The role of biotic forces in driving macroevolution: beyond the Red Queen

Kjetil L. Voje, Øistein H. Holen, Lee Hsiang Liow and Nils Chr. Stenseth

Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, PO Box 1066 Blindern, Oslo 0316, Norway

A multitude of hypotheses claim that abiotic factors are the main drivers of macroevolutionary change. By contrast, Van Valen’s Red Queen hypothesis is often put forward as the sole representative of the view that biotic forcing is the main evolutionary driver. This imbalance of hypotheses does not reflect our current knowledge: theoretical work demonstrates the plausibility of biotically driven long-term evolution, whereas empirical work suggests a central role for biotic forcing in macroevolution. We call for a more pluralistic view of how biotic forces may drive long-term evolution that is compatible with both phenotypic stasis in the fossil record and with non-constant extinction rates. Promising avenues of research include contrasting predictions from relevant theories within ecology and macroevolution, as well as embracing both abiotic and biotic proxies while modelling long-term evolutionary data. By fitting models describing hypotheses of biotically driven macroevolution to data, we could dissect their predictions and transcend beyond pattern description, possibly narrowing the divide between our current understanding of micro- and macroevolution.

1. Introduction

‘A fundamental question on how the physical world relates to evolution remains unsolved. Is physical change the necessary pacemaker of speciations and extinctions, or do living entities drive themselves to evolve and disappear even in its absence?’, Vrba [1, p. 24] asked two decades ago. The same question had previously been raised by Maynard Smith [2], and still lacks a satisfactory answer.

Many palaeobiologists argue that biotic interactions drive evolution only on shorter time scales, while changes in the physical environment are necessary to trigger macroevolutionary changes on geological time scales (e.g. [1,3–9]). Barnosky [6] compiled eight ‘Court Jester’ hypotheses from the literature, each claiming that abiotic change is the most important factor driving major evolutionary change. He contrasted this collection of Court Jester hypotheses with a class of ‘Red Queen’ hypotheses that claim that biotic factors are more important in driving macroevolution—a class that in practice boils down to a single hypothesis: Van Valen’s Red Queen [10]. Barnosky’s strong emphasis on abiotic factors driving macroevolution was echoed by Benton [7, p. 728], who claimed that ‘Competition, predation, and other biotic factors shape ecosystems locally and over short time spans, but extrinsic factors such as climate and oceanographic and tectonic events shape larger-scale patterns regionally and globally, and through thousands and millions of years.’ Likewise, many others have also stressed the dominance of abiotic factors. For instance, Eldredge [11, p. 28] stated that ‘the conclusion seems inevitable to me that nothing much in the way of adaptive evolutionary change takes place unless and until physical disturbance impacts ecosystems and that the amount of evolution is proportional to the amount of disturbance and extinction that affects ecological and genealogical systems’. Likewise, Lieberman [9, p. 182] argued that ‘abundant evidence has been gathered to suggest that it may be principally abiotic factors, such as geologic change, climatic change, and a taxon’s presence in a geographically complex region, which causes clades to undergo rampant diversification, rather than organismal adaptation’.
Although the evidence that abiotic change drive macro-evolutionary change is beyond dispute (e.g. [12–14]), many palaeobiologists disagree with the categorical idea that abiotic perturbations are the only drivers and allow for a significant role for biotic factors in driving macroevolutionary change [15–18]. For example, competition-driven clade–clade replacements have been hypothesized for diverse taxonomic groups, including cheilostome bryozoans replacing cyclostome bryozoans [19], angiosperms replacing gymnosperms [20] and bivalves replacing brachiopods [16]. Moreover, macroevolutionary turnover driven by interspecific interactions can be observed as diversity-dependent diversification (see review by Rabosky [21]).

For more than four decades, Van Valen’s Red Queen hypothesis [10] has been central in the discussion on the role of biotic factors in evolution (see review by Liow et al. [22]). In essence, Van Valen proposed that biotic interactions suffice to drive species to evolve indefinitely, independent of abiotic changes. He argued that evolutionary advantages gained by one species cause evolutionary disadvantages to other species in the community, an idea dating back to Darwin [23] and Fisher [24]. Furthermore, the negative effects adaptation in one species have on other interacting species result in a continuous evolutionary race in which species catch up or go extinct, resulting in a constant, age-independent risk of extinction. Thus, the Red Queen hypothesis invokes a macroevolutionary process to explain a macroevolutionary pattern.

