Mapping the zoonotic niche of Ebola virus disease in Africa

David M. Pigott[1†], Nick Golding[1†], Adrian Mylne[1], Zhi Huang[1], Andrew J. Henry[1], Daniel J. Weiss[1], Oliver J. Brady[1], Moritz U. G. Kraemer[1], David L. Smith[1,2], Catherine L. Moyes[1], Samir Bhatt[1], Peter W. Gething[1], Peter W. Horby[3], Isaac I. Bogoch[4,5], John S. Brownstein[6,7], Sumiko R. Mekaru[6], Andrew J. Tatem[8,9,10], Kamran Khan[4,11] and Simon I. Hay[1,9]*

[1] *Spatial Ecology & Epidemiology Group, Department of Zoology, University of Oxford, UK*

[2] *Sanaria Institute for Global Health and Tropical Medicine*

[3] *Epidemic Diseases Research Group, Centre for Tropical Medicine and Global Health, University of Oxford, UK*

[4] *Department of Medicine, Division of Infectious Diseases, University of Toronto, Toronto, Canada*

[5] *Divisions of Internal Medicine and Infectious Diseases, University Health Network, Toronto, Canada*

[6] *Children's Hospital Informatics Program, Boston Children's Hospital, Boston, MA, USA*

[7] *Department of Pediatrics, Harvard Medical School, Boston, MA, USA*

[8] *Department of Geography and Environment, University of Southampton, Highfield, Southampton, UK*

[9] *Fogarty International Center, National Institutes of Health, Bethesda, MD, USA*

[10] *Flowminder Foundation, Stockholm, Sweden*

[11] *Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada*

*corresponding author: simon.hay@zoo.ox.ac.uk

[†]joint first authors

Abstract

Ebola virus disease (EVD) is a complex zoonosis that is highly virulent in humans. The largest recorded outbreak of EVD is ongoing in West Africa, outside of its previously reported and predicted niche. We assembled location data on all recorded zoonotic transmission to humans and Ebola virus infection in bats and primates (1976-2014). Using species distribution models these occurrence data were paired with environmental covariates to predict a zoonotic transmission niche covering 22 countries across Central and West Africa. Vegetation, elevation, temperature, evapotranspiration and suspected reservoir bat distributions define this relationship. At-risk areas are inhabited by 22 million people, however the rarity of human outbreaks emphasises the very low probability of transmission to humans. Increasing population sizes and international connectivity by air since the first detection of EVD in 1976 suggest that the dynamics of human-to-human secondary transmission in contemporary outbreaks will be very different to those of the past. [150/150]

Introduction

Ebola viruses have for the last forty years been responsible for a number of outbreaks of Ebola virus disease (EVD) in humans *(Pattyn et al., 1977)*, with high case fatality rates typically around 60-70%, but potentially reaching as high as 90% *(Feldmann and Geisbert, 2011)*. The most recent outbreak began in Guinea in December 2013 *(Baize et al., 2014; Bausch and Schwarz, 2014)* and has subsequently spread to Liberia, Sierra Leone and Nigeria *(ECDC, 2014)*. The unprecedented size and scale of this ongoing outbreak has the potential to destabilise already fragile economies and healthcare systems *(Fauci, 2014)*, and fears of international spread of a Category A Priority Pathogen *(NIH, 2014)* have made this a massive focus for international public health *(Chan, 2014)*. This has led to the current outbreak being declared a Public Health Emergency of International Concern on the 8th August 2014 *(Briand et al., 2014; Gostin et al., 2014; WHO, 2014)*.

The *Filoviridae*, of which *Ebolavirus* is a constituent genus, belong to the order *Mononegavirales*. Two other genera complete the family: *Marburgvirus*, itself responsible for a number of outbreaks of haemorrhagic fever across Africa *(Conrad et al., 1978; Gear et al., 1975; Smith et al., 1982; Towner et al., 2006)* and *Cuevavirus*, recently isolated from bats in northern Spain *(Negredo et al., 2011)*. Five species of *Ebolavirus* have been isolated to date *(King et al., 2011; Kuhn et al., 2010)*; the earliest recognised outbreaks of EVD were reported in Zaire (now the Democratic Republic of the Congo (DRC)) and Sudan in 1976 *(International Commission, 1978; WHO International Study Team, 1978)*. The causative viruses were isolated *(Pattyn et al., 1977)* and later identified to be distinct species, *Zaire ebolavirus* (EBOV) and *Sudan ebolavirus* (SUDV). A third species of *Ebolavirus*, *Reston ebolavirus*, was isolated from Cynomologus monkeys imported from the Philippines to a facility in the United States, where they experienced severe haemorrhaging *(Jahrling et al., 1990)*. Whilst serological evidence of infection with this species has been reported in individuals in the Philippines *(Miranda et al., 1991)*, no pathogenicity has been reported beyond primates and porcids *(Barrette et al., 2009; Feldmann and Geisbert, 2011)*. In 1994 a fourth species, *Tai Forest ebolavirus* was isolated from a veterinarian who had autopsied a chimpanzee in Côte d'Ivoire *(Le Guenno et al., 1995)*, though the virus has not been detected subsequently. The final species, *Bundibugyo ebolavirus*, was responsible for an outbreak of EVD in Uganda in 2007 *(Towner et al., 2008)*, as well as a more recent outbreak in the DRC *(WHO, 2012)*.

Initial analysis suggested that the viruses isolated from the current outbreak, originating in Guinea, formed a separate clade within the five *Ebolavirus* species *(Baize et al., 2014)*. Subsequent re-analysis of the same sequences however, indicated that these isolates instead nest within the *Zaire ebolavirus* lineage *(Dudas and Rambaut, 2014)*, and diverged from Central Africa strains approximately ten years ago *(Gire et al., 2014)*.

Which reservoir species are responsible for maintaining Ebola transmission between outbreaks is not well understood *(Peterson et al., 2004)*, but over the last decade significant progress has been made in narrowing down the list of likely hosts *(Peterson et al., 2007)* (Figure 1). Primates have long been known to harbour filoviral infections, with the first Marburg strains identified in African green monkeys in 1967 *(Beer et al., 1999; Siegert et al., 1967)*. Significant mortality has also been reported in wild primate populations across Africa, most notably in gorilla (*Gorilla gorilla*) and chimpanzee (*Pan troglodytes*) populations *(Bermejo et al., 2006; Formenty et al., 1999; Rouquet et al., 2005)*. The high case fatality rates recorded in the great apes combined with their declining populations and limited geographical range, indicate they are likely dead-end hosts for the virus and not reservoir species *(Groseth et al., 2007)*. A large survey of small mammals in and around Gabon identified three species of bats which were infected with Ebola viruses – *Hypsignathus monstrosus*, *Epomops franqueti* and *Myonycetris torquata (Leroy et al., 2005)*. Subsequent serological surveys *(Hayman et al., 2010; Pourrut et al., 2009)* and evidence linking the potential source of human outbreaks to bats *(Leroy et al., 2009)* lend support to the hypothesis of a bat reservoir. This, coupled with repeated detection of *Marburgvirus* in the fruit bat *Rousettus aegypticus (Towner et al., 2009)* and the only isolations of *Cuevavirus* also from bats (specifically *Llovia virus (Negredo et al., 2011)*), all support the suspicion that Chiroptera play an important role in the natural life-cycle of the filoviruses.

Humans represent a dead-end host for the virus, with only stuttering chains of transmission reported between humans in the majority of previous outbreaks *(Chowell et al., 2004; Legrand et al., 2007)* and no indication that humans can reintroduce the virus back into reservoir species *(Karesh et al., 2012)*. The incubation period in humans ranges from two days to three weeks, after which a variety of clinical symptoms arise, affecting multiple organs of the body. At the peak of illness, haemorrhaging shock and widespread tissue damage can occur and can eventually lead to death within 6-16 days *(Feldmann and Geisbert, 2011)*. Human-to-human transmission is mainly through direct unprotected contact with infected individuals and cadavers, with infectious particles detected in a number of different body fluids *(Feldmann and Geisbert, 2011)*. The typical outbreak profile is defined by an index individual that has recently come into contact with the blood of another mammal through either hunting or the butchering of animal carcasses *(Pourrut et al., 2005)*. Whilst it has been difficult to identify the zoonotic source for the index cases of some outbreaks, a recurring theme of hunting and handling bushmeat is suspected (Table 1) *(Boumandouki et al., 2005; Leroy et al., 2009; Nkoghe et al., 2005; Nkoghe et al., 2011)*. For some outbreaks, including the most recent, the initial source of zoonotic transmission has not been identified. In subsequent human-to-human transmission, the highest risk activities are those that bring humans into close contact with infected individuals. These include medical settings where insufficient infection control precautions have been taken, as well as home care and funeral preparations carried out by families or close friends *(Baron et al., 1983; Boumandouki et al., 2005; Georges et al., 1999)*. As the conditions required for transmission are

culturally and contextually dependent, opportunities for sustained transmission are highly heterogeneously distributed. Typically, chains of infection do not exceed three or four sequential transmission events, although occasionally (and particularly in the early stages of infection) a single individual may be responsible for directly infecting a large number of others *(Brady et al., 2014)*. In the outbreak in Gabon in 1996, a single person was responsible for infecting ten other individuals *(Milleliri et al., 2004)* whilst in the 1995 outbreak in the DRC, thirty five cases resulted from one individual *(Khan et al., 1999)*. Secondary transmission can be restricted by effective case detection and isolation measures *(Shoemaker et al., 2012; WHO, 2014)*. Where this cannot be achieved, either due to a lack of infrastructure, poor understanding of the disease, or distrust of medical practices, secondary cases can continue to occur *(Hewlett et al., 2005; Khan et al., 1999; Larkin, 2003)*. As the number of infections grows, the ability of healthcare systems to control the further spread diminishes and the risk of a large outbreak increases.

The recent outbreak in Guinea and surrounding countries indicate that the previous paradigm for Ebola outbreaks is shifting *(Briand et al., 2014; Chan, 2014)*. The last forty years of EVD outbreaks were accompanied by considerable changes in demographic patterns throughout Africa. There has been a large increase in population size coupled with increasing urbanisation *(Cohen, 2004; Linard et al., 2013; Seto et al., 2012)*. African populations have also become better connected internally and internationally *(Huang and Tatem, 2013; Linard et al., 2012)*. Only recently have we begun to understand the dynamic nature of these travel patterns *(Garcia et al., 2014; Gonzalez et al., 2008; Simini et al., 2012; Wesolowski et al., 2013; Wesolowski et al., 2012)* which have been clearly demonstrated to influence disease transmission over different temporal and spatial scales *(Brockmann and Helbing, 2013; Hufnagel et al., 2004; Pindolia et al., 2014; Stoddard et al., 2009; Talbi et al., 2010; Yang et al., 2008)*. Changes in land use and penetration into previously remote areas of rainforest bring humans into contact with potential new reservoirs *(Daszak, 2000)*, while changes in human mobility and connectivity will likely have profound impacts on the dispersion of Ebola cases during outbreaks. These conditions are thought to have a major role in setting the stage for the current outbreak.

This paper aims to define the areas suitable for zoonotic transmission of *Ebolavirus* (*i.e.* those routes defined in Figure 1 excluding human-to-human transmission) through species distribution modelling techniques. The fundamental niche of a species can be conceptualised as the confluence of environmental conditions that support its presence in a particular location *(Franklin, 2009)*. Species distribution models quantitatively describe this niche based on known occurrence records of the organism and their associated environmental conditions, enabling predictions of the likely geographic distribution of the species in other regions *(Elith and Leathwick, 2009)*. The era of satellites and geographical information systems has made high resolution global data on environmental conditions increasingly available *(Hay et al., 2006; Weiss et al., 2014)*. Species distribution modelling using

145 flexible machine learning approaches have been successfully applied to map the global distributions
146 of disease vectors *(Sinka et al., 2012)* and pathogens such as dengue *(Bhatt et al., 2013)*, influenza
147 *(Gilbert et al., 2014)* and leishmaniasis *(Pigott et al., 2014)*.

148 Previous studies applied the GARP (Genetic Algorithm for Rule-set Production) species distribution
149 modelling approach *(Stockwell and Peters, 1999)* to the locations of twelve Ebola outbreaks in
150 humans between 1976 and 2002 to map the likely distribution of Ebola viruses *(Peterson et al., 2004)*
151 and as a mechanism to identify potential reservoir hosts *(Peterson et al., 2004; Peterson et al., 2007)*.
152 Here we update and improve the maps of the zoonotic transmission niche of EVD by: (i)
153 incorporating more recent outbreak data from outside the formerly predicted niche of EVD; (ii)
154 integrating for the first time data on outbreaks in primates and the occurrence of the virus in the
155 suspected Old World fruit bat (OWFB) reservoirs; (iii) using new satellite-derived information on
156 bespoke environmental covariates from Africa, including new distribution maps of the OWFB; and
157 (iv) using new increasingly flexible niche mapping techniques in the modelling framework. To
158 elucidate the relevance of these maps for transmission, we have also calculated the population at risk
159 of primary spillover outbreaks from the zoonotic niche of EVD in Africa, and we investigated the
160 changing nature of the populations within this niche.

161

Results

Reported EVD outbreaks

In total, 23 outbreaks of Ebola virus were identified in humans across Africa, consisting of a hypothesised 30 independent primary infection events (Table 1 and Figure 2). These outbreaks span the last forty years from the first outbreaks in 1976 to the five outbreaks that have occurred since 2010 (Table 1). The locations of the index cases span from West Africa, with the most westerly outbreak ongoing in Guinea, to Gabon, the Republic of Congo (ROC), the DRC, South Sudan and Uganda. Before December 2013, a total of 2,322 cases had occurred from *Ebolavirus* infections, a number already overtaken by the likely underreported current case count of the ongoing outbreak >2,250 *(WHO, 2014)*(Figure 2A). Of the four viruses circulating in Africa, *Zaire ebolavirus* has been responsible for the most outbreaks (13), followed by *Sudan ebolavirus* (7) and *Bundigbuyo ebolavirus* with just two outbreaks in 2007/8 and 2012. Tai Forest has caused one confirmed infection in humans, from which the patient recovered *(Formenty et al., 1999; Le Guenno et al., 1995)*. Although outbreaks have been reported since 1976, there was an absence of reported outbreaks in humans for 15 years between 1979 and 1994 (although antibodies in humans were identified over the period *(Kuhn, 2008)*) and the frequency of outbreaks has increased substantially post 2000 (Figure 2A).

Reported Ebola virus infections in animals

A total of 51 surveyed locations reporting infections in animals were identified in the literature since the discovery of the disease (Table 2 and Figure 3). These comprised 17 infections in gorillas (*Gorilla gorilla)*, nine infections in chimpanzees (*Pan troglodytes*), 18 in OWFB and two in duikers (*Cephalophus* spp.). A large proportion of the great ape cases originated from the ROC / Gabon border, coinciding with the main known distributions of both chimpanzees and gorillas *(Petter and Desbordes, 2013)* and representing a period of well-documented great ape Ebola outbreaks in and around the Lossi Animal Sanctuary *(Bermejo et al., 2006; Rouquet et al., 2005; Walsh et al., 2009)*. All animal isolations of Ebola viruses have come from countries that have also reported index cases of human outbreaks, with the exception of several seropositive bats from a survey in southern Ghana.

Predicted distribution of suspected reservoir species of bats.

Three species of bats, *Hypsignathus monstrosus*, *Myonycteris torquata* and *Epomops franqueti*, were identified as the most likely candidates to be reservoir species for Ebola viruses due to high seroprevalence and the isolation of RNA closely related to *Zaire ebolavirus (Leroy et al., 2005; Olival and Hayman, 2014)*. In total, 239 locations were identified from the Global Biodiversity Information Facility (GBIF) *(GBIF, 2014)*: 67 for *H. monstrosus* (Figure 4A), 52 for *M. torquata* (Figure 4B) and 120 for *E. franqueti* (Figure 4C). Distribution models for all three species demonstrated predictive

skill (indicated by an area under the curve (AUC) greater than 0.5) as follows: *H. monstrosus* AUC 0.63±0.04; *M. torquata* AUC=0.59±0.04; *E. franqueti* AUC=0.58±0.03, n=50 submodels for all three species. In addition, each species was broadly predicted within its considered expert opinion range (Figure 4A-C) *(Schipper et al., 2008)*. The marginal effect plots (not shown) were strongly influenced by land surface temperature (LST) and vegetation (as measured by the enhanced vegetation index (EVI)). The predicted combined distribution of these species (Figure 4D), covers West and Central Africa, specifically the moist forests of the northeastern, western and central Congo basin, and Guinea, as well as the Congolian coastal forest ecoregions *(WWF, 2014)*.

