

Rodent trapping studies as an overlooked information source for understanding endemic and novel zoonotic spillover.

David Simons Lauren A. Attfield Kate E. Jones Deborah Watson-Jones
Richard Kock

2022-08-24

Abstract

Rodents are important globally distributed reservoirs of known and novel zoonotic pathogens. Ongoing anthropogenic land use change is altering the composition of host species assemblages and modifying the risk of zoonoses spillover events. These changes mean that an understanding of the current distribution of rodent species is vital for accurately describing disease hazard and managing risk. However, available species distribution and host-pathogen association datasets (e.g. IUCN, GBIF, CLOVER) are often taxonomically and spatially biased. Here, we synthesise data from West Africa from 127 rodent trapping studies, published between 1964-2022, as an additional source of information to characterise the range and presence of important zoonotic pathogen host species in this region. We identify that these rodent trapping studies, although biased towards human dominated landscapes across West Africa, can usefully complement current rodent species distribution datasets and we calculate the discrepancies between these datasets. For five regionally important zoonotic pathogens (Arenaviridae spp., *Borrelia* spp., *Lassa mammarenavirus*, *Leptospira* spp. and *Toxoplasma gondii*), we identify host-pathogen associations that have not been previously reported in host-association datasets. These omissions have the potential for biasing estimates of current risk and drivers of zoonoses. Finally, for these five pathogen groups, we find that the proportion of a rodent hosts range that has been sampled remains small. A priority of future rodent trapping studies should be to sample rodent hosts across a greater geographic range to better characterise current and future risk. In the interim, future studies of spatial pathogen risk informed by rodent distributions must incorporate a measure of the current sampling biases. The current synthesis of contextually rich rodent trapping data enriches available information from IUCN, GBIF and CLOVER which can support a more complete understanding of the hazard of zoonotic spillover events.

Introduction

There is increasing awareness of the global health and economic impacts of novel zoonotic pathogen spillover, evidenced by the ongoing SARS-CoV-2 pandemic and previous HIV/AIDS and Spanish Influenza pandemics (Bernstein *et al.*, 2022). The number of zoonotic disease spillover events and the frequency of the emergence of novel zoonotic pathogens from rodents are predicted to increase under intensifying anthropogenic pressure driven by increased human populations, urbanisation, intensification of agriculture, climate change and wildlife defaunation (Allen *et al.*, 2017; Hassell *et al.*, 2017; McMahon, Morand and Gray, 2018; García-Peña *et al.*, 2021). The impact of endemic zoonoses meanwhile remains underestimated (Maudlin, Eisler and Welburn, 2009). Endemic zoonoses disproportionately affect those in the poorest sections of society, those living in close contact with their animals and those with limited access to healthcare (Molyneux *et al.*, 2011; Halliday *et al.*, 2015; Judson and Rabinowitz, 2021).

Rodents along with bats contribute the greatest number of predicted novel zoonotic pathogens and known zoonoses (Han *et al.*, 2015; Gibb, Gregory F. Albery, *et al.*, 2021b). Of 2,220 extant rodent species, 244 (10.7%) are described as reservoirs of 85 zoonotic pathogens (Han *et al.*, 2015), although many species provide

important and beneficial ecosystem services including pest regulation and seed dispersal (Fischer *et al.*, 2018). Rodents typically demonstrate “fast” life histories (Dobson and Oli, 2007) with traits such as early maturation and short gestation times (<4 days) further associated with being zoonotic reservoirs (Han *et al.*, 2015; Albery and Becker, 2021). Rodent species with “fast” life histories thrive in human dominated landscapes, displacing species less likely to be reservoirs of zoonotic pathogens (Gibb, Redding, *et al.*, 2020). The widespread occurrence of reservoir species and their proximity to human activity make the description of rodent species assemblages and host-pathogen associations vitally important to understanding the hazard of zoonotic disease spillover and novel zoonotic pathogen emergence (Han, Kramer and Drake, 2016).

Despite the importance of understanding these complex systems, current evidence on host-pathogen associations is considerably affected by taxonomic and geographical sampling biases (Gibb, Franklino, *et al.*, 2020; Gibb, Gregory F. Albery, *et al.*, 2021b). Curated biodiversity datasets such as the Global Biodiversity Information Facility (GBIF) and resources produced by the International Union for Conservation of Nature suffer from well described spatial and temporal sampling biases (Boakes *et al.*, 2010; Bowler *et al.*, 2022). These sampling biases can importantly distort produced species distribution models (Beck *et al.*, 2014). Datasets on host-pathogen associations also can suffer from biases introduced from literature selection criteria and taxonomic discrepancies. These biases are important because identification of potential geographic hotspots of zoonotic disease spillover and novel pathogen emergence are often based on these types of host species distributions and host-pathogen associations (Plowright *et al.*, 2019; Carlson *et al.*, 2021). 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 (Redding *et al.*, 2017; Wille, Geoghegan and Holmes, 2021; Gibb, Gregory F. Albery, *et al.*, 2021b). Predictions of zoonotic disease spillover and novel zoonotic pathogen emergence must account for these biases to understand the future hazard of zoonotic diseases (Carlson *et al.*, 2021).

West Africa has been identified as a region at increased risk for rodent-borne zoonotic disease spillover events, the probability of these events are predicted to increase under different projected future land-use change scenarios (Grace *et al.*, 2012; García-Peña *et al.*, 2021). Currently within West Africa, rodents are involved in the transmission of multiple endemic zoonoses with large burdens on human health, these pathogens include, Lassa fever, Schistosomiasis, Leptospirosis and Toxoplasmosis (Meerburg, Singleton and Kijlstra, 2009; Galeh *et al.*, 2020). Understanding of the distribution of these zoonoses are limited by biases in consolidated datasets. Rodent trapping studies provide contextually rich information on when, where and under what conditions rodents were trapped, potentially enriching consolidated datasets (Bovendorp, McCleery and Galetti, 2017). Studies have been conducted in West Africa to investigate the distribution of rodent species, their species assemblages, the prevalence of endemic zoonoses within rodent hosts (e.g., Lassa fever, Schistosomiasis) and to identify emerging and novel zoonotic pathogens (Fichet-Calvet *et al.*, 2009; Catalano *et al.*, 2020; USAID, 2021). However, individual level data from these studies have not previously been synthesised for inclusion in assessments of zoonotic disease spillover and novel zoonotic pathogen emergence.

Here, we synthesise rodent trapping studies conducted across West Africa published between 1964-2022. First, we use this dataset to investigate the geographic sampling biases in relation to human population density and land use classification. Second, we compare this to curated host datasets (IUCN and GBIF) to understand differences in reported host geographic distributions. Third, we compare identified host-pathogen associations with a consolidated dataset (CLOVER) to explore discrepancies in rodent host-pathogen associations and report the proportion of positive assays for pathogens of interest. Finally, within our dataset we investigate the spatial extent of current host-pathogen sampling to identify areas of sparse sampling of pathogens within their host ranges. We expect that rodent trapping studies provide an important additional source of high-resolution data that can be used to enrich available consolidated datasets to better understand the hazard of zoonotic disease spillover and novel zoonotic pathogen emergence across West Africa.

Methods

Data sources

Host and pathogen trapping data

To identify relevant literature, 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 exploded keywords: (1) Rodent OR Rodent trap* AND (2) West Africa, no date limits were set. We also 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 2022-05-01, and returned studies conducted between 1964-2021.

We included studies for further analysis 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, and 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 author screened titles, abstracts and full texts against the inclusion and exclusion criteria. At each stage; title screening, abstract screening and full text review, a random subset (10%) was reviewed by a second author.

We extracted data from eligible studies using a standardised tool that was piloted on 5 studies (Supplementary Table 1.). Data was abstracted into a Google Sheets document, which was archived on completion of data extraction [ref to zenodo on finalised]. We identified the aims of included studies, for example, whether it was conducted as a survey of small mammal species or specifically to assess the risk of zoonotic disease spillover. We extracted data on study methodology, such as, the number of trap nights, the type of traps used and whether the study attempted to estimate abundance. For studies not reporting number of trap nights we used imputation based on the number of trapped individuals, stratified by the habitat type from which they were obtained. This was performed by multiplying the total number of trapped individuals within that study site by the median trap success for study sites with the same reported habitat type. Stratification was used as trap success varied importantly between traps placed in or around buildings (13%, IQR 6-24%) compared with other habitats (3%, IQR 1-9%).

We also recorded how species were identified within a study and species identification was assumed to be accurate. The number of individuals of these species or genera was extracted with taxonomic names mapped to GBIF taxonomy (GBIF: The Global Biodiversity Information Facility, 2021b). We expanded species detection and non-detection records by explicitly specifying non-detection at a trap site if a species was recorded as detected at other trapping locations within the same study.

Geographic locations of trapping studies were extracted using GPS locations for the most precise location presented. Missing locations were found using the National Geospatial-Intelligence Agency GEOnet Names Server (National Geospatial-Intelligence Agency, no date) based on place-names and maps presented in the study. All locations were converted to decimal degrees. The year of rodent trapping was extracted alongside the length of the trapping activity to understand seasonal representativeness of trapping activity. The habitats of trapping sites were mapped to the IUCN Habitat Classification Scheme (Version 3.1). For studies reporting multiple habitat types for a single trap, trap-line or trapping grid, a higher order classification of habitat type was recorded.

For relevant studies we extracted data on all microorganisms and zoonotic pathogens tested and the method used (e.g., molecular or serological diagnosis). Where assays were able to identify the microorganism to species level this was recorded, non-specific assays higher order attribution was used (e.g. to family level). We recorded the species of rodent host tested, the number of individuals tested and the number of positive and negative results. For studies reporting summary results all testing data were extracted, this may introduce double counting of individual rodents, for example, if a single rodent was tested using both molecular and serological assays. Where studies reported indeterminate results, these were also recorded.

Out of 4,692 relevant citations, we identified 127 rodent trapping studies (Supplementary Table 2.). The earliest trapping studies were conducted in 1964, with a trend of increasing numbers of studies being performed annually since 2000. The median year of first trapping activity was 2007, with the median length of trapping activity being 1 year (IQR 0-2 years) (Supplementary Fig 1.). Studies were conducted in 14 West African countries, with no studies reported from The Gambia or Togo, at 1,611 trap sites (Fig 1A.).

Included studies explicitly reported on 601,184 trap nights, a further 341,445 trap nights were imputed from studies with no recording of trapping effort based on trap success, leading to an estimate of 942,629 trap nights (Fig 1B.). A minority of studies trapped at a single study site (30, 24%), with 46 (36%) trapping at between two and five sites, the remaining 51 studies (40%) trapped at between six and 93 study sites.

In total 76,275 small mammals were trapped with 65,628 (90%) identified to species level and 7,439 (10%) identified to genus, with the remaining classified to higher taxonomic level. The majority of the 132 identified species were Rodentia (102, 78%), of which Muridae (73, 72%) were the most common family. Soricomorpha were the second most identified order of small mammals (28, 21%). 57 studies tested for 32 microorganisms, defined to species or genus level that are known or potential pathogens. Most studies tested for a single microorganism (48, 84%). The most frequently assayed microorganisms were *Lassa mammarenavirus* or Arenaviridae (21, 37%), *Borrelia* sp. (9, 16%), *Bartonella* sp. (4, 7%) and *Toxoplasma gondii* (4, 7%). Most studies used Polymerase Chain Reaction (PCR) to detect microorganisms (37, 65%), with fewer studies using serology-based tests (11, 19%) or histological or direct visualisation assays (11, 21%). From 32,920 individual rodent samples we produced 351 host-pathogen pairs. With *Rattus rattus*, *Mus musculus*, *Mastomys erythroleucus*, *Mastomys natalensis* and *Arvicanthis niloticus* being assayed for at least 18 microorganisms.

Analysis

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

To investigate the extent of spatial bias in the rodent trapping data, we calculated trap-night (TN) density within each West African level-2 administrative region. The *sf* package in the R statistical language was used to manipulate geographic data, administrative boundaries were obtained from GADM 4.0.4 (Pebesma, 2018; R Core Team, 2020; Database of Global Administrative Areas, 2022). Density ($TN_{density}$) was calculated by dividing the number of trap nights by the area of a level-2 administrative area (R_{area}). For studies not reporting trap nights imputation was used as previously described. Human population density was obtained for the closest year (2005) to the median year of trapping (2007) from Socioeconomic Data and Applications Center (SEDAC) gridded population of the world v4 at ~1km resolution ($P_{density}$) (Socioeconomic Data and Applications Center, 2021). Median population density was then calculated for each level-2 administrative region. Land cover classification was obtained from the Copernicus climate change service at ~300m resolution (European Space Agency Climate Change Initiative, 2022). The proportion of cropland, shrubland, tree cover (ψ_{tree}) and urban land cover (ψ_{urban}) within a level-2 administrative region in 2005 was calculated.

