Introduction

Systematics and the origin of species: An introduction

Jody Hey*, Walter M. Fitch†, and Francisco J. Ayala‡‡

*Department of Genetics, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; and †Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Ernst Mayr, one of the 20th century’s greatest scientists and a principal author of the modern theory of evolution, passed away on February 3, 2005, at the age of 100. From December 16 to 18, 2004, before Mayr’s passing, a colloquium on “Systematics and the Origin of Species” sponsored by the National Academy of Sciences was held in his honor. The colloquium’s title was the same as that of Mayr’s 1942 book (1), generally considered one of the foundation books of the theory of evolution. The 17 papers that follow explore current knowledge about the main topics of Mayr’s book.

The modern theory of evolution embodies a complex array of biological knowledge centered around Darwin’s theory of evolution by natural selection couched in genetic terms. It is not one single theory with its corroborating evidence, but a multidisciplinary body of knowledge bearing on biological evolution: an amalgam of well-established theories and working hypotheses together with the observations and experiments that support accepted hypotheses (and falsify rejected ones), which jointly seek to explain the evolutionary process and its outcomes. These hypotheses, observations, and experiments originate in disciplines such as genetics, developmental biology, neurobiology, zoology, botany, paleontology, and molecular biology.

Darwin’s theory of evolution (2) argued that natural selection, the process accounting for the adaptation and diversity of organisms, emerges as a necessary conclusion from two premises: (i) the assumption that hereditary variations useful to organisms occur and (ii) the observation that more individuals are produced than can possibly survive. A serious difficulty facing Darwin’s evolutionary theory was the lack of an adequate theory of inheritance that would account for the preservation through the generations of the variations on which natural selection was supposed to act. Theories then current of “blending inheritance” proposed that offspring struck an average between the characteristics of their parents. As Darwin became aware, blending inheritance could not account for the conservation of variations, because differences among variant offspring would be halved each generation, rapidly reducing the original variation to the average of the preexisting characteristics.

The missing link in Darwin’s argument was provided by Mendelian genetics. Mendel’s paper published in 1866 (3) formulated the fundamental principles of a theory of heredity that accounts for biological inheritance through particulate factors (now called “genes”) inherited one from each parent that do not mix or blend but segregate in the formation of the sex cells, or gametes. Mendel’s discoveries, however, remained unknown to Darwin and, indeed, did not become generally known until 1900, when they were simultaneously rediscovered by several scientists.

The synthesis of Mendelian genetics with Darwin’s theory of natural selection was initially accomplished in the 1920s and 1930s through the theoretical work of several geneticists who used mathematical arguments to show, first, that continuous variation (in such characteristics as size, number of progeny, longevity, and the like) could be explained by Mendel’s laws and, second, that natural selection acting cumulatively on small variations could yield major evolutionary changes in form and function. Distinguished members of this group of theoretical geneticists were R. A. Fisher (4) and J. B. S. Haldane (5) in Great Britain and Sewall Wright (6) in the United States. Their work provided a theoretical framework for the integration of genetics into Darwin’s theory of natural selection but had a limited impact on contemporary biologists because (i) it was formulated in a mathematical language that most biologists could not understand, (ii) it was almost exclusively theoretical, with little empirical corroboration, and (iii) it was limited in scope, largely omitting many issues, like speciation, that were of great importance to evolutionists.

The synthesis accomplished by the theoreticians was greatly expanded in the following decades by biologists coming from various disciplines who enlarged the initial theoretical synthesis with relevant concepts and theories and provided supporting empirical evidence. Several books are considered emblematic of this original expansion of the theory in addition to Mayr’s Systematics and the Origin of Species (1), notably, Theodosius Dobzhansky’s Genetics and the Origin of Species, published in 1937 (7), George Gaylord Simpson’s Tempo and Mode in Evolution (8), and G. Ledyard Stebbins’ Variation and Evolution in Plants (9). Three earlier colloquia sponsored by the National Academy of Sciences were dedicated to current knowledge concerning the distinctive topics originally explored in these books (10–12).

