Towards an Ecosystem Model of Infectious Disease

James M. Hassell[1, 2*], Tim Newbold[3], Andrew P. Dobson[4], Yvonne Linton[5, 6, 7], Lydia H.V. Franklinos[3], Dawn Zimmerman[1, 2], Katrina M. Pagenkopp Lohan[8]

[1] Global Health Program, Smithsonian Conservation Biology Institute, Washington DC, USA

[2] Department of Epidemiology of Microbial Disease, Yale School of Public Health, New Haven, CT, USA

[3] Centre for Biodiversity & Environment Research (CBER), Department of Genetics, Evolution and Environment, University College London, London, UK

[4] Department of Ecology & Evolutionary Biology, Princeton University, Princeton, NJ, USA

[5] Walter Reed Biosystematics Unit (WRBU), Smithsonian Institution Museum Support Center, Suitland, MD, USA

[6] Department of Entomology, Smithsonian National Museum of Natural History, Washington, DC, USA

[7] Walter Reed Army Institute of Research (WRAIR), 503 Robert Grant Ave, Silver Spring MD 20910, USA.

[8] Marine Disease Ecology Laboratory, Smithsonian Environmental Research Center, Edgewater, MD, USA

* Corresponding author

Abstract

Increasingly intimate associations between human society and the natural environment are driving the emergence of novel pathogens, with devastating consequences for humans and animals alike. Prior to emergence, these pathogens exist within complex ecological systems that are characterized by trophic interactions between parasites, their hosts, and the environment. Predicting how disturbance to these ecological systems places people and animals at risk from emerging pathogens—and the best ways to manage this—remains a significant challenge. Predictive systems ecology models are powerful tools for the reconstruction of ecosystem function but have yet to be considered for modeling infectious disease. Part of this stems from a mistaken tendency to forget about the role that pathogens play in structuring the abundance and interactions of the free-living species favored by systems ecologists. Here, we explore how developing and applying these more complete systems ecology models at a landscape scale would greatly enhance our understanding of the reciprocal interactions between parasites, pathogens and the environment, placing zoonoses in an ecological context, while identifying key variables and simplifying assumptions that underly pathogen host switching and animal-to-human spillover risk. As well as transforming our understanding of disease ecology, this would also allow us to better direct resources in preparation for future pandemics.

Introduction

Emerging infectious diseases (EIDs) are increasing in frequency as global environmental and anthropogenic changes accelerate[1–3]. For animal-to-human (zoonotic) spillover and subsequent pathogen amplification to occur, a complex set of epidemiological, ecological and behavioral conditions that influence the composition, infection dynamics, contact rates and likelihood of

47 infection within and between host populations must align[4]. Mitigation of future pandemics will
48 rely on our ability to understand how these mechanisms converge to result in exposure of people
49 to novel pathogens, and identify areas at higher risk of pathogen spillover, so that limited
50 resources for animal and human surveillance and risk mitigation efforts can be proactively
51 directed to these sites[5].

52 Accurate forecasting of spillover risk requires a clear understanding of the pathogen
53 dynamics at play in differing global biomes. Interactions between parasites (throughout this
54 article we use the term parasite to describe all pathogenic (disease causing) and non-pathogenic
55 organisms that colonize and can be transmitted between hosts), their hosts, vectors and the
56 environment over defined geographic and temporal scales can be thought of as "episystems"[6,7]
57 (Figure 1). Pathogen communities are focal points of episystems, where competition and co-
58 existence between pathogens and commensal organisms for resources within hosts regulates
59 virulence and transmission, while exerting effects on host fitness and behavior that percolate
60 across trophic scales. The composition and function of these parasite communities are also
61 defined by the top-down impacts of environmental conditions on the fitness, distribution and
62 interactions between host populations. By linking host population dynamics to the composition
63 and turnover of parasite communities inhabiting these host 'patches', metacommunity theory can
64 be used to place zoonotic pathogens and their emergence into new host populations in an
65 ecological context (an approach we refer to as 'pathogen community ecology')[8,9]. While
66 empirical investigations can reveal important associations between host and parasite
67 communities (e.g. [10–13]), modeling of the fundamental processes underpinning these relationships
68 provides the only replicable opportunity to understand how natural and human-driven changes to
69 these systems modify the risks that pathogens pose to humans, and to forecast change in these

risks. The scale of this computationally intensive task—compounded by limited data, complex and often nonlinear relationships, and high levels of uncertainty—has so far eluded conventional epidemiological approaches. We propose that rescaling and novel structural reorganization of models for these systems now make this goal attainable.

Our understanding of infectious disease transmission has come a long way in the past 30 years[14,15]; modern epidemiological models facilitate more accurate predictions about pathogen transmission and disease risk than ever before. Being rooted within foundational concepts of single-agent, single-host systems (such as the basic reproductive number R0), most existing epidemiological models—including more recent frameworks such as stochastic metacommunity models and multi-pathogen SIR models—require significant modifications if they are required to explore the interactions and feedback loops that exist between multiple pathogens, hosts and their shared environment[8,16,17]. Statistical and machine learning methods that have been adapted from ecology (e.g. species distribution models, hierarchical spatio-temporal models, joint species distribution models) have made significant contributions to public health by mapping infectious disease risk and are capable of identifying relationships between zoonotic pathogens, parasite communities, macro fauna and ecosystem structure and function[18–20]. However, using these top-down approaches to extrapolate beyond existing conditions can be problematic, as they lack a mechanistic framework with which to test the impact of management changes and interventions on infectious diseases[21–23].

Whole systems approaches, akin to those used to forecast the world's weather, study biological regulation within the human body, and manage the World's fisheries, are increasingly applied in ecology to understand how anthropogenic forces (such as climate change) change the behavior of ecological systems. Predictive systems ecology[24] promotes the use of mechanistic,

process-based models, parameterized by observational and experimental data, to understand and predict the future state of ecological systems. Outputs are 'emergent properties' of these models – quantitative measures for how different components of the ecosystem change over time. Models of terrestrial and ocean ecosystems (e.g. dynamic global vegetation models, ocean ecosystem models, general ecosystem models)[25] have been used to generate estimates of primary production from forests, community structure of phytoplankton, and have recently been extended to model the World's ecosystems[26]. Unfortunately, none of these approaches consider hosts and their parasites, which exert a ubiquitous influence on all free-living species. We believe that now is the time to extend this approach into the fields of epidemiology and disease ecology[27].

Applying systems-level thinking to forecast disease emergence will necessitate a fundamental change in how we conceptualize infectious diseases. In much the same way that a mechanic working to improve the future performance of a race car requires complete knowledge of how its engineered components are assembled and interact during operation, practitioners looking to predict and affect the future state of episystems require models that capture the suite of biological and social mechanisms underpinning the behavior of host and pathogen communities. *Process-based* models, in which the fundamental ecological and epidemiological mechanisms determining disease risk are described in a mathematical framework, are ideally suited to this task. Recent efforts to simulate and predict the locations of historic and future Ebola virus and Lassa fever outbreaks in West Africa (from environmental, host and epidemiological data using 'environmental-mechanistic models') demonstrate the potential of systems models in forecasting emerging disease risk, but to date these are relatively limited in scope, focusing on single pathogens and omitting aspects of within-host pathogen dynamics[28,29].

We show the relevance of predictive systems ecology models to epidemiology by explaining how they could be developed and applied to forecast and ultimately improve our understanding of pathogen community ecology and how this translates to emerging disease risk. From these models—which we term 'General Episystem Models' (GEpMs)—the dynamics of functionally similar pathogens would emerge from the cumulative responses of parasites, their hosts and vectors to environmental inputs, rooted in ecological and evolutionary theory. To ground these efforts in real-world episystems, we propose model refinement and validation as part of a global experimental network representing replicates across a common set of anthropogenic environmental drivers for disease emergence (e.g., habitat fragmentation, agricultural intensification, pollution, urbanization) in terrestrial and marine environments. Experimental and observational data could be used to develop and validate standardized approximations for describing broad-scale levels of host and parasite organization (genetic, individual, population, community) and their interactions under different environmental conditions across spatial and biological scales.

System structure

Host, Pathogen and Vector Population Dynamics. Where possible, and in common with general ecosystem models, fundamental concepts and processes derived from ecological and epidemiological theory (many of which already exist and are backed up by data) should be used as general baselines with which to model host, parasite and vector population dynamics[24]. The complexity of microbial ecology and evolution, its relative infancy as a field of study, and our lack of knowledge on parasite diversity[30], mean that uncertainty will pose a major challenge in incorporating pathogen community ecology into predictive systems ecology models. While

GEpMs should be no more complex than is necessary to realistically represent episystems, sufficient information on the biological organization of parasites, their hosts and vectors, and the interactions and feedback between this triad and their abiotic and biotic environments, is required for emergent behaviors of pathogen communities and the risk that they pose to humans to be considered reliable. Applying simplifying assumptions as a means of reducing complexity in these models will therefore be central to achieving a balance between predictive accuracy, and methodological and computational feasibility (Figure 2).

