CHAPTER 6

SUBSPECIES AS A MEANINGFUL TAXONOMIC RANK IN AVIAN CLASSIFICATION

J. V. Remsen, Jr.[1]

Museum of Natural Science, Louisiana State University, Baton Rouge, Louisiana 70803, USA

Abstract.—Dissatisfaction with the subspecies unit of classification is, in part, a consequence of the failure of many of those who have described subspecies to follow the conceptual definition of the subspecies, namely that it should represent diagnosable units. The antiquity of the descriptions of most subspecies (median year of description of currently recognized subspecies estimated to be 1908–1909) means that the majority predated any statistical tools for assessing diagnosability. The traditional subspecies concept, as originally construed, identifies minimum diagnosable units as terminal taxa, and I suggest that it is thus essentially synonymous with the phylogenetic species concept. Therefore, both must deal with the fundamental difficulties inherent in using diagnosability as a criterion. Application of monophyly as a criterion for taxon rank at the population level has inherent difficulties. An advantage of the biological species concept is that it incorporates, in its classification of taxa, assessments of gene flow and reproductive isolation, which are critical components of the evolutionary process. Critics of the biological species concept persistently overlook the fact that it includes the subspecies rank as a necessary component of that concept for distinct populations within biological species. Analyses that require terminal taxa can, with care, be conducted under the biological species concept using subspecies plus monotypic species. Critics of the biological species concept with respect to its application have missed the biological and political disadvantages of treating minimum diagnosable units as the primary unit of conservation concern. Human perception is in accord with ranking such minimum diagnosable units below the species rank; socially and scientifically, humans consider diagnosable units of other humans as distinct groups but not separate species.

Key words: species concepts, species definitions, subspecies definitions.

Las Subespecies como un Rango Taxonómico Significativo en la Clasificación de las Aves

Resumen.—En parte, la insatisfacción con la unidad de clasificación de subespecie es consecuencia de que muchos de aquellos que han descrito subespecies no han seguido la definición conceptual de la subespecie como una unidad diagnosticable. La antigüedad de las descripciones de la mayoría de las subespecies (la mediana del año de descripción de las subespecies actualmente reconocidas se estima en 1908–1909) significa que la mayoría precedió a las herramientas estadísticas para evaluar la diagnosticabilidad. El concepto tradicional de subespecie, como se concibió originalmente, identifica unidades diagnosticables mínimas como taxones terminales, por lo que sugiero que esencialmente es sinónimo del concepto filogenético de especie. Por lo tanto, ambos deben lidiar con las dificultades fundamentales inherentes vinculadas con el uso del criterio de diagnosis. La aplicación de la monofilia como un criterio para la clasificación de los taxones al nivel poblacional tiene dificultades inherentes. Una ventaja del concepto biológico de especie es que incorpora, en su clasificación de los taxones, evaluaciones del flujo génico y del aislamiento reproductivo, que son componentes fundamentales del proceso evolutivo. Las críticas al concepto

---

[1]E-mail: najames@LSU.edu

---

biológico de especie persistentemente pasan por alto el hecho de que éste incluye el rango de sub-especie como un componente necesario para poblaciones diferentes dentro de la especie biológica. Los análisis que requieren taxones terminales pueden ser conducidos, con cuidado, bajo el concepto biológico de especie usando subespecies y especies monotípicas. Las críticas del concepto biológico de especie con respecto a su aplicación han pasado por alto las desventajas biológicas y políticas de tratar a las unidades diagnosticables mínimas como las unidades principales de preocupación con-servacionista. La percepción humana coincide en clasificar estas unidades diagnosticables mínimas por debajo del rango de especie; social y científicamente, los humanos consideran unidades diag-nosticables de otros humanos como grupos distintivos pero no como especies separadas.

WHETHER THE SUBSPECIES rank in classification is considered useful depends on whether one's concept of species includes room for geographi-cally non-overlapping, diagnosable units within a species. Can a species be subdivided into distinct, biologically meaningful units? Should such units be formally named? As noted by all who have written about the classification of organisms, im-posing a categorical scheme, such as the Linnaean system of classification, on the pattern of continu-ous variation produced by evolutionary processes is doomed to be unsatisfactory. As noted by Strese-mann (1936:157), "Whoever wants to hold firm rules, should give up taxonomic work. Nature is too disorderly for such a man." Empirical exam-ples can be mustered that defy the tidy either/or demands of any species concept. Yet human per-ception, dominated by categorical thinking, uses such schemes to produce the vocabulary of labels needed for communication. In short, biological classification attempts to inflict an unrealistic cat-egorical scheme on the patterns produced by a dis-orderly, fundamentally noncategorical process.

Controversy over the utility and definition of the subspecies rank in such a categorical classifica-tion has a long history, with episodic reappraisals (e.g., Wiens 1982), yet the category survives in al-most all modern classifications of birds. This sur-vival, since the mid-1800s, is presumably driven by a perception among most humans that the cat-egory that we term "species" can include within it named subpopulations to identify nonclinal geo-graphic variation. This, in turn, may follow from our own widespread, long-standing perception of the nature of the species *Homo sapiens*, in which pronounced, nonclinal geographic variation is included within that species rather than each dis-tinct group being considered a separate species.

What Is a Subspecies?
Conceptual Definitions

To recognize nonclinal intraspecific geographic variation in animals, some taxonomists have ap-plied trinomials as subspecies names since at least 1844 (*fide* Simpson 1961). The concept behind subspecies definitions centers on the existence of separate units or geographic units within the rank of species. Historically, dissatisfaction with rank-ing every distinctive geographic population as a species was the catalyst for the use of trinomials, which were regarded a century ago as a radical and progressive step in classification (Knox 2007). Definitions of subspecies extracted from standard references and textbooks (Table 1) are founded on the theme that a "species" may consist of subunits that differ from each other in diagnosable ways yet share the characters attributed to the species itself. I combine these ideas into the following definition: "geographic populations diagnosable by one or more phenotypic traits."

The theme that unifies these definitions of subspecies is that subspecific names identify dis-tinct population units: they are phenotypic pre-dictors of past or current genetic continuity, the phenotypic analogue of genetic markers. Hennig (1966:102) stated that the goal of species-level taxonomy was to relegate to subspecies rank "all vicarying reproductive communities."

That the subspecies category has biological meaning is reinforced by the observation that populations known to be reproductively isolated, and thus considered species by any definition, typically differ from close relatives in the same kinds of phenotypic characters and patterns, but to a greater degree (e.g., as seen among subspe-cies that intergrade where in contact). Whether such characters and patterns represent causation or correlation is an open question.

A definition of subspecies as "geographic pop-ulations diagnosable by one or more phenotypic traits" is a simple statement concerning the cur-rent geographic distribution of distinct pheno-typic traits. This definition makes no assumptions about whether the traits are adaptive or whether they represent populations that are incipient spe-cies, and thus makes no predictions concerning the future. Assuming that a phenotypic trait has a genetic rather than environmental basis, sub-species boundaries imply that all individuals of

TABLE 1. Subspecies definitions from textbooks and reference works.

Source	Definition
Mayr et al. 1953, Mayr 1963	Geographically defined aggregates of local populations which differ taxonomically from other such subdivisions of a species
Mayr and Ashlock 1991:43, 430	An aggregate of local populations of a species inhabiting a geographic subdivision of the range of the species and differing taxonomically [differing by sufficient diagnostic characters] from other populations of the species
Futuyma 1979	A set of populations of a species that share one or more distinctive features and occupy a different geographic area from other subspecies
Futuyma 2005:213, 356	A recognizably distinct population, or group of populations, that occupies a different geographic area from other populations of the same species; populations of a species that are distinguishable by one or more characteristics and are given subspecific names
Strickberger 2000	A taxonomic division of a species often distinguished by special phenotypic characters and by its origin or localization in a given geographic region

the subspecies share the genes responsible for the diagnostic trait, which arose in a common ancestor, and, thus, form a monophyletic group with respect to those genes. Whether they also form a monophyletic group with respect to other gene trees is an open question. As I discuss below, however, application of the term "monophyly" at the population level is problematic.

