1

Macroevolution for plant reproductive biologists

PAUL WILSON

1.1 From micro- to macroevolution

Just as there is a microevolutionary process that explains organismal adaptations, so is there a macroevolutionary process that explains biological diversity. Consider western North America's wildflowers. How is it that there are 246 penstemons that are hymenopteran pollinated, and 40 penstemons that have taken on hummingbirds, but no penstemon species has adapted to fly or butterfly or beetle pollination? How is it that there are 60 kinds of dudleyas, all with ranges emanating from the coastal mountains? And how about mariposa lilies, a group of 35 species varying in flower colors and petal hairs yet all pollinated by both beetles and bees via a highly generalized floral mechanism? The amounts of diversity and the patterns in which they are arranged are the products of a macroevolutionary process.

The microevolutionary process is more familiar. Mutations occur from time to time. They are undirected. Many are deleterious to the functioning of the organism in its environment. For a while they contribute to the genetic load, then eventually they are lost due to natural selection. Many other mutations are neutral or nearly neutral given the environment where the organism lives and the genetic state of the organism at other loci. Neutral alleles change in frequency due to genetic drift. A few new mutations are beneficial to the individuals that carry them, or to

---

Evolution of Plant–Pollinator Relationships, ed S. Patiny. Published by Cambridge University Press.

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

their close relatives, and these are selected up in frequency. The beneficence of these alleles may depend on the outside environment, for example on the kinds of animals that are pollinating those plants in their local population. Likewise, whether or not an allele is beneficial may depend on the genetic state of the rest of the organism. If the outside environment or the genetic background change, then what was once deleterious or neutral may become beneficial.

This dynamic that happens at the microevolutionary scale has an analogue at the macroevolutionary scale, what Stephen Jay Gould (2002) called a "grand analogy." The analogue to selection among individuals within a population is selection among clades in a biota. Clades with certain character states diversify more. For example, flowers with nectar spurs have higher rates of diversification than flowers without nectar spurs (Hodges 1997a). The analogue of mutation is the punctuation in punctuated equilibrium, the shift to a new adaptive state, such as when an isolated population shifts to a different pollinator. The analogue of genetic drift in allele frequencies at a locus is clade drift in the frequency of species having a particular trait in a region's biota. In Gould's hierarchical process, characters come to be fixed in a lineage through organismal adaptation, and then those fixed differences among lineages become the criteria for selection at a higher level.

Individual selection along with some other microevolutionary ingredients such as mutation and drift are mainly what is responsible for the adaptations of organisms: how a bee-pollinated penstemon has come to have purple vestibular flowers that make nectar of a certain sort and have a staminode for levering the anthers and stigma onto the bee's back in a certain way, etc. But there is more to explain about life than just the adaptations of this or that flower: there is the amount and pattern of biodiversity. Clade selection and other macroevolutionary ingredients are responsible for the diversity of organisms: how many species of penstemons there are, how many are specialized for pollination by bees versus birds, the size of penstemon geographic ranges, the way they remain clustered into groups nested within the larger penstemon clade, how each of those smaller groups is characterized, the disparity of specializations within the groups, etc. (Wilson et al. 2006). Microevolution and macroevolution work together and end up affecting one another. Together they constitute one unified machine that generates order out of history.[1]

---

[1] Those who are reading for pleasure should ignore my footnotes. My chapter is an introduction to hierarchical evolution aimed at people interested in pollination. As such, I have refrained from reviewing many philosophical distinctions and historical debates. For instance, I do not review the claims of Gould and associates circa 1980 and the criticisms of those early attempts. An improved and less controversial version of how hierarchical evolution works followed from a change in definitions announced in Gould and Lloyd (1999). Refinements to the grand analogy beyond Gould (2002) are continuing, and I here add some of my own. True, the logic of hierarchical evolution could use some

