*Paleobiology,* 30(2), 2004, pp. 179–202

# Macroevolution, hierarchy theory, and the C-value enigma

T. Ryan Gregory

*Abstract.*—For more than 60 years, evolutionary biologists have debated the issue of whether the processes of genetic change observable within populations (microevolution) can provide an adequate explanation for the large-scale patterns in the history of life (macroevolution). In general, population geneticists have argued in favor of microevolutionary extrapolation, whereas paleontologists have sought to establish an autonomous and hierarchical macroevolutionary theory based on the operation of selection at several levels of biological organization (especially species). The massive variation in eukaryotic genome sizes (haploid nuclear DNA contents, or ''C-values'') has similarly been a subject of debate for more than half a century, and it has become clear that no one-dimensional explanation can account for it. In this article, the basic concepts of macroevolutionary theory are reviewed and then applied to the long-standing puzzle of genome size variation (the ''C-value enigma''). Genome size evolution provides a clear example of hierarchy in action and therefore lends support to the theoretical approach of macroevolutionists. Perhaps more importantly, it is apparent that genome evolution cannot be understood without such a hierarchical approach, thereby providing an intriguing conceptual link between the most reductionistic and expansive subjects of evolutionary study.

*T. Ryan Gregory. Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at Seventy-ninth Street, New York, New York 10024. E-mail: rgregory@genomesize.com*

Accepted: 28 October 2003

*Punctuated equilibrium is but one pathway to the elaboration of hierarchy, and probably not the best or most persuasive; that role will probably fall to our new understanding of the genome and the need for gene-level selection embodied in such ideas as "selfish DNA."*

Stephen Jay Gould, 1992

*Although genes are made of DNA, much DNA is not genes, and it is not clear that we can so easily understand all of the structures and evolutionary behaviors of DNA without some further theoretical expansion.*

W. Ford Doolittle, 1989

## Introduction

In many ways, genomics and paleontology represent opposite ends of the professional spectrum in evolutionary biology. Some important recent interaction aside, the two fields have not generally been linked in any explicit way, and on occasion they have even become embroiled in an acrimonious contest of ''molecules versus morphology.'' But as the quotations given above suggest, there is agreement among prominent authors from the two disciplines in at least one area—the implications of a modern understanding of genome structure for evolutionary theory in the broadest sense.

This article borrows macroevolutionary concepts derived from paleontology and applies them to one of the longest-standing puzzles in genome biology, namely the evolution of genome size. In the process, some background information is provided regarding the case for a hierarchical theory of evolution (largely for the benefit of readers of the neontologist persuasion), followed by a brief outline of some of the forms that this theory has taken when discussing selection at the group and species levels. These approaches are then applied to the question of genome size, following an introduction to the basic concepts of the puzzle. Not only should paleontology and genomics be considered mutually enlightening from a conceptual standpoint, but, I argue here, the evolution of the genome cannot be fully understood without such integration.

## The Case for Hierarchy

The extent to which mathematically tractable processes observable within populations can be extrapolated to explain patterns of diversification in deep time has long been one of the most contentious issues in evolutionary bi-

 0094-8373/04/3002-0003/$1.00

ology. For obvious reasons, the protagonists in this debate have typically been divided along professional lines, particularly among population geneticists who measure and model small genetic changes within populations, and paleontologists whose purview consists of documenting the evolutionary patterns produced over geological timescales. Members of this latter group have generally resisted the implication that all evolutionary change is a product of small-scale intrapopulational processes, and they have spent the past 30 years developing and promoting a theory of large-scale evolutionary change distinct from the extrapolationist approach of population genetics.

Several arguments have been advanced against extrapolationism and in favor of a distinct macroevolutionary theory. On an empirical basis, it has been pointed out many times that there is no good evidence that natural selection operating within populations can precipitate species-level diversifications. Both laboratory and field observations of natural selection have been criticized as being too extreme to be applicable to evolution at large (e.g., Gould 2000; Lewontin 2000). In what he calls the ''paradox of the visibly irrelevant,'' Gould (2000: p. 343) notes that such observable changes ''are *vastly too rapid* to represent the general modes of change that build life's history through geological ages.'' From a philosophical perspective, it has been argued that even if macroevolution really is microevolution writ large, then it must still be studied at its own scale. As physicist Steven Weinberg (2001) pointed out with regard to a different set of scientific disciplines,

> Almost any physicist would say that chemistry is explained by quantum mechanics and the simple properties of electrons and atomic nuclei. But chemical phenomena will never be entirely explained in this way, and so chemistry persists as a separate discipline. Chemists do not call themselves physicists; they have different journals and different skills from physicists. It's difficult to deal with complicated molecules by the methods of quantum mechanics, but still we know that physics explains why chemicals are the way they are.

Similarly, microevolutionary explanations may not provide a sufficient account of macroevolution, even if such reductionism were justifiable in principle. Of course, most macroevolutionists argue the stronger case that distinct macroevolutionary processes operate that cannot be so reduced. These usually relate to the operation of natural selection at levels higher than the organism, especially ''groups'' and species. A brief review of the principles developed for these levels of biological organization is presented below before they are applied to the genome.

### Group Selection: New and Improved and No Longer Naïve

In 1962, V. C. Wynne-Edwards had proposed that individual birds limit their clutch sizes in times of strife to the intended benefit of the population as a whole. This proposition was challenged on empirical grounds by Lack (1966), and most forcefully from a theoretical perspective by Williams (1966). Williams's (1966) emphasis on individual reproduction as immensely more powerful than competition among groups was buttressed by the development of concepts such as kin selection (Hamilton 1964) and reciprocal altruism (Trivers 1971), which seemed to render unnecessary any claims for group-level adaptations in the explanation of apparently self-sacrificing behaviors. Most crucially, it was recognized that groups of altruistic organisms would be subject to infiltration by ''selfish'' individuals, which would then swamp out the altruists by virtue of their ill-gotten reproductive superiority.

According to Wilson (2000), ''Williams and others of his time were reacting to a form of bloated groupism that deserved to be rejected,'' but it now seems that when the bathwater of naïve group selectionism was purged, the baby went along for the ride. More-sophisticated models of group-level selection have since been generated in theory and tested in practice, and multilevel selection has been strongly implicated in the evolution of such features as reduced virulence in pathogens

(Bull 1994; Frank 1996; Sober and Wilson 1998), multicellularity (Michod 1997; Michod and Roze 2000), insect eusociality (Seeley 1997), female-biased sex ratios (Frank 1986), and human social behavior (Wilson and Sober 1994; Sober and Wilson 1998), to name a few. Notably, Darwin (1871) himself had invoked intergroup selection as an explanation for human sociality, and some of these examples have been accepted by even the staunchest early opponents of naïve group selectionism (e.g., Hamilton 1975; Williams 1992).

A particularly pertinent model dealing with the evolution of altruistic behaviors has been presented by Sober and Wilson (1998). In this case, the crucial feature of the model that distinguishes it from naïve formulations subject to invasion by cheaters is that the groups *combine* at some stage to form a metagroup. This allows the differential input of altruists to affect the overall population characteristics despite decreasing in relative abundance within each group. Specifically, groups containing more altruistic members will have a higher net production than groups dominated by selfish organisms, even though the selfish organisms out-reproduce the altruists within any given group. Thus, ''the altruists increase globally, despite decreasing in frequency within each group, because the two groups contribute different numbers of individuals to the global population'' (Sober and Wilson 1998: p. 23).

Like organism-level (microevolutionary) natural selection, group selection has been shown to be remarkably effective under experimental conditions (for reviews, see Goodnight and Stevens 1997; Sober and Wilson 1998). However, although these models are theoretically sound and supported by experimental evidence, the overall role of group selection under natural evolutionary conditions has yet to be determined. As outlined below, one of the most significant inputs of group selection to macroevolution may actually have come from its early operation at the subgenomic level.

### Selection, Sorting, and Species

*Species and Biological Individuality.*—In what he viewed as a ''radical solution to the species problem,'' Ghiselin (1974) proposed that species are not mere arbitrary snapshots of a constantly changing morphological continuum, as Darwin and other strict anageneticists had believed, but rather ''individuals'' in their own right, with properties akin to those of organisms. The link between this conceptual shift and the ''levels of selection'' debate was made a short time later by Hull (1976, 1980), who pointed out that, as individuals, species could potentially undergo their own process of higher-level natural selection.

Several authors have proposed criteria that must be met by any potential evolutionary ''individual'' (e.g., Mishler and Brandon 1987; Gould 1995, 2002; Baum 1998; Gould and Lloyd 1999). According to Gould (2002: p. 602), any true biological individual must possess three general features: ''[1] a discrete and definable beginning, or birth; [2] an equally discrete and definable ending, or death; and [3] sufficient stability (defined as coherence of substance and constancy of form) during its lifetime to merit continuous recognition as the same 'thing'.'' The application of these criteria to organisms clearly establishes them as individuals (and indeed, organisms remain the primary exemplars of individuality). In the most familiar example of all, humans are organisms that are born, remain discrete and recognizable during their lifetimes, and eventually die. However, even at the clearest level of human organisms, there must be some flexibility in definitions, given that conception and birth are separated by a nine-month gestation period, and that substantial change accrues during both ontogeny and senescence. Other types of organisms may present an even more substantial challenge to definitions of individuality (e.g., Santelices 1999; Heddi et al. 2001).

When applied to the notion of macroevolutionary theory, which is inherently hierarchical in its emphasis on evolutionary processes operating at several levels of biological organization, a second feature of individuals becomes important. This is the role that individuals play both as collectivities of lower-level units and as parts of higher-level conglomerations (Gould 1998, 2002). In the classic example, organisms are made up of cells at the level below and make up populations at the

level above. Moreover, this triad of part-individual-collectivity shifts along with the level under consideration. Moving up one level, we see that if populations are considered as individuals, then organisms become parts and species become collectivities. At an even higher level, it is not difficult to view species as composed of populations and constituting clades, and therefore falling in the middle position as individuals. It has, however, been considerably more problematic to establish the individuality of species on the basis of the more specific criteria outlined above.

Taking a strict anagenetic view of evolutionary change, in which the imperceptible transformation of one species into another is the dominant model, there can be no classification of species as individuals. Species defined in these anagenetic terms have neither discrete temporal boundaries (i.e., a clear beginning and end) nor a sufficient level of constancy during their evolutionary tenures. For this reason, it should come as little surprise that proponents of a hierarchical approach to macroevolution do not consider anagenesis to be the dominant mode of species transformation.

*Punctuated Equilibria and Species as Individuals.*—The punctuated view of macroevolutionary change began with Mayr's (1954) discussion of speciation in island birds of the New Guinea region, in which he argued that ''rapidly evolving peripherally isolated populations may be the place of origin of many evolutionary novelties. Their isolation and comparatively small size may explain phenomena of rapid evolution and lack of documentation in the fossil record, hitherto puzzling to the palaeontologist'' (see Mayr 1963, 1992 for additional discussion). Eldredge's (1971) study of the Devonian trilobite *Phacops rana* provided the first explicit application of Mayr's allopatric speciation model to paleontological data, and of course Eldredge and Gould (1972) formally introduced the theory of punctuated equilibria in the following year. The major contributions of this approach have been to emphasize the long-term stasis of species through evolutionary time and the formation of new species by branching, with significant temporal overlap between parent and daughter species rather than the gradual anagenetic transformation of one species into another.

Several causes of stasis have been proposed and discussed, including stabilizing selection (Eldredge 1971), habitat tracking (Eldredge 1995), and organismal integration and plasticity (e.g., Wake et al. 1983; Seaborg 1999), among others. The important implication of stasis, whatever its cause, is that it grants spaciotemporal boundaries to species that anagenetic theories do not. That is, punctuated equilibria provides species with a real ''birth'' (geologically rapid speciation), ''life span'' (long-term existence in stasis, with only moderate fluctuation), and ''death'' (extinction) (e.g., Gould 1982, 2002). According to Eldredge (1985), the status of individual afforded to species by such a process is, in itself, justification for the development of a hierarchical approach to evolution.

