Diversity and Distributions, (Diversity Distrib.) (2007) **13**, 561–572

BIODIVERSITY RESEARCH

A global gap analysis of infectious agents in wild primates

Mariah E. Hopkins[1]* and Charles L. Nunn[2,3]

[1]*Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720-3114, USA,* [2]*Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany,* [3]*Department of Integrative Biology, Berkeley, University of California, CA 94720-3140, USA*

*Correspondence: Mariah E. Hopkins, Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720-3114, USA. Tel.: 510-643-6992; Fax: 510-643-6992. E-mail: mhopkins@nature.berkeley.edu

ABSTRACT

A number of infectious diseases have emerged as threats to humans and wildlife. Despite the growing importance of georeferenced data for mitigating disease risk, information on parasite threat is patchily distributed at a global scale. In this paper, we explore the utility of gap analysis techniques to investigate the global geographical distribution of parasite sampling in non-human primates. Specifically, we identify geographical areas that are undersampled for parasites in relation to primate geographical distributions, primate taxonomic sampling, primate threat status, and parasite taxonomy. Our results reveal that East Asia (particularly China), South-East Asia, and the South American Amazon are the most deficient in sampling effort with respect to all criteria. We also identify sampling gaps based on several criteria in West and Central Africa. Future research aimed at filling these gaps is needed for both human health and primate conservation purposes.

Keywords

GIS, infectious disease, parasites, primates, sampling gaps, zoonoses.

INTRODUCTION

A number of infectious agents currently threaten both human and wildlife populations (Daszak *et al.*, 2000; Cleaveland *et al.*, 2001; Taylor *et al.*, 2001; Woolhouse & Gowtage-Sequeria, 2005). For example, recent outbreaks of Ebola haemorrhagic fever have drastically impacted populations of African apes (Walsh *et al.*, 2003; Leroy *et al.*, 2004; Bermejo *et al.*, 2006), and the global decline of amphibian populations has been attributed to disease outbreaks caused by a chytrid fungus (Berger *et al.*, 1998; Daszak *et al.*, 1999). Many of the diseases that threaten wildlife also cause significant threats to humans. In fact, approximately 60% of all infectious organisms known to infect humans — and 75% of all emerging pathogens — are zoonotic (Taylor *et al.*, 2001; Woolhouse *et al.*, 2001). Thus, understanding the factors that influence the geographical distribution of infectious agents in wildlife is important not only for animal conservation, but also for preventing disease outbreaks in humans.

Despite the growing importance of georeferenced data for mitigating disease threats to wildlife and humans (Smith *et al.*, 2002; Ostfeld *et al.*, 2005; Rouquet *et al.*, 2005), information on parasitism is patchily distributed at a global scale, with some areas better sampled than others. For example, countries that experience political instability generally will attract less interest from field biologists. Similarly, species living in relatively inaccessible areas, such as high mountain ranges or swamp forests, are less likely to be sampled for parasites than those that are found in locations closer to areas of high human density, established field sites and efficient transportation networks. While sampling gaps have been recognized and addressed in comparative studies of parasite ecology, such as patterns of parasite species richness (Gregory, 1990; Walther *et al.*, 1995; Poulin, 1998; Walther & Morand, 1998; Nunn *et al.*, 2003), studies of spatial epidemiology have yet to address similar issues when studying the geographical distribution of disease risk.

In this paper, we develop a method for identifying geographical gaps in our knowledge of global disease distributions, and for prioritizing the geographical areas most in need of further sampling effort. We then apply this method to identify sampling deficiencies in our understanding of parasites in non-human primates. Primates are an important group of mammals for considering these questions. Over 60% of primates are threatened (Hilton-Taylor, 2002), and infectious disease has caused devastating mortality in diverse primate species (Leendertz *et al.*, 2004; Chapman *et al.*, 2005; Bermejo *et al.*, 2006; Nunn & Altizer, 2006). Understanding infectious disease in non-human primates also provides important insights into many human diseases, partly because primates harbour many infectious agents that can spread to humans (Wolfe *et al.*, 1998; Peeters *et al.*, 2002; Woolhouse & Gowtage-Sequeria, 2005). For example, a recent study found that approximately 25% of the 415 parasites found in non-human primates also infect humans (Pedersen *et al.*, 2005), and disease outbreaks in human populations have previously been associated with human–wildlife contact (Rouquet *et al.*, 2005). Thus, understanding geographical biases in our knowledge of primate parasites is important for mitigating future disease-related threats to humans.

DOI: 10.1111/j.1472-4642.2007.00364.x
www.blackwellpublishing.com/ddi

The methodological approach developed here is based on the concept of gap analysis — a technique used in conservation biology to both outline geographical areas needed to protect threatened species and to identify areas with a large number of undescribed species for biodiversity surveys (Scott *et al*., 1993; Jennings, 2000; Funk *et al*., 2005). We used this same basic approach to identify areas that should be prioritized for parasite sampling, comparing and contrasting results when incorporating information on primate distributions (geographical ranges), host taxonomic sampling, host threat status, and parasite taxonomy.

METHODS

Primate parasite records

To investigate patterns of parasitism in primates, we obtained data from the Global Mammal Parasite Database (GMPD), which is an ecological informatics database that was constructed using information from the published literature on micro- and macroparasites in free-living primates (Nunn & Altizer, 2005). The GMPD was generated using bibliographic search engines covering both European and non-European languages and currently provides the most complete list of the parasites documented in primates. However, it should be considered a sample of the existing literature; in rare cases, an existing study of primate parasites may be missing from the database, and studies of parasites in primates may have been published since the database was finalized and first posted online (http://www.mammalparasites.org).

For each record of a host–parasite combination in the database, we identified records in which authors of the study provided information on the sampling locality or a nearby landmark that could be used to identify geographical coordinates. When latitude and longitude were not provided in the paper, we searched for these coordinates in online resources, especially the Geographic Names Database (National Geospatial-Intelligence Agency, 2005), and in other papers that studied hosts or parasites at the same locality. We excluded publications that identified localities only to the general region, state/province or country level. When it was not possible to identify geographical coordinates, or when searches resulted in multiple localities in obviously different areas, we eliminated that host–parasite record from the analysis. Because we were interested in where sampling effort was placed, we included localities where parasites were searched for but not found.

Primate ranges

Information on primate geographical ranges was obtained using geographically rectified range maps for each of the world's primate species, as compiled in a geographical information system (GIS) and used in previous studies of parasite richness in primates (Sechrest *et al*., 2002; Nunn *et al*., 2003, 2005). The host taxonomy used for the parasite database was developed around the taxonomy of Corbet & Hill (1992), while the taxonomy for the geographical range data followed Groves (1993). In combining these taxonomies, host lineages were generally split into more recognized species for the geographical range data than for data in the GMPD. We therefore combined geographical ranges of species in the range database that were recognized as synonyms for a species in the GMPD. Baboons (genus *Papio*) were an exception to this rule, with the range maps recognizing only one species of *Papio* (*Papio hamadryas*), compared to multiple species of *Papio* in the GMPD. In this case, we combined the parasite data into one record, which we identified as *P. hamadryas*. When lumping several synonymous species into a single species, we consulted the IUCN Red List (2006) to assign this species' threat status. If the synonymous species differed in their IUCN threat status score, we assigned the resulting combined species the highest of these scores.

After taxonomic information on geographical ranges was matched to the parasite database, we entered the data into a GIS (ARCGIS 8.3; Environmental Systems Research Institute, Redlands, CA, USA) and made one further check of the parasite data in relation to host geographical ranges. Specifically, we identified sampling localities for hosts that resided outside of the recognized range. We then referred to the publications and corrected the database appropriately, retaining the data (and fixing the host geographical range data) only when we had strong evidence that the locality was correct.

