Evolution in action: climate change, biodiversity dynamics and emerging infectious disease

Eric P. Hoberg and Daniel R. Brooks

Climatological variation and ecological perturbation have been pervasive drivers of faunal assembly, structure and diversification for parasites and pathogens through recurrent events of geographical and host colonization at varying spatial and temporal scales of Earth history. Episodic shifts in climate and environmental settings, in conjunction with ecological mechanisms and host switching, are often critical determinants of parasite diversification, a view counter to more than a century of coevolutionary thinking about the nature of complex host–parasite assemblages. Parasites are resource specialists with restricted host ranges, yet shifts onto relatively unrelated hosts are common during phylogenetic diversification of parasite lineages and directly observable in real time. The emerging Stockholm Paradigm resolves this paradox: Ecological Fitting (EF)—phenotypic flexibility and phylogenetic conservatism in traits related to resource use, most notably host preference—provides many opportunities for rapid host switching in changing environments, without the evolution of novel host-utilization capabilities. Host shifts via EF fuel the expansion phase of the Oscillation Hypothesis of host range and speciation and, more generally, the generation of novel combinations of interacting species within the Geographic Mosaic Theory of Coevolution. In synergy, an environmental dynamic of Taxon Pulses establishes an episodic context for host and geographical colonization.

1. Introduction: a nexus of climate and diversity

We exist at the nexus of cascading crises for biodiversity (species loss), accelerating climate warming along with attendant ecological perturbation and emerging infectious diseases (EIDs) (expansion of geographical and host ranges and modified interfaces for many pathogens). The biosphere is changing rapidly through landscape alteration, species invasions and ecological disruption, potentially driving development of new irreversible states largely attributable to anthropogenic factors, increasing connectivity and globalization [1]. Charles Elton, a founder of modern ecology, succinctly recognized the expanding human footprint across the world through his observation: ‘We must make no mistake; we are seeing one of the greatest historical convulsions in the world’s fauna and flora’ ([2], p. 31). Within this matrix of change, an epidemiological crisis emanates from the interactions between climate warming and the abiotic and biotic influences determining geographical distributions for diverse species assemblages and their associated pathogens [3–6]. A growing body of empirical evidence accords with predictions made by most models of climate change, which anticipate major shifts in the structure of ecosystems and the distribution of biodiversity [7–11]. These processes highlight the linkages for evolutionary and ecological mechanisms as a basis for the often broad geographical distributions of pathogens and the more limited and landscape mosaic patterns of emergent diseases in association with a breakdown in ecological isolation on varying spatial and temporal scales.
Accelerating climate warming and environmental perturbation constitute a critical threat to ecosystem integrity and sustainability, the distribution and continuity of biodiversity, socio-economic stability, and changing interfaces and echotones influencing patterns of disease [8,10–15]. The scope, scale and pervasive nature of anthropogenic climate warming anticipate substantial impacts across the biosphere and necessitate an integrative approach to understanding environmental change that incorporates historical and contemporary insights about the factors that have determined the structure and distribution of biodiverse systems.

During the past 10 000–15 000 years, agriculture, domestication and urbanization disseminated EID risk on a global scale as people and their interfaces with the environment were altered over time [4]. During the past 50 years, burgeoning human population, rapid and global transportation networks (connectivity), and now accelerating climate change have acted in synergy [6]. Although the temporal–spatial connections and the overall equivalence of events in evolutionary and ecological time are evident [16,17], the phenomenon of increasing frequency of EID is often posited to be an anomaly restricted to recent history and our species. Thus, emerging disease is often considered to be remarkable, and with EIDs as isolated events, response is usually based on reaction. Such demonstrates the need for a proactive capacity to explore pathogen diversity in anticipation of emergence (e.g. active survey and inventory of the global biota), in conjunction with a fundamental conceptual shift about the factors that determine and influence the distribution of pathogens in the continuum of landscape to regional and global systems [1].