We investigated the association between relative trapping effort, measured as TN density, and the proportion of urban, cropland, tree cover and human population density using Generalised Additive Models (GAM) incorporating a spatial interaction term (longitude and latitude) (Pedersen *et al.* (2019)). The models were constructed in the *mgcv* package (Wood (2017)). Selection of the most parsimonious model was based on Deviance explained and the Akaike Information Criterion for each model (Equations 1-5 below). Relative trapping effort was then predicted across West Africa using these covariates. We performed two sensitivity analyses, first, by removing sites with imputed trapping effort. Second rather than associating trap locations to level-2 administrative areas we assigned them to ~1km pixels.

$$TN_{density} \sim Tweedie(X * Y)$$

$$TN_{density} \sim Tweedie(P_{density} + (X * Y))$$

$$TN_{density} \sim Tweedie(P_{density} + R_{area} + (X * Y))$$

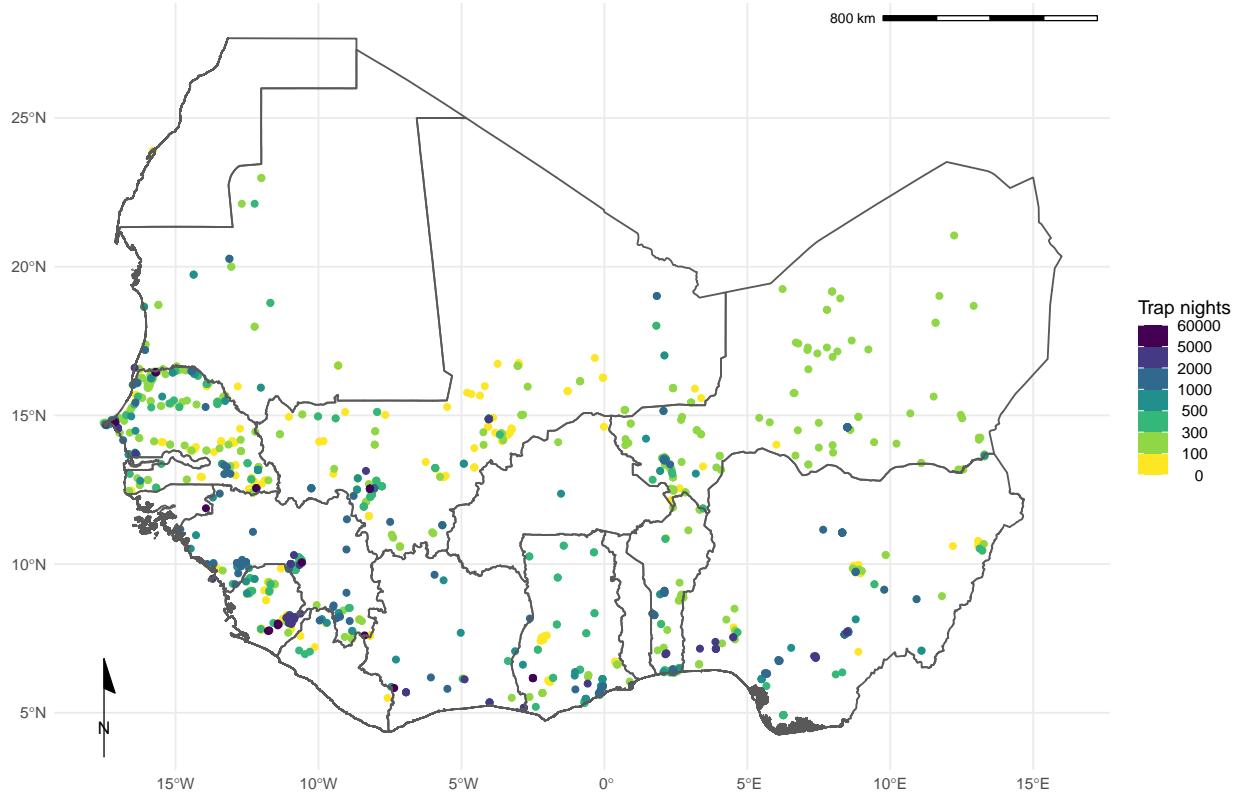
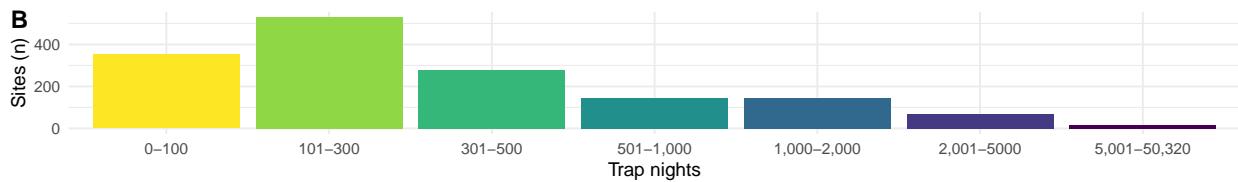
A**B**

Figure 1: Fig 1: A) The location of trapping sites in West Africa. No sites were recorded from Togo or The Gambia. Heterogeneity is 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) Histogram of trap nights performed at each study site, a median of 248 trap nights (IQR 116–500) was performed at each site

$$TN_{\text{density}} \sim \text{Tweedie}(P_{\text{density}} + \psi_{\text{tree}} + \psi_{\text{urban}} + (X * Y))$$

$$TN_{\text{density}} \sim \text{Tweedie}(Pop_{\text{density}} + R_{\text{area}} + \psi_{\text{urban}} + (X * Y))$$

What is the difference in rodent host distributions between curated datasets and rodent trapping studies?

We assessed the concordance of curated rodent host distributions from IUCN and GBIF with observed rodent presence and absence from rodent trapping studies for seven species with the most trap locations (*M. natalensis*, *R. rattus*, *M. erythroleucus*, *M. musculus*, *A. niloticus*, *Praomys daltoni* and *Cricetomys gambianus*). We obtained rodent species distribution maps as shapefiles from the IUCN red list and translated these to a ~20km resolution raster (IUCN, 2021). Distributions were cropped to the study region for globally distributed rodent species. We obtained rodent presence locations from GBIF as point data limited to the study region (GBIF: The Global Biodiversity Information Facility, 2021a). Presence locations were associated to cells of raster with a ~20km resolution produced for the study region.

For each of the seven species, we first calculated the area of the IUCN expected range, and then the percentage of this range covered by presence detections in GBIF, and from presence detections in the rodent trapping data. We then calculated the area of both types of detections outside of the IUCN expected range. For rodent trapping data, we calculated the area of non-detections within the IUCN expected area. Finally, we calculated the combined area of detection from both GBIF and rodent trapping data.

Are rodent trapping derived host-pathogen associations present in a consolidated zoonoses dataset?

To examine the usefulness of rodent trapping studies as an additional source of data we compared identified host-pathogen associations from trapping studies investigating zoonoses with a consolidated zoonoses dataset (CLOVER) (Gibb, Gregory F. Albery, *et al.*, 2021b; Gibb, Carlson and Farrell, 2021). CLOVER is a synthesis of four host-pathogen datasets (GMPD2, EID2, HP3 and Shaw, 2020) and was released in 2021, it contains more than 25,000 host-pathogen associations for Bacteria, Viruses, Helminth, Protozoa and Fungi. We compared the host-pathogen networks across the two datasets, where the CLOVER data was subset for host species present in the rodent trapping data.

For host-pathogen pairs with assay results consistent with acute or prior infection, we calculated the proportion positive and identify those absent from CLOVER. We expand the analysis to host-pathogen pairs with pathogens identified to genus level in Supplementary Fig 4.

What is the spatial extent of pathogen testing within host ranges?

We use the sampled area of three pathogen groups and two pathogens (Arenaviridae, Borreliaceae, Leptospiraceae, *Lassa mammarenavirus* and *Toxoplasma gondii*) to quantify the bias of sampling within their hosts ranges. For each pathogen, we first describe the number of host species assayed, for the five most commonly tested species we associate the locations of sampled individuals to ~20km pixels and calculate the proportion of the IUCN range of the host in which sampling has occurred. We compare this figure to the total area in which the host has been detected to produce a measure of relative completeness of sampling within the included rodent trapping studies.

All data and code to reproduce these analyses are available in an archived Zenodo repository [ref_zenodo].

Results

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

We found a non-random spatial distribution of rodent trapping locations across the study region, suggestive of underlying bias in the sampling or rodents across West Africa. Trap sites were situated in 256 of 1,450

(17.6%) level-2 administrative regions in 14 West African nations. The regions with highest TN density included the capitals and large cities of Niger (Niamey), Nigeria (Ibadan), Ghana (Accra), Senegal (Dakar), Ghana (Accra) and Benin (Cotonou). Outside of these cities, regions in, Northern Senegal, Southern Guinea, Edo and Ogun States in Nigeria and Eastern Sierra Leone had the greatest TN density (Fig 1A.).

The most parsimonious GAM model (adjusted R² = 0.3, Deviance explained = 48.7%) reported significant non-linear associations between relative trapping effort bias and human population densities (Effective Degrees of Freedom (EDF) = 7.13, $p < 0.001$), proportion of urban landscape (EDF = 1.92, $p < 0.002$) and region area (EDF = 3.63, $p < 0.001$), alongside significant spatial associations (EDF = 27.3, $p < 0.001$) (Supplementary table 3.1). Greatest trapping effort bias peaked at population densities between 5,000-7,500 individuals/km², urban landscape >10% and region areas < 1,000km². Increased trapping effort was found in North West Senegal, North and East Sierra Leone, Central Guinea and coastal regions of Nigeria, Benin and Ghana; in contrast South East Nigeria, Northern Nigeria and Burkina Faso had an observed bias towards a reduced trapping effort (Fig 2). In sensitivity analysis, excluding sites with imputed trap nights, Mauritania, Northern Senegal and Sierra Leone remained as regions trapped at higher rates, with Nigeria being trapped at lower than expected rates (Supplementary Fig 3a.). In pixel based sensitivity analysis spatial coverage was reduced with similar patterns of bias observed to the primary analysis (Supplementary Fig 3b.).

What is the difference in rodent host distributions between curated datasets and rodent trapping studies?

We found that for six of the seven rodent (*M. natalensis*, *R. rattus*, *M. erythroleucus*, *M. musculus*, *A. niloticus* and *P. daltoni*), trapping studies provided more distinct locations of detection and non-detection than were available from GBIF. For the endemic rodent species (*M. natalensis*, *M. erythroleucus*, *A. niloticus*, *P. daltoni* and *C. gambianus*) IUCN ranges had good concordance to both trapping studies and GBIF, however, individuals of *A. niloticus* and *P. daltoni* were detected outside of IUCN ranges. In contrast, the non-native species *R. rattus* and *M. musculus* were detected across much greater ranges than were expected from IUCN distributions. Comparisons for *M. natalensis*, *R. rattus* and *M. musculus* are shown in Fig 3, the remaining species are shown in Supplementary Fig 4.

Comparison of proportion of a species IUCN range in which detections and non-detections occurred showed that sampling locations of these seven species within GBIF covered between 0.09-0.26% of expected ranges (Table 1.). Detections occurred outside IUCN ranges for all species, most noticeably for *A. niloticus* and *R. rattus*. Sampling locations within rodent trapping data covered 0.03-0.24% of expected ranges with detections occurring outside IUCN ranges most noticeably for *A. niloticus* and *R. rattus*. Combining GBIF and rodent trapping data increased the sampled area by a mean of 1.65 times, suggesting limited overlap between the locations sampled. Non-detection of species occurred across species ranges (mean = 0.11%, SD = 0.03%), suggestive of spatial heterogeneity of presence within IUCN ranges.

| | IUCN | GBIF | |
|-------------------------------|-----------------------------------|--|---|
| Species | Range (1,000 km ²) | Area inside range (1,000 km ²) (% of IUCN) | Area outside range (1,000 km ²) |
| <i>Mastomys natalensis</i> | 3,257.11 | 6.83 (0.21%) | 0.19 |
| <i>Rattus rattus</i> | 1,018.71 | 2.61 (0.26%) | 0.52 |
| <i>Mastomys erythroleucus</i> | 3,735.48 | 4.48 (0.12%) | 0.04 |
| <i>Mus musculus</i> | | | 2.15 |
| <i>Arvicantis niloticus</i> | 1,829.14 | 1.69 (0.09%) | 2.41 |

| | IUCN | GBIF | |
|-----------------------------|-----------------------------------|--|---|
| Species | Range (1,000 km ²) | Area inside range (1,000 km ²) (% of IUCN) | Area outside range (1,000 km ²) |
| <i>Praomys daltoni</i> | 2,657.77 | 4.03 (0.15%) | 0.29 |
| <i>Cricetomys gambianus</i> | 2,475.97 | 5 (0.2%) | 0.17 |

Are rodent trapping derived host-pathogen associations present in a consolidated zoonoses dataset?