Upon publication, the Red Queen hypothesis received immediate criticism from both neontologists and palaeobiologists. One criticized aspect of the hypothesis (e.g. [25,26]) was the underlying assumption that an increase in fitness in one species should be exactly balanced by a loss of fitness in other species (the ‘zero-sum’ assumption), which Van Valen had based on the argument that the amount of energy in a community stays approximately constant. The inference that extinction probability should be constant was also criticized (e.g. [27,28]); the observation of this pattern was what had motivated Van Valen to formulate his hypothesis in the first place. It was also claimed that the Red Queen hypothesis predicts gradual evolution and would be incompatible with widespread observations of phenotypic stasis in the fossil record [26,29,30]. Despite widespread criticism, the Red Queen continued to attract attention, being the only major theory that gave biotic factors the central role in driving macro-evolution. As Hoffman [31, p. 2] succinctly observed, the Red Queen hypothesis ‘has been repeatedly dethroned and restored to power but the issue has not yet been settled’. The situation is much the same today, with the Red Queen still having an influential hold on our imaginations. The review by Liow et al. [22] presented a view of the Red Queen that was close to the one originally presented by Van Valen [10], while Vermeij & Roopnarine [32, p. 561] devoted an entire paper to dethroning the Red Queen, in the process repeating many of the above criticisms and concluding that ‘the evolutionary turf over which the Red Queen reigns is so small that the Red Queen is best eased into retirement’. Regardless of the view one takes, it is clear that the Red Queen hypothesis has been incredibly productive, in the sense that it has stimulated much research and served as a justification for questioning the view that abiotic perturbations are necessary for macroevolutionary change.

As we see it, the central proposition of the Red Queen hypothesis—that evolution will continue without changes in the physical environment—needs and deserves to be taken seriously. Abiotic changes in the form of volcanism, tectonic events, asteroid impacts and climate changes have been present throughout the history of life. Yet any theory of evolution that aims to describe the macroevolutionary consequences of abiotic change needs a notion of what long-term evolution would be like in its absence. The proposition that major evolutionary change will continue in the absence of abiotic perturbations therefore remains fundamental to evolutionary biology [1,2]. A more pluralistic approach that moves beyond the formulation of the original Red Queen hypothesis is needed: just as there are multiple Court Jester hypotheses that detail how abiotic forcing might initiate and drive macroevolution [6], we find it necessary to encourage development of alternative hypotheses for how biotic forces might drive macroevolution. Further testing and discussion of the specific assumptions of Van Valen’s original Red Queen (e.g. constant extinction rates or the zero-sum assumption) will not settle the issue. This point has perhaps been most convincingly demonstrated by the many theoretical models that exhibit intrinsically driven continuous trait evolution or repeated lineage branching and extinction [33–40]. Because these models are built on assumptions that are very different from those of the original Red Queen, their validity is unaffected by the criticisms raised against her. They in fact lend plausibility to the view that long-term evolution may occur independent of changes in the physical environment.

In this paper, we briefly summarize theoretical work that shows that long-term evolution driven by biotic interactions is plausible, and point to empirical evidence from the fossil record that indicates that biotic forcing is important in macroevolution. We also argue that evolution driven by persistent biotic forcing may be consistent with observed phenotypic stasis in the fossil record and with varying extinction probabilities. We end by discussing some hypotheses for how biotic forces shape macroevolutionary patterns, challenges that lie ahead in studying biotic factors in macroevolution and paths that may be challenging but fruitful to explore.

2. Models of long-term evolution driven by biotic interaction

The strictest interpretation of the claim that abiotic forces drive macroevolution is that abiotic change is absolutely necessary to drive evolution on longer time scales and that evolution would cease in its absence. This is sometimes referred to as the ‘stationary model’ [4,6,26], and implies that dinosaurs would be ruling the Earth morphologically unaltered if abiotic conditions ceased changing around 70 Ma.