*Species Selection in Principle and in Practice.*—If species can be characterized as legitimate biological individuals, then they possess at least some of the properties necessary for direct participation in the process of natural selection. In addition to birth, cohesion, and death, Darwinian individuals (that is, potential loci of selection) must also display the crucial features of a capacity to reproduce, a hereditary parent-offspring similarity, and variation among their fellow individuals in fitness-related traits (Lewontin 1970; Gould and Lloyd 1999). Again, under punctuated equilibria, species not only are born, remain static, and die, but they also produce offspring species by branching (versus transforming into new species themselves). Of course, the descendant species are more similar to their parental stocks than to unrelated species, which also provides the component of heredity. The final question in regard to whether a process of ''species selection'' (Stanley 1975, 1979) can operate is whether there are features of species that can make some of them more likely to produce offspring species, or (less importantly) more resistant to extinction, than others.

Two main arguments have been raised against species selection as an important evolutionary force. The first is that, although plausible in principle, species selection would be overshadowed by the more powerful action of

selection operating on the more numerous and faster-reproducing organisms of which species are composed (e.g., Fisher 1958; see Gould 1998, 2000 for additional discussion). The second major criticism of species selection as a causal process is that, even if one accepts the real-life operation of the process, the features that may contribute to increased ''reproduction'' or ''longevity'' of species are actually possessed by the organisms of which species are composed, and not of the species themselves. This is clearly of particular relevance to arguments for species selection that deal with species longevity instead of the differential production of daughter species, because extinction (species death) can almost always be viewed as the cumulative deaths of all component organisms (though mass extinctions may provide an exception). The main claim in this critical argument is that species have no ''emergent properties'' that necessitate an interpretation of species selection as anything other than microevolution extrapolated to a geological timescale. Two approaches have been taken to counteract these criticisms, one focusing on the identification of irreducible species-level characters, and the other emphasizing the necessity of a hierarchical view even in the absence of such features. The distinction is of relevance to the case of genome size evolution as well, and so will be treated in some detail.

*Aggregate versus Emergent Characters.*—Granting, for the sake of argument, that species are true individuals, we may ask what types of features might contribute to their differential survival and reproduction and allow selection to operate at this level. Any such traits that can be reduced to the summation of organismal properties, known as *aggregate characters,* may not qualify as contributing to true species selection (generation time, which can influence a lineage's evolutionary rate, provides a good example of an organismal feature with consequences at the species level). Instead, honest-to-goodness species selection, sensu stricto, has often been seen (by both supporters and detractors) as requiring the existence of *emergent characters* of species (e.g., Vrba 1983, 1989; Gilinsky 1986; Vrba and Gould 1986; Lieberman and Vrba 1995). Thus, in the first defense of species selection outlined above, there has been an effort to identify characteristics of species (or at least populations) that could serve as targets of selection at the higher level without being reduced to the properties of organisms taken in sum. Some likely candidates for relevant and distinctly higher-level traits include ''specific mate recognition systems'' (Paterson 1985; Lieberman 1992), sex ratios (Colwell 1981; Wilson and Colwell 1981; Wilson and Sober 1994), population size (Vrba and Eldredge 1984), and geographic distribution (Jablonski 1987).

Although each of the items on this list may be accepted as a characteristic of the population or species as a whole, there is still the issue of whether or not these have contributed to the differential success of real species during the course of evolution. In practice, this has proven a very difficult problem for species selectionists to resolve. The most commonly discussed trend potentially caused by species selection is that involving fossil (specifically, Tertiary and Cretaceous) gastropod molluscs with differing larval types (e.g., Hansen 1980; Arnold and Fristup 1982; Jablonski and Lutz 1983; Gilinsky 1986; Jablonski 1987; Lloyd and Gould 1993; Grantham 1995; Gould 2002). The trend in question is the increased representation of species with nonplanktonic larvae over those with a planktonic larval type, despite the higher susceptibility to extinction of the former. The suggested reason for this seemingly paradoxical increase in nonplanktonic forms involves their higher speciation rates relative to planktonic forms (which have motile larvae, more gene flow, and less opportunity to speciate), whereby more nonplanktonic species are produced despite their generally shorter evolutionary life spans.

However, even this literal textbook example of species selection (e.g., Ridley 1993; Futuyma 1998) may not provide the conclusive demonstration that macroevolutionists may wish it to. In a molecular phylogenetic study of extant turritellids (screw shells), Lieberman et al. (1993) showed that a nonplanktonic larval type evolved several times in different clades, with little or no reversal to a planktonic type. They conclude on this basis that species selec-

tion on larval type has not been the dominant process affecting diversification in these animals, and instead favor an interpretation based on developmental constraints. Likewise, a phylogenetic analysis of modern cone shells by Duda and Palumbi (1999) revealed repeated parallel shifts in larval type among species, suggesting that daughter species do not always share the larval type of their ancestors as assumed under the species selection model. As Duda and Palumbi (1999) put it, ''such results challenge the conclusion that increases in the number of nonplanktonic species relative to species with planktonic larvae over geologic time is necessarily a result of higher rates of speciation of nonplanktonic lineages and show that demonstration of species selection requires a phylogenetic framework.''

*Emergent Fitness versus the Effect Hypothesis (the Lloyd-Vrba Debate).*—Logically, three types of processes can affect fitness as examined at the level of species. These involve (1) emergent species-level traits, (2) irreducible effects of aggregate traits, and (3) effects of aggregate traits that can be completely reduced to organism-level processes (Grantham 1995). The first and third processes are uncontroversial, because examples of the first type of process would necessarily be considered species selection, whereas clearly the third would not. The second of these processes, on the other hand, has been a subject of debate among macroevolutionists.

According to Vrba (1980, 1983, 1984, 1989), cases in which nonrandom species diversification is caused by aggregate characters do not constitute species selection. Under her formulation, the properties of organisms may have legitimate effects on the success of species, but this process of upward causation must be distinguished from species selection as it does not involve emergent features at the species level. Instead, this would be considered an example of what Vrba (1980) dubbed the ''effect hypothesis.'' Differential species success caused by variation in the properties of their component organisms must still be analyzed from a macroevolutionary perspective in this view, but the causation is ultimately reducible to lower-level processes.

Other macroevolutionist authors have argued that this emphasis on emergent characters is too restrictive and needlessly enjoins a futile search for adaptations at the species level (e.g., Lloyd and Gould 1993; Grantham 1995; Gould 2002). In particular, Lloyd and Gould (1993) proposed that the focus should not be on emergent characters of species, but on *emergent fitness.* Thus, ''the emergent fitness approach requires only that a trait have a specified relation to fitness in order to support the claim that a selection process is occurring at that level'' (Lloyd and Gould 1993). In other words, even aggregate characters—if they contribute to the irreducible fitness of the species—can generate a legitimate process of species selection. Put more explicitly,

> In the ''emergent fitness'' approach, we do not inquire into the history of species-level traits that interact with the environment to secure differential proliferation. We do not ask where the traits originated in a structural or temporal sense—that is, whether such traits arose by emergence at the species level, or as aggregate features by summation of properties in component organisms or demes. We only require that these traits characterize the species and influence its differential rate of proliferation in interaction with the environment. In other words, we only demand that aspects of the fitness of the species be emergent and irreducible to the fitnesses of component organisms. (Gould 2002: pp. 659–660)

On the basis of this difference in interpretation, the argument over irreducible aggregate characters has been dubbed the ''Lloyd-Vrba debate'' (Grantham 1995). The primary point of disagreement in this case is on what should be considered true species selection, and what counts merely as ''effect macroevolution.''

Even under the more flexible ''emergent fitness'' approach, Vrba's strict definition based only on emergent characters is recognized as providing a ''best case'' of species selection (Gould 2002). Similarly, an antagonistic interaction among levels of selection provides the clearest opportunity to demonstrate higher-level processes in action. As Vrba (1989) ar-

gued, ''the acid test of a higher-level selection process is whether it can in principle oppose selection at the next lower level.'' For this reason, the existence of female-biased sex ratios (which are favored at the group level but opposed at the organism level) provides a good demonstration of higher-level selection in action (Wilson and Colwell 1981). It has also been pointed out that stasis at one level (e.g., species) may often reflect a dynamic tension between opposing pressures from above and below (Gould 1998, 2002). However, antagonism need not characterize the interplay among levels of selection, and in some (perhaps most?) cases, evolutionary processes operating simultaneously on several levels may do so synergistically (or orthogonally, with no direct interaction of processes operating at different levels) (e.g., Vrba 1989; Gould 1998, 2002). As macroevolutionists point out, there is a tendency to notice multilevel interactions only when antagonism dominates, and to otherwise eschew hierarchical explanations in the name of parsimony. But the maxim that if a selective outcome can be explained in terms of a low-level process then it should be done so automatically and exclusively (Williams 1966) represents a philosophical preference, and not necessarily an objective truth.

*Selection versus Sorting*.—The primary issue in the Lloyd-Vrba debate is one of identifying *patterns* versus *processes.* Under the more narrow ''effect hypothesis,'' differential species success not related to emergent characters must be interpreted as no more than a *pattern* produced by upward causation from organism-level interactions. The pattern in this case is macroevolutionary and must be understood as such, but its causal basis is not separate from microevolution. However, under Lloyd's ''emergent fitness'' view, not only the higher-level patterns, but also the *processes* that generate them, are distinctly macroevolutionary.

In traditional microevolutionary interpretations of evolution, the distinction between pattern and process may not be problematic, because natural selection among organisms is considered the dominant force that shapes macroevolutionary patterns. However, when a hierarchical approach is adopted, the separation of the two concepts becomes crucial, because the causal processes in question need not operate on the same level as the observed pattern. To clarify this issue, Vrba and Gould (1986) argue that the differential evolutionary success of entities at any level should be considered an expression of *sorting,* of which selection is one (but not the only) cause. As Lieberman and Vrba (1995) point out (under the restricted definition of the effect hypothesis): ''Sorting is the pattern of differential survival and/or reproduction of entities. It occurs at levels including genes, cells, organisms, groups, and species. In contrast, selection is the interaction between heritable, varying, emergent characters of individuals and the environment that causes differences in birth and/or death rates of those individuals. Selection is one of the many processes that can produce a pattern of sorting.''

When considering the issue of aggregate characters with implications for emergent fitness, the question is one of bottom-up sorting. That is, features of organisms (the lower level) affect the sorting of species (the higher level). Vrba considers this to be an example of ''effect sorting,'' because it lacks the input of emergent characters. Lloyd, on the other hand, would characterize this as a legitimate example of species selection because the fitness of the species is not reducible to the sum fitnesses of its component organisms. In this case, there is a process of top-down sorting, whereby selection among species influences the proportions of different types of organisms. As Eldredge (1985) has noted, ''sorting of lower-level individuals, in general, has a far less profound effect on upper-level individuals than the converse,'' making top-down effects potentially more important in discussions of macroevolutionary theory.

