50% within taxa and clades recovered by our ND2 phylogeny (Table 1). Within T. ignobilis, T. i. murinus exhibited the highest genetic divergence with respect to other T. ignobilis taxa (5.8–6.0%), genetic divergence among the other four subspecies of T. ignobilis and T. maranonicus was much lower (0.4–1.0%).

Five of out six G3PDH sequences of Andean subspecies T. i. ignobilis and T. i. goodfellowi had an insertion of 10 bp, which was absent in one T. i. goodfellowi individual (sample # 5) and in the rest of T. ignobilis and Turdus species analyzed. However, ML and BI analyses of the G3PDH data set recovered similarly unresolved trees (not shown). The low phylogenetic signal of this marker was evidenced by the existence of only four G3PDH haplotypes (Fig. 4), distanced by 1–2 mutational steps. The most common haplotype was shared among T. i. debilis (from the Amazon), T. i. arthuri and T. maranonicus, whereas the second most common haplotype was shared among T. i. ignobilis, T. i. goodfellowi, T. i. debilis (Orinoco) and T. i. arthuri.

The phylogeny built using the concatenated data was similar in topology and support values to the ND2 tree (Fig. 5). Relative to the ND2 tree, support for some clades such as that formed by T. i. ignobilis, T. i. arthuri and T. maranonicus was greater in the concatenated analysis, whereas some of the phylogeographic structure evidenced in ND2 for T. i. debilis (e.g. the existence of a clade including samples from the Amazon and Orinoco basins) was absent in the
concatenated tree. At deeper nodes, the main difference relative to the ND2 tree was the inclusion of *T. subalaris* within the clade conformed by *T. lawrencii*, *T. ignobilis* and other *Turdus* species (1.0/96%).

3.3. Coalescent-based species delimitation

We obtained similar results for algorithms 0 and 1 and among runs with different starting trees, which suggests that MCMC chains adequately converged; we report results obtained using algorithm 1 with fine-tune parameters $e = 2$, $z = 2$, and $m = 1$. We recovered high speciation probabilities (PP = 1.0; 0.99 for *T. i. debilis*; Fig. 6) in eight out of nine internal nodes across the three different priors that represent distinct populations sizes and divergence times among clades of the guide tree. The node joining *T. i. ignobilis* and *T. i. goodfellowi* had a high speciation probability (PP = 0.98) when using priors indicating small population size and shallow divergence, but probabilities were much lower when using the other two prior choices (PP = 0.77, 0.82). In sum, coalescent analyses of mitochondrial and nuclear DNA data reveal the existence of several lineages which may merit species status within *T. ignobilis* as currently circumscribed: *T. i. murinus*, *T. i. arthuri*, *T. i. debilis*, and a lineage formed by the Andean subspecies *T. i. ignobilis* and *T. i. goodfellowi*. This approach is also consistent with the recognition of *T. maranonicus* and *T. amaurochalinus* as distinct species.

4. Discussion

4.1. The geographic and elevational distribution of *T. ignobilis*

The data on the geographic origin of specimens and other records we compiled have partly clarified the geographic and elevational range of members of the *T. ignobilis* complex. Our maps show that *T. ignobilis* ranges widely in foothill and lowland areas east of the Andes in both forests and savannas in the Maracaibo, Orinoco and Amazon basins of Colombia, where two different taxa (*T. i. debilis* and *T. i. arthuri*) occur. This contrasts with the map shown by CEA (2016), which implied that *T. i. debilis* only occurs in the southern part of the Colombian Amazon in forests along the Amazonas (Solimões), Putumayo and upper Caquetá rivers, and that *T. i. arthuri* does not occur in the country except perhaps in savannas near the Venezuelan border in Vichada Department. The map based on expert knowledge (BirdLife International and NatureServe, 2015) shows a more extensive range for *T. ignobilis* in Colombia east of the Andes relative to CEA (2016), but also does not illustrate the occurrence of the species in the Andean foothills and through much of the Orinoco and Amazon basins despite the
existence of multiple records, several of which have been mentioned in the literature (Acevedo-Charry et al., 2014; Stiles and Beckers, 2016; Stiles et al., 1995; Villareal-Leal et al., 2009). In terms of elevational distribution, the records we compiled indicate that subspecies T. i. ignobilis and T. i. goodfellowi are not exclusively montane and that T. i. debilis is not exclusively from the lowlands as existing range maps may suggest (CEA, 2016); the former two taxa regularly occur at mid to low elevations in the Cauca and Magdalena valleys, whereas the latter ranges into mid elevations along the eastern slope of the Andes.

