The genomic revolution and species delimitation in birds (and other organisms): gene flow matters, but what about other evolutionary forces?

Carlos Daniel Cadena[1] **and Felipe Zapata**[2]

[1]*Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia.* Email: ccadena@uniandes.edu.co

[2]*Department of Ecology and Evolutionary Biology, University of California, Los Angeles.*

Abstract

Given the notion that species are population-level lineages and the availability of genomic data to identify separately evolving populations, researchers usually establish species limits based on gene flow or lack thereof. A strict focus on gene flow as the main –or only– criterion to delimit species involves two main complications in practice. First, approaches often used to apply this criterion to genome-wide data cannot by themselves distinguish species limits from within-species population structure, particularly in allopatric organisms. Second, recognizing as species only those lineages one can identify using such approaches fails to embrace the role of other evolutionary forces (i.e. various forms of selection) in defining evolutionary lineages. Using examples from various groups of birds, we call for the importance of considering evolutionary forces additional to gene flow in species delimitation and explain why genomic approaches commonly used in taxonomic studies may be insufficient by themselves to properly uncover species limits. By considering the processes that structure genotypic and phenotypic variation during speciation, we argue that rigorous analyses of phenotypic variation remain crucial for species delimitation in the genomics era because phenotypes uniquely inform us about the role of selection maintaining the cohesion of evolutionary lineages. Evolutionary theory describing the roles of gene flow, genetic drift and natural selection in the origin and maintenance of species calls for an integration of genomics with phenomics in avian species delimitation.

Key words: cohesion mechanism, demographic exchangeability, genetic exchangeability, integrative taxonomy multispecies coalescent, nature of species, syngameon.

Introduction

Few issues in systematics and evolutionary biology have been as contentious as what species are and which criteria should researchers use to delimit species. Ornithologists have been active participants in debates on the theory and practice of species delimitation, with advocates of reproductive isolation (Mayr 1963; Johnson et al. 1999) and diagnosability (Cracraft 1983; Zink and McKitrick 1995) as criteria for species recognition being particularly influential. Part of the controversies have settled based on the idea that one should separate the ontological question of what species are from the epistemological matter of how one goes about recognizing species (Hey 2006). By focusing on the issue of what kind of entities are species, and by examining elements common to all species definitions, de Queiroz (1998) formulated the generalized lineage species concept which defines species as segments of separately evolving metapopulation-level lineages (see also de Queiroz 2007).

With the adoption of the notion that species are population-level lineages and the practical availability of genomic data to identify separately evolving populations, researchers increasingly establish species limits based on gene flow or lack thereof. This is explicitly embodied in the use of the multispecies coalescent as a criterion for species delimitation, which has become popular over recent years (Yang and Rannala 2010; Fujita et al. 2012). We see two main complications associated with increasing reliance on this approach. First, lineages identified using genomic data analyzed with coalescent methods need not represent species but may solely reflect within-species population structure (Barley et al. 2017; Sukumaran and Knowles 2017; Leaché et al. 2018). Accordingly, coalescent analyses of genomic data often recover distinct lineages among organisms many researchers would consider conspecific including populations of *Drosophila* subject to intensive speciation research (Campillo et al. 2019) and even humans (Jackson et al. 2017).

A second complication relates to the fact that with ever increasing amount and quality of genomic data employed to address questions about patterns of genetic structure and to infer species limits, it has become apparent that species readily diagnosable based on phenotypic traits which biologists typically use to distinguish species in nature do not always form distinct genomic clusters (Mallet et al. 2016; Wang et al. 2019). Many researchers view such results as evidence that species limits require revision, with some even questioning the reality of species in particular groups (McKay and Zink 2015; Zink and Vázquez-Miranda 2019). Some authors have gone as far to suggest redefining the species concept altogether based on genomics (Jarvis 2016). However, as we shall describe below, the second complication is that there are biologically plausible scenarios in which good species (i.e. populations most of us would agree are separately evolving lineages) may not be distinguishable with approaches frequently used by genomicists.

Both complications above imply that genomic data as often analyzed in species delimitation studies cannot be used in isolation from other sources of information to establish species limits even if one favors gene flow as a central criterion for species delimitation. More broadly, recognizing species only as those lineages one can identify using approaches to analyze genome-wide variation may be problematic because it fails to embrace the role that other evolutionary forces – especially natural selection– play in defining evolutionary lineages (Van Valen 1976; Templeton 1989; Coyne and Orr 2004; Barraclough 2019). In particular, such a view is inconsistent with a genic perspective of speciation in which individual genes subject to selection and not complete genomes are the units underlying the differentiation of evolutionary lineages (Wu 2001).

We examine conceptual issues related to species delimitation using genetic data often overlooked by researchers interested in species limits in ornithology in the genomic era. We focus on evolutionary forces involved in maintaining species separate regardless of gene flow because we believe that information on the nature of species (i.e. on the processes that structure genotypic and phenotypic variation during speciation) should guide species delimitation in practice. This paper complements a recent review on avian species concepts in the light of genomics based on the criteria of diagnosability, phylogeny, and reproductive isolation (Ottenburghs 2019). Both that review and our perspective convey the message that genomic data should be integrated with –yet cannot replace– other types of data in species diagnoses (see also Winker 2009). In particular, we stress that there are firm theoretical bases implying that assessments of phenotypic variation remain central to studying species limits in the genomic era. While our focus is on birds, the ideas we discuss are applicable to a wide range of organisms.

Are all these lineages species?

Many ornithologists favor the view that species of birds are reproductively isolated populations (Mayr 1963; Johnson et al. 1999). However, the most common practice in avian taxonomy has been to delimit species based on diagnosability in phenotype and not on direct assessments of reproductive compatibility (Sangster 2014), partly because such assessments are impractical for allopatric populations. The long tradition of museum-based taxonomy focusing on plumage characters was outgrown by a new paradigm in which variation in traits presumed or known to be involved in reproductive isolation –notably vocalizations– is given precedence in species delimitation (Remsen 2005). Over recent years, information on morphological and behavioral phenotypes has been increasingly integrated with genetic data, allowing for comprehensive analyses of avian species limits, particularly in the Neotropics (e.g. Cadena and Cuervo 2010; Isler et al. 2012; Isler et al. 2020; Krabbe et al. 2020). However, the significance of patterns of variation revealed by genetic data for species delimitation is not always straightforward for reasons we describe with empirical examples below. We focus on cases employing methods based on the multispecies coalescent to analyze genetic variation given their increasing application and the implications of their use for areas in avian biology beyond taxonomy such as speciation and biogeography (Andersen et al. 2015; Smith et al. 2018).

The multispecies coalescent (MSC) uses genealogical histories of putatively neutral regions across the genome to infer the history of populations as reflected in parameters such as modern and ancestral populations sizes, divergence times, and past migration rates between populations (Rannala and Yang 2003). This model is grounded on well-established principles of population genetics (Kingman 1982). It states that under high effective migration rates, populations do not achieve genetic isolation and thus represent a single panmictic species (Wright 1931). In contrast, when populations experience low effective migration, genetic isolation is plausible and thus distinct species can be recognized (Yang and Rannala 2010).

Consider a case in which researchers sample broadly across geography to characterize patterns of genetic variation at a single locus (e.g. mtDNA) in a widespread group of birds, and then analyze data using an MSC approach for species delimitation. Such analysis reveals multiple distinct lineages, indicating that the number of species in the group might be more than 10 times greater than traditionally thought. This is precisely what a study of the Grey-breasted Wood-wren complex (*Henicorhina leucoprhys,* Troglodytidae) found. Current taxonomy recognizes three species in the group, yet coalescent analyses of mtDNA data collected across mountains from Mexico to Bolivia revealed that the complex consists of at least 35 distinct lineages (Cadena et al. 2019). Should all these lineages be recognized as species?

Because the MSC detects population structure and not species, lineages identified using this approach are best considered presumptive species, the limits of which ought to be tested with additional data (Sukumaran and Knowles 2017). In the case of *H. leucophrys*, a few distinct mtDNA lineages come into geographic contact, a rare opportunity to test whether they are indeed separately evolving populations worthy of species status. Results of such tests allowed by sympatry (Mayr 1963) are mixed: while some mtDNA lineages in the group are reproductively isolated from others as indicated by behavior and differentiation in phenotype and nuclear markers (Salaman et al. 2003; Caro et al. 2013), others seem to experience extensive gene flow (Halfwerk et al. 2016). Therefore, the question of how many species of grey-breasted wood-wrens should one recognize remains up for grabs given the data available.

