PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Conceptual Aspects of the Quantification of the Extent of Biological Diversity

Robert M. May

Phil. Trans. R. Soc. Lond. B 1994 **345**, 13-20
doi: 10.1098/rstb.1994.0082

Conceptual aspects of the quantification of the extent of biological diversity

ROBERT M. MAY

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

CONTENTS

SUMMARY

This paper begins by asking to what extent numbers of species are an adequate measure of biological diversity, either locally or globally; both for evolutionary understanding and for practical applications, biodiversity may often be better quantified at lower or higher levels, from genes to ecosystems. The subsequent discussion, however, focuses on species, and discusses questions that arise in estimating how many species there have ever been, how many there currently are in various taxonomic groups, and how we may quantify the differing degrees of 'independent evolutionary history' or 'taxonomic distinctiveness' in different species or groups. I conclude with opinions about how the practical task of identifying and recording species diversity might be better managed.

1. INTRODUCTION: FROM GENES TO ECOSYSTEMS

Biological diversity exists at many different levels, from the genetic diversity within local populations of a species, or between geographically distinct populations of the same species, all the way up to communities or ecosystems.

Depending on the context, any one of this nested hierarchy of levels can be of predominant importance. At the most fundamental evolutionary level, the genetic diversity within species is the raw stuff upon which evolutionary processes act. On shorter time-scales, such genetic variability enables a species to cope with old and new pathogens, environmental fluctuations, and so on. We still lack a clear understanding of how the long-term survival of many species is likely to be affected by recent and severe reduction in the sizes of their populations, either in the wild or as 'rescued' populations in captive-breeding programmes; Lande's '500 rule' was never intended as more than an initial and crude guess (Lande 1988).

At the opposite extreme, we do not have to embrace the wilder poetic flights of the Gaians to acknowledge that ecosystems can usefully be regarded as supra-organisms for many discussions of the way biological and physical processes entwine to maintain the biosphere as a place where life can flourish. For many discussions of the role of plants in cloud formation and structure, or in water or carbon dioxide balance more generally, we do best by dealing with functional aggregates. The same is true of discussions of soil formation and maintenance, where a diverse array of functional communities of organisms (rather than individual species) are the effective units to be studied.

Alongside the sweep from intrapopulation genetic diversity to ecosystem diversity there lies another rich spectrum of levels, ranging from interpopulation diversity or races within a defined species, through a hierarchy of taxonomic levels from genus to kingdom. Quantification of diversity at these distinct taxonomic levels is different from the hierarchy that ascends from species through communities to ecosystems. Taxonomic hierarchies generally emphasize evolutionary origins and relationships, often against the background of the 600 million year (Ma), or longer, span of the fossil record; species–community–ecosystem hierarchies tend to emphasize contemporary ecological similarities and differences in different environmental and geographical settings.

In short, biological diversity can be quantified in many different ways, at many different levels. Commonly, however, we choose numbers of species. This is sensible, both for practical purposes and for more fundamental reasons.

Take the practical reasons first. Effective action needs money, and money ultimately depends on widespread support among the general public. It is easier to recognize 'biodiversity' immanent in species – especially charismatic vertebrates or colourful plants – than in gene pools or ecosystems.

At a more operational level, we can preserve endangered species, either in captive-breeding programmes *ex situ* or, when possible, *in situ*. Preserving an endangered species' gene pool is not only a more abstract concept, but it involves a range of unanswered scientific questions. Today's populations of Pere David's deer are genetically greatly impoverished, yet seem healthy and viable. How would they fare if we could perform the imaginary experiment of reintroducing them into their original, human-free environment? Past whale bottlenecks and present cheetah homozygosity present parallel questions which challenge contemporary understanding.

At the opposite extreme, preserving ecosystems rather than species is simply more difficult (although correspondingly more important, in my opinion). Preserving the Guam rail in captive-breeding programmes is one thing; preserving Guam's ecosystems is another. More generally, operational assessment at the level of species survival is relatively straightforward, whereas assessment of the meaning of observed changes in, for example, the relative abundances of species in an ecosystem is fraught with uncertainty and ambiguity (witness the continuing controversies about the naturalness or otherwise of *Acanthaster* outbreaks in coral reef communities, or of the role of fire in many ecosystems).

