

*Original Contribution*

## Cross-Species Pathogen Transmission and Disease Emergence in Primates

Amy B. Pedersen<sup>1</sup> and T. Jonathan Davies<sup>2,3</sup>

<sup>1</sup>Centre for Immunity, Infection and Evolution, Institutes of Evolutionary Biology, Immunology and Infection Research, School of Biological Sciences, University of Edinburgh, Ashworth Labs, Edinburgh EH9 3JT, UK

<sup>2</sup>National Center for Ecological Analysis and Synthesis, 735 State Street, Santa Barbara, CA 93101

<sup>3</sup>Department of Biology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC H3A 1B1, Canada

**Abstract:** Many of the most virulent emerging infectious diseases in humans, e.g., AIDS and Ebola, are zoonotic, having shifted from wildlife populations. Critical questions for predicting disease emergence are: (1) what determines when and where a disease will first cross from one species to another, and (2) which factors facilitate emergence after a successful host shift. In wild primates, infectious diseases most often are shared between species that are closely related and inhabit the same geographic region. Therefore, humans may be most vulnerable to diseases from the great apes, which include chimpanzees and gorillas, because these species represent our closest relatives. Geographic overlap may provide the opportunity for cross-species transmission, but successful infection and establishment will be determined by the biology of both the host and pathogen. We extrapolate the evolutionary relationship between pathogen sharing and divergence time between primate species to generate “hotspot” maps, highlighting regions where the risk of disease transfer between wild primates and from wild primates to humans is greatest. We find that central Africa and Amazonia are hotspots for cross-species transmission events between wild primates, due to a high diversity of closely related primate species. Hotspots of host shifts to humans will be most likely in the forests of central and west Africa, where humans come into frequent contact with their wild primate relatives. These areas also are likely to sustain a novel epidemic due to their rapidly growing human populations, close proximity to apes, and population centers with high density and contact rates among individuals.

**Keywords:** host shift, HIV, phylogeny, host-specificity, parasite, virus

### INTRODUCTION

Emerging infectious diseases (EIDs) pose a serious and increasing threat to human health and welfare (Daszak et al., 2000; King et al., 2006; Jones et al., 2008). Diseases that have recently emerged in humans include SARS, HIV,

and swine flu. It is projected that 39.5 million people are currently infected with HIV/AIDS (WHO, 2006), and in 2007 alone, it was estimated that 18 billion dollars would be needed to prevent future HIV transmission and provide care for those already infected (WHO, 2007). In wildlife populations, EIDs also have resulted in recent and dramatic declines in mammals (e.g., Ebola in African apes; Walsh et al., 2003, canine distemper virus in wild dogs and lions; Roelke-Parker et al., 1996), amphibians (e.g., *Chytridi-*

Correspondence to: Amy B. Pedersen, e-mail: amy.pedersen@ed.ac.uk

*omyces* fungus in global amphibian populations; Daszak et al., 1999), insects (e.g., the mite *Varroa jacobsoni* in honey bees; Oldroyd, 1999), and birds (e.g., conjunctivitis due to *Mycoplasma gallisepticum* in passerines; Williams et al., 2002). Human EIDs originate from multiple sources (e.g., host shifts from animal reservoirs [zoonotics], evolution of existing organisms, and reemergence due to antimicrobial resistance), are increasing in number, and are globally distributed (Jones et al., 2008). Although there is not sufficient data to evaluate whether there is a similar trend for increasing frequency of disease emergence in wildlife populations, if similar drivers influence the process of disease emergence we might expect this to be the case (Daszak et al., 2000). The ability to predict when, where, and within which species pathogens are most likely to emerge is therefore increasingly urgent. Understanding the process of disease emergence will be critical for developing strategies to minimize risk and reduce the high cost of managing outbreaks once a disease has emerged.

Several review papers have explored the ecological and evolutionary drivers of disease emergence in human and wildlife populations and the conditions most likely to heavily impact new host species (Daszak et al., 2000; Osterhaus, 2001; Woolhouse et al., 2005; Jones et al., 2008; Parrish et al., 2008). Ecological drivers that favor emergence are most frequently associated with changes in host or pathogen ecology, specifically (1) an increase in host population density and contact rates, (2) environmental changes that effect host quality and demography, and (3) changes in host mobility and behavior. Evolutionary drivers of emergence include high genetic variability in the pathogen population (Morand et al., 1996; Gupta et al., 1998) or a lack of variability in the host (Daszak et al., 2000; Altizer et al., 2003b). However, sustainable emergence of a

pathogen on a host is only one outcome of a complex multistep process.

We focus on disease emergence after host shifts, defined as the movement of a pathogen to a new host species irrespective of the long-term consequences (Antonovics et al., 2002). Emergence after a host shift can result in rapid spread and high virulence because naïve hosts may lack appropriate immune responses (Osterhaus, 2001; Altizer et al., 2003b). The longer-term outcomes of a host shift can range from transient, unsustainable “spillover” events (e.g., Hantavirus; Hughes et al., 1993) to persistent, self-sustaining epidemics (e.g., HIV; Hahn et al., 2000; Fenton and Pedersen, 2005). There are three key stages required for successful emergence of an infectious disease on a “new” target host: (1) opportunity, (2) infection and transmission, and (3) establishment and sustainability (Fig. 1).

Close geographic proximity among hosts likely increases contact rates and, hence, opportunity for shifts (step 1, Fig. 1). Although geographic overlap may be important for determining which pathogens a host species is exposed to, successful infection and then transmission of the pathogen in the new host will depend on evolutionary and ecological factors determined by the biology of both host and pathogen (step 2, Fig. 1). For example, recent comparative analyses have demonstrated that viruses are the most likely group to cross species boundaries (Cleaveland et al., 2001; Pedersen et al., 2005), perhaps facilitated by high mutation rates and fast generation times providing the potential to rapidly overcome barriers to infection (Parrish et al., 2008). In contrast, within primates, half of the recorded helminth pathogens are specific to a single host species (Pedersen et al., 2005). Pathogen taxonomy and, by implication, pathogen biology, may then serve as an indicator of potential for host



**Figure 1.** Three key stages of disease emergence: (1) opportunity, (2) infection and transmission, and (3) establishment. The biogeography of both hosts and pathogens will contribute to the opportunity for cross-species transmission. For natural populations, opportunity for new pathogens may be restricted to neighboring species; however, wildlife trade, invasive species, and domestic animals also may be a source of new pathogens. The ability for a pathogen to infect and be

transmitted within the new host species is likely driven by ecological and evolutionary barriers, such as phylogenetic relatedness of the hosts and the evolutionary potential of the pathogen. Lastly, multiple demographic factors will affect the ultimate sustainability and establishment of a pathogen following a host shift, including population density, contact rates, and the rate of spread of the pathogen.

shifts. However, host characteristics also are important. Close evolutionary relationships among host species might translate into similar immunological responses and life-history traits (Pfennig, 2000; Ricklefs and Fallon, 2002; Perlman and Jaenike, 2003), and hence increase the likelihood of successful cross-species infection. The rate of spread in the new host and the progression through the stages of host shift to sustainable disease emergence will be affected by host and pathogen demographics (step 3, Fig. 1), for example, population density, contact rates, and the rate of spread of the pathogen in the new host ( $R_0$ ; number of new cases per infected individual) (Dobson and Foufopoulos, 2001; Fenton and Pedersen, 2005; Woolhouse et al., 2005). Mathematical models of host-pathogen interactions (e.g., traditional SIR—susceptible, infected and recovered—models) indicate that for diseases transmitted in a density dependent fashion (i.e., by direct contact), pathogens will only persist and become established in populations above a critical threshold density ( $R_0 > 1$ ; Anderson and May, 1992). Disease emergence may be inhibited at any stage, providing the opportunity for alternative management strategies; unfortunately, this complexity contributes to the difficulty in generating accurate predictive models.

