Systematics and Biodiversity **1** (1): 13–28
DOI: 10.1017/S1477200003001051 Printed in the United Kingdom

Issued 19 May 2003

Richard M. Bateman*
Department of Botany,
The Natural History Museum,
Cromwell Road, London,
SW7 5BD, UK

William A. DiMichele
Department of Paleobiology,
National Museum of Natural
History, Smithsonian
Institution, Washington,
DC 20560, USA

submitted March 2002
accepted November 2002

Genesis of phenotypic and genotypic diversity in land plants: the present as the key to the past

Abstract The evolutionary history of vascular plants is reviewed by extrapolation back through time from a wide range of data recently derived from the present flora, using as the central theme evolutionary inferences gained from phylogenies reconstructed as cladograms. Any region of the genome can be used to infer relationships, but only a combination of knowledge of morphology and the developmental genes that underpin morphology can allow evolutionary interpretation of macroevolutionary transitions; this in turn is necessary to identify *bona fide* evolutionary radiations and any putative causal key innovations. Such studies require clades to be delimited not by the inclusion of particular extant 'crown' species but rather by specific apomorphies, thereby giving important phylogenetic roles to extinct as well as extant species. Dating phylogenetic divergences via molecular clocks remains seriously inaccurate, and ultimately relies primarily on fossil benchmarks. First principles suggest that evolution of most regions of the genome is fundamentally gradual, whereas evolution of regions especially prone to strong selection pressure, and of the many facets of the phenotype, is punctuational, being characterized through time dominantly by stasis. Sequence data have proved valuable for inferring monophyletic groups, but within the now widely accepted context of monophyly the taxonomic hierarchy should primarily reflect degrees of morphological rather than molecular divergence. Incongruence among contrasting data sets is best explained by understanding the biological constraints operating on each type of phylogenetic information. The conventional 'uniformitarian' view of evolution has only limited applicability as one traces the history of land plants through time. Diversity increased in stepwise fashion, reflecting either attainments of complexity and/or fitness thresholds by the lineage (intrinsic) or the availability of unusually permissive environments, often following major perturbations (extrinsic). The Quaternary period demonstrates especially well the resilience, and ease of migration, of the Earth's vegetation. A higher frequency of generation of novel phenotypes in the deep past is possible, but a far higher frequency of their establishment is certain; together, these factors generate an evolutionary pattern of nested radiations that is fractal, as saturation of the resource space rendered the environment decreasingly permissive through time. In the immediate future, evolutionary–developmental genetics will have increasing value for testing homology, interpreting homoplasy and elucidating evolutionary constraints, and will become easier to pursue as whole-genome sequences of additional 'model' species further invigorate comparative genomics. Complexity of gene regulation, both by other genes and by the cellular and extra-cellular environment, appears a particularly fruitful area for further research. Nonetheless, environmental filtering of evolutionary novelties (whether instantaneously isolated mutant 'prospecies' or classic neoDarwinian 'selfish genes' selectively spreading through panmictic populations) can only be effectively understood by longer term monitoring of populations in the wild, to better capture rare evolutionary and ecological events and to better assess the efficacy of traditional microevolutionary processes. We believe

*Corresponding author
(email: R.Bateman@nhm.ac.uk).

that the resulting renaissance in macroevolutionary studies will encourage a broader systematic perspective – one that better encompasses the remarkable diversity of evolutionary processes that together generated the present diversity of life.

Key words adaptation, cladistics, co-evolution, DNA sequencing, ecology, evolutionary–developmental genetics, key innovation, macroevolution, migration, molecular clock, morphology, neobotany, non-adaptation, palaeobotany, phylogenetics, punctuated equilibrium, radiation, supraDarwinism, uniformitarianism

> Evolutionary biology today has an immensely broader perspective than 30 years ago; ... we recognize stasis, constraints, multiple levels of selection, differential clade diversification, and historical contingency as valid principles worthy of research.
>
> [But] because the chief implication of punctuated equilibrium is an autonomous macroevolutionary theory of trends, ... perhaps we should pose the empirical question of just how common trends are. If they can only be rarely discerned amidst the thickets of adaptive radiation, a selection theory beyond the neoDarwinian may prove more needed in principle than in practice.
>
> Futuyma (2002, pp. 662–663), reviewing S.J. Gould (2002), *The Structure of Evolutionary Theory*

Introduction

Having recently considered land-plant evolution from an explicitly forward-looking perspective that emphasized the historical continuity of palaeobotanical data (DiMichele & Bateman, 2003), we happily acknowledge the key role played by the fossil record in prompting many of the concepts discussed in this review. However, here we have deliberately chosen to reverse our temporal polarity, extrapolating backward from neobotanical data obtained from today's flora and using as our focal theme evolutionary phylogenetics.

Specifically, we consider the now widespread use of parsimony-based cladograms to identify and delimit clades (each being a group of species that is inferred to be monophyletic, incorporating a hypothetical ancestral species and all of its descendants), which have become the common currency of modern biodiversity studies. In doing so we make the philosophically questionable but pragmatically necessary assumption that a cladogram can be interpreted as a phylogeny. This assumption allows us to integrate phylogenetic information with recent insights into the modes of origin, and of subsequent ecological filtration, of phenotypic novelties. Our objective is not only to further our advocacy of a plurality of evolutionary mechanisms but also to speculatively identify research areas that could engender breakthroughs in our understanding of the genesis of biodiversity during the next decade.

Review of existing data on present vascular plant diversity

Neobotanists studying the present terrestrial flora are by definition dealing with a transient snapshot of history. The 90–95% of vascular plant species estimated to have gone extinct since they first appeared in the fossil record 420 my ago are beyond the reach of neobotany, as are most potential examples of ancestor–descendant relationships. Also, the ability to approximately date the first occurrences of particular clades (or, more precisely, of particular morphological synapomorphies: the all-important derived character states delimiting clades) using molecular methods is precluded, other than via decidedly unreliable molecular clocks based on relatively few extant descendants.

However, these deficiencies in the extant flora are more than compensated for by the potential for near-complete sampling of this one snapshot of the Earth's biota, rather than relying on the vicissitudes of preservation in the fossil record (Donovan & Paul, 1998); modern species-level sampling is made incomplete only by recent man-induced extinctions. More importantly, living plants are almost infinitely testable; this contrast with inert fossil plants is made even more stark by repeated failures to replicate early 'discoveries' of DNA in *c.* 18 my-old magnoliid leaves still retaining their green coloration (cf. Golenberg *et al.*, 1990; Austin *et al.*, 1997; Golenberg, 1999). Thus, the fossil record offers strong and at least approximately datable patterns of diversification and extinction, whereas the living flora allows process-based interpretations of particular plants ('experimental' in the broad sense) that can then become the basis of evolutionary and/or ecological speculations projected backward through time.

With the arguable exception of vertebrates, vascular plants are the most completely documented extant major clade; widely cited morphological extrapolations suggest that at least 90% of extant species have already been named, though more recent DNA studies are uncovering morphologically cryptic species at a rate that suggests this figure may underestimate the disparity between the number of phenotypically delimited species that reflect morphology only and the greater number of genotypically delimited species that largely reflect DNA

sequences. For example, a recent sequencing study of all fungi associated with the roots of a single plant of the grass *Arrhenatherum* suggested that only 15% of the 49 putative species identified had previously been sequenced (Van den Koornhuyse *et al.*, 2002). Moreover, the proportion of extant species of both plants and fungi whose biology and genetics are genuinely well-understood, and that consequently constitute prospective 'model organisms', remains minuscule.

One enormous advance over the past two decades has been the development of taxonomically broad phylogenies depicting the relationships among species, initially based entirely on morphological characters (many mixing extant and extinct representatives) but more recently dominated by DNA sequences (thus encompassing extant representatives only). Together, these studies have provided a strong interpretative framework, and increasingly monophyletic classifications, for land plants. They demonstrate that the estimated 270 000 extant species (Walter & Gillett, 1998) are apportioned very unevenly among major clades, the largest groups being angiosperms (231 000: dominantly eudicots), ferns (9000: dominantly derived leptosporangiate families), and, to a lesser degree, conifers (600: mostly Pinaceae–Cupressaceae and Podocarpaceae). Nonetheless, we are fortunate in having within the extant flora some low-diversity relictual lineages representing taxonomic classes and orders that are phylogenetically pivotal (and formerly were ecologically pivotal). Notable examples include Ginkgoales and Gnetales among the 'gymnosperms' (paraphyletic to the angiosperms: Chaw *et al.*, 2000) and Psilotales, Equisetales, Marattiales and Isoetales among 'pteridophytes' (paraphyletic to the seed-plants: Pryer *et al.*, 2001). However, many groups that were ecologically dominant in the past have no close living descendants, particularly among the less derived members of notoriously paraphyletic groups such as the pteridophytes and the gymnosperms.

Revisiting phylogenetic conventions

It has become increasingly popular in cladistic circles (e.g. Smith, 1994; Kitching *et al.*, 1999) to distinguish between a 'crown group' of extant species plus any extinct species nested within it (e.g. species A–H, including the extinct species D and G, in Fig. 1) and a paraphyletic 'stem group' of related but extinct species (species I and J in Fig. 1). In morphological phylogenetic analyses the 'stem group' often provides one or more taxa selected by the analyst as outgroups, since extinct species are judged on average more likely than extant species to exhibit suites of plesiomorphic character states (species J in Fig. 1). In our view, this conventional distinction between crown group and stem group is meaningless, serving primarily to render fossil species second-class citizens in the mind of the reader. It places undue emphasis on the most deeply divergent of the extant species under consideration (species H in Fig. 1), and is in any case an arbitrary relative concept, since species D could be viewed as the stem group to crown group A–C in Fig. 1. The key difference between extinct and extant species is practical rather than philosophical; specifically, the

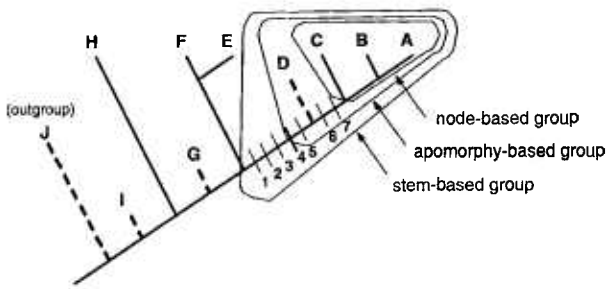

Figure 1 Hypothetical phylogeny of ten species, six extant (A, B, C, E, F, H: solid lines) and four extinct (D, G, I, plus outgroup J: dashed lines). The hypothetical ancestor of group A–F is separated from that of group A–C by seven apomorphies (1–7). See text for further explanation.

far greater testability of the living than the dead. Examples of such tests include the ability to directly observe both ontogeny and 'behaviour', to distinguish ecophenotypic from genuinely heritable influences on phenotype, to transfer genes among individuals in an attempt to better understand their function, and to explore crucial aspects of gene expression such as pleiotropy and epigenetics (Bateman *et al.*, 1998; Bateman & DiMichele, 2002).

