We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 9

Species Delimitation: A Decade After the Renaissance

Arley Camargo and Jack Jr. Sites

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/52664

1. Introduction

A decade ago Sites and Marshall [1] described the empirical practice of species delimitation as "a Renaissance issue in systematic biology". At the time there was an odd disconnect between the two frequently stated empirical goals systematic biology: the discovery of: (1) monophyletic groups (clades) and relationships within these at all hierarchical levels above species; and (2) lineages (species); compared to the actual practice of the discipline. While much of systematic biology had been devoted to the first goal, the second goal had until recently been largely ignored [2], despite the fact that species are routinely used as the basic units of analysis in biogeography, ecology, evolutionary biology, and conservation biology [3,4]. However, Sites and Marshall [1] noted "signs of a Renaissance" at the time of their review, which was precipitated in part by others emphasizing the need to distinguish between a non-operational, ontological definition of species, versus the empirical (operational) data needed to test their reality [5-7]. De Queiroz [7] (p. 60) noted that "All modern species definitions either explicitly or implicitly equate species with segments of population level evolutionary lineages." De Queiroz also noted that this was a revised version of Simpson's "evolutionary species concept", which defines a species as "a lineage (an ancestral- descendent sequence of populations) evolving separately from others and with its own evolutionary role and tendencies" ([8], p. 153), and called this a General Lineage Concept (GLC) of species ([7], p. 65). De Queiroz [9] further emphasized that the multiple empirical criteria simply reflect the many contingent properties (differences in genetic or morphological features, adaptive zones or ecological niches, mate-recognition systems, reproductive compatibility, monophyly, etc.) of diverging populations associated with different evolutionary processes operating in various geographic contexts [10,11]. Sites and Marshall [1] noted that the emerging consensus among systematists and evolutionary biologists was based on the utility of this distinction (ontological definition vs. empirical species delimitation [SDL] meth-

INTECH
open science | open minds

© 2013 Camargo and Sites; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ods), and as also noted by de Queiroz [12], due to the contingencies of speciation processes, any single criterion or data set will artificially reduce the complexity of evolving lineages.

The subject matter of these and other reviews [12,13] focused strictly on methods of detecting various lines of evidence for lineage independence (reproductive isolation, ecological distinctiveness, diagnosability, monophyly, etc.), and since then new methods continue to be described [14], as do studies comparing the performance of some of these [14,15]. In 2006, the Society of Systematic Biologists (SSB'06) organized the first symposium dedicated to the topic of species delimitation [2]; 11 papers were presented and six of those published, including an update by referenced de Queiroz [16], which emphasized the distinction between the GLC as "separately evolving metapopulation lineages, or more specifically, with segments of such lineages", versus secondary biological attributes or properties of organisms that can be quantified to empirically test for species status. This is a crucial distinction because it clearly separates the conceptual issue of defining the species category from the methodological issues of delimiting species; previously these had been conflated with the result that properties used to infer species boundaries (the empirical test) were also sometimes regarded as necessary for defining a species (a conceptualization issue). The advantage of the unified GLC is that no specific biological attributes of a species are considered necessary properties – species may exist as segments of metapopulations lineages regardless of our ability to empirically delimit them. Prior to this clarification and the realization that many different properties are relevant to the issue of species delimitation [17], the alternative species "concepts" in which various biological attributes had accumulated in diverging lineages required these same attributes to be necessary properties of species. This led to a confusing situation in which a different property was considered necessary under each alternative concept (22 such "concepts" were identified by Mayden [6]), and a long and ultimately non-productive debate about species definitions. Now most of these earlier "concepts" can be viewed as secondary species criteria that provide evidence of lineage separation.

Recently, Hausdorf [17] argued for an up-dated ontological species concept, based in new insights into speciation processes, particularly evidence that reproductive barriers are semipermeable to some gene flow, and that speciation may occur despite ongoing gene flow between diverging populations [18-23]. Two other lines of evidence are relevant to the point of re-visiting the GLC: (1) findings of polyphyletic species of animals, due to parallel speciation in which similar traits conferring reproductive isolation arise separately in closely related populations [24,25], or in plants, due to recurrent polyploidization in different populations of the ancestral species [26,27]; and (2) discoveries of uniparental organisms that can be characterized as distinct units resembling species of biparental organisms [28]. We cannot resolve all of these larger issues here, but we return to some of the general points raised by Hausdorf [17] in the discussion.

Empirically, species delimitation continues to be a topic of increasing interest in evolutionary biology. A reference search in the ISI Web of Science with the keyword 'species delimitation' retrieved 227 articles published since 2000, of which 60% were published after 2008. Less than 10 articles per year were published between 2000 and 2005; subse-

http://dx.doi.org/10.5772/52664

quently 10-20 articles per year between 2006 and 2008, and after 2008 the publication rate reached ~ 40 articles (Figure 1A). These increases include papers describing new SDL methods, or using existing methods with novel data sets and/or applications to new taxa. Because new SDL methods apply the same coalescent models developed for species tree estimation and usually lead to the discovery of morphologically 'cryptic' species, we also searched for references with the keywords 'species tree' and 'cryptic species'. During the same period of time, papers about 'species trees' were few until 2007, increased between 2008 and 2010 to 5-10 articles per year, and nearly doubled to >20 papers last year (Figure 1B). Publications referring to 'cryptic species' show a constant increase from 20 papers/year in 2000 to 90 papers/year in 2011 with the larger annual increase between 2010 and 2011 (Figure 1C). These publication trends suggest that the recent paradigm shift in phylogenetic systematics to incorporate species trees (29) is having a positive impact on the development of new SDL methods, which are gradually being incorporated into integrative taxonomic practices for the discovery of cryptic species diversity [30].

2. Body

2.1. Short history of some early methods

Sites and Marshall [1,13] separated SDL methods into non-tree and tree-based approaches, and included among the former (1) pairwise genetic distances that could be tested for either correlations with reproductive isolation [31,32], morphological distances [33], or geographic distances [34]; (2) gene flow statistics to estimate the extent of gene flow across hybrid zones [35]; (3) fixed alternative character states as an indicator of no gene flow in a "population aggregation analysis" (PAA; [36]); (4) the presence of heterozygous genotypes as an indicator of a "field for recombination" [37]; and (5) genotypic clusters [38].

Early tree-based methods included: (1) three versions of the phylogenetic species "concept" based on apomorphy, or lineage splitting, or node-based criteria, following the terminology of Brooks and McLennan [39]; (2) cladistic haplotype aggregation [40]; (3) molecular-morphological assessments using dichotomous flow charts [41]; (4) genealogical exclusivity [42]; and (5) an extension of the nested clade analysis [43] that includes tests of species boundaries [44]. The data sets in these early studies most often included genotypes resolved from multilocus isozymes [15], morphological (usually meristic) characters, and with few exceptions [45,46], mitochondrial DNA (mtDNA) sequences. An innovative phylogenetic method described by Pons *et al.* [14] was based on a likelihood analysis of the mtDNA gene tree that estimates the inflection point between species-level (speciation-extinction) and population-level (coalescent) evolutionary processes, and demonstrated that groups delimited by this approach were generally concordant with geographic distributions and morphologically recognized species. This was one of a small number of early studies comparing the performance of multiple SDL methods (see also [15, 40, 45,46]).