Predicted environmental suitability for zoonotic transmission of Ebola

The predicted environmental niche for zoonotic transmission of EVD is shown in Figure 5. All countries with observed index cases of EVD (n=7, hereafter Set 1) have areas of the highest environmental suitability (see list in Table 1). In addition, areas of high environmental suitability for zoonotic transmission are predicted in a further 15 countries where, to date, index cases of the four African species of *Ebolavirus* have not been recorded. These are Nigeria, Cameroon, Central African Republic (CAR), Ghana, Liberia, Sierra Leone, Angola, Tanzania, Togo, Ethiopia, Mozambique, Burundi, Equatorial Guinea, Madagascar and Malawi (hereafter Set 2).

The AUC for the Ebola model was relatively high (AUC=0.85±0.04, n=500 submodels) indicating that the model could strongly distinguish regions of environmental suitability for EVD. Enhanced vegetation index had the greatest impact on the distribution (relative contribution (RC) of 65.3%) followed by elevation (RC=11.7%), night-time land surface temperature (LST) (RC=7.7%), potential evapotranspiration (PET) (RC=5.7%) and combined bat distribution (RC=3.8%). Marginal effect plots are presented in Figure 5 – figure supplement 2.

In total, 22.2 million people are predicted to live in areas suitable for zoonotic transmission of Ebola. The vast majority, 21.7 million (approximately 97%), live in rural areas, as opposed to urban or peri-urban areas *(CIESIN/IFPRI/WB/CIAT, 2007; WorldPop, 2014)*. Of these, 15.2 million are in Set 1 and 7 million are in Set 2. In terms of ranked populations at risk, DRC, Guinea and Uganda are highest in Set 1 and Nigeria, Cameroon and CAR are top in Set 2. For a full listing of these populations living in areas of risk, see the stacked bar plot in Figure 5B.

National level demographic and mobility changes

Over the 40 year period since discovery of EVD, the total population living in those countries predicted to be within the zoonotic niche has nearly tripled (from 230 million to 639 million) and the proportion of the population in these countries living in an urban (rather than rural) setting has changed from 25.5% to 59.2% (Figure 6).

Data on the connectivity of human populations over this period were not available. We can infer however, intuitively, empirically and theoretically *(Simini et al., 2012; Zipf, 1946)* that rates of population movement within a country will scale directly in proportion to population growth.

International connectivity by airline traffic

Records of passenger seat capacity are available since 2000 and show substantive increases over the period in Set 1 (from 2.96 to 4.77 million, a fractional change of 1.61) and Set 2 (from 5.6 to 15.6 million, a change of 2.8) (Figure 7A). More specific data on passenger volumes show almost universally similar increases since 2005 with Set 1 nations changing from 2 million to 2.5 million, a fractional change of 1.22 and Set 2 changing from 5 million to 7.9 million, a change of 1.57 (Figure 7B).

Global analysis of airline passenger volumes demonstrates that international connectivity has increased amongst all global regions and national income strata (Figure 8). Total passenger volumes have increased by a third from 9.5 to over 14 million during the eight year window (2005-2012) where records are available. The largest increases have occurred in WHO regions *(WHO, 2014)* outside of the sub-Saharan African region (AFRO) (Figures 8A and B). In 2012, almost half of the final destinations of those travelling from these at-risk countries were to other AFRO nations (47%). Other frequent destinations were in Europe (EURO; 27%) and the Eastern Mediterranean (EMRO; 13%). Similarly, analysis of passenger volumes by World Bank national income groupings *(WHO, 2014)* (Figures 8C and D) show that in 2012 40% of all passenger final destinations were to low or low-middle income countries.

Discussion

Summary of the main findings

We have re-evaluated the zoonotic niche for EVD in Africa. In doing so we have (i) used all existing outbreaks to assemble an inventory of index cases (n=30); (ii) added to this all confirmed records of Ebola virus in animals (n=51); (iii) assembled more accurate and contemporary environmental covariates including new maps of the distribution of confirmed OWFB reservoirs of the disease; and (iv) used the latest niche modelling techniques to predict the geographic distribution of potential zoonotic transmission of the disease. Using these predictions we have estimated the populations at risk of EVD both in countries which have confirmed index cases (Set 1, n=7) and those for which we predict strong environmental suitability for outbreaks (Set 2, n=15). In all countries at risk we show that since the discovery of EVD in 1976, urban and rural populations have increased and have become more interconnected both within and across national borders. During the last 40 years the increasing size and connectivity of these populations may have facilitated the subsequent spread of EVD outbreaks. These factors underline a change in the way in which EVD interacts with human populations.

Interpreting the zoonotic niche

The remote and isolated nature of Ebola zoonotic transmission events, paired with the relatively poor diagnostics and understanding of the disease transmission routes in early outbreaks, mean that under-reporting of previous outbreaks is probable. An increasing understanding and description of a broader range of symptoms used in case definitions of EVD *(Feldmann and Geisbert, 2011; Leroy et al., 2000)* also increase the possibility that past outbreaks may have been misattributed to different diseases *(Tignor et al., 1993)*. This poor detectability of EVD also clearly limits capacity to accurately identify the locations and transmission routes of index cases *(Baize et al., 2014; Heymann et al., 1980)*. We must assume, as has been done previously *(Jones et al., 2008; Peterson et al., 2004)*, that the first reported cases are representative of the true location of the index cases. Where possible we have represented this geographic uncertainty by attributing the index case to a wide-area polygon which then incorporated this uncertainty into the mapping process (see Methods).

The relationship between the EVD niche and the environmental covariates (Figure 5 – figure supplement 2), particularly the high relative contribution of the vegetation index, underscore that there are clear environmental limits to transmission of the virus from animals to humans, and that ecoregions dominated by rainforest are the primary home of such zoonotic cycles. Our analysis has shown that the zoonotic niche of the pathogen is more widespread than previously predicted or appreciated *(Peterson et al., 2004)*, most notably in West Africa.

This analysis used information from all human outbreaks and animal infections to delineate the likely zoonotic niche of the disease. Further analysis, excluding the existing outbreak focussed in Guinea from the dataset used to train the model (Figure 5 – figure supplement 3), still resulted in prediction of high suitability in this region, with the presumed index village located within 5km of an at-risk pixel. This implies that the eco-epidemiological situation in Guinea is very similar to that in past outbreaks, mirroring phylogenetic similarity in the causative viruses *(Dudas and Rambaut, 2014; Gire et al., 2014)*. The ecological similarity between the past and current outbreaks also lends support to the notion that the scale of this outbreak is more heavily influenced by patterns of human-to-human transmission than any expansion of the zoonotic niche.

Interpreting Population at Risk

It is important to appreciate that this zoonotic niche map delineates areas in which populations are at-risk of zoonotic transmission of EVD (Figure 5B). It does not predict the likelihood of EVD spillover, the likelihood of an outbreak establishing, or its subsequent rate of spread within a population. Increasing human encroachment and certain cultural practices sometimes linked with poverty, such as bushmeat hunting, result in increasing exposure of humans to animals which may harbour diseases including Ebola *(Daszak, 2000; Wolfe et al., 2005; Wolfe et al., 2007)*. Increasing human population may accelerate the degree of risk through these processes but spatially refined information on these factors is not available comprehensively. It is hoped that as the understanding of the risk factors for zoonotic transmission of *Ebolavirus* to humans increases, it will be possible to incorporate this information into future risk mapping assessments.

Previous considerations of the geographic distribution of EVD have used human outbreaks alone. We have updated this work to include the last decade of outbreaks, as well as disaggregated outbreaks where evidence suggests multiple independent zoonotic transmission events overlap in space and time. Furthermore, our modelling process accommodates uncertainty in geopositioning of these index cases by utilising both point and polygon data. In addition, we include occurrence of infection in wildlife, important to the wider scale of zoonotic transmission (Figure1), which in total has increased the dataset used in the model to 81 occurrences. The rareness of EVD outbreaks and the prevalence of detectable Ebola virus in reservoir species suggests that there will always be a limited set of observation data when compared to mapping of more prevalent zoonoses *(Pigott et al., 2014)*. The results demonstrate predictive skill using a stringent validation procedure, however, indicating strong model performance even with this relatively limited observation dataset.

A broad zoonotic niche is predicted across 22 countries in Central and West Africa. Whilst several of these countries have reported index cases of EVD, others have not, although serological evidence in some regions points to possible underreporting of small-scale outbreaks *(Kuhn, 2008)*. With improved

316 ecological understanding, particularly with improvements to our knowledge of specific reservoir
317 species and their distributions, it may be possible to delineate areas not at risk due to the absence of
318 these species.

319 Despite relatively a large population living in areas of risk and the widespread practice of bushmeat
320 hunting in these predicted areas *(Brashares et al., 2011; Kamins et al., 2011; Mfunda and Røskaft,*
321 *2010; Wolfe et al., 2005)*, *Ebolavirus* is rare both in suspected animal reservoirs *(Leroy et al., 2005;*
322 *Olival and Hayman, 2014)* and in terms of human outbreaks (Table 1). There is some indication
323 however, that the frequency of Ebola outbreaks has increased since 2000, as shown in Figure 2A. We
324 have shown that the human population living within this niche is larger, more mobile and better
325 internationally connected than when the pathogen was first observed. As a result, when spillover
326 events do occur, the likelihood of continued spread amongst the human population is greater,
327 particularly in areas with poor healthcare infrastructure *(Briand et al., 2014; Fauci, 2014)*.

328 Whilst rare in comparison to other high burden diseases prevalent in this region, such as malaria
329 *(Gething et al., 2011; Murray et al., 2012)*, Ebola outbreaks can have a considerable economic and
330 political impact, and the subsequent destabilisation of basic health care provisioning in affected
331 regions increases the toll of unrecorded morbidity and mortality of more common infectious diseases
332 *(Murray et al., 2014; Wang et al., 2014)*, throughout and after the epidemic period. The number of
333 concurrent infections during the present outbreak represents a significant strain on healthcare systems
334 that are already poorly provisioned *(Briand et al., 2014; Chan, 2014; Fauci, 2014)* and many other
335 Set 1 and Set 2 countries rank amongst the lowest per capita healthcare spenders. These
336 considerations should be paramount when international organizations debate the financing
337 requirements for the improvement of healthcare needed in the region and the urgency with which new
338 therapeutics and vaccines can be brought into production *(Brady et al., 2014; Goodman, 2014)*.

339 Together, these considerations necessitate prioritisation of efforts to reinforce and improve existing
340 surveillance and control, and encourage the development of therapeutics and vaccines. The national
341 population at risk estimates presented here would be a strong rationale for improving, prioritising and
342 stratifying surveillance for EVD outbreaks and diagnostic capacity in these countries. We believe it
343 would be prudent to test OWFB species in Set 2 countries for Ebola virus *(Hayman et al., 2012)*,
344 particularly during the bat breeding season to maximise chances of isolation in order to clarify the
345 outbreak risk in these countries. In all Set 1 and Set 2 countries, raising awareness about the risk
346 presented by reservoir bats and incidental primate hosts and the modes of transmission of this disease
347 could be of value. Finally, increasing our capacity to rapidly map ever changing biological threats is
348 also a core need *(Hay et al., 2013)*.

349 Interpreting International Connectivity

The increasing connectedness of the Africa region means that EVD is now a problem of international concern. While most EVD secondary transmission occurs locally and is likely transported *via* ground transit *(Francesconi et al., 2003)*, the potential for international spread of infection is possible, as demonstrated by the importation of the disease from Liberia to Nigeria, culminating in further secondary transmission in Lagos *(WHO, 2014)*. The aetiology of EVD infection and disease progression means that an international outbreak propagated by air travel remains unlikely, particularly in high-income countries better able to handle EVD cases *(Fauci, 2014)*. Nevertheless, a non-negligible threat remains, particularly in the low and middle income destinations and the rapid increase in global connectivity of these at-risk regions indicates that international airports could see more imported cases *(Chan, 2014)*.

Future work

We have focussed on reanalysing the zoonotic niche for EVD transmission and the characterisation of the populations at risk to improve the landscape in which future risk and impact of EVD outbreaks can be discussed. During the current emergency much of the work will concentrate on routes of secondary transmission in the human population – conceptually the red arrow of the H box in Figure 1. An important task is to stratify the risk of EVD spread both within and between countries and identify the most likely pathways of spread for characterisation and surveillance. Our next priority therefore is to investigate aspects of secondary human-to-human transmission by documenting the rate of geographic spread of EVD during the past and ongoing epidemics to help understand changes in these patterns in the historical record. Simulating these movements in a real landscape of population movement patterns, inferred from population movements assessed by mobile phones and other data *(Garcia et al., 2014)*, as well as parametric movement models *(Simini et al., 2013)* is a logical next step, and can be used in future targeting of interventions and potential new treatments for both the current and future outbreaks *(Brady et al., 2014; Goodman, 2014)*.

As previously discussed, whilst there is the risk of human travel during the latent phase of infection, and therefore potential for international spread, the high pathogenicity during infectiousness (immobilising infected persons) and the likely rapid and effective isolation measures implemented in regions with strong health care systems, limit the pandemic potential of EVD. Nevertheless, improvement of international containment plans and informed discussions of potential risks to airline carriers and populations of other regions will be supported by knowledge of local, regional and international population flows. Assessing these flows by air traffic volumes is an ongoing priority.

There are several other zoonotic viral haemorrhagic fevers (for example *Marburgvirus*, Lassa fever, hantaviral infections and arenaviruses) that are of similar public health and biosecurity concern *(Bannister, 2010)*, due to their high virulence and mortality and their potential to cause outbreaks and

384 spread internationally. Despite this their geographical distributions are poorly understood *(Hay et al.,*
385 *2013)*. Many of the methods applied here can be adapted to these diseases and improve our
386 geographical understanding of the risk presented by these pathogens.

387 We are in the midst of a public health emergency that will likely last for many more months *(Chan,*
388 *2014)* and which has brought EVD to global attention. We emphasise that the maps of zoonotic
389 transmission presented here do not enable assessment of secondary transmission rates in human
390 populations, but they do act as an evidenced-based indicator of locations with potential for future
391 zoonotic transmission and thus outbreaks. Interestingly, early reports of another independent zoonotic
392 outbreak in the DRC *(MSF, 2014)* are in predicted at-risk areas. An improved understanding of the
393 spatial extent of the zoonotic niche can only help future efforts in biosurveillance.

394

Methods

Methodological overview

A boosted regression tree (BRT) modelling framework was used to generate predictive risk maps of the zoonotic Ebola virus niche in Africa. This methodology combines regression trees, where trees are built according to optimal decision rules based on how binary decisions best accommodate a given dataset *(De'ath, 2007; Elith et al., 2008)*, and boosting, which selects the tree that minimises the loss function. In doing so, a parameter space is defined which captures the greatest amount of variation present in the dataset. In order to train the model, four component datasets were compiled: (i) a comprehensive dataset of the reported locations of Ebola virus transmission from a zoonotic reservoir to a human; (ii) a dataset of the locations of Ebola virus infections in suspected reservoir and (non-human) susceptible host species (iii) a suite of ecologically relevant environmental covariates for Africa, including predicted distribution maps of suspected reservoir bat species and (iv) background (or pseudo-absence) records representing locations where zoonotic Ebola virus has not been reported. This study was limited to the African continent since no natural outbreaks of EVD have occurred outside the continent *(CDC, 2014)*. Only *Reston ebolavirus* has a distribution reported outside of Africa, focussed in the Philippines, but has never been reported as pathogenic in humans; as a result this species was not included in the analysis.

Identifying index cases and reconstructing zoonotic transmission events in space and time

Tables detailing proven outbreaks of Ebola virus, initially sourced from the scientific literature *(Kuhn, 2008)* and from health reporting organisations *(CDC, 2014)*, were used to coordinate searches of the formal scientific literature using Web of Science and PubMed for each specific outbreak. Relevant papers were abstracted and where possible outbreak-specific epidemiological surveys were sourced. The citations in these references were obtained in order to reconstruct the outbreak in detail and to identify the most probable index case. Index cases were defined as any human infection resulting from interaction with non-human sources of the disease. Some of these cases arose from presumed interactions with zoonotic reservoirs or hosts, such as primates and other mammals during hunting trips *(Boumandouki et al., 2005; Nkoghe et al., 2005; Nkoghe et al., 2011; WHO, 2003)* or butchering of bats *(Leroy et al., 2009)*. Any cases arising from existing human infections are considered as secondary infections rather than index cases. Similar to methodology employed elsewhere *(Messina et al., 2014)*, the site, or supposed site, of this zoonotic transmission event was geopositioned using Google Earth. For locations where precise geographic information (*e.g.* geographic coordinates of a hunting camp) was provided by the authors, these were used. Where precise geographic information could not be accurately geopositioned, a geographic area (termed a polygon) was defined covering the reported region. In several cases only the first reported patient could be identified, with the source of

infection unknown. With these outbreaks the location of the first patient was geopositioned under the assumption that an initial zoonotic spillover event occurred in the vicinity of this location. In two outbreaks multiple independent zoonotic transmission events were identified *(Nkoghe et al., 2005; Pourrut et al., 2005; WHO, 2003)*, and each unique event was geopositioned and included in the model as separate entities. Table 1 catalogues the outbreaks used in this study.

