SPECIAL FEATURE: Species Limits and Taxonomy in Birds

Integrative taxonomy and geographic sampling underlie successful species delimitation

Carla Cicero,[1,*,] **Nicholas A. Mason,**[1,2,] **Rosa Alicia Jiménez,**[1,3,4,] **Daniel R. Wait,**[1,3] **Cynthia Y. Wang-Claypool,**[1,3] **and Rauri C. K. Bowie**[1,3,]

[1] Museum of Vertebrate Zoology, University of California, Berkeley, California, USA
[2] Museum of Natural Science and Department of Biological Sciences, Louisiana State University, Baton Rouge, Louisiana, USA
[3] Department of Integrative Biology, University of California, Berkeley, California, USA
[4] Escuela de Biología, Universidad de San Carlos de Guatemala, Guatemala
* Corresponding author: ccicero@berkeley.edu

ABSTRACT

Species delimitation requires a broad assessment of population-level variation using multiple lines of evidence, a process known as integrative taxonomy. More specifically, studies of species limits must address underlying questions of what limits the distribution of populations, how traits vary in association with different environments, and whether the observed trait differences may lead to speciation through reproductive isolation. While genomic data have revolutionized the process of delimiting species, such data should be analyzed along with phenotypic, behavioral, and ecological traits that shape individuals across geographic and environmental space. The integration of multiple traits promotes taxonomic stability and should be a major guiding principle for species delimitation. Equally important, however, is thorough geographic sampling to adequately represent population-level variation—both in allopatry and across putative contact zones. We discuss the importance of both of these factors in the context of species concepts and traits and present different examples from birds that illustrate criteria for species delimitation. In addition, we review a decade of proposals for species-level taxonomic revisions considered by the American Ornithological Society's North American Classification Committee, and summarize the basis for decisions on whether to split or lump species. Finally, we present recommendations and discuss challenges (specifically permits, time, and funding) for species delimitation studies. This is an exciting time to be studying species delimitation in birds: many species-level questions remain, and methodological advances along with increased access to data enable new approaches to studying age-old problems in avian taxonomy.

Keywords: contact zones, ecology, genomics, geographic sampling, integrative taxonomy, phenotype, species delimitation

LAY SUMMARY

- Standardized taxonomy and nomenclature are fundamental to biodiversity research and conservation, with wide-reaching impacts for diverse stakeholders.
- Rigorous approaches to species delimitation require an integrative assessment of population-level variation using multiple lines of evidence and robust geographic sampling.
- Sampling should reflect phenotypic, ecological, and behavioral trait variation and include contact zones between divergent populations if relevant.
- We review basic concepts and criteria in species delimitation, drawing from examples across different avian lineages.
- We summarize the outcomes of a decade of species-level proposals considered by the American Ornithological Society's North American Classification Committee.
- We provide recommendations and discuss challenges for researchers interested in pursuing species-level taxonomic studies.
- Many species-level questions remain in ornithology, and this is an exciting time to be studying integrative species delimitation in birds.

La taxonomía integradora y el muestreo geográfico subyacen a la delimitación de especies exitosa

RESUMEN

La delimitación de especies requiere una evaluación amplia de la variación a nivel de poblaciones utilizando múltiples líneas de evidencia, un proceso conocido como taxonomía integradora. En específico, los estudios de límites de especies deben abordar preguntas subyacentes acerca de qué delimita la distribución de las poblaciones, cómo varían los rasgos en asociación con diferentes ambientes y si las diferencias observadas en los rasgos pueden conducir a la especiación a través del aislamiento reproductivo. Si bien los datos genómicos han revolucionado el proceso de delimitar especies, dichos datos deben analizarse en conjunto con los rasgos fenotípicos, de comportamiento y ecológicos que moldean a los individuos en el espacio geográfico y ambiental. La integración de múltiples rasgos promueve la estabilidad taxonómica y debería ser un principio rector importante para la delimitación de especies. Sin embargo, igualmente importante es el muestreo geográfico exhaustivo para representar adecuadamente la variación a nivel de poblaciones— tanto en alopatría como en potenciales zonas de contacto. Discutimos la importancia de ambos factores en el contexto de conceptos de especie y los atributos de las especies, también presentamos diferentes ejemplos de aves que ilustran los criterios para la delimitación de especies. Además, revisamos una década de las propuestas consideradas por el Comité de Clasificación de América del Norte de la Sociedad Americana de Ornitología y resumimos el fundamento utilizado para tomar las decisiones de separar o unir especies. Finalmente, presentamos recomendaciones de muestreo y discutimos los desafíos (específicamente permisos, tiempo y financiamiento) para realizar estudios de delimitación de especies. Es un momento emocionante para estudiar la delimitación de especies en aves: quedan muchas preguntas a nivel de especie y los avances metodológicos junto con un mayor acceso a datos permiten nuevos enfoques para estudiar problemas antiguos de la taxonomía de aves.

Palabras clave: zonas de contacto, ecología, genómica, muestreo geográfico, taxonomía integradora, fenotipo, delimitación de especies

INTRODUCTION

Researchers interested in understanding and describing species limits must have a firm grasp of what limits the distribution of populations, how traits vary in association with different environments, and whether the observed trait differences reflect pre- or post-zygotic barriers to reproduction (Price 2008, Hudson and Price 2014). For example, morphological differences may arise because of adaptation to local environmental niches (Pigot et al. 2020), and subsequent genomic changes may lead to lineage divergence and ultimately speciation (e.g., Hawaiian honeycreepers; Campana et al. 2020). Likewise, song differences may result from acoustic adaptation to specific habitats, morphological constraints, evolutionary history, and/or cultural drift (Derryberry 2009, Giraudeau et al. 2014, Potvin and Clegg 2015, Karin et al. 2018, Derryberry et al. 2020), which in turn may lead to species recognition and reinforcement (Pfennig 2016). "Integrative taxonomy," in which multiple lines of evidence from different suites of traits (e.g., ecology, genetics, phenotype, behavior, and/or geography) are analyzed to identify evolutionary lineages, is a particularly useful approach for delineating species (Padial et al. 2010, Schlick-Steiner et al. 2010, Fujita et al. 2012, Ribeiro et al. 2014, Venkatraman et al. 2019). Likewise, documenting how traits vary across space and time, and using that information for taxonomic decisions, is critical for conserving biological diversity (Sangster 2018).

The advent of Sanger and later genomic sequencing methods, along with associated analytical approaches (e.g., coalescent-based species delimitation; Fujita et al. 2012), has provided new tools for uncovering diversity and delimiting species. Such advances have led to significant findings regarding patterns and processes of genetic variation, and how hybridization and introgression between divergent populations affect species limits (Gowen et al. 2014, Toews et al. 2016, Billerman et al. 2019). While genomic data have revolutionized the field of systematics, studies aimed at delimiting species are best framed in the context of individual phenotypic, behavioral, and ecological traits across geographic and environmental space. A strong sampling scheme coupled with integrative approaches that examine different suites of traits will result in a clearer picture of geographic variation in phenotypes and genotypes, their association with the environment, the importance of biotic interactions in driving trait variation (e.g., character displacement), and species limits (Wang et al. 2019).

Given the extensive literature on species delimitation (Google Scholar search on the term "species delimitation" July 3, 2020 resulted in 126,000 papers), it is beyond the scope of this paper to review all of the concepts and methods that have been proposed to address this complex topic. Rather, we address species delimitation with 5 goals: (1) provide an overview of species concepts and traits as a framework for discussing species delimitation in birds; (2) present criteria for species delimitation, illustrated by avian examples; (3) emphasize the importance of robust geographic sampling and the consequences of biased sampling; (4) review species delimitation in practice, using the American Ornithological Society's (AOS) North American Classification Committee (NACC) as an example; and

(5) provide recommendations and discuss challenges for current and future researchers, especially students and early-career professionals who are interested in pursuing species-level taxonomic studies.

SPECIES CONCEPTS AND TRAITS

Species concepts have been the subject of much debate for decades (Mayr 1957, McKitrick and Zink 1988, Zink and McKitrick 1995, Johnson et al. 1999, De Queiroz 1998, 2007, Ottenburghs 2019). Although this paper does not seek to critique different species concepts, the topic is unavoidable when discussing species delimitation. Different concepts agree that species represent separate evolutionary lineages, but they disagree on the criteria (e.g., reproductive isolation, diagnosability, monophyly) used to delimit species along the divergence continuum (De Queiroz 2007, Ottenburghs 2019). The Biological Species Concept (BSC; Mayr 1940, 1942) and the Phylogenetic Species Concept (PSC; Cracraft 1983) have been applied most commonly in ornithology. The BSC is inherently attractive because it incorporates the property of interbreeding, with reproductive isolation as the main criterion used to delimit species. Arguments against the BSC (Zink and McKitrick 1995; but see counter-arguments by Johnson et al. 1999) focus on its difficult application to allopatric populations, hybridizing lineages, and non-historical groups. Although the BSC may be challenging to apply at times, it can be used to evaluate the taxonomic status of populations based on multi-trait analyses (Young et al. 2000) and estimated divergence times among closely related lineages relative to geographic breaks (Ramírez-Barrera et al. 2018, Venkatraman et al. 2019). Under pragmatic applications of the BSC, it is important to recognize that low levels of hybridization may occur even in well-delineated species (Toews et al. 2018, 2020). The incorporation of modern genomic data provides increased resolution for diagnosing species while uncovering the genetic basis of reproductive isolation (Ottenburghs 2019), and further enables characterization of the genomic landscape of introgression (Martin and Jiggins 2017). Proponents of the PSC argue that it is easier to apply operationally than the BSC because it relies on objective criteria of diagnosability and monophyly. While these are tangible considerations, critics of the PSC have focused on the arbitrariness of setting diagnosability levels, the need to consider multiple traits and their geographic variation, the attention on pattern vs. process, and the reality of species-level paraphyly or polyphyly due to reticulation (Johnson et al. 1999, Funk and Omland 2003, Freudenstein et al. 2017). Furthermore, shared traits may result from convergent evolution rather than phylogenetic history (Pigot et al. 2020). Other concepts such as the General Lineage Concept of Species (De Queiroz 1998, 2007) provide a unifying approach by separating the concept of species (a separately evolving metapopulation lineage) from the operational criteria used for species delimitation. It also asserts that any evidence of lineage divergence is relevant to determining species limits, and that species delineation is strengthened by corroboration among multiple lines of evidence (De Queiroz 2007). Likewise, an analysis of the application of species criteria in avian taxonomic studies concluded that different criteria can be viewed as complementary rather than as rival approaches to species delimitation (Sangster 2014). While species are viewed as historically connected populations, restricting the concept to lineages is insufficient without integrating their unique ecological roles and extended phenotypes (Freudenstein et al. 2017).

