Synthesising rodent trapping studies in West Africa to understand the hazard of zoonotic infectious disease outbreaks.

David Simons

Lauren A. Attfield

Kate E. Jones

Deborah Watson-Jones

Richard Kock

2022-03-08

# Abstract

**To complete after written**

# Introduction

Zoonotic infectious disease spillover events into human populations (animal-to-human transmission) can have important health and economic consequences at both the global and local level (e.g. SARS-CoV-2 and Ebola virus disease) (reference). The number of zoonotic disease spillover events and the emergence of novel zoonotic pathogens are projected to increase under intensifying anthropenic pressure driven by increased human populations, urbanisation, climate change and wildlife defaunation [1–4]. West Africa has been identified as a nexus for these processes and a hotspot for both the risk of large epidemics of known zoonotic diseases and novel zoonotic pathogen emergence and [1,5,6]. Further, two taxa - rodents (Rodentia) and bats (Chiroptera) - contribute the greatest number of predicted novel zoonotic pathogens and known zoonoses [7,8]. Of 2,220 extant rodent species, 244 (10.7%) are described as reservoirs of 85 zoonotic pathogens [6]. Specifically, West Africa is identified as a region at increased hazard for rodent-borne zoonotic disease spillover events under different projected future land-use change scenarios [3]. In West Africa rodents are involved in the transmission of multiple zoonotic diseases, including, Lassa fever, Schistosomiasis, Leptospirosis and Toxoplasmosis [9,10].

Rodent species form diverse assemblages, which provide important and beneficial ecosystem services including pest regulation and seed dispersal [11]. The role of rodent species in zoonotic disease spillover or novel zoonotic pathogens emergence are examples of ecosystem disservices. Rodents typically demonstrate “fast” life histories [12] with traits such as early maturation and short gestation further associated with being zoonotic reservoirs [7,13]. Traits associated with “fast” life histories are prevalent in species that thrive in human dominated landscapes, displacing species less likely to be reservoirs of zoonotic pathogens [14]. The widespread occurrence of reservoir species and their proximity to human activity make the description of rodent species assemblages and host-pathogen interactions vitally important to understanding the hazard of zoonotic disease spillover and novel zoonotic pathogen emergence [6].

Despite the importance of understanding these complex systems in the hazard of zoonotic spillover, current evidence on host-pathogen associations is affected by taxonomic and geographical sampling biases [8,15]. Studies identifying potential geographic hotspots of zoonotic disease spillover and novel pathogen emergence are based on global datasets of host species distributions and host-pathogen associations (e.g., IUCN, GBIF, GIDEON) that suffer from several biases [16,17]. For example, systematically increased sampling, over-representation of certain habitats and clustering around areas of high human population could lead to an apparent association between locations and hazard that is driven by these factors rather than underlying host-pathogen associations [8,18,19]. While other regions remain systematically undersampled (e.g., areas of sparse human populations), resulting in a reduced hazard of zoonotic spillover events or novel zoonotic pathogen emergence being attributed [20]. Predictions of zoonotic disease spillover and novel zoonotic pathogen emergence must account for these biases to understand the future hazard of zoonotic diseases [17].

Rodent trapping studies provide contextually rich information on when, where and under what conditions rodents were trapped, potentially enriching global datasets and counteracting inherent biases [21]. Rodent trapping provides a useful method to describe rodent population assemblages, their geographic distribution and host-pathogen associations beyond what is currently available in global datasets. Studies have been conducted in West Africa to investigate the distribution of rodent speciesm, their species assemblages, the prevalence of known zoonoses within rodent hosts (e.g., Lassa fever, Schistosomiasis) and to identify novel zoonotic pathogens [22–24]. However, data from these studies have not been synthesised for inclusion in assessments of zoonotic disease spillover and novel zoonotic pathogen emergence.

Here, we synthesise the rodent trapping studies conducted across West African from a search of literature between 1964-2019. First, we investigate the geographic biases in the rodent trapping dataset in relation to human population density and land use classification. Second, we compare this to global host datasets (IUCN and GBIF) to understand differences in reported host geographic distributions. Third, we compare identified host-pathogen associations with global datasets (CLOVER) to understand discrepencies in rodent host-pathogen associations and report the proportion positive for each pathogen. Finally, within our dataset we investigate host-pathogen spatial extent to identify areas of undersampling of pathogens within host ranges. We expect that rodent trapping studies will provide an important additional source of high resolution data that can be used to enrich global datasets to better understand the hazard of zoonotic disease spillover and novel zoonotic pathogen emergence across West Africa. We find that rodent trapping studies identify greater geographic ranges of important rodent hosts of zoonotic diseases than are reported from global datasets while also identifying locations and habitats in which further sampling is required. We also identify host-pathogen associations that are not reported in global host-pathogen association data, however, there remains systematic undersampling for these pathogens across the host ranges.

# Methods

## Data sources

### Host and pathogen trapping data

We conducted a search in Ovid MEDLINE, Web of Science (Core collection and Zoological Record), JSTOR, BioOne, African Journals Online, Global Health and the pre-print servers, BioRxiv and EcoEvoRxiv for the following terms as keywords, no date limits were set: (1) Rodent OR Rodent trap\* AND (2) West Africa (or the individual countries).

We searched other resources including the UN Official Documents System, Open Grey, AGRIS FAO and Google Scholar using combinations of the above terms. Searches were run on 2021-03-01, returning studies conducted between 1964-2019.

We included studies if they met all of the following inclusion criteria; i) reported findings from trapping studies where the target was a small mammal, ii) described the type of trap used or the length of trapping activity or the location of the trapping activity, iii) included trapping activity from at least one West African country, iv) recorded the genus or species of trapped individuals, v) were published in a peer-reviewed journal or as a pre-print on a digital platform or as a report by a credible organisation. We excluded studies if they met any of the following exclusion criteria: i) reported data that were duplicated from a previously included study, ii) no full text available, iii) not available in English. One reviewer screened titles, abstracts and full texts against the inclusion and exclusion criteria. At each stage, a random subset (10%) was reviewed by a second reviewer.

We extracted data from eligible studies using a standardised tool that was piloted on 5 studies (Supplementary 1.). Data was abstracted into a Google Sheets document, which was archived on completion of data extraction (Supplementary 2.). Extracted variables included i) study identifiers; ii) study aims; iii) trapping methodology; iv) geolocation data; v) method of speciation; vi) trapping locations and dates; vii) trapped species; viii) number of trap-nights and ix) microorganisms/pathogens of interest. We summarised the number of studies, the year in which trapping occurred and the country in which they were conducted.

