Journal of Mammalogy, 100(3):1008–1018, 2019
DOI:10.1093/jmammal/gyy174

Topic modeling of major research themes in disease ecology of mammals

Barbara A. Han*, and Richard S. Ostfeld

Cary Institute of Ecosystem Studies, Box AB, Millbrook, NY 12545, USA

* *Correspondent:* *hanb@caryinstitute.org*

Disease ecology is a rapidly growing subdiscipline, and mammals and their parasites feature prominently in both historical and more recent research efforts. Nevertheless, the diversity of topics explored, and those not well explored, has not been systematically assessed. We conducted a systematic review of the published scientific literature in disease ecology of mammals and subjected the collection of original and review articles identified to a topic modeling approach, which is based on the words used in the published texts and their contexts (i.e., the frequency and strength of their semantic relationships with one another). In addition to concept maps identifying the most prominent research themes, we identified eight (not mutually exclusive) subcategories of studies, including experimental, theoretical, comparative, behavioral, immunological–microbiological, biogeographic–macroecological, vector-focused (e.g., mosquitoes), and disturbance-focused. The most prominent themes arising in review papers included the ecology of zoonotic diseases transmitted from non-human mammals, comparisons of pathogen prevalence between mammalian species, and pathogen discovery–disease surveillance studies, particularly of marine mammals and bats. For the original articles, the most prominent themes included ecology of rodent-transmitted viral and bacterial diseases and the population biology of zoonotic hosts. Most studies used comparative or descriptive approaches to investigate mammal–pathogen–disease relationships at a local scale, focusing on vector-borne diseases. Experimental, modeling, immunological, and behavioral approaches were strikingly underrepresented. Topics of strong conceptual importance, but that are underrepresented in the current literature, include: 1) the effects of the population density of mammalian hosts, and manipulations of density, on pathogen transmission; 2) macroecological studies that quantify effects of mammalian host species on parasite abundance and prevalence; and 3) effects of climate change on physiological and behavioral processes relevant to mammal–parasite interactions.

La ecología de las enfermedades es una subdisciplina en rápido crecimiento, y los mamíferos y sus parásitos ocupan un lugar destacado en los esfuerzos de investigación tanto históricos como en los más recientes. Sin embargo, la diversidad de temas estudiados, y de aquellos que no han sido tan bien explorados, no ha sido evaluada sistemáticamente. Llevamos a cabo una revisión sistemática de la literatura científica publicada respecto a ecología de las enfermedades en mamíferos, y sometimos los artículos originales y revisiones encontradas a un análisis temático basado en las palabras utilizadas en las publicaciones y en el contexto de uso (es decir, la frecuencia y la intensidad de las relaciones semánticas). Además de definir mapas conceptuales que muestran los temas de investigación más destacados, identificamos ocho subcategorías de estudios (no mutuamente excluyentes), que incluyen estudios experimentales, teóricos, comparativos, de comportamiento, inmunológicos-microbiológicos, biogeográficos-macroecológicos, centrados en vectores (por ejemplo, mosquitos), y enfocados en perturbaciones. Los temas más relevantes que surgieron de los artículos de revisión incluyeron la ecología de las enfermedades zoonóticas transmitidas por mamíferos no humanos, las comparaciones de la prevalencia de patógenos entre especies de mamíferos, y los estudios de monitoreo para asociar patógenos y enfermedades, en particular de mamíferos marinos y murciélagos. En los artículos originales, los temas más destacados incluyeron la ecología de las enfermedades virales y bacterianas transmitidas por roedores, y la biología de poblaciones en huéspedes zoonóticos. La mayoría de los estudios utilizaron enfoques comparativos o descriptivos para investigar las relaciones entre mamíferos, patógenos y enfermedades a escala local, centrándose en las enfermedades transmitidas por vectores. Estudios con enfoques experimentales, utilizando modelos matemáticos, o estudios inmunológicos y de comportamiento fueron sorprendentemente más escasos. Los temas de gran importancia conceptual, pero que están poco representados en la literatura actual, incluyen: 1) los efectos de la densidad de la población de los mamíferos hospedadores, y de las manipulaciones de esta densidad, sobre la transmisión de patógenos; 2) estudios macroecológicos que cuantifican los efectos de los mamíferos hospedadores en la abundancia y prevalencia de parásitos; y 3) los efectos del cambio climático en los procesos fisiológicos y de comportamiento relevantes para las interacciones mamífero-parásito.

Key words: behavioral ecology, content mining, helminth parasites, machine learning, marine disease, vector-borne, zoonosis

The global rise of infectious diseases in one particular species of mammal (humans) has stimulated widespread research about the role that animals play in improving and compromising human health. Disease ecology, while a relatively young field, is now secure as a prominent subdiscipline because of the central role that ecology plays in the distribution and dynamics of infectious diseases. This is evidenced by increasing research outputs (in the form of scholarly publications) on both infectious diseases and zoonotic diseases (human infectious diseases with animal origins) over the last 20 years (Supplementary Data SD1 and SD2). Indeed, the number of published studies is increasing monotonically for mammals with no signs of slowing (Supplementary Data SD2). Although parasites and pathogens are ubiquitous in all vertebrates, and despite greater species richness in the other vertebrate classes (e.g., birds, amphibians), disease ecology of mammals has become and will likely remain a major research focus.

The trend of increasing interest in the disease ecology of mammals is evident from simple tallies of publications, but the diversity of topics explored, and the topics not well explored, is more difficult to assess. This is partly due to increasing difficulty in keeping pace with the volume of new literature produced each year. Since this high volume is not accompanied by proportional increases in human cognition or time, this also means that fewer papers are incorporated into qualitative synthetic reviews by experts. Even with the application of methods developed to systematically identify and select relevant articles (e.g., Lajeunesse 2016), selecting which articles to include in the final synthesis is inherently subjective. Similarly, mammalogists develop their own impressions of which topics are well-studied and which are neglected, but these are bound to be subjective and untested. We thought that an objective assessment of how published studies of disease ecology of mammals have been allocated to various topics would help to create mental maps (i.e., concept maps) of research spaces that are well-trodden versus those that have been neglected. Consequently, we collated and analyzed primary research (scholarly publications, including reviews) and applied a topic modeling approach (Stockwell et al. 2009; Blei 2012; Nunez-Mir et al. 2016). Topic modeling is a machine-learning method that can directly discern prominent trends from text using semantic learning algorithms. These methods go beyond simply identifying frequently used terms or the occurrence of particular terms (key words) by also leveraging information about surrounding words. For instance, words like "detected," "during," "collected," and "including" are typically not included in literature searches, which are usually limited to key words, but such words are crucial for providing context and meaning (the semantics). These groups of words form "concepts," which comprise key words and surrounding words that together form a meaningful unit whose frequency and co-occurrence with other concepts can be measured statistically across a collection of texts, and that can also be mapped graphically (concept map). Topic modeling approaches have been applied to diverse fields, for instance, to uncover dominant sentiments in text from social media, and for bias determination through the analysis of word groupings between genders or ethnicities. Here, we apply this method to literature on disease ecology in mammals to identify the most prominent research topics (concepts). Because these mining and machine-learning methods are applied directly to the text, prominent research themes are identified on the basis of the words themselves, as well as on their context (i.e., the frequency and strength of their semantic relationships with each other). We applied these methods to two subsets of scholarly literature in disease ecology, one exclusively for review articles and one exclusively for journal articles, with the expectation that these corpora would distinguish well-established research topics (journal articles on particular host–pathogen systems) from more complex, systems-level topics (e.g., the relationship between climate or land-use change and disease) that may emerge from review articles. For the most prominent themes identified from journal and review articles, we further examined the text according to particular subcategories in disease ecology to highlight underrepresented areas for novel contributions.

