Review article

Divergence with gene flow – the amphibian perspective

Krystyna   Nadachowska

Jagiellonian University, Institute of Environmental Sciences, Kraków, Poland

Advances in molecular and computational biology as well as in population genetics theory open new avenues in the study of speciation, for example enabling the explicit estimation of the amount of gene flow that has occurred during population divergence. Developments in two areas seem particularly important. First, novel coalescent-based methods can be applied to multilocus sequence data to infer the time of population divergence, long-term effective population sizes and their changes and the extent of gene flow between diverging populations. Second, the advent of ultra high-throughput sequencing technologies enable the inexpensive generation of vast amount of sequence data for any organism. Many amphibian species have been shown to be incompletely reproductively isolated and can hybridize for prolonged periods of time, making them ideal models to study the divergence of populations to form new taxa despite ongoing gene flow. Here I discuss the new findings emerging from multilocus DNA sequence-based approaches that have already been applied in amphibian population genetics. I also outline future directions of research, emphasizing the utility of parallel sequencing technologies together with methods of population genetic and phylogenetic inference, which are likely to provide a better understanding of the process of population differentiation and divergence to ultimately form new species.

Key words: amphibians, coalescent-based methods, divergence models, isolation with migration, ultra high-throughput sequencing

INTRODUCTION

There are almost 6,500 described amphibian species (Frost, 2009), a number which is rapidly increasing due to the discovery of new species in tropical areas and to a lesser extent as a result of taxonomic revisions (Hanken, 1999; Kohler et al., 2005; Fouquet et al., 2007). Some aspects of amphibian biology such as low dispersal abilities often result in a high degree of spatial structure, but widespread hybridization between species and the ability to exchange genes for prolonged periods of time may hamper the description of true relationships between populations and species. Thus, the delimitation of species often requires information on both morphology and multilocus DNA sequence variation throughout the geographic range. On the other hand, the features listed above also make amphibians a group that is well suited to study the initial stages of speciation, when reproductive isolation is not complete. Both species delimitation and understanding processes of population divergence over space and time are longstanding questions in evolutionary biology (Sites & Marshall, 2003; Hey, 2006).

The focus in speciation studies and the approaches to study the process of divergence have changed with the development of new species concepts, speciation models and advances in both laboratory and computational methods (Coyne & Orr, 2004; Noor & Feder, 2006; Schluter & Conte, 2009). The biological species concept (Mayr, 1942) defines species as reproductively isolated groups and focuses on the evolution of reproductive isolation. It is tightly linked to allopatric (isolation) models of speciation, in which two populations accumulate differences in complete isolation from each other. Reproductive isolation is seen as the result of the emergence of genomic incompatibilities *sensu* Bateson–Dobzhansky–Muller (Bateson, 1909; Dobzhansky, 1936; Muller, 1942), or related mechanisms (Coyne & Orr, 2004; Masly et al., 2006; Phadnis & Orr, 2009). In this model, given enough time, the development of reproductive isolation is an inevitable consequence of isolation and the process of speciation can be accelerated by divergent selection (Rice & Hostert, 1993; Fitzpatrick, 2002). Allopatric speciation has been uncontroversial and is believed to be the most common mechanism leading to the formation of new species in most animals (Coyne & Orr, 2004). At the other extreme, however, populations can also diverge in sympatry with no external barriers to gene flow, in a process driven by divergent selection (e.g. Rundle & Nosil, 2005; Barluenga et al., 2006; Savolainen et al., 2006; Bolnick & Fitzpatrick, 2007) or frequency dependent disruptive selection (Dieckmann et al., 2004).

Strict isolation and migration models can be linked by a continuum of scenarios with different levels of gene flow (Rice & Hostert, 1993; Butlin et al., 2008; Fitzpatrick et al., 2008, 2009a; Mallet et al., 2009). Gene flow can be restricted spatially or temporally and can be limited to certain parts of the genome (Wu, 2001; Wu & Ting, 2004; Mallet, 2007; Yatabe et al., 2007; Nosil et al., 2009; Via, 2009). Apparently natural selection prevents introgression of some parts of the genome between di-

verging populations, while other parts may introgress more easily (reviewed in Nosil et al., 2009). These non-introgressing parts are expected to contain genes responsible for reproductive isolation and species-specific adaptations, and have been termed "genomic islands of speciation" (Turner et al., 2005). There has been considerable recent interest in the characteristics of such regions in the genome and the mechanisms of their maintenance (Ting et al., 2000; Harr, 2006; Via & West, 2008). During a selective sweep in a single population, a spread of an adaptive allele together with neutral sites close to the selected locus produces a "valley of diversity". However, the effect of a selective sweep is transient, due to the accumulation of new mutations via drift and the levels of diversity in the valley that are rebuilt. In subdivided populations, islands of divergence can be maintained because selection acting in one subpopulation reduces gene flow and effective recombination between populations. This mechanism is termed "divergence hitchhiking" and can explain the maintenance of large genomic islands of differentiation (Via & West, 2008). However, empirical studies have found rather small genomic islands of differentiation (sunflowers: Yatabe et al., 2007; malaria mosquitoes: Turner et al., 2005).

One of the key questions in speciation studies is how much gene flow occurs during the initial phases of speciation. The increasing availability of DNA polymorphism data and advances in analytical and computational techniques now allows us to directly address questions about the patterns of gene flow during speciation. How often do the early stages of speciation occur in the face of gene flow? How common is speciation with gene flow? How many and which genes are involved in divergence and which parts of the genome can be easily exchanged between diverging populations? What is the role of ecological factors in divergence with gene flow? These are major areas of research in current speciation studies.