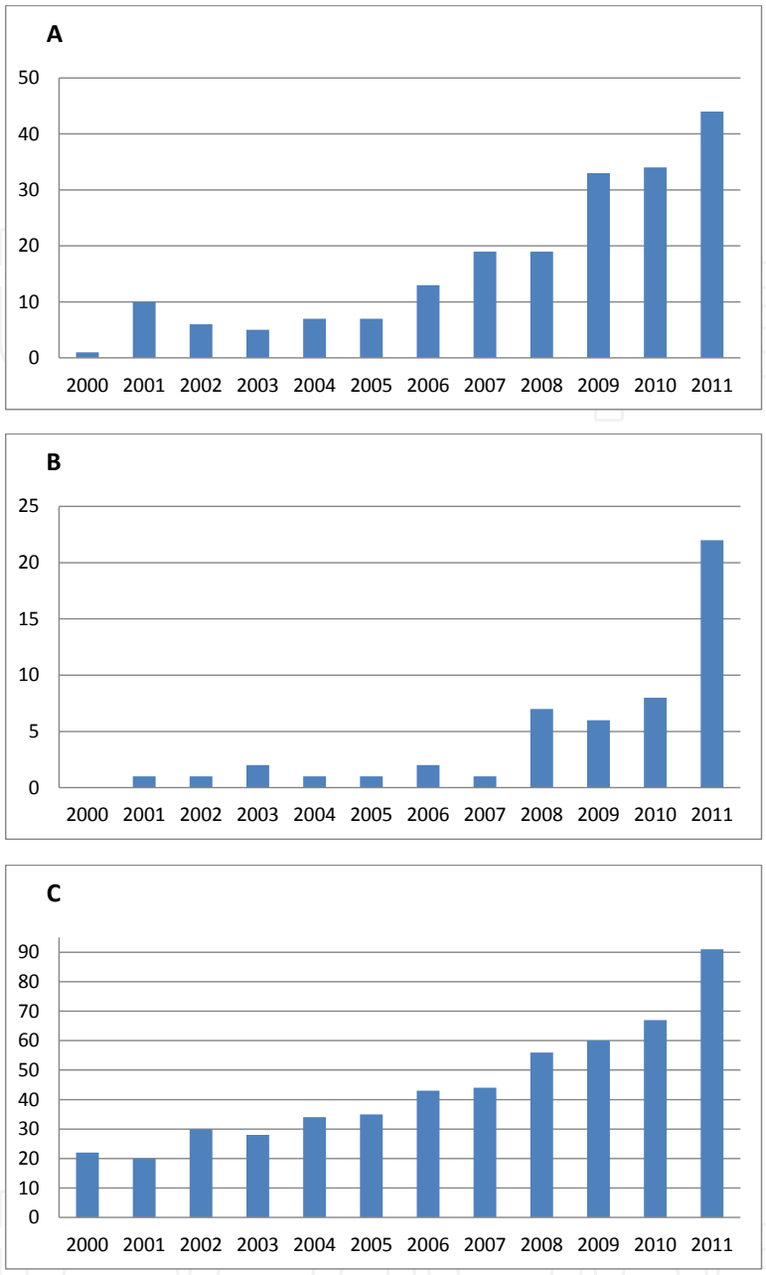


Figure 1. The number of papers with (A) "species delimitation"; (B) "species tree", or (C) "cryptic species" in the title, published from 2000 – 2011.

The published contributions of SSB'06 symposium [2] included several novel SDL methods, the first method [47] described a coalescent approach to estimating species boundaries based on multiple unlinked gene trees, and that does not require species to be characterized by reciprocal monophyly. This is an explicitly model-based approach that accommodates stochastic variance of the gene sorting process by linking estimates of two key parameters, a range of estimates of effective population sizes relative to possible divergence times. This type of gene tree-coalescence approach also directly links population genetic SDL methods to phylogenetic inference at deeper levels of divergence, which has been identified as a

"new paradigm" in systematics [29]. In this same issue, Shaffer and Thomson [48] introduced a population genetic SDL based on large sets of single nucleotide polymorphisms (SNPs), which would be most suited to delimiting very young species. Finally, this volume included two more novel SDL methods, both in this case using ecological and distributional data in novel ways to model "niche envelopes" that can augment molecular or morphological data in species delimitation [49-51].

2.2. Recent progress

2.2.1. New methods & new theory

New empirical SDL methods continue to be developed, based on multiple lines of evidence and multiple statistical methods. Among some of these is the approach of Bond and Stockman [52] that is especially relevant to highly geographically-structured populations in which traditional sequence-only data sets are likely to recover large numbers of well-defined, well-supported, and geographically concordant/genetically divergent-but-morphologically cryptic populations (species). These authors describe a framework for testing potential genetic and ecological exchangeability as a means of delimiting cohesion species [53], and present an example in trapdoor spiders of the *Aptostichus atomarius* complex. A completely different approach [54] is based on statistical tests of both population structure [48] and genealogical exclusivity [54] of nuclear loci, to test species provisionally identified from well-supported mtDNA haploclades; the focal taxa in this study were Malagasy mouse lemurs (55; genus *Microcebus*). As a third example, Puillandre et al. [56] described a four-step approach to "generating robust speciation hypotheses" in mollusk family Turridae (genus *Gemmula*) based on: (a) collection of the COI DNA barcode gene for GMYC [14] and ABGD (Automatic Barcode Gap Discovery; [57]) analyses; coupled with (b) nuclear gene (rRNA 28S), morphological, geographical and bathymetrical data, to redefine species boundaries in this clade. This protocol more than doubled the previously known species diversity in *Gemmula*, and may be useful for large-scale SDL in hyperdiverse groups. A few additional examples include genotype-based methods for dominant and co-dominant multi-locus markers [58], combined estimates of divergence times and gene flow to discriminate intraspecific from interspecific patterns [59], and an extension of the R package GENELAND to include genetic, phenotypic (morphometric), and geographic data for delimitation of populations and species [60].

The recent merge of coalescent theory with phylogenetics has driven a new generation of SDL methods and a new paradigm in systematics [29]. This new theoretical framework, and its derived analytical applications, was in part required as a solution for accommodating the observed conflict among genealogies from multiple loci (gene trees) with the underlying population-level genealogies (species trees) [61]. A multi-species or 'censored' model was formulated to account for this discordance by considering each branch of the species tree as a separate coalescent model and by connecting them into a population-level genealogy following the topology of the species tree [62,63]. Under this new approach, two major key innovations over the classic phylogenetic methods were achieved. First, multiple individual

samples can be assigned to a single species and the estimated phylogeny represents the speciation history of ancestral and descendant species-level lineages, in contrast to the gene genealogies estimated with individual samples. Second, because the coalescent process of each gene tree is dependent upon parameters of its containing species tree, this approach can co-estimate gene and species tree simultaneously, by-passing the task of calculating a consensus tree or estimating a phylogeny from a concatenated dataset. This new theoretical framework allows prediction of the probability distribution of gene trees given the species tree, and consequently, several methods were developed for estimating species trees from a collection of multiple gene trees under different algorithms [64,65]. Based on these new methods, a generation of fully-coalescent SDL methods was introduced that consisted of selecting the best species-tree model from a set of alternative models that represent different hypotheses of species limits. For instance, one approach finds the maximum-likelihood for the full species tree (all species are hypothesized as separate lineages) and for alternative species trees (two or more species at a given node are collapsed into one), and then selects the best model using Akaike information criteria, assuming fixed gene trees and constant population sizes along the species tree (SpeDeSTEM; [66]).