Materials and Methods

Literature collation.—Scientific literature on mammals and infectious diseases was collated via a literature search from Web of Science using the search string *mammal** AND *disease ecology* from 1980 to 2018. The resulting publications (n = 406) were divided into two groups: reviews and journal articles. Books, book chapters, and dissertations were excluded. These two groups were then checked via perusal to assess relevance. For example, publications that focused primarily on non-mammal taxa (e.g., birds, insects, fish) were removed (i.e., publications that were not about mammals); publications that described the development or usage of mammalian cell lines to assess pathogenesis of an infectious organism were removed (i.e., publications that were not ecological). Several articles were excluded because explicit investigation of parasites, pathogens, or infectious disease was not obvious. This manual filtering process resulted in 62 reviews and 76 journal articles for a total corpus of 138 publications. Although several articles were actively excluded, no attempt was made to manually augment this initial dataset. For instance, particular journal articles and reviews were notably missing from the results of this structured search in Web of Science. However, for reproducibility and consistency, our analyses proceeded with the subset of relevant papers that were discoverable using Web of Science. Articles were prepared for topic modeling by including the text of the abstract for each paper, but excluding all other identifiers (e.g., author names, journal title, article title).

Topic modeling.—Topic modeling is an approach that applies groups of algorithms for text mining and machine learning to discover themes within unstructured bodies of text (Grün and Hornik 2011; Blei 2012; Nunez-Mir et al. 2016). These themes are defined by concepts, which are sets of weighted terms that are found to frequently co-occur (or "travel" together) throughout the literature corpus. In our topic model, implemented via Leximancer software (www.leximancer.com), these concepts were extracted automatically from the literature corpus using unsupervised seeding in which the starting point of each concept (a single word, or concept seed) is identified based on the frequency of its occurrence in the text. More terms are added to each concept seed until the learning process forms a thesaurus of terms for each concept. The thesaurus is then used to determine if that concept is represented in particular segments of text across the literature corpus (both the frequency of occurrence and co-occurrences between concepts). In our model, sentence segment size was set to three sentences, and a text segment is scored as containing a particular concept if the cumulative weights of the thesaurus words contained in that segment cross a predefined threshold (Smith and Humphreys 2006; Leximancer 2017).

We modeled the corpus in three separate analyses: all articles together (reviews + journal articles, 138 articles total); review articles separately (62 articles); and journal articles separately (non-reviews; 76 articles). We generated a concept map depicting the results of each analysis. We further compared the results of the most important themes identified for reviews and journal articles by examining subcategories represented within all text segments comprising these major themes.

Research subcategories.—The most important themes in these analyses each contained 140–200+ text segments. To increase the resolution of these broad themes, we a priori identified eight subcategories of research in disease ecology. These subcategories were identified subjectively by authors (i.e., not by the topic models themselves) to represent more general research areas that cut across themes. We scored all of the three-sentence text segments that composed the most frequently represented themes for each analysis (the top two and the top four naturally emerging themes for journal and review articles, respectively). These subcategories were not mutually exclusive. The eight subcategories were defined as follows:

1. Experimental studies. Research was based on experimental manipulation.
2. Theoretical studies. Research was based on mathematical or other formal, predictive models; if models were used only to analyze empirical data (statistical models), then this was not considered to be "modeling." Modeling "experiments" (i.e., studies in which a model was used as the device for manipulating factors) were included here rather than in experimental studies.
3. Comparative or descriptive studies. Research involved non-experimental, empirical studies (e.g., cross-species comparisons [same pathogen, multiple hosts; same host, multiple pathogens] and monitoring of patterns through time [serosurveys through a season]).
4. Behavioral studies. Research involved some element of mammal behavior or behavioral ecology (e.g., studies that examine host behaviors that mitigate or exacerbate disease transmission).
5. Immunological or microbiome studies. Research involved immunological interactions between hosts and pathogens or vectors. In addition, studies are included involving mechanisms of pathogenicity, including host microbiomes, and -omics research (e.g., genomics, proteomics, metabolomics) that was relevant to disease ecology.
6. Biogeographic or macroecological studies. Research based on data from multiple study sites within a broad region or from multiple regions. Studies examining the geographic distributions of mammalian hosts or their pathogens are included.
7. Vector studies. Research focused specifically on vectors (typically arthropods) and their role in mammalian disease ecology.
8. Disturbance or environmental change. Research addressed the impact of rapid environmental change or disturbance (e.g., assessments of the effects of climate change, land-use change, pollution, or culling).

In some cases, text segments were not assigned to any category. This occurred only when the text segment entries provided insufficient information (e.g., when they included only introductory or contextual prose with no information on the approach or results of the study). Categorization was exclusively based on information in the text segments regardless of prior knowledge by the authors about a particular study or system.

All of the abstracts for reviews and journal articles, as well as Leximancer project configuration settings for the topic modeling analyses, are available online at figshare at the following DOIs—reviews: http://doi.org/10.6084/m9.figshare.5395066; journal articles: https://doi.org/10.6084/m9.figshare.5395093; reviews project configuration settings: https://doi.org/10.6084/m9.figshare.5395123; journal articles project configuration settings: https://doi.org/10.6084/m9.figshare.5395135; all articles project configuration settings: http://doi.org/10.6084/m9.figshare.5411608).

Results and Discussion

Below, we present three separate model runs: one for all papers combined, one for reviews, and one for journal articles. Since the words and sentence fragments for each corpus are different, the number of themes identified and concepts included in each theme are different for each analysis.

All papers combined.—Altogether, 138 articles (journal articles and review articles) were retained for topic modeling. Each theme (circles in Figs. 1–3) represents one or more concepts (words inside the circles in concept maps). These concepts are initially placed on the concept map randomly, but as the semantic evidence for associations between concepts increase, concepts cluster closer together. Concepts located in close proximity appear more frequently or co-occur more frequently across many text segments in the corpus. The colors on the concept map correspond to a heat map, whose colors are based on theme connectivity, which is the sum of co-occurrences among each of the concepts contained within the theme. Theme connectivity gives a measure of the relative importance of the themes within a data set, with warmer colors representing higher connectivity between the concepts within a theme. Although theme color indicates connectivity, the size and relative positions of the circles are not meaningful; the circles serve only as boundaries and do not reflect theme prevalence within the corpus. Prominent themes are identified and ranked by the total number of sentence fragments and the summed connectivity between concepts in a theme.

