Restoring the species status of *Catharus maculatus* (Aves: Turdidae), a secretive Andean thrush, with a critique of the yardstick approach to species delimitation

MATTHEW R. HALLEY1,4, JOHN C. KLiCKA2, PAUL R. SESINK CLEE3 & JASON D. WECKSTEIN1

1 Academy of Natural Sciences of Drexel University, Department of Biodiversity, Earth & Environmental Science, Drexel University, Philadelphia, Pennsylvania 19104, USA
2 Burke Museum of Natural History and Culture, Department of Biology, University of Washington, Seattle, Washington 98195, USA
3 Department of Biology, Drexel University, Philadelphia, Pennsylvania 19104, USA
4 Corresponding Author. E-mail: matthewhalley@gmail.com

Abstract

In the 1850s, two species of "Spotted" Nightingale-Thrush (*Aves: Catharus*) were independently described from montane rainforests of Guatemala, *C. dryas* (Gould, 1855) and Ecuador, *C. maculatus* (Sclater, 1858). However, due to similarities in plumage color, *C. maculatus* was reclassified as a subspecies of *C. dryas* in 1878, a decision that has been upheld for 137 years. We collected multiple lines of evidence including phylogenetic analysis of mitochondrial DNA sequences (ND2), discriminant and principal components analysis of morphometric and vocal data, and statistical modeling of ecological niches, that collectively indicate that *C. d. dryas* and *C. d. maculatus* are independent species. We recommend restoring species status to *C. maculatus* of South America and applying the common name Sclater's Nightingale-Thrush to this species.

Key words: *Catharus dryas*, Spotted Nightingale-Thrush, Neotropics, avian systematics

Introduction

During the 19th century, waves of European explorers shipped collections of bird skins from the Neotropics to their home countries, where they were identified by taxonomists who had never seen the birds in life, and knew nothing of their habits or habitats (Vuilleumier 2003). Many new taxa were described during this period because the type specimen(s) differed in body dimension and/or plumage color from previously known taxa (e.g., Ridgway 1872, 1886; Stone 1912), but often no quantitative method of diagnosis was used to determine species limits. By the early 20th century, myriads of closely related (and similarly sized) species and subspecies had been described on the sole basis of (often subtle) differences in plumage color, in a scientific literature that was increasingly difficult to access because papers were published in multiple outlets, countries, and languages (Vuilleumier 2003).

The convoluted taxonomy of the songbird genus *Catharus* (Aves: Turdidae) developed in this fashion, and most species were delimited on the basis of plumage color alone. Even today, many life history traits and behaviors of *Catharus* species remain poorly known, owing to their shy disposition and penchant for dense vegetation (e.g., Greeney et al. 2015; Heckscher et al. 2015; Halley et al. 2016). Twelve species are currently recognized, of which eleven (92%) inhabit Neotropical forests for all or most of the year (Clement 2000; Clements 2007). Four species...
are obligate long-distance migrants, spending about four months of each year in North America during the Nearctic summer, where they breed and molt, and the remainder of the year in the Neotropics (Clement 2000). Seven species are permanent residents of the Neotropics and are known individually and collectively as "nightingale-thrushes". Finally, a single species—Hermit Thrush *Catharus guttatus* (Pallas, 1811)—is a short-distance migrant with some populations that remain in the temperate zone all year.

Resident and migrant *Catharus* were long thought to comprise two reciprocally monophyletic lineages (e.g., Winker 2000), and they were classified in different genera until the mid-20th century, but phylogenetic analysis of DNA sequences has now shown that they do not form monophyletic groups (Outlaw et al. 2003; Winker & Pruett 2006; Voelker et al. 2013). This means that either distant lineages have converged on nearly identical plumage phenotypes, or the spotted ventral plumage of the migratory species is ancestral and has been lost multiple times in the non-migratory lineages. Either way, plumage color, which was the primary (and sometimes only) justification underlying *Catharus* taxonomy, is actually a poor indicator of evolutionary relationships among *Catharus* species. This pattern has been demonstrated for many other birds (e.g., Armenta et al. 2005).

Parallel or convergent evolution of plumage coloration is also evident in both resident and migratory species pairs. For example, resident species *C. occidentalis* Sclater, 1859 and *C. frantzii* Cabanis, 1861 were once considered a single species because of their similar coloration (Hellmayr 1934; but see Ridgway 1907), but in fact they belong to relatively distant genetic lineages, and differ in voice, behavior, and egg color (Rowley & Orr 1964; Phillips 1969; Raitt & Hardy 1970; Outlaw et al. 2003). Similarly, the brown, spotted phenotypes of *C. guttatus* and *C. fuscescens* (Stephens, 1817), distantly related migratory taxa that were considered conspecific until the early 19th century, are either the product of convergent evolution, or the retention of an ancestral phenotype. Each species is more closely related to an unspotted Neotropical congener than to the other. In another example, sister taxa *C. minimus* (LaFresnaye, 1848) and *C. bicknelli* (Ridgway, 1882) cannot be reliably distinguished by sight alone, but exhibit diagnostic differences in song, and occupy allopatric breeding and wintering grounds (Bent 1949; Ouellet 1993; Marshall 2000). Because plumage color is evolutionarily labile in *Catharus*, and apparently not critical to the maintenance of reproductive isolation among species (Dilger 1956), any taxonomic approach requiring that species-level taxa be visually diagnosable in the field (i.e., with no vocal or other cue) risks systematically underestimating the true biodiversity of the genus and confounding our understanding of its evolution. Instead, the logical step is to integrate multiple lines of evidence (e.g., genetics, voice, morphometrics, ecology) to test the veracity of the plumage-based taxonomy, and in so doing, expose any cryptic lineages that have been obscured by homoplasy (e.g., Armenta et al. 2005; Cadena & Cuervo 2010; O'Neill et al. 2011).

