THE 2013 AGA KEY DISTINGUISHED LECTURE

Models of Speciation: Where Are We Now?

SERGEY GAVRILETS

From the Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996 (Gavrilets); the Department of Mathematics, University of Tennessee, Knoxville, TN 37996 (Gavrilets); and the National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, TN 37996 (Gavrilets).

Address correspondence to Sergey Gavrilets at the address above, or e-mail: gavrila@tiem.utk.edu

Sergey Gavrilets is a Distinguished Professor at the Department of Ecology and Evolutionary Biology and the Department of Mathematics at the University of Tennessee, Knoxville. He obtained his PhD in physics and mathematics from the Moscow State University in the USSR in 1987. He subsequently worked at the Vavilov Institute of General Genetics in Moscow, the Institute National de la Recherche Agronomique in Toulouse, France, and the University of California at Davis before moving to Knoxville in 1995. He has received the President’s Award from the American Society of Naturalists (1999) and the Guggenheim Fellowship (2008). He has published a monograph on the mathematical theory of speciation (in 2004) and over 100 papers. His current research interests are in speciation, human origins, and human social and societal evolution.

Abstract

Theory building is an integral part of biological research, in general, and of speciation research, in particular. Here, I review the modeling work on speciation done in the last 10 years or so, assessing the progress made and identifying areas where additional effort is required. Specific topics considered include evolutionary dynamics of genetic incompatibilities, spatial and temporal patterns of speciation, links to neutral theory of biodiversity, effects of multidimensionality of phenotype, sympatric and parapatric speciation, adaptive radiation, speciation by sexual conflict, and models tailored for specific biological systems. Particularly challenging questions for future theoretical research identified here are 1) incorporating gene regulatory networks in models describing accumulation of genetic incompatibilities; 2) integrating models of community ecology with those developed in speciation theory; 3) building models providing better insights on the dynamics of parapatric speciation; 4) modeling speciation in multidimensional ecological niches with mating preferences based on multidimensional mating cues and sexual characters; 5) linking microevolutionary processes with macroevolutionary patterns as observed in adaptive radiations and paleontological record; 6) modeling speciation in specific systems studied by empirical biologists; and 7) modeling human origins. The insights from dynamic models of speciation should be useful in developing statistical tools that would enable empiricists to infer the history of past evolutionary divergence and speciation from genomic data.

Subject areas: Population structure and phylogeography; Quantitative genetics and Mendelian inheritance

Key Words: diversification, mathematical modeling, patterns, speciation

The diversity of species on earth continues to astonish and inspire scientists and the public alike. Its evolutionary origins were viewed by Darwin as the “mystery of mysteries.” It is not surprising therefore that the process of speciation has received an enormous amount of attention of evolutionary biologists. What makes this process extremely complex and
difficult to understand from the theoretical point of view is that different evolutionary factors controlling the dynamics of speciation (such as mutation, random genetic drift, recombination, natural selection, and sexual selection) act simultaneously and often have opposite effects.

Metaphors play an important role in science greatly simplifying thinking about complex processes. (In the words of Per Bak (1996), a brilliant theoretical physicist, “perhaps our ultimate understanding of scientific topics is measured in terms of our ability to generate metaphorical pictures of what is going on. Maybe understanding is coming with metaphorical pictures.”) Within the context of speciation, I find it useful to think of an individual as a point in the multidimensional genotype (or phenotype) space. A biological population can be imagined as a cloud of points. The size, structure, and location of the cloud are changed by different evolutionary factors and forces. Speciation can then be thought of as cluster formation accompanied by reproductive isolation between the emerging clusters (Dobzhansky 1937; Mallet 1995).

Within this metaphor, there are several necessary steps for speciation to occur. First, the population needs to accumulate and maintain enough genetic variation (i.e., the initial cluster has to be sufficiently spread in the genotype space). Then the population has to split into 2 (or more) subpopulations (clusters). This is usually thought to be done by natural and/or sexual selection. Finally, one needs to eliminate (or at least reduce significantly) the gene flow between the emerging clusters and prevent the reappearance of intermediate/hybrid genotypes. This is usually thought to be accomplished by the evolution of premating reproductive isolation. The eventual outcome of evolutionary divergence depends on a delicate balance of competing forces which also vary in space and time. As a result, “speciation can occur in different ways” (Dobzhansky et al. 1977) and “there are multiple possible answers to every aspect of speciation” (Mayr 1982).

Making sense of this complexity requires that mathematical modeling becomes an integral part of speciation research.

How can mathematical models help us better understand speciation? There are many questions that can be answered theoretically including those about the conditions for speciation, the probability of speciation, the waiting time to speciation, the duration of speciation, the degree of genetic and phenotypic divergence between the emerging species, the way different resources (including space) are partitioned between the sister species, and the effects of different parameters and factors (such as the strength of selection, rates of mutation, recombination, migration, population size, number of loci, distribution of allelic effects, etc.) on the dynamics of speciation. These questions have been at the center of the theoretical research on speciation for a long time.

One can identify 3 different periods in the process of building the mathematical theory of speciation. The first is the work of the founders of theoretical population genetics—Fisher, Wright, and Haldane. This work is relatively limited in scope. It included Fisher’s verbal model of runaway evolution caused by sexual selection and a verbal model of reinforcement (Fisher 1930), Wright’s verbal shifting balance theory (Wright 1931, 1982) and a model of assortative mating (Wright 1921), as well as studies of stochastic peak shifts (Haldane 1931; Wright 1941) and clines (Haldane 1948; Fisher 1950). A likely reason for the lack of more interest in speciation during that period was that the founders were still fighting “Darwin’s battle” focusing on selection and adaptation.