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

1.2 Four forms of clade selection

A key innovation is a derived feature of a lineage that leads to greater diversity than would otherwise arise. Typically, this is detected by finding more species in the clade with the innovation than in a sister clade without the innovation (Kay et al. 2006). Nectar spurs seem to have led the groups that possess them to be species-rich compared to sister clades. Such a key innovation can work by either increasing the rate of speciation or decreasing the rate of extinction, and new statistics are starting to allow people to tease apart the two (FitzJohn et al. 2009). Evolutionary biologists have gotten used to thinking of speciation and extinction, but I shall ease into my developing argument by using slightly different language. I invoke *cladogenesis* (which is like speciation without focusing on the point when reproductive barriers become permanent) and *persistence* (which describes a lineage before its extinction). There are two causal paths for an innovation to be favored: it may be favored via increasing the likelihood of cladogenesis or increasing the likelihood of a clade persisting. Innovations may also be disfavored via lowering the likelihood of subsequent cladogenesis, or more generally, lowering the time that lineages are likely to persist. Consider all four cases (Fig 1.1).

(1) An innovation that favors cladogenesis is bilateral symmetry in flowers. Sargent (2004) found that groups with bilaterally symmetric flowers had more species in them than sister groups. Kay et al. (2006) worry that six of 22 sister groups show the reverse pattern, with radially symmetric flowers being more species rich. Nevertheless, the pattern seen in most cases is that bilateral symmetry increases the rate of cladogenesis. How would this work? Flowers that guide their pollinators to visit in a receiving line place pollen on the pollinator more accurately (Armbruster et al. 2009b). That tends to promote speciation in the form of a reproductive isolating barrier whereby different kinds of pollinators are employed or different areas of the pollinator's body are used by different plants. In addition to presenting numerous species to pollinators, having high rates of cladogenesis might multiply the chances that a sub-lineage of a

---

scholarly help, but that should be done elsewhere. I add only a very few footnotes to help readers who are of a more critical mind.

The overarching semantic debate would be whether it is better to draw a grand analogy or to use verbiage at the level of clades distinct from the verbiage of microevolution. Many would prefer to not use loaded terms like clade selection as an analogue to individual selection because by their definitions there is only one kind of selection. They use other terminology to write about the phenomenon of some clades being more successful than others because of the traits of those clades. I choose to put as much as possible into a theory of hierarchical evolution. I do, however, believe the hierarchy should be presented with some exploration of how the two levels are not parallel.

