The fall and rise of Dr Pangloss: adaptationism and the *Spandrels* paper 20 years later

Massimo Pigliucci and Jonathan Kaplan

Twenty years have passed since Gould and Lewontin published their critique of 'the adaptationist program' – the tendency of some evolutionary biologists to assume, rather than demonstrate, the operation of natural selection. After the 'Spandrels paper', evolutionists were more careful about producing just-so stories based on selection, and paid more attention to a panoply of other processes. Then came reactions against the excesses of the anti-adaptationist movement, which ranged from a complete dismissal of Gould and Lewontin's contribution to a positive call to overcome the problems. We now have an excellent opportunity for finally affirming a more balanced and pluralistic approach to the study of evolutionary biology.

Massimo Pigliucci is at the Depts of Botany and of Ecology and Evolutionary Biology (pigliucci@utk.edu); Jonathan Kaplan is at the Dept of Philosophy, University of Tennessee, Knoxville, TN 37996-1100, USA (jkaplan@utk.edu).

Twenty years ago, Gould and Lewontin published *The Spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme*[1]. In it, they described 'the adaptationist program' as an attempt to explain the existence and the particular forms of any phenotypic trait as the result of natural selection. The paper's exotic title derived from the example with which the piece began. The authors noted that the tapered spaces (the 'spandrels') between the archways supporting the domed roof of the basilica of St Mark's in Venice were beautifully decorated in a way that made admirable use of the triangular space. They claimed that although this space was put to great artistic use, it was an architectural by-product of employing arches to support a domed room, and was not designed for that artistic use. The lesson to learn is that when faced with a 'trait' that is being put to good use, one ought not to jump to the conclusion that the particular use is the reason the trait is present. Gould and Lewontin argued that a particularly sloppy form of evolutionary thinking did just that, and that its practitioners assumed that any trait had to have a good use to explain its presence, even if none was obvious. Gould and Lewontin proceeded to criticize the assumptions that such a program depended upon (Box 1), and laid out an alternative approach that they felt better accounted for the complexities of the evolutionary process.

Over the past 20 years, *Spandrels* has been much cited and criticized.

Criticisms have ranged from the significant – for example, that the role of constraints in evolution is over-stated[2] – to the irrelevant (i.e. that the architectural feature they referred to is really called a 'pendentive' not a spandrel[3]). Today, it is commonplace to pay lip service to the sorts of difficulties involved in demonstrating that a trait is 'adaptive', which Gould and Lewontin pointed out. Twenty years is a long enough time for the dust to have settled and to ask to what extent the difficulties raised by the original paper can now be addressed. How have conceptual and empirical advances helped us to move away from adaptationist story-telling and towards testable hypotheses involving adaptation?

Whence *Spandrels*?

Before trying to answer these questions, it is worth thinking about what Gould and Lewontin's attack on the adaptationist program was motivated by. In large part, it was a reaction to two related phenomena: the rise of sociobiology, attributable in large part to E.O. Wilson's 1975 book by the same name[4]; and the popularization of the narrowly 'gene-centric' approach favored by Dawkins and defended explicitly in his 1976 book *The Selfish Gene*[5]. Those features that Gould and Lewontin claim describe the theory and the practice of adaptationism bear directly upon the styles of reasoning used in these arenas.

As Lewontin has always been quick to point out, sociobiology has primarily been concerned with telling stories about (for example) human behavior that

insinuate that such behavior is adaptive. Adaptationists regard the consistency of the trait with the hypothesis that it is an adaptation shaped by natural selection as good evidence for the conclusion that it is. Arguments such as Dawkins', which were meant to demonstrate that the focus of selective forces are individual genes made visible to selection through their direct relationship with particular phenotypic features, were then an integral part of the adaptationist program. These assumptions, when applied to human behaviors, quickly yield conclusions with dramatic political and social implications[6]. The attack on adaptationism expressed in *Spandrels*, and which Lewontin and Gould each pursued in many other works, would probably have been far less aggressive if the adaptive significance of the variation in the color of snail-shells were the only thing at stake.