The second period started in the mid 1960s and lasted through the mid 1990s. During that time, some pioneering work laying the foundation for the mathematical theory of speciation was done by Maynard Smith (1962, 1966), Bazykin (1965), Bazykin (1969), Crosby (1970), Balkau and Feldman (1973), Dickinson and Antonovics (1973), Udovic (1980), Felsenstein (1981), Nei et al. (1983), Kondrashov (1983a), Kondrashov (1983b), Wu (1985), Kondrashov and Mina (1986), Diehl and Bush (1989), Liou and Price (1994), and others. Most of that work focused on sympatric speciation. Although many of the published models aimed to show its plausibility, their conclusions were not convincing enough to change the perception of the majority of evolutionary biologists that sympatric speciation was highly unlikely. Partially to “blame” for this was a brilliantly written theoretical paper by Felsenstein (1981) which clearly identified major obstacles to sympatric speciation. Another popular idea—founder effect speciation (Mayr 1942; Templeton 1980; Carson and Templeton 1984)—was strongly criticized on theoretical grounds by Barton and Charlesworth (1984). I think an unfortunate side effect of Felsenstein’s and Barton and Charlesworth’s papers was that studying speciation temporarily lost an appeal among theoreticians, delaying further theoretical advances.

The situation changed dramatically in the mid 1990s when modeling speciation started to move to the forefront of theoretical research in evolutionary biology. The third period, still ongoing, has seen a lot of theoretical advances. The work done during its first decade was summarized in Kirkpatrick and Ravigné (2002), Coyne and Orr (2004), Dieckmann et al. (2004), and Gavrilets (2003, 2004). My goal here is to evaluate theoretical advances over the last 10 years and to revisit some old controversies. Specific topics to be considered include evolutionary dynamics of genetic incompatibilities, spatial and temporal patterns of speciation, links to neutral theory of biodiversity, effects of multidimensionality of phenotype, sympatric speciation, adaptive radiation, speciation by sexual conflict, and models tailored for specific biological systems. Due to space, time, and expertise limitations, my analysis is not comprehensive. In particular, I do not discuss the theoretical work on reinforcement (Servedio and Noor 2003; Servedio et al. 2009; Olofsson et al. 2011) or issues of hypothesis testing and statistical inference about evolutionary processes from data. Although I have tried to give a balanced overview, some biases are probably unavoidable, especially when I give a historical context of certain controversies. My major focus will be on the tools developed by theoreticians, that is, the mathematical models, but I will also address some biological insights achieved using these tools. Some of the topics I discuss below were reviewed to a certain extent in recent publications, but having the pieces and bits currently
spread across the literature in a single place might be useful for the reader.

Recent Progress

The Bateson–Dobzhansky–Muller Model

One of the stimuli for a renewed interest in building a quantitative theory of speciation was the development of mathematical models (Orr 1995; Gavrilets and Hastings 1996; Orr and Orr 1996) formalizing Dobzhansky’s old idea (Dobzhansky 1936, 1937) on the appearance of genetic incompatibilities and the evolution of reproductive isolation as a side effect of genetic divergence.

Following Dobzhansky (1937), consider a 2-locus, 2-allele diploid population initially monomorphic for a genotype, say AAbb. Let us assume that this population is split into 2 geographically isolated parts. In one part, evolutionary forces cause substitution of B for b and a subpopulation ABBB is formed. In the other part, there is a substitution of a for A, giving rise to a subpopulation aabb. Assume there is no reproductive isolation among genotypes AAbb, Aabb, and aabb and among genotypes AAbb, Aabb, and ABBB. In contrast, let the cross of AABB and aabb be difficult or impossible because alleles a and B are incompatible in the sense that their interaction “produces one of the physiological isolating mechanisms” (Dobzhansky 1937). In this model, substitutions of B for b in one population and a for A in the other population will result in the emergence of 2 reproductively isolated populations.

A very similar (practically identical) description of this idea was given by Muller (1939, 1942) (without citing Dobzhansky) who also provided some refinements. Much earlier, Bateson (1909) put forward a similar verbal argument as discovered by Orr (1996). The underlying mechanism is now referred to as the Dobzhansky model, the Dobzhansky–Muller model, or the Bateson–Dobzhansky–Muller (BDM) model of genetic incompatibilities.

The idea that genes brought together in hybrids are incompatible is very intuitive. Indeed, the loss in fitness to species hybrids is no more surprising than the fact that a part from one car manufacturer does not function in a car from another manufacturer (Charlessworth 1990). However, it took 60 years for this idea to start playing an important role in speciation research. One reason for this delay was the dominance of Wright’s metaphor of rugged fitness landscapes in evolutionary biology (Wright 1932, 1988; Provine 1986). Within this metaphor, speciation is imagined as a peak shift, that is, a transition from one fitness peak to another fitness peak and implies evolution across a valley of maladaptation. Theoretical work has shown that such transitions are very difficult to accomplish (Barton and Charlesworth 1984, Coyne et al. 2000; Gavrilets 2004). Note that Wright’s shifting balance theory (Wright 1932, 1982) was put forward as a possible solution to the problem of peak shifts, but later work has shown this solution is not quite satisfactory (Coyne et al. 2000). However, within the framework of the Dobzhansky model, going through the valley of maladaptation is not needed as the population can move to a reproductively isolated state along a ridge of high fitness values (Gavrilets 1997). Ironically, for many years Dobzhansky was promoting Wright’s rugged landscapes without realizing that his favored mechanism of speciation implies a very different structure of fitness landscapes.

