

## Natural selection and history

John Beatty · Eric Cyr Desjardins

Received: 19 November 2008 / Accepted: 26 November 2008 / Published online: 14 January 2009  
© Springer Science+Business Media B.V. 2009

**Abstract** In “Spandrels,” Gould and Lewontin criticized what they took to be an all-too-common conviction, namely, that adaptation to *current* environments determines organic form. They stressed instead the importance of *history*. In this paper, we elaborate upon their concerns by appealing to other writings in which those issues are treated in greater detail. Gould and Lewontin’s combined emphasis on history was three-fold. First, evolution by natural selection does not start from scratch, but always refashions preexisting forms. Second, preexisting forms are refashioned by the selection of whatever mutational variations happen to arise: the historical order of mutations needs to be taken into account. Third, the order of environments and selection pressures also needs to be taken into account.

**Keywords** Adaptationism · History · Historicity · Chance · Contingency

### Introduction

It is common to explain the prevalence of a trait within a lineage by demonstrating how natural selection would favor the trait in the environment inhabited. Two prominent evolutionary biologists, Stephen Gould and Richard Lewontin, criticized such explanations for being insufficiently *historical*. A proper selectionist account, they argued, would focus on the past as well as the present, and the path from then-and-there to here-and-now. It would feature, among other things, the *ancestral* state

---

J. Beatty (✉) · E. C. Desjardins  
Department of Philosophy, University of British Columbia, 1866 Main Mall,  
Buchanan E-370, Vancouver, BC V6T 1Z1, Canada  
e-mail: john.beatty@ubc.ca

E. C. Desjardins  
e-mail: ericd@interchange.ubc.ca

of the lineage, and/or the *order of environments* inhabited by the lineage, and/or *the order of mutational variations* that arose within the lineage.

The most famous of their critiques, “The Spandrels of San Marco,” was aimed more specifically at what they called “the adaptationist program,” followers of which supposedly advocated:

the near omnipotence of natural selection in forging organic design and fashioning the best of all possible worlds. This program regards natural selection as so powerful and the constraints upon it so few that direct production of adaptation through its operation becomes the primary cause of nearly all organic form, function, and behavior. (Gould and Lewontin 1979, pp. 584–585)

There are two elements of the adaptationist program that are especially worth emphasizing in connection with the importance or unimportance of history. First, there is the view that natural selection leads to the production of “*optimal*” organic form (“the best of all possible worlds,” Gould and Lewontin 1979, pp. 581, 584, 585, 593, 597). If organic form is optimal, then considerations of (current) adaptive fit are sufficient in a way that they would not be if organic forms were suboptimal and some (historical) factors other than (current) adaptive utility had to be invoked.

Second, there is the view that natural selection leads to the “*direct* production of adaptation,” or as Gould and Lewontin also put it, natural selection leads to “*immediate* adaptation to local conditions” (1979, pp. 583, 584, 587, 590, 593; our emphasis). By “immediate” they did not mean instantaneous. Rather, they were referring to what they believed was the all-too-common conviction that adaptation to the *current* environment determines organic form. Thus they also objected to the exclusive consideration of “current utility” in explanations of form (1979, p. 581).

Has anyone ever defended such a position? Another prominent evolutionary biologist, John Maynard Smith, came close to endorsing it, at least on one (perhaps careless) occasion:

most populations have had time to come close to the *optimum for the environment in which they live*. (Maynard Smith 1975, 1993, pp. 11–12; our emphasis; but see also Maynard Smith et al. 1985)

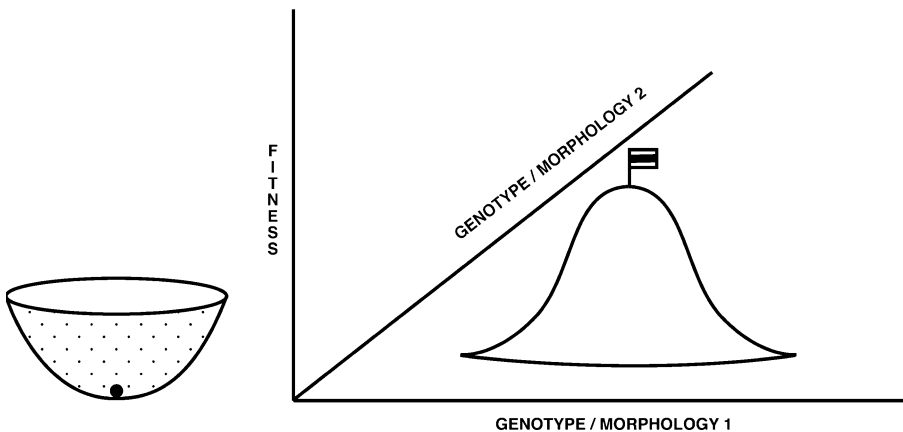
But Gould and Lewontin were for the most part not concerned to quote their colleagues. Their point was not that too many evolutionary biologists *explicitly* advocate the adaptationists’ conception of evolution by natural selection, but rather that too many evolutionary biologists *practice* evolutionary biology as if it were true, by ignoring history. If current utility *were* sufficient to explain the present genetic or morphological state of a population, then evolutionary biologists could justifiably ignore the history of that population. Its previous environments would be irrelevant. As would the ancestral states that had prevailed in previous environments, and the evolutionary paths taken *en route* to the current form. Nothing other than the current form and the current environment, and the respects in which the former is suited to the latter, would bear on the explanandum.