WHAT IS A SUBSPECIES? OPERATIONAL DEFINITIONS

Mayr et al. (1953) provided objective, quantitative definitions of subspecies based on degree of overlap that can be applied across taxa. They outlined why using simple linear overlap in measurements, for example, overemphasizes extreme individuals in a population and overestimates true population overlap. They also discussed various interpretations of the "75% rule" as the threshold for naming subspecies. Although one interpretation is that only 75% of the individuals of each sample have to be correctly classified, the rule as defined by Amadon (1949), Mayr et al. (1953), and Patten and Unitt (2002) is based on standard deviations from the mean of normally distributed data. Depending on which metric is applied, in essence these definitions mean that 90–97% of the individuals of one population must be distinguishable from the equivalent percentage of the other population to be considered subspecies under the somewhat misleadingly named 75% rule.

As emphasized by Mayr et al. (1953) and Patten and Unitt (2002), defining subspecies solely on the basis of statistically significant differences in population means is an unfortunate misinterpretation of the conceptual definition. Given large enough sample sizes, the means of any two populations likely differ significantly (>95%), even though actual overlap can be nearly complete, and so statistically significant differences in the means alone provide almost no information on how distinctive two populations are in terms of diagnosability, the key theme of the conceptual definitions of subspecies. The problem is that the conceptual definitions emphasize the population as a whole, not the individuals that constitute it, and so statistically significant differences between means can be interpreted as diagnosability if the population is the unit of analysis.

Although the 75% rule has a long history in ornithology, its application has been erratic at best. For example, it is generally not mentioned as a criterion for recognizing subspecies in classifications (e.g., American Ornithologists' Union 1957, Dickinson 2003) or in any of the *Handbook of the World* series (del Hoyo et al. 1992–2008). It is not possible to tell how many of the subspecies currently recognized in such sources would qualify as subspecies under the 75% rule, but it is certain that many subspecies, especially in North America, would not qualify as valid taxa under this rule, particularly those defined by mensural differences. From personal experience in attempting to use subspecies diagnoses, such as the keys in the *Birds of North and Middle America* series (Ridgway and Friedmann 1901–1950), I predict that more than 75% of North American subspecies

taxa delimited by mensural data would not survive application of the 75% rule.

Although Patten and Unitt (2002) used 75% as their target level of degree of diagnosability in deference to tradition, they advocated a higher level, 95%, as a standard for diagnosability. I also propose that this level of diagnosability become the operational definition of diagnosable. McKitrick and Zink (1988) gave reasons why aiming for 100% diagnosability for phylogenetic species is conceptually and methodologically unreasonable. Also, geographic sampling to determine diagnosability in the case of parapatric populations must exclude any zones of intergradation in statistical treatments because heavy sampling from that zone would eliminate any potential diagnosability of two populations. To avoid circularity, delimitation of zones of intergradation must be objective (e.g., Harrison 1993), because, hypothetically, one could expand the area considered the zone of intergradation until diagnosability reaches 95% for populations on either side of it. Of course, the amount of intergradation occurring at parapatric zones of contact is of considerable biological and taxonomic interest, especially in determining whether taxa are subspecies or full species (Mayr and Ashlock 1991).

By incorporating a quantitative operational definition into the conceptual historical definition, I produced the following definition, modified from Futuyma (2005), including an explicit statement that diagnosability refers to individuals that comprise the population: A subspecies is a distinct population, or group of populations, that occupies a different breeding range from other populations of the same species; individuals are distinguishable from those other populations by one or more phenotypic traits at the 95% level of diagnosability.

Application of this rigorous definition would result in the synonymization of many subspecies names in North America and elsewhere where broad geographic patterns of smoothly clinal differences in coloration and, especially, morphometrics have been artificially categorized as subspecies. Quantitative analyses of this geographic variation typically found that much or most of this variation, at least in terms of morphometrics, cannot be partitioned into diagnosable units (e.g., Power 1969, Behle 1973, Tacha et al. 1985, Aldrich and James 1991, Wood 1992, Rising et al. 2001, Rising et al. 2009). The implied agenda of much of the work in the first half of the 20th century was that all geographic variation had to be described in a categorical way, namely by use of subspecies names (see Knox 2007). However, patterns of geographic variation in phenotype provide valuable insights into population structure and the process of evolution, regardless of whether the variation can be apportioned into diagnosable units (James 1970, Zink and Remsen 1986). Note also that this definition is based on phenotypic traits in plumage and morphology with an assumed genetic basis, not on other phenotypic traits such as behavior or physiology, nor on genetic markers not expressed in the phenotype; one could make a case for recognizing any diagnosable, geographically distinct genetic unit with a subspecies label. Note also that this definition is incomplete with respect to distinguishing subspecies from species, which I address below. To refer to populations or individuals that represent extremes of clinal variation or populations that do not meet statistical thresholds of diagnosability, for convenience one could use an informal vocabulary using the formerly recognized subspecies names in quotes, followed by "grade" (e.g., "the '*nigrideus*' grade" for the darker northeasternmost populations of the American Robin (*Turdus migratorius*). Note that use of a 95% diagnosability criterion applies only to two-way comparisons, as in the 75% rule, not to multiple simultaneous comparisons. It also applies only to populations in the breeding ranges; an empirical consequence of such a rigorous standard (if individuals from intergrade zones were excluded) would also allow assignment of individuals from the nonbreeding range to breeding population with a high level of statistical certainty.

A persistent criticism of the subspecies concept is that analysis of different characters may produce different subdivisions (e.g., Wilson and Brown 1953). In other words, the characters are not distributed in a concordant geographic manner; for example, three characters might show geographic variation, but each character could show three different patterns that would delimit subspecies boundaries in three conflicting ways. The existence of such conflict is inevitable, and application of the subspecies concept in such cases is unwarranted. However, my impression, based on examining many hundreds of primarily Neotropical bird species over the past 30 years with respect to distinct plumage characters, is that such conflicts are greatly outnumbered by examples of concordance. For example, if three traits

show nonclinal geographic variation, all three may not show breaks at the same point, but actual conflict in where the breaks are is infrequent. For exceptionally thorough quantifications of patterns of geographic variation in morphometrics and plumage that showed strong concordance, see Johnson (1980) and Cicero (1996). By contrast, smoothly clinal variation regularly shows conflicting patterns, with trends in, for example, tail length showing a different pattern from that of, for example, back color.

Ridley (2004) used lack of concordance in geographic variation as grounds for essentially dismissing the importance of the entire subspecies unit of classification. He used the example of clinal size variation in North American House Sparrow (*Passer domesticus*) to illustrate clinal variation and then noted that there was no reason to expect clines in other characters to match the body-size cline. However, he did not point out that no subspecies have been described or recognized in North American House Sparrows because the variation is smoothly clinal rather than discrete, so his example is inappropriate, and his premise is flawed. If the geographic pattern of variation in distinct characters produces conflicting patterns, then this implies complex underlying population structure that may not be amenable to diagnoses and, therefore, subspecies names. However, Johnston and Selander (1964:549) found that "color differences between samples are both marked and consistent, permitting 100 percent separation of specimens from two localities" but did not name any subspecies because they did not think that the situation was temporally stable.

Subspecies versus Phylogenetic Species

Given that conceptual definitions of subspecies have always emphasized diagnosable units, how do they differ from phylogenetic species other than that, under the biological species concept, many diagnosable units are ranked as species? Cracraft (1983:170) defined phylogenetic species as "the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent." Cracraft did not define how these differ from subspecies but emphasized the heterogeneous nature of the results of applying the subspecies concept. He urged abandoning the subspecies rank in classification, without detailing how it differs from phylogenetic species. The methodological difficulties

that produced the heterogeneity in units called subspecies are assumed to disappear if Cracraft's phylogenetic species concept is adopted—when, in fact, delimitation of "diagnosable clusters" entails all the methodological problems that complicate subspecies delimitation. Any renaming of all minimum diagnosable units as species would require determining what units are actually diagnosable and at what statistical thresholds of diagnosability.

