Towards an Eco-phylogenetic Framework for Infectious Disease Ecology

Nicholas M. Fountain-Jones
University of Minnesota

William D. Pearse
Utah State University

Luis E. Escobar
University of Minnesota

Ana Alba-Casals
University of Minnesota

Scott Carver
University of Tasmania

T. Jonathan Davies
McGill University

See next page for additional authors

Follow this and additional works at: https://digitalcommons.usu.edu/biology_facpub

Part of the Biology Commons

Recommended Citation
Fountain-Jones, Nicholas M.; Pearse, William D.; Escobar, Luis E.; Alba-Casals, Ana; Carver, Scott; Davies, T. Jonathan; Kraberger, Simona; Papes, Monica; Vandegrift, Kurt; Worsley-Tonks, Katherine; and Craft, Meggan E., “Towards an Eco-phylogenetic Framework for Infectious Disease Ecology” (2017), Biology Faculty Publications. Paper 1573.
https://digitalcommons.usu.edu/biology_facpub/1573
Towards an eco-phylogenetic framework for infectious disease ecology

Nicholas M. Fountain-Jones¹*, William D. Pearse², Luis E. Escobar¹,³, Ana Alba-Casals¹, Scott Carver⁴, T. Jonathan Davies⁵, Simona Kraberger⁶, Monica Papeş⁷, Kurt Vandegrift⁸, Katherine Worsley-Tonks¹ and Meggan E. Craft¹

¹Department of Veterinary Population Medicine, University of Minnesota, St Paul, Minnesota 55108, USA
²Ecology Center and Department of Biology, Utah State University, Logan, Utah, 84321, USA
³Department of Fisheries, Wildlife and Conservation Biology University of Minnesota, St Paul, Minnesota 55108, USA
⁴School of Biological Sciences, University of Tasmania, Hobart 7001, Australia
⁵Department of Biology, McGill University, Montreal, Quebec H3A 1B1, Canada
⁶Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO 80523, USA
⁷Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA
⁸The Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA
ABSTRACT

Identifying patterns and drivers of infectious disease dynamics across multiple scales is a fundamental challenge for modern science. There is growing awareness that it is necessary to incorporate multi-host and/or multi-parasite interactions to understand and predict current and future disease threats better, and new tools are needed to help address this task. Eco-phylogenetics (phylogenetic community ecology) provides one avenue for exploring multi-host multi-parasite systems, yet the incorporation of eco-phylogenetic concepts and methods into studies of host pathogen dynamics has lagged behind. Eco-phylogenetics is a transformative approach that uses evolutionary history to infer present-day dynamics. Here, we present an eco-phylogenetic framework to reveal insights into parasite communities and infectious disease dynamics across spatial and temporal scales. We illustrate how eco-phylogenetic methods can help untangle the mechanisms of host–parasite dynamics from individual (e.g. co-infection) to landscape scales (e.g. parasite/host community structure). An improved ecological understanding of multi-host and multi-pathogen dynamics across scales will increase our ability to predict disease threats.

Key words: co-infection, ecological niche modelling, multi-host, multi-parasite, pathogens, phylodynamics, phylogenetic community ecology, spill-over, transmission.
(1) Phylogenetic ecology

(2) A guide to eco-phylogenetic methods and complementary tools

(a) Bayesian phylodynamics

(b) Co-phylogenetics

(c) Distance-based methods

(d) Ecological niche modelling

(e) Phylo-diversity

(f) Phylogenetic Generalised Linear Mixed Models (PGLMM)

II. Co-infection

III. Transmission dynamics and pathogen gene flow

(1) Eco-phylogenetic approaches

(2) Application of an eco-phylogenetic framework to analyse drivers of pathogen transmission

IV. Multi-host complex

V. Parasite community structure

VI. Challenges

(1) Data

(2) Pattern, process, and functional characterisation

VII. Future directions

VIII. Conclusions

IX. Acknowledgements

X. References
I. INTRODUCTION

(1) Phylogenetic ecology

Merging community and disease ecology is increasingly considered necessary to understand infectious disease dynamics (Seabloom et al., 2015; Rynkiewicz, Pedersen & Fenton, 2015; Johnson, de Roode & Fenton, 2015). Hosts are frequently infected by multiple parasites (even multiple genotypes of the same parasite), and parasites often circulate among multiple hosts (both intra- and inter-host species) (Cleaveland, Laurenson, & Taylor, 2001; Vaumourin et al., 2015). Potential parasite–parasite and host–parasite synergisms and antagonisms are not only important for host recovery or mortality (e.g. Druilhe, Tall & Sokhna, 2005), but also can be a driving force in parasite evolution and diversification (e.g. Vandegrift et al., 2010). There has, therefore, been a shift in disease-driven research towards encompassing multi-host and multi-parasite (MH–MP) interactions; advancing the ‘one host one parasite’ paradigm, which has provided basic theory essential for the field, but which is limited in scope. While it is now increasingly recognised that community ecology is critical for our understanding of host–pathogen systems (Pedersen & Fenton, 2007; Rigaud et al., 2010; Johnson et al., 2015), a new framework is required to embrace the ecological and evolutionary complexity of host–pathogen dynamics across scales.

Over the last decade, a subdiscipline of ecology has emerged with the potential to address the challenge of MH–MP dynamics in disease biology. Advancement of sequencing and phylogenetic methods, coupled with evolving bioinformatic tools, has helped integrate community ecology and phylogenetic biology, giving rise to the new field of eco-phylogenetics (Webb et al., 2002). The merging of these two previously separate fields has enabled a new and greater understanding of the complex mechanisms driving non-parasitic communities and is now
a central component of macroecology (Cavender-Bares et al., 2009). While population genetics (e.g. phylodynamics, see Grenfell et al., 2004; Archie, Luikart & Ezenwa, 2009) has facilitated major inroads into understanding single-pathogen dynamics, here we explain how eco-phylogenetics can further aid progress in understanding MH–MP systems, from individual to landscape scales (Fitzpatrick & Keller, 2015; Gilbert & Parker, 2016).

The integration of phylogenetics within community ecology has been transformative (Vamosi et al., 2009; Cavender-Bares et al., 2009). Eco-phylogenetics was originally predicated on the simple assumption that more closely related species will tend to resemble each other more closely in their ecologies, physiologies, and life histories than more distantly related species (i.e. phylogenetic niche conservatism; Wiens & Graham, 2005. Under this assumption, patterns of phylogenetic clustering (co-occurrence of closely related species) have been interpreted as evidence for environmental filtering on evolutionarily conserved traits, whereas patterns of phylogenetic over-dispersion (co-occurrence of more distantly related species) may be indicative of competitively structured communities (Webb et al., 2002). At a broader scale, the phylogenetic structure of regional species pools provides additional information on historical biogeography and lineage diversification (Davies & Buckley, 2011, 2012) and helps to explain gradients in biodiversity (Wiens & Donoghue, 2004). Even though there is debate about the strength of phylogenetic niche conservatism for free-living species across scales (e.g. Cavender-Bares et al., 2009; Mayfield & Levine, 2010), there is now a large and rapidly growing body of evidence suggesting that phylogenetic conservatism is an important feature of host–parasite systems (e.g. Davies & Pedersen, 2008; Streicker et al., 2010; Poulin, Krasnov & Mouillot, 2011; Huang et al., 2014; Parker et al., 2015; Gilbert & Parker, 2016). Nonetheless, new eco-phylogenetic approaches have been developed that relax the assumption of strict niche.
conservatism (Ives & Helmus, 2011), and there is now a wide range of metrics and models applicable to numerous systems (see Pearse et al., 2014; Tucker et al., 2017). Eco-phylogenetic metrics can be grouped into three broad components of phylogenetic diversity: richness (sum of accumulated phylogenetic difference in a community), divergence (average phylogenetic difference of taxa in a community), and regularity (how regular the phylogenetic differences are in a community) (Tucker et al., 2017). Each phylo-diversity (see Section 1.2e) component can be applied to examine patterns within sets of communities (α diversity) and between sets of communities (β diversity), with each component suited to different types of ecological questions (see Tucker et al., 2017). Such metrics are now commonplace in the plant and animal ecology literature; however, the application of eco-phylogenetic approaches into other fields, such as disease ecology, has lagged behind. Gilbert & Parker (2016) review the current understanding of phylogenetic patterns of disease risk in plants, but a more general synthesis across scales and taxa is needed.