That wood-wrens with highly divergent mtDNA are not reproductively isolated likely reflects that barriers to genetic exchange of presumably neutral regions of the genome do not accumulate predictably over time (Roux et al. 2016), notwithstanding that in some cases mitonuclear interactions may result in isolation (Hill et al. 2019). In practice, the wood-wren example illustrates that limits of lineages identified using a single locus are not reliable surrogates of species limits, but we of course have known this for years (Edwards et al. 2005). Are multilocus data sets the solution?

A natural step to take given problems with single-locus species delimitation is to extend sampling to multiple regions in the genome. Nuclear markers have not been extensively assayed in wood-wrens, but research on a similar case, that of the Rufous Antpitta complex (*Grallaria rufula,* Grallariidae), provides some insight. As with wood-wrens, antpittas in this group exhibit remarkably strong population structure in the montane Neotropics, with mtDNA data analyzed using an MSC approach revealing 16-17 distinct lineages, the majority of which are allopatric (Chesser et al. 2020). Again, the question is how many of these lineages should be considered distinct species. Adding sequences of three nuclear introns to the mtDNA data produced remarkably congruent results: almost all of the mtDNA lineages were also distinct in nuclear loci (Chesser et al. 2020). Furthermore, because lineages identified using genetic data corresponded tightly with vocally distinct populations, the integration of molecular and vocal data sets resulted in a revised taxonomy recognizing 16 different species (Isler et al. 2020). Reaching a seemingly robust conclusion about species limits in the *G. rufula* complex was no easy task; integrative taxonomic analyses took decades to complete, with input from many researchers and contributors in multiple countries.

Given that implementing efforts similar to those in *G. rufula* (Chesser et al. 2020; Isler et al. 2020) and other groups (e.g. the *Scytalopus magellanicus* complex; Krabbe et al. 2020) across all birds appears hardly feasible, researchers may have the hope that genomics will expedite avian species delimitation (Jarvis 2016). Such hope is not unfounded because ornithology entered a new era over the past few years, in which various questions we used to address with one locus or a handful of loci can now be tackled with genome-wide markers assayed with high-throughput sequencing (Toews et al. 2016a). However, despite widespread application of such approaches to assess species limits in other vertebrate groups like amphibians and reptiles (Hillis 2019), they have seldom been implemented in analyses of avian species delimitation. Nonetheless, empirical work on other taxa as well as simulations indicate that coalescent approaches for species delimitation with multilocus data may also oversplit species by recognizing allopatric populations as distinct lineages (Leaché et al. 2018). This implies that increasing the number of loci assayed need not solve the issue of too many lineages identified by single markers like mtDNA.

Of course, we do not deny the enormous power of genomics to aid species delimitation in birds, allowing the elucidation of species boundaries not observable with few genetic markers. For example, phenotypically distinct hummingbirds from the northern Andes in the genus *Coeligena* are not distinguishable with mitochondrial genomes nor with variable regions flanking nuclear ultraconserved elements (Palacios et al. 2019; Palacios et al. in review), yet complete genomes clearly separate taxa in line with traditional plumage-based taxonomy (Palacios et al. unpubl. data). However, in *Coeligena* and other similar cases (Campagna et al. 2017; Aguillon et al. 2018), the existence of distinct species was already clear because discrete phenotypes occur in sympatry, an issue we shall return to in detail below. It follows, then, that arguably the real promise of genomics in avian species delimitation lies in its potential to detect previously unknown lineages, i.e. cryptic species (Fišer et al. 2018). The best avian example of such potential we are aware of involves Darwin's finches (Thraupidae), in which complete genomes revealed that morphologically similar forms in the genus *Geospiza* from different islands in the Galápagos are not each other's closest relatives (Lamichhaney et al. 2015). Such result prompted reconsideration of phenotypic and behavioral traits, which led to recognition of additional species in the group (Remsen et al. 2018).

The iconic Darwin's finches exemplify the power of genomics to reveal unrecognized species, but as we show later on in this paper they also serve to illustrate a limitation of the standard genomics toolkit. Namely, genomic approaches usually employed by avian taxonomists may often fail to detect boundaries among distinct evolutionary lineages. We turn to this crucial yet often underappreciated issue in the following section.

"Good species" overturned by genomics?

Recent genomic analyses of various organisms have revealed that multiple traditionally defined species are less genetically differentiated from others than one may have expected under the premise that species-level lineages are kept separate from others owing to lack of gene flow (Mallet et al. 2016; Wang et al. 2019). Examples of phenotypically distinct organisms exhibiting extensive gene flow with others as evidenced in reticulate gene genealogies now abound across the tree of life, including various cases in groups like plants (Novikova et al. 2016; Hipp et al. 2019), invertebrates (Martin et al. 2013; Fontaine et al. 2015), and vertebrates (Cui et al. 2013; Kumar et al. 2017; Gopalakrishnan et al. 2018; Rogers et al. 2019; Barth et al. 2020).

There are two possible interpretations accounting for phenotypically distinct yet genomically admixed and thereby molecularly undiagnosable "species". One is that traditional means to identify evolutionary lineages have often failed, mistreating within-species phenotypic variation as if it were evidence of species limits. Under this view, evidence of extensive gene flow revealed by genomic analyses is evidence of conspecifity. Alternatively, equating species only with lineages one can identify using tools to analyze genomic variation like the MSC and various clustering approaches may be problematic because it fails to recognize the role that other evolutionary forces, especially natural selection, play in defining evolutionary lineages. We illustrate this latter interpretation based on case studies of birds.

In scenarios where differences between species exist at few loci of critical phenotypic effect while much of the genome is undifferentiated or shows evidence of admixture, methods used to detect genomic clusters which many researchers may favor as *prima facie* evidence of distinct species may often be unsuccessful in uncovering species limits. For example, two European crow (Corvidae) species differing distinctly in plumage coloration are largely genetically homogeneous owing to genome-wide introgression; however, differences in regulatory regions influencing the expression of genes involved in the melanogenesis pathway maintain phenotypically distinct lineages in the face of gene flow (Poelstra et al. 2014). Likewise, species of *Vermivora* wood-warblers (Parulidae) are indistinguishable across most of their genomes, with marked phenotypic differences residing in a handful of genes influencing plumage traits important for species recognition and likely targets of sexual selection (Toews et al. 2016b). Another example are redpolls in the genus *Acanthis* (Fringillidae), in which the coexistence of three forms differing in plumage and bill morphology suggests more than one species may be involved yet a variety of analyses of genomic data obtained using a reduced-representation approach failed to detect distinct genetic clusters (Mason and Taylor 2015). Such result could well be interpreted as evidence of a single polymorphic lineage bound together by gene flow. Intriguingly, however, transcriptomes revealed that polygenic patterns of gene expression correlate with redpoll phenotypes in sympatry, suggesting that differences among forms in plumage and bill morphology may be controlled by regulatory elements and are potentially maintained by selection in the face of genetic exchange (Mason and Taylor 2015). This latter result is interpretable as evidence of species boundaries.

Other scenarios in which good species may not be easily distinguished using genomic markers are those in which speciation may have proceeded in the face of gene flow, as evidenced by historical introgression events observable in reticulate genealogies. Recently studied avian examples include *Catharus* thrushes across the Americas (Everson et al. 2020) and *Dendrocincla* woodcreepers in Amazonia (Pulido-Santacruz et al. 2020). The question of how much gene flow between named species are taxonomists willing to tolerate while maintaining their status as distinct taxonomic entities is not new, yet it remains a challenge even with genomic information and tools to analyze such data. Back to the MSC framework, determining the threshold of gene flow below which species can be detected is critical and a matter of unresolved debate (Zhang et al. 2011; Jackson et al. 2017; Leaché et al. 2018). Regardless of whether and how such debate is settled, widespread gene exchange evidently limits the ability of the MSC to detect distinct lineages (but see Flouri et al. 2019). Other approaches to analyzing genomic data such as clustering analysis based on principal components (Patterson et al. 2006) or Structure (Pritchard et al. 2000) are also prone to miss detecting species when gene flow between lineages is high (Sorenson in review).