Another SDL method consists of sampling from the Bayesian posterior distribution of species delimitation models using reversible-jump Markov chain Monte Carlo (rjMCMC) with the program BP&P 2.1 [67]. This approach accomodates gene tree uncertainty and variable population sizes, but a "known" species tree must be provided *a priori*. In addition, heuristic and/or semi-parametric approaches have been developed for: resolving the boundary between coalescent and speciation processes using single gene trees (generalized mixed Yule-coalescent, [14]), finding both the optimum species tree and species limits via minimization of gene tree conflict and intraspecific structure (Brownie; [68]), and selection of SDL models using approximate Bayesian computation (ABC) [69]. Other tree-based [54] and non-tree-based [58] SDL methods that can handle multiple loci with limited variation have been applied with success. In addition, there has been also a resurgence of morphology-based SDL using multivariate techniques in a hypothesis-driven statistical framework [60,70].

2.2.2. New kinds of data

The development of new multi-species/multi-locus SDL methods was also in part due to the demand of efficient analytical tools to handle the rapidly increasing amounts of molecular data collected with modern techniques. New SDL methods should be able to handle tens of loci for multiple individuals derived from the development and screening of anonymous nuclear loci (ANL), introns, and protein-coding loci using genomic resources [71-73]. However, these new SDL methods are inadequate to analyze the influx of whole-genome data that have started to be collected for non-model organisms via next-generation sequencing (NGS) technologies ([74-76]; e.g, genome of the lizard *Anolis carolinensis*; [77]). NGS technologies have been recently applied to development of thousands of gene regions spanning multiple divergence times [78], or loci targeted for "shallow-scale" phylogenetic/phylogeographic studies [79], and microsatellites [80] or SNPs [81] for extremely shallow phylogeo-

graphic histories [82]. The microsatellite or SNP data should be useful for genotyping individuals for SDL studies of very young species [48].

More efficient and less costly whole-genome sequencing is becoming available on a regular basis, a trend that started with the first-generation technology (Sanger capillary-sequencing), followed by the second-generation (i.e., SOLiD 454, Illumina, Solexa, etc; [83]), and continuing today with the recently introduced third-generation 'nanopore' sequencing [84,85]. A significant by-product of these single-molecule sequencing methods is their ability to automatically resolve the allelic phases of heterozygotes, in contrast to the time-consuming phase estimation and/or cloning required after direct dideoxy-sequencing [86]. In addition, the uniform sampling of hundreds of loci across the genome can help identifying "outlier" loci via genome scans, which can represent candidate genes with fitness value, subject to selection and linked to processes such as ecological speciation [87].

2.2.3. Advantages of Multi-Species Coalescent-Based Methods (MSCM)

Model-based.–Because these SDL are based in the multi-species coalescent model, the likelihood of the data can be evaluated to find maximum-likelihood and posterior probability estimates of parameters and testing alternative SDL models under different criteria (e.g., likelihood-ratio test, Akaike information criterion, Bayes factors [46,88]). More importantly, these methods implement SDL in a hypothesis-testing framework, and taking into account uncertainty due to genetic processes and insufficient sampling [89,90]. In addition, coalescent simulations generated under a null hypothesis of no-speciation and the alternative hypothesis of speciation can be used for evaluating the performance of these methods based on estimations of inferential errors (type I and II errors, see [91]). For example, the accuracy of three coalescent-based SDL (SpeDeSTEM, BP&P, and ABC) has been compared using simulations under a model of speciation for variable sampling densities and parameter values to estimate type II error (i.e., failing to reject no-speciation when it is false) across a range of conditions [66,69,92]. When there is no migration, SpeDeSTEM can delimit species that have diverged as recently as $0.5N_e$ generations ago using only 5 loci and 5 alleles per species [66] while BP&P could detect speciation at shorter divergence times ($0.4N_e$ generations ago) with the same sampling design [67,92]. In agreement with these results, a comparison under identical simulation conditions showed that BPP outperformed SpeDeSTEM (and also ABC) when speciation takes place with or without gene flow [69]. In spite of these simulations covering different speciation scenarios, sampling designs, and SDLs, the practical question of the appropriate balance between number of loci and number of alleles sequenced has not been explicitly explored until now. Below, in the last section of this chapter, we performed some simulations for a preliminary evaluation of the relative benefits of sampling more alleles vs. loci for accurate species delimitation.

Neutral loci.–These markers should be insensitive to 'phenotypic plasticity', the phenotypic response to environmental variation that is not genetically-based (in contrast to adaptive variation), which could bias morphological-based taxonomy. Environmental variation in different parts of the range can lead to a plastic phenotypic response, which can be revealed and distinguished from local adaptation via reciprocal transplant or 'common garden' ex-

periments [93]. In these cases, morphological variation as a result of this plastic response could be used as a criterion to delimit species, while neutral markers would indicate that there is no genetic differentiation [94,95]. In contrast, in cases of morphologically-cryptic species due to for example to niche conservatism [96], genetic divergence and lineage sorting is expected to occur in neutral markers due to independent evolution in isolation, and those markers with higher mutation rates and smaller effective population size (e.g. mtDNA) should be ideal for species delimitation [97,98]. Moreover, it has been suggested that neutral loci will also differ in their usefulness for species delimitation since those with higher rates of intra-specific gene flow will be less sensitive to the effect of inter-specific introgression [99]. However, the mitochondrial locus does not always meet assumptions of neutrality [100], and it frequently introgresses across species boundaries [101], so in our view it should be used to identify "candidate species" [102], which can then be verified with independent lines of evidence [103].

Repeatability.–The results of a SDL analysis can be replicated exactly when using the same data and the same analytical methods, which eliminates much of the subjectivity and/or investigator bias for/against certain kinds of data (morphology vs. molecular, etc.). Because these methods rely on explicit predictions about genealogical patterns under alternative models of lineage divergence, it is possible to carry out species delimitation in a more objective and bias-free fashion compared to diagnosability-based SDL methods [90]. In addition, because inferences are dependent upon a specific sampling design and the method used, one can make explicit statements about how robust a given species delimitation method is to variation in these parameters, and to violations of the method's assumptions.

Universality.–The same SDL method and the same kind of data (i.e, DNA sequences or gene trees from homologous regions of the genome) can be used for SDL across different taxa, making these approaches comparable across all parts of the Tree of Life, as long as the assumptions of the method are reasonable for the taxon under study (see below). Another advantage associated with the use of neutral markers in coalescent-based SDLs is related to the standard criterion used for assigning species status across a variety of taxa when using the same markers and analyses [90], assuming that these markers offer similar resolving power. This is a desirable property for a SDL method since a uniform criterion implies that the species level could be compared readily among different higher-level taxa, thereby allowing meaningful analyses of species diversity among communities typical of ecological studies [91].