Identification of sampling gaps

We developed four goals for analysing the geographical data on parasite sampling:

Quantify geographical biases with respect to host diversity

In our first analysis, we investigated whether geographical biases in the distribution of sampling localities exist in relation to host diversity. Simply plotting the occurrence of parasite samples provides some information on where past effort has been placed (Fig. 1a; see Appendix S1 in Supplementary Material for a list of countries where primates have not been sampled for parasites at georeferenced locations using this version of the database). However, these plots tell us little about the distribution of primates, which is important because gaps can only exist in relation to where members of a host species occur. Thus, all sampling localities were mapped in ARCGIS 8.3 and examined in relation to the distribution and densities of free-living primate species. Grids of several resolutions (1, 4, 9, 12, and 25 square-degrees) were overlaid on Africa, Asia, and the Americas and used to calculate the total number of primate species and the number of sampling localities per grid cell. Ordinary Least Squares (OLS) and spatial autoregression models were used to determine if the number of sampling localities per cell was associated with the number of primate species present in that cell. Resulting residuals were mapped using ARCGIS, with negative residuals indicating sampling gaps.

We also performed an analysis based on quantile subtraction. A neighbourhood contour analysis using Cook's criteria was applied to the map of sampling localities for spatial smoothing. In this analysis, a moving window was applied in which the value of each grid cell was designated as the mean value of all cells within a distance of two cells from the focal cell. Cell values for mean smoothed densities of sampling localities and the original

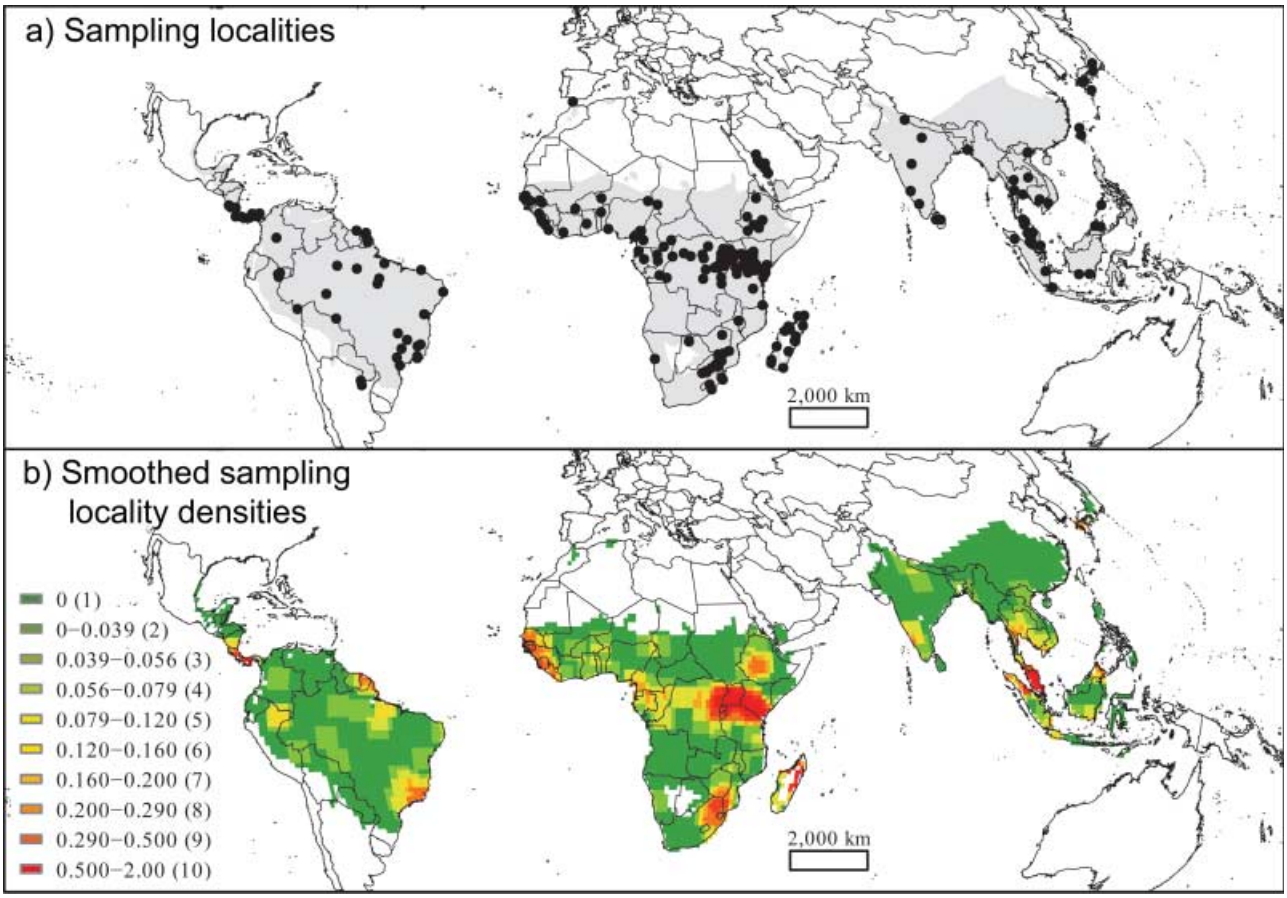

Figure 1 Global distribution of sampling localities: (a) Sampling localities at which at least one primate species has been sampled for any type of parasite; (b) Smoothed sampling densities measured as the mean number of sampling localities per 1 square-degree cell, with quantile values in parentheses. Red values indicate areas with the highest sampling effort. Maps are displayed using Aitoff's projection at a scale of 1 : 200,000,000.

primate species density map were divided into 10 quantiles. For each grid cell, the quantile value (ranging from 1 to 10) for parasite sampling density was subtracted from the quantile value for primate abundance (also ranging from 1 to 10), the primary assumption being that the highest sampling effort should occur in areas with the highest primate species diversity in order to accurately capture the full range of parasites that infect non-human primates. Individual cell difference values were then mapped according to how many standard deviations this difference value fell from the mean difference value for all cells with at least one primate species. Negative values indicate areas that are relatively oversampled in relation to primate abundance, while positive values indicate that areas are relatively undersampled in relation to primate abundance.

Quantify geographical gaps with respect to primate taxonomic sampling

Although the first goal identifies geographical biases in the distribution of sampling localities in relation to host diversity, it does not incorporate information on which primates were sampled at localities. To account for host taxonomy, a list of all unsampled primate species was generated (see Appendix S2 in Supplementary Material), as was a list of all unsampled primate species for which no other primate has been sampled within their entire range. The ranges of the resulting species from both lists were mapped individually and the density of these species per grid cell was calculated. The proportion of unsampled species in each grid cell was calculated by dividing the number of unsampled species per grid cell by the number of species present in that cell, and the results were mapped.

To obtain a quantitative measure of the number of species needed to be sampled in a given cell in order to reach mean sampling levels, we calculated the 'sampling gap' as $S_i = (m–p_i)t_i$, where m is the mean proportion of sampled species across all cells, p is the proportion of sampled species in cell i, and t is the total species present in cell i. Positive values of S_i indicate the number of species that need to be sampled in order to reach the worldwide mean proportion of species sampled per cell. Negative values indicate the number of species sampled above mean proportions. A list of protected areas with the highest S_i values (see Appendix S3 in Supplementary Material) was also generated by overlaying a map of world protected areas (UNEP-WCMC, 2005) upon the analysis grid.