Eco-phylogenetic approaches coupled with population genetic and macroecological approaches such as ecological niche modelling (see Section 1.2d), have the potential to expand insights into host–parasite dynamics across spatial scales. In non-parasite communities, the fusion of phylogenetics and ecological niche modelling has enabled investigations of environmental niche conservatism (Peterson, Soberón & Sánchez-Cordero, 1999; Peterson et al., 2011; Wiens & Graham, 2005; Losos, 2008), speciation in relation to environmental heterogeneity, biogeographic barriers, refugia, and non-analogous climates and communities (Bonaccorso, Koch & Peterson, 2006; Peterson & Nyári, 2008; Serra-Varela et al., 2015). Emergent methods in ecological niche modelling include adding phylogenetic dimensions (e.g. Morales-Castilla et al., 2017) with the goal of understanding the role of environmental and geographic distances in
diversification patterns (Rissler & Apodaca, 2007; Alvarado-Serrano & Knowles, 2014; Lira-Noriega & Manthey, 2014). Phylogenetic niche conservatism additionally suggests potential for reciprocal niche estimation of species that are phylogenetically close, or that have co-evolved (e.g. host and parasite). Eco-phylogenetic tools and ecological niche modelling, however, have rarely been used together to gain insights into ‘infra-communities’ of parasites and symbionts (but see Poulin et al., 2011).

While phylogenetic patterns of hosts and parasites can greatly increase our understanding of infectious disease dynamics (Archie et al., 2009), there are no established frameworks that can unify both components across scales. For example, in some cases host–parasite associations can have a strong phylogenetic structure, such that host phylogeny can predict disease spill-over and prevalence (Streicker et al., 2010; Parker et al., 2015). Experimental inoculations of tropical forest trees with fungal parasites have shown that the fungi have restricted phylogenetic host ranges, such that the likelihood of a parasite infecting two tree species declines with increasing phylogenetic distance (Gilbert & Webb, 2007). For some primates and carnivores, it has been documented that the similarity in some parasite communities among species is predicted well by their hosts’ phylogenetic relationships, whereby close relatives tend to harbour similar suites of parasites (Davies & Pedersen, 2008; Huang et al., 2014). In addition, parasite phylogeny can provide a rich source of information about infectious disease dynamics by providing insights into parasite biogeography and spread, particularly for viruses and bacteria (e.g. Biek et al., 2012).

For example, identifying areas of high pathogen phylo-diversity may be important for understanding novel disease emergence (Barton et al., 2014) and more broadly can be used to infer the evolutionary mechanisms underlying parasite community assembly (Krasnov et al., 2014; Scordato & Kardish, 2014).
Here, we highlight how eco-phylogenetics can play a key role in advancing our knowledge of disease ecology across a hierarchy of scales (Fig. 1). Such advances are urgently needed as emerging and re-emerging infectious diseases are increasing in frequency and distribution, and constitute a major public, animal, and plant health threat (Daszak, Cunningham & Hyatt, 2000; Jones et al., 2008; Tompkins et al., 2015), however our ability to forecast accurately where and when these emergence events will occur is limited (for examples see Jones et al., 2008; Pedersen & Davies, 2009). We suggest, in agreement with Johnson et al. (2015), that the lack of accurate predictions is, in part, because infectious diseases are often considered in isolation, when in reality they are embedded within a complex and interactive community. Importantly, we argue that eco-phylogenetic techniques can help address four linked challenges in infectious disease ecology (Fig. 1), spanning from within-host to landscape scales.

The first challenge to address is to understand co-infection better within hosts. Co-infections introduce at least one additional layer of complexity as parasites can interact with each other directly or indirectly via the immune system, and co-infecting parasites may have inhibitory (e.g. Blackwell et al., 2013) or facilitative effects upon one another (e.g. Syller, 2012). Co-infections are ubiquitous, yet remain understudied (Cattadori, Boag & Hudson, 2008; Susi et al., 2015).

The second challenge concerns parasite gene flow and transmission mode where critical questions include: what are the underlying landscape, environmental, and host determinants of parasite gene flow; and which transmission pathway (e.g. horizontal or vertical) is the most important for the parasite? The third challenge is to incorporate multi-host complexes (i.e. when a parasite infects multiple host species) into our understanding of host–pathogen dynamics. This will be critical for predicting when and where the risks of spill-over are highest, and for extrapolating epidemiologically relevant parameters across multiple host species. It is also
possible that understanding the phylogenetic dimensions of host community assembly could help
gain insights into dilution effects (e.g. Keesing, Holt & Ostfeld, 2006). The final challenge deals
with parasite community structure across the landscape, and is aimed at identifying hotspots and
drivers of parasite diversity, and parasite community assembly rules, which are in part driven by
host community assemblages. Each of these challenges and associated questions are of
fundamental importance to understanding host–pathogen dynamics across scales, yet they remain
largely unaddressed (Mouillot et al., 2005; Scordato & Kardish, 2014; Seabloom et al., 2015).

Here, we show how eco-phylogenetic approaches can contribute to addressing these challenges,
and link scales from individual to landscape in a unified framework.

(2) A guide to eco-phylogenetic methods and complementary tools.

(a) Bayesian phylodynamics

Bayesian phylodynamics methods are commonly employed to understand parasite transmission
dynamics and phylogeography predominantly using coalescent-based estimation and Bayesian
Markov chain Monte Carlo (MCMC) methods. Useful software packages include BEAST
software architecture (Drummond & Rambaut, 2007). In addition, SERAPHIM (Dellicour et al.,
2016a) can analyse how a landscape can alter parasite movement by extracting spatio-temporal
information from phylogenetic reconstructions.

(b) Co-phylogenetics

A set of methods that enable host–parasite co-speciation and host shifts to be inferred by
comparing host and parasite phylogenetic relationships. If both host and parasite phylogenies are
congruent, co-speciation is likely to be the dominant evolutionary process, whereas mismatches
between the two may be evidence for a host shift (i.e. when a parasite population adapts to a new host and speciates). Co-phylogenetic methods can be either ‘event based’ where the number and frequency of co-speciation events is estimated (e.g. if there are more co-speciation events compared with a null model), or ‘topology or distance based’ methods where the symmetry or similarity of the phylogenies is compared (e.g. high similarity may indicate co-speciation events). See de Vienne et al. (2013) for a comprehensive review of co-phylogenetic methods.

(c) Distance-based methods

A very broad array of analyses (often non-parametric) that can be applied to phylogenetic distance or phylogenetic weighted community (dis)similarity matrices and use permutation or Monte-Carlo methods to estimate significance of effects. Includes tests such as permutational ANOVA (PERMANOVA) that tests for differences between groups (Anderson, 2001) and multiple regression approaches such as distance-based linear modelling (DISTLM) and generalised dissimilarity modelling (GDM; Ferrier et al., 2007; Rosauer et al., 2014). These methods are particularly useful if the aim of the study is to understand what factors shape parasite phylogenetic patterns. Useful software packages include: R package ‘vegan’ (Oksanen et al., 2013) and PRIMER PERMANOVA+ (Anderson, Gorley & Clarke, 2008).

(d) Ecological niche modelling

An approximation of species’ ecological niches based on analytical methods that link the geographic presences (or presences and absences) of a species with abiotic constraints (environmental conditions like climate and landscape structure). For example, Maxent is a presence–background correlative method that uses species’ presences and large samples of the
background of the study area (Phillips, Anderson & Schapire, 2006); other methods conceptually
use presence–absence data: General Linear Models (GLM), General Additive Models (GAM),
and Boosted Regression Trees (BRT); these methods are available in the R package
‘dismo’(Hijmans et al., 2017). Other strictly presence-only methods, based on the position of the
presences in environmental dimensions, include NicheA (Qiao et al., 2016), Marble (Qiao et al.,
2015), and hypervolume kernel density estimation (Blonder et al., 2014).