Because the maps we produced are not free of problems, they should not be considered definitive illustrations of the distribution ranges of taxa in the T. ignobilis complex; rather, these maps should be treated more as an example of how different data sources may differ in their depictions of the geographic ranges of Neotropical birds. The information in databases suffers both from erroneous identifications or mistakes in georeferencing (e.g. note the existence of clearly out-of-range observations in Minas Gerais, Brazil; Fig. 1) and from paucity of fieldwork in many areas (e.g. we suspect that apparent gaps in the distribution range of this taxon in the Amazon are not true absences but simply reflect lack of information). In addition, existing data do not allow one to accurately delimit the distribution of subspecies in many areas; this is not possible at all based on data from sound recordings or field observations, and even specimen data may be problematic if subspecific identification has not been carefully examined. Despite these shortcomings, our mapping exercise serves to illustrate the importance of considering various sources of information when characterizing the geographic distribution of Neotropical organisms and designing biogeographic studies.

4.2. Phylogeography and diversification of T. ignobilis

By considering two previously unsampled subspecies of T. ignobilis from the Colombian Andes (T. i. ignobilis and T. i. goodfellowi), and by extending geographic sampling to include Colombian populations from the Orinoco and Amazon basins (T. i. arthuri and T. i. debilis), we provide a new perspective on the biogeography and diversification of this group of Neotropical thrushes. Shallow genetic distances indicate that the ancestor of the clade containing Amazonian/Orinocan (T. i. arthuri, T. i. debilis) and Andean taxa (T. i. ignobilis, T. i. goodfellowi, T. maranonicus) probably diversified within the past million years, a period in which members of the group expanded their range across the Eastern Cordillera to the west and occupied a variety of habitats including savannas, different kinds of lowland forests, and montane environments. Within this clade, the monophyly of the widely distributed T. i. arthuri is presumably linked to the occurrence of this taxon in patches of savanna-type vegetation and white sands; these narrow habitat preferences likely result in relatively small effective population sizes and limit dispersal as observed in other savanna white-sand forest birds (Capurucho et al., 2013; Matos et al., 2016). In contrast, the more generalist and widely distributed T. i. debilis does not consist of a single lineage; most Amazonian samples of this taxon formed a well-supported ND2 clade, but a few were more similar to samples from the Orinoco and Maracaibo basins. An Amazon-Orinoco relationship has been previously documented in fishes (Willis et al., 2010), but not in other vertebrates likely due to the low representativeness of Orinoco populations in phylogeographic studies of tetrapods. Our limited sampling (in terms of loci and individuals) of T. i. ignobilis and T. i. goodfellowi, the trans-Andean populations in the complex, did not allow us to clearly establish their affiliates (i.e. it is unclear whether they are sister groups) and is insufficient to determine whether they are reciprocally monophyletic. Additional sampling in the Cauca and Magdalena valleys and the west slope of the Eastern Cordillera, in addition to considering other markers, is necessary to understand genetic differentiation among populations in this region.

Our increased taxonomic and geographic sampling allowed us to corroborate some of the results of CEA (2016), including the polyphyly of T. ignobilis with respect to other species of Turdus (T. subalaris, T. lawrencii, T. eremita, T. amaurochalinus, and T. maranonicus) and the evolutionary distinctiveness of the Guianan endemic T. i. murinus. However, our analyses suggested that Andean taxa (T. i. ignobilis and T. i. goodfellowi) are not close relatives

Table 1
Pairwise genetic distances (uncorrected p) based on ND2 gene sequences, between and within different lineages of Turdus ignobilis and closely related species of Turdus. Numbers in parentheses correspond to the number of individuals compared per taxon. Bold numbers refer to within taxon mean genetic distance.