We investigated genetic, phenotypic, and ecological variation in the Spotted Nightingale-Thrush complex, *C. dryas* (Gould, 1855), which consists of two widely disjunct populations (>1000 km apart) that inhabit montane rainforests, and as such are separated by numerous lowland barriers to gene flow (Fig. 1). One population is found in mountainous regions of southern Mexico, Guatemala, El Salvador, Honduras, and northern Nicaragua. In South America, the other population occurs in the northern Andes of Colombia and Venezuela, along the eastern spine of the Andes to Ecuador, and in the mountains of Peru, Bolivia, and northern Argentina (Clement 2000). There are no representatives of the *C. dryas* complex in the intervening countries of Panama and Costa Rica.

**Taxonomic History.** Gould (1855) erected a new genus with type species *Malacocichla dryas* (= *C. dryas*), from a single Guatemalan specimen, to which Sclater (1858) added a second species from Ecuador, *M. maculata*, which he argued could be distinguished from *M. dryas* by its dark slaty dorsum (versus green-olive in *M. dryas*). Gould (1855) neglected to specify a holotype for *M. dryas*, but there is a type in the Natural History Museum in Tring (NHMUK) that was, as per the label, "collected by Skinner and purchased by J. Gould" (NHMUK 1854.12.20.17; type locality: Vera Paz, Guatemala). This is the single skin listed as the "Type of *Malacocichla dryas*, Gould" by Seebohm (1881). A type of *M. maculata* (= *C. maculata*) is also present (NHMUK 1858.2.4.33; type locality: "Rio Napo, Quijos Province, east Ecuador"), and was apparently one of a series when it was acquired by Maison Verreaux, the family business of the French ornithologist Jules Verreaux (1807–1873). One specimen was sold as a "type" to the United States National Museum (USNM 32683, type locality: Ecuador, Province of Napo-Pastaza, "Rio Napo"), but whether this specimen was actually a part of Sclater's type series is unknown (Deignan 1961). According to Sclater (1858:64), Verreaux suggested the name "*Catharus maculatus*" when the specimen was transferred to the NHMUK, thereby recognizing the new taxon's affinity to Bonaparte's (1850) genus *Catharus*. Sclater (1858) honored the epithet *maculatus*, but placed the species (not without hesitation) not in *Catharus* but in a new genus (*Malacocichla*) because of its spotted ventral surface. However, the following year, Sclater (1859) changed his mind, moving both species to *Catharus*, and in so doing, corrected the gender of the epithet *maculata* to *maculatus*.
Baird (1872) found a Guatemalan specimen of *C. dryas* to be "easily distinguished from *C. maculatus* of Ecuador, by the much more olive back (not dark ash) and ochry yellow breast." However, Salvin (1866) was not convinced, and Salvin & Godman (1879) advocated subsuming *C. maculatus* into *C. dryas* after correctly noting that the yellow underparts suffered from post-mortem fading ("one of the chief differences relied upon fails to differentiate the species"). However, Ridgway (1907) expressed uncertainty about the merger, noting that the South American taxon is "so conspicuously different that it can hardly be the same form." Nevertheless, the taxonomic treatment of Salvin & Godman (1879)—one species, two disjunct populations—was adopted by later authorities and has now stood for 137 years, although no additional populations were discovered in the intervening regions of Central America as they had anticipated (Clement 2000). We combined four independent lines of evidence to test Salvin & Godman’s (1879) null hypothesis that the differences between *C. d. dryas* (Mesoamerica) and *C. d. maculatus* (South America) are not of a magnitude sufficient for the treatment of both as species.

**Methods**

**Criteria for Species Delimitation.** Following the Phylogenetic Species Concept (PSC; Cracraft 1983), we required reciprocal monophyly with strong support (>70% in bootstrap analysis) for recognition of species-level taxa in our phylogenetic analysis. Such a result would indicate that the taxa in question represent the "smallest diagnosable cluster[s] of individual organisms within which there is a parental pattern of ancestry and descent" (Cracraft 1983). We also required that phenotypic divergences between species be diagnosable (>95% in discriminant analysis) and greater in magnitude than differences within each of those species (i.e., species are more divergent than subspecies; Huxley 1942). In the case of *C. d. dryas* and *C. d. maculatus*, application of the Biological Species Concept (BSC; Dobzhansky 1937; Mayr 1942) is inhibited by allopatry. Some authors (e.g., Helbig et al. 2002; Tobias et al. 2010) have recommended that divergences between allopatric taxa should be compared to sympatric undisputed congeners, *i.e.*, a threshold beyond which species rank can be applied (hereafter
"yardstick approach"). Helbig et al. (2002), referring to both phenotypic and genetic divergence, advised: "The likelihood that allopatric taxa will remain distinct can only be judged by the degree of their divergence, preferably in comparison with taxa that are closely related to the group under investigation and that are known to coexist in sympathy."

We do not think that the yardstick approach is appropriate or useful for species delimitation, for the following reasons: (1) genetic and/or phenotypic divergence alone does not make a species; (2) there is no standard threshold for divergence for species level taxa; (3) not all allopatric taxa have appropriate yardsticks and the process of choosing one is biased by sampling error that stems from the vagaries of extinction and incomplete sampling; (4) the timing of speciation following geographic isolation is not expected to be identical, even within the same genus, and so yardsticks cannot be properly calibrated; (5) the yardstick approach is founded on the uncertain premise that currently recognized sister taxa (under the BSC) are in fact correctly identified species that are indeed each other's closest extant relatives; and (6) the expectation (or burden of proof required) is unreasonably high—why should allopatric species necessarily be as, or more different than, sympatric species?; (7) the stipulated standard yardstick value is likely to vary dramatically from clade to clade, and study to study; and (8) if divergences between allopatric taxa are less than the yardstick divergence, nothing definitive can be concluded about allopatric taxa that would otherwise be "good" species under any species concept.

