

From Phenomenology to First Principles: Toward a Theory of Diversity

Geerat J. Vermeij

*Department of Geology, University of California at Davis, One Shields Avenue,
Davis, CA 95616, Email: vermeij@geology.ucdavis.edu*

Diversity is a phenomenon that biologists typically associate with species, but in fact it characterizes all economic systems in which there is competition for locally limited resources. Patterns in species diversity — the equatorward increase in species number, the tendency for larger areas to support greater numbers of species, the association between high species richness and sexual selection, and a general increase in diversity through time, among others — are accounted for by two fundamental principles. First, the potential range of available phenotypes increases from conditions of low temperature, near the freezing point of water, to conditions where tropical climates prevail (30 to 40°C), because interactions ranging from the intermolecular to those among species speed up as temperature rises. Diversity also tends to increase from systems with low availability, predictability, and productivity of resources to those in which resources are plentiful, reliably available, yet still subject to local competition. Second, and more important, competition for locally limiting resources promotes division of labor and therefore functional differentiation and specialization. Levels of competition rise and evolutionary responses to competition proceed further in larger economic systems, where many phenotypic options are energetically available. Moreover, the number of resources for which economic entities compete also increases as economies expand and as rates of supply of those resources become increasingly controlled and enhanced by the entities comprising the system. In other words, competition creates positive feedbacks between potential and realized diversity. High diversity in an economic context thus arises when potential differences in performance — success in competition in its broadest sense — are large and when the stakes for survival and reproduction (the benefits of success versus the costs of failure) are high.

To many scientists, there is something seductively attractive about a concept that is at once abstract — devoid of the messy details of reality — and quantifiable. Such a concept is diversity, an expression of variety that has been widely embraced by ecologists, conservationists, paleontologists, sociologists, and even politicians. Perhaps because diversity is an abstract epiphenomenon, a more or less context-independent property of groups not possessed by any single member of that group, its scientific study has very largely been phenomenological, or descriptive. We describe how diversity — the number of species, adaptive types, languages, and so on — varies in space and time at scales ranging from the local community to the global biosphere. Mostly, we think of diversity as a snapshot of variety. We may compare diversity as it varies from place to place or through time, but inferences of process or cause from pattern have generally accomplished little more than to elevate phenomenology to a higher level.

I would point to two primary reasons for our collective failure to understand biological diversity at the species level in a theoretical framework that transcends descriptive phenomenology. First, whereas there is an overwhelming amount of information on the geography of diversity in terrestrial settings — that is, about diversity of creatures on the dry land — studies of diversity in such other environments as the sea, the soil, and the microbial world lag far behind. At the purely descriptive level, therefore, we are missing the richly comparative data that might constrain our musings about the “causes” of diversity as inferred from observed patterns among birds, land plants, mammals, and insects. Second, insufficient attention has been focused on first principles, the laws that describe how variety arises and diminishes. These rules — the economic and physical factors that promote diversity, the categories of calamity that decrease it — are mostly well known; the problem is that they reside in parts of science that are unfamiliar to most scientists who are most directly concerned with diversity. Moreover, the principles governing diversity explicitly introduce context, the very concept that many quantifiers of diversity sought to purge.

I have come to the subject of diversity as both a paleontologist and marine biologist with a taste for understanding the context that creates diversity and that is created by diversity. My aim in this essay is to catalogue briefly the major patterns of biological diversity as revealed by comparative studies in present-day and past oceans and lands, and to identify the principles that account for the phenomenology of diversity.

PATTERNS OF DIVERSITY

Patterns of diversity have been the subject of such a vast literature that I confine myself here to a statement of the main patterns. These are: (1) an increase in species numbers from the poles to the equator, from high to low altitudes on land, and from deep to shallow water in the sea; (2) an increase in diversity with area of habitat, where the number of species rises as area to the 0.2 to 0.3 power, meaning that islands and island-like environments have fewer species than continents or large bodies of water; (3) higher diversity in regions where geographical opportunities for genetic isolation are most numerous, that is, where many barriers exist and where these barriers change frequently in position and permeability; (4) higher diversity in clades (evolutionary units comprising an ancestor and all its descendants) characterized by internal fertilization than in those in which the union of male and female gametes takes place outside the parental bodies; and (e) a general increase in diversity, through time, punctuated briefly by extinction events.

Interested readers should consult Rosenzweig's (1995) comprehensive book for detailed documentation and discussion of these patterns. Here I wish to draw attention to a few potential exceptions, some problems of how to evaluate the relative contribution of factors that have been identified in empirical studies of diversity, and to redirect such studies to variables that more precisely identify the phenomena that cause or promote variety.

Perhaps the best known spatial pattern of diversity is the poles-to-equator increase in species number. Most clades of plants and animals on land and in the sea conform to this latitudinal gradient (for a recent comprehensive review, see Hillebrand 2004). So do human languages and cultures (Pagel and Mace 2004).