Similarly, conventional wisdom is correct to say that a cladogram cannot be used to demonstrate ancestor–descendant relationships, but in most cases it is reasonable to assume that an ancestor and its immediate descendant will, if sampled, be perceived phylogenetically as sisters, unless the speciation event occurred by combining the genomes (reticulation: Fig. 2b) of non-sister species. There is much merit in contemplating what mode of speciation would be consistent with a particular pattern of morphological, ecological and genetic change summarized by a cladogram (Fig. 2), particularly as the more informed debates on evolutionary mechanisms are gradually becoming less qualitative and more quantitative. Specifically, evolutionary biologists are less inclined to ask whether the many different theoretical modes of evolution (e.g. adaptation, drift, shifting balance, dichotomous and reticulate saltation: Bateman & DiMichele, 1994*a*, 2002) occur at all, but instead are more likely to assess what proportion of the surprisingly small total of well-founded case-studies support each of these different modes.

Another phylogenetic convention also in practice incorporates the 'tyranny of the Present'. It is rarely explicitly noted that paraphyletic groups were monophyletic until the origination of the more derived group nested within them (e.g. gymnosperms were monophyletic during the Devono-Carboniferous: DiMichele & Bateman, 2003), and polyphyletic groups were monophyletic until a particular biologically correlated set of apomorphies had evolved in more than one lineage, conferring superficial similarity between those lineages (cf. Brummitt, 2002). Good examples include the similar patterns in the loss of photosynthetic function and associated vegetative simplification evident in the repeated origins of parasitic and mycoheterotrophic ('saprophytic') syndromes among angiosperms (Nickrent *et al.*, 1998) and repeated origins of localized autogamous species from within a more widespread allogamous

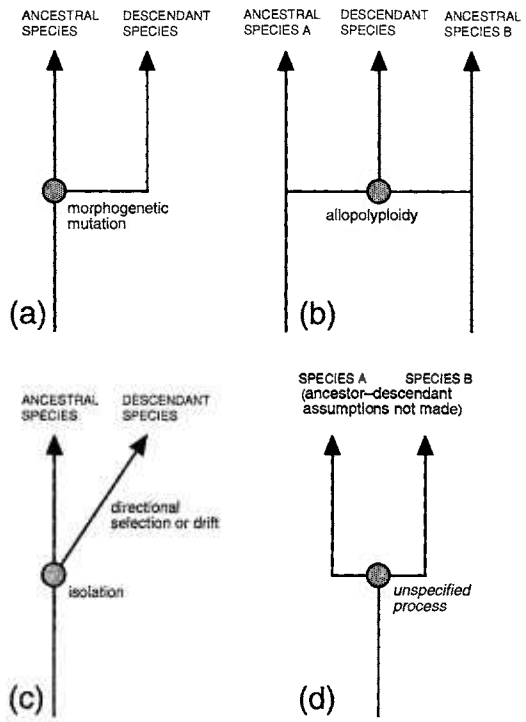

Figure 2 Comparison of contrasting evolutionary patterns. Dichotomous saltation (a) shows instantaneous divergence of descendant from ancestor via a mutation in a key gene controlling morphogenesis. Reticulate saltation (b) also occurs instantaneously via allopolyploidy but by definition blends two parental genomes. Directional selection or drift (c) cause gradual divergence of the descendant following its isolation. Cladistic representations of speciation (d) traditionally assume no ancestor–descendant relationships and specify no underlying process (after Bateman & DiMichele, 2002, fig. 1).

species (cf. Takebayashi & Morrell, 2001; Hollingsworth *et al.*, 2003).

Extending this theme, the current revolution in plant evolutionary–developmental genetics ('evo-devo': Cronk, 2001; Cronk *et al.*, 2002) is helping to untangle many previously intractable problems of phenotypic homoplasy (multiple transitions between the same pair of states for a single character). Specifically, comparative genetic investigations of extant species can at a causal level discriminate among apparently similar apomorphies in different lineages; even a parallel loss of a feature or suite of features can be distinguished as two separate evolutionary events if the precise genetic change underpinning that loss differs between the lineages under comparison (cf. Bateman, 1996). Studies of extant species and extinct species can both be undermined by insufficiently holistic approaches. An increasing proportion of phylogenetic studies of extant species pays no more than lip-service to morphological data, which are now on average much more time-consuming to gather than molecular phylogenetic data (Bateman, 1999); the temptation is always to 'borrow' a few widely used characters from the existing literature. For their part, many palaeontologists have an understandable habit of viewing characteristic assemblages of fossil organs as assemblages of morphological character states, rather than attempting to painstakingly reconstruct dissociated organs into specific conceptual whole plants. Although whole-plant reconstruction is not currently favoured by funding bodies, it is an essential prerequisite to identify a more narrowly delimited set of phenotypic features that is likely to reflect the expression of an integrated genome representing a single biological species.

Although the many polemical attacks on the frequent use of cladograms to 'prove' evolutionary hypotheses have some theoretical justification, phylogenies are nonetheless invaluable for inferring evolutionary process if viewed in the light of one or more of the many modes of 'disproof' – that is, of falsification. Demonstrating that apparently functionally connected sets of character states did not appear simultaneously in the same lineage disproves (or at least reduces the scope of) saltational evolutionary hypotheses (Bateman & DiMichele, 1994*a*). Similarly, demonstrating that a particular character state and putatively correlated biological function did not appear simultaneously in the same lineage disproves adaptational hypotheses (Donoghue, 1989). Moreover, phylogenies of extant species that combine morphological and developmental genetic data have the potential to refute or support putative cause (genetics or epigenetics) and effect (phenotype) in specific hypotheses of heterochrony and heterotopy (Rudall & Bateman, 2002).

Phylogenetic insights into diversification: radiations and key innovations

As noted by Bateman *et al.* (1998), evolutionary radiations pose three primary challenges: satisfactorily defining a radiation; distinguishing among the 'unholy trinity' of clade origination, radiation and migration; and identifying the underlying cause(s) of the radiation.

There are many definitions of radiations, despite the fact that most authors use no explicit definition. Bateman (1999) sought a definition independent of any specific evolutionary process, arguing that the commonplace manacling of 'adaptive' to 'radiation' is both unjustified and highly prejudicial to any subsequent attempts at causal interpretation. He concluded that a radiation can be defined most effectively using one or both of two properties: species diversity and phenotypic character diversity, the latter being more strongly positively correlated with higher taxonomic diversity than with species-level diversity. Using either criterion, the best measure of a radiation is the net surfeit of rate of gain over rate of loss in a specified clade during a specified time interval. Obviously, the practical applicability of this definition depends on our ability to at least approximately date evolutionary events (see 'Dating and interpreting major diversification events'), and on the existence of a critical mass of both genes and phenotypic expressions of those genes; in other words, on attaining a minimal degree of genetic, morphological and functional complexity (Bateman *et al.*, 1998). These data similarly allow us to define a catastrophic

decline in diversity. Such an event, essentially the converse of a radiation, is often described as a mass extinction, though this term also is rarely adequately defined.

Once radiations (and mass extinctions) are effectively and consistently defined, different kinds of radiations can be distinguished (Erwin, 1992). For example, the increase in species number during the early Devonian land-plant radiation is proportionally great, but the total numbers involved are small relative to the concomitant radical increase in phenotypic diversity (measuring phenotypic diversity among fossils, especially disarticulated plant fossils, is of course challenging: cf. Wagner, 1996: Bateman *et al.*, 1998; Southwood & Henderson, 2000; DiMichele *et al.*, 2001*b*; Wills, 2001). Rapid, highly divergent increases in phenotypic (and thus presumably genotypic) complexity accompanied by relatively low speciation rates constitute a novelty radiation *sensu* Erwin (1992). The resulting coarse niche differentiation contrasts strongly with a much finer niche-filling adaptive radiation, which also occurs within a single clade but typically involves slower and less profound phenotypic diversification and much greater species diversification. Note that this contrast between arguably the two most interesting modes of radiation relies on decoupling of phylogenetic disparity and species diversity, the latter correlating more closely with the degree of niche differentiation.

Another current philosophical debate in phylogeny has practical relevance to this discussion. The controversial 'Phylocode' (cf. Benton, 2000; Cantino & de Queiroz, 2000; Nixon & Carpenter, 2000; Brummitt, 2002; Bryant & Cantino, 2002; Forey, 2002) seeks to establish formal, node-based definitions of monophyletic groups as an alternative to the long-established Linnaean hierarchy. Although we have yet to be convinced of the overall merit of this classificatory scheme, its advent has usefully rekindled debates regarding the best protocol for delimiting monophyletic groups in cladograms. The three alternatives, based on nodes, stems or apomorphies, are summarized with regard to the status of fictitious group A–D in Fig. 1. A node-based group arbitrarily excludes all seven character-state transitions occurring on the branch immediately subtending the most recent common ancestor of the group A–D, whereas a stem-based group equally indiscriminately includes all those transitions.

In contrast, an apomorphy-based concept forces the systematist to define the derived group on the basis of a single apomorphy considered by the systematist to be most diagnostic; for example, one could use double fertilization leading to a functional triploid endosperm to delimit the angiosperms (e.g. selecting apomorphy 4 to delimit group A–D in Fig. 1). If the group is to be readily identified, that apomorphy must be expressed phenotypically, though of course that chosen character state carries much greater conviction if the genetic transition that led to its first appearance can also be determined. This approach negates two controversies: whether to include in phylogenetic analyses fossil species (e.g. species D in Fig. 1), and whether to give the node subtending a fossil species equal standing with a node subtending an extant species, since the answer to both questions must by definition be affirmative in an apomorphy-based classification. In the hypothetical example of Fig. 1, fossil species D performs a vital role in demonstrating that apomorphies 1–5 arose in the lineage before apomorphies 6–7. Most importantly, the apomorphy-based approach allows us to distinguish between the point of origin of a clade and any subsequent radiation within that clade, an especially crucial distinction for any biologist seeking ever-popular 'key innovations' (which may or may not include the character state chosen by the systematist as being most diagnostic of the clade).

Most authors studying putative radiations of specific taxonomic groups have identified one or more supposed key innovations in an attempt to explain the evolutionary and/or ecological 'success' of the group, however that success is assessed. Recognizing that it is another vital term subject to many definitions, we view a key innovation as a synapomorphy that was acquired immediately prior to a *bona fide* evolutionary radiation and that can be shown to have been a far greater stimulus to that radiation than any other synapomorphies acquired on the same phylogenetic branch (Sanderson & Donoghue, 1996; Bateman *et al.*, 1998; Bateman, 1999; Schluter, 2000). This is extremely difficult to demonstrate in practice, as it requires access to a far wider range of evolutionary and/or environmental information than that explicitly incorporated in the actual phylogenetic analysis. As in many major issues addressed through the fossil record, falsification of a hypothesis is a weaker but more practical test to apply when using fossil data to explore radiations (P. Kenrick, pers. comm., 2002).