Divergence population genetics aims at inferring the evolutionary history of populations or recently diverged species, and is mostly used to study divergence with gene flow. Sequence data can be used in the framework of divergence models to extract information on long-term effective population sizes, times of divergence and patterns and times of gene flow between populations. The mathematical framework applied in divergence models is based on the coalescent theory (Wakeley, 2009), which describes the ancestry of a random sample of sequences drawn from a contemporary population. One of the most important consequences of the stochastic nature of the coalescent is the commonly observed discordance between species trees and individual gene trees due to shared ancestral polymorphisms and incomplete lineage sorting (Nichols, 2001). Different genealogical histories of various parts of the genome will be particularly common when studying closely related species in which both incomplete lineage sorting and interspecific gene flow may contribute to a conflict between gene trees and species trees. Multilocus sequence data allow us to distinguish between patterns arising due to the stochastic nature of the coalescent process and differential gene flow. Thus, it

is essential to use information from many unlinked loci to infer demographic parameters and the timing of migration between diverging species, and to evaluate the possible role of selection.

Here I discuss new findings from DNA sequence-based multilocus studies on divergence population genetics and speciation in amphibians, with a special emphasis on the divergence of populations despite gene flow between them. The focus will be placed on closely related, often incompletely reproductively isolated species. I will also briefly discuss some peculiar features of amphibians that make them an interesting but also difficult group to study. Finally, I will try to outline future directions of research that I consider most promising and exciting.

AMPHIBIANS – CRYPTIC STRUCTURE AND WIDESPREAD HYBRIDIZATION

Within amphibians, there is often no consistent relationship between rates of species diversification and morphological evolution, with evidence for rapid diversification occurring with little morphological change and vice versa (Adams et al., 2009, based on studies on plethodontid salamanders). Indeed, many amphibian species exhibit extreme levels of population structure and genetic differentiation over relatively small geographical areas with no morphological divergence (Highton, 1995; Kozak et al., 2006), and many have been found to be morphologically "cryptic" species (Hanken, 1999; Chek et al., 2001; Fouquet et al., 2007; Lemmon et al., 2007). On the other hand, amphibians are also known to retain the ability to exchange genes for a very long time (Wilson et al., 1974; Prager & Wilson, 1975; Vences & Wake, 2007; but see Sasa et al., 1998) and hybridization is a well known phenomenon with many examples of various kinds of hybrid zones (e.g. Szymura, 1993; Babik et al., 2003; Fitzpatrick & Shaffer, 2004; Sequeira et al., 2005; Arntzen et al., 2009). New molecular methods of revealing cryptic and hybridizing species can also indicate species groups that may have been diverging in the face of gene flow, and give an exciting opportunity to study several aspects of speciation and the role of gene flow in population divergence.

MULTILOCUS STUDIES ON AMPHIBIANS

The majority of amphibian DNA sequence studies published so far have been based on mtDNA data alone. However, inferences about species/population histories based on one locus are often misleading, as they do not reveal whole genome patterns and cannot describe complex evolutionary histories (Zhang & Hewitt, 2003). The up to four times smaller mtDNA effective population size in comparison to nuclear loci may cause an oversimplification of evolutionary history (the process of lineage sorting, when alleles in the two diverging species/populations become reciprocally monophyletic, is expected to proceed faster in mtDNA), a distortion due to

selection on mtDNA (Ballard & Whitlock, 2004; Hurst & Jiggins, 2005), or an underestimation of genetic diversity (stronger action of drift on mtDNA, with the possibility of recurrent selective sweeps; Bazin et al., 2006).

Another widely used class of markers, microsatellites, may provide a large number of loci that are evenly distributed across the genome. However, repetitive DNA is often characterized by a high mutation rate, departures from the strict stepwise mutation model and a high frequency of homoplasy, resulting in difficulties in describing relationships between alleles. These features make microsatellites of limited use in phylogenetics and divergence population genetics. Moreover, developing new microsatellites in amphibians is often a challenging task due to their huge genomes (Garner, 2002; Gregory, 2005).

Single copy nuclear sequence markers provide much more information about genealogical relationships among alleles. Patterns of genetic differentiation contain information on times of divergence and population sizes, as well as current and past gene flow, and can be applied in the framework of divergence models. There have been just a few studies on closely related amphibian species utilizing nuclear multilocus sequence data (Weisrock et al., 2006; Gamble et al., 2008; Fitzpatrick et al., 2009b; Themudo et al., 2009) and even fewer studies that made explicit attempts to study the process of population divergence in the face of gene flow over time (Niemiller et al., 2008; Nadachowska & Babik, 2009).