There are, however, exceptions. Small zooplankters such as foraminifers and copepods in the pelagic realm reach peak diversities not in the tropics, but at the middle latitudes (McGowan and Walker 1993; Rutherford et al. 1999). According to some interpretations (Rutherford et al. 1999), the water column is more finely and more predictably divided into depth zones at middle than at low latitudes, so that the zooplankton species can specialize to particular depths more in the temperate zones than in either the tropics or the polar regions. The number of zooplankton species is, however, vastly greater than the number of definable depth zones, much as the number of strata in

terrestrial vegetations (three to four layers in temperate and tropical forests; see A.R. Smith 1973; Terborgh 1985) and the number of zones in the intertidal belt do not begin to account for the patterns of diversity among trees or shore animals.

On rocky shores, the diversity of multicellular seaweeds and of the invertebrates that graze them likewise seems to reach a peak at middle northern and southern latitudes (for a review, see Bakus 1969). This pattern does not, however, apply to photosynthesizing animals — corals, gorgonians, some sponges, and even some bivalves — or to herbivorous fishes, which are most diverse in the tropics. The apparent exception of seaweeds and their consumers may disappear if we consider photosynthesizing animals and their consumers instead of seaweeds and their herbivores as the relevant group. My guess is that the diversity of attached multicellular marine primary producers and of their consumers will be found to conform to the typical equatorward increase in species as seen in most other clades and ecological groups.

Latitude is easy to measure, but it is unlikely to be the variable to which diversity or the organisms whose species number is being analyzed respond. Instead, latitude is a proxy for some other variable that is more directly linked to the lives of organisms and to the opportunities and constraints on species formation and loss. There is, for example, a general (if not wholly consistent) rise in temperature, incoming solar radiation, and net primary productivity as one moves from the poles to the equator, that is, with decreasing latitude. Incoming radiation amounts to about 60 W/m² (watts per meter squared) at 80°N, 120 to 125 W/m² at 60°N, and about 200 W/m² at the equator. Poleward of 40° north and south latitude, there is a net loss of radiation. These numbers will vary according to cloudiness and other factors, and may not accurately reflect the ability of organisms to absorb and to use incoming radiation. On land, productivity tends to increase with decreasing latitude, but again there is considerable variation in productivity, higher values occurring in areas of higher rainfall. For this reason, evapo-transpiration is a better proxy for energy in terrestrial ecosystems than incoming radiation (DeAngelis 1992), and is a better predictor of the number of plant and bird species (Wright 1983; Wright et al. 1993).

Temperature, another measure of energy availability, likewise is positively associated with high diversity in both terrestrial and aquatic environments (Vermeij 1978; Rosenzweig 1995; Fraser and Currie 1996; Guégan et al. 1998; Rohde 1999). Both diversity and temperature (mean average, extreme low, and extreme high temperatures) tend to rise toward the equator, but there appear to be important thresholds of temperature that affect diversity regardless of latitude. In the sea, the diversity of bivalves and other bottom-dwelling animals falls sharply south of the transition between the tropics and the temperate zones (Crame 2000, 2002). In Peru, this transition occurs near 5°S, whereas in Australia it is much further south, between 25 and 30°S. Hence, latitudinal gradients are often stepped instead of gradual, and their shape depends very much on where major thermal transitions occur (see also Vermeij 1996). The position of these thermal transitions, in turn, depends on the pattern of oceanic and atmospheric circulation. In Peru, the edge of the biological tropics is determined by the northward extent of the cold Peru Current. In the northwestern Atlantic, the transition zone between the subtropical Carolinian and the cool-temperate Virginian biogeographic regions occurs off North Carolina, and is determined by the interaction between the cold Labrador Current and the warm Gulf Stream. On land, frost eliminates large numbers of warm-adapted species, so the geography of winter frost determines the latitudinal pattern of diversity (Ricklefs and Latham 1993).

Temperature, energy availability, and evapo-transpiration all affect net primary productivity, the rate of production of biomass that is available for consumption by animals and decomposers. In the deep sea below a depth of 500 m, water temperatures are uniformly low (below 4°C), but there is still an equatorward increase in the diversity of bivalves and other bottom-dwelling animals

in the North Atlantic (Culver and Buzas 2000; Rex et al. 2000). In this cold, generally nutrient-poor environment, the best empirical predictor of diversity is net primary productivity in surface waters (Rex 1973, 1976; Rex et al. 1993, 2000). Low diversity at abyssal and greater depths compared to shallower zones on the ocean floor similarly reflect lower availability of usable food. A drop in species numbers toward higher altitudes (Rosenzweig 1995) similarly may be attributable to both lower temperature and lower productivity.

The comparative study of islands, pioneered by MacArthur and Wilson (1967), provides the basis for the generalization that habitat area — or, more generally, the size of ecosystems or communities in which species are embedded — is the most important single factor controlling diversity. Not only does diversity increase with area raised to the 0.2 to 0.3 power, but larger areas can support types of species that smaller areas cannot. In particular, species with low population density but high per-capita demands require large habitats for feeding and are, therefore, excluded from small regions.