We extracted GPS locations for the most precise location presented, converting to decimal degrees as required. For studies not using standardised habitat recording (e.g., IUCN Habitat Classification Scheme (Version 3.1)), the explicit description of the habitat in which the trap was placed was extracted. For studies reporting multiple habitat types for a single trap, trap-line or trapping grid, a higher order classification of habitat type was recorded.

We mapped genus and species names to those in the Global Biodiversity Information Facility (GBIF) taxonomy [25]. We extracted information on the detection, non-detection and number of trapped individuals.

We extracted data on all pathogens assayed. The number of rodents tested and the number of positive or negative samples were recorded alongside the type of assay used. If studies reported indeterminate results this was noted. Pathogens were identified to species level, however, where an assay only allowed for attribution to a family of viruses or bacteria, the higher order grouping was used.

### Environmental covariates

### Rodent species ranges

### Rodent presence

### Host pathogen associations

CLOVER a synthesis of four host-pathogen datasets (GMPD2, EID2, HP3 and Shaw 2020) was released in 2021 and represents the most comprehensive global repository of host-pathogen associations for Bacteria, Viruses, Helminth, Protozoa and Fungi [clover\_paper]. Data were obtained from an archived version of the dataset [clover\_dataset].

## Included studies

We identified 4,282 relevant citations, with 126 rodent trapping studies included. The earliest trapping studies were conducted in 1964, with increasing numbers of studies being performed annually. The median year of first trapping activity was 2005, with the median length of trapping activity being 1 year (IQR 0-2 years) (Supplementary figure 1.). Studies were conducted in 14 West African countries with no studies reported from Gambia or Togo, at 1,331 sites (Figure 1A.). A minority (31, 25%) of studies trapped at a single study site, with 46 (37%) trapping at between two and five sites, the remaining 49 studies (38%) trapped at between six and 93 study sites. There were 581,426 reported trap nights, with 384,983 trap nights imputed for studies with no recording of trapping effort (Figure 1B.).

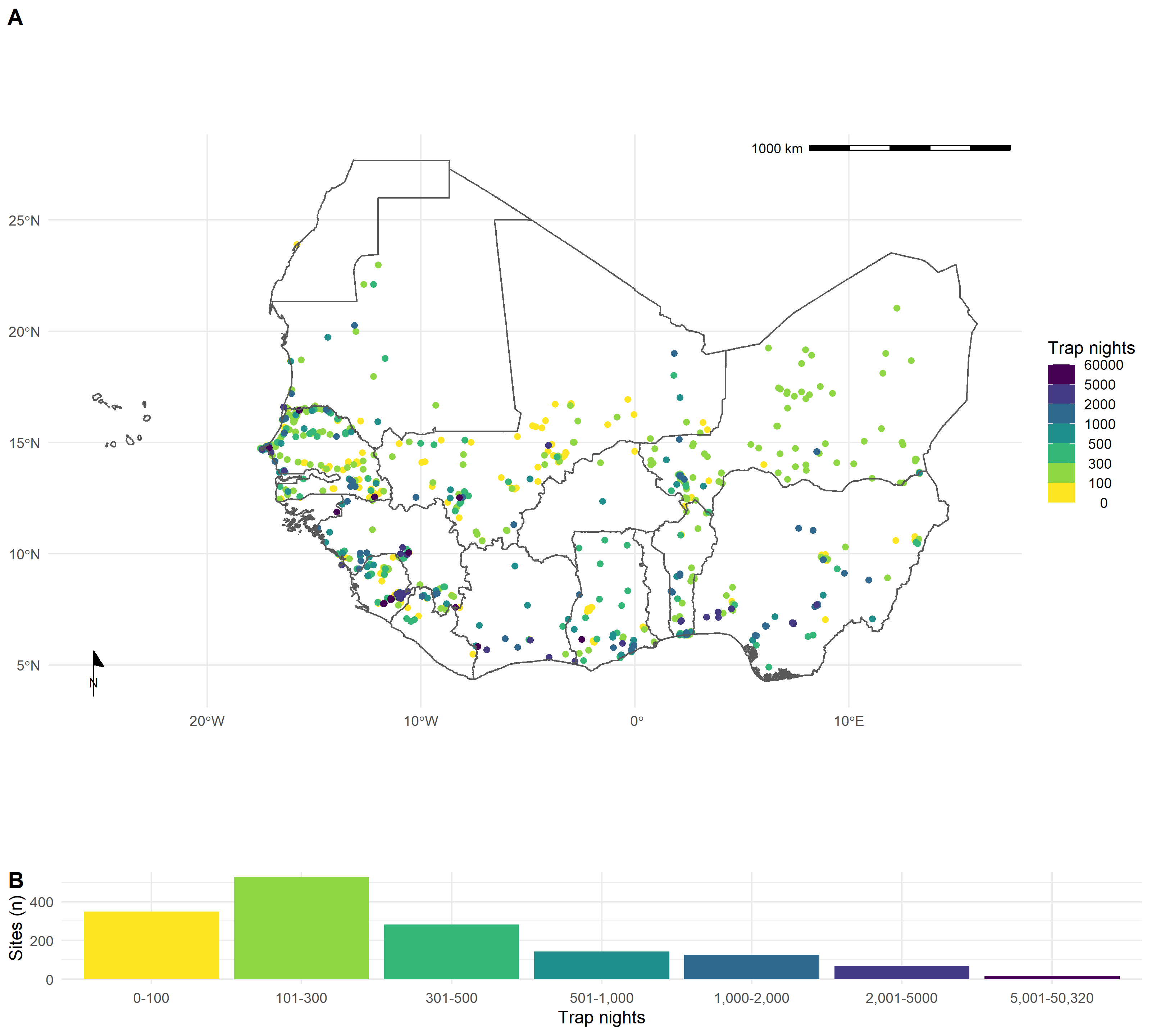


Figure 1. A) The location of trapping sites in West Africa. No sites were recorded from Togo or Gambia. There is important heterogeneity observed in the coverage of each country by trap night (colour) and location of sites. For example Senegal, Mali and Sierra Leone have generally good coverage compared to Guinea and Burkina Faso. B) The distribution of trap nights performed at each study site on a log scale with a median of 244 trap nights (IQR 115-494) performed at each site

Within the included studies trap success varied importantly between traps placed in or around buildings (13%, IQR 6-24%) compared to other habitats (3%, IQR 1-9%). In total 73,164 small mammals were trapped with 62,574 (85%) identified to species level and 7,760 (11%) identified to genus. The majority of the 147 identified species were Rodentia (112, 76%) of which Muridae (82, 73%) were the most common family.