Species boundaries and shallow phylogenies

The case of Mexican ambystomatid salamanders exemplifies the utility of a multilocus approach in a recently and rapidly radiating clade (Weisrock et al., 2006). The *Ambystoma tigrinum* complex is a recently derived and extremely morphologically diverse group, with species undergoing metamorphosis as well as species that are facultatively or obligately paedomorphic. The complex has been extensively studied for over 20 years. Obligate paedomorphosis has been shown to evolve multiple times from a facultatively paedomorphic ancestor (Shaffer, 1984; Shaffer & Voss, 1996); however, even the most morphologically divergent species retain the ability to hybridize (Voss & Shaffer, 1996; Riley et al., 2003; Fitzpatrick & Shaffer, 2004). Weisrock et al. (2006) aimed at determining species boundaries in a facultative paedomorph, *Ambystoma ordinarium*, which has been shown to comprise at least two different mtDNA lineages indicating that the group might not be a monophyletic taxon (Shaffer & McKnight, 1996; Highton, 2000). The authors analysed mtDNA data sets and eight unlinked nuclear loci derived from an EST library. The mtDNA based phylogenetic tree revealed no distinct clades and general non-monophyly of species, in contrast to most nuclear loci, two of which reveal exclusive monophyly, two reveal paraphyly and one reveals polyphyly. The authors argued that the observed discordance of gene trees can be explained by both historical introgression and incomplete lineage sorting due to the retention of ancestral polymorphisms in descendent populations (Pamilo & Nei,

1988). This study constitutes a good example of how a combination of past hybridization and retention of ancestral polymorphisms can lead to striking differences in gene and species trees even among extremely morphologically divergent and currently allopatric species.

Due to the limited opportunity for a phylogenetic signal to accumulate between successive speciation events, the reconstruction of the evolutionary history of young species pairs is difficult. Themudo et al. (2009) used mtDNA and five nuclear loci to resolve the relationships of the rapidly radiated crested newt superspecies (*Triturus*, Salamandridae). The radiation was dated to 8–11 million years ago. The authors applied a new hierarchical Bayesian approach to construct the species tree (Edwards et al., 2007; Liu & Pearl, 2007; see below). The resulting species tree was almost fully resolved, and showed previously undetected branching events and deep differentiation in two lineages of *T. karelini*, which the authors proposed to elevate to full species.

The last example in this section concerns morphologically very similar species with substantial levels of genetic differentiation. There are many examples of changes in taxonomic practice in "cryptic" amphibian species after applying molecular tools. In the vast majority of cases, these judgments are based on mtDNA alone (e.g. Austin et al., 2002; Moriarty & Cannatella, 2004). Only a few studies have used mtDNA and allozymes (e.g. Pauly et al., 2007), or mtDNA and one nuclear locus (e.g. Crawford, 2003; Timpe et al., 2009). Gamble et al. (2008) used three nuclear loci and mtDNA to study the genealogical relationships between the subspecies of two species of cricket frogs (*Acris crepitans* and *A. gryllus*). They found discordances between morphological and vocalization data and the distribution of evolutionary lineages as inferred from genetic markers. Based on combined mitochondrial and nuclear analyses, the authors proposed to elevate one of the subspecies to species status (*A. blanchardi*).

Divergence with gene flow

Only two studies have so far explicitly attempted to study divergence in the face of gene flow over time in amphibians. Niemiller et al. (2008) studied the Tennessee cave salamander complex (*Gyrinophilus palleucus*, all obligately paedomorphic) and its terrestrial and metamorphosing sister species, the spring salamander (*G. porphyriticus*), to find out whether cave salamanders arose in allopatry or by divergence with gene flow. Several coalescent-based methods were used to study levels of gene flow in mtDNA and a RAG-1 gene fragment and to address potential discordances between the species and gene trees. Specifically, the authors applied an "isolation with migration" model, which assumes that a single ancestral population split into two descendent populations *t* generations ago; since then, the populations may (or may not) have been subject to gene exchange. They used the recently updated version of the program IM (IMa, Hey & Nielsen, 2007) to obtain the distribution of mean time of migration events and to distinguish between scenarios of secondary contact and hybridization versus

continuous divergence with gene flow. In contrast to the traditional view, the authors found strong evidence that Tennessee cave salamanders originated from spring salamanders via divergence with gene flow, indicating an ecological mode of speciation.

A similar approach was used by Nadachowska & Babik (2009) to study the evolutionary history of two subspecies of the smooth newt (*Lissotriton vulgaris vulgaris* and *L. v. kosswigi*). The authors used sequence variation at eight nuclear loci and mtDNA to study diversity, divergence and patterns of gene flow. They found asymmetric introgression of mtDNA from *vulgaris* to *kosswigi*. Phylogenetic trees of nuclear loci revealed clearly distinct clades in four studied markers and haplotype-sharing between subspecies in one anonymous nuclear locus. Despite deep, pre-Pleistocene divergence, IMa analyses provided evidence for ongoing migrations from *vulgaris* to *kosswigi*. By comparing temporal distributions of migration between subspecies with the estimated divergence time, the authors found signatures of more ancient gene flow, indicating that initial divergence might have occurred in the face of gene flow.

Heterogeneous introgression in hybrid zones

Multilocus approaches provide the opportunity to study heterogeneities in the frequencies of introgressed alleles across the genome and the dynamics of secondary contact hybrid zones. Genetic interactions of hybridizing species can result in hybrid dysfunction, increased hybrid vigour, or rapid introgression. These scenarios were traditionally viewed as mutually exclusive. However, natural selection can act differentially on different genes, preventing introgression of some parts of the genome while other parts may introgress more readily (Nosil et al., 2009). Thus, variation at different loci within the same hybrid population can be shaped by various population genetic processes. Fitzpatrick et al. (2009b) studied admixture dynamics in recently initiated hybrid populations of barred tiger salamanders (*Ambystoma mavortium*) and Californian tiger salamanders (*A. californiense*). The authors analysed 64 EST markers and found variable levels of introgression across studied loci and populations. Few, strongly selected alleles displaced or almost displaced native alleles within hybrid populations. This study constitutes convincing evidence that in recently established hybrid zones some loci can undergo rapid selective sweeps and quickly go to fixation. In long-established hybrid zones, such loci can mistakenly be considered as ancestral shared states that never diverged. The system studied provided a rare opportunity to unravel the history of rapid selective sweeps, because good pre-contact reference populations still exist and the hybrid zone is young enough to study genes before they have gone to fixation.