We approach the issue of species delimitation from the perspective of the General Lineage Concept of Species (De Queiroz 1998, 2007), but follow the NACC in viewing evidence for *essential* (not necessarily complete) reproductive isolation per the BSC (Mayr 1940, 1942) as important for assessing species limits. We also adhere to the importance of integrating multiple traits—per the extended phenotype (Freudenstein et al. 2017)—to corroborate species limits (De Queiroz 2007). Our fundamental premise is that interpretation of species limits requires an understanding of their basic natural history, ecology, and factors influencing their distribution and geographic variation. This reflects several long-standing principles (Mayr 1940, Grinnell 1943) underlying species and their relationship to the environment: (1) species consist of populations that are separated geographically and/or ecologically and that have the potential for not interbreeding where they come into contact (Mayr 1940); (2) species' distributions are limited strongly by their environment, reflecting adaptations to different abiotic as well as biotic conditions (Grinnell 1917); (3) traits that characterize species are shaped through geographic and environmental isolation; this holds true whether their populations occur on oceanic islands (Johnson 1972, Baier and Hoekstra 2019, Hanna et al. 2019), islands of montane habitat (Bowie et al. 2004, Tennessen and Zamudio 2008, Manthey and Moyle 2015), or ecologically variable habitats on continents (Pitelka 1951, Bardwell et al. 2001, Cicero and Koo 2012); and (4) ecoregions and associated "faunal districts" (Grinnell 1915, 1935) represent important areas for evolutionary change driven by environmental factors. Furthermore, ecotones between different habitats may be a source of evolutionary novelty (Smith et al. 1997) and often represent contact or suture zones between divergent lineages (Harrison 1993, Cicero 2004).

A fifth and more controversial principle is that subspecies are critical stages in evolution, forming part of a continuum from limited differentiation among populations to reproductive isolation per the General Lineage Concept of Species (De Queiroz 2007). The utility of subspecies

has been debated for decades, and many subspecies require reexamination to assess their validity (Wiens 1982, Zink 2004, Phillimore and Owens 2006, Winker and Haig 2010). In North American birds, for example, most avian subspecies were described decades ago based on few specimens and characters, and names have been applied to populations that are barely diagnosable, differ clinically, or show strong phenotypic differences (Johnson 1982, Cicero 2010). Furthermore, molecular and phenotypic characters often conflict at the subspecies level, and subspecies may represent local adaptation (Luttrell et al. 2015, Walsh et al. 2017) that does not necessarily lead to speciation. Nonetheless, subspecies bring attention to recognized geographic variation in traits and often exhibit genetic differentiation as well (Phillimore and Owens 2006, Winker 2009, Patten and Remsen 2017). Thus, they can be useful for elucidating biological differences that are potentially meaningful for species delimitation (Cicero and Johnson 2006, Winker and Haig 2010, and papers therein).

CRITERIA FOR SPECIES DELIMITATION

Given the importance of rigorously defining species for studying and conserving biodiversity (Wiens 2007, Stanton et al. 2019), what criteria can be applied broadly for species delimitation? Sangster (2014) provided a thorough review of different criteria and how those relate to different species concepts. Although the process of identifying species limits is inherently arbitrary in some cases, especially for allopatric populations, efforts have been made to develop quantitative criteria based on levels of differentiation in phenotype (Tobias et al. 2010) or genetics (Hebert et al. 2004, Moritz and Cicero 2004, Roux et al. 2016). Multispecies coalescent methods for analyzing phylogenetic and genomic data (Yang and Rannala 2010, Jiang et al. 2020) also have advanced taxonomic studies, although questions have been raised about their application to species delimitation (Sukumaran and Knowles 2017). Regardless of the methods used to study species, speciation occurs along a continuum of genetic and phenotypic divergence (De Queiroz 2007, Roux et al. 2016). At one extreme, species may show weak genetic divergence but strong divergence in other traits. For example, Red-breasted and Red-naped sapsuckers (*Sphyrapicus ruber* and *S. nuchalis*) have low levels of divergence in allozymes (Johnson and Zink 1983), mitochondrial DNA (mtDNA, Cicero and Johnson 1995), and across thousands of single nucleotide polymorphisms acquired through a reduced genomic approach (Billerman et al. 2019). Nonetheless, these species are clearly diagnosable in plumage, and they maintain phenotypic boundaries in the face of extensive admixture, likely due to strong assortative mating where they contact (Johnson and Johnson 1985, Billerman et al. 2019). Similarly, Hooded and Carrion

crows (*Corvus cornix* and *C. corone*) show low levels of genomic differentiation and exhibit introgression across most of the genome, yet also demonstrate divergence concentrated in a relatively small number of genes involved in pigmentation (gray vs. black feathers, respectively) and visual perception. Like the sapsuckers, these crows show assortative mating preferences based on plumage differences, thus maintaining phenotypic differences in spite of extensive gene flow (Poelstra et al. 2014). A third but more controversial case involves the Common, Hoary, and Lesser redpolls (*Acanthis flammea*, *A. hornemanni*, and *A. cabaret*), which exhibit continuous variation in plumage and morphology, show substantial niche overlap, and have largely undifferentiated genomes (Mason and Taylor 2015). Evidence regarding mating behavior (i.e. assortative or mixed pairs, hybrid offspring) is mixed (Mason and Taylor 2015) and requires further study. While these data call into question species boundaries in the redpoll complex and suggest a single, polymorphic species that has undergone rapid phenotypic evolution, whether they represent 1, 2, or 3 species remains unsettled.

At the other extreme, species may show deep molecular divergence but little phenotypic differentiation. One classic example involves sibling species of *Empidonax* flycatchers, which are renowned for their morphological stasis in the face of genetic, vocal, and ecological differences (Johnson and Cicero 2002). Another involves morphologically cryptic species where molecular markers have been used to reveal divergent lineages that also differ in behavioral and ecological traits (e.g., Brazilian tyrant-flycatchers, Chaves et al. 2008; American and Pacific Golden-Plovers, Withrow and Winker 2014; Amazonian bird species pairs, Pulido-Santacruz et al. 2018; Andean antpittas, Chesser et al. 2020). Likewise, phenotypic convergence may lead to broad-scale similarity that masks deep divergences or polyphyly, as in the case of *Sporophila* seedeaters (Mason et al. 2018). Pronounced ecological differences are not a strict prerequisite for reproductive isolation to evolve (Rundell and Price 2009). Rather, mutations and genomic incompatibilities can accrue in allopatry independently from ecological variation, leading to reduced reproductive compatibility upon secondary contact (Orr and Turelli 2001). Cryptic species occur across diverse groups of plants and animals in both temperate and tropical regions and are continuously being reported; therefore, they contribute importantly to understanding patterns and processes of diversification and can provide novel insights for assessing species boundaries that should be incorporated into biodiversity science and conservation (Fujita et al. 2012, Espíndola et al. 2016, Fišer et al. 2018).

Because different types of data may show conflicting patterns, species delimitation studies that incorporate multiple traits in an integrative taxonomic framework

provide a more comprehensive view of lineage diversification and species limits. An excellent recent example involves the Unicolored Jay (*Aphelocoma unicolor*), which has 5 subspecies found in different cloud forest regions of Mexico and Central America (Venkatraman et al. 2019). Analyses of genomic divergence (thousands of ultraconserved elements), phenotype (morphology and plumage), and ecological niche (temperature and precipitation) showed that populations of each subspecies represent 5 lineages that occupy isolated forest patches; 2–4 of these lineages may represent distinct species based on comparisons with other species-level divergences in *Aphelocoma* Jays (Venkatraman et al. 2019), although there is some debate on whether these lineages should be elevated to species status (https://americanornithology.org/nacc/current-prior-proposals/2020-proposals/proposals-2020-c/#2020-C-17). One concern is the lack of vocal data, which the authors acknowledged would add value to their integrative dataset—the limitation being the availability of samples to address this question quantitatively.