Climatological variation and ecological perturbation have been pervasive drivers of faunal assembly, structure and diversification for parasites and pathogens through often recurrent events of geographical and host colonization at all scales of Earth history [16,17]. Processes for expansion and invasion are equivalent irrespective of observations in evolutionary or ecological time, thus a deeper understanding of historical events and processes provides a window (or analogue) to reveal potential outcomes of perturbation in contemporary systems. Episodic shifts in climate and environmental settings, in conjunction with ecological mechanisms and host switching, are often critical determinants of parasite associations and speciation, a view counter to over a century of coevolutionary thinking about the nature of complex host–parasite assemblages. This conceptual adherence to a paradigm of coevolution and cospeciation (association by descent of host–parasite lineages) has directly influenced our perception of the importance and potential for host colonization (and dispersal) in explaining the structure of otherwise complex faunal assemblages. If cospeciation often has limited explanatory power and host switching is not rare, there are considerable consequences for our understanding of the nature of EID in a regime of accelerating environmental perturbation [1].

2. Cospeciation, host colonization and emerging infectious disease

The expanding interaction of climate and EID is fundamentally an evolutionary and ecological issue, a predictable consequence of species that evolved in isolation being brought into close contact following breakdown in mechanisms for biogeographic and ecological isolation. Our current era differs from the Quaternary and earlier in that human activity accelerates the rate of introductions [18], so outbreaks may occur more frequently and over wider geographical ranges. One reason, however, for a general belief that emerging diseases will be rare is the recognition that emerging diseases are often the result of pathogens switching hosts, and the conventional wisdom in evolutionary biology has been that host switches are difficult to achieve [19]. One of the most studied features of parasitism is pronounced conservatism (often termed specificity) in the range of hosts used [20–25]. Most parasites appear to be resource specialists and the overwhelming majority of parasites use only a tiny fraction of the available host species in the habitat. That parasites are resource specialists with restricted host ranges, and yet shifts into relatively unrelated hosts are common in the phylogenetic diversification of parasite lineages and directly observable in real time, has direct consequences for defining the potential for EID.

Host shifts in large part result in EID. Each host shift must begin with colonization or host range expansion. This combines the capacity to use both the ancestral and newly colonized host. Multiple host exploitation following colonization may be brief or prolonged. Additional hosts are assumed to be inferior alternatives to the original host, to which the parasite is supposedly co-adapted, and special circumstances should be needed to incorporate such a host into the repertoire. And yet, host shifts and host range expansions occur often and can happen rapidly [26]. How parasites can be highly specialized and often shift to novel hosts constitutes the Parasite Paradox [27].

Resolving this paradox requires an understanding of how completion of a shift to a novel host is possible if specialization results only from coevolution that constrains parasites to their current hosts. In this case, a full host shift will require more or less simultaneous correlated evolution across a number of traits. In order to successfully colonize a novel host, a parasite will need to modify traits that enable it to locate the new resource, identify it as a possible host and ensure reproductive continuity in association with the new host. In addition, offspring finding themselves on this novel resource will need to be able to sustain themselves nutritionally, and their metabolic system will have to be able to digest the new resource and overcome its chemical defense (or immune system). Each new host may also come with a different set of external enemies requiring new methods of defense or evasion and a different micro-habitat requiring novel physiological adaptations. The correlated changes occurring simultaneously across these suites of characters ought to be so unlikely as to preclude host shifts. Yet phylogenetic comparative studies of hosts and parasites demonstrate that (i) host range is narrow for most parasite species, and (ii) there is substantial evidence of host switching and in some cases host colonization seems to have been the primary driver of diversification [17,23–30].

Resolving the parasite paradox is based on assuming that host shifts comprise two different phases: (1) host range expansion followed by (2) loss of the ancestral host. For a host shift to be completed, there must be mechanisms for generalization (increased host breadth) and for specialization (decreased host breadth). These mechanisms must be at least partly independent, so specialists maintain the potential to
become generalists and generalists maintain the potential to become specialists. Otherwise, specialization becomes an evolutionary ‘dead-end’.

Episodic shifts in climate and environmental settings, in conjunction with ecological mechanisms and host switching, are often critical determinants of parasite diversification [28,31–33], a view counter to more than a century of coevolutionary thinking about the structure and history of host–parasite assemblages (for comprehensive reviews, see [23,24,34]). A new conceptual insight (termed the Stockholm Paradigm because of the core principles emanating from four academic generations of researchers at Stockholm University; for a review, see [25,27,34] and references therein) resolves the parasitological paradox and long-standing perceptions of cospeciation and the nature of host–parasite diversification.