Such an a-historical explanation would be similar to the account we might give of a marble resting at the bottom of a bowl (see Fig. 1, left; our treatment of this issue follows Sober 1988, pp. 3–4).

The marble could have rolled there from anywhere inside the rim. Its equilibrium position is independent of its initial position, which is to say that history is causally irrelevant to the outcome. Moreover, history is not only irrelevant here; it is also difficult to recover. Once the marble has reached equilibrium, unless there has been some record of its trajectory (some tracks it left on the inside of the bowl perhaps?), then its history will have been lost; it cannot be inferred or reconstructed from its present state. Elliott Sober refers to the sort of process illustrated by the bowl example as “information-destroying” (1988, p. 3–4).

The analogous evolutionary situation can be somewhat similarly (if invertedly) represented with a so-called “adaptive landscape” (see Fig. 1, right). Here the  $x$  and  $z$  axes represent the range of genotypes or morphologies under investigation, and the  $y$  axis represents the adaptive values or fitnesses of the various genotypes or morphologies in a particular environment. According to the adaptationist assumption, the population under study would have inhabited the environment long enough to ascend by natural selection to the peak of the landscape, where it would then rest, regardless of where it began its climb. If we had no record of the process, then we would have no idea what route it took to that outcome.

Let us now turn to Gould’s and Lewontin’s arguments for the importance of history. Our sources will include not only “Spandrels,” but also work that they published prior to that collaboration. These more expansive discussions are important for illuminating the necessarily abbreviated treatment of history in their 1979 article. This also gives us the opportunity to discuss, in our conclusion, an important if problematic difference between the way that Gould and Lewontin treated history in these other publications, and the way that they treated it in



**Fig. 1** Examples of processes whose outcomes retain no trace of the past (see text). *Left*, a marble dropped from the edge of a bowl with one well will always end up in the same final state, regardless of its initial position. *Right*, an adaptive landscape with one peak that is eventually occupied regardless of the starting point

“Spandrels.” In the latter article, they portrayed historical approaches as worthy alternatives to the strictly adaptationist view. Whereas in the earlier works they portrayed evolution by natural selection as being by its very nature historical.

### When and why history matters—Lewontin

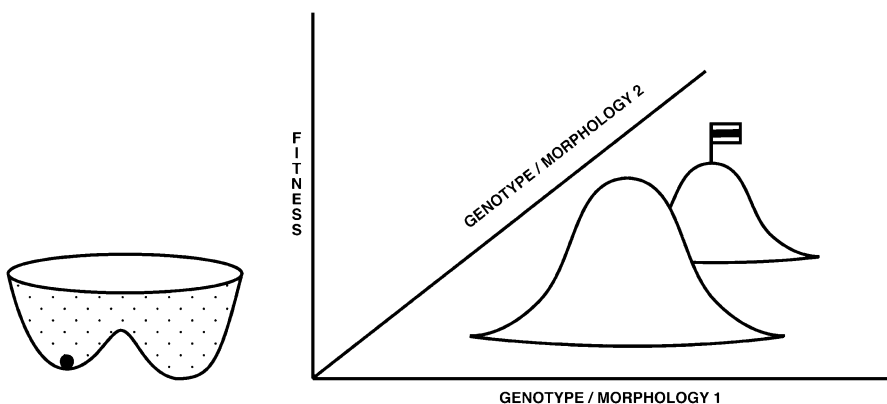
Lewontin had earlier criticized the tendency of evolutionary biologists to appeal to optimal equilibrium states in explanations of evolutionary outcomes. His concern was based on the fact that “equilibria annihilate history.”

It is the nature of an equilibrium point that all paths in the dynamical space lead to it (at least locally), so that the particular history of change is irrelevant and, once the system is at equilibrium, there is no trace left of historical information. (Lewontin 1974, p. 269)

By invoking optimal equilibrium states, then, investigators excuse themselves from the need to take history seriously. As Lewontin added, this is especially convenient because there is often no reliable record of the process anyway. “The banishing of history is almost a necessity because there is no historical record.... Although [history may be relevant in a particular case] there is no way to find out.” However, as he proceeded to remark, “that is not a sufficient reason to reject the [historical] explanation” (p. 269).

Actually, stable equilibria, in and of themselves, do not render history irrelevant, nor unrecoverable. When there are *multiple* equilibria, history can be quite important. Consider the position of a marble in a bowl with two wells (see Fig. 2, left).

The marble ends up on the left-hand side, but might have ended-up on the right instead, depending on where inside the rim it began its descent. Its final location



**Fig. 2** Examples of processes whose outcomes reflect their past (see text). *Left*, a marble dropped from the edge of a bowl with two wells will come to rest in one rather than the other depending on its initial position. *Right*, an adaptive landscape with multiple peaks, one rather than the other of which will eventually be occupied, depending on the starting point

would then reflect its initial position (or at least a limited range of initial positions). Sober refers to this sort of process as “information-preserving” (Sober 1988, pp. 3–4). Similarly, for a particular population in a particular environment, there might be two equilibrium states—two genotypes or morphologies of relatively high adaptive value, separated in genotypic or morphological space by forms of lower value that would not be traversable by natural selection (see Fig. 2, right). The evolutionary outcome would then depend on the genotypic or morphological starting point, together with the historical sequence of mutations that happened to arise. A population might rest on one rather than the other peak for the simple reason that mutational changes in that direction happened to arise and were selected, whereas mutational changes in the direction of the other peak did not occur and could not be selectively accumulated.