Rather than viewing organisms as collections of more or less optimized individual traits, Gould and Lewontin called for the acknowledgment of other possibilities and for predictions, made on the basis of these alternative models, to be compared (Boxes 2 and 3). They argued that this entailed recognizing the legitimacy of different approaches to evolution in which constraints upon optimization through natural selection would play a more central role. Of course, Gould and Lewontin freely admitted that adaptation by natural selection has been, and still is, a powerful force in shaping the phenotypic traits of organisms. The question they wished to raise was: what kind of evidence is necessary to support the hypothesis that a trait is an adaptation formed by natural selection and what evidence could point towards some other cause of the current state of the trait in question?

Conceptual advances since the *Spandrels*

Since the publication of the *Spandrels* paper, theoretical evolutionary biology has made several advances that have helped researchers to question a purely adaptationist approach to the study of phenotypic evolution. In particular, the role of constraints, tradeoffs and costs in evolution has been widely discussed and generally acknowledged[7–9]. Although the terminology of constraints is still vague, most people seem to agree that at least genetic and developmental (sometimes referred to as epigenetic) constraints are a reality, and can be both measured in practice and accounted for in theory.

The common perception of the Fisherian idea that selection is an omnipowerful force, which always brings a

population to the maximum fitness peak available in the adaptive landscape (although Fisher's contribution was actually more subtle than that and did not make use of the landscape metaphor[10]), has been dealt a fatal blow by more recent advances in population genetics theory. Fisher's 'fundamental theorem of natural selection'[11], often perceived as the quintessential Panglossian paradigm, was originally developed for the special case of simple one locus–two allele systems (although it can be generalized for multiple alleles). However, even slightly more complex models, such as two loci with epistasis, fecundity selection, linkage disequilibrium and frequency dependence, will often (albeit not necessarily) result in adaptive landscapes characterized by maladaptive evolution in which selection drives the population 'downhill'[12]. More complex and more realistic models have proven too difficult for easy mathematical tractability. The currently accepted generalization is that complex genotypic architectures are less, not more, amenable to being altered by natural selection[13], that is, constraints are indeed a common and inescapable feature of living systems[14]. However, the debate still rages around how much genetic information is necessary to include in models of evolutionary trajectories, with population geneticists and optimality theorists discussing the possibility of a convergence of the two approaches – the so-called 'streetcar' model[15].

Another major (and slow in coming) conceptual shift in evolutionary biology theory, which has contributed to a widespread questioning of the adaptationist program, is the recognition of epigenesis as a central, and largely misunderstood, player in mediating the genotype–phenotype mapping function[16]. Perhaps in part because of the recent advances in molecular developmental genetics[17], not an original component of the adaptationist program, the simplistic idea of selection acting more or less directly on genes is being set aside as a serious contender in the evolutionary arena[18]. Even some sociobiologists have backed down from straight gene selection talk to embrace perhaps more vague, but certainly more realistic, 'epigenetic rules'[19]. However, a new brand of sociobiology, renamed 'evolutionary psychology', seems to be retracing its ancestor's path of mistakes by making wild claims on the genetic basis of human behavior while ignoring the two decades of debate about the adaptationist program[20].

These observations suggest that another way to avoid slipping into the adaptationist habit (suggested by Gould

and Lewontin themselves) is to think not about the current state of the trait, but how, given its phylogenetic history and functional significance, that trait should be. Although a purely engineering approach is not informative because it fails to account for historical pathways, once we take enough of the organism's basic developmental features into account, an analysis in terms of optimization theory can be revealing[21]. This can then lead to the specification of appropriate null hypotheses to test selective and constraint scenarios.