<table>
<thead>
<tr>
<th>Taxon 1</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. subalaris (2)</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. lawrencii (3)</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. eremita (1)</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. amaurochalinus (3)</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. maranonicus (2)</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. i. murinus (19)</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. i. arthuri (30)</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. i. debilis (21)</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. i. goodfellowi (3)</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T. i. ignobilis (3)</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Median-joining haplotype networks for ND2 and G3PDH showing relationships among haplotypes in the clade including populations of T. ignobilis (except for murinus) and T. maranonicus, indicating shallow divergence among taxa. Note also that a relatively common mtDNA haplotype of T. i. debilis is shared between the Amazon and Orinoco basins. Nuclear DNA variation consists of a few common haplotypes shared across taxa.
Fig. 5. Bayesian phylogeny of Turdus ignobilis and near relatives based on ND2 gene and G3PDH intron sequences. Colored branches represent different subspecies of T. ignobilis, whereas black branches correspond to other Turdus taxa. Tip labels correspond with the geographic location and ID of specimens included in analyses (see map of Fig. 3 and Supplementary material). Numbers on nodes are Bayesian posterior probabilities and maximum-likelihood bootstrap values, respectively. Note that the pattern of phylogenetic relationships and support values are similar to those found in the ND2 phylogeny (Fig. 3). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
of *T. i. murinus* as assumed by CEA (2016). Instead, our ND2 genealogy and the phylogeny using concatenated data indicated that Andean taxa form a well-supported clade with taxa from the Amazon and Orinoco regions (*T. i. debilis, T. i. arthuri*), and from the Marañón inter-Andean valley of Peru and adjacent Ecuador (*T. maranonicus*). Although CEA (2016) referred to similarities in plumage to infer a close relationship between *T. i. murinus* and Andean taxa, phenotypic characters used to delimit subspecies in this group do not vary in parallel across taxa, and the distinctions mentioned by these authors do not appear clear cut (see Table 3 and Fig. S2 in CEA, 2016). For example, brown streaking on the throat is not uniquely shared by Andean taxa and *T. i. murinus*; because it is also present in *T. i. debilis* despite the seemingly more extensive white in the background in the latter. In addition, *T. i. ignobilis* and *T. i. debilis* are more similar to each other in the brown coloration of the upperparts than either of them is to both *T. i. goodfellowi* and *T. i. murinus*. We suggest that quantitative analyses involving larger samples of specimens are necessary to properly establish patterns of plumage variation in the complex. Moreover, studies on Neotropical *Turdus* have shown that morphologically similar populations may well represent genetically and vocally divergent taxa (Montaño-Rendón et al., 2015; Núñez-Zapata et al., 2016; O’Neill et al., 2011; Ridgely and Tudor, 1989; Voelker et al., 2007), suggesting that plumage variation alone is insufficient to properly infer relationships and delimit species in the group. However, CEA (2016) did not rely only on plumage variation to support their inference of close relationship between *T. i. murinus* and Andean populations of *T. ignobilis*; these authors also argued for ecological similarity among these populations due to their presumably restricted distributions to “highland environments, usually above 1000 m”. However, the data we compiled indicate that the elevational ranges of *T. i. ignobilis* and *T. i. goodfellowi*, despite occurring in the Andean region, frequently reach warm environments at low elevations (i.e., below 1000 m); the elevational range of these populations thus overlap considerably with lowland taxa such as *T. i. debilis*. Finally, CEA (2016) indicated that biogeographic studies have found a sister relationship between populations of the Pantepui and the Andes of Colombia and the Cordillera de la Costa in Venezuela (Bonaccorso and Guayasamin, 2013; Mauck and Burns, 2009; Pérez-Emán, 2005). However, given the limited number of existing studies, we caution against generalizing about close relationships between Andean populations and Pantepui elements because some of the latter likely derived from the Amazonian lowlands (Chapman, 1931; Mayr and Phelps, 1967; Pavan et al., 2016) or are relicts of formerly more extensive ranges (Brumfield and Edwards, 2007). Our analyses reveal that the Pantepui taxon *T. i. murinus* has no close living relatives in the Andes nor in the surrounding lowlands and thus little can be said about its potential geographic origin.

4.3. Reevaluating species limits in the *T. ignobilis* complex

CEA (2016) proposed to split *T. ignobilis* in three distinct species: *T. ignobilis* (including subspecies *ignobilis, goodfellowi* and *murinus*), *T. arthuri*, and *T. debilis*. Our analyses allowed us to confirm that *murinus* is a distinctive evolutionary lineage, which is distinctly related to other taxa traditionally considered conspecific with it. Unfortunately, available information on the songs and calls of *murinus* is too scarce to determine whether it may also differ vocally from other taxa in the *T. ignobilis* complex (we found only two recordings of calls putatively assignable to this taxon in Xeno-canto). Nonetheless, diagnostic plumage characters and its ecological restriction to montane areas in the tepuis of Guyana, Venezuela, and Brazil are consistent with its treatment as a distinct species relative to *T. i. debilis* and *T. i. arthuri*. However, our work demonstrates that *T. i. murinus* is not closely related to nominate *T. i. ignobilis* nor to *T. i. goodfellowi* as assumed by CEA (2016). Therefore, although we agree that *T. i. murinus* is not conspecific with other populations in the complex, our data suggest the taxon *T. ignobilis* as circumscribed by CEA (2016) including *murinus*, nominate *ignobilis* and *goodfellowi* is an artificial group, which should not be recognized.