62 studies tested for 32 microorganisms, defined to species or genus leve that are known or potential pathogens. The majority of studies tested for a single microorganism (39, 62%). The most frequently assayed microorganisms were *Lassa mammarenavirus* or Arenaviridae (27, 43%), *Borrelia sp.* (8, 13%), *Bartonella sp.* (4, 6%) and *Toxoplasma gondii* (4, 6%). The majority of studies used Polymerase Chain Reaction to detect microorganisms (32, 52%), with fewer studies using molecular based tests (11, 18%) or histological or direct visualisation assays (8, 13%).

From 32,014 individual rodent samples we produced 339 host-pathogen pairs. With *Rattus rattus*, *Mus musculus*, *Mastomys erythroleucus*, *Mastomys natalensis* and *A. niloticus* being assayed for at least 18 microorganisms.

## Analysis

### What is the extent of spatial bias in the rodent trapping data?

We used the GPS coordinates of single trap, trapping grid/line or study site and the number of trap nights to calculate trapping effort (trap night density) within level 2 administrative regions in West Africa. The median trap success rate for all rodents at a defined trap site was calculated separately for trap sites which included built environments and non-built environments due to important differences in trap success between these different environments. The number of rodents trapped at a trap site was then used in combination with these values of trap success to impute the number of trap nights for trap sites where there was no record of the number of trap nights. We summarised the habitat types of trap sites based on information reported in the study; the dictionary used to match habitat types is available in Supplementary 3.

To better understand the potential bias associated with studies for zoonotic disease spillover or novel zoonotic pathogen emergence we limited further analyses to zoonotic disease studies only. We used a Generalised Additive Model (GAM) incorporating a spatial interaction term to investigate the association between number of trap nights, proportion of urban, cropland and forest land classification and human population density [26]. Covariates were obtained for 2005, the median year studies were conducted, proportions of habitat types were obtained from Copernicus 2005 land cover classification, global population density estimates were obtained from SEDAC [27]. We performed sensitivity analysis by removing sites with imputed trapping effort. The model structure was specified as:

The model was used to identify locations of high and low trapping effort relative to population density and land use classification. Predictions were limited to regions around observed trap sites.

### What is the difference between global datasets rodent host distributions and distributions from trapping studies?

We assessed the taxonomic coverage of the trapped taxa and for the seven species (*M. natalensis*, *R. rattus*, *M. erythroleucus*, *M. musculus*, *Arvicanthis niloticus*, *Praomys daltoni* and *Cricetomys gambianus*) with the most trap locations we mapped the detection and non-detection of rodent species and compared this to detection data obtained from GBIF and species ranges from IUCN [28,29]. We identified rodent locations from GBIF and the rodent trapping studies within and outside the IUCN expected range. We located these to level 2 administrative regions to calculate the discrepencies between observed detections and non-detections with IUCN distributions.

### Are rodent trapping derived host-pathogen associations present in global zoonoses datasets?

For studies investigating rodent zoonoses we produced host-pathogen pairs. For each host-pathogen pair we compared our records with those available in CLOVER reporting concordance and discordance. We identify host-pathogen pairs that are not present in CLOVER and contain assay results consistent with acute or prior infection.

### What is the spatial extent of pathogen testing within a hosts range?

For the four pathogens sampled at the most locations (Borreliaceae, Leptospiraceae, *Lassa mammarenavirus* and *Toxoplasma gondii*) we identify their host-pathogen associations. We identified the five species with most spatially complete sampling and associated these sampling locations with level 2 administrative regions. We compared the area of these regions to their IUCN range and detection range to calculate sampling coverage.

# Results

Further descriptive information from the included studies including geolocated trapping of species, their detection and non-detection alongside microorganism data has been made available online [30].

## What is the extent of spatial bias in the rodent trapping data?

Trap sites were situated in 273 (19.3%) of 1414 level 2 administrative regions in 14 West African nations. The areas with highest trap night density included the capital and large cities of Niger (Niamey), Sierra Leone (Freetown), Senegal (Dakar), Mali (Bamako), Ghana (Accra), Côte d’Ivoire (Abidjan), and Benin (Cotonou). Outside of these cities, regions in, Northern Senegal, Southern Guinea, Edo and Osun States in Nigeria and Eastern Sierra Leone had the highest density of trap nights (Figure 1A.).

Trapping sites were biased towards human modified landscapes (i.e., cropland, grassland and mosaic habitats), with under-representation of forest and bare habitats. However, urban habitats were undersampled.

Trapping effort, measured through trap nights, was biased towards areas with higher population density and predominantly urban and cropland land classifications (Figure 2.). In particular Northern Senegal, Eastern Sierra Leone, Eastern Guinea and South West Nigeria had relatively high trapping efforts while South East Nigeria, Northern Nigeria, Liberia, Côte d’Ivoire, Ghana, Niger, and Burkina Faso were trapped at lower trapping effort than would be expected. Excluding sites with imputed trap nights reduced the number of trapping sites contributing data. Despite this limitation, Mauritania, Northern Senegal and Sierra Leone remained as regions trapped at higher than expected rates with Nigeria as being trapped at lower than expected rates (Supplementary Figure 2.).