2.2.4. Disadvantages of MSCM

Many of the advantages listed above also impose some limitations of MSCM and other SDL methods for different reasons. First, these are *model-based* methods, and any violations of assumptions of the standard coalescent are expected to introduce inference errors. For instance, and most relevant to the SDL problem, while the standard coalescent assumes panmixis within populations, it is clear that in most natural populations, there is almost always some degree of population structure (i.e, demes connected by limited gene flow). In fact, a recent study using the Brownie's SDL method found that more dense sampling in-

http://dx.doi.org/10.5772/52664

creased the chances of detecting population structure, supporting more species boundaries, and consequently, inflating estimates of the number of species [104]. Thus, MSCM could be more prone to split a single real species into multiple lineages due to intra-specific population structure alone, increasing type I error (i.e., rejecting a true hypothesis of a single species), and leading to 'taxonomic inflation' [91]. Fortunately, some flexible MSCM methods allow incorporating population structure within species via coalescent simulation of island, stepping-stone, and other potential models, and subsequent comparison of SDL hypotheses with ABC approaches [69].

Another frequent assumption of most MSCM is that species have diverged from a common ancestral species without gene flow even though speciation with gene flow seems to be rather common in nature, especially in cases of ecological speciation [22,95,105]. While these methods ignore the effects of gene flow, simulation testing has shown that some of them are relatively robust to low levels of gene flow [66,92], and that its impact on delimitation accuracy is ameliorated when gene flow is explicitly incorporated in the speciation model [69]. This result supports the suggestion that, in order to distinguish between species- and population-level differentiation, it is necessary to jointly consider the two components of the divergence process: time since splitting and gene flow after divergence [59].

Second, *sampling effort* is well known to strongly impact coalescent-based and other SDL methods. A number of studies evaluating the accuracy of several MSCM methods suggest that limited sampling of loci and sequences will decrease the probability of detecting speciation when this hypothesis is correct [66,69,92] and consequently, increasing type II error [91]. In addition, these simulation results also support the intuitive idea that the problem of insufficient sampling becomes more serious when SDL is more difficult: shorter divergence times, larger population size, and increasing inter-specific gene flow. However, more simulations are necessary to evaluate the appropriate balance between sampling intensity and design (e.g., geographic vs. genealogical dimensions, [91]) for different parameter configurations, in a power analysis context to provide further guidance to empirical studies [106]. In addition to limited geographic sampling, the collected sequence data also impose a limit to the amount of genetic data available for analysis. In the next section, we explore how accuracy in species delimitation responds to variable sampling of loci and alleles for a fixed sequencing effort.

Third, coalescent-based SDL approaches assume *selective neutrality* of gene regions used, but divergent selection on ecological traits, across habitats or along an environmental gradient, can lead to local adaptation and correlated reproductive isolation in a process of ecological speciation [95,107]. Phenotypic divergence can be so fast that mutation rates could produce little or no differentiation at all in neutral markers used in SDL approaches. Only those "outlier loci" under selection revealed by genomic scans, which are potentially associated with the selected traits, would be appropriate markers under these scenarios [87].

Fourth, as in other methods *data conflict* may be evident when multiple data sets are used. These SDL methods are not expected to resolve the discordance among different kinds of data sets (i.e., morphological, behavioral, ecological, molecular, etc.) since they typically use sequence data or gene trees from presumably neutral loci. However, Bayesian approaches

have the potential of incorporating previous information about species limits derived from non-molecular data into prior distributions of genetic-based analyses [67].

Fifth, there may be conflicts with *traditional taxonomic practices.* The discovery of new cryptic species with coalescent-based SDL in a statistical framework, is still insufficient for formal taxonomic descriptions, since nomenclatural rules still require traditional morphology-based diagnoses [108,109]. While these methods will help diagnosing new cryptic diversity, many taxonomists will be reluctant to formally describe new species based on molecular-data alone, which ultimately will further expand the 'taxonomy-phylogeny' gap [91]. While the description of cryptic species is complicated by the lack of morphological diagnostic characters, another difficulty relies in the inability of MSCM to assign newly collected specimens to species (i.e., taxonomic determination) unless new analyses are carried out to re-evaluate species limits.

3. Future directions

Statistical testing of SDL.–The ongoing surge in the new generation of SDL methods will probably encourage many taxonomists to apply these methods empirically, especially for recently evolved, cryptic taxa that cannot be delimited with other data. The ability to frame species limits as statistical hypotheses that can be tested objectively with multi-locus and multi-species analyses make these new SDL methods very appealing for empirical systematists in the context of an 'integrative taxonomy' [4,110,111]. In addition to empirical application to real data sets, we also expect that more simulation studies will be carried out to compare the performance of different data sets, under different methods/assumptions, and for variable sampling designs, using statistical power analyses. Previous studies have compared methods for a limited set of parameter conditions (e.g., usually population size has been assumed to remain constant) and have examined the effect of increased sampling effort for loci or sequences separately. However, performance of these SDL methods has not been evaluated for a variable sampling design and a fixed sampling effort; in other words, what should be the optimal balance between number of loci and number of sequences when the total number of sequenced base pairs (bp) is the same?

In order to provide a preliminary evaluation of the impact of sampling design on performance of new SDM, we simulated coalescent genealogies with the program ms [112] and sequence data with the program Seq-Gen [113] for a speciation model between species A and B for three increasing divergence times: 0.25, 0.5, and $1N_e$ (Figure 2A). We assumed a constant θ per site = 0.01, 500 bp per locus, and ~50 variable sites per locus. For each divergence time, we simulated 5 combinations of number of loci (1, 2, 4, 10, and 20) and number of sequences per species (1, 2, 5, 10, and 20) while keeping the total sequencing effort constant (20 sequences per species). We simulated 100 replicates for each sampling treatment which were analyzed with BP&P to calculate the mean speciation probability between species A and B across replicates, which represents the accuracy of the method (i.e., the probability of detecting speciation when it is the true hypothesis). We also simulated a no-speciation model

http://dx.doi.org/10.5772/52664

where sequences from species A and B were collapsed into a single lineage, and repeated the same sampling and analytical procedure to examine the performance of the method based on a plot of true positive and false positives rates (i.e., ROC plot; [114]).

The results show that under the conditions examined, more sequences per species is better than more loci at least in the range of 1-20 loci and sequences per species (Figure 2B). The ROC plots for the 5 sampling treatments at a divergence time of $0.5N_e$ show that performance is higher (i.e., area under ROC curve is larger) when sampling 20 sequences for 1 locus or 10 sequences for 2 loci, but performance gradually decreased with more loci and fewer sequences (Figure 2C). These results are congruent with the impact of sampling design on the accuracy of species-tree methods (STM) at shallow divergence times [115,116], which is an expected outcome because both STM and SDL methods share the same basic multispecies coalescent model [67,117]. However, our results are contingent upon the conditions simulated, in particular the assumptions of panmixia within species, and a constant θ across the species tree. This second assumption is a critical parameter of coalescent models, which can be estimated more accurately with a larger sample of loci [118]. Our attempt with this simulation example was to show how we can evaluate the performance of a SDL method under a variety of sampling conditions based on a power analysis, and that this same approach can be applied for comparisons across different SDL methods and more complex speciation scenarios than those that have been examined so far.