Another comprehensive example of integrative taxonomy involves a pair of incipient shorebird species with parapatric breeding ranges: the Kentish Plover (*Charadrius alexandrinus*) and White-faced Plover (*C. dealbatus*) (Wang et al. 2019). Prior work confirmed differences in morphology, plumage, ecology, and behavior, but found no evidence of genetic differentiation (Rheindt et al. 2011), thus leading to uncertainty about their taxonomic status and mixed treatment in checklists of avian diversity (Clements 2019, Gill et al. 2020, del Hoyo et al. 2020). Wang et al. (2019) revisited this case by sampling broadly (*n* = 454 individuals from 19 populations, including a contact zone) with a focus on obtaining data for mitochondrial, nuclear, and microsatellite markers, in combination with morphological measurements, ecological niche models, and stable isotope (dietary) profiles, to assess species limits and gene flow. Evidence of character displacement and ecological niche differentiation suggested that divergent selection associated with ecological differences led to incipient speciation in the face of historical symmetric gene flow. Although the authors detected only a small number of hybrids and did not find a broad hybrid zone where these taxa overlap, they acknowledged that a narrow hybrid zone may be observed with more fine-grained sampling. Taken together, the data by Wang et al. (2019) provide evidence that Kentish and White-faced plovers are distinct lineages that deserve full species status under the General Lineage Concept of Species (De Queiroz 2007). However, strict adherence to the criterion of reproductive isolation under the BSC, and desire for an increased sampling of the contact zone, may lead some taxonomists to oppose the split, highlighting the nuances of many species-level taxonomic proposals.

While integrative taxonomy is an important guiding principle for species delimitation, there are cases where a complex evolutionary history and behavioral observations of mating preferences present a taxonomic challenge despite diagnosability across multiple traits. One classic but controversial example involves the Yellow-rumped Warbler (*Setophaga coronata*) complex. This species currently consists of 4 subspecies (*S. c. coronata*, *S. c. auduboni*, *S. c. nigrifrons*, and *S. c. goldmani*; Clements et al. 2019) that are diagnosable phenotypically, genetically, ecologically, vocally, and in migratory behavior (Hubbard 1969, Milá et al. 2007, 2008, Brelsford et al. 2011, Toews et al. 2013, 2016). Data thus far have resulted in a recommendation to split the Yellow-rumped Warbler into 3 species (Toews et al. 2016), which is followed by some taxonomic authorities (Gill et al. 2020): Goldman's Warbler (*S. goldmani*), Myrtle Warbler (*S. coronata*), and Audubon's Warbler (*S. auduboni*, including *S. a. auduboni* and *S. a. nigrifrons*). However, several factors complicate species delimitation in this complex. First, Brelsford et al. (2011) provided evidence that *S. auduboni* originated through hybridization between long-diverged *coronata* and *nigrifrons*; Jacobsen and Omland (2011) likewise acknowledged the role that gene flow plays in speciation, but suggested that other possible scenarios may be elucidated with more thorough geographic sampling. Second, Toews et al. (2016) noted the need for a further sampling of *auduboni* and *nigrifrons* near putative contact zones to assess levels of gene flow and reproductive isolation between those taxa (also see Milá et al. 2011). Third, a close study of the narrow contact zone between *S. c. coronata* and *S. c. auduboni* (Hubbard 1969, Brelsford and Irwin 2009) has shown that assortative mating is weak or absent, leading to extensive hybridization within the zone but strong differences away from the zone. The limited width and temporal stability of the zone, along with evidence for fixed differences at 2 nuclear loci and concordance with plumage patterns, have been invoked to argue for postmating reproductive isolation, selection against hybrids, and species status (Brelsford and Irwin 2009). However, a genome-wide analysis of the complex (Toews et al. 2016) did not include samples from the contact zone, and those authors note that further study is needed to elucidate how divergent genomic regions relate to phenotypic differences and reproductive barriers in areas of sympatry. While the *S. coronata* complex appears to be in the process of incipient speciation (Brelsford and Irwin 2009), the debate continues over whether these represent biological species and how many species to recognize. Twice over the past decade, the NACC has reviewed proposals to split this complex into 2–4 species, and the majority consensus has been to retain the status quo for now; see comments for proposals 2010-A-4 (https://americanornithology.org/nacc/current-prior-proposals/2010-proposals/comments-

2010-a) and 2017-A-9 (https://americanornithology.org/nacc/current-prior-proposals/2017-proposals/comments-2017-a).

GEOGRAPHIC SAMPLING

Species delimitation based on integrative taxonomy requires robust geographic sampling to capture trait variation across the distributional and environmental range of populations and any described subspecies. Likewise, more intensive sampling is often necessary where phenotypically and/or ecological divergent populations contact. Developing a strategy for sampling requires an understanding of species traits as outlined previously. For this reason, species delimitation should incorporate targeted field sampling where possible rather than relying solely on existing material in museum collections (although existing samples may be used to supplement field-collected material). Familiarity with organisms in their natural environment, and first-hand experience with geographical differences in ecology, behavior, and phenotype, is indispensable for interpreting patterns of divergence in genomic or other traits.

Although exhaustive sampling is often logistically difficult, especially for taxa that occupy large or remote geographic regions, incomplete sampling may lead to biased findings. Simulations supported by empirical analyses have shown that sampling can have a clear effect on species delimitation scenarios, especially where there are strong patterns of isolation by distance (Mason et al. 2020). In a study of the *Empidonax difficilis*/*E. occidentalis* complex, dense sampling from a previously unsampled portion of the range in Mexico uncovered paraphyly and a cryptic lineage based on genomic data, thus highlighting the consequences of biased geographic sampling for species delimitation (Linck et al. 2019). Another study of Sagebrush and Bell's sparrows (*Artemisiospiza nevadensis* and *A. belli*, respectively) with extensive sampling (Figure 1) revealed both concordant and discordant differences among populations and subspecies (Johnson and Marten 1992, Cicero and Johnson 2007, Cicero and Koo 2012, Karin et al. 2018). Dense sampling and analyses of genetic, morphological, vocal, and ecological niche variation in different environments and across a narrow contact zone between *A. nevadensis* and *A. belli canescens* showed concordant patterns and abrupt transitions in all of these traits. The sampling also uncovered discordance in 2 populations of *A. b. canescens* (southern San Joaquin Valley and Mojave Desert of California) that are geographically separated by the Tehachapi Mountains. Although these 2 populations are similar in body size, plumage coloration, song, and ecology, they are genetically distinct (Cicero and Koo 2012, Karin et al. 2018). Findings show that *A. b. canescens* from

the San Joaquin Valley is allied genetically with *A. b. belli* from the California Coast Ranges, even though *A. b. belli* differs in phenotype (smaller and darker), song (more rapid with a distinct structure), and environment (denser chaparral habitat, different bioclimate) (Johnson and Marten 1992, Cicero and Koo 2012, Karin et al. 2018). This discordance, which requires further study, would not have been uncovered if *A. b. canescens* was represented only by samples from the Mojave Desert, which is the primary ecoregion that it occupies.

While sampling in contact zones is important for inferring reproductive isolation and species limits on the basis of different traits (per the BSC), this is not possible for allopatric populations. Despite the challenge of delimiting species in these cases, sampling that covers the geographic distribution of allopatric populations, and that integrates multiple traits, should remain a driving principle for these kinds of studies (Fujita et al. 2012). Quantitative methods for comparing genetic and phenotypic traits among allopatric populations provide metrics of divergence that can be useful when determining species status and identifying cryptic species (Campbell et al. 2016). Likewise, the incorporation of Geographic Information System (GIS) analyses—along with genetics, morphology, and behavior—can be used to rigorously examine how traits are distributed across environmental space and to understand what limits the geographic and ecological range of populations (Raxworthy et al. 2007, Wiens 2007).