The examples above are not meant to imply that clusters corresponding to species will not be recovered in all cases in which differences among bird species may have a simple genetic basis amid largely undifferentiated or introgressed genomes (Campagna et al. 2017; Stryjewski and Sorenson 2017; Aguillon et al. 2018). However, they serve to make the point that some analyses of genomic data by themselves may be insufficient tests of species limits, particularly if one accepts that gene flow is not the only force setting the limits of separately evolving lineages (Van Valen 1976; Templeton 1989; Barraclough 2019).

Gene flow matters, but what about other evolutionary forces?

We have no interest in arguing in favor or against any particular species definition, but we find that the rationale underlying the “cohesion species concept” (Templeton 1989) is a useful theoretical framework to guide thinking about the nature of species and the forces maintaining them apart. Moreover, such framework provides insight into how different types of data (i.e. genomes, phenotypes) can best be brought to bear on practical issues involving species delimitation in birds and other organisms. Templeton (1989) was quite correct when he argued that species definitions based on reproductive isolation and recognition "*elevated a single micro evolutionary force –gene flow– into the conclusive and exclusive criterion for species status"*, while failing to consider that "*... gene flow is not the only micro evolutionary force that defines the boundaries of an evolutionary lineage"* and that "*genetic drift and natural selection play a far more potent and universal role..."*. In this section we describe how research programs with a narrow focus on gene flow restrict studies of species limits because they overlook other evolutionary forces involved in the origin and maintenance of species boundaries.

Why do species exist? From a population genetics standpoint, species of birds and other sexually reproducing organisms may exist as separately evolving entities (1) because reproduction maintains coherence within lineages and breeding isolation causes divergence among lineages, or (2) because divergent selection prevents lineages from merging (Templeton 1981, 1989; Coyne and Orr 2004; Barraclough 2019). Work employing genetic data to delimit species of birds has largely focused on point (1), often missing to consider point (2). Such focus on genetic isolation (i.e. on the assessment of gene flow or lack thereof) in species delimitation as ornithology moves into the genomics era fails to fully connect taxonomic practice with evolutionary theory. As summarized by Barraclough (2019), genetic analyses of single loci allow one to identify species as separate arenas for genetic drift, while analyses of multilocus markers point at species as separate arenas for reproductive isolation. However, focusing only on these perspectives hampers the assessment of species as separate arenas for natural selection, which requires a shift towards studying phenotypic traits and their genetic bases (Barraclough 2019).

Templeton (1989) conceptualized species as groups of organisms having the potential for genetic or demographic exchangeability via cohesion mechanisms. Genetic exchangeability refers to factors limiting the spread of genetic variants through gene flow and is appropriately studied with data allowing one to determine whether organisms share fertilization and developmental systems promoting genetic identity relative to other groups of organisms with which there is no genetic exchange (Templeton 1989). For this task, tools designed to assess gene flow such as the MSC applied to the genomic datasets ornithologists now routinely collect (Toews et al. 2016a) are well suited. On the other hand, demographic exchangeability refers to factors defining ecological niches and limiting the spread of genetic variants via genetic drift and natural selection (Templeton 1989). The signature of genetic drift promoting genetic identity of organisms sharing common ancestors is also properly studied using population-genetic and tree-based tools to delimit lineages based on genome-wide markers (Sites and Marshall 2004; Fujita et al. 2012). In turn, assessing how natural selection sets the limits of lineages requires examining its role in promoting genetic identity by favoring the fixation of genetic variants and in promoting adaptations that affect demographic exchangeability (Templeton 1989; see also Stockman and Bond 2007; Bond and Stockman 2008). We argue below that approaches in the standard genomic toolbox applied to species delimitation do not allow one to fully address the crucial question of whether organisms are demographically exchangeable or not, particularly when the genetic basis of adaptations is unknown.

Our thesis is that to determine whether individuals are interchangeable with one another relative to factors which regulate population dynamics and thereby set the boundaries of lineages based on niche differences (Templeton 1989), researchers require information that genomic approaches by themselves can hardly provide. This not to say that genomes cannot give insights about demographic exchangeability (Harvey et al. 2019), but understanding what genomics might or might not allow requires considering the genetic basis of species differences.

Under species definitions regarding gene flow (or lack thereof) as the main evolutionary force setting species boundaries, isolating mechanisms are often thought to maintain the cohesiveness of species across their entire genomes. For example, Ernst Mayr referred to species as "protected gene pools", called for studying the "unity of the genotype", and emphasized the importance of "coadapted gene complexes" in speciation (Mayr 1963, 1970). Accordingly, he suggested (see also Dobzhansky 1937) that "harmful gene flow" would break up the network of interactions among genes which determine species identity, creating "disharmonious combinations" and thereby resulting in the collapse of species as evolutionary units (Mayr 1992). Such view has been repeatedly challenged with evidence that species boundaries may persist despite extensive gene flow owing to divergent selection acting on few loci. This implies that the identity of distinct lineages is often governed by a small fraction of the genome involved in adaptation, i.e. in determining whether individuals are demographically interchangeable (Templeton 1989). This idea supports a genic –as opposed to genomic– view of speciation (Wu 2001; Wang et al. 2019) as well as the metaphor of genomic islands of divergence underlying species boundaries (Turner et al. 2005), which have received substantial empirical support. Hence, because boundaries among many species in nature may be set by few genomic regions which resist the homogenizing effect of gene flow, analyses measuring genome-wide, largely neutral divergence may fail at detecting distinct species. Therefore, approaches for species delimitation not focusing exclusively on estimates of gene flow to identify lineages are required.

By explicitly considering the role that forces like gene flow, selection or changes in population size influencing genetic drift may play in speciation, process-based approaches to species delimitation using genomic data appear especially promising to connect the above theory with taxonomic practice (Smith and Carstens 2020). While such approaches continue to be developed and used, however, we point out that researchers already have at their disposal rich data and tools allowing for meaningful species delimitation given that (1) gene flow is not the only force setting the limits of evolutionary lineages and (2) we often lack information on loci underlying adaptations that alter demographic exchangeability. In the next section, we argue for the fundamental role of the quantitative study of phenotypes under this framework.

Phenotypic variation and statistical species delimitation

A century ago, R.A. Fisher (1918) formulated an evolutionary model which underlies the use of data on continuous phenotypic traits in studies of species delimitation. Under the assumptions of random mating and polygenic inheritance, the model states that phenotypic variation within a single species can be reasonably described by a normal distribution. Conversely, when a sample of organisms includes two or more species, phenotypic variation may be best described by two or more distinct normal distributions. Therefore, evidence of distinct normal distributions in a sample of individuals may allow one to infer that the sample comprises more than one species, provided that distinct normal distributions do not result from age- or sex-related polymorphisms, or from phenotypic plasticity within a species.

Theory in the Fisher (1918) model can be applied to delimit species using phenotypic data in various ways. For example, one may fit normal mixture models (NMMs; Pearson 1894; Fraley and Raftery 2002) to measurements of continuous traits to gauge the evidence supporting putative species boundaries defined *a priori* (Zapata and Jiménez 2012). Alternatively, one may use NMMs to determine the most probable number of species in a sample given phenotypic data through statistical model selection with no *a priori* information about groups (Edwards and Knowles 2014; Pearson and Ezard 2014; Cadena et al. 2018). Some examples of the use of approaches based on NMMs to inform species delimitation in birds using morphometric data involve *Henicorhina* wood-wrens replacing each other with elevation in the Sierra Nevada de Santa Marta, Colombia (Caro et al. 2013), migratory and sedentary Fork-tailed Flycatchers (*Tyrannus savana*) coexisting seasonally in the llanos of northern South America (Gómez-Bahamón et al. 2020), and Torrent Ducks (*Merganetta armata*) ranging along the Andes from Colombia to Argentina (Gutiérrez-Pinto et al. 2019).