Population and species delimitation.–The application of coalescent-based SDM, which can delimit species at very shallow levels of divergence [66,69,92] should reduce the 'taxonomy-phylogeny' gap and help decrease the type I error of biological-species criteria that often fail to detect species, when reproductive isolation is not yet complete [91]. Thus, coalescent-based SDL methods will probably help to delimit entities, name taxonomic units, and give appropriate conservation priority to the increasing amounts of cryptic diversity being discovered in nature [91]. On the other hand, MLCM should be used with caution to avoid confusing species-level divergence with intra-specific population structure and therefore, over-splitting lineages, with serious consequences for conservation science since limited resources would be potentially wasted due to bad taxonomy [91].

A potential protocol for an informed species delimitation approach that takes into account population structure, could consist of first applying a clustering/population aggregation method to identify the smaller clusters of individuals under a population genetics criterion based on genotype or allele frequency data ('e.g., Structure 48, 58, 60). Subsequently, a SDL method can be applied to test if these clusters also represent independent evolutionary lineages based on the pattern of allele coalescence in gene genealogies (e.g. BP&P). Because initial population divergence starts with differentiation in allele frequencies and secondly, with random lineage sorting and mutation that further differentiates lineages during speciation [59], population genetics approaches are expected to detect lineages earlier than SDL approaches. For example, an empirical analysis of West African forest geckos (*Hemidactylus fasciatus*) found ~10 populations with Structure, which were considered as 'candidate' species in a subsequent BPP analysis that collapsed them into 4 species [119]. This two-stage approach would provide a consistent and standard criterion for distinguishing between popu-

lation- and species-level divergence, a threshold that has been difficult to resolve with genetic parameters measuring amounts of evolutionary differentiation [59].

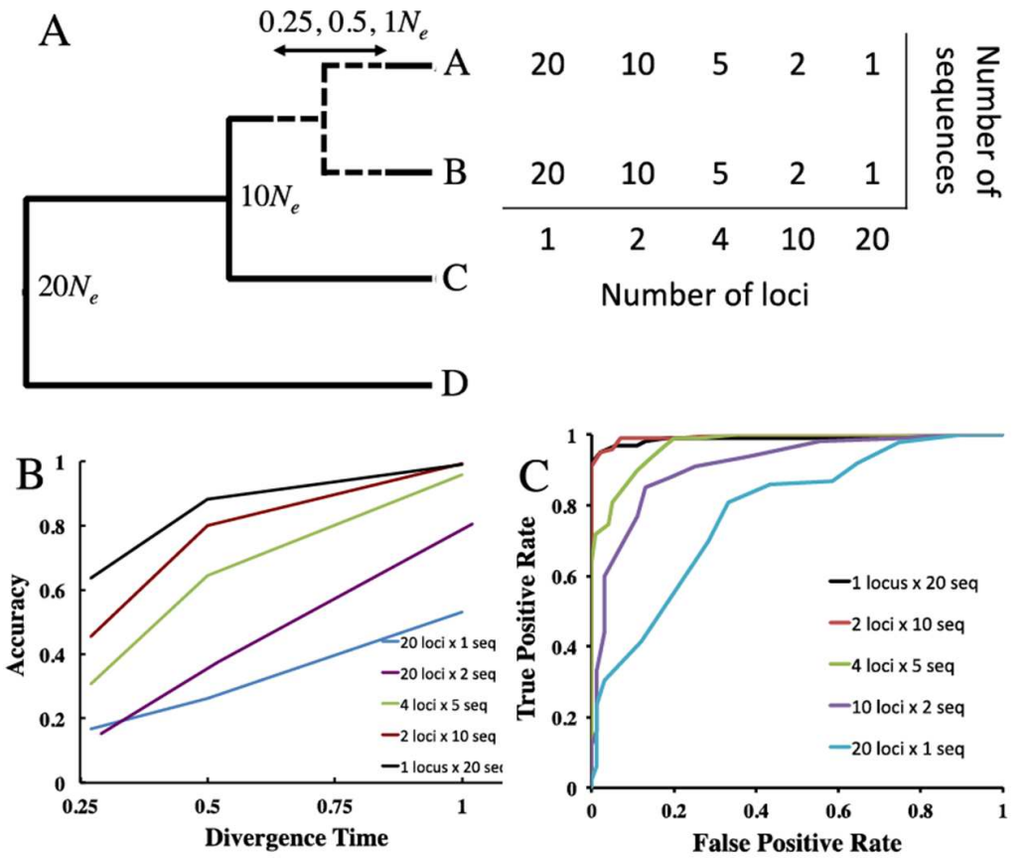


Figure 2. Simulation-based testing of the accuracy of BP&P to detect speciation between species A and B using five alternative sampling designs with the same sequencing effort and at three increasing divergence times (0.25, 0.5, and 1Ne) (A). Plot of accuracy and divergence time for each sampling design (B).ROC plot for each sampling design when divergence time = 0.5Ne (C).

The next generation of SDL methods.–We have emphasized that species delimitation should take into account the speciation processes that have shaped the patterns of trait divergence in genetic, morphological, and ecological data [89]. In a process-oriented classification of modes of speciation, we can distinguish between 'passive' modes driven by random divergence associated with the classic allopatric models, and the 'adaptive' modes of speciation. The formulation of a null hypothesis of speciation due to stochastic forces (i.e., 'passive divergence' or 'drift-only' model) should facilitate testing this mode of speciation, because rejecting this hypothesis is probably easier than demonstrating 'adaptive' speciation due to deterministic processes [120]. In nature, both speciation models appear to interact and work in concert during diversification of closely related lineages [121,122]. Adaptive speciation in turn can be subdivided into 'ecological' speciation, reproductive isolation due to disruptive natural selection operating on ecological traits [95], and speciation due to sexual selection that results in divergent mating prefer-

http://dx.doi.org/10.5772/52664

ences and assortative mating [123]. In theory, both kinds of selection seem to be necessary to drive speciation to completion [124], and limited empirical data supports the role of this interaction during diversification [125]. Due to this variety in speciation processes, we should expect different patterns of trait divergence, and consequently, different kinds of data would be more appropriate for species delimitation under each speciation scenario. Therefore, relying on any single kind of trait could potentially miss the detection of a speciation event, for example using exclusively morphological data will fail to recognize cryptic species. Similarly, if we use only the typical neutral genetic markers of phylogeography and population genetics, we could miss many instances of ecological speciation that takes place in contemporary time scales [126], and/or without divergence in neutral loci [127].

4. Conclusion

There is an ongoing genomics revolution for the study of adaptation in ecological and evolutionary non-model organisms derived from (NGS) technologies [76,128]. Decreasing sequencing costs and new protocols for discovering and screening thousands of markers scattered throughout the genome [79], is now allowing application of population genomics approaches to identifying the candidate loci underlying adaptive traits with ecological significance [87]. In fact, recent studies have found genomic regions and/or specific loci related to repeated local adaptation, population divergence, and reproductive isolation between ecotypes in different habitats or hosts [129,130]. We anticipate that these 'speciation genomics' approaches will become more common in non-model organisms and will provide a basis for species delimitation in scenarios of adaptive speciation SDL methods, complementing current SDL methods. Moreover, this plurality of criteria for species delimitation based on multiple kinds of traits is consistent with the GLC of species that views these organismal traits as evolving in different temporal order depending on how speciation has actually taken place [9,12]. In addition, it is also compatible with the more recent 'differential fitness' concept, which is based on those organismal features of one species that have negative fitness effects in other species and cannot be exchanged upon contact [17].