Explanatory power for the pervasive role of host (and geographical) colonization in faunal assembly and emerging disease is based on an integration of four key concepts: (i) Ecological Fitting (EF) [35]—phenotypic flexibility and phylogenetic conservatism in traits related to resource use, most notably host preference—provides substantial opportunities for rapid host switching prior to the evolution of novel host-utilization capabilities. Host shifts via EF allow host range expansion by specialists, which set the stage for the evolution of generalists that then evolve into new specialists described by (ii) the Oscillation Hypothesis [30,36] and, more generally, the generation of novel combinations of interacting species within (iii) the Geographic Mosaic Theory of Coevolution [21]. Host–parasite assemblages—whether micro- and macroparasites of vertebrates or phytophagous insects—exist and persist in a crucible of accelerating change and demonstrate the equivalence of processes across spatial scales and through evolutionary and ecological time [6,16,17]. This supports the conclusion that (iv) Taxon Pulses [37,38], driven by climate change and large-scale ecological perturbation, are drivers of biotic mixing, which leads to episodes of rapid host switching, including outbreaks of EIDs [1,17,27,31,39].

A traditional paradigm of coevolution/cospeciation predicts that the more intensive the co-adaptive responses by hosts and parasites, the less likely the chances of host switching. In a sense, the cospeciation process itself should provide a high degree of protection against emerging diseases. A considerable body of empirical evidence, however, demonstrates that cospeciation is relatively rare and is only one among a number of processes involved in host–parasite diversification; in any event, the idea of specificity (a microevolutionary phenomenon) is decoupled from macroevolutionary mechanisms in coevolution [28,31–33,40].

Significantly, the model case for cospeciation, based on explorations of ectoparasitic lice (Pthiraptera) and pocket gophers (Rodentia: Geomyidae) can best be described under the dynamic of EF, oscillation and the taxon pulse (for detailed discussion see [41]). Contrary to serving as an exemplar for cospeciation, processes linking evolution, ecology and biogeography accommodate insights about the age and history for the assemblage of contemporary pocket gophers and the nature of diversification in the gopherlouse model. We can explore some pertinent details of this system—for example, Geomyidae is an endemic Nearctic family comprising an assemblage of highly sedentary rodents exhibiting considerable stability in geographical range with numerous species and subspecies that are strongly partitioned at local scales [42]. Contemporary diversity among pocket gophers is limited to the temperate zone, and the group appears to have been restricted south of the Laurentide– Cordillera continental glaciers in North America during sequential glacial maxima over the extent of the Late Pliocene and Quaternary, with a primary radiation being limited to a relatively brief temporal window between 4.2 and 1.8 Ma [43]. A burst of diversification for genera and species of pocket gophers, and presumably their louse parasites, coincided with a substantial regime of episodic variation in climate and habitat perturbation. Cyclical shifts in climate have been identified as primary drivers for expansion/contraction, isolation (often in restricted refugia) and secondary contact as a principal influence on patterns of geographical and host colonization among assemblages of mammals (other vertebrates) and parasites since the Pliocene [17,32,44]. Re-analysis of the pocket gophers and lice using a method [45] that does not assume maximum co-speciation produced an evolutionary story involving alternating periods of host colonization and cospeciation. This result is consistent with the general climatological and regional chronology in which episodes of dynamic climate change have strongly determined patterns of diversification and distribution [16], that is, within the expectations of the Stockholm Paradigm.

In this regard, a prevailing maximum cospeciation model has provided an oversimplified view of macroevolutionary processes leading to ecosystem assembly and structure of the biosphere in evolutionary and ecological time. Empirical studies indicate that few parasite groups conform to the phylogenetic patterns of host–parasite associations expected if opportunities for EF were relatively rare. Clades such as ectoparasitic arthropods that have been proposed as the exemplars of limited host switching [46,47], although interesting to evolutionary biologists and ecologists, cannot form the general conceptual framework for dealing with EID because they are rare. Indeed, the primary classical model has now been shown to involve extensive patterns of host colonization [41]. More broadly, the majority of cases indicate substantial host switching throughout history, and extensive diversification through cospeciation appears to be circumscribed [24,28,48].