SPECIES DELIMITATION IN PRACTICE: THE NACC

Standardized taxonomy and nomenclature are critical for research and conservation, and impact diverse stakeholders including scientists, educators, students, wildlife managers, government agencies, and the general public. In ornithology, this extends to the millions of birdwatchers who use field guides and apps for species identification or who contribute to citizen science efforts such as eBird. Thus, it is the responsibility of biologists to approach species delimitation scientifically and objectively and make use of the best available published data. The AOS NACC is responsible for reviewing data that relate to the systematics, taxonomy, and nomenclature of birds from the North Pole to the boundary between Panama and Colombia, including adjacent islands. Thus, the committee serves all of these diverse interest groups who rely on its decisions for species-level taxonomy. The NACC operates on a proposal basis, whereby proposals may be submitted by anyone (either on the committee or not) who considers that a change is warranted. There is an open process for proposal submission, and the NACC's Early Career Systematics Group has helped to engage a new and more diverse generation of scientists in taxonomic and nomenclatural issues. The
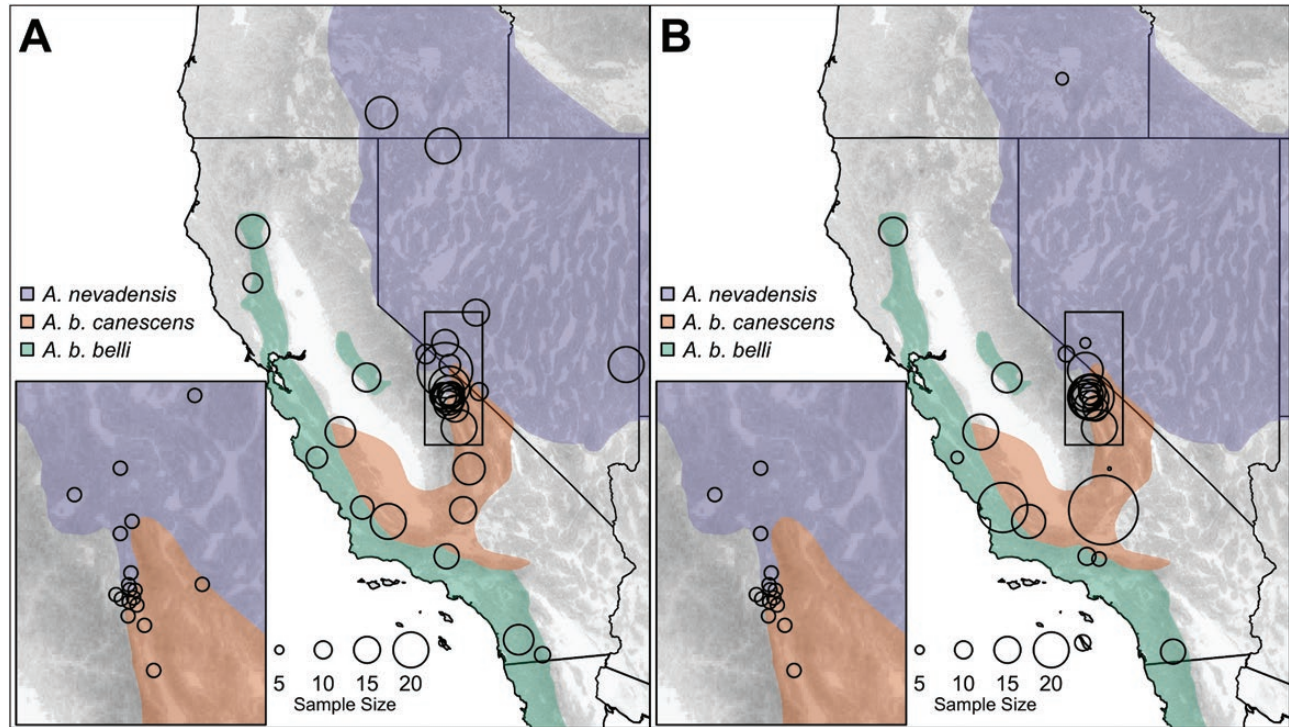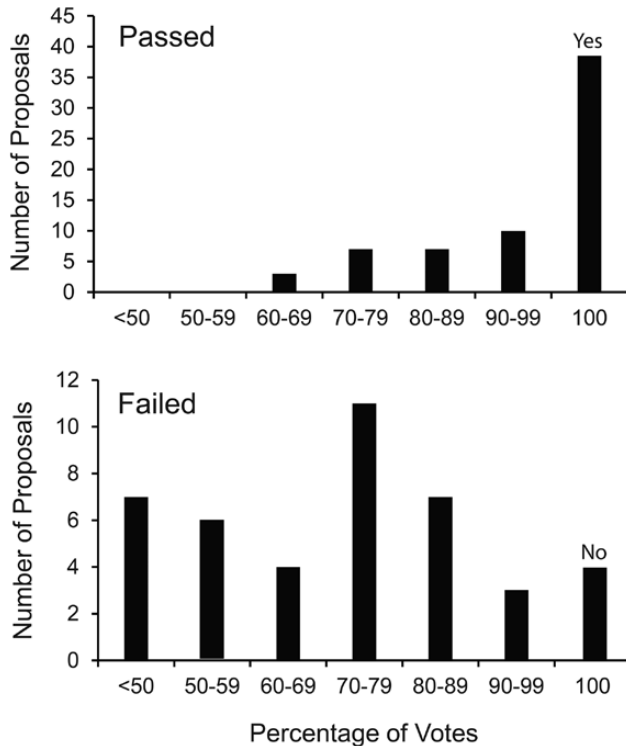
FIGURE 1. Maps showing the distribution and sample sizes for a species delimitation study of *Artemisiospiza nevadensis* and *A. belli*. (A) Samples used in morphological and genetic analyses (allozymes, mtDNA, microsatellites). (B) Samples used in song analyses. Ranges of the continental USA taxa (*A. nevadensis*, *A. b. belli*, *A. b. canescens*) are shown. The subspecies *A. b. clementeae*, endemic to the California Channel Islands, was studied for song and genetic variation. A fifth subspecies from northern Baja California (*A. b. cinerea*) was not studied. The figure illustrates the multiple traits used for species delimitation, and the emphasis on broad sampling with more intensive coverage where *A. nevadensis* and *A. belli* contact in eastern California.

content of proposals ranges from species delimitation to changes in higher-level classification or linear sequence, English name changes, and distributional records. All proposals are reviewed critically by committee members, and external input may be solicited to provide additional perspectives on specific issues. In addition, information and resources not mentioned in the proposal (e.g., other literature, audiovisual repositories such as xeno-canto [https://www.xeno-canto.org/] or Macaulay Library [https://macaulaylibrary.org/], museum specimens, personal field experiences) also may be brought to bear on the proposal topic. Members then provide discussion and comments internally, which sometimes involve deliberation and debate on certain topics, and ultimately vote on each proposal. Successful proposals (i.e. proposals that pass) must receive at least two-thirds majority vote. All votes and comments, as well as guidelines for proposals and English name changes, are posted for transparency on the AOS's website. For more information about how the committee operates and to view proposals, visit the AOS website (https://americanornithology.org/nacc).

Historically, the NACC has followed the BSC in theory and practice (American Ornithologists' Union 1998). Furthermore, in the interest of taxonomic stability, the committee tends to operate conservatively and requires strong justification for the change. As noted in the preface to the Check-list (American Ornithologists' Union 1998), the committee favors multiple lines of evidence and independent datasets that provide an integrative perspective. To investigate how the NACC implements species delimitation in practice, we reviewed NACC proposals that evaluated species-level issues from 2008 to 2018 and assessed the basis for why a proposal passed or not. We considered a total of 138 proposals (Supplementary Material Table S1), but excluded 19 that involved extralimital or exotic taxa (e.g., where other authorities take priority) or where the proposal was later withdrawn. We also excluded 10 proposals where the author recommended rejection, with the intent to get a NACC decision on record. For the remaining 109 proposals, we recorded the following information: (1) Traits mentioned in the proposal (and references therein) as the basis for a recommended species split or lump. These included genetics, morphology, plumage, voice, other behavior, and habitat. Typically, this is provided in the "New Information" section of a proposal, but also may include information in the "Background" section. (2) The total number of genetic samples analyzed, if applicable (regardless of the geographic coverage, which

FIGURE 2. Distribution of votes for AOS North American Classification Committee proposals (2008–2018) that passed vs. failed. Data reflect the number of proposals that received a certain percentage of votes in favor of ("Yes") or against ("No") a proposal.

TABLE 1. Summary of traits mentioned in AOS North American Classification Committee proposals that passed or failed. Values indicate the number of proposals that mentioned a particular type of trait.

Trait mentioned in proposal	Passed ($n = 67$)	Failed ($n = 42$)
Genetics	47	39
Morphology	45	25
Plumage	48	25
Voice	44	24
Other behavior	18	12
Habitat	16	15
Number of traits mentioned (mean)	3.3	3.3
Number of traits mentioned (range)	1–6	1–6

or range of traits mentioned between successful and unsuccessful proposals; however, more successful proposals referenced genetic data and other kinds of traits compared to unsuccessful proposals (Table 1). Furthermore, congruence among traits was viewed as supporting evidence for the proposed species delimitation—thus reflecting an integrative taxonomic approach in practice.

The total number of traits examined is not an adequate measure of a successful proposal because of differences in (1) the types of data presented, (2) the number and geographic distribution of samples, and (3) whether the information is qualitative or quantitative (Supplementary Material Table S1). For example, of the 69 proposals that mentioned vocalizations, most involved qualitative assessments while a small number ($n = 29$) was based on quantitative analysis of acoustic traits and even fewer ($n = 5$) referenced formal playback experiments. Likewise, of the 94 proposals that included published genetic data, there was wide variation in the number and type of loci examined and in the number of individuals sampled. The scale of genetic datasets ranged from single-locus mitochondrial gene sequences to multi-locus Sanger sequencing datasets to more recent genomic data with thousands of loci. Similarly, sample sizes varied from a few individuals (*Rhynchopsitta pachyrhyncha/R. terrisi*, proposal 2014-A-9, part of a larger phylogeny) to nearly 700 in the case of *Aphelocoma californica/A. woodhouseii* (proposal 2016-A-1).

Another important consideration was the extent to which contact zones were sampled for parapatric populations. Of the 109 proposals considered here, 41 involved taxa with populations that were known or thought to occur in parapatry; the other 68 involved allopatric populations. As noted previously, sampling in contact zones is important for assessing species limits in the context of the BSC (the operational concept for the NACC). Although this only applies to parapatric populations, the relative paucity of data from contact zones was noteworthy.

is clearly important but often not specified in proposals). (3) The known distribution of the proposed species (i.e. whether populations are allopatric or parapatric). (4) Whether the contact zone was sampled, if applicable (i.e. for parapatric populations). (5) The number of "yes" and "no" votes for each proposal. (6) A summary of the comments in favor of, or against, a proposed split or lump.

Slightly more than half (60, 55%) of the 109 proposals concerned passerines compared to non-passerines (49, 45%). One striking result was that only 67 (61%) of the proposed changes in species delimitation were successful. Furthermore, successful proposals were much more likely to get a unanimous vote compared to unsuccessful proposals; 39 (58%) of the 67 successful proposals received a unanimous "yes" vote, while 4 (10%) of the 42 unsuccessful proposals received a unanimous "no" vote. The distribution of the number of votes for proposals that passed vs. those that failed (Figure 2) provides a metric for evaluating the level of agreement in favor of, or against, a proposal. Strong, convincing proposals typically received majority or unanimous votes supporting the change, while proposals with more mixed votes reflect greater contention or disagreement about the strength of the argument being made. Interestingly, there was no difference in the mean number

Of the 41 proposals involving parapatric populations, only about half (*n* = 22) reported some level of sampling in the contact zone. Furthermore, contact zone data varied from minimal sampling (e.g., *n* = 3, *Poecile gambeli*, proposal 2010-A-3) to more focused efforts (e.g., *Troglodytes troglodytes/T. pacificus*, proposal 2009-A-3; *Artemisiospiza belli/A. nevadensis*, proposal 2013-A-8). Despite the fact that parapatric populations present an opportunity to directly assess species limits, relatively few proposals reflected a concerted effort to sample in contact zones.