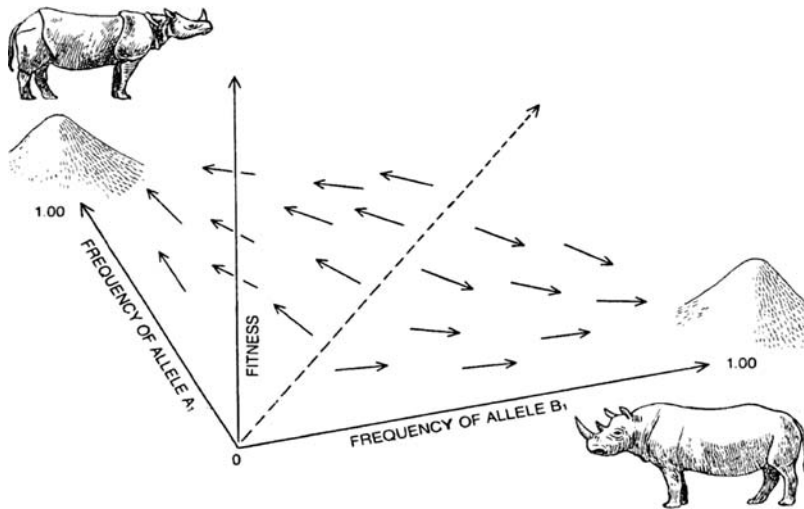
Lewontin employed precisely this line of reasoning in his 1978 paper, “Adaptation:”

The theory of population genetics predicts that if more than one gene influences a character, there will often be several alternative stable equilibria of genetic composition even when the force of natural selection remains the same. Which of these adaptive peaks in the space of genetic composition is eventually reached by a population depends entirely on chance events at the beginning of the selective process. (Lewontin 1978, p. 228)

So for example, two closely related and initially very similar populations or lineages, inhabiting very similar environments, might have different mutational histories. Thus, evolution by natural selection might result in two different gene combinations, both of which are more adaptive than the ancestral state. This difference would not reflect differences in the environments to which they adapted, but rather just differences in the historical order in which mutations arose and were selected.

Lewontin suggested that this might account for the fact that Indian rhinoceroses have one horn, while African rhinoceroses have two (see Fig. 3). That is, it is easy enough to imagine the protective advantages of horns in both environments, and how evolution by natural selection would have resulted in horns of one sort or another in both lineages. But, given that the two environments are very similar, it is difficult to imagine that the Indian lineage acquired one horn rather than two because of the advantages of this structure in this environment, while the African lineage acquired two horns rather than one because of the advantages of that trait in that environment. An alternative scenario would be that one-horned rather than two-horned variants first arose and were subsequently selected in the Indian lineage, whereas two-horned rather than one-horned variants first arose and were subsequently selected in the African branch. In this case, history would matter.

Actually, this is just one way that history might matter in making sense of the rhinoceroses’ horns. As Jonathan Coddington pointed out, there seems to be a crucial, unexamined assumption in Lewontin’s suggestion, namely that the ancestor of the Indian and African rhinoceroses had no horns (Coddington 1988). But the best evidence suggests that their immediate ancestor was two-horned. Why then would one horn come to prevail over the two-horned ancestral state in the Indian



**Fig. 3** Adaptive landscape showing how two lineages of rhinoceros might evolve in different directions, depending on which mutations arise, and are selected, in each lineage (see text Lewontin 1978, p 225)

lineage? This historical information about the ancestry of the lineages suggests that the two traits might be adaptations to different environments after all (however difficult it might be to conceive what advantages one trait would have over the other in the other environment). So while Lewontin focused on the importance of taking into account the historical order of mutations within a lineage, Coddington focused on the importance of taking into account the genealogical history of the lineages in question. (See the discussion of this case in Griffiths 1996, pp. 522–523; this essay is an excellent treatment of issues concerning the importance of history, mostly post-“Spandrels,” and especially with regard to the importance of cladistic analysis.)

Elsewhere, Lewontin stressed the importance of not only genetic history but also environmental history. In his 1967 article, “The principle of historicity in evolution,” he demonstrated that, in a fluctuating environment (and what environment does not fluctuate?), the *order* of fluctuations affects the evolutionary dynamics of a population inhabiting that environment. Imagine an environment in which some days/months/years are cooler than average, some warmer, some wetter, some drier, etc., and that selection for or against some genotype or morphology fluctuates accordingly. Imagine another environment in which there are exactly the same numbers of cooler, warmer, wetter and drier days, but in which the fluctuations above or below the average occur in a different order. Now imagine two initially identical populations inhabiting those two environments. As it turns out, their evolutionary trajectories may diverge considerably, as Lewontin showed by way of a simulation. In the first run, he randomly varied the selection values associated with three genotypes of a two-allele locus. In the second run, he applied the very same selection values, but in the reverse order.