One key development of the theory of evolution is the replacement of “population thinking” by “typological thinking.” Darwin had postulated that hereditary variations occur in organisms that are useful to the organisms themselves. Natural selection could only occur if such variations were pervasive. In genetics, populational thinking gave rise to a new branch of genetics that, as Dobzhansky (7) noted, “has as its province the processes taking place in groups of individuals—populations—and therefore is called the genetics of populations . . . The rules governing the genetic structure of a population are distinct from those governing the genetics of individuals.”

Mayr’s Systematics and the Origin of Species (1) represents a self-conscious effort to explicate the significance of population variation in the understanding of evolutionary processes and the origin of new species. “It is true that the change from the static species concept of Linnaeus to the dynamic species concept of the modern systematist has not entirely escaped the attention of progressive students of genetics and evolution. However, the whole significance of the polytypic species, of the phenomena of geographic variation, of the differences between geographic and other forms of isolation are by no means as widely appreciated . . . as they deserve” (1).

This paper serves as an introduction to the following papers, which result from the Arthur M. Sackler Colloquium of the National Academy of Sciences, “Systematics and the Origin of Species: On Ernst Mayr’s 100th Anniversary,” held December 16–18, 2004, at the Arnold and Mabel Beckman Center of the National Academies of Science and Engineering in Irvine, CA.

To whom correspondence should be addressed at: Department of Ecology and Evolutionary Biology, University of California, 321 Steinhaus Hall, Irvine, CA 92697-2525. E-mail: fjayala@uci.edu.

© 2005 by The National Academy of Sciences of the USA.
Mayr would later write: “Systematics, contrary to widespread misconceptions . . . was not at all in a backward and static condition during the first third of the 20th century. . . . Population thinking was widely adopted, and, as a consequence, variation within and between populations was actively studied, which led to the development of the biological species concept, to the widespread adoption of polytypic species taxa, and to the study of species in space and time as adapted systems . . . [T]he experimental geneticists, with few exceptions, were quite unaware that a populational species concept had been widely adopted by naturalists” (13). “The biological species concept emphasizes the species as a community of populations, reproductive isolation . . . and the ecological interactions of sympatric populations that do not belong to the same species” (13).

Species as taxonomic entities and, most of all, as populations and units of evolution have remained Mayr’s supreme subject of intellectual engagement. In his most recent book, Mayr writes: “The species is the principal unit of evolution. A sound understanding of the biological nature of species is fundamental to writing about evolution and indeed about almost any aspect of the philosophy of biology . . . I define biological species as ‘groups of interbreeding natural populations that are reproductively (genetically) isolated from other such groups.’ The emphasis of this definition is . . . on genetic relationship. This new interpretation of species of organisms emphasizes that biological species are something very different from the natural kinds of inanimate nature” (14).

Ernst Mayr was born on July 5, 1904, in Kempton, Bavaria, Germany. On July 1, 1926, he became an assistant at the University Museum in Berlin “but left for New Guinea and the Salomon Islands in February 1928. I did not return until the end of April 1930” (15). He came to the United States in 1931 to be curator for birds at the American Museum of Natural History in New York. In 1953, he became Alexander Agassiz Professor of Zoology at Harvard University, where from 1961 to 1970 he was director of the Museum of Comparative Zoology and retired from the faculty in 1974.

Mayr’s scholarly publications span >80 years, starting with his first two scientific papers, published in 1923, and reaching to 2004. Walter J. Bock, who has written a fairly comprehensive overview of Mayr’s career, divides his contributions into three major periods. “The first period (1923 until 1953 when he left the American Museum of Natural History) was devoted mainly to avian systematics and the theory of systematics. This work formed the foundation for the second period (beginning in 1942 but becoming more dominant in the latter part of the 1940s and lasting until his formal retirement from Harvard University in 1974), which was devoted largely to evolutionary theory. His systematic and evolutionary contributions, in turn, provided the basis for the last period (beginning in the early 1970s), devoted chiefly to the history and philosophy of biology” (16).


The colloquium honoring Mayr’s book and its legacy featured 17 presentations, including an essay by E. O. Wilson that appears immediately after this introduction (17). The essay is titled “Systematics and the Future of Biology,” and it makes the point that the tremendous growth in molecular and cellular biology will be hampered if it is not balanced by similar progress in our understanding of biological diversity. Wilson argues forcefully for growth in systematics and biodiversity research and for the establishment and increasing use of Internet-based virtual collections so that researchers and laypersons can freely access the resources of museums worldwide.