Because contact zone sampling does not apply to allopatric populations, it is worth examining the extent to which those proposals involved multi-trait studies. Corroboration of multiple lines of evidence provides a stronger justification for species limits than single-trait analyses (Sangster 2018), especially in the absence of direct evidence from contact zones. Of the 68 proposals involving allopatric populations, the majority (61, 90%) examined 2 or more traits and 68% reported on 3 or more traits. A higher proportion of successful proposals for allopatric taxa mentioned multiple traits. There were no proposals that passed on the basis of genetic data only.

Given these variables, what made a successful proposal to split or lump species? Our review revealed wide variation in the content of proposals, including the number and types of traits examined, whether the evidence was quantitative or qualitative, and the extent of sampling (total number of samples and whether they extended into contact zones for parapatric populations). Thus, a direct comparison of proposals is difficult. Nonetheless, features of successful proposals can be summarized in 4 main points, as revealed by our review and proposal comments (Supplementary Material Table S1): (1) Multiple sets of traits were diagnosed, showed congruent patterns, and could be used to infer likely reproductive isolation (or lack thereof), regardless of whether the taxa occurred in allopatry. In general, diagnosability followed by reproductive isolation is the most frequently applied criteria for species delimitation (Sangster 2014), although the latter is often a judgment call in the case of allopatric populations; in cases where reproductive isolation cannot be tested directly, the key is in evaluating concordance in different diagnosable traits among populations (Padial et al. 2010). (2) Data were based on robust geographic sampling, including the availability of samples from putative contact zones; the lack of samples from contact zones was cited often as a missing component in proposals that failed, where relevant. (3) Information for or against assortative mating, gene flow, and introgression provided direct evidence of species boundaries in cases where delimitation involved parapatric taxa. (4) Comparisons with other congeneric taxa provided a framework for evaluating whether traits have diverged to the level of species status.

Finally, it is worth noting that many proposals submitted to the NACC are not written by authors of the original species delimitation studies on which a proposal is based. In the past decade, the NACC has made a concerted effort to engage the broader community in proposal submission. As a result, the number of proposals submitted by non-committee members—including study authors—has increased. However, more direct contact with study authors encouraging them to submit a proposal would be beneficial. Authors benefit by getting recognition for their work, which has value for broad stakeholders, and can cite NACC contributions on their curriculum vitae for professional advancement (this is especially important for students and early-career professionals). Likewise, the NACC and taxonomic revision process benefit by reviewing proposals that are written by the person most familiar with the case.

RECOMMENDATIONS AND CHALLENGES

Recommendations

Every species-delimitation study is unique and requires close consideration of the geography, ecology, phenotype, and behavior of the populations under consideration. Moreover, it also requires evaluation of gaps in knowledge, including areas where samples are lacking but suitable habitat occurs. However, some basic protocols can be established. To begin with, samples should combine material and data available from museum collections with new field-collected samples to fill in geographic and/or ecological gaps—especially across putative contact zones. Fortunately, online resources now provide ready access to information that can guide sampling strategies; these include (1) aggregators of voucher-based species data such as VertNet (http://vertnet.org), Integrated Digitized Biocollections (https://www.idigbio.org), the Global Biodiversity Information Facility (https://gbif.org), and Map of Life (https://mol.org); (2) phylogenetically based projects such as OpenWings (https://www.openwings.org), Open Tree of Life (https://tree.opentreeoflife.org), Tree of Life (http://tolweb.org), and VertLife (https://vertlife.org); (3) sound archives including Macaulay Library (https://www.macaulaylibrary.org) and xeno-canto (https://www.xeno-canto.org); and (4) citizen science initiatives that document observational occurrences including eBird (https://ebird.org) and iNaturalist (https://www.inaturalist.org), which complement voucher-based datasets (especially for GIS analyses) with the caveat that identification from observations may be difficult for some taxa or in contact zones. Although all of these efforts are indispensable, they cannot take the place of targeted field-based sampling. Researchers who conduct field work and collect specimens will have a firmer grasp of how

populations vary geographically in phenotype, behavior, and ecology compared to those who rely solely on existing resources for specimens and data.

Sampling for species delimitation studies should minimally include (1) high-quality tissues with associated voucher specimens that can be measured for phenotypic traits, and (2) locality information with coordinate data (latitude, longitude, datum, error) taken by GPS to pinpoint distributional occurrences for GIS data and ecological niche modeling. Moreover, sampling should capture the range of variation exhibited by populations from different environments, with denser sampling across putative contact zones to better capture evolutionary patterns and taxonomic limits. Value can be added to these samples through the integration of (3) audio recordings and playback experiments that show behavioral differences important in species recognition (McEntee 2015, Freeman and Montgomery 2017, Nwanko et al. 2018), although playback responses may be context-dependent (McEntee 2015); (4) other types of samples such as cloacal/buccal/ feather swabs, guts, or parasites to examine specificity in microbiomes (Hird et al. 2015) or hosts (Sari et al. 2012); (5) photographs and/or vegetation surveys to document habitat; (6) field notes that describe ecological and behavioral observations (Grinnell 1910, Herman 1986); and (7) plant and insect material collected as a control for stable isotope ratios obtained from specimens sampled at the same sites (Hobson 1999, English et al. 2018).

Researchers today are fortunate to have modern field and analytical equipment at their disposal and thus can address questions of species delimitation more comprehensively and definitively than several decades ago. However, even with these new tools, it is important to consider the fundamental principles outlined earlier. Accordingly, sampling strategies should be based on an assessment of known variation across populations (e.g., in phenotypic traits, ecology, behavior, or geography), with an exploration into previously unsampled areas to capture unknown variation and/or putative trait transitions. Although subspecies descriptions can help to guide this assessment and provide a framework for analysis, sampling must be objective and should not be dictated by previously established subspecies boundaries (Cicero 1996), which may or may not reflect biological reality (Winker and Haig 2010). Studies that involve short or long-distance migrants also must ensure that samples represent individuals from breeding populations in analyses of geographic variation (Cicero and Johnson 2006, Wang et al. 2019). Finally, researchers should think not only about the immediate question that they are investigating, but also about the legacy of making material available to the "student of the future" (Grinnell 1910) by properly archiving samples, specimens, photographs, recordings, and field notes into established collections. Likewise, journals

should require the deposition of vouchers for species delimitation as well as ecological and other kinds of studies (Salvador and Cunha 2020). This is essential for replicating data across time as well as space, either by incorporating new methods and/or examining temporal changes in species interactions, allele frequencies, and/or phenotype. For instance, significant insight has been gained into the evolutionary process by investigating changes in hybrid zone placement over time (Carling and Zuckerberg 2010, Krosby and Rohwer 2010, Taylor et al. 2014, Billerman et al. 2019, Wang et al. 2019). Although studies based on non-vouchered samples are useful and may be unavoidable in cases where taking vouchers is not possible (e.g., due to permitting or species status), they limit researchers' ability to reexamine data or assess new traits that were not part of the original study.

Challenges

The standards for species delimitation outlined here represent an ideal approach to this problem. However, there are various challenges that make it difficult to consistently apply these standards from a practical perspective. First is the challenge of obtaining all of the necessary permits for collecting samples and, if necessary, transporting them internationally. The permit landscape has become increasingly complex, with permits needed at different jurisdictional levels. In one case, the proposed split of eastern and western Bell's Vireo (*Vireo belli belli/V. b. medius* and *V. b. pusillus/V. b. arizonae*, respectively; NACC proposal 2017-C-14) failed in part because of a lack of samples from New Mexico where *V. b. medius* and *V. b. arizonae* come into contact. This proposal was based on strong molecular divergence as well as differences in plumage characters, morphology, behavior, song (qualitative), and ecological niche (Greaves et al. 2006, Klicka et al. 2016, Kus et al. 2020). Species delimitation, in this case, has important conservation implications because all 4 subspecies are state and/or federally threatened or endangered. According to comments on this proposal, permission was denied to collect any material from New Mexico, thus precluding the study of the contact zone. Furthermore, efforts to extract mtDNA from historic specimens were unsuccessful. Although permits are undoubtedly more difficult to obtain for species of conservation concern, and it may not be possible to collect in some areas (e.g., if there is political unrest), the complexity of navigating the permit process can be a deterrent to studies of even common species. Nonetheless, while molecular data from the putative contact zone of Bell's Vireos would be valuable, other types of data such as quantitative analysis of song and playback experiments in the contact zone also could be used as evidence for whether eastern and western Bell's Vireos should be considered separate species.

Another major challenge for studies of species delimitation is time and expertise. Comprehensive studies that involve field collecting, quantitative analyses of multiple traits (genomics, morphology, plumage, vocalizations), playback experiments, niche models, and other types of data all require a significant investment of time as well as expertise in many different methods. The latter can be solved through collaborations, but the amount of time it takes to do this type of work is often limiting—especially for students, early-career professionals, and others in positions with constraints on their research time. For example, the split of Sagebrush and Bell's sparrows (*Artemisiospiza nevadensis* and *A. belli*, respectively; Chesser et al. 2013) was based on museum specimens with tissues that were collected over a period of nearly 30 years. Likewise, our currently unpublished studies of Steller's Jays (*Cyanocitta stelleri*) and Spotted Towhees (*Pipilo maculatus*) have taken over 20 years to collect extensive material for analyses of species limits. While not all studies of species delimitation take this long to complete, and the length of time for a study depends on the scope and complexity of the problem, it can be a daunting consideration. For this reason, and because of academic incentive structures, students and early-career professionals may not see the value in pursuing species-level taxonomic work. Collaborations that divide the work as well as the expertise are increasingly common and are perhaps the best way to ensure that studies are completed in a timely but thorough manner.