In Davies and Pedersen (2008), we demonstrated a significant relationship between the geographical distribution and evolutionary relatedness of primate host species, and the similarity of their pathogen communities. This relationship reflects both the opportunity for host–host contact (geographic overlap), and the barriers to infection (host biology) (steps 1 and 2, Fig. 1). We expand on our previous analysis to explore the potential for future pathogen host shifts and disease emergence within and between wild primates and humans. First, we use the relationship between host relatedness, geographic overlap, and pathogen community similarity to quantify the risk of future host shifts that each species faces. Second, we evaluate how past host shifts between primates might have shaped pathogen communities by analyzing patterns of host specificity. Third, we explore whether the phylogenetic risk of host shifts may have shaped host geographic distributions over time. Finally, we identify regions of the globe where we predict host shifts between nonhuman primates and humans may be most frequent. This analysis provides the first quantitative attempt to assess the risk of pathogens host-shifting to humans from wildlife populations, a critical step toward predicting disease emergence.

## METHODS

### Primate Host and Pathogen Data

Pathogen species occurrences were obtained from the Global Mammal Parasite Database (Nunn and Altizer, 2005; [www.mammalparasites.org](http://www.mammalparasites.org)), comprising 2,462 records representing 415 pathogen species (including viruses, bacteria, helminths, protozoa, arthropods, and fungi) across 117 of the 232 wild primate species recognized in the phylogenetic tree of Bininda-Emonds et al. (2007). The human disease database from Taylor et al. (2001) was used to measure pathogen sharing between primates and humans. We use the term “pathogen” broadly to include both microparasites (i.e., viruses, bacteria, and protozoans) and macroparasites (i.e., helminths, fungi, and arthropods). For all pair-wise primate–primate combinations we estimated pathogen community similarity as:  $a/(a + b + c)$ , where  $a$  is the number of pathogen species found on two host species,  $X$  and  $Y$ ;  $b$  is the number of pathogen species on host  $X$  that are not found on host  $Y$ ; and  $c$  is the number of pathogen species on host  $Y$  that are not found on host  $X$  (for further details see Davies and Pedersen, 2008). Critically, this metric is not biased by differences in sampling intensity or the relative sizes of the pathogen species pools of  $X$  and  $Y$ .

We focus on general patterns of pathogen sharing, thus our estimate of pathogen community similarity was calculated for all pathogen types combined (helminths, protozoa, and viruses). The relationship between community similarity and evolutionary divergence may differ across pathogen taxonomy; however, we currently lack comprehensive data on pathogen occurrences (and absences), which limits the scope of more taxon-specific analyses. In theory, our approach could be applied to any pathogen subset given sufficient data.

### Generalized Linear Model of Pathogen Sharing

Following Davies and Pedersen (2008), we derived the relationship between evolutionary divergence (representing time to most recent common ancestor from the dated phylogenetic tree of Bininda-Emonds et al., 2007), and pathogen community similarity (as described above) between each primate pair using generalized linear modeling (GLM) with binomial errors and a logit link function in the statistical package R (R: a programming environment for data analysis and graphics, v. 2.7.1; <http://www.r-project.org/>). We used a GLM approach because this enabled the shape of the rela-

relationship between pathogen sharing with evolutionary relatedness to be characterized, and followed a similar protocol to that in previously published studies (Gilbert and Webb, 2007). Because we explore the combined risk of cross-species transmission between co-occurring primate hosts separately (see below), we did not include range overlap in the GLM directly.

For our model, we assume that pathogen community similarity (pathogen sharing) reflects the frequency of host shifts between any pair of primate species; the greater the frequency of host shifts, the more similar the pathogen communities. However, we note that pathogen community sharing will also reflect co-inheritance of similar pathogen communities from a common primate ancestor. Although the relative importance of co-inheritance versus host-shifts in structuring pathogen communities remains unresolved, experimental cross-infection of fungal pathogens among tropical plant species (Gilbert and Webb, 2007) revealed a qualitatively similar relationship between host evolutionary relationships and cross-infection success, to that between host relatedness and pathogen sharing in primates (Davies and Pedersen, 2008). We therefore suggest that, although co-inheritance will be important, the shape of the relationship may be determined by host shifts. Nonetheless, disentangling co-inheritance from host shifts should be a focus of future studies but will require detailed information on pathogen phylogeny for many species.

### Phylogenetic Risk of Host Shifts

For each primate species, we calculated a metric encapsulating the summed risk of cross-species pathogen transmission (potential for host shifts) posed by all geographically co-occurring primates. This metric was calculated from the GLM of pathogen community similarity against divergence times of each primate pair. We refer to this metric as the phylogenetic risk of host shifts (PRHS). Phylogenetic risk for species X ( $PRHS_X$ ) is then calculated as:

$$PRHS_X = \sum_{i=1}^S (PRHS_{Y_i})$$

where  $PRHS_X$  is the summation of the phylogenetic risk ( $PRHS_Y$ ) for each  $S$  host species that overlaps in geographic range with the target species X. The phylogenetic risk from each overlapping species (Y) to the target species X ( $PRHS_Y$ ) is derived from the GLM given the phylogenetic relatedness of species X and Y.

### Host Shifts and Pathogen Community Structure

We hypothesize that primate species with high PRHS will share a larger proportion of their pathogens with their geographic neighbors and close relatives. We therefore predict that: (1) pathogen species richness will correlate positively with PRHS, and (2) the proportion of host-specific pathogens will decrease with increasing PRHS.

Our metric (PRHS) assumes that each overlapping species contributes additively to the total risk of a pathogen shifting to the target species. Because the number of primate pathogens is finite, it is possible that the target species could accumulate all pathogens found in the local community of hosts. Although we do not have complete data on the total pathogen community size of any one host or community, the large number of species-specific pathogens (>30% in wild primates; Pedersen et al., 2005) indicates that complete pathogen sharing within a community is unlikely. Our model therefore predicts a positive relationship between the number of overlapping host species and within-host pathogen species richness. Correlates of pathogen species richness in primates have been explored elsewhere (Nunn et al., 2003, Nunn et al., 2005) and may be confounded by variation in sampling intensities (Altizer and Pedersen, 2008). We focused on testing our second prediction, relating to host range, which is the primary focus of this paper.