Cracraft (1983) pointed out the biological species concept lacks equivalency among the units called species. However, the same problem pervades species defined under the phylogenetic species concept, in which, for example, species reproductively isolated from all other lineages, including syntopic sister taxa, are treated as the same taxonomic unit as populations that differ only in the possession of a single diagnostic character and cannot coexist syntopically with sister taxa. By contrast, use of the subspecies rank within biological species as the unit of analysis reduces the problems of heterogeneity because population units diagnosed only by minor plumage differences are not treated as the same unit as lineages known or inferred to be sealed from other lineages by reproductive isolation.

Another criticism of the biological species concept is that biological species are not the appropriate unit for biogeographic and speciation analyses (Cracraft 1983). I agree. The appropriate units are indeed minimum diagnosable units—that is, subspecies under the biological species concept. That subspecies can be used productively for such analyses is shown inadvertently by none other than Cracraft (1983), who used subspecies names in outlining his methods for determining areas of endemism. Cracraft (1985) later also used lengthy lists of trinomials to demarcate and name areas of endemism in the Neotropics. Those areas, defined by the terminal taxonomic unit of the biological species, namely subspecies, are still the standard nomenclature for Neotropical biogeographic analyses, thereby demonstrating the utility of the subspecies unit of classification.

Other definitions of phylogenetic species repeat the essence of Cracraft's phylogenetic species concept, with the emphasis on diagnosability and common ancestry, and they do not address how this definition differs from that of the subspecies. Futuyma (2005) and Freeman and Herron (2007) also reported the definition of the phylogenetic species concept without explaining how it differs

from the rank of subspecies within the biological species concept. The explicit conceptual definition of phylogenetic species is that they represent monophyletic units, whereas subspecies are not defined explicitly with respect to monophyly. In practice, however, when phenotypic characters define phylogenetic species, the issue of monophyly is often ignored. In fact, the phylogenetic species concept's method of the minimum diagnosable unit, when applied to phenotypes, can be applied to inanimate objects and is not inherently phylogenetic (Johnson et al. 1999). Further, when genetic criteria are used to define monophyly, these criteria are typically just one or two loci, typically non-recombining mitochondrial DNA (mtDNA) genes, and monophyly with respect to other loci is not addressed (see below).

Furthermore, using unique or even multiple characters to identify minimum diagnosable units does not guarantee monophyly of the taxa in question if three or more populations are involved. If rates of character evolution are unequal, then some populations will become diagnosable before others, leading to paraphyletic groupings of populations that have not become diagnosable, similar to the "metaspecies" and "plesiospecies" problems (Donoghue 1985, Olmstead 1995, Willmann and Meier 2000). If three populations have a known history C+(A+B), if B is the first population to acquire a diagnostic character, leaving C and A with nothing but ancestral character states, then even if the true history were known, there would be no way to avoid a paraphyletic taxon A+C if the strict rules of diagnosability are followed. If their geographic ranges are linear, the taxa are sedentary, and the central taxon acquires an apomorphy first, then at least we would be suspicious that the character distribution represents unequal rates of character evolution. In fact, such a linear array provided a clue suggesting that many such cases in Andean birds represented cases of unequal character acquisition that would potentially mislead phylogeny (Remsen 1984). In many cases, however, the populations' ranges are not linear, and in such cases, geography cannot provide hints that the populations without diagnostic characters form a paraphyletic taxon.

Advocates of the phylogenetic species concept often promote its adoption because it makes the fundamental unit of classification "historical taxa" (Zink and McKitrick 1995), whereas in the biological species concept non-sisters can be treated as a single species. As noted above, morphology-based applications of the phylogenetic species concept do not necessarily produce historical taxa. History is a continuum, and the exercise of recognizing which historical units within this continuum are named taxa is inherently arbitrary. Worse, at the population level, defining historical units depends on which characters or which loci are thought to represent the true history. As acknowledged by Zink and McKitrick (1995), it is well known that use of any one set of markers can lead to misrepresentations of history (Tateno et al. 1982, Neigel and Avise 1986, Pamilo and Nei 1988). Only by knowing the gene trees of a large number of polymorphic loci can the true population history be reconstructed, and even then, incomplete lineage-sorting may complicate resolving a single history even if entire genomes are sequenced (Pollard et al. 2006). Further, for all populations with topographically and climatologically heterogeneous ranges, this history likely dates no farther back than the most recent pulse in the cycle of fragmentation and secondary contact.

Using diagnosability as a criterion for naming taxa has inherent methodological problems that affect phylogenetic species and subspecies (under the biological species concept) equally, for four reasons. (1) Any diagnosability level is arbitrary. Because diagnosability is a continuum, from 0 to 100%, any cutoff is inherently arbitrary and cannot be defended conceptually (Johnson et al. 1999). Setting the threshold at 95% is a reasonable level because of the widespread use of that arbitrary level for statistical "significance." Nonetheless, the consequence is that two populations that are, for example, 95% diagnosable are given taxon status, whereas those at 94% are not and are included in the same unnamed category as those population samples diagnosable at 0%. (2) A corollary of arbitrary diagnosability is that the outcome is driven in part by sampling. The closer the diagnosability approaches the threshold, the higher the chance that an increase of one additional individual in the sample will determine the outcome; thus, such an addition to the sample could change the ranking from unnamed taxon to phylogenetic species or subspecies without any true change in the biology and history of the populations. (3) The geography of sampling is critical to the outcome if the character assessed shows any geographic variation (Zink and Remsen 1986). Past gene flow or residual geographic variation in the once-continuous populations makes it essential that

sampling be focused on geographically proximate populations. And (4) diagnosability is driven by the resolution of the technique used (see Collar 1997, Avise 2004).

Zink (2006) objected to my definition of subspecies and phylogenetic species as synonyms (Remsen 2005) because, in essence, some minimum diagnosable units under the biological species concept would be ranked as species if reproductively isolated from other such units. Zink's (2006) argument is largely semantic, because a biological species that is monotypic (contains no units ranked as subspecies) would still be treated as equivalent to the subspecies unit in those analyses for which minimum diagnosable units are the appropriate unit of analysis. In other words, an analysis using minimum diagnosable units under the biological species concept would include all taxa ranked as subspecies plus all monotypic species. The difference between the biological species concept and the phylogenetic species concept is not in defining minimum diagnosable units but in the ranking of some of those units as species. Under the biological species concept, 4,677 (48%) of the 9,722 species in Dickinson are monotypic (D. Lepage pers. comm.)

Monophyly at the Population Level?

The original conceptual theme of the phylogenetic species concept is that minimum diagnosable units are not only diagnosable, but monophyletic (Cracraft 1983). As is now well known, the problem is that at the population level, monophyly is difficult to define and determine (e.g., de Queiroz and Donoghue 1990, Wheeler and Nixon 1990, Davis and Nixon 1992). Only if all gene trees within a series of populations that share a common ancestor have topologies that do not conflict can a single population be labeled unambiguously monophyletic. Genetic data (e.g., Avise 1989) confirm what common sense predicts: the turbulent history and complex population genetics of real-world situations are often unlikely to produce true monophyly because of incomplete lineage-sorting and gene flow among populations that are not reproductively isolated. Gene tree topologies, superimposed, probably look more like a tangled net than a tree (Degnan and Rosenberg 2009). Further, it is now understood that under some circumstances the gene trees of independently segregating loci are not expected to recover the true species tree (Rosenberg

and Tao 2008) and that postdivergence gene flow may make reconstructing species trees from gene trees particularly problematic (Takahata and Slatkin 1990, Eckert and Carstens 2008). Add to this the historical likelihood of repeated phases of expansion, range fragmentation, and secondary contact, and the use of the term "monophyly" becomes problematic. For a particularly well documented example of how a single gene tree can misrepresent species trees of buntings in the genus *Passerina*, see Carling and Brumfield (2008). In part because of this, Hennig (1966: fig. 4) recognized and illustrated this problem graphically, did not apply the term "monophyly" below the species level, and used the reasoning of the biological species concept in his definition of species. Although Hennig used characters to label species in his diagrams illustrating cladistic methodology and is thus widely cited as an advocate of the phylogenetic species concept, Hennig clearly considered reproductive isolation the essential component of speciation (e.g., Hennig 1966:54). Reproductive isolation is the necessary first step toward true monophyly.