The differences between the two pathways are striking (see Fig. 4); the second trajectory is by no means just the reverse of the first. For example, in the first run, the frequency of the charted allele was below 0.5 for most of the duration of the simulation, whereas in the second run, it was above 0.5 most of the time.

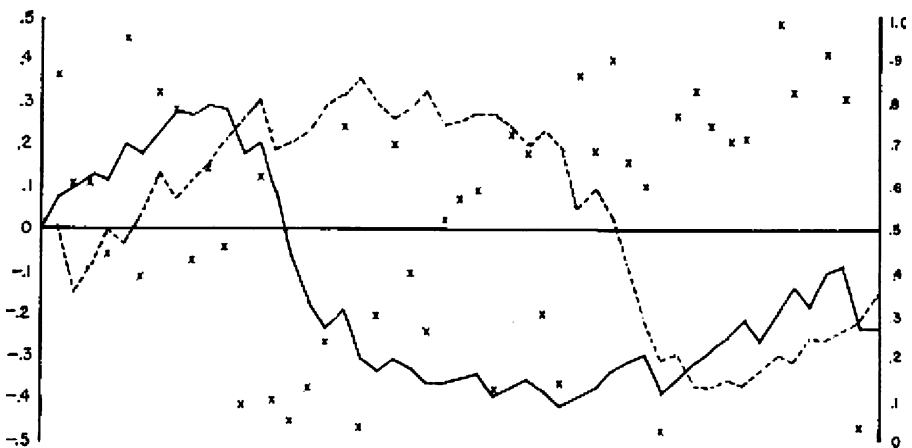
The main reason why the order of selection values makes such a difference resides in elementary population genetic theory, which describes evolutionary dynamics under Mendelian inheritance. It has to do in particular with the way in which the *rate of change* in gene frequency itself changes in relation to gene frequency. Consider the following two straightforward consequences of the Hardy–Weinberg principle, first for haploid and then for diploid populations. In each case we are considering one locus with only two alleles.

$$\Delta p = p(1 - p) \frac{W_{A1} - W_{A2}}{\bar{W}}$$

$$\Delta p = p(1 - p) \frac{p(W_{A1A1} - W_{A1A2}) + (1 - p)(W_{A1A2} - W_{A2A2})}{\bar{W}}$$

In the haploid case,  $p$  is the frequency of allele  $A1$ ,  $W_{A1}$  is the relative fitness of genotype  $A1$ ,  $W_{A2}$  is the relative fitness of genotype  $A2$ , and  $\bar{W}$  is the average fitness. In the diploid case,  $p$  is again the frequency of allele  $A1$ ,  $W_{A1A1}$  is the relative fitness of genotype  $A1A1$ ,  $W_{A1A2}$  is the relative fitness of genotype  $A1A2$ ,  $W_{A2A2}$  is the relative fitness of genotype  $A2A2$ , and  $\bar{W}$  is again the average fitness.

What is important to note here is that in both cases,  $\Delta p$  is proportional to the product  $p(1 - p)$ , which means that as the frequency of an allele approaches 1.0 or 0.0, its rate of change slows to a near standstill, responding less and less to the selective advantages or disadvantages associated with it. Consequently, a selection pressure of a particular magnitude will result in a larger or smaller change in allele



**Fig. 4** Evolutionary pathways with order of selection values reversed (see text). *Solid line*, allele frequency (*right-hand y axis*) changes in response to fluctuating environments. *Cross-marks* represent selection values (*left-hand y axis*) for and against the allele. *Dashed line*, allele frequency changes in response to the reverse order of selection values (Lewontin 1967, p 86)

frequency depending on the magnitude of the allele frequency at the time when the selection pressure is applied (and also depending on whether the selection pressure is—at the time it is applied—in the direction of an increase or a decrease in  $\Delta p$ ). Thus, for instance, a small positive selection pressure applied when the frequency of an allele is low (and when  $\Delta p$  is not very responsive), followed by a strong positive selection pressure (when  $p$  and also  $\Delta p$  is somewhat higher), will yield a greater change in the frequency of that allele than would a strong positive selection pressure applied when the frequency is low, followed by a small positive selection pressure.

Now look back at the solid line of simulated gene frequency changes. The starting frequency is 0.5. Recall that any particular selection pressure will result in a greater gene frequency change if the initial frequency is intermediate. Note also that the initial selection pressures included many large, positive values. Accordingly,  $\Delta p$  increased responsively, until several large negative selection pressures pulled  $p$  down so far that it hardly responded to subsequent, increasingly strong selection pressures, and ended up considerably lower than 0.5. The dashed-line trajectory of gene frequency changes (due to the same selection values applied in the opposite order) dropped responsively following an initial, very negative selection value. It then climbed in response to mostly positive selection pressures, to the point where it was only slowly dragged down by a sustained period of strong, negative selection. And it failed to bounce back in response to quite high, positive selection pressures toward the end.