FUTURE DIRECTIONS
New sequencing technologies

The advent of ultra high-throughput sequencing has opened up unprecedented opportunities to study the divergence between species and populations at the genomic level. There are several technologies currently available; the most popular are Roche 454FLX, Solexa Illumina GA and ABI Solid (Shendure & Ji, 2008), which all sequence $4 \times 10^8 - 5 \times 10^{10}$ basepairs of DNA in a single analysis, and the throughput increases rapidly, although the individual reads are usually shorter that in classical Sanger sequencing. The 454 Roche technology is probably the most interesting in this respect, as it provides the longest read length (currently 400–500 bp).

A promising method that takes advantage of new sequencing technologies is CRoPS (Complexity Reduction of Polymorphic Sequences; van Orsouw et al., 2007), enabling rapid scoring in population studies of hundreds of sequence markers from any genome. The method combines AFLP technology (Amplified Fragments Length Polymorphisms), which is based on selective PCR amplification of DNA fragments obtained from restriction enzyme digestion of total genomic DNA, with next generation sequencing technologies (454). The reduced representation of the genome obtained (AFLP fragment libraries sequenced and aligned) is scanned for SNPs (Single Nucleotide Polymorphisms) using fully-automated bioinformatics tools. The method is especially efficient in organisms for which large-scale genotyping may be difficult due to lack of extensive genome sequence information, high proportions of highly repetitive DNA fragments in large genomes and/or low levels of polymorphisms. Amphibians are known to have large genomes (Gregory, 2005), with a high proportion of noncoding and repetitive DNA. Furthermore, to date extensive genome sequence information is available only for *Xenopus* (*Silurana*) *tropicalis* (http://www.ncbi.nlm.nih.gov/genome/guide/frog/).

Another promising prospect is provided by the study of transcriptomes, which should provide a wealth of data including variation of adaptive significance. Using new sequencing technologies and new automated bioinformatic pipelines, it is possible to sequence whole transcriptomes and assemble transcripts *de novo* (Vera et al., 2008; Meyer et al., 2009). Transcriptomes can be scanned for divergent genes, genes with fixed nonsynonymous mutations and genes with signals of positive selection resulting in the identification of candidate genes for adaptations. Genome scans can be particularly informative when applied to closely related populations with ongoing hybridization (Schluter & Conte, 2009), where identification of parts under divergence selection and genes that easily flow between two species is crucial for understanding the dynamics of hybrid zones and the divergence process in more detail.

Coalescent-based methods of inference from DNA sequences

There has been enormous theoretical progress in population genetics and coalescent theory in recent years (Kuhner, 2009; Wakeley, 2009). Despite existing theoretical models showing that speciation can occur in the face of gene flow (for a review, see Bolnick & Fitzpatrick, 2007), convincing empirical evidence has been scarce (Coyne & Orr, 2004). One of the most serious problems has been to distinguish divergence with gene flow in the early stages of speciation from gene flow after secondary contact. The

two most popular methods currently available are IM and IMa (Hey & Nielsen, 2004, 2007) and MIMAR (Becquet & Przeworski, 2007, 2009). Further methods based on ABC (Approximate Bayesian Computation) were developed by Ross-Ibarra and colleagues (RI-ABC; Ross-Ibarra et al., 2009). IMa and MIMAR both use multilocus sequence data to infer effective population sizes (ancestral population size and descendent population sizes), time of ancestral population split and levels of gene flow between descendent populations. IMa analyses full polymorphism spectra and is sensitive to intra-locus recombination. In contrast, MIMAR uses summary statistics to estimate the model parameters and can be applied to data with signs of intra-locus recombination. Both programs have been used to study patterns of gene flow between differentiating populations (Bull et al., 2006; Kronforst et al., 2006; Carling & Brumfield, 2008; Niemiller et al., 2008; Stadler et al., 2008; Nadachowska & Babik, 2009), although only IMa provides distributions of the number and mean time of migration events. This method has been used to study divergence in cave salamanders (Niemiller et al., 2008) and smooth newts (Nadachowska & Babik, 2009). In both cases, the authors argued that initial divergence might have occurred in the face of gene flow (see above).

The third method, RI-ABC (Ross-Ibarra et al., 2009), may use data with or without intralocus recombination to test alternative models of population divergence explicitly, including an isolation model, a migration model, an isolation with migration model in sympatry where gene flow occurs only during initial stages of population differentiation and is followed by complete isolation, and an allopatric model followed by secondary contact and gene flow. This method was introduced only recently and so far has been used to infer historical divergence and gene flow in the genus *Zea* (Ross-Ibarra et al., 2009); it gives the exciting prospect of explicitly testing divergence with ongoing gene flow against divergence in isolation with subsequent gene flow after secondary contact.