Even with respect to a single gene, monophyly at the population level differs fundamentally from monophyly at higher levels because it can be ephemeral, perhaps typically persisting only during the refugial phase of range expansion and contraction cycles, and even then being vulnerable to dispersal-generated gene flow. Therefore, the objection to subspecies or biological species because they are not monophyletic (e.g., McKitrick and Zink 1988) is not condemning. Paraphyly and polyphyly at the population level are predicted, and empirically demonstrated, to be widespread (Funk and Omland 2003). For example, Hull et al. (2008) showed that Swainson's Hawk (*Buteo swainsoni*) is paraphyletic with respect to Galapagos Hawk (*B. galapagoensis*) in terms of mtDNA; however, there is no other biological support for merging *B. galapagoensis* into *B. swainsoni* or for recognizing two or more species within traditionally defined *B. swainsoni*. Further, possession of a diagnostic character, the criterion needed for phylogenetic species rank, is no guarantee of monophyly with respect to other genes. For example, Swainson's Hawk has a suite of diagnostic phenotypic characters despite its being a paraphyletic unit with respect to Galapagos Hawk. Labeling clusters of populations as species on the basis of monophyly with respect to single gene trees indicates monophyly only with respect

to that gene tree, and not necessarily with respect to the the population or species tree (Edwards et al. 2005, 2007). Additionally, in practice, far too few individuals are typically sampled to determine whether two populations are monophyletic with respect to the loci surveyed (for an example of how differences in sample size can affect conclusions concerning population monophyly, see Brumfield 2005). In fact, an earlier study of Swainson's and Galapagos hawks (Riesing et al. 2003) had reported that the two were reciprocally monophyletic only because too few individuals had been sampled (Hull et al. 2008). In summary, the putative advantage of the phylogenetic species concept in establishing monophyletic units as the fundamental unit of taxonomy is appealing rhetoric but elusive reality. Hennig's (1966) restriction of the term "monophyly" to levels of classification above that of species under the biological species concept reflects remarkable wisdom given the state of knowledge of population genetics at that time.

What Is a Species?

One cannot discuss subspecies without also defining species. The controversy over species concepts is obviously too large and complex to treat here; see Coyne and Orr (2004) for a comprehensive review. De Queiroz (2005a, b) pointed out that all species concepts share the property, explicit or implicit, that the unit called "species" represents the uniquely biological property of a separately evolving metapopulation lineage. The problem is how to apply that concept and which criteria are used to delimit species. Although de Queiroz (2005a, b) tried to present his broadly defined species concept as a solution, he offered no real operational definition with respect to explicitly defining such a unit; in fact, at one extreme, a pair of individuals colonizing an island and successfully reproducing could fit the definition of "separately evolving metapopulation lineage" after a single generation. Nonetheless, de Queiroz's (2005b) simple diagram of the splitting and subsequent divergence of populations crisply illustrates the underlying problem of setting criteria to demarcate species boundaries. His use of continuous shading aptly emphasizes the continuum of degrees of divergence and the inherently arbitrary decisions necessary. The island example above would represent the first point past divergence on his time axis. Therefore, some level of subjectivity inevitably influences one's choice of criteria.

As expressed more fully elsewhere (Johnson et al. 1999, Remsen 2005), I favor definitions of species based on a fundamental process of evolution at the population level, namely gene flow or lack of it; that is the essence of the biological species concept. My support for process-based definitions—rather than being "blind allegiance" to the biological species concept, the accusation leveled by Peterson et al. (2006)—is based on recognition that severe diminishment or cessation of gene flow is clearly critical to diversification. Personally, I regard the biological species concept as an imperfect attempt at inflicting a typology on a continuum; however, I dislike even more any other categorical scheme proposed so far (e.g., various versions of the phylogenetic species concept). Rather than become disillusioned at the failures, I recommend rejoicing in the underlying complexity that the failures reveal.

The primary operational problem of the biological species concept, as emphasized by Ernst Mayr from the outset (e.g., Mayr 1942b), is in dealing with ranking allopatric differentiated populations. Here, I note that human cognition deals directly with this problem in recognizing differentiated but reproductively fully compatible units within *Homo sapiens* as conspecifics. This predates science, much less the Modern Synthesis, in that even the earliest historians treated allopatric differentiated populations of humans as "people," rather than as some other type of species. Territoriality and combat, typical manifestations of intraspecific competition but relatively rare in interspecific competition, were expressions of that cognitive framework. Therefore, in treating distinct interpopulational differences as part of the same species, the biological species concept has a subjective appeal that the phylogenetic species concept lacks. The phylogenetic species concept could also produce some unknown number of species within *Homo sapiens*, a result refuted by human behavior long before modern societal influences.

The problem of assigning rank to differentiated allopatric populations is not as intractable as is often portrayed. By placing the degree of differentiation in a comparative phylogenetic framework, namely comparing degree of differentiation in the allopatric form to that seen in closely related sympatric or parapatric populations, a reasonable and testable hypothesis can be

made concerning whether the allopatric form has or has not differentiated to the degree shown by related forms that do or do not freely interbreed (Miller 1955; Mayr 1969, 1996; Mayr and Ashlock 1991; Helbig et al. 2002; Futuyma 2005). Any arbitrariness involved in assignment of taxon rank through this process is no greater than that inherent in assessing minimum diagnosable units under the phylogenetic species concept.

Reproductive Isolation

The importance of reproductive isolation in guaranteeing independent evolutionary lineages has been emphasized by many authors, including Hennig (1966) and Cracraft (1983). Proponents of the phylogenetic species concept can seem schizophrenic toward reproductive isolation, first acknowledging its importance, then dismissing its importance. For example, McKitrick and Zink (1988:6) stated that "the 'closure' or sealing of a gene pool is therefore an important evolutionary event." Yet they explicitly denied a role to reproductive isolation in ranking taxa because interbreeding is a "primitive trait" or "ancestral character" (e.g., Zink 2006). The ability to interbreed could perhaps be construed as an ancestral character, but empirical evidence in birds suggests a severe limit to interbreeding in terms of time since divergence: Price and Bouvier's (2002) survey indicated that postzygotic incompatibilities begin to originate by ~2 million years after divergence. Moreover, free interbreeding (i.e., nonassortative mating with hybrids having equal fitness to pure parental) provides a highly reliable indicator of a close relationship. Empirically, it is limited in birds to populations that have diverged to a limited degree; if not sisters, such populations are members of a lineage that abruptly replace each other geographically. In other words, the ancestral component of free interbreeding is highly restricted to parapatric representatives of a single lineage. Zink and McKitrick (1995) reiterated the importance of reproductive isolation and considered studies of it valuable, but they also argued that it should not have a role in delimiting species. Missed altogether is that reproductive isolation or its absence governs the distribution of characters that delimit the phylogenetic species concept's minimum diagnosable units in sympatric and parapatric taxa; therefore, the pattern of diagnosability is a product of the process dismissed as an "ancestral character." As noted

previously (Avise and Wollenberg 1997, Remsen 2005), denying a role in classification to the most important threshold in the history of a lineage seems incongruous if that classification is supposed to be based on the history of a lineage.