The fact that a selection pressure of a particular magnitude will have a greater or lesser influence, depending on the gene frequencies at the time, means that the order of selection pressures can have a considerable impact on evolutionary *pathways*. In the case of diploids, the order of selection values can also have considerable impact on evolutionary *outcomes*. Playing with the following equation,

$$p_{t+1} = \frac{p_t^2 W_{A1A1} + p_t q_t W_{A1A2}}{\bar{W}}$$

(where  $\bar{W} = p_t^2 W_{A1A1} + 2p_t(1-p_t)W_{A1A2} + (1-p_t)^2 W_{A2A2}$ ) you can satisfy yourself that it requires a special combination of initial gene frequencies and fitness values in order to guarantee that even a *single, one-generation* change from one selection regime to another will give exactly the same outcome if reversed. (The effects of reversing a single change are small to be sure; the point is that *even* the reversal of a single change makes a difference; multiple changes over extended generations lead to greater differences.)

Implicit in all this is the perhaps unintuitive point that the evolutionary state of a population is sometimes a better reflection of its more distant than its more immediate environmental history. For example, the fact that the solid line trajectory of gene frequencies in Fig. 4 ended up well under 0.5 was certainly not due to the fact that there had just occurred a long stretch of strong selection in its favor! This was much more a matter of the preceding stretch of negative selection that had reduced its frequency to such a non-responsive point.

This can be more easily seen by looking at just one change in environments or selection pressures. Imagine a population with three genotypes *A1A1*, *A1A2*, and *A2A2*. Suppose that the population has long inhabited an environment in which



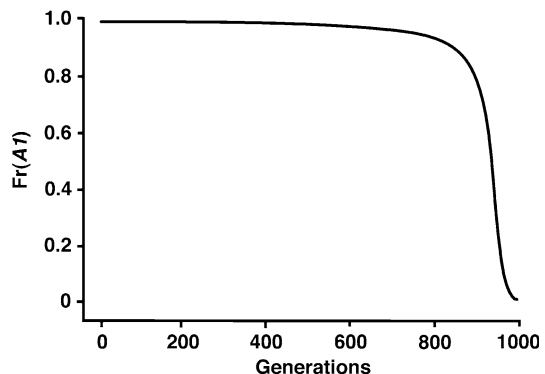
*A1A1* and *A1A2* have been favored over *A2A2*, so that the frequency of *A1* now hovers near 1.0. Now suppose that the environment changes substantially, so that *A1A1* and *A1A2* are selected against. For a long time, the frequency of *A1* will continue to hover near 1.0, suggesting falsely to an observer who notes only their static prevalence, that *A1A1* and *A1A2* are quite fit relative to *A2A2* (see Fig. 5). In similar fashion, a population starting at a very low gene frequency may remain in that state for a long time before undergoing a significant increase, regardless of strong selection in favor of the genotypes that include it. On the other hand, if the allele frequency does not reach such extreme values (i.e.,  $\sim 0.0$  or  $\sim 1.0$ ), then its rate of change will better reflect the recent selection regime.

### When and why history matters—Gould

Prior to their collaboration on “Spandrels,” Gould had also emphasized the importance of history, in several respects. For example, he had stressed the possible existence of something like multiple adaptive equilibria—or at least multiple, adaptively adequate responses to the same environment. Which of these outcomes would be realized depended, for him, not only on the order of mutations that arose for natural selection to act upon, but also—and perhaps most critically from Gould’s point of view—on the ancestral state that was being modified in the process and that acted as a “constraint” on the outcome. He liked to stress that this combination of indeterminacy and historical constraint allowed for, or almost guaranteed, sub-optimal results. One of the examples for which he was most famous was the panda’s “thumb,” a workable, opposed digit that was no digit at all, but rather an evolutionarily modified wrist bone. The lesson taught was that,

An engineer’s best solution is debarred by history. The panda’s true thumb is committed to another role, too specialized for a different function to become an opposable, manipulating digit. So the panda must use parts on hand and settle for an enlarged wrist bone and a somewhat clumsy, but quite workable, solution. (Gould 1978a, 1980, p. 24)

**Fig. 5** Changes in frequency of allele *A1* over 1,000 generations when the starting frequency of *A1* is 0.99, and the fitnesses of *A1A1*, *A1A2*, and *A2A2* are 0.9, 0.9, and 1.0, respectively (and when the population size is infinite; see text; simulated using *AlleleA1* software by Jon Herron)



As was Gould's style, he traced the lineage of his own reasoning back to Darwin, and especially to Darwin's work on orchids (Gould 1978a, 1980, *passim*). It is worth elaborating on this not-so-well-known work of Darwin in order to show how it illuminates Gould's own views of the importance of history.

In his book, *On the Various Contrivances by which British and Foreign Orchids are Fertilized by Insects* (1877), Darwin argued that as diverse as orchid flowers are, they are all three-petal, three-sepal devices serving basically the same function, namely to enlist flying insects in their cross-pollination (and thus to avoid inbreeding). These various "contrivances" for intercrossing had evolved, he believed, under virtually the same environmental circumstances, e.g., the same range of available insects. Sometimes one part of the flower had been modified to entice insects in the vicinity, by mimicry or by scent, or to trap them; sometimes another part had been modified to do the same job. Once the insects had arrived, the pollen had to be attached. Some flowers were so constructed as to catapult pollen at the visiting insects; some catapult the insects against the pollen; some simply induce the visitors to travel past and brush-up against the pollen, etc. Thus cross-pollination is accomplished in very different ways, the different outcomes being due in large measure—Darwin argued—to differences in the historical order of mutational variation among different lineages.