Funding is a third major challenge for studies of species delimitation. The funding required for field and lab work can be prohibitive without a large grant to support this work, and funding outlets for basic studies are limited and increasingly competitive. Studies that employ "integrative taxonomy" approaches, and especially those that integrate novel data such as microbiomes and their connection to social behavior (Archie and Tung 2015) in the context of species limits, may be more likely to be funded. Genomic methods that get at the underlying mechanisms for reproductive isolation and species boundaries (Pulido-Santacruz et al. 2018, Cowles and Uy 2019) also provide exciting avenues for potential research funding. Apart from grants for particular studies, funding to support museum collections is also challenging but critically important. Natural history museums have been impacted heavily by insufficient and declining financial support, yet they provide the foundation for biodiversity research and for training the next generation of museum-based scientists (Suarez and Tsutsui 2004, Cook et al. 2014, Buerki and Baker 2016, Hiller et al. 2017, Bakker et al. 2020). Students and early career professionals interested in species-level work are at an advantage if they are affiliated with a museum collection, yet support for their training and research is often hampered by budget constraints (Hiller et al. 2017). Efforts aimed at increasing financial support for museum collections and collection-based training will have a host of benefits, from ensuring proper care for biological specimens (Nowogrodzki 2016) to supporting taxonomic research that is crucial for understanding and preserving biodiversity.

Despite the aforementioned challenges, this is an exciting time to be studying species delimitation in birds. Many questions remain, and methodological advances along with increased access to data enable new approaches to studying long-standing questions in avian taxonomy. In particular, innovative collections-based research, especially in the genomic era (Buerki and Baker 2016), provides new opportunities that foster cross-disciplinary studies. Species delimitation studies that examine multiple traits using both traditional and novel methods, and that are based on robust geographic and ecologic sampling, are most likely to stand the test of time.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Ornithology* online.

LITERATURE CITED

American Ornithologists' Union (1998). Check-list of North American Birds, seventh edition. American Ornithologists' Union, Washington, DC, USA.

Archie, E. A., and J. Tung (2015). Social behavior and the microbiome. Current Opinion in Behavioral Sciences 6:28–34.

Baier, F., and H. E. Hoekstra (2019). The genetics of morphological and behavioral island traits in deer mice. Proceedings of the Royal Society B 286:20191697.

Bakker, F. T., A. Antonelli, J. A. Clarke, J. A. Cook, S. V. Edwards, P. G. P. Ericson, S. Faurby, N. Ferrand, M. Gelang, R. G. Gillespie, et al. (2020). The Global Museum: Natural history collections and the future of evolutionary science and public education. PeerJ 8:e8225.

Bardwell, E., C. W. Benkman, and W. R. Gould (2001). Adaptive geographic variation in Western Scrub-Jays. Ecology 82:2617–2627.

Billerman, S. M., C. Cicero, R. C. K. Bowie, and M. D. Carling (2019). Phenotypic and genetic introgression across a moving woodpecker hybrid zone. Molecular Ecology 28:1692–1708.

Bowie, R. C. K., J. Fjeldså, S. J. Hackett, and T. M. Crowe (2004). Systematics and biogeography of Double-Collared Sunbirds from the Eastern Arc Mountains, Tanzania. The Auk 121:660–681.

Brelsford, A., and D. E. Irwin (2009). Incipient speciation despite little assortative mating: the yellow-rumped warbler hybrid zone. Evolution 63:3050–3060.

Brelsford, A., B. Milá, and D. E. Irwin (2011). Hybrid origin of Audubon's warbler. Molecular Ecology 20:2380–2389.

Buerki, S., and W. J. Baker (2016). Collections-based research in the genomic era. Biological Journal of the Linnean Society 117:5–10.

Campana, M. G., A. Corvelo, J. Shelton, T. E. Callicrate, K. L. Bunting, B. Riley-Gillis, F. Wos, J. DeGrazia, E. D. Jarvis, and R. C. Fleischer (2020). Adaptive radiation genomics of two ecologically divergent Hawai'ian honeycreepers: The 'akiapōlā'au and the Hawai'i 'amakihi. The Journal of Heredity 111:21–32.

Campbell, K. K., T. Braile, and K. Winker (2016). Integration of genetic and phenotypic data in 48 lineages of Philippine birds shows heterogeneous divergence processes and numerous cryptic species. PLoS One 11:e0159325.

Carling, M. D., and B. Zuckerberg (2010). Spatio-temporal changes in the genetic structure of the *Passerina* bunting hybrid zone. Molecular Ecology 20:1166–1175.

Chaves, A. V., C. L. Clozato, D. R. Lacerda, E. H. Sari, and F. R. Santos (2008). Molecular taxonomy of Brazilian tyrant-flycatchers (Passeriformes: Tyrannidae). Molecular Ecology Resources 8:1169–1177.

Chesser, R. T., R. C. Banks, F. K. Barker, C. Cicero, J. L. Dunn, A. W. Kratter, I. J. Lovette, P. C. Rasmussen, J. V. Remsen, Jr., J. D. Rising, et al. (2013). Fifty-fourth supplement to the American Ornithologists' Union Check-list of North American Birds. The Auk 130:558–571.

Chesser, R. T., M. L. Isler, A. M. Cuervo, C. D. Cadena, S. G. Galen, L. M. Bergner, R. C. Fleischer, G. A. Bravo, D. F. Lane, and P. A. Hosner. 2020. Conservative plumage masks extraordinary phylogenetic diversity in the *Grallaria rufula* (Rufous Antpitta) complex of the humid Andes. The Auk: Ornithological Advances 137:1–25.

Cicero, C. (1996). Sibling species of titmice in the *Parus inornatus* complex. University of California Publications in Zoology 128:1–217.

Cicero, C. (2004). Barriers to sympatry between avian sibling species (Paridae: *Baeolophus*) in local secondary contact. Evolution 58:1573–1587.

Cicero, C. (2010). The significance of subspecies: A case study of Sage Sparrows (Emberizidae, *Amphispiza belli*). In Avian Subspecies (K. Winker and S. M. Haig, Editors). Ornithological Monographs 67:103–113.

Cicero, C., and N. K. Johnson (1995). Speciation in sapsuckers (*Sphyrapicus*): III. Mitochondria-DNA sequence divergence at the cytochrome-*b* locus. The Auk 112:547–563.

Cicero, C., and N. K. Johnson (2006). Diagnosability of subspecies: Lessons from Sage Sparrows (*Artemisiospiza belli*) for analysis of geographic variation in birds. The Auk 123:266–274.

Cicero, C., and N. K. Johnson (2007). Narrow contact of desert Sage Sparrows (*Amphispiza belli nevadensis* and *A. b. canescens*) in Owens Valley, eastern California: Evidence from mitochondrial DNA, morphology, and GIS-based niche models. In Festschrift for Ned K. Johnson: Geographic Variation and Evolution in Birds (C. Cicero and J. V. Remsen, Jr., Editors). Ornithological Monographs 63: 78–95.

Cicero, C., and M. S. Koo (2012). The role of niche divergence and phenotypic adaptation in promoting lineage diversification in the Sage Sparrow (*Artemisiospiza belli*, Aves: Emberizidae). Biological Journal of the Linnean Society 107:332–354.

Cicero, C., N. A. Mason, R. A. Jiménez, D. R. Wait, C. Y. Wang-Claypool, and R. C. K. Bowie (2021). Data from: Integrative taxonomy and geographic sampling underlie successful species delimitation. Ornithology 138:1–15. https://doi.org/10.6078.D1Z69F

Clements, J. F., T. S. Schulenberg, M. J. Iliff, S. M. Billerman, T. A. Fredericks, B. L. Sullivan, and C. L. Wood (2019). The eBird/Clements Checklist of Birds of the World: v2019. https://www.birds.cornell.edu/clementschecklist/download

Cook, J. A., S. V. Edwards, E. A. Lacey, R. P. Guralnick, P. S. Soltis, D. E. Soltis, C. K. Welch, K. C. Bell, K. E. Galbreath, C. Himes, et al. (2014). Natural history collections as emerging resources for innovative education. BioScience 64:725–734.

Cowles, S. A., and J. A. C. Uy (2019). Rapid, complete reproductive isolation in two closely related *Zosterops* White-eye bird species despite broadly overlapping ranges. Evolution 73:1647–1662.

Cracraft, J. (1983). Species concepts and speciation analysis. Current Ornithology 1:159–187.

del Hoyo J., A. Elliot, J. Sargatal, D. A. Christie, and G. Kirwan (Editors) (2020). Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona, Spain. https://www.hbw.com

De Queiroz, K. (2007). Species concepts and species delimitation. Systematic Biology 56:879–886.

Derryberry, E. P. (2009). Ecology shapes birdsong evolution: Variation in morphology and habitat explains variation in white-crowned sparrow song. The American Naturalist 174:24–33.

Derryberry, E. P., J. N. Phillips, G. E. Derryberry, M. J. Blum, and D. Luther (2020). Singing in a silent spring: Birds respond to a half-century soundscape reversion during the COVID-19 shutdown. Science 370:575–579.

English, P. A., D. J. Green, and J. J. Nocera (2018). Stable isotopes from museum specimens may provide evidence of long-term change in the trophic ecology of a migratory aerial insectivore. Frontiers in Ecology and Evolution 6:14.

Espíndola, A., M. Ruffley, M. L. Smith, B. C. Carstens, D. C. Tank, and J. Sullivan (2016). Identifying cryptic diversity with predictive phylogeography. Proceedings of the Royal Society B 283:20161529.