Zink and McKitrick (1995) implied that some proponents of the biological species concept place theoretical emphasis on reproductive isolation because the lack of it, namely hybridization, means that the two populations may eventually homogenize. Similarly, Zink (2006) portrayed the biological species concept as placing importance on the "potential future outcome of current interbreeding." Rather than making such predictions, the classification of two differentiated, freely interbreeding populations as one biological species represents only a statement concerning the current interaction of the two populations, namely that in terms of mate selection and recognition, individuals of both populations treat each other as equivalents, regardless of any previous history of differentiation. It does not necessarily predict the future (although considered by some a hallmark of a mature research field, not speculation), nor does it necessarily group historical taxa. However, it represents important information concerning the current situation in terms of individual behavior and its consequences for population genetics. In summary, such population interactions provide taxonomist-free data on whether (or to what degree) two populations consider themselves "the same" or "different."

At least some of the controversy over the importance of reproductive isolation is caused by disagreement over, or misrepresentation of, the definition of reproductive isolation. Mayr's definition of the biological species concept emphasizes free interbreeding, widely interpreted as nonassortative mating in contact with no reduction in hybrid fitness. In contrast to any criteria based on diagnosability, the advantages of these criteria are (1) that ranking depends on the biological behavior of the individuals involved and (2) that any change in that behavior has consequences for gene flow. That patterns of mate choice may change temporally or geographically is inevitable, and these differences will generate problems for anyone who expects a typological categorization scheme to nimbly handle all real-world variation. Populations that interbreed but still mate assortatively (e.g., no hybrid swarm in contact zone) are treated as separate species under most interpretations of the biological species

concept. If new data reveal that existing taxa classified by the biological species concept are actually freely interbreeding, that classification should be changed.

Subspecies as Straw Men and Phylogenetic Species

Finding that existing classifications of subspecies are defective at some level is not an indictment of the subspecies concept itself, no more than a reanalysis of a phylogenetically based classification that found problems with a previous classification would be an indictment of the phylogenetic species concept. Those who attack subspecies as a taxon rank consistently miss the distinction between a concept and the correct application of that concept. An everyday analogy would be to blame the car, not the mechanic, for a botched repair job. For example, McKitrick and Zink (1988:11) advocated abandonment of the subspecies rank largely because of the historical inconsistency in its application and admitted that properly characterized subspecies—namely, in their words, those "distinct from other populations in one or more characters"—"would be called [phylogenetic] species by our criteria" (as echoed by Zink 2006).

Using existing subspecies classifications as an indictment of anything is disingenuous. The vast majority of such classifications has not been subjected to a modern, quantitative analysis since their original presentation, often the *Check-list of the Birds of the World* series (Peters 1934–1987), many dating as far back as the 1930s. More recent synopses, such as Dickinson (2003) and the *Handbook of the Birds of the World* series (del Hoyo et al. 1992–2008), largely repeat the initial classifications in Peters's *Check-list* unless subsequent studies have altered them. Although the 1960s and early 1970s saw a wave of quantitative studies, particularly in North America, few such studies have been published since then. Thus, the vast majority of subspecies-level classifications remain mostly unchanged from those of Peters's *Check-list* and are maintained largely by historical inertia, a diminishment in this type of biodiversity science, and a lack of adequate material to readdress historical hypotheses.

However, many critiques of the subspecies concept seem to assume that these classifications undergo some sort of constant, modern, quantitative scrutiny. As pointed out previously (Remsen

2005), the majority of subspecies were described in a prestatistical era. In fact, the term "statistics" and even the simplest statistical analyses, such as the *t*-test, postdate the majority of subspecies descriptions. The percentage of subspecific classifications in the Peters's *Check-list* that have ever been subjected to statistical evaluation is minute, perhaps <1%. Therefore, the chances that any of these classifications would not require modification after a modern reanalysis are also minute. I am unaware of any quantitative reanalyses of existing subspecies designations that have not produced modifications of existing subspecies classifications. For example, see Cicero (1996), who found that 4 of the 10 subspecies in the *Baeolophus inornatus* complex were not diagnosable, and Patten and Pruett (2009), who found that only 25 of 51 subspecies of *Melospiza melodia* represented diagnosable units.

To illustrate these points, I plotted (Fig. 1) the date of the type descriptions of all subspecies currently recognized by Dickinson (2003) for two bird families, Parulidae and Pycnonotidae, of similar size but contrasting features. The family Parulidae is restricted to the New World, much of its diversity is at temperate latitudes, and many species are highly migratory. The family Pycnonotidae is restricted to the Old World tropics and includes no highly migratory species. Despite the differences, the chronology and pattern of subspecies descriptions are remarkably similar. Fifty percent of all descriptions predate the first publication of Student's *t*-test (1908), much less its widespread use in ornithology, 70% predate Fisher's (1930) seminal work on population genetics, and 79% predate Huxley's (1942) book on the Modern Synthesis. Therefore, to use such classifications as ammunition to attack subspecies as a concept is a classic straw-man approach that is counterproductive to elucidating the patterns of diversification and the processes that produce them. Any critique of the subspecies unit as a concept using empirical results should start by determining which named subspecies fit the conceptual definition. Failure to apply such a conceptual definition to subspecies designations over the past century has, in my opinion, directly catalyzed the origin of the phylogenetic species concept.

In contrast to subspecies designations, the phylogenetic species concept benefits from having few empirical applications to examine on any large scale. A reasonable prediction is that if all
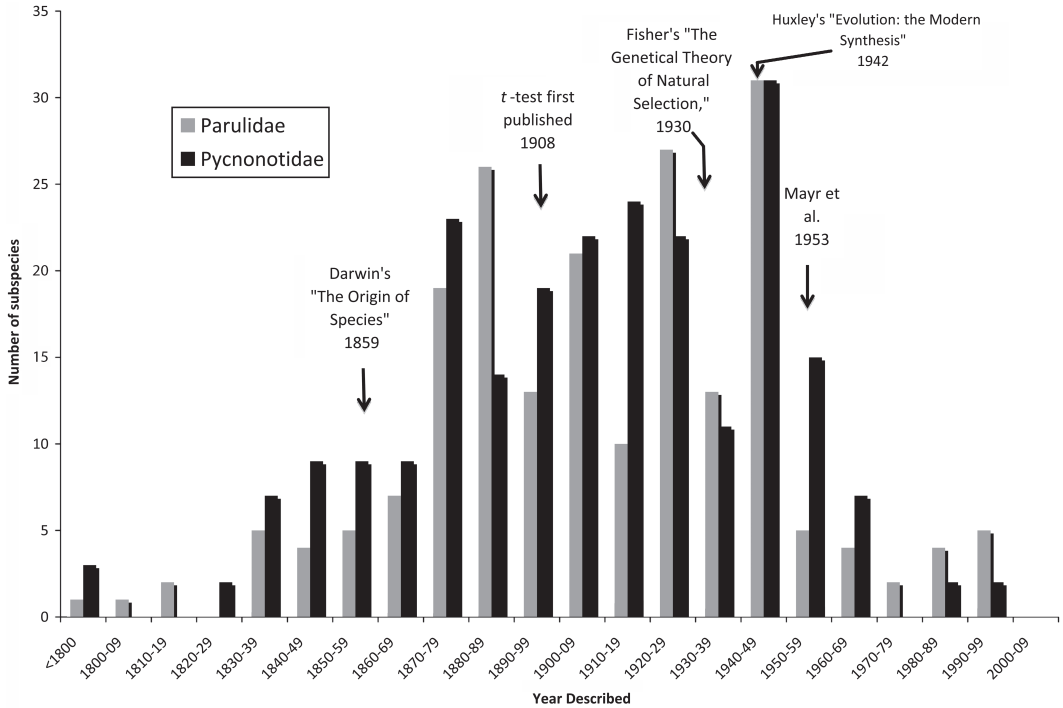
Fig. 1. Historical pattern of dates of descriptions of subspecies in the Parulidae and Pycnonotidae. The data plotted are the publication years of the type descriptions for all subspecies currently recognized by Dickinson (2003), not including, of course, the type description of the species.

classifications started using a phylogenetic species classification as of today, then 100 years from now those results would be viewed with the same disdain directed at current subspecies classifications. Published analyses using phylogenetic species as units already provide ample fodder for criticism, with a near absence of quantitative rigor in determining whether their units actually represent minimum diagnosable units. Given the importance in many analyses of using minimum diagnosable units, whether called subspecies or phylogenetic species, the first step requires a rigorous determination of what those units are.