A simple but effective form of dimension reduction commonly used in community ecology, and favored for predictive systems ecology models, involves grouping organisms that share life history traits. These similarities dictate that they interact with one another and their environment in a similar manner, so that they are considered identically for modelling purposes. For example, by grouping organisms into functional groups, the Madingley Model has been able to capture global patterns in broad ecosystem structure with a reasonable degree of accuracy[26]. Similarly, trait-based grouping of parasites has been identified as an approach that would contextually simplify modelling of complex within- and between-host pathogen dynamics, and being more directly relevant to ecosystem function, provide greater deterministic and predictive power than taxonomic groupings[9,31,32]. Representing parasites, hosts and vectors as cohorts that share common resource mechanisms and functional traits (e.g., immune evasion strategies for pathogens, and reproductive and feeding preferences for pathogens, commensal organisms, hosts and vectors), could therefore provide much-needed simplification to overcome data paucity and the logistical challenges of trying to model all individuals in large and complex episystems (Box 1, Table 1)[26,33]. By simplifying and compartmentalizing GEpMs in this way, these models would not be able to make predictions about the behavior or emergence of specific pathogens. Rather,

161 they would possess the predictive power to model how the relative abundance of functionally
162 related groups of pathogens (e.g., reverse-transcribing RNA viruses, extracellular drug-resistant
163 bacteria, intracellular apicomplexans) changes across space and time, while reproducing the
164 cross-scale biological processes that are responsible for this variation (Table 1).

165 Since ecosystem structure and stability is predominantly governed by consumer-resource
166 interactions between species – extending, for example, from cellular invasion of viruses within
167 bats, to the impact of bats on arthropod herbivory of the tropical rainforests that they inhabit[34] –
168 identifying generalizations for these interactions ("food webs") will greatly simplify mechanistic
169 models of the ecological processes that link cohorts of parasites, their hosts, vectors and the
170 environment. Lafferty et al.[35] demonstrated how classical models of food web structure
171 (including predator-prey, pathogen, autotroph, decomposer and scavenger models) could be used
172 to generate a general consumer-resource model, capturing all forms of species interaction and
173 revealing new insights into the commonalities of different consumer-resource interactions.
174 Recent studies suggest that complex microbial community dynamics can also be predicted by a
175 relatively simple set of rules expressed as species functional traits and metabolic properties of the
176 environment (such as nutrient availability)[36,37].

177 Because interactions between parasites, hosts, vectors and the environment occur across
178 and between a multitude of microscopic and macroscopic scales, course-grained statistical laws
179 such as allometric scaling rules will also be crucial to identify commonalities that can be used to
180 resolve the underlying interactions between parasite, host and vector communities at a
181 computationally feasible resolution[38,39]. Body mass scaling laws are widely used in ecology, and
182 represent simple predictors of metabolism, abundance, growth and mortality across taxa[39].
183 Recent work has explored these four scaling laws across all eukaryotes, and found that a scaling

184 regime based on the ontogenic and reproductive growth of individuals holds consistently across
185 all species, and could therefore be considered a general basis for the assembly of biological
186 communities[39]. Unsurprisingly, scaling rules also apply to microorganisms – a 'dominance'
187 scaling law (representing the number of individuals belonging to the most abundant species in a
188 defined space) predicts microbial diversity from individual plants and animals to the entire
189 ocean's sediment[40], and log-log scaling rules link gut microbial diversity and animal mass across
190 mammals and birds[41]. With next-generation deep sequencing data being generated at an
191 exponential rate, further unifying principles for biological scaling across eukaryotes and
192 prokaryotes are likely to emerge. Recent work shows that by incorporating allometric scaling of
193 hosts (and other correlative biological relationships) into mechanistic disease transmission, the
194 influence of changes in host communities (such as biodiversity) on pathogen dynamics can be
195 predicted – causal relationships that are difficult to measure directly[42,43]. Collaboration between
196 landscape ecologists, mathematical epidemiologists, immunologists, parasitologists, and disease
197 ecologists who are advancing our understanding of pathogen community ecology, will be
198 required to extend scaling rules to consumer resource models that describe host-pathogen
199 dynamics in multi-agent, multi-host systems across local and regional scales[43–45].
200
201 **Evolution.** GEpMs should also incorporate evolutionary change into parasite and vector
202 population dynamics, as rapid generation times that vary widely between microorganisms
203 (bacteria, viruses fungi), macroparasites and vectors are likely to outpace the duration of model
204 projections. In the simplest terms, parasites could be grouped by evolutionary traits that take into
205 account rates of recombination – for example as clonal or non-clonal organisms[46] (Box 1, Table
206 1). At a finer resolution, Gorter et al.[47] propose a general framework to predict the effects of

evolutionary changes on microbial communities, and develop a cellular automaton model for the positive or negative fitness effects of mutations on the composition of a simple, spatially structured microbial community. Others have developed simulation models for the effects of individual-level microbe fitness and host selection on microbiome diversity and the composition of beneficial, commensal, and pathogenic microorganisms[48,49]. How mutualistic or antagonistic interspecific interactions that are conferred by mutation scale to more complex microbial communities is an area of great uncertainty, but there is evidence to suggest that the general form of such interactions at the community level is responsible for shaping microbial assemblages[50–52]. Carefully controlled experimental studies that improve our understanding of how specific traits (gained through mutation or recombination and that are thought to drive the interaction between species) impact fitness, are required to refine these models so that their predictive power can be tested against real-world parasite and vector communities[47] (Figure 2).

Stochastic evolutionary processes (i.e., random genetic variation of pathogens such as genetic drift) will be particularly difficult to model mechanistically and might be best approached using correlative models that generate simple statistical relationships (such as power laws[53]) between patterns of genetic variation within parasite assemblages, community structure and the environment. Recent studies that have successfully predicted evolutionary processes in microbial communities using knowledge of community architecture and environmental conditions provide evidence that microbial community structure can be forecast without requiring a detailed mechanistic understanding of evolutionary processes[54,55]. The increasingly large data sets provided by next-generation, high-throughput sequencing provide a rich resource that can be mined for biologically significant relationships that link pathogen genetics and ecology using machine learning approaches[56]. Parameters derived from correlative models can

then be used to simplify, and parameterize, semi-mechanistic models for parasite evolution and fitness described above[54] (Figure 2).

Parameterizing GEpMs with data

Once a prototype GEpM has been defined from existing knowledge, a large amount of data would be required to refine and validate the system's structure. Because of the extensive scales at which episystems operate, data gathering efforts – both experimental and observational – would need to be undertaken as part of an ambitious cooperative approach that takes place across spatial and temporal scales relevant to the processes being modeled (Figure 2). For such an effort to be practical and cost-effective, experimental design would need to be an iterative process, in which the model is used to highlight data gaps and develop hypotheses, which in turn inform study design and generate results which are utilized to further simplify and constrain the GEpM (Figure 2)[57,58]. By closely mimicking specific microbiological processes of interest, single-site experimental trials conducted in animal models provide a practical and targeted way of studying the fundamental dynamics (e.g., competition, mutualism, evolution) of parasite communities within the host environment, and identifying feedback loops between parasite communities and their hosts (e.g., via the immune system). Under carefully controlled field conditions, animal models would also be appropriate for studying the mechanisms by which specific abiotic drivers impacting hosts (such as nutritional and psychological stress) and host population dynamics influence the accumulation and turnover of parasite communities.

For GEpMs to be parameterized with simplifying assumptions that can account for how environmental inputs (such as land-use and climate) structure parasite, host and vector populations, observational and experimental field data will need to be collected under 'real-

253 world' conditions. In the first instance, incorporating parasite communities into well-established,
254 long-term studies of intact ecosystems would be an excellent way to test how baseline parasite
255 community dynamics scale across relatively stable ecosystems. For example, sites such as
256 Yellowstone National Park where long-term studies have been conducted on elk, bison, wolves
257 and bears and their interactions within the park provide opportunities to compare the parasitic
258 fauna of predators and prey, seasonal variation in these, and also their interactions with well-
259 studied pathogens such as *Brucella* spp. in bison and elk and scabies and canine distemper in
260 wolves[59,60]. The diets of grizzly and black bears have been well characterized, as they have for
261 most species in the park, so temporal studies could be applied to examine how life history traits
262 like annual hibernation impact mammalian microbiomes[61,62]. Studies in Yellowstone could be
263 expanded to include data from the Yellowstone to Yukon Conservation Initiative (Y2Y) that has
264 set up experimental sites along a vast longitudinal gradient[63]. This would allow examination of
265 how parasite communities change along a climate gradient that spans multiple ecosystems.
266 The effects of anthropogenic environmental change, which manifests on pathogen
267 community ecology at both fine and broad spatial scales, would need to be studied
268 experimentally and by observation under differing levels of anthropogenic stress. Consider a
269 pastoral grassland system for example. Here, controlled experimental trials in grasslands can
270 provide insight into how local-scale forces (such as agricultural practices) shape host and
271 parasite populations and their interactions with the environment within and between plots[64,65].
272 Upscaling to landscapes, where the effects of environmental filtering and dispersal on host and
273 vector populations are greatest, observational studies conducted using remote monitoring devices
274 along gradients of human activity (such as the 'Biome Health Project'
275 https://www.biomehealthproject.com/) can be used to estimate how anthropogenic environmental

change impacts the spatial distribution of host and vector populations (e.g., ungulate wildlife, livestock, mosquitos, ticks)[66]. When paired with metagenomic and metatranscriptomic sequencing, associations between hosts and their environment can be related to pathogens and their functional roles within parasite communities, through blood-meal or gut content analysis[67]. Collecting these 'real-world' observations over time will be especially important to elucidate evolutionary processes, and perturbations that can disrupt competition between parasites, leading to pathogen colonization[51,68,69].