This is not to say that Vrba's interpretation of macroevolutionary effects cannot include the all-important component of top-down sorting. Indeed, Vrba (1989) has developed a useful model of ''context dependent sorting'' that allows such a top-down process to operate. In this case, constraints from above may dictate the limits of selection at levels below. By way of analogy, Vrba (1989) describes the downward control of sorting among citizens residing in different political states: ''To the

extent that a national ruler or law dictates that members of the population with certain characteristics may have more children than others, or must die at different ages, sorting among humans depends on whether they live in that nation or in another more liberal one.'' In reference to this analogy, Vrba (1989) has labeled this top-down process as ''Mustapha Mond sorting'' (à la *Brave New World*). When applied to real-life scenarios, it is apparent that the

> existence of an organism within a species (or context dependence) implies a pattern different from selection strictly at the organismal level without the existence of groups, because patterns at a lower level cannot be smoothly extrapolated to a higher level. . . Thus, even if species selection does not operate, the existence of biological units of organization above the level of the individual organism can have evolutionary effects. (Lieberman and Vrba 1995)

To summarize, true species selection has often been seen as necessitating the existence of emergent characters exclusive to the species level and irreducible to lower-level properties. For the most part, these must relate to an increased capacity to speciate, and not to an improved resistance to extinction. Aggregate characters that do not contribute to the differential success of species are not considered relevant to macroevolutionary theory. In between these extremes are aggregate characters that exert significant effects at the species level, though they themselves remain reducible to the summed properties of the constituent organisms. Whether viewed in the context of Vrba's ''effect hypothesis'' (no true species selection without emergent characters) or Lloyd's ''emergent fitness'' approach (species selection so long as the aggregate character exerts an effect on species fitness), it is necessary in principle to consider processes operating at the species level. Species selection has been difficult to establish in practice, however, and it is therefore prudent to seek examples at alternative levels in the biological hierarchy.

### The C-value Enigma

In 1948, Vendrely and Vendrely reported ''a remarkable constancy in the nuclear DNA content of all the cells in all the individuals within a given animal species'' [my translation], which they took as evidence that DNA rather than proteins served as the hereditary material. Watson and Crick's (1953) elucidation of the structure of DNA a few years later settled this debate once and for all, but interest remained regarding bulk amounts of DNA in different species (known as ''C-values,'' in reference to the haploid ''class'' of DNA [Swift 1950]). Despite the apparent constancy of most genome sizes, it was noted early on that DNA content bore no relationship to intuitive notions of organismal complexity (e.g., Mirsky and Ris 1951). In other words, DNA was thought to be constant because it is the stuff of genes, but expected gene number is unrelated to DNA amount. By the early 1970s, this (apparent) contradiction had become known as the ''C-value paradox'' (Thomas 1971).

It is undoubtedly true that genome size and gene number are decoupled. Genome sizes vary more than 200,000-fold among eukaryotes, with both the largest and smallest found among protists. Even among animals the range in genome size is greater than 3000-fold (Fig. 1). Nevertheless, this is no longer the least bit paradoxical: most eukaryotic DNA is

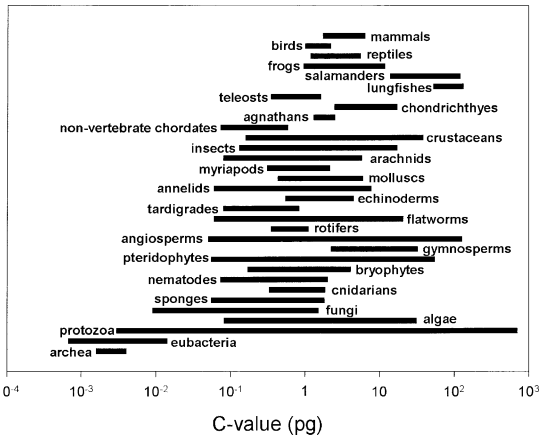


FIGURE 1. The ranges in haploid genome sizes (''C-values'' in picograms) so far observed in different groups of organisms, showing that genome size is clearly unrelated to intuitive notions of organismal complexity. Based on Raff and Kauffman 1983, using the most recent C-value data available (from Biderre et al. 1995; Renzaglia et al. 1995; Li 1997; Baumann et al. 1998; Bennett et al. 2000b; Voglmayr 2000; Bennett and Leitch 2001a,b; Gregory 2001c; Leitch et al. 2001; T. R. Gregory unpublished data).

noncoding, so a large genome does not imply a large number of genes (and, as it turns out, neither does high organismal complexity [International Human Genome Sequencing Consortium 2001]). However, the existence of vast quantities of noncoding DNA raises several questions of its own, specifically in relation to its evolutionary origin and maintenance, potential effects on the organismal phenotype, and pronounced quantitative variation among species. Together, these make up the complex puzzle known as the ''C-value enigma'' (Gregory 2001a).

As with many complex questions in evolutionary biology, there has been a counterproductive tendency to seek a unitary solution to the C-value enigma. Thus, numerous general functions have been proposed for noncoding DNA, including gene regulation, providing a storehouse of potential genetic sequences, buffering against mutations, promoting mutations, and numerous others. Taking an opposite tack, other authors have characterized noncoding DNA as entirely functionless ''junk'' that simply accumulates until it finally begins to impose costs on the host cell. In one sense, this search for a single solution can be blamed on the persistent use of the term ''C-value paradox,'' which clearly frames the problem of genome size evolution in a simple one-dimensional context. The ''paradox,'' properly defined, was solved long ago with the discovery of noncoding DNA, but if one thing has been learned in 50 years of genome size study, it is that no such easy answer exists for the broader C-value enigma.

Because it is composed of several distinct subquestions, the C-value enigma will necessarily require a pluralistic outlook capable of incorporating various explanatory approaches. Moreover, the evolution of genome size is a macroevolutionary question, so it is important to consider how macroevolutionary theory may shed light on this still-unresolved puzzle. However, even before the macroevolutionary theory developed for application to species is translated to the genome, some interesting arguments can be presented in favor of taking a hierarchical approach to genome evolution.

## The Necessity of Hierarchy Theory for Understanding Genomes (and Vice Versa)

*Group Selection and the Origin of the Genome.*- The evolutionary origin of organized genomes, like that of cellular life itself, remains an unresolved puzzle in biology. In fact, these two issues are probably part of a single larger question. As Maynard Smith and Szathmáry (1995: p. 114) note, ''During the evolution of protocells, unlinked genes must have given rise to linked ones. It is thus essential to discern the selective force as well as the molecular mechanisms enabling this transition to take place.''

As part of their explanation for the origin of genomes organized into chromosomes, Maynard Smith and Szathmáry (1993, 1995) point out that the primary problem is one of overcoming the disadvantage that the linkage of genes would have generated for the early genomic constituents:

> Imagine a simple protocell with two essential genes. Under what conditions does a primitive chromosome, with the two genes linked, increase in frequency in the population? First, one has to take into account that chromosomes will have a competitive disadvantage within the cell, since it takes longer to replicate them than unlinked genes (Maynard Smith and Szathmáry 1995: p. 114)

To explain how such a disadvantage might be overcome, Maynard Smith and Szathmáry (1995: p. 114) present the following scenario:

> Let *A* and *B* be two complementary genes [i.e., both are needed by the protocell], and *AB* the chromosome. Is an *A* gene better off on its own, or as part of a chromosome? On its own, it replicates faster within a cell. However, an *A* gene that is part of a chromosome is certain to find itself, in the next generation, in a cell that also contains a *B* gene: that is, it is certain to be in a fit cell, whereas an isolated *A* gene may find itself in an unfit cell, with no *B* gene. A second advantage of linkage is the synchronization of replication, and hence the elimination of within-cell competition between different genes.

This gene's-eye view is certainly informative, but it is not the whole story. The protocells in this scenario are also dividing, and their relative success in the next generation is influenced by the behavior of the genes contained within them. Maynard Smith (2002) continues with the gene-based approach in this context as well by arguing that ''selection will favor an AB 'chromosome' over independently replicating A and B genes, even if the independent genes replicate more rapidly than the chromosome within a cell, provided that a cell containing both A and B genes grows substantially faster than one lacking one or the other gene.''

However, there is another way to conceive of this problem. That is, it is possible to view each protocell as a compartment housing members of a metapopulation of genes, some of which are autonomous and fast replicating whereas others are linked in a cooperative protochromosome and therefore divide more slowly. Unlinked genes hold the within-protocell advantage but may hinder the ability of the cell to divide by failing to cooperate (or even competing) with the other necessary genes. Linked genes may have lower fitness within a protocell, but the cells containing them would boast a higher rate of reproductive success than cells containing more egoistic genes. If, by their cooperative behavior, the linked genes contribute more to the metapopulation of genes (i.e., the entire population of protocells) than do the unlinked genes, then organized chromosomes would evolve despite the replicative disadvantage that this cooperation entails. In this case, it is a matter not only of preventing two necessary genes from being separated, but also of promoting the cooperation of genes to the benefit of the protocell in which they reside.

Obviously, this revised scenario is essentially a genome-level application of the model of group-level selection for altruism proposed by Sober and Wilson (1998) and described above. To put it bluntly, the origin of integrated genomes, and therefore of cellular life itself, may be owed to the operation of hierarchical selection during the earliest stages of evolution on the Earth.

*Selfish DNA and Its Hierarchical Implications.*– Not all DNA sequences are as cooperative as those in the hypothetical scenario outlined above for the origin of genomes. At the genic level, the best known example of uncooperative sequences are segregation distorters that lever themselves into more than their share of descendant cells. The first noncoding sequences identified as being self-centered were B chromosomes, which coexist alongside the usual A chromosomes in many groups (see Camacho et al. 2000 for review). As early as 1945, Östergren argued that ''reasonable support may be given to the view that in many cases these chromosomes have no useful function at all to the species carrying them, but that they often lead an exclusively parasitic existence.'' More common examples of ''parasitic'' or ''selfish'' DNA include the various types of transposable elements (TEs) found within all eukaryotic genomes (for recent reviews, see e.g., Hurst and Werren 2001, Kidwell and Lisch 2001, and Feschotte et al. 2002).

For more than 60 years (but especially over the past 20), the existence of selfish DNA has been considered, at least in general terms, as a demonstration of hierarchical selection in action (Östergren 1945; Orgel and Crick 1980; Sapienza and Doolittle 1981; Gould 1983; Vrba and Eldredge 1984; Doolittle 1987, 1989). As Doolittle (1989) put it, ''much of the data of modern molecular biology might better be understood as revealing the operation of selection at several levels in a real biological hierarchy, and that failure to recognize this may lead to nonsensical statements about the functions of genomic structures.'' In some cases, selfish DNA has been considered the best example of hierarchy of all, trumping any hypothetical considerations of group or species selection (e.g., Gould 1992).

From the outset, selfish DNA clearly provided strong logical evidence for the principle of hierarchical selection, but logical consistency does not, in itself, prove actual importance in evolution. As Doolittle (1987) pointed out, ''many logically possible evolutionary processes do not actually often occur, and it was important to the selfish DNA argument (not in terms of the logic but in terms of the biology) to show that some real DNAs actually are

selfish.'' Similarly, Gould (1983: p. 176) asked: ''How much repetitive DNA is self-centered DNA? If the answer is 'way less than one percent' because conventional selection on bodies almost always overwhelms selection among genes, then self-centered DNA is one more good and plausible idea scorned by nature. If the answer is 'lots of it,' then we need a fully articulated hierarchical theory of evolution.'' Transposable elements, the prime example of selfish DNA, are now known to make up about half of the human genome (International Human Genome Sequencing Consortium 2001), indicating that selfish DNA is indeed extraordinarily common in nature. And yet, these elements are not amplified *ad infinitum,* which suggests that constraints at the genomic and/or cellular level act to prevent uncontrolled spread. As such, the very existence, in limited quantities, of selfish elements provides compelling evidence for multilevel selection (Gould 1983, 1992).