In a topic model on all articles (reviews and non-review journal articles), a total of 12 themes emerged (Fig. 1). For concepts in each theme, some words (e.g., vectors, habitat) have clearer meanings than do others (e.g., using, different, study). In Leximancer, it is possible to exclude such terms from the model (dropping the terms using, different, study). However, the relative importance of themes, and the composition of concepts included in the top most important themes, remained unchanged with minor variations in the thesaurus; thus, we elected not to engage in secondary pruning of concepts. Unsurprisingly, studies on mammalian species was the most prominent theme, followed by the "disease" theme, particularly the ecology of viral pathogens that pose an important health risk to humans. Additional major themes included studies related to the control of both wild and some domesticated populations (dogs [*Canis familiaris*], in particular), and studies emphasizing the geographical patterns (abundance, distribution, and population counts) of mammalian hosts and vectors.

Reviews versus journal articles.—A subset of 62 articles was categorized by Web of Science as review articles about disease ecology of mammals. Out of 14 major themes, four themes

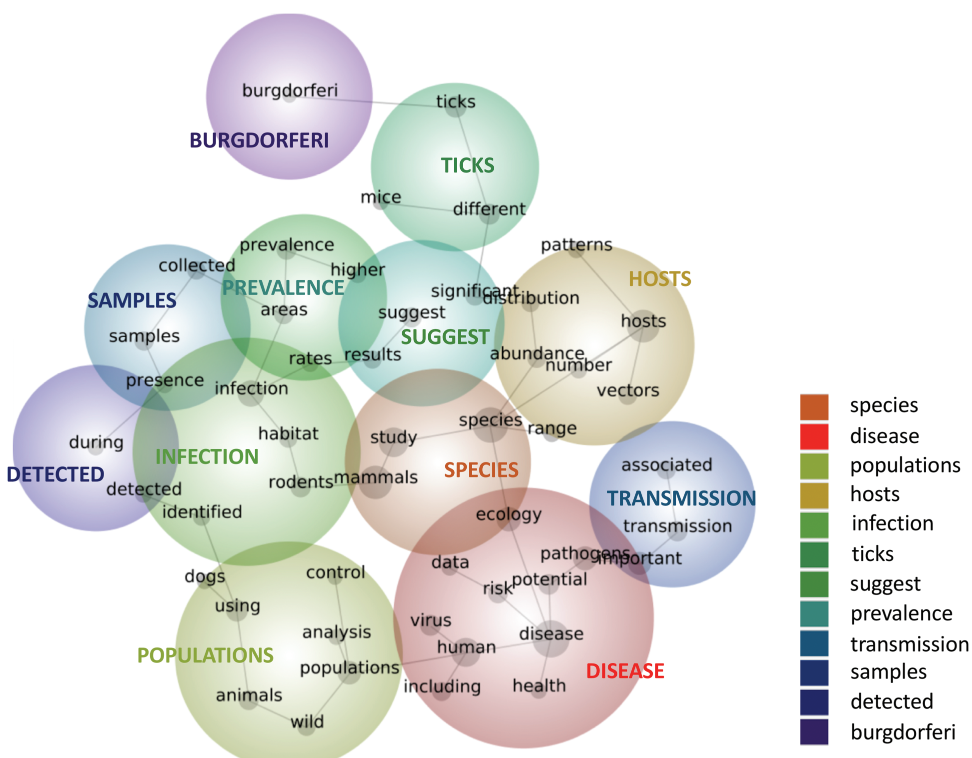

Fig. 1.—A concept map depicting the major research themes represented in the academic literature on disease ecology of mammals (including both reviews and journal articles). Words in black text are the concepts comprising larger themes. The size of the gray dot underlying each concept reflects its connectivity to other concepts on the map (i.e., how often that concept co-occurs with every other concept on the map); concepts with larger gray dots have greater connectivity to other concepts. The spanning tree (gray lines) draws further attention to the strongest connections between particular concepts. The larger, colored circles depict prominent themes, which are made up of highly connected concepts that appear together frequently within the same text segments. Concepts that "travel together" in this way also tend to appear near each other in map space. Occasional overlaps are meaningless. The name of each theme is determined by the most highly connected concept within the theme. These theme names are displayed within each circle as well as in the inset legend, which is organized in descending order of connectivity.

were the most important (quantified by the number of text segments in each theme; Fig. 2). Review articles are strongly associated with the ecology of diseases with the potential to pose important health risks to humans. Another major theme of reviews included comparative studies of pathogen prevalence across mammal species, with particular emphasis on quantifying infection in marine mammals and bats.

Prominent themes among the 76 journal articles (Fig. 3) represented studies conducted on rodents examining the ecology of rodent-borne diseases, and vector-borne diseases such as Lyme disease (caused by *Borrelia burgdorferi*, transmitted by ticks). Additional themes among journal articles included studies of natural populations of mammals that carry zoonotic pathogens (pathogens with a vertebrate origin that are infectious to humans), investigating their abundance, as well as disease risk and prevalence in both humans and wild species.

Further examination of the text segments composing the top themes in journal articles and reviews, separately, revealed that for journal articles, comparative or descriptive studies were far more common (over 80%) than were either experimental or modeling studies, which together comprised ~7% of the total text segments. The preponderance of comparative or descriptive studies was apparent for both of the top themes ("mammals" and "species"; Figs. 4 and 5). For both themes, about 40% of text segments concerned particular arthropod vectors. Rather than being macroecological, the great majority of studies were undertaken at a single locality, although some of these studies examined more than one habitat type. Studies focusing on immunological, microbiological, and behavioral factors were represented poorly, and fewer than 10% of studies examined the effects of anthropogenic disturbance or environmental change. Taken together, the results indicate that most studies represented local-scale, comparative, or descriptive approaches to mammal–pathogen–disease relationships, with a strong focus on vector-borne diseases. Experimental, modeling, immunological, and behavioral approaches were strikingly underrepresented.