GEpMs need not be restricted to terrestrial settings, as a similar theory and data gathering approach could be used to develop them for aquatic systems, where the risk posed by infectious diseases is high (such as coastal shorelines). However, in contrast to terrestrial systems, GEpMs would need to be refined to account for differences in aquatic systems that impact the dispersal of pathogens[70]. Experimental trials that focus on aquaculture species could elucidate the dynamics between parasite and host communities, while observational studies conducted at a broader scale could determine the mechanisms that cause certain aquatic habitats, such as marshes[71] and seagrasses[72], to remove and potentially destroy human pathogens that invade these habitats. In both terrestrial and aquatic systems, sentinel interfaces deemed important for inter-species disease transmission and zoonotic pathogen spillover would make particularly useful study sites where the experimental approaches outlined above could be used to link patterns of parasite diversity to host and vector population dynamics, and the environment.

System dynamics and spillover risk

Once built, a GEpM would simulate how functional groups of pathogens behave under varying environmental and anthropological inputs (e.g., spatially explicit data on climate change, habitat,

299 socioeconomics and human distribution), generating results that can be used to evaluate human
300 disease risk across land or seascapes. To achieve this, system structure – comprising cohorts of
301 parasites, their hosts and vectors, each defined by functional traits – would be modelled within
302 grid cells that represent a layer of spatially heterogeneous environmental and anthropological
303 conditions across the land or seascape under consideration[26] (Box 1, Figure 3). In line with
304 existing general ecosystem models, it wouldn't be unreasonable to expect a process-based GEpM
305 to be capable of simulating episystem dynamics within any ecosystem and at any level of spatial
306 resolution. Properties of pathogen communities (e.g., the relative abundance and biomass of
307 different functional groups) would manifest within each grid cell over consecutive model
308 iterations, emerging from macro-scale processes at the level of individual host and vector
309 cohorts, and in accordance with their responses to environmental and anthropogenic conditions
310 within that grid cell (Figure 3). Comparison of pathogen functional group abundance (and host,
311 and vector abundance and distribution) with empirical data collected within sentinel land and
312 seascapes, would enable validation of the model's results under different environmental
313 scenarios.

314 Incorporating human behavior into GEpMs will be critical to account for the impacts of
315 human activities on pathogen community ecology and generate meaningful estimates of human
316 disease risk. With the exception of administering medical treatments to livestock, we would
317 expect anthropological effects to manifest indirectly on parasite communities through changes in
318 the distribution and composition of host and vector populations resulting from the top-down
319 impacts of climate change, human-mediated introduction of invasive species, land-use change
320 and fragmentation, and variation in livestock-keeping or aquaculture practices. As such, rather
321 than including humans and their activities as agents within the model, GEpMs could follow

general ecosystem models in accounting for human impacts as exogenous factors, incorporated into climatic, land-use, socioeconomic or human demographic layers that are inputs for the model[26]. For example, a discrete harvesting parameter based upon socioeconomic data could be used to constrain the growth of livestock cohorts with the model. Socioeconomic determinants of livestock keeping are relatively well understood, and models pairing social, economic and ecological systems show that the impacts of humans on the environment and vice-versa can be modelled in a predictive fashion[73,74].

To estimate human spillover risk, predictions for the abundance and distribution of pathogen functional groups made by GEpMs would need to be expressed in terms of human risk. The risk of disease outbreaks in people can be quantitatively expressed by the following equation: Risk = Hazard x (Vulnerability x Exposure), where hazard is the availability of pathogens to infect a human at any given time and space, exposure is people's contact with these pathogens, and vulnerability is the likelihood of infection occurring upon contact[75]. General mathematical expressions that use this framework to measure animal-to-human spillover risk have been proposed[4,76], and in generating estimates of abundance for pathogen cohorts, GEpMs could be used to predict hazard for groups recognized as emergent threats (such as negative-strand RNA viruses, or drug-resistant bacteria) within these models (Figure 3; Box 1).

Control and design

We think that GEpMs could radically improve our understanding of epidemiological processes occurring in human-modified landscapes, directing surveillance and control efforts for emerging diseases, and ultimately identifying the stability of parasite communities within landscapes. Since forecasting of disease emergence is primarily informed by phenomenological studies[77],

GEpMs could ensure that health policy decisions are guided by an understanding of how epidemiological systems actually function. For example, applied to ecological systems under anthropogenic stress (we use the examples of a grassland ecosystem in Figure 3 and coastal ecosystems in Suppl. Figure 1), GEpMs could be used to create dynamic risk maps for priority groups of pathogens (e.g., negative-strand RNA viruses which include zoonotic viruses responsible for Ebola, hantaviruses, influenza, and rabies), and forecast how these might change in response to climate change, land-use change, population and socioeconomic trends. Because pathogen dynamics would emerge from spatially explicit environmental and socioeconomic data, computers of the future could run these models at broad spatial scales to provide real-time forecasting for priority groups of pathogens.

Once armed with a more detailed quantitative and mechanistic understanding of the role of parasites in natural ecosystems, a key question remains how progress can be made towards preventing and controlling outbreaks of infectious agents, or breakdowns in ecosystem services. The best way to confront this might be to 'reverse engineer' these problems. For example, we know that vital ecosystem services such as the cleansing of air and water are driven by a diversity of species within the ecosystem. If these ecosystem functions could be characterized as outputs from general ecosystem or episystem models, it would be possible to examine the ways in which their relative production declines as the abundance and diversity of species that drive the pathways changes (*sensu* Dobson et al.[78]). Applying these principles to emerging infectious diseases, where the primary drivers of animal-to-human spillover are known to be the wildlife trade, and destruction and fragmentation of tropical forests, GEpMs could be used to identify species that carry significant burdens of pathogens with characteristics that would make their appearance in the wildlife trade particularly problematic (low specificity, unusual range of

368 hosts). What would this then tell us about minimizing species loss and reductions in abundance
369 in ways that minimize loss of ecosystem function and reduce risk of human exposure to
370 emerging pathogens? Armed with knowledge of the ecological mechanisms that systematically
371 control the state of host and pathogen communities, novel targets for mitigating spillover risk
372 could be identified and tested[9] – such as creating spatial buffers between hosts, managing habitat
373 to control host and vector populations[79], or encouraging changes in livestock-keeping practices
374 and other behavioral risk factors for disease emergence[80]. In this way, strategies to modify
375 epidemiological processes and thereby disrupt pathogen spillover, could be designed on the basis
376 of '*in-silica*' simulation.

377 The considerable challenges associated with developing these models, and their
378 limitations, should be recognized. As is the case for general ecosystem models, acquiring
379 sufficient data to parameterize and validate GEpMs represents a significant obstacle to their
380 development. We therefore suggest that initial efforts focus on developing GEpMs for areas
381 where long-term studies of free-living species are ongoing, and where concerns are increasingly
382 expressed that pathogens play a crucial but only partially understood role in structuring
383 communities of hosts. For example, longstanding ecological monitoring projects in ecosystems
384 such as Yellowstone[81,82], The Serengeti[83], Gorongosa[84] and the Galápagos National Parks, where
385 rich historical datasets of pathogen prevalence exist from different trophic guilds of hosts, would
386 provide valuable resources with which to begin parameterizing and validating GEpMs[85–87]. To
387 scale predictions beyond well-characterized sentinel landscapes and achieve the impact we
388 envisage relating to predicting emerging disease risk, a coordinated global effort will be
389 required. Although daunting, the challenge of conducting and connecting studies that scale from
390 individual hosts, to host populations in experimental plots and across landscapes, could be met

by a distributed experimental network – a collaborative effort between scientists, consisting of multifactorial studies replicated across many sites, and conducted using standardized protocols that enable comparison and sharing of data[88]. This form of collaboration across sites is not without precedent in ecology – for example the US National Science Foundation's National Ecological Observatory Network (NEON)—which is now collecting data on host and parasite communities)[89,90]—and the Smithsonian's Forest Global Earth Observatory (ForestGEO)[91] and Marine Global Earth Observatory (MarineGEO) networks, apply rigorous, standardized data collection protocols across sites to monitor long-term ecological change. The availability of high-resolution geospatial observations, coupled with rapid advances in autonomous biosensing technology, promise the ability to collect large quantities of biological data across spatial and ecological scales, and at relatively low cost.