To illustrate the subjectivity of this approach, we have chosen two congeners—C. fuscescens and C. minimus—for standard comparisons of intra- and interspecific genetic distances in our phylogenetic analysis. We do not apply the yardstick in our other analyses. Of the 12 species currently recognized in Catharus, they are the most closely related pair that are field diagnosable by sight alone, and they maintain their phenotypic distinctiveness in southwestern Newfoundland where they breed in sympathy (Marshall 2000; MRH, pers. obs.). If we had used C. bicknelli and C. minimus as our yardstick species, sister taxa that are not field diagnosable by sight alone, the minimum genetic distance required for assigning species rank would have been reduced; already the subjectivity of the yardstick approach is apparent. Nevertheless, in a genus renowned for taxonomic uncertainty, there has never been disagreement among ornithologists that C. fuscescens and C. minimus are good species, whereas C. bicknelli was long considered to be a subspecies of C. minimus, and was only recently elevated to species rank (Ouellet 1993), and not without dissenters (e.g., Marshall 2000). Therefore, for those philosophically tied to a yardstick approach, C. fuscescens and C. minimus would seem to provide a reasonable comparison, although they are not sister species (Voelker et al. 2013), adding yet more subjectivity.

**Phylogenetic Analysis.** We used DNeasy tissue extraction kits and the manufacturer's protocol (Qiagen, Inc., Hilden, Germany) to extract whole-genomic DNA from tissue samples (N = 17). We amplified and sequenced the mitochondrial NADH dehydrogenase subunit 2 gene (ND2) using methods described by Voelker et al. (2013), aligned the 17 nucleotide sequences with the MUSCLE algorithm in MEGA 5.0 (Edgar 2004; Tamura et al. 2011), and then truncated positions lacking data for ≥ 1 species from the beginning and end of the alignment. All new sequences were deposited in the NCBI GenBank (http://www.ncbi.nlm.nih.gov) and accession numbers are provided below. We included: (1) four sequences from the published range of C. d. maculatus in South America (KY564234–KY564237); (2) four sequences from Guatemala near the type locality of C. d. dryas (KY564238–KY564241); (3) two sequences of C. fuscescens: one each from Nova Scotia (AY049519; C. f. fuscescens) and Minnesota (KY564242; presumably C. f. levii, Phillips, 1991); (4) two sequences from C. minimus: one from Newfoundland (AY049526; C. m. minimus) and a passenger migrant taken in New York (KY564243; C. m. aliciae); (5) one sequence each from C. fuscateter (AY049518; Lafresnaye, 1845). C. occidentalis (KY564245), C. guttatus (KY564244), and C. ustulatus ustulatus (Nuttall, 1840); and (6) as an outgroup, one sequence from the Wood Thrush Hylocichla mustelina (AY049525; Gmelin, 1789), sister taxon to Catharus (Voelker et al. 2013). We were unable to acquire genetic data from C. d. ovandensis Brodkorb, 1938, C. d. harrisoni Phillips & Rook, 1965, C. d. ecuadorensis Carriker, 1935, and C. d. blakei Olrog, 1973.

We imported the resulting 822 bp alignment into PAUP* 4.0a146 (Swofford 2002), and performed an automated test of 56 substitution models. We used Bayesian Information Criterion (BIC) to select the model that best fit the data, a general time reversible model with invariant sites: GTR+I (TrN+I submodel); base frequencies: A=0.31583, C=0.34759, G=0.11106, T=0.22552; proportion of invariant sites: 0.590621). We conducted a heuristic search using both Maximum Parsimony (MP) and Maximum Likelihood (ML) methods to reconstruct phylogenetic trees, and estimated support for each topology via bootstrapping (1000 pseudoreplicates). One testable prediction of Salvin & Godman's (1879) null hypothesis is that C. d. dryas and C. d. maculatus are not
reciprocally monophyletic and divergent from one another. Thus, strongly supported reciprocal monophony (>70% bootstrap support) and a high mean genetic distance between C. d. dryas and C. d. maculatus, relative to genetic distances within each taxon, would imply that the two disjunct populations are actually two independent species. We used Minitab 16 (v.16.2.2, Microsoft Corporation, Inc., Redmond, WA) to test uncorrected pairwise genetic distances (p) for normality (Anderson-Darling test, α = 0.05) and equal variances (Levene's test, α = 0.05). Because distances were not normally distributed, we used Mann-Whitney U-tests (α = 0.05) to determine whether between-taxa and within-taxon pairings were drawn from an identical distribution (i.e., whether their medians differed).

**Morphometric Analysis.** MRH used dial calipers and a metric ruler to measure study skins (N = 57) in the collections of the American Museum of Natural History (AMNH), Academy of Natural Sciences of Drexel University (ANSP), and Delaware Museum of Natural History (DMNH), as follows: *Mesoamerica*: AMNH: 39090, 39091, 396399, 784661, 793479–793482; ANSP: 175658, 175659; DMNH: 53345–53349. *South America*: AMNH: 130156, 130160, 130161, 167895–167897, 167899, 172119–172123, 172125, 172126, 173495, 180630, 181089, 439498, 503900, 820707; ANSP: 43253, 91628, 119622, 143246–143250, 143252, 143254, 143257, 170138, 171214, 177003, 177643, 184102, 184825, 185733, 187143; DMNH: 67730, 67731. The following measurements were taken: (1) *bill shape*, equal to bill width, measured at the anterior edge of the nares, divided by bill length, measured from the anterior edge of the nare to the tip; lower values indicate a relatively long thin bill, and higher values represent a relatively shallow, wide bill; (2) *wing length*, the distance between the carpal joint and the tip of the longest flattened primary; (3) *tarsometatarsus length*; and (4) *tail length*. We tested each variable for normality (Anderson-Darling test, α = 0.05) and equal variances (Levene's test, α = 0.05), and then used two-sample t-tests (α = 0.05) to determine if there were significant differences between the means of the sexes within each region (Mesoamerica vs. South America), and between the two regions within each sex. When data were non-normally distributed, we used a Mann-Whitney U-test (α = 0.05). We performed a principal components analysis (PCA) of the four variables, and a discriminant function analysis (DFA) within each sex class, to determine how well the morphometric data predicted the geographic origin of the specimen (Mesoamerica vs. South America). We also tested for differences between the sexes, to determine whether the taxa exhibit similar patterns of sexual size dimorphism. Statistics for the morphometric analysis were performed with the program R-Studio (R Core Team 2014).