A similar approach that can be used together with methods implemented in the programs mentioned above is to run coalescent simulations (e.g. SIMCOAL – Laval & Excoffier, 2004; SIMDIV – http://lifesci.rutgers.edu/~heylab/HeylabSoftware.htm) to test the fit of the data to various models of population divergence. Assuming scenarios with different levels and times of gene flow after a population split, one can compare summary statistics (such as nucleotide diversity, p or Fst) obtained from simulated data under an assumed model (scenario) with the values of summary statistics calculated from real data. With multilocus sequence data increasingly easy to obtain, these techniques and their modifications should find widespread application in the study of the complex evolutionary histories of amphibian species.

Bridging phylogenetics and population genetics

The true relationships of closely related species can often be obscured by incomplete sorting of ancestral polymorphisms (see above). Incomplete lineage sorting can be especially pronounced in species with certain evolutionary histories (such as recent divergences) and can

be observed for a longer period of time in cases of high population sizes (Hudson & Coyne, 2002). Both recent divergences (e.g. Weisrock et al., 2006) and large population sizes can be found in amphibians (Crawford, 2003; Carnaval & Bates, 2007; Fouquet et al., 2007; Nadachowska & Babik, 2009). In such cases, discordance between gene trees from different loci may be common. There is an ongoing discussion of how researchers should deal with such data (Edwards, 2009). The commonly used concatenation method, where sequences from different genes are combined into one "supersequence" to construct a phylogeny, is likely to be misleading because unlinked loci throughout the genome have different histories (Mossel & Vigoda, 2005; Carstens & Knowles, 2007; Kubatko & Degnan, 2007; Matsen & Steel, 2007). An increasing number of newly developed coalescent-based methods now incorporate independent loci in phylogenetic inference, taking population genetic processes underlying species divergence into account (Maddison, 1997; Degnan & Salter, 2005; Maddison & Knowles, 2006; Ane et al., 2007; Carstens & Knowles, 2007; Edwards et al., 2007; Liu & Pearl, 2007; Oliver, 2008; Kubatko et al., 2009). These methods either identify the species tree that maximizes the probability of a set of genealogies given the species tree (e.g. Maddison 1997; COAL – Degnan & Salter, 2005, ESP-COAL – Carstens & Knowles, 2007), minimize the amount of deep coalescence (deep coalescence measures the discordance between a gene tree and a species/population tree by assuming that all discordances between trees are just due to incomplete lineage sorting and counting extra gene lineages) to estimate species phylogeny (e.g. MDC – Maddison, 1997; Maddison & Knowles, 2006; implemented in the Mesquite package, Maddison & Maddison, 2008), or apply Bayesian hierarchical models and use joint distribution of gene trees for many loci given a species tree as a prior (e.g. BEST – Edwards et al., 2007; Liu & Pearl, 2007; Liu, 2008). For more examples and more detailed discussion of multilocus molecular systematics, see Edwards (2009).

Incomplete lineage sorting at the initial stages of speciation results in discordance between gene trees, although the same pattern will be observed when there is ongoing gene flow between diverging populations. Thus, it is often difficult to conclude whether shared polymorphisms are the result of the retention of ancestral polymorphisms or recent migration. Undoubtedly, estimating species phylogenies in complex systems with unclear patterns of gene flow is one of the most challenging tasks in evolutionary biology. To my knowledge, there has been only one simulation study to date which examined the influence of gene flow on phylogenetic signal comparing three methods (concatenation, ESP-COAL and MDC) of species multilocus tree inference under four different models (n-island, stepping stone, parapatric, allopatric) and different levels of gene flow (low, moderate, high) (Eckert & Carstens, 2008). The simulations showed that, in contrast to the coalescent-based methods, the concatenation method performed poorly, even in the face of gene flow (with the exception of n-island and stepping stone models with moderate to high gene flow).

When more than four loci were sampled, the probability of finding the true species tree increased substantially. Nonetheless, care is needed when interpreting the resulting trees. Further work is certainly required, but the substantial development in recent years is promising and will have an impact on future research (Edwards, 2009).

CONCLUSIONS

Several aspects of amphibian biology, such as cryptic structure, widespread hybridization and the ability to exchange genes for long periods of time, make them an excellent system for the study of longstanding issues in evolutionary biology, including divergence in the face of gene flow. So far there have only been a limited number of studies on amphibians using multilocus sequence data; nevertheless these studies have already advanced our understanding of the mechanisms of divergence and historical gene flow in animals. The development of new molecular and computational tools is expected to boost multilocus studies of nonmodel organisms. New sequencing technologies and coalescent-based methods of inference from multilocus sequence data are of special interest and can be extremely useful in the field of amphibian divergence and speciation.

ACKNOWLEDGEMENTS

REFERENCES

Adams, D.C., Berns, C.M., Kozak, K.H. & Wiens, J.J. (2009). Are rates of species diversification correlated with rates of morphological evolution? *Proceedings of the Royal Society B* 276, 2729–2738.

Ane, C., Larget, B., Baum, D.A., Smith, S.D. & Rokas, A. (2007). Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution* 24, 412–426.

Arntzen, J.W., Jehle, R., Bardacki, F., Burke, T. & Wallis, G.P. (2009). Asymmetric viability of reciprocal-cross hybrids between crested and marbled newts (*Triturus cristatus*, *T. marmoratus*). *Evolution* 63, 1191–1202.

Austin, J.D., Lougheed, S.C., Neidrauer, L., Chek, A.A. & Boag, P.T. (2002). Cryptic lineages in a small frog: the post-glacial history of the spring peeper, *Pseudacris crucifer* (Anura: Hylidae). *Molecular Phylogenetics and Evolution* 25, 316–329.