Advancements over the past two decades in the understanding of genome-level dynamics have only increased the scope (and necessity) of this hierarchical perspective on transposable elements. These elements do not operate in a void, but rather live and reproduce in a genomic ecosystem along with large numbers of other transposable elements (both ''conspecifics'' and unrelated sequences). Like the members of any ecosystem, cohabiting transposable elements may interact in a variety of ways. Thus, transposable elements may, in some cases, be required to compete with one another for ''resources'' (e.g., preferred insertion sites, materials for replication). At other times, they may contribute to each other's spread, as with the generation of processed pseudogenes by long interspersed nuclear elements (LINEs) (Brosius 1999; Esnault et al. 2000). At the other extreme, some elements may be seen to parasitize their fellow parasites, a prime example of which is the hitchhiking of short interspersed nuclear elements (SINEs) on LINEs (Zeyl and Bell 1996; Luning Prak and Kazazian 2000).

Moving from the analogy of ecosystems to that of hosts, another series of important interactions becomes apparent. As with all symbionts and their hosts, there is a dynamic coevolutionary interaction between TEs and the genomes in which they reside. Under the simplified version of the selfish DNA theory, this interaction is strictly one of parasitism, although there may of course be selection for reduced virulence (e.g., Sapienza and Doolittle 1981; Doolittle et al. 1984; Charlesworth and Langley 1986). The modern perspective on transposable element evolution is considerably more flexible: ''Rather than labeling TE-host associations as either selfish or parasitic, we prefer the idea of a continuum, ranging from aggressive parasitism at one extreme, through a neutral middle ground, to mutualism at the other extreme'' (Kidwell and Lisch 2001). SINEs, for example, are considered by many authors to be at least partly functional in host genomes (e.g., Makalowski 1995; Schmid and Rubin 1995).

In a more specific example, transposable elements of various types are now known to have been co-opted to serve regulatory functions, such that they now play a vital role in the host genome (Britten 1996, 1997; Brosius 1999; Kidwell and Lisch 2001). (It is interesting to note in this regard that transposable elements were initially dubbed ''controlling elements'' by Barbara McClintock in the late 1940s.) In terms of gene structure, transposable elements are capable of broad-scale mutator activity during both insertion and excision events (Kidwell and Lisch 1997, 2002). Most importantly from a macroevolutionary perspective, this can include the generation of regulatory mutations of major developmental effect (Finnegan 1989; McDonald 1990, 1995). Various genomic mechanisms are known for the suppression of selfish DNA activity, but in many cases this is not necessary, as some TEs preferentially insert into noncoding heterochromatic regions (e.g., Dimitri and Junakovic 1999; Hutchison et al. 1999). In short, even if they begin as strictly selfish, transposable elements may experience a variety of evolutionary fates, ranging from active suppression by the host to exaptation and integration into the regulatory mechanisms of the genome itself. Which of the available outcomes befalls any given TE will be dependent on its impacts at each of several levels of organization and the

net result of the hierarchical interactions it engenders.

At the level above genes and other DNA sequences, transposable elements exert important effects on the chromosomes. In this case, the influence proceeds in both directions. For example, transposable elements may incite chromosomal rearrangements by virtue of their transpositional activity (Gray 2000; Kidwell and Lisch 2000, 2001) and have also recently been implicated in the process of double-strand break repair (Labrador and Corces 1997; Eickbush 2002; Morrish et al. 2002). In a more general sense, transposable elements may contribute to the very construction of chromosomes, as in the case of telomeres in *Drosophila*, which are maintained directly by the serial insertion of transposable elements rather than by the usual activity of telomerase (Levis et al. 1993; Pardue et al. 1997; Casacuberta and Pardue 2002). More broadly, it seems that telomerase itself may have evolved from transposable elements (Eickbush 1997; Pardue et al. 1997). Of course, the reverse scenario is also possible, namely that the transposable elements in question evolved from the telomerases, but currently available evidence puts the origins of the transposable elements as the earlier event (Malik et al. 1999; Smit 1999). Centromeres, too, may be the byproducts of transposable element activity, and it is notable in this regard that newly formed centromeres may act as segregation distorters—that is, they may not yet have abandoned their selfish leanings (Kidwell and Lisch 2001; Kidwell 2002). If transposable elements exert an influence on gene structure and expression, then these effects will be played out at the levels of the cellular and organismal phenotypes. However, there are additional points of interaction besides these obvious genic effects, and once again these operate bidirectionally between the different levels of organization. In terms of upward causation, transposable elements may have ''accidental'' deleterious effects by causing deletions in coding regions that lead to various genetic diseases (Labuda et al. 1995; Kidwell and Lisch 2001).

From the perspective of the C-value enigma, it is more interesting to consider the influence of downward causation on transposable element evolution. In the most obvious example of such a process, organism-level (or perhaps one may say population-level) traits such as sexual versus asexual reproduction may directly affect the ability of selfish elements to spread and be maintained (e.g., Zeyl and Bell 1996; Burt and Trivers 1998; Wright and Schoen 1999; Arkhipova and Meselson 2000; Schön and Martens 2000). Other organismal happenings, such as the invasion of new habitats, can also have downward effects on transposable element activity (e.g., Biémont et al. 2001; Vieira et al. 2002).

Not only does selfish DNA provide the most persuasive demonstration of hierarchy in action, but it is also best understood by using the hierarchical principles developed for use with other levels in the hierarchy. As discussed in more detail below, the differential sorting of transposable elements within genomes is subject to the same principles as that of organisms within species. This is particularly true when considering the impacts of transposable elements and other noncoding sequences on the cellular and organismal phenotypes.

### The Genome, the Cell, and the Organism

*Genome Size and Cell Size.*—According to Cavalier-Smith (1982), ''the most reliably established fact about genome evolution is that C-values are generally positively correlated with cell and nuclear volumes.'' This observation is most clearly evident in the best-studied case of vertebrate erythrocytes, but it has been shown to apply to various other cell types and taxa as well (for reviews, see Gregory 2001a,b). Although the exact mechanistic basis for this relationship has not yet been elucidated, diverse circumstantial evidence strongly suggests a causal (''nucleotypic'') influence of bulk DNA content on cell division rate and cell and nucleus sizes (reviewed in Gregory 2001a). Under one recent model, a large amount of DNA causally generates a larger nucleus and a slower cell division rate, which in turn produce a larger cell (Gregory 2001a). Whatever the explanation, it is apparent that genome size is tied in an important way to variation in cell size and division rates, and therefore to any organismal parameters affected by these.

*Genome Size and Body Size.*—A correlation between genome size and body size is most likely to be observed in organisms with reduced body sizes (and therefore less variation in cell number), and especially those with determinate growth and cell number constancy. Such a relationship has been reported in various plants and many invertebrate groups, including aphids, copepod crustaceans, flatworms, and others (e.g., Finston et al. 1995; Gregory et al. 2000). Nematodes and copepods are perhaps the best-known examples of animals with determinate growth, and in the former, body size appears to be particularly influenced not by genome size per se, but by the level of polyploidy reached in the somatic cells (Flemming et al. 2000). Cell numbers are highly variable in most vertebrates and body size is therefore not greatly affected by differences in cell sizes. Nevertheless, cell and genome sizes do correlate positively with body size in birds and rodents (Gregory 2002a,b).

*Genome Size and Metabolism.*—It is a well-known principle in physics that while an object's volume increases as the cube of its radius, its area increases only as the square. The result is that larger objects have lower surface-area-to-volume ratios than smaller ones. In biological terms, this is of relevance to processes such as heat loss at the organismal level, and gas exchange at the cellular level. Larger cells have relatively lower surface areas for gas exchange and are therefore metabolically less active than small cells.

Erythrocyte size and genome size are both inversely correlated with mass-corrected metabolic rate in mammals and birds (Vinogradov 1995; Gregory 2002a), suggesting that this surface-area-to-volume-ratio effect on gas and ion exchange translates directly to the organism level in endotherms. Because both groups require high metabolic rates, it should come as no surprise that mammals and birds have the smallest red blood cells among vertebrates (e.g., Hawkey et al. 1991). More tellingly, birds and bats—the only vertebrates capable of powered flight—appear to have particularly constrained genome sizes (Burton et al. 1989; Gregory 2002a), and flightless birds tend to have larger genomes than strong flyers (Hughes 1999). In mammals, erythrocyte reduction has been taken to an extreme by the ejection of nuclei from mature cells. In other words, mature mammalian erythrocytes contain no genomes, a fact that has probably allowed their genome sizes to expand beyond those of other amniotes while maintaining the smallest erythrocytes (Cavalier-Smith 1978; Gregory 2000). In birds, genome size reduction seems to be the means by which small cells have been achieved (Gregory 2002a).

A lack of metabolic data has made this association more difficult to assess in most ectothermic vertebrates, and little is known of any possible relationship between genome size and metabolic rate in fishes and reptiles. Amphibians have far more variable cell and genome sizes than endotherms, but it is clear that genome size is of little or no significance in considerations of metabolic rate in this group (Licht and Lowcock 1991; Gregory 2003). On the surface this may be somewhat surprising, especially because ''frugal'' metabolic rates have long been invoked to explain the corpulent genomes of aquatic salamanders (e.g., Szarski 1976, 1983; Cavalier-Smith 1985, 1991), but it is evident that red blood cells (and therefore genome sizes) play, at best, only a minor role in determining metabolic parameters in these animals (Gregory 2003).

Clearly, the implications of cell and genome size vary considerably according to the biology of the group under consideration. In endotherms, metabolic rate constraints may have been very important in the evolution of genome size, whereas in amphibians this is of little relevance. However, cell size is not the only parameter influenced by genome size. Cell division rate is also affected by DNA content, and indeed a delay in cell division may generate the cell size correlation and explain the persistent correlation in mammals despite erythrocyte enucleation (Gregory 2001a). In this regard, it is interesting to consider the potential effects of a division rate correlation on the organismal phenotype.

*Genome Size and Development.*—Because growth in multicells proceeds primarily by cell division, the most obvious expression of the relationship between genome size and cell division rate will be with growth rate. Negative correlations between genome size and

growth rate have indeed been identified in various plants, amphibians, crustaceans, and other taxa (reviewed in Gregory 2002c). Interestingly, no such relationship can be found between genome size and any measures of developmental rate in either mammals or birds (Gregory 2002b), in direct contrast to the situation with metabolic rate. In plants, genome size is linked to developmental lifestyle, with annual plants necessarily having small genomes whereas perennials may possess very large C-values (Bennett 1987). In many cases, plant species may shift from perennial to annual lifestyle as they encounter harsh new environments, and this is found to occur concomitantly with a reduction in genome size (e.g., Watanabe et al. 1999).

More recently, it has been pointed out that developmental rate is only one side of the coin in considerations of genome size evolution. That is, correlations with developmental rate are visible only when developmental complexity—the amount of developing to be done in a given amount of time—is held constant (Gregory 2002c). When the time available for differentiation is limited, developmental complexity becomes the potentially visible correlate of cell division rate and genome size. The best example of a time-limited period of intensive morphological differentiation is metamorphosis, and it appears that the presence/absence and intensity of metamorphosis is a major determinant of genome size distributions in both amphibians and insects. Thus, the smallest amphibian genomes are found in frogs that inhabit ephemeral pools and must complete their metamorphosis very quickly. Direct-developing and normal biphasic frogs have slightly larger genomes, followed by biphasic, then direct-developing, and finally facultatively and then obligately neotenic salamanders (Gregory 2002c). In insects, members of those orders with holometabolous development (complete metamorphosis involving distinct larval, pupal, and adult stages) invariably possess genomes below a threshold that is easily breached by species in hemimetabolous and ametabolous orders, which do not undergo complete metamorphosis (Gregory 2002c).