(e) Phylo-diversity
Metrics of diversity that extend species-level diversity (species richness, divergence, and
evenness) to consider phylogenetic non-independence of species: species do not equally
represent all fractions of the tree of life. Common α phylo-diversity metrics, i.e. that measure
phylogenetic richness in a community, include Faith’s Phylo-diversity (PD - the sum of all
branch lengths on a phylogeny; Faith, 1992), whereas Mean Phylogenetic Distance (MPD; the
mean pairwise distance among all species on a phylogeny) and Mean Nearest Taxonomic
Distance (MNTD; the mean nearest-neighbour distance of all species on a phylogeny) are
common phylogenetic divergence metrics. Common β-diversity metrics, comparing phylogenetic
richness between communities, such as UniFrac and PhyloSor (Bryant et al., 2008) are based
around shared branch lengths, and phylogenetic community dissimilarity (PCD; Ives & Helmus,
2010) is notable for decomposing diversity into phylogenetic and non-phylogenetic components.
There are almost as many reviews of phylo-diversity as there are metrics; recent approaches have
grouped metrics according to the kinds of data they require (Pearse et al., 2015) and what kinds
of ecological questions they can answer (Tucker et al., 2017).
Phylogenetic Generalised Linear Mixed Models (PGLMM)

An extension of mixed-effects modelling approaches (Bolker et al., 2009), regressing species (parasite) abundance or presence/absence as a function of environmental conditions and species’ traits (Ives & Helmus, 2011). Random effect parameters can detect phylogenetic patterns in species’ distributions, and can be extended to model both hosts and parasites (see Rafferty & Ives, 2013).

II. CO-INFECTION

Understanding within-host infection patterns is a fundamental challenge for disease ecology (Buhnerkempe et al., 2015). Diverse organisms co-occur and interact within the same host, but within-host parasite communities are not commonly investigated using eco-phylogenetic methods (Cox, 2001; Steinmann & Du, 2010; Seabloom et al., 2015; Rynkiewicz et al., 2015). We define co-infections as multiple infections of unrelated pathogens or between distinct taxonomic units of the same pathogen in the same host. Numerous studies have emphasised the importance of co-infections, showing that host immune response to parasites and ecological interactions between parasites can facilitate or limit infections with other parasites (Cox, 2001; Rynkiewicz et al., 2015), and that the effects can be reciprocal or asymmetrical. For example, in field voles (Microtus agrestis), Telfer et al. (2010) described both positive and negative associations among micro-parasite species (a virus, a protozoan, and two bacteria) and showed that the co-infections were stronger predictors of infection risk than other variables such as exposure risk and host condition. Parasite interactions are commonly explored by assessing ‘top-down’ (host immune system) and ‘bottom-up’ (host resources) control mechanisms (Pedersen & Fenton, 2007) using patterns of exposure/infection or co-occurrence. Here, we suggest that an
eco-phylogenetic framework can expand this approach by incorporating phylogenetic relatedness of the within-host community across scales.

Patterns of phylo-diversity have been particularly important in understanding the mechanisms behind species coexistence of non-parasite communities, and thus have the potential to reveal new insights into the patterns of co-infection. For example, plant communities that have high phylogenetic divergence and regularity were much more stable over time compared to plant communities with low divergence and regularity (e.g. Cadotte, Dinnage & Tilman, 2012).

Conversely, resource stability may lead to high phylo-diversity, as has been shown in a variety of plant and animal systems (e.g. Graham et al., 2009; Whitfeld et al., 2012). Furthermore, insights from invasion ecology demonstrate that invasion success of novel species is less likely in communities with high phylogenetic richness (Davis, Grime & Thompson, 2000; Whitfeld et al., 2014). We argue that the same type of phylo-diversity relationships may apply to co-infected hosts (Fig. 2). To our knowledge, the only study that has used phylo-diversity metrics to examine parasite co-occurrence found that ectoparasitic co-infections were more common between related flea species (phylogenetic clustering) (Krasnov et al., 2014). Experimental work has also shown that more-related parasite strains were more likely to co-infect a plant host (Koskella, Giraud & Hood, 2006), although this effect was not quantified. Whether these finding are generalizable to other systems and other locations in the host is unknown. Analysing phylo-diversity patterns across multiple locations within the host (e.g. the large versus small intestine, Fig. 2) for a cross section of the host population may allow for generalisations for how species that form co-infections assemble. For example, the lower nutrient content of the large intestine may hypothetically lead to co-infections with high parasite relatedness (low divergence) that are similar across a host population or group of host species. The species that constitute the co-
infection could be different but the environmental conditions could lead to the same pattern.

Future empirical work should examine and manipulate the phylogenetic structure of co-infecting parasite communities to determine if such generalities exist, although such experimental manipulations can be difficult to achieve (see Section VI).

What organisms to include in these analyses, and at what taxonomic scale, are important considerations, as eco-phylogenetic inferences are sensitive to both phylogenetic and spatial scale (e.g. Cavender-Bares et al., 2009). For example, including unrelated symbiotic bacteria (the microbiome), as well as nematodes from the two locations in the gut (Fig. 2), or increasing the spatial scale by lumping both intestinal communities together, would likely influence clustering patterns, and thus change the eco-phylogenetic inference drawn. These are general and well-recognised issues in phylogenetic ecology, and whilst there are no simple answers as to which scale is the most appropriate, there are some general guidelines that can help. In essence, the location sampled has to be small enough for individuals to interact, and the taxonomic units related enough that competition is a plausible explanation (the Darwin–Hutchinson zone, see Vamosi et al., 2009 for a more detailed discussion). In poorly understood host co-infection systems, identifying the Darwin–Hutchinson zone may be challenging, but being aware of potential biases associated with scale-sensitivity might not only help guide study design, but also highlight the relative importance of different processes operating across scales (see Cadotte & Davies 2016).

III. TRANSMISSION DYNAMICS AND PATHOGEN GENE FLOW

(1) Eco-phylogenetic approaches
Transmission dynamics among individuals, more broadly, pathogen and parasite gene flow at a landscape scale, are still poorly understood, mostly because data are sparse and difficult to collect, particularly for wild animal populations. Here, we define ‘transmission’ as the process of transmitting a pathogen from one individual to another, ‘spread’ as the geographic dispersal of the pathogen (Lemey et al., 2009), and ‘gene flow’ as the signature of past transmission events across a landscape. Previous studies of transmission in wildlife have relied on intensive field studies (e.g. construction of transmission networks by studying contact networks; Grear, Luong & Hudson, 2013) or have been restricted to purely theoretical explorations (e.g. network simulations and mechanistic modelling; Clauset, Moore & Newman, 2008). Evolutionary approaches are starting to be used to help fill this gap (e.g. for Ebola; Carroll et al., 2015). In particular, due to high rates of evolution, viral genetic data collected from host populations can provide fundamental insights into the transmission dynamics and gene flow of parasites. With greater sequencing effort, however, evolutionary approaches can also be utilised with non-viral agents that evolve at slower rates. For example, high-resolution whole-genome sequencing can reveal transmission dynamics of bacteria (Biek et al., 2012; Kao et al., 2014; Kamath et al., 2016), but this can be prohibitively expensive for large numbers of samples (see Section VI). Integrated population genetic approaches, such as the field of phylodynamics (Grenfell et al., 2004), can link demography and evolution of a parasite, thereby providing important insights into the gene flow of directly transmitted parasites (Biek et al., 2007; Biek & Real, 2010; Lee et al., 2012). This approach, however, is technically challenging (Dellicour et al., 2016b). Eco-phylogenetic approaches are already well developed, but still remain under-utilised in tackling the drivers of parasite gene flow.
Eco-phylogenetic approaches, particularly distance-based methods (see Section 1.2c), could provide insights into two aspects of transmission, specifically: how gene flow is shaped by multiple interacting factors and transmission mode across multiple scales. Approaches commonly employed to examine the environmental and biotic determinants of phylo-β diversity (see Section 1.2c) could be readily applied to parasite patristic distance (i.e. pair-wise sum of branch lengths) or phylogenetic dissimilarity matrices. Machine-learning techniques, such as gradient forests (Ellis, Smith & Pitcher, 2012), also offer possibilities to reconstruct pathogen gene flow by analysing the host genetic data [single nucleotide polymorphism (SNP) matrix] (Fitzpatrick & Keller, 2015). Gradient forests are well suited to analyse how large sets of complex predictor variables shape gene flow as the technique is less sensitive to variable collinearity than more traditional approaches (Ellis et al., 2012). Furthermore, techniques such as gradient forests or GDM can be used to make spatial predictions about pathogen gene flow (Fitzpatrick & Keller, 2015) that can be used to forecast parasite spread in a landscape.