For these reasons, the bulk of this paper deals with conceptual issues surrounding the assessment of how many species there have ever been on earth, how many there are today, and how we might quantify the relative 'taxonomic distinctiveness' of different species. In so doing, I follow most of the other authors in this volume. Before proceeding, however, I stress some concrete examples of practical problems that arise from too mindless a focus on simple species counting.

A variety of factors, including concern for the conservation of wildflowers and insects, is leading to some changes in agricultural practice in parts of Britain. These include wider field margins, where 'trials have shown that sowing a 2 m perennial grassy sward with a wildflower mix will establish a weed-fighting combination that controls and chokes troublesome annuals like sterile brome and cleavers ... the border will also harbour insects which eat cereal pests' (Anon 1994). This theme is expanded by Smith *et al.* (1994). The use of currently available commercial wildflower seed mixtures for these purposes, and also for distribution along road margins and so on, is increasingly advocated both for the practical purpose of weed and pest control, as just noted, but also on general grounds of enhancing or restoring biodiversity. Unfortunately, much of this praiseworthy but naïve enthusiasm for biodiversity is based on a Victorian-chocolate-box-top vision of wildflowers in the British countryside. The British countryside in fact exhibits great local variation, and different regional populations of particular wildflower species can show large differences in morphology and genetic composition (A. T. Jones, private communication). Restoration programmes need to take account of such local variability, and to ensure it is conserved. This requires working with local farmers, using Sites of Special Scientific Interest (SSSI's), nature reserves, and rich meadows. Elevated Personages scattering commercial wildflower mixtures from speeding cars along road margins may make good magazine copy, but whether the long-term contribution to plant diversity in Britain is positive or negative is not clear.

Another instance of well-intentioned practice, guided, however, by visual rather than scientific considerations, is the widespread recent planting of hedges along newly widened roads in Britain. Such hedgerow preservation is praiseworthy, but usually uses non-native hawthorns and other alien species, mainly from eastern Europe (where seed-collecting costs are significantly lower than in Britain). But there is evidence (Jones 1994) that alien forms perform less well than native ones as hedgerow material, probably owing to poorer ecological adaptations to local environments. Here is an example where, visual impressions to the contrary, the interests of biological diversity are not being well-served (and the long-term costs of preserving these hedges may well offset the initial savings).

The basic message of this extensive introduction is that biological diversity has many dimensions. Summarizing it by a simple species count, as is done in the rest of my chapter, can often obscure conceptual understanding, and can sometimes do harm in practice.

2. HOW MANY SPECIES HAVE EVER LIVED?

As a background to quantifying present species richness, it is useful to examine the fossil record of plants and animals over the past 600 Ma or so, since the Cambrian. Estimates of the lifespans of species in the fossil record, from origination to extinction, are mainly indirect; the best are based on computer analyses of large numbers of cohorts of fossil genera (Raup 1978). May *et al.* (1994) have summarized a variety of such estimates, which suggest the average species' lifespan is around 5–10 Ma. There is, however, much variability both within and among groups. Thus the characteristic lifespan of mammal species in the fossil record is roughly 1 Ma, whereas there are suggestions that insect species may be unusually long-lived, at least in north-temperate regions (Coope 1994; Labandeira & Sepkoski 1993). Even within marine invertebrates, for which Raup (1978) assigned an average species lifespan of 11 Ma, there are significant variations from group to group; Mesozoic ammonoids, for instance, have an average species' lifespan of only 1–2 Ma.

Sepkoski (1992) suggests that species diversity in the fossil record has, very roughly and with severe fluctuations, increased linearly over the 600 Ma of the Phanerozoic. If we combine this with an estimated species' lifespan of 5–10 Ma before extinction, we conclude that roughly 2–4% of all species of plants and animals ever to have lived are alive today.

Around 95% of the roughly 250 000 species in the fossil record are, however, marine animals (Sepkoski 1992; Raup 1976). This contrasts greatly with the situation today, where only 15% or so of recorded species of plants and animals are found in the sea; most (56%) are terrestrial insects. So conclusions drawn from average patterns in the fossil record could be misleading. Although the fossil record for insects is much more fragmentary than for shallow-water invertebrates, the indication is that insect diversity has risen steadily – very roughly, linearly – over the past 450 Ma or so (Labandeira & Sepkoski 1993). If we take 10 Ma as the average lifespan of an insect species, from origination to extinction, we would thus conclude that at least 5% of all plant and animal species are alive today; this figure could be larger if average insect species' lifespans exceeded 10 Ma and/or if insect species' diversity increased faster than linearly throughout the Mesozoic (May *et al.* 1994). To put this in a different (and more trenchant) way, the fossil record, being so dominated by marine invertebrates, may be unrepresentative of the history of life on earth over the past 400 Ma or so, when terrestrial insects may have dominated (Briggs 1994).