For the review articles, results were qualitatively similar to those for journal articles, but showed greater diversity across subcategories. For example, although comparative or descriptive reviews were more frequent than were experimental or modeling approaches, reviews generally emphasized macroecological and immunological or microbiological approaches more than did non-review articles (Fig. 4). This suggests that researchers reviewing published literature disproportionately emphasize the importance of these two approaches relative to researchers who publish non-review articles. The considerably greater emphasis in review than non-review papers on

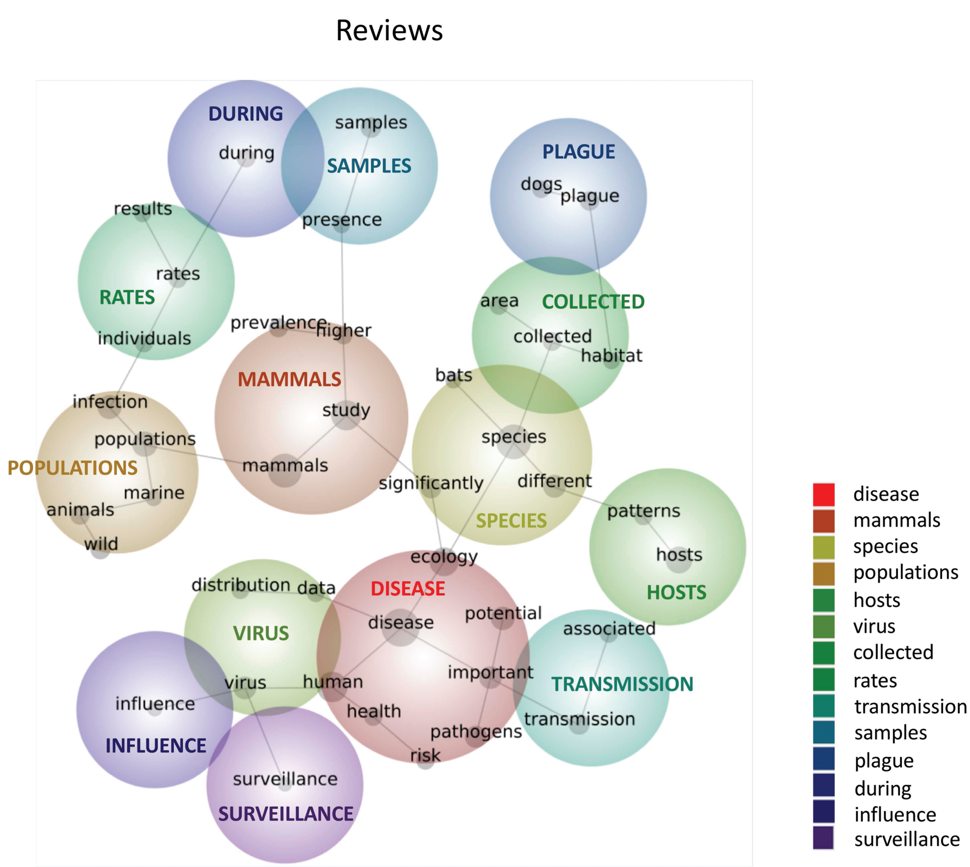

Fig. 2.—A concept map depicting the major research themes represented in academic review articles on disease ecology of mammals (excluding non-review journal articles). Each word represents a concept, with the size of gray dots beneath each concept scaling with the degree to which each concept is connected to other concepts. The spanning tree (gray lines) depicts the strongest connections between particular concepts across the corpus. Themes are collections of concepts that are identified by colored circles, with warmer colors corresponding to the most prominent concept within the theme. Theme labels within circles match colors in the legend, which is organized in descending order of connectivity.

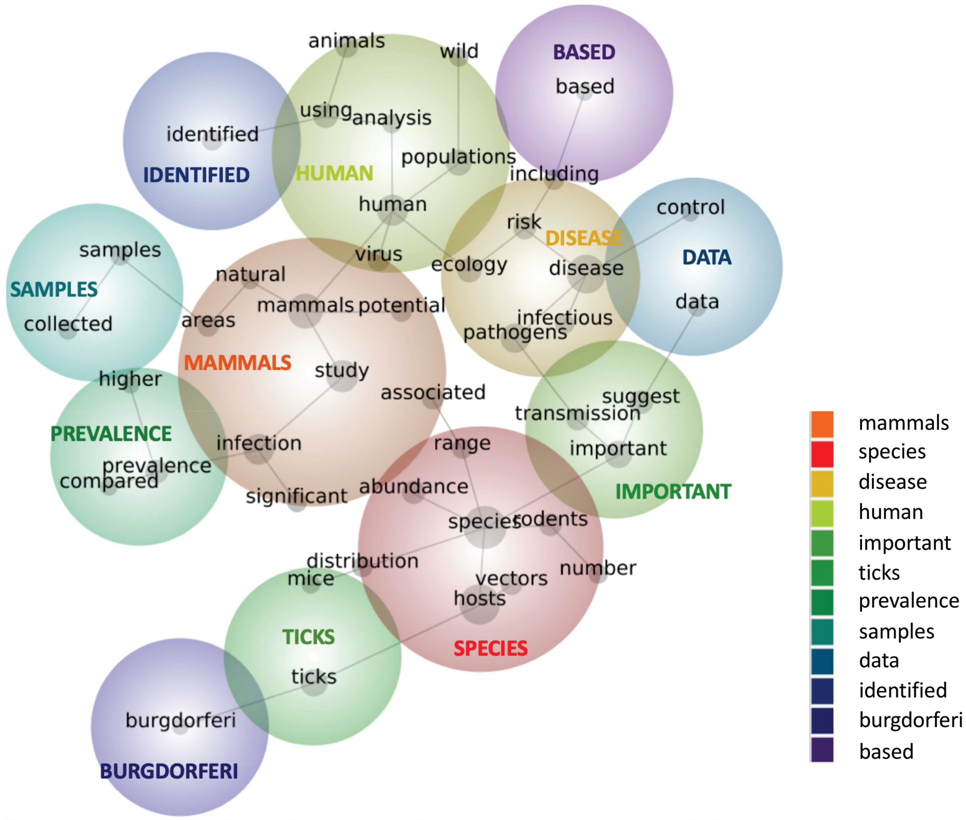

Fig. 3.—A concept map depicting the major research themes represented in academic journal articles on disease ecology of mammals (excluding reviews). Multiple concepts (individual words) comprises themes whose prominence is indicated by warmer colors. The gray dots scale with the connectivity of the individual concept, and gray lines indicate concepts that are the most strongly connected to each other in the corpus.

macroecological and immunological or microbiological approaches suggests that authors writing overviews and syntheses have called for more primary research in these areas, but so far, these urgings have not come to fruition. Across the four major themes identified in review articles, few clear differences characterized the sub-categorization of text segments that represent these themes (Fig. 6). One exception is the apparent emphasis on the "disease" theme in the macroecological, vector, and disturbance categories, whereas the "populations" theme was uncommon across all categories.