Assembling a database of reported infections in animals

A literature search was conducted in Web of Science using the search term "Ebola" that returned 8,635 citations. The abstracts were examined and for those that contained possible data on animal Ebola infection, the full text was obtained. The sampling site or location of the animal in the study was identified and geopositioned using Google Maps. These locations were recorded either as precise locations or as polygons, as with human index cases. Records for which local transmission of Ebola virus was deemed unlikely (*e.g.* seropositive primates tested in containment facilities several years after their capture) were excluded from the study. The non-human Ebola virus occurrence data collected are detailed in Table 2, including the diagnostic methods used.

GenBank isolates

The open access sequence database GenBank *(NCBI, 2014)* was searched using MESH Umbrella search terms for Ebola virus, returning 181 results. These were then manually cross-referenced with the existing human and animal Ebola information, collected above, and 30 duplicates were removed. For the remaining isolates, original references and GenBank information fields were examined, but as there was insufficient information to establish precise location of isolation and/or whether the isolate represented an index case for any of these data sources, they were excluded from subsequent analyses.

Covariates assembled and used in the analyses

A suite of ecologically relevant gridded environmental covariates for Africa was compiled, each having a nominal resolution of 5km x 5km. The environmental covariates used in this analysis were: elevation (from the shuttle radar topography mission *(ORNL DAAC, 2000)*); the mean value, and a measure of spatial variation (range, described below) between 2000 and 2012 of Enhanced Vegetation Index (EVI), daytime Land Surface Temperature (LST) and night-time LST; and mean potential evapotranspiration from 1950-2000 *(Trabucco and Zomer, 2009)* (Figure 5 – figure supplement 1).

The EVI and LST datasets were derived from satellite imagery collected by NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing platform *(Tatem et al., 2004)*. EVI is a metric designed to characterise vegetation density and vigour based on the ratio of absorbed photosynthetically active radiation to near infrared radiation *(Huete et al., 2002)*. LST is a modelled product derived from emissivity as measured by the MODIS thermal sensor *(Wan and Li, 1997)*,

which is correlated, though not synonymous with air temperature, and effective for differentiating landscapes based on a combination of thermal energy and properties of surface types. The MODIS datasets utilized in this research (EVI was derived from the MCD43B4 product and the MOD11A2 LST product was used directly) were acquired as composite datasets created using imagery collected over multiple days, a procedure that results in products with eight-day temporal resolutions. Despite compositing, the EVI and LST datasets contained gaps due to persistent cloud cover found in forested equatorial regions, and these gaps were filled using a previously described approach *(Weiss et al., 2014)*. The EVI and LST datasets were then aggregated from their native 1km x 1km spatial resolution to a final 5km x 5km resolution, calculating both the mean and the range of the values of the subpixels making up each larger pixel. These spatial summaries therefore characterise both the mean temperature in each location as well as the degree of spatial heterogeneity within that pixel. This is of interest as humans and susceptible species are more likely to come into contact in transitional areas (*e.g.* boundary areas between areas of highly suitable susceptible species habitat and areas heavily utilised by humans). The final covariate production step consisted of summarising temporally across the 13-year data archive to produce synoptic datasets devoid of annual or seasonal anomalies *(Weiss et al., 2014)*.

Implicated bat reservoir distributions

Over recent years, significant research has been undertaken in investigating the role bats have to play in the transmission cycle of Ebola viruses *(Olival and Hayman, 2014)* and evidence of asymptomatic infection in fruit bats has been documented to varying extents *(Hayman et al., 2010; Hayman et al., 2012; Leroy et al., 2005; Pourrut et al., 2007; Pourrut et al., 2009)*. In order to incorporate this potential driver of Ebola virus transmission into the model we developed predicted distribution maps for three species of fruit bat implicated as primary reservoirs of the disease: *Hypsignathus monstrosus*, *Epomops franqueti* and *Myonycteris torquata*. The evidence was strongest for these three species having a reservoir role as Ebola virus RNA (all nested within the *Zaire ebolavirus* phylogeny *(Leroy et al., 2005)*) has been detected in all three. Whilst a handful of other bat species have been found to be seropositive, no further viral isolations have been recorded *(Olival and Hayman, 2014)*.

Whilst expert opinion range maps for these species exist *(Schipper et al., 2008)*, there is some disagreement with independently-sourced occurrence data (all archived in the Global Biodiversity Information Facility). As a result, a predictive modelling approach was used to create a continuous surface of habitat suitability for these species which we then included as a predictor in the model. Occurrence data for all Megachiroptera in Africa was extracted from GBIF *(GBIF, 2014)* using the R packages dismo *(Hijmans et al., 2014)* and taxize *(Chamberlain et al., 2014)*. To remove apparently erroneous records in the GBIF archive all occurrence records more than 100km from the species known ranges, as determined by expert-opinion range maps *(Schipper et al., 2008)*, were excluded, as

were duplicate records and those located in the ocean. This resulted in a total dataset of 1341 unique occurrence records.

The occurrence database was then used to train separate boosted regression tree species distribution models *(Elith et al., 2008)* to predict the likely distribution of each of these suspected reservoir species. For each model, occurrence records for the target species (*H. monstrosus*, n=67; *E. franqueti*, n=120; and *M. torquata*, n=52) were considered presence records and occurrence records of all other species were used as background records. This procedure is designed to account for the potentially biasing effect of spatial variation in recording of Megachiroptera occurrences *(Phillips et al., 2009)*.

For each species we ran fifty submodels each trained to a randomly selected bootstrap of this dataset, subject to the constraint that each bootstrap contained a minimum of ten occurrence and ten background records. Each submodel was fitted using the gbm.step subroutine *(Elith et al., 2008)* in the dismo R package. In each submodel the background records were down weighted so that the weighted sum of presence records equalled the weighted sum of background records *(Barbet-Massin et al., 2012)* in order to maximise the discrimination capacity of the model. We generated a prediction map from each of these submodels and calculated both the mean prediction and 95% confidence interval around the prediction for each 5km x 5km pixel for each species.

Model accuracy was assessed by calculating the mean area under the curve (AUC) statistic for each submodel under a stringent ten-fold cross validation for each submodel and obtaining the mean and standard deviation across all fifty submodels. Under this procedure the dataset was split into ten subsets, each containing approximately equal numbers of presence and background points. The ability of a model trained on each subset to predict the distribution of the other 90% of records was assessed by AUC and the mean value taken. As so few presence records were used to train each fold model (*i.e.* around five presence records for *M. torquata* up to twelve for *E. franqueti*), this represents a very stringent test of the model's predictive capacity. Additionally, to prevent inflation of the accuracy statistics due to spatial sorting bias, these statistics were estimated using a pairwise distance sampling procedure *(Hijmans, 2012)*. Consequently, the AUC statistics presented here are lower than would be returned by standard procedures but gives a more realistic quantification of the model's ability to extrapolate predictions to new regions *(Wenger and Olden, 2012)*. We also generated marginal effect plots with associated uncertainty intervals and relative contribution statistics (how often each covariate was selected during the model fitting process) as quantification of the sensitivity of the model to the different covariates. These allow us to make inferences about the ecological relationship between each species and its environment as well as to identify where this relationship is most uncertain.

To generate a single surface describing the distribution of the bat reservoir species to be used as a covariate in the subsequent Ebola modelling, the three mean prediction distribution maps were merged by taking the average habitat suitability for each of the three bat species at each pixel.

Ebola distribution modelling

The Ebola virus occurrence dataset was supplemented with a background record dataset generated by randomly sampling 10,000 locations across Africa, biased towards more populous areas as a proxy for reporting bias *(Phillips et al., 2009)*. We fitted 500 submodels to bootstraps of this dataset. To account for uncertainty in the geographic location of those occurrences reported as polygons, for each submodel one point was randomly selected from each of these occurrence polygons. This Monte Carlo procedure enabled the model to efficiently integrate over the environmental uncertainty associated with imprecise geographic data. A bootstrap sample was then taken from each of these datasets and used to train the BRT model using the same procedure and weighting of background records as for the bat distribution models. Similarly, we generated a prediction map from each of these models and calculated both the mean prediction and corresponding 95% confidence intervals for each pixel and analysed prediction accuracy using the same stringent cross validation and sensitivity analysis procedure as for the bat distribution models (detailed above).

The predicted distribution map produced by this approach represents the environmental suitability of each pixel for zoonotic Ebola virus transmission. This may be interpreted as a relative probability of presence in the sense that more suitable pixels are more likely to contain zoonotic transmission than less suitable pixels, though not an absolute probability that transmission occurs in a given pixel. Similarly, the presence of zoonotic transmission increases the risk of transmission to a human, though this is also contingent on how humans interact with these zoonotic pools, through hunting or other activities.

Population living in areas of environmental suitability for zoonotic transmission.

In order to identify areas which are likely to be at risk of transmission of *Ebolavirus* from zoonotic reservoir hosts to humans, the continuous map of the predicted environmental suitability for zoonotic transmission (shown in Figure 5) was converted into a binary map classifying pixels as either at risk or not at risk. A pixel was assumed to be at risk if its predicted environmental suitability for zoonotic Ebola virus transmission was greater than 0.673, the lowest suitability value predicted at the locations of known transmission to humans (point records of human index cases). Countries containing at least one at-risk pixel are shown in Figure 5B – those countries that previously report an index case were defined as Set 1; countries with at least one at-risk pixel with no previous index cases of EVD were categorised as Set 2. The number of people living in at-risk areas in each of these countries was calculated by summing the estimated population of at-risk pixels using population density maps from

564 the AfriPop project *(Linard et al., 2012; WorldPop, 2014)* and the proportion of those living in urban,
565 periurban and rural areas was evaluated using the Global Rural Urban Mapping Project
566 *(CIESIN/IFPRI/WB/CIAT, 2007)*.

567 The R code used for all of the analysis has been made available on an open source basis
568 (https://github.com/SEEG-Oxford/ebola_zoonotic).

569 ## National level demographic and mobility data

570 For three separate years (1976, 2000 and 2014), total national populations were retrieved and the
571 proportion of rural to urban populations noted from World Bank statistics *(World Bank, 2014)*. To
572 describe global air travel patterns from Set 1 and Set 2 countries, flight schedules data from the
573 Official Airline Guide, reflecting an estimated 95% of all commercial flights worldwide, were
574 analysed between 2000 and 2013 to calculate the annual volume of seats on direct flights that depart
575 from each predicted country and which have an international destination. Complementing these seat
576 capacity data, worldwide data on anonymised, individual passenger flight itineraries from the
577 International Air Transport Association (2012) *(IATA, 2014)* were analysed between 2005 and 2012 to
578 calculate the annual volume of international passenger departures out of each Set 1 and Set 2 country.
579 The IATA dataset represents an estimated 93% of the world's commercial air traffic at the passenger
580 level and includes points of departure and arrival and final destination information for travellers as
581 well as their connecting flights.

Acknowledgments

We thank Katherine Battle, Maria Devine and Kirsten Duda for proof-reading and Jane Messina for creating Figure 1. We also thank Andrew Rambaut for his comments on the final draft.

Author contributions

DMP – advised and assembled the outbreak and animal infection data and wrote the manuscript

NG – extracted the bat data and conducted all of the analysis

AM – assembled and geo-positioned the outbreak and animal infection data

ZH – geo-positioned the outbreak and animal infection data and provided all maps

AJH – analysed international flight data and helped assemble maps

DJW – assembled the covariate layers

OJB – screened GenBank data, provided Figure 2A and edited drafts of the manuscript

MUGK – edited drafts of the manuscript

DLS – edited drafts of the manuscript

CLM – edited drafts of the manuscript

SB – extracted Ebola information from GenBank

PWG – assembled the covariate layers

PWH – advised on international public health context and edited the final draft

IIB – assisted with international transportation analysis and edited the final draft

JSB – advised on international public health context and edited the final draft

SRM – assisted in geopositioning and contributed to the manuscript

AJT – provided information on urban change and migration data

KK – provided data and conducted all analyses on international air traffic patterns

SIH – conceived the work and analysis, wrote content and edited the manuscript at all stages of development. He acts as guarantor of the paper.

Funding

DMP is funded by a Sir Richard Southwood Graduate Scholarship from the Department of Zoology at the University of Oxford. NG is funded by a grant from the Bill & Melinda Gates Foundation (#OPP1053338). ZH and AH are funded by the Bill & Melinda Gates Foundation (#OPP1106023). OJB is funded by a BBSRC studentship. PWG is a Medical Research Council (UK) Career Development Fellow (#K00669X) and receives support from the Bill & Melinda Gates Foundation (#OPP1068048) which also supports DJW and SB. MUGK is funded by the German Academic Exchange Service (DAAD) through a graduate scholarship. JSB and SRM acknowledge funding from NIH National Library of Medicine (R01LM010812) and from the Bill & Melinda Gates Foundation (#OPP1093011). AJT is supported by funding from NIH/NIAD (U19AI089674) and the Bill & Melinda Gates Foundation (#OPP1106427, #OPP1032350). PWH is funded by the EU FP7 project PREPARE (#602525) and the Li Ka Shing Foundation. KK and IIB acknowledge the support of the Canadian Institutes of Health Research. SIH is funded by a Senior Research Fellowship from the Wellcome Trust (#095066) which also supports AM and a grant from the Bill & Melinda Gates Foundation (#OPP1093011) which also supports CLM. SIH and AJT would also like to acknowledge funding support from the RAPIDD program of the Science & Technology Directorate, Department of Homeland Security, and the Fogarty International Center, National Institutes of Health. Funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

The authors have declared that no competing interests exist.

Abbreviations

AFRO	African Region (WHO)
AMRO	Region of the Americas (WHO)
AUC	Area under the curve
BDBV	*Bundibugyo ebolavirus*
CAR	Central African Republic
DRC	Democratic Republic of the Congo
EBOV	*Zaire ebolavirus*
EMRO	Eastern Mediterranean Region (WHO)
EURO	European Region (WHO)

EVD	Ebola virus disease
EVI	Enhanced vegetation index
LST	Land surface temperature
OWFB	Old World fruit bat
PCR	Polymerase chain reaction
PET	Potential evapotranspiration
RC	Relative contribution
ROC	Republic of Congo
SEARO	South-East Asia Region (WHO)
SUDV	*Sudan ebolavirus*
TAFV	*Tai Forest ebolavirus*
WPRO	Western Pacific Region (WHO)

References

Amblard J, Obiang P, Edzang S, Prehaud C, Bouloy M, et al. 1997. Identification of the Ebola virus in Gabon in 1994. *Lancet* **349**: 181-182. doi:10.1016/S0140-6736(05)60984-1.

Baize S, Pannetier D, Oestereich L, Rieger T, Koivogui L, et al. 2014. Emergence of Zaire Ebola virus disease in Guinea - preliminary report. *N Engl J Med*: doi:10.1056/NEJMoa1404505.

Bannister B. 2010. Viral haemorrhagic fevers imported into non-endemic countries: risk assessment and management. *Brit Med Bull* **95**: 193-225. doi:10.1093/Bmb/Ldq022.

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol Evol* **3**: 327-338. doi:10.1111/j.2041-210X.2011.00172.x.

Baron RC, McCormick JB, Zubeir OA. 1983. Ebola virus disease in southern Sudan - hospital dissemination and intrafamilial spread. *Bull World Health Organ* **61**: 997-1003.

Barrette RW, Metwally SA, Rowland JM, Xu LZ, Zaki SR, et al. 2009. Discovery of swine as a host for the Reston ebolavirus. *Science* **325**: 204-206. doi:10.1126/science.1172705.

Bausch DG, Schwarz L. 2014. Outbreak of Ebola virus disease in Guinea: where ecology meets economy. *PLoS Negl Trop Dis* **8**: e3056. doi:10.1371/journal.pntd.0003056.

Beer B, Kurth R, Bukreyev A. 1999. Characteristics of *Filoviridae*: Marburg and Ebola viruses. *Naturwissenschaften* **86**: 8-17. doi:10.1007/s001140050562.

Bermejo M, Rodriguez-Teijeiro JD, Illera G, Barroso A, Vila C, et al. 2006. Ebola outbreak killed 5000 gorillas. *Science* **314**: 1564-1564. doi:10.1126/science.1133105.

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, et al. 2013. The global distribution and burden of dengue. *Nature* **496**: 504-507. doi:10.1038/Nature12060.

Boumandouki P, Formenty P, Epelboin A, Campbell P, Atsangandoko C, et al. 2005. [Clinical management of patients and deceased during the Ebola outbreak from October to December 2003 in Republic of Congo]. *Bull Soc Pathol Exot* **98**: 218-223.