When quantitative approaches described above detect distinct phenotypic groups (i.e., distinct normal distributions) among individuals occurring in allopatry, groups may plausibly represent phenotypically distinct populations within a species. In such cases, spatially explicit statistical tools may allow one to examine whether phenotypic variation is best explained as intra- or interspecific (Zapata and Jiménez 2012). On the other hand, when quantitative approaches to analyze continuous phenotypic traits detect distinct phenotypic groups in sympatry, species status for these groups is granted under a wide range of species definitions (de Queiroz 1998; Coyne and Orr 2004; Mallet 2008; Cadena et al. 2018). The power of such quantitative analyses is especially telling when other types of data, including complete genomes, seemingly provide no evidence of species boundaries in sympatry as described in the example that follows. The example further serves as evidence that phenotypes may reveal critical information about demographic exchangeability of individuals in populations connected by genetic exchange, an aspect central to the theory described above implying that species delimitation should embrace evolutionary mechanisms additional to gene flow.

Darwin's finches and phenotypes in species delimitation in the genomic era

The Darwin's finches of the Galápagos Islands are a textbook example of how new species originate and of the role of natural selection in clade diversification, with much of what we know about evolution in the group deriving from detailed work on ground-finches in the genus *Geospiza* (Grant and Grant 2008, 2014). Although ground-finches are well established models for the study of adaptation and speciation, their species level-taxonomy has been contentious. Controversy scaled to the point that the evidence supporting the existence of distinct species in the group was questioned because hybridization is extensive, species cannot be told apart with presumably neutral markers, and phenotypic traits (i.e. bill morphology) show seemingly continuous variation among individuals (Zink 2002). Furthermore, analyses based on complete genomes failed to recover all species as distinct genomic clusters (Lamichhaney et al. 2015). Given genetic and genomic data, various lines of evidence revealing hybridization, and patterns of morphological variation which seemed to imply no diagnosably distinct phenotypic groups existed, McKay and Zink (2015) advanced the hypothesis that phenotypic variation in ground-finches represented transient morphs within a species maintained by gene flow as a single lineage. A subsequent analysis highlighted topological incongruences among phylogenetic trees reconstructed using genomics (Lamichhaney et al. 2015; Lamichhaney et al. 2016; Lamichhaney et al. 2018) and interpreted such data as evidence that species limits in Darwin's finches require revision (Zink and Vázquez-Miranda 2019).

If gene flow were the only force setting the boundaries of evolutionary lineages and hence the most critical criterion for species delimitation, then we might agree with claims that there is a single species of *Geospiza* ground-finch. However, genome-wide evidence that fails to recover distinct clusters or clades corresponding to phenotypically defined species is an insufficient test of species limits in the group because it does not allow one to ascertain whether there are distinct lineages maintained separate from others by selection despite widespread genetic exchange (Nosil 2008; Pinho and Hey 2010). Phenotypes, however, can get at this very issue.

A reanalysis of morphometric data implementing NMMs as described above revealed there are multiple, phenotypically distinct groups of ground-finches across the Galápagos (Cadena et al. 2018). Most importantly, several of these groups coexist within islands. Given the assumption that continuous morphometric traits exhibit polygenic inheritance, the evidence of distinct phenotypic groups in sympatry points strongly towards the existence of multiple species (Fisher 1918; Coyne and Orr 2004; Mallet 2008). Because such species hybridize extensively as evidenced by behavioral work and by shared genetic variation across much of their genomes, it follows that species are most likely maintained as separate lineages by selection. Furthermore, the rich natural history data available for ground-finches reveals correspondence between form and function (Grant and Grant 2008, 2014). Bill morphology determines how organisms interact with food resources in a way implying that populations with distinct phenotypes are not demographically interchangeable because they have different niches (Templeton 1989; Bond and Stockman 2008), hence they are best considered different species. In sum, ground-finches serve as a prime example of the power of phenotypes to inform species delimitation even when comprehensive genomic data are at hand.

Based on the purported lack of genetic differences and dubious evidence of morphological differentiation among ground-finches available at the time, McKay and Zink (2015) suggested that speciation in the group had been repeatedly initiated but was never completed due to fluctuating selective pressures and extensive gene flow. They wittily coined the term "Sisyphean evolution" to describe such process in reference to Sisyphus, the character in Greek mythology condemned to eternally push a boulder up a mountain only to see it roll back down to where it started. In contrast, based on reanalyses of morphological data revealing distinct phenotypic groups in sympatry, Cadena et al. (2018) referred to evolution in the ground-finches as "Atlantean", signifying that despite genetic exchange, selection as evidenced by phenotypes imposes barriers keeping species as separately evolving entities just like the shoulders of Atlas prevent the Earth and the sky from merging. In observing that extensive evidence of gene flow across species boundaries need not undermine the idea of distinct lineages and in keeping with the tongue-in-cheek tradition of making reference to Greece, we concur with Mallet et al. (2016) in that species of ground-finches and other organisms "*...are like the Ship of Theseus in philosophy, which can progressively but almost completely be rebuilt with new wood, and yet remain the same ship*". Like pieces of wood in a ship, genes in a species may come and go while other evolutionary forces maintain species as separately evolving lineages.

Having said the above, we must clarify that phenotypically distinct species sharing much of their genomes owing to hybridization and introgression can indeed be distinguished using genomic approaches to assay genetic variation. Given that bill morphology in Darwin's finches is highly heritable (Keller et al. 2001), phenotypic differences among species should be observable in their genomes. Accordingly, distinct species and morphologically differentiated populations in the group evidently differ in regions of the genome that collectively control bill size and shape as demonstrated via comparative analyses of gene expression and development (Abzhanov et al. 2004; Abzhanov et al. 2006; Mallarino et al. 2011), reduced-representation approaches to assess genomic variation in relation to quantitative phenotypes (Chaves et al. 2016), and whole-genome resequencing (Lamichhaney et al. 2015; Lamichhaney et al. 2016; Lamichhaney et al. 2018). Thus, our point is not to critique the use of genomics to delimit species when there is gene flow, but rather to stress that analytical approaches seeking to detect clusters or clades using genome-wide data (e.g. Zink and Vázquez-Miranda 2019) are not appropriate tests of species limits in such cases.

Quantitative phenotypic analysis in avian species delimitation

We have emphasized the importance of quantitative analyses of phenotypic variation as a robust tool to inform studies of species limits. In doing so, we of course realize that quantitative approaches to examine phenotypes are not new to taxonomy in ornithology. For example, statistical approaches guided the development of frameworks establishing "rules" to recognize subspecies taxa based on morphometric traits (Amadon 1949; Patten 2010). Likewise, tests designed to assess phenotypic diagnosability (Isler et al. 1998) have been instrumental in advancing our knowledge of species limits based on variation in vocalizations (Remsen 2005). More recently, quantitative assessments of phenotypic variation based on an ordinal scaling system have been used as a heuristic to advance revision of species limits in birds by comparing the degree of differentiation between taxa of uncertain status with that between well-established species, with the latter serving as a "yardstick" to gauge evidence in favor of species status (Tobias et al. 2010). Moreover, most analyses of species limits in birds involve assessments of phenotypic distinctiveness (Sangster 2014), so what exactly are we calling for?

First, we call for the continued use of phenotypic data in analyses of species limits even as genomic data sets become ubiquitous. Second, our focus on the Fisher (1918) model as a conceptual basis for species delimitation and the practical implementation of approaches based on NMMs to analyze phenotypic data is deliberate because such a framework has potential to overcome limitations of other analytical approaches which ornithologists commonly use. In the interest of space, we defer readers to Cadena et al. (2018) for a detailed description of pitfalls and potential solutions, but note here that species delimitation using conventional approaches to examine phenotypic data are prone to problems related to (1) the use of graphical analyses conveying limited information on the frequency of phenotypes, (2) the reduction of dimensionality of phenotypic data sets resulting in the exclusion of informative characters, and (3) the consideration of statistical measures of central tendency (e.g. means, effect sizes) as indicators of phenotypic distinctiveness. Just as theory calls for considering phenotypic data in species delimitation as we argued above, theory should also inform which approaches to analyzing phenotypic variation are more appropriate given the framework we outlined.