Quantify geographical gaps with respect to host threat status

Threatened species of primates could be especially vulnerable to disease threats, particularly spill-over of generalist parasites that

can be harboured in more abundant non-threatened mammalian hosts (McCallum & Dobson, 1995; Daszak *et al.*, 2000; Gog *et al.*, 2002). Incorporating host threat status is important for identifying the most critical geographical areas and species to sample. To incorporate threat status, each species in our analyses was cross-referenced with the IUCN 2006 Red List, and its conservation status classified as: 1 — Not Listed (i.e. species of least concern or those not evaluated); 2 — Threatened (includes 'near threatened'); 3 — Vulnerable; 4 — Endangered; or 5 — Critically Endangered. The analyses described for Goals 1 and 2 were then repeated for all primate species with IUCN classifications of 2 and higher. Endangered species were also considered separately by mapping the density of unsampled 'endangered' primate species (classifications of 4 and 5) per grid cell.

Quantify geographical gaps with respect to parasite taxonomy

Certain types of parasites, such as bacteria, protozoa, and viruses, are more likely to be emerging zoonotic diseases (Taylor *et al.*, 2001; Woolhouse & Gowtage-Sequeria, 2005), raising the question as to whether sampling for particular parasite types is biased geographically. We conducted analyses to address this question by separating each geographically referenced parasite record by taxonomic group (arthropods, bacteria, fungi, helminths, protozoa, and viruses). We then mapped and compared locality patterns across continents for each parasite group. In addition, we separated all records in which a primate was sampled for a zoonotic or emerging zoonotic disease and mapped the smoothed mean density of these localities per cell.

RESULTS

Distribution of sample points

A total of 1473 records with unique host-parasite-author-location combinations could be georeferenced, constituting 43.9% of all records present in the GMPD. Sampling took place at 248 different locations, with the most sampling effort concentrated in Africa and the least effort in Asia (Fig. 1a). When smoothed neighbourhood values were considered (Fig. 1b), East Africa emerged as the region with the highest sampling effort (up to nine sites per 1 square-degree area), while large portions of East and South-East Asia, south-west Africa, and the South American Amazon remain unsampled. These patterns were consistent when using different grid resolutions; thus, all subsequent results are presented for a 5017 grid-cell network of 1 square-degree resolution.

Identification of sampling gaps

Quantification of geographical biases with respect to host distributions

Primate species density ranges from one primate species per 1 square-degree cell in Mexico and Japan, to 23 primate species per 1 square-degree cell in Central Africa and interior portions of the Amazon (Fig. 2a). In total, 87 of the 189 primate species considered have been sampled at georeferenced locations in the GMPD, with 48.0% of Old World primate species and 43.5% of New World primate species sampled for at least one parasite at some location across their geographical ranges. Among Old World primates, African primates have received more than twice as much attention as Asian primates (61.3% of African species have been sampled vs. 28.8% of Asian species). The best-studied primate genus is *Papio*, which accounts for 36.4% of all georeferenced sampling records in GMPD (537 of 1473 records, representing 69 localities).

OLS regression analyses revealed a significant positive relationship between the number of primate species per cell and the number of sampling localities per cell ($b = 0.0052$, $P < 0.001$ for 1 square-degree resolution), but the overall explanatory value of the model was poor ($r^2 = 0.0081$) and residuals from this model demonstrated significant spatial autocorrelation (Lagrange Multiplier = 173.2, $P < 0.001$, neighbourhood matrix based on first-order Rook Contiguity). Spatial autoregressive models accounting for neighbourhood trend effects had slightly better overall explanatory power (e.g. SAR: $b = 0.0067$, $P < 0.001$, $R^2 = 0.05$), but 95% of the variation in sampling effort still could not be explained by the number of host species present in a given area.

When primate densities were separated into 10 quantiles and compared to smoothed sampling locality quantiles, areas with high primate species diversity often failed to correspond to areas with high densities of parasite sampling in the GMPD. This effect is easily seen by comparing Figs 1b and 2a. When quantile subtraction was performed (Fig. 2b), areas with the highest overall primate species density (i.e. West/Central Africa and the Amazon) remained those with the greatest discrepancy between species density and sampling point density. Thus, areas with the greatest primate diversity are understudied with respect to parasites.

Quantification of geographical gaps with respect to primate taxonomy

When the density of unsampled primate species was calculated across the world, large portions of the Amazon, West-Central Africa, and South-East Asia demonstrated the highest values (Fig. 2c). In these areas, the number of primate species that needs to be sampled to reach mean levels can be as high as six species (Fig. 2d). In addition, in areas of Asia and South America, many species still exist for which no primate of any species has been sampled for parasites at a georeferenced locality within their entire range in the GMPD (Fig. 3a), and in many areas, the proportion of unsampled species in a given cell reaches 100% (Fig. 3b).

Quantification of geographical gaps with respect to host threat status

Of the unsampled primates, 65.7% are present on the IUCN Red List, while only 46.0% of sampled species have any type of threat status. The odds of being sampled were significantly greater for non-threatened primates (1.34) than for threatened primates