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny

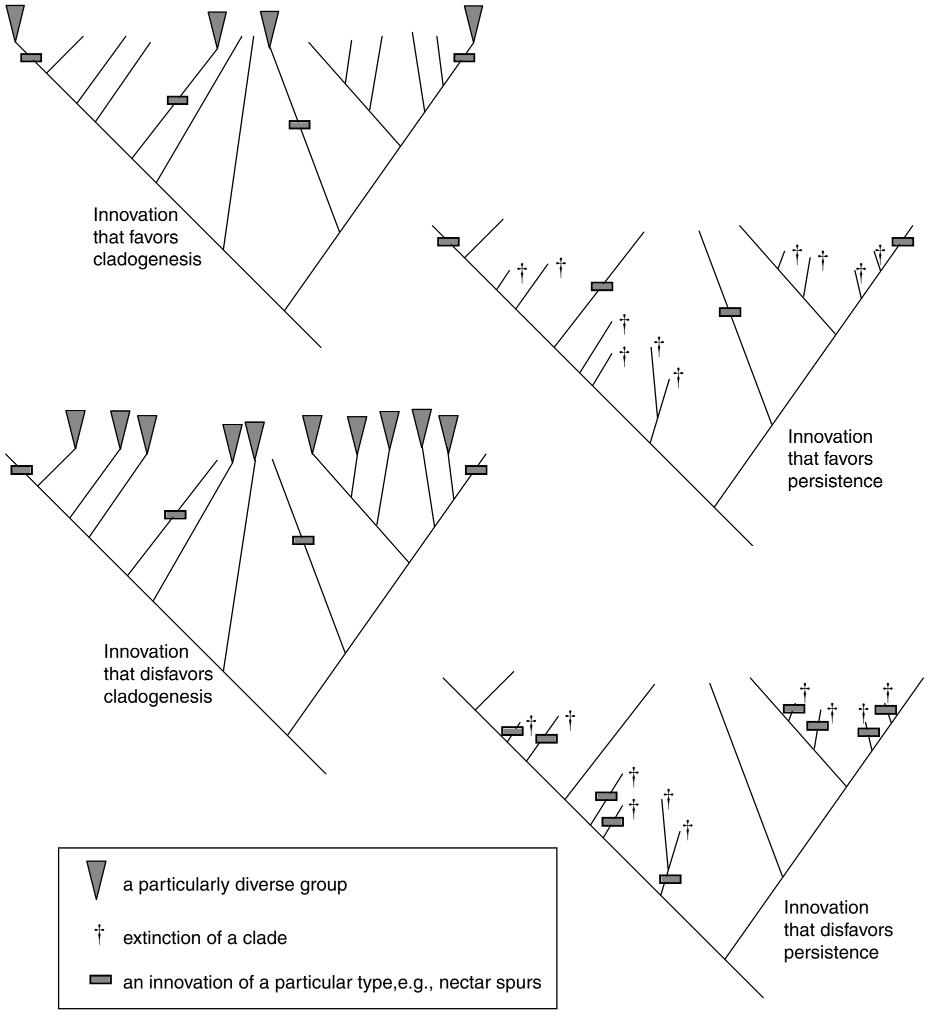

Fig 1.1 Four forms of clade selection affecting diversity.

clade will survive through catastrophes. Thus, an innovation that favors cladogenesis has two effects: in a snapshot in time, the groups with the innovation have many species; and in the long term, as ecological divergence proceeds, the clade as an aggregate is likely to have varied chances of surviving.

(2) An innovation that favors long persistence by means other than the multiplication of its clades might be the evolution of seed dormancy. Seed dormancy allows seeds to survive in a seed bank for longer than the seeds of cousins that lack seed dormancy. I know of no phylogenetic analysis that shows this pattern, but a bit of inspiration can be drawn from work done on an ecological time scale. Kalisz et al. (1997) have used population data on blue-eyed

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

mary to parameterize a model showing how a seed bank buffers a population against the vagaries of bad years. Stöcklin and Fischer (1999), reporting on a grassland community, found that species with seeds that live for more than five years are less likely to go to extirpation than species with short-lived seeds. On a longer time scale, clade selection in favor of seed dormancy seems likely. If one had quantitative measures of seed dormancy for a group of species and a phylogeny relating the species, one could test for a phylogenetic effect of seed longevity on how deeply rooted the dormant clades are. It would also be worthwhile to see if the effect was contingent on the life history of the plants involved. The clade selection might be stronger in annuals than in perennials. It might also be stronger in biomes with highly stochastic weather than in biomes where rainfall is relatively constant. Comparative tests of hypotheses about seed dormancy will surely be complicated (Baskin and Baskin 1998).

(3) An innovation that disfavors cladogenesis is the shift to abiotic pollination (Dodd et al. 1999). Most major lineages of flowering plants were once animal pollinated and those that are still animal pollinated have high rates of cladogenesis, but those that have gone over to wind or water pollination have lower rates of cladogenesis. Why? First, animal pollinators tend to be picky about the appearance of the flowers they visit. For example, individual bees are prone to becoming temporarily constant to a particular color or appearance of flower (Gegear and Laverty 2005). In a local community, flowers that evolve to be distinctive in appearance compared to co-flowering species encourage such constancy and thereby have their pollen moved with less wastage and purer delivery (Wilson and Stine 1996). This may even cause ecological sorting allowing species with distinctive colors to become abundant in their community (McEwen and Vamosi 2010). In addition to appearance, the mechanical fit of flowers around pollinator bodies is probably selected to be as efficient as possible (Castellanos et al. 2003). Second, pollinators differ discontinuously in physical dimensions, so flowers pollinated by different types of animals could be experiencing diversifying selection (Wilson and Thomson 1996). Finally, aside from being an organ of local differentiation (Johnson 2006), when divergent flowers come back together in sympatry, the functional variation is grist for positive assortative mating whereby similar flowers mate with one another. The assortative mating is caused by animals having a behavioral tendency to categorize (Jones 2001), and the assortative mating maintains or even adds to genetic correlations. Genetic correlations in turn predispose lineages to evolve reproductive isolating barriers (Kondrashov and Shpak 1998). When a clade changes from animal pollination to pollination by wind or water, there is then a relaxation of the tendency towards subsequent cladogenesis.