However, the role of constraints in evolution was probably oversold in the immediate aftermath of the *Spandrels*. This has led to an almost comic proliferation of types of constraints, resulting

in a rather confusing terminology and literature[22]. Recently, Schlichting and Pigliucci[23] have attempted to simplify and clarify the issue by proposing that all categories of constraints (with the exception of absolute constraints imposed by fundamental laws of physics) can be considered genetic or epigenetic in nature. This leaves selection and constraints as the two major deterministic players on the evolutionary stage; sometimes these are in opposition, but other times they work together to push a population in one direction or another in an adaptive landscape. It is this synthesis of constraints (spandrelism) and selection (panglossianism) that is the key to a more sober and realistic understanding of phenotypic evolution.

Box 1. The objects of the attack

According to Gould and Lewontin[1], the adaptationist program is characterized by the telling of stories involving natural selection, which account for the presence and particular forms of traits by reference to their hypothesized adaptive significance. Adaptationists assume the following:

Ontological assumptions
- Organisms can be usefully considered as assemblages of traits, the adaptive nature of which can be considered independently of the others.
- Constraints on the power of natural selection can be assumed to be minor, thus adaptation by natural selection is the proper null-hypothesis for particular traits, either behavioral or physical.

Methodological practices
- Consistency between the observed trait and an explanatory story told in terms of natural selection is sufficient for the preliminary acceptance of the hypothesis that the trait is adaptive and evolved in just that way.
- The failure of one adaptive story leads immediately to the search for another story of the same sort.
- Any failure of particular traits to be optimal is accounted for by evoking 'trade-offs' with other adaptive traits.
- Because of these features, adaptive stories are easy to create and hard to falsify – hallmarks, Gould and Lewontin argue, of poor scientific hypotheses. Indeed, they claim that it is difficult to ever reject the hypothesis that the trait is adaptive (in some way or other) within the context of the adaptationist program.

Box 2. Alternatives to adaptationism

Gould and Lewontin[1] suggested that at least the following alternative hypotheses be considered in attempts to understand the etiology of particular traits:
- No adaptation and no selection: the trait in question might be the result of genetic drift.
- Indirect selection: the trait in question is not the subject of selection – its features are the result of its association with another trait (which might or might not have been the subject of natural selection).
- Selection without adaptation: a trait might increase in frequency owing to natural selection but might not be 'adaptive' as generally understood (Lewontin's example involves a resource-limited species and a genetic mutation that doubles fecundity: this does not increase the population's mean fitness, it only alters the population dynamics).
- Adaptation without selection: the trait itself might be adaptive but not explicitly the product of selection for that particular form of the trait. Some types of phenotypic plasticity might be a special case of this, as is behavioral flexibility. Although these abilities might be selected for, the traits themselves (which are the result of the plasticity/flexibility) were not necessarily the result of selection for that form of the trait.
- Adaptation and selection, but no basis for distinguishing between adaptations: although the trait might be adaptive and have been selected for, there might be no way of distinguishing between different forms of a trait on the basis of their adaptive significance (the problem of multiple adaptive peaks).
- Adaptation and selection, but the particular adaptation represents a 'secondary' use of a trait already present for other (generally historical) reasons. This is the case to which Gould and Vrba[43] later referred to as 'exaptation'.

The other factor that has been increasingly taken into consideration by theoretical evolutionary biologists in the past decade is the fact that the environment (and therefore the adaptive landscape) is far from constant. Although models of environmental heterogeneity had been proposed before the *Spandrels* paper[24], the effects of environmental heterogeneity and the evolutionary strategies or outcomes that it can elicit have been investigated with a wider range of theoretical tools, including quantitative genetic and optimization models[25,26]. The general agreement is that the external environment poses as much of a limitation to adaptive evolution as internal (genetic–developmental, or epigenetic) constraints, and that they both constitute moving targets for natural selection[23]. Of course, this does not mean that organisms do not evolve in response to selection, but the currently available theory leads us away from the idea of survival of the fittest and towards a model of survival of the barely tolerable; this model is analogous to the concept of 'satisficing' in foraging theory[27].

Some of these ideas were actually expressed before the publication of the *Spandrels* by Francois Jacob[28], who elaborated a model of evolution as 'tinkering' (bricoleur, in the original French). According to Jacob, natural selection only works with the materials available and within the constraints present at a particular time in a particular place (see the metaphor of the builder using stones that fall from a cliffside in Darwin[29]). Jacob might have struck closest to the balance between spandrelism and adaptationism that we describe in this article.