If pathogen communities are shaped by host shifts, we predict that host species with high PRHS would have fewer host-specific pathogens because they would be shared with other members of the host community. To test this prediction, we constructed a further GLM to evaluate the relationship between PRHS and the proportion of host-specific pathogens in each host's parasite community. Because the number of recorded pathogens varies considerably between hosts (see above), some estimates of host specificity will be derived from very small sample sizes. To explore model sensitivity, we therefore repeated the GLM weighting each datum by the logarithm of the total number of pathogens sampled for each species. Previous studies found that viruses are the least likely to be specific to a single primate host (Pedersen et al., 2005). Therefore, we ran a final GLM of PRHS against the proportion of host specific nonviral pathogens only.

## Geographic Range Overlap and Phylogenetic Risk of Host Shifts

We hypothesize that phylogenetic risk of host shifts might limit geographic co-occurrence between closely related host species. This process would lead to the prediction that the observed  $PRHS_X$  should be lower than expected from a null model of drawing neighbors at random from the regional species pool.

To evaluate whether PRHS across wild primates varies nonrandomly or simply reflects the geographical distribution of host species within regional communities, we compared observed  $PRHS_X$  for each X primate host against a null distribution generated by randomly drawing “neighbors” from the phylogeny. For each primate host, we performed 1,000 random draws, with sample size of neighbors,  $S$ , matching the empirical number of hosts that overlaps with each X species’ range. Because of the large geographic distance separating the New and Old Worlds, New and Old World primates will never have the opportunity to overlap, thus we performed our random draw of neighbors only from within the regional primate communities.

## Geographic Distribution of Phylogenetic Risk—Hotspots for Cross-Species Pathogen Transmission

To generate a geographical representation of the potential for host shifts within wild primates, we first constructed a weighted distribution map of primates, using our PRHS metric as weights. Primate species distributions were obtained from Grenyer et al. (2006). Because the probability of pathogen sharing was influenced only by whether two species were found in sympatry across any part of their range, but not by the magnitude of range overlap (Davies and Pedersen, 2008), we assumed PRHS to be uniform across each target host species range. Spatial data manipulation was performed in ArcView (GIS 3.3, Environmental Systems Research Institute Inc.) and ArcMap (v9.2 Environmental Systems Research Institute Inc.) at a grid cell size of  $1^\circ \times 1^\circ$ . Cell weights ( $cell_{w1}$ ) summed the phylogenetic risk of host shifts across co-occurring species and were calculated as:

$$cell_{w1} = \sum_{i=1}^k (PRHS_{xi})$$

where  $k$  is the number of primate hosts with ranges intersecting with  $cell_w$ . Hotspots represent regions with high diversity of closely related primate species—where we predict host shifts will be most frequent.

Next, to provide an estimate of the cross-species pathogen transmission risk from wild primates to humans, we constructed a second hotspot map, weighting each primate distribution in proportion to its evolutionary distance from humans, using the nonlinear transformation determined from the GLM model coefficients described above. Here, cell weights ( $cell_{w2}$ ) are calculated as:

$$cell_{w2} = \sum_{i=1}^k (PRHS_{hi})$$

where  $PRHS_h$  is the phylogenetic risk of host shifts from species X to *Homo sapiens*.

Our hotspot maps provide a coarse estimate of the potential for host shifts, accounting for opportunity (geography) and biology (phylogeny). However, successful establishment of a pathogen in a novel host will depend on several additional factors, including those that influence contact rates among and within host species (Fenton and Pedersen, 2005; Wolfe et al., 2005). For example, habitat disturbance and transformation, resulting from human growth, may increase the probability of humans encountering new pathogens from wildlife. As a first approximation, we use estimates of human population growth (1990–2000; SEDAC: <http://sedac.ciesin.columbia.edu/gpw/>) to represent human expansion into previously untransformed or isolated habitats, and hence potential for increased human–wildlife contact. We map the product of PRHS to humans ( $cell_{w2}$ ) and an index of human population growth (increase in population density during 10 years) to highlight areas of both high population growth and many closely related primates. Here, cell weights ( $cell_{w3}$ ) are calculated as:

$$Cell_{w3} = cell_{w2} \cdot (\ln[\text{population size 2000}] - \ln[\text{population size 1990}]).$$

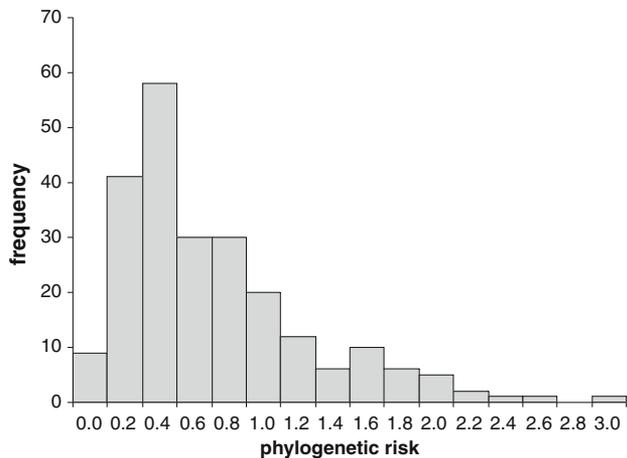
Last, we consider minimum threshold population densities required for sustained transmission and successful establishment of a disease after a host shift and map human population centers (cities > 25,000) on the weighted human risk map. Population centers in close proximity to regions with high phylogenetic risk of host shifts and human population growth are likely to be foci of disease emergence.

## RESULTS

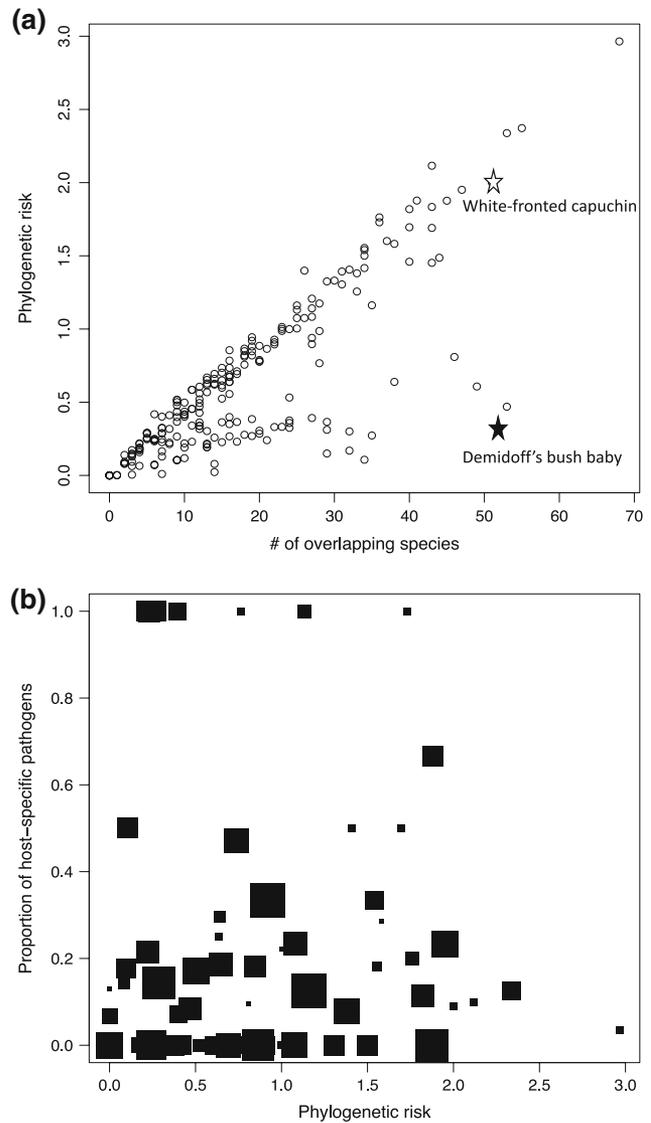
A third of the variation in pathogen communities can be explained by phylogenetic relatedness; more closely related