Brady OJ, Hay SI, Horby P. 2014. Estimating vaccine and drug requirements for Ebola outbreaks. *Nature* **512**: 7514.

Brashares JS, Golden CD, Weinbaum KZ, Barrett CB, Okello GV. 2011. Economic and geographic drivers of wildlife consumption in rural Africa. *Proc Natl Acad Sci U S A* **108**: 13931-13936. doi:10.1073/pnas.1011526108.

Briand S, Bertherat E, Cox P, Formenty P, Kieny MP, et al. 2014. The international Ebola emergency. *N Engl J Med*: doi:10.1056/NEJMp1409858.

Brockmann D, Helbing D. 2013. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**: 1337-1342. doi:10.1126/science.1245200.

Caillaud D, Levrero F, Cristescu R, Gatti S, Dewas M, et al. 2006. Gorilla susceptibility to Ebola virus: the cost of sociality. *Curr Biol* **16**: R489-491. doi:10.1016/j.cub.2006.06.017.

CDC. Chronology of Ebola hemorrhagic fever outbreaks. Available: http://www.cdc.gov/vhf/ebola/resources/outbreak-table.html. Accessed: August 2014

Chamberlain S, Szocs E, Boettiger C, Ram K, Bartomeus I, et al. taxize: taxonomic information from around the web. R package. Available: https://github.com/ropensci/taxize. Accessed: August 2014

Chan M. 2014. Ebola virus disease in West Africa - no early end to the outbreak. *N Engl J Med*: online. doi:10.156/NEJMp1409859.

Chowell G, Hengartner NW, Castillo-Chavez C, Fenimore PW, Hyman JM. 2004. The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda. *J Theor Biol* **229**: 119-126. doi:10.1016/j.jtbi.2004.03.006.

CIESIN/IFPRI/WB/CIAT. Global Rural Urban Mapping Project (GRUMP) : Gridded Population of the World, version 3. Available: http://sedac.ciesin.columbia.edu/gpw. Accessed: December 2013

Cohen B. 2004. Urban growth in developing countries: a review of current trends and a caution regarding existing forecasts. *World Dev* **32**: 23-51. doi:10.1016/j.worlddev.2003.04.008.

Conrad JL, Isaacson M, Smith EB, Wulff H, Crees M, et al. 1978. Epidemiologic investigation of Marburg virus disease, southern Africa, 1975. *Am J Trop Med Hyg* **27**: 1210-1215.

706 Daszak P. 2000. Emerging infectious diseases of wildlife - threats to biodiversity and human health.
707 *Science* **287**: 1756-1756. doi:10.1126/science.287.5452.443.
708 De'ath G. 2007. Boosted trees for ecological modeling and prediction. *Ecology* **88**: 243-251.
709 doi:10.1890/0012-9658(2007)88[243:Btfema]2.0.Co;2.
710 Dudas G, Rambaut A. 2014. Phylogenetic analysis of Guinea 2014 EBOV ebolavirus outbreak. *PLoS*
711 *Curr* **6**: ecurrents.outbreaks.84eefe85ce43ec89dc80bf0670f0677b0678b0417d.
712 doi:10.1371/currents.outbreaks.84eefe5ce43ec9dc0bf0670f7b8b417d.
713 ECDC. 2014. Outbreak of Ebola virus disease in West Africa. Stockholm: ECDC. 17 p.
714 Elith J, Leathwick JR. 2009. Species distribution models: ecological explanation and prediction across
715 space and time. *Annu Rev Ecol Evol Syst* **40**: 677-697.
716 doi:10.1146/annurev.ecolsys.110308.120159.
717 Elith J, Leathwick JR, Hastie T. 2008. A working guide to boosted regression trees. *J Anim Ecol* **77**:
718 802-813. doi:10.1111/j.1365-2656.2008.01390.x.
719 Fauci AS. 2014. Ebola - underscoring the global disparities in health care resources. *N Engl J Med*:
720 doi:10.1056/NEJMp1409494.
721 Feldmann H, Geisbert TW. 2011. Ebola haemorrhagic fever. *Lancet* **377**: 849-862.
722 doi:10.1016/S0140-6736(10)60667-8.
723 Formenty P, Boesch C, Wyers M, Steiner C, Donati F, et al. 1999. Ebola virus outbreak among wild
724 chimpanzees living in a rain forest of Cote d'Ivoire. *J Infect Dis* **179**: S120-S126.
725 doi:10.1086/514296.
726 Francesconi P, Yoti Z, Declich S, Onek PA, Fabiani M, et al. 2003. Ebola hemorrhagic fever
727 transmission and risk factors of contacts, Uganda. *Emerg Infect Dis* **9**: 1430-1437.
728 doi:10.3201/eid0911.030339.
729 Franklin J. 2009. Mapping species distributions. Cambridge: Cambridge University Press. 320 p.
730 Garcia AJ, Pindolia DK, Lopiano KK, Tatem AJ. 2014. Modeling internal migration flows in sub-
731 Saharan Africa using census microdata. *Migrat Stud*: in press.
732 doi:10.1093/migration/mnu036.
733 GBIF. Global biodiversity information facility. Available: http://www.gbif.org/. Accessed: August
734 2014
735 Gear JSS, Cassel GA, Gear AJ, Trappler B, Clausen L, et al. 1975. Outbreak of Marburg virus disease
736 in Johannesburg. *Br Med J* **4**: 489-493. doi:10.1136/bmj.4.5995.489.
737 Georges-Courbot MC, Sanchez A, Lu CY, Baize S, Leroy E, et al. 1997. Isolation and phylogenetic
738 characterization of Ebola viruses causing different outbreaks in Gabon. *Emerg Infect Dis* **3**:
739 59-62. doi:10.3201/eid0301.970107.
740 Georges AJ, Leroy EM, Renaut AA, Benissan CT, Nabias RJ, et al. 1999. Ebola hemorrhagic fever
741 outbreaks in Gabon, 1994-1997: epidemiologic and health control issues. *J Infect Dis* **179**:
742 S65-S75. doi:10.1086/514290.
743 Gething PW, Patil AP, Smith DL, Guerra CA, Elyazar IRF, et al. 2011. A new world malaria map:
744 *Plasmodium falciparum* endemicity in 2010. *Malar J* **10**: 378. doi:10.1186/1475-2875-10-
745 378.
746 Gilbert M, Golding N, Zhou H, Wint GR, Robinson TP, et al. 2014. Predicting the risk of avian
747 influenza A H7N9 infection in live-poultry markets across Asia. *Nat Commun* **5**: 4116.
748 doi:10.1038/ncomms5116.
749 Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, et al. 2014. Genomic surveillance elucidates
750 Ebola virus origin and transmission during the 2014 outbreak. *Science*: in press.
751 Gonzalez MC, Hidalgo CA, Barabasi AL. 2008. Understanding individual human mobility patterns.
752 *Nature* **453**: 779-782. doi:10.1038/Nature06958.
753 Goodman JL. 2014. Studying "secret serums" - toward safe, effective Ebola treatments. *N Engl J*
754 *Med*: 10.1056/NEJMp1409817.
755 Gostin LO, Lucey D, Phelan A. 2014. The ebola epidemic: a global health emergency. *JAMA*: E1-2.
756 doi:10.1001/jama.2014.11176.
757 Grard G, Biek R, Tamfum JJM, Fair J, Wolfe N, et al. 2011. Emergence of divergent Zaire Ebola
758 virus strains in Democratic Republic of the Congo in 2007 and 2008. *J Infect Dis* **204**: S776-
759 S784. doi:10.1093/infdis/jir364.

760 Groseth A, Feldmann H, Strong JE. 2007. The ecology of Ebola virus. *Trends Microbiol* **15**: 408-416.
761 doi:10.1016/j.tim.2007.08.001.
762 Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, et al. 2013. Global mapping of infectious
763 disease. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120250. doi:10.1098/Rstb.2012.0250.
764 Hay SI, George DB, Moyes CL, Brownstein JS. 2013. Big data opportunities for global infectious
765 disease surveillance. *PLoS Med* **10**: e1001413. doi:10.1371/journal.pmed.1001413.
766 Hay SI, Tatem AJ, Graham AJ, Goetz SJ, Rogers DJ. 2006. Global environmental data for mapping
767 infectious disease distribution. *Adv Parasit* **62**: 37-77. doi:10.1016/S0065-308x(05)62002-7.
768 Hayman DTS, Emmerich P, Yu M, Wang LF, Suu-Ire R, et al. 2010. Long-term survival of an urban
769 fruit bat seropositive for Ebola and Lagos Bat viruses. *PLoS One* **5**: e11978.
770 doi:10.1371/journal.pone.0011978.
771 Hayman DTS, Yu M, Crameri G, Wang LF, Suu-Ire R, et al. 2012. Ebola virus antibodies in fruit
772 bats, Ghana, West Africa. *Emerg Infect Dis* **18**: 1207-1209. doi:10.3201/eid1807.111654.
773 Hewlett BS, Epelboin A, Hewlett BL, Formenty P. 2005. Medical anthropology and Ebola in Congo:
774 cultural models and humanistic care. *Bull Soc Pathol Exot* **98**: 230-236.
775 Heymann DL, Weisfeld JS, Webb PA, Johnson KM, Cairns T, et al. 1980. Ebola hemorrhagic fever:
776 Tandala, Zaire, 1977-1978. *J Infect Dis* **142**: 372-376. doi:10.1093/infdis/142.3.372.
777 Hijmans RJ. 2012. Cross-validation of species distribution models: removing spatial sorting bias and
778 calibration with a null model. *Ecology* **93**: 679-688.
779 Hijmans RJ, Phillips S, Leathwick J, Elith J. dismo. R package. Available: http://cran.r-
780 project.org/web/packages/dismo/dismo.pdf. Accessed: August 2014
781 Huang ZJ, Tatem AJ. 2013. Global malaria connectivity through air travel. *Malar J* **12**: 269.
782 doi:10.1186/1475-2875-12-269.
783 Huete A, Didan K, Miura T, Rodriguez EP, Gao X, et al. 2002. Overview of the radiometric and
784 biophysical performance of the MODIS vegetation indices. *Remote Sens Environ* **83**: 195-
785 213. doi:10.1016/S0034-4257(02)00096-2.
786 Hufnagel L, Brockmann D, Geisel T. 2004. Forecast and control of epidemics in a globalized world.
787 *Proc Natl Acad Sci USA* **101**: 15124-15129. doi:10.1073/pnas.0308344101.
788 IATA. International Air Transport Association. Available: http://www.iata.org/Pages/default.aspx.
789 Accessed: August 2014
790 International Commission. 1978. Ebola haemorrhagic fever in Zaire, 1976. *Bull World Health Organ*
791 **56**: 271-293.
792 Jahrling PB, Geisbert TW, Dalgard DW, Johnson ED, Ksiazek TG, et al. 1990. Preliminary report -
793 isolation of Ebola virus from monkeys imported to USA. *Lancet* **335**: 502-505.
794 doi:10.1016/0140-6736(90)90737-P.
795 Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, et al. 2008. Global trends in emerging
796 infectious diseases. *Nature* **451**: 990-994. doi:10.1038/Nature06536.
797 Kamins AO, Restif O, Ntiamoa-Baidu Y, Suu-Ire R, Hayman DT, et al. 2011. Uncovering the fruit bat
798 bushmeat commodity chain and the true extent of fruit bat hunting in Ghana, West Africa.
799 *Biol Conserv* **144**: 3000-3008. doi:10.1016/j.biocon.2011.09.003.
800 Karesh WB, Dobson A, Lloyd-Smith JO, Lubroth J, Dixon MA, et al. 2012. Ecology of zoonoses:
801 natural and unnatural histories. *Lancet* **380**: 1936-1945. doi:10.1016/S0140-6736(12)61678-
802 X.
803 Khan AS, Tshioko FK, Heymann DL, Le Guenno B, Nabeth P, et al. 1999. The reemergence of Ebola
804 hemorrhagic fever, Democratic Republic of the Congo, 1995. *J Infect Dis* **179**: S76-S86.
805 doi:10.1086/514306.
806 King AMQ, Adams MJ, Carstens EB, Lefkowitz E, editors (2011) Virus taxonomy: ninth report of the
807 International Committee on Taxonomy of Viruses. London: Elsevier. 1338 p.
808 Kuhn JH. 2008. Filoviruses: a compendium of 40 years of epidemiological, clinical, and laboratory
809 studies. Vienna: Springer-Verlag. 413 p.
810 Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, et al. 2010. Proposal for a revised
811 taxonomy of the family *Filoviridae*: classification, names of taxa and viruses, and virus
812 abbreviations. *Arch Virol* **155**: 2083-2103. doi:10.1007/s00705-010-0814-x.

813 Lahm SA, Kombila M, Swanepoel R, Barnes RF. 2007. Morbidity and mortality of wild animals in
814 relation to outbreaks of Ebola haemorrhagic fever in Gabon, 1994-2003. *Trans R Soc Trop*
815 *Med Hyg* **101**: 64-78. doi:10.1016/j.trstmh.2006.07.002.
816 Lamunu M, Lutwama JJ, Kamugisha J, Opio A, Nambooze J, et al. 2004. Containing a haemorrhagic
817 fever epidemic: the Ebola experience in Uganda (October 2000-January 2001). *Int J Infect*
818 *Dis* **8**: 27-37. doi:10.1016/j.ijid.2003.04.001.
819 Larkin M. 2003. Ebola outbreak in the news. *Lancet Infect Dis* **3**: 255. doi:10.1016/S1473-
820 3099(03)00584-X.
821 Le Guenno B, Formenty P, Wyers M, Gounon P, Walker F, et al. 1995. Isolation and partial
822 characterization of a new strain of Ebola virus. *Lancet* **345**: 1271-1274. doi:10.1016/S0140-
823 6736(95)90925-7.
824 Legrand J, Grais RF, Boelle PY, Valleron AJ, Flahault A. 2007. Understanding the dynamics of Ebola
825 epidemics. *Epidemiol Infect* **135**: 610-621. doi:10.1017/S0950268806007217.
826 Leroy EM, Baize S, Volchkov VE, Fisher-Hoch SP, Georges-Courbot MC, et al. 2000. Human
827 asymptomatic Ebola infection and strong inflammatory response. *Lancet* **355**: 2210-2215.
828 doi:10.1016/S0140-6736(00)02405-3.
829 Leroy EM, Epelboin A, Mondonge V, Pourrut X, Gonzalez JP, et al. 2009. Human Ebola outbreak
830 resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007.
831 *Vector Borne Zoonotic Dis* **9**: 723-728. doi:10.1089/vbz.2008.0167.
832 Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, et al. 2005. Fruit bats as reservoirs of
833 Ebola virus. *Nature* **438**: 575-576. doi:10.1038/438575a.
834 Leroy EM, Souquiere S, Rouquet P, Drevet D. 2002. Re-emergence of ebola haemorrhagic fever in
835 Gabon. *Lancet* **359**: 712-712. doi:10.1016/S0140-6736(02)07796-6.
836 Linard C, Gilbert M, Snow RW, Noor AM, Tatem AJ. 2012. Population distribution, settlement
837 patterns and accessibility across Africa in 2010. *PLoS One* **7**: e31743.
838 doi:10.1371/journal.pone.0031743.
839 Linard C, Tatem AJ, Gilbert M. 2013. Modelling spatial patterns of urban growth in Africa. *Appl*
840 *Geogr* **44**: 23-32. doi:10.1016/j.apgeog.2013.07.009.
841 MacNeil A, Farnon EC, Wamala J, Okware S, Cannon DL, et al. 2010. Proportion of deaths and
842 clinical features in Bundibugyo Ebola virus infection, Uganda. *Emerg Infect Dis* **16**: 1969-
843 1972. doi:10.3201/eid1612.100627.
844 Messina JP, Brady OJ, Pigott DM, Brownstein JS, Hoen AG, et al. 2014. A global compendium of
845 human dengue virus occurrence. *Scientific Data* **1**: 140004. doi:10.1038/sdata.2014.4.
846 Mfunda IM, Røskaft E. 2010. Bushmeat hunting in Serengeti, Tanzania: an important economic
847 activity to local people. *Int J Biodivers Conserv* **2**: 263-272.
848 Milleliri JM, Tevi-Benissan C, Baize S, Leroy E, Georges-Courbot MC. 2004. [Epidemics of Ebola
849 haemorrhagic fever in Gabon (1994-2002). Epidemiologic aspects and considerations on
850 control measures]. *Bull Soc Pathol Exot* **97**: 199-205.
851 Miranda MEG, White ME, Dayrit MM, Hayes CG, Ksiazek TG, et al. 1991. Seroepidemiological
852 study of filovirus related to Ebola in the Philippines. *Lancet* **337**: 425-426. doi:10.1016/0140-
853 6736(91)91199-5.
854 MSF. Ebola epidemic confirmed in Democratic Republic of Congo: MSF sends specialists and
855 material to the epicentre. Available: http://www.msf.org/article/ebola-epidemic-confirmed-
856 democratic-republic-congo-msf-sends-specialists-and-material. Accessed: August 2014
857 Murray CJ, Ortblad KF, Guinovart C, Lim SS, Wolock TM, et al. 2014. Global, regional, and national
858 incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic
859 analysis for the Global Burden of Disease Study 2013. *Lancet*: doi:10.1016/S0140-
860 6736(14)60844-8.
861 Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, et al. 2012. Disability-adjusted life years
862 (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the
863 Global Burden of Disease Study 2010. *Lancet* **380**: 2197-2223. doi:10.1016/S0140-
864 6736(12)61689-4.
865 Muyembe T, Kipasa M. 1995. Ebola haemorrhagic fever in Kikwit, Zaire. *Lancet* **345**: 1448.
866 doi:10.1016/S0140-6736(95)92640-2.
867 NCBI. GenBank - Ebola. Available: http://www.ncbi.nlm.nih.gov/genbank/. Accessed: August 2014