Which traits matter? The nagging problem of allopatry, or the need to study the function of phenotypes

As we described above, evidence of phenotypically distinct groups in polygenic traits in sympatry is strong evidence of species limits, but distinct phenotypic groups in allopatry may or may not merit recognition as separate species. The treatment of allopatric populations is, of course, one of the major challenges faced by taxonomists, with ornithologists championing the idea that one can use phenotypic differences between sympatric species as a yardstick to assess the status of allopatric forms (Mayr 1969). Under this framework, treating allopatric populations as species is justified if their degree of divergence is at or beyond that of reproductively isolated taxa (Isler et al. 1998; Remsen 2005; Tobias et al. 2010). While this approach is often valuable, interpreting patterns of variation comparatively is not always straightforward because some allopatric populations may differ more than good species in some traits and less in others, and in many cases we lack information about which traits are involved in lineage isolation in sympatry and which simply diverge owing to lack of gene flow (i.e. we cannot distinguish causes and consequences of speciation; Cadena and Cuervo 2010).

An alternative to the yardstick approach to examine species limits in cases involving allopatric populations using phenotypic data is to simulate sympatry by conducting experiments in which individuals are presented with signals from allopatric populations and their response is compared to that exhibited towards their own local signals. Playback experiments using vocalizations have been extensively used in this regard (Lanyon 1978; Freeman and Montgomery 2017), with lack of response to foreign songs often considered strong evidence of species boundaries. This is justifiable because voices are often crucial for species recognition and speciation in birds (Price 2008), but extending such experiments to other dimensions of the phenotype is likely important for reasons we describe in the next example.

Passerine birds in the *Arremon torquatus* complex (Passarellidae) vary extensively in plumage, vocalizations, ecology, and molecular markers in the Neotropics, which led to a substantial revision of species limits (Cadena and Cuervo 2010). In such revision, the presence or absence of pectoral bands was one of the traits considered potentially important for species delimitation given its marked influence on the overall appearance of birds and because it seemed plausible that this plumage trait may play a role in species recognition. In a series of field experiments, however, males in both a population having (*A. basilicus*) and in a population lacking (*A. assimilis*) pectoral bands responded equally to specimen mounts with and without such bands (Avendaño and Cadena in review). This result implies that pectoral bands are likely unimportant for species recognition by males and therefore that hypotheses attributing a role for social selection via male-male interactions are unlikely drivers of patterns of geographic variation (Cadena et al. 2011). We note this because presence or absence of plumage signals like pectoral bands were not only considered informative for species delimitation in this particular group, but more generally are among the types of traits employed when assessing the degree of plumage differentiation of allopatric populations to inform species limits based on scoring systems under the yardstick approach (Tobias et al. 2010). Furthermore, it remains possible that pectoral bands do play a role in species recognition (and hence actually be informative about species limits) in *Arremon* if birds do not perceive them in isolation but rather integrate information from various traits in species recognition (Uy and Safran 2013). Therefore, we suggest that multimodal signaling (i.e. the joint use of plumage and voice) should be considered when examining whether or not particular traits may be informative about avian species limits. More generally, if we are to meaningfully use phenotypes to guide decisions about species limits in cases involving allopatric populations, then we need studies of the function of phenotypic traits traditionally employed in taxonomy.

The importance of thoughtfully selecting phenotypic traits for species delimitation does not only apply to cases involving allopatric populations. In the example of Darwin's ground-finches described above, the existence of distinct normal distributions in sympatry is relevant for species delimitation because body size and bill morphology are quantitative traits with polygenic inheritance (Abzhanov et al. 2004; Abzhanov et al. 2006; Lamichhaney et al. 2015; Chaves et al. 2016; Lamichhaney et al. 2016). However, bimodal or multimodal phenotypic distributions within a population need not imply more than one species is involved. For example, discrete bill morphologies associated with alternative food sources in *Pyrenestes* seed crackers in Africa are not evidence of distinct species but rather Mendelian polymorphism within populations (Smith 1993; vonHoldt et al. 2018). Likewise, bimodal distributions of bill size in Hook-billed Kites

(*Chondrohierax uncinatus*) in America likely reflect Mendelian traits under disruptive selection dictated by prey size (Smith and Temple 1982), while work on Snail Kites (*Rostrhamus sociabilis*) indicates that changes in bill morphology caused by changes in prey base reflect phenotypic plasticity (Cattau et al. 2018). The approaches for analyzing morphometric variation to infer species boundaries we described are not appropriate for such cases.

Conclusion

We echo calls for integrative taxonomy in which genomic and phenotypic data are considered on equal footing when delimiting species (Winker 2009; Padial et al. 2010; Chambers and Hillis 2020). If genotypes and phenotypes are examined together, then one may envision four possible scenarios, three of which need not cause much confusion (Winker 2009). First, if sympatric organisms are both genetically distinct and phenotypically distinct, researchers would most likely agree that more than one species is involved. Alternatively, if organisms do not differ genetically nor phenotypically, then they most likely belong to a single species. Finally, sympatric organisms differing genetically but lacking overt phenotypic differences are often considered distinct cryptic species (Fišer et al. 2018; Leaché et al. 2018).

Our review placed special emphasis on the remaining case, that in which organisms do not appear to differ genetically but are distinct in phenotype. This scenario may apply to pairs of populations and also to multiple groups of organisms which readily exchange genes via hybridization yet remain distinct in traits relevant for ecological or sexual isolation because selection counteracts the homogenizing effects of gene flow. Such "syngameons" (Lotsy 1925) occur widely across the tree of life (Seehausen 2004; Mallet et al. 2016), and given genomic patterns of variation one could well apply this concept to avian radiations including *Sporophila* capuchinos (Campagna et al. 2017), *Lonchura* munias (Stryjewski and Sorenson 2017) and of course Darwin's ground-finches (Lamichhaney et al. 2015). Given the now well-established insight that speciation and diversification are often spurred by introgressive hybridization involving loci under selection (Seehausen 2004; Marques et al. 2019; Gillespie et al. 2020), additional examples of species one may not be able tell apart based on neutral variation across the genome will continue to surface.

The inability to diagnose particular species using genomic approaches need not be transitory because not all species are expected to evolve along a similar trajectory in which the completion of speciation predictably leads to cessation of gene flow across the genome. Studies on birds indicate that hybrid incompatibilities shutting down opportunities for gene flow can take millions of years to accrue (Price 2008), whereas work on other organisms reveals that neutral genomic markers may be unable to readily distinguish species maintained as distinct phenotypic lineages by selection over ca. 10-20 million years of hybridization (Hipp et al. 2019; Barth et al. 2020).

We suggest that in cases where species seem not to differ genetically yet appear phenotypically distinct, researchers should turn to the issue of whether cohesion mechanisms maintaining lineages apart may be in place. In particular, given the framework we described, the question of whether organisms are demographically exchangeable (sensu Templeton 1989) is central to species delimitation (Bond and Stockman 2008). We have argued that analyses of phenotypic traits in their proper functional context is essential to adequately answer such question.

We close with a personal anecdote. Three years ago, while preparing to publish a perspective piece on the role of phenotypic data in species delimitation which involved reanalyses of morphological measurements of Darwin's finches (Cadena et al. 2018), we received a note from a prominent evolutionary ornithologist commenting on our work. We quote such note with the sole purpose of illustrating what we believe is a shared view among several researchers in the field:

> "*... I am not a fan of identifying species by statistical clustering of morphological data. I think examination of morphology was all evolutionary biologists could do fifty or 100 years ago, but now that we can sequence genes, we have key information on gene flow. The genomic data for Darwin's finches present strong evidence that we are not dealing with distinct species...*"
>
> (Anonymous, April 6, 2017).

We hope to have conveyed two central take-home messages germane to examining the above view and to reflecting on how we go about delimiting species more broadly. First, gene flow is not the only evolutionary force setting species boundaries and this must be incorporated in our thinking about best practices for species delimitation. Second, just as they were relevant decades or a century ago (Fisher 1918), rigorous analyses of phenotypic variation remain crucial for species delimitation in the genomics era because phenotypes uniquely inform us about the role of natural selection maintaining the cohesion of evolutionary lineages. Evolutionary theory describing the roles of gene flow, genetic drift and natural selection in the origin and maintenance of species calls for an integration of genomics with phenomics in species delimitation.