Across all topic models and identified themes, 'mosquitoes' was notably absent as a concept. One reason for this may be that studies focusing on mosquito-borne diseases of mammals are subsumed in the prominent "human disease" theme, although if this were the case, we might expect that mosquitoes, or the mosquito genera *Aedes* and *Anopheles*, would arise as a concept in the topic modeling of journal articles, similar to the appearance of ticks (Fig. 3). We posit that the absence of mosquitoes as a theme or as an identifiable concept in the disease ecology corpus may highlight a public health (as opposed to an ecological) bias for conducting research towards the development of control measures against mosquito-borne diseases (for instance, vaccine and drug development) as opposed to basic research on the ecology of mosquito vectors or mosquito-borne diseases, even zoonotic diseases with important sylvatic cycles in mammals (e.g., flaviviruses, such as Zika and Yellow Fever viruses, and some *Plasmodium* parasites).

Although the concept "virus" appeared consistently in top themes across all analyses, none of the other types of parasites or pathogens did so. This was somewhat surprising given that there are more mammal-borne bacterial pathogens that are capable of causing illness in humans (Taylor et al. 2001; Han et al. 2016), and the majority of human disease outbreaks (those that are increasing in incidence of geographic range) have been caused by bacterial pathogens (Smith et al. 2014). However, among zoonoses categorized as emerging (increasing in incidence, or expanding geographically), viruses are more common than bacteria, and more people are impacted by common viral diseases globally (e.g., viral gastroenteritis, dengue virus, influenza, HIV—Marston et al. 2014). This is accompanied by an outsized perception of disease risk for highly virulent zoonotic viruses (Ebola virus, Lassa fever) compared to bacterial diseases, and especially compared to more chronic infections caused by helminths. Helminth parasites are a major source of disease and poverty in one-third of the global human population (Hotez et al. 2008), and cause diseases in several economically important livestock species (Cleaveland et al. 2001). Helminth parasites play an important role in regulating some host populations and affecting broader ecological communities (Dobson and Hudson 1986). Despite this, we found that studies on the ecology of helminth parasites in mammals were few, suggesting that helminth diseases are a neglected research topic in humans (Hotez et al. 2009) as well as other mammals.

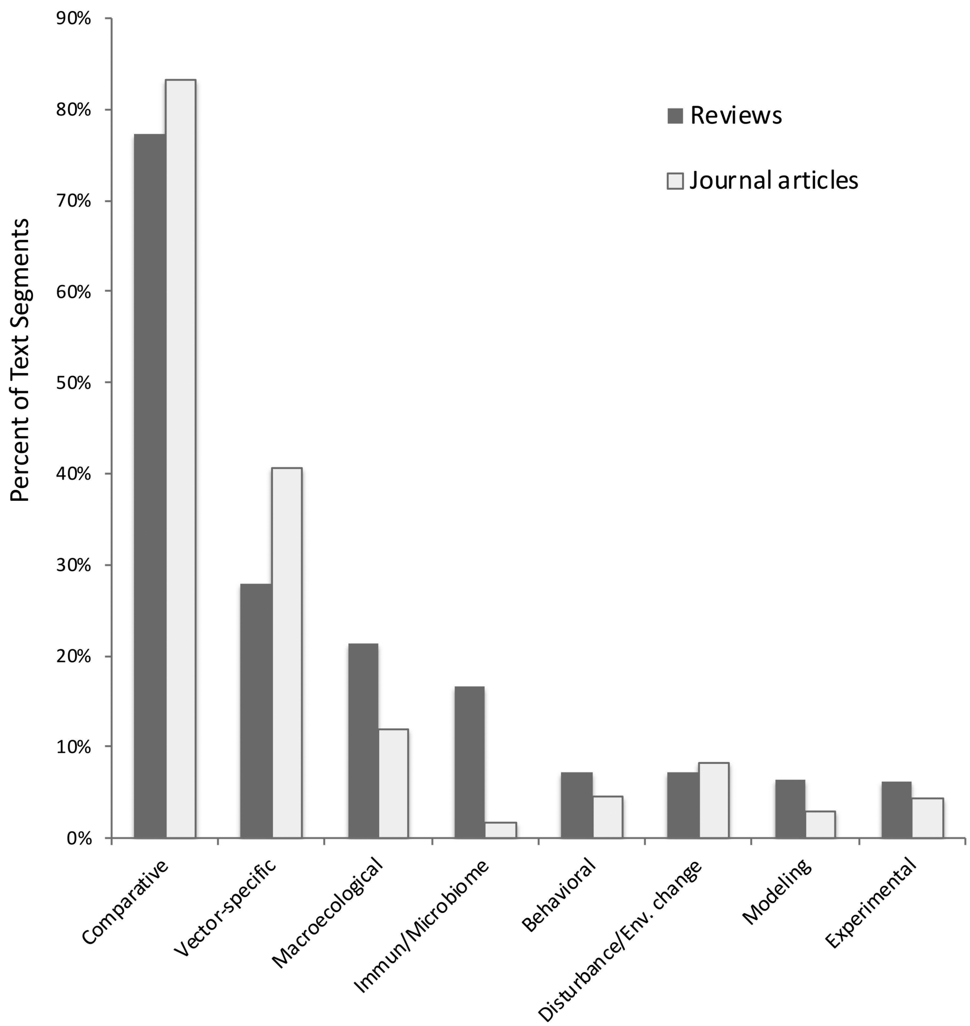

Fig. 4.—The percentage of text segments that fell into eight non-mutually exclusive subcategories of research in disease ecology found in review articles and non-review journal articles. For reviews, sub-categorizations of the text segments from the top four most important themes were averaged together. For journal, text segments from the top two themes were averaged.

Behavioral ecology and animal behavior in general are fundamentally important to studies of infectious disease dynamics and risk (Hawley et al. 2011; Dizney and Dearing 2016; Friesen and Roth 2016). Nonetheless, the number of studies explicitly investigating behavioral aspects of disease ecology represented only 5% and 7% of the top themes in journal articles and reviews, respectively. This could reflect long-standing challenges in behavioral ecology, which include the need to collect large amounts of data to account for high variation among individuals, populations, environments, or seasons. In contrast to areas like molecular ecology, where the fast pace of technological developments has made diagnostic tools more affordable, basic empirical work in behavioral ecology continues to be an expensive endeavor in terms of time and funding. However, advances in understanding disease dynamics rely on basic knowledge about host movement and behaviors (Giles et al. 2016), and may be especially important for more accurately forecasting spillover events of interest for zoonotic diseases in human populations (Plowright et al. 2015). Technological advances to studying animal movement, such as those incorporated through the International Cooperation for Animal Research Using Space (ICARUS) Initiative (ICARUS, 2002), could catalyze new research exploring disease-related questions that were previously intractable due to data limitations.