A number of theoretical models have investigated evolutionary dynamics in the absence of extrinsic change. Some early models were directly motivated by Van Valen’s Red Queen. These models investigated the evolutionary dynamics of evolutionary lag, or ‘lag load’, a measure of the difference in fitness between the fittest possible genotype in a population and the current mean fitness of the same population, keeping all other species constant (e.g. [23,26,41]). Maynard Smith [25] showed that the zero-sum assumption was consistent with a constant rate of evolution, just as Van Valen claimed, but also argued that the zero-sum assumption was untenable. Stenseth & Maynard Smith [26] demonstrated that Red Queen evolution could be produced without the zero-sum assumption and that evolutionary stasis was also
a possible outcome (the stationary model). In retrospect, it is clear that the definition of lag load used in these papers only makes sense because there was no intraspecific frequency dependence of fitness, a limitation these models shared with other coevolutionary models of that time (discussed in [37,41,42]). Other early models featured trait evolution, and showed (not surprisingly) that predators and victims would evolve more and more extreme arms levels if the improvements were always both beneficial and cost-free and the trait space was unbounded (e.g. [41]).

The next generation of Red Queen models focused solely on trait evolution instead of lag loads and incorporated intraspecific frequency-dependence of fitness [33–35,40]. These models assumed that evolution is driven by mutations of small effect, and showed that relatively simple ecological interactions might lead to cyclic evolutionary trait dynamics [33,34,40] or repeated cycles of branching and extinction in lineages [35]. The zero-sum assumption played no role in these models, and they did not require an unbounded trait space. Their major limitation from a Red Queen perspective was their focus on pairwise interacting species only.

The last decade has seen a series of important theoretical papers that explore evolving food webs and their non-equilibrium dynamics [36–39,43]. These papers make no reference to the Red Queen. The models are based on trait evolution and show how complex food webs that contain many species at several trophic levels can evolve and be dynamically maintained through evolutionary branching and extinction, driven by biotic processes such as predation and competition. They are consistent with the observation that diversification is diversity-dependent. For instance, they show that in the absence of extrinsic forcing, biotic interactions can create coevolutionary trait dynamics in which the number of species fluctuates widely in a periodic or chaotic fashion [36]. Biotically driven mass extinctions or extinction cascades are also demonstrated in these models [36,39,43]. Takahashi et al.’s [39] model features food webs that alternate between two community states: they repeatedly build up and evolve towards a quasi-stable state that is very vulnerable to perturbation, where demographic stochasticity triggers an extinction cascade, starting the cycle again. In Allhoff et al.’s [43] model, periods of increasing specialization in the food web are interrupted by extinction cascades; the trigger for the cascades is the arrival of a new mutant form that escapes specialist predation, and therefore can increase to great numbers and displace other species. Together, the food-web models demonstrate that non-equilibrium evolutionary dynamics in food webs can be very rich, and that evolution can last for a very long time in the absence of abiotic forcing, taking the community through many cycles of expansion and contraction. The models are simulation-based and not analytically tractable, and no direct claim is made in any of them that evolution never ceases (indeed this is not their focus). They are nevertheless highly relevant for understanding the processes shaping macroevolution. Other models produce food webs with a very stable structure [44], and there are ongoing efforts to identify the factors (e.g. functional responses and trophic interactions) that stabilize and destabilize evolving food webs (see discussions in [39,43]).

It is interesting to note here that while Van Valen [10] phrased his Red Queen scenario in terms of a struggle for resources, which emphasizes the role of competition, predator–prey interactions seem to be an important driver of the continuous coevolutionary dynamics in the food-web models. Dieckmann et al. [37] presented a different and more tractable model of community coevolution based on competition alone. In this model, a handful of species within a single trophic level coevolve, and exhibit perpetual turnover of species through branching and extinction. The patterns of extinction seem almost regular (underdispersed) in the community coevolution model by Dieckmann et al. [37] when compared with the more complex patterns seen in the food-web models.