The remaining papers in the volume fall under four themes: The Origins of Species Barriers, Discerning Recent Divergence, The Nature of Species and the Meaning of “Species,” and Genomic Approaches and New Insights on Diversity.

The Origins of Species Barriers

Mayr was well known for his championship of the biological species concept and for asserting a predominant role for the geographic separation of populations in the diversification process that gives rise to separate species. The genetic version of this perspective is, to a good approximation, the Dobzhansky–Muller model of divergence, in which genes that have been the site of adaptive fixations within separate populations may also be the site of negative epistatic interactions in species hybrids and cause inviability or sterility when hybridization occurs (18–20). But what are these adaptations that accumulate within separate populations and give rise to reproductive barriers? This question is addressed by Allen Orr (21), and he explains what we know from the growing handful of cases in which the actual genes that contribute to low hybrid fitness have been isolated. The most striking commonality to emerge from these studies is that these genes have extraordinarily rapid rates of adaptive amino acid replacement. This finding makes sense, for if we suppose that some fraction of amino acid replacements are prone to negatively epistatic interactions when placed in a hybrid background, then those genes that have the highest rates of amino acid replacement will also tend to be those that cause these types of Dobzhansky–Muller incompatibilities.

This discovery of rapidly evolving genes that contribute to reproductive barriers also necessarily focuses our attention on the kinds of phenotypes and genes that are particularly prone to evolve rapidly. Classically, much of the discussion on rapid population divergence tended to focus on the kinds of environmental or geographic circumstances that might promote rapid evolution. However, in recent years, the attention has shifted to situations where intraspecies and intragenomic conflicts can lead to rapid evolution of genes (22). Such conflicts arise whenever natural selection favors alleles that are penetrant under some circumstances (such as in one sex), even though those same alleles may reduce other components of fitness that are manifest in other contexts (such as in the other sex). Genomic conflicts can lead to tit-for-tat, or arms-race, evolution between groups of genes within the same genome. William Rice et al. (23) explore this issue directly by developing a model evolutionary system with *Drosophila melanogaster*. In this system, individual haploid chromosome complements (hemiclones) are drawn, using genetic tricks that are possible with *Drosophila*, from a long-standing laboratory population. Once isolated, each hemicline can be measured, by replicating in combination with other hemiclones, for its net effect upon fitness of particular phenotypes. The approach allows a careful assessment of the selection gradient and additive genetic variance for traits that enhance fitness in males but reduce fitness in their female mates.

The findings that rapid evolution of genes, including that caused by genomic conflict, can lead to the formation of
reproductive barriers notwithstanding, there remains the ques-
tion of how much gene flow can be tolerated between diverging
populations if speciation is to occur (24). Not even populations
with many rapidly evolving genes can be expected to become
reproductively isolated from other populations if gene flow rates
are high. With this point in mind, Francisco Ayala and Mario
Coluzzi (25) explore models in which recombination suppressors,
such as chromosomal inversions, can enhance the oppor-
tunity for adaptive divergence in the face of gene flow between
parapatric populations (26–28). These models are more plau-
sible than those in which chromosomal inversions enable diver-
gence by causing low hybrid fitness, and they are supported
particularly by recent evidence from Drosophila and Anopheles.

Regardless of the rates and roles of genetic changes that
contribute to divergence, there remain very large questions
about the phenotypic manifestation of these changes. Mary Jane
West-Eberhard (29) addresses these questions and stresses that
phenotypes may derive from genes in ways that are highly
contingent upon other genes and upon environmental circum-
stances. Thus, single new alleles or new environmental circum-
stances alone, in the absence of genetic changes, may trigger
large changes in the phenotype. Furthermore, if genetic accom-
modation is common, such as by the Baldwin effect (30, 31), then
phenotypic variation that arises primarily by environmental
causes may play a driving role in divergence.