Accumulation of Genetic Incompatibilities and Reproductive Isolations

Two major general predictions emerged from earlier work on genetic incompatibilities. The first is that genetic divergence between allopatric populations should result in a faster than linear growth in the number of genetic incompatibilities (“snowball effect,” Orr 1995; Orr and Orr 1996). The second is that as a result of this process, there should be a rapid transition in the degree of reproductive isolation from low to high (“threshold effect,” Gavrilets 2004). The threshold effect is most pronounced when strong reproductive isolation requires multiple complex incompatibilities. It follows from the fact that the probability of reproductive compatibility between 2 genotypes decreases exponentially with the number of genetic incompatibilities which itself grows polynomially with the number of substitutions. The “snowball effect” and the “threshold effect” were predicted for both simple (i.e., dyadic) and complex (e.g., including combinations of multiple genes) incompatibilities. Genetic incompatibilities were studied within the scenarios of founder effect speciation (Gavrilets and Hastings 1996) and sympatric speciation by drift and mutation (Wu 1985). A mathematical generalization of the BDM model has led to the emergence of the concept of holey fitness landscapes (Gavrilets 1997; Gavrilets and Gravner 1997). The latter were independently developed within studies of RNA secondary structures where they became known as neutral and nearly neutral networks (Reidys 1997; Reidys et al. 1997; Reidys and Stadler 2001). The threshold effect can be used to justify a simple model of reproductive isolation in which 2 populations are reproductively compatible as long as the genetic distance between them remains below a particular constant $K$. Once genetic distance reaches $K$, the populations become reproductively isolated. This model has been used for studying the conditions for, and the time to, parapatric speciation with or without selection for local adaptation (Higgs and Derrida 1991; Manzo and Pelti 1994; Gavrilets et al. 1998; Gavrilets 1999a, 2000b; Gavrilets et al. 2000b; Gavrilets 2004).

More recently, the Dobzhansky model has been extended in different directions. Allopatric divergence was investigated in several papers using various assumptions about genetic architecture. For example, Fierst and Hansen (2010) focused on multiple genes controlling a trait under stabilizing selection. Barton and de Cara (2009) studied the accumulation of incompatibilities in the presence of assortment. Gourbiere and Mallet (2010) questioned the generality of the “snowball effect” and develop alternative models predicting a linear increase in reproductive isolation with time. Fraisse et al. (2014) showed that reproductive isolation due to simple and complex incompatibilities evolve at similar rates (see also the earlier work by Welch 2004). Wang et al. (2013) extended the BDM model to multispecies clades by describing the
mathematical relationship between tree topology and the number of shared incompatibilities among related pairs of species. Parapatric divergence in a deterministic model was studied by Bank et al. (2012) who concluded that “substitutions involved in a parapatric [BDM incompatibilities] must be adaptive” for divergence to be stably maintained. Yamaguchi and Iwasa (2013) studied a stochastic model of accumulation of incompatibilities between populations exchanging migrants at a low rate. In their model, reproductive isolation (and speciation) occurs once the genetic distance between the populations reaches a prespecified threshold. Using the diffusion approximation, Yamaguchi and Iwasa (2013) developed analytical predictions for the time to speciation.

Although rather different biologically, accumulation of genetic incompatibilities affecting viability and/or fertility and divergence in mating preferences can be treated theoretically within a single theoretical framework of fitness landscapes (Gavrilets 2004, Chap. 2). Uyeda et al. (2009) used numerical simulations to argue that allopatric population can readily diverge in mating preferences by random genetic drift. In their model, populations can evolve along a single phenotypic axis. Increasing the dimensionality of phenotype (or genotype) space will make divergence by drift easier.

In most earlier models, each gene could be potentially incompatible with each other gene (but see Johnson and Porter 2000). Rather than using standard population genetic models, Palmer and Feldman (2009) explicitly modeled gene regulatory networks. Intriguingly, in their simulations there is no snowball effect while populations can drift in and out of incompatibility with reproductive isolation evolving in a punctuated manner. Livingstone et al. (2012) studied the rate of accumulation of genetic incompatibilities in a model adapted for Saccharomyces cerevisiae protein–protein interaction network. In an intriguing paper, ten Tusscher and Hogeweg (2009) explicitly modeled transcription factors while allowing for multiple ecological niches and assortative mating dependent on genetic distance between individuals. A major claim of their paper is that sympatric speciation can happen easily. I note that moving from classical population genetics models to gene network models promises to greatly increase our understanding of the dynamics of speciation. More work in this direction is definitely needed. Other largely neglected areas are the accumulation of mutually incompatible advantageous mutations (Church and Taylor 2002; Kondrashov 2003; Gavrilets 2004) and spatially explicit models.

Isolation by Genetic Distance and the Neutral Theory of Biodiversity

Earlier work used the threshold model of reproductive isolation to study the dynamics of diversification in spatially distributed populations (Gavrilets et al. 1998, 2000a, 2000b). In particular, it was shown that the interaction of the isolation by spatial distance and by genetic distance can lead to rapid emergence of multiple species. This process generates diversity by mutation and random genetic drift.