primate hosts have more similar pathogen communities (GLM of pathogen community similarity against divergence time:  $P$  value  $< 0.01$ ,  $Z = -37.69$ , degrees of freedom ( $df$ ) = 6785, deviance explained = 33%; see also Davies and Pedersen, 2008). Phylogenetic risk of host shifts (PRHS), our metric summarizing the likelihood of pathogens shifting between neighboring hosts, varied considerably across primates (Fig. 2). Nine primates had geographic ranges that overlapped with no other primate species, hence a PRHS of zero, including three species of Old World macaques (Japanese macaque, *Macaca fuscata*; Taiwan macaque, *Macaca cyclopis*; Barbary macaque, *Macaca sylvanus*) and two species of titi monkeys from the New World (dubious titi, *Callicebus dubius*; chestnut bellied titi, *Callicebus caligatus*). Most primates had low-to-intermediate phylogenetic risk (PHRS; 0.2–0.8, median = 0.42).

Phylogenetic risk was highly correlated with the number of geographically overlapping primates ( $R = 0.81$ ,  $N = 233$ ; Fig. 3a), as expected, given that our metric, PRHS, is a summation of all overlapping primates. The brown capuchin (*Cebus apella*), with a geographic range extending from Colombia and Venezuela to Paraguay and northern Argentina, has the greatest PRHS (2.96), and overlaps with more than 60 other New World primate species across its large range. Critically, however, phylogenetic risk of host shifts frequently differs even between primates overlapping with the same number of neighbors. For example, Demidoff’s bushbaby (*Galagoides demidoff*)



**Figure 2.** Frequency histogram of the phylogenetic risk of host shifts (PRHS: summation of the number of overlapping primate species, weighted by the relationship between divergence time and pathogen sharing, see main text for details). Phylogenetic risk varied from 0 to 2.96, with most species having intermediate to low risk (median = 0.42).



**Figure 3.** **a** Scatter plot of phylogenetic risk against the number of overlapping species for each primate host. Risk correlated positively with overlap; however, there was still large variation in risk among species with the same number of neighbors. Variation in phylogenetic risk is a product of both number of neighbors and phylogenetic relationships in the primate tree. **b** Scatter plot of the proportion of host-specific pathogens against phylogenetic risk; symbol size is proportional to the total number of recorded pathogens for each host species respectively. Hosts with high phylogenetic risk have significantly fewer host-specific pathogens, supporting our hypothesis that host shifts are more frequent among high-risk species.

and the white-fronted capuchin (*Cebus albifrons*) both have geographic distributions that overlap with 53 other primate species, but the former has a PRHS of 0.47, and the latter 2.34. Across primates, our randomization procedure indicates a tendency for PRHS to be greater than predicted from a random draw of neighbors (sign test of the number

of times observed PRHS is greater than predicted from random draws:  $P < 0.01$ ), indicating frequent geographical overlap among close relatives.

Primates with higher phylogenetic risk had a significantly lower proportion of host specific pathogens (all pathogens combined), consistent with an increased frequency of host shifts in species with lots of closely related neighbors, but this was found only in the weighted GLM ( $P$  value  $< 0.05$ ,  $Z = -2.13$ ,  $df = 103$ ;  $P = 0.29$  for the unweighted GLM). When excluding viruses, the relationship between PRHS and the proportion of host-specific pathogens is significant with and without model weights (weighted GLM:  $P < 0.01$ ,  $Z = -3.13$ ,  $df = 95$ ; unweighted GLM:  $P < 0.05$ ,  $Z = -2.03$ ,  $df = 95$ ). Unsurprisingly, there was large scatter in the relationship between PRHS and the proportion of host-specific pathogens recorded for each primate host (Fig. 3b), suggesting other factors also are important in shaping pathogen communities. There are much fewer host-specific pathogens for primates with very high PRHS ( $> 2$ ), although the sample size is relatively small. In contrast, primates with low PRHS ( $< 1$ ) have varying proportions of host-specific and host-generalist pathogens (Fig. 3b).

### Hotspot Maps for Cross-Species Pathogen Transmission

We show central Africa and western Amazonia to be hotspots of phylogenetic risk for host shifts between wild primates (Fig. 4a). These regions contain many closely related primate species that overlap in their geographical ranges. We interpret these hotspots to indicate regions where the frequency of pathogen transmission between wild primates is likely to be the highest, and thus we additionally predict increased pathogen community similarity, high within-primate pathogen species richness and a low frequency of host specific pathogens within these localities.

Centers of high phylogenetic risk of host shifts from wild primates to humans include central and western Africa. Areas of low risk include much of Brazil, southern Africa below  $15^{\circ}\text{S}$ , and northern India into Nepal (Fig. 4b). Therefore, despite the high frequency of predicted host shifts between wild primates within Amazonia, it is not identified as a high risk region for shifts to humans. New World primates are more distantly related to humans than Old World primates, and hence pose lower phylogenetic risk.