Negredo A, Palacios G, Vazquez-Moron S, Gonzalez F, Dopazo H, et al. 2011. Discovery of an *Ebolavirus*-like Filovirus in Europe. *PLoS Pathog* **7**: e1002304. doi:10.1371/journal.ppat.1002304.
NIH. NIAID emering infectious diseases/pathogens. Available: Accessed: August 2014
Nkoghe D, Formenty P, Leroy EM, Nnegue S, Edou SY, et al. 2005. [Multiple Ebola virus haemorrhagic fever outbreaks in Gabon, from October 2001 to April 2002]. *Bull Soc Pathol Exot* **98**: 224-229.
Nkoghe D, Kone ML, Yada A, Leroy E. 2011. A limited outbreak of Ebola haemorrhagic fever in Etoumbi, Republic of Congo, 2005. *Trans R Soc Trop Med Hyg* **105**: 466-472. doi:10.1016/j.trstmh.2011.04.011.
Okware SI, Omaswa FG, Zaramba S, Opio A, Lutwama JJ, et al. 2002. An outbreak of Ebola in Uganda. *Trop Med Int Health* **7**: 1068-1075. doi:10.1046/j.1365-3156.2002.00944.x.
Olival KJ, Hayman DTS. 2014. Filoviruses in bats: current knowledge and future directions. *Viruses* **6**: 1759-1788. doi:10.3390/V6041759.
Onyango CO, Opoka ML, Ksiazek TG, Formenty P, Ahmed A, et al. 2007. Laboratory diagnosis of Ebola hemorrhagic fever during an outbreak in Yambio, Sudan, 2004. *J Infect Dis* **196**: S193-S198. doi:10.1086/520609.
ORNL DAAC. Shuttle radar topography mission near-global digital elevation models. Available: http://webmap.ornl.gov/wcsdown/wcsdown.jsp?dg_id=10008_1. Accessed: August 2014
Pattyn S, Jacob W, Vandergroen G, Piot P, Courteille G. 1977. Isolation of Marburg-like virus from a case of hemorrhagic fever in Zaire. *Lancet* **1**: 573-574. doi:10.1016/S0140-6736(77)92002-5.
Peterson AT, Bauer JT, Mills JN. 2004. Ecologic and geographic distribution of filovirus disease. *Emerg Infect Dis* **10**: 40-47. doi:10.3201/eid1001.030125.
Peterson AT, Carroll DS, Mills JN, Johnson KM. 2004. Potential mammalian filovirus reservoirs. *Emerg Infect Dis* **10**: 2073-2081. doi:10.3201/eid1012.040346.
Peterson AT, Papes M, Carroll DS, Leirs H, Johnson KM. 2007. Mammal taxa constituting potential coevolved reservoirs of filoviruses. *J Mammal* **88**: 1544-1554. doi:10.1644/06-Mamm-a-280r1.1.
Petter J-J, Desbordes F. 2013. Primates of the world. Princeton: Princeton University Press. 186 p.
Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A, et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl* **19**: 181-197. doi:10.1890/07-2153.1.
Pigott DM, Bhatt S, Golding N, Duda KA, Battle KE, et al. 2014. Global distribution maps of the leishmaniases. *eLife*: e02851. doi:10.7554/eLife.02851.
Pindolia DK, Garcia AJ, Huang ZJ, Fik T, Smith DL, et al. 2014. Quantifying cross-border movements and migrations for guiding the strategic planning of malaria control and elimination. *Malar J* **13**: 169. doi:10.1186/1475-2875-13-169.
Pourrut X, Delicat A, Rollin PE, Ksiazek TG, Gonzalez JP, et al. 2007. Spatial and temporal patterns of *Zaire ebolavirus* antibody prevalence in the possible reservoir bat species. *J Infect Dis* **196**: S176-S183. doi:10.1086/520541.
Pourrut X, Kumulungui B, Wittmann T, Moussavou G, Delicat A, et al. 2005. The natural history of Ebola virus in Africa. *Microbes Infect* **7**: 1005-1014. doi:10.1016/j.micinf.2005.04.006.
Pourrut X, Souris M, Towner JS, Rollin PE, Nichol ST, et al. 2009. Large serological survey showing cocirculation of Ebola and Marburg viruses in Gabonese bat populations, and a high seroprevalence of both viruses in *Rousettus aegyptiacus*. *BMC Infect Dis* **9**: e159. doi:10.1186/1471-2334-9-159.
Rouquet P, Froment JM, Bermejo M, Kilbourn A, Karesh W, et al. 2005. Wild animal mortality monitoring and human Ebola outbreaks, Gabon and Republic of Congo, 2001-2003. *Emerg Infect Dis* **11**: 283-290. doi:10.3201/eid1102.040533.
Schipper J, Chanson JS, Chiozza F, Cox NA, Hoffmann M, et al. 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science* **322**: 225-230. doi:10.1126/science.1165115.
Seto KC, Guneralp B, Hutyra LR. 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc Natl Acad Sci USA* **109**: 16083-16088. doi:10.1073/pnas.1211658109.

923 Shoemaker T, MacNeil A, Balinandi S, Campbell S, Wamala JF, et al. 2012. Reemerging Sudan
924 Ebola virus disease in Uganda, 2011. *Emerg Infect Dis* **18**: 1480-1483.
925 doi:10.3201/eid1809.111536.
926 Siegert R, Shu H-L, Slenczka W, Peters D, Mueller G. 1967. Zur Ätiologie einer unbekannten, von
927 Affen ausgegangenen menschlichen Infektionskrankheit. *Dtsch med Wochenschr* **92**: 2341-
928 2343. doi:10.1055/s-0028-1106144.
929 Simini F, Gonzalez MC, Maritan A, Barabasi AL. 2012. A universal model for mobility and
930 migration patterns. *Nature* **484**: 96-100. doi:10.1038/nature10856.
931 Simini F, Maritan A, Neda Z. 2013. Human mobility in a continuum approach. *PLoS One* **8**: e60069.
932 doi:10.1371/journal.pone.0060069.
933 Sinka ME, Bangs MJ, Manguin S, Rubio-Palis Y, Chareonviriyaphap T, et al. 2012. A global map of
934 dominant malaria vectors. *Parasit Vectors* **5**: 69. doi:10.1186/1756-3305-5-69.
935 Smith DH, Isaacson M, Johnson KM, Bagshawe A, Johnson BK, et al. 1982. Marburg virus disease in
936 Kenya. *Lancet* **1**: 816-820. doi:10.1016/S0140-6736(82)91871-2.
937 Stockwell D, Peters D. 1999. The GARP modelling system: problems and solutions to automated
938 spatial prediction. *Int J Geogr Inf Sci* **13**: 143-158. doi:10.1080/136588199241391.
939 Stoddard ST, Morrison AC, Vazquez-Prokopec GM, Soldan VP, Kochel TJ, et al. 2009. The role of
940 human movement in the transmission of vector-borne pathogens. *PLoS Negl Trop Dis* **3**:
941 e481. doi:10.1371/journal.pntd.0000481.
942 Talbi C, Lemey P, Suchard MA, Abdelatif E, Elharrak M, et al. 2010. Phylodynamics and human-
943 mediated dispersal of a zoonotic virus. *PLoS Pathog* **6**: e1001166.
944 doi:10.1371/journal.ppat.1001166.
945 Tatem AJ, Goetz SJ, Hay SI. 2004. Terra and aqua: new data for epidemiology and public health. *Int J*
946 *Appl Earth Obs Geoinf* **6**: 33-46.
947 Tignor GH, Casals J, Shope RE. 1993. The Yellow Fever epidemic in Ethiopia, 1961-1962 -
948 retrospective serological evidence for concomitant Ebola or Ebola-like virus infection. *Trans*
949 *R Soc Trop Med Hyg* **87**: 162-162. doi:10.1016/0035-9203(93)90471-2.
950 Towner JS, Amman BR, Sealy TK, Carroll SAR, Comer JA, et al. 2009. Isolation of genetically
951 diverse Marburg viruses from Egyptian fruit bats. *PLoS Pathog* **5**: e1000536.
952 doi:10.1371/journal.ppat.1000536.
953 Towner JS, Khristova ML, Sealy TK, Vincent MJ, Erickson BR, et al. 2006. Marburgvirus genomics
954 and association with a large hemorrhagic fever outbreak in Angola. *J Virol* **80**: 6497-6516.
955 doi:10.1128/JVI.00069-06.
956 Towner JS, Sealy TK, Khristova ML, Albarino CG, Conlan S, et al. 2008. Newly discovered Ebola
957 virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog* **4**: e1000212.
958 doi:10.1371/journal.ppat.1000212.
959 Trabucco A, Zomer RJ. Global Aridity Index (Global-Aridity) and Global Potential Evapo-
960 Transpiration (Global-PET) Geospatial Database. Available: http://www.csi.cgiar.org.
961 Accessed: August 2014
962 Walsh PD, Bermejo M, Rodriguez-Teijeiro JD. 2009. Disease avoidance and the evolution of primate
963 social connectivity: Ebola, bats, gorillas, and chimpanzees. In: Huffman MA, Chapman CA,
964 editors. Primate parasite ecology: the dynamics and study of host-parasite relationships.
965 Cambridge: Cambridge University Press. pp. 183-197.
966 Wamala JF, Lukwago L, Malimbo M, Nguku P, Yoti Z, et al. 2010. Ebola hemorrhagic fever
967 associated with novel virus strain, Uganda, 2007-2008. *Emerg Infect Dis* **16**: 1087-1092.
968 doi:10.3201/eid1607.091525.
969 Wan ZM, Li ZL. 1997. A physics-based algorithm for retrieving land-surface emissivity and
970 temperature from EOS/MODIS data. *IEEE Trans Geosci Remote Sens* **35**: 980-996.
971 Wang H, Liddell CA, Coates MM, Mooney MD, Levitz CE, et al. 2014. Global, regional, and national
972 levels of neonatal, infant, and under-5 mortality during 1990-2013: a systematic analysis for
973 the Global Burden of Disease Study 2013. *Lancet*: doi:10.1016/S0140-6736(14)60497-9.
974 Weiss DJ, Atkinson PM, Bhatt S, Mappin B, Hay SI, et al. 2014. An effective approach for gap-filling
975 continental scale remotely sensed time-series. *Remote Sens Environ*: accepted manuscript.

976 Weiss DJ, Bhatt S, Mappin B, Van Boeckel TP, Smith DL, et al. 2014. Air temperature suitability for
977 *Plasmodium falciparum* malaria transmission in Africa 2000-2012: a high-resolution
978 spatiotemporal prediction. *Malar J* **13**: 171. doi:10.1186/1475-2875-13-171.
979 Wenger SJ, Olden JD. 2012. Assessing transferability of ecological models: an underappreciated
980 aspect of statistical validation. *Methods Ecol Evol* **3**: 260-267. doi:10.1111/j.2041-
981 210X.2011.00170.x.
982 Wesolowski A, Buckee CO, Pindolia DK, Eagle N, Smith DL, et al. 2013. The use of census
983 migration data to approximate human movement patterns across temporal scales. *PLoS One* **8**:
984 e52971. doi:10.1371/journal.pone.0052971.
985 Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, et al. 2012. Quantifying the impact of
986 human mobility on malaria. *Science* **338**: 267-270. doi:10.1126/science.1223467.
987 WHO. 2001. Outbreak of Ebola haemorrhagic fever, Uganda, August 2000-January 2001. *Wkly*
988 *Epidemiol Rec* **76**: 41-46.
989 WHO. 2003. Outbreak(s) of Ebola haemorrhagic fever in the Republic of the Congo, January-April
990 2003. *Wkly Epidemiol Rec* **78**: 285-289.
991 WHO. 2003. Outbreak(s) of Ebola haemorrhagic fever, Congo and Gabon, October 2001-July 2002.
992 *Wkly Epidemiol Rec* **78**: 223-228.
993 WHO. 2005. Outbreak of Ebola haemorrhagic fever in Yambio, south Sudan, April - June 2004. *Wkly*
994 *Epidemiol Rec* **80**: 370-375.
995 WHO. Ebola in Uganda - 17 November 2012. Available:
996 http://www.who.int/csr/don/2012_11_17/en/. Accessed: August 2014
997 WHO. Ebola in Uganda - 29 July 2012. Available: http://www.who.int/csr/don/2012_07_29/en/.
998 Accessed: August 2014
999 WHO. Ebola outbreak in Democratic Republic of Congo - 17 August 2012. Available:
1000 http://www.who.int/csr/don/2012_08_18/en/. Accessed: August 2014
1001 WHO. Ebola virus disease update - west Africa. Available:
1002 http://www.who.int/csr/don/2014_08_15_ebola/en/. Accessed: August 2014
1003 WHO. Health statistics and information systems - definition of region groupings. Available:
1004 http://www.who.int/healthinfo/global_burden_disease/definition_regions/en/. Accessed:
1005 August 2014
1006 WHO. 2014. Interim manual - Ebola and Marburg virus disease epidemics: preparedness, alert,
1007 control, and evaluation. Geneva: World Health Organization. 123 p.
1008 WHO. WHO statement on the meeting of the International Health Regulations Emergency Committee
1009 regarding the 2014 Ebola outbreak in West Africa. Available:
1010 http://www.who.int/mediacentre/news/statements/2014/ebola-20140808/en/. Accessed:
1011 August 2014
1012 WHO International Study Team. 1978. Ebola haemorrhagic fever in Sudan, 1976. *Bull World Health*
1013 *Organ* **56**: 247-270.
1014 Wittmann TJ, Biek R, Hassanin A, Rouquet P, Reed P, et al. 2007. Isolates of *Zaire ebolavirus* from
1015 wild apes reveal genetic lineage and recombinants. *Proc Natl Acad Sci U S A* **104**: 17123-
1016 17127. doi:10.1073/pnas.0704076104.
1017 Wolfe ND, Daszak P, Kilpatrick AM, Burke DS. 2005. Bushmeat Hunting, deforestation, and
1018 prediction of zoonotic disease emergence. *Emerg Infect Dis* **11**: 1822-1827.
1019 doi:10.3201/eid1112.040789.
1020 Wolfe ND, Dunavan CP, Diamond J. 2007. Origins of major human infectious diseases. *Nature* **447**:
1021 279-283. doi:10.1038/Nature05775.
1022 World Bank. World Bank open dataset. Available: http://data.worldbank.org/. Accessed: July 2014
1023 WorldPop. WorldPop project. Available: http://www.worldpop.org.uk/. Accessed: August 2014
1024 WWF. List of terrestrial ecoregions. Available:
1025 http://wwf.panda.org/about_our_earth/ecoregions/ecoregion_list/. Accessed: August 2014
1026 Yang Y, Atkinson P, Ettema D. 2008. Individual space-time activity-based modelling of infectious
1027 disease transmission within a city. *J R Soc Interface* **5**: 759-772. doi:10.1098/rsif.2007.1218.
1028 Zipf GK. 1946. The $P_1 P_2$/D hypothesis: on the intercity movement of persons. *Am Sociol Rev* **11**:
1029 677-686.