These contraptions are hardly optimal cross-pollination devices, constructed out of whatever parts would have been most suitable to the task; rather, they are workable modifications of the ancestral state, based on the selection/rejection of whatever variations on the ancestral theme happened to present themselves. They are, as Gould put it, "jury-rigged" or improvised, more in the manner of a handyman than God. As he quoted Darwin, it was "[as] if a man were to make a machine for some special purpose, but were to use old wheels, springs, and pulleys, only slightly altered.... (Darwin 1877, pp. 283–284, quoted in Gould (1978a, 1980), p. 26)

An example that particularly struck Darwin involved the position of the so-called "labellum" petal in the bog orchid, *Malaxis paludosa*. Gould did not discuss this example, but it optimally (!) illustrates the message he drew from Darwin's work, and so we present it here. In most fully formed orchid flowers, the labellum is the lowermost of the three petals. In that position, it often serves as a landing pad or a trap. But interestingly, during the development of the flower, the labellum actually starts out on top and then arrives at the bottom through a 180-degree twisting of the flower's stem (usually including the ovarium). Darwin reckoned that the position of the labellum in the ancestral orchid had been uppermost, presumably on the grounds that this is also the original position in development, and assuming more generally that the order of development reflects the order of ancestry. He understood the now-typical, lowermost position of the labellum to be an outcome of evolution by natural selection of the more twisted variations that had happened, by chance, to arise (Darwin 1877, p. 284).

But Darwin was especially intrigued by cases where the labellum had resumed its uppermost position, which in some cases had resulted from the selection of less and less twisted variations (so that the flower did not twist when emerging), and in other cases like *Malaxis paludosa* had come about as the result of selection for more and

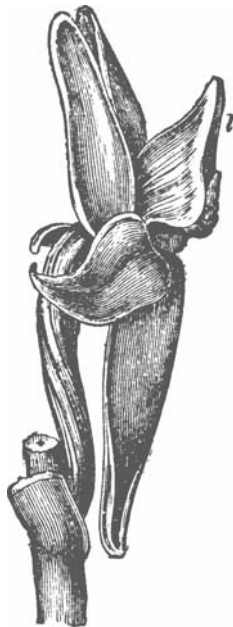
more twisted forms. Flowers of the latter sort twist a full 360 degrees just in order to resume their starting position (see Fig. 6)!

As Darwin described the situation,

... in many Orchids the ovarium (but sometimes the foot-stalk) becomes for a period twisted, causing the labellum to assume the position of a lower petal, so that insects can easily visit the flower; but ... it might be advantageous to the plant that the labellum should resume its normal position on the upper side of the flower, as is actually the case with *Malaxis paludosa*, and some species of *Catasetum*, &c. This change, it is obvious, might be simply effected by the continued selection of varieties which had their ovaria less and less twisted, but if the plant only afforded variations with the ovarium more twisted, the same end could be attained by the selection of such variations, until the flower was turned completely on its axis. This seems to have actually occurred with *Malaxis paludosa*, for the labellum has acquired its present upward position by the ovarium being twisted twice as much as is usual. (pp. 284–285)

So it had apparently become advantageous for *Malaxis* and some species of *Catasetum* to have their labellae uppermost. But due to differences in the order of mutational variations that happened to occur in the two lineages, evolution by natural selection had resulted in very different means of serving this end: a 360-degree twist in the first case, and no twist in the second (see also Lennox 1993; Beatty 2004).

Darwin came to believe that much of the diversity of life could be understood in this way: as different means to the same ends—alternative modifications of an ancestral form, brought about by natural selection of whatever variations happen to



**Fig. 6** *Malaxis paludosa*, with its 360-degree twisted flower stalk. The labellum petal is labelled “1” (see text; Darwin 1877, p. 130)

arise (see also later editions of the *Origin*, e.g., Darwin 1872, pp. 153–165). From Gould’s point of view, what Darwin’s work on orchids also demonstrated was the fact that, and the reason why, evolutionary outcomes are far from optimal. This is due not only to the possibility of multiple adaptively adequate solutions being arrived at by the selection of whatever variations happen to arise, but also to what Gould called “historical constraints.” The patent sub-optimality of *Malaxis paludosa*’s 360-degree twist, for example, is partly a matter of evolution by natural selection always having to depend on whatever mutations present themselves—in this case greater-than-180-degree instead of less-than-180-degree variations. But it is also a matter of evolution by natural selection always having to refashion existing forms rather than being able to start from scratch with whatever materials would be best for the occasion. The novel forms selected are always variants of preexisting forms. In the case of *Malaxis paludosa*, the starting point is the already twisted ancestor. Evolution by natural selection is in this sense “constrained by history,” as Gould liked to say.

Gould also liked to point out that Darwin had used the suboptimality of organic form to make the case for evolution by natural selection *rather than divine design*, and in this respect the orchid book was one of Darwin’s most pointed critiques of special creation. If orchids were the work of a creator, then he was more like an ordinary handyman than the Judeo-Christian God. As Gould himself summarized the point,

Odd arrangements and funny solutions are the proof of evolution [by natural selection]—paths that a sensible God would never have tread but that a natural process, constrained by history, follows perforce. (1979, p. 21)

Gould loved to tweak adaptationists for making the same mistake as special creationists (see also Gould 1978b, 1980, 1978c, 1980).