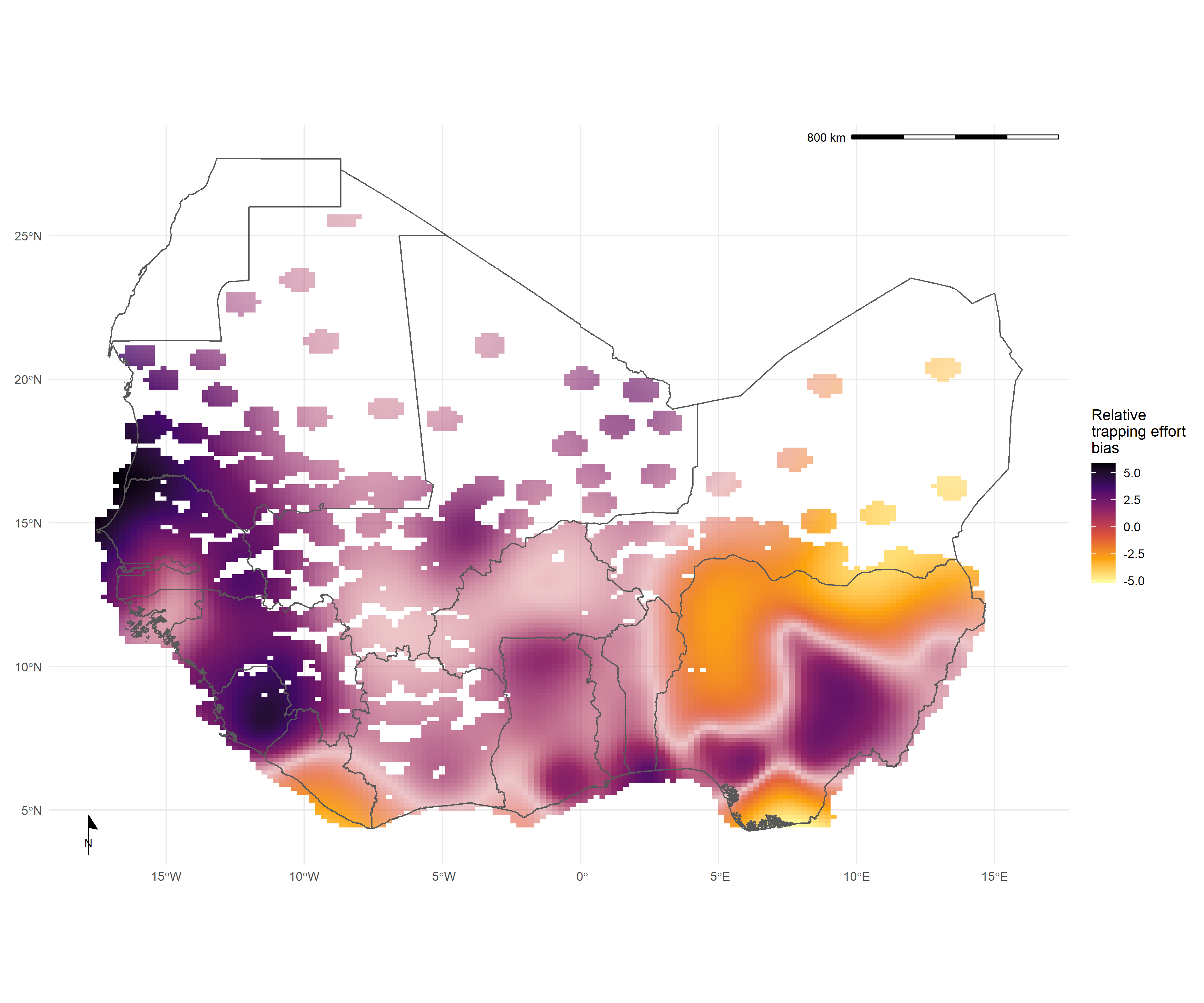


Figure 2. Relative trapping effort bias across West Africa from included studies adjusted for human population density, proportion cropland cover, proportion forest cover and proportion urban habitat. Predictions are limited to areas around trap sites (coloured areas), uncertainty in the estimate of bias is represented by colour transparency/pallor. Purple regions represent areas with higher than expected trapping effort, yellow regions represent areas lower than expected trapping effort

## What is the difference between global datasets rodent host distributions and distributions from trapping studies?

For our species of interest, except *C. gambianus*, trapping studies provided more locations of detection and non-detection than were available from GBIF. For the West African endemic species IUCN ranges showed generally good concordance to both trapping studies and GBIF, although the range of *A. niloticus* and *P. daltoni* are greater than expected. The invasive species *R. rattus* and *M. musculus* are present in a much larger range than is expected (Figure 3A. and 3B.).

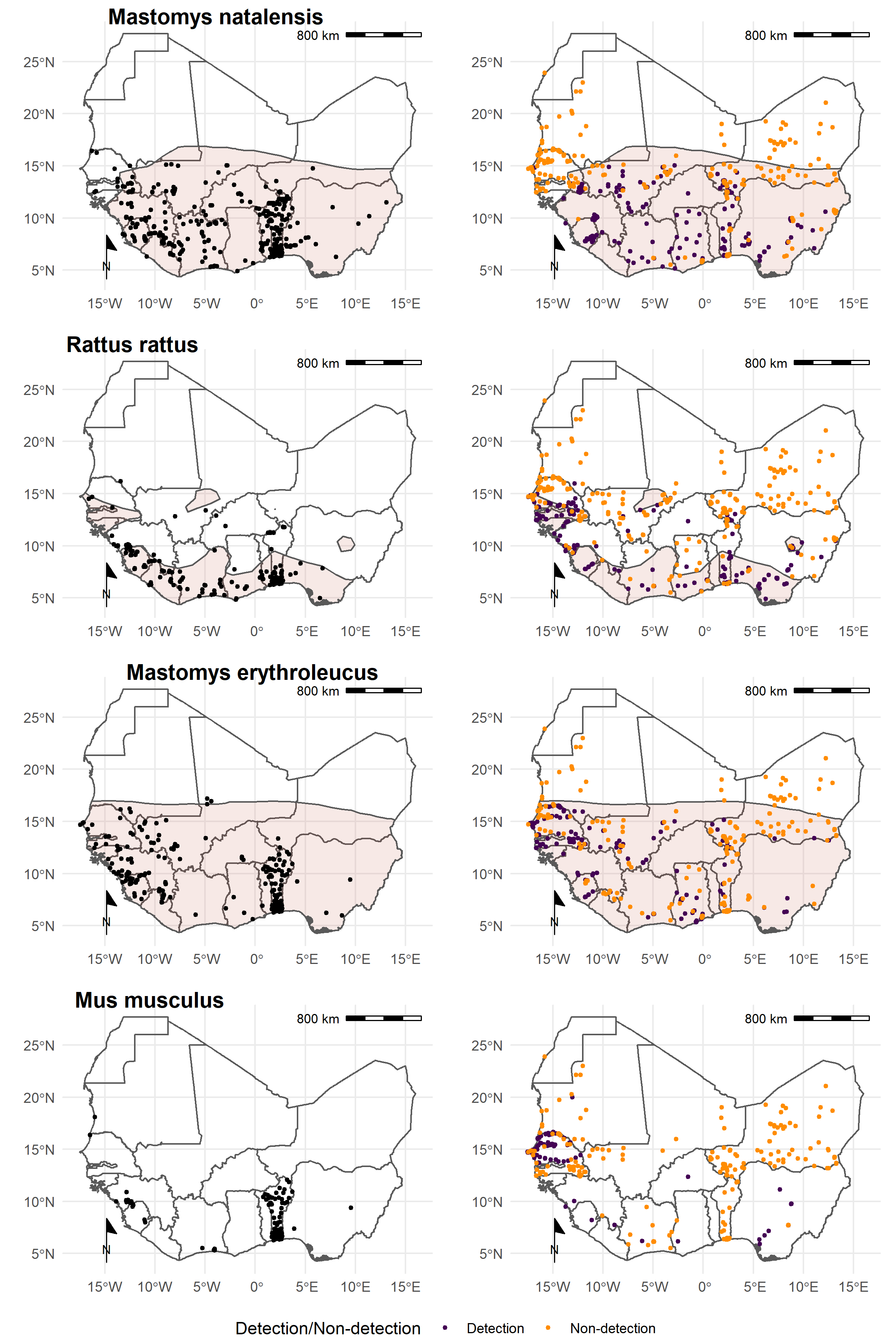


Figure 3A. Each row corresponds to a single rodent species. L) Presence recorded in GBIF overlaid on IUCN species range. R) Detection and non-detection from rodent trapping studies overlaid on IUCN species ranges. No IUCN range for *M. musculus* is available for the study region.