However, Navarro-Sigüenza and Peterson's (2004) listing of bird species for Mexico based on diagnosable units is quantitatively inferior to that of Robert Ridgway's volumes from the early 1900s that cover the same area (Remsen 2005), although Peterson and Navarro-Sigüenza (2006) assured us that unpublished analyses supported their designations. Peterson and Navarro-Sigüenza (2006:886) also assured us that their 2004 classification was a "consistent taxonomy" that was

"based on the same criteria as in all other clades," yet their methodology and criteria remain unspecified. Likewise, Peterson's (2006) synopsis of diagnosable units in Philippine birds rests on unspecified sample sizes (noted as "woefully small" for many populations) and qualitative assessments. See Collar (2007a) for a full critique of Peterson's (2006) approach. Simply dismissing all trinomial nomenclature and then labeling as species all populations that by qualitative inspection appear diagnosable is not an acceptable program for assessing biodiversity; see Collar (1997) for similar comments on fundamental problems with Cracraft's (1992) revision of a single family, the Paradisaeidae, based on the phylogenetic species concept. In fairness to all these attempts, reevaluation of the diagnosability of currently described subspecies, especially for a rich avifauna such as that of Mexico, is a daunting, monumental task that will require detailed research, and Navarro-Sigüenza and Peterson have made a noble start. Unfortunately for biodiversity assessment, these kinds of baseline analyses of geographic variation

are not considered groundbreaking research. Nonetheless, casual, qualitative inspection of study skins is no longer an acceptable practice for taxonomic revisions, whether the taxa are labeled "subspecies" or "phylogenetic species."

Some biogeographic analyses using phylogenetic species as units start with the assumption that certain described subspecies accurately represent minimum diagnosable units, declare them to be species, and then proceed with the analysis. These analyses typically do not report sample sizes or the geography of their sampling distribution, much less character analyses, diagnosability indices, or anything else that would permit replication. Notable recent exceptions are the analyses of McKay (2008) and D'Horta et al. (2008). Prior to any phylogeographic analysis, they began with a quantitative analysis of geographic variation in plumage characters to define minimum diagnosable units. As noted previously, the antiquity of most subspecies names makes it inevitable that many will fail diagnosability tests. However, analyses that do not include the characters used to diagnose the taxa are unlikely to address diagnosability adequately. For example, Drovetski et al. (2009) quantitatively analyzed geographic variation in breast plumage in the currently recognized species of North American rosy-finches in the genus *Leucosticte* but omitted those plumage characters (face pattern) formally (e.g., Ridgway 1901, MacDougall-Shackleton et al. 2000, Johnson 2002, Johnson et al. 2002) used to diagnose the taxa. Snow (1997) pointed out that incomplete geographic sampling and small sample sizes for many taxa make it necessary to study geographic variation and taxonomy in detail before determining what constitutes minimum diagnosable units.

Environmental Induction

In his widely used textbook, Gill (2007:575) stated that "geographical differences in size or color may be due directly to environmental differences rather than evolved genetic differences among populations" but provided no further details or citations. One likely source of such statements is a tiny number of studies that have documented minor environmental effects on body size and shape in relation to the genetic component (James 1983, Larsson and Forslund 1992, Leafloor et al. 1998), although most such studies have not found an environmental component

(see Merilä and Fry 1998). Many subspecies have been described on the basis of measurements that reflect overall body size. Regardless of whether such differences have an environmental component, I suspect that many or most of these subspecies will be shown to fail diagnosability tests. The vast majority of such subspecies have not been analyzed using any test of degree of overlap, and their validity often rests on differences between means and various qualitative assessments of the ranges.

The other potential source of Gill's (2007) statements concerning environmental effects is the relationship between diet, or other measures of condition, and feather pigmentation or structure. Environmental effects on the ability to express appropriate coloration are widely known, in that poor condition or disease may effect the coloration of individuals within a bird population. Coloration based on carotenoids can be affected strongly by diet because carotenoid pigments must be acquired from food (reviewed by McGraw 2006a), and the expression of carotenoid-based coloration can be altered by environmental conditions, such as parasite load (reviewed by Hill 2006). As for melanin-based coloration, the most widespread source of coloration in birds, documentation of environmental effects is not as clear-cut, and some experiments have failed to find an effect of diet on melanin production or expression (e.g., Gonzalez et al. 1999, Buchanan et al. 2001, McGraw et al. 2002). Nonetheless, because melanin is synthesized from amino acid precursors using metabolic energy and their deposition is influenced by at least four classes of hormones (McGraw 2006b), the potential remains for an effect on its production owing to general health and nutrition. As for structural colors, limited experimental data suggest that nutrition during molt may affect their expression (Hill 2006), but such experiments are limited mainly to glossy black species, are largely correlational, and have not addressed potential confounding influences of age (Prum 2006). Nonetheless, given the complex pathways involved and the extraordinary structural precision required to produce normal coloration (reviewed by Prum 2006), the potential for environmental effects would seem large. Whether coloration is based on nanostructure or pigments, environmental effects on individuals within a population are highly likely. (And this is discounting the ways in which the environment can alter plumage through time, such as through

fading and wear.) That environmental differences might also produce region-wide effects on certain populations of a species remains a potential explanation for geographic variation in coloration and body size in birds. However, data that document a link between environmental effects and among-population differences in coloration are lacking.

Evidence for natural selection on color shades and patterns is reasonably strong (e.g., Burtt 1981, Rohwer and Ewald 1981, Prum 1997, Negro et al. 1998, Dumbacher and Fleischer 2001, Mumme 2002, Tickell 2003). Certainly, the strong associations between patterns of coloration and various ecological and social factors (for reviews, see Bortolotti 2006, Dale 2006) imply that natural selection on the underlying genetic basis of these patterns is widespread. Dramatic seasonal changes in plumage coloration in some species, typically associated with changes in social system, also imply strong selection (although environmental effects caused by a seasonal shift in food supply are not necessarily ruled out by such correlations).

Subspecies and Conflicts with Gene-based Phylogenies

Several recent papers have attacked the utility of subspecies by comparing current subspecies-level classification with patterns of diversification shown by mtDNA (e.g., Zink et al. 2001). A mismatch between units defined by mtDNA versus subspecies is then proclaimed as evidence that subspecies mask patterns of diversity or obscure analyses of the process of historical diversification (Zink 2004).