1030

1031

1032 **Table 1: Locations of outbreaks of Ebola virus disease in humans.** DRC = Democratic Republic of the Congo, ROC = Republic of Congo

Outbreak	Countries	Date range	Location	Species	Reference
1	South Sudan	Jun – Nov 1976	Nzara	SUDV	*(WHO International Study Team, 1978)*
2	DRC	Sep – Oct 1976	Yambuku	EBOV	*(International Commission, 1978)*
3	DRC	Jun 1977	Bonduni	EBOV	*(Heymann et al., 1980)*
4	South Sudan	Jul – Oct 1979	Nzara	SUDV	*(Baron et al., 1983)*
5	Côte d'Ivoire	Nov 1994	Tai Forest	TAFV	*(Formenty et al., 1999; Le Guenno et al., 1995)*
6	Gabon	Nov 1994 – Feb 1995	Mekouka and Andock mining camps	EBOV	*(Amblard et al., 1997; Georges et al., 1999; Milleliri et al., 2004)*
7	DRC	Jan – Jul 1995	Mwembe Forest	EBOV	*(Khan et al., 1999; Muyembe and Kipasa, 1995)*
8	Gabon	Jan – Mar 1996	Mayibout 2	EBOV	*(Georges et al., 1999; Milleliri et al., 2004)*
9	Gabon	Jul 1996 – Jan 1997	Booue	EBOV	*(Georges et al., 1999; Milleliri et al., 2004)*
10	Uganda	Oct 2000 – Feb 2001	Rwot-Obillo	SUDV	*(Lamunu et al., 2004; Okware et al., 2002; WHO, 2001)*
11	Gabon & ROC	Oct 2001 – Mar 2002	Memdemba Entsiami, Abolo and Ambomi Ekata Oloba Etakangaye Grand Etoumbi	EBOV	*(Milleliri et al., 2004; Nkoghe et al., 2005; Pourrut et al., 2005; WHO, 2003)*
12	ROC	Dec 2002 – Apr 2003	Yembelangoye Nearby hunting camp Mvoula	EBOV	*(Pourrut et al., 2005; WHO, 2003)*

13	ROC	Oct – Dec 2003	Mbandza	EBOV	*(Boumandouki et al., 2005)*
14	South Sudan	Apr – Jun 2004	Forests bordering Yambio	SUDV	*(Onyango et al., 2007; WHO, 2005)*
15	ROC	Apr – May 2005	Odzala National Park	EBOV	*(Nkoghe et al., 2011)*
16	DRC	May – Nov 2007	Mombo Mounene 2 market	EBOV	*(Leroy et al., 2009)*
17	Uganda	Aug – Dec 2007	Kabango	BDBV	*(MacNeil et al., 2010; Towner et al., 2008; Wamala et al., 2010)*
18	DRC	Nov 2008 – Feb 2009	Luebo	EBOV	*(Grard et al., 2011)*
19	Uganda	May 2011	Nakisamata	SUDV	*(Shoemaker et al., 2012)*
20	DRC	July – Nov 2012	Isiro	BDBV	*(CDC, 2014; WHO, 2012)*
21	Uganda	July - Oct 2012	Nyanswiga	SUDV	*(CDC, 2014; WHO, 2012)*
22	Uganda	Nov 2012 – Jan 2013	Luwero District	SUDV	*(CDC, 2014; WHO, 2012)*
23	Guinea	Dec 2013 -	Meliandou	EBOV	*(Baize et al., 2014; Bausch and Schwarz, 2014)*

1033

1034

1035 **Table 2: Locations of reported infections with Ebola virus in animals.** ROC = Republic of Congo

Site	Country	Date range	Location	Species	Diagnosis	Reference
1	Côte d'Ivoire	Oct – Nov 1994	Tai Forest	Chimpanzee	Serology	*(Formenty et al., 1999)*
2	Gabon	Jan 1996	Mayiboth 2	Chimpanzee	PCR	*(Lahm et al., 2007)*
3	Gabon	Jul 1996	Near Booue	Chimpanzee	Serology	*(Georges-Courbot et al., 1997)*
4	Gabon	Sept 1996	Lope National Park	Chimpanzee	PCR	*(Lahm et al., 2007)*
5	Gabon & ROC	Aug 2001	Mendemba / Lossi Animal Sanctuary	Chimpanzee	PCR	*(Lahm et al., 2007)*
6	Gabon & ROC	Aug 2001	Mendemba / Lossi Animal Sanctuary	Gorilla	PCR	*(Lahm et al., 2007)*
7	Gabon & ROC	Aug 2001	Mendemba / Lossi Animal Sanctuary	*Cephalophus dorsalis*	PCR	*(Lahm et al., 2007)*
8	Gabon	Nov 2001	Zadie	Gorilla	PCR	*(Rouquet et al., 2005)*
9	Gabon	Nov 2001	Ekata	Gorilla	PCR	*(Wittmann et al., 2007)*
10	Gabon	Dec 2001	Medemba and neighbouring villages	Chimpanzee and Gorilla	PCR	*(Leroy et al., 2002)*
11	Gabon	Feb 2002	Zadie	Gorilla	PCR	*(Rouquet et al., 2005)*
12	Gabon	Feb 2002	Ekata	Various bat species	Serology	*(Leroy et al., 2005)*
13	Gabon	Mar 2002	Zadie	Gorilla	PCR	*(Rouquet et al., 2005)*
14	Gabon	Mar 2002	Grand Etoumbi	Gorilla	PCR	*(Wittmann et al., 2007)*
15	Gabon	Apr 2002	Ekata	Gorilla	PCR	*(Wittmann et al., 2007)*
16	ROC	May 2002	Oloba	Chimpanzee	PCR	*(Lahm et al., 2007)*
17	ROC	Dec 2002	Lossi Animal Sanctuary	Gorilla	PCR	*(Rouquet et al., 2005)*

18	ROC	Dec 2002	Lossi Animal Sanctuary	Gorilla	PCR	*(Rouquet et al., 2005)*
19	ROC	Dec 2002	Lossi Animal Sanctuary	Chimpanzee	Serology	*(Rouquet et al., 2005)*
20	ROC	Dec 2002	Lossi Animal Sanctuary	Gorilla	PCR	*(Rouquet et al., 2005)*
21	ROC	Dec 2002	Lossi Animal Sanctuary	Gorilla	PCR	*(Rouquet et al., 2005)*
22	ROC	Dec 2002	Lossi Animal Sanctuary	*Cephalophus* spp.	PCR	*(Rouquet et al., 2005)*
23	Gabon	Feb 2003	Mbomo	Various bat species	PCR	*(Leroy et al., 2005)*
24	ROC	Feb 2003	Lossi Animal Sanctuary	Gorilla	Serology	*(Rouquet et al., 2005)*
25	Gabon	Feb 2003	Lossi Animal Sanctuary	Chimpanzee	PCR	*(Wittmann et al., 2007)*
26	Gabon	Jun 2003	Mbomo	Various bat species	PCR and serology	*(Leroy et al., 2005)*
27	ROC	Jun 2003	Near Mbomo and Ozala National Park	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*
28	ROC	Jun 2003	Near Mbomo and Ozala National Park	*Hypsignathus monstrosus*	Serology	*(Pourrut et al., 2009)*
29	ROC	Jun 2003	Near Mbomo and Ozala National Park	*Myonycteris torquata*	Serology	*(Pourrut et al., 2009)*
30	ROC	Jun 2003	Mbanza	Gorilla	PCR	*(Rouquet et al., 2005)*
31	ROC	Jan – Jun 2004	Lokoué	Gorilla	Reported	*(Caillaud et al., 2006)*
32	ROC	May 2004	Lokoué	Gorilla	PCR	*(Wittmann et al., 2007)*
33	Gabon	Feb 2005	Near Franceville	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*
34	Gabon	Feb 2005	Near Franceville	*Myonycteris torquata*	Serology	*(Pourrut et al., 2009)*
35	Gabon	Apr 2005	Near Lambarene	*Epomops franqueti* and *Hypsignathus monstrosus*	Serology	*(Pourrut et al., 2007)*
36	ROC	May 2005	Near Mbomo and Ozala National Park	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*
37	ROC	May 2005	Near Mbomo and Ozala	*Hypsignathus monstrosus*	Serology	*(Pourrut et al., 2009)*

			National Park			
38	ROC	May 2005	Near Mbomo and Ozala National Park	*Myonycteris torquata*	Serology	*(Pourrut et al., 2009)*
39	ROC	Jun 2005	Odzala National Park	Gorilla	PCR	*(Wittmann et al., 2007)*
40	Gabon	Feb 2006	Near Tchibanga	Various bat species	Serology	*(Pourrut et al., 2009)*
41	ROC	May 2006	Near Mbomo and Ozala National Park	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*
42	ROC	May 2006	Near Mbomo and Ozala National Park	*Hypsignathus monstrosus*	Serology	*(Pourrut et al., 2009)*
43	ROC	May 2006	Near Mbomo and Ozala National Park	*Myonycteris torquata*	Serology	*(Pourrut et al., 2009)*
44	Gabon	Oct 2006	Near Franceville	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*
45	Ghana	May 2007	Sagyimase	*Epomops franqueti*	Serology	*(Hayman et al., 2012)*
46	Ghana	May 2007	Sagyimase	*Hypsignathus monstrosus*	Serology	*(Hayman et al., 2012)*
47	Ghana	May 2007	Adoagyir	*Epomophorus gambianus*	Serology	*(Hayman et al., 2012)*
48	Ghana	May 2007	Adoagyir	*Epomops franqueti*	Serology	*(Hayman et al., 2012)*
49	Ghana	Jun 2007	Oyibi	*Epomophorus gambianus*	Serology	*(Hayman et al., 2012)*
50	Ghana	Jan 2008	Accra	*Eidolon helvum*	Serology	*(Hayman et al., 2010)*
51	Gabon	Mar 2008	Near Franceville	*Epomops franqueti*	Serology	*(Pourrut et al., 2009)*

1036

Figure 1 – The epidemiology of Ebola virus transmission in Africa. Of the suspected reservoir species, 1, 2 and 3 represent the three bat species from which Ebola virus has been isolated (*Hypsignathus monstrosus*, *Myonycteris torquata* and *Epomops franqueti*) and *n* represents unknown reservoirs of the disease yet to be discovered. Of the susceptible species, A represents *Pan troglodytes*, B *Gorilla gorilla* and *m* represents other organisms susceptible to the disease, such as duikers. H represents humans. Blue arrows indicate unknown transmission cycles or infection routes and red arrow routes have been confirmed or are suspected. Adapted from *Groseth et al. (2007)*.

Figure 2 –The locations of Ebola virus disease outbreaks in humans in Africa. Panel A illustrates the 23 reported outbreaks of Ebola virus disease through time, with the area of each circle and its position along the y-axis representing the number of cases. The onset year is represented by the colour as per Panel B. Panel B shows a map of the index cases for each of these outbreaks. Panels C through H show these outbreaks over a series of time periods. Numbers refer to outbreaks as listed in Table 1. In panels B-H the species of *Ebolavirus* responsible for the outbreak is illustrated by the symbol shape, the number of resulting cases and onset date by symbol colour. The most recent outbreak (#23) is indicated in orange. Countries in which zoonotic transmission to humans has been reported or is assumed to have occurred are coloured in blue. In each map the Democratic Republic of Congo is outlined for reference.

Figure 3 –The locations of reported Ebola virus infection in animals in Africa. Panel A shows the locations of reported Ebola virus infection in animals. Panels B through D show these records in animals over three different time periods. Numbers refer to records as listed in Table 2. In all panels, the species in which infection was detected is given by symbol shape and the year recorded by symbol colour. Blue countries represent locations where zoonotic transmission to humans has been reported or is assumed to have occurred. In each map the Democratic Republic of Congo is outlined for reference.

Figure 4 – Predicted geographical distribution of the three species of Megachiroptera suspected to reservoir Ebola virus. Panel A shows the distribution of the hammer-headed bat (*Hypsignathus monstrosus*), panel B the little collared fruit bat (*Myonycteris torquata*) and panel C Franquet's epauletted fruit bat (*Epomops franqueti*). In each map, the locations of reported observations of each species, extracted and curated from the Global Biodiversity Information Facility *(GBIF, 2014)* and used to train each model are given as grey points (*H. monstrosus*, n=67; *E. franqueti*, n=120 and *M. torquata*, n=52). Expert opinion maps of the known range of each species, generated by the IUCN *(Schipper et al., 2008)*, are outlined in grey. The colour legend represents a scale of the relative probability that the species occurs in that location from 0 (white, low) to 1 (green, high). Area under the curve statistics, calculated under a stringent ten-fold cross validation procedure, are 0.63±0.04,

0.59±0.04 and 0.58±0.03 for *H. monstrosus*, *M. torquata* and *E. franqueti* respectively. Panel D is a composite distribution map giving the mean, relative probability of occurrence from panels A-C.

Figure 5 – Predicted geographical distribution of the zoonotic niche for Ebola virus. Panel A shows the total populations living in areas of risk of zoonotic transmission for each at-risk country. The grey rectangle highlights countries in which index cases of Ebola virus disease have been reported (Set 1); the remainder are countries in which risk of zoonotic transmission is predicted, but in which index cases of Ebola have not been reported (Set 2). These countries are ranked by population at risk within each set. The population at risk figure in 100,000s is given above each bar. Panel B shows the predicted distribution of zoonotic Ebola virus. The scale reflects the relative probability that zoonotic transmission of Ebola virus could occur at these locations; areas closer to 1 (red) are more likely to harbour zoonotic transmission than those closer to 0 (blue). Countries with borders outlined are those which are predicted to contain at-risk areas for zoonotic transmission based on a thresholding approach (see Methods). The area under the curve statistic, calculated under a stringent ten-fold cross-validation procedure is 0.85±0.04. Solid lines represent Set 1 whilst dashed lines delimit Set 2. Areas covered by major lakes have been masked white.

Figure 5 – figure supplement 1 – Covariates used in predicting zoonotic transmission niche of Ebola. Panel A displays elevation across Africa measured in metres, relative to sea level. Panels B and C show enhanced vegetation index (EVI) values (mean and spatial range respectively) on a scale from 0 to 1. Panels D through G display land surface temperature (LST) (mean and spatial range for day and night respectively) measured in degrees Celsius. Panel H shows potential evapotranspiration (PET) for Africa, in millimetres per month and Panel I gives the composite, relative probability of occurrence of the three suspected reservoir bat species. For details of how each of these covariate layers was derived see Methods.

Figure 5 – figure supplement 2 – Marginal effect plots for each covariate used in the Ebola virus distribution model. Each panel illustrates the marginal effect (averaging over the effects of other covariates) that changes in each of the covariates has on the predicted relative probability of occurrence of zoonotic Ebola virus transmission. Grey regions and solid lines give the 95% confidence region (a metric of uncertainty) and mean value calculated across all 500 submodels. The mean relative contribution of the covariate to the model (the proportion of iterations in which the covariate was selected by the model-fitting algorithm, indicating sensitivity to the covariates) is given as an inset number. The dependency plots are ordered by mean relative contribution of the covariate. EVI = enhanced vegetation index, LST = land surface temperature and PET = potential evapotranspiration.