Discerning Recent Divergence

One of Mayr’s achievements that is not always counted is that he
made the difficult questions on evolutionary divergence seem
accessible. By laying out clear scenarios for a seemingly intrac-
table process in which evolutionary factors interact with the
georaphic circumstances of populations to cause divergence, he
fueled the interest and enthusiasm of generations of evolution-
ary biologists. The modern fruits of this enthusiasm are the many
studies that detail reconstruction of recent cases where evolution
has given rise to new species. The progress, and rapid pace of
progress, in this field is clearly shown in the paper by Scott
Edwards et al. (32) that outlines how present-day studies on
speciation in birds are gaining the genetic and theoretical
sophistication that had formerly only been associated with the
model Drosophila systems. Topics such as the role of sexual
selection and the frequency of sympatric speciation are now
being addressed genetically in a number of avian systems.

Among the most intricate of speciation puzzles are those
involving obligate mutualistic relationships. If one species of a
mutualistic assemblage diverges into two, must the other species
of the assemblage follow along? This is the question that arises
for figs and their pollinating wasps, and it is addressed by Carlos
Machado, Nancy Robbins, Tom Gilbert, and Allen Herre (33).
Each of the 750 or so species of fig depends upon fig wasps for
pollination; the wasps, in turn, require the ovaries of the figs as
oviposition sites (34). Strong reciprocal species specificity sug-
gests that when individual fig species or individual wasp species
undergo speciation, they do so in tandem with their mutualistic
partner (cospeciation). But Machado et al. ’s (33) phylogenetic
and population genetic study shows that the history has not been
straightforward and that host switches by wasps, and possibly
species hybridization by figs, have created partly independent
phylogenetic histories of figs and wasps.

In Mayr’s world view, new species arise under allopatry, and,
after that, as divergence accrues, the geographic ranges of
related species may later come to overlap. In this way, related
but divergent species may be sympatric, in contrast to most
closely related species, which are expected to have disjunct,
allopatric distributions. This sequence of events was outlined
explicitly by Mayr in a 1954 paper on the biogeography of sea
urchins (35). Stephen Palumbi and Harilaos Lesissios have
returned to this same Echinoid system and reconsidered
Mayr’s synthesis using DNA sequence data (36). They find that
although the pattern described by Mayr still largely applies,
rapidly evolving gamete recognition proteins play a strong role
in reproductive isolation. In contrast, Mayr had envisioned the
evolution of reproductive isolation by a more genome-wide
steady accumulation of substitutions.

For many biologists, the question of whether geographic
separation is strictly necessary for speciation (i.e., the question
of whether sympatric or parapatric speciation occurs) comes into
sharpest focus with the case of Rhagoletis pomonella. This is the
apple maggot fly that has diverged into two host races (apple
and hawthorne), apparently under geographic sympathy and aided by
the different fruited times of the two hosts (37). Mayr’s former
student Guy Bush discovered the history of sympatric divergence in
Rhagoletis, and it has long been a standard component of the
debates on the prevalence of sympatric speciation. Now we learn
from Guy Bush’s former student Jeffrey Feder and his col-
leagues (38) that the sympatric divergence that occurred within
U.S. populations may have been facilitated by genetic variation
that came in by means of gene flow from largely separated
populations in Mexico.

The question of sympatric speciation has also been much
discussed in the context of the highly speciose cichlid fishes from
the great African lakes: Victoria, Malawi, and Tanganyika (39).
Particularly in the cases of Lakes Malawi and Victoria, which are
relatively young, it is a wonder how hundreds of species could
form within confined bodies of water within <1 million years.
Yong-Jin Won, Arjun Sivasundar, Yong Wang, and Jody Hey
(40) take a close look at a group of rock-dwelling species from
Lake Malawi. To gain resolution, they used a new type of genetic
marker that includes a microsatellite and linked sequence and a
new Bayesian method for fitting complex models of divergence
(41). The results suggest that some of these species have formed
within the past few thousand years and that gene exchange is
ongoing between species at some loci.

The Nature of Species and the Meaning of “Species”

When Mayr outlined several species concepts, including the
biological species concept in his 1942 book, he started a new era
in species-problem debate. From that point on, Mayr was the
major figure in both the biological and philosophical compon-
ents of the debate (42, 43). In this volume, we have three papers
that address, from widely different perspectives, the very nature
of species. The many biologists who, like Mayr, take a primarily
zoological perspective, will appreciate the case studies presented
by Anne Yoder et al. (44) for vertebrate species complexes that
are endemic to Madagascar. The authors describe, with exam-
ples, a protocol that begins with field collections and existing
taxonomic resources and proceed to devise hypotheses of species
boundaries and priorities for additional collecting and exper-
imental work.