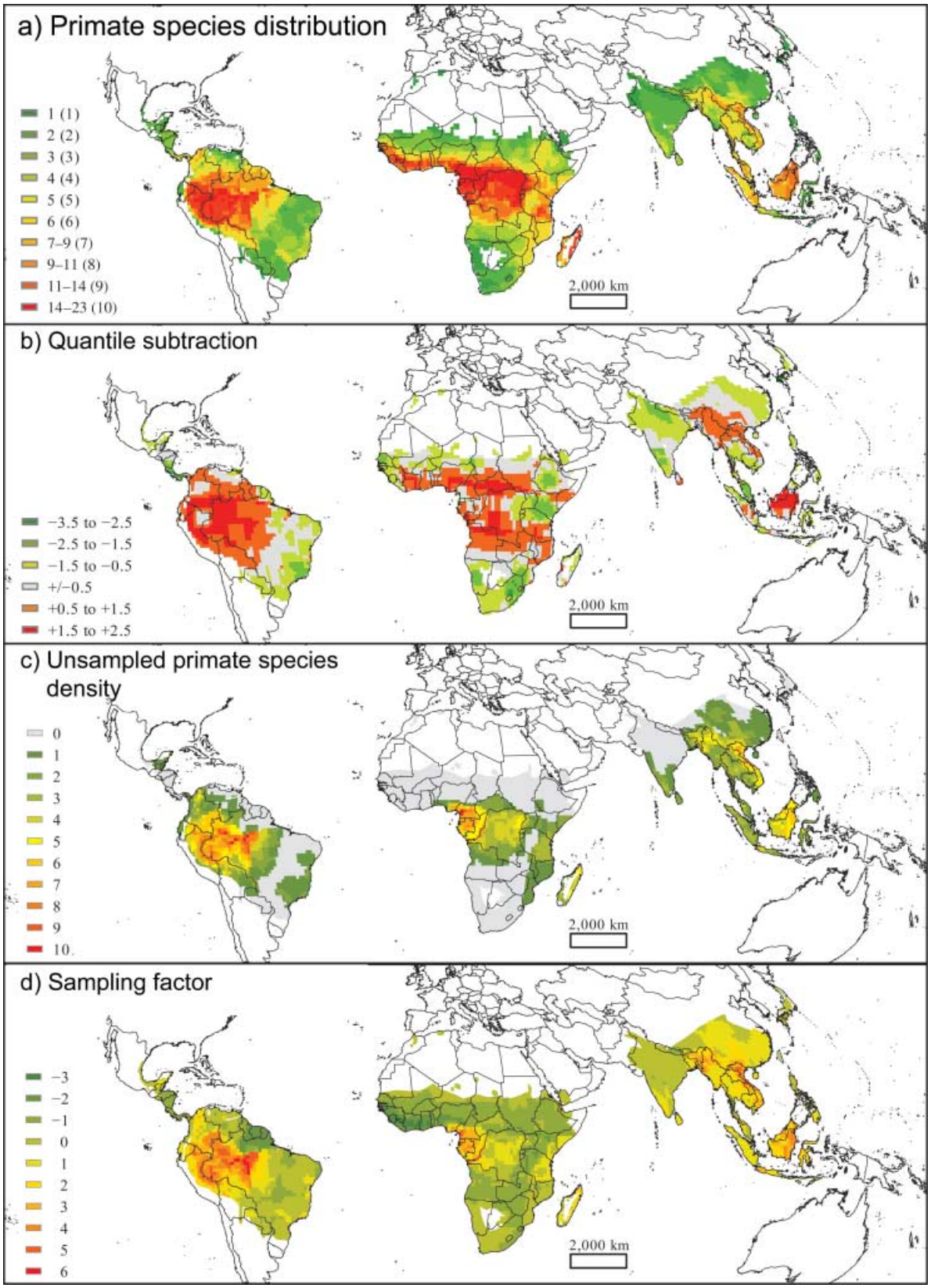

Figure 2 Quantification of geographical biases with respect to host diversity and primate taxonomic sampling: (a) Species density of primates, with quantile values in parentheses; (b) Cell differences in quantile values between primate species density and smoothed sampling locality densities, classified according to standard deviations from the means; (c) Number of unsampled primate species; (d) The number of primate species that needs to be sampled per cell in order to reach mean sampling levels. Maps are displayed using Aitoff's projection at a scale of 200,000,000.

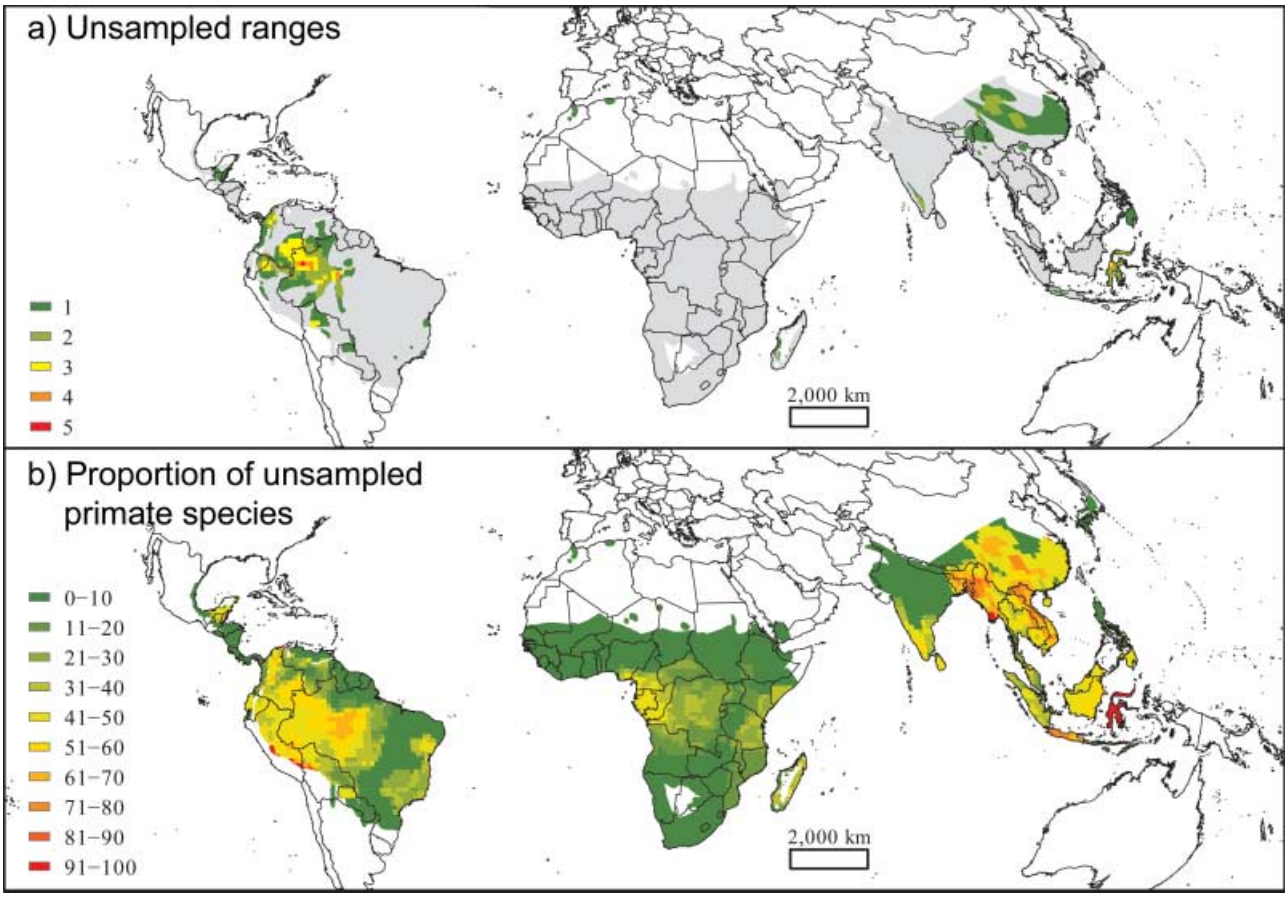

Figure 3 Unsampled primates: (a) Cell densities of primate species with geographical ranges in which no primate parasite sampling of any kind has occurred; (b) Proportion of unsampled species within a cell. Maps are displayed using Aitoff's projection at a scale of 1 : 200,000,000.

(0.597) ($\chi^2 = 7.43$, $P < 0.01$, d.f. = 1). However, spatial regression analyses taking into account neighbourhood effects indicated that the number of sampling localities per cell was significantly correlated with the total number of threatened species per cell (SAR: b = 0.010, $P < 0.001$, $R^2 = 0.05$).

The majority of threatened primates are concentrated in West Africa, with East to South-East Asia also having high numbers (Fig. 4a). When quantile subtraction was performed, Asia emerged as the geographical area with the greatest discrepancy between geographical distributions of parasite sampling and distributions of threatened primate species (particularly parts of China, Malaysia, Sri Lanka and Indonesia, Fig. 4b). Portions of Africa (specifically West Africa and Madagascar) and the Amazon also emerged as sampling gaps, but had smaller overall area. The number of unsampled threatened primate species was greatest in West Africa, with Asia and parts of the South American Amazon having fewer unsampled threatened species (Fig. 4c). However, in order to reach mean sampling levels, more threatened primates need to be sampled in South America and Asia than in Africa (Fig. 4d). Analyses considering solely the number of unsampled 'endangered' or 'critically endangered' primates (IUCN classifications 4 and 5) exhibited the same spatial trends as those for analyses where all threatened species were considered together (classifications 2–5).

Quantification of geographical gaps with respect to parasite taxonomy

One hundred and eighty-three parasite genera were represented in the georeferenced records, and sample points for all taxa were heavily concentrated in Africa — especially East Africa (Fig. 5). Of the six parasite types studied, helminths were sampled most frequently, at the most locations, and in the most host species (Fig. 6). Fungi and bacteria were the least sampled taxonomic groups, both in terms of the number of parasite species sampled and the number of locations at which primates were sampled for these parasites. Only three primate species have been sampled for fungal parasites at georeferenced locations in the GMPD, with none of these samples from the Americas or Asia. Bacterial parasites have been sampled in eight primate species at georeferenced locations, with only one sampling site located in the Americas and two in Asia.