(4) An innovation that disfavors persistence might be the evolution of separate sexes. Dioecious clades have been found to have fewer species than their sister clades that are co-sexual (Heilbuth 2000), and they tend to have more endangered and threatened species (Vamosi and Vamosi 2005). Compared to hermaphroditic lineages, dioecious plants are likely to be inferior at establishing new sub-populations in the meta-population dynamic. Also, dioecious plants have more of a seed-shadow handicap whereby seedlings are clumped around mother plants and compete with each other to the detriment of the population. Finally, dioecious plants have a stronger reliance upon pollinators in the face of stochastic variation in pollinator services. Not only is self-pollination impossible, but there is more of a chance that neighboring plants will be of the same gender. The phylogenetic patterns could be because dioecious clades have low rates of cladogenesis, but it seems more likely that they have a higher rate of extinction than co-sexual clades. In other words, dioecy evolves from time to time but tends to be an evolutionary dead end.[2]

1.3 Other macroevolutionary ingredients

Just as clade selection is an analog of allelic selection, so there is a process of clade drift that is analogous to genetic drift. Gould suggested that at the macroevolutionary level drift might be more important than it is at the level of sexual individuals adapting to their surroundings. More generally, there is the possibility that ingredients that are most important for macroevolution might not be parallel to ingredients that are most important for microevolution. At any rate, several ingredients other than selection need to be introduced as I proceed to layer my argument for recognizing a hierarchical evolutionary process.

---

[2] I choose to use clade selection generally and species selection as a special case. Gould used the word “species” a great deal, as in species selection, species drift, and directional speciation. He defended punctuated equilibrium at the species level; at levels above the species level, he would say the dynamic was punctuational. I am reluctant to extend this usage (Mishler 2010). It seems particularly odd to speak of species selection resulting from characters acquired in a lineage making it more prone to subsequent cladogenesis and through that proliferation to the clade’s extended life than if the characters were otherwise. I am more comfortable speaking of species selection when its mechanism is to delay extinction by some means other than favoring additional cladogenesis, but the term clade selection works in all cases. I probably picked up my usage by taking a class from George Williams, who considered it a fallacy to tie much of anything to the taxonomic species level (1992 starting on p. 118). “Clade” is actually also objectionable because, unlike an individual, a clade includes its descendants (Okasha 2003). A clade is like a family, so the strict analog of clade selection would be clone selection.

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

As a way of appreciating clade drift, consider the role of founder effects in the Hawaiian Islands. The original colonizers dispersed a fantastic distance and were not absolute outcrossers, yet aside from these traits (which need not have been innovations), once a founding colonist got to Hawaii and established a population, it had a very good chance of undergoing an adaptive radiation. Ricklefs and Renner (2000) sought explanations for Hawaii's radiations and found little more than the usual tendency for animal-pollinated groups to diversify. Thus, it would seem that local populations have adapted, each guided by individual selection, while at the level of the larger clades, those clades that diversified and those that did not have been largely drawn at random with respect to their traits. An appreciation for both levels in the hierarchy greatly aids understanding of the process.

Founder effects are not the only form of clade drift. The frequency of a trait among species in a biota on a phylogeny naturally staggers up and down even without the trait of interest causing the changes in its frequency. Differences might *seem* big in terms of numbers of species, but from a purely statistical standpoint, all possible partitions of species richness into two groups are equally likely (Slowinski and Guyer 1993). Maybe the number of petals has experienced clade drift. Vast swaths of angiosperm diversity have five-parted flowers; fewer, such as mustards, have four-parted flowers. The monocots are the main group that has three-parted flowers. Even if the number of parts has not causally affected the diversification process, the frequencies of five-, four- and three-parted flowers would still have changed as the tree of life has grown. So far as I know, clade drift could have accounted for the way in which biogeographic provinces around the world have different proportions of their floras that are five-, four- and three-parted.