Empirical advances since the *Spandrels*

Since the publication of the *Spandrels*, the roles played by constraints in evolution have been investigated empirically. We are now slowly gaining data on how common constraints actually are and on how they evolve when studied in the context of phylogenetic hypotheses. Of course, one must keep in mind that the fallacy of spandrelism is as easy to commit to as that of adaptationism. In particular, Gould has posited constraints for the land snail *Cerion*[30], whereas a computer simulation of the shell morphospace has actually indicated that such space is essentially saturated, with no evidence of a limitation on form[31].

The empirical study of constraints has been approached from different perspectives, but one of the most intriguing from an evolutionary standpoint focuses on research on the stability (or lack thereof) of genetic variance–covariance matrices. An example of this is the work of Conner and Via[32] on natural selection on body size in the flour beetle *Tribolium*. They estimated the intensity and the direction of selection, concluding that it would favor an increase in pupal weight and a decrease in width in male beetles. However, they detected a positive genetic covariance between the two characters. When they used this information in a model predicting the probable evolutionary response of the population, they obtained a small change in pupal weight (in spite of the strong selection coefficient) and a large increase in width (in spite of selection acting in the opposite direction). Several other studies of the stability of variance–covariance matrices have been published[33–36].

Even though constraints exist, is it possible to somehow 'break' them? In other words, how stable are the genetic covariances through evolutionary time? Several approaches have been used to answer this question. For example, Steppan[37] has reconstructed the evolution of variance–covariance matrices using a phylogenetic hypothesis describing the historical branching of the leaf-eared mice of the genus *Phyllotis*. He concluded that the matrix describing the evolution of cranial characters is variable within species, and that this variation sorts out taxa at higher phylogenetic levels. Therefore, Steppan concludes that there is 'no support (for) attempts to extrapolate structure to explain or predict macroevolutionary change'. However, experimental manipulations of genetic matrices by mutation-selection studies have also pointed towards the stability of at least some constraints in the crucifer plant *Arabidopsis thaliana*[38].

Immediately after the publication of the *Spandrels*, Lande and Arnold[39] developed a conceptually straightforward way to measure natural selection and to determine its magnitude and its type. The idea is to use a regression analysis of the trait(s) of interest against a reasonable measure of fitness. Many variations and improvements on this theme have been proposed since, including the use of genotypic rather than phenotypic data[40], path analysis[41] and nonparametric techniques[42]. Of course, demonstrating selection currently acting on a character is not the same as demonstrating that the character is an adaptation in the historical process sense of the term[43], but it is a necessary step to build the case for current adaptive value.

Such a case can then be expanded by the use of manipulative experiments, both under controlled and field conditions. These experiments have a long and venerable history in evolutionary biology[44], and can be used to bridge the gap between simply demonstrating the action of selection and actually identifying the causes underlying the observed selection gradients. Along similar lines, experimental evolutionary research has been conducted on a series of model organisms[45]. In these cases, one starts

out with a base population that is subjected to a novel environment. The evolutionary trajectories that yield an increased adaptation to that environment are then followed and compared with predictions based on *a priori* functional and quantitative genetic models. All of these approaches are yielding more satisfactory insights into the dynamics and constraints of adaptive phenotypic evolution. This information is crucial to assess the actual balance between non-adaptive forces and selection in natural populations.

The new synthesis between spandrelism and adaptationism

The difficulty of assessing where the true middle lies between spandrelism and panglossianism is to be found in the fact that biology is partly an experimental science, but partly a historical one. It is this historical component that makes it difficult, albeit not impossible, to deduce the underlying processes by observing the patterns they produce, because similar patterns can be the outcome of different processes. For example, we have mentioned the study of natural selection by the use of regression techniques. Even when successful, such techniques tell us little or nothing about the nature of the observed selection[46]. What aspects of the environment are exerting the selective force? How does variation in the genetic architecture of the traits under selection allow the population to respond? What other, unmeasured, traits might be driving the dynamics of the system? The answers to these questions do not come easily, and they can be sought only through a complex feedback between field studies and manipulative experiments, combined with mechanistic knowledge of the epigenetics of the traits themselves.