Babik, W., Szymura, J.M. & Rafinski, J. (2003). Nuclear markers, mitochondrial DNA and male secondary sexual traits variation in a newt hybrid zone (*Triturus vulgaris* x *T. montandoni*). *Molecular Ecology* 12, 1913–1930.

Ballard, J.W.O. & Whitlock, M.C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology* 13, 729–744.

Barluenga, M., Stoelting, K.N., Salzburger, W., Muschick, M. & Meyer, A. (2006). Evolutionary biology – evidence for sympatric speciation? Reply. *Nature* 444, E13–E13.

Bateson, W. (1909). Heredity and variation in modern lights. In *Darwin and Modern Science*, 85–101. Seward, A.C. (ed.). Cambridge: Cambridge University Press.

Bazin, E., Glemin, S. & Galtier, N. (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* 312, 570–572.

Becquet, C. & Przeworski, M. (2007). A new approach to estimate parameters of speciation models with application to apes. *Genome Research* 17, 1505–1519.

Becquet, C. & Przeworski, M. (2009). Learning about models of speciation by computational approches. *Evolution* 63, 2547–2562.

Bolnick, D.I. & Fitzpatrick, B.M. (2007). Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution and Systematics* 38, 459–487.

Bull, V., Beltran, M., Jiggins, C.D., McMillan, W.O., Bermingham, E. & Mallet, J. (2006). Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biology* 4, 11.

Butlin, R.K., Galindo, J. & Grahame, J.W. (2008). Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philosophical Transactions of the Royal Society B* 363, 2997–3007.

Carling, M.D. & Brumfield, R.T. (2008). Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. *Genetics* 178, 363–377.

Carnaval, A.C. & Bates, J.M. (2007). Amphibian DNA shows marked genetic structure and tracks Pleistocene climate change in northeastern Brazil. *Evolution* 61, 2942–2957.

Carstens, B.C. & Knowles, L.L. (2007). Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Systematic Biology* 56, 400–411.

Chek, A.A., Lougheed, S.C., Bogart, J.P. & Boag, P.T. (2001). Perception and history: molecular phylogeny of a diverse group of neotropical frogs, the 30-chromosome *Hyla* (Anura: Hylidae). *Molecular Phylogenetics and Evolution* 18, 370–385.

Coyne, J.A. & Orr, H.A. (2004). *Speciation*. Sunderland: Sinauer Associates.

Crawford, A.J. (2003). Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. *Molecular Ecology* 12, 2525–2540.

Degnan, J.H. & Salter, L.A. (2005). Gene tree distributions under the coalescent process. *Evolution* 59, 24–37.

Dieckmann, U., Doebeli, M., Metz, J.A.J. & Tautz, D. (eds) (2004). *Adaptive Speciation*. Cambridge: Cambridge University Press.

Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* 21, 113–135.

Eckert, A.J. & Carstens, B.C. (2008). Does gene flow destroy phylogenetic signal? The performance of three methods for estimating species phylogenies in the presence of gene flow. *Molecular Phylogenetics and Evolution* 49, 832–842.

Edwards, S.V. (2009). Is a new and general theory of

molecular systematics emerging? *Evolution* 63, 1–19.

Edwards, S.V., Liu, L. & Pearl, D.K. (2007). High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences of the United States of America* 104, 5936–5941.

Fitzpatrick, B.M. (2002). Molecular correlates of reproductive isolation. *Evolution* 56, 191–198.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. (2008). What, if anything, is sympatric speciation? *Journal of Evolutionary Biology* 21, 1452–1459.

Fitzpatrick, B.M., Fordyce, J.A. & Gavrilets, S. (2009a). Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology* 22, 2342–2347.

Fitzpatrick, B.M., Johnson, J.R., Kump, D.K., Shaffer, H.B., Smith, J.J. & Voss, S.R. (2009b). Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evolutionary Biology* 9, 176.

Fitzpatrick, B.M. & Shaffer, H.B. (2004). Environment-dependent admixture dynamics in a tiger salamander hybrid zone. *Evolution* 58, 1282–1293.

Fouquet, A., Vences, M., Salducci, M.D., Meyer, A., Marty, C., Blanc, M. & Gilles, A. (2007). Revealing cryptic diversity using molecular phylogenetics and phylogeography in frogs of the *Scinax ruber* and *Rhinella margaritifera* species groups. *Molecular Phylogenetics and Evolution* 43, 567–582.

Frost, D.R. (2009). *Amphibian Species of the World: an Online Reference. Version 5.3* (12 February, 2009). Electronic database accessible at http://research.amnh.org/herpetology/amphibia/. New York: American Museum of Natural History.

Gamble, T., Berendzen, P.B., Shaffer, H.B., Starkey, D.E. & Simons, A.M. (2008). Species limits and phylogeography of North American cricket frogs (Acris: Hylidae). *Molecular Phylogenetics and Evolution* 48, 112–125.

Garner, T.W.J. (2002). Genome size and microsatellites: the effect of nuclear size on amplification potential. *Genome* 45, 212–215.

Gregory, T.R. (2005). *Animal Genome Size Database*. http://www.genomesize.com.

Hanken, J. (1999). Why are there so many new amphibian species when amphibians are declining? *Trends in Ecology and Evolution* 14, 7–8.