Our consideration of scientific publications demonstrates a high and increasing level of research productivity in the realm of mammalian disease ecology and suggests a promising future for this field (Supplementary Data SD1 and SD2). In general, research undertaken is dominated by descriptive and comparative approaches for studying mammal host–vector–parasite associations within local study sites. Such studies are fundamental to the development of a strong empirical foundation in natural history and population biology of disease. Mammalian disease ecologists are poised to capitalize on this foundation and make rapid advancements in developing research areas of strong conceptual and practical importance. In particular, the pursuit of mechanisms underlying associations between mammals and their pathogens and vectors would be enhanced by experimental approaches that strengthen causal inference. For example, experiments can be designed to test molecular-level controls on pathogen invasion of host tissues, or immunological and endocrinological influences on host invasibility

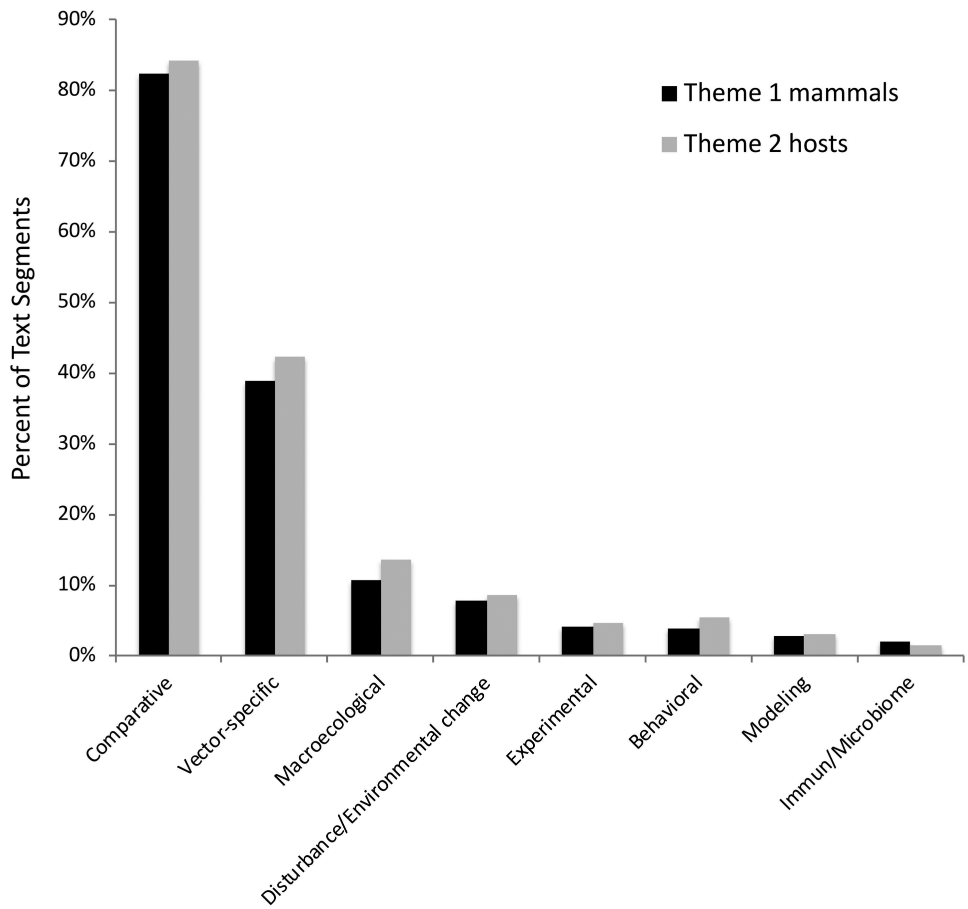

Fig. 5.—The percentage of text segments from the top two most important themes identified from a topic model on non-review journal articles on disease ecology of mammals.

by particular pathogens (Gervasi et al. 2017). Experiments can reveal which host community parameters, anthropogenic disturbances, or host behaviors strongly influence the dynamics of pathogen transmission and disease spread or decline.

Although host behaviors, by impacting pathogen transmission, play a fundamental role in disease ecology, several widespread animal behaviors remain poorly studied, including sociality (e.g., Nunn and Altizer 2006), aggression (e.g., Hamede et al. 2013), individual dispersal (e.g., Buskirk and Ostfeld 1998), and behaviors that span vast distances (e.g., nomadism, seasonal migration—Altizer et al. 2011). Given their importance and the inherent difficulty in studying host behavior in natural settings, hypotheses generated by a combination of empirical and theoretical approaches may help identify situations in which causality is empirically tractable. Simple epidemiological models can incorporate empirical data patterns from manipulative or observational studies explicitly to form an iterative model-data-model loop that begins to draw boundaries on realistic disease patterns and the mechanisms that drive them in nature (Han et al. 2015; Stephens et al. 2016). Similarly, theoretical approaches can identify important characteristics that are causally linked to host–pathogen–vector associations, for instance, by incorporating biting and pathogen incubation rates within vectors (extrinsic incubation) to enhance predictive understanding of the risk of pathogen invasion into naive host populations (Caminade et al. 2017). Biogeographic patterns of mammalian host–pathogen–vector interactions are less amenable to experimental approaches but are critical for understanding zoonotic disease risk to human populations and to the conservation of mammals worldwide (Stephens et al. 2016).

Our results can be compared to those of a recent review of the application of disease ecology concepts to wildlife disease management (Joseph et al. 2013). Of the 12 theoretical concepts arising from disease ecology that they selected, Joseph et al. (2013) found that only density-dependent transmission was applied in >10 of the reviewed papers, and only four other concepts (i.e., seasonal disease dynamics, individual-level variation, spatial transmission dynamics, and multi-host dynamics) were applied in more than five papers. The approaches taken by Joseph et al. (2013) are distinct from those we take in important ways. They a priori identified a set of concepts from disease ecology, whereas we allowed the topic modeling algorithm to identify themes, including theoretical concepts. In addition, Joseph et al. (2013) explored the conceptual drivers of wildlife management, whereas we focused more broadly on mammalian biology. Nevertheless, some lessons emerge from comparing the two efforts. One is that the literatures on wildlife disease management and mammalian biology overlap in emphasizing host population density as either a management endpoint or a