From the standpoint of theory, quantitative genetics and optimization theories have contributed substantially to our understanding of the interplay between selection and constraints. However, neither approach is without severe limitations[47]. As far as quantitative genetics is concerned, several authors have demonstrated that the relationship between the observable genetic variance–covariance matrix and the underlying genetic architecture is not as simple as it was once thought[48,49]. One cannot deduce the existence (or inexistence) of a constraint simply because a genetic correlation has been observed (or not). Optimization theory has shown its own limits by demonstrating that the more complex and 'multitasking' a phenotype is, the more likely one is to find many alternative 'optimal' designs[21]. Although

these limits are arguably a fact of life, it is clear that our current mathematical tools are sufficient to highlight them, but not to provide a comprehensive satisfactory theory of evolution on complex adaptive landscapes.

Of course, these are not reasons to despair, just inspiration for more work to come. However, the considerations submitted here do underscore not only that the world is neither made wholly of spandrels nor of panglossianisms, but also that simply saying that the true answer lies in the proverbial middle is not really saying much – unless one is prepared to roll up one's sleeves and thoroughly investigate that middle ground for decades to come.

References

1 Gould, S.J. and Lewontin, R.C. (1979) The Spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. London B Biol. Sci.* 205, 581–598

2 Rose, M.R. and Lauder, G.V. (1996) Post-Spandrel adaptationism. In *Adaptation* (Rose, M.R. and Lauder, G.V., eds), pp. 1–8, Academic Press

3 Mark, R. (1996) Architecture and evolution. *Am. Sci.* 84, 383–389

4 Wilson, E.O. (1975) *Sociobiology: The New Synthesis*, Harvard University Press

5 Dawkins, R. (1976) *The Selfish Gene*, Oxford University Press

6 Lewontin, R.C. (1992) *Biology as Ideology: The Doctrine of DNA*, HarperPerennial

7 Gould, S.J. (1980) The evolutionary biology of constraint. *Daedalus* 109, 39–52

8 Cheverud, J.M. (1984) Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* 110, 155–171

9 Maynard Smith, J. *et al.* (1985) Developmental constraints and evolution. *Q. Rev. Biol.* 60, 265–287

10 Frank, S.A. and Slatkin, M. (1992) Fisher's fundamental theorem of natural selection. *Trends Ecol. Evol.* 7, 92–95

11 Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*, Clarendon

12 Hartl, D.L. and Clark, A.G. (1989) *Principle of Population Genetics*, Sinauer

13 Cheverud, J.M. (1988) The evolution of genetic correlation and developmental constraints. In *Population Genetics and Evolution* (de Jong, G.D., ed.), pp. 94–101, Springer-Verlag

14 Baatz, M. and Wagner, G.P. (1997) Adaptive inertia caused by hidden pleiotropic effects. *Theor. Popul. Biol.* 51, 49–66

15 Marrow, P. and Johnstone, R.A. (1996) Riding the evolutionary streetcar: where population genetics and game theory meet. *Trends Ecol. Evol.* 11, 445–446

16 Alberch, P. (1991) From genes to phenotypes: dynamical systems and evolvability. *Genetica* 84, 5–11

17 Wray, G.A. (1994) Developmental evolution: new paradigms and paradoxes. *Dev. Genet.* 15, 1–6

18 Dawkins, R. (1998) *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder*, Houghton Mifflin Co.

19 Wilson, E.O. (1998) *Consilience: The Unity of Knowledge*, Knopf

20 Daly, M. and Wilson, M.I. (1999) Human evolutionary psychology and animal behaviour. *Anim. Behav.* 57, 509–519

21 Niklas, K.J. (1997) Adaptive walks through fitness landscapes for early vascular land plants. *Am. J. Bot.* 84, 16–25

22 Antonovics, J. and van Tienderen, P.H. (1991) Ontoecogenophyloconstraints? The chaos of constraint terminology. *Trends Ecol. Evol.* 6, 166–168