Although a sizeable initial grant would be required to establish such a network on an international scale, the necessary expansion would be constrained by hypotheses generated by the model, and costs could be offset through the contribution of these efforts towards mitigation of disease emergence and future pandemics[92]. An experimental network based on voluntary participation, in which contributors benefit from the results of the model by submitting their data to help improve it, would reduce costs and extend its reach into under-resourced areas, paying dividends over the long-term. Finally, to scale predictions of spillover risk beyond well-characterized sentinel landscapes, detailed global inventories of hosts, vectors and their parasites will be required. Large-scale data-gathering programs already exist for phenotypic and genetic diversity of vertebrates, vectors and their pathogens (e.g. PanTHERIA, ViPR (Virus Pathogen resource), NCBI GenBank, VectorBase, Barcode of Life Database (BOLD)) and proposed

initiatives such as the Global Virome Project[93] and a Global Parasite Project[30] will be central to these global efforts.

Progress in linking complex parasite-host-environment systems with elegant mathematical expressions would represent huge advances in the fields of disease ecology, and success should therefore not be assumed. The computational power required to simulate complex systems is a major hurdle. Nevertheless, the development of global general ecosystem models has proven to be achievable by reducing dimensionality (grouping organisms into functional groups, and cohorts within functional groups)[26]. Because GEpMs would necessarily simplify episystems into trait-based groups of pathogens, they will not possess the predictive power to model the behavior of specific pathogens, or determine exactly where and when new pathogens will emerge. For this reason, where the goal is to inform management of the risk associated with specific diseases, we recommend that GEpMs are coupled with more traditional epidemiological models/approaches. By unlocking broader principles that underlie epidemiological processes (*sensu* Lafferty et al.[35]), GEpMs could lead to breakthroughs in the design of more detailed, accurate statistical or agent-based models of specific diseases, while identifying areas that require further investigation.

In the midst of a global pandemic of wildlife origin, the need for models that consider the full ecological and anthropological contexts of disease transmission is clear. By challenging scientists to reconstruct epidemiological processes from the bottom-up and on the basis of ecological principles, systems models could form a new frontier in epidemiology, uncovering new processes and ultimately improving our understanding of disease emergence, and ability to target surveillance activities and interventions at a global scale. The potential benefits to

435 understanding health across species, communities and ecosystems across the planet are
436 enormous.
437
438 **References cited**
439

440 1. Allen, T. *et al.* Global hotspots and correlates of emerging zoonotic diseases. *Nat.*
441 *Commun.* **8**, (2017).
442 2. Gibb, R. *et al.* Zoonotic host diversity increases in human-dominated ecosystems. *Nature*
443 **584**, 398–402 (2020).
444 3. Hassell, J. M., Begon, M., Ward, M. J. & Fèvre, E. M. Urbanization and Disease
445 Emergence: Dynamics at the Wildlife–Livestock–Human Interface. *Trends Ecol. Evol.* **32**,
446 55–67 (2017).
447 4. Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nat. Rev. Microbiol.* **15**, 502–510
448 (2017).
449 5. Holmes, E. C., Rambaut, A. & Andersen, K. G. Pandemics: Spend on surveillance, not
450 prediction comment. *Nature* **558**, 180–182 (2018).
451 6. Tabachnick, W. J. Challenges in predicting climate and environmental effects on vector-
452 borne disease episystems in a changing world. *J. Exp. Biol.* **213**, 946–954 (2010).
453 7. Franklinos, L. H. V., Jones, K. E., Redding, D. W. & Abubakar, I. The effect of global
454 change on mosquito-borne disease. *The Lancet Inf. Dis.* **19**, e302–e312 (2019).
455 8. Seabloom, E. W. *et al.* The community ecology of pathogens: Coinfection, coexistence
456 and community composition. *Ecol. Lett.* **18**, 401–415 (2015).
457 9. Johnson, P. T. J., De Roode, J. C. & Fenton, A. Why infectious disease research needs
458 community ecology. *Science* vol. 349 1259504 (2015).
459 10. Parker, I. M. *et al.* Phylogenetic structure and host abundance drive disease pressure in
460 communities. *Nature* **520**, 542–544 (2015).
461 11. Telfer, S. *et al.* Species interactions in a parasite community drive infection risk in a
462 wildlife population. *Science* **330**, 243–6 (2010).
463 12. Dallas, T. A., Laine, A.-L. L. & Ovaskainen, O. Detecting parasite associations within
464 multi-species host and parasite communities. *Proc. R. Soc. B Biol. Sci.* **286**, 20191109

(2019).
13. Weinstein, S., Titcomb, G., Agwanda, B., Riginos, C. & Young, H. Parasite responses to large mammal loss in an African savanna. *Ecology* **98**, 1839–1848 (2017).
14. Anderson, R. & May, R. *Infectious diseases of humans: dynamics and control*. doi:10.1016/0966-842x(93)90094-8.
15. Keeling, M. J. & Rohani, P. *Modeling infectious diseases in humans and animals*. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2011). doi:10.1016/s1473-3099(08)70147-6.
16. Buhnerkempe, M. G. *et al.* Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics* **10**, 26–30 (2014).
17. Cross, P. C., Prosser, D. J., Ramey, A. M., Hanks, E. M. & Pepin, K. M. Confronting models with data: the challenges of estimating disease spillover. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180435 (2019).
18. Johnson, E. E., Escobar, L. E. & Zambrana-Torrelio, C. An Ecological Framework for Modeling the Geography of Disease Transmission. *Trends Ecol. Evol.* **34**, 655–668 (2019).
19. Warton, D. I. *et al.* So Many Variables: Joint Modeling in Community Ecology. *Trends Ecol. Evol.* **30**, 766–779 (2015).
20. Carlson, C. J. *et al.* The global distribution of Bacillus anthracis and associated anthrax risk to humans, livestock and wildlife. *Nat. Microbiol.* **4**, 1337–1343 (2019).
21. Sutherland, W. J. Predicting the ecological consequences of environmental change: A review of the methods. *J.Applied Ecol.* **43**, 599–616 (2006).
22. Getz, W. M. *et al.* Making ecological models adequate. *Ecol. Lett.* **21**, 153–166 (2018).
23. Carlson, C. J., Chipperfield, J. D., Benito, B. M., Telford, R. J. & O'Hara, R. B. Species distribution models are inappropriate for COVID-19. *Nature Ecology and Evolution*, **4** 770–771 (2020).
24. Evans, M. R. *et al.* Predictive systems ecology. *Proc. R. Soc. B Biol. Sci.* **280**, 20131452–20131452 (2013).
25. Purves, D. & Pacala, S. Predictive models of forest dynamics. *Science* **320**,1452–1453 (2008).
26. Harfoot, M. B. J. *et al.* Emergent Global Patterns of Ecosystem Structure and Function

from a Mechanistic General Ecosystem Model. *PLoS Biol.* **12**, (2014).
27. Lafferty, K. D. *et al.* Parasites in food webs: The ultimate missing links. *Ecol. Lett.* **11**, 533–546 (2008).
28. Redding, D. W. *et al.* Impacts of environmental and socio-economic factors on emergence and epidemic potential of Ebola in Africa. *Nat. Commun.* **10**, 1–11 (2019).
29. Redding, D. W., Moses, L. M., Cunningham, A. A., Wood, J. & Jones, K. E. Environmental-mechanistic modelling of the impact of global change on human zoonotic disease emergence: a case study of Lassa fever. *Methods Ecol. Evol.* **7**, 646–655 (2016).
30. Carlson, C. J., Dallas, T. A., Alexander, L. W., Phelan, A. L. & Phillips, A. J. What would it take to describe the global diversity of parasites? *Proc. R. Soc. B Biol. Sci.* **287**, 20201841 (2020).
31. Rynkiewicz, E. C., Pedersen, A. B. & Fenton, A. An ecosystem approach to understanding and managing within-host parasite community dynamics. *Trends Parasitol.* **31**, 212–221 (2015).
32. Lello, J. & Hussell, T. Functional group/guild modelling of inter-specific pathogen interactions: A potential tool for predicting the consequences of co-infection. *Parasitology* vol. 135 825–839 (2008).
33. Purves, D. *et al.* Time to model all life on Earth. *Nature* **493**, 295–297 (2013).
34. Kalka, M. B., Smith, A. R. & Kalko, E. K. V. Bats limit arthropods and herbivory in a tropical forest. *Science* vol. 320 71 (2008).
35. Lafferty, K. D. *et al.* A general consumer-resource population model. *Science.* **349**, 854–857 (2015).
36. Friedman, J., Higgins, L. M. & Gore, J. Community structure follows simple assembly rules in microbial microcosms. *Nat. Ecol. Evol.* **1**, 109 (2017).
37. Goldford, J. E. *et al.* Emergent simplicity in microbial community assembly. *Science.* **361**, 469–474 (2018).
38. Hatton, I. A. *et al.* The predator-prey power law: Biomass scaling across terrestrial and aquatic biomes. *Science.* **349**, 6252 (2015).
39. Hatton, I. A., Dobson, A. P., Storch, D., Galbraith, E. D. & Loreau, M. Linking scaling laws across eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21616–21622 (2019).
40. Locey, K. J. & Lennon, J. T. Scaling laws predict global microbial diversity. *Proc. Natl.*

Acad. Sci. U. S. A. **113**, 5970–5975 (2016).