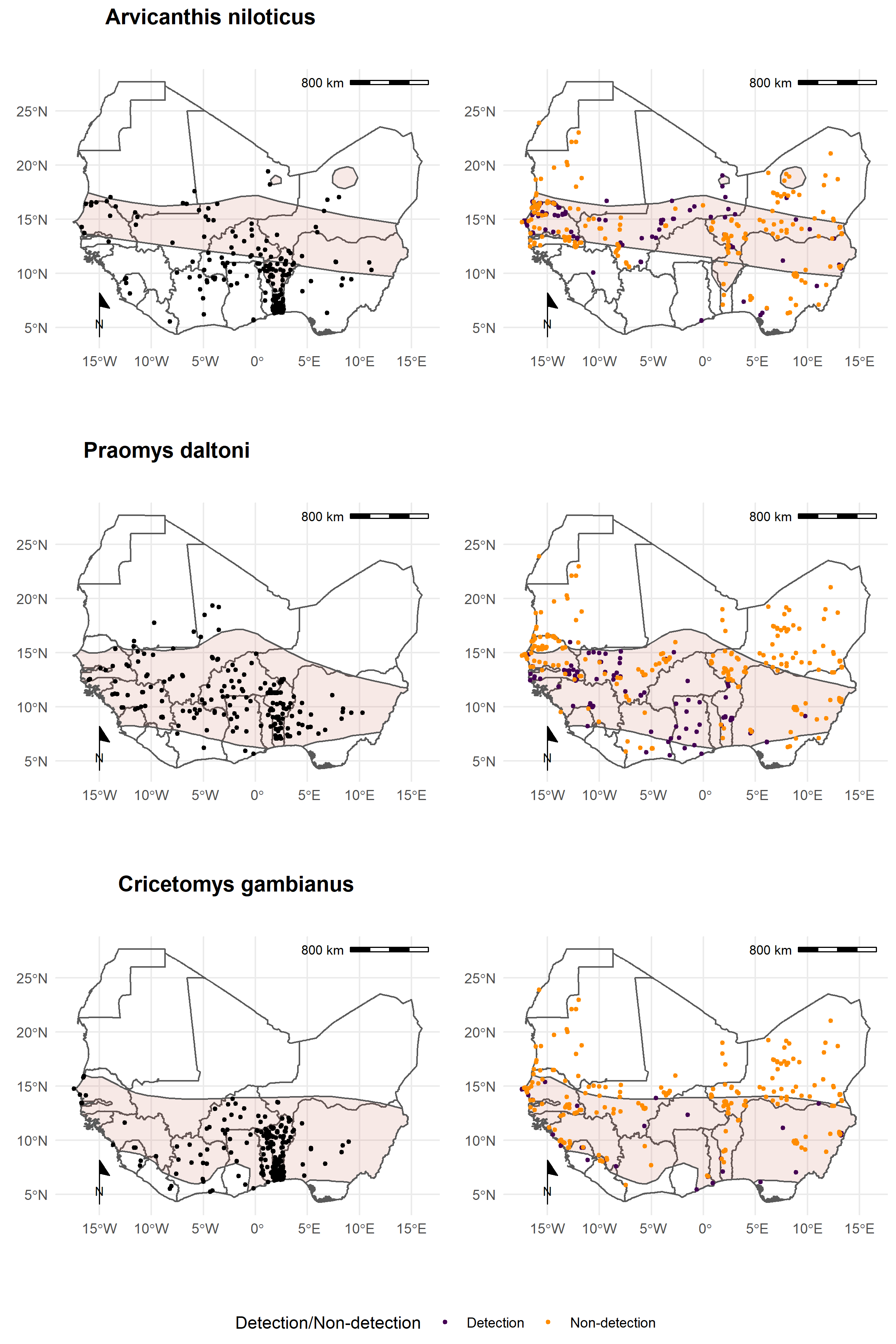


Figure 3B. Each row corresponds to a single rodent species. L) Presence recorded in GBIF overlaid on IUCN species range. R) Detection and non-detection from rodent trapping studies overlaid on IUCN species ranges.

We identified the proportion of a rodents expected range in which detections and non-detections occurred (Table 1.). GBIF sampling occurred over > 15% of a rodents expected range with detections occuring outside of IUCN ranges for all species, most noticeably for *A. niloticus* and *P. daltoni*. Trapping studies occurred over > 13% for all species except *C. gambianus* with detections occurring outside of IUCN ranges most noticeably for *A. niloticus* and *R. rattus*. Combining GBIF and trapping data increased coverage by a mean of 9% of the species range suggesting limited overlap between the regions studied. Non-detections of rodents occurred across large areas of their ranges with *A. niloticus* and *M. erythroleucus* not being detected in >15% of their expected range.

|  | IUCN | GBIF | | Trapping studies | | | Combined |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Species | Range   (1,000 km2) | Range covered  (%) | Area outside range   (1,000 km2) | Range covered  detection (%) | Area outside range   (1,000 km2) | Range covered  non-detection (%) | Range covered  (%) |
| *Mastomys natalensis* | 3,257 | 26.0 | 42 | 17.6 | 12 | 10.3 | 34 |
| *Rattus rattus* | 1,019 | 22.1 | 112 | 17.2 | 217 | 6.5 | 32 |
| *Mastomys erythroleucus* | 3,736 | 15.7 | 100 | 13.9 | 19 | 16.9 | 23 |
| *Arvicanthis niloticus* | 1,830 | 21.0 | 805 | 24.2 | 801 | 22.2 | 41 |
| *Praomys daltoni* | 2,658 | 28.6 | 679 | 15.5 | 45 | 9.4 | 35 |
| *Cricetomys gambianus* | 2,476 | 17.1 | 34 | 4.3 | 23 | 12.8 | 20 |

## Are rodent trapping derived host-pathogen associations present in global zoonoses datasets?

When limited to confirmed pathogens and classification of both pathogen and host to species level we identified 21 host-pathogen pairs among 11 rodent and 6 pathogen species (Figure 4.). We identified negative associations for 41 host-pathogen pairs among 31 rodent and 7 pathogen species. CLOVER contained 8 (38%) of our identified host-pathogen associations and positive associations for 6 (15%) of the negative associations.