Epilogue, or the illusion of the genotype vs. phenotype dichotomy in species delimitation

A central thread through this paper is the argument that phenotypes should be considered alongside genotypes in studies of species limits, but to finalize we take a step back to reflect upon purported distinctions between genotypes and phenotypes and the implications for species delimitation. If one considers the phenotype of an organism as the totality of gene expression and metabolic networks modulated by responses to environmental cues (Wray 2013; Duncan et al. 2014; Dunn and Munro 2016), then knowing how variation in functional attributes of genomes gives rise to diversity in development, morphology, physiology, and other phenotypic dimensions is relevant to understanding the nature of species and thus delimiting them appropriately. As shown by some of the avian examples we discussed (e.g. Poelstra et al. 2014; Lamichhaney et al. 2015; Mason and Taylor 2015), evidence for species limits may be revealed by considering emerging properties of genomes and their interactions with the environment to produce phenotypes in ways not readily predictable from the gene sequences or SNP data that researchers interested in species limits typically collect. Progress towards a comprehensive comparative framework to investigate how gene flow, genetic drift, and selection shape functional genomic traits (Pease et al. 2016; Catalán et al. 2019; Smith et al. 2020) should enhance our understanding of genotype-phenotype associations underlying species boundaries and move us closer to a truly integrative approach to species delimitation.

Acknowledgements

We thank Pamela Rasmussen and Kevin Winker for their kind invitation to participate in the symposium and to contribute this paper. For helpful discussions on the theory and practice of species delimitation we thank Iván Jiménez. Seema Sheth and CDC's laboratory group provided helpful comments on the manuscript.

Literature Cited

Abzhanov, A., W. P. Kuo, C. Hartmann, B. R. Grant, P. R. Grant, and C. J. Tabin (2006). The calmudolin pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442:563-567.

Abzhanov, A., M. Protas, B. R. Grant, P. R. Grant, and C. J. Tabin (2004). *Bmp4* and morphological variation of beaks in Darwin's finches. Science 305:1462-1465.

Aguillon, S. M., L. Campagna, R. G. Harrison, and I. J. Lovette (2018). A flicker of hope: Genomic data distinguish Northern Flicker taxa despite low levels of divergence. Auk 135:748-766.

Amadon, D. (1949). The seventy-five per cent rule for subspecies. Condor 51:250-258.

Andersen, M. J., H. T. Shult, A. Cibois, J.-C. Thibault, C. E. Filardi, and R. G. Moyle (2015). Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (Aves: Alcedinidae: *Todiramphus*). Royal Society Open Science 2:140375.

Avendaño, J. E., and C. D. Cadena (in review). Territorial males do not discriminate between local and novel plumage phenotypes in a tropical songbird species complex: implications for the role of social selection in signal evolution.

Barley, A. J., J. M. Brown, and R. C. Thomson (2017). Impact of model violations on the inference of species boundaries under the multispecies coalescent. Systematic Biology 67:269-284.

Barraclough, T. G. (2019). The evolutionary biology of species. Oxford University Press, Oxford, UK.

Barth, J. M. I., C. Gubili, M. Matschiner, O. K. Tørresen, S. Watanabe, B. Egger, Y.-S. Han, E. Feunteun, R. Sommaruga, R. Jehle, and R. Schabetsberger (2020). Stable species boundaries despite ten million years of hybridization in tropical eels. Nature Communications 11:1433.

Bond, J. E., and A. K. Stockman (2008). An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. Systematic Biology 57:628-646.

Cadena, C. D., Z. A. Cheviron, and W. C. Funk (2011). Testing the molecular and evolutionary causes of a 'leapfrog' pattern of geographical variation in coloration. Journal of Evolutionary Biology 24:402-414.

Cadena, C. D., and A. M. Cuervo (2010). Molecules, ecology, morphology, and songs in concert: How many species is "*Arremon torquatus*" (Aves, Emberizidae)? . Biological Journal of the Linnean Society 99:152-176.

Cadena, C. D., J. L. Pérez-Emán, A. M. Cuervo, L. N. Céspedes, K. L. Epperly, and J. T. Klicka (2019). Extreme genetic structure and dynamic range evolution in a montane passerine bird: implications for tropical diversification. Biological Journal of the Linnean Society 126:487-506.

Cadena, C. D., F. Zapata, and I. Jiménez (2018). Issues and perspectives in species delimitation using phenotypic data: Atlantean evolution in Darwin's finches. Systematic Biology 67:181-194.

Campagna, L., M. Repenning, L. F. Silveira, C. S. Fontana, P. L. Tubaro, and I. J. Lovette (2017). Repeated divergent selection on pigmentation genes in a rapid finch radiation. Science Advances 3:e1602404.

Campillo, L. C., A. J. Barley, and R. C. Thomson (2019). Model-based species delimitation: Are coalescent species reproductively isolated? Systematic Biology.

Caro, L. M., P. C. Caycedo-Rosales, R. C. K. Bowie, H. Slabbekoorn, and C. D. Cadena (2013). Ecological speciation along an elevational gradient in a tropical passerine bird? Journal of Evolutionary Biology 26:357-374.

Catalán, A., A. D. Briscoe, and S. Höhna (2019). Drift and directional selection are the evolutionary forces driving gene expression divergence in eye and brain tissue of *Heliconius* butterflies. Genetics 213:581-594.

Cattau, C. E., R. J. Fletcher Jr, R. T. Kimball, C. W. Miller, and W. M. Kitchens (2018). Rapid morphological change of a top predator with the invasion of a novel prey. Nature Ecology & Evolution 2:108-115.

Chambers, E. A., and D. M. Hillis (2020). The multispecies coalescent over-splits species in the case of geographically widespread taxa. Systematic Biology 69:184-193.

Chaves, J. A., E. A. Cooper, A. P. Hendry, J. Podos, L. F. De León, J. A. M. Raeymaekers, W. O. MacMillan, and J. A. C. Uy (2016). Genomic variation at the tips of the adaptive radiation of Darwin's finches. Molecular Ecology 25:5282-5295.

Chesser, R. T., M. L. Isler, A. M. Cuervo, C. D. Cadena, S. C. 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. Auk in press.

Coyne, J. A., and H. A. Orr (2004). Speciation. Sinauer Associates, Sunderland, Massachusetts.

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

Cui, R., M. Schumer, K. Kruesi, R. Walter, P. Andolfatto, and G. G. Rosenthal (2013). Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. Evolution 67:2166-2179.

de Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. In Endless forms: species and speciation (Howard, D. J., and S. H. Berlocher, Howard, D. J., and S. H. BerlocherHoward, D. J., and S. H. Berlochers). Oxford University Press, New York 57-75.

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

Dobzhansky, T. (1937). Genetics and the origin of species. Columbia University Press, New York.

Duncan, E. J., P. D. Gluckman, and P. K. Dearden (2014). Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 322:208-220.

Dunn, C. W., and C. Munro (2016). Comparative genomics and the diversity of life. Zoologica Scripta 45:5-13.

Edwards, D. L., and L. L. Knowles (2014). Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proceedings of the Royal Society of London B 281:20132765.

Edwards, S. V., S. B. Kingan, J. D. Calkins, C. N. Balakrishnan, W. B. Jennings, W. J. Swanson, and M. D. Sorenson (2005). Speciation in birds: genes, geography, and sexual selection. Proceedings of the National Academy of Sciences of the United States of America 102:6550-6557.

Everson, K. M., J. F. McLaughlin, I. A. Cato, M. M. Evans, A. R. Gastaldi, K. K. Mills, K. G. Shink, S. M. Wilbur, and K. Winker (2020). Speciation, gene flow, and seasonal migration in *Catharus* thrushes (Aves:Turdidae). Molecular Phylogenetics and Evolution 139:106564.

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.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52:399-433.

Flouri, T., X. Jiao, B. Rannala, and Z. Yang (2019). A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. Molecular Biology and Evolution 37:1211-1223.

Fontaine, M. C., J. B. Pease, A. Steele, R. M. Waterhouse, D. E. Neafsey, I. V. Sharakhov, X. Jiang, A. B. Hall, F. Catteruccia, E. Kakani, S. N. Mitchell, et al. (2015). Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 347:1258524.