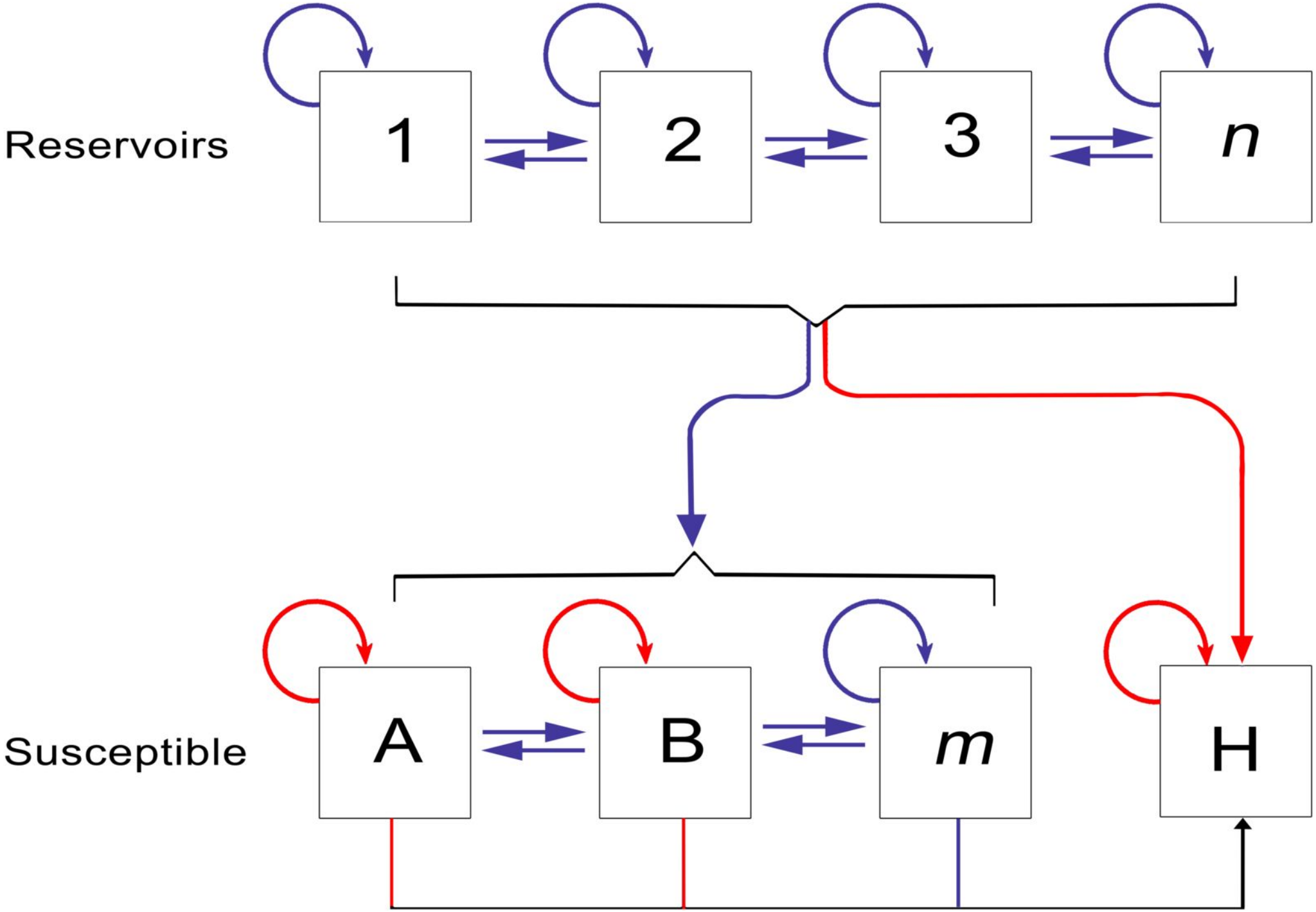
Figure 5 – figure supplement 3 – Comparison of predictions for zoonotic niche of Ebola virus excluding the Guinea outbreak. Panel A shows the predicted zoonotic niche excluding the index

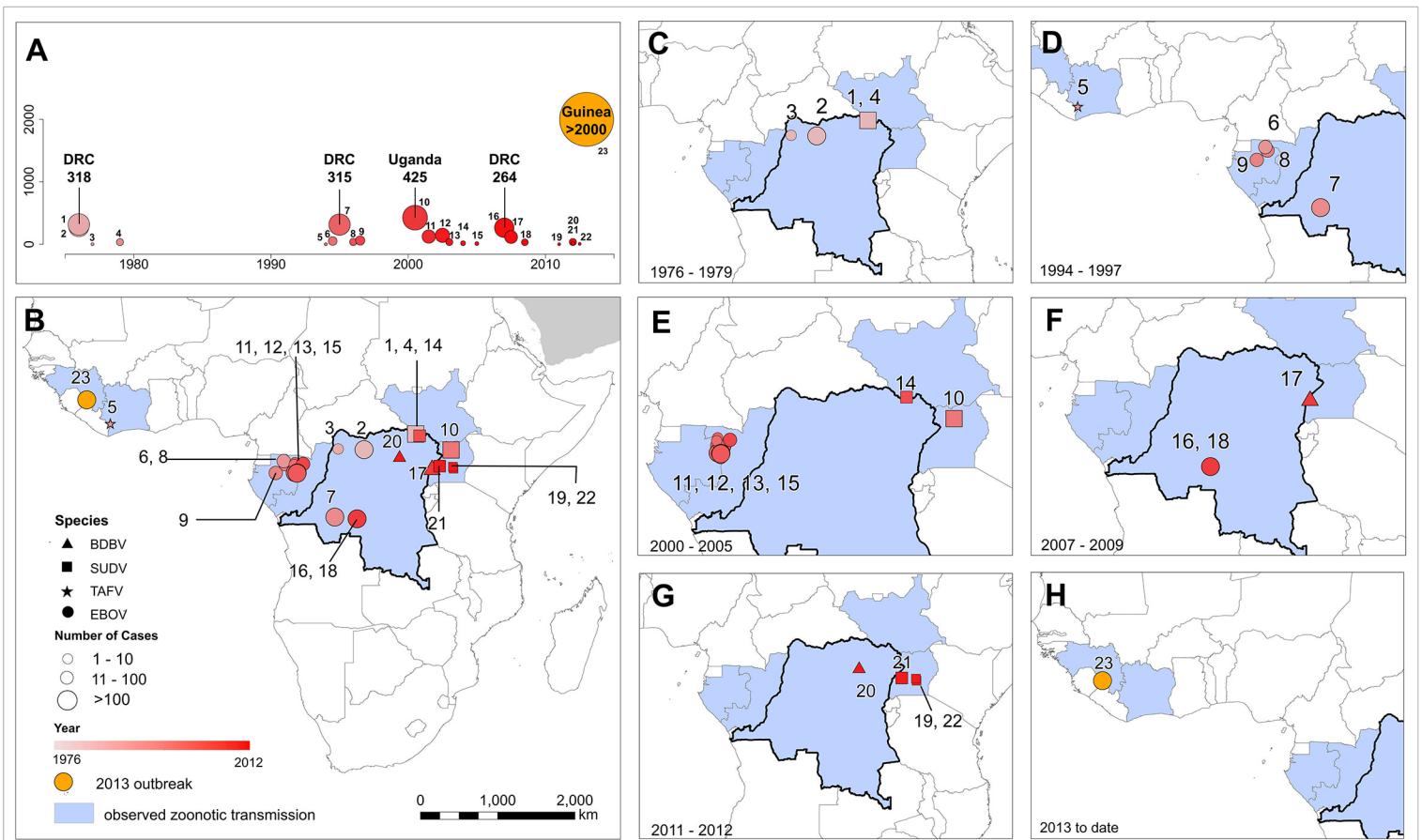
case for the Guinea outbreak from the dataset used to train the model. Panel B shows the prediction when including the Guinea data in the model (the model presented in Figure 5). The circle depicts the location of the Guinean index case (#23 in Table 1). As per Figure 5, the scale reflects the relative probability that zoonotic transmission of Ebola virus could occur at these locations; areas closer to 1 (red) are more likely to harbour zoonotic transmission than those closer to 0 (blue).

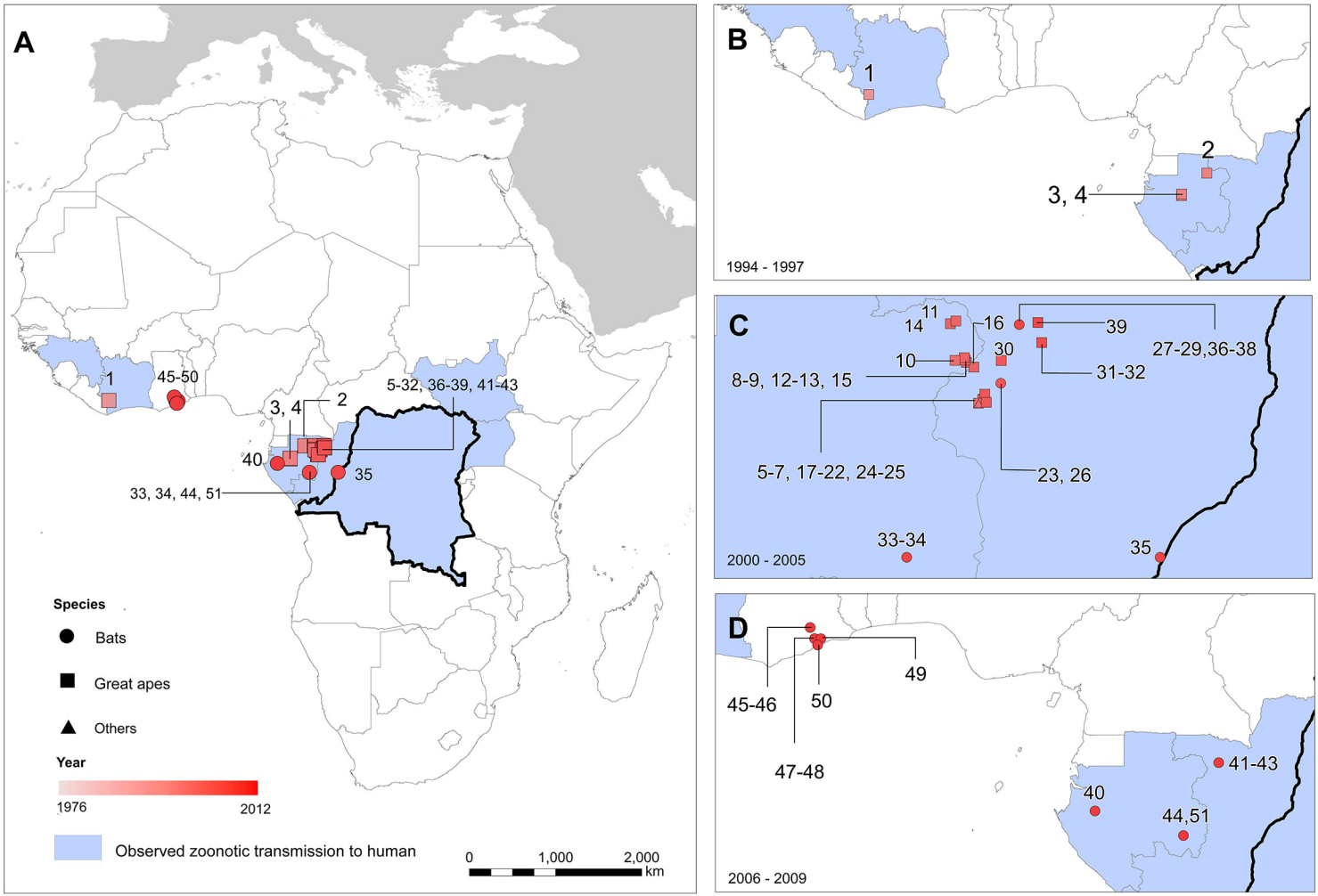
Figure 6 – Changes in national population for countries predicted to contain areas at-risk of zoonotic Ebola virus transmission. For each country the population (in millions) is presented for three time periods (1976, 2000 and 2014) as three bars. Each stacked bar gives the rural (green) and urban (blue) populations of the country. The grey rectangle highlights countries in which index cases of Ebola virus diseases have been reported (Set 1); the remainder are countries in which risk of zoonotic transmission is predicted, but where index cases have not been reported (Set 2). The fractional change in population between 1976 and 2014 is given above each set of bars.

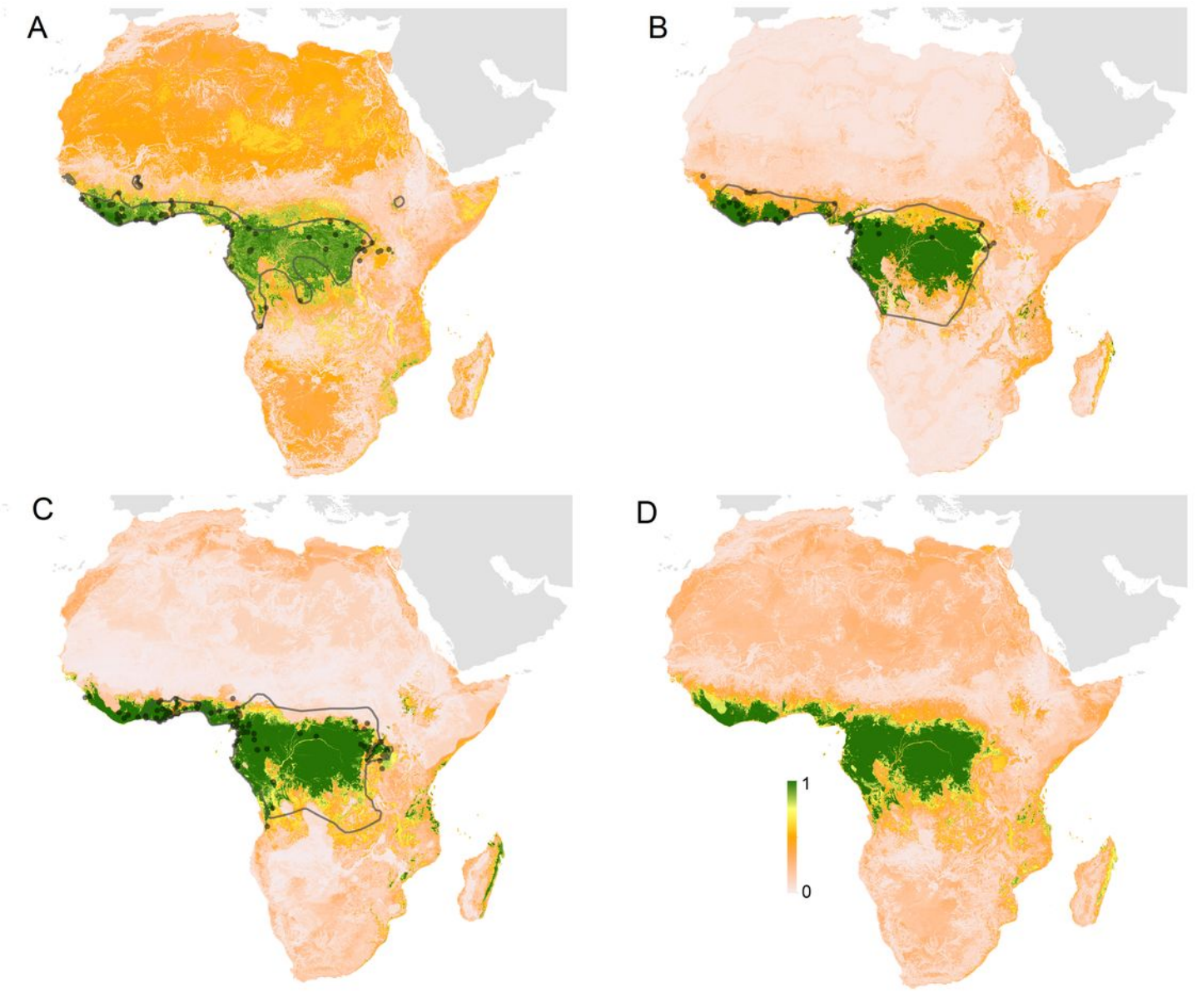
Figure 7 – Changes in international flight capacity and traveller volumes for countries predicted to contain areas at-risk of zoonotic Ebola virus transmission. The grey rectangle highlights countries in which index cases of EVD have been reported (Set 1). The remainder are countries in which risk of zoonotic transmission is predicted, but where index cases have not been reported (Set 2). Panel A shows changes in annual outbound international seat capacity (between 2000 in red and 2013 in blue). Panel B depicts changes in annual outbound international passenger volume by country (between 2005 in red and 2012 in blue). For each country, the fractional change in volume is given above each set of bars. Note that only one bar is presented for South Sudan as data for this region prior to formation of the country in 2011 were unavailable

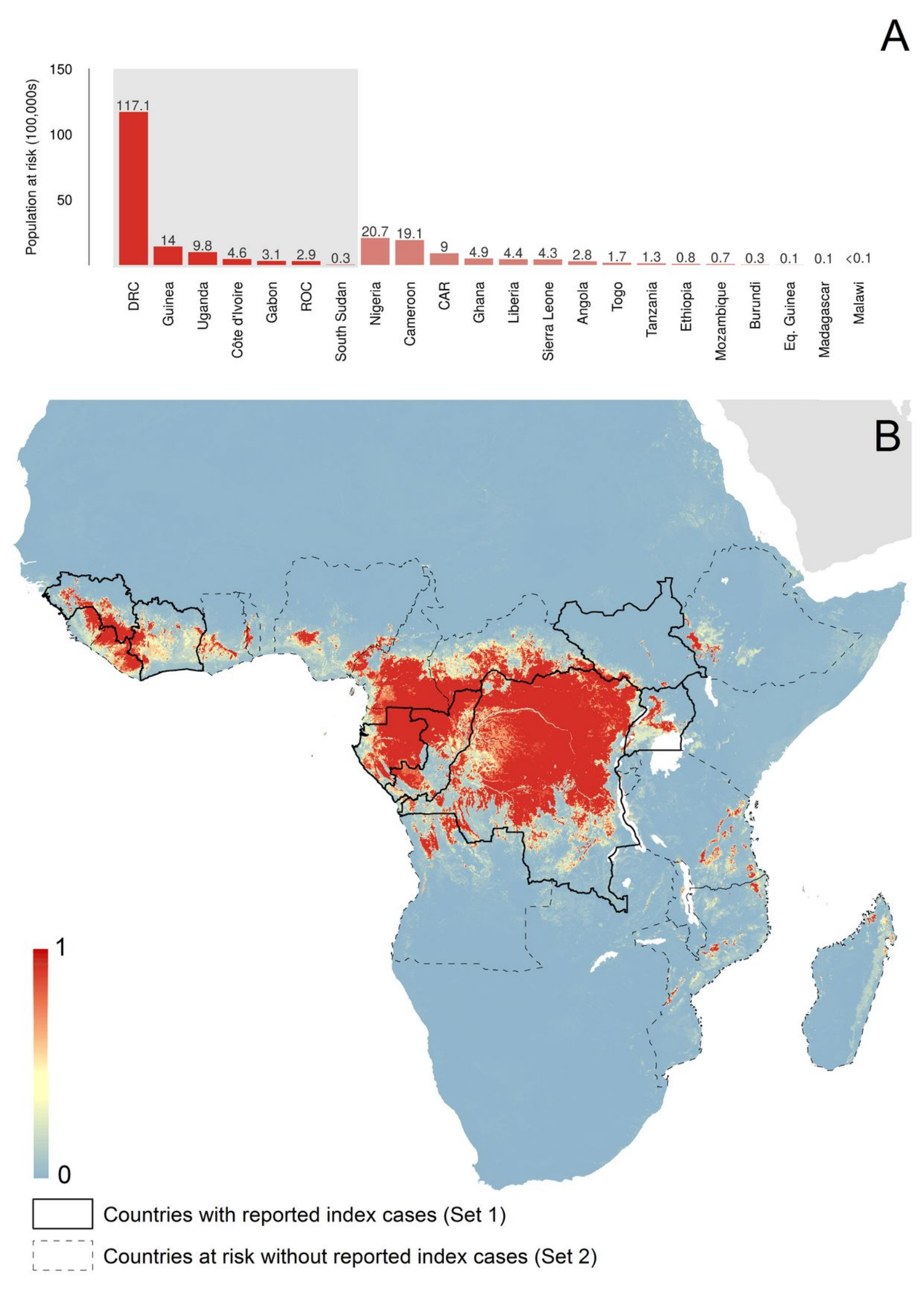
Figure 8 – Numbers of airline passengers arriving from at-risk countries to other countries stratified by major geographic regions and national income groups. Panel A shows the locations of WHO regions (AFRO - African Region; AMRO – Region of the Americas; EMRO - Eastern Mediterranean Region; EURO - European Region; SEARO – South-East Asian Region; WPRO - Western Pacific Region). Panel B displays the numbers of passengers arriving in each of these regions from countries predicted to contain areas at risk of zoonotic Ebola virus transmission (Sets 1 and 2) in 2005 and 2012. Panel C shows the income tiers of all countries as defined by the World Bank. Panel D displays the total numbers of passengers arriving in countries in each of these income strata from at-risk countries in 2005 and 2012. The number above each pair of bars indicates the fractional change in these numbers of incoming passengers between 2005 and 2012.