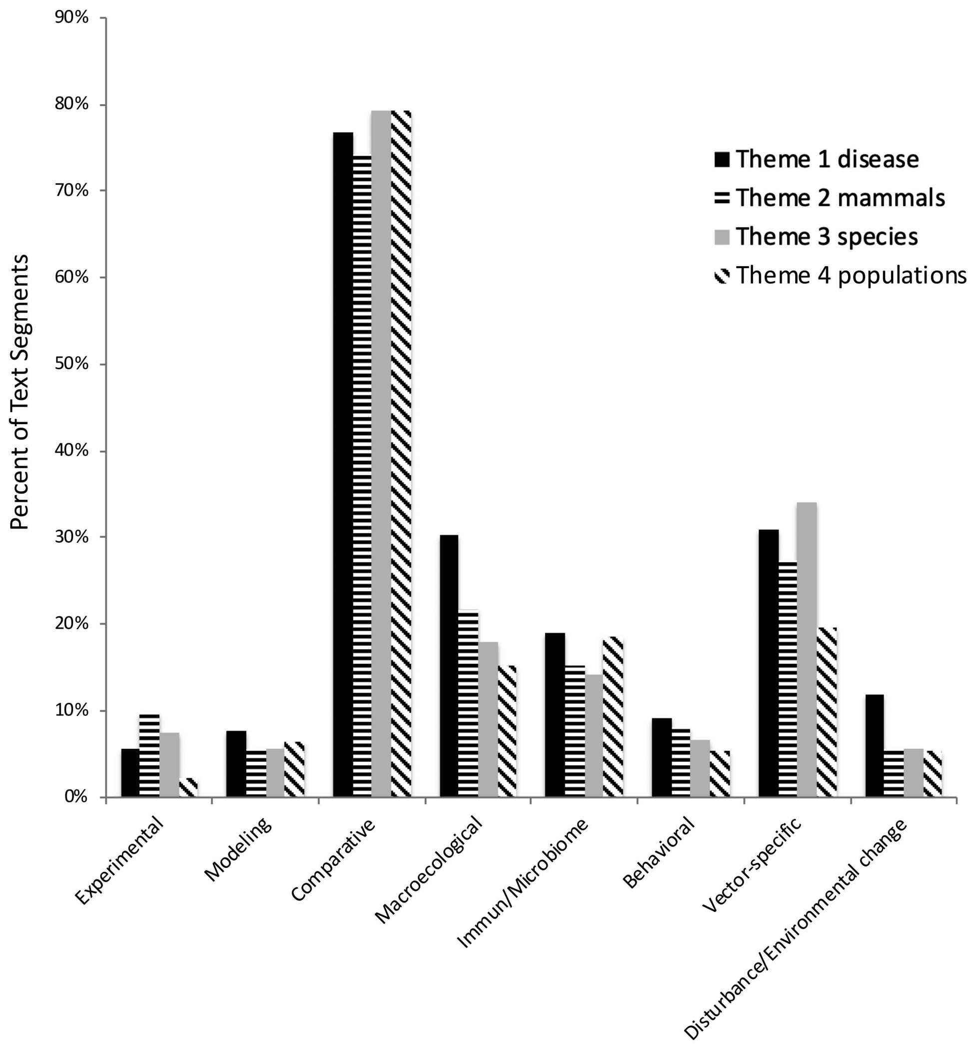

Fig. 6.—The percentage of text segments from each of the four most important themes identified from a topic model on review papers on disease ecology of mammals.

driver of disease dynamics. A second is that the literature on mammalian disease ecology emphasizes studies of rodents and bats, which are perhaps less frequently the focus of efforts by wildlife biologists to manage wildlife disease, leading to somewhat distinct taxonomic foci by the two disciplines. Finally, the results of systematic reviews can differ based on initial search strategies and analytical approaches. Future work could investigate these broad patterns by comparing major research themes identified in literature compared to those arising from more narrowly defined search terms. Comparing such results may shed light on whether our understanding of some topics in disease ecology are driven primarily by research arising from particular animal taxa (e.g., specific mammal taxa, or specific classes of vertebrates), or from particular systems (terrestrial versus aquatic, managed versus protected versus wild).

Conclusion

The field of disease ecology has a well-developed theoretical and conceptual framework that guides much of the empirical science, including that conducted by mammalogists (Ostfeld et al. 2010). Rather than relying on expert opinion or other subjective approaches to understand and synthesize recent efforts in disease ecology involving mammals, we chose to apply topic modeling to enhance transparency, repeatability, and objectivity. We close by suggesting a few research areas that would help develop some of the research areas that our analyses indicate are uncommon.

First, although many pathogens experience enhanced rates of transmission with increasing host density (density-dependent transmission), the importance of transmission processes that are independent of host density (frequency-dependent transmission) is poorly understood. Frequency dependence can occur when behaviors structure intraspecific interactions so that contact rates and opportunities for transmission are independent of density, and also can occur when vectors are required for transmission. Because of the prominence of culling hosts to manage disease transmission, under the assumption of density dependence, determining when transmission is less dependent on density is critical in developing both theory and application (Borremans et al. 2017).

Second, publicly available databases on mammals and their parasites are allowing explorations of key topics in mammalian disease ecology (Stephens et al. 2016, 2017). However, determining which parasites infect which hosts remains deceptively difficult in some cases. For instance, some mammals might be listed as a host for a particular parasite on the basis of the detection of antibodies to that parasite, but some mammals might be able to clear infections rapidly after seroconversion, whereas others might not. Even the binary detection of the parasite itself, or its nucleic acids, in certain host species might mask large differences between hosts in their ability to support infection and shed propagules. Quantification of the roles played by mammalian host species in maintaining and transmitting parasites, versus absorbing and destroying them, would permit researchers to determine the consequences of population or community dynamics and anthropogenic impacts on disease prevalence. Comparisons of the macroecology of host exposure to parasites with the macroecology of host immune responses could constitute the foundation for a new discipline, macroimmunology.

Third, disease ecologists increasingly focus on the impacts of global climate change on transmission, prevalence, and disease dynamics (Harvell et al. 2009; Rohr et al. 2011; Altizer et al. 2013). However, more attention is devoted to ectothermic species, including the parasites, arthropod vectors, and many of the hosts, than to endotherms. Although the body temperatures of endotherms are buffered against a range of environmental conditions, environmental temperatures can affect immunological and endocrinological functions that influence host susceptibility and behavior (Hawley et al. 2011; MacGillivray and Kollmann 2014). As the geographic ranges of both endotherms and ectotherms change with climate, new host–parasite–vector associations can develop, potentially altering transmission, infection, and disease (Kutz et al. 2005; Angert et al. 2013). Integrated studies of how climate change affects physiological and behavioral processes relevant to mammal–parasite interactions and alters interactions with vectors could help to develop mitigation strategies that protect mammals and their ecosystems in a rapidly changing world.

Acknowledgments

The authors thank V. Ramesh for data collation and programming assistance and M. Gillespie for editorial assistance. We thank Paula Pappalardo and Alonso Aguirre for translation and assistance with the Spanish summary. This work was supported by grants from the National Science Foundation, Ecology and Evolution of Infectious Diseases program (DEB-1717282 to BH, and DEB-1619072 to BH and RO).

Supplementary Data

Supplementary data are available at *Journal of Mammalogy* online.

Supplementary Data SD1.—Bar charts representing the annual number of articles catalogued in Web of Science through 2016. Results are from a search using the string [<*vertebrate class*>] AND *infectious disease*.

Supplementary Data SD2.—Bar charts representing the annual number of articles catalogued in Web of Science through 2016. Results are from a search using the string [<*vertebrate class*>] AND *zoono* disease*.

Literature Cited

Altizer, S., R. Bartel, and B. A. Han. 2011. Animal migration and infectious disease risk. Science (New York, N.Y.) 331:296–302.