Given the above, to which of the species as defined by CEA (2016), if any, should nominate *ignobilis* and *goodfellowi* be assigned? More generally, is there sufficient evidence to accord species rank to other taxa in this complex beyond *murinus*? Phylogenetic resolution among members of the clade including *T. i. arthuri, T. i. debilis, T. i. goodfellowi, T. i. ignobilis* and *T. maranonicus* is lacking in our analyses, and genetic divergence among these taxa is shallow. Shallow divergence by itself does not imply these populations are conspecific, considering the unquestioned species status of the phenotypically distinct *T. maranonicus*, which is only 0.4–1.0% divergent in ND2 sequences from other members of this clade (Table 1). Moreover, CEA (2016) presented evidence from coalescent analyses of sequences of three unlike loci (mtDNA and two nuclear introns) consistent with the hypothesis that *arthuri, debilis* and *maranonicus* are distinct species, and showing that *debilis* is more closely allied to *maranonicus* than to *arthuri*. These results were corroborated by our BPP analyses, which recovered high speciation probabilities for most taxa of *T. ignobilis* and their allies, with the exception of the Andean forms *T. i. ignobilis* and *T. i. goodfellowi* for which there is no strong evidence to reject the hypothesis that they are conspecific. Although the Bayesian modeling method employed by CEA (2016) and in our study has been said to identify reproductively isolated populations and hence to delimit species under the biological species concept, it in fact delimits groups that have experienced no recent gene flow and requires no other evidence of reproductive isolation, and is thus most useful for identifying cryptic species in sympathy (Yang and Rannala, 2010). Because in cases involving allopatric populations lack of recent gene flow may reflect true reproductive
isolates (i.e. the existence of biological barriers to interbreeding) but also merely the effects of geographic isolation, we believe the approach is more consistent with delineating species under genealogical or evolutionary species concepts, and should be coupled with other kinds of data including morphology, behavior, and geography/ecology to gauge evidence in favor of reproductive isolation under the biological species concept (Yang and Rannala, 2010). In fact, a recent paper has demonstrated that the multi-species coalescent as implemented in BPP& detects population structure associated with population isolation and not species boundaries, implying that units delimited under this approach should be considered only tentative hypotheses of species, to be confirmed or rejected through subsequent analysis or with other types of data (Sukumaran and Knowles, 2017).

We think that *T. i. arthuri* is a serious candidate to merit species rank under the biological species concept because in addition to its diagnostic plumage and its genetic differentiation relative to other members of *T. ignobilis*, this taxon appears to be ecologically distinct, being restricted to white-sand vegetation and savannas of the Amazon and Orinoco basins (Borges et al., 2016; Collar, 2015). In Colombia, *T. i. arthuri* has been recorded in patches of open savanna in the Orinoce region and in white-sand forest in the Serranía de Chiribiquete in the Amazon Basin, whereas *T. i. debilis* occurs in close geographic proximity in Amazonia (e.g. in Caquetá and Guainía; Fig. 1) likely associated to gallery and terra firme forests (Stiles et al., 1995; Villareal-Leal et al., 2009). Although these patterns are suggestive, we believe that determining whether *T. i. arthuri* and *T. i. debilis* may represent distinct biological species in true parapatry would need to confirm these apparent ecological differences in areas of potential contact, ideally in combination with analyses of variation in phenotype and molecular markers in such areas including a larger sample of individuals than those currently available. Likewise, a reassessment of the status of Andean taxa would require a thorough population-level study integrating genetic analyses with assessments of phenotypic variation in potential contact zones in Colombia (i.e. between *T. i. ignobilis* and *T. i. debilis* along the Eastern Cordillera, and between *T. i. ignobilis* and *T. i. goodfellowi* in the middle Cauca Valley). Given the importance of songs and calls in ascertaining species limits among Neotropical *Turdus* (O’Neill et al., 2011), devoting attention to examining vocal variation is especially critical to assess the status of all the above populations. In sum, we conclude that although CEA (2016) have provided compelling evidence that the *T. ignobilis* complex likely consists of several genealogical/evolutionary species, the existing data are insufficient to propose any additional changes in species-level classification under the biological species concept.