41. Godon, J. J., Arulazhagan, P., Steyer, J. P. & Hamelin, J. Vertebrate bacterial gut diversity: Size also matters. *BMC Ecol.* **16**, 12 (2016).

42. Faust, C. L. *et al.* Null expectations for disease dynamics in shrinking habitat: dilution or amplification? *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160173 (2017).

43. De Leo, G. A. & Dobson, A. P. Allometry and simple epidemic models for microparasites. *Nature* **379**, 720–722 (1996).

44. Strauss, A. T., Shoemaker, L. G., Seabloom, E. W. & Borer, E. T. Cross-scale dynamics in community and disease ecology: relative timescales shape the community ecology of pathogens. *Ecology* **100**, (2019).

45. Handel, A. & Rohani, P. Crossing the scale from within-host infection dynamics to between-host transmission fitness: A discussion of current assumptions and knowledge. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).

46. Tibayrenc, M. & Ayala, F. J. Reproductive clonality of pathogens: A perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proc. Natl. Acad. Sci. USA* **109**, E3305–E3313 (2012).

47. Gorter, F. A., Manhart, M. & Ackermann, M. Understanding the evolution of interspecies interactions in microbial communities. *Phil. Trans. Royal Soc. B: Biol. Sci*. **375**, 1798 (2020).

48. Zeng, Q., Wu, S., Sukumaran, J. & Rodrigo, A. Models of microbiome evolution incorporating host and microbial selection. *Microbiome* **5**, 127 (2017).

49. Zeng, Q., Sukumaran, J., Wu, S. & Rodrigo, A. Neutral Models of Microbiome Evolution. *PLOS Comput. Biol.* **11**, e1004365 (2015).

50. Liautaud, K., van Nes, E. H., Barbier, M., Scheffer, M. & Loreau, M. Superorganisms or loose collections of species? A unifying theory of community patterns along environmental gradients. *Ecol. Lett.* **22**, 1243–1252 (2019).

51. Coyte, K. Z., Schluter, J. & Foster, K. R. The ecology of the microbiome: Networks, competition, and stability. *Science.* **350**, 663–666 (2015).

52. Wright, E. S. & Vetsigian, K. H. Inhibitory interactions promote frequent bistability among competing bacteria. *Nat. Commun.* **7**, 1–7 (2016).

53. Hekstra, D. R. & Leibler, S. Contingency and statistical laws in replicate microbial closed

ecosystems. *Cell* **149**, 1164–1173 (2012).
54. Lopatkin, A. J. & Collins, J. J. Predictive biology: modelling, understanding and harnessing microbial complexity. *Nat. Rev. Microbiol.* **18**, 507–520 (2020).
55. Sanchez, A. & Gore, J. Feedback between Population and Evolutionary Dynamics Determines the Fate of Social Microbial Populations. *PLoS Biol.* **11**, e1001547 (2013).
56. Wu, F. *et al.* A unifying framework for interpreting and predicting mutualistic systems. *Nat. Commun.* **10**, 1–10 (2019).
57. Restif, O. *et al.* Model-guided fieldwork: practical guidelines for multidisciplinary research on wildlife ecological and epidemiological dynamics. *Ecol. Lett.* **15**, 1083–94 (2012).
58. Herzog, S. A., Blaizot, S. & Hens, N. Mathematical models used to inform study design or surveillance systems in infectious diseases: A systematic review. *BMC Infect. Dis.* **17**, 1–10 (2017).
59. Cotterill, G. G. *et al.* Winter feeding of elk in the Greater Yellowstone Ecosystem and its effects on disease dynamics. *Philos. Trans. R. Soc. B Biol. Sci.* **373**, 20170093 (2018).
60. Cross, P. C. *et al.* Estimating distemper virus dynamics among wolves and grizzly bears using serology and Bayesian state-space models. *Ecol. Evol.* **8**, 8726–8735 (2018).
61. Hopkins, J. B., Ferguson, J. M., Tyers, D. B. & Kurle, C. M. Selecting the best stable isotope mixing model to estimate grizzly bear diets in the Greater Yellowstone Ecosystem. *PLoS One* **12**, e0174903 (2017).
62. Schwartz, C. C. *et al.* Body and diet composition of sympatric black and grizzly bears in the Greater Yellowstone Ecosystem. *J. Wildl. Manage.* **78**, 68–78 (2014).
63. Chester, C. C. Yellowstone to Yukon: Transborder conservation across a vast international landscape. *Environ. Sci. Policy* **49**, 75–84 (2015).
64. Young, H. S. *et al.* Interacting effects of land use and climate on rodent-borne pathogens in central Kenya. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160116 (2017).
65. Sitters, J., Kimuyu, D. M., Young, T. P., Claeys, P. & Olde Venterink, H. Negative effects of cattle on soil carbon and nutrient pools reversed by megaherbivores. *Nat. Sustain.* **3**, 360–366 (2020).
66. Sethi, S. S., Ewers, R. M., Jones, N. S., Orme, C. D. L. & Picinali, L. Robust, real-time and autonomous monitoring of ecosystems with an open, low-cost, networked device.

Methods Ecol. Evol. **9**, 2383–2387 (2018).
67. Alfano, N., Dayaram, A. & Tsangaras, K. Non-invasive surveys of mammalian viruses using environmental DNA. *bioRxiv Mol. Biol.* 2020.03.26.009993 (2020)
68. Coyte, K. Z. & Rakoff-Nahoum, S. Understanding Competition and Cooperation within the Mammalian Gut Microbiome. *Current Biology* vol. 29 R538–R544 (2019).
69. Murray, M. H. *et al.* Gut microbiome shifts with urbanization and potentially facilitates a zoonotic pathogen in a wading bird. *PLoS One* **15**, (2020).
70. McCallum, H. I. *et al.* Does terrestrial epidemiology apply to marine systems? *Trends Ecol. Evol.* **19** 585–591 (2004).
71. Wu, S., Carvalho, P. N., Müller, J. A., Manoj, V. R. & Dong, R. Sanitation in constructed wetlands: A review on the removal of human pathogens and fecal indicators. *Sci. Total. Environ.* **541** 8–22 (2016).
72. Lamb, J. B. *et al.* Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. *Science.* **355**, 731–733 (2017).
73. Janssen, M. A., Walker, B. H., Langridge, J. & Abel, N. An adaptive agent model for analysing co-evolution of management and policies in a complex rangeland system. *Ecol. Modell.* **131**, 249–268 (2000).
74. Ngonghala, C. N. *et al.* General ecological models for human subsistence, health and poverty. *Nat. Ecol. Evol.* **1**, 1153–1159 (2017).
75. Hosseini, P. R. *et al.* Does the impact of biodiversity differ between emerging and endemic pathogens? The need to separate the concepts of hazard and risk. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160129 (2017).
76. Washburne, A. D. *et al.* Percolation models of pathogen spillover. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180331 (2019).
77. Olival, K. J. *et al.* Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
78. Dobson, A. *et al.* HABITAT LOSS, TROPHIC COLLAPSE, AND THE DECLINE OF ECOSYSTEM SERVICES. *Ecology* **87**, 1915–1924 (2006).
79. Faust, C. L. *et al.* Pathogen spillover during land conversion. *Ecol. Lett.* **21**, 471–483 (2018).
80. Sokolow, S. H. *et al.* Ecological interventions to prevent and manage zoonotic pathogen