Combined, eco-phylogenetics and phylodynamics have significant scope to extend insights into pathogen gene flow and transmission mode dynamics between hosts more broadly. We envision a landscape phylodynamic approach where Bayesian phylodynamic methods (Section 1.2a) estimate pathogen transmission and spread through time and space (Lemey et al., 2009, 2010) and how landscape factors shape parasite movement (Dellicour et al., 2016a). Eco-phylogenetic approaches could provide evidence related to how landscape and host variables interact to influence pathogen gene flow (Fig. 3) and forecast where areas of gene flow, and thus future transmission events, are likely to occur in the future. For example, in Fig. 3, phylodynamic techniques can be used to derive spatio-temporal estimates for the two likely transmission events (between individuals a–c and c–b) and to show how fast the parasite is spreading in the
landscape (Dellicour et al., 2016b). Eco-phylogenetic techniques could show that an interaction between host relatedness and forest cover best explained gene flow and these variables could be used to predict future gene flow (Fig. 3).

Using similar tool sets but different types of data, eco-phylogenetic analyses can also be used to untangle transmission mode (i.e. how a parasite is transmitted). For example, distance-based regression techniques could be used to assess if parasite phylogenetic distance between known host parent–offspring relationships is less than expected by chance. If geographic distance or the host contact network (Fig. 3) is a more important predictor of parasite phylogenetic distance, as compared to host genetic distance, this could provide evidence for horizontal transmission.

Below we provide a case study where similar techniques were applied to untangle disease transmission dynamics and gene flow in a bobcat (Lynx rufus) population in California. A similar approach could be employed to assess the impact of social networks on disease. For example, for directly transmitted parasites, the impact of social organisation could be assessed by comparing within-group parasite phylogenetic distance to that between groups. Distance-based regression methods could also augment what has been done to understand the social component of *Escherichia coli* transmission in multiple wildlife species (Chiyo et al., 2014; VanderWaal et al., 2014). Bayesian discrete trait analysis (Drummond & Rambaut, 2007) could be further employed to assess how frequently between-group transmission events were likely to have occurred.

(2) Application of an eco-phylogenetic framework to analyse drivers of pathogen transmission

Fountain-Jones et al. (in press) applied a gradient forest approach (Fitzpatrick & Keller, 2015) coupled with Bayesian phylogeography to disentangle the effects of urbanisation on pathogen
gene flow and transmission in a bobcat populations around Los Angeles, California. As part of this study 106 bobcats were sampled and tested for feline immunodeficiency virus (FIV), of these 19 were positive. Positive samples were then sequenced (see Lee et al., 2012), and gradient forest methods were applied to relate host traits (e.g. sex), relatedness (based on microsatellite loci), and landscape features with FIV SNPs, a surrogate for virus gene flow (Fitzpatrick & Keller, 2015). Fig. 4 provides a summary of the data and eco-phylogenetic tools used to answer each linked question.

This analysis revealed that landscape variables that shaped host gene flow, such as grassland and forest, combined with host relatedness, also shaped FIV gene flow. Finer scale Bayesian phylogeographic estimates of transmission event locations confirmed that transmission was more likely in areas with less human activity, even in this heavily urbanized system (Fountain-Jones et al., in press).

Whilst this study provides new insights into transmission of a pathogen in a secretive species’ population, that would have been hard to determine otherwise, it also highlights some of the methodological difficulties that still need to be resolved. For instance, the spatial data used in the models were based on individual capture location which does not necessarily represent individual movement, and this could bias phylogeographic estimates. In addition, non-independence of genetic data arising from linkage disequilibrium can complicate interpretation of analyses using methods such as gradient forest (Fitzpatrick & Keller, 2015). Nonetheless, methodological improvement is occurring rapidly, and will greatly increase the efficacy of these methods in the future.
While there has been increased attention on understanding co-infection, it is now also widely recognised that most parasites are multi-host, i.e. they infect and are transmitted within several host species that may differ in both susceptibility and infectiousness (Woolhouse, Taylor & Haydon, 2001; Cleaveland et al., 2001). One key challenge is the identification of reservoir species hosts that maintain parasite populations, which may then spill over into other host species (Haydon et al., 2002; Streicker, Fenton & Pedersen, 2013; Viana et al., 2015; Johnson et al., 2015; Han et al., 2016). To address this challenge often requires long-term observational, experimental (Johnson et al., 2015), and/or theoretical modelling studies (e.g. O’Hare et al., 2014). The process of pathogen spill-over to a new host is a major concern as most emerging infectious diseases in humans have zoonotic origins (Woolhouse & Gowtage-Sequeria, 2005) including strains of pandemic influenza (Vandegrift et al., 2010), West Nile virus (Kilpatrick et al., 2006) and Ebola (Leroy et al., 2005). Spill-over between species is often not phylogenetically random (Kuiken et al., 2006; Streicker et al., 2010), and thus eco-phylogenetic methods should also provide insights into the processes underlying these spill-over events.

Host–parasite communities can persist over evolutionary periods and thus comparing host and parasite phylogenies directly can help to understand the temporal dimension of pathogen spill-over. One of the most intuitive approaches is direct mapping of the parasite tree onto the host tree or co-phylogenetics (see Section I.2b), with the goal of identifying co-speciation and spill-over events (e.g. Charleston, & Robertson, 2002; Switzer et al., 2005; Brooks & Ferrao, 2005; Ramsden, Holmes & Charleston, 2008; Firth et al., 2009). Co-phylogenetic approaches have been applied to a wide variety of pathogen–host systems and have revealed important insights into host–pathogen co-evolution at a macroevolutionary scale. For example, co-phylogenetic
analysis of hantavirus–host relationships revealed little evidence for co-speciation with rodent
and insectivore hosts and it is likely that recent spill-over events have shaped hantavirus
evolution (Ramsden et al., 2008). By contrast, co-speciation was found to be important for
Simian foamy viruses in primates with highly congruent branching orders and divergence times
(Switzer et al., 2005). Overall, however, spill-over could be the dominant force shaping host–
pathogen evolution (see de Vienne et al., 2013 for a review of the literature). Co-phylogeny
alone, however, cannot help predict spill-over events or provide direct insight into the
mechanisms driving host shifts and spill-over (Penczykowski, Laine & Koskella, 2016).
Differences in parasite host breadth may provide information on the mechanisms underlying
parasite host shifts and spill-over. The probability of cross-species transmission may be largely
determined by host contact rates, but the outcome of such events can vary widely, from single
infection spill-over events (e.g. West Nile Virus in humans; Lloyd-Smith et al., 2009), to
transient outbreaks (Wölfe, Dunavan & Diamond, 2007; Lloyd-Smith et al., 2009), to sustained
outbreaks in the new host (e.g. rabies; Streicker et al., 2010). The taxonomic host breadth of a
parasite is likely a product of attributes intrinsic to both parasite and host biology. Key parasite
traits include transmission mode, genetic diversity, and evolutionary rates (e.g. high genetic
diversity or rapid evolutionary rates; Fenton & Pedersen, 2005, but see Streicker et al., 2010).
Whether a parasite is transmitted via a vector, through the environment, or by close contact
might also impact host breadth, although there is likely an interaction between parasite type and
transmission mode (Fenton & Pedersen, 2005). On the other hand, a number of empirical and
theoretical studies have shown that certain host species have specific behavioural, physiological,
and genetic traits that allow the parasite to spread rapidly and persist within communities
(Cleaveland et al., 2001; Kilpatrick et al., 2006). However, measurements of such traits are often
lacking (see Section VI). Phylogenetic structure of host communities can not only provide a surrogate for traits, but also provide insights into the evolutionary mechanisms shaping parasite–host interactions.