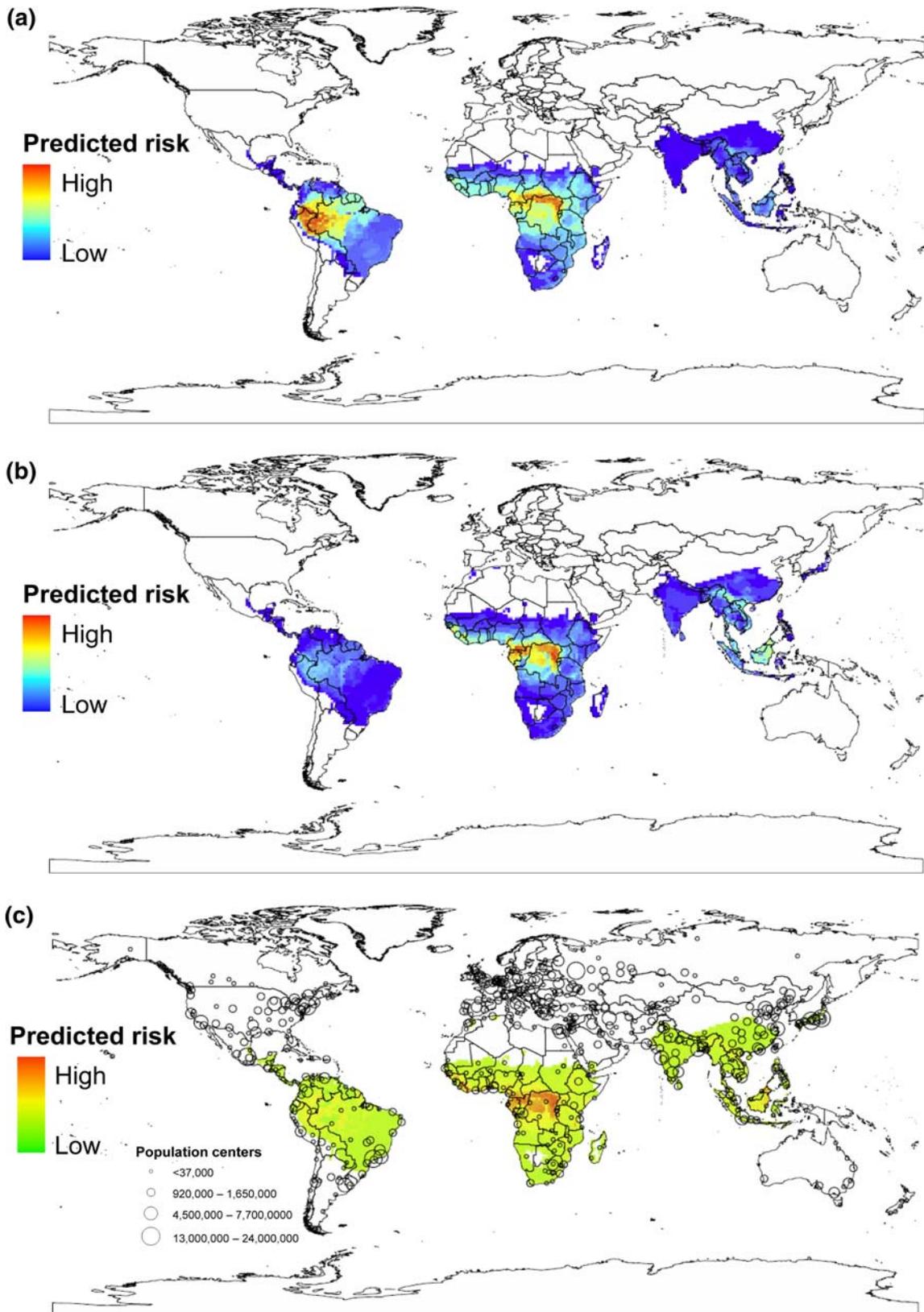
The likelihood of successful establishment and sustained within-host transmission of a novel pathogen will be affected by several ecological and demographic factors. In humans, a critical factor is likely to be human-wildlife contact rates, population connectivity, and human population density. To stimulate discussion, we map one such factor—the intersection between human population growth (as a proxy for increasing human-wildlife contact; see Methods), phylogenetic risk of host shifts to humans, and the distribution of human population centers of size  $> 25,000$  (Fig. 4c). We do not have accurate data to model these terms directly; in particular, it is not clear how the likelihood of disease emergence decays with distance from population centers, or how human movement facilitates disease emergence.

## DISCUSSION

---

Disease emergence is a complex, multistage process, but our results show that we can begin to make predictions about when and where diseases are likely to emerge. By understanding the factors contributing to each stage: opportunity, infection, and ultimately, establishment of a novel pathogen, we can identify which lineages and within which geographic areas disease emergence may be most likely. Previous work has mapped where diseases have emerged in humans in the past (Jones et al., 2008) and which pathogens likely pose the biggest risk of host shifts to humans (Cleaveland et al., 2001). Our study is the first to make predictions about the likelihood of host shifts in the future. Specifically, we use a robust evolutionary relationship between host relatedness and pathogen sharing to quantify the phylogenetic risk of host shifts within primates. We identify species between which host shifts are likely to be frequent, and map geographic areas in which they are distributed. In addition, we generate a hotspot map to highlight regions where humans may be most at risk of host shifts from wild primates. Last, we explore recent demographic trends in human populations that might contribute to disease emergence following successful cross-species transmission.

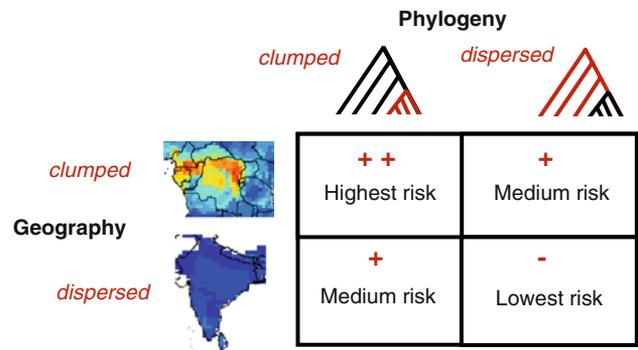
Cross-species pathogen transmission events can heavily impact both human and wildlife populations (Daszak et al., 2000; Woolhouse et al., 2005; Parrish et al., 2008). However, limited data on disease in natural populations, in particular, historical information on when and where a pathogen first moved between host species, has made it



◀ **Figure 4.** **a** Hotspot map of the summed phylogenetic risk for host shifts across wild primates. Central Africa and west Amazonia represent regions of high geographical overlap among many closely related primate species. We predict host shifts among primates to be frequent in these localities. **b** Phylogenetic risk of pathogens host shifting to humans from wild primates. West central Africa is a hotspot of high risk to humans, due to the overlapping ranges of many of our closest relatives. **c** The intersection between high phylogenetic risk and an index of human population growth (increase in density from 1990–2000), revealing regions where we expect high rates of human contact with primate species that pose the greatest risk of cross-species pathogen transmission to humans. Open circles indicate human population centers (> 25,000) that may facilitate disease emergence following host shift events.

difficult to generate predictive models (Hopkins and Nunn, 2007). In a recent paper, we demonstrated that pathogen community similarity among primate species-pairs was correlated closely with their geographical proximity and phylogenetic relatedness (Davies and Pedersen, 2008). Close geographical proximity provides the opportunity for host–host contact and hence pathogen transfer, whereas phylogenetic distance might be indicative of the ease with which a pathogen can infect a novel host given the opportunity. In this paper, we extrapolate the observed relationship between pathogen sharing and host evolutionary relatedness to explore the likely frequency of host shifts within communities of primate hosts. We predict that host shifts will be most frequent among co-occurring and closely related hosts. Therefore, the frequency of host shifts will reflect both the geographical distribution of host species across the landscape, as well as the shape of the phylogenetic tree that connects them (Fig. 5). For example, the white fronted capuchin (*Cebus albifrons*) overlaps in its distribution with many other New World primates and is nested within the recent New World radiation of capuchin monkeys (genus: *Cebus*). We predict the risk of cross-species pathogen transmission to be high for this species. In contrast, Demidoff’s bushbaby (*Galagoides demidoff*) is common across central and western Africa, and also overlaps with many other primates in its distribution, but has only few close relatives. We predict intermediate risk of cross-species pathogen transmission for this species. The lowest risk of cross-species pathogen transmission is predicted for evolutionary distinct species in regions of low primate diversity (Fig. 5).