Another macroevolutionary ingredient is clade hitchhiking. If you were taught evolution by focusing on quantitative traits being optimized in sexually reproducing organisms with selection among individuals, then you are not likely to think hitchhiking is a very important feature of evolution. If you were taught evolution by comparing DNA sequences for a particular gene where there is little chance of recombination, you are more likely to be familiar with the idea. Hitchhiking is when one character's frequency is dragged quickly up or down by linkage to another character that is undergoing a selective sweep (Barton 2000). Orchids have inferior ovaries, and there are many orchids, so maybe inferior ovaries favor cladogenesis. But maybe not. For orchids, it is possible that having inferior ovaries is merely coincidental with other traits that favor diversification: like having bilaterally symmetric flowers, stigma and anthers on a rigid column, pollen dispersed in pollinia, tiny seeds, ectomycorrhizae, the proclivity to grow as epiphytes, etc. The inferior ovary might have been dragged to high frequency by clade hitchhiking. Because there is usually no recombination between clades, clade selection is similar to selection in a strictly asexual organism, or selection on a stretch of DNA that does not recombine, and hitchhiking is to be expected.

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

Examples of clade drift and clade hitchhiking bring up the possibility that nondeterministic factors can explain many macroevolutionary patterns. A particular ancestor was at the right place at the right time and so it gave rise to a diverse lineage. It had a peculiar feature, not generally superior to alternative character states, and that feature was lucky to get to go along for the ride in a group that for other reasons diversified. Gould gave the name “contingency” to the way such arbitrariness can be propagated, and he suspected that many of the great successes and failures in the pageant of life were contingent turns of history, rather than determined by the features of the clades.

Let’s say, the contingent bit of luck is dispersal to a different biogeographic province. That dispersal event might have the effect of favoring cladogenesis, favoring clade persistence, disfavoring cladogenesis, and/or disfavoring clade persistence. Moreover, luck and innovation could interact: the value of an innovation could be contingent upon the dispersal event. Moore and Donoghue (2007) considered both dispersal and innovation, looking to see if either or both might affect diversification in the Adoxaceae and Valerianaceae. They looked for changes in diversification rate anywhere in their phylogenies, and then they tested for associations with characters and with dispersal events. Of seven phylogenetic segments where diversification rates shifted into high gear, three were associated with dispersal to a new province. One of those also might have been associated with a decrease in stamen number. No specific reason was found for the remaining four increases in diversification rate.

The final evolutionary ingredient to be introduced early in my chapter is less whimsical and more tractable than drift or hitchhiking. It is transitional drive. Transitional drive corresponds to what is called mutational drive at the level of the gene. Mutational drive is an inequality in the direction of mutations. For example, if mutations from allele *Violet* to allele *White* are very frequent but mutations from allele *White* to allele *Violet* are nearly impossible, then over time a population of violet flowers will be converted to a population of mostly white flowers (if selection is negligible). At the level of clades, transitional drive is an inequality in shifts from adaptive mode *A* to *B* versus from adaptive mode *B* to *A*. An example of transitional drive is found among penstemons. Hummingbird pollination has evolved from hymenopteran pollination many times, and there is no evidence for any reversals, nor have there been shifts to any other pollination syndrome (Wilson et al. 2007).

Transitional drive may figure largely in macroevolution, whereas at the level of organismal evolution it is rarely treated as a very strong ingredient. Within populations, mutation–selection balance on flower color seems to keep albino-flowered individuals very rare despite mutational drive that eliminates floral pigments (Waser and Price 1981). At the macroevolutionary level, eudicots started out having purplish anthocyanin-pigmented flowers, but a great many lineages have transitioned to warmer or paler colors and few have transitioned back to purple

(Rausher 2006, 2008). Perhaps color has not had a consistent effect on cladogenesis or clade persistence, so transitional drive would seem to greatly explain why now there are untold numbers of eudicot species with flowers that are red, orange, yellow, or white.[3]

1.4 A full-blown example: the evolution of selfing

With these evolutionary ingredients in mind, contemplate an extended example: the rise and fall of self-pollination in the flora of a region. Self-pollination is when pollen on an individual plant ends up siring seeds on the same individual. It is not asexual: meiosis and fertilization remain part of the genetic lifecycle. However, as for asexuality, some of the inefficiency of outcrossing and the cost of male function can be saved by selfing. For example, a population of selfers would be expected to have a growth rate higher than a population of outcrossers if all other things were equal.