Limited explanatory power for a paradigm of cospeciation indicates that it is a conceptually inappropriate model for understanding the nuances of pathogen distribution and emerging disease. Further, reliance on assumptions about cospeciation leads to two logical conclusions: first, host switches should be rare; and second, when host switches occur, there must be some underlying genetic change that increases the ability to be associated with a new host [49]. The focus of discussions about climate change and emerging diseases then centres on possible mechanisms by which climate change can lead to such novel genetic changes. In cases involving shifts among hosts or regions, we often uncritically assume that demonstrated genetic change or new variation in a pathogen is an adaptation to a new array of hosts, when it may well represent only the genetic consequences of introduction and invasion on population structure. The expectation remains, however, that because novel genetic innovations must lead the way, emerging diseases will be rare, even during periods of climate change; they may increase relative to background, but they will not be common.

A paradigm encompassing the pervasive nature of colonization predicts that emerging diseases—in the form of
parasites of humans, livestock, crops (we include in this novel pest phytophagous insects and parasitoids insects of beneficial insects) and wildlife—will be common rather than rare events during episodes of climate change. This is because host switching is initially driven by EF, and that is based on genetic capabilities already in the system. The paradigm assumes that there is a large space (Sloppy Fitness Space: [26,27,50,51]) of potential hosts from which most pathogens are precluded by circumstances of time, space and origin. Climate change and the associated biotic expansion events make much more of that space available, in which case switches are expected to occur rapidly and often. Concurrently, biotic contractions could also be predicted to increase the rate with which host switching may occur, as ranges and distributions of species become restricted or are compressed into smaller biogeographic areas.

3. Invasion pathways, expansion and colonization in ecological time

A context for geographical colonization and EF is also evident in contemporary systems and may represent a general phenomenon and potential model, as demonstrated in rapidly expanding ranges for helminth parasites at high latitudes in the central Canadian Arctic. Interacting factors of climate warming (increasingly permissive environments), host migration and dispersal appear responsible for geographical expansion and establishment of two lungworm nematodes (Protostrongylidae) on Victoria Island [52]. These parasites, with gastropod intermediate and ungyulate definitive hosts (Umingmakstrongylus in muskoxen; Varestrongylus in caribou and muskoxen) were previously restricted to mainland habitats and were unknown in the low Arctic islands until 2008 and 2010 [52]. Appearance of these parasites coincided with accelerated warming across the region, and relatively cool climatic conditions before the 2000 may have restricted establishment of lungworms on the island through limitations on development and population dynamics for larval parasites and ectothermic gastropod intermediate hosts. Recent climate warming has likely relaxed constraints for establishment and could further be driving the potential for host switching from caribou to muskoxen (for Varestrongylus) in areas of sympathy [52]. Thus, parasites, acquired by migratory caribou during the winter on the mainland have likely been introduced repeatedly through annual migrations over possibly decades but often without successful establishment. By contrast, muskoxen are not strongly migratory and only stochastic events of dispersal lead to movements between mainland and island habitats. Further, Umingmakstrongylus is host-specific in muskoxen and had previously been reported only from the central Canadian Arctic mainland [53], where temperature constraints limited distribution [54]. Converging conditions of warming climate and a high prevalence and intensity of infection in adult muskoxen appear to have facilitated initial introduction coinciding with dispersal for hosts from the mainland. Additionally, prior to expansion, populations of U. pallikuukensis appear to have crossed a tipping point in transmission from a biannual to annual pattern, coincidental with a reduction in generation time and amplification of parasite populations and infection pressure across its core range on the mainland [52,54]. Thus, in the Canadian Arctic, the contemporary long range invasion of two protostrongylids has occurred under contrasting mechanisms of recurrent migration (Varestrongylus in caribou) versus sporadic dispersal (U. pallikuukensis in muskoxen). Climate warming, in both cases, is a central driver in expanding distributions and successful establishment of the parasites on the island. Direct insights into the dynamic processes linking climate, parasite developmental biology and host population ecology with the invasion and establishment of macroparasites are apparent.