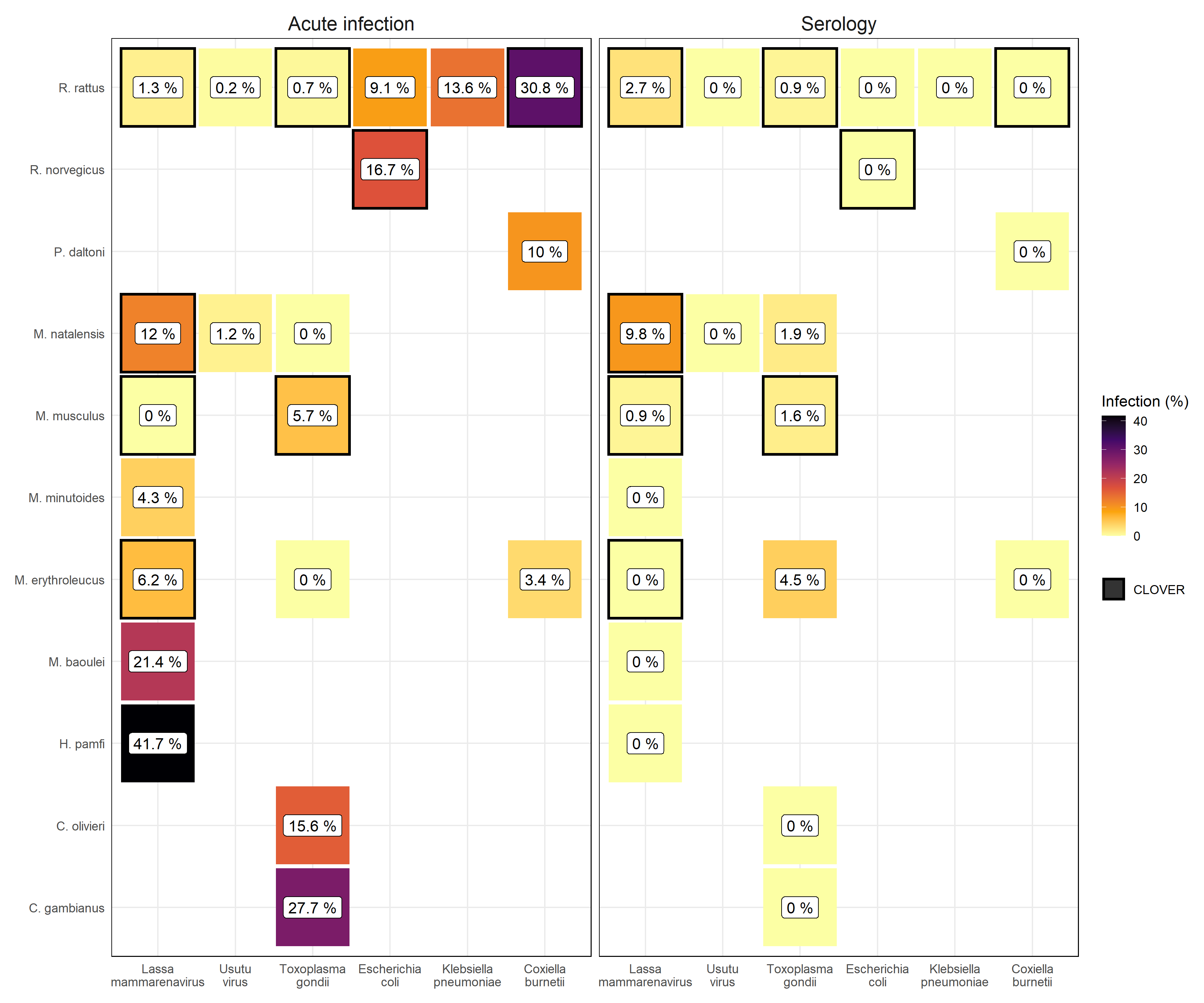


Figure 4. L) Identified species level host-pathogen associations through detection of acute infection (i.e. PCR, culture). Percentages and colour relate to the proportion of all assays that were positive. Associations with a black border are present in the CLOVER dataset. R) Identified species level host-pathogen associations through serological assays (i.e. ELISA). Percentages and colour relate to the proportion of all assays that were positive. Associations with a black border are present in the CLOVER dataset.

CLOVER included 399 host-pathogen associations we do not observe in rodent trapping studies among our included rodent species. The majority of these 329 (82%) pairs are from species with global distributions (*M. musculus*, *R. rattus* and *R. norvegicus*), or from wide ranging distributions in sub-Saharan Africa (31, 9%) (i.e., *A. niloticus*, *M. natalensis* and *A. albiventris*).

For pathogens not identified to species level (i.e. family or lower taxa) we identified 128 host-pathogen pairs among 32 rodent species and 25 pathogen families (Supplementary figure 3.), with CLOVER containing 52 (41%) of these associations.

Rodent trapping studies identified rodent host species for six pathogens, *Toxoplasma gondii* (4), *Lassa mammarenavirus* (3), Usutu virus (2), *Coxiella burnetii* (2), *Escherichia coli* and *Klebsiella pneumoniae* (both 1), that were not present in this global host-pathogen association dataset.

## What is the spatial extent of pathogen testing within a hosts range?

The four most widely sampled pathogen species/families were were Borreliaceae, *Lassa mammarenavirus*, Leptospiraceae and *Toxoplasma gondii* (Table 2.). Infection with species of Borreliaceae was assessed in 42 species, with evidence of infection in 17 rodent species. Among species with more than 500 samples the highest rates of infection were among *A. niloticus* (16%), *M. huberti* (11%) and *M. erythroleucus* (8%). The five most widely assayed species were sampled from a relative small area of their species ranges (7.3-16%), however, there was good representation when compared to where rodent trapping studies have occurred (24-85%). Infection with *Lassa mammarenavirus* was assessed in 28 species, with evidence of infection in 7 rodent species. Among species with more than 100 samples the highest rates of infection were among *M. natalensis* (22%), *M. erythroleucus* (6%) and *R. rattus* (4%). Species sampling occurred from limited areas within IUCN ranges (2-3%), however, sampling was more representative for rodent trapping studies (3-24%). Infection with species of Leptospiraceae was assessed in 8 species, with evidence of infection in 5 rodent species. Among species with more than 100 samples *M. natalensis* (19%), *R. rattus* (11%) and *Crocidura olivieri* (10%) were most frequently infected. Species sampling occurred across less than 1% of these species IUCN ranges and for all except *R. norvegicus* for less than 1% of their detection range in included studies. *Toxoplasma gondii* infection was assessed in 8 species, with evidence of infection in 6 rodent species. Among species with more than 100 samples *M. musculus* (7%), *M. erythroleucus* (4%) and *M. natalensis* (2%) were most commonly infected. Sampling occurred across less than 6% of their IUCN ranges, however, for *M. musculus*, *R. rattus* and *M. erythroleucus* sampling occurred across greater than 12% of their detection range.