But all other things are not equal. There are often fitness advantages to outcrossing. If the particular species has been outcrossing for a large number of generations, then deleterious recessive alleles will have built up in the gene pool. This is the dominance genetic load that is carried around by diploid organisms such as poppies and people. If, from this state, a plant self-pollinates, homozygocity will increase. Selfing brings together recessive deleterious alleles, and this makes for seeds and seedlings plagued with genetic disease. Across much of angiosperm diversity, outcrossing has been maintained as the norm (Stebbins 1957).

Nevertheless and despite the norm, selfing has arisen over and over in monkeyflowers and lupines, gilias and lotuses, and collinsias and drabas. In scattered species, selfing becomes habitual (although not necessarily obligate). Anthers and stigmas evolve to mature at the same moment, to have no positional separation, and to be self-compatible. Nectar evolves to nothing. Petals become diminutive. Pollen production declines. All these traits tend to evolve together or as a close cascade (Cruden 1977; Aarssen 2000).

Near the microevolutionary scale, there are many circumstances that can favor selfing. Perhaps a population finds itself in a situation where pollinators are scarce. Perhaps selfing allows the plants to set seed quickly over a growing season that has

---

[3] Vrba and Gould (1986) distinguish upward versus downward causation in the hierarchical evolutionary process. Transitional drive upwardly causes patterns among clades. Selection downwardly causes patterns among gene frequencies. In this passage, I presume flower color is acted on by individual selection and is adaptive at the level of the organism. This translates into drive among clades. I entertain the possibility that although flower color is selected upon at the individual level, at the clade level its frequency distribution might be determined by transitional drive (caused by selection at a lower level) plus clade drift, and not by clade selection.

Cambridge University Press
978-0-521-19892-9 - Evolution of Plant–Pollinator Relationships: The Systematics Association Special: Volume 81
Edited by Sébastien Patiny
Excerpt
More information

become compressed compared to the growing season of ancestors. Perhaps the costs of putting on a show for pollinators and rewarding them with nectar have become exorbitant because the plants, though they once lived in good conditions, are now living in harsh conditions. All these reasons can be considered aspects of the “efficiency of selfing.”

Another class of reason for the evolution of selfing is often called “the two-fold advantage.” This is not a way in which selfing is advantageous to the health of the organism, rather it is a way in which selfish alleles tend to spread in the population. If, in a population of outcrossers, a mutant arises that makes its bearers put their pollen on their own stigmas and still present about as much pollen to pollinators for outcrossing, then that allele will nearly double its success via male function. There are, however, complications that might make the two-fold advantage less than two-fold.

(1) The mechanism by which the flowers self-pollinate might be that pollination happens quickly and pre-empts outcrossing. In this case, as the population becomes dominated by plants that never present fresh stigmas, which are available to receive outcrossed pollen, the two-fold advantage goes to zero.

(2) By selfing, the plant may use up some of the pollen that would have been available for outcrossing. If so, then the two-fold advantage would be reduced by pollen discounting (Holsinger and Thomson 1994; Harder and Wilson 1998).

Presumably because of the efficiency of selfing and the up-to-two-fold allelic advantage, there are many species that have transitioned to become selfers, but it is hypothesized that selfing clades tend to be dead ends (Stebbins 1957). The dead-end-selfers hypothesis asserts first, that there is transitional drive from outcrossing to selfing, and second, that it is counteracted by clade selection favoring outcrossing clades.

The first assertion, that of transitional drive, is expected since, as selfing becomes the norm for a species, deleterious recessive alleles ought to be purged, inbreeding depression ought to be lessened, and there would then be less of an immediate microevolutionary selective reason for outcrossing (Lande and Schemske 1985). This diminution of selection-for-outcrossing ought to remain even after the species enters better environments where selfing would not have been favored in the first place. The reasons why a species evolved from outcrossing to habitual selfing do not work in reverse.

The second assertion, that of clade selection favoring outcrossing, is suggested by the facts that most selfing species are closely related to outcrossing species and few large genera consist only of selfers. Eventual extinction of selfing lines compared to outcrossing clades can proceed by either of two mechanisms.