Our model assumes that primate hosts will differ in their vulnerability to host shifts. We predict that frequent host shifts will increase pathogen species richness and de-



**Figure 5.** Representation of the interaction between phylogeny and geography that may determine the frequency of host shifts within host communities. In the maps, warmer colors represent areas with many overlapping species ranges (geographically clumped distributions). In the phylogeny, red branches represent samples of closely related taxa (phylogenetically clumped). We predict the greatest frequency of host shifts will be observed among species which are both phylogenetically clumped (those with many close relatives) and geographically clumped (species with many overlapping neighbors) (*top left*). Species that are neither geographically clumped nor phylogenetically clumped are predicted to have the lowest risk (*bottom right*). We expect intermediate risk of host shifts for species that are phylogenetically clumped but geographically dispersed, or geographically clumped but phylogenetically dispersed.

crease the proportion of host-specific pathogens per host, as pathogens are shared with an increasing fraction of the host community. However, evaluating pathogen species richness is confounded by sampling artifacts (Altizer and Pedersen, 2008; but see Nunn et al., 2003; Lindenfors et al., 2007), which may vary both taxonomically and geographically (Hopkins and Nunn, 2007). We therefore focus on variation in host specificity. Although it is possible that a systematic geographic bias in the distribution of pathogen species richness (e.g., as suggested for protozoan parasites; Nunn et al., 2005) might additionally influence the frequency of host shifts, we do not believe that this trend would be sufficient to affect the strong geographical patterns found here. The results strongly support our hypothesis: primates with many close phylogenetic relatives and geographical neighbors have significantly fewer host-specific pathogens. This relationship becomes stronger when we exclude viral pathogens, which tend to be less constrained by taxonomic barriers (Pedersen et al., 2005; Davies and Pedersen, 2008). To our knowledge, our study is the first to explore variation in host specificity among within-host pathogen communities across an entire mammalian order. In addition, we suggest that future studies

investigating correlates of pathogen species richness should consider the frequency of host shifts.

Our results suggest that the frequency of host shifts between primates is likely highest in central Africa and western Amazonia—both centers of primate diversity. These findings are important because current estimates indicate that almost 49% of primates are considered threatened (IUCN, 2008), and infectious diseases, especially those that have resulted from a host shift, have been recognized recently as a significant driver of extinction risk in wild animals (Pedersen et al., 2007; Smith et al., 2006, 2009). Primates suffer from multiple threats, including habitat fragmentation and transformation, and direct exploitation (IUCN, 2008), which may increase susceptibility to disease-mediated declines (Smith et al., 2009). Direct exploitation through hunting for bushmeat, necessitating frequent contact between humans and wild primates, is already associated with emerging infectious diseases moving to humans (i.e., SIV, Ebola, STLV-1; Wolfe et al., 2005; Nunn and Altizer, 2006). More recently, there has been increasing awareness of the risk for diseases moving from humans to wildlife populations. In Tai National Park, Ivory Coast endangered chimpanzee populations have suffered dramatic declines from pathogens shared with humans, including three outbreaks from pandemic human viruses (human respiratory syncytial virus (hRSV) and human metapneumovirus (hMPV); Köngden et al., 2008). Significant declines in both chimpanzee and gorilla populations also have been documented across central Africa due to infectious diseases (Leroy et al., 2004). Ominously, it is likely that current risk scenarios are underestimates, based on limited pathogen sampling, especially in wild primates (Hopkins and Nunn, 2007).

Geographic regions with high primate species richness, particularly, high Hominidae diversity, such as west Africa, pose a disproportionate risk of host shifts from wild primates to humans. Notably, Amazonia is not a hotspot for risks of shifts to humans, despite high primate diversity. New World monkeys diverged from Old World monkeys and apes  $\sim 40$ MYA, hence represent only distant evolutionary relatives to humans. Although our model indicates host shifts from wild primates to humans to be most frequent within central and west Africa, the likelihood of a host shift will be affected by additional factors influencing host–host contact rates. Predicting successful host shifts will therefore require information on these additional factors. In this study, we used the rate at which humans are transforming natural habitat, indexed by human popula-

tion growth, as a proxy for encounter rates between wildlife species and their associated pathogens. As human populations grow, encroach into new areas, and transform habitats, we may be at greater risk of contacting new pathogens (Daszak et al., 2000; Dobson and Foufopoulos, 2001; Weiss and McMichael, 2004; Wolfe et al., 2005). This is of particular concern, given the positive correlation between human population growth and our modeled risk of host shifts (Spearman's  $R = 0.328$ ,  $P = 0.049$ , adjusting degrees of freedom to account for spatial autocorrelation), suggesting that in many parts of central Africa, where host shifts are predicted to be frequent, human population growth also is high.

Global emergence of human infectious diseases frequently originate not from the point of first infection but from human population centers, where contact rates among individuals are high and where international travel and dissemination of the pathogen to other cities is facilitated (Woolhouse, 2002; Tatem et al., 2006; Sharp and Hahn, 2008). In epidemiological models, directly transmitted pathogens are not likely to persist in populations below a critical size threshold (a function of both host and pathogen characteristics; Anderson and May, 1992). For example, measles is thought to go locally extinct in populations comprising less than 250,000 individuals (Keeling and Grenfell, 1997). Many host shift events may therefore never become established. However, large population centers in or near areas of rapid human expansion can provide the demographic component necessary to convert a host shift into a global emergence event. Here, we map the distribution of the major human population centers across the globe. Population centers that have the potential to facilitate the emergence of new diseases shifting from wild primates to humans are those in west and east Central Africa, including: Kinshasa, Democratic Republic of Congo; Brazzaville, Congo; and Yaounde, Cameroon to the west, and Kampala, Uganda; Kigali, Rwanda; and Nairobi, Kenya to the east. Infectious disease detection and control, before emergence events, should focus on these regions.

Possibly the world's most devastating pandemic, HIV-1, provides a remarkable case study, closely matching our model predictions. Resulting from a host shift from chimpanzees, one of our closest living relatives (Korber et al., 2000), HIV-1, was first described in 1981, but was introduced to human populations several times during the past 100 years—two of which were from SIVcpzPtt (simian immunodeficiency virus) of chimpanzees (*Pan troglodytes troglodytes*) from west central Africa (Korber et al., 2000).