We will now briefly consider the three large-scale radiations most familiar to us: the primary radiation of land plants in the mid-Palaeozoic (Bateman *et al.*, 1998), the late Palaeozoic radiation of the rhizomorphic lycopsids (DiMichele *et al.*, 1992; Phillips & DiMichele, 1992) and the largely Cenozoic radiation of the Orchidaceae (Chase, 2001; Rudall & Bateman, 2002). The overall impression gained from each of these case studies is that no single character can be accused of having engendered that particular radiation. Even when attempts are made to tease apart each radiation into a nested sequence of smaller-scale radiations, key innovations still cannot be readily identified. Rather, it seems likely that a critical mass of functional phenotypic characters accumulated in each clade (or, in the case of the primary terrestrial radiation, in several clades), eventually offering sufficient phenotypic and 'behavioural' flexibility to define and divide many niches. Only in the case of smaller-scale, more recent radiations are testable key innovations likely to be identified with any confidence (Baldwin *et al.*, 1998; Carroll *et al.*, 2001; Richardson *et al.*, 2001*a*). Success is most likely when the system under scrutiny is a relatively closed 'island', a term used here to cover not only the classic cases of oceanic islands but also for example mountain-tops, lakes and discrete igneous or metamorphic hard-rock bodies generating physiologically challenging soils.

Also lacking has been the null hypothesis of equally careful studies of groups that have failed to radiate (Bateman & DiMichele, 2002). For example, the 12 orchid species currently recognized in the Macaronesian islands represent 11 genera, and each species has a close analogue in the

adjacent regions of the Mediterranean and/or north-west Africa. Thus, the Macaronesian species apparently reflect anagenesis: gradual speciation without post-immigration dichotomy of the lineage. Indeed, the majority of immigrants to oceanic islands located over volcanic hot-spots have failed to dichotomously speciate, let alone radiate (Baldwin *et al.*, 1998), suggesting that fresh terrain and low competition is a prerequisite for, but is by no means always a stimulus of, radical innovation in form. In the case of the Macaronesian orchids, the potential for post-immigration radiation is probably constrained not only by the need for fidelity from associated pollinating insects but also from at least two cohorts of mycorrhizal fungi: the first is essential to initiate germination of the minute embryos in the airborne seeds and the second is neccesary to supply nutrition to the mature tubers or rhizomes. Such interpretations reinforce the commonly held view that co-evolution has been a major driver of species-level diversity in angiosperms (Crane, 1989; Behrensmeyer *et al.*, 1992; Dilcher, 2000).

Dating and interpreting major diversification events: roles of morphological and molecular data

Defining radiations on the basis of relative rates of change, most commonly via calculations of rates of birth and death (strictly, mortality and natality) of species, obviously requires an effective measure of time. Setting aside laboratory-based breeding experiments, which involve overly simple organisms, overly simple ecosystems and observations across too few generations, there are two viable methods of assessing biodiversity levels through time: direct fossil dating and molecular clocks.

First, we can simply document first and last appearances of species in the fossil record. Although a valuable exercise, this approach is fraught with difficulties. Both the fossil record and the fossil plants themselves are highly fragmentary. Fragmentation of the plants means that only the most skilfully reconstructed whole-plant species can be compared confidently with their modern descendants in a morphological phylogenetic analysis. The temporally and geographically sporadic nature of the terrestrial fossil record means that the observed origin and extinction dates of a species are liable to incur considerable statistical errors. Such temporal resolution will be especially inadequate if (as is argued here) most major evolutionary radiations are extremely rapid. This realization led some authors to argue that the timing of speciation events may be more accurately predicted indirectly, via cladograms, than observed directly in the fossil record (Smith, 1994).

Together with the recent proliferation of molecular data, this has prompted the development (and, more recently, the widespread implementation) of molecular clock techniques, inevitably transferring the focus from extinct to extant species capable of yielding usable strands of DNA. In theory, once the preferred most-parsimonious tree based on nucleic acid sequences has been identified, we can determine the number of base substitutions separating each speciation event as a measure of molecular disparity. We must in addition date at least one, and preferably at least two, branch points in order to relate the divergence point to the present-day time-line. Although this can be done by approximately dating the separation or collision of continents, or the emergence of oceanic islands, most commonly the chosen reference datum is the perceived point of origin of a relevant species or clade in the fossil record, thus effectively appealing once again to the 'first appearance' method described above. We can then interpolate dates of the remaining branch points in the phylogeny by assuming a constant rate of base substitution for the targeted fragment of genome throughout the evolutionary history of the clade under scrutiny.

However, the burgeoning database of molecular phylogenies has made it abundantly clear that the molecular clock is also a rash assumption, subject to many modifying factors (Avise, 1994; Page & Holmes, 1998). Thus, statistical tests for adjusting transition rates within molecular phylogenies have recently been developed, most popularly non-parametric rate smoothing (Sanderson, 1997; Sanderson & Doyle, 2001; Wikström *et al.*, 2001). Two important conclusions can be drawn from applications of this method. First, the error bars on specific origination dates for specific clades are disconcertingly broad; they are on average far greater than those associated with first appearances in the fossil record (at least, in those groups that actually *possess* a fossil record). Second, it has become clear that the longevity of a major clade cannot be equated with the longevity of its component species.

For example, the extant genus *Huperzia* (Lycopodiaceae) closely resembles in morphology the well-characterized 420 my old fossil *Asteroxylon*, and its divergence from its sister genus, *Lycopodium*, has been dated as *c.* 320 Ma (Wikström & Kenrick, 2001). However, the considerable species-level diversity of *Huperzia* (manifested mainly as subtly differing tropical epiphytes) was dated at an estimated 80–100 Ma by Wikström & Kenrick (2001) and, given the negligible sequence divergence among the species, it could in fact be much more recent. Similarly, a more derived line of lycopsids, the isoetaleans *s.s.*, dates back at least 330 Ma (DiMichele & Bateman, 2002) but molecular divergence among extant species of *Isoetes* is remarkably low (Rydin & Wikström, 2002). Moving on to more recent higher taxa, the European terrestrial orchid genera *Ophrys* and *Serapias* are subtended by the longest rDNA branches of any genera of the subtribe Orchidinae, yet their many constituent species have by far the shortest mean infrageneric divergence (Bateman, 2001). This suggests a decoupling of rapid species-level turnover (which probably most commonly represents adaptive variations on a highly phenotypically constrained theme) from the more profound phenotypic shifts that delimit higher taxa (Bateman & DiMichele, 1994*a*, 2002; Kellogg, 2002).

Another interesting pattern commonly evident (though rarely commented on) in molecular phylogenies is that, irrespective of the size and taxonomic rank of the putative clade under scrutiny, they tend to resolve most effectively (i.e. yield stronger bootstrap support values for) intermediate taxonomic levels rather than relatively deep or relatively shallow

divergences. At present, the most common response to this conundrum is to attempt to resolve the weaker intermediate branches by sequencing additional (and often disparate) regions of the genome and then, in most cases, generating a single set of cladograms from the amalgamated DNA data. In effect, the 'tyranny of the Present' is compounded by the 'tyranny of the bootstrap value'. This 'kitchen sink' approach squanders the opportunity to contrast phylogenetic signals from different parts of the plant genome (e.g. nuclear versus plastid versus mitochondrial, coding versus non-coding, physiologically expressed versus morphologically expressed, upstream versus downstream), where incongruences might be explained in terms of specific molecular processes. It also side-steps the issue of whether weak phylogenetic resolution in intermediate portions of trees may be a genuine by-product of specific modes of evolution.

In this context, the occurrence of, and mechanism(s) underlying, evolutionary radiations are especially relevant. Bateman (1999) argued that morphology describes a genuinely punctuational pattern through the fossil record, changing rapidly during speciation events irrespective of underlying mechanism (spanning a few tens or hundreds of generations for directional or disruptive selection through to a single generation for saltation *sensu stricto*), but then fluctuating only slightly in each lineage through time, constrained by stabilizing selection. By contrast, the non-expressed or physiologically expressed gene regions that are preferentially employed for molecular phylogenetics have no influence over an organism's morphology, which is more likely to constitute the front line of fitness-mediated selection. Hence, Bateman (1999) concluded that if morphological evolution does indeed describe a punctuational pattern, but molecular evolution describes a far more gradual pattern, then the vast majority of morphological character-state transitions occur *during* evolutionary events and the vast majority of molecular character-state transitions occur *between* them (cf. Chase *et al.*, 2000; Bateman & DiMichele, 2002).

Moreover, slowly mutating regions of the genome are likely to fail to register a radiation at all, unless the observer successfully targets a specific genetic change that generated a key innovation and thus prompted the radiation. Rapidly mutating regions are soon over-saturated if the radiation occurred deep in time, and like slowly mutating regions are unable to capture a radiation that is very recent (Harris *et al.*, 2000; Bateman, 2001; Richardson *et al.*, 2001*b*). In other words, the absence of a morphological clock gives the phenotype a clear advantage over the genotype in resolving *bona fide* radiations, by weight of phylogenetic evidence.

The one approach that could ultimately bring greater evolutionary meaning to DNA sequence data is to identify the precise substitution or insertion–deletion event in the key gene that actually caused a specific morphological transition – a realistic expected outcome of the 'evo-devo' revolution over the next decade (Cronk *et al.*, 2002). Even at this new level of enquiry, rates of change will differ radically among taxa. For example, recent studies show that chromosomal rearrangements and gene acquisitions occur 2000 times more frequently in free-living bacteria than in their obligately endosymbiotic brethren (Tamas *et al.*, 2002), which are insulated from the vicissitudes of the external environment.

Diversification versus migration as factors determining plant distributions

We now turn to more strictly ecological aspects of vascular plant diversification, considering briefly the origin of today's biomes from 300 Ma to the Present.

It seems likely that the gymnosperm-dominated Permian to mid-Cretaceous flora owed its origins to the accumulation of a critical mass of functional phenotypic characters in several clades, typically while they occupied relatively low-competition, extra-basinal habitats. Sufficient phenotypic diversity and flexibility accumulated to define and divide many niches, generating the threshold number of niches necessary to form communities that exhibited broadly modern ecological dynamics (if not modern species diversity: DiMichele & Bateman, 2003). The subsequent, angiosperm-dominated flora of the Cenozoic then further refined and partitioned those niches, permitting a quantum leap in species-level diversity but adding relatively little to the larger-scale repertoire of roles and functions.

The broad patterns evident in today's flora when species-level plant diversity is mapped geographically are well-known; if major hot deserts are factored out of the equation, species density (i.e. species per km^2) increases steeply from the poles to the equator (O'Brien *et al.*, 1998, 2000), where the three-dimensionality of forests and mesoclimates conducive to epiphytes together encourage even finer niche partitioning. In islands, both oceanic (e.g. mid-ocean hot-spots) and topographic (e.g. isolated mountains and lakes), taxonomic diversity shows a positive correlation with areal extent (Rosenzweig, 1995) and range of altitude (or, in the case of water bodies, of depth). Total percentages of narrow endemics are negatively correlated with ease of inward migration, which in turn reflects both geographic and climatic factors.