Host specificity, or the extent to which a parasite can exploit different hosts (Poulin et al., 2011), is one of the few areas where eco-phylogenetic metrics have been employed to understand host–parasite systems. Studies incorporating host specificity have provided novel insights into primate–parasite co-evolution; for example, the majority of parasites infecting primates were phylogenetic generalists that infected more distantly related primate hosts than was expected by chance (Cooper et al., 2012). By contrast, for Drosophila sigma viruses, host specificity was predicted by host phylogeny, with host switching only between related hosts (phylogenetic clustering, low divergence; Longdon et al., 2011). In both cases, however, taxonomic richness (i.e. number of hosts a parasite can exploit) would not have provided resolution of the evolutionary drivers of these multi-host complexes. By gaining a better understanding of host specificity and differences in host-specific virulence, we can then generate more accurate models of parasite persistence within multi-host communities (Poulin et al., 2011; Fenton et al., 2015). The phylo-diversity of the pathogen can also help identify the direction of spill-over events from previously collected sequence data, information that may be particularly valuable when sample size of pathogens is low and nucleotide substitutions rates are variable. If sample sizes across multiple hosts are comparable, Bayesian discrete trait methods (Drummond & Rambaut, 2007) can accurately trace multi-host dynamics (e.g. Kamath et al., 2016). However, if sample sizes are uneven, these methods are likely to assign the species with the most samples as the reservoir host in a particular system (Beerli, 2004; Viana et al., 2014). Standardised phylo-diversity indices may be helpful in identifying host reservoirs in complex communities. It is likely that a parasite
will have a greater evolutionary history in the adapted (reservoir) host, therefore reservoir species may have much higher pathogen phylo-diversity (greater phylogenetic dispersion) compared to spill-over hosts (non-reservoir hosts receiving spill-over pathogens from a reservoir host), where phylogenetic clustering is more likely (for an example where badgers (Meles meles) are the reservoir of bovine tuberculosis, see Fig. 5). As measures like MPD and MNTD (see Section I.2e) are standardized by number of samples (i.e. they compare the observed phylo-diversity to the diversity expected by chance for a given sample size), they may offer a useful solution for unbalanced data sets, though empirical testing is required.

At a host community level, the phylogenetic structure and diversity of hosts may provide valuable insights into pathogen–host diversity relationships, although the phylogenetic dimensions of diversity have rarely been considered in this context (Gilbert & Parker, 2016). The dilution effect hypothesis, for example, is based on the premise that a host community with high α diversity dilutes individual infection risk by providing a greater density of incompetent hosts (Ostfeld & Keesing, 2000; Keesing et al., 2006); thus biodiversity loss (or community disassembly) may increase disease prevalence assuming that the most competent hosts remain in the community. Even though recent work has shown some support for the dilution effect (e.g., Civitello et al., 2015), it is not clear how generalisable it is and there are documented cases illustrating that diversity can increase (amplification effect) disease risk (e.g. Mitchell, Tilman & Groth, 2002). Host relatedness is likely to play a role in multi-host systems (Parker et al., 2015), and thus eco-phylogenetic metrics may help elucidate the mechanisms behind the dilution and amplification effects (for a hypothetical example of avian influenza, see Fig. 6; Gilbert & Parker, 2016). For example, two communities may have identical species richness, but high phylogenetic clustering of hosts in one community may lead to amplification, as related hosts might be more
likely to be competent vectors and may be more likely to come into contact with each other (Fig. 6, Scenario 1). The opposite could be true for a phylogenetically dispersed host community as prevalence could be diminished via the dilution effect (Fig. 6, Scenario 2); and, of course, transmission between distant hosts would be less likely if the parasites had co-evolved with their hosts. Even though these ideas have yet to be fully tested on empirical data, it seems likely that eco-phylogenetic patterns can provide an important tool set to understand the mechanisms of biodiversity–disease relationships.

V. PARASITE COMMUNITY STRUCTURE

Although parasites are non-randomly distributed across landscapes and host species, we have a poor understanding of the drivers of parasite diversity across space and hosts. Of the various aspects of diversity, phylo-diversity is often linked with ecosystem performance (e.g. Srivastava et al., 2012), implying that hotspots of phylo-diversity might be associated with disease outbreaks. For example, locations harbouring genetically diverse parasites (measured by Simpson’s Index, or other measures) are linked to greater risk of disease emergence, particularly for viruses, due to the increased likelihood of evolutionary innovations (Barton et al., 2014).

Because phylo-diversity might capture features of diversity better than simple measures of species richness (Faith, 1992), we might also predict that phylo-diversity would be a better predictor of disease emergence via evolutionary innovation. If a parasite community has high phylogenetic divergence we hypothesise that the possible reassortment/recombination event may be novel. However, if a viral community, for example, is both highly divergent and irregular, it is possible that biological limitations could be imposed on recombination (Pérez-Losada et al., 2015). Eco-phylogenetic metrics and models, such as the PGLMM (see Section I.2f), allow for
the estimation of where highly phylogenetically divergent parasite communities could be, and
what particular set of host or environmental variables underlie these diversity patterns. However,
eco-phylogenetic methods are currently reliant on the ability to detect species consistently, which
can often be a large challenge for parasites, particularly when a high proportion remains
undiscovered (see Section VI).

Phylo-β diversity (variation in phylogenetic composition of a parasite community) among sites
or host populations in a geographic area (Legendre, Borcard & Peres-Neto, 2005; Graham &
Fine, 2008) can also reveal insights into how abiotic and biotic gradients (biotic interactions
among hosts and within hosts) shape parasite communities. Patterns of taxonomic β diversity
have helped untangle the role of geographic distance versus abiotic niche in structuring parasite
communities across the world (Poulin & Mouillot, 2003; Krasnov et al., 2005; Timi, Lanfranchi,
& Luque, 2010; Scordato & Kardish, 2014; Warburton, Kohler & Vonhof, 2016). For example,
helminth taxonomic β diversity in big brown bat (Eptesicus fuscus) populations was better
predicted by abiotic variables compared to distance between bat populations or host variables
(sex and age), likely because bat populations that experience the same abiotic variables were
more likely to be exposed to the same helminth communities (Warburton et al., 2016).

Incorporating measures of phylo-β diversity might provide additional information on drivers of
parasite–host interactions (Ives & Helmus, 2010). Phylo-β diversity approaches have led to
major advances in our understanding of the microbial, plant and animal communities (e.g. Ley et
al., 2006; Wang et al., 2013; Burbrink et al., 2015; Herrera, 2016). For example, phylo-β
diversity of soil microbial communities can be highly deterministic, with different habitats
harbouring phylogenetically related microbe communities, indicating strong environmental
selection of all microbe groups (Wang et al., 2013). Indeed, obesity has been found to be
associated with phylogenetically distinct microbe communities (Ley et al., 2006). Phylo-β
diversity metrics are well developed in the non-parasite literature (Bryant et al., 2008), but have
rarely been used for exploring host–parasite systems (Scordato & Kardish, 2014). Model-based
approaches such as PGLMM (see Section I.2) statistically model the likelihood of co-occurrence
as a function of species’ traits, environment, and their interactions, while simultaneously
accounting for and measuring phylogenetic independence. By incorporating phylogenetic
distance into their fitting procedure, PGLMMs can inform predictions for poorly studied (e.g.
emerging) diseases, assuming that data on close relatives are available. PGLMMs have already
been applied to interaction network data (Rafferty & Ives, 2013), thus their use in MH–MP
systems would seem straightforward.

To illustrate how one could apply these methods, we provide a hypothetical application of an
eco-phylogenetic framework to understand multi-host/multi-pathogen dynamics at a macroscale.
For important zoonotic parasites, such as avian influenza or rabies, there are large-scale
molecular sequence databases that can be exploited to help answer some of the challenges
proposed herein using eco-phylogenetic methods. For example, the Influenza Research Database
(Zhang et al., 2017) has hundreds of thousands of avian influenza sequences from hundreds of
subtypes from across the globe, often from bird communities that are sampled intensively in one
location over a period of years. These data have already been used to test and generate
hypotheses about the role of taxonomic α diversity in the formation of novel influenza strains
(Barton et al., 2014). Eco-phylogenetic models and metrics can expand this research further, and
we both α and phylo-β diversity metrics and eco-phylogenetic models can be used to predict
diversity hotspots and to understand the environmental and evolutionary dimensions of parasite
community coexistence.
Phylogenetic reconstruction of parasite sequence data (e.g. using BEAST) can be used to create molecular operational taxonomic units (OTUs) based on a model-derived threshold value (e.g. Zhang et al., 2013). The geographic data attached to each OTU can be used to estimate the environmental niche, and the niche estimates can be overlaid to assess areas where there could be a high degree of overlap (Fig. 7). To test if areas with high niche overlap are also diversity hotspots, phylogenetic divergence can be calculated for the community of OTUs at each location using the MPD metric (weighted by OTU abundance at each location) $MPD = \frac{\sum_{i,j \neq j} \delta_{i,j}}{n(n-1)}$, where $\delta_{i,j}$ is the pairwise distance between OTU $i$ and $j$ and $n$ is the number of OTUs. As host sampling is likely to be uneven across locations (different numbers of samples at each location), this metric, and all other phylo-diversity metrics, is standardised by comparing the observed community MPD with a value generated from a null model generated by a randomisation method (SES$MPD$) (see Webb et al., 2002) to generate Standardized Effect Sizes (SES): $SES_{MPD} = \frac{MPD_{observed} - mean(MPD_{null})}{sd(MPD_{null})}$. Positive SES$MPD$ values with high quantiles (≥0.95) is evidence for phylogenetic dispersion, whereas negative SES$MPD$ values with low quantiles (≤ 0.05) is evidence for phylogenetic clustering (Kembel et al., 2010). These values can be compared against the degree of niche overlap and then regressed against landscape or host variables (e.g. host community MPD) to provide insights into drivers of these patterns. Furthermore, they can provide landscape-level insight into the mechanisms generating co-infection patterns (Fig. 7).