**Song Analysis.** A total of 108 audio recordings were obtained from the Macaulay Library of Natural Sounds (ML; Cornell University, Ithaca, NY) and the xeno-canto archive (XC; http://www.xeno-canto.org), and visualized and quantified with Raven Pro 1.5 software (Cornell University, Ithaca, NY). The vocalizations of this species have never been systematically studied. We restricted our analysis to "songs", which like other *Catharus* species (Stein 1956), consist of repeated phrases of whistled notes with little or no spectral content above the fundamental frequency (*note* defined here as a continuous trace on the sonogram). We sampled one song per individual (the first whistled song that was not masked by any other sound), and removed duplicate individuals from the dataset (birds with the same repertoire recorded on the same day). We also removed recordings in which all songs were fully or partially masked by insect noise, other birds, etc. Therefore, each data point in our analysis was statistically independent, and the influence of extraneous sounds was mitigated. The following recordings (N = 82) met the above criteria and were included in our analysis: *Mesoamerica*: ML: 55452, 86869, 90085, 95060; XC: 185904, 185905, 335815. *South America*: ML: 11071, 13899, 18335, 28503, 28510, 33761, 40228, 41226, 58066, 58872, 66344–66347, 89352, 110634, 116117, 120868, 120872, 120901; XC: 1681, 4034, 4220, 4666, 5457, 6703, 10681, 16695, 22498, 23043, 32474, 32476, 35147, 38763, 44231, 58066, 61427, 61428, 86577, 98322, 106632, 120901, 143257, 170138, 171214, 177003, 177643, 184102, 184825, 185733, 187143; DMNH: 67730, 67731. We used the selection box feature to calculate four variables for each song: (1) *Notes*, the number of discrete traces on the sonogram, (2) *Center Frequency*, which divides the selection into two frequency intervals of equal energy, (3) *Duration 90%,* the interval of time containing 90% of the energy in the selection, (4) *Bandwidth 90%,* the range of frequencies that contain 90% of the energy. These characters are considered 'robust' because they are unaffected by small differences in the size of the selection box. We tested each variable for normality (Anderson-Darling test, α = 0.05) and equal variances (Levene's test, α = 0.05), and then used two-sample t-tests (α = 0.05) to determine whether there were significant differences between the means (Mesoamerica vs. South America), or, when data were not normal, we used Mann-Whitney U-tests (α = 0.05). We then performed PCA and DFA to determine to what extent the songs from the two regions are statistically diagnosable.
Environmental Niche Models (ENMs). Evidence of divergence or conservation of ecological niche characteristics can be useful for species delimitation, especially for sedentary allopatric taxa (reviewed by Wiens & Graham 2005). Although the migratory species in *Catharus* inhabit different climatic niches on their breeding and wintering grounds, populations of *C. dryas* and its sister taxon *C. fuscater* are sedentary and inhabit the same climatic niche year round (Clement 2000). We used VertNet (http://portal.vertnet.org/search) to search the databases of major museum collections for specimen records (including eggs) with latitude and longitude data. The resulting list included specimen localities from Mesoamerica (*N* = 76, including specimens from Mexico, Guatemala, and El Salvador) and South America (*N* = 34, including specimens from Colombia, Ecuador, Peru, and Bolivia). Using ENMTools (Warren et al. 2010), we performed Pearson correlation tests between all original environmental layers (see Supplementary Table 1; Hijmans et al. 2005; Farr et al. 2007; Hansen et al. 2013; ESRI 2014) to identify layers that showed very similar trends (-0.85 > *p* > 0.85). The remaining records were projected and visually inspected in ArcMap 10.3 (ESRI 2014) to ensure that no more than one specimen record occurred in each environmental data grid cell or fell outside of the study region. We used ENMTools (Warren et al. 2010) to perform Pairwise Pearson correlation tests between all original environmental layers (see Supplementary Table 1; Hijmans et al. 2005; Farr et al. 2007; Hansen et al. 2013; ESRI 2014) to identify layers that showed very similar trends (-0.85 > *p* > 0.85). Initial Maxent runs were used to identify which of the highly correlated layers to remove based on percent contribution to the models. Layers that scored less than one percent contribution to initial models were also removed from final model runs. Average ENMs were produced using 100 replicate iterations of the data for both the Mesoamerican population models, and the robustness of each ENM was estimated with the 'area under the curve' (AUC) value. Values above 0.9 are considered "very good" descriptions of the ecological niche of a population (Baldwin 2009). The distinction between ‘suitable’ and ‘not suitable’ habitat was calculated as the minimum estimated suitability at a location of known presence for each population.

Results

Molecular Phylogenetics. *Catharus d. dryas* and *C. d. maculatus* are reciprocally monophyletic with 100% and 98% bootstrap support in MP and ML trees, respectively (Fig. 2). The mean (± SD) uncorrected divergence between *C. d. dryas/C. d. maculatus* pairwise comparisons (*N* = 16, *p* = 0.041 ± 0.002) was an order of magnitude greater than the mean distance within each taxon (within *C. d. dryas*: *N* = 6, *p* = 0.003 ± 0.001; within *C. d. maculatus*: *N* = 6, *p* = 0.003 ± 0.002; Mann-Whitney U-test adjusted for ties, *W* = 312.0, *P* < 0.001). As with *C. d. dryas* and *C. d. maculatus*, intraspecific genetic distances in the yardstick species were smaller than the mean distance between them (within *C. fuscescens*: *N* = 2, *p* = 0.009; within *C. minimus*: *N* = 2, *p* = 0.013).