spillover. *Philos. Trans. R. Soc. B Biol. Sci.* **374**, 20180342 (2019).
81. Kauffman, M. J. *et al.* Landscape heterogeneity shapes predation in a newly restored predator-prey system. *Ecol. Lett.* **10**, 690–700 (2007).
82. Smith, D. W., Peterson, R. O. & Houston, D. B. Yellowstone after wolves. *BioScience.* **53,** 330–340 (2003).
83. McNaughton, S. J. Ecology of a Grazing Ecosystem: The Serengeti. *Ecol. Monogr.* **55**, 259–294 (1985).
84. Atkins, J. L. *et al.* Cascading impacts of large-carnivore extirpation in an African ecosystem. *Science.* **364**, 173–177 (2019).
85. Cross, P. C., Edwards, W. H., Scurlock, B. M., Maichak, E. J. & Rogerson, J. D. Effects of Management and Climate on Elk Brucellosis in the Greater Yellowstone Ecosystem. *Ecol. Appl.* **17**, 957–964 (2007).
86. Almberg, E. S., Cross, P. C. & Smith, D. W. Persistence of canine distemper virus in the Greater Yellowstone Ecosystem's carnivore community. *Ecol. Appl.* **20**, 2058–2074 (2010).
87. Holdo, R. M. *et al.* A Disease-Mediated Trophic Cascade in the Serengeti and its Implications for Ecosystem C. *PLoS Biol.* **7**, e1000210 (2009).
88. Borer, E. T. Ecological Scaling in Time and Space. in *Unsolved Problems in Ecology* (eds. Dobson, A. P., Holt, R. D. & Tilman, D.) 3–15 (Princeton University Press, 2020).
89. Kao, R. H. *et al.* NEON terrestrial field observations: designing continental-scale, standardized sampling. *Ecosphere* **3**, art115 (2012).
90. Springer, Y. P. *et al.* Tick-, mosquito-, and rodent-borne parasite sampling designs for the National Ecological Observatory Network. *Ecosphere* **7**, (2016).
91. Anderson-Teixeira, K. J. *et al.* CTFS-ForestGEO: A worldwide network monitoring forests in an era of global change. *Glob. Chang. Biol.* **21**, 528–549 (2015).
92. Dobson, A. P. *et al.* Ecology and economics for pandemic prevention. *Science* **369**, 379–381 (2020).
93. Kress, W. J., Mazet, J. A. K. & Hebert, P. D. N. Opinion: Intercepting pandemics through genomics. *Proc. Natl. Acad. Sci.* **117**, 202009508 (2020).
94. Pedersen, A. B. & Fenton, A. Emphasizing the ecology in parasite community ecology. *Trends Ecol. Evol.* **22**, 133–139 (2007).

95. Plowright, R. K. *et al.* A call to action: Understanding land use-induced zoonotic spillover to protect environmental, animal, and human health. *EcoEvoRxiv* (2020) doi:10.32942/OSF.IO/CRU9W.
96. Barychka, T., Mace, G. & Purves, D. The Madingley General Ecosystem Model predicts bushmeat yields, species extinction rates and ecosystem-level impacts of bushmeat harvesting. *Cold Spring Harb. Lab. Cold Spring Harb. NY, USA.* 2020.03.02.959718 (2020) doi:10.1101/2020.03.02.959718.
97. Durmuş, S. & Ülgen, K. Comparative interactomics for virus–human protein–protein interactions: DNA viruses versus RNA viruses. *FEBS Open Bio* **7**, 96–107 (2017).
98. Becker, D. J. & Albery, G. F. Expanding host specificity and pathogen sharing beyond viruses. *Mol. Ecol.* **29**, 3170–3172 (2020).
99. Rittershaus, E. S. C., Baek, S. H. & Sassetti, C. M. The normalcy of dormancy: Common themes in microbial quiescence. *Cell Host Microbe.* **13**, 643–651 (2013).
100. Schmid-Hempel, P. Immune defence, parasite evasion strategies and their relevance for 'macroscopic phenomena' such as virulence. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 85–98 (2009).

Acknowledgements

J.H., D.Z. & Y.M.L. were supported by the US Army Medical Research and Development Command under Contract No.W81XWH-21-C-0001 and the Armed Forces Health Surveillance Division - Global Emerging Infections Surveillance branch (AFHSC-GEIS) award P0031_21_WR. The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation. T.N. is supported by two grants from the Leverhulme Trust (RPG-2015-073 and RPG-2018-069).

Contributions

J.M.H. conceptualized the structure and content of the manuscript and wrote an initial draft. J.M.H., T.N., A.P.D., Y.L., L.V.H.F., D.Z. and K.M.P. expanded upon the ideas contained within this initial draft, and engaged in discussion and editing of the final manuscript.

Corresponding author

Correspondence to James M. Hassell (hasselljm@si.edu)

Box 1: Modeling parasites as cohorts

Grouping individuals by their ecological traits is the principal form of dimension reduction used in General Ecosystem Models (GEMs), and an approach that we propose could also be applied when developing GEpMs. In terrestrial GEMs, autotrophs (plants) and heterotrophs (herbivorous, omnivorous, and carnivorous animals) are grouped by nutrition source, mobility, leaf strategy (autotrophs), mobility, reproductive strategy, and thermoregulation mode (heterotrophs). GEpMs would extend GEMs, adding parasites as a second group of heterotrophs that are modelled differently to their hosts (see Harfoot et al.[26] for a detailed description of how autotrophs and heterotrophs are modelled in GEMs). Drawing on generalized frameworks developed by Pedersen & Fenton[94], Lafferty et al.[35] and Lello & Hussell[32] we propose six categorical traits that represent the ecological processes conducted by parasites, and their interactions with hosts (Table 1). Once grouped by these traits, the resource exploitation strategies of individual parasites within each cohort would be modelled using the same mathematical expressions that represent; *(i)* consumption strategy and impact on host fitness; *(ii)* immune stimulation and immune evasion (e.g., quiescence); *(iii)* reproduction; *(iv)* mortality resulting from the host immune system, or as a result of background mortality processes such as

senescence; and *(v)* dispersion from their current grid cell to another grid cell (Figure 3). The impact of parasites on host fitness (e.g., through consumer strategies that either reduce host fitness to zero or have a density dependent reduction on the reproductive performance of hosts) would feed back into the modelling of host heterotroph cohorts, and their effects on autotroph biomass.

Case study: hazard posed by negative-strand RNA viruses in changing terrestrial systems. Human-mediated ecosystem change is considered an important driver of animal-to-human pathogen spillover, but the macro-ecological processes by which this occurs are rarely studied and poorly understood[95]. GEpMs would offer a unique opportunity to simulate the impacts of ecosystem changes (e.g., land use change, harvesting of wild animals) on host populations, and emerging pathogens. Using this as a scenario to demonstrate the potential application of GEpM's, we describe how a prototype model could be used to study the dynamics of negative-strand (NS)-RNA viruses in wild animals, generate predictions of the hazard they pose to humans, and design interventions to protect human health. Following the functional groupings in Table 1, models could target parasites described using the categorical traits 'Pathogen | Intracellular-RNA-reverse transcription | Horizontal-direct | Cellular/Humoral/T-helper cell'. By specifying these classifications, important zoonotic viral families such as orthomyxoviruses, paramyxoviruses and filoviruses would be targeted.

Figure 3 depicts how modeling studies conducted across grid cells at different resolutions could assess the GEpM's capacity to simulate ecosystem-scale dynamics across trophic levels from which (NS)-RNA virus properties emerge, and generate high-resolution predictions of the relative abundance/biomass of (NS)-RNA viruses at specific sites undergoing ecosystem

changes. By sourcing environmental input data from closely monitored sites experiencing changes in land use over a defined period, and aligning this to the time steps over which simulations occur, the predicted responses of host and parasite cohorts could be evaluated against empirical data on vegetation, host and parasite abundance. A term that simulates harvesting of certain wild animal host cohorts could then be added to the model to investigate how specific changes in trophic structure influence parasite dynamics[96]. As an emergent property of the GEpM, the relative abundance and biomass of the (NS)-RNA virus cohort could estimate 'pathogen pressure' for each grid cell on which the model is run – representing the quantity of (NS)-RNA viruses in wildlife to which humans could be exposed at a given point in space and time. Over multiple grid cells, these predictions would represent the distribution of wild animals carrying these pathogens, and the intensity with which they are infected and shedding them (i.e., persistence and transmission within wild animal populations). When combined with information on human-wildlife interactions and human susceptibility to infection, this data could be used to predict spillover risk at local, national and global scales. Including livestock hosts would increase the accuracy of these models, and we demonstrate how this could be achieved in Figure 3.

Furthermore, these models could permit "*in-silico*" design and testing of interventions aimed at maintaining stable population dynamics of species and their pathogens and mediating human behavior in a way that minimizes the impact of land-use change on biodiversity and human health. For example, a GEpM that describes changes in the predator-prey dynamics of non-human primates in response to fragmentation of tropical forests, and predicts how this impacts their exposure to zoonotic viruses, could be used to forecast the human health risks posed by hunting these species within a given area, and target educational campaigns at

communities who rely on non-human primates as a food source. As new empirical findings emerge, GEpMs could be used to scale and test competing hypotheses for how ecosystem stressors impact host assemblages and the (NS)-RNA viruses they carry, identifying critical processes that require further investigation.