The unified neutral theory of biodiversity also postulates that the differences between individuals of the same species or different species at the same trophic level are “neutral,” or irrelevant for their birth, death, and dispersal rates (Hubbell 2001; McGill 2003; Volkov et al. 2003; Kopp 2010; Rosindell et al. 2011; Ricklefs and Renner 2012). This assumption greatly simplifies mathematical derivations, allowing one to make a number of specific predictions about various characteristics of populations, species, and communities, such as species abundances and species-area relationships. Although this theory provided explanations for a number of observed patterns and some of its predictions are well supported by data, the theory remains extremely controversial (Rosindell et al. 2011; Ricklefs and Renner 2012).

One of the crucial parameters of this model is the “speciation rate” which, in this theory, is simply the probability that an offspring of an individual belongs to a different species. In the original formulation of the unified neutral theory of biodiversity, “speciation rate” is treated as an exogenously specified constant parameter. The value of this parameter has been shown to have extremely strong effects on relative species abundances and other measurable characteristics. In practice, “speciation rate” is estimated by maximizing the fit between predictions and observations. Given all that we know about speciation, treating species origin as an instantaneous process analogous to mutation looks to be an extremely suspicious approximation which casts uncertainty on the conclusions reached.

Recently, however, researchers started to use the threshold model of reproductive isolation (see above) within the context of the neutral theory of biodiversity (Hoelzer et al. 2008; de Aguiar et al. 2009; Melián et al. 2012; Desjardins-Proulx and Gravel 2012a, 2012b; Baptestini et al. 2013a, 2013b; Martins et al. 2013). This work alleviates some of the problems of the earlier approaches, while confirming the general ability of the unified neutral theory to predict empirical patterns of biodiversity (Kopp 2010). A fuller integration of the studies of community ecology with evolutionary biology and speciation research is a very challenging but necessary next step.

Sympatric Speciation

The topic of the plausibility and generality of sympatric speciation has been a source of controversy for several generations of evolutionary biologists. E. Mayr was very skeptical about this mode of speciation and argued forcibly against its biological relevance throughout his long career. But he also knew that the issue was not going to disappear. In 1963, he commented that “one would think it should no longer be necessary to devote much time to this topic, but past experience permits one to predict with confidence that the issue will be raised again at regular intervals. Sympatric speciation is like the Lernaean Hydra which grew two new heads whenever one of its old heads was cut off.” Mayr (1963). Mayr’s prediction was proved to be correct almost immediately as first Maynard Smith (1966) and then others (e.g., Dickinson and Antonovics 1973; Udovic 1980) published a series of theoretical papers arguing that sympatric speciation is not difficult at all. These papers led Felsenstein (1981) to note that “while these authors have largely been concerned with
showing that sympatric speciation is possible, one might come away from some of these papers with the disturbing impression that it is all but inevitable.” Felsenstein then went to highlight a number of theoretical obstacles to sympatric speciation. His paper was very influential in restoring the Mayrían status quo, although some dissenters remained (Kondrashov 1983a, 1983b; Wu 1985; Kondrashov and Mina 1986; Diehl and Bush 1989). The situation changed again a generation later after the journal Nature published 2 back-to-back theoretical papers (Dieckmann and Doebeli 1999; Kondrashov and Kondrashov 1999) arguing once again that sympatric speciation is very plausible. These papers generated both a lot of excitement in the field and intense critique with the latter being focused on the biological relevance and plausibility of the assumptions and parameter values used in the models. Simultaneously, there has been an explosion of empirical work on speciation driven by ecological factors (Schluter 2000; Nosil 2012) which play a crucial role in models of sympatric speciation.

The resulting controversy was very beneficial for the field as it led to a flurry of theoretical and empirical advances. Earlier theoretical work relied almost exclusively on numerical simulations, which are always more difficult to generalize from than analytical results. However, conditions for sympatric speciation were found analytically (Gavrilets 2003, 2004) in at least 11 different models including the classical models of Maynard Smith (1966), Udovic (1980), Felsenstein (1981), and Diehl and Bush (1989). For example, the classical Udovic model considers 2 loci, one of which is under disruptive natural selection and another controls nonrandom mating. If the loci are unlinked, sympatric speciation occurs if

\[ s + a > 1, \]

where \( s \) is the fitness loss of hybrids and \( a \) is the probability of assortative mating (Gavrilets 2003). The above inequality implies that the joint effect of disruptive selection (\( s \)) and nonrandom mating (\( a \)) have to be sufficiently large. Mathematically similar conditions emerge in other models. There was also a very large number of numerical studies of more complex and realistic models of sympatric speciation (van Doorn and Weissing 2001; Dieckmann et al. 2004). The debate about sympatric speciation also introduced and popularized a powerful analytical technique for studying evolutionary diversification—adaptive dynamics (Geritz et al. 1998; Dieckmann et al. 2004; Waxman and Gavrilets 2005; Doebeli 2011).

By the mid 2000s, all this work has led to a set of generalizations about the conditions for sympatric speciation (Gavrilets 2004, 2005). These are: 1) strong joint effects of disruptive selection and nonrandom mating, 2) high levels of initial genetic variation (or very high mutation rates, presence of other mechanisms promoting increase in genetic variance, etc.), 3) close association of traits experiencing disruptive selection and those controlling nonrandom mating (which can be achieved by a close linkage of the corresponding genes, the “habitat choice” mechanism, or the “magic trait” mechanism), and 4) the absence of costs of being choosy.