In drawing comparisons across geological epochs, and – even more – across taxonomic groups, we must worry whether we are comparing like with like. There certainly are problems with 'taxonomic inflation' over the years; what was a genus to Linnaeus might be a family, or even an order, today (Sepkoski 1992). These instabilities, applying differently to different groups, and depending to some degree on the different attention given to different groups, create problems when we try to make comparisons or to draw inferences about the history of particular groups of species from their superspecific taxonomy (Patterson & Smith 1989). Such methodological differences among taxonomic groups show up in other fundamental ways. For example, Selander (1985) has observed that different strains of what is currently classified as a single bacterial species, *Legionella pneumophila*, have nucleotide sequence homologies (as revealed by DNA hybridization) of less than 50%; this is as large as the characteristic genetic distance between mammals and fishes. In short, one of the basic conceptual issues in quantifying biological diversity is the extent to which a 'species' does or does not represent the same unit of evolutionary currency for a bacterium, a protozoan, a mite, and a bird.

3. ESTIMATING CONTEMPORARY SPECIES NUMBERS

Other papers in this volume deal with estimates of species numbers, either for particular groups in particular places, or more generally (see also Hammond 1992; May 1994). So what follows is only a sketchy guide to what I see as some of the conceptual problems. I have chosen to list these conceptual issues under the heading of particular taxonomic groups, rather than more abstractly.

(*a*) *Terrestrial insects*

Erwin's (1982) provocative assessment of numbers of insect species, by a chain of argument anchored to numbers of beetle species in the canopy of particular tropical tree species, brings several conceptual issues into sharp focus.

First, it raises questions about extending comparisons from one geographical location to another. Even if roughly 20% of canopy beetles are effectively specialized to *Luehea seemannii* at a particular study site, how do we know these same beetle species do not effectively specialize on other tree species at other sites? Or that *L. seemannii* has a different effectively specialized beetle fauna elsewhere? Examples which illustrates such complications, and cut across any simple 'scaling-up', are discussed by Thomas (1990) and May (1990a).

Second, a related question is what we mean by 'effectively specialized', and how we assess it (May 1990*a*). What proportion (p_i) of the beetle species found on *L. seemannii* is found on i other tree species? What general guidelines can basic ecological theory offer us? (Not many!) Even relatively simple, technical questions have received surprisingly little attention: to what extent will sampling problems obscure efforts to assess p_i from comprehensive field studies, and how large need samples be so that ecological signals are not overwhelmed by sampling noise.

Third, to what extent can we trust figures for canopy beetle species to be broadly representative of, say, ants? Hammond discusses these issues more fully elsewhere in this collection (see also Hammond 1992, 1994; May 1990*a*).

Fourth – and to sound for the first time a tocsin that will ring throughout this section – to what extent are Erwin's extrapolations likely to be consistent with the fraction of species in his collections which are previously unrecorded (remember, Erwin's samples have not yet been 'keyed out')? Erwin's chain of argument suggests 30 million species of (tropical) insects. But only around 1 million insect species have been recorded. Thus, on average, we might expect only around 3% of Erwin's as-yet-unidentified beetle species (or less than 40 of his 1100+ canopy beetle species from *L. seemannii*) to be already known. I will be amazed if this is the case (and I would chance a guess that 30% is more likely). This form of 'check' seems to me to be an important constraint on any extrapolation or indirect assessment of species richness. Of course, the fraction of a newly studied flora or fauna that has been previously recorded is likely to vary greatly from place to place and group to group (Hammond 1992, 1994 and this volume). And these kinds of crude numerical estimates of fractions previously recorded ignore complications that can arise from significant disparities in the relative abundances of different species. But even so, the most extreme of such figures from recent studies rarely exceed 50% new species (May 1994).

(*b*) *Marine macrofauna*

Grassle & Maciolek (1992) proposed a global total of 10 million or more marine molluscs, crustaceans, polychaete worms and other benthic macrofauna. These estimates were extrapolations from 'box-core' samples from the ocean floor. Because fewer than 200 000 such marine species are currently recorded, this estimate would seem broadly to suggest 2% or fewer recorded species in samples from really new places; no such extreme figures have been found (May 1992, 1993; but see Poore & Wilson 1993). This re-echoes the point made in the preceding paragraph.