Resource Use	Reproductive Strategy			Metabolism	Immune Response	Evolution
Consumer Strategy [35]	Location [97]	Dispersal	Host Breadth [98]	Dormancy/Cellular Quiescence [99]	Type of Immune Response [94,100]	Clonality [46]
Castrator	Intracellular, DNA reverse transcription	Horizontal - direct	Composite measure for each pathogen functional group based on databases of host-parasite associations.	No dormant phase	Cellular	Clonal
Macroparasite	Intracellular, DNA non-reverse transcription	Horizontal - indirect		Can perform dormancy	Humoral	Not clonal
Pathogen	Intracellular, RNA reverse transcription	Vertical			T-helper cell	
Parasitoid	Intracellular, RNA non-reverse transcription					
	Intracellular, binary fission / horizontal gene transfer					
	Extracellular, within-host, asexual					
	Extracellular, within-host, sexual					
	Extracellular, environmental, asexual					
	Extracellular, environmental, sexual					

Table 1. Parasite functional groups. To simplify the process of modeling diverse parasite communities, we propose splitting parasites into functionally related groups that represent their consumer strategies, reproductive and metabolic processes, interaction with the host's immune response and evolutionary traits. These classifications represent how parasites *i*) use host resources *(what they eat and how this impacts host fitness)*, *ii*) reproduce *(how they reproduce, and the mode and extent of their dissemination to other hosts)*, *iii*) respond to stressors *(whether they are capable of entering dormancy or not)*, *iv*) activate the host immune response *(components of the host immune system that are stimulated by each pathogen functional group)*, and *v*) evolve *(as differentiated by the levels of genetic recombination that parasites undergo)*.

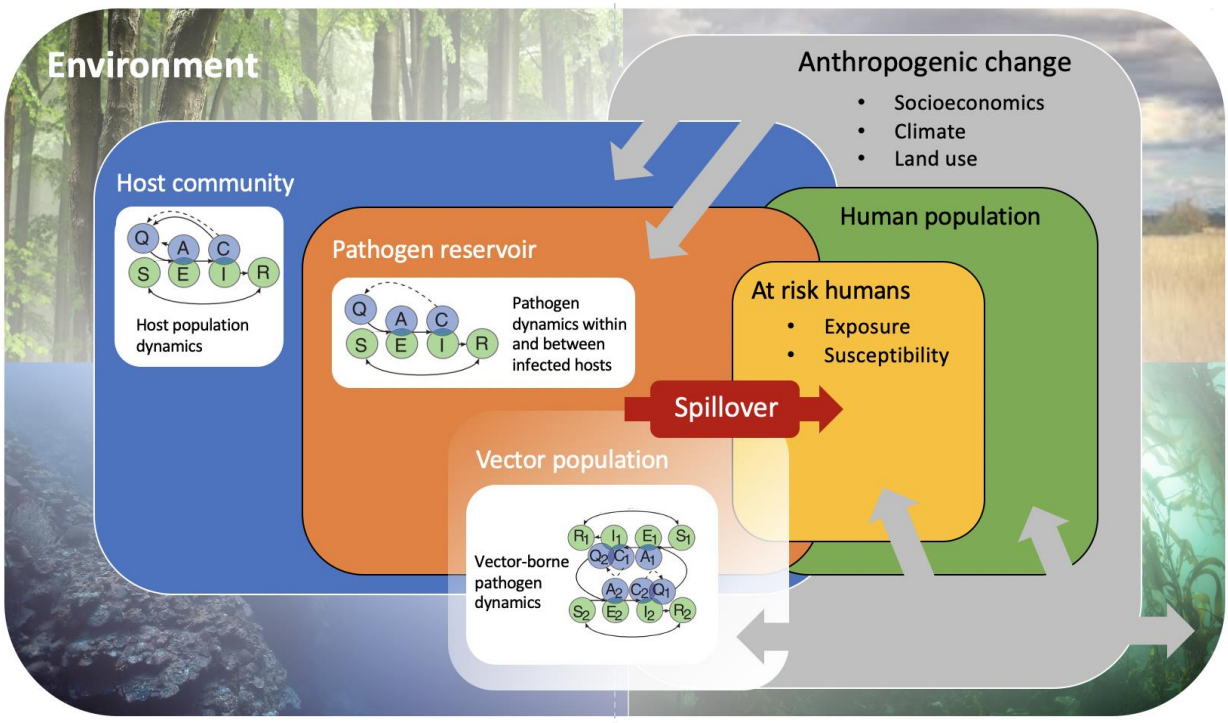


Figure 1. Diagrammatic representation of a disease episystem, depicting interactions between pathogens, their hosts and the environment, and the interface for spillover into people. Pictures represent four terrestrial and marine biomes (forest, grassland, coral reef and

kelp forest), and colored boxes nested within this represent host (animal and human), vector and pathogen populations. Anthropogenic factors that drive changes in environment, host and vector populations are depicted in grey, with arrows showing directionality of these effects. White boxes within animal host and vector compartments represent classic consumer-resource models, depicting host-environment, host-pathogen and vector-pathogen interactions (adapted from Lafferty et al.[21]). Circles within boxes are state variables for questing (Q), attacking (A), and consuming (C) consumers (blue – predators, or pathogens) and susceptible (S), exposed (E), ingested (I), and resistant (R) resources (green – autotrophs, or hosts). Per Lafferty et al.[21], arrows represent transitions (of individuals or biomass) among states – a dashed line represents production or conversion (e.g., births), whereas a solid line is a transition from one state to another (implying no change in numbers from one state to the next). Circles numbered "1" for the model of vector-borne pathogen dynamics represent processes occurring in the vector, and those numbered "2" represent processes occurring in the host.