| Pathogen | Species | Administrative regions (n) | Tested (n) | Positive (n) | Area sampled  (1,000 km2) | IUCN area (%) | Trapped area (%) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Borrelia sp. |  |  |  |  |  |  |  |
|  | *M. erythroleucus* | 65 | 1,563 | 140 | 273.1 | 7.3 | 50.8 |
|  | *A. niloticus* | 36 | 1,528 | 244 | 293.2 | 16 | 23.6 |
|  | *M. natalensis* | 42 | 731 | 54 | 250.5 | 7.7 | 42.8 |
|  | *M. huberti* | 14 | 718 | 81 | 58.6 | 10.1 | 84.5 |
|  | *P. daltoni* | 30 | 264 | 19 | 234.2 | 8.8 | 51.3 |
| Lassa mammarenavirus |  |  |  |  |  |  |  |
|  | *M. natalensis* | 45 | 3,113 | 713 | 106.9 | 3.3 | 18.3 |
|  | *M. musculus* | 2 | 141 | 0 | 6.2 | \* | 3.1 |
|  | *M. erythroleucus* | 7 | 139 | 9 | 55.9 | 1.5 | 10.4 |
|  | *R. rattus* | 12 | 88 | 2 | 34.5 | 3.4 | 8.8 |
|  | *H. pamfi* | 1 | 12 | 5 | 0.6 | 1.8 | 24.3 |
| Leptospira sp. |  |  |  |  |  |  |  |
|  | *R. rattus* | 54 | 626 | 65 | 3.3 | < 1 | < 1 |
|  | *A. niloticus* | 1 | 221 | 10 | 0.2 | < 1 | < 1 |
|  | *M. natalensis* | 34 | 126 | 24 | 1.6 | < 1 | < 1 |
|  | *C. olivieri* | 39 | 123 | 14 | 1.6 | < 1 | < 1 |
|  | *R. norvegicus* | 27 | 65 | 17 | 1.4 | \* | 17.9 |
| Toxoplasma gondii |  |  |  |  |  |  |  |
|  | *M. musculus* | 32 | 1,548 | 115 | 77.5 | \* | 38.1 |
|  | *R. rattus* | 18 | 428 | 8 | 61.7 | 6.1 | 15.7 |
|  | *M. erythroleucus* | 18 | 272 | 10 | 68.2 | 1.8 | 12.7 |
|  | *C. gambianus* | 3 | 47 | 13 | 2.4 | < 1 | 1.8 |
|  | *C. olivieri* | 2 | 32 | 5 | 0.5 | < 1 | < 1 |

# Discussion

We synthesised data from 126 rodent trapping studies containing information on more than 72,000 trapped rodents across 1,331 trapping sites producing an estimated 966,000 trap-nights. Trapping occurred across 14 West African countries, studies investigating known or potential zoonotic pathogens occurred within 19% of level 2 administrative regions across West Africa with important bias towards areas with high population density and human dominated landscapes. For several important rodent species we identified discordance between IUCN species ranges, GBIF presence data and detections and non-detections from rodent trapping studies, particularly among non-native species (*R. rattus* and *M. musculus*). We compared host-pathogen associations from rodent trapping studies with a global dataset and identified missing associations for several important zoonotic infectious diseases including, Lassa fever, Toxoplasmosis and Q fever in commonly occurring rodent species. Finally, we assessed the spatial coverage of testing for important zoonotic pathogens and found low representativeness across their host species range.

Rodent trapping studies provide important information on the distribution of rodent species and their assemblages, we have identified important bias in the locations in which studies to data have been performed with large regions of countries including Nigeria, Guinea and Cote d’Ivoire being understudied. Trapping sites are clustered in areas with relatively high human population density and anthropogenic land use which can importantly bias our understanding of the distribution of rodent species and their pathogens across West Africa. Further effort is required to minimise this bias by sampling currently undersampled regions.

Global datasets of rodent host distribution (e.g., IUCN) and presence (e.g., GBIF) provide useful resources to model the hazard of zoonotic disease spillover and novel pathogen emergence. We show here that non-detection data from rodent trapping studies can provide additional information to describe the heterogeneity of rodent assemblages across diverse habitats in West Africa. Combining presence data from GBIF and detection/non-detection data from rodent trapping studies can substantially increase coverage across a rodents range, allowing improved estimates of rodent host prevalence for zoonotic infectious disease studies.

Host-pathogen association datasets are becoming increasing comprehensive alongside increased access to molecular discovery techniques for potential pathogens. We show that currently available data from rodent trapping studies can supplement these resources to better understand the potential hosts and reservoirs of known zoonotic pathogens. Rodent trapping data can support studies using these global datasets to predict novel pathogen emergence and to quantify the risk of known zoonotic disease spillover based on host distributions.

We have identified important limitations in the sampling of rodent zoonoses for known diseases across West Africa. For example, studies of *Lassa mammarenavirus* have taken place across 3% of *M. natalensis*’s range in West Africa. Lassa fever is known to be endemic in eight West African countries, although it is potentially more widely spread. Further studies are required to investigate this pathogen across a wider geographic range. Other important zoonotic diseases including Leptospirosis and Toxoplasmosis have been assessed over very limited regions of their hosts distribution.

We propose that rodent trapping studies can provide important data to supplement global datasets to quantify the hazard of zoonotic disease spillover and novel pathogen emergence. They can reduce some of the bias in global datasets however, important limitations remain. Further effort is required to survey rodent populations in non-human dominated landscapes and to assay pathogens across more of a hosts range.

# Supplementary

## Supplementary figure 3



Figure 3. L) Identified host-pathogen associations at pathogen family level through detection of acute infection (i.e. PCR, culture). Percentages and colour relate to the proportion of all assays that were positive. Associations with a black border are present in the CLOVER dataset. R) Identified host-pathogen associations at pathogen family level through serological assays (i.e. ELISA). Percentages and colour relate to the proportion of all assays that were positive. Associations with a black border are present in the CLOVER dataset.

[1] Allen T, Murray KA, Zambrana-Torrelio C, Morse SS, Rondinini C, Di Marco M, et al. Global hotspots and correlates of emerging zoonotic diseases. Nature Communications 2017;8:1124. <https://doi.org/10.1038/s41467-017-00923-8>.