In summary, theoretical models based on many different sets of assumptions have shown that biotic interactions have the potential to drive long-term evolution. The food-web models in particular have become increasingly sophisticated and relevant for understanding the evolution of ecosystems, and several exhibit diversity-dependent turnover dynamics. Few (if any) of these theoretical developments are discussed in the palaeobiological literature. It is still the Red Queen models from the 1970s and 1980s that tend to be discussed (e.g. [6,7,32,45,46]), probably because these were motivated directly by Van Valen’s [10] thoughts on the Red Queen, and are more easy to relate to macroevolution than later work. However, the food-web models differ in their detailed dynamics and suggest that biotically driven lineage extinctions do not produce a single ‘signature’ pattern. It is time to exchange the constantly running Red Queen and its constant extinction probability with a more pluralistic view of biotically driven evolution. The theories of ecosystems and food webs, and the palaeobiological literature, already present us with a diversity of hypotheses (table 1).

3. Studies of long-term evolution driven by biotic factors

A more moderate interpretation of the claim that abiotic forces drive macroevolution is that abiotic forces (the ‘Court Jester’) have played such a dominant role throughout the history of life that the effects of biotic forces on macroevolution are hard, or next to impossible, to detect [3,6,7,9]. A variant of this claim is that biotic factors are restricted to affecting evolutionary change on small temporal and spatial scales, whereas abiotic factors are responsible for change on larger scales [6,7].

The hypothesis that abiotic factors have been the dominant macroevolutionary force throughout the history of life may seem most plausible if we restrict our attention to taxonomic turnover. This is especially true for mass extinctions (e.g. [12]). For extinction in general, however, a role for biotic factors cannot be ruled out merely by showing that abiotic factors can explain some of the variation in turnover (e.g. [50]), or vice versa (e.g. [51]). Only a handful of studies of taxonomic turnover simultaneously consider data from both abiotic and biotic proxies when investigating which factor is the stronger driver of macroevolutionary change: Aberhan et al. [52] reported that the changes in sea level, climate and seawater chemistry could not explain Mid-Mesozoic relative abundance of different marine guilds, while increasing levels of predation and biotic ‘bulldozer’ explained the decline in the abundance of vulnerable epifaunal taxa. Likewise, Ezard et al. [53] used morphological traits as proxies for species ecology, and found that while both biotic and abiotic factors contributed to species diversity changes, extinction was more strongly affected by climate while speciation was more strongly
Table 1. Hypotheses of biotically driven macroevolution and their predictions. A number of hypotheses are compatible with the idea that macroevolution is driven by biotic forces, although most discussions tend to focus on Van Valen’s Red Queen. Other plausible hypotheses of biotically driven macroevolution can be formulated based on ecosystem and food-web theories, or the literature on phylogenetic diversification, and these yield predictions that differ strikingly from that of the original Red Queen.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Rate of Biotically Driven Extinction</th>
<th>Biotically Driven Extinction Cascades</th>
<th>Related Species Tend to Go Extinct Together</th>
<th>Periods of True Phenotypic Stasis</th>
<th>Rate of Biotically Driven Speciation</th>
<th>Evolution Ceases in Absence of Abiotic Forcing</th>
<th>Qualitative Changes of Dynamics Due to Abiotic Forcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Valen’s Red Queen [10]: evolutionary advance in one species exactly balanced by deterioration in others</td>
<td>Constant</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Constant</td>
<td>No</td>
<td>None; species are constantly evolving regardless of abiotic forcing</td>
</tr>
<tr>
<td>‘expansion and crash’ hypothesis: slow evolutionary build-up and occasional collapse of inherently fragile ecosystems</td>
<td>Varying, depending on stage of community build-up</td>
<td>Yes</td>
<td>No</td>
<td>No clear</td>
<td>Varying, depending on stage of community build-up</td>
<td>No</td>
<td>Abiotic forcing, just like biotic forcing, may cause fragile ecosystems to collapse, thus setting the stage for gradual evolutionary build-up of a new community</td>
</tr>
<tr>
<td>‘key innovations’ hypothesis: macroevolution dominated by key innovations followed by radiations, and clade replacements</td>
<td>Varying</td>
<td>No</td>
<td>Yes</td>
<td>No clear</td>
<td>Occasional bursts</td>
<td>No</td>
<td>Changes in the abiotic environment may create new niches that can be accessed by key innovations</td>
</tr>
<tr>
<td>The Escalation hypothesis [15,47]: natural enemies drive morphological diversification; ecosystem stability due to evolved interdependencies across taxa</td>
<td>Negligible</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No clear</td>
<td>No clear</td>
<td>Abiotic forcing is the predominant cause of extinction, which opens up new evolutionary opportunities</td>
</tr>
</tbody>
</table>