Based on the above, we propose the following taxonomic treatment and English names, recognizing provisionally two distinct biological species:

**Turdus ignobilis** Sclater, P. L. 1858
Black-billed Thrush

**Turdus debilis** (Archer, 1859)
Andes Thrush

**Turdus goodfellowi** (Archer, 1859)
Goodfellow Thrush

**Turdus arthuri** (Sclater, 1863)
Arthuri Thrush

**Turdus murinus** Salvin 1885
Pantepui Thrush

Tepuis of south Venezuela (Mounts Duida, Auyán-Tepui, Urtani, Marahuaca and Roraima), Guyana (Mounts Roraima and Merumé), and Brazil (Frontera Sierra Urtani). Monotypic.

4.4. Conclusions

Our reassessment of the diversification and species limits in the *T. ignobilis* complex demonstrates that phylogeographic studies that do not sample sufficiently across geography and taxa are prone to obtaining results that are at best incomplete and potentially spurious. The importance of judicious sampling to address questions in phylogenetics (Zwick and Hillis, 2002), phylogeography (Buckley, 2009), and species delimitation (Ahrens et al., 2016) using molecular data has been thoroughly stressed in the literature, so we will not reiterate general concerns related to taxonomic and geographic coverage. However, we argue that given what we have learned about the complex and idiosyncratic evolutionary histories of Neotropical organisms (Turchetto-Zolet et al., 2013), studies in this region are especially prone to pitfalls related to insufficient sampling given the scope of the questions asked. It follows that inferences about the affinities of species and populations based on phenotypic variation, ecology, or biogeography should not be considered substitutes to directly examining genetic variation in phylogeographic analyses of Neotropical organisms.

Historically, access to biological specimens from Colombia has been limited due to reasons related to political instability and difficulty in obtaining permits, which has resulted in Colombian biodiversity being underrepresented in various kinds of studies (Arbeláez-Cortés, 2013). However, times have changed over the past few years, and molecular work to address questions in systematics, taxonomy and biogeography involving Colombian populations is now clearly possible (Arbeláez-Cortés et al., 2015). A case in point is that of *T. ignobilis*: all the samples reported on in this paper were available at the time that the study by CEA (2016) was submitted for publication. Thus, although CEA (2016) mentioned that they unsuccessfully tried to obtain sequence data from Colombian specimens, their goal of having sufficient for their study could have been accomplished within a reasonable time frame by consulting with researchers based in Colombia either to directly collaborate with the research or to have them facilitate legal export of samples, which is now possible. To conclude, then, we can only reiterate a point made in this journal a decade ago: “Due to its geographic position at the crossroads between Central and South America and the expected effects of its complex geography on population structure, it comes as no surprise that the results of this study imply that analyses of patterns of differentiation involving detailed sampling in Colombia should be considered essential to understanding the biogeographic history of many Neotropical taxa” (Cadena et al., 2007). More generally, because many of the most interesting questions in tropical evolutionary biology involve taxa with wide distributions that do not recognize political boundaries, we believe that the best way to move the field forward is to always strive for the most complete geographic and taxonomic sampling possible given research questions, reaching out to colleagues in different countries whenever necessary. Recent work indicates that even large-scale, multi-taxon and multi-country projects can be satisfactorily
completed (e.g., Smith et al., 2014), but researchers, administrators, and governments should work together to foster and facilitate efficient international collaboration for biodiversity research.

Acknowledgments

We thank the collectors and institutions that provided tissue samples used in this study, especially Socorro Sierra, Fernando Forero, Mauricio Álvarez-R. and Claudia Medina (Instituto Alexander von Humboldt), Andrea Morales (Museo de Historia Natural Unillanos), Andrés M. Cuervo (Universidad de los Andes), and F. Gary Stiles, Juan Pablo López, Orlando Acevedo and Diego Caranton (Instituto de Ciencias Naturales). Nick Athanas provided the photograph for the graphical abstract. This study was financially supported by the Universidad de los Andes – Colombia and the Colección de Tejidos of the Instituto Alexander von Humboldt. We thank members of the Laboratorio de Biología Evolutiva de Vertebrados, A. Aleixo, and J.V. Remsen, Jr. for discussions and comments on the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.03.020.

References