Fišer, C., C. T. Robinson, and F. Malard (2018). Cryptic species as a window into the paradigm shift of the species concept. Molecular Ecology 27:613–635.

Freeman, B. G., and G. A. Montgomery (2017). Using song playback experiments to measure species recognition between geographically isolated populations: A comparison with acoustic trait analyses. The Auk: Ornithological Advances 134:857–870.

Freudenstein, J. V., M. B. Broe, R. A. Folk, and B. T. Sinn (2017). Biodiversity and the species concept—lineages are not enough. Systematic Biology 66:644–656.

Fujita, M. K., A. D. Leaché, F. T. Burbrink, J. A. McGuire, and C. Moritz (2012). Coalescent-based species delimitation in an integrative taxonomy. Trends in Ecology & Evolution 27:480–488.

Funk, D. J., and K. E. Omland (2003). Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 34:397–423.

Gill F., D. B. Donsker, and P. C. Rasmussen (Editors) (2020). IOC World Bird List (v10.1). doi:10.14344/IOC.ML.10.1

Giraudeau, M., P. M. Nolan, C. E. Black, S. R. Earl, M. Hasegawa, and K. J. McGraw (2014). Song characteristics track bill morphology along a gradient of urbanization in house finches (*Haemorhous mexicanus*). Frontiers in Zoology 11:83.

Gowen, F. C., J. M. Maley, C. Cicero, A. T. Peterson, B. C. Faircloth, T. C. Warr, and J. E. McCormack (2014). Speciation in Western Scrub-Jays, Haldane's rule, and genetic clines in secondary contact. BMC Evolutionary Biology 14:135.

Greaves, J., L. Chadwick, and J. Morlan (2006). Apparent eastern Bell's Vireo in San Francisco. Western Birds 37:119–122.

Grinnell, J. (1910). The methods and uses of a research museum. Popular Science Monthly 77:163–169.

Grinnell, J. (1915). A distributional list of the birds of California. Pacific Coast Avifauna 11:5–217.

Grinnell, J. (1917). The niche-relationships of the California Thrasher. The Auk 34:427–433.

Grinnell, J. (1935). A revised life-zone map of California. University of California Publications in Zoology 40:327–329.

Grinnell, J. (1943). Joseph Grinnell's Philosophy of Nature: Selected Writings of a Western Naturalist. University of California Press, Berkeley, CA, USA.

Hanna, Z. R., C. Cicero, and R. C. K. Bowie (2019). Molecular evidence that the Channel Islands populations of the orange-crowned warbler (*Oreothlypis celata*; Aves: Passeriformes: Parulidae) represent a distinct evolutionary lineage. PeerJ 7:e7388.

Harrison, R. G. (Editor) (1993). Hybrid Zones and the Evolutionary Process. Oxford University Press, New York, NY, USA.

Hebert, P. D., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis (2004). Identification of birds through DNA barcodes. PLoS Biology 2:e312.

Herman, S. (1986). The Naturalist's Field Journal: A Manual of Instruction Based on a System Established by Joseph Grinnell. Buteo Books, Vermillion, SD, USA.

Hiller, A. E., C. Cicero, M. J. Albe, T. L. W. Barclay, C. L. Spencer, M. S. Koo, R. C. K. Bowie, and E. A. Lacey (2017). Mutualism in museums: A model for engaging undergraduates in biodiversity science. PLoS Biology 15:e2003318.

Hird, S. M., C. Sánchez, B. C. Carstens, and R. T. Brumfield (2015). Comparative gut microbiota of 59 neotropical bird species. Frontiers in Microbiology 6:1403.

Hobson, K. A. (1999). Stable-carbon and nitrogen isotope ratios of songbird feathers grown in two terrestrial biomes: Implications for evaluating trophic relationships and breeding origins. The Condor 101:799–805.

Hubbard, J. P. (1969). The relationships and evolution of the *Dendroica coronata* complex. The Auk 86:393–432.

Hudson, E. J., and T. D. Price (2014). Pervasive reinforcement and the role of sexual selection in biological speciation. The Journal of Heredity 105:821–833.

Jacobsen, F., and K. E. Omland (2011). Increasing evidence of the role of gene flow in animal evolution: Hybrid speciation in the yellow-rumped warbler complex. Molecular Ecology 20:2236–2239.

Jiang, X., S. V. Edwards, and L. Liu (2020). The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. Systematic Biology 69:795–812.

Johnson, N. K. (1972). Origin and differentiation of the avifauna of the Channel Islands, California. The Condor 74:295–315.

Johnson, N. K. (1982). Retail subspecies—at least for the time being. The Auk 99:605–606.

Johnson, N. K., and C. Cicero (2002). The role of ecologic diversification in sibling speciation of *Empidonax* flycatchers (Tyrannidae): Multigene evidence from mtDNA. Molecular Ecology 11:2065–2081.

Johnson, N. K., and C. B. Johnson (1985). Speciation in sapsuckers (*Spyrapicus*): II. Sympatry, hybridization, and mate preference in *S. ruber daggetti* and *S. nuchalis*. The Auk 102:1–15.

Johnson, N. K., and J. A. Marten (1992). Macrogeographic patterns of morphometric and genetic variation in the Sage Sparrow complex. The Condor 94:1–19.

Johnson, N. K., J. V. Remsen, Jr., and C. Cicero (1999). Resolution of the debate over species concepts in ornithology: A new comprehensive biological species concept. In Proceedings of the 22 International Ornithological Congress, Durban (N. J. Adams and R. H. Slowtow, Editors). Birdlife South Africa, Johannesburg, South Africa.

Johnson, N. K., and R. M. Zink (1983). Speciation in sapsuckers (*Sphyrapicus*): I. Genetic differentiation. The Auk 100:871–884.

Karin, B. R., C. Cicero, M. S. Koo, and R. C. K. Bowie (2018). The role of history and ecology as drivers of song divergence in Bell's and Sagebrush sparrows (*Artemisiospiza*, Aves: Passerellidae). Biological Journal of the Linnean Society 125:421–440.

Klicka, L. B., B. E. Kus, P. O. Title, and J. J. Burns (2016). Conservation genomics reveals multiple evolutionary units within Bell's Vireo (*Vireo bellii*). Conservation Genetics 17:455–471.

Krosby, M., and S. Rohwer (2010). Ongoing movement of the hermit warbler X Townsend's warbler hybrid zone. PLoS One 5:e14164.

Kus, B., S. L. Hopp, R. R. Johnson, and B. T. Brown (2020). Bell's Vireo (*Vireo bellii*), version 1.0. In Birds of the World (A. F. Poole, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. doi:10.2173/bow.belvir.01

Linck, E., K. Epperly, P. Van Els, G. M. Spellman, R. W. Bryson, J. E. McCormack, R. Canales-Del-Castillo, and J. Klicka (2019). Dense geographic and genomic sampling reveals paraphyly and a cryptic lineage in a classic sibling species complex. Systematic Biology 68:956–966.

Luttrell, S. A. M., S. T. Gonzalez, B. Lohr, and R. Greenburg (2015). Digital photography quantifies plumage variation and salt marsh melanism among Song Sparrow (*Melosopiza melodia*) subspecies of the San Francisco Bay. The Auk: Ornithological Advances 132:277–287.

Manthey, J. D., and R. G. Moyle (2015). Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. Molecular Ecology 24:3628–3638.

Martin, S. H., and C. D. Jiggins (2017). Interpreting the genomic landscape of introgression. Current Opinion in Genetics & Development 47:69–74.

Mason, N. A., N. K. Fletcher, B. A. Gill, W. C. Funk, and K. R. Zamudio (2020). Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. Systematics and Biodiversity 18:269–280.

Mason, N. A., A. Olvera-Vital, I. J. Lovette, and A. G. Navarro-Sigüenza (2018). Hidden endemism, deep polyphyly, and repeated dispersal across the Isthmus of Tehuantepec: Diversification of the White-collared Seedeater complex (Thraupidae: *Sporophila torqueola*). Ecology and Evolution 8:1867–1881.

Mason, N. A., and S. A. Taylor (2015). Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. Molecular Ecology 24:3009–3025.

Mayr, E. (1940). Speciation phenomena in birds. The American Naturalist 74:249–278.

Mayr, E. (1942). Systematics and the Origin of Species. Columbia University Press, New York, NY, USA.

Mayr, E. (1957). Species concepts and definitions. In The Species Problem (E. Mayr, Editor). American Association for the Advancement of Science, Washington, DC, USA.

McEntee, J. (2015). Reciprocal territorial responses of parapatric African sunbirds: Species-level asymmetry and intraspecific geographic variation. Behavioral Ecology 25:1380–1394.

McKitrick, M. C., and R. M. Zink (1988). Species concepts in ornithology. The Condor 90:1–14.

Milá, B., T. B. Smith, and R. K. Wayne (2007). Speciation and rapid phenotypic differentiation in the yellow-rumped warbler *Dendroica coronata* complex. Molecular Ecology 16:159–173.

Milá, B., D. Toews, and T. B. Smith (2011). A cryptic contact zone between divergent mitochondrial DNA lineages in southwestern North America supports past introgressive hybridization in the Yellow-rumped Warbler complex (Aves: *Dendroica coronata*). Biological Journal of the Linnean Society 103:696–706.

Milá, B., R. K. Wayne, and T. B. Smith (2008). Ecomorphology of migratory and sedentary populations of the Yellow-rumped Warbler (*Dendroica coronata*). The Condor 110:335–344.

Moritz, C., and C. Cicero (2004). DNA barcoding: Promise and pitfalls. PLoS Biology 2:e354.

Nowogrodzki, A. (2016). Biological specimen troves threatened by funding pause. Nature 531:561.