*The ‘expansion and crash’ hypothesis is loosely based on ideas from ecosystem theory [48] and food-web models (e.g. [39]).

*The ‘key innovations’ hypothesis is loosely based on ideas on phylogenetic diversification (e.g. [16,21,49]).
influenced by diversity dependence in the last 65 Myr. Mayhew et al. [54] reported that both temperature and turnover rates in previous time intervals have effects on marine invertebrate richness in 10 Myr time intervals. Meachen & Samuels [55] reported that trait evolution in Pleistocene coyotes was probably attributable to biotic interactions rather than climatic changes. While we applaud these studies for considering biotic and abiotic proxies simultaneously, additional studies investigating other taxonomic groups and temporal scales are needed before any firm conclusions can be drawn on the relative contributions of biotic and abiotic drivers to taxonomic turnover in the fossil record. There are also many technical challenges in estimating taxonomic turnover and diversity on geological time scales, not least because of incomplete sampling in the fossil record, the lack of application of statistical frameworks for addressing causality, and the auto-correlation of time-series data. It is necessary to address these challenges with appropriate statistical tools to undertake a more robust evaluation of biotic and abiotic factors on evolutionary processes.

Macrovolution in the broader sense also includes the evolution of traits and novel structures (e.g. [56,57]). The origin of eyes, the evolution of wings and the transition from fins to limbs are classic macroevolutionary transitions that have largely been discussed without evoking abiotic factors (the ‘Court Jester’) as probable primary driving forces. Such evolutionary novelties allow new functions, may open new adaptive zones [58], and even trigger adaptive radiations [21,49]. Similarly, the long-term temporal trends in increasing complexity of adaptations as well as ecospace seen in the fossil record are usually linked to biotic factors [15,47,59]. For instance, the escalation of more complex traits due to predation pressure has been thought to have taken place during both the Cambrian Radiation [60] and the Mesozoic Marine Revolution (e.g. [47]). Moreover, intraspecific competition is key to understanding evolutionary trends in sexual dimorphism, like Rensch’s rule [61]. Biotic forces are important in our understanding of trait evolution on long time scales.

4. Biotically driven evolution and the paradox of stasis

It has been argued that the Red Queen is incompatible with observations of stasis in the fossil record (e.g. [26,29,30,32]). Indeed, the original Red Queen metaphor suggests that all species have to continuously evolve so as not to lag behind other species and that they have constant (age-independent) probabilities of extinction. Together, this suggests a uniformity of process that is clearly incompatible with zero evolution (e.g. [4,26]). It would be grossly misleading, however, to generalize from this and conclude that biotically driven evolution is incompatible with observations of stasis on geological time scales.

First, it is necessary to acknowledge that phenotypic stasis is itself a matter of interpretation. Stasis can be operationally defined as a state of minimal net rate of evolutionary change over some period of time, usually modelled as temporal variation around a constant mean phenotype (e.g. [62,63]). If stasis truly represents negligible evolution, the small-scale temporal fluctuations in phenotype observed during periods defined as stasis may simply reflect a combination of preservation artefacts, varying sampling probabilities, and other types of noise and measurement error. Under this interpretation, stasis may be caused by factors like stabilizing selection and evolutionary constraints (reviewed in [64]). Understood in this way, phenotypic stasis is difficult to reconcile with the high levels of additive genetic variance [65,66] and strength of selection (e.g. [67]) typically observed in contemporary populations. This is known as the ‘paradox of stasis’ [68], and it applies with equal force to biotically and abiotically driven evolution.