We found potentially important differences between the host-pathogen networks produced from included rodent trapping studies and the consolidated CLOVER dataset. When limited to taxonomic classification of both pathogen and host to species level we identified 25 host-pathogen pairs among 14 rodent and 6 pathogen species (Fig 4.). We identified negative associations (non-detection through specific assays) for 45 host-pathogen pairs among 35 rodent and 7 pathogen species. CLOVER contained 10 (40%) of our identified host-pathogen associations, the remaining 15 (60%) were not found to be present in CLOVER, additionally CLOVER recorded positive associations for 4 (9%) of the negative associations produced from the rodent trapping data.

CLOVER included an additional 492 host-pathogen associations we do not observe in rodent trapping studies. The majority of these 392 (80%) pairs are from species with global distributions (*M. musculus*, *R. rattus* and *R. norvegicus*), or from those with wide ranging distributions in sub-Saharan Africa (38, 8%) (i.e., *A. niloticus*, *M. natalensis* and *Atelerix albiventris*).

For pathogens not identified to species level (i.e. family or higher taxa only), we identified 148 host-pathogen pairs among 32 rodent species and 25 pathogen families (Supplementary Fig 4.), with CLOVER containing 66 (45%) of these associations.

Rodent trapping studies identified additional rodent host species for six pathogens; *Lassa mammarenavirus* (5), *Toxoplasma gondii* (4), Usutu virus (2), *Coxiella burnetii* (2), *Escherichia coli* and *Klebsiella pneumoniae* (both 1), that were not present in this consolidated host-pathogen association dataset.

What is the spatial extent of pathogen testing within a host's range?

The five most widely sampled pathogen species/families in included studies were Arenaviridae, Borreliaceae, *Lassa mammarenavirus*, Leptospiraceae and *Toxoplasma gondii* (Table 2.). Assays to identify Arenaviridae infection were performed in 44 rodent species with evidence of viral infection in 15 species. *Lassa mammarenavirus* was specifically tested for in 43 species with 10 showing evidence of viral infection. The most commonly infected species for both Arenaviridae, generally, and *Lassa mammarenavirus* specifically, were *M. natalensis* and *M. erythroleucus*. These species were assayed across between 10-20% of their trapped area, equating to ~0.02% of their IUCN range (Table 2.).

Infection with species of Borreliaceae was assessed in 42 species, with evidence of infection in 17 rodent species. The greatest rates of infection were among *A. niloticus* (16%), *Mastomys huberti* (11%) and *M. erythroleucus* (9%). Testing was more widespread than for Arenaviruses with coverage between 15-34% of their trapped area, however, this remains a small area in relation to their IUCN ranges (<0.05%). Leptospiraceae and *Toxoplasma gondii* was assessed in 8 species, with evidence of infection in 5 and 6 rodent species respectively. The spatial coverage of testing for these pathogens was more limited within IUCN host species ranges (~0.01%).

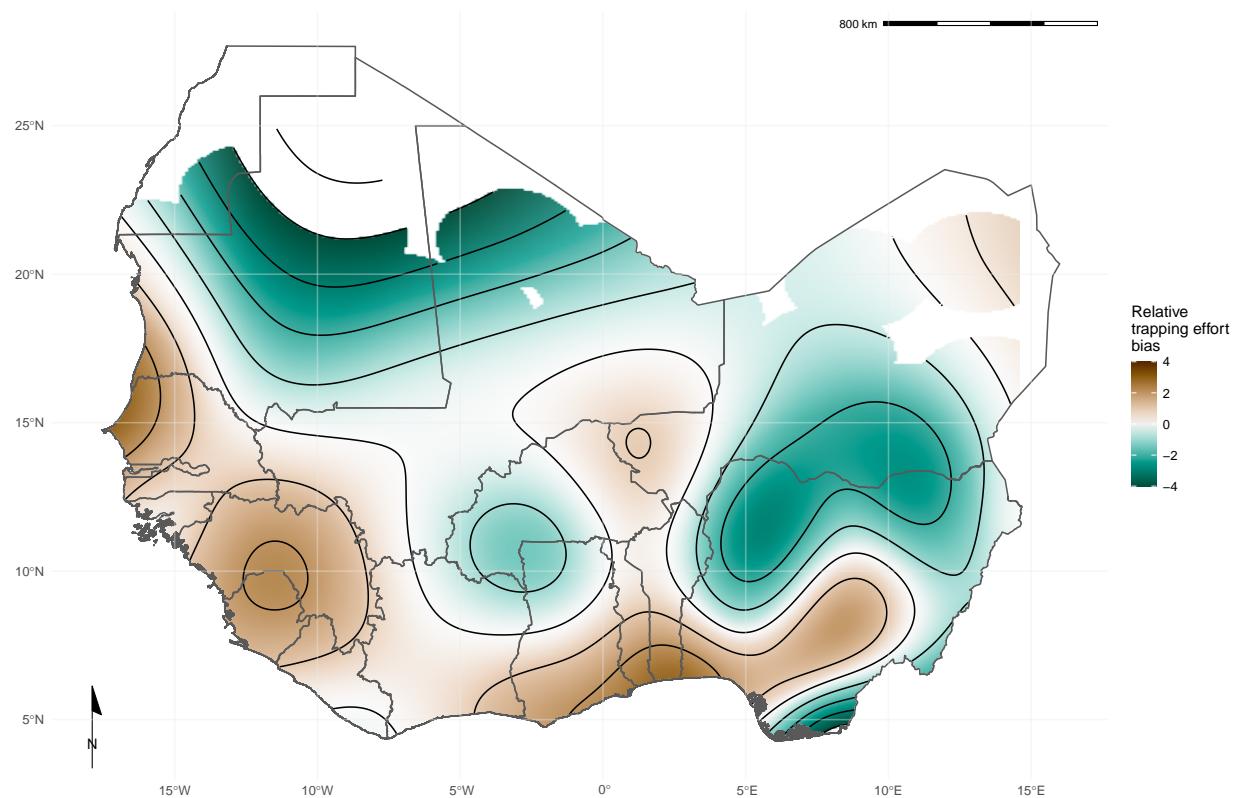
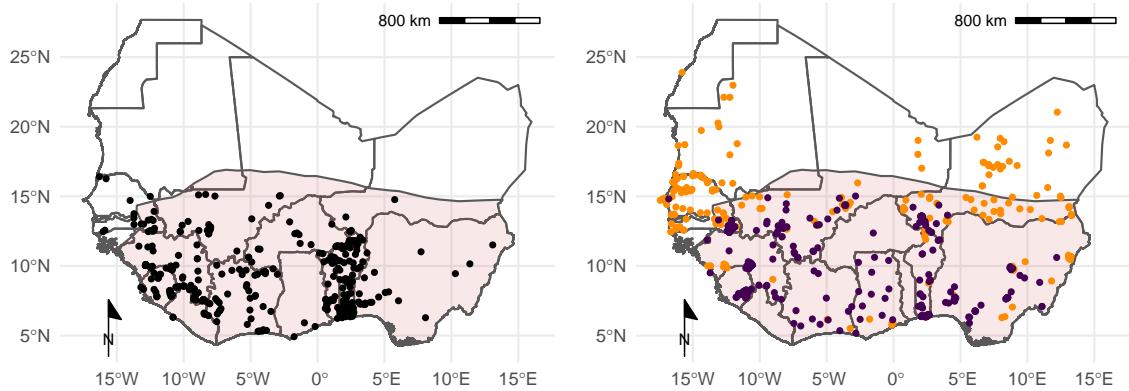
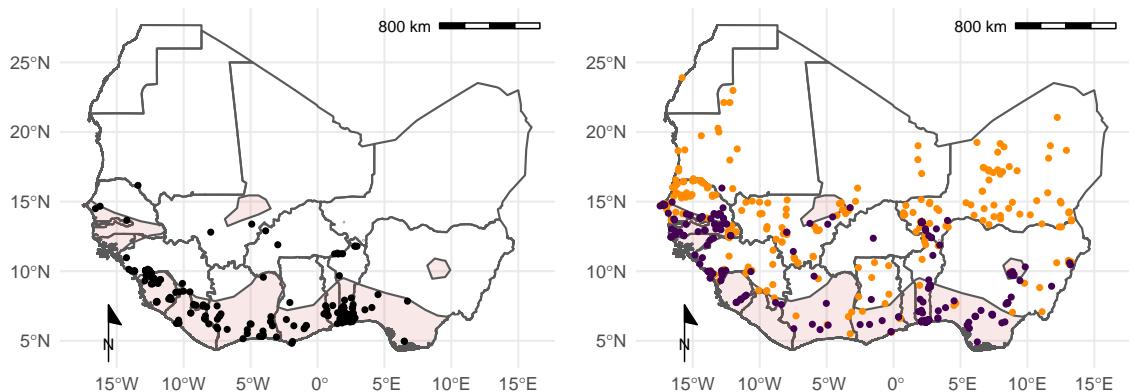


Figure 2: Fig 2. Relative trapping effort bias across West Africa from included studies adjusted for human population density, proportion urban land cover and area of the administrative region. Yellow regions represent areas with a bias towards increased trapping effort (e.g., North West Senegal), purple regions represent areas with a bias towards reduced trapping effort (e.g., Northern Nigeria).

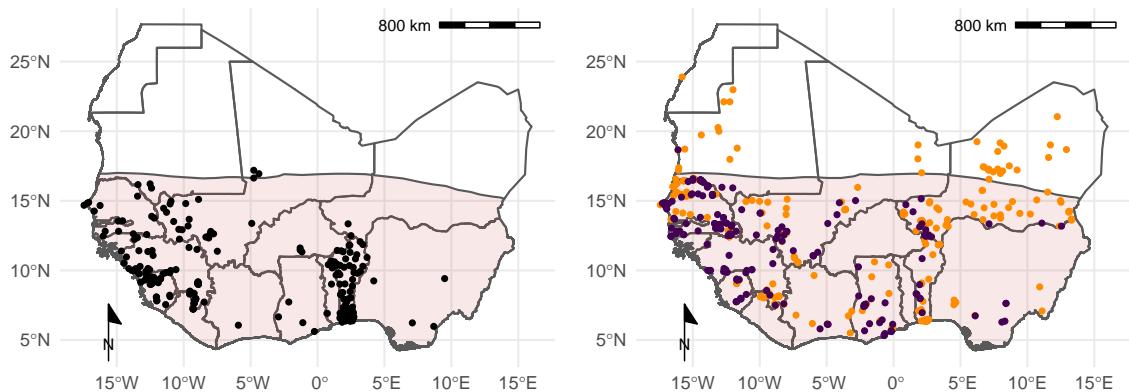
Mastomys natalensis



Rattus rattus



Mus musculus



Detection/Non-detection ● Detection ● Non-detection

Figure 3: Fig 3. Each row corresponds to a single rodent species. L) Presence recorded in GBIF overlaid on IUCN species range (red-shaded area). R) Detection and non-detection from rodent trapping studies overlaid on IUCN species ranges. M. musculus has no IUCN West African range.