Gould invoked basically the same two notions of history—the historical order of mutational variations and historical constraints—in another of his typically historical-conceptual analyses, this time his treatment of the so-called “irreversibility” of evolutionary change, where he focused on the work of the turn-of-the-century paleontologist, Louis Dollo (Gould 1970; see also Gould and Robinson 1994). We discussed microevolutionary irreversibility (and its Mendelian basis) above. But the question of evolutionary reversibility or irreversibility usually concerns macroevolution. More specifically, it concerns the question what would happen if a lineage came to occupy an environment similar to one it had previously inhabited and to which it was previously, but no longer adapted. Would it reacquire the same adaptations it once had? Dollo posited what has since become known as “Dollo’s law,” namely that,

An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those in which it previously lived.” (Dollo 1905, p. 443, cited in Gould 1970, p. 196)

Dollo mostly presented empirical evidence, but he was also concerned to provide theoretical grounds for his generalization. He offered at least two different reasons. First, in order for reversibility to occur, the exact same variations would have to

reappear and be selected in reverse order. And this becomes more and more improbable in proportion to the number and complexity of the traits in question. As Gould quoted him,

The irreversibility of evolution is not simply an empirical law resting on facts of observation, as many have believed. It has deeper causes which lead it, in the last analysis, to a question of probabilities as with other natural laws. In effect, evolution is a summation of perfectly determined individual variations in a perfectly determined order. In order for it to be reversible, we would have to admit the intervention of causes exactly inverse to those which gave rise to the individual variations which were the source of the first transformation and also to their fixation in an exactly inverse order—a circumstance so complex that we cannot imagine that it has ever occurred. Otherwise, we might as well maintain that by throwing into the air the characters necessary for printing the Iliad, the poem would be completely composed by the simple fall of these little metallic blocks. (Dollo 1913, p. 59; translated by, and cited in Gould 1970, p. 199)

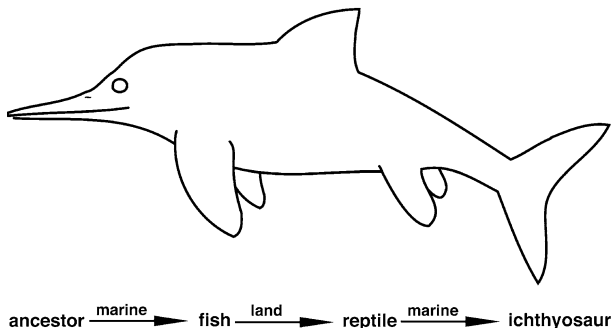
This is clearly related to the idea that the outcomes of evolution by natural selection depend on the historical order in which mutational variations arise.

A second reason for the irreversibility of evolution has to do with historical constraints. For a lineage to evolve from an initial state to a second state, and from the second state to a third, and then back to the second, this would involve reaching the same state, namely the second, from two different starting points, namely the initial and third states. But the more divergent the initial and third states are, the more unlikely this becomes, because natural selection would be accumulating variations on the initial type in the first transformation, and variations on the third type in the last transformation. The variations on the initial type would bear its historical stamp, and the variations on the third type would bear its historical stamp, and the outcomes would inevitably differ somewhat. As Gould quoted Dollo, this time continuing from the generalization concerning irreversibility,

An organism never returns exactly to a former state, even if it finds itself placed in conditions of existence identical to those in which it previously lived. But by virtue of the indestructibility of the past ... it always keeps some trace of the intermediate stages through which it has passed. The ichthyosaur, its lungs—*Zostera*, its flowers.” (Dollo 1905, p. 443, translated by and cited in Gould 1970, p. 196, and in Gould and Robinson 1994, pp. 368–369)

In other words, the marine reptilian ichthyosaurs were presumably derived from land reptiles, which were in turn derived from fish. And the ichthyosaurs do resemble fish in certain respects (see Fig. 7). But evolution was not completely reversed because the ichthyosaurs still had lungs and four rudimentary limbs, the virtually indestructible sign of their intermediate reptilian ancestry.

Similarly, the marine plants of the genus *Zostera* were presumably derived from land flowering plants, which were in turn derived from marine plants. And *Zostera* do resemble other marine plants in important respects. But again evolution has not been completely reversed, because *Zostera* still has the flowers acquired on land.



**Fig. 7** The ichthyosaur as an example of evolutionary irreversibility, due largely to historical constraints, but also to the happenstance of mutational variation (see text)

## Conclusion

In “The Spandrels of San Marco,” Gould and Lewontin said much less about the relevance of history than they had previously said separately. But what they did say was importantly—if also problematically—different. In “Spandrels,” they promoted what they called a “pluralistic” approach to evolution, in the same spirit as Darwin, who never relied exclusively on evolution by natural selection to explain organic form; he had also invoked, for instance, evolution by the inheritance of acquired characters. To be sure, some of Darwin’s alternatives had in the meantime been discredited. But Gould and Lewontin appreciated Darwin’s willingness to consider other possibilities, as opposed to the allegedly narrow outlook of the adaptationists, for whom natural selection is so powerful that it leads immediately to optimal adaptation and thus renders alternative evolutionary factors irrelevant.

In Darwin’s pluralistic spirit, we present ... alternatives to immediate adaptation for the explanation of form, function and behavior (p. 156).