Grassle & Maciolek's (1992) influential and stimulating paper raises another important conceptual issue, relevant to biodiversity assessment. They suggest that the underlying cause of the great ocean-floor species richness which they project is, first, that the input of nutrients to sediments is inherently patchy and ephemeral and, second, that sediment-dwellers themselves create small-scale disturbances which further increase environmental heterogeneity. Such spatio-temporal heterogeneity and disturbance is recognized by ecologists as a powerful promoter of diversity. All this is, they suggest, compounded by the lack of barriers (compared with terrestrial environments) to long-distance dispersal, which allows distant migrants to contribute to reshuffling patterns of local diversity. My worry about these ideas is that, if long-distance movement is an important cog in the machinery maintaining overall diversity, then you cannot extrapolate a 'local' rate of adding species with area (as Grassle & Maciolek do) beyond the characteristic distance scale on which the dispersal/diversity mechanism operates. More generally, any kind of scaling-up or extrapolation should be based on a clear understanding of the distance scales which characterise underlying ecological processes. This will often be an unhelpful counsel of perfection, but it is nevertheless a conceptual issue which I think invalidates several estimates based on such scaling-up.

(*c*) *Fungi*

Another interesting and dramatic upward revision is by Hawksworth (1991) for fungi. Currently some 70 000 species of fungi, *sensu lato*, are recognized. But for Britain's comparatively well-studied species of fungi and vascular plants, the ratio is around 6 : 1. If this ratio applies globally to the quarter million or so plant species, we arrive at an estimated 1.5 million species of fungi. And this estimate is in some ways conservative; for example, it has not allowed for fungi associated with insects and other animals.

For one thing, this estimate implies that, on average, about 95% of the fungal species found in a newly studied region should be previously unrecorded ones. In fact, the proportion of new species found in most of such studies is typically 15–30%, and rarely higher. This proportion, however, depends to some extent on the intensity and length of time over which studies are conducted; Hawksworth (1993) cites tropical examples where 50–70% of even the larger fungi have proved to be undescribed. Here, again, I am appealing to the check of any extrapolated estimate against secure facts about what fraction of species in previously unstudied areas are new to science.

For another thing, Hawksworth's estimate raises questions about extrapolating from Britain, which from an Australian viewpoint is a damp and fungal place, to other regions. This issue is covered more fully by Hammond (this volume; see also Thomas 1990).

(*d*) *Microorganisms*

The many questions surrounding the biological diversity of microorganisms – a group which created the oxygen-rich biosphere, and which plays a crucial role in maintaining soils and other ecosystem services – are discussed more fully elsewhere in this collection (see also Hammond 1994; May 1990*a*, 1994). To my mind, the central conceptual issues are the operational differences between definitions of a species for some groups of microorganisms (especially viral 'quasispecies') versus, say, birds and mammals.

(*e*) *Parasite diversity*

By the same token, it could reasonably be argued that for each species of metazoan or vascular plant there is at least one specialized species of parasitic nematode and protozoan, along with at least one species of bacterium and virus. Thus any estimate of plant and animal diversity can be multiplied by five, at a stroke. Even if we relegate bacteria and viruses to a different category of biodiversity, we still have a multiplicative factor of three. I am unsure what the conceptual or practical consequences are. Certainly there will be no lobby to save threatened nematodes and protozoans.

(*f*) *Differences between biodiversity on land and in the sea*

As noted earlier, only 15% or fewer of recorded species inhabit the marine realm. But the sea is increasingly represented as we move to higher taxonomic levels, from genus to phylum. Indeed, at the level of phylum, or basic body-plan, diversity is much greater in the sea (32 of 33 phyla in the sea, versus 12 of 33 on land, by one classification; or, at the level of class, 73 animal classes in the sea, 35 in freshwater, and 33 on land (Nicol 1971)). Possible reasons for this are listed by May (1994), but there is no generally agreed understanding. This is a major 'conceptual issue'.