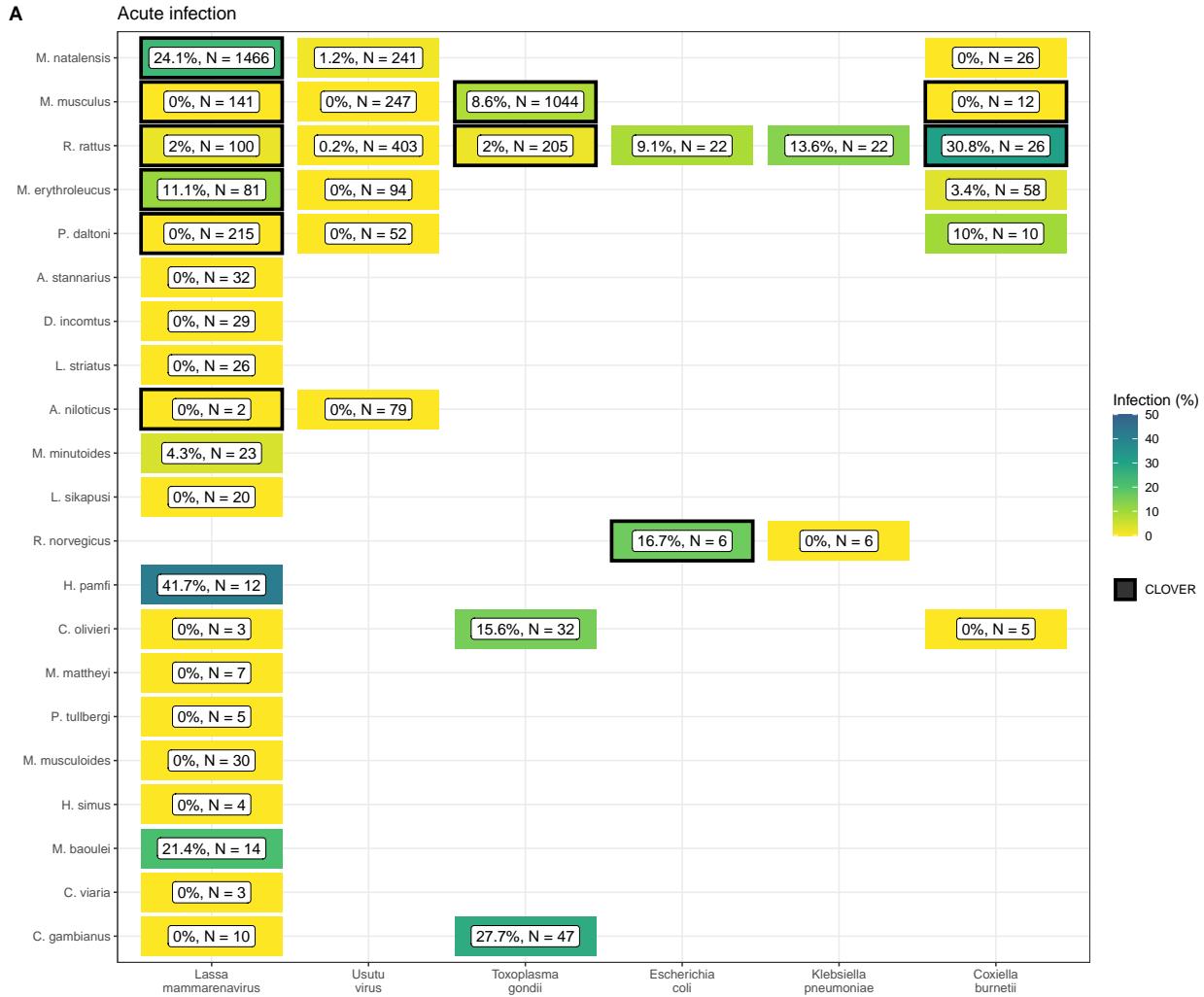


Figure 4: Fig 4. A) Identified species level host-pathogen associations through detection of acute infection (i.e. PCR, culture). Percentages and colour relate to the proportion of all assays that were positive, the number of individuals tested for the pathogen is labelled N. Associations with a black border are present in the CLOVER dataset.

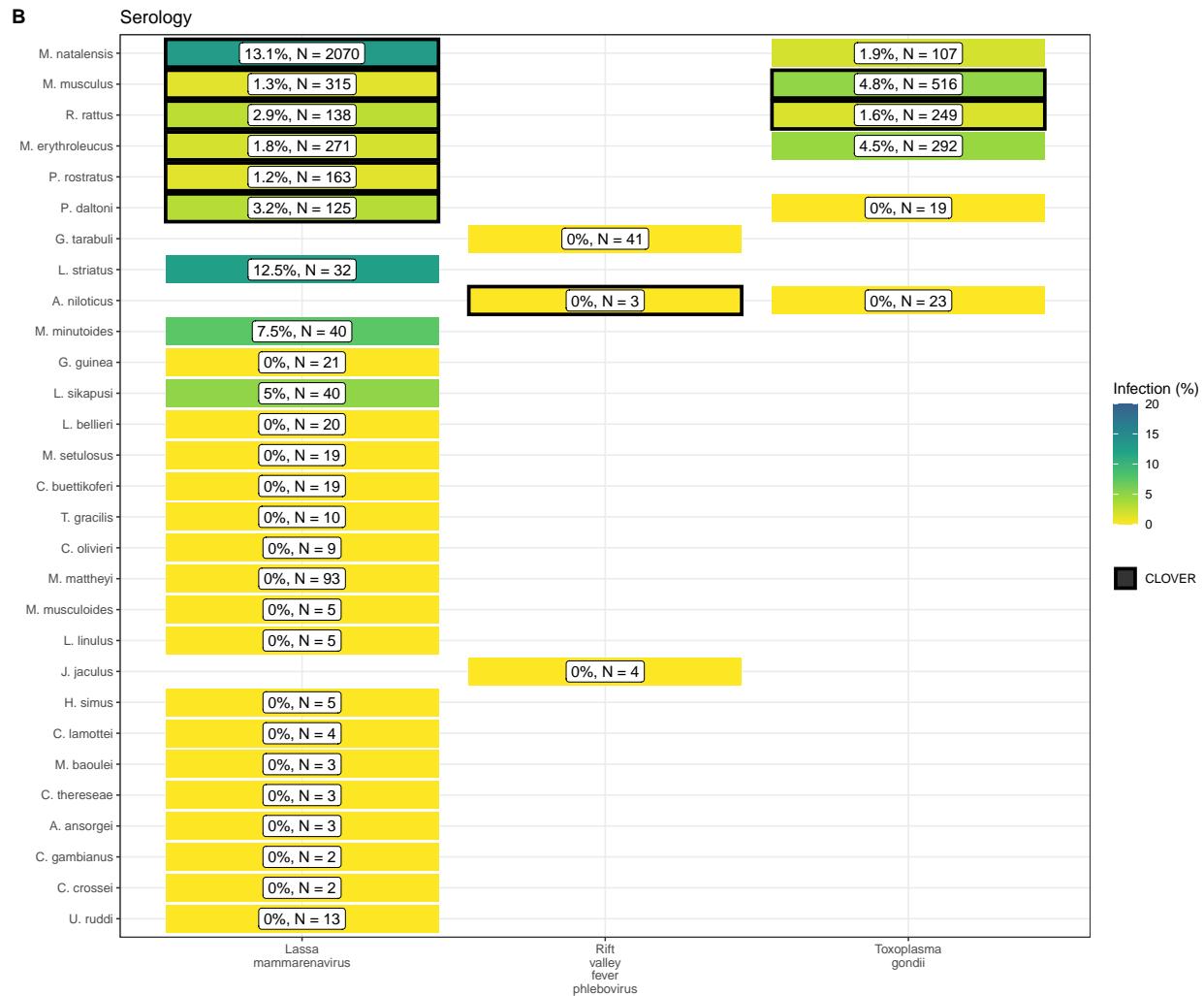


Figure 5: Fig 4. B) Identified species level host-pathogen associations through serological assays (i.e. ELISA). Percentages and colour relate to the proportion of all assays that were positive, the number of individuals tested for the pathogen is labelled N. Associations with a black border are present in the CLOVER dataset.

Table 2: Comparison of pathogen sampling ranges for 5 most widely sampled pathogens and the 5 most sampled rodent host species. * no IUCN range in West African

| Pathogen | Species | Tested | Positive | Pathogen testing area (1,000 km ²) |
|----------------------|-------------------------------|--------|-----------|--|
| Arenaviridae sp. | | | | |
| | <i>Mastomys natalensis</i> | 2,841 | 104 (4%) | 0.61 |
| | <i>Praomys daltoni</i> | 854 | 6 (1%) | 0.42 |
| | <i>Mastomys erythroleucus</i> | 398 | 20 (5%) | 0.40 |
| | <i>Rattus rattus</i> | 396 | 4 (1%) | 0.38 |
| | <i>Praomys rostratus</i> | 310 | 5 (2%) | 0.13 |
| Borrelia sp. | | | | |
| | <i>Mastomys erythroleucus</i> | 1,586 | 140 (9%) | 1.14 |
| | <i>Arvicanthis niloticus</i> | 1,551 | 253 (16%) | 0.66 |
| | <i>Mastomys natalensis</i> | 733 | 54 (7%) | 0.69 |
| | <i>Mastomys huberti</i> | 731 | 83 (11%) | 0.23 |
| | <i>Mus musculus</i> | 686 | 26 (4%) | 0.45 |
| Lassa mammarenavirus | | | | |
| | <i>Mastomys natalensis</i> | 3,199 | 580 (18%) | 1.03 |
| | <i>Mastomys erythroleucus</i> | 352 | 14 (4%) | 0.36 |
| | <i>Rattus rattus</i> | 177 | 2 (1%) | 0.34 |
| | <i>Praomys rostratus</i> | 163 | 2 (1%) | 0.27 |
| | <i>Mus musculus</i> | 147 | 0 (0%) | 0.04 |
| Leptospira sp. | | | | |
| | <i>Rattus rattus</i> | 646 | 65 (10%) | 0.40 |
| | <i>Arvicanthis niloticus</i> | 221 | 10 (5%) | 0.02 |
| | <i>Crocidura olivieri</i> | 141 | 14 (10%) | 0.34 |
| | <i>Mastomys natalensis</i> | 136 | 26 (19%) | 0.36 |
| | <i>Rattus norvegicus</i> | 79 | 19 (24%) | 0.21 |
| Toxoplasma gondii | | | | |
| | <i>Mus musculus</i> | 1,548 | 115 (7%) | 0.62 |
| | <i>Rattus rattus</i> | 428 | 8 (2%) | 0.36 |
| | <i>Mastomys erythroleucus</i> | 292 | 13 (4%) | 0.37 |
| | <i>Mastomys natalensis</i> | 107 | 2 (2%) | 0.08 |
| | <i>Cricetomys gambianus</i> | 47 | 13 (28%) | 0.06 |

Discussion

Endemic rodent zoonoses and novel pathogen emergence from rodent hosts are predicted to have an increasing burden in West Africa and globally (Han *et al.* (2015)). Here we have synthesised data from 126 rodent trapping studies containing information on more than 72,000 rodents across 1,611 trap sites producing an estimated 942,669 trap nights across 14 West African countries. Locations studied are complementary to curated datasets (e.g. IUCN, GBIF), incorporation of our synthesised dataset when assessing zoonosis risk based on host distributions could counteract some of the biases inherent to these curated datasets (Boakes *et al.*, 2010). We identified 25 host-pathogen pairs reported from included studies, 15 of these were not included in a consolidated host-pathogen dataset. Generally, the number of different species tested for a pathogen and the spatial extent of these sampling locations were limited. These findings highlight a further component of sampling bias, supporting calls for further pathogen sampling across diverse species in zoonotic hotspots (Harvey and Holmes, 2022).

We found that rodent trapping data, similar to biodiversity data, showed important spatial biases (Beck *et al.*, 2014). Relative trapping effort bias was greater in Benin, Guinea, Senegal and Sierra Leone driven by long-standing research collaborations investigating the invasion non-native rodent species (*M. musculus* and *R. rattus*) and the hazard of endemic zoonosis outbreaks (e.g., *Lassa mammarenavirus*). Much of the region remains relatively undersampled, particularly Burkina Faso, Côte d'Ivoire, Ghana and Nigeria, despite these countries facing many of the same challenges. For example, annual outbreaks of Lassa fever are reported in Nigeria and there are potentially 60,000 unrecognised cases of Lassa fever every year in Côte d'Ivoire and Ghana (Basinski *et al.*, 2021). Rodent sampling should be targeted towards currently under sampled regions to reduce this bias and improve our understanding of the risk from endemic and novel zoonoses.

Rodent trapping studies provide geographic and temporally contextualised data on both species detection and non-detection which are not available from curated datasets. Non-detection data can improve models of species distributions, unfortunately, high levels of missing data on trapping effort will continue to confound the allocations of non-detections as true absences (Václavík and Meentemeyer, 2009). Models of host species occurrence and abundance, improved by incorporating species absence, are important to assess the effect of land use and climate change on endemic zoonosis spillover to human populations and direct limited public health resources towards regions at greatest risk (Zeimes *et al.*, 2012; Judson *et al.*, 2018).