Acknowledgments

AC acknowledges a postdoctoral fellowship from CONICET (Argentina). For financial support we thank thank NSF awards OISE 0530267 and AToL 0334966 to JWS, as well as BYU graduate research and graduate mentoring awards, and student research awards from the Society of Systematic Biologists and the Society for the Study of Amphibians and Reptiles, to AC. We both also received support from the BYU Dept. of Biology and the Bean Life Science Museum.

Author details

Arley Camargo[1] and Jack Jr. Sites[2*]

*Address all correspondence to: jack_sites@byu.edu

1 Unidad de Diversidad, Sistemática y Evolución, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas, Puerto Madryn, Chubut, Argentina

2 Department of Biology and Bean Life Science Museum, Brigham Young University, Provo, Utah, USA

References

[1] Sites JW, Jr., Marshall JC. Delimiting species: a Renaissance issue in systematic biology. Trends in Ecology and Evolution 2003;18:462-420.

[2] Wiens JJ. Species delimitation: new approaches for discovering diversity. Systematic Biology 2007;55: 875-878.

[3] Agapow PM, Bininda-Edmonds ORP, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A. The impact of species concept on biodiversity studies. Quarterly Reviews in Biology 2004;79:161-179.

[4] Padial JM, Castroviejo-Fisher S, Köhler J, Vilá C, Chaparro JC, De la Riva I. Deciphering the products of evolution at the species level: the need for an integrative taxonomy. ZoologicaScripta 2009;38:431-447.

[5] Frost DE, Kluge AG. A consideration of epistemology in systematic biology, with special reference to species. Cladistics 1994;10:259-294.

[6] Mayden RL. A hierarchy of species concepts: the denoument in the saga of the species of the species problem. In: Claridge, M.F., H.A. Dawah, and M.R. Wilson (eds.) Species: The Units of Biodiversity. Chapman and Hall; 1997. p381-424.

[7] deQueiroz K. The general lineage concept of species, species criteria, and the process of speciation. In: DJ Howard, Berlocher SH (eds.) Endless Forms: Species and Speciation. Oxford University Press;1998. p.57-75.

[8] Simpson, G.G. 1961. Principles of Animal Taxonomy. Columbia Univ. Press, 247 pp.

[9] deQueiroz K. Ernst Mayr and the modern concept of species. Proceedings of the National Academy of Sciences of the USA 2005;102:6600-6607.

[10] Coyne JA, Orr HA. Speciation. Sunderland, MA: Sinauer Associates; 2004.

[11] Price TD. Speciation in Birds. Greenwood Village, CO: Roberts & Co.; 2008.

[12] deQueiroz K. A unified concept of species and its consequences for the future of taxonomy. Proceedings of the California Academy of Sciences 2005;56:195-215.

[13] Sites JW, Jr., Marshall JC. Operational criteria for delimiting species. Annual Review of Ecology, Evolution, and Systematics 2004;35:199-227.

[14] Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, and Vogler AP. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 2006;55:595-609.

[15] Marshall JC, Arévalo E, Benavides E, Sites JL, Sites, Jr. JW. Delimiting species: comparing methods for Mendelian loci using lizards of the *Sceloporus grammicus* complex (Phrynosomatidae). Evolution2006; 60:1050-1065.

[16] de Queiroz K. Species concepts and species delimitation. Systematic Biology 2007;56:879-886.

[17] Hausdorf B. Progress toward a general species concept. Evolution 2011;65:923-931.

[18] Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG. Understanding and confronting species uncertainty in biology and conservation. Trends in Ecology and Evolution 2003;18:597-603.

[19] Wu C-I. The genic view of the process of speciation. Journal of Evolutionary Biology 2001;14:851-856.

[20] Wu C-I, Ting CT. Genes and speciation. Nature Reviews Genetics 2004;5:114-122.

[21] Butlin R, The Marie Curie SPECIATION Network. What do we need to know about speciation? Trends in Ecology and Evolution 2011;27:27-39.

[22] Nosil P. Speciation with gene flow could be common. Molecular Ecology 2008;17:2103-2106.

[23] Nosil P, Feder JL. Genomic divergence during speciation: causes and consequences. Philosophical Transactions of the Royal Society, Series B 2012;367:332-342.

[24] Nosil P, Crespi BJ, Sandoval CP. Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 2002;417:440-443.

[25] Rundle HD, Nagel L, Boughman JW, Schluter D. Natural selection and parallel speciation in sympatric sticklebacks. Science 2000;287:306-308.

[26] Soltis DE, Soltis PS. Polyploidy: recurrent formation and genome evolution. Trends in Ecology and Evolution 1999;14:348-352.

[27] Soltis PS, Soltis DE. The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences of the USA 2000;97:7051-7057.

[28] Fontaneto D, Herniou EA, Boschetti C, Caprioli M, Melone G, Ricci C, Barraclough TG. Independently evolving species in asexual bdelloid rotifers. PLoS Biology 2007;5:914-921.

[29] Edwards SV. Is a new and general theory of molecular systematics emerging? Evolution 2009;63:1-19.

[30] Padial JM, Miralles A, De la Riva I, Vences M. The integrative future of taxonomy. Frontiers in Zoology 2010;7:16.

[31] Highton RR. Biochemical evolution in the slimy salamanders of the *Plethodonglutinosus* complex in the eastern United States. Part 1. Geographic protein variation. Illinois Biological Monographs 1989;57: 1-78.

[32] Highton RR. Taxonomic treatment of genetically differentiated populations. Herpetologica 1990;46:114-121.

[33] Puorto G, Salomão MG, Theakston RDG, Thorpe RS. Combining mitochondrial DNA sequences and morphological data to infer species boundaries: phylogeography of lanceheadpitvipers in the Brazilian Atlantic forest, and the status of *Bothrops pradoi* (Squamata: Serpentes: Vipiridae). Journal of Evolutionary Biology 2001;14:527-538.

[34] Good DA, Wake DB. Geographic variation and speciation in the torrent salamanders of the genus *Rhyacotriton* (Caudata: Rhyacotritonidae). University of California Publications in Zoology 1992;126:1-91.

[35] Porter AH. Testing nominal species boundaries using gene flow statistics: taxonomy of two hybridizing admiral butterflies (*Limenitis*: Nymphalidae). Systematic Zoology 1990;39:131-147.

[36] Davis JI, Nixon KC. Populations, genetic variation, and the delimitation of phylogenetic species. Systematic Biology 1992;41:421-435.

[37] Doyle J. The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. Systematic Botany 1995;20:574-588.

[38] Mallet J. A species definition for the modern synthesis. Trends in Ecology and Evolution 1995;10:294-299.

[39] Brooks DR, McLennan DA. Species: turning a conundrum into a research program. Journal of Nematology 1999;31:117-133.

[40] Brower AVZ. Delimitation of phylogenetic species with DNA sequences: a critique of Davis and Nixon's population aggregation analysis. Systematic Biology 1999;48:199-213.

[41] Wiens JJ, Penkrot TA. Species delimitation in systematics: inferring diagnostic differences between species. Proceedings of the Royal Society of London, Series B 2002;267:631-636.