## Selection, Sorting, and Genome Size

As a short review of the arguments presented earlier with regard to species selection, it should be recalled that under Vrba's ''effect hypothesis,'' patterns generated at the species level by aggregate characters are not considered to represent true species selection, although the lower level nevertheless has important bottom-up effects that must be examined at the higher level. Under Lloyd's ''emergent fitness'' approach, on the other hand, it is irrelevant whether the characters related to species fitness are emergent or aggregate, so long as the fitness effect on the species cannot be reduced to the summed fitnesses of individual organisms. In either case, it is important to distinguish between the observed patterns (sorting) and the process(es) responsible for generating them (such as selection). Sorting can proceed in either a bottom-up (e.g., effect sorting) or top-down (e.g., Mustapha Mond sorting) fashion, even in the absence of true higher-level selection in the strictest sense, and top-down sorting can be expected to have a greater influence on the fate of individuals at the lower level than does bottom-up sorting on higher-level individuals. Having outlined the basic elements of the C-value enigma and the interaction of DNA content with the cellular and organismal levels, it is now possible to apply these macroevolutionary concepts to the question of genome size evolution.

*Top-Down Sorting.*—The most important top-down effect on transposable elements in the context of the C-value enigma is provided by selection operating at the cellular/organismal level. As outlined above, the accumulation of DNA (by any mechanism) has significant effects on cell size and division rate, which may extend to important organismal features like metabolism, development, and morphology, depending on the group in question. Selection operating on the resulting organismal phenotypes may impose a limit on the degree to which subgenomic components like transposable elements can be multiplied. This is the crux of the hierarchical interpretation of the C-value enigma.

In terms of Vrba's (1989) model of Mustapha Mond sorting, the physiological and developmental properties of the cell/organism may be considered to define the context in which transposable elements must evolve. For example, the loss of metamorphosis at the organismal level may generate what can be considered a liberal genomic society in which the subgenomic citizens enjoy unparalleled freedom to travel and reproduce. This certainly appears to be the case in neotenic salamanders, for example (Gregory 2002c). In the event that larger cells are required as part of an organism-level adaptation (e.g., a lower metabolic rate), the proliferation of the citizenry may even be state sponsored. Highly constrained genomes such as those of birds represent the opposite extreme, in which the subgenomic populace suffers under a totalitarian regime. To stretch Vrba's (1989) political analogy a little further, sometimes rulers can be overthrown by a bottom-up revolution, resulting in the development of a more tolerant administration (see below).

As Vrba (1989) points out, Mustapha Mond sorting ''at a given level most often occurs together with selection at that level,'' but this means only that the downward effect—the *context* of context-dependence—is usually produced by selection at the higher level. Importantly, this is a model of context-dependent *sorting,* and therefore need not be restricted to cases where the lower-level units are driven by selection. As such, this process of sorting can apply to the accumulation (or if the regime is sufficiently strict, even the extermination) of *any* subgenomic elements, regardless of their mechanism of spread. Transposable elements (selfish DNA) are amplified by a process of intragenomic selection, whereas gene duplication and extinction (junk DNA) is more driftlike in nature. However, the exact mechanism(s) in operation to alter the quantity of DNA in the genome are not relevant to the effects that this has on the cell, and by extension the selective consequences experienced by the organism. In the language of hierarchy theory, these processes generate an *aggregate character*—genome size—which has consequences for the *emergent fitness* of cells and organisms.

*Bottom-Up Effects.*—Once again, it has been suggested that the ''acid test'' of hierarchical selection is the demonstration of selection at a higher level defeating opposing selective pressures at a lower level (Vrba 1989). This has been very difficult to demonstrate at the species-versus-organism levels, but it is incredibly easy to show at the organism-versus-cells and organism-versus-selfish DNA levels. The existence of selfish DNA in large but finite quantities does necessitate a multilevel perspective, but this is a questionable example of passing the ''acid test'' because one would expect organism-level selection to be a powerful force against selfish DNA spread in any case. Likewise, it is hardly considered overwhelming support for hierarchy that, except in such cases as cancer, selection at the organism level swamps out selection at the cell level. In reality, the ''acid test'' is not whether selection at a higher level can supercede a lower level per se, but rather whether selection at *any* other level, above or below, is shown to overpower the organism level. Thus, when dealing with levels below the standard organism level, what must be demonstrated is bottom-up selection in which the accumulation of transposable elements and other DNA sequences overpowers the expected top-down pressure of selection acting on the organism.

Amphibians have the most-variable genome sizes among vertebrates, and only the lungfishes surpass the largest salamander genomes (Gregory 2002c). As Roth et al. (1994) point out, ''small genome size is plesiomorphic in amphibians, and large genomes have evolved independently at least twice within frogs and salamanders.'' Genome size increase is a derived feature in lungfishes as well, as shown by fossil osteocyte volume data from both groups (Thomson 1972; Thomson and Muraszko 1978). This extensive diversity in genome size is strongly linked to erythrocyte size in amphibians as it is in all vertebrates, but it bears no clear relationship to metabolic parameters (Gregory 2001b, 2003). Developmental complexity with regard to the presence and intensity of metamorphosis seems to define the distributions of genome size in the Amphibia, but there is another sense in which complexity is of relevance in

this group. Specifically, neuron sizes and differentiation rates are also correlated with genome size, with the result that a large genome means larger and more slowly dividing neurons. Brain volume is limited, which means that a secondarily increased genome size produces secondarily simplified brains composed of a lower number of large and poorly differentiated neurons (Roth et al. 1994, 1997).

The impacts of this brain simplification are particularly notable in miniaturized salamanders of the family Plethodontidae. These animals are direct-developers and have large genomes, but they have also undergone a pronounced reduction in overall body (and therefore brain case) size. At one time, these plethodontids were active visual predators, but the reduction in brain complexity (particularly of the visual processing center) has made this impossible. What is intriguing is that instead of simply limiting the spread of noncoding DNA (probably composed primarily of transposable elements), or indeed reducing genome size along with body size (as may have occurred in birds), these salamanders have evolved compensatory features at the organismal level to accommodate the effects of their large genomes. Not only have they switched to a lie-in-wait predation strategy, but they have also evolved specialized projectile tongues to serve in this new feeding mode (Roth et al. 1997).

Many plethodontids also possess enucleated erythrocytes (Villolobos et al. 1988), which may represent a secondary compensation to the structural problem of circulating very large cells through tiny blood vessels (R. L. Mueller unpublished data). In this case, enucleation would not be a metabolic adaptation, nor would it have allowed genomes subsequently to increase in size as in mammals, but rather it would represent another clear response to pressures on the organism generated by the spread of elements within the genome.

**Punctuated Equilibria at the Genomic Level**

An underappreciated feature of genome size distributions among animals is that most higher taxonomic groups contain only a few members whose genomes have expanded greatly in size. That is, with only a few notable exceptions, most species in a given higher taxon have relatively constrained genome sizes. This is especially notable among vertebrates (with the exceptions being salamanders, cartilaginous fishes, and lungfishes), crustaceans (decapods), and insects (orthopterans), for example. According to Waltari and Edwards (2002), ''when viewed across the entire vertebrate tree, genome size evolution is highly punctuated, with relative stasis and low CVs [coefficients of variation] within clades but dramatic differences and high CVs between clades.'' With the exception of amphibians, the major groups of vertebrates do indeed appear quite subdued in their variation in genome size. Bird genome sizes vary only about twofold, mammals fourfold, reptiles fivefold, and even the hyper-diverse teleosts differ only about tenfold among species (less if ancient polyploids are excluded) (Gregory 2001c) (Fig. 1). So, even despite a large amount of speciation, genome sizes have remained remarkably stable within these classes.

Relative stasis at higher taxonomic levels is obviously of interest in the present context, but of more importance is the pattern of stasis found at the species level. Constancy in genome size within species was reported at the very beginning of genome size study, and it remains an important methodological assumption when calculating genome sizes from relative density or fluorescence measurements by comparison against a known standard (e.g., Hardie et al. 2002). Nevertheless, numerous examples of intraspecific variation in genome size have been presented, sometimes relating to chromosomal polymorphisms, differences between the sexes, or the differential presence of B chromosomes. On the other hand, several previous reports of intraspecific variation have been dismissed as experimental error (e.g., Greilhuber 1998), and impressive stability in genome size has been recorded across vast geographical ranges in some species (e.g., Bennett et al. 2000a). It could be that only certain groups display substantial intraspecific variation in genome size. For example, all of the examples in fishes come from groups that are known to tolerate other large-scale modulations of DNA content

such as polyploidy. Overall, the issue of whether genome sizes are most often stable within species and over long periods of time is a subject of ongoing investigation. However, it does appear that in at least a large fraction of species, genome sizes remain remarkably static.

In viewing this question from the perspective of hierarchy theory, one can find obvious parallels with the debate surrounding gradualistic versus punctuational models of species formation. Under punctuated equilibria, species remain in relative stasis for millions of years and only exhibit significant morphological change during the comparatively rapid period associated with speciation. Interestingly, Gold and Amemiya (1987) have suggested that genome size change may be concentrated in episodes of speciation in fishes. However, given the limited variation in teleost C-values, it is clear that extensive speciation can occur without substantial changes in genome size. Additionally, the approximately normal distributions of genome sizes in mammals and birds bespeak more gradual shifts in these groups (Gregory and Hebert 1999). As with the question of speciation itself, the argument of gradual versus punctuated genome size change is one of relative frequency, not exclusivity.

Evaluations of genome size change in geological time are few and far between and must rely on inferences from fossil cell size data. In the handful of cases currently available for animals, namely lungfishes (Thomson 1972), amphibians (Thomson and Muraszko 1978), and conodonts (Conway Morris and Harper 1988), long-term stasis was found to be the rule. On the flipside of this, rapid change in genome size has been documented in many groups, as exemplified by apparent quantum shifts in certain taxa (Gregory and Hebert 1999). Some authors attribute these punctuations (one may even say ''saltations'') to a shift from one ''stable equilibrium state'' to another (e.g., Narayan 1998). In some cases, the mechanistic basis of these large-scale shifts in genome size can be identified, as with the doubling of the maize genome in only a few million years by a surge in transposable element activity (SanMiguel and Bennetzen 1998).

Assuming for the sake of argument that in most cases C-values are relatively stable within species or even higher taxa, what might be the cause of this genomic stasis? More specifically, how might a hierarchical perspective address this issue? As Gould (1995) points out, ''in the world of hierarchical selection, stable systems usually represent balances of negative feedback between adjacent levels.'' The present discussion is essentially a claim about just such a process of interaction among DNA elements and the cell/organism via the intermediate of genome size. Notably, the example most often provided by hierarchy theorists to demonstrate this dynamic selective equilibrium is none other than the balance between intragenomic and organismal selection in determining the amount of selfish DNA in genomes (Gould 1982, 1983, 1995, 1998).

In terms of the genome, as with higher levels of organization, it is easiest to see hierarchy in action when selection on the higher lever (the organism or the cell) pushes down against pressures generated from below (e.g., the upward mutation pressure of intragenomic selection). A prime example would be the small and mostly invariant genomes of birds (Gregory 2002a). But this process is no less hierarchical when cell- and organism-level selection are relaxed and allow (or, in some cases, perhaps even favor) both larger cells and larger genomes, as following the emergence of enucleated erythrocytes in mammals (Cavalier-Smith 1978; Gregory 2000). And indeed, the process can be one of a shifting balance, because different levels of selection can act synergistically via a feedback loop, as with the coevolution of neoteny and large genome size in certain salamanders (Gregory 2002c). Hierarchical interactions are also apparent when subgenomic elements successfully push up on the organism and force change at the higher level, as shown by the example of brain simplification in amphibians (Roth et al. 1997; Gregory 2002c, 2003).

This notion of multilevel interaction can also help to explain why otherwise static genomes may sometimes change substantially in size. This can occur by internal mechanisms (e.g., whole-scale duplication events, massive transposable element activity), in which case

the organism may be left to adjust to the newly established nucleotype. Perhaps more commonly, external (or at least, higher-level) factors may determine the patterns of genome size change. Such is the case when adaptations at the organism level require the institution of constraints on the genome level (e.g., flight and small cells), or when the genomic analogue of an ''ecological release'' of DNA elements is prompted by a relaxation of these constraints (e.g., following erythrocyte enucleation or the evolution of neoteny).