Currently available consolidated datasets on host-pathogen associations (e.g. CLOVER, EID2 and GMPD2) do not include spatial or temporal components (Gibb, Gregory F. Albery, *et al.*, 2021a). The current synthesis of rodent trapping studies has highlighted that pathogens have been sparsely sampled within a host's range. Current zoonosis risk models dependent on these sources of data are therefore not able to incorporate spatial heterogeneity in pathogen prevalence across the host range. Additional uncertainty in current models of zoonotic disease risk arises from host-pathogen associations that have not been reported in these consolidated datasets. For example *Hylomyscus pamfi* infected with *Lassa mammarenavirus* and *R. rattus* infected with *Coxiella burnetii*, will not be included when solely dependent consolidated host-pathogen datasets. Further, detection of zoonotic pathogens in multiple co-occurring host species supports the adoption of multi-species to better understand the potential range of endemic zoonoses (Wilkinson *et al.*, 2019).

Few studies stratified detection and non-detection of hosts or pathogen prevalence by time, therefore limiting inference of changes in host and pathogen dynamics. This limitation prevents calculation of incidence of infection and the abundance of infectious rodents which potentially varies by both time and space (Fichet-Calvet *et al.*, 2016). Understanding of temporal changes in viral burden and shedding for endemic zoonoses is required to accurately predict current and future risk of pathogen spillover.

Finally, due to data sparsity, we were unable to account for temporal change over the six decades of rodent trapping studies. Land use change and population density have changed dramatically over this period in West Africa (Herrmann *et al.*, 2020). We attempted to mitigate against this by using the median year of trapping to understand the spatial and land use biases in trapping activity. It is possible that land use and population density at trapping sites varied importantly between when rodent trapping was conducted and the conditions in 2005. Despite this limitation, the finding that trapping is biased towards high density, human dominated landscapes is unlikely to substantially change.

We have shown that synthesis of rodent trapping studies to supplement curated rodent distributions can counteract some of the inherent biases in these data and that they can add further contextual data to host-pathogen association data. Together this supports their inclusion in efforts to model endemic zoonotic risk and novel pathogen emergence. Contribution of rodent trapping studies as data sources can be improved by adopting reporting standards and practices consistent with Open Science, namely sharing of disaggregated datasets alongside publication (Foster and Deardorff, 2017).

Future rodent trapping studies should be targeted towards regions that are currently under-studied. Further information on rodent presence and abundance across West Africa will assist in modelling changing endemic zoonosis risk and the potential for novel pathogen emergence. Sharing of disaggregated data alongside research publications should be supported with adoption of data standards to support further data synthesis. Inclusion of exact locations of trapping sites, trapping effort and the dates at which trapping occurred would support more detailed understanding of the spatio-temporal dynamics of host populations and the risk of endemic zoonosis spillover events. Despite these challenges we propose that rodent trapping studies can provide an important source of data to supplement curated datasets on rodent distributions to quantify the risk of endemic zoonosis spillover events and the hazard of novel pathogen emergence.

References

- Albery, G. F. and Becker, D. J. (2021) ‘Fast-lived hosts and zoonotic risk’, *Trends in Parasitology*, 37(2), pp. 117–129. doi: 10.1016/j.pt.2020.10.012.
- Allen, T. et al. (2017) ‘Global hotspots and correlates of emerging zoonotic diseases’, *Nature Communications*, 8(1), p. 1124. doi: 10.1038/s41467-017-00923-8.
- Basinski, A. J. et al. (2021) ‘Bridging the gap: Using reservoir ecology and human serosurveys to estimate lassa virus spillover in west africa’, *PLOS Computational Biology*. Edited by A. Wesolowski, 17(3), p. e1008811. doi: 10.1371/journal.pcbi.1008811.
- Beck, J. et al. (2014) ‘Spatial bias in the GBIF database and its effect on modeling species’ geographic distributions’, *Ecological Informatics*, 19, pp. 10–15. doi: 10.1016/j.ecoinf.2013.11.002.
- Bernstein, A. S. et al. (2022) ‘The costs and benefits of primary prevention of zoonotic pandemics’, *Science Advances*, 8(5), p. eabl4183. doi: 10.1126/sciadv.abl4183.
- Boakes, E. H. et al. (2010) ‘Distorted views of biodiversity: Spatial and temporal bias in species occurrence data’, *PLOS Biology*, 8(6), p. e1000385. doi: 10.1371/journal.pbio.1000385.
- Bovendorp, R. S., McCleery, R. A. and Galetti, M. (2017) ‘Optimising sampling methods for small mammal communities in neotropical rainforests’, *Mammal Review*, 47(2), pp. 148–158. doi: 10.1111/mam.12088.
- Bowler, D. E. et al. (2022) ‘Temporal trends in the spatial bias of species occurrence records’, *Ecography*, n/a, p. e06219. doi: 10.1111/ecog.06219.
- Carlson, C. J. et al. (2021) ‘The future of zoonotic risk prediction’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1837), p. 20200358. doi: 10.1098/rstb.2020.0358.
- Catalano, S. et al. (2020) ‘Multihost transmission of schistosoma mansoni in senegal, 2015-2018’, *Emerging Infectious Diseases*, 26(6), pp. 1234–1242. doi: 10.3201/eid2606.200107.
- Database of Global Administrative Areas (2022) *GADM*. Available at: <https://gadm.org/index.html> (Accessed: 25 April 2021).
- Dobson, F. S. and Oli, M. K. (2007) ‘Fast and slow life histories of rodents’, *Rodent societies: an ecological and evolutionary perspective*, pp. 99–105. Available at: <https://www.jstor.org/stable/42902037>.
- European Space Agency Climate Change Initiative (2022) *Land cover classification gridded maps from 1992 to present derived from satellite observations*. Available at: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover?tab=overview> (Accessed: 3 March 2022).
- Fichet-Calvet, E. et al. (2009) ‘Diversity and dynamics in a community of small mammals in coastal guinea, west africa’, *Belgian Journal of Zoology*, 139(2), pp. 93–102. Available at: <http://www.vliz.be/nl/open-marien-archief?module=ref&refid=204614>.
- Fichet-Calvet, E. et al. (2016) ‘Spatial and temporal evolution of lassa virus in the natural host population in upper guinea’, *Scientific Reports*, 6, p. 21977. doi: 10.1038/srep21977.
- Fischer, C. et al. (2018) ‘Ecosystem services and disservices provided by small rodents in arable fields: Effects of local and landscape management’, *Journal of Applied Ecology*, 55(2), pp. 548–558. doi: 10.1111/1365-

- 2664.13016.
- Foster, E. D. and Deardorff, A. (2017) ‘Open science framework (OSF)’, *Journal of the Medical Library Association : JMLA*, 105(2), pp. 203–206. doi: 10.5195/jmla.2017.88.
- Galeh, T. M. et al. (2020) ‘Global status of toxoplasma gondii seroprevalence in rodents: A systematic review and meta-analysis’, *Frontiers in Veterinary Science*, 7. Available at: <https://www.frontiersin.org/article/10.3389/fvets.2020.00461> (Accessed: 2 March 2022).
- García-Peña, G. E. et al. (2021) ‘Land-use change and rodent-borne diseases: Hazards on the shared socioeconomic pathways’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1837), p. 20200362. doi: 10.1098/rstb.2020.0362.
- GBIF: The Global Biodiversity Information Facility (2021a) ‘Occurrence download’. The Global Biodiversity Information Facility. doi: 10.15468/DL.S52MEH.
- GBIF: The Global Biodiversity Information Facility (2021b) *The global biodiversity information facility*. Available at: <https://www.gbif.org/> (Accessed: 25 April 2021).
- Gibb, R., Franklinos, L. H. V., et al. (2020) ‘Ecosystem perspectives are needed to manage zoonotic risks in a changing climate’, *BMJ*, 371, p. m3389. doi: 10.1136/bmj.m3389.
- Gibb, R., Redding, D. W., et al. (2020) ‘Zoonotic host diversity increases in human-dominated ecosystems’, *Nature*, 584(7821), pp. 398–402. doi: 10.1038/s41586-020-2562-8.
- Gibb, R., Albery, Gregory F., et al. (2021a) ‘Data proliferation, reconciliation, and synthesis in viral ecology’, *BioScience*, 71(11), pp. 1148–1156. doi: 10.1093/biosci/biab080.
- Gibb, R., Albery, Gregory F., et al. (2021b) ‘Mammal virus diversity estimates are unstable due to accelerating discovery effort’, *Biology Letters*, 18(1), p. 20210427. doi: 10.1098/rsbl.2021.0427.
- Gibb, R., Carlson, C. J. and Farrell, M. J. (2021) *Viralemegance/clover: Preprint + zenodo*. Zenodo. doi: 10.5281/zenodo.4435128.
- Grace, D. et al. (2012) ‘Mapping of poverty and likely zoonoses hotspots’.
- Halliday, J. E. B. et al. (2015) ‘Endemic zoonoses in the tropics: A public health problem hiding in plain sight’, *The Veterinary Record*, 176(9), pp. 220–225. doi: 10.1136/vr.h798.
- Han, B. A. et al. (2015) ‘Rodent reservoirs of future zoonotic diseases’, *Proceedings of the National Academy of Sciences*, 112(22), pp. 7039–7044. doi: 10.1073/pnas.1501598112.
- Han, B. A., Kramer, A. M. and Drake, J. M. (2016) ‘Global patterns of zoonotic disease in mammals’, *Trends in Parasitology*, 32(7), pp. 565–577. doi: 10.1016/j.pt.2016.04.007.
- Harvey, E. and Holmes, E. C. (2022) ‘Diversity and evolution of the animal virome’, *Nature Reviews Microbiology*, 20(6), pp. 321–334. doi: 10.1038/s41579-021-00665-x.
- Hassell, J. M. et al. (2017) ‘Urbanization and disease emergence: Dynamics at the wildlife–livestock–human interface’, *Trends in Ecology & Evolution*, 32(1), pp. 55–67. doi: 10.1016/j.tree.2016.09.012.
- Herrmann, S. M. et al. (2020) ‘Accelerating land cover change in west africa over four decades as population pressure increased’, *Communications Earth & Environment*, 1(1), pp. 1–10. doi: 10.1038/s43247-020-00053-y.
- IUCN (2021) *The IUCN red list of threatened species*. Available at: www.iucnredlist.org.
- Judson, S. D. et al. (2018) ‘Translating predictions of zoonotic viruses for policymakers’, *EcoHealth*, 15(1), pp. 52–62. doi: 10.1007/s10393-017-1304-3.
- Judson, S. D. and Rabinowitz, P. M. (2021) ‘Zoonoses and global epidemics’, *Current Opinion in Infectious Diseases*, 34(5), pp. 385–392. doi: 10.1097/QCO.0000000000000749.
- Maudlin, I., Eisler, M. C. and Welburn, S. C. (2009) ‘Neglected and endemic zoonoses’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1530), pp. 2777–2787. doi: 10.1098/rstb.2009.0067.
- McMahon, B. J., Morand, S. and Gray, J. S. (2018) ‘Ecosystem change and zoonoses in the anthropocene’, *Zoonoses and Public Health*, 65(7), pp. 755–765. doi: 10.1111/zph.12489.
- Meerburg, B., Singleton, G. and Kijlstra, A. (2009) ‘Rodent-borne diseases and their risks for public health’, *Critical reviews in microbiology*, 35, pp. 221–70. doi: 10.1080/10408410902989837.
- Molyneux, D. et al. (2011) ‘Zoonoses and marginalised infectious diseases of poverty: Where do we stand?’, *Parasites & Vectors*, 4(1), p. 106. doi: 10.1186/1756-3305-4-106.
- National Geospatial-Intelligence Agency (no date) *NGA: GNS home*. Available at: <https://geonames.nga.mil/gns/html/> (Accessed: 22 March 2021).
- Pebesma, E. (2018) ‘Simple features for r: Standardized support for spatial vector data’, *The R Journal*, 10(1), pp. 439–446. Available at: <https://journal.r-project.org/archive/2018/RJ-2018-009/index.html> (Accessed: 24 May 2022).