(*g*) *Geographical distributions and ranges*

Implicit in several of the questions raised above are comparisons among the characteristic geographical ranges of different species and different groups. Fenchel (1993; see also Hammond 1994) has recently suggested that – at least in the sea, and possibly more generally – small organisms (roughly 10^{-4} m and below) may typically have wider geographical distributions than those of intermediate size (as exemplified by insects and mites). He has documented this for some marine protozoans and invertebrates, and conjectures that it might explain the humped-shape distribution of numbers of species versus physical size, found in studies of particular groups and more generally by May (1978). If Fenchel is right, this could explain why the total number of protozoan species is an order-of-magnitude less than the global number of invertebrate species, yet typically more protozoan species than invertebrate species are found in any one pond: the protozoan species have larger geographical distributions than the invertebrate species.

Here again is a central ecological and evolutionary question, which clearly relates to many of the questions about scaling-up and extrapolation that are raised above, and by other authors in this collection. Lacking centralized inventories and codifications of such information about ranges, across different taxonomic groups and different characteristic sizes of individuals, we are only just beginning to deal with these questions.

(*h*) *Sibling species*

Knowlton (1993) has recently surveyed evidence which suggests sibling species – species which are difficult or impossible to distinguish based on their available morphological characters – are common in all major marine groups and habitats. She argues that such widespread misidentification of a group of truly distinct species as being a single species arises partly from inadequate study of morphological features and partly because such groups diverge in habitat, life history, and chemical recognition systems without parallel divergence in morphology. As reviewed by Knowlton, a large number of abundant, well-studied, and/or economically important taxa have recently been shown to be complexes of sibling species.

The sibling species phenomenon clearly poses problems both for taxonomic research on certain groups, and for ecological and evolutionary understanding that is based on such taxonomy. As Knowlton (1993) writes: 'consider a world where birds are only occasionally seen alive by the handful of scientists who study their alpha taxonomy. They arrive in museums either as colourless corpses in jars of formalin, or as skeletal material alone. The bills are often delicate structures whose normal shape cannot be reliably inferred from preserved material. Growth is often indeterminate, and weather can affect both the size and shape of the skeleton. Field observations are generally limited to a few hours a day, and identification keys, where they exist, generally lack information on colour pattern and bill shape. Communication between individuals probably occurs via pheromones, as there are few auditory or visual displays. Contact chemicals or micrometereological conditions appear to shape preferences for nesting and feeding sites.' Underlining the consequent problems, she asks whether, under these circumstances, we would not see Darwin's finches or MacArthur's warblers as single species, with obvious implications for our ecological understanding.

4. QUANTIFYING THE 'TAXONOMIC DISTINCTIVENESS' OF A SPECIES

As we move from the furries and featheries, down through the innumerable species of insects, and on down to bacteria and viruses, sentimental concern does not merely wane. It changes sign. We mourn extinction of bird and mammal species, whereas we are about to celebrate the deliberate extinction of the smallpox virus. These are facts, but they lack any conceptual underpinning.

Vane-Wright *et al.* (1991) were the first to suggest that, for conservation purposes, we should quantify the relative values we attach to different species. The scheme they proposed attempts, in essentials, to assign an objective value to the 'taxonomic distinctiveness' or degree of 'independent evolutionary history' (IEH) that is vested in a given species. Such a weighting is made relative to other members of a group. Although currently there is much activity in this area, no one (to my knowledge) has yet made a start on extending IEH valuation across disparate taxonomic groups.

The tuatara (*Sphenodon*), for example, is a large, iguana-like reptile which is the sole survivor of a group that flourished in the Triassic. Today it survives as two species on a few islets off the coast of New Zealand. The tuatara branched off from the main stem of the reptile's phylogenetic tree so long ago, and is so distinctive, that it comes close to being a two-species subclass of its own (Daugherty *et al.* 1990). How do we value the tuatara against any other species of reptile? At one democratic extreme, we could regard all species as equally important, each a unique evolutionary product; in this view, the tuatara is no more important than any other among the roughly 6000 species of reptiles. At the opposite extreme, we might give equal weight to each 'sister group' in the phylogenetic tree of reptiles; on this basis, the two species of tuatara would be weighed equally with the sum of all 6000 other reptile species. Vane-Wright *et al.* (1991) propose a sensible middle way, based on the topology of the phylogenetic branching diagram, which seeks to value species according to some rough measure of their taxonomic distinctiveness or IEH, and which gives results intermediate between the two extremes just outlined (the tuatara, on this scheme, would represent something like a few percent of the taxonomic distinctiveness found among reptiles, intermediate between the 0.03% of the democratic extreme and the 50% of the opposite extreme; May, 1990*b*).