Statistical analyses have persuaded many biogeographers that area explains more of the variation in species diversity than do other variables such as temperature, longitude, latitude, or productivity (see e.g., Rosenzweig 1995; Bellwood and Hughes 2001). It is crucial to realize, however, that the contribution a particular variable makes to the range of values of diversity depends entirely on the scale of the analysis. For example, Bellwood and Hughes's (2001) conclusion that area explains more of the regional variation in reef fish and coral diversity than do latitude and longitude arises from the fact that, although a large range of values of area was considered, the range of latitudes was restricted to the 60° of latitude in which coral reefs flourish. With a large range of latitude, or a smaller range of areas, the contribution that each variable makes to diversity would change, with latitude assuming a more prominent role. Estimates of how much variation in a dependent variable is explained by each of several independent "controlling" variables depend on the ranges of the controlling variables. Such estimates must, therefore, be treated with great interpretative caution and should not be taken too seriously. Furthermore, latitude and diversity may be relatively independent of each other in the immediate vicinity of the equator or within the Caribbean basin, but be strongly linked elsewhere. In other words, the nature of the relationship among variables may change from place to place and over time, making any estimates of the relative contribution of a given variable to an epiphenomenon like diversity suspect and unreliable.

Many tectonically active or topographically complex regions harbor large numbers of species, presumably because barriers have formed and shifted frequently. Terrestrial hotspots of diversity include the Himalayas and adjacent parts of China and India, and the Andean region of South America among many others. In the sea, places like the Philippines and the western Pacific generally come to mind. Unaccountably high species numbers of plants occur in southwestern Australia and the Cape region of southern Africa (reviewed in Rosenzweig 1995; Linder et al. 2003), areas that are neither equatorial nor productive nor, it would seem, tectonically active.

As has been noted by many previous authors, some clades are vastly more diverse than others arising at the same time. In general, plant and animal groups that practice internal fertilization (either directly or with the use of animal vectors) are far richer in species than those in which the union between male and female gametes takes place unsupervised away from the parents (see e.g., West-Eberhard 1983). A related pattern is that some geologically younger clades appear to be richer in species, and to be more concentrated in equatorial regions, than older clades (Crame 2000, 2002; Magallón and Sanderson 2001; Davies et al. 2004). Thus, in the sea, veneroidean bivalves are more diverse and more tropical in distribution than are protobranchs (Crame 2002). Host-specialized parasites are vastly more diverse than herbivores and predators with broader diets (Price 1980; Mitter et al. 1988).

Finally, diversity at the regional level seems to have increased over the course of geological time. The best evidence comes from studies of shallow-water communities on the seafloor (Bambach 1977), with large increases occurring during the Cambrian and Ordovician, and during the Cretaceous to Recent interval. Debate rages about whether this local or regional pattern also applies to global diversity. The issue is beset with daunting methodological and statistical problems. Given the difficulty of measuring global diversity in the modern biota, I am skeptical that an empirical pattern in global diversity is meaningful either statistically or in a biological context (Vermeij 1987; Vermeij and Leighton 2003). Nonetheless, there are reasons to expect that diversity in most environments has risen over time. I return to this matter toward the end of this essay.

BUILDING A THEORY OF DIVERSITY

These and many other examples are empirical generalizations of pattern. They are descriptions, phenomenological statements that invite explanation. Variables such as latitude, temperature, productivity, area, and clade membership seem empirically to be related to diversity, but none is perfectly correlated with species number, and the variables are neither independent of each other nor consistently linked. In order to arrive at a more fundamental understanding of diversity, we, therefore, need a theory that accounts for the patterns and that explains exceptions.

The task of creating such a theory requires an understanding of why diversity should arise in the first place, as well as knowledge of the factors that promote and depress it. To gain perspective on this matter, it is important to realize that diversity characterizes all complex systems. Although this essay is devoted mainly to the diversity of species, the notion of diversity can also be applied to human occupations, languages, linguistically distinguishable phonemes, cultures, sources of food, brand names, architectural styles, chemical compounds, cell types, and thousands of other manifestations of variety. Any theory of diversity must therefore not be so specific that it can deal with only one or two of these manifestations. The generality of the phenomenon of diversity suggests that some very fundamental, broadly applicable principles are at work in complex systems that govern not just the diversity of species, but other expressions of variety in the nonliving universe, in the organization of life, and in human affairs.

Elements of such a theory must at minimum include the following: (1) the division of labor, the result of the trade-off principle, which asserts that competition among entities for locally limiting resources promotes specialization because entities cannot perform all functions equally well; (2) the range of possible states, the universe of phenotypes that are potentially available for occupation by some entity in a system; (3) a mechanism for creating change, a necessity if diversity arises from uniformity; (4) opportunity, the set of conditions that enable entities to explore and occupy the universe of possibilities and to create new ones; and (5) constraint, the set of conditions that set limits to, or prevent occupation of, the universe of possibilities.