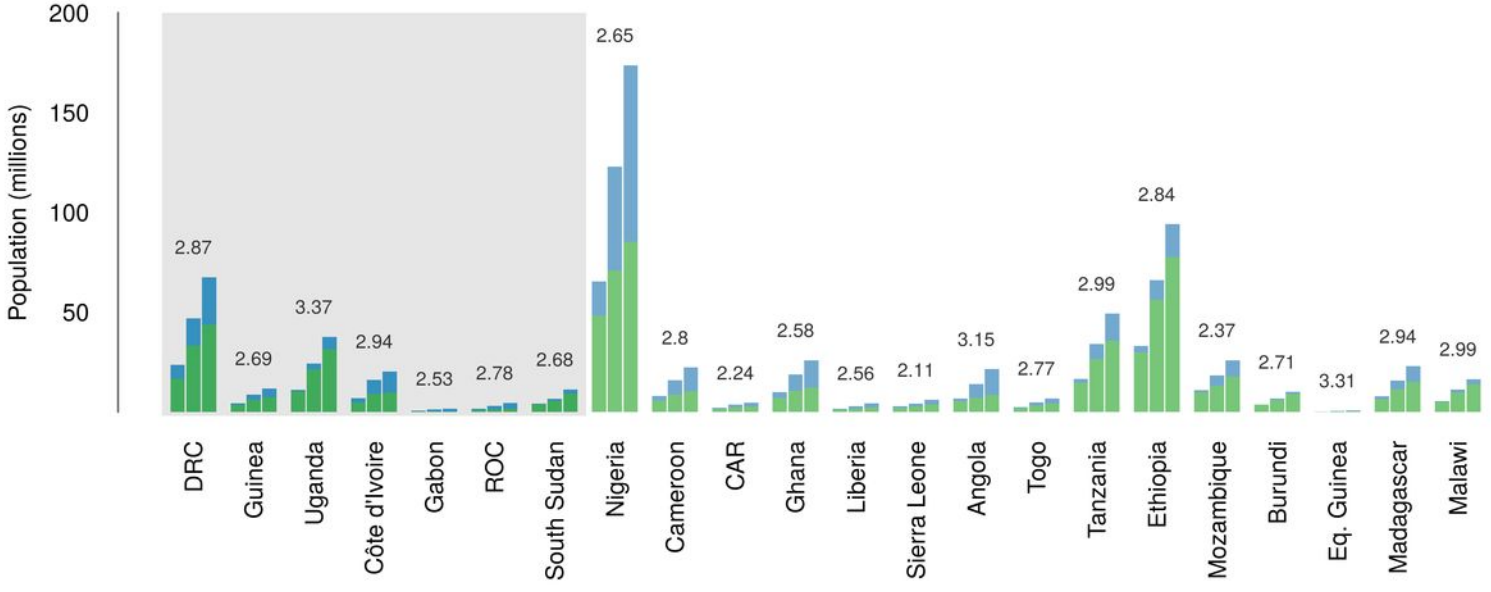

Reservoirs
1
2
3
n
Susceptible
A
B
m
H

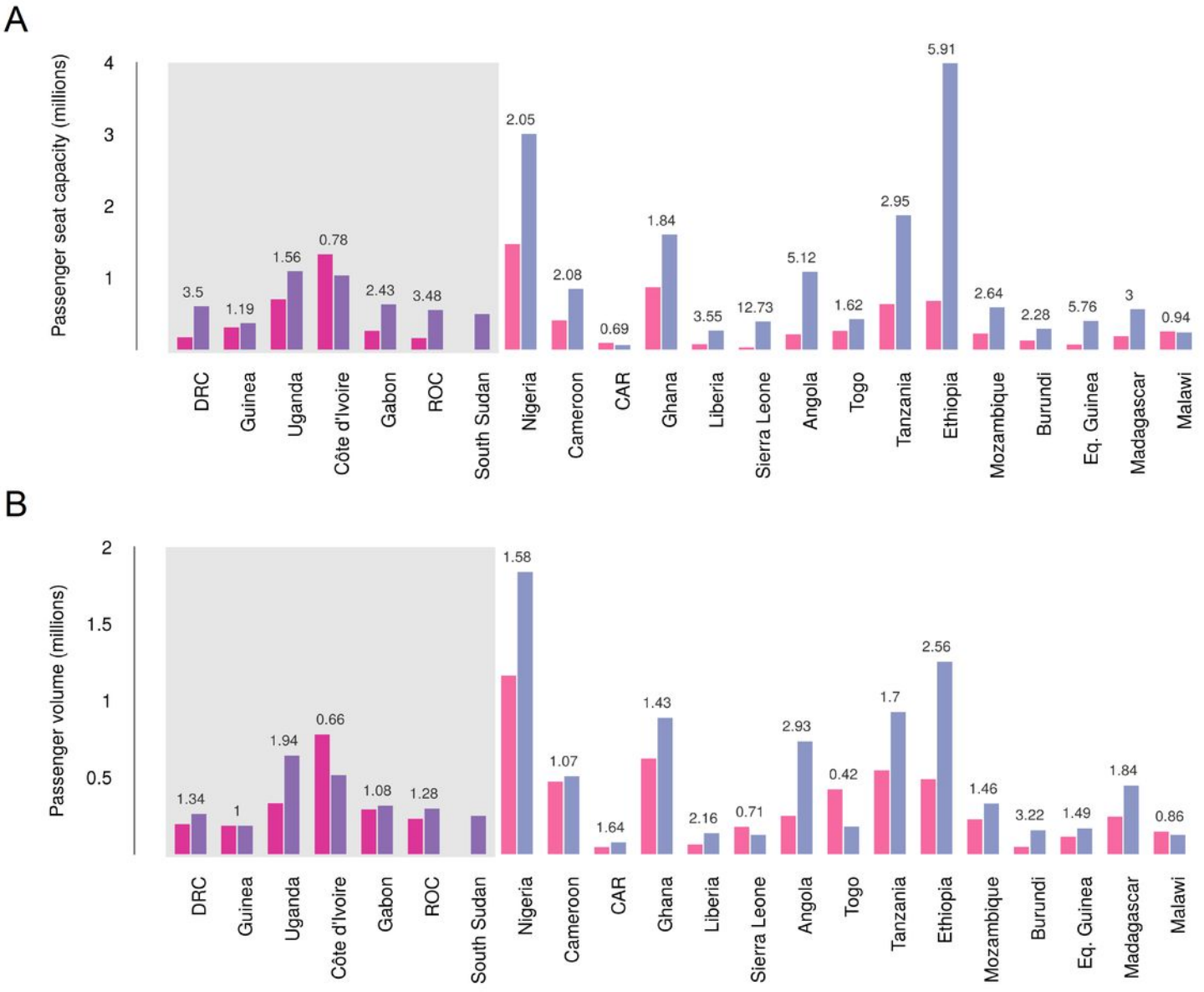

A
2000
1000
0
DRC
318
DRC
315
Uganda
425
DRC
264
Guinea
>2000
23
1980
1990
2000
2010
B
11, 12, 13, 15
1, 4, 14
23
5
3
2
20
10
6, 8
17
19, 22
7
9
21
16, 18
Species
BDBV
SUDV
TAFV
EBOV
Number of Cases
1 - 10
11 - 100
>100
Year
1976
2012
2013 outbreak
observed zoonotic transmission
0
1,000
2,000
km
C
3
2
1, 4
1976 - 1979
D
5
6
9
8
7
1994 - 1997
E
14
10
11, 12, 13, 15
2000 - 2005
F
17
16, 18
2007 - 2009
G
21
20
19, 22
2011 - 2012
H
23
2013 to date

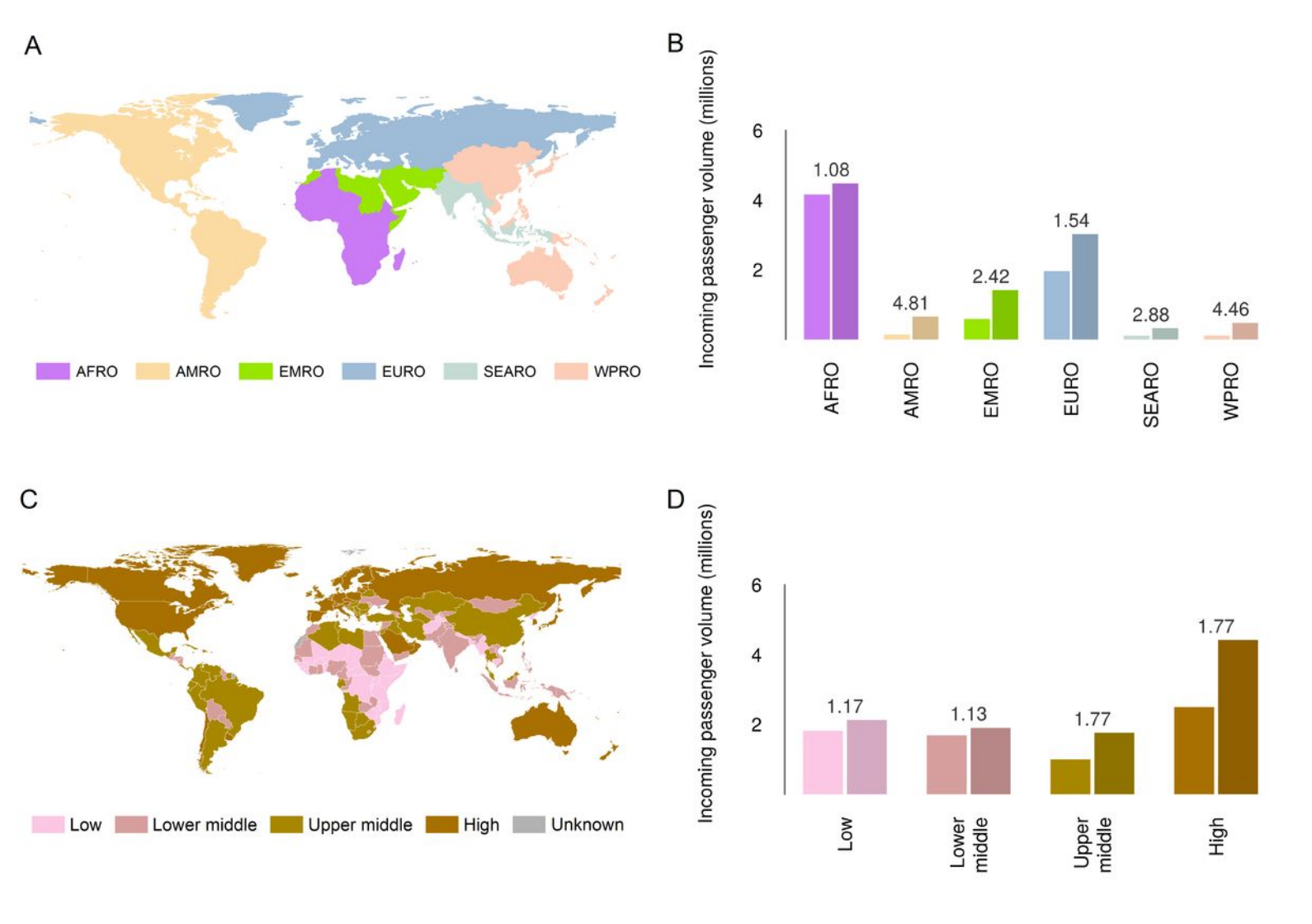

A
1
45-50
3, 4
2
5-32, 36-39, 41-43
40
35
33, 34, 44, 51
Species
Bats
Great apes
Others
Year
1976
2012
Observed zoonotic transmission to human
0
1,000
2,000
km
B
1
2
3, 4
1994 - 1997
C
11
14
16
39
27-29,36-38
10
30
31-32
8-9, 12-13, 15
5-7, 17-22, 24-25
23, 26
33-34
35
2000 - 2005
D
49
45-46
50
47-48
41-43
40
44,51
2006 - 2009


A
B
C
D
1
0


A
Population at risk (100,000s)
150
100
50
117.1
14
9.8
4.6
3.1
2.9
0.3
20.7
19.1
9
4.9
4.4
4.3
2.8
1.7
1.3
0.8
0.7
0.3
0.1
0.1
<0.1
DRC
Guinea
Uganda
Côte d'Ivoire
Gabon
ROC
South Sudan
Nigeria
Cameroon
CAR
Ghana
Liberia
Sierra Leone
Angola
Togo
Tanzania
Ethiopia
Mozambique
Burundi
Eq. Guinea
Madagascar
Malawi
B
1
0
Countries with reported index cases (Set 1)
Countries at risk without reported index cases (Set 2)


Population (millions)
200
150
100
50
2.87
2.69
3.37
2.94
2.53
2.78
2.68
2.65
2.8
2.24
2.58
2.56
2.11
3.15
2.77
2.99
2.84
2.37
2.71
3.31
2.94
2.99
DRC
Guinea
Uganda
Côte d'Ivoire
Gabon
ROC
South Sudan
Nigeria
Cameroon
CAR
Ghana
Liberia
Sierra Leone
Angola
Togo
Tanzania
Ethiopia
Mozambique
Burundi
Eq. Guinea
Madagascar
Malawi


A
Passenger seat capacity (millions)
4
3
2
1
3.5
1.19
1.56
0.78
2.43
3.48
2.05
2.08
0.69
1.84
3.55
12.73
5.12
1.62
2.95
5.91
2.64
2.28
5.76
3
0.94
DRC
Guinea
Uganda
Côte d'Ivoire
Gabon
ROC
South Sudan
Nigeria
Cameroon
CAR
Ghana
Liberia
Sierra Leone
Angola
Togo
Tanzania
Ethiopia
Mozambique
Burundi
Eq. Guinea
Madagascar
Malawi
B
Passenger volume (millions)
2
1.5
1
0.5
1.34
1
1.94
0.66
1.08
1.28
1.58
1.07
1.64
1.43
2.16
0.71
2.93
0.42
1.7
2.56
1.46
3.22
1.49
1.84
0.86
DRC
Guinea
Uganda
Côte d'Ivoire
Gabon
ROC
South Sudan
Nigeria
Cameroon
CAR
Ghana
Liberia
Sierra Leone
Angola
Togo
Tanzania
Ethiopia
Mozambique
Burundi
Eq. Guinea
Madagascar
Malawi


A
AFRO
AMRO
EMRO
EURO
SEARO
WPRO
B
Incoming passenger volume (millions)
6
4
2
1.08
4.81
2.42
1.54
2.88
4.46
AFRO
AMRO
EMRO
EURO
SEARO
WPRO
C
Low
Lower middle
Upper middle
High
Unknown
D
Incoming passenger volume (millions)
6
4
2
1.17
1.13
1.77
1.77
Low
Lower middle
Upper middle
High