[42] Baum DA, Shaw KL. Genealogical perspectives on the species problem. In: Hoch PC, Stephenson AG (eds.) Molecular and experimental approaches to plant biosystematics. St. Louis: Missouri Botanical Garden; 1995. p.289-303.

[43] Templeton AR, Routman E, Phillips CA. Separating population structure from history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotype in the tiger salamander, *Ambystoma tigrinum*. Genetics 1995;140:767-782.

[44] Templeton AR. Using phylogeographic analyses of gene trees to test species status and boundaries. Molecular Ecology 2001;10:779-791.

[45] Dettman JR, Jacobson DJ, Taylor JW. A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. Evolution 2003;57:2703-2720.

[46] Dettman JR, Jacobson DJ, Turner E, Pringle A, Taylor JW. Reproductive isolation and phylogenetic divergence in *Neurospora*: comparing methods of species recognition in a model eukaryote. Evolution 2003;57: 2721-2741.

[47] Knowles LL, Carstens BC. Delimiting species without monophyletic gene trees. Systematic Biology 2007;56:887-895.

[48] Shaffer HB, Thomson RC. Delimiting species in recent radiations. Systematic Biology 2007;56:896-906.

[49] Raxworthy C, Ingram CM, Rabibisoa N, Pearson RG. Applications of ecological niche modeling for species delimitation: A review and empirical evaluating using day geckos (*Phelsuma*) from Madagascar. Systematic Biology 2007;56:907-923.

[50] Rissler LJ, Apodaca JJ. Adding more ecology into species delimitation: Ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). Systematic Biology 2007;56:924-942.

[51] Leaché AD, Koo MS, Spencer CL, Papenfuss TJ, Fisher RN, McGuire JA. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proceedings of the National Academy of Sciences of the USA 2009;106:12418-12423.

[52] Bond JE, Stockman AK. An integrative method for delimiting cohesion species: Finding the population-species interface in a group of California trapdoor spiders with extreme genetic divergence and geographic structuring. Systematic Biology 2008;57:628-646.

[53] Templeton AR. Species and speciation: Geography, population structure, ecology, and gene trees. In: (Howard DJ, Berlocher SH (eds.) Endless Forms: Species and Speciation. New York, NY:Oxford University Press;1989. p.32-41,

[54] Cummings MP, Neel MC, Shaw KL. A genealogical approach to quantifying lineage divergence. Evolution 2008;62:2411-2422.

[55] Weisrock DW, Rasoloarison RM, Fiorentino I, Ralison JM, Goodman SM, Kappeler PM, Yoder AD. Delimiting Species without Nuclear Monophyly in Madagascar's Mouse Lemurs. PLoS ONE 2010;5: e9883. doi:10.1371/journal.pone.0009883

[56] Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S. Large-scale species delimitation method for hyperdiverse groups. Molecular Ecology 2012;21:2671-2691.

[57] Puillandre N, Lambert A, Brouillet S, Achaz G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular Ecology 2012;21(8):1864-1877.

[58] Hausdorf B, Hennig C. Species delimitation using dominant and codominantmultilocus markers. Systematic Biology 2010;59:491-503.

[59] Hey J, Pinho C. Population genetics and objectivity in species diagnosis. Evolution 2012;66:1413-1429.

[60] Guillot G, Renaud S, Ledevin R, Michaux J, Claude J. A unifying model for the analysis of phenotypic, genetic, and geographic data. Systematic Biology 2012; doi:10.1093/sysbio/sys038.

[61] Maddison WP. Gene trees in species trees. Systematic Biology 1997;46:523-536

[62] Rannala B, Yang Z. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 2003;164:1645-1656.

[63] Degnan JH, Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology and Evolution 2009;24:332-340.

[64] Kubatko LS, Carstens BC, Knowles LL. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 2009;25:971-973.

[65] Liu L. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 2008;24:2542-2543.

[66] Ence DD, Carstens BC. SpedeSTEM: a rapid and accurate method for species delimitation. Molecular Ecology Resources 2011;11:473-480.

[67] Yang Z, Rannala B. Bayesian species delimitation using multilocus sequence data. Proceedings of the National Academy of Sciences of the USA 2010;107:9264-9269.

[68] O'Meara BC. New heursitic methods for joint species delimitation and species tree inference. Systematic Biology 2010;59:59-73.

[69] Camargo A, Morando M, Avila LJ, Sites, Jr. JW. Species delimitation with ABC and other coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwinii* complex (Squamata: Liolaemidae). Evolution 2012;doi:66:2834-2849

[70] Zapata F, Jiménez I. Species delimitation: inferring gaps in morphology across geography. Systematic Biology 2012;61:179-194.

[71] Thomson RC, Wang IJ, Johnson JR. Genome-enabled development of DNA markers for ecology, evolution and conservation. Molecular Ecology 2010;19:2184,àí2195.

[72] Townsend TM, Alegre RE, Kelley ST, Wiens JJ, Reeder TW. Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles. Molecular Phylogenetics and Evolution 2008;47:129-142.

[73] Portik DM, Wood Jr., PL, Grismer JL, Stanley EL, Jackman TR. Identification of 104 rapidly-evolving nuclear protein-coding markers for amplification across scaled reptiles using genomic resources. Conservation Genetics Resources 2012;4:1-10.

[74] Holsinger KE. Next generation population genetics and phylogeography. Molecular Ecology 2010;19:2361-2363.

[75] Rokas A, Abbot P. Harnessing genomics for evolutionary insights. Trends in Ecology and Evolution 2009;24:192-200.

[76] Tautz D, Ellegren H, Weigel D. Next generation molecular ecology. Trends in Ecology and Evolution 2010;19(Suppl. 1):1-3.

[77] Alföldi J, Di Palma F, Grabherr M, Williams C, Kong L, Mauceli E, Russell P, Lowe CB, Glor RE, Jaffe JD, Ray DA, Boissinot S, Shedlock AM, Botka C, Castoe TA, Colbourne JK, Fujita MK, Godinez-Moreno R, ten Hallers BF, Haussler D, Heger A, Heiman, D Janes DE, Johnson J, de Jong PJ, Koriabine MY, Lara M, Novick PA, Organ CL, Peach SE, Poe S, Pollock DD, de Queiroz K, Sanger T, Searle S, Smith JD, Smith Z, Swofford R, Turner-Maier J, Wade J, Young S, Zadissa A, Edwards SV, Glenn TC, Schneider CJ, Losos JB, Lander ES, Breen M, Ponting CP, Lindblad-Toh K. The genome of the green anole lizard and a comparative analysis with birds and mammals. Nature 2011;477:587-591.

[78] Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 2012;61:717-726.

[79] Lemmon AR, Emme S, Lemmon EM. Anchored hybrid enrichment for massively high-throughput phylogenomics. Systematic Biology 2012;61:727-744.

[80] Castoe TA, Poole AW, Gu W, Jason de Koning AP, Daza JM, Smith EN, Pollock DD. Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. Molecular Ecology Resources 2010;10:341-347.

[81] Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 2008;3(10):e3376./journal.pone.0003376.

[82] Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM. Resolving post-glacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences of the USA 2010;107:16196-16200.

[83] Metzker ML. Sequencing technologies the next generation. Nature Reviews Genetics 2010;11:31-46.

[84] Schneider GF, Dekker C. DNA sequencing with nanopores. Nature Biotechnology 2012;30: 326-328.