A prerequisite for diversity is that there exists a universe in which many states are not only possible, but realized. In biology, we might refer to states as phenotypes; in the nonliving world we think of elements, compounds, thermodynamic states, and the like. The universe of possibilities is typically vastly larger than the realized diversity, but at the same time it depends on what already exists. In a world of single-celled prokaryotes, for example, a warm-blooded multicellular mammal would be essentially inconceivable, because there is no mechanism to transform one prokaryote or even a collection of prokaryotes into a mammal in one step or even a few steps. Potential diversity — the range of possibilities — is, thus, dictated in part by the possibilities that have already been realized (Kauffman 2000). The higher the existing diversity, the greater is the universe of possibilities. Key questions are how potential states are generated, how these states come to be

occupied, and which factors limit both the universe of attainable states and the realization of the potential to occupy those states.

In the realm of life, where evolution — descent with modification — prevails as the mechanism of change, diversity arises ultimately as error. Variation arises as an alteration in the genetic code, as an error in replication or as insertion or deletion of code. Genetic error may be phenotypically silent, as it typically is if it involves the substitution of a nucleotide in the third position of the three-nucleotide code for an amino acid; or it may be expressed phenotypically, that is, when the genetic code is translated into physiology or morphology. Meaningful diversity arises only when initial error comes to be associated with a phenotypically meaningful, functional change, that is, with a change that affects the survival or reproduction of the entity bearing it. At the species level, diversity becomes meaningful only when the population in which a change arises becomes genetically isolated from the parent population and from potential sister populations. The creation of diversity is thus all about the emergence of meaning, of phenotypic function that is genetically isolated and distinct from functions in progenitors and contemporaries.

The generation of change and the emergence of meaningful differences depend both on genetic architecture and on the environment in which entities make a living. It is far beyond the scope of this essay to review these matters in detail, but a few points deserve emphasis. First, a genetic and developmental architecture in which several semi-independent modules are loosely linked is critical not just for the origin of mutations or other genetic alterations, but also for shielding such variants from selective removal for a time. Raw variation is, thus, generated and conserved long enough to be available for adaptation (Wagner and Altenberg 1996; Kirschner and Gerhart 1998; Newman and Müller 2001). In a functioning system of parts, most changes are deleterious. If the parts are highly integrated and interdependent, the change would affect the entire system and would therefore jeopardize that system, with the result that the alteration will be lost. If, on the other hand, the parts or modules comprising the system are loosely linked, and thus semiautonomous, a change with short-term, mildly deleterious effects can persist for a longer time, because its damaging effects are spatially limited. The variant may be incorporated in later generations as part of a new beneficial order. Most organisms, ecosystems, and societies at every level from cells to civilizations are organized as semiautonomous, interacting modules, which are variously specialized in function. This style of organization — itself the product of adaptive evolutionary processes — promotes the generation and conservation of variation. With a larger number of types of modules and with greater independence among them, the range of adaptive possibilities — the adaptive versatility of the system — increases (Vermeij 1973; Kirschner and Gerhart 1998; Raff 1996; Newman and Müller 2001; Galis et al. 2001; Galis and Metz 2003).

The environment in which variation arises has three fundamental roles in determining the potential and realized range of phenotypic space. First, it affects the potential range through temperature and the availability of resources. Second, it affects the likelihood of isolation of populations, a necessary condition for the generation of variation at the species level. Third, the realized range of possibilities is affected by competition between entities for locally limited resources. These three fundamental roles are intimately connected through feedbacks between living things and the environments to which life responds and which life creates (Vermeij 2003, 2004).

Almost every physical characteristic and chemical reaction important to living organisms depends on temperature. As temperatures rise from the freezing point of water to values between 30 and 40°C, rates of reaction increase, the activation energy for enzymatically catalyzed reaction decreases, the viscosity of water falls, and the solubility of gases and minerals changes. Accordingly, many functions become less expensive as temperatures rise (e.g., production of a calcareous skeleton, rapid locomotion in water, circulation or conduction of fluids). The range of pos-

sible metabolic rates, growth rates, locomotor speeds, and morphologies related to these rates is much greater when environmental or body temperatures are toward the high end of the thermal range of plant and animal life than when temperatures are 100°C or lower (Vermeij 1978, 2003). As a result, temperature regimes typical of the tropics or of warm-temperate summers are conducive to a much larger universe of phenotypes and lifestyles than are the thermal regimes characteristic of polar regions or the deep ocean.

Similarly, conditions in which material resources are plentiful and predictably available permit a wider range of adaptive possibilities than do conditions of chronic, unpredictable, or periodic scarcity (Clarke 1980, 1983). High metabolic rates and adaptive lifestyles simply cannot be maintained at times of scarcity, and are excluded from the universe of adaptive options for those entities that have no means of overcoming or provisioning against fluctuations in resource supply. Just as low temperatures prevent effective exploitation of abundant resources, so the absence of light for photosynthesizing plants or the unavailability of food for animals will severely constrain life processes even if the thermal environment is highly favorable. Moreover, newly isolated, small populations that would survive under a regime of plentiful resources might perish under conditions of scarcity, and therefore would be prevented from forming the nucleus of new daughter species (Allmon 1992, 2001; Allmon and Ross 2001; Vermeij 1995).