- Pedersen, E. J. *et al.* (2019) ‘Hierarchical generalized additive models in ecology: An introduction with mgcv’, *PeerJ*, 7, p. e6876. doi: 10.7717/peerj.6876.
- Plowright, R. K. *et al.* (2019) ‘Sampling to elucidate the dynamics of infections in reservoir hosts’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1782), p. 20180336. doi: 10.1098/rstb.2018.0336.
- R Core Team (2020) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Redding, D. W. *et al.* (2017) ‘Evaluating bayesian spatial methods for modelling species distributions with clumped and restricted occurrence data’, *PLOS ONE*, 12(11), p. e0187602. doi: 10.1371/journal.pone.0187602.
- Socioeconomic Data and Applications Center (2021) *Gridded population of the world (GPW), v4 / SEDAC*. Available at: <https://sedac.ciesin.columbia.edu/data/collection/gpw-v4> (Accessed: 4 February 2021).
- USAID (2021) *One health surveillance. PREDICT project*. Available at: <https://p2.predict.global/surveillance> (Accessed: 27 September 2021).
- Václavík, T. and Meentemeyer, R. K. (2009) ‘Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions?’, *Ecological Modelling*, 220(23), pp. 3248–3258. doi: 10.1016/j.ecolmodel.2009.08.013.
- Wilkinson, D. P. *et al.* (2019) ‘A comparison of joint species distribution models for presence-absence data’, *Methods in Ecology and Evolution*, 10(2), pp. 198–211. doi: 10.1111/2041-210X.13106.
- Wille, M., Geoghegan, J. L. and Holmes, E. C. (2021) ‘How accurately can we assess zoonotic risk?’, *PLOS Biology*. Edited by A. P. Dobson, 19(4), p. e3001135. doi: 10.1371/journal.pbio.3001135.
- Wood, S. N. (2017) *Generalized additive models: An introduction with r*. 2nd edn. Chapman; Hall/CRC.
- Zeimes, C. B. *et al.* (2012) ‘Modelling zoonotic diseases in humans: Comparison of methods for hantavirus in sweden’, *International Journal of Health Geographics*, 11(1), p. 39. doi: 10.1186/1476-072X-11-39.

Supplementary

Supplementary Table 1

Table 3: Supplementary Table 1: Data extraction tool for studies meeting inclusion criteria

| Extraction tool | Variable | Description |
|-----------------|------------------|--|
| Study data | | |
| | link | link to manuscript |
| | year_publication | year of publication |
| | title | title of manuscript |
| | journal_name | journal |
| | aim_1 | stated aim of study |
| | aim_2 | stated aim of study |
| | aim_3 | stated aim of study |
| | first_author | first author of the study |
| | reference_uid | DOI/ISSN/ISBN of the publication |
| | unique_id | unique ID for the current study |
| | metric | measurement of species presence abundance/presence |
| | trap_types | type of rodent traps used |
| | trapping_method | construction of the sampling grid |

| Extraction tool | Variable | Description |
|-----------------|-----------------------|--|
| | repeated_visit | whether there were multiple study visits to the same sites |
| | geolocation_level | the level of geolocation reported |
| | speciation | the level of speciation of trapped rodents |
| | aim | aim of the study dichotomised to Ecology or Zoonotic risk |
| | aim_detail | categorisation of study aims |
| | species_accumulation | whether a species accumulation curve to describe trapping effort |
| | diversity_measurement | whether there is a measure of rodent species diversity reported |
| | trapping_effort | whether pathogens are assayed |
| | pathogen | completeness of reported trapping effort |
| Rodent data | | |
| | unique_id | unique ID for the current study |
| | year_trapping | year rodent trapping occurred (range) |
| | month_trapping | months trapping occurred (range) |
| | country | country trapping occurred within |
| | region | region trapping occurred within |
| | town_village | name of towns or villages trapping occurred within |
| | latitude_DMS_N | latitude of trapping site in degrees minutes seconds (North) |
| | longitude_DMS_W | longitude of trapping site in degrees minutes seconds (West) |
| | latitude_D_N | latitude of trapping site in decimal degrees (North) |
| | longitude_D_E | longitude of trapping site in decimal degrees (East) |
| | UTM_coordinates | location of trapping site in UTM coordinates |
| | habitat | habitat type of trapping site |
| | intensity_use | the intensity of human disturbance in the trapping site |
| | genus | reported genus of trapped rodent/small mammal species |
| | species | reported species of trapped rodent/small mammal species |
| | number | number of trapped individuals |
| | trap_nights | number of trap nights reported |
| | capture_rate | rate of capture if reported |
| | trap_night_unit | the unit of trap night measurement |
| | study_nights | the number of study nights completed at the trap site |
| Pathogen data | | |
| | unique_id | unique ID for the current study |
| | year_trapping | year rodent trapping occurred (range) |
| | month | months trapping occurred (range) |
| | country | country trapping occurred within |

| Extraction tool | Variable | Description |
|-----------------|--------------------|--|
| | region | region trapping occurred within |
| | town_village | name of towns or villages trapping occurred within |
| | habitat | habitat type of trapping site |
| | genus | reported genus of trapped rodent/small mammal species |
| | species | reported species of trapped rodent/small mammal species |
| | pathogen_x | pathogens tested for, 1-7 possible columns |
| | latitude_DMS_N | latitude of trapping site in degrees minutes seconds (North) |
| | longitude_DMS_W | longitude of trapping site in degrees minutes seconds (West) |
| | latitude_D_N | latitude of trapping site in decimal degrees (North) |
| | longitude_D_E | longitude of trapping site in decimal degrees (East) |
| | UTM_coordinates | location of trapping site in UTM coordinates |
| | path_x_tested | number of individuals assayed for the corresponding pathogen, 1- |
| | pcr_x_positive | number of individuals PCR positive for the corresponding patho |
| | ab_ag_x_positive | number of individuals with positive serological assays for the cor |
| | culture_x_positive | number of individuals culture positive for the corresponding path |
| | histo_x_positive | number of individuals histologically/histopathologically positive |

Supplementary Table 2

| Year publication | Author | Title |
|------------------|------------------|--|
| 1974 | D. C. D. Happold | The small rodents of the forest-savanna-farmland association no |
| 1974 | Thomas Monath | Lassa virus isolation from <i>Mastomys natalensis</i> rodents during |
| 1975 | Herta Wulff | Recent isolations of Lassa virus from Nigerian rodents |
| 1975 | Malcolm Coe | Mammalian ecological studies on Mount Nimba, Liberia. |
| 1977 | D. C. D. Happold | A population study on small rodents in the tropical rain forest |
| 1977 | Sonia Jeffrey | Rodent Ecology and Land Use in Western Ghana |
| 1979 | J P Dedet | Isolation of <i>Leishmania major</i> from <i>Mastomys erythroleucus</i> an |
| 1982 | A Diallo | Bacteriological survey of leptospirosis in Zaria, Nigeria. |
| 1983 | C Robbins | <i>Mastomys</i> (rodentia: muridae) species distinguished by hemogl |
| 1987 | Joseph McCormick | A prospective study of the epidemiology and ecology of Lassa f |
| 1988 | J. Iyawe | Distribution of small rodents and shrews in a lowland rain fore |
| 1991 | J Trape | Tick-borne Borreliosis in West Africa |
| 1992 | Brian Mahy | Maintenance support of a field station in Sierra Leone, West A |

| Year publication | Author | Title |
|------------------|-------------------------|---|
| 1993 | Laurent Granjon | Social structure in synanthropic populations of a murid rodent |
| 1994 | Bruno Godeluck | A longitudinal survey of <i>Borrelia crocidurae</i> prevalence in rodent |
| 1994 | G Diatta | A comparative study of three methods of detection of <i>Borrelia</i> |
| 1995 | E Ikeh | <i>Mastomys natalensis</i> and <i>Tatera gambiana</i> as probable reservoir |
| 1997 | C Mafiana | Gastrointestinal helminth parasites of the black rat (<i>Rattus rattus</i>) |
| 1999 | Holger Meinig | Notes on the mammal fauna of the southern part of the Republic |
| 1999 | Jan Decher | Diversity and structure of terrestrial small mammal communities |
| 2000 | Adrian Barnett | Ecology of rodent communities in agricultural habitats in eastern |
| 2000 | J Duplantier | Rodents as reservoir hosts in the transmission of <i>Schistosoma</i> n |
| 2000 | James Ryan | Mammal fauna of the Muni-Pomadze Ramsar site, Ghana. |
| 2001 | Austin Demby | Lassa Fever in Guinea: {II}. Distribution and Prevalence of La |
| 2001 | Khalilou Ba | Preliminary study on some rodents of southern Mauritania as re |
| 2002 | Gauthier Dobigny | A cytotaxonomic survey of Rodents from Niger: implications for |
| 2002 | Laurent Granjon | The small mammal community of a coastal site of south-west Ma |
| 2003 | D Attuquayefio | A study of bushfires in a Ghanaian coastal wetland. Impact on |
| 2003 | Laurent Granjon | The importance of cytotaxonomy in understanding the biogeog |
| 2004 | B. Sicard | Effects of climate and local aridity on the latitudinal and habit |
| 2004 | Jan Decher | A rapid survey of terrestrial small mammals (shrews and rodent |
| 2004 | Sara Churchfield | First results on the feeding ecology of sympatric shrews (Insect |
| 2005 | Ara Monadjem | Preliminary biodiversity assessment (herpetofauna and mamma |
| 2005 | D Attuquayefio | Patterns of specific diversity and population size in small mamm |
| 2005 | Francesco Angelici | Rapid assessment of Small Mammals at Draw River, {BoiTano} |
| 2005 | Jan Decher | Population dynamics of the multimammate rat <i>Mastomys huber</i> |
| 2005 | Laurent Granjon | A Rapid Survey of Small Mammals from the Atewa Range For |
| 2006 | Natalie Weber | A rapid biological assessment of three classified forests in south |
| 2006 | Patrick Barriere | Lassa Virus-Infected Rodents in Refugee Camps in Guinea: A I |
| 2006 | Ryan Norris | Diversity and distribution of murid rodent populations between |
| 2007 | Joseph Fair | Biodiversity assessment (rodents and avifauna) of five forest res |
| 2008 | Ayodeji Olayemi | Cutaneous leishmaniasis in the Volta district of Ghana: An und |
| 2008 | D Attuquayefio | Annual flooding, survival and recruitment in a rodent populatio |
| 2008 | G Racznik | New data on the taxonomy and distribution of Rodentia (Mam |
| 2008 | Laurent Crespin | Diversity and dynamics in a community of small mammals in c |
| 2009 | Christiane Denys | Studies on the seasonal variations and prevalence of helminth fa |
| 2009 | Elisabeth Fichet-Calvet | |
| 2009 | Ivoke Njoku | |