More recent work has extended and clarified these findings (de Cara et al. 2008; Kopp and Hermisson 2008; Otto et al. 2008; Pennings et al. 2008; Thibert-Plante and Hendry 2009, 2011; Peischl and Schneider 2010; Doebeli 2011; Kisdi and Priklopil 2011; Rettelbach et al. 2011, 2013). For example, Bürger and Schneider (2006) studied numerically the case of strong selection and multiple loci. They showed a possibility of up to 5 species emerging sympatrically. They also observed limit cycles in the the population structure dynamics. Pennings et al. (2008) found analytically the conditions for sympatric speciation in an one-locus model accounting for stabilizing selection, competition, and nonrandom mating. Barton (2010) developed analytical methods for studying speciation by local adaptation and habitat preference (see also the earlier work by Fry 2003). Kisdi and Priklopil (2011) studied branching in a magic trait (i.e., phenotypic traits involved in the existence of discrete ecological niches and from competition within each niche. Disruptive ecological selection favors the evolution of sexual preferences for ornaments that signal local adaptation. van Doorn et al. (2009) showed that the interaction of natural and sexual selection can lead to local adaptation and reproductive isolation without the divergence in mating preferences. The latter is commonly assumed to be necessary for speciation. Equally unexpectedly, M’Gonigle et al. (2012) showed that long-term coexistence of species within overlapping ranges is possible in absence of ecological differentiation. In their model, coexistence requires spatial variation in local carrying capacity, and mate-search costs in females.

Mathematical models clearly show that under certain biologically reasonable conditions, sympatric speciation, in general, and â™ adaptive speciation, â™™ in particular, are indeed possible. How biologically plausible these conditions are is an empirical rather than theoretical question, remaining largely open.

Empirical papers arguing that sympatric speciation happens continue to be published (e.g., Crow et al. 2010; Papadopolos et al. 2011; Hadid et al. 2013, 2014). An important question that has recently received substantial attention is what one actually means by sympatric speciation and how it can be demonstrated in nature (Butlin et al. 2008; Fitzpatrick et al. 2008, 2009; Mallet et al. 2009). In the literature, different investigators have used different definitions of sympatric speciation and different criteria for diagnosing cases of sympatric speciation. In particular, empiricists typically use much broader definitions than theorists. Some authors now argue that sympatric speciation is virtually impossible to demonstrate in an uncontroversial way and that, in any case, testing whether a particular case fits a particular definition of sympatric speciation is less informative than evaluating the biological processes affecting divergence (Fitzpatrick et al. 2008, 2009). These authors
believe that geographical context of speciation can be better understood by modeling and measuring quantities, such as gene flow and selection, rather than assigning cases to discrete categories like sympatric and allopatric speciation.

Environmental Gradients

The majority of natural populations experience spatially heterogeneous selection. Theoretical studies of environmental gradients in selection and their effects on genetic variation have a long history (Haldane 1948; Fisher 1950; Endler 1977). The classical work has mostly focused on the shape and steepness of clines in allele frequencies or the values of a quantitative trait emerging from the interaction of selection and local dispersal. One of the major theoretical findings emerging from that work was that smooth changes in selection across a species range can cause abrupt genetic (or phenotypic) discontinuities, hybrid zones, and sharp species boundaries (Endler 1977; Slatkin 1978; Kirkpatrick and Barton 1997; Barton 1999).

In 2003, Doebeli and Dieckmann published an extension of their earlier model (Dieckmann and Doebeli 1999) to the case of spatially distributed populations experiencing spatially heterogeneous stabilizing selection with a linearly changing optimum (see also the earlier work by Balkau and Feldman 1973; Caisse and Antonovics 1978). Their numerical results demonstrated the formation of spatially separated, genetically distinct, and reproductively isolated clusters. These results have led the authors to conclude that parapatric speciation can occur easily. They also observed that the most conducive conditions for speciation were when the selection gradients had intermediate slopes.

Initially these results were interpreted as an artifact of edge effects (Gavrilets 2004; Polechová and Barton 2005) and specific genetic architecture (Polechová and Barton 2005). However, more recent work has shown that edge effects are not crucial and that for certain parameter values, genetic clustering and speciation are a robust outcome of the interaction of spatially heterogeneous selection and local migration (Kawata et al. 2007; Leimar et al. 2006; Heinz et al. 2009; Ispolatov and Doebeli 2009; Rettelbach et al. 2013). Moreover, the effect of intermediate selection slopes on diversification has been observed in a number of other models (e.g., Gavrilets et al. 2007; Rettelbach et al. 2013). There is mounting evidence that this effect is a general feature of adaptive diversification (Gavrilets and Losos 2009, see below).

More recently, in a related work, Payne et al. (2011) showed that parapatric speciation can also be caused by the evolution of conditional dispersal, while Irwin (2012) extended the modeling framework using an explicit consideration of the phylogenetic relationship between emerging species. I note that parapatric speciation is the least theoretically studied geographic mode of speciation.

Multiple Speciation Events and the Dimensionality of Niche and Phenotype Space

Earlier models of speciation focused on the first speciation event leading to the increase in the number of species from 1 to 2. The appearance of multiple species was either not possible by modeling design or was not studied explicitly. More recently, modeling work has started to focus explicitly on multiple speciation events. There are 3 general sets of models describing the emergence of multiple species: 1) models of neutral divergence in a spatially distributed population with isolation by spatial and genetic distance, 2) models of diversification within a single ecological niche driven by frequency-dependent selection due to competition, and 3) models with multiple discrete ecological niches.