Sampling for zoonotic diseases exhibited a pattern similar to all diseases considered together (Fig. 7a). However, sampling sites for 'emerging' zoonotic diseases were comparatively under-represented in Asia (Fig. 7b). Only two Asian sites where primates have been sampled for emerging zoonotic diseases were represented in the GMPD, and both of these sites were located in southern India. Among records in the GMPD, no primates in East or South-East Asia have been sampled for an emerging zoonotic disease at a georeferenced location (Fig. 7b). (See Appendix S4 in Supplementary Material

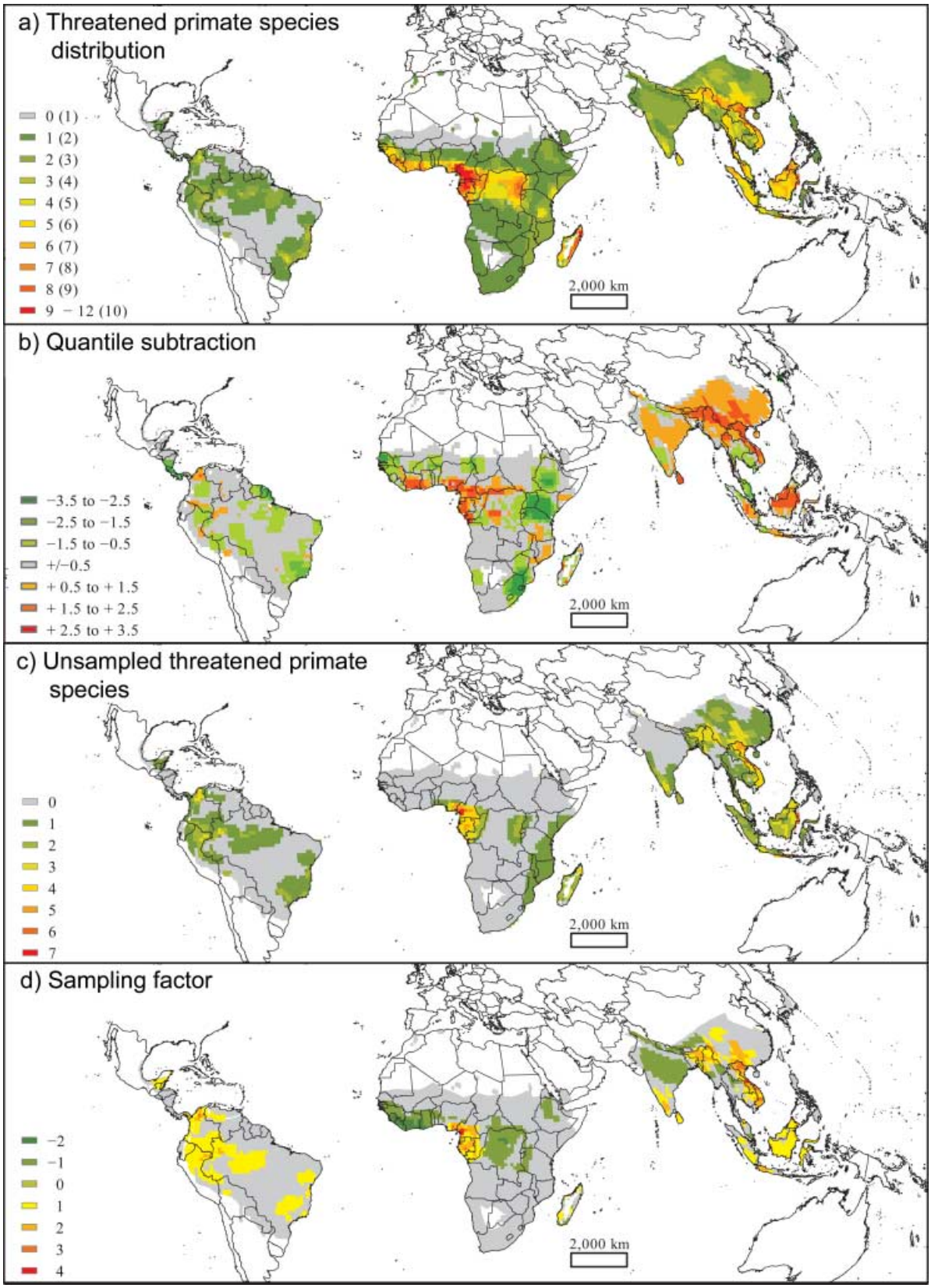

Figure 4 Quantification of geographical gaps with respect to host threat status and primate taxonomic sampling: (a) Threatened primate species density; (b) Cell differences in quantile values between threatened primate species density and smoothed sampling locality densities, classified according to standard deviations from the mean; (c) Number of unsampled threatened primate species per cell; (d) The number of threatened primate species that needs to be sampled per cell in order to reach mean sampling levels for threatened species. Maps are displayed using Aitoff's projection at a scale of 1 : 200,000,000.

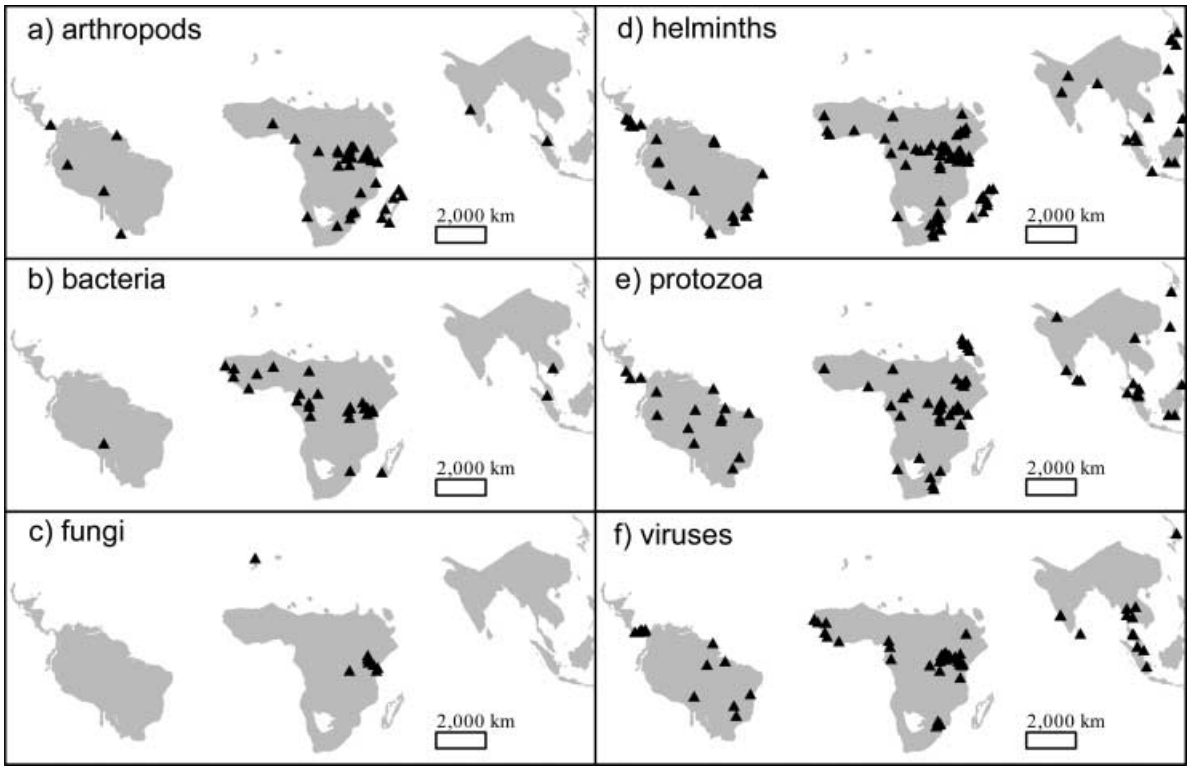

Figure 5 Geographical distribution of sampling localities at which at least one primate has been sampled for parasites belonging to the following groups: (a) arthropods; (b) bacteria; (c) fungi; (d) helminths; (e) protozoa; (f) viruses. Maps are displayed using Aitoff's projection at a scale of 1 : 300,000,000.