Both the actual fossil evidence and evolutionary–ecological theory suggest that biotic diversity increases through time in a stepwise fashion (Valentine, 1980; Vermeij, 1987). Within particular clades, the periods of more rapid diversification may ultimately reflect an unusually permissive environment (e.g. recovery from a widespread environmental perturbation) or the attainment of a fitness threshold (e.g. the acquisition and (epi)genetic stabilization of a genuine key innovation). Where multiple clades diversify simultaneously, a co-evolutionary explanation could be invoked, but if there is no obvious ecological interaction among the clades an environmental stimulus is inevitably suspected.

However, our especially strong palaeobotanical record for the Quaternary period indicates that global vegetation has acquired remarkable resilience in the face of major environmental perturbations, reflecting well-documented profound fluctuations in global climate (Bennett, 1997; Huntley, 1999; Taberlet & Cheddadi, 2002) and, more recently, the depredations of mankind (Parmesan *et al.*, 2001). Although ecological

succession in a particular area varies considerably according to local environmental factors, even given similar initial conditions (Hughes & Dumayne-Peaty, 2002), true global extinctions in plant species (as opposed to local or regional 'extinctions', better termed extirpations) appear surprisingly uncommon.

Rather, in the temperate zone, rapid migrations are well-documented in both the pollen record and the residual levels of genetic diversity evident along probable migration routes. For example, most of the British flora has colonized from mainland Europe since the last major (Devensian) glaciation. Although the ice-sheet melted rapidly about 14 700 years ago, western Europe suffered a further brief period of periglacial conditions during the Loch Lomond Stadial (= Younger Dryas: Hughen *et al.*, 2000) of *c.* 13 000–11 700 years ago, probably reflecting temporary diversion northward of the warm oceanic currents of the North Atlantic Drift (cf. Visbeck, 2002). There has been much debate as to whether the current global climate is inter-glacial or post-glacial, the probability of the latter perhaps being greatly increased by anthropogenic greenhouse gases (Parmesan *et al.*, 2001). The post-Younger Dryas sequence of (re-)colonization of the more common species of the present flora has been well-documented in fossil pollen sequences (Huntley & Birks, 1983; Birks, 1989). Moreover, the dominantly northward routes taken by those species to reach Britain are clearly evident in surveys of plant genetic diversity across Europe, which reliably show a repeated pattern of three diversity 'hot-spots' in trans-Mediterranean migration routes through Iberia, Italy and the Balkans (Hewitt, 1996; Ferris *et al.*, 1999). Surprisingly little attention has been paid to the likelihood of 'post-glacial' speciation, though credible case-studies exist (e.g. Bateman & DiMichele, 2002).

More recently, major controversies have intensified regarding the degree of ice-age disturbance experienced by equatorial zone tropical forests, most notably the especially species-rich communities in Amazonia where a substantial number of closely related species now occur sympatrically (cf. Colinvaux & De Oliviera, 2001; Richardson *et al.*, 2001*b*). The main competing theories have been conveniently summarized as the 'museum' and 'cradle' hypotheses, the former stating that the diversity reflects steady production of co-adapted assemblages of species over many tens of millions of years under stable environmental conditions, the latter arguing that much of the diversity has arisen during the last few million years and may reflect increased rates of speciation in smaller populations temporarily retreating to refugia during glacial periods. The surprisingly limited molecular phylogenetic evidence accumulated thus far offers stronger support to recent, rapid radiations (Richardson *et al.*, 2001*b*; N. C. Garwood & T. Parrish, pers. comm., 2002), suggesting that at least the recent history of global vegetation has been strikingly dynamic. Earlier parts of the plant fossil record that offer opportunities to test this hypothesis similarly suggest relatively rapid changes, notably the 'great drying' approximating the end of the Carboniferous Period (also a story of polar glaciations: DiMichele *et al.*, 2001*a*; Rees, 2002) and the reassembly of tropical rain forests in Colorado a mere 1.4 my after the devastation wrought by the K–T boundary asteroid impacts (Johnson & Ellis, 2002).

Hybridization in general, and allopolyploidy in particular, have also played important roles in this botanical diversification (Rieseberg, 1997), though of course gene flow between populations also constitutes one of the main barriers to speciation in plants (Arnold, 1997) and is also increasingly seen as important in animals (Grant & Grant, 2002).

Tracing plant evolution back through time

The initial supraDarwinian perspective

In a review of supraDarwinian (put simply, non-selective) evolutionary mechanisms in plants written 10 years ago (Bateman & DiMichele, 1994*a*), we offered the following hypothetical account of macroevolution and its consequences through time.

Uniformitarianism – the constancy of processes through time – has long been a fundamental principle of geology. Most physical contraventions of the principle, such as long-term changes in atmospheric composition and in terrestrial weathering rates, can be attributed largely to non-uniformitarianism in the Earth's biota – the former prompted primarily by the evolution of photosynthesis, the latter by the advent of soil-binding roots. Many neontologists apply uniformitarianism to evolutionary theory, often unconsciously and without serious consideration of the vast time-span that is at their disposal; in other words, periods of time that can be measured in generations are more tangible to humans than the approximately 420 my of vascular plant evolution (Kenrick & Crane, 1997; Bateman *et al.*, 1998).

Indeed, uniformitarianism is generally applicable to evolution at the molecular level – the generation of variation through mutation and the subsequent history of that genetic variation. However, if one could trace plant phylogeny back through geological time, there can be little doubt that the pool from which that variation is drawn would change, showing an overall decrease in genotypic diversity and a decrease in the phenotypic complexity of the most derived clades present in each time-slice. Reproduction would become on average increasingly simple: first the seed habit would be lost, then heterospory, then heterothally (Bateman & DiMichele, 1994*b*). Self-fertilization would become more common, increasing the probability that non-lethal mutations will persist. At some point early in the history of land plants, one would encounter early 'pre-vascular' plants, such as *Aglaophyton* and *Cooksonia*, that appear to have possessed near-isomorphic, independent haploid gametophytic and diploid sporophytic generations (Kenrick & Crane, 1997). The haploid phase of the life history may have been as long as the diploid, offering a far greater probability of incurring mutations unbuffered by the continued presence of a second, non-mutant copy of the allele.

This reproductive simplification back through time undoubtedly followed a stepwise pattern rather than continuously declining in sophistication; there were long periods of relative stability with little meaningful change. Nonetheless, the overall trend is paralleled by developmental simplification.

Although the number of key developmental genes is modest even in the more derived extant clades (Cronk *et al.*, 2002), it was probably still fewer in primitive fossil groups such as the 'rhyniophytes', though the transitions from gymnosperms to the more derived groups within angiosperms required loss-of-function mutations as well as gene duplications in genes controlling key reproductive characters (Albert *et al.*, 2002; Cubas, 2002; Frohlich, 2002; Theißen *et al.*, 2002). In many cases, the relevant genes would have been duplicated *en masse* through polyploidy. Bateman & DiMichele (1994*a*) argued that gene diversification would decrease the average phenotypic effect of any mutation to a developmental gene, by increasing the probabilities (*a*) of utilizing alternative developmental pathways and (*b*) of epigenetic readjustment being sufficient to compensate for the genomic change (admittedly, both assertions remain highly speculative).

However, the most profound contraventions of biological uniformitarianism are ecological. Neobotanical ecology inevitably focuses heavily on angiosperms, tacitly acknowledging their remarkable species-level diversity. Conventional wisdom argues that this diversity largely reflects unusually fine niche partitioning. This in turn reflects unusually high frequencies of intimate co-evolutionary relationships (e.g. with pollinators, herbivores and mycorrhizae), rather than the greater genomic 'rigidity' of non-angiosperms invoked by some authors (e.g. van Steenis, 1976). However, passing back through time, the small-scale niche partitioning of the angiosperms would gradually give way to less diverse communities, with a sharp break evident between angiosperm-dominated assemblages and those dominated by non-angiosperms (Behrensmeyer *et al.*, 1992). Individual interspecific evolutionary and ecological links may remain strong, but the potential total number of links is inevitably greatly reduced.

Thus, we speculate that during one's journey back from the Present to the mid- to late Devonian, one would observe stepwise changes in the degree of opportunity that existed for partitioning existing niches and for increasing connectivity among the species occupying those niches, the overall trend being one of average increase. Moreover, many more niches (perhaps better termed evolutionary–ecological opportunities), were vacant at any one moment in time, and this vacancy extended to entire habitats. Opportunities for temporary release from stabilizing selection, facilitating the establishment of novel phenotypes, were far more common. In short, a higher frequency of generation of novel phenotypes in the past is possible, but a far higher frequency of their establishment is certain. The further back in time one travelled, the greater would be the significance of non-adaptive evolutionary mechanisms relative to adaptive selection, which has probably reached its maximal intensity to date, albeit still operating alongside many other contrasting modes of evolution (Bateman & DiMichele, 1994*a*, 2002).

This observation encourages one additional prediction, namely that evolutionary patterns through time were fractal, each successive radiation within a lineage leading to smaller relative degrees of phenotypic divergence than the preceding radiation. Once plants had invaded the land and a basic toolkit of terrestrial adaptations had evolved in the Superdivision Polysporangiomorpha (the traditional Division Tracheophyta plus *Aglaophyton* and *Cooksonia*: see Kenrick & Crane, 1997), taxa that we traditionally classify as classes and orders, at least partly on the basis of our perception that they show greater overall phenotypic dissimilarity, should appear earlier in the fossil record than co-ordinal families, and families should appear earlier than co-familial genera. Even if one viewed higher taxa as an artefact of the desire for nested hierarchies, the synchroneity of appearance of two or more higher taxa should have some biological significance. For example, if the controversial decision is taken to treat the angiosperms as an order of the Class Gymnospermopsida (rather than, for example, accepting the artificially inflated Superclass Magnolidra of Kenrick & Crane, 1997), it can be argued that all eutracheophyte classes originated over a remarkably short period of approximately 70 my during the Late Silurian and Devonian. Thus, we can readily identify an apparent 'Golden Age' of plant macroevolution, which can be tested further using measures of phenotypic and/or phylogenetic disparity.

A modified supraDarwinian perspective

Following the Golden Age, land-plant radiations typically form a nested sequence through time. The relative phenotypic 'breadth' of radiations (especially novelty radiations) decreases (Bateman *et al.*, 1998), and adaptive radiations become more likely than novelty radiations. This scenario is receiving increasing support from some ecological models, most notably the 'unified neutral theory of biodiversity' recently constructed by Hubbell (2001) and Bell (2001) on foundations laid by Levinton (1979, 1988; see also Silvertown & Antonovics, 2001; Whitfield, 2002). This paradigm views the key phylogenetic unit not as the species but as the individual organism; the rationale applies to any level in the demographic hierarchy, from population through metapopulation to species. In addition, the paradigm rejects the (essentially untestable) hypothesis that the number and diversity of niches in a given volume of ecospace approximates a predetermined equilibrium. Any equilibrium is set not by the number of species present but by the number of individuals that can be sustained by the available resources.