Distributions and central (core) ranges for host–parasite assemblages are determined by interactions defined by history, biotic structure (interconnectivity within ecosystems and particular life history/cycle parameters for multi-host parasites), developmental rates, thermal tolerances, resilience, degree of adaptations, host and pathogen longevity and vagility for their component parts [32,55]. Rapidly changing environmental regimes, particularly temperature and strongly related factors such as humidity, will then be predicted to have substantial influence on the continuity and overall future outlook (contracting, stable and constrained, or expanding) at landscape to regional spatial scales [52,56,57]. If this is correct, then the applicability of concepts linked to hard tipping points, and Shifting Balancing Points, become generalities for predicting the potential range of responses in complex host–parasite (or pathogen) systems under accelerated warming and environmental perturbation.

Shifts between permissive and non-permissive environments have played out in high-latitude systems (and in the context of altitude) related to increasing variation and perturbation contained with longer-term incremental warming. Consequently, there is interplay between long-term and short-term (ephemeral or extreme) processes. Within short-term events, a shifting balance occurs between points that are either permissive (‘allowing’ establishment on an ephemeral basis) or those that dampen the potential for successful geographical colonization. Shifts in distribution and abundance would be predicted and linked to this range in variation for environmental settings occurring on the periphery of ranges (consider expansion dynamics, and the properties involved with population thresholds, their density and distribution). A permanent change in distribution, however, would not be achieved outside of a new steady state associated with a regime shift or tipping point.

There is a contrast between what is happening within the core geographical range and the variables responsible, and what is occurring on the peripheries of an expanding range. These processes can be extrapolated from landscape to regional scales and further provide an analogue that links processes in ecological and evolutionary time; consider the comparison of geographical colonization and mosaic faunal dynamics at the Beringian nexus under recurrent glacial cycles and the outcomes for the ongoing expansion from the mainland to Victoria Island [32]. Hard tipping points are defining boundaries within core ranges where an assemblage has historically been established (e.g. Umingmakstrongylus). Shifting balancing points occur on the periphery of core range and constitute the potential for expansion in the context of climate variation (wobbling climate) over short to long timeframes (both Umingmakstrongylus and Varestrongylus). Expansion and persistence then are dynamic and play out as ephemeral processes (of episodic establishment and extirpation) in a regime of balances in a shifting climate that are either conducive (permissive) to introduction or are neutral, or are negative. Episodes of
recurrent expansion into negative habitat, for example associated with migration, will not support establishment or introduction [58]; recurrent expansion and introduction may be negated in the short term by secondary development of non-permissive environments. Although infection pressure in the sense of geographical expansion may occur with migratory host populations (such as caribou or birds at high northern latitudes), only a permissive environmental setting will result in introduction. A tipping point related to reductions in generation time, population amplification and increasing infection pressure in core range may be the antecedent for expansion under a positive shifting balance in peripheral habitats and environments.

In the Northern Hemisphere, northward expansion is predicted as a generality for an array of temperate host–parasite systems, especially under the driver for expanding ranges occupied by diverse assemblages of vertebrate species in terrestrial (and aquatic) systems [32,55,59]. Consequences of expansion relate to a continuum in which EF has a prominent role: (i) host expands, parasite dies; (ii) host and parasite expand and persist (exposing naive host populations and species to parasites through switching (EF)); (iii) host expands, introduces parasite into new assemblage (EF) and original host does not persist; (iv) host expands (leaves parasites), acquires new arrays of parasites from endemic assemblage (EF). Episodes of expansion lead to development of mosaic assemblages over space and time as a consequence of processes driven substantially through geographical colonization, EF and host switching [6,32].

As an example, historical processes across the Beringian nexus at the cross-roads of the Palearctic and Nearctic, and such routes linking North America—South America, and Africa—Eurasia, involved the interaction of tipping points (within and from core range) and shifting balancing points (into peripheral range) given the cyclical nature of climate and environmental change over the past 3–5 Myr of Earth history. These may constitute important mechanisms associated with parasite/pathogen persistence, amplification, geographical expansion and acquisition of novel host assemblages in the context of range shifts for contemporary systems in ecological time that may increasingly characterize distributions for complex parasite assemblages and disease complexes [60].

4. Synthesis and conclusion

Host–parasite systems are ubiquitous. Understanding the factors that generate, maintain and constrain these associations has implications for broad ecological and environmental issues including the dynamics of EIDs [29,39,61], biological control, biological introductions and invasions, and biotic responses to climate change [25].