Higher temperatures and a more prolific and predictable resource supply create conditions favorable to the generation and retention of the variants that are essential as raw material for diversity. In human society, this kind of permissiveness is promoted by social tolerance of difference, the acceptance of freedom as a desirable state for individuals and groups, and — in order to realize the potential of individuals' abilities — a high predictability and prosperity of the economy that supports society. Conditions that allow for experimentation, expansion, and freedom without immediate sanction are, thus, essential for generating and nurturing variety (Vermeij 2002a, 2004).

Biologists studying speciation — the evolutionary process that generates diversity at the species level — have emphasized the isolation of populations from parent and contemporary populations as a necessary condition for the creation of new species. Isolation arises either when a previously single, genetically unified population is divided into mutually isolated fragments by some disruptive process, or when individuals from a parent population disperse to an environment not previously occupied by that population. Environmental conditions that increase spatial heterogeneity or promote intermittent dispersal, therefore, favor isolation.

But potential is not enough to explain diversity. There must be factors at work that do not merely permit difference, but that favor divergence. As noted by Ehrlich and Raven (1969), isolation may be a necessary condition for the creation of new, genetically and evolutionarily independent species; but it is not a sufficient one. In addition to isolation, the formation of species entails divergence, which, in turn, implies a regime of selection and adaptation that differs from the regime prevailing in populations with which the isolate could potentially exchange genes. Divergence may require genetic isolation, but it is not an inevitable consequence of isolation.

Competition for locally limiting resources would seem to be a critical agency for selection, and therefore is one factor — perhaps the most important factor — propelling populations to diverge. It provides the key that translates meaningless, often invisible variation into meaningful function. Making a living entails the acquisition and retention of resources. This activity of life takes place in settings where many entities are engaged in the same pursuit. Individuals, or groups of individuals, thus compete. Mechanisms of competition and the resources for which entities compete are themselves highly diverse, and of course are the products of variation and the selective processes acting on that variation. Cooperation, for example, is a highly effective means of competition; so are predation, parasitism, rapid incorporation of nutrients, preemption, aggression, passive defense,

and a hundred other evolved mechanisms. Evolutionarily highly derived plants and animals compete for resources such as mates or dispersal agents that for many other kinds of organisms simply have no meaning.

Because variants tend to differ in how effectively they gather resources or prevent other entities from taking them, a very common evolutionary response to competition is to specialize, either by minimizing competition with other entities or by beating others at their own game. Specialization is enforced by the trade-off principle, which asserts that improvement in one function engenders less effective performance in other functions. As a result, individuals may confine themselves to one or a few resources, or they may specialize to particular occupations, modes of defense, methods of resource acquisition, enemies on which defenses are concentrated, places to live, and so on. The greater the diversity of resources and of enemies, the greater are the scope and pressures of specialization. In short, competition-driven division of labor is the agency that promotes functional diversity; it is the agency that translates potential into realized variety. As Adam Smith (1776) pointed out long ago, greater wealth — that is, greater access to predictable and plentiful energy — makes possible more intense competition and is therefore associated with greater division of labor and with a larger number of occupations. As the thermally- and resource-based potential for the generation of phenotypes increases, so do the advantages of competition-driven division of labor and specialization.

Specialization is not simply an adaptive response to competition; it creates environmental heterogeneity where previously there was a perceived sameness. The specialization of an herbivorous insect species, to one particular host species, for example, means that one plant is not like another. If locally limiting resources provide conditions favorable to host-specific specialization, the environment for the specialist is patchier than that same environment would be for the specialist's more generalized ancestor. There is, in other words, a positive feedback between the generation of variety — through heterogeneity and isolation in this case — and the realization of potential variation, in this example through competition-driven divergence and specialization. In short, environmental heterogeneity and the potential for isolation are not just manifestations of the physical and chemical heterogeneity of the world; to an important degree, they are created by organisms themselves.

Competition creates feedback between realization and potential in still another way. The “exploration” of phenotypic space is not merely a passive process or a passive consequence of randomly generated variation; it reflects successful, adaptive, realized solutions to challenges that are ultimately imposed by competitors. The envelope of possibilities thereby enlarges, and this, in turn, provides the raw phenotypic material for still further charting of new phenotypic territory. For example, higher metabolic rates provide substantial competitive advantages to entities as long as resources are plentiful and predictable enough to sustain them. Rapid metabolism, in turn, permits physiologies and morphologies that are unavailable to organisms with low metabolic rates (Vermeij 2002a). These states would be left unexplored and unrealized in the absence of intense competition and the absence of the resource infrastructure necessarily for the evolution of high metabolic rates.

High diversity in an economic context thus arises when potential differences in performance are large and when the stakes for success — the difference between the benefits of success and the cost of failure — are high. In nature and in human society, competition and diversity appear to be highest when the environment permits the generation and the nurturing of variation (Vermeij 2002b, 2004).

It is important to emphasize that neither competition nor potential is by itself sufficient for diversity. Intense competition will reinforce the status quo and will purge novelty if conditions prevent entities from expanding in size or numbers (Vermeij 1995, 2002b). In order to promote diversity, strong selection exerted by competitors must be accompanied by an ability of entities to

respond to selection. Without that ability — that is, without environmental permissiveness — competition merely stultifies and imposes constraint. Similarly, in the absence of competition, much potential remains unrealized, and the potential universe of phenotypic options will remain largely empty.