Phylo-β diversity models can be used to make inference about what drives OTU occurrence between sites and can provide valuable insights into the degree to which phylogenetically related OTUs co-occur. Phylogenetic generalized linear mixed models (PGLMMs) (Box 1) is one method to do this and can be used to predict where OTUs could occur (Ives & Helmus, 2011). PGLMMs can be formulated in different ways, but one model that could be used to untangle the...
phylogenetic and environmental drivers of OTU coexistence through a series of logistic regressions (Ives & Helmus, 2011) is:

\[
Pr(Y_i = 1) = \mu_i
\]

\[
\mu_i = \logit^{-1}(a_{sp[i]} + b_{sp[i]}x_{site[i]} + c_{site[i]})
\]

\[
b_{sp[i]} = B + e_{slope[i]} + g_{phyto[i]}
\]

\[
e\text{Gaussian}(0, \sigma_{slope}^2 I_n)
\]

\[
g\text{Gaussian}(0, \sigma_{phyto}^2 \sum sp)
\]

\[
c\text{Gaussian}(0, \sigma_{site}^2 I_n)
\]

where \(Y_i\) is the presence (1) or absence (0) of the OTU, and the probabilities \(\mu_i\) are random effects with the distribution \(\logit(\mu_i)\), \(a_{sp[i]}\) is a categorical variable for each species, \(I\) is a identity matrix. \(\logit(\mu_i)\) in this model is assumed to be linearly dependent on \(x\) (an environmental or host variable) through the coefficient \(b_{sp[i]}\). The covariance matrices \(e, g,\) and \(c\) account for species-specific differences independent of phylogeny, phylogenetic relatedness (assuming a Brownian or some other model of evolution), and site-specific differences in community structure, respectively (Ives & Helmus, 2011).

The abiotic and biotic drivers of co-infection and parasite community assembly can thus be investigated through a combination of phylogenetic metric and model approaches, and MH–MP ecological niche modelling (Fig. 7, Section I.2d). Ecological niche modelling could identify where co-infections could occur and phylo-diversity indices could be used to explore the pattern of likely and realized co-infections (Fig. 7). If the ecological niche dimensions of parasites are governed by phylogenetic relatedness (Peterson et al., 1999), phylogenetic placement of parasites may be important for predicting parasite occurrence and persistence. For example, if two subtypes of influenza have phylogenetically conserved niches, they could be found in co-
infections more often than would be expected by chance. The phylogenetic structure of
coeexisting parasites can also reveal insights into the mechanisms underlying co-infection
patterns. Significant phylogenetic clustering of co-infecting parasites (or genotypes of the same
pathogen) in a host or population may be evidence for environmental filtering (i.e. environmental
conditions restrictive to other species; Webb et al., 2002, Fig. 7). Environmental filtering of
parasites could be due to external abiotic limits, or within-host immunological responses filtering
less-related pathogens. If co-infecting parasites are phylogenetically dispersed (e.g. co-infection
with distantly related parasites), this could be evidence for biotic interactions such as resource
competition or facilitation (e.g. Webb et al., 2002, Fig. 7). Such process-based interpretations
are, however, predicated on assumptions of niche conservatism (see Section I), and other
processes might give rise to similar patterns (see Mayfield & Levine, 2010).
An emerging class of eco-phylogenetic methods are exploring the possibility of quantifying the
processes that generate diversity (Pearse et al., 2015). These methods, while still in their infancy,
have the potential to help predict the kinds of environmental conditions and ecological processes
that will lead to novel parasite diversity. For example, the DAMOCLES approach places
observed phylo-diversity within the context of null hypotheses generated from speciation and
extinction patterns (Pigot & Etienne, 2015). This approach may allow us to determine whether
parasite diversity is being created de novo within a system, or whether it is moving from one
local patch to another. Such information has consequences for predicting the relevance of
vaccine development versus spatial isolation, and could be powerful in linking short-term and
longer-term management interventions in systems where diseases are endemic.
VI. CHALLENGES

(1) Data

Data challenges represent one of the most limiting factors in broad-scale application of eco-phylogenetic techniques to disease ecology. Despite the decreasing cost of sequencing and advances in technology, most wildlife disease studies still rely on serological testing (i.e. testing for antibodies for micro-parasites) or presence/absence or count data (for macro-parasites) rather than parasite isolation and genetic sequencing to investigate infection. Because current phylogenetic resolution is insufficient for most parasites, there are few parasite groups where detection can be linked to known phylogenetic relationships. Detection and sequencing can be complicated, particularly for micro-parasites; viruses, unlike bacteria and fungi, do not share a conserved gene or genes across all members and thereby designing specific probes can be difficult, especially for those taxa where few sequence data are available. Furthermore, identification of a specific pathogen can, for example, rely on successful culturing which is variable among lineages and requires significant laboratory protocols to be developed (Mohiuddin & Schellhorn, 2015). Conventional PCR (polymerase chain reaction) methods are also often parasite specific, therefore co-infections may be missed because they are not amplified (Shi et al., 2016; Simmonds et al., 2017). This problem is compounded as parasite diversity may be very large: for mammals alone there are an estimated ~320,000 undiscovered viruses (Anthony et al., 2013), >11,500 helminth species (Dobson et al., 2008), and likely vast numbers of undiscovered bacterial parasites. However, new molecular methods provide a promising platform for identifying greater parasite diversity. For example, advances in high-throughput metagenomic techniques, such as environmental DNA (eDNA; Venter et al., 2004) and virus-like particle (VLP) fraction (Melcher et al., 2008; Mohiuddin & Schellhorn, 2015), can enable
rapid inventory of a broad range of parasites from within-host and environmental samples (Shi et al., 2016; Simmonds et al., 2017). Results from this approach can be used for distinguishing OTUs based on a threshold of nucleotide sequence identity and for phylogenetic reconstruction.

(2) Pattern, process, and functional characterisation

Obtaining suitable data is, however, just the first challenge. A consideration only rarely dealt with by eco-phylogenetic studies is how to account for phylogenetic uncertainty. Most of the methods discussed herein do not account explicitly for phylogenetic uncertainty, even though differences in tree topology and branch lengths could impact measures of phylo-diversity (Cadotte et al., 2010). Population genetic approaches, such as BEAST, incorporate phylogenetic uncertainty directly by sampling trees from a posterior distribution, and similar strategies should be implemented for eco-phylogenetic methods. It is straightforward (if time-consuming), however, to incorporate phylogenetic uncertainty if a posterior-predictive approach is used, fitting models and calculating indices across a set of plausible candidate phylogenies (in a Bayesian setting, a posterior distribution of trees) and reporting estimates across that set (Bollback, 2005).

How to link eco-phylogenetic patterns to processes in infectious disease communities is a major conceptual challenge. Community ecologists often infer the mechanisms structuring a community based on observed patterns, but the utility of doing so is fiercely debated (see Gotelli & McCabe, 2002; Gotelli & Ulrich, 2012; Connor, Collins & Simberloff, 2013). This is particularly the case within the field of eco-phylogenetics where the processes that can be inferred from phylogenetic clustering and dispersion have recently been questioned based on modern coexistence theory (Mayfield & Levine, 2010; Kraft, Godoy & Levine, 2015; Gerhold et
Throughout this review, we have emphasised the ways in which phylo-diversity metrics and models can contribute to our understanding of eco-evolutionary processes, moving beyond the use of phylogenetic structure as a (controversial) proxy for ecological processes. However, we stress that, within a public health perspective, any metric or approach that can improve predictive power is valuable, irrespective of its contributions towards the separate goal of trying to understand the processes taking place. While it is possible to identify general trends, the combination of processes that shape any one parasite community are likely to be unique due to the interaction with the host immune system and host mobility and differences between parasite types (e.g. macro- versus micro-parasites).