Conflicts between mtDNA trees and subspecies units may result from faulty delineation of subspecies boundaries, because most have not been critically or quantitatively examined (see above). However, if the subspecies boundaries represent diagnosable units, then I am unaware of any model of evolution that predicts perfect concordance between diagnosable phenotypic units and any single gene tree, particularly those of presumably neutral loci. Lost in the discussion of such conflicts is that (except for populations without any history of fragmentation and secondary contact) gene trees and population trees not only differ, but also are expected to do so because of the influences of incomplete lineage-sorting and gene flow (for review, see Coyne and

Orr 2004). This is especially true for the most frequently analyzed genes, those of mtDNA, which are matrilinearly inherited as a single linkage unit. Empirically, mtDNA markers may do as well as any in tracking population history (Zink and Barrowclough 2008), but to uncritically treat an mtDNA gene tree as equivalent to the true population history should be termed "mtDNA myopia." For example, Zink (2004:563) stated that "subspecies should be judged to fail as meaningful units if they do not predict the evolutionary history of the populations they represent," but in Zink's view mtDNA phylo-groupings represent the only history worth recognizing taxonomically, without recognizing that an mtDNA phylogeny is merely a gene tree. A population marked by a phenotypically diagnosable character, provided that character has a genetic basis, also shares a common history but on a different time-scale. For subspecies units to show perfect concordance with an mtDNA gene tree, each subspecies would also have to have a unique haplotype (or haplotype lineage), an unrealistic expectation. Even so, Phillimore and Owens (2006) showed that Zink's estimates were an order of magnitude too low because of sampling bias and that broader sampling indicated that more than a third of the taxa ranked as subspecies were monophyletic even by the highly restrictive and unrealistic criterion of mtDNA haplotypes. The title of Zink's (2004) paper proclaimed that subspecies obscured biological diversity; however, one could also make a case that using mtDNA phylogroups as taxonomic units obscures biodiversity because it ignores biologically important, phenotypic markers of recent population history. Under Zink's extreme view, some diversity even at the species level would be erased, with most Galápagos finches merged into a few monotypic species because their mtDNA gene trees are not reciprocally monophyletic (Zink 2002). Described as "an unfortunate reliance on a single, potentially misleading molecule" by Grant and Grant (2006), such a treatment as single species would ignore the reproductive isolation and divergence of multiple lineages within this radiation.

Researchers who do not find concordance between genetic data and subspecies boundaries often proclaim that such subspecies are not genetically distinct. Two fundamental problems beset such statements. First, such studies typically analyze one or two genes, often mitochondrial—that is, a tiny fraction of the genome. The

appropriate qualifier for such statements would be that a subspecies is not genetically distinct with respect to whatever number of genes was analyzed. Second, if a subspecies is diagnosable by phenotypic characters (external manifestations of genetic characters), then indeed it is also likely genetically distinct, but the gene(s) that control those characters have not been located or analyzed. If two or more populations share the same phenotypic characters that have arisen by common selection pressure, then their grouping into a single taxon would mislead phylogenetic classification; this is where mtDNA or other genetic markers can elucidate the true population history that phylogenetic classification requires.

Subspecies as Impediments to Conservation

Some (e.g., Hazevoet 1996, Sangster 2000, Peterson 2006) have claimed that ranking diagnosable units as species under the phylogenetic species concept or a similar concept rather than as subspecies under the biological species concept benefits conservation. See Collar (1996, 1997, 2007a), Garnett and Christidis (2007), and Winker et al. (2007) for opposing views. A benefit of the biological species concept to conservation is that it provides a degree of triage in terms of prioritizing resources at the global level. Restricting the species rank to populations known to be reproductively isolated or to have diverged to a level comparable to that shown by reproductively isolated populations (i.e., species by anyone's definition) allows limited conservation resources to be concentrated on those populations. For example, if one had a limited amount of funding to be divided evenly among Caribbean parrot species in the genus *Amazona*, using the classification based on the biological species concept would divide those funds among species that all differ strongly from one another and are species by any reasonable criterion. By contrast, elevating all diagnosable subspecies to species rank under the phylogenetic species concept would give equivalent taxonomic rank and funding, for example, to *Amazona leucocephala hesterna* (endemic to Cayman Brac and differing from nearby *A. l. caymanensis* of Gran Cayman only in having a larger patch of red in the belly plumage) as to the bizarrely plumaged, highly distinctive *A. guildingii* of St. Vincent. Advocates for conservation on Cayman Brac naturally would be pleased

with such an outcome, and so it is no surprise that among the most vocal advocates for the phylogenetic species concept are those devoted to the conservation of small areas or islands (e.g., Hazevoet 1996), whose cause benefits from raising every endemic subspecies to species rank. A more global view, however, would be that a prioritization scheme based in part on taxon rank is beneficial in that populations diagnosable only by characters that do not impede on gene flow, i.e., taxa ranked as subspecies under the biological species concept, do not receive the resources allocated to taxa ranked as species under the that concept.

The other criticism of the use of subspecies in defining conservation units is that many do not correspond to "historically significant groups" (Zink 2004). However, these groups are typically delimited only by patterns of shared mtDNA haplotypes (e.g., Zink et al. 2001). Whether such groups are the only historically significant groups, however, is open to discussion. Because these genetic markers are assumed to be neutral, by definition they have no biologically meaningful manifestation. Further, because of their matrilineal pattern of descent and because of the widely recognized problem of incomplete lineage sorting, these haplogroups represent only the history of perhaps one or two non-recombining genes (Edwards and Bensch 2009). Although such markers are useful tools for tracking aspects of population history, phenotypic markers also have the potential to do the same. Moreover, in contrast to haplotype differences, phenotypic markers have the potential to be biologically meaningful and should thus be of greater conservation concern (Crandall et al. 2000). Differences in pattern and coloration, for example, frequently correspond to abrupt discontinuities in gene flow in birds, a taxonomic class in which sexual selection has played a key role in diversification; their more subtle manifestation as diagnosable characters that mark subspecies boundaries gave rise to the phrase "incipient species" for some subspecies. To ignore this aspect of geographic variation and population biology only because of lack of correspondence to neutral mtDNA markers in vogue today should be regarded as myopic by those interested in patterns of biodiversity or the identification of units of conservation concern—or, indeed, the process of evolution.

A particularly disingenuous criticism of the biological species concept as an impediment to

conservation is the claim that it masks biodiversity. For example, Peterson (2006) denounced the biological species concept for overlooking numerous distinct populations but did not mention that under this concept all of those populations are named, as subspecies, and overlooked only if one restricts an analysis to the species rank. Thus, Peterson (2006) found much higher levels of species richness and unrecognized or underappreciated patterns of endemism by application of a diagnosability-based species concept; however, he did not point out that an analysis that included subspecies would have revealed the same patterns that he "discovered."

Application of the phylogenetic species concept produces two potentially severe problems for conservation. First, opponents of conservation would quickly discover that the definition of species had been changed to elevate more taxa to higher threat levels, with accusations of manipulation of the rules. Changing the definition would only fuel the suspicions of conservation opponents that scientists have abandoned objectivity in favor of a pro-conservation agenda. Second, elevating to species rank many taxa diagnosable only by characters that conservation opponents, the general public, and most biologists would justifiably label as trivial could diminish confidence in conservation science, undermine the credibility of taxonomists, and erode support for programs to protect threatened species.

Subspecies Are Overlooked as a Component of Biological Species

De Queiroz and Donoghue (1988:334) concluded that "no one species concept can meet the needs of all comparative biologists." I suggest that use of a biological species concept that identifies minimum diagnosable units as subspecies spans more of those needs than is appreciated. Debates over the merits of species concepts based on whether they emphasize reproductive isolation or minimum diagnosable units overlook that subspecies, an integral part of the biological species concept, are its minimum diagnosable units. Criticizing the biological species concept for not allowing analyses of basal evolutionary units overlooks that the subspecies rank is an integral part of the concept. The biological species concept encompasses units that fit the conceptual definition of phylogenetic species but calls these minimum diagnosable units subspecies rather

than species (if they are not ranked as biological species). Proponents of the phylogenetic species concept would point out that regardless of conceptual definitions, in practice many subspecies are not diagnosable units. As discussed above, this (1) is largely the consequence of incorrect application of the definition and (2) has to be dealt with regardless of whether these units are called subspecies or species. In fact, Phillimore et al. (2007) showed that analyses of subspecies as an index of intraspecific geographic differentiation within a species yield sensible results with respect to biogeographic influences on intraspecific variation. Under the biological species concept, classification with diagnosable units provides two levels of information: one that emphasizes genetic discontinuities (species) and another that emphasizes geographic units within the species identified by diagnostic characters (subspecies). Analyses that require terminal taxa can use populations ranked as subspecies (e.g., Cracraft 1985), whereas analyses based on active or potential barriers to gene flow can use the species rank. Geographic variation not partitioned into diagnosable units may occur within taxa ranked either as species or subspecies under the biological species concept.