The first set of models were already discussed above (within the context of the threshold model of reproductive isolation and the unified neutral theory of biodiversity). The second set of models builds on classical models in which frequency-dependent competition based on a single quantitative character results in “branching” (i.e., the emergence of discrete phenotypic and/or genetic clusters), if competition is stronger than stabilizing selection (Roughgarden 1972; Christiansen and Loeschcke 1980; Dieckmann and Doebeli 1999; Dieckmann et al. 2004; Doebeli 2011). For example, Bürger et al. (2006), Bolnick (2006), and Doebeli et al. (2007) showed that frequency-dependent competition based on a single quantitative character can produce multiple clusters/species. Multiple clusters were also observed in models explicitly describing competition mediated by 2 quantitative characters (Vukics et al. 2003; Ito and Dieckmann 2007, 2012). In an intriguing recent paper, Doebeli and Ispolatov (2010) show that the probability of generating and maintaining diversity by competition increases with the dimensionality of the phenotype space. This suggests that earlier studies which focused on a single quantitative character (mostly for mathematical simplicity) may significantly underestimate the power of competition to cause adaptive diversification. Note that some of this work concerns only the emergence of “ecological species” (i.e., discrete clusters of genotypes or phenotypes adapted to different ecological niches) as the evolution of premating isolation was not studied explicitly.

Multiple species can naturally emerge if there are multiple discrete ecological niches. In Gavrilets and Vose (2005) model focusing on local adaptation and habitat preference, a number of emerging “ecological species” equals the number of ecological niches. Later, Gavrilets and Vose (2009) extended their model by introducing additional mating characters and the possibility of premating reproductive isolation. They observed the emergence of up to 3 reproductively isolated “sexual species” (i.e., discrete clusters of genotypes differing in mating characters only) per each ecological niche. For example, 8 ecological niches would produce 24 different species, reproductively isolated by mating preferences and/or habitat preferences and ecological traits. Birand et al. (2012) showed that even without any differences in mating traits, the number of “ecological species” can be much larger than the number of basic ecological niches. For example, in their model with 4 different resources, it is possible to observe the emergence of up to 15 different species: 4 single-resource specialists, 6 species utilizing 2 resources, 4 species utilizing 3 resources, and 1 species utilizing all 4 resources. Some biological implications of these theoretical results are discussed in the next section.
While theoretical work on speciation into multiple ecological niches has been relatively broad, not much has been done so far on models of nonrandom mating based on multiple mating cues and secondary sexual traits. Recently, Thibert-Plante and Gavrilets (2013) studied a series of models of ecological speciation with multiple types of phenotypic traits which can be used for mating decisions. Their work shows that certain traits that are under direct natural selection are more likely to be co-opted as mating cues, leading to the appearance of “magic traits.” They also observed that multiple mechanisms of nonrandom mating can interact so that trait coevolution enables the evolution of nonrandom mating mechanisms that would not evolve alone. Given that data clearly show that multiple sensory modalities and multiple cues control mate choice in a wide range of animal taxa (Hohenlohe and Arnold 2010; Oh and Shaw 2013), explicitly modeling multidimensional mating characters can lead to new insights on the dynamics of speciation and large-scale diversification.

Patterns of Adaptive Radiation

The term “adaptive radiation” refers to evolutionary groups that exhibit an exceptional extent of adaptive diversification into a variety of ecological niches, with such divergence often occurring extremely rapidly (Givnish and Sytsma 1997; Schluter 2000; Losos and Miles 2002). Classic examples of adaptive radiation include Darwin’s finches on the Galápagos islands, Anolis lizards on Caribbean islands, Hawaiian silverswords, and cichlids of the East African Great Lakes. Theoretical studies of adaptive radiation using population genetic models have not started until relatively recently. The analysis of different mathematical models describing the emergence of multiple species (outlined in the previous section) has led to a number of generalizations regarding patterns of adaptive radiation (Gavrilets and Losos 2009). These patterns are:

1) Early burst of evolutionary divergence: Typically, there is a burst of speciation and morphological diversification happening soon after the onset of the radiation rather than divergence at similar rates through time.

2) Overshooting: There is an early increase in species diversity which is followed by its decline.

3) Stages of radiation: All else being equal, a particular sequence of the diversification events is expected: (i) divergence with respect to macrohabitat; (ii) evolution of microhabitat choice and divergence with respect to microhabitat; (iii) divergence with respect to “magic traits” (i.e., traits that simultaneously control local adaptation and nonrandom mating); and (iv) divergence with respect to other traits controlling survival and reproduction.

4) Area effects: Speciation driven by ecological factors is promoted by larger geographic areas (e.g., of islands or lakes).

5) Nonallopatric diversification: Speciation during adaptive radiation can occur in the presence of some gene flow between diverging populations.

6) Selection gradient effect: Parapatric speciation is promoted by selection gradients of intermediate slopes.

7) Spatial dimensionality effect: Geographic areas that can approximately be viewed as one-dimensional (such as rivers or shores of lakes and oceans) promote more speciation and tend to maintain higher species richness and phenotypic and genetic diversity per unit area than geographic areas that are 2-dimensional (such as lakes, oceans, and continental areas).

8) Least action effect: Speciation occurring after the initial burst of diversification usually involves a minimum genetic change.

9) Effect of the number of loci: Rapid and extensive diversification is most likely if the number of loci underlying the traits under selection is small.