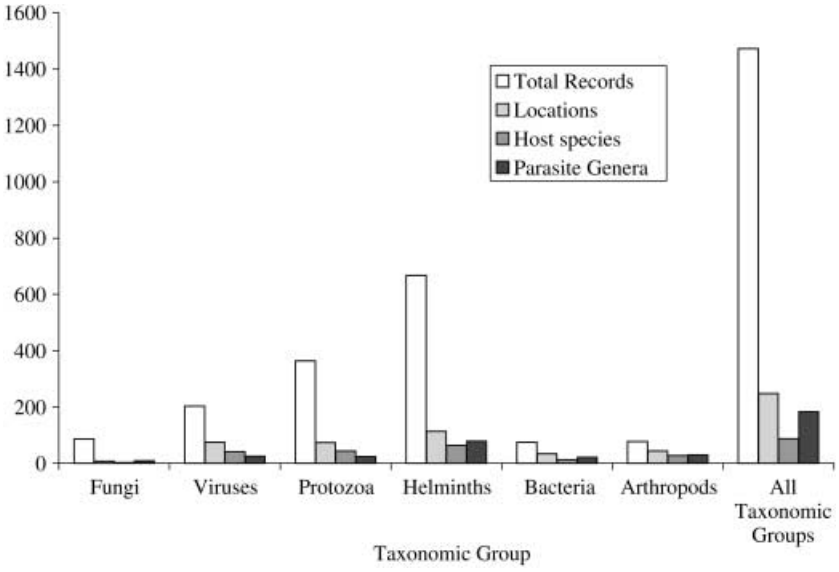

Figure 6 Parasite sampling separated by primate taxonomic group. Bars represent the total number of records in the GMPD for each parasite type, the number of locations at which each parasite type was sampled, the number of host species sampled for each parasite type, and the number of parasite genera of each parasite type documented in geo-referenced records.

for a list of emerging zoonotic diseases for which no primate has been sampled at a georeferenced location in the GMPD.)

DISCUSSION

Major epidemics in wild primates, such as those resulting from Ebola (Walsh *et al.*, 2003), have attracted worldwide attention and emphasize the potential negative consequences of infectious disease for wildlife. Despite this attention, our results show that major geographical and taxonomic sampling gaps exist in our understanding of parasites in wild primate populations. Specifically, we found that regardless of criteria used to define sampling gaps, Asia, South America, and small portions of West Africa emerged as areas that require greater sampling effort while East and South

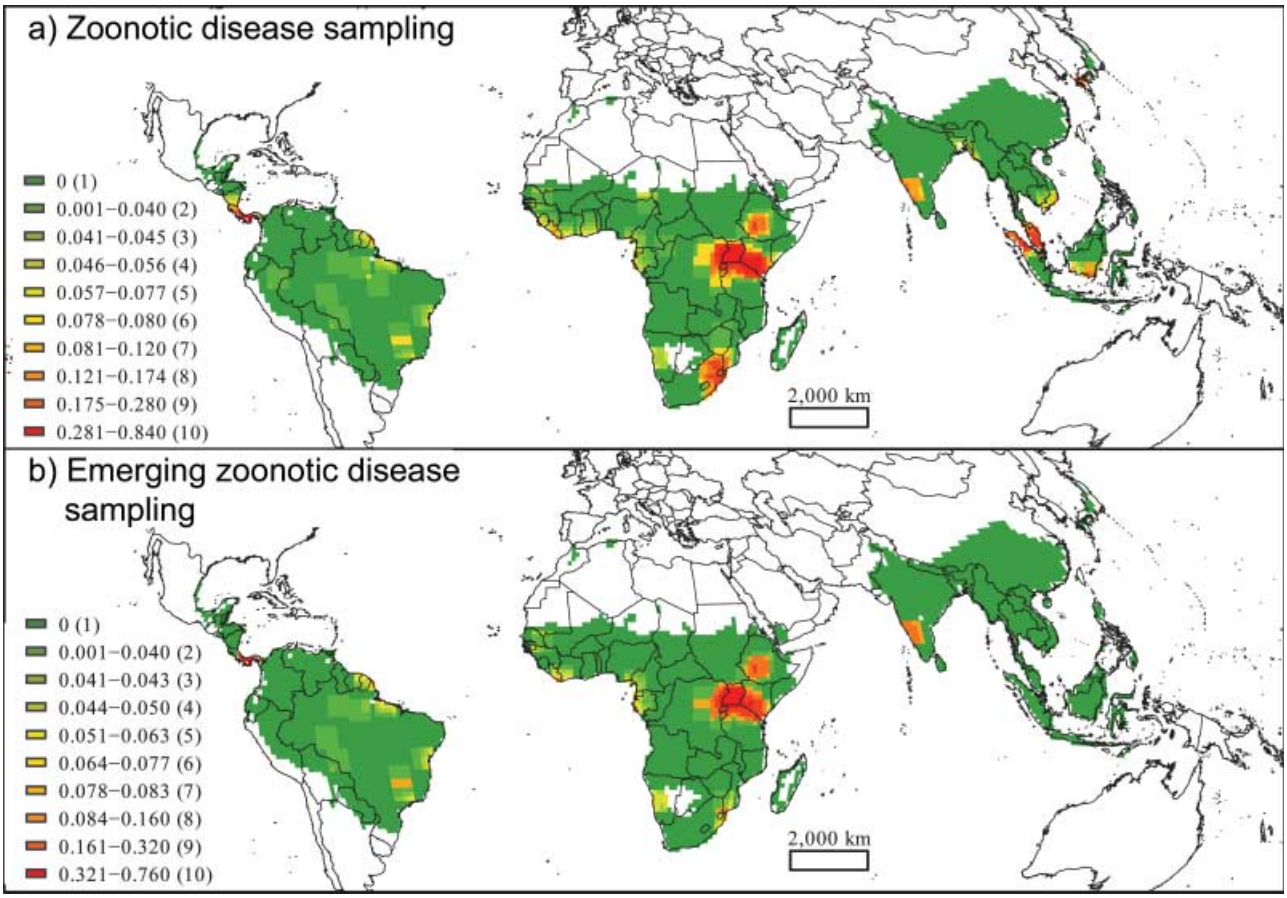

Figure 7 Zoonotic disease sampling: (a) Smoothed sampling densities measured as the mean number of sampling localities per 1 square-degree cell at which a primate was sampled for a zoonotic disease, with quantile values in parentheses; (b) Smoothed sampling locality densities measured as the mean number of sampling localities per 1 square-degree cell at which a primate was sampled for an emerging zoonotic disease, with quantile values in parentheses. Red values indicate areas with the highest sampling effort, and maps are displayed using Aitoff's projection at a scale of 1 : 300,000,000.

Africa have been sampled most intensively. We highlighted specific locations that require the most sampling effort to maximize our understanding of worldwide geographical patterns of disease risk in non-human primates, including the Amazon, western Africa, China, and South-East Asia. In addition, we demonstrated that bacteria and fungi — which account for approximately 60% of the infectious agents known to cause disease in humans (Taylor *et al.*, 2001) — are undersampled numerically and geographically in studies of non-human primate parasites. Addressing these geographical and taxonomic biases in parasite sampling should be an important component of future primate conservation efforts.