The alternative interpretation of observed stasis is that the temporal small-scale fluctuations observed during periods of stasis represent adaptive evolutionary change towards a constrained but non-static fitness optimum on the adaptive landscape. The rate of evolution around such an optimum may be fast or slow, depending on the evolvability of the population and the tempo of the movement of that fitness optimum. Recent theoretical and empirical work indeed indicates that bounded but fluctuating changes are the norm (e.g. [69–71]), which indicates that phenotypic change during stasis could be more than error and noise. Selection is inevitable as long as there is a mismatch between the average phenotype in a population and an optimum in the adaptive landscape, regardless of whether it is caused by abiotic and biotic forces. Just as climate changes have led to adaptive and genetically based responses in phenotypes on microevolutionary time scales in a wide range of organisms (e.g. [72–74]; reviewed by Merilä [75]), so have biotic factors, including competition for food resources, interactions between predator and prey, frequency and density dependence, variability of mate choice driven by sexual selection, and sexual conflict (e.g. [49,76–78]). Even migration alone can displace populations off the adaptive hilltop as gene flow counteracts local adaptation (see review by Lenormand [79]).

Second, the broader proposition that biotic factors drive macroevolution does not in any sense rely on the original Red Queen in all its detail. In particular, there is absolutely no reason to assume that biotically driven evolution should be a uniform process. As Vrba [4, p. 441] stated, ‘One might well ask whether the Red Queen could lead to punctuated equilibria. Perhaps the metaphorical lady is not running constantly without getting anywhere, but instead is an occasional sprinter to new places where she rests at length before the next sprint (which would of course not be Lewis Carroll’s Red Queen).’ In other words, perhaps biotic factors trigger minor change most of the time but major change some of the time. Selection pressures may be constant but varying, and species may frequently be subject to various genetic or developmental constraints that must be overcome before large responses can be made. This is likely to hold equally true for evolution driven by biotic and abiotic forces.

The recent models of food-web evolution (e.g. [36–39]) show clearly that extinctions can be diversity dependent and may occur in cascades. The fact that the models feature phylogenetic gradualism rather than punctuated equilibrium is likely to be a consequence of the models being purely phenotypic and allowing only mutations of small effect: evolution slows down only when the fitness landscape flattens (e.g. near a fitness peak or a branching point) or when the population size becomes very small. The incorporation of a more realistic genetic architecture might allow for periods of stasis that are punctuated by fast transitions when the organisms finally break free from developmental and genetic constraints [80]. Superior phenotypes that are not mutationally accessible from current genotypes may become accessible at some later
point via nearby genotypes that are attainable by several selectively neutral steps (i.e. they give rise to the same phenotype as the current one). Punctuations could also represent rare jumps across fitness valleys (i.e. between the basins of attraction of different evolutionary attractors [81]).

5. Macroevolution and ecosystem theory: the need for a synthesis

To develop a theory of biotic forcing in macroevolution, it might be useful to contrast observations and predictions from related and complementary fields. Historically, comparing data and theory from palaeontology and neontology has indeed helped shape our current understanding of tempo and mode in evolution (e.g. [82–84]). Moving beyond the Red Queen, one may argue that the most clearly formulated and well-developed grand theory on how biotic forces drive macroevolution and biological complexity is the escalation hypothesis [15,47]. It claims that natural enemies (i.e. strong competitors and consumers) are the main driving forces behind the evolution of morphological diversification and hence taxa. Moreover, it posits that coevolution tends to create an increasing number of interdependencies (including mutualisms) among species that render ecosystems more robust and stable. Although the escalation hypothesis is largely a theory of biotic evolution, it shares elements with the Court Jester hypotheses: Vermeij [85, p. 222] argued that ‘Although the directions of evolution are determined largely by organisms, the timing of evolutionary events is dictated by extrinsic causes related to climate, sea level, tectonic movements, and mass extinctions.’