One shift led to the global AIDS pandemic (HIV-1 group M), and the other resulted in a localized endemic in Cameroon (HIV-1 group N) (Keele et al., 2006). Although SIV has been found in more than 30 African primate species, chimpanzees are the only species that harbor the strain most closely linked to pandemic HIV-1 (Sharp and Hahn, 2008; Keele et al., 2006). Recent analysis of HIV-1 group M suggests that several variants were circulating in the 1960s, and molecular analysis now proposes that HIV-1 likely emerged from *SIVcpzPtt* at the turn of the 20<sup>th</sup> century (1902–1921) but had a period of relatively slow growth during the next 50 years (Worobey et al., 2008). It is thought that HIV-1 group M was first transmitted to humans from wild chimpanzees in the southwest corner of Cameroon (Keele et al., 2006), and its establishment and spread was likely facilitated by the rise of cities, specifically Kinshasa, Democratic Republic of Congo (formerly Leopoldville, Zaire). In 1910, no population centers in the area of the cross-species transmission event were greater than 10,000 people; however, Kinshasa grew rapidly as a trading center and had a population between 100,000–400,000 during the middle of the twentieth century when HIV-1 can be found in human samples (Sharp and Hahn, 2008; Worobey et al., 2008). Kinshasa is 700 km from the location of the original host shift event; however, it was a key transportation hub through river travel (Sharp and Hahn, 2008). The successful emergence of HIV occurred in our predicted hot spot areas of high human risk. This example highlights the complex multistage process that drives disease emergence: a region with high risk of shifts to humans; a host shift from our closest relative; delayed emergence after the local growth of the human population at a transportation center.

Whereas human population density and encroachment into new habitats may affect the likelihood of the successful emergence of a primate pathogen into humans, the risk map of human disease emergence also highlights geographic locations where wild primates may be at the greatest risk of contracting human diseases. Many pathogens from humans have been found in wild primate populations (e.g., polio, syphilis, influenza A, and measles) and pose a significant threat to their persistence (Nunn and Altizer, 2006). Almost half of wild primates face the risk of extinction (IUCN, 2008) and pathogens have been linked to substantial population declines (Formenty et al., 1999, Leroy et al., 2004, K ngden et al., 2008); understanding the patterns driving cross-species transmission and host shifts in wild populations will therefore be crucial for their

conservation. Within wildlife populations, ecological factors may be more important in determining rates of cross-species transmission and host shifts within wild primates. Host life history, ecological niche, feeding and activity patterns, sexual behavior, group size, and territoriality are strong determinants of pathogen transmission rates within a host population (Altizer et al., 2003a; Nunn and Altizer, 2006), and therefore may be important for determining when and to whom cross-species transmission events are likely to be successful. Although a comprehensive exploration of these ecological factors is beyond our current scope, knowledge of how host ecology is likely to interact with pathogen ecology, host phylogeny, and geography will be key for developing a predictive framework of wildlife disease emergence.

In our analyses, we focused on primates because they are perhaps the best studied animal group and because of their close evolutionary relationship to humans. Primates represent a significant disease reservoir for humans; approximately 25% of zoonotic human EIDs are shared with primate host species; however, other taxa (e.g., bats, rodents, birds, and domestic animals) also pose a threat of host shifts to humans (Woolhouse and Gowtage-Sequeria, 2005). Nonetheless, we suspect that our results, specifically that phylogenetic tree shape and geographic dispersion are likely to affect the phylogenetic risk of a host shift, will likely extrapolate across taxa (see Gilbert and Webb, 2007). Last, we note that the relationship between pathogen taxonomy, host phylogeny, and host geography may be complex and vary across the different pathogen groups. For example, we might predict a greater frequency of host shifts among more distant relatives for viruses than for helminths, which suggests that the risk of viral emergence may be even greater than predicted in our framework.

Although data on pathogens from natural populations is incomplete, and there may be many factors that affect the likelihood of a host shift, our results suggest that a broad biogeographical approach can help predict cross-species pathogen transmission. By combining information across multiple species and communities, it is possible to detect general trends that are not apparent from single case studies. In contrast with traditional epidemiological studies, our analysis explores macroevolutionary and macroecological patterns in the distribution of pathogens across species and the landscape. Our study makes direct predictions about how much variation is expected in pathogen species richness among hosts and the frequency of which host specific pathogens are likely to be distributed across

hosts. Detailed species-specific data on both the presence and absence of a pathogen will be essential for evaluating model predictions and for future analyses.

## ACKNOWLEDGMENTS

The authors thank K. Smith and S. Scheiner for the invitation to participate in the symposium “The Biogeography of Disease: Examining the Forces that Drive Disease Distributions and Emergence” at the 2009 International Biogeography Society biennial meeting. A.B.P. was funded by a Royal Society Incoming Research Fellowship and a Wellcome Trust Centre of Infection, Immunity & Evolution Advanced Fellowship. T.J.D. was supported as a postdoctoral associate at the National Center for Ecological Analysis and Synthesis, a center funded by National Science Foundation grant DEB-0072909, and the University of California, Santa Barbara.

## REFERENCES

- Altizer S, Nunn CL, Thrall P, Gittleman JL, Antonovics J, Cunningham A, et al. (2003a) Social organization and disease risk in mammals: integrating theory and comparative studies. *Annual Review of Ecology, Evolution and Systematics* 34:517–547
- Altizer S, Harvell D, Friedle E (2003b) Rapid evolutionary dynamics and threats to biodiversity. *Trends in Ecology and Evolution* 18:589–596
- Altizer S, Pedersen AB (2008) Host–pathogen evolution, biodiversity and disease risks for natural populations. In: *Conservation Biology: Evolution in Action*, Carroll SP, Fox CW (editors), New York, NY: Oxford University Press, pp 259–277
- Anderson RM, May RM (1992) *Infectious Diseases of Humans: Dynamics and Control*, New York, NY: Oxford University Press
- Antonovics J, Hood M, Partain J (2002) The ecology and genetics of a host shift: *Microbotryum* as a model system. *American Naturalist* 160:S40–S53
- Bininda-Emonds ORP, Cardillo M, Jones KE, MacPhee RDE, Beck RMD, Beck RMD, et al. (2007) The delayed rise of present-day mammals. *Nature* 446:507–512
- Daszak P, Berger L, Cunningham AA, Hyatt AD, Green DE, Speare R (1999) Emerging infectious diseases and amphibian population declines. *Emerging Infectious Diseases* 5:735–748
- Daszak P, Cunningham AA, Hyatt AD (2000) Emerging infectious diseases of wildlife: threats to biodiversity and human health. *Science* 287:443–449
- Davies TJ, Pedersen AB (2008) Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proceedings of the Royal Society London B* 275:1695–1701
- Cleaveland S, Laurenson ML, Taylor LH (2001) Disease of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philosophical Transactions of the Royal Society of London B* 356:991–999
- Dobson A, Foufopoulos J (2001) Emerging infectious pathogens of wildlife. *Philosophical Transactions of the Royal Society of London B* 356:1001–1012
- Fenton A, Pedersen AB (2005) Community epidemiology in theory and practice: A conceptual framework for describing transmission dynamics in multiple hosts. *Emerging Infectious Disease* 11:1815–1821
- Formenty P, Boesch C, Wyers M, Steiner C, Donati F, Dind F, et al. (1999) Ebola virus outbreak among wild chimpanzees living in a rain-forest of Cote d’Ivoire. *Journal of Infectious Diseases* 179(S1):120–126
- Gilbert GS, Webb CO (2007) Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Science USA* 104:4979–4983
- Grenyer R, Orme CDL, Jackson SF, Thomas GH, Davies RG, Davies TJ, et al. (2006) The global distribution and conservation of rare and threatened vertebrates. *Nature* 444:93–96
- Gupta S, Ferguson N, Anderson RM (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280:607–614
- Hahn BH, Shaw GM, DeCock KM, Sharp PM (2000) AIDS as a zoonosis: scientific and public health implications. *Science* 287:607–614
- Hopkins ME, Nunn CL (2007) A global gap analysis of infectious agents in wild primates. *Diversity and Distribution* 13:561–572
- Hughes JM, Peters CJ, Cohen ML, Mahy BWJ (1993) Hantavirus pulmonary syndrome—An emerging infectious disease. *Science* 262:850–851
- IUCN, International Union for the Conservation of Nature (2008) IUCN Red List of threatened species. <http://www.redlist.org>
- Jones KE, Patel N, Levy M, Storeygard A, Balk D, Gittleman JL, et al. (2008) Global trends in emerging infectious diseases. *Nature* 451:990–993
- Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, et al. (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 313:523–526
- Keeling MJ, Grenfell B (1997) Disease extinction and community size: modeling the persistence of measles. *Science* 275:65–67
- King DA, Peckham C, Waag JK, Brownlie J, Woolhouse MEJ (2006) Infectious diseases: preparing for the future. *Science* 313:1392–1393
- Köngden S, Kühl H, N’Goran PK, Walsh PD, Schenk S, Ernst N, et al. (2008) Pandemic human viruses cause decline of endangered great apes. *Current Biology* 18:260–264
- Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, et al. (2000) Timing the ancestor of HIV-1 pandemic strains. *Science* 288:1789–1796
- Leroy EM, Rouquet P, Formenty P, Souquiere S, Kilbourne A, Froment JM, et al. (2004) Multiple Ebola virus transmission events and rapid decline of Central African wildlife. *Science* 303:387–390
- Lindenfors P, Nunn CL, Jones KE, Cunningham AA, Sechrest W, Gittleman JL (2007) Parasite species richness in carnivores: effects of host body mass, latitude, geographic range and population density. *Global Ecology and Biogeography* 16:496–509
- Morand S, Manning SD, Woolhouse MEJ (1996) Parasite-host coevolution and geographic patterns of parasite infectivity and host susceptibility. *Proceedings of the Royal Society London B* 263:119–128