To reiterate, the claim is not being made here that intraspecific variation and gradual change do not occur in the evolution of genome size. Not even the most ardent supporters of punctuated equilibria make such a bold claim for exclusivity in evolutionary mode. However, if, as with the species level, evidence continues to suggest a high relative frequency of stasis in C-values, then this will clearly be a phenomenon in need of explanation regardless of the number of exceptions that may be found. A pluralistic, hierarchical perspective on genome size evolution would seem to be the most effective way to account for it. Of course, the same has been said many times of stasis at the species level.

### Genomes in the Evolutionary Hierarchy

*Are Genomes "Individuals"?*—To reiterate, qualification as an evolutionary individual is generally considered to involve the criteria of spatiotemporal boundedness (discrete physical borders as well as a birth, stable existence through time, and death) and the capacity for reproduction of high (but not perfect) fidelity. In the simplest application of hierarchy theory to the genomic level, we may consider whether genomes meet these criteria for status as individuals. Certainly, genomes are spatially bounded (into chromosomes within nuclei), they are born (by replication), remain cohesive, and then perish (e.g., at cell death). Moreover, genomes and their components are the most capable level in the hierarchy in terms of reproduction. On the surface, genomes therefore appear to meet all the necessary criteria, and to warrant inclusion in the macroevolutionary hierarchy.

This conference upon genomes of individual status is not without conceptual challenges, however. Most simply, there is the question of what is meant by the term ''genome'': Does it refer to each individual set of chromosomes? The assembled ''genome'' of each individual organism? The collective ''genome'' of a species? These are questions that must be resolved if the concept of genomes as individuals is to make a fruitful contribution to the understanding of evolutionary processes. Unfortunately, there are no simple solutions to this issue, and it may be that all of these definitions are partially valid.

There is also the complication that, in sexually reproducing organisms, recombination disrupts and reshuffles any particular ''genome,'' and fertilization creates an entirely new combination through merger. Of course, populations and species undergo significant reshufflings and mergers of their constituent parts, so this is not a problem unique to the genome level. Again, the solution to this issue will depend heavily on how one defines a ''genome'': if defined as the ''genome'' of an organism, this is a potentially serious issue, but if defined as the collective ''genome'' of a species, it may be of little consequence.

Finally, it is also apparent that the genomes-as-individuals concept is somewhat problematic when moving to the secondary criterion of individuality involving the existence of a discernable trinity of part-individual-collectivity. That is, although genomes are unquestionably composed of identifiable parts (namely chromosomes or even individual DNA sequences), it is not so obvious what a collectivity of genomes would constitute. Cells do indeed make up organisms, which in turn constitute populations, but collections of genomes do not have an obvious counterpart to this scheme. There is no obvious answer to this dilemma, but several solutions do seem possible. At the simplest extreme, it may not be necessary for genomes to exist in physical collectivities to be valid members of the hierarchy. In this regard, it is noteworthy that ''genes'' are almost always ranked directly below ''cells'' in evolutionary hierarchies (see below), even though these do not make up cells (McShea 2002). (Note that ''genes'' do not form ''ge-

nomes'' in their aggregate either, because the genomes of all organisms contain some, if not a great majority, of noncoding sequences). Alternatively, we may postulate a separate hierarchy culminating in genomes in the sense that the global biota consists of clades, but does not combine into any larger entity. Divided (but interconnected) hierarchies have been postulated before (e.g., Eldredge 1998), so this may not be unrealistic despite being rather inconvenient.

*Genomes and the Rest of the Hierarchy.*—Obviously, before the position of genomes in a hierarchy can be contemplated, it must be established what the other members of the hierarchy are to be. As it turns out, there is little agreement on this issue even in terms of the higher-level components of the macroevolutionary hierarchy (see Eldredge 1985 for a review of numerous proposals). In fact, it is not even clear whether one (Gould 2002) or multiple (Eldredge 1998) hierarchies should be constructed. Eldredge (1985, 1998), in particular, envisions two parallel hierarchies, including an ''ecological hierarchy'' composed of entities whose business (literally, in an economic sense) it is to interact with the environment (organisms, avatars, local ecosystems, regional ecosystems) and a ''genealogical hierarchy'' consisting of those entities concerned with ''more-making'' (germ-line genome, organisms, demes, species, monophyletic taxa). Natural selection, in this view, is the process whereby organismal variation generated in the genealogical hierarchy is filtered according to its success in the ecological hierarchy. The organism enjoys a special status in the scheme, because it is the only ''individual'' claiming membership in both hierarchies, although this does not preclude the operation of evolutionary processes at other levels.

Vrba and Eldredge (1984) point out a third possibility in addition to the ecological and genealogical hierarchies, although its independent status may be somewhat questionable:

> A third hierarchy commonly found in biology texts may be called the ''somatic hierarchy'' because it consists of the familiar proteins, organelles, cells, tissues, organs, and organ systems that make up the bodies of individual organisms. These entities are spatiotemporally localized and nested in the same fashion as are the individuals of the genealogical and ecological hierarchies. The somatic hierarchy is a hierarchy of phenotype and as such it is related to, but not properly an aspect of, the genealogical hierarchy. One could perhaps argue that it is more appropriately considered a lower division of the ecological hierarchy.

In most cases, the genome is conspicuously absent from the proposed hierarchy, and even when subcellular components are considered, the leap is usually from ''genes'' to ''cells.'' Some authors do consider ''genomes,'' although this is usually meant in reference to the collected set of genes (but again, this equivocation is inaccurate). In part, this omission of genomes as a whole could stem from the conceptual quagmire involved in attempting to fit them into the standard hierarchical scheme. For one thing, genomes play a dual role in the evolutionary process, acting as both genotypes in a coding and regulatory sense and nucleotypes in a holistic sense, thereby complicating even the simplest hierarchical classifications (Gregory and Hebert 1999). As the hierarchy grows more complex, so does the task of integrating genomes into it. For example, one may argue that the genome must, like the organism, reside in both the economic and genealogical hierarchies in Eldredge's (1985, 1998) scheme. This is because genomes are not only the chief players in genealogy, but they are also responsible for producing the proteins and ultimately the cells used in somatic functions. In the simplest sense, they could be divided along genic lines into ''germline'' and ''somatic'' genomes, but even this would be only the beginning of the necessary expansion because it ignores the nucleotypic effects that genomes exert on cells (and indirectly on organisms) independently of gene action by virtue of their physico-chemical properties.

Speaking of the subgenomic level, Eldredge and Salthe (1984) suggest that chromosomes are not part of the genealogical hierarchy, be-

cause their reproduction can be reduced directly to the replication of their component DNA sequences. However, the highly organized structure of chromosomes, and the role of this structure in replication, suggests that this interpretation is oversimplified. Vrba and Eldredge (1984) exclude chromosomes and cells from any relevant evolutionary hierarchies on different grounds: ''We have difficulty in seeing chromosomes and cells per se engaging in evolutionarily important birth and death processes that need to be considered separately from the dynamics at genome and phenotype levels.'' The prominence of selfish B chromosomes as independent evolutionary actors would seem to refute this assertion as well (e.g., Camacho et al. 2000), as does the continued abundance of unicellular organisms. Selection among self-replicating organelles within cells was probably also a driving force in the extreme genome size reduction of mitochondria (Selosse et al. 2001), which adds an additional twist to this discussion.

Even if one admits that selection among chromosomes and cells is not very common (anymore), there is another reason for including them in the macroevolutionary hierarchy. Specifically, it is relevant not only whether the level in question displays emergent properties that undergo selection, but also whether the level contributes to the emergent fitness of higher levels undergoing sorting. So, although the integrated nature of organisms usually precludes the action of intercellular selection, cells as ''individuals'' nevertheless contribute to sorting at the organism level. Moreover, the ghosts of intercellular selection past do remain in the form of cancer. Likewise, chromosomes are entities whose structure, size, and number can contribute to the cellular phenotype (and, by extension, to sorting at the organism level). Subgenomic elements are also subject to selection on their own level and by their aggregate effects on the cellular/organismal phenotype. As such, it is apparent that cells, genomes, chromosomes, and DNA sequences (genes or otherwise) must all be part of any comprehensive biological hierarchy, despite the conceptual challenges involved.

## Concluding Remarks

It certainly seems remarkable that insights from the study of subgenomic elements, the most reductionistic focal point in biology, could provide the best evidence for a necessary expansion of evolutionary theory at the highest levels. Likewise, that an understanding of genome evolution would require a hierarchical framework developed by paleontologists is surprising. Of course, nature has never been ashamed of presenting complexities that the human mind deems as paradoxical. Such was the case, after all, with the macroevolutionary questions of both genome size and species stasis from which the present discussion has arisen.

Several important implications can be identified from the present discussion. The most obvious is that the genome itself must be recognized as a legitimate level of biological organization with its own evolutionary dynamics and capacity for interaction with other levels. In particular, genomes simultaneously exhibit ''phenotypes'' generated by basic evolutionary processes acting at levels below, and ''genotypes'' (i.e., sources of variation) for the same processes operating at levels above. From this realization follows the implication that the basic *principles* of the Modern Synthesis—mutation, selection, drift—remain valid at each level of biological organization, even though their *expression* at any one level cannot be extrapolated to explain evolutionary phenomena as observed at all levels.

One-dimensional, or even single-level, explanations are clearly insufficient to resolve the complex puzzle of genome size evolution. Even at the genetic level, hierarchical interactions must be invoked in order to achieve a full understanding of the observed patterns. In this light, it seems unfathomable that the same would not also be true of the even more complex issue of diversification at the highest taxonomic levels. In both cases, it is essential to recognize the importance of evolutionary interactions among different levels of organization. Perhaps the most important lesson for macroevolutionists, be they specialists in genomics or paleontology, is that there is great utility in promoting conceptual interactions

among disparate but mutually enlightening fields of scientific inquiry.

## Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and University of Guelph alumni doctoral scholarships, an NSERC postdoctoral fellowship, and the NSERC Howard Alper Post-Doctoral Prize. Sincere thanks to S. Adamowicz, T. Crease, N. Eldredge, P. Hebert, D. Lynn, and two anonymous reviewers who provided helpful comments on an earlier draft of the paper.

## Literature Cited

Arkhipova, I., and M. Meselson. 2000. Transposable elements in sexual and ancient asexual taxa. Proceedings of the National Academy of Sciences USA 97:14473–14477.

Arnold, A. J., and K. Fristup. 1982. The theory of evolution by natural selection: a hierarchical expansion. Paleobiology 8: 113–129.

Baum, D. A. 1998. Individuality and the existence of species through time. Systematic Biology 47:641–653.

Baumann, C., M. Judex, H. Huber, and R. Wirth. 1998. Estimation of genome sizes of hyperthermophiles. Extremophiles 2: 101–108.

Bennett, M. D. 1987. Variation in genomic form in plants and its ecological implications. New Phytologist 106(Suppl.):177–200.

Bennett, M. D., and I. J. Leitch. 2001a. Nuclear DNA amounts in pteridophytes. Annals of Botany 87:335–345.

———. 2001b. Plant DNA C-values Database. Royal Botanic Gardens, Kew, U.K. http://www.rbgkew.org.uk/cval/homepage.html

Bennett, M. D., P. Bhandol, and I. J. Leitch. 2000a. Nuclear DNA amounts in angiosperms and their modern uses—807 new estimates. Annals of Botany 86:859–909.

Bennett, M. D., S. Johnston, G. L. Hodnett, and H. J. Price. 2000b. *Allium cepa* L. cultivars from four continents compared by flow cytometry show nuclear DNA constancy. Annals of Botany 85:351–357.