Harr, B. (2006). Genomic islands of differentiation between house mouse subspecies. *Genome Research* 16, 730–737.

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

Hey, J. & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167, 747–760.

Hey, J. & Nielsen, R. (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America* 104, 2785–2790.

Highton, R. (1995). Speciation in eastern North American salamanders of the genus *Plethodon*. *Annual Review of Ecology, Evolution and Systematics* 26, 579–600.

Highton, R. (2000). Detecting cryptic species using allozyme data. In *The Biology of Plethodontid Salamanders*, 215–241. Bruce, R.C., Jaeger, R. & Houck, L.D. (eds). New York: Kluwer Academic/Plenum Publishers.

Hudson, R.R. & Coyne, J.A. (2002). Mathematical consequences of the genealogical species concept. *Evolution* 8, 1557–1565.

Hurst, G.D.D. & Jiggins, F.M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B* 272, 1525–1534.

Kohler, J., Vieites, D.R., Bonett, R.M., Garcia, F.H., Glaw, F., Steinke, D. & Vences, M. (2005). New amphibians and global conservation: a boost in species discoveries in a highly endangered vertebrate group. *Bioscience* 55, 693–696.

Kozak, K.H., Weisrock, D.W. & Larson, A. (2006). Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae: *Plethodon*). *Proceedings of the Royal Society B* 273, 539–546.

Kronforst, M.R., Young, L.G., Blume, L.M. & Gilbert, L.E. (2006). Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. *Evolution* 60, 1254–1268.

Kubatko, L.S., Carstens, B.C. & Knowles, L.L. (2009). STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25, 971–973.

Kubatko, L.S. & Degnan, J.H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56, 17–24.

Kuhner, M.K. (2009). Coalescent genealogy samplers: windows into population history. *Trends in Ecology and Evolution* 24, 86–93.

Laval, G. & Excoffier, L. (2004). SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. *Bioinformatics* 20, 2485–2487.

Lemmon, E.M., Lemmon, A.R., Collins, J.T., Lee-Yaw, J.A. & Cannatella, D.C. (2007). Phylogeny-based delimitation of species boundaries and contact zones in the trilling chorus frogs (*Pseudacris*). *Molecular Phylogenetics and Evolution* 44, 1068–1082.

Liu, L. (2008). BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 21, 2542–2543.

Liu, L. & Pearl, D.K. (2007). Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* 56, 504–514.

Maddison, W.P. (1997). Gene trees in species trees. *Systematic Biology* 46, 523–536.

Maddison, W.P. & Knowles, L.L. (2006). Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55, 21–30.

Maddison, W.P. & Maddison, D.R. (2008). *Mesquite: A Modular System for Evolutionary Analysis.* http://mesquiteproject.org.

Mallet, J. (2007). Hybrid speciation. *Nature* 446, 279–283.

Mallet, J., Meyer, A., Nosil, P. & Feder, J.L. (2009). Space, sympatry and speciation. *Journal of Evolutionary Biology* 22, 2332–2341.

Masly, J.P., Jones, C.D., Noor, M.A.F., Locke, J. & Orr, H.A. (2006). Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* 313, 1448–1450.

Matsen, F.A. & Steel, M. (2007). Phylogenetic mixtures on a single tree can mimic a tree of another topology. *Systematic Biology* 56, 767–775.

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

Meyer, E., Aglyamova, G.V., Wang, S., Buchanan-Carter, J., Abrego, D., Colbourne, J.K., Willis, B.L. & Matz, M.V. (2009). Sequencing and de novo analysis of a coral larval transcriptome using 454 GSFlx. *BMC Genomics* 10, 19.

Moriarty, E.C. & Cannatella, D.C. (2004). Phylogenetic relationships of the North American chorus frogs (*Pseudacris*: Hylidae). *Molecular Phylogenetics and Evolution* 30, 409–420.

Mossel, E. & Vigoda, E. (2005). Phylogenetic MCMC algorithms are misleading on mixtures of trees. *Science* 309, 2207–2209.

Muller, H.J. (1942). Isolating mechanisms, evolution and temperature. *Biological Symposia* 6, 71–125.

Nadachowska, K. & Babik, W. (2009). Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). *Molecular Biology and Evolution* 26, 829–841.

Nichols, R. (2001). Gene trees and species trees are not the same. *Trends in Ecology and Evolution* 16, 358–364.

Niemiller, M.L., Fitzpatrick, B.M. & Miller, B.T. (2008). Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Molecular Ecology* 17, 2258–2275.

Noor, M.A.F. & Feder, J.L. (2006). Speciation genetics: evolving approaches. *Nature Reviews Genetics* 7, 851–861.

Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 18, 375–402.

Oliver, J.C. (2008). AUGIST: inferring species trees while accommodating gene tree uncertainty. *Bioinformatics* 24, 2932–2933.

Pamilo, P. & Nei, M. (1988). Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5, 568–583.

Pauly, G.B., Piskurek, O. & Shaffer, H.B. (2007). Phylogeographic concordance in the southeastern United States: the flatwoods salamander, *Ambystoma cingulatum*, as a test case. *Molecular Ecology* 16, 415–429.

Phadnis, N. & Orr, H.A. (2009). A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science* 323, 376–379.

Prager, E.M. & Wilson, A.C. (1975). Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. *Proceedings of the National Academy of Sciences of the United States of America* 72, 200–204.

Rice, W.R. & Hostert, E.E. (1993). Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* 47, 1637–1653.