Many important questions remain about the role of infectious disease in species extinctions (Smith *et al.*, 2006). Future attempts to understand and mitigate wildlife disease outbreaks will require epidemiological research using a combination of two approaches: one based on detailed study of infection dynamics in a particular population over time, and the other using broader knowledge of parasitism at regional and global scales. Detailed studies at well-established sites have provided (and will continue to provide) essential data on the impact of infection on host survival and reproductive success, ecological and behavioural factors that promote disease spread, and the spread of parasites among hosts of the same and different species (Milton, 1996; Gillespie & Chapman, 2006; Nunn & Altizer, 2006). In contrast, studies that focus on understudied hosts, parasites, or geographical areas — such as those we identify here — are likely to reveal previously unknown parasites and pathogens, many of which could be essential for understanding future disease outbreaks in wildlife and humans in those areas. In addition, these broader-scale studies add to our knowledge of infectious disease distributions and thus contribute significantly to understanding the ecological and evolutionary drivers of parasitism (Nunn *et al.*, 2003, 2005). Ultimately, a balance of these two approaches will lead to the greatest progress.

In identifying which understudied geographical areas to target for future study sites, consideration should be given to factors that could impact the sampling patterns we documented. Although we cannot completely rule out the possibility of sampling biases arising from differential representation of foreign language literature in the bibliographic databases used to create the GMPD, such biases seem unlikely to account for the large spatial gaps revealed by our analyses. The geographical gaps we pinpointed are not restricted to areas where non-European

languages are spoken. Moreover, as mentioned in the Methods, the search engines used to construct the GMPD cover all major foreign languages. The bibliographic database 'PrimateLit' (http://primatelit.library.wisc.edu/), for instance, includes languages spoken in areas we have pinpointed as major sampling gaps (e.g. Portuguese, Chinese, and Japanese language papers had 962, 895 and 1777 hits, respectively).

Possible anthropogenic and environmental factors in specific areas seem more likely to explain observed sampling patterns. For example, central portions of the Amazon have probably remained relatively untouched due to the lack of infrastructure in such a remote area. Likewise, certain portions of western Africa and South-East Asia have also been characterized by a lack of infrastructure and, in specific cases, political instability that could deter research effort. Yet, while many obstacles to studies of primate parasites remain in specific areas, the maps we provide often show massive spatial gaps in our knowledge of primate parasites that cover several countries and span diverse landscapes. As a result, it is likely that there are sites within these gaps where research could feasibly begin.

In addition to site-specific research, studies conducted at broad taxonomic or geographical scales are needed in order to identify host traits and environmental factors that impact disease distributions, possibly leading to predictive maps of disease emergence. Global analyses of disease risk in a host clade, such as primates, will be best achieved using a combination of GIS and comparative approaches. For example, Nunn *et al.* (2005) combined data on parasites in primates with geographical information on median latitude and found a greater diversity of protozoa from wild primates that live in closer proximity to the equator. These results suggest that climatic and biotic conditions closer to the equator could favour the establishment and spread of protozoa, possibly by impacting the abundance of arthropods that spread parasitic protozoa. Vector-borne protozoa cause some of the most virulent diseases known to humans (Ewald, 1994), including malaria and trypanosomiasis, and many of these parasites also infect non-human primates. Future analyses could build upon current knowledge by combining information on primate distributions with data on the environmental factors that drive parasite abundance to better predict geographical patterns of disease risk. However, our results highlight the need to control for geographical sampling effort in global analyses of parasitism, or at least to consider how sampling gaps might influence the results of comparative and geographical studies.

Examinations of the role of multi-host parasites are also fundamental to attempts to understand and mitigate emerging infectious disease outbreaks, as multihost parasites can (by definition) move easily among host species, particularly in the context of anthropogenic changes that alter host ranging patterns, resource use and contact with domesticated animals (McCallum & Dobson, 1995; Daszak *et al.*, 2000; Woolhouse *et al.*, 2001; Gog *et al.*, 2002; Chapman *et al.*, 2005; Woolhouse & Gowtage-Sequeria, 2005). Research is needed to assess the risk of pathogen spill-over to threatened species, focusing on (1) parasite characteristics that lead to host shifts or increase host range; (2) host traits that increase exposure to new sources of infection; (3) compatibility between hosts and parasites, which is important for assessing the susceptibility of different host species and predicting the outcome of disease spill-over; and (4) ecological conditions that promote spill-over events.

In conclusion, understanding the global distribution of disease risk is likely to be increasingly important for understanding disease emergence in humans and wildlife (Daszak *et al.*, 2000; Ostfeld *et al.*, 2005). Here we provided a simple tool to identify sampling gaps in wild primate populations and to examine patterns in relation to host threat status. Future research could tie these approaches to other host characteristics, such as body mass or group size, and to factors that influence research effort devoted to parasite sampling in wild mammal populations, including climate, human population density, and political instability. Similarly, greater understanding will emerge if studies of sampling gaps are linked with details on the biotic, anthropogenic, and climatic factors that lead to disease establishment in natural populations, including factors that lead to uneven distributions of parasites themselves.

ACKNOWLEDGEMENTS

This research was supported by an NSF grant to CLN (#DEB-0212096), the Center for Applied Biodiversity Science at Conservation International, and the Max Planck Society. We are very grateful to Wes Sechrest, Jonathan Davies, and John Gittleman for providing primate geographical range data and associated details. This research benefited from helpful discussion with Sonia Altizer, Vanessa Ezenwa, Nikki Patel and Christophe Boesch. We would also like to thank Jonathan Davies, Rich Grenyer, Hjalmar Kuehl, and especially Wes Sechrest for helpful comments on early drafts of the manuscript.

REFERENCES

Berger, L., Speare, R., Daszak, P., Green, D.E., Cunningham, A.A., Goggin, C.L., Slocombe, R., Ragan, M.A., Hyatt, A.D., McDonald, K.R., Hines, H.B., Lips, K.R., Marantelli, G. & Parkes, H. (1998) Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences USA*, **95**, 9031–9036.

Bermejo, M., Rodriguez-Teijeiro, J.D., Illera, G., Barroso, A., Vila, C. & Walsh, P.D. (2006) Ebola outbreak killed 5000 gorillas. *Science*, **314**, 1564–1564.

Chapman, C.A., Gillespie, T.R. & Goldberg, T.L. (2005) Primates and the ecology of their infectious diseases: how will anthropogenic change affect host–parasite interactions? *Evolutionary Anthropology*, **14**, 134–144.

Cleaveland, S., Laurenson, M.K. & Taylor, L.H. (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.

Corbet, G.B. & Hill, J.E. (1992) *A world list of mammalian species*. Oxford University Press, Oxford.

Daszak, P., Berger, L., Cunningham, A.A., Hyatt, A.D., Green, D.E. & Speare, R. (1999) Emerging infectious diseases and amphibian population declines. *Emerging Infectious Diseases*, **5**, 735–748.

Daszak, P., Cunningham, A.A. & Hyatt, A.D. (2000) Threats to biodiversity and human health. *Science*, **287**, 443–449.

Ewald, P.W. (1994) *Evolution of infectious disease*. Oxford University Press, Oxford.

Funk, V.A., Richardson, K.S. & Ferrier, S. (2005) Survey-gap analysis in expeditionary research: Where do we go from here? *Biological Journal of the Linnean Society*, **85**, 549–567.

Gillespie, T. & Chapman, C. (2006) Prediction of parasite infection dynamics in primate metapopulations based on attributes of forest fragmentation. *Conservation Biology*, **20**, 441–448.