<p>| TABLE 1. Mean values of four morphometric characters recorded from male and female specimens of <em>C. dryas</em> from Mesoamerica and South America. Bill shape is the quotient of the width of the bill at the anterior edge of the nares, and the length from that edge to the tip. For all variables except bill shape, the unit of measurement is mm. For each character, the difference (in mm) between the means of the Mesoamerican and South American samples (Δ) is followed by the test statistic (<em>T/W</em>), degrees of freedom (df), and <em>P</em>-value (<em>α</em> ≤ 0.05). Two-sample <em>t</em>-tests (TST) were used for all but one character, for which a Mann-Whitney U-test (MWU) was used. For means that differed significantly between the sexes within each population, the mean for males is shown in boldface. |
|---|---|---|---|---|---|---|</p>
<table>
<thead>
<tr>
<th>Sex</th>
<th>Character</th>
<th>Mesoamerica</th>
<th>South America</th>
<th>Δ</th>
<th>Test</th>
<th><em>T/W</em></th>
<th>df</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>bill shape</td>
<td>10</td>
<td>0.362 ± 0.03</td>
<td>23</td>
<td>0.395 ± 0.03</td>
<td>-0.03</td>
<td>TST</td>
<td>-2.98</td>
</tr>
<tr>
<td></td>
<td>tarsus length</td>
<td>9</td>
<td>35.7 ± 1.16</td>
<td>25</td>
<td>34.14 ± 0.98</td>
<td>1.56</td>
<td>TST</td>
<td>3.92</td>
</tr>
<tr>
<td></td>
<td>wing length</td>
<td>10</td>
<td>101.5 ± 3.14</td>
<td>25</td>
<td>90.88 ± 2.95</td>
<td>10.62</td>
<td>MWU</td>
<td>304.5</td>
</tr>
<tr>
<td></td>
<td>tail length</td>
<td>10</td>
<td>74.9 ± 4.58</td>
<td>25</td>
<td>68.88 ± 4.44</td>
<td>6.02</td>
<td>TST</td>
<td>3.59</td>
</tr>
<tr>
<td>F</td>
<td>bill shape</td>
<td>4</td>
<td>0.408 ± 0.01</td>
<td>15</td>
<td>0.412 ± 0.03</td>
<td>-0.004</td>
<td>TST</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>tarsus length</td>
<td>4</td>
<td>34.31 ± 1.39</td>
<td>16</td>
<td>33.02 ± 0.92</td>
<td>1.29</td>
<td>TST</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>wing length</td>
<td>4</td>
<td>93.75 ± 2.75</td>
<td>16</td>
<td>86.38 ± 3.18</td>
<td>7.37</td>
<td>TST</td>
<td>4.24</td>
</tr>
<tr>
<td></td>
<td>tail length</td>
<td>4</td>
<td>65.75 ± 4.86</td>
<td>16</td>
<td>63.38 ± 3.34</td>
<td>2.37</td>
<td>TST</td>
<td>1.17</td>
</tr>
</tbody>
</table>
**Plumage.** Previous authors have noted that the dorsal plumage of South American specimens is dark slaty gray, whereas there is a subtle olive wash to the dorsal plumage of the Mesoamerican specimens (Sclater 1858; Baird 1872; Ridgway 1907; but see Salvin 1866). However, the yellow coloration of the ventral surface is prone to post-mortem fading (Salvin & Godman 1879; Phillips 1991), and the yellow component of the olive hue of the upper dorsal surface of *C. d. dryas* specimens likewise fades, rendering them slaty-gray like *C. d. maculatus* specimens. The "colors are exceptionally fugitive" and "the races are identifiable for a very short time only, after collecting, and must be immediately painted or compared to a very detailed color guide" (Phillips 1991). The collection dates of the skins we studied ranged from 1873–1993, and the birds varied in how fresh the plumage was at the time of collection. For these reasons, the color of the dorsal and ventral plumage of these specimens is now unreliable for taxonomic study. Notwithstanding these difficulties, we noticed that skins from South America, labeled *C. d. maculatus* and *C. d. ecuadoreanus*, had small dark streaks or spots that crossed their throats *sensu stricto*, forming a dark band that was faint in some specimens and bold in others. By contrast, specimens of Mesoamerican forms *C. d. dryas, C. d. ovandensis*, and *C. d. harrisoni* sometimes have a small streak or two in the malar region, but the center of their throats are generally white and unmarked (Fig. 3).
**South America — Streaked or Spotted Throats**

A  

B  

C  

D  

**Central America — Unmarked White Throats**

E  

F  

G  

H  

**FIGURE 3.** Digital photograph series showing the upper breast and throat of eight museum skins in the collection of the American Museum of Natural History (AMNH): (A–B) *Catharus dryas ecuadorensis*, males collected in Ecuador in 1920 and 1921, respectively (167879, 172129); (C–D) *C. d. maculatus*, males collected in Ecuador in 1968 and 1923, respectively (820707, 180630); (E–F) *C. d. dryas*, male and female collected by Salvin in 1843 in Guatemala (39090, 39091); (G) *C. d. harrisoni*, female collected in Oaxaca, Mexico, 1967 (793482); (H) *C. d. ovandensis*, male collected in Chiapas, Mexico, 1965 (784661).

**Morphometrics.** Because reciprocal monophyly was strongly supported by a combination of genetic and plumage characters, we pooled morphometric data from each region and compared all South American (SA) specimens to all Mesoamerican (MA) specimens within each sex. Statistical analysis of morphometric differences among the various subspecies within each region was precluded by small sample sizes of some taxa. The first two principal components accounted for 82% of the total variation among males, and 76% among females (Fig. 4), and there were significant differences between the putative species in both sexes (Table 1; MANOVA: males: Wilks’ $\lambda = 0.131$, $P < 0.001$; females: Wilks’ $\lambda = 0.441$, $P = 0.016$). Discriminant analysis with cross-validation correctly classified 100% of males ($N = 32$, MA:SA, 9:23), 86.7% of South American females ($N = 19$), and 75% of Mesoamerican females ($N = 4$), to the correct region based on the four morphometric characters. Mesoamerican males were significantly larger than South American males in all characters; and Mesoamerican females were significantly larger than South American females in two of four characters.
FIGURE 4. Scatterplots of principal component scores separated by sex, derived from four morphometric characters from Mesoamerican (C. d. dryas, C. d. harrisoni, C. d. ovandensis) and South American (C. d. maculatus) specimens. Ellipses encompass ≥95% of measured individuals.
Sexual Size Dimorphism. Specimens from Mesoamerica (C. d. dryas, C. d. ovandensis, and C. d. harrisoni) exhibited different patterns of sexual size dimorphism from those of South America (C. d. maculatus and C. d. ecuadoreanus; Table 1). In South America, males were significantly larger than females in wing length (two-sample $t$-test, $T = -4.63, df = 39, P < 0.001$), tarsometatarsus length ($T = -3.66, df = 39, P = 0.001$), and tail length ($T = -4.24, df = 39, P < 0.001$). The shape of the bill did not differ between the sexes in South America ($T = 1.73, df = 36, P = 0.092$). In Mesoamerica, males were significantly larger than females in mean wing length (two-sample $t$-test, $T = -4.30, df = 12, P = 0.001$) and tail length ($T = -3.32, df = 12, P = 0.006$), but the sexes did not differ in tarsometatarsus length ($T = -1.88, df = 11, P = 0.086$). Unlike South American specimens, the shape of the bill differed significantly between the sexes in Mesoamerica ($T = 3.26, df = 12, P = 0.007$), with females having shorter and wider bills than males on average.