Riley, S.P.D., Shaffer, H.B., Voss, S.R. & Fitzpatrick, B.M. (2003). Hybridization between a rare, native tiger salamander (*Ambystoma californiense*) and its introduced congener. *Ecological Applications* 13, 1263–1275.

Ross-Ibarra, J., Tenaillon, M. & Gaut, B.S. (2009). Historical divergence and gene flow in the genus *Zea*. *Genetics* 181, 1397–1409.

Rundle, H.D. & Nosil, P. (2005). Ecological speciation. *Ecology Letters* 8, 336–352.

Sasa, M.M., Chippindale, P.T. & Johnson, N.A. (1998). Patterns of postzygotic isolation in frogs. *Evolution* 52, 1811–1820.

Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N. & Baker, W.J. (2006). Sympatric speciation in palms on an oceanic island. *Nature* 441, 210–213.

Schluter, D. & Conte, G.L. (2009). Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9955–9962.

Sequeira, F., Alexandrino, J., Rocha, S., Arntzen, J.W. & Ferrand, N. (2005). Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander. *Chioglossa lusitanica*. *Molecular Ecology* 14, 245–254.

Shaffer, H.B. (1984). Evolution in a paedomorphic lineage. I. An electrophoretic analysis of the Mexican ambystomatid salamanders. *Evolution*, 1194–1206.

Shaffer, H.B. & McKnight, M.L. (1996). The polytypic species revisited: genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. *Evolution* 50, 417–433.

Shaffer, H.B. & Voss, S.R. (1996). Phylogenetic and mechanistic analysis of a developmentally integrated character complex: alternate life history modes in ambystomatid salamanders. *American Zoologist* 36, 24–35.

Shendure, J. & Ji, H.L. (2008). Next-generation DNA sequencing. *Nature Biotechnology* 26, 1135–1145.

Sites, J.W. & Marshall, J.C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* 18, 462–470.

Stadler, T., Arunyawat, U. & Stephan, W. (2008). Population genetics of speciation in two closely related wild tomatoes (*Solanum section lycopersicon*). *Genetics* 178, 339–350.

Szymura, J.M. (1993). Analysis of hybrid zones with *Bombina*. In *Hybrid Zones and the Evolutionary Process*, 261–389. Harrison, R.G. (ed.). New York: Oxford University Press.

Themudo, G.E., Wielstra, B. & Arntzen, J.W. (2009). Multiple nuclear and mitochondrial genes resolve the branching order of a rapid radiation of crested newts

(*Triturus*, Salamandridae). *Molecular Phylogenetics and Evolution* 53, 359–359.

Timpe, E.K., Graham, S.P. & Bonett, R.M. (2009). Phylogeography of the brownback salamander reveals patterns of local endemism in southern Appalachian springs. *Molecular Phylogenetics and Evolution* 52, 368–376.

Ting, C.T., Tsaur, S.C. & Wu, C.I. (2000). The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. *Proceedings of the National Academy of Sciences of the United States of America* 97, 5313–5316.

Turner, T.L., Hahn, M.W. & Nuzhdin, S.V. (2005). Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology* 3, 1572–1578.

van Orsouw, N.J., Hogers, R.C.J., Janssen, A., Yalcin, F., Snoeijers, S., Verstege, E., Schneiders, H., van der Poel, H., van Oeveren, J., Verstegen, H. & van Eijk, M.J.T. (2007). Complexity Reduction of Polymorphic Sequences (CRoPS™): a novel approach for large-scale polymorphism discovery in complex genomes. *PloS ONE*, e1172. doi: 10.1371/journal.pone.0001172

Vences, M. & Wake, D.B. (2007). Speciation, species boundaries and phylogeography of amphibians. In *Amphibian Biology 7*, 2613–2671. Heatwole, H. (ed.). Chipping Norton: Surrey Beatty & Sons.

Vera, J.C., Wheat, C.W., Fescemyer, H.W., Frilander, M.J., Crawford, D.L., Hanski, I. & Marden, J.H. (2008). Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology* 17, 1636–1647.

Via, S. (2009). Natural selection in action during speciation. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9939–9946.

Via, S. & West, J. (2008). The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology* 17, 4334–4345.

Voss, S.R. & Shaffer, H.B. (1996). What insights into the developmental traits of urodeles does the study of interspecific hybrids provide? *International Journal of Developmental Biology* 40, 885–893.

Wakeley, J. (2009). *Coalescent Theory*. Greenwood Village: Roberts and Company Publishers.

Weisrock, D.W., Shaffer, H.B., Storz, B.L., Storz, S.R. & Voss, S.R. (2006). Multiple nuclear gene sequences identify phylogenetic species boundaries in the rapidly radiating clade of Mexican ambystomatid salamanders. *Molecular Ecology* 15, 2489–2503.

Wilson, A.C., Maxson, L.R. & Sarich, V.M. (1974). Two types of molecular evolution. Evidence from studies of interspecific hybridization. *Proceedings of the National Academy of Sciences of the United States of America* 71, 2843–2847.

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

Wu, C.I. & Ting, C.T. (2004). Genes and speciation. *Nature Reviews Genetics* 5, 114–122.

Yatabe, Y., Kane, N.C., Scotti-Saintagne, C. & Rieseberg, L.H. (2007). Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* 175, 1883–1893.

Zhang, D.X. & Hewitt, G.M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* 12, 563–584.