Species emerge because members of species can recognize each other. The species is perpetuated through reproduction, which in many forms of life requires mating between two individuals whose genes come together and then recombine to form individuals of the next generation. Correct pairing — that is, pairing of individuals belonging to the same species — requires that individuals distinguish their kind from other kinds.

Species in which mating takes place outside the body of one of the members often have recognition expressed at the level of proteins in the gametes; species with internal fertilization typically have elaborate mechanisms by which adult individuals make the distinctions. With only a few species coexisting, methods of recognition need not be very elaborate; but as the number of co-occurring entities rises, and as population densities for individual species fall, the cost of making the wrong choice increases, and mechanisms ensuring correct pairing can become complex. This is notably the case with internally fertilizing groups. Diversity in these groups — many arthropods, molluscs, vertebrates, and flowering plants — is staggeringly high. As noted by West-Eberhard (1983), even small changes in the reproductive system or the recognition system can lead to genetic isolation of a daughter population from its parent or sister, and so set the stage for speciation if opportunities for the perpetuation of the population exist. The creation of more diversity as recognition systems become more elaborate is one of several mechanisms by which positive feedbacks allow diversity to increase.

All these factors conspire to give diversity a self-propagating quality. Every species is potentially a resource on which some other species can in principle specialize or to which another species must adapt. Moreover, species have the ability to regulate resources, enabling other species to specialize on that resource. Through adaptation, those species that either increase or stabilize the supply of resources create an environment in which the greater commonwealth permits and promotes higher competition-driven diversity (Leigh and Vermeij 2002). The positive feedback that species have on each other therefore amplifies and exaggerates the stimulatory effects that such factors as higher temperatures and greater productivity by themselves have on diversity.

It is the self-propagating property of diversity, together with a general rise in productivity and in competition, that yields the expectation that diversity has increased through time. The argument, as laid out in detail by Leigh and Vermeij (2002), is that species that create conditions favorable to other species as well as to themselves will more often succeed than selfish species. These species will have the effect of increasing the productivity of the ecosystem they inhabit. Increased productivity, in turn, “creates jobs” for other species. Moreover, as species collectively create a more highly regulated physical and chemical environment that is increasingly stable and resilient in the face of external disruptions, rates of extinction of clades tend to decrease through time (for a detailed discussion and documentation of decreasing rates of extinction through time, see MacLeod 2003). Clades with sophisticated recognition systems proliferate relative to clades in which competition for mates is less intense. Environments and regions in which competition and adaptation are least constrained produce the species with the highest competitive, defensive, and reproductive performance. From these centers of innovation and diversification, clades spread out, bringing to more outlying environments and regions increases in productivity and regulation that already existed in the centers. All of these factors conspire to favor increases in productivity and diversity. Occasional reversals at times of major extinction briefly interrupt, but do not fundamentally alter, this expected trend (Vermeij 1999, 2004).

A final factor contributing to the positive feedback between productivity and a rise in diversity through time is mobility, which enables rare species to persist. As Leigh (1999) has emphasized, the great majority of species in highly diverse rain forests are rare, being represented by widely scattered individuals. This is equally true of species on reefs. In Leigh's interpretation, rarity of many rain-forest tree species may be enforced by specialized enemies, whose effectiveness greatly decreases as the distance between nearest neighbors increases (Janzen 1970). In general, small populations of widely scattered individuals are susceptible to extinction because, in sexually reproducing species, individuals cannot easily find mates and so may be unable to reproduce. Rarity is made possible in part by a combination of high mobility and the capacity to identify and locate members of the same species at a distance. For flowering plants, which as adults do not move, mating is often facilitated by highly mobile pollinators. Without these animal helpers, flowering plants could not maintain populations of low density, because mating through wind- or water-pollination becomes highly inefficient as the distance between nearest neighbors increases (Raven 1977; Regal 1977). Animal mobility entails relatively high metabolic rates, which are sustainable only if the availability and reliability of primary production are sufficient to support substantial populations of consumers. As productivity generally rises through geological time thanks to intense competition and collective, ecosystem-level regulation of resources, an increased emphasis on mobility enables more and more species to coexist and to maintain small populations of widely scattered, sexually reproducing individuals. Altogether, then, there exist powerful positive feedbacks among productivity, enemies, mobility, and diversity.

The foregoing thoughts can hardly be considered a full-fledged theory of diversity, but I hope they will pave the way to a more formal proposal. Diversity is a phenomenon rich in patterns and complex in its causes and consequences. What we need now is a fundamental understanding of the phenomenon based on the laws of economics, organization, and evolution.