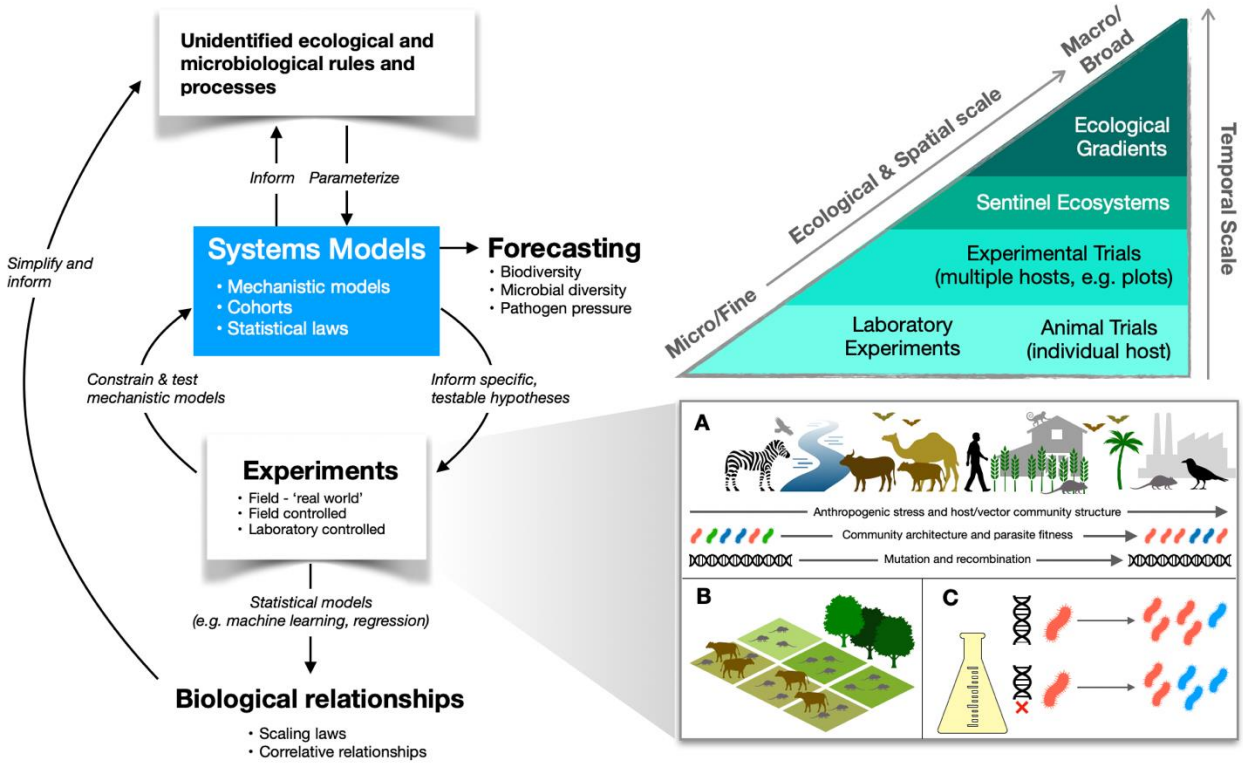
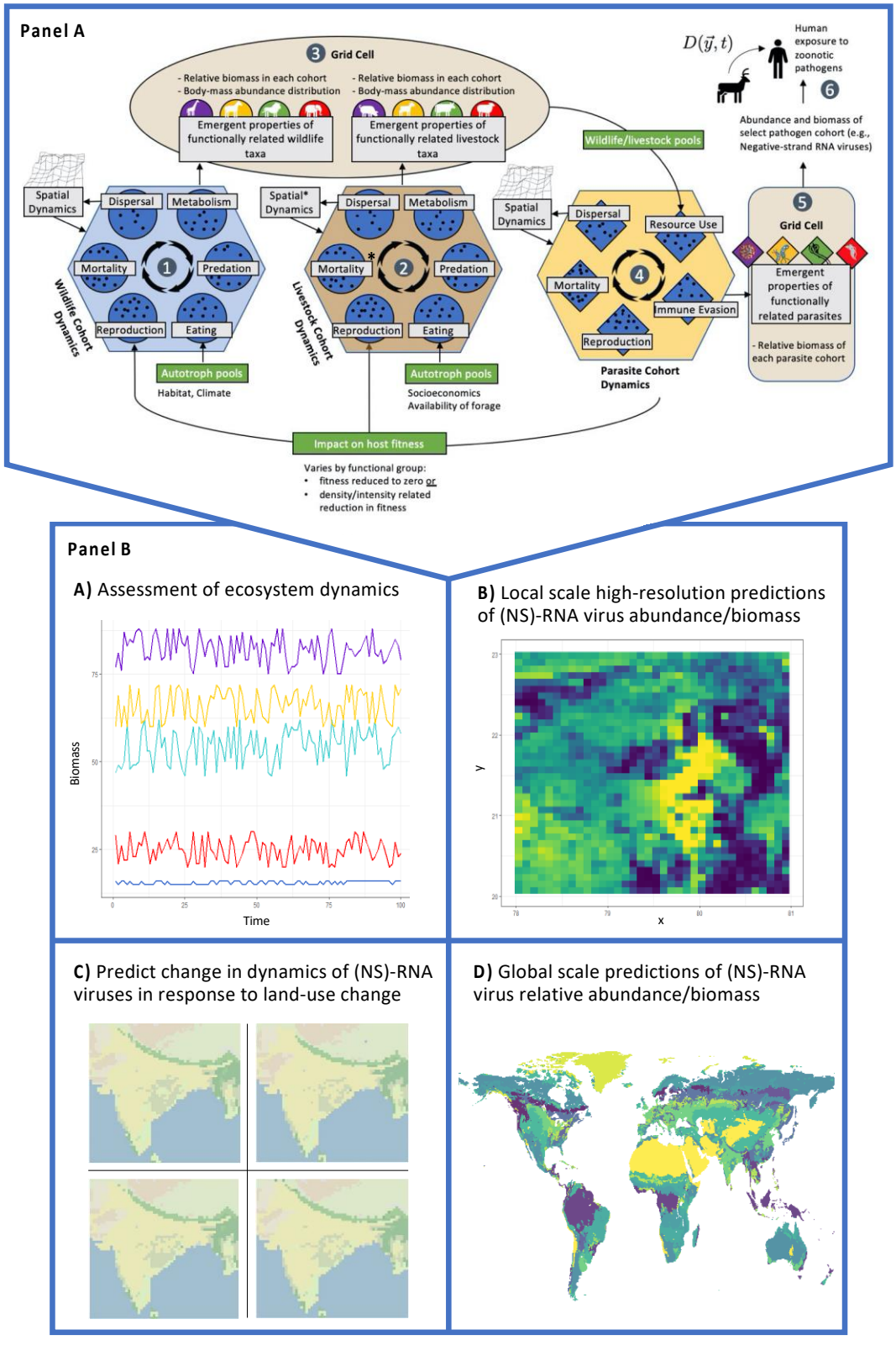


Figure 2. Iterative development of an ecosystem model for infectious disease (General Episystems Model - GEpM). Panel 1: Development of an ecosystem model for infectious disease would be an iterative process, in which systems models (collections of interacting models representing the GEpM) are constrained and tested through field and laboratory experiments conducted over varying spatial and temporal scales. In this way, statistical models that explain complex but important relationships could be incorporated into a mechanistic modeling framework, as a means of decreasing complexity while maintaining predictive power. Types of experiment depicted represent **a**) 'real world' field experiments, where studies investigate species turnover and related evolutionary processes along gradients of anthropogenic stress in ecosystems; **b**) controlled field trials, where conditions that closely mimic the ecological processes of interest are simulated to improve model accuracy; **c**) controlled laboratory trials,

where conditions that closely mimic the microbiological (both ecology and evolutionary) processes of interest are simulated to improve model accuracy. To capture the multitude of ecological scales across which parasites interact with one-another and their hosts, and these interactions are filtered by environmental variables, experiments would need to take place across spatial and temporal scales. Together, these experiments also serve to address unanswered questions in ecology and microbiology—as identified during model development—improving predictive capability and simplifying model structure. **Panel 2:** Initial steps that could be taken towards the development of GEpMs are outlined in this table, along with some of the key challenges facing development of these models.


Panel A
3 Grid Cell
- Relative biomass in each cohort
- Body-mass abundance distribution
Emergent properties of functionally related wildlife taxa
- Relative biomass in each cohort
- Body-mass abundance distribution
Emergent properties of functionally related livestock taxa
$D(\vec{y}, t)$
Human exposure to zoonotic pathogens
6
Abundance and biomass of select pathogen cohort (e.g., Negative-strand RNA viruses)
Wildlife/livestock pools
Spatial Dynamics
Dispersal
Metabolism
Mortality
1
Predation
Reproduction
Eating
Wildlife Cohort Dynamics
Autotroph pools
Habitat, Climate
Spatial* Dynamics
Dispersal
Metabolism
Mortality
2
Predation
Reproduction
Eating
Livestock Cohort Dynamics
Autotroph pools
Socioeconomics
Availability of forage
Spatial Dynamics
Dispersal
Resource Use
Mortality
4
Immune Evasion
Reproduction
Parasite Cohort Dynamics
5
Grid Cell
Emergent properties of functionally related parasites
- Relative biomass of each parasite cohort
Impact on host fitness
Varies by functional group:
• fitness reduced to zero or
• density/intensity related reduction in fitness
Panel B
A) Assessment of ecosystem dynamics
Biomass
Time
B) Local scale high-resolution predictions of (NS)-RNA virus abundance/biomass
y
x
C) Predict change in dynamics of (NS)-RNA viruses in response to land-use change
D) Global scale predictions of (NS)-RNA virus relative abundance/biomass

854

Figure 3. Schematic of a General Episystems Model (GEpM) as applied to predict the hazard posed by Negative-strand RNA viruses. Panel A: Following the Madingley model[26], wildlife are modelled as individuals within cohorts, defined by categorical and quantitative traits. Autotroph biomass (derived from spatially explicit land use per grid cell and climatic variables, economic data and the availability of forage) are used as input data into the wildlife **(1)** and livestock **(2)** models. Each grid cell is stocked with initial densities of wildlife, livestock and their parasites, which could be negatively scaled to body masses randomly drawn from a designated range for each cohort[26]. A term that simulates commercial harvesting of livestock could be included in livestock models **(2*)**. Allometric relationships, combined with spatial models in 1 and 2 lead to emergent properties of wildlife and livestock cohorts across a grid cell **(3)**. Parasites are also modelled as cohorts of functionally related taxa. Emergent properties of wildlife and livestock cohorts ('host pools') in each grid cell inform allometric relationships between parasites and their hosts, and models which capture transmission between hosts **(4)**. Emergent properties of parasite models feed back to impact host dynamics, and result in measures of parasite community structure that can be projected across grid cells – including the abundance/biomass of pathogen cohorts **(5)**. Mathematical expressions couple changes in host and pathogen dynamics with socioeconomic and behavioral models to predict zoonotic spillover risk **(6)**. **Panel B:** The GEpM is used to **A)** make basic assessments of ecosystem dynamics across trophic scales from which (NS)-RNA virus properties emerge, and assess whether these dynamics reach an equilibrium (colors represent different host and parasite cohorts); **B)** make high-resolution predictions of the relative abundance/biomass of (NS)-RNA viruses at specific sites, where empirical data on vegetation, mammalian and parasite abundance or biomass exist; **C)** extend these predictions to forecast changes in relative abundance/biomass of (NS)-RNA viruses in response to land-use change or harvesting of certain host cohorts at specific sites, and **D)** make global, lower-resolution predictions of the relative abundance/biomass of (NS)-RNA viruses