Various refinements of these basic ideas are being actively pursued (Faith 1992, 1993, this volume; Williams *et al.* 1991; Crozier 1992). Ideally, if we had some quantitative measure of the branch lengths within the phylogenetic tree of the group in question, we could unambiguously quantify the amount of IEH vested in a species, by adding up the lengths of the branches which connect it to the base of the tree and appropriately discounting all shared branches (Faith 1992, 1993; May & Nee 1994). If we could preserve only, say, half the species in the group, the optimum choice would then be found by maximizing the summed branch length that was preserved. However, generally we have only the topology of the tree, without quantitative measures of the various branch lengths; in this case, the best procedure would be to assign the branches the lengths that are, on average, most likely for this particular topology, and then go forward on this basis. Such a procedure will, of course, often in fact be suboptimal, because the underlying evolutionary tree differs from the statistically 'expected' one. In general, however, extensive theoretical simulations of choices made on a topological basis, from artificially generated trees whose underlying branch lengths are known, suggest that values assigned in this way are close to the 'true' ones (May & Nee 1994). Ultimately, our question is how much of the IEH within a group will be preserved if we can only save, say, 10 of 20 species? The simulations referred to above suggest that, for the 10 of 20 cases, we can on average preserve 82% of the group's IEH if we have quantitative information about branch lengths, 77% if we have only topological information about the branching structure of the phylogenetic tree, and 63% if we must choose at random (May & Nee 1994).

Real situations will obviously involve many other important considerations, including other measures of the relative values of species (for example, in preserving 'ecosystem services', or in possessing unusual behavioural or ecological properties that are not captured by crude measures of genetic distances), and political and economic constraints on which areas may be preserved. Any programme of assessment and quantification of biological diversity needs to go beyond mere species counting, and move towards developing a 'calculus of diversity' along the lines just sketched.

5. THE WORKFORCE AND THE DATABASE

No survey of conceptual aspects of the assessment of biological diversity is complete without consideration of how the effort is being deployed, and how the emerging information will be organized. These questions are, of course, addressed elsewhere, but a few of the conceptual issues deserve emphasis.

Systematic information about how the taxonomic workforce is deployed among the various taxonomic groups is hard to get. This fact is itself revealing, and unfortunate. One survey, based on information from Australia, the U.S.A. and the U.K. (Gaston & May 1992), can be broadly summarized as saying that the taxonomic workforce in each of these three countries is roughly evenly divided among vertebrates, invertebrates, and vascular plants; microorganisms typically account for less than 5%. Taking a very conservative estimate of 3 million invertebrate species as the global total, this means the ratio of taxonomists to species is an order-of-magnitude greater for vertebrates than for plants, and two orders-of-magnitude greater for vertebrates than for invertebrates. These disparities are mirrored in publications per species (May 1988). This is no way to run a business.

Gaston & May (1992) also make the indirect estimate that only about 6% of practising taxonomists (professionals and serious amateurs) are based in developing countries in Africa, Asia and Latin America. This 6% figure is similar to the corresponding estimates for other scientific subjects. Such a similarity is, of course, to be expected, but the figure has a different significance than it has, for example, for numbers of mathematicians or chemists: the greatest part of the planet's terrestrial diversity – both recorded and unrecorded – is in just those tropical countries where only such a small fraction of the taxonomic workforce is based

This mismatch between the geographical location of the workforce and its workload gives special point to how we organize the information we have.

First, we simply need to be moving faster to coordinate the information that already exists, on file cards and computers, scattered around the world's major and minor museums and other collections. This requires money, but on nothing like the scale we currently fund library catalogues or astronomical enterprises (the taxonomy and systematics of stars, in effect): many of the planet's species look to have a shelf life shorter than our richly and carefully preserved books, much less the average star. Conceptual questions arise in the design of such coordinated databases. They need, where possible, to go beyond basic taxonomic information, codifying information about species distributions, vegetation types, climate, and the physiographic variables which influence where a species may live. A leading example of such a database is Australia's Environmental Resources Information Network (ERIN). The result is that ERIN can predict, for example, where new populations of an endangered species with a limited known range might be expected to be found, or what regions are likely to become trouble spots of endangered species in the near future. ERIN has also been a pioneer in dealing with vexed questions of ownership of data, and who should pay how much to acquire it, and under what circumstances. I do realize the difficulties here: Stork & Hine (1994), for instance, estimate that 40% of the roughly 400 000 recorded species of beetles are known from only one site. But we must do the best we can, using informed guesses (and calibrating information with indices to indicate its reliability; see Hammond, this volume).