LITERATURE CITED

- ALLMON, W.D. 1992. A causal analysis of stages in allopatric speciation. *Oxford Surveys in Evolutionary Biology* 8:219–257.
- ALLMON, W.D. 2001. Nutrients, temperature, disturbance, and evolution: A model for the Late Cenozoic marine record of the western Atlantic. *Palaeogeography, Palaeoclimatology, Palaeoecology* 166:2–26.
- ALLMON, W.D., AND R.M. ROSS. 2001. Nutrients and evolution in the marine realm. Pages 105–148 in W.D. Allmon and D. Bottjer, eds., *Evolutionary Paleocology: The Ecological Context of Macroevolution*. Columbia University Press, New York, New York, USA.
- BAKUS, G.J. 1969. Energetics and feeding in shallow marine waters. *International Review of General and Experimental Zoology* 4:175–369.
- BAMBACH, R.K. 1977. Species richness in marine benthic habitats through the Phanerozoic. *Paleobiology* 3:152–167.
- BELLWOOD, D.R., AND T.P. HUGHES. 2001. Regional-scale assembly rules and biodiversity of coral reefs. *Science* 292:1532–1534.
- CLARKE, A. 1980. A reappraisal of the concept of metabolic cold adaptation in polar marine invertebrates. *Biological Journal of the Linnean Society* 14:77–92.
- CLARKE, A. 1983. Life in cold water: The physiological ecology of polar marine ectotherms. *Oceanography and Marine Biology Annual Review* 21:341–453.
- CRAME, J.A. 2000. Evolution of taxonomic diversity gradients in the marine realm: Evidence from the composition of Recent bivalve faunas. *Paleobiology* 26:187–214.
- CRAME, J.A. 2002. Evolution of taxonomic diversity gradients in the marine realm: A comparison of Late Jurassic and Recent bivalve faunas. *Paleobiology* 28:184–207.
- CULVER, S.J., AND M.A. BUZAS. 2000. Global latitudinal species diversity gradient in deep-sea benthic Foraminifera. *Deep-Sea Research I* 47:259–275.

- DAVIES, T.J., T.G. BARRACLOUGH, M.W. CHASE, P.S. SOLTIS, D.E. SOLTIS, AND V. SAVOLAINEN. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences USA* 101:1904–1909.
- DEANGELIS, D.L. 1992. *Dynamics of Nutrient Cycling and Food Webs*. Chapman and Hall, London, England, UK. 270 pp.
- EHRlich, P.H., AND P.H. RAVEN. Differentiation of populations. *Science* 165:1228–1231.
- FRASER, R.H., AND D.J. CURRIE. 1996. The species richness-energy hypothesis in a system where historical factors are thought to prevail: Coral reefs. *American Naturalist* 148:138–159.
- GALIS, F., AND J.A.J. METZ. 2003. Anti-cancer selection as a source of developmental and evolutionary constraints. *BioEssays* 25:1035–1038.
- GALIS, F., J.J.M. VAN ALPHEN, AND J.A.J. METZ. 2001. Why five fingers? Evolutionary constraints on digit numbers. *Trends in Ecology and Evolution* 16:637–646.
- GUÉGAN, S-F., S. LEK, AND T. OBERDORFF. 1998. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature* 391:382–384.
- HILLEBRAND, H. 2004. On the generality of the latitudinal diversity gradient. *American Naturalist* 163:192–211.
- JANZEN, D.H. 1970. Herbivores and the number of tree species in tropical forests. *American Naturalist* 104:501–528.
- KAUFFMAN, S.A. 2000. *Investigations*. Oxford University Press, Oxford, England, UK. 287 pp.
- KIRSCHNER, M., AND J. GERHART. 1998. Evolvability. *Proceedings of the National Academy of Sciences USA* 95:8420–8427.
- LEIGH, E.G., JR. 1999. *Tropical Forest Ecology: A View from Barro Colorado Island*. Oxford University Press, New York, New York. 245 pp.
- LEIGH, E.G., JR., AND G.J. VERMEIJ. 2002. Does natural selection organize ecosystems for the maintenance of high productivity and diversity? *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 357:709–718.
- LINDER, H.P., P. ELDENÄS, AND B.G. BRIGGS. 2003. Contrasting patterns of radiation in African and Australian Restionaceae. *Evolution* 57:2688–2702.
- MACARTHUR, R.H., AND E.O. WILSON. 1967. *The Theory of Island Biogeography*. Princeton University Press, Princeton, New Jersey, USA. 203 pp.
- MACLEOD, N. 2003. The causes of Phanerozoic extinctions. Pages 253–277 in L.J. Rothschild and A. Lister, eds., *Evolution on Planet Earth: The Impact of the Physical Environment*. Academic Press, Amsterdam, The Netherlands.
- MCGOWAN, J.A., AND P.A. WALKER. 1993. Pelagic diversity patterns. Pages 203–214 in R.E. Ricklefs and D. Schluter, eds., *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. University of Chicago Press, Chicago, Illinois, USA.
- MAGALLÓN, S., AND M.J. SANDERSON. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762–1780.
- MITTER, C., B. FARRELL, AND B. WIEGMANN. 1988. The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification? *American Naturalist* 132:107–128.
- NEWMAN, S.A., AND G.B. MÜLLER. 2001. Epigenetic mechanisms for character origination. Pages 559–579 in G.P. Wagner, ed., *The Character Concept in Evolutionary Biology*. Academic Press, San Diego, California.
- PAGEL, M., AND R. MACE. 2004. The cultural wealth of nations. *Nature* 428:275–278.
- PRICE, P.W. 1980. *Evolutionary Biology of Parasites*. Princeton University Press, Princeton, New Jersey, USA. 237 pp.
- RAFF, R.A. 1996. *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. University of Chicago Press, Chicago, Illinois, USA. 520 pp.
- RAVEN, P.H. 1977. A suggestion concerning the Cretaceous rise to dominance of the angiosperms. *Evolution* 31:451–452.
- REGAL, P.J. 1977. Ecology and evolution of flowering plant dominance. *Science* 196:622–629.
- REX, M.A. 1973. Deep-sea species diversity: Decreased gastropod diversity at abyssal depths. *Science* 181:1051–1053.