A recurring misconception in some recently published papers is that the phylogenetic species concept reveals diversity and the biological species concept obscures it. For example, Reddy's (2008) application of the phylogenetic species concept to *Pteruthius*, currently considered to consist of 5 species under the biological species concept, first required determining which of the 23 recognized subspecies were diagnosable units; that is the same procedure that would be necessary under a modern reevaluation of the genus. Although the geography of sampling and sample sizes were not reported, Reddy found that 19 taxa were diagnosably distinct. She then claimed that this was "almost a four-fold increase in recognized diversity"; in fact, all of that diversity was recognized under the biological species concept, 5 as species and the other 14 as subspecies of those species. In terms of overall taxonomic diversity, this application of the phylogenetic species concept actually reduced the number of recognized taxa by some 15%. Once nondiagnosable taxa are identified and eliminated (a problem shared by all species concepts), the differences are not in diversity per se but in the ranks assigned to those units of diversity.

In summary, the biological species concept provides two levels of information, whereas the phylogenetic species concept provides one. The biological species concept incorporates the acquisition of diagnostic characters into its classification by ranking diagnosable populations minimally as subspecies. The biological species concept also incorporates reproductive isolation, acknowledged even by many proponents of the phylogenetic species concept as an important evolutionary step in the history of any lineage, by ranking such populations as species.

Subspecies and Human Perception

What would happen if the phylogenetic species concept's minimum diagnosable units were applied to *Homo sapiens*? Certainly, until recent decades, humans classified one another into racial groups thought to have diagnostic characters (e.g., Hall and Kelson 1959), and even today, one's race is a data field in many nonscientific categorization schemes. Research has shown that such schemes fail to classify individuals reliably and that, at the genetic level, ≤95% of all genetic variation is among-individual, not among-group. Nonetheless, despite rampant ongoing gene flow and the relatively recent origin of *Homo sapiens*, the residual variation may accurately predict region of origin and show strong geographic structuring. For example, even different groups of Native Americans differ strongly in haplotype frequencies (Malhi et al. 2003). Research on the genetic basis of human diseases has spawned an interest in ancestry-informative markers that predict the geographic origin of individual humans. Although complex computations are required to identify unique combinations of alleles, the geographic structure of this variation can identify individuals with respect to continent of origin (Rosenberg et al. 2002, Collins-Schramm et al. 2004, Mao et al. 2007, Li et al. 2008) and subregion (Tian et al. 2008a). Recently, Tian et al. (2008b), using a sample of European Americans categorized according to "self-reported" region of European descent, showed that principal component analysis of single nucleotide polymorphisms allowed accurate discrimination of individuals as either northern vs. southern European ancestry and found further evidence of structure within the northern European sample.

If geographic variation in *Homo sapiens* were sampled in the same limited way that it is in most birds, then application of the phylogenetic species concept to *Homo sapiens* would certainly produce "minimum diagnosable units" that are neither biologically nor socially acceptable as "species." However, the detailed structure of this variation, both phenotypic and genotypic, is sufficiently well studied that we can be sure that few if any character states analogous to those used in bird taxonomy would unambiguously diagnose any subpopulations of humans. Even today, after much global movement and genetic mixing, our own genetic and morphological (e.g., Shriver et al. 2003) diversity could be partitioned into an unknown number of diagnosable units by use of unique combinations of characters and allele frequency differences. By contrast, although cultural barriers prevent full application of the biological species concept to humans, this concept would consider all humans conspecific (*Homo sapiens*). In terms of perception and the absence of biologically based reproductive isolation, humans clearly think of themselves as belonging to one species, as defined by the biological species concept, despite marked geographic variation within *Homo sapiens*. Given that species definitions are scientifically untestable matters of taste (Brookfield 2002), human perception has spoken with resounding clarity that "species" are not minimum diagnosable units.

Common Ground

The debate over species and subspecies concepts is healthy, particularly in forcing a reevaluation of currently recognized subspecies names. I strongly agree with critics of the biological species concept that terminal taxa should be used in analyses of, for example, biogeography and biodiversity. The uncertainty of the diagnosability of many subspecies, especially in temperate North America, requires that anyone undertaking an analysis using terminal taxa must carefully scrutinize their diagnosability. Empirically, however, using named subspecies from Peters's *Check-list* series, even without critical evaluation (e.g., Cracraft 1985), successfully demarcates areas of endemism. So, if the sample is large enough and the error rate (nondiagnosable taxa) small enough, real patterns should emerge even if current subspecies names are taken as is.

I also strongly concur with McKitrick and Zink (1988) and others that subjective notions of whether a character is too trivial to use to diagnose

a taxon are unscientific. What matters is whether that character is a marker for a cohesive evolutionary unit, regardless of any known functional significance. If that character is "one extra hooklet on a barb of the seventh primary" (McKitrick and Zink 1988:9), and it passes the 95% diagnosability test, then it defines an entity worthy of a name, in my opinion.

Some defenders of the biological species concept worry that adoption of the phylogenetic species concept would lead to too many species (e.g., Mayr [1993] as cited by Zink and McKitrick [1995]). Preconceived notions of how many species there ought to be are scientifically indefensible. I echo McKitrick and Zink (1988) and Zink and McKitrick (1995) on the importance of letting the data determine the number of populations ranked as species. Even under the biological species concept, the number of species is increasing dramatically, particularly in the tropics, where many taxa formerly ranked as subspecies are elevated to species rank through careful study of vocalizations and population interactions at contact zones. For example, field studies of polytypic species of antbirds (Thamnophilidae), many using the comparative framework of Isler et al. (1998), have already elevated 31 taxa previously treated as subspecies to species rank. These 31 species, ranked as subspecies either by Peters (1951) or by Meyer de Schauensee (1970), were subsumed under 16 species names, including one, *Myrmeciza castanea*, considered a synonym of an existing subspecies. They include *Frederickena fulva*, *Cymbilaimus sanctaemariae*, *Thamnophilus zarumae*, *T. tenuepunctatus*, *T. cryptoleucus*, *T. atrinucha*, *T. stictocephalus*, *T. sticturus*, *T. pelzelni*, *T. ambiguus*, *Thamnomanes schistogynus*, *Dysithamnus leucostictus*, *Epinecrophylla spodionota*, *Myrmotherula ignota*, *M. multostriata*, *M. pacifica*, *Herpsilochmus atricapillus*, *H. motacilloides*, *H. dugandi*, *Drymophila rubricollis*, *Cercomacra laeta*, *Hypocnemis flavescens*, *H. peruviana*, *H. subflava*, *Hypocnemis ochrogyna*, *H. striata*, *Schistocichla humaythae*, *S. brunneiceps*, *S. rufifacies*, *S. saturata*, and *Myrmeciza castanea* (for references, see Zimmer and Isler 2003, Remsen et al. 2009). Species richness has thus increased by 88% in the 18 cases studied so far. Ongoing studies of other groups of thamnophilids will undoubtedly increase this tally, perhaps by as many as 50 species (M. L. Isler pers. comm.). If the thamnophilid results can be extrapolated to tropical avifaunas as a whole, many hundreds of subspecies will be elevated to species rank under the guidelines of the biological species concept when critical data become available. This does not represent a shift toward the phylogenetic species concept but, rather, an increase in data on reproductive isolation.

McKitrick and Zink (1998) provided a protocol, based in part on Zink and Remsen (1986), for applying the phylogenetic species concept to real-world situations. I suspect that they would also share my concern that few studies undertake the necessary steps to determine diagnosability. I disagree only semantically. I call the diagnosable units revealed by such analyses subspecies, not species.