10) Porous genome effect: Species can stably maintain their divergence in a large number of selected loci for very long periods despite substantial hybridization and gene flow that decreases or removes differentiation in neutral markers.

11) Short duration of speciation: The time interval during which intermediate forms are present is relatively short (Gavrilets 2004).

12) Disparity versus diversity: Morphological disparity increases most rapidly early in the clade history at low levels of species diversity (Gavrilets 1999b).

Some of these patterns are strongly supported by empirical work (Gavrilets and Losos 2009), whereas for others, empirical support is tentative. In almost all cases, more empirical data are needed. A huge theoretical challenge that still remains is to link microevolutionary processes with macroevolutionary patterns observed by paleontologists (Eldredge et al. 2005; Futuyma 2010; Uyeda et al. 2011).

Speciation by Sexual Conflict

Sexual conflict is present if the interests of the sexes with regard to certain aspects of reproduction differ. There are numerous examples including sexual conflict over the mating rate, parental care, offspring size, use of sperm, and epigenetic control of development (Rice and Holland 1997; Holland and Rice 1998; Rice 1998; Haig 2000; Smith and Haxby 2000; Barra et al. 2002; Ball and Parker 2003; Rice et al. 2012). There are 2 general types of sexual conflict (Chapman and Partridge 1996; Parker and Partridge 1998). One is within-locus conflict which occurs when the locus controls a trait expressed in both sexes and the optimum trait values differ between the sexes. Another is between-locus conflict which occurs when there are different traits each expressed in one sex only but affecting the fitness of both sexes in opposite directions. Studies of sexual conflict and sexually antagonistic coevolution moved to the forefront of experimental and theoretical research in evolutionary biology by the end of the last century (Rice and Holland 1997; Holland and Rice 1998; Rice 1998). A particularly exciting idea that has emerged from these studies is that sexual conflict can be an important “engine of speciation” (Rice 1996, 1998; Howard et al. 1998; Parker and Partridge 1998).

Initial work—experimental (Martin and Hosken 2003), comparative (Arntqvist et al. 2000), and mathematical
(Gavrilets 2000a)—provided strong support for this assertion. In particular, mathematical models have shown that between-locus sexual conflict can lead to rapid genetic changes in both sexes (a “coevolutionary chase” scenario) within local populations which can lead, as a by-product, to reproductive isolation between different allopatric populations and speciation. It was also shown that sexual conflict can affect the success of reinforcement of preexisting isolation (Parker and Partridge 1998). Moreover, it can play an important role in the evolution and diversification of ecological traits and strategies through the co-optation of sexual traits for viability-related functions (Bonduriansky 2011).

However, further modeling work revealed not only the power of sexual conflict but also constraints on the evolution of reproductive isolation as driven by sexual conflict. From theoretical studies, we now know that speciation is only one of several possible evolutionary outcomes of sexual conflict. Overall, there are at least 6 different dynamics observed in models of between-locus sexual conflict:

1) continuous “coevolutionary chase” between the sexes which can lead, as a by-product, to reproductive isolation between diverging allopatric populations (Gavrilets 2000a; Gavrilets et al. 2001; Gavrilets and Waxman 2002; Rowe et al. 2003; Gavrilets and Hayashi 2006);
2) evolution toward an equilibrium (Kondoh and Higashi 2000; Gavrilets et al. 2001; Kimura and Ihara 2009);
3) evolution toward a line of equilibria with subsequent random drift along this line (Gavrilets 2000a; Gavrilets and Hayashi 2005; Hayashi et al. 2007);
4) cyclic evolution (Gavrilets et al. 2001; Haygood 2004);
5) “Buridan’s Ass” regime involving extensive sympatric diversification in female alleles without comparable diversification in male alleles (Frank 2000; Gavrilets and Waxman 2002; Gavrilets and Hayashi 2005; Hayashi et al. 2007);
6) extensive diversification in both male and female alleles (Frank 2000; Gavrilets and Hayashi 2005; Hardling and Bergsten 2006; Hayashi et al. 2007; Hårdling and Karlsson 2009), including a possibility of sympatric speciation.

Mathematical models (Gavrilets and Hayashi 2005; Hayashi et al. 2007) also show that 1) different dynamic regimes can be observed with the same set of parameter values but under different initial conditions; 2) a population can switch from one regime to another as a result of stochastic perturbations due to, say, random genetic drift; 3) different loci controlling mating and fertilization in the same population can follow different dynamic regimes. Overall, the evolutionary dynamics observed in models of sexual conflict is extremely varied and complex (Rowe et al. 2003; Gavrilets and Hayashi 2005). These models have focused on isolated, well-mixed populations. Adding spatial structure and limited dispersal is naturally expected to complicate their dynamics even further.

In line with theoretical predictions, both experimental evolution studies and comparative analyses of fertilization proteins and of species richness show that sexual conflict leads to, or is associated with, reproductive isolation and speciation only in some cases but not in others (reviewed in Gavrilets 2014). Sexual conflict is an engine of speciation. But as with other engines of speciation—ecological selection, sexual selection, and accumulation of incompatibilities between diverging lineages by mutation and drift—it is most efficient when operating under optimum conditions and can stall otherwise.