Second, these databases must be widely available and 'customer friendly'. We need to accelerate current efforts for international cooperation and coordination, so that common formats are increasingly agreed and used.

Third, we should be forward-looking in the compilation and use of these databases. Using CD-ROM, we can store images of type-specimens; three-dimensional hologram images should be the norm of the future. Here is one way partly to remedy the mismatch between where the information is – in the collections of the major museums of natural history in Europe and the U.S.A., legacies of an imperial past – and where the biodiversity is: in the developing countries of tropical Africa, Asia and America. More than this, we should be aiming to combine these synoptic databases with computerized keys. In this way, the laborious and time-consuming task of identifying species, and of assessing which species among a new collection have previously been recorded, could be greatly speeded.

Fourth, much is written elsewhere of the need for 'parataxonomy' or quick-and-dirty ways of assessing biological diversity, in all its dimensions, in a representative range of carefully chosen sites. I have nothing to add, except strongly to emphasize the overriding importance of such innovative approaches. Time-honoured methods have their defenders, but we literally lack 'world enough and time'.

Much of this will require not so much a refinement of conceptual issues as a revolution in attitudes, both inside and outside the profession. Currently, in Britain, an estimated 1400 professional conservators care for the humanities' collections, in vivid contrast to fewer than 20 safeguarding natural history collections (Seymour 1994). This reverence for the handiworks of humans, and disregard for the handiworks of the natural world, is deep-rooted in the fashions and accidents of the intellectual history of our species.

How will the enterprise of quantifying biological diversity look, 50 years from now? I hope we will have had the vision and persuasive force to put in place the kind of codification of existing information which I have outlined above, and which others have discussed more fully elsewhere in this collection. And I hope this database will be complete with associated information about the ecology, evolution and conservation status of individual species, even if this is very often a guess (and identified as such, with some indication of the precision). So many of the conceptual questions which I have raised in earlier sections can only be answered if we indeed have such a synoptic and coherent database. The problem is a nonlinear one: good assessment of the true magnitude and distribution of biological diversity hinges on fundamental research, based on better codification of the scattered information we already have; but that codifying effort is unlikely to happen unless we put more intellectual and financial resources into basic taxonomic research.

So, in 50 years I expect us to have a fairly good idea of roughly how many species there currently are on earth, and even of why it is that number, rather than many more or less. At current rates of advance, it is even conceivable that species identifications and subsequent taxonomic assignments will be based primarily on automated analyses of appropriately chosen DNA or other molecular material, keyed out against synoptic molecular cladograms. But the addition of new species, and the movement toward a comprehensive account of the treasurehouse of biological diversity that billions of years of evolution have bequeathed us, will, I believe, necessarily accumulate more slowly.

REFERENCES

Anon. 1994 Wildlife-friendly weed control. *Eng. Nat.* **12** (March), 12.

Briggs, J.C. 1994 The great extinction myth. *Nature, Lond.* (In the press.)

Coope, G.R. 1994 The response of insect faunas to glacial-interglacial climate fluctuations. *Phil. Trans. R. Soc. Lond.* B **343**, 19–26.

Crozier, R.H. 1992 Genetic diversity and the agony of choice. *Biol. Consv.* **61**, 11–15.

Daugherty, C.H., Cree, A., Hay, J.M. & Thompson, M.B. 1990 Neglected taxonomy and continuing extinctions of tuatara. *Nature, Lond.* **47**, 177–179.

Erwin, T.L. 1982 Tropical forests: their richness in *Coleoptera* and other arthropod species. *Coleopt. Bull.* **36**, 74–82.

Faith, D.P. 1992 Conservation evaluation and phylogenetic diversity. *Biol. Consv.* **61**, 1–10.

Faith, D.P. 1993 Systematics and conservation: on predicting the feature diversity of subsets of taxa. *Cladistics* **8**, 361–373.

Fenchel, T. 1993 There are more small than large species? *Oikos* **68**, 375–378.

Gaston, K.J. & May, R.M. 1992 The taxonomy of taxonomists. *Nature, Lond.* **356**, 281–282.

Grassle, J.F. & Maciolek, N.J. 1992 Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. *Am. Nat.* **139**, 313–341.

Hammond, P.M. 1992 Species inventory. In *Global diversity: status of the Earth's living resources* (ed. B. Groombridge), pp. 17–39. London: Chapman & Hall.