[2] Hassell JM, Begon M, Ward MJ, Fèvre EM. Urbanization and disease emergence: Dynamics at the wildlife–livestock–human interface. Trends in Ecology & Evolution 2017;32:55–67. <https://doi.org/10.1016/j.tree.2016.09.012>.

[3] García-Peña GE, Rubio AV, Mendoza H, Fernández M, Milholland MT, Aguirre AA, et al. Land-use change and rodent-borne diseases: Hazards on the shared socioeconomic pathways. Philosophical Transactions of the Royal Society B: Biological Sciences 2021;376:20200362. <https://doi.org/10.1098/rstb.2020.0362>.

[4] McMahon BJ, Morand S, Gray JS. Ecosystem change and zoonoses in the anthropocene. Zoonoses and Public Health 2018;65:755–65. <https://doi.org/10.1111/zph.12489>.

[5] Grace D, Mutua F, Ochungo P, Kruska R, Jones K, Brierley L, et al. Mapping of poverty and likely zoonoses hotspots 2012.

[6] Han BA, Kramer AM, Drake JM. Global patterns of zoonotic disease in mammals. Trends in Parasitology 2016;32:565–77. <https://doi.org/10.1016/j.pt.2016.04.007>.

[7] Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. PNAS 2015;112:7039–44. <https://doi.org/10.1073/pnas.1501598112>.

[8] Gibb R, Albery GF, Mollentze N, Eskew EA, Brierley L, Ryan SJ, et al. Mammal virus diversity estimates are unstable due to accelerating discovery effort. Biology Letters 2021;18:20210427. <https://doi.org/10.1098/rsbl.2021.0427>.

[9] Meerburg B, Singleton G, Kijlstra A. Rodent-borne diseases and their risks for public health. Critical Reviews in Microbiology 2009;35:221–70. <https://doi.org/10.1080/10408410902989837>.

[10] Galeh TM, Sarvi S, Montazeri M, Moosazadeh M, Nakhaei M, Shariatzadeh SA, et al. [Global status of toxoplasma gondii seroprevalence in rodents: A systematic review and meta-analysis](https://www.frontiersin.org/article/10.3389/fvets.2020.00461). Frontiers in Veterinary Science 2020;7.

[11] Fischer C, Gayer C, Kurucz K, Riesch F, Tscharntke T, Batáry P. Ecosystem services and disservices provided by small rodents in arable fields: Effects of local and landscape management. Journal of Applied Ecology 2018;55:548–58. <https://doi.org/10.1111/1365-2664.13016>.

[12] Dobson FS, Oli MK. [Fast and slow life histories of rodents](https://www.jstor.org/stable/42902037). Rodent Societies: An Ecological and Evolutionary Perspective 2007:99–105.

[13] Albery GF, Becker DJ. Fast-lived hosts and zoonotic risk. Trends in Parasitology 2021;37:117–29. <https://doi.org/10.1016/j.pt.2020.10.012>.

[14] Gibb R, Redding DW, Chin KQ, Donnelly CA, Blackburn TM, Newbold T, et al. Zoonotic host diversity increases in human-dominated ecosystems. Nature 2020;584:398–402. <https://doi.org/10.1038/s41586-020-2562-8>.

[15] Gibb R, Franklinos LHV, Redding DW, Jones KE. Ecosystem perspectives are needed to manage zoonotic risks in a changing climate. BMJ 2020;371:m3389. <https://doi.org/10.1136/bmj.m3389>.

[16] Plowright RK, Becker DJ, McCallum H, Manlove KR. Sampling to elucidate the dynamics of infections in reservoir hosts. Philos Trans R Soc Lond B Biol Sci 2019;374:20180336. <https://doi.org/10.1098/rstb.2018.0336>.

[17] Carlson CJ, Farrell MJ, Grange Z, Han BA, Mollentze N, Phelan AL, et al. The future of zoonotic risk prediction. Philosophical Transactions of the Royal Society B: Biological Sciences 2021;376:20200358. <https://doi.org/10.1098/rstb.2020.0358>.

[18] Wille M, Geoghegan JL, Holmes EC. How accurately can we assess zoonotic risk? PLoS Biol 2021;19:e3001135. <https://doi.org/10.1371/journal.pbio.3001135>.

[19] Redding DW, Lucas TCD, Blackburn TM, Jones KE. Evaluating bayesian spatial methods for modelling species distributions with clumped and restricted occurrence data. PLOS ONE 2017;12:e0187602. <https://doi.org/10.1371/journal.pone.0187602>.

[20] Beck J, Böller M, Erhardt A, Schwanghart W. Spatial bias in the GBIF database and its effect on modeling species’ geographic distributions. Ecological Informatics 2014;19:10–5. <https://doi.org/10.1016/j.ecoinf.2013.11.002>.

[21] Bovendorp RS, MCCleery RA, Galetti M. Optimising sampling methods for small mammal communities in neotropical rainforests. Mammal Review 2017;47:148–58. <https://doi.org/10.1111/mam.12088>.

[22] Fichet-Calvet E, Lecompte E, Veyrunes F, Barrière P, Nicolas V, Koulémou K. [Diversity and dynamics in a community of small mammals in coastal guinea, west africa](http://www.vliz.be/nl/open-marien-archief?module=ref&refid=204614). Belgian Journal of Zoology 2009;139:93–102.

[23] Catalano S, Leger E, Fall CB, Borlase A, Diop SD, Berger D, et al. Multihost transmission of schistosoma mansoni in senegal, 2015-2018. Emerging Infectious Diseases 2020;26:1234–42. <https://doi.org/10.3201/eid2606.200107>.

[24] USAID. One health surveillance. PREDICT project 2021. <https://p2.predict.global/surveillance> (accessed September 27, 2021).

[25] GBIF: The Global Biodiversity Information Facility. The global biodiversity information facility 2021. <https://www.gbif.org/> (accessed April 25, 2021).

[26] Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: An introduction with mgcv. PeerJ 2019;7:e6876. <https://doi.org/10.7717/peerj.6876>.

[27] Socioeconomic Data and Applications Center. Gridded population of the world (GPW), v4 SEDAC 2021. <https://sedac.ciesin.columbia.edu/data/collection/gpw-v4> (accessed February 4, 2021).

[28] GBIF: The Global Biodiversity Information Facility. Occurrence download 2021. <https://doi.org/10.15468/DL.S52MEH>.

[29] IUCN. [The IUCN red list of threatened species](https://www.iucnredlist.org). 2021.

[30] Simons D. Rodents and potential zoonotic diseases in west africa 2022. <https://diddrog11.shinyapps.io/scoping_review_app/> (accessed March 2, 2022).