- Nunn CL, Altizer SM, Jones KE, Sechrest W (2003) Comparative tests of parasite species richness in primates. *American Naturalist* 162:597–614
- Nunn CL, Altizer SM (2005) The global mammal parasite database: an online resource for infectious disease records in wild primates. *Evolutionary Anthropology* 14:1–2
- Nunn CL, Altizer SM, Sechrest W, Cunningham AA (2005) Latitudinal gradients of parasite species richness in primates. *Diversity and Distributions* 11:249–256
- Nunn CL, Altizer S (2006) *Infectious Diseases in Primates: Behaviour Ecology and Evolution*, New York, NY: Oxford University Press
- Oldroyd BP (1999) Coevolution while you wait: *Varroa jacobsoni*, a new parasite of western honey bees. *Trends in Ecology and Evolution* 14:312–315
- Osterhaus A (2001) Catastrophes after crossing species barriers. *Philosophical Transactions of the Royal Society of London B* 356:791–793
- Parrish CR, Holmes EC, Morens DM, Park E, Burke DS, Calisher CH (2008) Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews* 72:457–470
- Pedersen AB, Poss M, Nunn CL, Cunningham AA, Altizer S (2005) Patterns of host specificity and transmission among parasites of free-living primates. *International Journal for Parasitology* 35:647–657
- Pedersen AB, Jones KE, Nunn CL, Altizer S (2007) Infectious diseases and extinction risk in wild mammals. *Conservation Biology* 21:1269–1279
- Perlman SJ, Jaenike J (2003) Evolution of multiple components of virulence in *Drosophila*-nematode associations. *Evolution* 57:1543–1551
- Pfennig DW (2000) Effect of predator-prey phylogenetic similarity on the fitness consequences of predation: a trade-off between nutrition and disease? *American Naturalist* 155:335–345
- Ricklefs RE, Fallon SM (2002) Diversification and host switching in avian malaria parasites. *Proceedings of the Royal Society London B* 269:885–892
- Roelke-Parker ME, Munson L, Packer C, Kock R, Cleaveland S, Carpenter M, et al. (1996) A canine distemper virus epidemic in Serengeti lions (*Panthera leo*). *Nature* 376:441–445
- Sharp PM, Hahn BH (2008) Prehistory of HIV-1. *Nature* 455:605–606
- Smith KF, Sax DF, Lafferty KD (2006) Evidence for the role of infectious disease in species extinction. *Conservation Biology* 20:1349–1357
- Smith KF, Acevedo-Whitehouse K, Pedersen AB (2009) The role of infectious diseases in contemporary and future wildlife extinctions. *Animal Conservation* 12:1–12
- Taylor LH, Latham SM, Woolhouse MEJ (2001) Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society of London B* 356:983–989
- Tatem AJ, Hay SI, Rogers DJ (2006) Global traffic and disease vector dispersal. *Proceedings of the National Academy of Science USA* 103:6242–6247
- Walsh PD, Abernethy KA, Bermejo M, Beyersk R, De Wachter P, Akou ME, et al. (2003) Catastrophic ape decline in western equatorial Africa. *Nature* 422:611–614
- Weiss RA, McMichael AJ (2004) Social and environmental risk factors in the emergence of infectious diseases. *Nature Medicine* 10:S70–S76
- WHO, World Health Organization (2006) AIDS Epidemic Update 2005, Joint United Nations Programme on HIV/AIDS (UNAIDS) and WHO [http://data.unaids.org/pub/EpiReport/2006/2006\\_EpiUpdate\\_en.pdf](http://data.unaids.org/pub/EpiReport/2006/2006_EpiUpdate_en.pdf)
- WHO, World Health Organization (2007) Financial resources required to achieve universal access to HIV prevention, treatment, care and support. Joint United Nations Programme on HIV/AIDS (UNAIDS) [http://data.unaids.org/pub/Report/2007/20070925\\_advocacy\\_grne2\\_en.pdf](http://data.unaids.org/pub/Report/2007/20070925_advocacy_grne2_en.pdf)
- Williams ES, Yuill T, Artois A, Fischer J, Haigh SA (2002) Emerging infectious diseases in Talyorwildlife. *Revue Scientifique et Technique de l'OIE* 21:139–157
- Wolfe ND, Daszak P, Kilpatrick AM, Burke DS (2005) Bushmeat hunting, deforestation, and prediction of zoonoses emergence. *Emerging Infectious Diseases* 11:1822–1827
- Woolhouse ME (2002) Population biology of emerging and re-emerging pathogens. *Trends in Microbiology* 10:S3–S7
- Woolhouse MEJ, Haydon DT, Antia R (2005) Emerging pathogens: the epidemiology and evolution of species jumps. *Trends in Ecology and Evolution* 20:239–244
- Woolhouse MEJ, Gowtage-Sequeria S (2005) Host range and emerging and re-emerging pathogens. *Emerging Infectious Diseases* 11:1842–1847
- Worobey M, Gemmel M, Teuwen DE, Haselkorn T, Kunstman K, Bunce M, et al. (2008) Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature* 455:661–664