Songs. The first two principal components accounted for 71% of the total variation among songs (Fig. 5), and there were significant differences between the putative species (MANOVA; males: Wilks’ $\lambda = 0.198, P < 0.001$). Discriminant analysis with cross-validation correctly classified 100% of songs ($N = 82$) to the correct population based on the four variables. This pattern was entirely driven by differences in Notes and 90% Duration, both of which were significantly different between regions (Table 2). Sonograms of Mesoamerican songs show a pattern of 6–10 short notes in a shape resembling a caret grapheme (‘^’), whereas songs from South America are structurally simpler, consisting of 2–5 relatively longer notes (Fig. 6).
**FIGURE 6.** Audio spectrograms showing variation in whistled songs from Mesoamerica (*C. d. dryas*), and South America (*C. d. maculatus*). Two songs each are shown (a, b) from the repertoires of 7 singers (1–7), representing Mexico (1–4), Peru (5), Bolivia (6), and Ecuador (7).

**FIGURE 7.** Maxent projected models of estimate ecological niche distributions for Mesoamerican (A) and South American (B) populations of *C. dryas*. Dotted lines indicate the approximate southern and northern range boundaries of the populations in Mesoamerica and South America, respectively. Black circles represent collection localities of genetic samples used in this study.
TABLE 2. Mean values (± SD) of four variables calculated from recorded songs of the *C. dryas* complex, from Mesoamerica (*N* = 7) and South America (*N* = 75). Results of two-sample *t*-tests (TST) and Mann-Whitney U-tests (MWU) of the null hypothesis (i.e., no difference between the regions) are shown via test statistics (*T*/*W*), degrees of freedom (df), and *P*-values (*α* ≤ 0.05).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mesoamerica mean ± SD</th>
<th>South America mean ± SD</th>
<th>Test</th>
<th><em>T</em>/<em>W</em></th>
<th>df</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes</td>
<td>8 ± 1.15</td>
<td>3.01 ± 0.69</td>
<td>MWU</td>
<td>9553.0</td>
<td>80</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Center Frequency</td>
<td>3052 ± 425 Hz</td>
<td>2816.3 ± 395.6 Hz</td>
<td>TST</td>
<td>1.5</td>
<td>80</td>
<td>0.138</td>
</tr>
<tr>
<td>Duration 90%</td>
<td>0.24 ± 0.05 s</td>
<td>0.48 ± 0.15 s</td>
<td>MWU</td>
<td>78.5</td>
<td>80</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Bandwidth 90%</td>
<td>1111 ± 418 Hz</td>
<td>854.4 ± 432.6 Hz</td>
<td>MWU</td>
<td>389.5</td>
<td>80</td>
<td>0.102</td>
</tr>
</tbody>
</table>

Environmental Niche Models. High AUC values (SA: 0.959, CA: 0.994) indicated that the models performed well in characterizing the environmental variables at presence sites. The ENM generated from South American specimen localities predicted suitable habitat in Mesoamerica within the range of *C. d. dryas* (Fig. 7A). However, the ENM generated from Mesoamerican specimen localities found very little suitable habitat in South America, and almost none in the Colombian Andes (Fig. 7B). The difference between the two models was significant (Schoener's *D*: observed mean = 0.22, null mean ± SD = 0.77 ± 0.04, *P* < 0.001; Schoener's *I*: observed = 0.48, null mean = 0.95 ± 0.02, *P* < 0.001), even though the top environmental variables contributing to the predictive power of each model were mostly the same. In the South American model, the top environmental variables, contributing a combined 93.2% of predictive power of the model, were: (1) slope (32.2%), (2) tree cover in 2015 (29.8%), (3) precipitation of the warmest quarter (11.2%), (4) elevation (10.3%), and (5) isothermality (9.7%). In the Mesoamerican model, the top environmental variables, contributing a combined 94.2% of the predictive power of the model, were: (1) slope (21.8%), (2) elevation (19.4%), (3) tree cover in 2015 (18.5%), (4) tree cover gain since year 2000 (12.5%), (5) precipitation seasonality (9.1%), (6) temperature seasonality (7.3%), and (7) isothermality (5.6%).

Discussion

Molecular Phylogenetics. The reciprocal monophyly of *C. d. dryas* and *C. d. maculatus* was strongly supported in the phylogenetic analysis. The mean uncorrected genetic distance between *C. d. dryas* and *C. d. maculatus* (4.1%) was an order of magnitude greater than the mean distances within each taxon (0.3% for both). Similarly, the mean uncorrected distance between *C. fuscescens* and *C. minimus* (2.7%) was also greater than within each of those taxa (0.9% and 1.3%, respectively). However, that one pair differed by 4.1% and the other by 2.7% does not provide any meaningful insight about whether any of the taxa are independent species, nor would comparison to other taxon-pairs in the genus (e.g., *C. u. ustulatus* and *C. u. swainsoni* (Cabanis in Tschudi, 1845); Ruegg 2008), because speciation does not occur at the same rate in every system (reviewed by Price 2008). More pertinent to species delimitation is the result that *C. d. dryas* and *C. d. maculatus* are reciprocally monophyletic and deeply divergent from each other, relative to mean genetic distances within each of those taxa (i.e., species are more divergent than subspecies, Huxley 1942). That distinction is at present clearer in the case of *C. d. dryas* and *C. d. maculatus* than in the two undisputed species *C. fuscescens* and *C. minimus*, but that is the extent to which such a comparison is useful or reasonable.