Altizer, S., R. S. Ostfeld, P. T. Johnson, S. Kutz, and C. D. Harvell. 2013. Climate change and infectious diseases: from evidence to a predictive framework. Science (New York, N.Y.) 341:514–519.

Angert, A. L., S. L. LaDeau, and R. S. Ostfeld. 2013. Climate change and species interactions: ways forward. Annals of the New York Academy of Sciences 1297:1–7.

Blei, D. M. 2012. Probabilistic topic models. Communications of the ACM 55:77–84.

Borremans, B., J. Reijniers, N. Hens, and H. Leirs. 2017. The shape of the contact-density function matters when modelling parasite transmission in fluctuating populations. Royal Society Open Science 4:171308.

Buskirk, J. V., and R. S. Ostfeld. 1998. Habitat heterogeneity, dispersal, and local risk of exposure to Lyme disease. Ecological Applications 8:365–378.

Caminade, C., et al. 2017. Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño. 2015. Proceedings of the National Academy of Sciences of the United States of America 114:119–124.

Cleaveland, S., M. K. Laurenson, and L. H. Taylor. 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 356:991–999.

Dizney, L., and M. D. Dearing. 2016. Behavioural differences: a link between biodiversity and pathogen transmission. Animal Behaviour 111:341–347.

Dobson, A. P., and P. J. Hudson. 1986. Parasites, disease and the structure of ecological communities. Trends in Ecology & Evolution 1:11–15.

Friesen, O. C., and J. D. Roth. 2016. Alternative prey use affects helminth parasite infections in grey wolves. The Journal of Animal Ecology 85:1265–1274.

Gervasi, S. S., S. C. Burgan, E. Hofmeister, T. R. Unnasch, and L. B. Martin. 2017. Stress hormones predict a host superspreader phenotype in the West Nile virus system. Proceedings of the Royal Society B: Biological Sciences 284:20171019.

Giles, J. R., R. K. Plowright, P. Eby, A. J. Peel, and H. McCallum. 2016. Models of Eucalypt phenology predict bat population flux. Ecology and Evolution 6:7230–7245.

Grün, B., and K. Hornik. 2011. Topicmodels: an R package for fitting topic models. Journal of Statistical Software 40:1–30.

Hamede, R. K., H. McCallum, and M. Jones. 2013. Biting injuries and transmission of Tasmanian devil facial tumour disease. The Journal of Animal Ecology 82:182–190.

Han, B. A., A. M. Kramer, and J. M. Drake. 2016. Global patterns of zoonotic disease in mammals. Trends in Parasitology 32:565–577.

Han, B. A., A. W. Park, A. E. Jolles, and S. Altizer. 2015. Infectious disease transmission and behavioural allometry in wild mammals. The Journal of Animal Ecology 84:637–646.

Harvell, D., S. Altizer, I. M. Cattadori, L. Harrington, and E. Weil. 2009. Climate change and wildlife diseases: when does the host matter the most? Ecology 90:912–920.

Hawley, D. M., R. S. Etienne, V. O. Ezenwa, and A. E. Jolles. 2011. Does animal behavior underlie covariation between hosts' exposure to infectious agents and susceptibility to infection? Implications for disease dynamics. Integrative and Comparative Biology 51:528–539.

Hotez, P. J., P. J. Brindley, J. M. Bethony, C. H. King, E. J. Pearce, and J. Jacobson. 2008. Helminth infections: the great neglected tropical diseases. The Journal of Clinical Investigation 118:1311–1321.

Hotez, P. J., A. Fenwick, L. Savioli, and D. H. Molyneux. 2009. Rescuing the bottom billion through control of neglected tropical diseases. Lancet (London, England) 373:1570–1575.

International Cooperation for Animal Research Using Space (ICARUS). 2002. ICARUS Global Monitoring with Animals. https://icarusinitiative.org/. Accessed 15 November 2017.

Joseph, M. B., J. R. Mihaljevic, A. L. Arellano, J. G. Kueneman, D. L. Preston, P. C. Cross, and P. T. J. Johnson. 2013. Taming wildlife disease: bridging the gap between science and management. Journal of Applied Ecology 50:702–712.

Kutz, S. J., E. P. Hoberg, L. Polley, and E. J. Jenkins. 2005. Global warming is changing the dynamics of Arctic host–parasite systems. Proceedings of the Royal Society B: Biological Sciences 272:2571–2576.

Lajeunesse, M. J. 2016. Facilitating systematic reviews, data extraction and meta-analysis with the metagear package for R. Methods in Ecology and Evolution 7:323–330.

Leximancer. 2017. User guide (No. Release 4.5). https://paperpile.com/shared/SkoxET

MacGillivray, D. M., and T. R. Kollmann. 2014. The role of environmental factors in modulating immune responses in early life. Frontiers in Immunology 5:434.

Marston, H. D., G. K. Folkers, D. M. Morens, and A. S. Fauci. 2014. Emerging viral diseases: confronting threats with new technologies. Science Translational Medicine 6:253ps10.

Nunez-Mir, G. C., B. V. Iannone, B. C. Pijanowski, N. Kong, and S. Fei. 2016. Automated content analysis: addressing the big literature challenge in ecology and evolution. Methods in Ecology and Evolution 7:1262–1272.

Nunn, C. L., and S. Altizer. 2006. Infectious diseases in primates: behavior, ecology and evolution. Oxford University Press, New York.

Ostfeld, R. S., F. Keesing, and V. T. Eviner. 2010. Infectious disease ecology: effects of ecosystems on disease and of disease on ecosystems. Princeton University Press, Princeton, New Jersey.

Plowright, R. K., et al. 2015. Ecological dynamics of emerging bat virus spillover. Proceedings of the Royal Society B: Biological Sciences 282:20142124.

Rohr, J. R., et al. 2011. Frontiers in climate change-disease research. Trends in Ecology & Evolution 26:270–277.

Smith, A. E., and M. S. Humphreys. 2006. Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. Behavior Research Methods 38:262–279.

Smith, K. F., et al. 2014. Global rise in human infectious disease outbreaks. Journal of the Royal Society, Interface 11:20140950.

Stephens, P. R., et al. 2016. The macroecology of infectious diseases: a new perspective on global-scale drivers of pathogen distributions and impacts. Ecology Letters 19:1159–1171.

Stephens, P. R., et al. 2017. Global mammal parasite database version 2.0. Ecology 98:1476.

Stockwell, P., R. M. Colomb, A. E. Smith, and J. Wiles. 2009. Use of an automatic content analysis tool: a technique for seeing both local and global scope. International Journal of Human-Computer Studies 67:424–436.

Taylor, L. H., S. M. Latham, and M. E. Woolhouse. 2001. Risk factors for human disease emergence. Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences 356:983–989.

Submitted 27 November 2018. Accepted 19 December 2018.

Special Issue Editor was Michael R. Willig.