Fraley, C., and A. E. Raftery (2002). Model-based clustering, discriminant analysis, and density estimation. Journal of the American Statistical Association 97:611-631.

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. Auk 134:857-870.

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

Gillespie, R. G., G. M. Bennett, L. De Meester, J. L. Feder, R. C. Fleischer, L. J. Harmon, A. P. Hendry, M. L. Knope, J. Mallet, C. Martin, C. E. Parent, et al. (2020). Comparing adaptive radiations across space, time, and taxa. Journal of Heredity 111:1-20.

Gómez-Bahamón, V., R. Márquez, A. E. Jahn, C. Y. Miyaki, D. T. Tuero, O. Laverde-R, S. Restrepo, and C. D. Cadena (2020). Speciation associated with shifts in migratory behavior in an avian radiation. Current Biology 30:1312-1321.e1316.

Gopalakrishnan, S., M.-H. S. Sinding, J. Ramos-Madrigal, J. Niemann, J. A. Samaniego Castruita, F. G. Vieira, C. Carøe, M. d. M. Montero, L. Kuderna, A. Serres, V. M. González-Basallote, et al. (2018). Interspecific gene flow shaped the evolution of the genus *Canis*. Current Biology 28:3441-3449.e3445.

Grant, P. R., and B. R. Grant (2008). How and why species multiply: the radiation of Darwin's finches. Princeton University Press, Princeton, New Jersey, USA.

Grant, P. R., and B. R. Grant (2014). 40 Years of Evolution: Darwin's Finches on Daphne Major Island. Princeton University Press, Princeton, New Jersey, USA.

Gutiérrez-Pinto, N., K. G. McCracken, P. Tubaro, C. Kopuchian, A. Astie, and C. D. Cadena (2019). Molecular and morphological differentiation among Torrent Duck (*Merganetta armata*) populations in the Andes. Zoologica Scripta 48:589-604.

Halfwerk, W., C. Dingle, D. M. Brinkhuizen, J. W. Poelstra, J. Komdeur, and H. Slabbekoorn (2016). Sharp acoustic boundaries across an altitudinal avian hybrid zone despite asymmetric introgression. Journal of Evolutionary Biology 29:1356-1367.

Harvey, M. G., S. Singhal, and D. L. Rabosky (2019). Beyond reproductive isolation: demographic controls on the speciation process. Annual Review of Ecology, Evolution, and Systematics 50:75-95.

Hey, J. (2006). On the failure of modern species concepts. Trends in Ecology and Evolution 21:447-450.

Hill, G. E., J. C. Havird, D. B. Sloan, R. S. Burton, C. Greening, and D. K. Dowling (2019). Assessing the fitness consequences of mitonuclear interactions in natural populations. Biological Reviews 94:1089-1104.

Hillis, D. M. (2019). Species delimitation in herpetology. Journal of Herpetology 53:3-12.

Hipp, A. L., P. S. Manos, M. Hahn, M. Avishai, C. Bodénès, J. Cavender-Bares, A. A. Crowl, M. Deng, T. Denk, S. Fitz-Gibbon, O. Gailing, et al. (2019). Genomic landscape of the global oak phylogeny. New Phytologist doi:10.1111/nph.16162.

Isler, M. L., R. T. Chesser, M. B. Robbins, A. M. Cuervo, C. D. Cadena, and P. A. Hosner (2020). Taxonomic evaluation of the *Grallaria rufula* (Rufous Antpitta) complex (Aves: Passeriformes: Grallariidae) distinguishes sixteen species. Zootaxa in press.

Isler, M. L., A. M. Cuervo, G. A. Bravo, and R. T. Brumfield (2012). An integrative approach to species-level systematics reveals the depth of diversification in an Andean thamnophilid, the Long-tailed Antbird. Condor 114:571-583.

Isler, M. L., P. R. Isler, and B. M. Whitney (1998). Use of vocalizations to establish species limits in antbirds (Passeriformes:Thamnophilidae). Auk 115:577-590.

Jackson, N. D., B. C. Carstens, A. E. Morales, and B. C. O'Meara (2017). Species delimitation with gene flow. Systematic Biology 66:799-812.

Jarvis, E. D. (2016). Perspectives from the Avian Phylogenomics Project: Questions that can be answered with sequencing all genomes of a vertebrate class. Annual Review of Animal Biosciences 4:45-59.

Johnson, N. K., J. V. Remsen, Jr., and C. Cicero (1999). Resolution of the debate over species concepts in ornithology: a new comprehensive biologic species concept. Proc.22 Int. Ornithol. Congr., Durban:1470-1482.

Keller, L. F., P. R. Grant, B. R. Grant, and K. Petren (2001). Heritability of morphological traits in Darwin's finches: misidentified paternity and maternal effects. Heredity 87:325-336.

Kingman, J. F. C. (1982). The coalescent. Stochastic Processes and Their Applications 13:235-248.

Krabbe, N. K., T. S. Schulenberg, P. A. Hosner, K. V. Rosenberg, T. J. Davis, G. H. Rosenberg, D. F. Lane, M. J. Andersen, M. B. Robbins, C. D. Cadena, T. Valqui, et al. (2020). Untangling cryptic diversity in the High Andes: Revision of the *Scytalopus* [*magellanicus*] complex (Rhinocryptidae) in Peru reveals three new species. The Auk in press.

Kumar, V., F. Lammers, T. Bidon, M. Pfenninger, L. Kolter, M. A. Nilsson, and A. Janke (2017). The evolutionary history of bears is characterized by gene flow across species. Scientific Reports 7:46487.

Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C.-J. Rubin, C. Wang, N. Zamani, B. R. Grant, et al. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371-375.

Lamichhaney, S., F. Han, J. Berglund, C. Wang, M. S. Almén, M. T. Webster, B. R. Grant, P. R. Grant, and L. Andersson (2016). A beak size locus in Darwin's finches facilitated character displacement during a drought. Science 352:470-474.

Lamichhaney, S., F. Han, M. T. Webster, L. Andersson, B. R. Grant, and P. R. Grant (2018). Rapid hybrid speciation in Darwin's finches. Science 359:224-228.

Lanyon, W. E. (1978). Revision of the *Myiarchus* flycatchers of South America. Bulletin of the American Museum of Natural History 161:427-628.

Leaché, A. D., T. Zhu, B. Rannala, and Z. Yang (2018). The spectre of too many species. Systematic Biology 68:168-181.

Lotsy, J. P. (1925). Species or Linneon? Genetica 7:487-506.

Mallarino, R., P. R. Grant, B. R. Grant, A. Herrel, W. P. Kuo, and A. Abzhanov (2011). Two developmental modules establish 3D beak-shape variation in Darwin's finches. Proceedings of the National Academy of Sciences of the USA 108:4057-4062.

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. Philosophical Transactions of the Royal Society of London B 363:2971-2986.

Mallet, J., N. Besansky, and M. W. Hahn (2016). How reticulated are species? BioEssays 38:140-149.

Marques, D. A., J. I. Meier, and O. Seehausen (2019). A combinatorial view on speciation and adaptive radiation. Trends in Ecology & Evolution 34:531-544.

Martin, S. H., K. K. Dashamapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Research 23:1817-1828.

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. (1963). Animal species and evolution. Harvard University Press, Cambridge, Massachusetts.

Mayr, E. (1969). Principles of systematic zoology. McGraw-Hill, New York.

Mayr, E. (1970). Populations, species, and evolution. Harvard University Press, Cambridge, Massachusetts.

Mayr, E. (1992). Species concepts and their application. In The Units of Evolution: Essays on the Nature of Species (Ereshefsky, M., Ereshefsky, M.)Ereshefsky, M.s). MIT Press, Cambridge, Massachusetts 15-26.

McKay, B. D., and R. M. Zink (2015). Sisyphean evolution in Darwin's finches. Biological Reviews 90:689-698.

Nosil, P. (2008). Speciation with gene flow could be common. Molecular Ecology 17:2103-2106.