[85] Pennisi E. Search for Pore-fection. Science 2012;336:534-537.

[86] Brito PH, Edwards SV. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 2009;135:439-455.

[87] Nosil P, Funk DJ, Ortiz-Barrientos D. Divergent selection and heterogeneous genomic divergence. Molecular Ecology 2009;18:385-402.

[88] Carstens BC, Dewey TA. Species delimitation using a combined coalescent and information theoretic approach: An example from North American *Myotis* bats. Systematic Biology 2010;59:400-414.

[89] Camargo A, Sinervo B, Sites, Jr., JW. Lizards as model organisms for linking phylogeographic and speciation studies. Molecular Ecology 2010;19:3250-3270.

[90] Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. Coalescent-based species delimitation in an integrative taxonomy. Trends in Ecology and Evolution 2012;27:480-488.

[91] Bernardo J. A critical appraisal of the meaning and diagnosability of cryptic evolutionary diversity, and its implications for conservation in the face of climate change. In: Hodkinson TR, Jones MB, Waldren S, Parnell JAN (eds.) Climate Change, Ecology and Systematics. Cambridge University Press:The Systematics Association; 2011.p380-438.

[92] Zhang C, Zhang D-X, Zhu T, Yang Z. Evaluation of a Bayesian coalescent method of species delimitation. Systematic Biology 2011;60:747-761.

[93] Kawecki TJ, Ebert D. Conceptual issues in local adaptation. Ecology Letters 2004;7:1225-1241.

[94] Schluter D. Ecology and the origin of species. Trends in Ecology and Evolution 2001;16:372-380

[95] Schluter D. Evidence for ecological speciation and its alternative. Science 2009;323:737-741.

[96] Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, Damschen EI, Davies TJ, Grytnes J-A, Harrison SP, Hawkins BA, Holt RD, McCain CM, Stephens PR. Niche conservatism as an emerging principle in ecology and conservation biology. Ecology Letters 2010;13:1310-1324.

[97] Avise JC. Molecular markers, natural history, and evolution. Second edition. Sinauer Associates; 2004.

[98] Zink RM, Barrowclough GF. Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology 2008;17:2107-2121.

[99] Petit RJ, Excoffier L. Gene flow and species delimitation. Trends in Ecology and Evolution 2009;24:386-393.

[100] Galter N, Nabholz B, Glémin S, Hurst GDG. Mitochondrial DNA as a marker of molecular diversity. Molecular Ecology 2009;18:4541-4550.

[101] Funk DJ, Omland KE. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 2003;34:397-423.

[102] Morando M, Avila LJ, Sites, Jr. JW. Sampling strategies for delimiting species: genes, individuals and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaeminidae) in Andean Patagonian South America. Systematic Biology 2003.52:159-185.

[103] Funk WC, Caminer M, Ron SR. High levels of cryptic species diversity uncovered in Amazonian frogs. Proceedings of the Royal Society, Series B, 2012;279:1806-1814.

[104] Niemiller ML, Near TJ, Fitzpatrick BM. Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish, *Typhlichthys subterraneous* (Teleostei: Amblyopsidae). Evolution 2012;66:846-866.

[105] Agrawal AF, Feder JL, Nosil P. Ecological divergence and the origins of intrinsic postmating isolation with gene flow. International Journal of Ecology, 2011;doi: 10.1155/2011/435357

[106] Case LD, Ambrosius WT. Power and sample size. Methods in Molecular Biology 2007;404:377-408.

[107] Rundle HD, Nosil P. Ecological speciation. Ecology Letters 2005;8:336-352.

[108] Bauer AM, Parham JF, Brown RM, Stuart BL, Grismer L, Papenfuss TJ, Böhme W, Savage JM, Carranza S, Grismer JL, Wagner P, Schmitz A, Ananjeva NB, Inger RF. Availability of new Bayesian-delimited gecko names and the importance of character-based species descriptions. Proceedings of the Royal Society Series B 2011;278:490-492.

[109] Fujita MK, Leaché AD. A coalescent perspective on delimiting and naming species: a reply to Bauer et al. Proceedings of the Royal Society Series B 2011;278:493-495.

[110] Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH. Integrative taxonomy: a multisource approach to exploring biodiversity. Annual Review of Entomology 2010;55:421-438.

[111] Yeates DK, Seago S, Nelson L, Cameron SL, Joseph L, Trueman JWH. Integrative taxonomy, or iterative taxonomy? Systematic Entomology 2011;36:209-217.

[112] Hudson RR. Generating samples under a Wright-Fisher neutral model. Bioinformatics 2002;18:337-338.

[113] Rambaut A, Grassly NC. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computer Applications in the Biosciences 1997;13:235-238.

[114] Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 1997;24:38-49.

[115] McCormack JE, Huang H, Knowles LL. Maximum likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. Systematic Biology 2009;58:501-508.

[116] Knowles LL. Sampling strategies for species tree estiation. In: Knowles LL, Kubatko LS (eds.) Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell;2010. p163-174.

[117] Knowles LL, Kubatko LS. Estimating species trees: An introduction to concepts and models. In: Knowles LL, Kubatko LS (eds.) Species Trees: Practical and Theoretical Aspects. Wiley-Blackwell;2010. p.1-14.

[118] Felsenstein J. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? Molecular Biology ad Evolution 2006;23:691-700.

[119] Leaché AD, Fujita M. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proceedings of the Royal Society Series B: Biological Sciences 2010;277:3071-3077.

[120] Futuyma DJ. Progress on the origin of species. PLoS Biology 2005;3:e62.

[121] Thorpe RS, Surget-Groba Y, Johansson H. Genetic test for ecological and allopatric speciation in anoles on an island archipelago. PLoS Genetics 2010;6:e1000929.

[122] Pereira RJ, Wake DB. Genetic leakage after adaptive divergence and nonadaptive divergence in the *Ensatina eschscholtzii* ring species. Evolution 2011;63:2288-2301.

[123] Weissing FJ, Edelaar P, van Doorn GS. Adaptive speciation theory: a conceptual review. Behavioral Ecology and Sociobiology 2011;65:461-480.

[124] Van Doorn GS, Edelaar P, Weissing FJ. On the origin of species by natural and sexual selection. Science 2009;326:1704-1707.

[125] Maan ME, Seehausen O. Ecology, sexual selection, and speciation. Ecology Letters 2011,14:591-602.

[126] Hendry AP, Nosil P, Rieseberg LH. The speed of ecological speciation. Functional Ecology 2007;21:455-464.

[127] Thibert-Plante X, Hendry AP. When can ecological speciation be detected with neutral loci? Molecular Ecology 2010;19:2301-2314.

[128] Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate J. Adaptation genomics: the next generation. Trends in Ecology and Evolution 2010;25:705-712.

[129] Roesti M, Hendry AP, Salzburger W, Berner D. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. Molecular Ecology 2012;21:2852-2862.

[130] Smadja CM, Canbäck B, Vitalis R, Gautier M, Ferrari J, Zhou J-J, Butlin RK. Large-scale candidate gene scan reveals the role of chemoreceptor genes in host plant specialization and speciation in the pea aphid. Evolution, 2012;doi: 66:2723-2738

IntechOpen

IntechOpen

IntechOpen