Hammond, P.M. 1994 Described and estimated species numbers: an objective assessment of current knowledge. In *Microbial biodiversity and ecosystem function* (ed. D. Allsopp, D. L. Hawksworth & R. R. Colwell), pp. 11–25. Wallingford: CAB International.

Hawksworth, D.L. 1991 The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycol. Res.* **95**, 441–456.

Hawksworth, D.L. 1993 The tropical fungal biota: census, pertinence, prophylaxis, and prognosis. In *Aspects of tropical mycology* (ed. S. Isaac, J. S. Frankland, A. J. S. Whalley & R. Watling), pp. 265–293. Cambridge University Press.

Jones, A.T. 1994 Hedge planting: the implications of using alien rather than native genotypes.

Knowlton, N. 1993 Sibling species in the sea. *A. Rev. Ecol. Syst.* **24**, 189–216.

Labandeira, C.C. & Sepkoski, J.J. 1993 Insect diversity in the fossil record. *Science, Wash.* **261**, 310–315.

Lande, R. 1988 Genetics and demography in biological conservation. *Science, Wash.* **241**, 1455–1460.

May, R.M. 1978 The dynamics and diversity of insect faunas. In *Diversity of insect faunas* (ed. L. A. Mound & N. Waloff), pp. 188–204. Oxford: Blackwell.

May, R.M. 1988 How many species are there on earth? *Science, Wash.* **241**, 1441–1449.

May, R.M. 1990*a* How many species? *Phil. Trans. Roy. Soc. Lond.* B **330**, 292–304.

May, R.M. 1990*b* Taxonomy as destiny. *Nature, Lond.* **347**, 129–130.

May, R.M. 1992 Bottoms up for the oceans. *Nature, Lond.* **357**, 278–279.

May, R.M. 1993 Marine species richness. *Nature, Lond.* **361**, 598.

May, R.M. 1994 Biological diversity: differences between land and sea. *Phil. Trans. Roy. Soc. Lond.* B **343**, 105–111.

May, R.M. & Nee, S. 1994 Making conservation choices: towards a calculus of biodiversity. (In preparation).

May, R.M., Lawton, J.H. & Stork, N.E. 1994 Assessing extinction rates. In *Extinction rates* (ed. J. H. Lawton & R. M. May). Oxford University Press. (In the press.)

Nicol, D. 1971 Species, class, and phylum diversity of animals. *Q. Jl Fla. Acad. Sci.* **34**, 191–194.

Patterson, C. & Smith, A.B. 1989 Periodicity in extinction: the role of systematics. *Ecology* **70**, 802–811.

Poore, G.C.B. & Wilson, G.D.F. 1993 Marine species richness. *Nature, Lond.* **361**, 597–598.

Raup, D.M. 1976 Species diversity on the Phanerozoic: a tabulation. *Paleobiology* **2**, 279–288.

Raup, D.M. 1978 Cohort analysis of generic survivorship. *Paleobiology* **4**, 1–15.

Selander, R.K. 1985 Protein polymorphism and the genetic structure of natural populations of bacteria. In *Population genetics and molecular evolution* (ed. T. Ohta & K. Aoki), pp. 85–106. Berlin: Springer-Verlag.

Sepkoski, J.J. 1992 Phylogenetic and ecologic patterns in the Phanerozoic history of marine biodiversity. In *Systematics, ecology, and the biodiversity crises* (ed. N. Eldridge), pp. 77–100. New York: Columbia University Press.

Seymour, J. 1994 No way to treat a national treasure. *New Scient.* 12 March, 32–35.

Smith, H., Feber, R.E., Johnson, P.J., McCallum, K., Jensen, S.P., Younes, M. and Macdonald, D.W. 1994 *The conservation and management of arable field margins.* Peterborough: English Nature.

Stork, N.E. & Hine, S.J. 1994 Patterns of distribution and abundance of described species as deduced from recent taxonomic revision (Submitted.)

Thomas, C.D. 1990 Estimating the number of tropical arthropod species. *Nature*, **347**, 237.

Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. 1991 What to protect: systematics and the agony of choice. *Biol. Conserv.* **55**, 235–254.

Williams, P.H., Humphries, C.J. & Vane-Wright, R.I. 1991 Measuring biodiversity: taxonomic relatedness for conservation priorities. *Aust. J. syst. Bot.* **4**, 665–679.