| Year publication | Author | Title |
|------------------|-------------------------|--|
| 2010 | Adam Konecny | Indications of higher diversity and abundance of small rodents |
| 2010 | David Safronetz | Detection of Lassa Virus, Mali |
| 2010 | Edward Omudu | A survey of rats trapped in residential apartments and their ecology |
| 2010 | Elisabeth Fichet-Calvet | Diversity, dynamics and reproduction in a community of small mammals |
| 2010 | Gauthier Dobigny | Molecular survey of rodent-borne Trypanosoma in Niger with sequence analysis |
| 2010 | Jan Decher | Small mammal survey in the upper Seli River valley, Sierra Leone |
| 2010 | Lies Durnez | Terrestrial Small Mammals as Reservoirs of Mycobacterium ulcerans |
| 2010 | Mary Reynolds | A Silent Enzootic of an Orthopoxvirus in Ghana, West Africa: A seroepidemiological study |
| 2010 | R Sall-Drame | Variation in cestode assemblages of Mastomys and Arvicantis |
| 2010 | Violaine Nicolas | Terrestrial small mammal diversity and abundance in central Burkina Faso |
| 2011 | David Coulibaly-N'golo | Novel Arenavirus Sequences in Hylomyscus sp. and Mus (Nannospalax) leucodon |
| 2011 | Karmidine Hima | Extensive Robertsonian polymorphism in the African rodent Gerbillus |
| 2011 | Laurent Granjon | Guinean biodiversity at the edge: Rodents in forest patches of savanna |
| 2011 | M. Thiam | Capacity for water conservation in invasive (Gerbillus nigeriae) |
| 2012 | Amawulu Ebenezer | Effects of Urbanization and Agricultural Expansion on the Uptake of Lead |
| 2012 | Christiane Denys | On a new species of Dendromus (Rodentia, Nesomyidae) from Ivory Coast |
| 2012 | Khalilou Ba | Ecology of a typical West African Sudanian savannah rodent community |
| 2012 | Laurent Crespin | Demographic aspects of the island syndrome in two Afrotropical rodent species |
| 2012 | Tom Schwan | Endemic Foci of the Tick-Borne Relapsing Fever Spirochete Borrelia |
| 2013 | Adam Konecny | Invasion genetics of the introduced black rat (<i>Rattus rattus</i>) in Europe |
| 2013 | Blaise Kadjo | Assessment of terrestrial small mammals and a record of the crabs |
| 2013 | Gualbert Houemou | Leptospira spp. Prevalence in Small Mammal Populations in Côte d'Ivoire |
| 2013 | Jean-Francois Trape | The epidemiology and geographic distribution of relapsing fever |
| 2013 | Joshua Kamani | Prevalence and diversity of Bartonella species in commensal rodents |
| 2013 | Karl Kronmann | Two Novel Arenaviruses Detected in Pygmy Mice, Ghana |
| 2013 | Reuben Garshong | Effect of Habitat Change through Infrastructural Development on Small Mammals |
| 2014 | Benjamin Ofori | Preliminary checklist and aspects of the ecology of small mammals |
| 2014 | Elisabeth Fichet-Calvet | Lassa Serology in Natural Populations of Rodents and Horizontally Transmitted Viruses |
| 2014 | Madougou Garba | Spatial Segregation between Invasive and Native Commensal Rodents |
| 2015 | Charles Narh | Source Tracking Mycobacterium ulcerans Infections in the Ashanti Region of Ghana |
| 2015 | Christelle Dassi | Detection of Mycobacterium ulcerans in Mastomys natalensis and other small mammals |
| 2015 | Georges Diatta | Borrelia infection in small mammals in West Africa and its relation to human disease |
| 2015 | Pilar Foronda | Serological survey of antibodies to Toxoplasma gondii and Coxiella burnetii |
| 2015 | R Mol | Small terrestrial mammal and amphibian survey BoaCô region, Brazil |
| 2015 | Thomasz Leski | Sequence variability and geographic distribution of Lassa virus, West Africa |

| Year publication | Author | Title |
|------------------|---------------------|--|
| 2016 | Ayodeji Olayemi | Arenavirus Diversity and Phylogeography of Mastomys natalensis |
| 2016 | Ayodeji Olayemi | New Hosts of The Lassa Virus |
| 2016 | Benjamin Ofori | Spatio-temporal variation in small mammal species richness, re |
| 2017 | Alexis Ribas | Whipworm diversity in West African rodents: a molecular appr |
| 2017 | C Lippens | Genetic structure and invasion history of the house mouse (<i>Mu</i> |
| 2017 | Christophe Diagne | Ecological and sanitary impacts of bacterial communities associ |
| 2017 | Christophe Diagne | Serological Survey of Zoonotic Viruses in Invasive and Native C |
| 2017 | Daniel Attuquayefio | Impact of mining and forest regeneration on small mammal bio |
| 2018 | Ayodeji Olayemi | Widespread arenavirus occurrence and seroprevalence in small m |
| 2018 | Benjamin Ofori | Urban green area provides refuge for native small mammal biodi |
| 2018 | Carine Brouat | Seroprevalence of <i>Toxoplasma gondii</i> in commensal rodents sam |
| 2018 | Isaac Clement | Endoparasites of Small Mammals in Edo State, Nigeria: Public |
| 2018 | Joshua Kamani | Prevalence of Hepatozoon and <i>Sarcocystis</i> spp. in rodents and t |
| 2018 | Katharina Schaufler | Clinically Relevant {ESBL}-Producing <i>K. pneumoniae</i> {ST}30 |
| 2018 | Kouame Akpatou | Terrestrial small mammal diversity and abundance in Taï National |
| 2018 | Lokman Galal | Diversity of <i>Toxoplasma gondii</i> strains shaped by commensal co |
| 2018 | Marien Joachim | Movement Patterns of Small Rodents in Lassa Fever-Endemic V |
| 2018 | Stefano Catalano | Rodents of Senegal and their role as intermediate hosts of Hydat |
| 2019 | Agnes Yadouleton | Lassa Virus in Pygmy Mice, Benin, 2016–2017 |
| 2019 | Gualbert Houemou | Pathogenic Leptospira in Commensal Small Mammals from the |
| 2019 | Joachim Marien | Evaluation of rodent control to fight Lassa fever based on field |
| 2019 | Karmidine Hima | Native and Invasive Small Mammals in Urban Habitats along th |
| 2019 | Karmidine Hima | Population Dynamics and Genetics of <i>Gerbillus nigeriae</i> in Cen |
| 2019 | Kouame Akpatou | Assessment of Terrestrial Small Mammals in an Agro-industrial |
| 2019 | L Karan | Lassa Virus in the Host Rodent <i>Mastomys Natalensis</i> within U |
| 2019 | Moussa Diagne | Usutu Virus Isolated from Rodents in Senegal |
| 2019 | Natalie Weber | New records of bats and terrestrial small mammals from the Se |
| 2019 | Safianu Rabiu | Demographic response of the Gambian Gerbil to seasonal chang |
| 2019 | Stefano Catalano | <i>Plagiorchis</i> sp. in small mammals of Senegal and the potential o |
| 2020 | Adama Diarra | Molecular Detection of Microorganisms Associated with Small |
| 2020 | Adama Zida | <i>Mastomys natalensis</i> , <i>Cricetomys gambianus</i> and <i>Taterillus</i> sp. |
| 2020 | Christophe Diagne | Association between temporal patterns in helminth assemblages |
| 2020 | Claire Stragier | Interplay between historical and current features of the cityscape |
| 2020 | Handi Dahmana | Rodents as Hosts of Pathogens and Related Zoonotic Disease R |
| 2020 | Henri-Joel Dossou | Invasive rodents and damages to food stocks: a study in the Au |

| Year publication | Author | Title |
|------------------|------------------|---|
| 2020 | Joachim Marien | Households as hotspots of Lassa fever? Assessing the spatial distribution of Lassa virus circulation in Guinea |
| 2020 | Laurent Ahissa | Species composition and community structure of terrestrial small mammals in a West African savanna |
| 2020 | Leonce Kouadio | Detection of possible spillover of a novel hantavirus in a Natal bushpig |
| 2020 | Stefano Catalano | Multihost Transmission of Schistosoma mansoni in Senegal, 2016 |
| 2020 | Violaine Nicolas | Small mammal inventory in the Lama forest reserve (south Benin) |
| 2021 | Chibuisi Alimba | Wild black rats (<i>Rattus rattus</i> Linnaeus, 1758) as zoomonitor of tick-borne relapsing fever Borrellosis, a major public health problem in Malawi |
| 2021 | El Hadji Ndiaye | Tick-borne relapsing fever Borrellosis, a major public health problem in Malawi |
| 2021 | Mnqobi Mamba | Small mammals of a West African hotspot, the Ziama-Wonegizi Forest Reserve, Mali |
| 2021 | Umaru Bangura | Lassa Virus Circulation in Small Mammal Populations in Bo Duri, Sierra Leone |

Supplementary Table 3

Table 5: Supplementary Table 3.1: GAM model Trap night density
 $\sim \text{Tweedie}(\text{Longitude} * \text{Latitude})$

| Component | Term | Estimate | Std Error | t-value | p-value |
|----------------------------|-------------|----------|-----------|---------|---------|
| A. parametric coefficients | (Intercept) | -2.062 | 0.126 | -16.403 | *** |
| Component | Term | edf | Ref. df | F-value | p-value |
| B. smooth terms | s(x,y) | 30.431 | 39.000 | 7.290 | *** |

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05 < '}' < 0.1 < " < 1

Adjusted R-squared: 0.0261, Deviance explained 0.349

fREML : 4252.745, Scale est: 11.991, N: 1450

Table 6: Supplementary Table 3.2: GAM model Trap night density
 \sim Tweedie(Population density + (Longitdue * Latitude))

| Component | Term | Estimate | Std Error | t-value | p-value |
|----------------------------|-------------|-----------------|------------------|----------------|----------------|
| A. parametric coefficients | (Intercept) | -2.230 | 0.119 | -18.704 | *** |
| Component | Term | edf | Ref. df | F-value | p-value |
| B. smooth terms | s(pop_2005) | 2.818 | 11.000 | 7.231 | *** |
| | s(x,y) | 27.640 | 39.000 | 4.545 | *** |

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05 < .' < 0.1 < " < 1

Adjusted R-squared: 0.161, Deviance explained 0.422

fREML : 3924.187, Scale est: 9.836, N: 1450

Table 7: Supplementary Table 3.3: GAM model Trap night density
 \sim Tweedie(Population density + Region area + (Longitdue * Latitude))

| Component | Term | Estimate | Std Error | t-value | p-value |
|----------------------------|-------------|-----------------|------------------|----------------|----------------|
| A. parametric coefficients | (Intercept) | -2.371 | 0.120 | -19.699 | *** |
| Component | Term | edf | Ref. df | F-value | p-value |
| B. smooth terms | s(pop_2005) | 3.182 | 11.000 | 9.055 | *** |
| | s(area_km2) | 3.052 | 9.000 | 2.294 | *** |
| | s(x,y) | 26.819 | 39.000 | 4.843 | *** |

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05 < .' < 0.1 < " < 1

Adjusted R-squared: 0.148, Deviance explained 0.443

fREML : 3941.346, Scale est: 9.432, N: 1450

Table 8: Supplementary Table 3.4: GAM model Trap night density ~ Tweedie(Proportion cropland + Proportion shrubland + Proportion tree cover + Proportion urban + (Longitdue * Latitude))

| Component | Term | Estimate | Std Error | t-value | p-value |
|----------------------------|---------------|-----------------|------------------|----------------|----------------|
| A. parametric coefficients | (Intercept) | -1.775 | 0.123 | -14.379 | *** |
| Component | Term | edf | Ref. df | F-value | p-value |
| B. smooth terms | s(cropland) | 0.000 | 9.000 | 0.000 | |
| | s(shrubland) | 0.000 | 8.000 | 0.000 | |
| | s(tree_cover) | 1.579 | 9.000 | 0.795 | ** |
| | s(urban) | 4.772 | 9.000 | 19.703 | *** |
| | s(x,y) | 1.837 | 19.000 | 1.499 | *** |

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05 < '.' < 0.1 < " < 1

Adjusted R-squared: 0.0599, Deviance explained 0.306

fREML : 4198.018, Scale est: 12.895, N: 1450

Table 9: Supplementary Table 3.5: Final GAM model Trap night density ~ Tweedie(Population density + Region area + Proportion urban + (Longitdue * Latitude))

| Component | Term | Estimate | Std Error | t-value | p-value |
|----------------------------|-------------|-----------------|------------------|----------------|----------------|
| A. parametric coefficients | (Intercept) | -2.599 | 0.123 | -21.112 | *** |
| Component | Term | edf | Ref. df | F-value | p-value |
| B. smooth terms | s(pop_2005) | 7.131 | 11.000 | 6.398 | *** |
| | s(area_km2) | 3.629 | 9.000 | 3.450 | *** |
| | s(urban) | 1.921 | 9.000 | 1.235 | ** |
| | s(x,y) | 27.252 | 39.000 | 4.563 | *** |

Signif. codes: 0 <= '***' < 0.001 < '**' < 0.01 < '*' < 0.05 < '.' < 0.1 < " " < 1