More specifically, the world is populated by a relatively small number of widespread, ecologically dominant species, each frequently generating prospective descendant species by a wide range of biological mechanisms. Each ancestral species competes with its descendant species (and in turn with their descendants) for limited resources in an ecospace that, with the exception of local perturbations, is saturated with organisms. The large number and size of the metapopulations of the ancestral species confer strong resistance to extinction. As more and more species are added to the ecospace, without any corresponding increase in resources, the average metapopulation size of each species inevitably declines, increasing the extinction rate until it eventually reaches an equilibrium with the speciation rate, thus defining a steady-state level of species richness.

Other consequences of this model are that species longevities vary enormously, and the closer to the base of a radiation

a species diverges, the greater is its average metapopulation size and thus its average longevity. Both its larger size and longer lifespan combine to greatly increase its probability of generating further new species. The model thereby explains increases in species-level diversity through time, and the fractal nature of evolution that has led to smaller average phenotypic divergences between ancestor and descendant through time. This in turn lends at least partial biological reality to the taxonomic hierarchy imposed on organisms by the increasingly sophisticated phylogenetic methods that reveal sister-species (Bateman & DiMichele, 2002; Kellogg, 2002). More reassuringly from our viewpoint, it is consistent with the model of a pot pourri of supraDarwinian evolutionary mechanisms long advocated by us (Bateman & DiMichele, 1994*a*, fig. 7). Moreover, we view the intrinsic driving force of diversification outlined by Hubbell's (2001) model as both more credible and more testable than the less constrained niche-partitioning arguments outlined above.

Promising directions for future inquiry into plant and animal macroevolution

Inferring relationships

The morphological phylogenetic, and subsequent molecular phylogenetic, revolutions have given botany a much stronger and scientifically based framework for understanding relationships and developing more natural classifications (Angiosperm Phylogeny Group, 1998; Hollingsworth *et al.*, 1999; Judd *et al.*, 1999; Qiu *et al.*, 1999; Chase *et al.*, 2000). The underlying principle of monophyly is now well-established but, somewhat perversely, the recent exponential increase in DNA sequence data has fragmented the morphological community, leading to a great diversity of opinions regarding the relative value of, and most appropriate uses for, morphological phylogenetic data (cf. Bateman, 1999; Scotland *et al.*, 2003).

There is general agreement that higher taxa are best delimited at relatively long branches in phylogenies, but we would argue that those phylogenies should include morphological characters capable of quantifying the degrees of phenotypic (as opposed to genotypic) disparity that are the primary evolutionary legacy of higher taxa. In other words, molecular data are useful for delimiting monophyletic groups, but taxonomic ranking should be determined primarily according to the relative lengths of morphological rather than molecular branches, longer branches denoting higher ranks. Also, morphological characters are essential if the vast tracts of the plant kingdom that lack close living relatives (for example, the many extinct higher taxa that alone bridge the vast phenotypic gap between extant ferns and extant cycads) are to play appropriate roles in understanding higher-level plant diversification. Without such taxa, homology assessment occurs by extrapolation backward through time of extant character combinations, making extant higher taxa appear artificially morphologically distinct and leading to a false sense of security regarding the 'accuracy' (which is in truth untestable) of the resulting phylogenies.

There are particular dangers in building phylogenies using sequence data, then 'mapping' selected morphological characters across cladograms and claiming to be studying evolution. Most notably, the causal arguments become circular, as the study considers only those morphological features viewed as benign directly involved in that particular evolutionary transition. Also, there is no opportunity to discover undescribed or poorly known characters capable of providing stronger information regarding phylogenetic relationships and major evolutionary transitions.

The practical intellectual rewards are far greater if the analyst views phylogeny reconstruction as a probability statement (Rieppel, 1988), and elects to use biology as a positive guide to assessing the likelihood that any particular type of data will give a credible result in a particular circumstance. We have already discussed the benefits of morphology for resolving radiations, but in the case of ecological shifts in lineages that simplify morphology in convergent ways (e.g. to aquatic habitats or to parasitizing other vascular plants, fungi or bacteria), sequence data are likely to be more informative. But in the case of, for example, non-photosynthesizing mycoheterotrophic ('saprophytic') plants, the regions of the genome that are directly involved in photosynthesis and are popularly used for phylogeny reconstruction are clearly unsuitable (Nickrent *et al.*, 1998).

We can only conclude that the frequent ideological conflicts between morphologists and molecular biologists are profoundly misconceived; rather, the identification and interpretation of evolutionary radiations cannot be satisfactorily achieved without constructing morphological *and* molecular matrices for the same range of analysed species and comparing the resulting phylogenies. Much valuable biological interpretation can be obtained by attempting to understand incongruence between different sources of phylogenetic data before combining the disparate data sources to generate 'maximum available evidence' trees (Bateman, 1999). In other words, progress will be made most rapidly by reciprocal illumination, both between contrasting data-sets and between data and the hypotheses that they are gathered to address.

Understanding relationships

The best arguments for pursuing carefully targeted studies that integrate morphological and sequence data are emanating from the current revolution in evolutionary–developmental genetics ('evo-devo': Bharathan *et al.*, 2002; Cronk *et al.*, 2002; Hofer & Ellis, 2002). However, this statement should be qualified by noting that evo-devo is emerging at a time when our perception of functionality of different components of the genome is becoming ever more complex (Avise, 2001), and we are increasingly accepting arguments that contrasting genetic processes undermine the long-cherished goal of a unified species concept for both prokaryotes and eukaryotes.

Existing phylogenies usefully highlight evolutionary transitions of particular importance or mechanistic interest, allowing us to transfer our current developmental genetic research focus from largely non-comparative studies of model

organisms to genuinely comparative studies of 'model transitions' between morphologies. Comparisons may be made between species or between distinct morphs within species, the latter being the products of either experimental or natural mutation (Theißen, 2000; Rudall & Bateman, 2002, 2003). The primary challenge is now to identify specific mutations in key developmental genes that engender corresponding phenotypic transitions, and then to fully understand their modes of expression. This will provide not only an ultimate test of homology but also deeper understanding of both homoplasy and functionality.

For example, functional constraints are presumably responsible for the fact that there is greater similarity between the DNA replication/repair enzymes of plants and humans than those of fruit flies and humans (Arabidopsis Genome Initiative, 2000). In contrast, increasing evidence is accumulating for the existence of 'spandrels': non-functional features that emerge as a consequence of functional optimization of features elsewhere on the bauplan of the organism (Gould & Lewontin, 1979; Gould, 2002). Examples include aspects of the much discussed double fertilization in angiosperms (W.E. Friedman, pers. comm., 2002) and single actinomorphic terminal flowers in lamialean inflorescences that otherwise bear many zygomorphic lateral flowers (Rudall & Bateman, 2003). Also, genuinely causal explanations will emerge for patterns of transitions in form between putative ancestor and descendant, including categories much discussed in the evolutionary literature such as heterochrony (Zelditch & Fink, 1996) and heterotopy (including homeosis: Baum & Donoghue, 2002).

Perhaps the 'hottest' topic of all at the time of writing is whole-genome sequencing of 'model' species. The fanfare associated with the release of the human genome sequence in 2000 overshadowed the publication in the same year of the sequence for the most widely used model plant, *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000). This herculean empirical effort immediately generated many biological insights (Walbot, 2000; Cronk, 2001; Bateman & DiMichele, 2002), beginning with the prediction that *Arabidopsis* possessed *c.* 26 000 genes, compared with *c.* 35 000 for humans (Arabidopsis Genome Initiative, 2000; Walbot, 2000; Bennetzen, 2002). This figure has since been explored further by expressed sequence tagging (Seki *et al.*, 2002). The 155 000 full-length cDNA clones obtained were eventually rationalized to 14 700 non-redundant cDNA groups; in other words, the bulk of the 26 300 genes originally estimated by the Arabidopsis Genome Initiative were located by the EST method. The 44% discrepancy may in part reflect the presence of pseudogenes or genes whose expression is weak and localized (Seki *et al.*, 2002).

Even more excitingly, *Arabidopsis*, a relatively derived eudicot angiosperm, has just become the subject of comparative genomics, as a consequence of the release of the whole-genome sequence for a monocot grass, specifically two subspecies of rice, *Oryza sativa* subsp. *indica* (Yu *et al.*, 2002) and subsp. *japonica* (Goff *et al.*, 2002). Monocots and dicots supposedly diverged 150–200 my ago (Bennetzen, 2002), so significant genetic differences should be expected. Not surprisingly, the alignability across the genome of key developmental genes evident among grasses (synteny: Gale & Devos, 1998) breaks down. Other molecular differences detected, including a distinctive GC gradient in rice and contrasts in codon and amino-acid usages, still awaitdetailed explanations.

However, more striking by far are the similarities between the two model genomes. Those similarities between *Arabidopsis* and rice, and between the two subspecies of rice, increase our confidence when making inter-kingdom genetic comparisons. The sedentary lifestyle of plants is reflected in a much higher proportion of genes that influence cell-wall formation, water and hormonal transport, together acting as proxys for the signal transduction systems that dominate higher animals. Higher copy numbers of many developmental gene families in plants relative to animals probably reflect ancient multiple polyploidy events (Bennetzen, 2002). Also, the rice genome sequence was generated rapidly, by reciprocal illumination between meticulous approaches and techniques that are cruder but cheaper and faster; these technical advances are likely to result in exponential increases in whole-genome sequences generated during the next decade.

Tens of thousands of genes may be a smaller number than was expected by most pre-genomics geneticists, but it is still a substantial number, potentially allowing phenomenal evolutionary complexity (Glazier *et al.*, 2002). However, it is also rapidly becoming clear that a relatively small number of genes, and an even smaller number of genes of profound phenotypic effect, underlie particular phenotypic transitions (Walbot, 2000; Cronk, 2001; Bateman & DiMichele, 2002).

Good examples of both the potential simplicity of gene control of important traits and of functional conservation across major groups have been elucidated in insects. For example, the gene *forager* strongly influences food-seeking behaviour in fruit flies but also dictates foraging behaviour in social bees (Ben-Shakar *et al.*, 2002). A more cautionary tale was presented by Marshall *et al.* (1999), who began their investigation with a hypothesis that the Hox gene *ultrabithorax* regulated the fundamental distinction between two-winged and four-winged insects, but later discovered that this gene was in fact highly conserved and that the actual culprit lay downstream.

Comparison of developmental control in higher plants and higher animals shows some ancient gene families and chromatin-based processes in common but the higher level families (e.g. Hox, MADS-box) evolved independently to fulfill similar mechanisms (Meyerowitz, 2002), as did cell-to-cell signalling. One major difference lies in the additional genes acquired by plants from the commensal cyanobacterium that was eventually to evolve via endosymbiosis into the green plant plastid.