Novikova, P. Y., N. Hohmann, V. Nizhynska, T. Tsuchimatsu, J. Ali, G. Muir, A. Guggisberg, T. Paape, K. Schmid, O. M. Fedorenko, S. Holm, et al. (2016). Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. Nature Genetics 48:1077-1082.

Ottenburghs, J. (2019). Avian species soncepts in the light of genomics. In Avian Genomics in Ecology and Evolution (Kraus, R. H. S., Kraus, R. H. S.)Kraus, R. H. S.s). Springer Nature, Switzerland 211-234.

Padial, J. M., A. Miralles, I. De la Riva, and M. Vences (2010). The integrative future of taxonomy. Frontiers in Zoology 7:216.

Palacios, C., L. Campagna, J. L. Parra, and C. D. Cadena (in review). Complete mitochondrial genomes do not distinguish phenotypically distinct lineages of Andean *Coeligena* hummingbirds.

Palacios, C., S. García-R, J. L. Parra, A. M. Cuervo, F. G. Stiles, J. E. McCormack, and C. D. Cadena (2019). Shallow genetic divergence and distinct phenotypic differences between two Andean hummingbirds: Speciation with gene flow? The Auk 136.

Patten, M. A. (2010). Null expectations in subspecies diagnosis. Ornithological Monographs 67:35-41.

Patterson, N., A. L. Price, and D. Reich (2006). Population structure and eigenanalysis. PLoS Genetics 2:e190.

Pearson, K. (1894). Contributions to the theory of mathematical evolution. Philosophical Transactions of the Royal Society of London A 185:71-110.

Pearson, P. N., and T. H. G. Ezard (2014). Evolution and speciation in the Eocene planktonic foraminifer *Turborotalia*. Paleobiology 40:130–143.

Pease, J. B., D. C. Haak, M. W. Hahn, and L. C. Moyle (2016). Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLoS Biology 14:e1002379.

Pinho, C., and J. Hey (2010). Divergence with gene flow: models and data. Annual Review of Ecology, Evolution and Systematics 41:215-230.

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.

Price, T. (2008). Speciation in birds. Roberts and Company Publishers, Greenwood Village, Colorado.

Pritchard, J. K., M. Stephens, and P. Donnelly (2000). Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Pulido-Santacruz, P., A. Aleixo, and J. T. Weir (2020). Genomic data reveal a protracted window of introgression during the diversification of a neotropical woodcreeper radiation. Evolution doi.org/10.1111/evo.13902.

Rannala, B., and Z. Yang (2003). Improved reversible jump algorithms for Bayesian species delimitation. Genetics 194:245-253.

Remsen, J. V., Jr. (2005). Pattern, process, and rigor meet classification. Auk 122:403-413.

Remsen, J. V., Jr., J. I. Areta, C. D. Cadena, S. Claramunt, A. Jaramillo, J. F. Pacheco, M. B. Robbins, F. G. Stiles, D. F. Stotz, and K. J. Zimmer (2018). A classification of the bird species of South America, version July 7 2018. American Ornithologists' Union. http://www.museum.lsu.edu/~Remsen/SACCBaseline.htm.

Rogers, J., M. Raveendran, R. A. Harris, T. Mailund, K. Leppälä, G. Athanasiadis, M. H. Schierup, J. Cheng, K. Munch, J. A. Walker, M. K. Konkel, et al. (2019). The comparative genomics and complex population history of *Papio* baboons. Science Advances 5:eaau6947.

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

Salaman, P., P. Coopmans, T. M. Donegan, M. Mulligan, A. Cortés, S. L. Hilty, and L. A. Ortega (2003). A new species of wood-wren (Troglodytidae: *Henicorhina*) from the Western Andes of Colombia. Ornitología Colombiana 1:4-21.

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.

Seehausen, O. (2004). Hybridization and adaptive radiation. Trends in Ecology and Evolution 19:198-207.

Sites, J. W., Jr., and J. C. Marshall (2004). Operational criteria for delimiting species. Annual Review of Ecology, Evolution and Systematics 35:199-227.

Smith, B. T., J. Bryson, Robert W., W. M. Mauck, J. Chaves, M. B. Robbins, A. Aleixo, and J. Klicka (2018). Species delimitation and biogeography of the gnatcatchers and gnatwrens (Aves: Polioptilidae). Molecular Phylogenetics and Evolution 126:45-57.

Smith, M. L., and B. C. Carstens (2020). Process-based species delimitation leads to identification of more biologically relevant species*. Evolution 74:216-229.

Smith, S. D., M. W. Pennell, C. W. Dunn, and S. V. Edwards (2020). Phylogenetics is the new genetics (for most of biodiversity). Trends in Ecology & Evolution in press.

Smith, T. B. (1993). Disruptive selection and the genetic basis of bill size polymorphism in the African finch *Pyrenestes*. Nature 363:618-620.

Smith, T. B., and S. A. Temple (1982). Feeding habits and bill polymorphism in Hook-billed Kites. Auk 99:197-207.

Sorenson, M. D. (in review). The utility and significance of population genomic data in assessing avian species limits.

Stockman, A. K., and J. E. Bond (2007). Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. Molecular Ecology 16:3374-3392.

Stryjewski, K. F., and M. D. Sorenson (2017). Mosaic genome evolution in a recent and rapid avian radiation. Nature Ecology and Evolution 1:1912–1922.

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

Templeton, A. R. (1981). Mechanisms of speciation - a population genetic approach. Annual Review of Ecology and Systematics 12:23-48.

Templeton, A. R. (1989). The meaning of species and speciation: a genetic perspective. In Speciation and its consequences (Otte, D., and J. A. Endler, Otte, D., and J. A. EndlerOtte, D., and J. A. Endlers). Sinauer Associates, Sunderland, Massachussetts 3-27.

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., L. Campagna, S. A. Taylor, C. N. Balakrishnan, D. T. Baldasarre, P. E. Deane-Coe, M. G. Harvey, D. M. Hooper, D. E. Irwin, C. D. Judy, N. A. Mason, et al. (2016a). Genomic approaches to understanding population divergence and speciation in birds. Auk 133:13-30.

Toews, D. P. L., S. A. Taylor, R. Vallender, A. Brelsford, B. G. Butcher, P. W. Messer, and I. J. Lovette (2016b). Plumage genes and little else distinguish the genomes of hybridizing warblers. Current Biology 26:2313-2318.

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin (2005). Genomic islands of speciation in *Anopheles gambiae*. PLoS Biology 3:e285.

Uy, J. A. C., and R. J. Safran (2013). Variation in the temporal and spatial use of signals and its implications for multimodal communication. Behavioral Ecology and Sociobiology 67:1499-1511.

Van Valen, L. (1976). Ecological species, multispecies, and oaks. Taxon 25:233-239.

vonHoldt, B. M., R. Y. Kartzinel, C. D. Huber, V. Le Underwood, Y. Zhen, K. Ruegg, K. E. Lohmueller, and T. B. Smith (2018). Growth factor gene IGF1 is associated with bill size in the black-bellied seedcracker *Pyrenestes ostrinus*. Nature Communications 9:4855.

Wang, X., Z. He, S. Shi, and C.-I. Wu (2019). Genes and speciation – Is it time to abandon the Biological Species Concept? National Science Review doi/10.1093/nsr/nwz220/5691349.

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

Wray, G. A. (2013). Genomics and the evolution of phenotypic traits. Annual Review of Ecology, Evolution, and Systematics 44:51-72.

Wright, S. (1931). Evolution in Mendelian populations. Genetics 16:97-159.

Wu, C.-I. (2001). The genic view of the process of speciation. Journal of Evolutionary Biology 14:851-865.

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

Zapata, F., and I. Jiménez (2012). Species delimitation: Inferring gaps in morphology across geography. Systematic Biology 61:179-194.

Zhang, C., D.-X. Zhang, T. Zhu, and Z. Yang (2011). Evaluation of a Bayesian coalescent method of species delimitation. Systematic Biology 60:747-761.

Zink, R. M. (2002). A new perspective on the evolutionary history of Darwin's finches. Auk 119:864-871.

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

Zink, R. M., and H. Vázquez-Miranda (2019). Species limits and phylogenomic relationships of Darwin’s finches remain unresolved: Potential consequences of a volatile ecological setting. Systematic Biology 68:347-357.