One way to disentangle process from pattern is to conduct controlled manipulative trials, for example, co-infecting a host or environment with either phylogenetically related or unrelated parasites (or knocking out components of the immune system) and then assessing changes in parasite load and host health. Gilbert & Webb (2007) provide one example, in this instance inoculating related host trees with a matching suite of (fungal) parasites. Host–parasite systems provide an ideal experimental system to assess the phylogenetic patterns of community assembly, because they represent rapidly evolving ‘island’ systems with diverse assemblages that can be easily manipulated (Mouillot et al., 2005). However, experiments can be unfeasible in some systems due to ethical or practical considerations (i.e. most parasite treatment methods are broad spectrum), but even if imperfect, they can be informative (e.g. Ferrari et al., 2009; Knowles et al., 2013; Pedersen & Antonovics, 2013). More broadly, eco-phylogenetic data coupled with information on functional traits provide another way to disentangle process from pattern (Baraloto et al., 2012). For example, morphological traits of the Dactylogyrus parasite species reaffirmed eco-phylogenetic trends indicating parasite clustering (Mouillot et al., 2005).
Functional traits can be defined as morphological, biochemical, physiological, and phenological characteristics of organisms that are linked to fitness (Violle et al., 2007). However, as with phylogenetic data, functional trait data for parasites are rare and mostly restricted to macro-parasites (Morand, 1996; Mouillot et al., 2005). Developing techniques to record standardised functional trait data for all parasites remains a fundamental challenge for parasite community ecology, particularly for micro-parasites.

VII. FUTURE DIRECTIONS

We have illustrated how eco-phylogenetic methods can be used to understand how host and parasite communities shape infectious disease dynamics across scales. Such information will be invaluable for making predictions on how human activities might alter infectious disease dynamics. Anthropogenic forces are fundamentally altering host–parasite systems (e.g. Gottdenker et al., 2014), and the frequent movement of humans and animals across the globe increases the possibility for hosts and parasites to be introduced into new regions (Daszak et al., 2000). The phylogenetic context of these introductions is likely to be important in determining which parasites and hosts successfully colonise. If an invader is distantly related to native species, invasion success may be greater due to less competition (e.g. Park & Potter, 2013), but predictions regarding parasite spread are not straightforward. Virulence might be lower because pathogens are less able to infect the native host species, or virulence might be much higher because native hosts are more susceptible (Lymbery et al., 2014). Similarly, climate change and land use play key roles in modifying the geographical distribution of host and vector species (Altizer et al., 2013; Estrada-Peña & de la Fuente García, 2014; Gottdenker et al., 2014), shifting the phylogenetic structure of both host and pathogen communities (Medeiros, Hamer & Ricklefs,
2013). Therefore, it is important to understand how phylogenetic community structure influences pathogen transmission, for example, whether a shift towards more phylogenetically clustered host communities will result in an increase in the transmission of their associated infectious agents.

Human, domestic, and wild animal interactions are becoming more frequent, and landscapes are becoming more urbanised and fragmented, creating a greater potential for parasite spill-over, spillback, and host switching (Suzan et al., 2015). Mixing of parasites among common and novel hosts may allow infectious diseases to persist in the environment for longer periods, and may increase the chance of strains recombining into new, potentially more virulent forms (Engering, Hogerwerf & Slingenbergh, 2013). In addition, to predict disease emergence better, there is a need to understand how parasites diversify and circulate in wild animal populations and to evaluate how the control of one parasite strain influences the prevalence and virulence of co-infecting strains and parasites (Barclay et al., 2014). Eco-phylogenetic techniques that start to untangle the processes that generate diversity (such as DAMOCLES; Pigot & Etienne, 2015) are likely to be particularly useful in this endeavour.

Eco-phylogenetic methods can play an important role in enhancing wildlife and human health-surveillance systems (Farrell et al., 2013). Research initiatives such as the Global Virome Project (http://www.globalviromeproject.org/) will provide a large volume of data that will improve our understanding about viral phylogenetic relationships, and linking this information to landscape and hosts over a variety of scales will enable a better understanding of pathogen dynamics. For example, identifying landscape and host variables that correlate with parasite phylo-diversity hotspots, where spill-over or co-infection are likely, may help to identify susceptible populations, contributing to the planning of risk-mitigation actions. Moreover, identification of phylo-
diversity hotspots can be used to predict the probability of parasite translocation, establishment, and magnitude of health hazards, essential for risk-based analysis and resource prioritisation.

VIII. CONCLUSIONS

(1) Eco-phylogenetic approaches have revolutionised our understanding of free-living communities, yet formal application of this framework to disease ecology and epidemiology has lagged behind. Studies that have applied eco-phylogenetic techniques to understand host–parasite systems have led to advances in our understanding of pathogen spill-over and the evolutionary relationships between parasites and hosts in particular.

(2) An eco-phylogenetic framework coupled with Bayesian phylodynamics and ecological niche modelling can help answer questions central to disease ecology from within-host to landscape scales. This unified framework can help us to understand the antagonistic and synergistic relationships between hosts and parasites that may enable better prediction of disease threats.

(3) We suggest that experimental studies coupled with increased parasite genomic sampling within individuals and among populations and communities will help develop the conceptual foundations of this nascent field. Techniques that help characterise functional traits across parasites will assist in understanding the processes underpinning complex relationships between hosts and parasites, particularly in a world with increased anthropogenic change.

(4) As our knowledge of global diversity improves, and in particular with the expansion of the multi-host multi-pathogen framework, eco-phylogenetic methods are likely to become an increasingly important component of infectious disease research.
IX. ACKNOWLEDGEMENTS

We thank the Minnesota Institute on the Environment Mini-Grant that sponsored the workshop that formed the basis of this paper. N.M.F.-J. and M.E.C. were funded by National Science Foundation (DEB-1413925); M.E.C. was funded by the University of Minnesota’s Office of the Vice President for Research and Academic Health Center Seed Grant. K.V. was funded by NSF grant 1619072 Ecology and Evolution of Infectious Diseases program.

X. REFERENCES


intestinal parasites in humans: the importance of co-infection for infection risk and recovery dynamics. *Proceedings of the Royal Society of London B: Biological Sciences* **280**.


behavior and parasite life cycle make social networks meaningful in disease ecology.


LERoy, E.M., KUMULUNGUI, B., POURRUT, X., ROUQuest, P., HASSANIN, A., YABA, P., DÉLICat,


Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences.


TUCKER, C.M., CADOTTE, M.W., CARVALHO, S.B., DAVIES, T.J., FERRIER, S., FRITZ, S.A.,
GREENYER, R., HELMUS, M.R., JIN, L.S., MOOERS, A.O., PAVoine, S., PURSCHKE, O.,
metrics for conservation, community ecology and macroecology. *Biological Reviews 92*,
658–715.

592.

113–128.


multiparasitism: examining the consequences of co-infections for human and animal health.
*Parasites & Vectors 8*, 545.

VENTER, J.C., REMINGTON, K., HEIDELBERG, J.F., HALPERN, A.L., RUSCH, D., EISEN, J.A., WU,
D., PAULSEN, I., NELSON, K.E., NELSON, W., FOUTS, D.E., LEVY, S., Knap, A.H., LOMAS,
M.W., NEALSON, K. ET AL. (2004). Environmental genome shotgun sequencing of the

VIANA, M., CLEAVELAND, S., MATTHIOPoulos, J., HALLIDAY, J., PACKER, C., CRAFT, M.E.,
HAMPSON, K., CZUPRYNA, A., DOBSON, A.P., DUBOVI, E.J., ERNEST, E., FYUMAGAWA, R.,


FIGURE LEGENDS

Fig. 1. Conceptual schema illustrating how an eco-phylogenetic framework can be applied to understand infectious disease dynamics. The example system used is the Ngoronogoro Crater (Tanzania), across scales: (A) within host; (B) among hosts of the same species; (C) multi-host complex; and (D) landscape scale. Colour-coded and lettered symbols below each panel indicate what data (squares) and statistical tools (circles) could be used to address each challenge (see Section I.2 for model and other tool details). White ovals contain hypothetical parasite communities within a host and different parasite colours and shapes (nematodes or viruses) represent different parasite species or genotypes. PGLMM: phylogenetic generalised linear mixed model.