Just like Van Valen’s Red Queen, the escalation hypothesis is also founded upon a number of uncertain assumptions—a fact that is rarely discussed. For instance, the escalation hypothesis suggests that communities in the absence of abiotic perturbations will evolve towards robustness and greater stability via evolved interdependences between species [47]. Within the field of ecology, in contrast, the stability of complex ecosystems has long been a matter of controversy and is regarded as a puzzle to be explained. This dates back to May [48], who developed a number of mathematical models to show that complexity in itself does not cause ecosystems to be stable; complexity may in fact be destabilizing. This prediction seems to contrast with the claimed and often observed positive relationship between complexity and stability (e.g. [86–89]). However, May reconciled his results with empirical observations by suggesting that stable abiotic environments may permit the build-up of ecosystem complexity. The resulting complex ecosystems may be stable within only very narrow parameter ranges, making them inherently fragile and prone to disintegration and simplification in response to disturbance [48]. Incidentally, this fragility seems analogous to the complex, vulnerable community state that a food-web model may reach before it goes through an extinction cascade (cf. [39]). Ever since May [48], the hunt has been on in ecology to identify different mechanisms that may help stabilize complex ecosystems, and thus explain parts of the puzzle (e.g. [90–92]). Many of these findings are at odds with assumptions underlying the escalation hypothesis, which proposes that stability arises mainly from the evolution of an increasing number of interdependencies (including mutualisms) that form among species [15,47]. For instance, complex food webs have been predicted to be more stable if they are highly compartmentalized, as the compartments limit the number of interdependencies between species, which helps contain extinction cascades [90,91]. Also, the presence of mutualisms has been predicted to destabilize rather than stabilize ecosystems [48,92].

Independent of what causes complex ecosystems to form, factors that destabilize such systems will be important for setting the pace for species turnover and ecosystem evolution. Abiotic disturbances are commonly argued to be the major causes of extinction by palaeontologists [1,3,4,6,11,32]. This view is shared by Vermeij’s escalation hypothesis, which claims that strong competitors rarely drive their enemies to complete extinction, as poor competitors usually are able to sustain viable populations in physiologically marginal environments. Identifying the crucial factors leading to extinction in nature is a difficult exercise when studying modern populations (e.g. [93]), and more so for the fossil record. From the perspective of ecosystem evolution, however, the processes that cause populations to decline to the point where the stability of ecosystems is at risk are more important than the processes that cause the last individual of the last population of a species to disappear. The death of a species may occur long after that species has ceased playing an important ecological role. Biotic factors are known to have potentially devastating effects on local populations (e.g. [2,94,95]), and this is also reflected in many of the recent food-web models, where local extinctions happen solely due to biotic interactions [36,39,43].

A satisfactory theory of biotically driven macroevolution requires a synthesis of ideas from palaeobiology, ecology and evolution. Inconsistencies must be confronted head-on, and relevant processes not conveniently dismissed as operating on different time scales [3,6,7]. We should use our current understanding of ecosystem stability and community assembly (e.g. [90,92]) to further develop theories on the role of biotic factors in driving macroevolution, including factors causing ecosystem collapse. It is possible to study the relationship between complexity and stability on macroevolutionary time scales [89], and investigations of the strength of ecosystem-stabilizing mechanisms using fossil data will be invaluable in resolving the roles of biotic interactions in community evolution.