Models Tailored for Specific Biological Systems

So far, I have focused on models of speciation built from basic principles of evolutionary dynamics and aiming for both generality and mathematical simplicity. These models are very useful and insightful in uncovering general rules and patterns of speciation, adaptive radiation, and biological diversification as discussed above. However, their generality almost necessarily implies that these models are very difficult to apply to specific biological systems studied by empirical biologists. Therefore, it is very important to supplement simple general models of speciation with those tailored for specific biological case studies. A small number of such models have been developed only relatively recently for some of the best studied systems including those aiming to capture the dynamics of nonallopatric speciation of cichlids in a lake (Lande et al. 2001; Gavrilets et al. 2007; Kawata et al. 2007; Aguillé et al. 2011) and palms on an oceanic island (Gavrilets and Vose 2009), hybrid speciation in butterflies in Central America (Duenez-Guzman et al. 2009), and eco-morph formation in marine snails in Sweden (Sadédin et al. 2009). Yamamichi and Sasaki (2013) developed a model of speciation in pulmonate snails driven by changes in a single gene controlling “handedness” (i.e., the direction of coiling of their shells; see also earlier work by Orr 1991). Such projects necessarily require close collaboration between theorists and empiricists which can be highly rewarding for both.

There are several useful lessons of these attempts. First, mathematical models emerging from these projects do lead to a better understanding of the evolutionary dynamics of the studied specific systems. Second, although the relevant models are case-specific, they contribute toward building the general theory of speciation, for example, by supporting or undermining the generality of particular observations and patterns (as discussed above). Third, the process of building a mathematical model even for a particularly well-studied empirical system usually reveals the lack of biological understanding or crucial empirical data needed to make appropriate modeling assumptions or specify parameters. This can greatly stimulate further empirical studies. Due to the specificity of empirical work, a theoretician’s “wish list” can remain unfulfilled for a long time. This however is not fatal for the success of modeling efforts as the effects of the differences in assumptions and parameters on the model’s predictions can often be explored numerically.

There is also a possibility for some disappointment and frustration for empiricists as their expectations and intuitions about their systems and the processes they study are not
always supported by modeling. For example, spatial heterogeneity in the model of diversification of palms (Gavrilets and Vose 2007) and isolation by distance in the model of speciation in cichlids (Gavrilets et al. 2007) appear necessary for speciation which weakens the case for sympatric speciation made in the 2 original empirical papers (Barluenga et al. 2006; Savolainen et al. 2006). The hybrid species emerging in the model of hybrid speciation in butterflies (Duenez-Guzman et al. 2009) was transient rather than stable (as was implied in the empirical paper which introduced this system, Mavarez et al. 2006). The overall effect of collaboration between empiricists and theoreticians is nevertheless overwhelmingly positive. More studies of this type are definitely needed.

Conclusions

Any modeling approach obviously has its own limitations, and the true insight will most likely arise from a comparison among different approaches and different mathematical models that reveals deeper generalities. Therefore, more models and more applications of these models are needed. Looking back, one can clearly see significant progress in theoretical speciation research over the last 10 years. During that period, some controversies have largely been resolved, at least for now. At the same time, theory is still incomplete.

Butlin et al. (2012) have recently published a review entitled “What do we need to know about speciation?” arranged around 13 specific questions. Six of these concern the origin and build-up of reproductive isolation, 4 are about genetic and genomic signatures of speciation, and 3 are on the connection between speciation and biodiversity. The authors did not discuss explicitly the role of theory in answering these questions but existing models can already provide both a general framework and theoretical guidance for answering the majority of these questions.

To the Butlin list I would add one particular entry. Decades of intensive work by generations of evolutionary biologists have led to a dramatic increase in our understanding of how new species arise. I believe that the time is ripe for attacking the ultimate speciation event—the origin of our own species (Darwin 1871). Understanding human origins will obviously have enormous implications for both science and society.

The models discussed above describe the dynamics of speciation forward in time. Given a set of assumptions about particular evolutionary forces and factors in play one can use these models to predict certain features of diverging populations and emerging species. The current explosion of genomic data generated by next generation DNA sequencing makes it particularly important to be able to predict the patterns of divergence at the genomic level (Seehausen et al. 2014). Some predictions can be made using well-established theory. For example, the properties of “islands of genomic divergence” in FST genomic scans (Via 2012; Feder et al. 2013) can be predicted using the classical theory of barriers to gene flow built by Bengtsson (1985), Barton and Bengtsson (1986), and others (see Juric, Gavrilets S, unpublished data). Other questions, for example, those about genome rearrangements (Renaut et al. 2013; Yeaman 2013), require new models and theory.

Being able to predict what should happen under certain circumstances is not enough though. As was already articulated by Dobzhansky and Mayr (see the quotes at the beginning) and is well illustrated by the modeling work discussed above, particular speciation outcomes and patterns can result from rather different, alternative scenarios, factors, and forces. Therefore, having appropriate theoretical tools for solving the reverse problem—how to estimate and evaluate different evolutionary forces and factors that have led to observed patterns of (genomic) divergence and how to test various hypotheses about ongoing or completed speciation—is essential (Butlin et al. 2012; Seehausen et al. 2014). A clear understanding of the links between speciation processes and their phenotypic, genetic, and genomic consequences studied in the theoretical models discussed above should definitely help in developing sound statistical methodologies to identify and evaluate speciation factors and processes from phenotypic, genetic, and genomic data.

Funding

National Institute for Mathematical and Biological Synthesis through NSF Award (#EF-0830858), with additional support from The University of Tennessee, Knoxville.

Acknowledgments

I thank K. L. Shaw for inviting me to the symposium. I also thank J. A. Fordyce, J. Hermisson, K. Rooker, K. L. Shaw; and reviewers for comments and suggestions.

References