Gog, J., Woodroffe, R. & Swinton, J. (2002) Disease in endangered metapopulations: the importance of alternative hosts. *Proceedings of the Royal Society of London Series B, Biological Sciences*, **269**, 671–676.

Gregory, R.D. (1990) Parasites and host geographic range as illustrated by waterfowl. *Functional Ecology*, **4**, 645–654.

Groves, C.P. (1993) Order primates. *Mammal species of the world* (ed. by D.E. Wilson and D.M. Reeder), pp. 243–278. Smithsonian Institution Press, Washington D.C. & London.

Hilton-Taylor, C. (2002) *IUCN red list of threatened species*. IUCN, Gland, Switzerland.

IUCN (2006) *IUCN red list of threatened species*. <www.iucnredlist.org>. Downloaded on 15 January 2007.

Jennings, M.D. (2000) Gap analysis: concepts, methods, and recent results. *Landscape Ecology*, **15**, 5–20.

Leendertz, F.H., Ellerbrok, H., Boesch, C., Couacy-Hymann, E., Matz-Rensing, K., Hakenbeck, R., Bergmann, C., Abaza, P., Junglen, S., Moebius, Y., Vigilant, L., Formenty, P. & Pauli, G. (2004) Anthrax kills wild chimpanzees in a tropical rainforest. *Nature*, **430**, 451–452.

Leroy, E.M., Rouquet, P., Formenty, P., Souquiere, S., Kilbourne, A., Froment, J.M., Bermejo, M., Smit, S., Karesh, W., Swanepoel, R., Zaki, S.R. & Rollin, P.E. (2004) Multiple Ebola virus transmission events and rapid decline of central African wildlife. *Science*, **303**, 387–390.

McCallum, H. & Dobson, A. (1995) Detecting disease and parasite threats to endangered species and ecosystems. *Trends in Ecology & Evolution*, **10**, 190–194.

Milton, K. (1996) Effects of bot fly (*Alouattamyia baeri*) parasitism on a free ranging howler (*Alouatta palliata*) population in Panama. *Journal of Zoology*, **239**, 39–63.

National Geospatial-Intelligence Agency. (2005) Geographic Names Database. http://gnshttp://www.nga.mil/geonames/GNS/index.jsp.

Nunn, C.L. & Altizer, S. (2005) The Global Mammal Parasite Database: An online resource for infectious disease records in wild primates. *Evolutionary Anthropology*, **14**, 1–2.

Nunn, C. & Altizer, S. (2006) *Infectious disease in primates: behavior ecology and evolution*. Oxford University Press, Oxford.

Nunn, C.L., Altizer, S., Jones, K.E. & Sechrest, W. (2003) Comparative tests of parasite species richness in primates. *The American Naturalist*, **162**, 597–614.

Nunn, C.L., Altizer, S.M., Sechrest, W. & Cunningham, A. (2005) Latitudinal gradients of disease risk in primates. *Diversity and Distributions*, **11**, 249–256.

Ostfeld, R.S., Glass, G.E. & Keesing, F. (2005) Spatial epidemiology: an emerging (or reemerging) discipline. *Trends in Ecology & Evolution*, **20**, 328–335.

Pedersen, A.B., Poss, M., Altizer, S., Cunningham, A. & Nunn, C. (2005) Patterns of host specificity and transmission ammong parasites of wild primates. *International Journal for Parasitology*, **35**, 647–657.

Peeters, M., Courgnaud, V., Abela, B., Auzel, P., Pourrut, X., Bibollet-Ruche, F., Loul, S., Liegeois, F., Butel, C., Koulagna, D., Mpoudi-Ngole, E., Shaw, G.M., Hahn, B.H. & Delaporte, E. (2002) Risk to human health from a plethora of Simian immunodeficiency viruses in primate bushmeat. *Emerging Infectious Diseases*, **8**, 451–457.

Poulin, R. (1998) Comparison of three estimators of species richness in parasite component communities. *Journal of Parasitology*, **84**, 485–490.

Rouquet, P., Froment, J.M., Bermejo, M., Kilbourn, A., Karesh, W., Reed, P., Kilbourn, A., Karesh, W., Reed, P., Kumulungui, B., Yaba, P., Delicat, A., Rollin, P.E. & Leroy, E.M. (2005) Wild animal mortality monitoring and human Ebola outbreaks, Gabon and Republic of Congo, 2001–03. *Emerging Infectious Diseases*, **11**, 283–290.

Scott, J.M., Davis, F., Cusuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., Derchia, F., Edwards, T.C., Ulliman, J. & Wright, R.G. (1993) Gap analysis — A geographic approach to the protection of biological diversity. *Wildlife Monographs*, **123**, 1–41.

Sechrest, W., Brooks, T.M., da Fonseca, G.A.B., Konstant, W.R., Mittermeier, R.A., Purvis, A., Rylands, A.B. & Gittleman, J.L. (2002) Hotspots and the conservation of evolutionary history. *Proceedings of the National Academy of Sciences USA*, **99**, 2067–2071.

Smith, D.L., Lucey, B., Waller, L.A., Childs, J.E. & Real, L.A. (2002) Predicting the spatial dynamics of rabies epidemics on heterogenous landscapes. *Proceedings of the National Academy of Sciences USA*, **99**, 3668–3672.

Smith, K., Sax, D. & Lafferty, K. (2006) Evidence for the role of infectious disease in species extinction and endangerment. *Conservation Biology*, **20**, 1349–1357.

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

United Nations Environment Programme World Conservation Monitoring Centre (UNEP-WCMC). (2005) World Database on Protected Areas. http://sea.unep-wcmc.org/wdbpa/.

Walsh, P.D., Abernethy, K.A., Bermejo, M., Beyersk, R., De Wachter, P., Akou, M.E., Huljbregis, B., Mambounga, D.I., Toham, A.K., Kilbourn, A.M., Lahm, S.A., Latour, S., Maisels, F., Mbina, C., Mihindou, Y., Obiang, S.N., Effa, E.N., Starkey, M.P., Telfer, P., Thibault, M., Tutin, C.E.G., White, L.J.T. & Wilkie, D.S. (2003) Catastrophic ape decline in western equatorial Africa. *Nature*, **422**, 611–614.

Walther, B.A., Cotgreave, P., Price, R.D., Gregory, R.D. & Clayton, D.H. (1995) Sampling effort and parasite species richness. *Parasitology Today*, **11**, 306–310.

Walther, B.A. & Morand, S. (1998) Comparative performance of species richness estimation methods. *Parasitology*, **116**, 395–405.

Wolfe, N.D., Escalante, A.A., Karesh, W.B., Kilbourn, A., Spielman, A. & Lal, A.A. (1998) Wild primate populations in emerging infectious disease research: The missing link? *Emerging Infectious Diseases*, 149–158.

Woolhouse, M.E.J. & Gowtage-Sequeria, S. (2005) Host range and emerging and reemerging pathogens. *Trends in Ecology & Evolution*, **11**.

Woolhouse, M.E.J., Taylor, L.H. & Haydon, D.T. (2001) Population biology of multihost pathogens. *Science*, **292**, 1109–1112.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 List of countries that are not represented for parasite sampling in the Global Mammal Parasite Database (Nunn & Altizer, 2005) at the level of georeferenced locations.

Appendix S2 List of primate species that have not yet been sampled for parasites at georeferenced locations in the Global Mammal Parasite Database (Nunn & Altizer, 2005).

Appendix S3 Protected areas within primate geographical ranges that are most in need of primate parasite sampling.

Appendix S4 Emerging zoonotic diseases listed in Taylor *et al*. (2001) for which primates have not been sampled at georeferenced locations [according to the Global Mammal Parasite Database (Nunn & Altizer, 2005)].

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1472-4642.2007.00364.x

(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.