Much recent evo-devo discussion has concerned the significance of gene duplication events, either individually or *in toto* as a result of a polyploidy event. Gene duplication followed by division of existing functions between the two copies or acquisition of new functions by one copy are potentially powerful drivers of evolution (Lynch, 2002). However, current evidence suggests that, in most cases, one copy is silenced within a few million years. Nonetheless,

preferential preservation of the duplicated copy still constitutes a chromosomal rearrangement capable of altering the subtleties of function (Lynch, 2002). And the loss of one copy *after* it has acquired a distinct function can prompt further evolutionary transitions. Most notably, loss of function in one of two gymnospermous copies of the *LEAFY* gene has been directly implicated in the origin of the angiosperm flower (Frohlich & Parker, 2000; Albert *et al.*, 2002; Frohlich, 2002). And reciprocal losses of function in contrasting, duplicated regulators of a likely protein-coding gene provide an effective isolation mechanism, prompting the duplication–degeneration–complementation model recently promoted by developmental biologists such as J.H. Postlethwaite and D.L. Stern (pers. comm., 2002). Moreover, fusion of genes that allow maintenance of both enzyme activities in a single protein are also viable, sometimes providing useful kingdom-scale synapomorphies (Stechmann & Cavalier-Smith, 2002).

However, it is essential that the context of expression of these genetic modifications is explored with care. Alternate genotypic changes yielding similar phenotypic changes are proving to be frequent; they are most clearly demonstrable in relatively simple bacterial systems (P.B. Rainey, pers. comm., 2002). And even when a particular point mutation can unequivocally cause a particular morphological transition *in vitro*, this does not necessarily mean that it caused such an evolutionary transition *in vivo*. A good example is the presence or absence of wings in genetically identical ants according to caste, which simply reflects differential environmental cues (polyphenism: Abouheif & Wray, 2002). Thus, evo-devo studies need to consider not only functional linkages with other genes but also epigenetic constraints on expression (see papers in Jablonka & Lamb, 1995) and environmental cues, potentially spawning yet another discipline that has playfully been titled 'eco-devo'.

Most critically for evolutionary theory, evo-devo studies are increasingly suggesting that subtly expressed genetic polymorphisms within species may not be the primary cause of macroevolution (Nordborg & Innan, 2002). The overall picture now emerging is one of protein-coding genes undergoing rare but profound mutations, but with the localization of the resulting morphogenetic effects being determined by the greater subtleties of more complex sets of regulators moderated by the cellular and extra-cellular environment.

Whatever the details of the molecular evolution of key developmental genes, it is clear that whole-genome comparison will yield rich rewards; those beneficiaries listed by Bennetzen (2002) included investigators of gene expression, protein structures, quantitative traits, allelic variability and genetic fingerprinting, as well as greatly enhancing the feasibility of candidate gene comparisons among much wider ranges of taxa.

Value of longer term observations

As well as understanding the generation of evolutionary novelty, we need to understand differential filtering of contrasting morphs, not only in experimental plots (Glover & Martin, 2002) but also as 'prospecies' in the wild (Bateman & DiMichele, 2002; Rudall & Bateman, 2002). Thus can we accurately gauge ecological success, and test evolutionary hypotheses that require multiple phenotypic transitions, such as pre-adaptation where by definition the origin of the first function is followed by the switch to the subsequent function. This is where the great value of longer term field observations becomes especially apparent.

First, it is now clear that, despite increasingly well-defined limits to the maximum rate of adaptive responses (Etterson & Shaw, 2001; Silvertown *et al.*, 2002), plants can adjust their behaviour rapidly to accommodate environmental changes. A 50-year project monitoring 385 vascular plant species in the British Isles (Fitter & Fitter, 2002) revealed that the first flowering date of these species has advanced by an average of 4.5 ± 0.4 days over the past decade relative to the previous four decades, presumably in response to global warming (spring flowering, insect pollinated annuals are most prone to change). Moreover, the 16% of those species that showed the largest shifts in flowering period advanced by an average of 15 days, a period sufficient to either induce or eliminate reproductive isolation among closely related but formerly phenologically differentiated species.

This inference is of particular importance because recent evolutionary studies have further emphasized the significance of even very limited gene flow in influencing selective traits at a fine temporal scale (Grant & Grant, 2002), and at a coarser temporal scale of maintaining the long periods of stasis that are the most fundamental component of the punctuated equilibrium model of evolution (Futuyma, 1987; Gould, 2002). A good example is provided by that archetypal model system, Darwin's finches, which have recently been demonstrated to be monophyletic and phylogenetically nested within Caribbean groups also prone to rapid speciation (Burns *et al.*, 2002). Grant & Grant's (2002) 30-year field study of two species of Darwin's finches, *Geospiza fortis* and *G. scandens*, that utilize different food sources on the Galápagos island of Daphne Major showed that mean body size and beak shape changed significantly over the study period in both species, *G. scandens* converging on *G. fortis*. In *G. fortis*, body and beak size rose and subsequently fell, while beak shape became more acute. The only synchronous selective event in the two co-existing species was increased size in response to drought conditions in the late 1970s. Gene flow engendering little or no fitness loss occurred between the two species in response to either drought or El Niño rains, the rate of flow being proportionally greater from *G. fortis* to *G. scandens* due to male-biased sex ratios favouring intraspecific fidelity in *G. scandens*; this species also reliably provided the partners for the resulting F_1 hybrids.

Thus, as noted by Grant & Grant (2002), unusually long-term studies are needed to detect reversals in trends in selection (or, for that matter, drift) and to identify the phenotypic effects of infrequent environmental perturbations. Moreover, the environmental perturbations often invoked as the initiators of speciation may also be the most effective barriers to speciation, as they encourage both oscillating selection trends and re-integration of recently separated lineages (Futuyma, 2002). As noted by Gould (2002), much depends on our ability to infer whether putative ancestor and descendant co-exist

long after the latter originated in a cladogenetic model of speciation (Fig. 2a, b) or whether the descendant is geographically or ecologically isolated, permitting an anagenetic model (Fig. 2c).

Finally, it should be noted that decades, periods almost unimaginably long to evolutionary biologists, are almost unimaginably short to palaeobiologists, who with very few exceptions (e.g. varved lakes) cannot hope to resolve time with anything close to annual precision. And where this high resolution *is* possible, as when a particular horizon yields a single autumnal layer of leaf litter, the horizon is difficult to trace laterally over sufficient distance to sample enough closely related individuals or species to monitor the evolution of macroscopic plants (Wing & DiMichele, 1995). Although greater opportunities may exist for comparable studies of aquatic microfossils, we must conclude that most transient phenotypic excursions of the kind documented in Darwin's finches cannot be traced through the fossil record.

Conclusions

In our opinion, current data suggest that microevolutionary excursions in particular phenotypic traits (irrespective of whether the excursions reflect directional selection or drift) rarely lead to macroevolution; in other words, to speciation.

Although current understanding of comparative gene expression is undoubtedly grossly simplistic, the problems at least appear tangible; we may still lack most of the answers but the questions now being asked are far better formulated and more relevant. We have gained access to a new panoply of potential explanations for constraints on, and iteration in, evolution, and a better comprehension of process-based interactions between different demographic tiers: in other words, in the hierarchy of individual > population > metapopulation > species that was especially beloved of S.J. Gould (most recently Gould, 2002). Moreover, the scales of both environmental perturbations and mutation-driven phenotypic transitions vary enormously, the most subtle changes being the most common and the most radical being the least common. All scales of change have the potential to cause macroevolution, but their relative importance in the history of life rightly remains highly contentious.

The two quotes that began this essay, both taken from Futuyma's (2002) critique of Gould's (2002) *magnum opus* (and swansong) that places evolutionary theory in a historical context, contrast qualitative acknowledgement of the existence of a wide range of evolutionary processes with quantitative scepticism regarding the importance of those processes lying outside the traditional boundaries of neoDarwinism. We believe that the mechanism traditionally invoked to explain our current remarkable levels of plant diversity – the gradual spread of successive, functionally linked mutations through large panmictic populations – may over the next decade give way to a less narrowly focused perspective that better encompasses the diversity of evolutionary processes now recognized (however grudgingly by evolutionary traditionalists) as generating the present diversity of life.

Acknowledgements

We dedicate this paper to organismal biology's most provocative essayist of the late 20th century, Stephen Jay Gould (1941–2002), and we thank Paul Kenrick, Brian Rosen, Paula Rudall, Dick Vane-Wright and an anonymous reviewer for their critical appraisals.

References

ABOUHEIF, E. & WRAY, G.A. 2002. Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**, 249–252.

ALBERT, V.A., OPPENHEIMER, D.G. & LINDQVIST, C. 2002. Pleiotropy, redundancy and the evolution of flowers. *Trends in Plant Science* **7**, 297–301.

Angiosperm Phylogeny Group 1998. An ordinal classification for the families of flowering plants. *Annals of the Missouri Botanical Garden* **85**, 531–553.

Arabidopsis Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.

ARNOLD, M.L. 1997. *Natural Hybridization and Evolution*. Oxford University Press, Oxford.

AUSTIN, J.J., SMITH, A.B. & THOMAS, R.J. 1997. Palaeontology in a molecular world: the search for authentic ancient DNA. *Trends in Ecology and Evolution* **12**, 303–306.

AVISE, J.C. 1994. *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, London.

AVISE, J.C. 2001. Evolving genomic metaphors: a new look at the language of DNA. *Science* **294**, 86–87.

BALDWIN, B.G., CRAWFORD, D.J., FRANCISCO-ORTEGA, J., KIM, S.C., SANG, T. & STEUSSY, T.F. 1998. Molecular phylogenetic insights on the origin and evolution of oceanic island plants. In Soltis, D.E., Soltis, P.S. & Doyle, J.J. (eds), *Molecular Systematics of Plants 2*. Chapman & Hall, London, pp. 410–441.

BATEMAN, R.M. 1996. Non-floral homoplasy and evolutionary scenarios in living and fossil land-plants. In Sanderson, M.J. & Hufford, L. (eds), *Homoplasy and the Evolutionary Process*. Academic Press, London, pp. 91–130.

BATEMAN, R.M. 1999. Integrating molecular and morphological evidence for evolutionary radiations. In Hollingsworth, P.M., Bateman, R.M. & Gornall, R.J. (eds), *Molecular Systematics and Plant Evolution*. Taylor & Francis, London, pp. 422–471.

BATEMAN, R.M. 2001. Evolution and classification of European orchids: insights from molecular and morphological characters. *Journal Europäischer Orchideen* **33**, 33–119.

BATEMAN, R.M. & DIMICHELE, W.A. 1994*a*. Saltational evolution of form in vascular plants: a neoGoldschmidtian synthesis. In Ingram, D.S. & Hudson, A. (eds), *Shape and Form in Plants and Fungi*. Academic Press, London, pp. 63–102.

BATEMAN, R.M. & DIMICHELE, W.A. 1994*b*. Heterospory: the most iterative key innovation in the evolutionary history of the plant kingdom. *Biological Reviews* **69**, 345–417.

BATEMAN, R.M. & DIMICHELE, W.A. 2002. Generating and filtering major phenotypic novelties: neoGoldschmidtian saltation revisited. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds), *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 109–159.