23 Schlichting, C.D. and Pigliucci, M. (1998) *Phenotypic Evolution. A Reaction Norm Perspective*, Sinauer

24 Levene, H. (1953) Genetic equilibrium when more than one ecological niche is available. *Am. Nat.* 87, 331–333

25 Gabriel, W. and Lynch, M. (1992) The selective advantage of reaction norms for environmental tolerance. *J. Evol. Biol.* 5, 41–59

26 Zhivotovsky, L.A. *et al.* (1996) Fitness patterns and phenotypic plasticity in a spatially heterogeneous environment. *Genet. Res.* 68, 241–248

27 Ward, D. (1992) The role of satisficing in foraging theory. *Oikos* 63, 312–317

28 Jacob, F. (1977) Evolution and tinkering. *Science* 196, 1161–1166

29 Darwin, C. (1868) *The Variation of Animals and Plants Under Domestication*, D. Appleton and Co.

30 Gould, S.J. (1989) A developmental constraint in *Cerion*, with comments on the definition and interpretation of constraint in evolution. *Evolution* 43, 516–539

31 Stone, J.R. (1996) Computer-simulated shell size and shape variation in the Caribbean land snail genus *Cerion*: a test of geometrical constraints. *Evolution* 50, 341–347

32 Conner, J. and Via, S. (1992) Natural selection on body size in *Tribolium*: possible genetic constraints on adaptive evolution. *Heredity* 69, 73–83

33 Brodie, E.D., III (1993) Homogeneity of the genetic variance-covariance matrix for antipredator traits in two natural populations of the garter snake *Thamnophis ordinoides*. *Evolution* 47, 844–854

34 Shaw, F.H. *et al.* (1995) Changes in genetic variances and covariances: G whiz! *Evolution* 49, 1260–1267

35 Johnson, N.A. and Wade, M.J. (1996) Genetic covariances within and between species: indirect selection for hybrid inviability. *J. Evol. Biol.* 9, 205–214

36 Roff, D.A. and Mousseau, T.A. (1999) Does natural selection alter genetic architecture? An evaluation of quantitative genetic variation among populations of *Allenomobius socius* and *A. fasciatus*. *J. Evol. Biol.* 12, 361–369

37 Steppan, S.J. (1997) Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. *Evolution* 51, 587–594

38 Camara, M. and Pigliucci, M. Mutational contributions to genetic variance/covariance matrices: an experimental approach using induced mutations in *Arabidopsis thaliana*. *Evolution* (in press)

39 Lande, R. and Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution* 37, 1210–1226

40 Rausher, M.D. (1992) The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* 46, 616–626

41 Crespi, B.J. and Bookstein, F.L. (1989) A path-analytic model for the measurement of selection on morphology. *Evolution* 43, 18–28

42 Schluter, D. and Nychka, D. (1994) Exploring fitness surfaces. *Am. Nat.* 143, 597–616

43 Gould, S.J. and Vrba, E.S. (1982) Exaptation – a missing term in the science of form. *Paleobiology* 8, 4–15

44 Clausen, J. *et al.* (1940) *Experimental Studies on the Nature of Plant Species. I. Effect of Varied Environment on Western North American Plants*, Carnegie Institution

45 Travisano, M. (1997) Long-term experimental evolution in *Escherichia coli*. VI. Environmental constraints on adaptation and divergence. *Genetics* 146, 471–479

46 Stratton, D.A. and Bennington, C.C. (1996) Measuring spatial variation in natural selection using randomly-sown seeds of *Arabidopsis thaliana*. *J. Evol. Biol.* 9, 215–228

47 Pigliucci, M. and Schlichting, C.D. (1997) On the limits of quantitative genetics for the study of phenotypic evolution. *Acta Biotheoretica* 45, 143–160