Nwanko, E. C., C. T. Pallari, L. Hadjioannou, A. Ioannou, R. K. Mulwa, and A. N. G. Kirschel (2018). Rapid song divergence leads to discordance between genetic distance and phenotypic characters important in reproductive isolation. Ecology and Evolution 8:716–731.

Orr, H. A., and M. Turelli (2001). The evolution of postzygotic isolation: Accumulating Dobzhansky–Muller incompatibilities. Evolution 55:1085–1094.

Ottenburghs, J. (2019). Avian species concepts in the light of genomics. In Avian Genomics in Ecology and Evolution (R. Kraus, Editor). Springer, Cham, Switzerland.

Padial, J. M., A. Miralles, I. De la Riva, and M. Vences (2010). The integrative future of taxonomy. Frontiers in Zoology 7:16.

Patten, M. A., and J. V. Remsen (2017). Complementary roles of phenotype and genotype in subspecies delimitation. Journal of Heredity 108:462–464.

Pfennig, K. S. (2016). Reinforcement as an initiator of population divergence and speciation. Current Zoology 62:145–154.

Phillimore, A. B., and I. P. F. Owens (2006). Are subspecies useful in evolutionary and conservation biology? Proceedings of the Royal Society of London B 273:1049–1053.

Pigot, A. L., C. Sheard, E. T. Miller, T. P. Bregman, B. G. Freeman, U. Roll, N. Seddon, C. H. Trisos, B. C. Weeks, and J. A. Tobias (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. Nature Ecology & Evolution 4:230–239.

Pitelka, F. A. (1951). Speciation and ecological distribution in American jays of the genus *Aphelocoma*. University of California Publications in Zoology 50:195–464.

Poelstra, J. W., N. Vijay, C. M. Bossu, H. Lantz, B. Ryll, I. Müller, V. Baglione, P. Unneberg, M. Wikelski, M. G. Grabherr, and J. B. W. Wolf (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science 344:1410–1414.

Potvin, D. A., and S. M. Clegg (2015). The relative roles of cultural drift and acoustic adaptation in shaping syllable repertoires of island bird populations change with time since colonization. Evolution 69:368–380.

Price, T. D. ((2008). Speciation in Birds. Roberts & Company, Greenwood Village, CO, USA.

Pulido-Santacruz, P., A. Aleixo, and J. T. Weir (2018). Morphologically cryptic Amazonian bird species pairs exhibit strong postzygotic reproductive isolation. Proceedings of the Royal Society B 285:20172081.

Ramírez-Barrera, S. M., B. E. Hernández-Baños, J. P. Jaramillo-Correa, and J. Klicka (2018). Deep divergence of Red-crowned Ant Tanager (*Habia rubica*: Cardinalidae), a multilocus phylogenetic analysis with emphasis in Mesoamerica. PeerJ 6:e5496.

Raxworthy, C. J., C. M. Ingram, N. Rabibisoa, and R. G. Pearson (2007). Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. Systematic Biology 56:907–923.

Rheindt, F. E., T. Székely, S. V. Edwards, P. L. M. Lee, T. Burke, P. R. Kennerley, D. N. Bakewell, M. Alrashidi, A. Kosztolányi, M. A. Weston, et al. (2011). Conflict between genetic and phenotypic differentiation: The evolutionary history of a 'lost and rediscovered' shorebird. PLoS One 6:e26995.

Ribeiro, A. M., P. Lloyd, W. R. J. Dean, M. Brown, and R. C. K. Bowie (2014). The ecological and geographical context of morphological and genetic divergence in an understory-dwelling bird. PLoS One 9:e85903.

Roux, C., C. Fraisse, J. Romiguier, Y. Anciaux, N. Galtier, and N. Bierne (2016). Shedding light on the grey zone of speciation along a continuum of genetic divergence. PLoS Biology 14:e2000234.

Rundell, R. J., and T. D. Price (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. Trends in Ecology & Evolution 24:394–399.

Salvador, R. B., and C. M. Cunha (2020). Natural history collections and the future legacy of ecological research. Oecologia 192:641–646.

Sangster, G. (2014). The application of species criteria in avian taxonomy and its implications for the debate over species concepts. Biological Reviews 89:199–214.

Sangster, G. (2018). Integrative taxonomy of birds: The nature and delimitation of species. In Bird Species: How they Arise, Modify, and Vanish (D. T. Tietze, Editor). Springer, Cham, Switzerland.

Sari, E. H. R., H. Klompen, and P. G. Parker (2012). Tracking the origins of lice, haemosporidian parasites and feather mites of the Galapagos flycatcher (*Myiarchus magnirostris*). Journal of Biogeography 40:1082–1093.

Schlick-Steiner, B. C., F. M. Steiner, B. Seifert, C. Stauffer, E. Christian, and R. H. Crozier (2010). Integrative taxonomy: A multisource approach to exploring biodiversity. Annual Review of Entomology 55:421–438.

Smith, T. B., R. K. Wayne, D. J. Girman, and M. W. Bruford (1997). A role for ecotones in generating rainforest biodiversity. Science 276:1855–1857.

Stanton, D. W. G., P. Frandsen, R. K. Waples, R. Heller, I. M. Russo, P. A. Orozco-terWengel, C. Tingskov Pederson, H. R. Siegismund, and M. W. Bruford (2019). More grist for the mill? Species delimitation in the genomic era and its implications for conservation. Conservation Genetics 20:101–113.

Suarez, A. V., and N. D. Tsutsui (2004). The value of museum collections for research and society. BioScience 54:66–74.

Sukumaran, J., and L. L. Knowles (2017). Multispecies coalescent delimits structure, not species. Proceedings of the National Academy of Sciences USA 114:1607–1612.

Taylor, S. A., R. L. Curry, T. A. White, V. Ferretti, and I. Lovette (2014). Spatiotemporally consistent genomic signatures of reproductive isolation in a moving hybrid zone. Evolution 68:3066–3081.

Tennessen, J. A., and K. R. Zamudio (2008). Genetic differentiation among mountain island populations of the Striped Plateau Lizard, *Sceloporus virgatus* (Squamata: Phrynosomidae). Copeia 3:558–564.

Tobias, J. A., N. Seddon, C. N. Spottiswoode, J. D. Pilgrim, L. D. C. Fishpool, and N. J. Collar (2010). Quantitative criteria for species delimitation. Ibis 152:724–746.

Toews, D. P. L., A. Brelsford, C. Grossen, B. Milá, and D. E. Irwin (2016). Genomic variation across the Yellow-rumped Warbler species complex. The Auk: Ornithological Advances 133:698–717.

Toews D. P. L., G. R. Kramer, A. W. Jones, C. L. Brennan, B. E. Cloud, D. E. Anderson, I. J. Lovette, and H. Streby (2020). Genomic identification of intergeneric hybrids in New World wood-warblers (Aves: Parulidae). Biological Journal of the Linnean Society 131:183–191.

Toews, D. P. L., M. Mandic, J. G. Richards, and D. E. Irwin (2013). Migration, mitochondria, and the Yellow-rumped Warbler. Evolution 68:241–255.

Toews, D. P. L., J. Walsh, and L. Campagna (2018) Population genomics of birds: Evolutionary history and conservation. In Population Genomics (O. P. Rajora, Editor). Springer, Cham, Switzerland.

Venkatraman, M. X., D. A. Deraad, W. L. E. Tsai, E. Zarza, A. J. Zellmer, J. M. Maley, and J. E. McCormack (2019). Cloudy with a chance of speciation: Integrative taxonomy reveals extraordinary divergence within a Mesoamerican cloud forest bird. Biological Journal of the Linnean Society 126:1–15.

Walsh, J., I. J. Lovette, V. Winder, C. S. Elphick, B. J. Olsen, G. Shriver, and A. J. Kovach (2017). Subspecies delineation amid phenotypic, geographic and genetic discordance in a songbird. Molecular Ecology 26:1242–1255.

Wang, X., P. Que, G. Heckel, J. Hu, X. Zhang, C. Chiang, N. Zhang, Q. Huang, S. Liu, J. Martinez, et al. (2019). Genetic, phenotypic, and ecological differentiation suggests incipient speciation in two *Charadrius* plovers along the Chinese coast. BMC Evolutionary Biology 19:135.

Wang, S., S. Rohwer, K. Delmore, and I. Irwin (2019). Cross-decades stability of an avian hybrid zone. Journal of Evolutionary Biology 32:1242–1251.

Wiens, J. A. (1982). Forum: Avian subspecies in the 1980's. The Auk 99:593.

Wiens, J. J. (2007). Species delimitation: New approaches for discovering diversity. Systematic Biology 56:875–878.

Winker, K. (2009). Reuniting phenotype and genotype in biodiversity research. BioScience 59:657–665.

Winker, K. and S. M. Haig (Editors) (2010). Avian Subspecies. Ornithological Monographs, no. 67. American Ornithologists' Union, Washington, DC, USA.

Withrow, J. J., and K. Winker (2014). Genetics of a high-latitude cryptic speciation event: American and Pacific Golden-Plovers. The Wilson Journal of Ornithology 126:429–442.

Yang, Z., and B. Rannala (2010). Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences USA 107:9264–9269.

Young, J. R., C. E. Braun, S. J. Oyler-McCance, J. W. Hupp, and T. W. Quinn (2000). A new species of Sage-Grouse (Phasianidae: *Centrocercus*) from southwestern Colorado. The Wilson Journal of Ornithology 112:445–453.

Zink, R. M. (2004). The role of subspecies in obscuring avian biological diversity and misleading conservation policy. Proceedings of the Royal Society of London B 272:71–78.

Zink, R. M., and M. C. McKitrick (1995). The debate over species concepts and its implications for ornithology. The Auk 112:701–719.