6. Challenges ahead

We have emphasized the valid reasons for expecting biotic factors to be important in determining macroevolutionary processes and hold that they should be included more often in our analyses of macroevolutionary data than is currently the case. This is easier said than done. Biotic interactions are notoriously difficult to study in deep time and come with many caveats. However, there is no way around facing this challenge if we want to disentangle the effects of biotic and abiotic drivers of macroevolutionary change. Palaeoenvironmental proxies that purportedly reflect past temperature, acidity and other climatic or environmental variables are often used in studies of macroevolution. An analogous solution for the study of biotic factors requires identification of proxies that meaningfully represent biotic variables. If specific morphological traits indicate functional differences among species, such traits may be used as proxies for the ecology of different taxa, as did Ezard et al. [53]. However, such proxies must always be chosen and estimated with care.
Developing testable hypotheses that make up cohesive theories of biotically driven macroevolution is challenging. To promote discussion, we list in Table 1 different hypotheses that assert a clear role for biotic forces in shaping macroevolution. While the Red Queen and escalation hypotheses can be directly ascribed to Van Valen [10] and Vermeij [15,47], respectively, the ‘expansion and crash’ and ‘key innovations’ hypotheses are inspired by a larger body of literature. Statistical modelling of specific empirical systems will be a major tool in investigating their predictions. Their investigation should also be informed by theoretical modelling. Suppose, for example, that biotic forces indeed tend to build up complex, fragile food webs that sooner or later will collapse by themselves. Minor abiotic perturbations might hijack such biotically driven processes by triggering premature collapses, thus synchronizing biotically driven extinction cascades with abiotic change. On a wider geographical scale, one might observe ‘pulses’ of turnover where climatic perturbations cause fragile ecosystems to collapse simultaneously [4]. In such a scenario, it might be relatively straightforward to show that the observed extinction cascades correlate with abiotic change, but a very strong link between theory and data would be needed to demonstrate that biotic factors drove the communities to the brink of collapse in the first place.

Detection of discrepancies and overlap in how key concepts are understood may pave the way for scientific progress and help us ask relevant questions [96]. For example, ecosystem theory [90] and Vermeij’s [15,47] escalation hypothesis make different predictions regarding the effects of mutualistic interactions in stabilizing ecosystems. These predictions can be tested, given that reliable estimates of the prevalence of mutualisms can be obtained for ecosystems that either survived or collapsed.

As we continue to study the effects of biotic (and abiotic) factors on macroevolutionary change, we must be mindful of the following:

— Proving the underlying assumptions of Van Valen’s Red Queen hypothesis wrong does not rule out biotic factors as important macroevolutionary drivers, nor does it show that evolution will cease in the absence of abiotic change.

— The temporal/spatial correlation of abiotic variables with turnover rates cannot be used as evidence that biotic factors are negligible when the latter are not directly examined in the same analyses.

— If ecosystems commonly evolve towards fragile states, the effect of an abiotic change may just be to trigger a collapse that was already imminent.

— Observation of no net evolution over millions of years (stasis) is not in itself an argument for or against macroevolution being driven by biotic factors.

A more mechanistic understanding of macroevolutionary change may help us reduce the micro- and macroevolutionary divide. However, this demands sufficiently complex, yet testable and clear hypotheses regarding the roles of biotic and abiotic factors in driving long-term evolutionary change, not least because biotic and abiotic forces are likely to act in an enmeshed manner.

We think it is necessary to confront conflicting theories in traditionally distant fields such as theoretical ecology and palaeobiology in order to progress in understanding the relative importance of the major drivers of macroevolution. Robust observations from both empirical and theoretical work need to be embraced, even when they seem to be incompatible, because the resolution of conflicts through the merging of theory and data from distant fields may bring forth a more cohesive theory of macroevolutionary change. Such merging has already successfully contributed to our understanding of how traits evolve on micro- and macroevolutionary time scales [69–71], and represents a way forward for a better understanding of the role of the biota in its own long-term evolution.

Acknowledgements. This manuscript benefited from discussions with colleagues in the macroevolution group at the Centre for Ecological and Evolutionary Synthesis, and from thoughtful and constructive comments by Torbjørn Ergon, Barbara Fischer, Thomas F. Hansen, Paul G. Hannik, Michael Matschiner, Jostein Starrfelt and Thomas O. Svennungsen. Wolfgang Kiesling and one anonymous reviewer provided helpful and insightful comments on the manuscript.

Funding statement. The work was supported by grant numbers 227860 and 235073 from the Research Council of Norway.

Authors’ contributions. K.L.V., Ø.H.H., L.H.L. and N.C.S. designed the study. K.L.V., Ø.H.H. and L.H.L. wrote the draft manuscript, with K.L.V. having the lead on the writing. All authors contributed to discussions and subsequent revisions, and approved the final version of the manuscript.

References


85. Elton CS. 1958 The ecology of invasions by animals and plants. Chicago, IL: University of Chicago Press.