Plumage. We identified throat markings as taxonomically informative, a trait mentioned briefly by previous taxonomists but apparently not as a diagnostic feature. The unmarked throats of specimens from Guatemala were noted by Gould (1855), and the trait is evident in the lithograph that accompanied his description (Plate LXXV). Sclater (1858) mentioned (in Latin) that the throats of his type series of *M. maculata* were marked ("...gula et pectore summo triangulariter negro maculatis...," italics ours), but his subsequent English description focused only on dorsal coloration as a diagnostic feature. The presence or absence of throat markings is apparently sufficient to visually distinguish even faded specimens of the South American taxa (*C. d. maculatus*, *C. d. ecuadorianus*) from those of Mesoamerica (*C. d. dryas*, *C. d. ovandensis*, *C. d. harrisoni*).

Morphometrics. Mesoamerican and South American males differed significantly in all four body measurements, and discriminant analysis of morphometric characters classified 100% of males into the correct
population. The taxa also differed in which traits were sexually dimorphic; in Mesoamerica (but not South America), the shape of the bill differed significantly between the sexes, and in South America (but not Mesoamerica), tarsometatarsus length was sexually dimorphic.

**Songs.** Vocal differences between the populations were 100% diagnostic and in agreement with reconstructed phylogenetic relationships. Songs from Mesoamerica were shorter in duration and contained significantly more notes than those of South America (Table 2). Like the songs of *C. fuscens*, the sonograms of *C. d. dryas* resemble a caret grapheme (^), whereas songs from South America were structurally similar to the central phrases in songs of the Wood Thrush (see Stein 1956), consisting of 2–5 ‘whistled’ notes in an ascending, descending or alternating frequency pattern. Ortiz-Ramírez et al. (2015) recently showed a considerable degree of geographic variation in song structure among *C. frantzii* subspecies, coincident but not necessarily correlated with deep genetic divergences. That finding may, like our own, reflect the existence of cryptic species in the *C. frantzii* complex, a matter that deserves further study.

Unfortunately, at the present time, the relative influence of cultural drift, sexual selection, and ecological selection on the evolution of song in *Catharus* is not clear, complicating interpretation of acoustic data with respect to species delimitation. To our knowledge, isolation experiments to determine whether and to what extent normal song development depends on exposure to conspecific song have not yet been performed on any species in *Catharus*, but Lanyon (1979) and Whitney & Miller (1987) established that some aspects of central phrases in songs of the Wood Thrush, which are structurally similar to songs in the *C. dryas* complex from South America, are shaped by a learning process. Wild Wood Thrushes and those raised in isolation showed "similarities in basic structure, syntax, and rhythm of songs, as well as in the use of two simultaneous signals with unrelated frequencies," and isolated birds developed normal repertoire sizes without a tutor (Lanyon 1979). However, the central phrases of Wood Thrush songs were shorter in duration and "less musical, more discordant, and simpler in pattern" in experimental isolates compared to wild-type songs (Lanyon 1979), and this difference was sufficient to prompt a differential response from wild territorial birds to speaker playback of wild-type songs (strong response) and isolates' songs (no response). Lanyon (1979) interpreted this result as evidence that central phrases of Wood Thrush songs "transmit information required for species-recognition," but his experiments also showed that some components of Wood Thrush songs are apparently innate (e.g., syntax and repertoire size)—and thus potentially taxonomically informative. Indeed, because songs of the Wood Thrush are highly resistant to change, showing extreme temporal and spatial stability (Whitney 1989, 1992), we find Ortiz-Ramírez et al.'s (2015) statement that "other factors acting in a much faster time scale, such as sexual selection or cultural drift, that could act over the course of a few years, could be molding their songs," to be premature with respect to *Catharus* species. Nevertheless, we found that the songs of *C. d. dryas* and *C. d. maculatus* are structurally dissimilar and fully diagnosable, satisfying our criteria to delimit them as independent species.

**Environmental Niche Models.** Comparison of ENMs indicated that environmental niche characteristics of the regions inhabited by *C. d. dryas* and *C. d. maculatus* were significantly different, with elevation, temperature seasonality, and precipitation seasonality as major contributors to the models. These data suggest that the allopatric taxa inhabiting the two regions are ecologically divergent to some extent. In Mesoamerica, *C. d. dryas* inhabits humid evergreen forests at 1200–3000 m elevation (Howell & Webb 2005). By comparison, *C. d. maculatus* has more variance in its elevational range, being widely distributed across the Andes, but it is generally found at lower elevations than *C. d. dryas*. In parts of Peru and Ecuador, where *C. d. maculatus* is usually encountered between 750–1800 m, some populations occur as low as 400–650 m (Ridgely & Greenfield 2001; Schulenberg et al. 2007).

In the northern Andes of Venezuela and Colombia, *C. d. maculatus* occurs as high as 2100 m (Hilty & Brown 1986; Hilty 2003), and at the southern extent of its range in Bolivia and Argentina it occurs as high as 2600 m (Ridgely & Tudor 2009). At the very least, the significant differences between the ENMs of *C. d. dryas* and *C. d. maculatus* are suggestive that, were these taxa to come into secondary contact in the future, they would not occupy the same environments.