Biderre, C., M. Pagès, G. Méténier, E. U. Canning, and C. P. Vivarès. 1995. Evidence for the smallest nuclear genome (2.9 Mb) in the microsporidium *Encephalitozoon cuniculi*. Molecular and Biochemical Parasitology 74:229–231.

Biémont, C., C. Vieira, and N. Borie. 2001. Éléments transposables et évolution du génome d'une espèce invasive: le cas de *Drosophila simulans*. Genetics Selection Evolution 33(Suppl. 1): S107–S120.

Britten, R. J. 1996. Cases of ancient mobile element DNA insertions that now affect gene regulation. Molecular Phylogenetics and Evolution 5:13–17.

———. 1997. Mobile elements inserted in the distant past have taken on important functions. Gene 205:117–182.

Brosius, J. 1999. RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory sequences. Gene 238:115–134.

Bull, J. J. 1994. Virulence. Evolution 48:1423–1437.

Burt, A., and R. Trivers. 1998. Selfish DNA and breeding system in flowering plants. Proceedings of the Royal Society of London B 265:141–146.

Burton, D. W., J. W. Bickham, and H. H. Genoways. 1989. Flow-cytometric analyses of nuclear DNA content in four families of neotropical bats. Evolution 43:756–765.

Camacho, J. P. M., T. F. Sharbel, and L. W. Beukeboom. 2000. B-chromosome evolution. Philosophical Transactions of the Royal Society of London B 355:163–178.

Casacuberta, E., and M.-L. Pardue. 2002. Coevolution of the telomeric retrotransposons across *Drosophila* species. Genetics 161:1113–1124.

Cavalier-Smith, T. 1978. Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA C-value paradox. Journal of Cell Science 34:247–278.

———. 1982. Skeletal DNA and the evolution of genome size. Annual Review of Biophysics and Bioengineering 11:273–302.

———. 1985. Cell volume and the evolution of eukaryotic genome size. Pp. 104–184 *in* T. Cavalier-Smith, ed. The evolution of genome size. Wiley, Chichester, U.K.

———. 1991. Coevolution of vertebrate genome, cell, and nuclear sizes. Pp. 51–86 *in* G. Ghiara, ed. Symposium on the evolution of terrestrial vertebrates. Mucchi, Modena.

Charlesworth, B., and C. H. Langley. 1986. The evolution of self-regulated transposition of transposable elements. Genetics 112:359–383.

Colwell, R. K. 1981. Group selection is implicated in the evolution of female-biased sex ratios. Nature 290:401–404.

Conway Morris, S., and E. Harper. 1988. Genome size in conodonts (Chordata): inferred variations during 270 million years. Science 241:1230–1232.

Darwin, C. 1871. The descent of man, and selection in relation to sex. John Murray, London.

Dimitri, P., and N. Junakovic. 1999. Revisiting the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. Trends in Genetics 15:123–124.

Doolittle, W. F. 1987. What introns have to tell us: hierarchy in genome evolution. Cold Spring Harbor Symposia on Quantitative Biology 52:907–913.

———. 1989. Hierarchical approaches to genome evolution. Canadian Journal of Philosophy 14(Suppl.):101–133.

Doolittle, W. F., T. B. L. Kirkwood, and M. A. H. Dempster. 1984. Selfish DNAs with self-restraint. Nature 307:501–502.

Duda, T. F., and S. R. Palumbi. 1999. Developmental shifts and species selection in gastropods. Proceedings of the National Academy of Sciences USA 96:10272–10277.

Eickbush, T. H. 1997. Telomerase and retrotransposons: which came first? Science 277:911–912.

———. 2002. Repair by retrotransposition. Nature Genetics 31: 126–127.

Eldredge, N. 1971. The allopatric model and phylogeny in Paleozoic invertebrates. Evolution 25:156–167.

———. 1985. Unfinished synthesis. Oxford University Press, Oxford.

———. 1995. Reinventing Darwin. Wiley, New York.

———. 1998. The pattern of evolution. W. H. Freeman, New York.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pp. 82–115 *in* T. J. M. Schopf, ed. Models in paleobiology. Freeman Cooper, San Francisco.

Eldredge, N., and S. N. Salthe. 1984. Hierarchy and evolution. Oxford Surveys in Evolutionary Biology 1:184–208.

Esnault, C., J. Maestre, and T. Heidmann. 2000. Human LINE retrotransposons generate processed pseudogenes. Nature Genetics 24:363–367.

Feschotte, C., N. Jiang, and S. R. Wessler. 2002. Plant transposable elements: where genetics meets genomics. Nature Reviews Genetics 3:329–341.

Finnegan, D. J. 1989. Eukaryotic transposable elements and genome evolution. Trends in Genetics 5:103–107.
Finston, T. L., P. D. N. Hebert, and R. B. Foottit. 1995. Genome size variation in aphids. Insect Biochemistry and Molecular Biology 25:189–196.
Fisher, R. A. 1958. The genetical theory of natural selection, 2d ed. Dover, New York.
Flemming, A. J., Z.-Z. Shen, A. Cunha, S. W. Emmons, and A. M. Leroi. 2000. Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. Proceedings of the National Academy of Sciences USA 97:5285–5290.
Frank, S. A. 1986. Hierarchical selection theory and sex ratios. I. General solutions for structured populations. Theoretical Population Biology 29:312–342.
———. 1996. Models of parasite virulence. Quarterly Review of Biology 71:339–344.
Futuyma, D. J. 1998. Evolutionary biology, 3d ed. Sinauer, Sunderland, Mass.
Ghiselin, M. T. 1974. A radical solution to the species problem. Systematic Zoology 23:536–544.
Gilinsky, N. L. 1986. Species selection as a causal process. Evolutionary Biology 20:249–273.
Gold, J. R., and C. T. Amemiya. 1987. Genome size variation in North American minnows (Cyprinidae). II. Variation among 20 species. Genome 29:481–489.
Goodnight, C. J., and L. Stevens. 1997. Experimental studies of group selection: what do they tell us about group selection in nature? American Naturalist 150(Suppl.):S59–S79.
Gould, S. J. 1982. The meaning of punctuated equilibrium and its role in validating a hierarchical approach to macroevolution. Pp. 83–104 *in* R. Milkman, ed. Perspectives on evolution. Sinauer, Sunderland, Mass.
———. 1983. Hen's teeth and horse's toes. Norton, New York.
———. 1992. Punctuated equilibrium in fact and theory. Pp. 54–84 *in* A. Somit and S. A. Peterson, eds. The dynamics of evolution. Cornell University Press, Ithaca, NY.
———. 1995. The Darwinian body. Neues Jarbuch für Geologie und Paläontologie Abhandlungen 195:267–278.
———. 1998. Gulliver's further travels: the necessity and difficulty of a hierarchical theory of selection. Philosophical Transactions of the Royal Society of London B 353:307–314.
———. 2000. The lying stones of Marrakech. Harmony Books, New York.
———. 2002. The structure of evolutionary theory. Harvard University Press, Cambridge.
Gould, S. J., and E. A. Lloyd. 1999. Individuality and adaptation across levels of selection: how shall we name and generalize the unit of Darwinism. Proceedings of the National Academy of Sciences USA 96:11904–11909.
Grantham, T. A. 1995. Hierarchical approaches to macroevolution: recent work on species selection and the ''effect hypothesis.'' Annual Review of Ecology and Systematics 26:301–321.
Gray, Y. H. M. 2000. It takes two transposons to tango: transposable-element-mediated chromosomal rearrangements. Trends in Genetics 16:461–468.
Gregory, T. R. 2000. Nucleotypic effects without nuclei: genome size and erythrocyte size in mammals. Genome 43:895–901.
———. 2001a. Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biological Reviews 76: 65–101.
———. 2001b. The bigger the C-value, the larger the cell: genome size and red blood cell size in vertebrates. Blood Cells, Molecules, and Diseases 27:830–843.
———. 2001c. Animal genome size database. http://www.genomesize.com
———. 2002a. A bird's-eye view of the C-value enigma: genome size, cell size, and metabolic rate in the class Aves. Evolution 56:121–130.
———. 2002b. Genome size and developmental parameters in the homeothermic vertebrates. Genome 45:833–838.
———. 2002c. Genome size and developmental complexity. Genetica 115:131–146.
———. 2003. Variation across amphibian species in the size of the nuclear genome supports a pluralistic, hierarchical approach to the C-value enigma. Biological Journal of the Society 78:329–339.
Gregory, T. R., and P. D. N. Hebert. 1999. The modulation of DNA content: proximate causes and ultimate consequences. Genome Research 9:317–324.
Gregory, T. R., P. D. N. Hebert, and J. Kolasa. 2000. Evolutionary implications of the relationship between genome size and body size in flatworms and copepods. Heredity 84:201–208.
Greilhuber, J. 1998. Intraspecific variation in genome size: a critical reassessment. Annals of Botany 82(Suppl. A):27–35.
Hamilton, W. D. 1964. The genetical evolution of social behaviour, I and II. Journal of Theoretical Biology 7:1–52.
———. 1975. Innate social aptitudes of Man: an approach from evolutionary genetics. Pp. 133–155 *in* R. Fox, ed. Biosocial anthropology. Wiley, New York.
Hansen, T. A. 1980. Influence of larval dispersal and geographic distribution on species longevity in neogastropods. Paleobiology 6:193–207.
Hardie, D. C., T. R. Gregory, and P. D. N. Hebert. 2002. From pixels to picograms: a beginners' guide to genome quantification by Feulgen image analysis densitometry. Journal of Histochemistry and Cytochemistry 50:735–749.
Hawkey, C. M., P. M. Bennett, S. C. Gascoyne, M. G. Hart, and J. K. Kirkwood. 1991. Erythrocyte size, number and haemoglobin content in vertebrates. British Journal of Haematology 77:392–397.
Heddi, A., H. Charles, and C. Khatchadourian. 2001. Intracellular bacterial symbiosis in the genus *Sitophilus*: the 'biological individual' concept revisited. Research in Microbiology 152:431–437.
Hughes, A. L. 1999. Adaptive evolution of genes and genomes. Oxford University Press, Oxford.
Hull, D. L. 1976. Are species really individuals? Systematic Zoology 25:174–191.
———. 1980. Individuality and selection. Annual Review of Ecology and Systematics 11:311–332.
Hurst, G. D. D., and J. H. Werren. 2001. The role of selfish genetic elements in eukaryotic evolution. Nature Reviews Genetics 2: 597–606.
Hutchison, C. A., S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith, and J. C. Venter. 1999. Global transposon mutagenesis and a minimal *Mycoplasma* genome. Science 286:2165–2169.
International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. Nature 409:860–921.
Jablonski, D. 1987. Heritability at the species level: analysis of geographic ranges of Cretaceous mollusks. Science 238:360–363.
Jablonski, D., and R. A. Lutz. 1983. Larval ecology of marine benthic invertebrates: paleobiological implications. Biological Reviews 58:21–89.
Kidwell, M. G. 2002. Tranposable elements and the evolution of genome size in eukaryotes. Genetica 115:49–63.
Kidwell, M. G., and D. Lisch. 1997. Transposable elements as sources of variation in animals and plants. Proceedings of the National Academy of Sciences USA 94:7704–7711.
———. 2000. Transposable elements and host genome evolution. Trends in Ecology and Evolution 15:95–99.
———. 2001. Transposable elements, parasitic DNA, and genome evolution. Evolution 55:1–24.
———. 2002. Transposable elements as sources of genomic var-

iation. Pp. 59–90 *in* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz, eds. Mobile DNA II. ASM Press, Washington, D.C.

Labrador, M., and V. G. Corces. 1997. Transposable element-host interactions: regulation of insertion and excision. Annual Review of Genetics 31:381–404.