Among the alternative factors, they included “history” in the sense of the “steps” (presumably including the order of mutational variations) leading a population in the direction of one rather than another adaptive peak, and also historical constraints in the sense of ancestral forms (no mention in “Spandrels” of the history of environments and selection pressures, which would have been more difficult to cover quickly). But in recommending these as “alternatives” to the adaptationist point of view, they seemed to endorse the idea that there is a non-historical kind of evolution by natural selection on the one hand, and on the other hand there is history, when what they had previously seemed at pains to point out was that evolution by natural selection *just is* historical, and deeply so.

First, evolution by natural selection is historical because it does not start from scratch, but always *refashions preexisting forms*. Evolutionary outcomes are dependent on ancestral starting points. Second, if we are talking about evolution by natural selection of *chance variations*, then the process is doubly historical, in that preexisting forms are refashioned by the selection of *whatever mutational variations happen to arise*. Evolutionary outcomes are dependent on the historical order of

mutation. Third, if we are talking about evolution by natural selection of chance *Mendelian* variations, then the process is triply historical. Because now the order of environments and selection pressures needs to be taken into account as well, as indicated by population genetic theory.

One might argue that in any particular case, the outcomes of evolution do not depend on historical factors. For example, two very closely related and hence genetically and morphologically similar lineages may come to inhabit different environments and diverge evolutionarily. In this case, the different outcomes would not be due to different historical constraints, and may as a matter of fact also not be due to differences in mutational history since the same mutations may have occurred in each lineage, and in the same order. Nonetheless, one would need to know something about the relevant ancestral and mutational histories to make this case. Or at least one would have to assume something about those histories.

Recall that Lewontin had sarcastically, but pointedly suggested that while evolution by natural selection may be an historical process, nonetheless the relevant history may be unrecorded and unrecoverable (perhaps Gould would have been less skeptical, at least with respect to historical constraints). Hence the temptation to wish that current utility is all that matters. However, wishing doesn't make it so.

**Acknowledgments** Thanks to Patrick Forber for a very careful reading and extensive comments. And thanks also to Richard Lewontin for clarifying some of the issues discussed here.

## References

- Beatty J (2004) Chance variation: Darwin on orchids. *Philos Sci* 73(5):629–641. doi:[10.1086/518332](https://doi.org/10.1086/518332)
- Coddington JA (1988) Cladistic tests of adaptational hypotheses. *Cladistics* 4:3–22. doi:[10.1111/j.1096-0031.1988.tb00465.x](https://doi.org/10.1111/j.1096-0031.1988.tb00465.x)
- Darwin C (1872) *On the origin of species by means of natural selection*, 6th edn. Murray, London
- Darwin C (1877) *On the various contrivances by which orchids are fertilised by insects*, 2nd edn. Murray, London
- Dollo L (1905) Les Dinosauriens adaptés à la vie quadrupède secondaire. *Bull Soc belge Geol Pal Hydr* 19:441–448
- Dollo L (1913) *Podocnemius congolensis*, tortue fluviatile nouvelle du Montien (Paléocène inférieur) du Congo et l'évolution des cheloniens fluviatiles. *Ann Mus Congo belge, Géol Pal Miner, serie 3, Bas et Moyen Congo* 1:47–65
- Gould SJ (1970) Dollo on Dollo's law: irreversibility and the status of evolutionary laws. *J Hist Biol* 3:189–212. doi:[10.1007/BF00137351](https://doi.org/10.1007/BF00137351)
- Gould SJ (1978a, 1980) The panda's thumb. In: *The panda's thumb: more reflections in natural history*. Norton, New York
- Gould SJ (1978b, 1980) Senseless signs of history. In: *The panda's thumb: more reflections in natural history*. Norton, New York
- Gould SJ (1978c, 1980) Double trouble. In: *The panda's thumb: more reflections in natural history*. Norton, New York
- Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc R Soc Lond B Biol Sci* 205:581–598
- Gould SJ, Robinson BA (1994) The promotion and prevention of recoiling in a maximally snaillike vermetid gastropod: a case study for the centenary of Dollo's law. *Paleobiology* 20:368–390
- Griffiths PE (1996) The historical turn in the study of adaptation. *Br J Philos Sci* 47:511–532. doi:[10.1093/bjps/47.4.511](https://doi.org/10.1093/bjps/47.4.511)
- Lennox J (1993) Darwin was a teleologist. *Biol Philos* 8:409–421. doi:[10.1007/BF00857687](https://doi.org/10.1007/BF00857687)

- Lewontin RC (1967) The principle of historicity in evolution. In: Moorhead PS, Kaplan MM (eds) *Mathematical challenges to the neo-Darwinian interpretation of evolution*. The Wistar Institute Press, Philadelphia, pp 81–88
- Lewontin RC (1974) *The genetic basis of evolutionary change*. Columbia University Press, New York
- Lewontin RC (1978) Adaptation. *Sci Am* 239(3):156–169
- Maynard Smith J (1975,1993) *The theory of evolution*, 3rd edn. Cambridge University Press, Cambridge
- Maynard Smith J, Burian R, Kaufmann S, Alberch P, Campbell J, Goodwin B, Lande R, Raup D, Wolpert L (1985) Developmental constraints and evolution. *Q Rev Biol* 60:285–287
- Sober E (1988) *Reconstructing the past: parsimony, evolution, and inference*. MIT Press, Cambridge