**Taxonomic Implications.** Regardless of the species concept applied, or one's philosophy about the yardstick approach, these data provide compelling evidence that *C. d. dryas* and *C. d. maculatus* are in fact, two independent species. They are 100% diagnosable in genetic, vocal, morphometric, and plumage characters. Likewise, ENMs indicate that the taxa occupy divergent ecological niches, which according to Peterson (2011) evolve after speciation is complete (but see Warren et al. 2008). Whether the magnitude of those differences are greater than in other species and subspecies pairs in *Catharus* is irrelevant, because as we have already mentioned, comparisons of
this nature are fraught with subjectivity and depend on the unreasonable assumption that speciation rates are uniform. Because of the extreme degree of geographic isolation of these sedentary taxa (over 1000 km apart), and the ongoing uplift of the Andes that is driving them still farther apart (Garzione et al. 2008), it is highly unlikely that they will come into secondary contact in the foreseeable future. Multiple lines of evidence indicate that they have ancient independent evolutionary histories and divergent evolutionary trajectories.

For these reasons, we advocate revising the taxonomy to reflect this conclusion, i.e., elevating the South American taxa to species status as *C. maculatus* (Sclater, 1858). Following Hellmayr (1934), we advocate using the English name Sclater's Nightingale-Thrush for this species, and for the Mesoamerican taxa retained within *C. dryas* (Gould, 1855), the English name Gould's Nightingale-Thrush. The long-standing English descriptor "Spotted" is not useful when applied to only one (or two) species in a genus that contains five other spotted species (in adult plumage), and in which all juveniles are spotted, so we advocate for discontinuation of its use, as well as to avoid confusion with past usage for these taxa. With our proposed revisions, *C. d. ecuadoreanus* Carriker, 1935 from western Ecuador and *C. d. blakei* Olor, 1973 from southern Bolivia and Argentina become *C. m. ecuadoreanus* and *C. m. blakei*, respectively; and the Mesoamerican subspecies, *C. d. ovandensis* Brodkorb, 1938 and *C. d. harrisoni* Phillips & Rook, 1965, retain their current names. We lacked DNA sequences from any location south of Cusco, Peru, and therefore are not able to determine whether the population in El Rey National Park, Argentina (Auer et al. 2007), belongs to *C. m. blakei* or the southern periphery of *C. m. maculatus*. We also have no data for *C. d. sztolcmani* (Domaniewski, 1918) from Peru, which was declared a synonym of *C. d. maculatus* (Chapman 1926; Hellmayr 1934). The validity of these subspecies has been doubted (e.g., Robbins & Ridgely 1990; Clement 2000) and should be tested via further study of vocal and genetic data. Notwithstanding, our data strongly support the reciprocal monophyly of *C. dryas* and *C. maculatus*, independent and divergent evolutionary lineages that should be treated as separate species.

**Acknowledgments**

For loans of tissues used in this study, we thank Jorge Pérez-Emán (Universidad Central de Venezuela, Instituto de Zoología y Ecología Tropical and Colección Ornitológia Phelps), Frederick H. Sheldon, Robb T. Brumfield, and Donna L. Dittmann (Louisiana State University Museum of Natural Science), Jon Fjeldså and Jan Kristensen (Zoological Museum—University of Copenhagen), and Shannon Hackett and John Bates (Field Museum of Natural History). We are grateful to Paul R. Sweet, Brian K. Schmidt, Christopher M. Milensky, Nathan H. Rice, and Jean L. Woods for assistance with museum collections, and to data scribes Sammi Ali, +Aziz, Ben Frederick, and Rachel Drane. Richard M. McCourt, two anonymous reviewers, and Pamela Rasmussen provided helpful comments on the manuscript. Valentina Gómez-Bahamón edited the Spanish abstract. We thank the many scientists and archivists who have built and maintained collections at the Academy of Natural Sciences of Drexel University, American Museum of Natural History, Delaware Museum of Natural History, University of Washington Burke Museum, Macaulay Library of Natural Sounds, and xeno-canto.org, for making specimen data and audio recordings publicly available online via VertNet.

**References**


ESRI (2014) ArcMap 10. ESRI (Environmental Systems Resource Institute), Redlands, California. [software]


Hellmayr, C.E. (1934) Catalogue of birds of the Americas and the adjacent islands in Field Museum of Natural History, including all species and subspecies known to occur in North America, Mexico, Central America, the West Indies, and islands of the Caribbean Sea, the Galapagos Archipelago and other islands which may be included on account of their faunal affinities. Field Museum of Natural History, Zoological Series, 13 (7), 1–531.


**SUPPLEMENTARY TABLE 1.** Table of environmental predicting factors used in ecological niche modeling procedure.

<table>
<thead>
<tr>
<th>Variable Type</th>
<th>Variable Name</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climatic Factors</td>
<td>Bio1—Annual Mean Temperature</td>
<td>WorldClim; (Hijmans <em>et al.</em> 2005)</td>
</tr>
<tr>
<td></td>
<td>Bio2—Mean Diurnal Range</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio3—Isothermality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio4—Temperature Seasonality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio5—Max. Temp. of the Warmest Month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 6—Min. Temp. of the Coldest Month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 7—Temperature Annual Range</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 8—Mean Temp. of the Wettest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 9—Mean Temp. of the Driest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio10—Mean Temp. of the Warmest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio11—Mean Temp. of the Coldest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio12—Annual Precipitation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio13—Precipitation of the Wettest Month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 14—Precipitation of the Driest Month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 15—Precipitation Seasonality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 16—Precipitation of the Wettest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio 17—Precipitation of the Driest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio18—Precipitation of the Warmest Quarter</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bio19—Precipitation of the Coldest Quarter</td>
<td></td>
</tr>
<tr>
<td>Topographic Factors</td>
<td>Elevation</td>
<td>NASA SRTM; (Farr <em>et al.</em> 2007)</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>Derived from above In ArcMap 10.3 (ESRI, 2014)</td>
</tr>
<tr>
<td>Vegetation</td>
<td>Percent Tree Cover</td>
<td>(Hansen <em>et al.</em> 2013)</td>
</tr>
<tr>
<td></td>
<td>Tree Cover Gain</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tree Cover Loss</td>
<td></td>
</tr>
</tbody>
</table>