Labuda, D., E. Zietkiewicz, and G. A. Mitchell. 1995. Alu elements as a source of genomic variation: deleterious effects and evolutionary novelties. Pp. 1–24 *in* Maraia 1995.

Lack, D. 1966. Population studies of birds. Oxford University Press, Oxford.

Leitch, I. J., L. Hanson, M. Winfield, J. Parker, and M. D. Bennett. 2001. Nuclear DNA C-values complete familial representation in gymnosperms. Annals of Botany 88:843–849.

Levis, R. W., R. Ganesan, K. Houtchens, L. A. Tolar, and F. M. Sheen. 1993. Transposons in place of telomeric repeats at a *Drosophila* telomere. Cell 75:1083–1093.

Lewontin, R. C. 1970. The units of selection. Annual Review of Ecology and Systematics 1:1–18.

———. 2000. The triple helix. Harvard University Press, Cambridge.

Li, W.-H. 1997. Molecular evolution. Sinauer, Sunderland, Mass.

Licht, L. E., and L. A. Lowcock. 1991. Genome size and metabolic rate in salamanders. Comparative Biochemistry and Physiology 100B:83–92.

Lieberman, B. S. 1992. An extension of the SMRS concept into a phylogenetic context. Evolutionary Theory 10:157–161.

Lieberman, B. S., and E. S. Vrba. 1995. Hierarchy theory, selection, and sorting. BioScience 45:394–399.

Lieberman, B. S., W. D. Allmon, and N. Eldredge. 1993. Levels of selection and macroevolutionary patterns in the turritellid gastropods. Paleobiology 19:205–215.

Lloyd, E. A., and S. J. Gould. 1993. Species selection on variability. Proceedings of the National Academy of Sciences USA 90:595–599.

Luning Prak, E. T. L., and H. H. Kazazian. 2000. Mobile elements and the human genome. Nature Reviews Genetics 1:134–144.

Makalowski, W. 1995. SINEs as a genomic scrap yard: an essay on genomic evolution. Pp. 81–104 *in* Maraia 1995.

Malik, H. S., W. D. Burke, and T. H. Eickbush. 1999. The age and evolution of non-LTR retrotransposable elements. Molecular Biology and Evolution 16:793–805.

Maraia, R. J., ed. 1995. The impact of short interspersed elements (SINEs) on the host genome. Springer, New York.

Maynard Smith, J. 2002. The major transitions in evolution. Pp. E17–E22 *in* M. Pagel, ed. Encyclopedia of evolution, Vol. 1. Oxford University Press, Oxford.

Maynard Smith, J., and E. Szathmáry. 1993. The origin of chromosomes. I. Selection for linkage. Journal of Theoretical Biology 164:437–466.

———. 1995. The major transitions in evolution. Oxford University Press, Oxford.

Mayr, E. 1954. Change of genetic environment and evolution. Pp. 157–180 *in* J. S. Huxley, A. C. Hardy, and E. B. Ford, eds. Evolution as a process. Allen and Unwin, London.

———. 1963. Animal species and evolution. Harvard University Press, Cambridge.

———. 1992. Speciational evolution or punctuated equilibria. Pp. 21–53 *in* A. Somit and S. A. Peterson, eds. The dynamics of evolution. Cornell University Press, Ithaca, NY.

McDonald, J. F. 1990. Macroevolution and retroviral elements. BioScience 40:183–191.

———. 1995. Transposable elements: possible catalysts of organismic evolution. Trends in Ecology and Evolution 10:123–126.

McShea, D. W. 2002. A complexity drain on cells in the evolution of multicellularity. Evolution 56:441–452.

Michod, R. E. 1997. Evolution of the individual. American Naturalist 150(Suppl.):S5–S21.

Michod, R. E., and D. Roze. 2000. Cooperation and conflict in the evolution of multicellularity. Heredity 86:1–7.

Mirsky, A. E., and H. Ris. 1951. The desoxyribonucleic acid content of animal cells and its evolutionary significance. Journal of General Physiology 34:451–462.

Mishler, B. D., and R. N. Brandon. 1987. Individuality, pluralism, and the phylogenetic species concept. Biology and Philosophy 2:397–414.

Morrish, T. A., N. Gilbert, J. S. Myers, B. J. Vincent, T. D. Stamato, G. E. Taccioli, M. A. Batzer, and J. V. Moran. 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nature Genetics 31:159–165.

Narayan, R. K. J. 1998. The role of genomic constraints upon evolutionary changes in genome size and chromosome organization. Annals of Botany 82(Suppl. A):57–66.

Orgel, L. E., and F. H. C. Crick. 1980. Selfish DNA: the ultimate parasite. Nature 284:604–607.

Östergren, G. 1945. Parasitic nature of extra fragment chromosomes. Botaniska Notiser 2:157–163.

Pardue, M. L., O. N. Danilevskaya, K. L. Traverse, and K. Lowenhaupt. 1997. Evolutionary links between telomeres and transposable elements. Genetica 100:73–84.

Paterson, H. E. H. 1985. The recognition concept of species. Transvaal Museum Monographs 4:21–29.

Raff, R. A., and T. C. Kaufman. 1983. Embryos, genes, and evolution. Macmillan, New York.

Renzaglia, K. S., E. M. Rasch, and L. M. Pike. 1995. Estimates of nuclear DNA content in bryophyte sperm cells: phylogenetic considerations. American Journal of Botany 82:18–25.

Ridley, M. 1993. Evolution. Blackwell Science, Boston.

Roth, G., J. Blanke, and D. B. Wake. 1994. Cell size predicts morphological complexity in the brains of frogs and salamanders. Proceedings of the National Academy of Sciences USA 91: 4796–4800.

Roth, G., K. C. Nishikawa, and D. B. Wake. 1997. Genome size, secondary simplification, and the evolution of the brain in salamanders. Brain, Behavior and Evolution 50:50–59.

SanMiguel, P., and J. L. Bennetzen. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Annals of Botany 82(Suppl. A):37–44.

Santelices, B. 1999. How many kinds of individual are there? Trends in Ecology and Evolution 14:152–155.

Sapienza, C., and W. F. Doolittle. 1981. Genes are things you have whether you want them or not. Cold Spring Harbor Symposia on Quantitative Biology 45:177–182.

Schmid, C. W., and C. M. Rubin. 1995. Alu: what's the use? Pp. 105–123 *in* Maraia 1995.

Schön, I., and K. Martens. 2000. Transposable elements and asexual reproduction. Trends in Ecology and Evolution 15: 287–288.

Seaborg, D. M. 1999. Evolutionary feedback: a new mechanism for stasis and punctuated evolutionary change based on integration of the organism. Journal of Theoretical Biology 198: 1–26.

Seeley, T. D. 1997. Honey bee colonies are group-level adaptive units. American Naturalist 150(Suppl.):S22–S41.

Selosse, M.-A., B. Albert, and B. Godelle. 2001. Reducing the genome size of organelles favours gene transfer to the nucleus. Trends in Ecology and Evolution 16:135–141.

Smit, A. F. A. 1999. Interspersed repeats and other momentos of transposable elements in mammalian genomes. Current Opinion in Genetics and Development 9:657–663.

Sober, E., and D. S. Wilson. 1998. Unto others. Harvard University Press, Cambridge.

Stanley, S. M. 1975. A theory of evolution above the species level.

Proceedings of the National Academy of Sciences USA 72: 646–650.

———. 1979. Macroevolution: pattern and process. W. H. Freeman, San Francisco.

Swift, H. 1950. The constancy of desoxyribose nucleic acid in plant nuclei. Proceedings of the National Academy of Science USA 36:643–654.

Szarski, H. 1976. Cell size and nuclear DNA content in vertebrates. International Review of Cytology 44:93–111.

———. 1983. Cell size and the concept of wasteful and frugal evolutionary strategies. Journal of Theoretical Biology 105: 201–209.

Thomas, C. A. 1971. The genetic organization of chromosomes. Annual Review of Genetics 5:237–256.

Thomson, K. S. 1972. An attempt to reconstruct evolutionary changes in the cellular DNA content of lungfish. Journal of Experimental Zoology 180:363–372.

Thomson, K. S., and K. Muraszko. 1978. Estimation of cell size and DNA content in fossil fishes and amphibians. Journal of Experimental Zoology 205:315–320.

Trivers, R. 1971. The evolution of reciprocal altruism. Quarterly Review of Biology 46:35–57.

Vendrely, R., and C. Vendrely. 1948. La teneur du noyau cellulaire en acide désoxyribonucléique à travers les organes, les individus et les espèces animales: techniques et premiers résultats. Experientia 4:434–436.

Vieira, C., C. Nardon, C. Arpin, D. Lepetit, and C. Biémont. 2002. Evolution of genome size in *Drosophila*: is the invader's genome being invaded by transposable elements? Molecular Biology and Evolution 19:1154–1161.

Villolobos, M., P. León, S. K. Sessions, and J. Kezer. 1988. Enucleated erythrocytes in plethodontid salamanders. Herpetologica 44:243–250.

Vinogradov, A. E. 1995. Nucleotypic effect in homeotherms: body mass-corrected basal metabolic rate of mammals is related to genome size. Evolution 49:1249–1259.

Voglmayr, H. 2000. Nuclear DNA amounts in mosses (Musci). Annals of Botany 85:531–546.

Vrba, E. S. 1980. Evolution, species and fossils: how does life evolve? South African Journal of Science 76:61–84.

———. 1983. Macroevolutionary trends: new perspectives on the roles of adaptation and incidental effect. Science 221:387–389.

———. 1984. What is species selection? Systematic Zoology 33: 318–328.

———. 1989. Levels of selection and sorting with special reference to the species level. Oxford Surveys in Evolutionary Biology 6:111–168.

Vrba, E. S., and N. Eldredge. 1984. Individuals, hierarchies and processes: towards a more complete evolutionary theory. Paleobiology 10:146–171.

Vrba, E. S., and S. J. Gould. 1986. The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. Paleobiology 12:217–228.

Wake, D. B., G. Roth, and M. H. Wake. 1983. On the problem of stasis in organismal evolution. Journal of Theoretical Biology 101:211–224.

Waltari, E., and S. V. Edwards. 2002. Evolutionary dynamics of intron size, genome size, and physiological correlates in archosaurs. American Naturalist 160:539–552.

Watanabe, K., T. Yahara, T. Denda, and K. Kosuge. 1999. Chromosomal evolution in the genus *Brachyscome* (Asteraceae, Astereae): statistical tests regarding correlation between changes in karyotype and habit using phylogenetic information. Journal of Plant Research 112:145–161.

Watson, J. D., and F. H. C. Crick. 1953. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature 171:737–738.

Weinberg, S. 2001. Can science explain everything? Anything? New York Review of Books, May 31.

Williams, G. C. 1966. Adaptation and natural selection. Princeton University Press, Princeton, N.J.

———. 1992. Natural selection: domains, levels, and challenges. Oxford University Press, New York.

Wilson, D. S. 2000. Nonzero and nonsense: group selection, nonzerosumness, and the human Gaia hypothesis. Skeptic 8:84–89.

Wilson, D. S., and R. K. Colwell. 1981. Evolution of sex ratio in structured demes. Evolution 35:882–897.

Wilson, D. S., and E. Sober. 1994. Re-introducing group selection to the human behavioral sciences. Behavioral and Brain Sciences 17:585–654.

Wright, S. I., and D. J. Schoen. 1999. Transposon dynamics and the breeding system. Genetica 107:139–148.

Wynne-Edwards, V. C. 1962. Animal dispersion in relation to social behavior. Oliver and Boyd, Edinburgh.

Zeyl, C., and G. Bell. 1996. Symbiotic DNA in eukaryotic genomes. Trends in Ecology in Evolution 11:10–15.