48 Houle, D. (1991) Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. *Evolution* 45, 630–648

49 Gromko, M.H. (1995) Unpredictability of correlated response to selection: pleiotropy and sampling interact. *Evolution* 49, 685–693

50 Steppan, S.J. (1997) Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution* 51, 571–586

The mid-domain effect: geometric constraints on the geography of species richness

Robert. K. Colwell and David C. Lees

Geographic patterns of species richness are influenced by many factors, but the role of shared physiographical and physiological boundaries in relation to range-size distributions has been surprisingly neglected, in spite of the fact that such geometric constraints lead to mid-domain richness peaks even without environmental gradients (the mid-domain effect). Relying on null models, several recent studies have begun to quantify this problem using simulated and empirical data. This approach promises to transform how we perceive geographic variation in diversity, including the long unresolved latitudinal gradient in species richness. The question is not whether geometry affects such patterns, but by how much.

Robert Colwell is at the Dept of Ecology and Evolutionary Biology, U-43, University of Connecticut, Storrs, CT 06269-3043, USA (colwell@uconnvm.uconn.edu); David Lees is at the Dept of Palaeontology and the Biogeography and Conservation Laboratory (Dept of Entomology), Natural History Museum, Cromwell Road, South Kensington, London, UK SW7 5BD (dcl@nhm.ac.uk).

The latitudinal gradient in species richness, perhaps the most conspicuous feature of global biogeography, has intrigued ecologists and biogeographers continuously since the times of de Candolle (Ref. 1, pp. 1270–1276) and Wallace (Ref. 2, pp. 65–68). In spite of this long history, consensus on the causes of tropical peaks in species richness remains elusive[3–5]. The overwhelming range of hypotheses proposed to account for changes of species richness (Palmer[6] lists 120 named hypotheses for variation in species richness or coexistence, and Rohde[7] identifies 28 specifically applied to the latitudinal gradient) makes the task of sorting out which factors are causal and which are incidental a daunting one. Increasingly, using computationally intensive methods, ecologists and biogeographers are looking for answers at a regional or a global scale, putting models and hypotheses[7–15] to the test against increasingly comprehensive distributional data encompassing the still poorly known tropics[15–19].

This search has taken an unexpected turn. In spite of the plethora of climatic, ecological, evolutionary and historical explanations proposed to explain biogeographic diversity patterns, something fundamental has, until recently, been almost completely ignored: the geometry of species ranges in relation to geographical boundaries. Based both on simulations[11,12,14] and on analytical null models[13,15], it is now clear that a mid-domain peak or plateau in species richness is inevitable for virtually any set of ranges, theoretical or empirical, when these ranges are randomly placed within a bounded geographical domain, in the complete absence of any supposition of environmental gradients within the domain. Qualitatively, this result survives a wide range of changes in model details and assumptions. Quantitatively, it explains a surprisingly large proportion of geographic variation in species richness for diverse groups of organisms, for the empirical cases so far examined. In spite of conceptual antecedents in the ecological literature dating back more than 40 years[20], and a substantial literature on formally analogous problems in niche overlap and phenology (Box 1), this important result seems to have gone unnoticed until 1994 (Ref. 11).

What do the geometric models predict?

There is a growing consensus that the regional pool of species is the key determinant of the species composition of more or less unsaturated local communities[21,22]. Based on regional biotas, geometric models[11,13–15] demonstrate that the stochastic placement of species ranges between shared geographic boundaries can generate precise predictions of species richness at points between the boundaries. For example, in the case of random, one-dimensional placement of ranges between two boundaries, the null models predict a convex, symmetrical pattern of species richness: this pattern is either parabolic[11,13] or quasi-parabolic[11,15] depending upon alternative distributions of range sizes and of midpoints (Fig. 1; Boxes 2 and 3). An instructive way to grasp this idea quickly is to experiment with a null model that simulates range size and randomizes range placement within differently defined boundaries[11] (Fig. 1). (A graphical simulation program, that not only implements the principal stochastic range model variants, but also allows input of empirical range-size frequency