But, unlike animals, and indeed most eukaryotes, prokaryo-
tes have always presented special species-related challenges because
of the absence of regular gene exchange. Yet is it possible that
lateral gene transfer (LGT) that does occur among bacteria,
often across wide taxonomic chasms, can provide perspective on
the species question in bacteria? This topic is addressed by
Howard Ochman, Emmanuelle Lerat, and Vincent Daubin (45),
who take a whole-genome approach to ask about the historical
and phylogenetic distribution of LGT. They find that although
LGT generally can obscure older phylogenetic histories, the
subset of LGT that leads to homologous recombination is largely
limited to closely related bacteria. This finding supports a view
of bacterial species that resembles, in some respects, the biolog-
ical species concept.

On the matter of the multiplicity of species concepts, Kevin de
Queiroz (46) has contributed an article that directly targets one
of the main sources of confusion that arises in species concept

debate. That confusion lies between species criteria, as articulated in various species concepts, which are actually contingent properties of species, and the necessary properties of species as they are understood in the general sense of being evolutionary lineages. Biologists who disagree about which contingent properties of species are the most useful for identification and classification should be able to find common ground by recognizing the contingent, as opposed to necessary, aspect of the features they prefer to study.

**Genomic Approaches and New Insights on Diversity**

Because Mayr was not a geneticist, we do not count among his direct legacies our current era of genomics. But, in some respects, genomic studies of biological diversity are just the next step on a ladder that Mayr helped to hoist. Furthermore, it is fair to ask whether genomic tools are changing our view of biological diversity. One example of the way our view has changed is provided in the article by Ochman et al. (45) that is mentioned above. Another example lies in the paper by James Lake and colleagues (47), who use genomic data to reconstruct the process by which eukaryotes arose from prokaryotes. Unlike typical phylogenetic events, such as the splitting of lineages, eukaryotes appear to have arisen by the fusion of genomes. The authors describe the development and application of a new phylogenetic method, called “conditioned reconstruction,” which is designed to detect fusion events.

As the number of sequenced genomes grows, so will the number and availability of tools for identifying the genes responsible for important variation. This is a major point of the paper by Scott Edwards et al. (32) that was discussed above. Two other papers in this volume demonstrate some of the latest techniques for finding genes responsible for phenotypes of interest. Stuart Macdonald and Anthony Long (48) describe a new method for reducing the number of single nucleotide polymorphisms that are required in association mapping studies for genes that contribute to traits that have recently been under natural selection. The idea follows from the expectation that recent selection will have shaped divergence, and especially polymorphism, in and around the relevant sites. Using population genetic predictions of the response to selection of linked sites, it should be possible to conduct genomic scans of variation and divergence to identify the subset of polymorphic sites upon which to base an association mapping study. They demonstrate the method by looking at polymorphism within and divergence between *Drosophila* species.

Trudy Mackay et al. (49) also used a *Drosophila* model to identify genomic sites with interesting functions—mating behavior, in this case (49). Traditionally, genes that are directly involved in reproduction are not the easiest to study genetically, simply because mutants often have low reproductive success. These authors took the artificial selection approach and generated, over the course of 20 generations, two lines of *Drosophila melanogaster* that had high and low mean values for mating speed. They then conducted a microarray study to see which genes differed in expression level between the two divergently selected lines of flies.

The final paper in the volume takes an explicitly forward look and describes the ongoing and future changes that are happening to the biological sciences. With genomic sequences for many organisms having been available for several years, many biologists are turning to the highly integrated study of cellular processes and networks, a field that is called Systems Biology (50). Mónica Medina (51) writes about how this nascent field is being shaped by the availability of genome sequences throughout the tree of life and of the kinds of questions about the evolution of networks that we can anticipate. Surely, just as Systems Biology emerges and qualitatively new kinds of insights emerge about how cells function, so too will emerge the field of Evolutionary Systems Biology with concomitant insights on the evolution of cell function.