The Stockholm Paradigm postulates that parasite specialists can shift rapidly to novel (naive) hosts via EF. EF between hosts and parasites occurs with high enough frequency to influence host range dynamics and the diversity of species and interactions among species. Although no quantitative statement of this importance can yet be made, it is clear from the above discussion that shifts onto relatively unrelated hosts appear routinely in phylogenetic analyses and are observed readily in contemporary time. These observations are fundamental for EID studies: EIDs arise when parasite species begin infecting and causing disease in host species with which they have no previous history of association. If the nature of host specificity is such that the potential for EF is small, host shifts are likely to be rare and attention can be focused on managing each EID as it emerges. Little attention need be paid to its origins, beyond a search for the taxonomic identity of the parasite acting as the pathogen, and its immediate reservoir. If the potential for EF is large, however, then host shifts are likely to be common, and a more predictive, pre-emptive framework for managing EID will be needed, greatly increasing the challenge of an already difficult problem.

Humanity has tended to react to emerging diseases as they occur, using our understanding of epidemiology in an attempt to mitigate the damage done. If the Stockholm Paradigm reflects a fundamentally correct explanation of the evolution of interspecific associations, then reactive management policies for dealing with emerging diseases cannot be economically sustainable. This implies that an additional strategy that could be employed in conjunction with those reactive tactics is being proactive. We can use our knowledge of what has happened in the past to help us anticipate the future. It is a kind of evolutionary risk assessment. Just as we cannot stop climate change, we cannot stop these emerging diseases. We believe, however, that proactive risk management [36,62] is less expensive and thus more effective, than responding after the crisis.

A broader macroevolutionary picture for general processes of expansion and invasion is emerging, which links historical and contemporary systems. Historical conservatism is pervasive, and it is evident that equivalent mechanisms have structured faunal assembly in the biosphere and that episodes of expansion and isolation have alternated over time. Fine-scale (landscape) processes as a mosaic within larger events, while important, are idiosyncratic and more strongly influenced by chance and founder events. Thus, in contemporary associations, under the influence of accelerating change, we cannot always predict which components of the biota will come into proximity or contact, the duration of these events or the temporal order in which faunal mixing occurs. In these instances, the importance of adaptation may be diminished, whereas the persistence of parasites and pathogens through broad sloppy fitness space can be seen as the capacity to use rapidly changing resources without narrow restriction to any particular ecological/environmental setting.

Climate and disturbance-driven taxon pulses coupled with oscillations in host range can be expected to influence the frequency of EID, because they create episodes of geographical range shifts and expansions. The episodes, in turn, increase biotic mixing and the opportunities for EF to occur. The current EID crisis is ‘new’ only in the sense that this is the first such event that scientists have witnessed directly. Previous episodes through earth history of global climate change and ecological perturbation, broadly defined, have been associated with environmental disruptions that led to EID [16,17,62]. From an epidemiological standpoint, episodes of global climate change should be expected to be associated with the origins of new host–parasite associations and bursts of EID. The combination of taxon pulses and EF suggests that host and parasite species with the greatest ability to disperse should be the primary sources of EID [58,62–64]. Palaeontological studies suggest that species with large geographical ranges and with high ability to disperse are most successful at surviving large-scale environmental perturbation and mass extinctions [65]. Thus, the species most successful at surviving global climate changes will be the primary sources of EID, so host extinction will
not limit the risk of EID. The planet is thus an evolutionary and ecological minefield of EID through which millions of people, their crops and their livestock wander daily.

Acknowledgements. D.R.B. and E.P.H. express thanks for critical discussion and conceptual development to C. Wiklund, S. Nylin, N. Janz (Stockholm University), S. J. Agosta (Virginia Commonwealth University) and W.A. Boeger (Universidade Federal do Paraná). E.P.H. thanks S. J. Kutz (University of Calgary) for 20 years of collaborations and insights exploring biodiversity and the effects of climate among lungworm faunas in Arctic ungulates.

Funding statement. Some aspects discussed herein represent contributions of the Beringian Coevolution Project and the Integrated Inventory of Biomes of the Arctic supported by the National Science Foundation (DEB-Biodiversity Discovery and Analysis-1258010–1256832–1256493) with funding to J. A. Cook (Museum of Southwestern Biology), E.P.H., K. E. Galbraith (Northern Michigan University), E. DeChaine (Western Washington University).

References