Fig. 2. Schematic diagram showing three differing hypothetical phylo-diversity patterns in co-infection in three locations in the human gastro-intestinal tract. (a) Co-infection of the oral cavity involving three viruses, each from phylogenetically distinct clades indicating that these viruses may have a relatively long evolutionary relationship with the host (Thézé et al., 2015). We hypothesize that the viral community in this case is likely to be stable due to competition for resources as there is high phylogenetic divergence and regularity. New viral invasions might be less likely. (b) Highly phylogenetically divergent but irregular co-infection in the small intestine involving four nematode species. Even though there is high phylogenetic divergence, the low phylogenetic regularity of this co-infection may indicate that this co-infection is not stable, yet infection by other nematode species may be less likely due to resource competition. (c) Phylogenetically clustered (low divergence and regularity) co-infection in the large intestine,
which could be evidence for environmental filtering (e.g. lower nutrient levels in this organ), or
unoccupied niche space. If the clustering represents unoccupied niche space, invasion by
phylogenetically distinct species may be more likely.

**Fig. 3.** A theoretical example of how eco-phylogenetic techniques combined with Bayesian
phylogeography (BEAST) could be used to analyse pathogen transmission dynamics in wildlife
based on pathogen phylogenetic relationships. (A) An example of how both eco-phylogenetic
and phylodynamic methods could be used to incorporate spatio-temporal data, landscape
variables, host relatedness, host variables, and contact network data to understand pathogen
transmission dynamics between pumas (*Puma concolor*) (individual pumas labelled a to e).
Statistical techniques are italicised (see Section I.2 for method details). Stars represent an
estimated transmission event [within a period defined by high probability density (HPD)
intervals]. (B) Inferences on transmission dynamics of the pathogen in the *P. concolor*
population. In this example, two transmission events were based on phylogeographic estimates (a
to c for 2004–2007 and c to b in 2007 (both based on 95% HPD estimates). Circles around the
stars in the landscape indicate 95% HPD interval for the location of each transmission event in
the landscape. Dotted line connecting puma d and e indicates that there was pathogen gene flow
in the past but transmission directly between these individuals may not have occurred due to
uncertainty in temporal estimates (i.e. it is possible that other individuals were intermediaries in
the transmission chain). Overall, based on the multivariate distance-based methods, forest cover
and host relatedness were the best predictors of gene flow in this hypothetical scenario.
Fig. 4. Summary of data and eco-phylogenetic techniques used to understand two interlinked questions. (A) Feline immunodeficiency virus (FIV) single nucleotide polymorphisms (SNPs) were predicted by bobcat relatedness, and demographic and landscape variables. This was achieved using gradient forest regression models in order to assess how these factors shape gene flow. See Section III.1 for information about gradient forests multiple regression. The box on the far right shows the urban landscape south west of Los Angeles, California with grey shading representing degree of urban development (% impervious surface). The black line indicates the position of a highway that restricts host gene flow in this system (see Lee et al., 2012). (B) FIV phylogeography using bobcat spatial location and sampling dates as well as demographic factors. This was achieved using BEAST and SERAPHIM in order to assess how, where and when FIV was transmitted. See Section I.2a for information on BEAST phylogeography and SERAPHIM. Coloured symbols reflect different individual bobcats.

Fig. 5. Schematic diagram illustrating a hypothetical scenario where standardized phylo-diversity measures could be used to help infer reservoir species and the direction of spill-over events for a parasite data set with unbalanced sampling. The hypothetical phylogenetic trees represent the relationships between bovine tuberculosis (bTB) isolates from badgers, ferrets, and a cattle herd, where more badgers were sampled than other host species. Phylogenetic tree (A) is a scenario where both cattle and badger samples have high values of phylogenetic richness, so there is no evidence for either species being the reservoir (bottom panel). By contrast, phylogenetic tree (B) illustrates a scenario where cattle samples had a much lower relative phylogenetic richness compared to the badger samples, indicating that badgers may be the reservoir host (bottom
panel). In both scenarios ferrets are unlikely to be the reservoir as they have low phylogenetic richness in both trees.

**Fig. 6.** The potential for an influenza strain to persist in two hypothetical avian communities that differ in the degree of host relatedness: (A) Phylogenetic tree of five avian orders. (B) Two host communities (Scenario 1 and Scenario 2) that differ in the degree of host phylogenetic divergence, and how this might influence pathogen persistence and prevalence. Host species richness is the same in both scenarios, however, host species are more closely related in Scenario 1 than in Scenario 2. The low divergence (phylogenetic clustering) in Scenario 1 may increase influenza prevalence over time in this community as contacts between related species (and thus between competent hosts) might happen more frequently than between distantly related species. If there are strong co-evolutionary relationship(s) between a particular host and parasite, transmission of that parasite to unrelated hosts may also be less likely. Conversely, in Scenario 2 high host divergence may reduce prevalence as contacts between related individual hosts would occur less often.

**Fig. 7.** Eco-phylogenetic and ecological niche modelling approaches can be used to understand co-infection patterns of parasite genotypes at a macro-scale for two hypothetical bird species in South America. Data for this approach could be derived, for example, from databases such as the Influenza Research Database (Zhang et al., 2017). In this case, the phylogeny (A) is constrained by the ecological niche of the parasite genotypes (B) (i.e. niche conservatism). Coloured ovals in B represent the ecological niche of parasites in environmental dimensions; those ecological niches are then projected onto geographic space to generate parasite genotype geographic
distributions (C). Dashed lines denote the distribution of each host bird. For the purposes of this example, populations of bird species 1 were found to be co-infected by three parasite genotypes (a, b and c), and these genotypes represented one clade in the phylogeny (A), indicating phylogenetic clustering, which might suggest that environmental filtering is important for this co-infection pattern. Populations of species 2 were found to be infected by three genotypes (j, f & o), and in contrast to Species 1, these genotypes were all from distinct clades (A), showing phylogenetic dispersion. This may indicate that biotic interactions or possibly allopatric speciation history (Pigot & Etienne, 2015) are important for this co-infection scenario.
(A) Co-infection

(B) Transmission pathways

(C) Multi-host complex

(D) Pathogen community structure

<table>
<thead>
<tr>
<th>Data</th>
<th>Phylo-diversity components</th>
<th>Models and other tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>Parasite phylogeny</td>
<td>PGLMM</td>
</tr>
<tr>
<td>HP</td>
<td>Host phylogeny</td>
<td>Distance-based regression methods</td>
</tr>
<tr>
<td>HT</td>
<td>Host traits</td>
<td>Bayesian phylodynamics</td>
</tr>
<tr>
<td>LV</td>
<td>Landscape variables</td>
<td>Ecological niche modelling</td>
</tr>
<tr>
<td>Phylogenetic divergence</td>
<td>(a)</td>
<td>(b)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phylogenetic regularity</th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>
(A) Data and analysis

Pathogen phylogeny

Spatio-temporal data

Individual

e.g. sample dates

Landscape variables

e.g. distance from landscape feature

Host relatedness

Host variables

e.g. host age, body mass.

Host contact network

Predictors used for multivariate distance-based methods

BEAST phylodynamic methods

e.g. pathogen patristic distance

2005

2010

(B) Transmission dynamics

Landscape

Best gene flow predictors: e.g., forest (green) together with host relatedness

2004–7

2000–2007
**Data (A)**

- **FIV SNPs**
  - Host relatedness
  - Gradient forest multiple regression

- **Host variables**
  - Host age and sex

- **Landscape and spatial data**
  - E.g. distance from urban edge

**Data (B)**

- **FIV phylogeny**
  - 2005-2010

- **Spatio-temporal data**
  - Sample dates and locations
  - Host age and sex

**Eco-phylogenetic techniques**

- BEAST phylogeography using continuous traits (latitude and longitude of sampled individual)
  - SERAPHIM

**Question**

What landscape and host factors shape FIV gene flow?

How, where and when of FIV transmission?
### Standardized phylogenetic richness values

<table>
<thead>
<tr>
<th></th>
<th>Badger</th>
<th>Ferret</th>
<th>Cattle herd</th>
<th>Inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Little evidence that badgers (or cattle) are the reservoir</td>
</tr>
<tr>
<td>(B)</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Evidence that badgers are the reservoir</td>
</tr>
</tbody>
</table>
(A) Parasite phylogenetic relationships

(B) Ecological niche of each parasite genotype

(C) Geographic distribution of host and parasite

Sp. 1.
Possible co-infections a, b, c, g, e, f & h
Sampled co-infections a, b & c
Divergence pattern: Clustering

Sp. 2.
Possible co-infections f, h, i, j, o & d
Sampled co-infections j, f & o
Divergence pattern: Dispersion