- REX, M.A. 1976. Biological accommodation in the deep-sea benthos: Comparative evidence on the importance of predation and productivity. *Deep Sea Research* 23:975–986.
- REX, M.A., C.T. STUART, AND G. COYNE. 2000. Latitudinal gradients of species richness in the deep-sea benthos of the North Atlantic. *Proceedings of the National Academy of Sciences USA* 97:4082–4085.
- REX, M.A., C.T. STUART, R.R. HESSLER, J.A. ALLEN, H.L. SANDERS, AND G.D.F. WILSON. 1993. Global-scale latitudinal patterns of species diversity in the deep-sea benthos. *Nature* 365:636–639.
- RICKLEFS, R.E., AND R.E. LATHAM. 1993. Global patterns of diversity in mangrove floras. Pages 215–229 in R.E. Ricklefs and D. Schluter, eds., *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. University of Chicago Press, Chicago, Illinois, USA.
- ROHDE, K. 1999. Latitudinal gradients in species diversity and Rapoport's Rule revisited: A review of recent work and what can parasites teach us about the causes of the gradients? *Ecography* 22:593–613.
- ROSENZWEIG, M.L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge, England, UK. 436 pp.
- RUTHERFORD, S., S. D'HONDT, AND W. PRELL. 1999. Environmental controls on the geographic distribution of zooplankton diversity. *Nature* 400:749–753.
- SMITH, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations* (1971 edition, with introduction by E. Cannon and notes by M. Lerner). Random House, New York, New York. 976 pp.
- SMITH, A.R. 1973. Stratification of temperate and tropical forests. *American Naturalist* 107:671–683.
- TERBORGH, J. 1985. The vertical component of plant species diversity in temperate and tropical forests. *American Naturalist* 126:760–776.
- VERMEIJ, G.J. 1973. Adaptation, versatility, and evolution. *Systematic Zoology* 22:466–477.
- VERMEIJ, G.J. 1978. *Biogeography and Adaptation: Patterns of Marine Life*. Harvard University Press, Cambridge, Massachusetts, USA. 332 pp.
- VERMEIJ, G.J. 1987. *Evolution and Escalation: An Ecological History of Life*. Princeton University Press, Princeton, New Jersey, USA. 527 pp.
- VERMEIJ, G.J. 1995. Economics, volcanoes, and Phanerozoic revolutions. *Paleobiology* 21:125–152.
- VERMEIJ, G.J. 1996. Marine biological diversity: Muricid gastropods as a case study. Pages 355–375 in D. Jablonski, D.H. Erwin, and J.H. Lipps, eds., *Evolutionary Paleobiology: In Honor of James W. Valentine*. University of Chicago Press, Chicago, Illinois, USA.
- VERMEIJ, G.J. 1999. Inequality and the directionality of history. *American Naturalist* 153:243–253.
- VERMEIJ, G.J. 2002a. The geography of evolutionary opportunity: hypothesis and two cases in gastropods. *Integrative and Comparative Biology* 42:935–940.
- VERMEIJ, G.J. 2002b. Characters in context: molluscan shells and the forces that mold them. *Paleobiology* 28:41–54.
- VERMEIJ, G.J. 2003. Temperature, tectonics, and evolution. Pages 209–232 in L.J. Rothschild and A. Lister, eds., *Evolution on Planet Earth: The Impact of the Physical Environment*. Academic Press, Amsterdam, The Netherlands.
- VERMEIJ, G.J. 2004. *Nature: An Economic History*. Princeton University Press, Princeton, New Jersey, USA.
- VERMEIJ, G.J., AND L.R. LEIGHTON. 2003. Does global diversity mean anything? *Paleobiology* 29:3–7.
- WAGNER, G.P., AND L. ALTENBERG. 1996. Complex adaptations and the evolution of evolvability. *Evolution* 50:967–976.
- WEST-EBERHARD, M.J. 1983. Sexual selection, social competition, and speciation. *Quarterly Review of Biology* 58:155–183.
- WRIGHT, D.J. 1983. Species-energy theory: An extension of species-area theory. *Oikos* 41:496–506.
- WRIGHT, D.J., D.J. CURRIE, AND B.A. MAURER. 1993. Energy supply and patterns of species richness on local and regional scales. Pages 66–74 in R.E. Ricklefs and D. Schluter, eds., *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. University of Chicago Press, Chicago, Illinois, USA.