BATEMAN, R.M., CRANE, P.R., DIMICHELE, W.A., KENRICK, P., ROWE, N.P., SPECK, T. & STEIN, W.E. 1998. Early evolution of land plants: phylogeny, physiology, and ecology of the primary terrestrial radiation. *Annual Review of Ecology and Systematics* **29**, 263–292.

BAUM, D.A. & DONOGHUE, M.J. 2002. Transference of function, heterotopy and the evolution of plant development. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds),

Developmental Genetics and Plant Evolution. Taylor & Francis, London, pp. 52–69.

BEHRENSMEYER, A.K., DAMUTH, J.D., DIMICHELE, W.A., POTTS, R., SUES, H.-D. & WING, S.L. 1992. *Terrestrial Ecosystems Through Time*. Chicago University Press, Chicago.

BELL, G. 2001. Neutral macroecology. *Science* **293**, 2413–2418.

BENNETT, K.D. 1997. *Evolution and Ecology: the Pace of Life*. Cambridge University Press, Cambridge.

BENNETZEN, J. 2002. Opening the door to comparative plant biology. *Science* **296**, 60–63.

BEN-SHAKAR, Y, ROBICHON, A., SOKOLOWSKI, M.B. & ROBINSON, G.E. 2002. Influence of gene action across different time scales on behavior. *Science* **296**, 741–744.

BENTON, M.J. 2000. Stems, nodes, crown clades, and rank-free lists: is Linnaeus dead? *Biological Reviews* **75**, 633–648.

BHARATHAN, G., GOLIBER, T.E., MOORE, C., KESSLER, S., PHAM, T. & SINHA, N.R. 2002. Homologies in leaf form inferred from *KNOX1* gene expression during development. *Science* **296**, 1858–1860.

BIRKS, H.J.B. 1989. Holocene isochrone maps and patterns of tree spreading in the British Isles. *Journal of Biogeography* **16**, 503–540.

BRUMMITT, R. 2002. How to chop up a tree. *Taxon* **51**, 31–41.

BRYANT, H.N. & CANTINO, P.D. 2002. A review of criticisms of phylogenetic nomenclature: is taxonomic freedom the fundamental issue? *Biological Reviews* **77**, 39–55.

BURNS, K.J., HACKETT, S.J. & KLEIN, N.K. 2002. Phylogenetic relationships and morphological divergence in Darwin's Finches and their relatives. *Evolution* **56**, 1240–1252.

CANTINO, P.D. & DE QUEIROZ, K. 2000. *PhyloCode*. http://www.ohiou.edu/phylocode/

CARROLL, S.B., GRENIER, J.K. & WEATHERBEE, S.D. 2001. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Blackwell, Oxford.

CHASE, M.W. 2001. The origin and biogeography of Orchidaceae. In Pridgeon, A.M., Cribb, P.L., Chase, M.W. & Rasmussen, F.N. (eds), *Genera Orchidacearum, 2. Orchidoideae, Part 1*. Oxford University Press, Oxford, pp. 1–5.

CHASE, M.W., FAY, M.F. & SAVOLAINEN, V. 2000. Higher-level classification in the angiosperms: new insights from the perspective of DNA sequence data. *Taxon* **49**, 685–704.

CHAW, S.-M., PARKINSON, C.L., CHENG, Y., VINCENT, T.M. & PALMER, J.D. 2000. Seed-plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origins of Gnetales from conifers. *Proceedings of the National Academy of Sciences of the USA* **97**, 4086–4091.

COLINVAUX, P.S. & DE OLIVIERA, P.E. 2001. Amazon plant diversity and climate through the Cenozoic. *Palaeogeography, Palaeoclimatology, Palaeoecology* **166**, 51–63.

CRANE, P.R. 1989. Patterns of evolution and extinction in vascular plants. In Allen, K.C. & Briggs, D.E.G. (eds), *Evolution and the Fossil Record*. Belhaven, Chichester, pp. 153–187.

CRONK, Q.C.B. 2001. Plant evolution and development in a post-genomic context. *Nature Reviews, Genetics* **2**, 607–619.

CRONK, Q.C.B., BATEMAN, R.M. & HAWKINS, J.A. (eds) 2002. *Developmental Genetics and Plant Evolution*. Taylor & Francis, London.

CUBAS, P. 2002. Role of TCP genes in the evolution of morphological characters in angiosperms. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds), *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 247–266.

DILCHER, D. 2000. Toward a new synthesis: major evolutionary trends in the angiosperm fossil record. In Ayala, F.J., Fitch, W.M. & Clegg, M.T. (eds), *Variation and Evolution in Plants and Microorganisms*. National Academy Press, Washington DC, pp. 255–270.

DIMICHELE, W.A. & BATEMAN, R.M. 2003. Evolution of land-plant diversity: Major innovations and lineages through time. In Kress, W.J. & Krupnik, G. (eds), *Plant Conservation: a Natural History Approach*. Smithsonian Institution Press, Washington DC.

DIMICHELE, W.A., HOOK, R.W., BEERBOWER, R., BOY, J.A., GASTALDO, R.A., HOTTON, N. III, PHILLIPS, T.L., SCHECKLER, S.E., SHEAR, W.A. & SUES, H.-D. 1992. Paleozoic terrestrial ecosystems. In Behrensmeyer, A.K., Damuth, J.D., DiMichele, W.A., Sues, H.-D. & Wing, S.L. (eds), *Terrestrial Ecosystems Through Time*. Chicago University Press, Chicago, pp. 204–325.

DIMICHELE, W.A., PFEFFERKORN, H.W. & GASTALDO, R.A. 2001*a*. Response of late Carboniferous and early Permian plant communities to climate change. *Annual Reviews of Earth and Planetary Sciences* **29**, 461–487.

DIMICHELE, W.A., STEIN, W.E. & BATEMAN, R.M. 2001*b*. Ecological sorting of vascular plant classes during the Paleozoic evolutionary radiations. In W.D. Allmon & D.J. Bottjer (eds), *Evolutionary Paleoecology: the Ecological Context of Macroevolutionary Change*. Columbia University Press, New York, pp. 285–335.

DONOGHUE, M.J. 1989. Phylogenies and the analysis of evolutionary sequences, with examples from seed plants. *Evolution* **43**, 1137–1156.

DONOVAN, S.K. & PAUL, C.R.C. (eds) 1998. *The Adequacy of the Fossil Record*. John Wiley & Sons, New York.

ERWIN, D.H. 1992. A preliminary classification of evolutionary radiations. *Historical Biology* **6**, 133–147.

ETTERSON, J.R. & SHAW, R.G. 2001. Constraint to adaptive evolution in response to global warming. *Science* **294**, 151–154.

FERRIS, C., KING, R.A. & HEWITT, G.M. 1999. Isolation within species and the history of glacial refugia. In Hollingsworth, P.M., Bateman, R.M. & Gornall, R.J. (eds), *Molecular Systematics and Plant Evolution*. Taylor & Francis, London, pp. 20–34.

FITTER, A.H. & FITTER, R.S.R. 2002. Rapid changes in flowering time in British plants. *Science* **296**, 1689–1691.

FOREY, P.L. 2002. *Phylocode* – no pain, no gain. *Taxon* **51**, 43–54.

FROHLICH, M.W. 2002. The Mostly Male theory of flower origins: summary and update regarding the Jurassic pteridosperm *Pteroma*. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds), *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 85–108.

FROHLICH, M.W. & PARKER, D.S. 2000. The mostly male theory of flower evolutionary origins: from genes to fossils. *Systematic Botany* **25**, 155–170.

FUTUYMA, D.J. 1987. On the role of species in anagenesis. *American Naturalist* **130**, 465–473.

FUTUYMA, D.J. 2002. Stephen Jay Gould à la recherche du temps perdu. *Science* **296**, 661–663.

GALE, M.D. & DEVOS, K.M. 1998. Plant comparative genetics after 10 years. *Science* **282**, 656–659.

GLAZIER, A.M., NADEAU, J.H. & ALTMAN, T.J. 2002. Finding genes that underlie complex traits. *Science* **297**, 2345–2349.

GLOVER, B.J. & MARTIN, C. 2002. Evolution of adaptive petal cell morphology. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds), *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 160–172.

GOFF, S.A. and 54 co-authors. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100.

GOLENBERG, E.M. 1999. Isolation, identification and authentication of DNA sequences derived from fossil material. In Jones, T.P. & Rowe, N.P. (eds), *Fossil Plants and Spores: Modern Techniques*. Geological Society, London, pp. 156–160.

GOLENBERG, E.M., GIANNASI, D.E., CLEGG, M.T., SMILEY, C.J., DURBIN, M., HENDERSON, D. & ZURAWSKI, G. 1990. Chloroplast DNA sequence from a Miocene *Magnolia* species. *Nature* **344**, 656–658.

GOULD, S.J. 2002. *The Structure of Evolutionary Theory*. Belknap, Cambridge, MA.

GOULD, S.J. & LEWONTIN, R.C. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London B* **205**, 581–598.

GRANT, P.R. & GRANT, B.R. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711.

HARRIS, D.J., POULSEN, A.D., FRIMODT-MOLLER, C., PRESTON, J. & CRONK, Q.C.B. 2000. Rapid radiation in *Aframomum* (Zingiberaceae): evidence from nuclear ribsosomal DNA internal transcribed spacer (ITS) sequences. *Edinburgh Journal of Botany* **57**, 377–395.

HEWITT, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247–276.

HOFER, J. & ELLIS, N. 2002. Conservation and diversification of gene function in plant development. *Current Opinion in Plant Biology* **5**, 56–61.

HOLLINGSWORTH, P.M., BATEMAN, R.M. & GORNALL, R.J. (eds) 1999. *Molecular Systematics and Plant Evolution*. Taylor & Francis, London.

HOLLINGSWORTH, P.M., BATEMAN, R.M. & SQUIRRELL, J. 2003. Molecular phylogenetics of subtribe Neottieae (Orchidaceae): iterative origins of allogamy and mycoheterotrophy. *American Journal of Botany* [Submitted].

HUBBELL, S.P. 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, NJ.

HUGHEN, K.A., SOUTHON, J.R., LEHMAN, S.J. & OVERPECK, J.T. 2000. Synchronous radiocarbon and climate shifts during the last deglaciation. *Science* **290**, 1951–1954.

HUGHES, P.D.M. & DUMAYNE-PEATY, L. 2002. Testing theories of mire development using multiple successions at Craymlyn Bog, West Glamorgan, South Wales, UK. *Journal of Ecology* **90**, 456–471.

HUNTLEY, B. 1999. The dynamic responses of plants to environmental change and the resulting risks of extinction. In Mace, G.M., Balmford, A. & Ginsberg, J.R. (eds), *Symposium of the Zoological Society of London 72*. Cambridge University Press, Cambridge, pp. 69–85.

HUNTLEY, B. & BIRKS, H.J.B. 1983. *An Atlas of Past and Present Pollen Maps for Europe: 0–13 000 years ago*. Cambridge University Press, Cambridge.

JABLONKA, E. & LAMB, M.J. 1995. *Epigenetic Inheritance and Evolution*. Oxford University Press, Oxford.

JOHNSON, K.R. & ELLIS, B. 2002. A tropical rainforest in Colorado 1.4 million years after the Cretaceous–Tertiary boundary. *Science* **296**, 2379–2383.

JUDD, W.S., CAMPBELL, C.S., KELLOGG, E.A. & STEVENS, P.F. 1999. *Plant Systematics: a Phylogenetic Approach*. Sinauer, Sunderland, MA.

KELLOGG, E.A. 2002. Are macroevolution and microevolution qualitatively different? Evidence from Poaceae and other families. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds) *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 70–84.

KENRICK, P. & CRANE, P.R. 1997. *The Origin and Early Diversification of Land Plants: a Cladistic Study*, Smithsonian Series in Comparative Evolutionary Biology. Smithsonian Institution Press, Washington, DC.

KITCHING, I.J., FOREY, P.L., HUMPHRIES, C.J., SIEBERT, D.J. & WILLIAMS, D.M. 1999. *Cladistics: a Practical Course in Systematics* (2nd edn). Oxford University Press, Oxford.

LEVINTON, J. 1979. A theory of diversity equilibrium and morphological evolution. *Science* **204**, 335–336.

LEVINTON, J. 1988. *Genetics, Paleontology, and Macroevolution*. Cambridge University Press, Cambridge.

LYNCH, M. 2002. Gene duplication and evolution. *Science* **297**, 945–947.

MARSHALL, C.R., ORR, H.A. & PATEL, N.H. 1999. Morphology, innovation and developmental genetics. *Proceedings of the National Academy of Sciences of the USA* **96**, 9995–9996.

MEYEROWITZ, E.M. 2002. Plants compared to animals: the broadest comparative study of development. *Science* **295**, 1482–1485.

NICKRENT, D.L., DUFF, R.J., COLWELL, A.E., WOLFE, A.D., YOUNG, N.D., STEINER, K.E. & DEPAMPHILIS, C.W. 1998. Molecular phylogenetic and evolutionary studies of parasitic plants. In Soltis, D.E., Soltis, P.S. & Doyle, J.J. (eds), *Molecular Systematics of Plants 2*. Chapman & Hall, London, pp. 211–241.

NIXON, K.C. & CARPENTER, J.M. 2000. On the other 'phylogenetic systematics.' *Cladistics* **16**, 298–318.

NORDBORG, M. & INNAN, H. 2002. Molecular population genetics. *Current Opinion in Plant Biology* **5**, 69–73.

O'BRIEN, E.M., WHITTAKER, R.J. & FIELDS, R. 1998. Climate and woody plant diversity in southern Africa: relationships at species, genus and family level. *Ecography* **21**, 495–509.

O'BRIEN, E.M., WHITTAKER, R.J. & FIELDS, R. 2000. Climatic gradients in woody plant (tree and shrub) diversity: water–energy dynamics, residual variation, and topography. *Oikos* **89**, 588–600.

PAGE, R.D.M. & HOLMES, E.C. 1998. *Molecular Evolution: a Phylogenetic Approach*. Blackwell, Oxford.

PARMESAN, C. *et al.* 2001. *Impacts, Adaptations and Vulnerabilities*. Intergovernmental panel on climate change (IPCC). Third Assessment Report (WGII). Cambridge University Press, Cambridge.

PHILLIPS, T.L. & DIMICHELE, W.A. 1992. Comparative ecology and life-history biology of arborescent lycopsids in Late Carboniferous swamps of Euramerica. *Annals of the Missouri Botanical Garden* **79**, 560–588.

PRYER, K.M., SCHNEIDER, H., SMITH, A.R., CRANFILL, R., WOLF, P.G., HUNT, J.S. & SIPES, S.D. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives of seed plants. *Nature* **409**, 618–622.

QIU, Y.-L., LEE, J., BERNASCONI-QUADRONI, F., SOLTIS, D.E., SOLTIS, P.S., ZANIS, M., ZIMMER, E.A., CHEN, Z.-D., SAVOLAINEN, V. & CHASE, M.W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* **402**, 404–407.

REES, P. MCA. 2002. Land-plant diversity and the end-Permian mass extinction. *Geology* **30**, 827–830.

RICHARDSON, J.E., WEITZ, F.M., FAY, M.F., CRONK, Q.C.B., LINDER, H.P., REEVES, G. & CHASE, M.W. 2001*a*. Rapid and recent origin of species richness in the Cape flora of South Africa. *Nature* **412**, 181–183.

RICHARDSON, J.E., PENNINGTON, R.T., PENNINGTON, T.D. & HOLLINGSWORTH, P.M. 2001*b*. Rapid diversification of a species-rich genus of Neotropical rain forest trees. *Science* **293**, 2242–2245.

RIEPPEL, O. 1988. *Fundamentals of Comparative Biology*. Birkhäuser, Basel.

RIESEBERG, L.H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**, 359–389.

ROSENZWEIG, M.L. 1995. *Species Diversity in Space and Time*. Cambridge University Press, Cambridge.

RUDALL, P.J. & BATEMAN, R.M. 2002. Synorganisation, zygomorphy and heterotopy in the evolution of flowers: the gynostemium and labellum of orchids and other lilioid monocots. *Biological Reviews* **77**, 403–441.

RUDALL, P.J. & BATEMAN, R.M. 2003. Evolutionary change in flowers and inflorescences: evidence from naturally occurring terata. *Trends in Plant Science* **8**, 76–82.

RYDIN, C. & WIKSTRÖM, N. 2002. Phylogeny of *Isoetes* (Lycopsida): resolving basal relationships using *rbcL* sequences. *Taxon* **51**, 83–90.

SANDERSON, M.J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. *Molecular Biology and Evolution* **14**, 1218–1231.

SANDERSON, M.J. & DONOGHUE, M.J. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends in Ecology and Evolution* **11**, 15–20.

SANDERSON, M.J. & DOYLE, J.A. 2001. Sources of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. *American Journal of Botany* **88**, 1499–1516.

SCHLUTER, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

SCOTLAND, R.W., OLMSTEAD, R.G. & BENNETT, J.R. 2002. Role of morphology in phylogeny reconstruction. *Systematic Biology* **51** [in press].

SEKI, M. and 19 co-authors. 2002. Functional annotation of a full-length *Arabidopsis* cDNA collection. *Science* **296**, 141–145.

SIEGAL, M.L. & BERGMAN, A. 2002. Waddington's canalization revisited: developmental stability and evolution. *Proceedings of the National Academy of Sciences USA* **99**, 10523–10528.

SILVERTOWN, J. & ANTONOVICS, J. (eds) 2001. *Integrating Ecology and Evolution in a Spatial Context*. Blackwell, Oxford.

SILVERTOWN, J., MCCONWAY, K.J., HUGHES, Z., BLISS, P., MACNAIR, M. & LUTMAN, P. 2002. Ecological and genetic correlates of long-term population trends in the Park Grass experiment. *American Naturalist* **160**, 409–420.

SMITH, A.B. 1994. *Systematics and the Fossil Record: Documenting Evolutionary Patterns*. Blackwell, Oxford.

SOUTHWOOD, J.R.E. & HENDERSON, P.A. 2000. *Ecological Methods*. Blackwell, Oxford.

STECHMANN, A. & CAVALIER-SMITH, T. 2002. Rooting the eukaryote tree by using a derived gene fusion. *Science* **297**, 89–91.

TABERLET, P. & CHEDDADI, R. 2002. Quaternary refugia and persistence of biodiversity. *Science* **297**, 2009–2010.

TAKEBAYASHI, N. & MORRELL, P.L. 2001. Is self-fertilization an evolutionary dead-end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *American Journal of Botany* **88**, 1143–1150.

TAMAS, I., KLASSON, L., CANBÄCK, B., NÄSLUND, A.K., ERIKSSON, A.-S., WERNEGREEN, J.J., SANDSTRÖM, J.P., MORAN, N. A. & ANDERSSON, S. G. E. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **296**, 2376–2379.

THEIßEN, G. 2000. Evolutionary developmental genetics of floral symmetry: the revealing power of Linnaeus' monstrous flower. *BioEssays* **22**, 209–213.

THEIßEN, G., BECKER, A., WINTER, K.-U., MÜNSTER, T., KIRCHNER, C. & SAEDLER, H. 2002. How the land plants learned their floral ABCs: the role of MADS-box genes in the evolutionary origin of flowers. In Cronk, Q.C.B., Bateman, R.M. & Hawkins, J.A. (eds), *Developmental Genetics and Plant Evolution*. Taylor & Francis, London, pp. 173–205.

VALENTINE, J.W. 1980. Determinants of diversity in higher taxonomic categories. *Paleobiology* **6**, 444–450.

VAN DEN KOORNHUYSE, P., BALDAUF, S.L., LEYVAL, C., SRTACZEK, J. & YOUNG, J.P.W. 2002. Extensive fungal diversity in plant roots. *Science* **295**, 2051.

VAN STEENIS, C.G.G.J. 1976. Autonomous evolution in plants: differences in plant and animal evolution. *Gardens' Bulletin, Singapore* **29**, 103–126.

VERMEIJ, G.J. 1987. *Evolution and Escalation*. Princeton University Press, Princeton, NJ.

VISBECK, M. 2002. The ocean's role in Atlantic climatic variability. *Science* **297**, 2223–2224.

WAGNER, P.J. 1996. Contrasting the underlying pattern of active trends in morphologic evolution. *Evolution* **50**, 990–1007.

WALBOT, V. 2000. A green chapter in the book of life. *Nature* **408**, 794–795.

WALTER, K.S. & GILLETT, H.J. (eds) 1998. *1997 IUCN Red List of Threatened Plants*. WCMC, Cambridge, and IUCN, Gland.

WHITFIELD, J. 2002. Neutrality versus the niche. *Nature* **417**, 480–481.

WIKSTRÖM, N. & KENRICK, P. 2001. Evolution of Lycopodiaceae (Lycopsida): estimating divergence times from *rbcL* gene sequences by use of nonparametric rate smoothing. *Molecular Phylogenetics and Evolution* **19**, 177–186.

WIKSTRÖM, N., SAVOLAINEN, V. & CHASE, M.W. 2001. Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London B* **268**, 2211–2220.

WILLS, M.A. 2001. Morphological disparity: a primer. In Adrain, J.M., Edgecombe, G.D. & Lieberman, B.S. (eds), *Fossils, Phylogeny and Form: an Analytical Approach*. Kluwer, New York, pp. 55–144.

WING, S.L. & DIMICHELE, W.A. 1995. Conflict between local and global changes in plant diversity through geologic time. *Palaios* **10**, 551–564.

YU, J. and 99 co-authors. 2002. A draft sequence of the rice genome *(Oryza sativa* L. ssp. *indica)*. *Science* **296**, 79–92.

ZELDITCH, M.L. & FINK, W.L. 1996. Heterochrony and heterotopy: stability and innovation in the evolution of form. *Paleobiology* **22**, 247–250.