Adjusted R-squared: 0.301, Deviance explained 0.487

fREML : 3907.563, Scale est: 8.408, N: 1450

Supplementary Fig 1

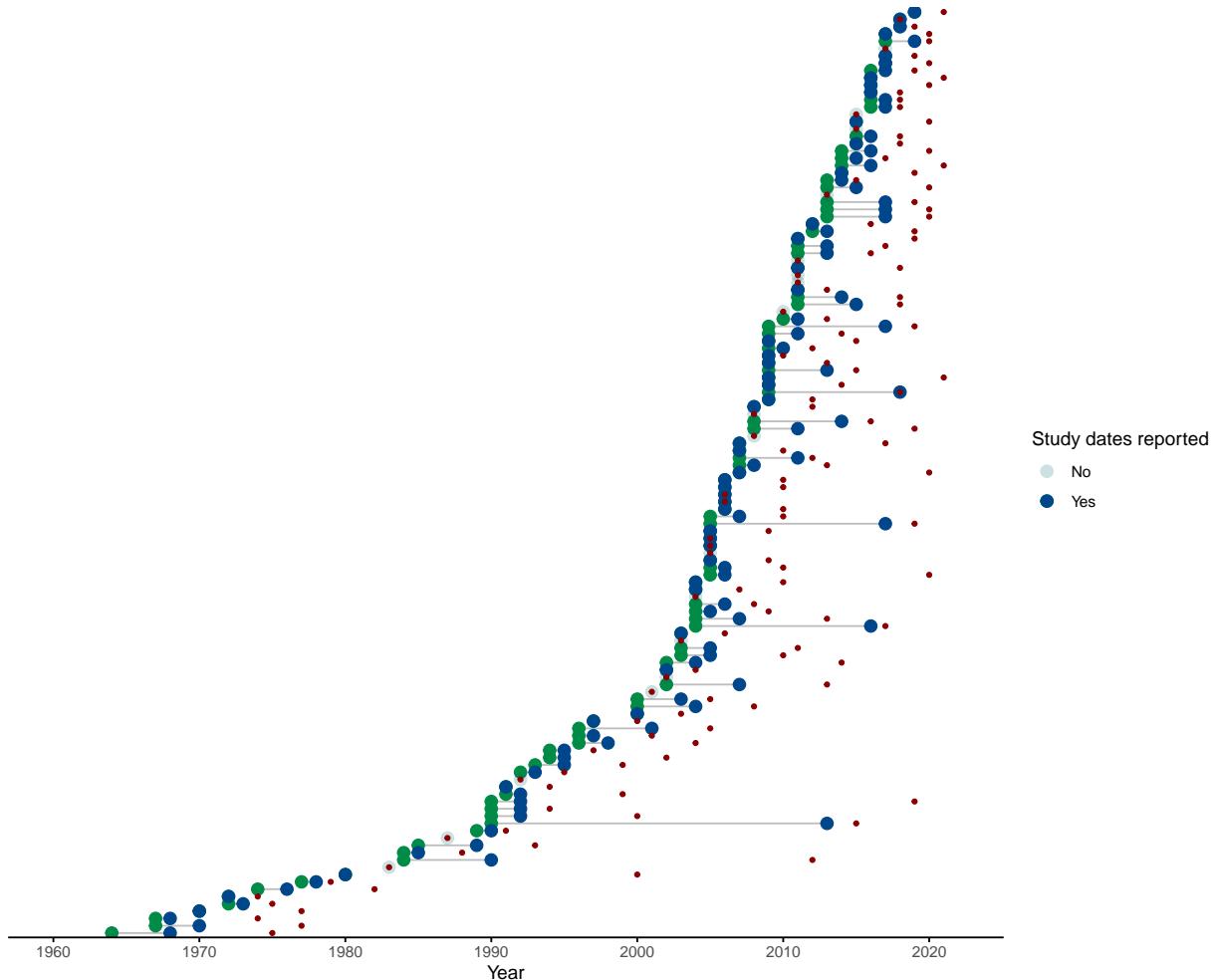


Figure 6: Supplementary Fig 1. Green points represent the start date of rodent trapping studies, blue points representing the final trapping activity. Red points indicate the publication of studies. Increasing numbers of studies have been published since 2000 with more studies being conducted over repeated visits.

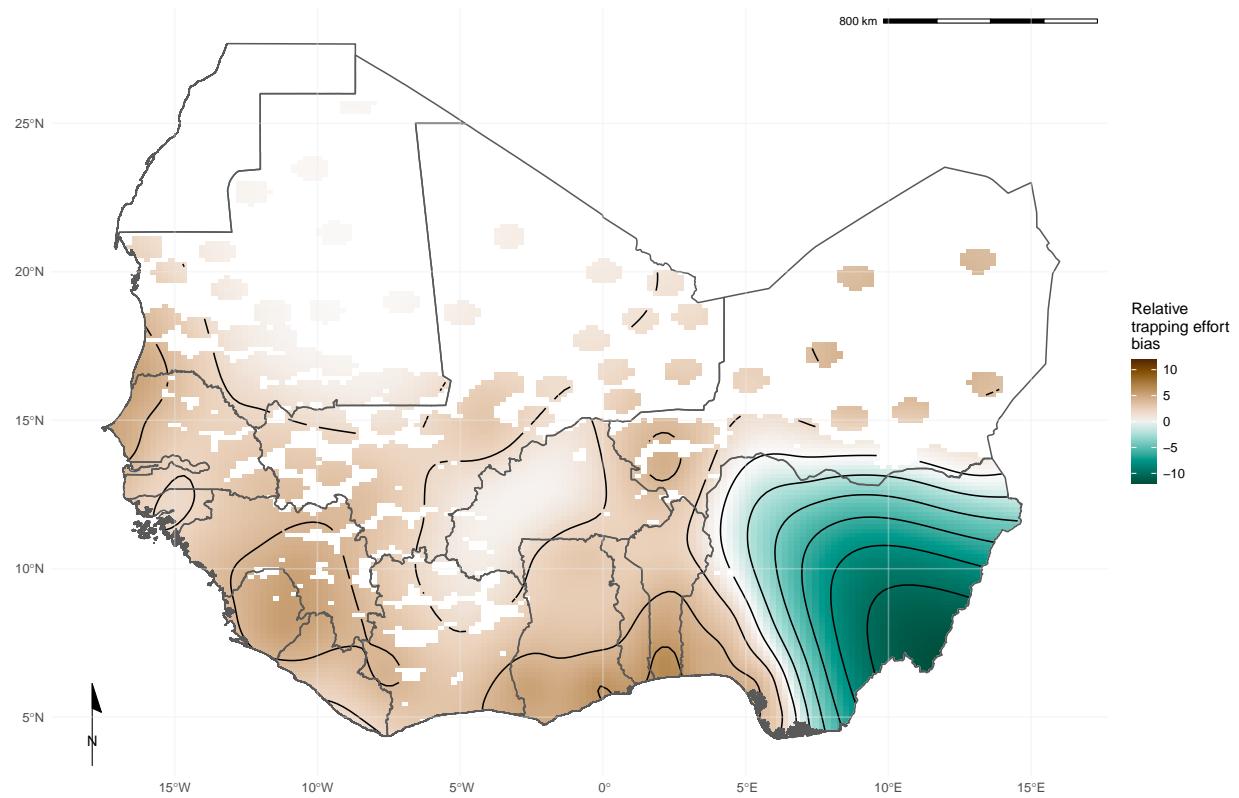


Figure 7: Supplementary Fig 2.

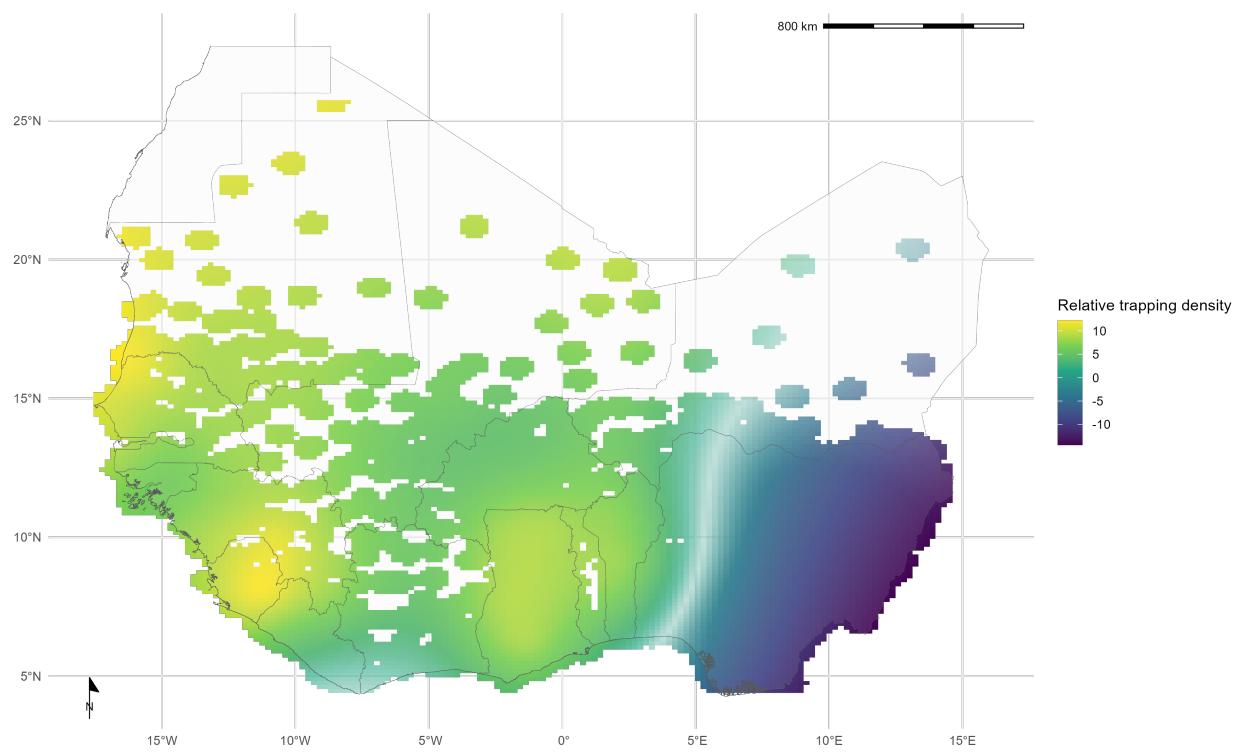
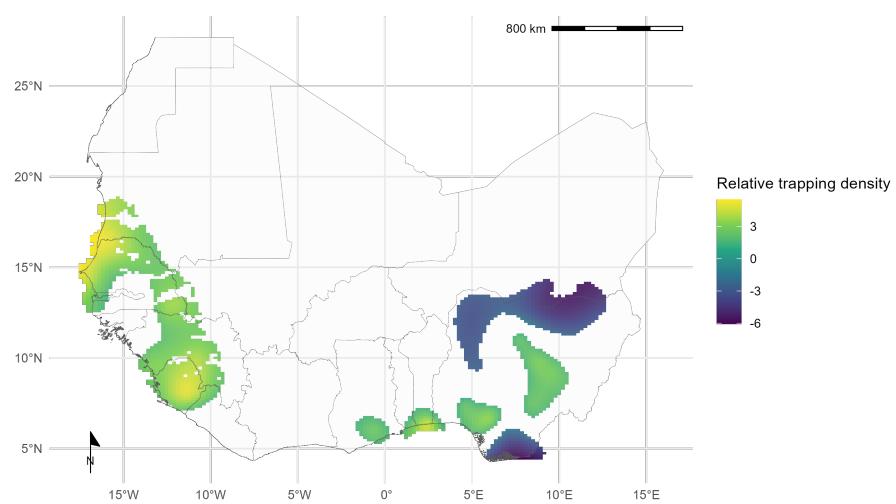


Figure 8: Supplementary Fig 3A. Relative trapping effort bias across West Africa from the subset of included studies reporting trapping effort adjusted for proportion urban land cover and proportion tree cover. Predictions are limited to areas around trap sites (coloured areas), uncertainty in the estimate 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

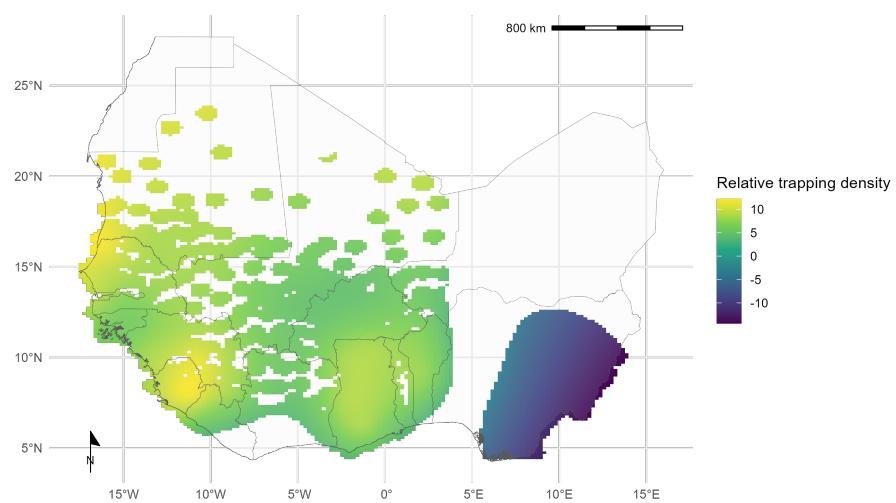
Supplementary Fig 2

Supplementary Fig 3

A

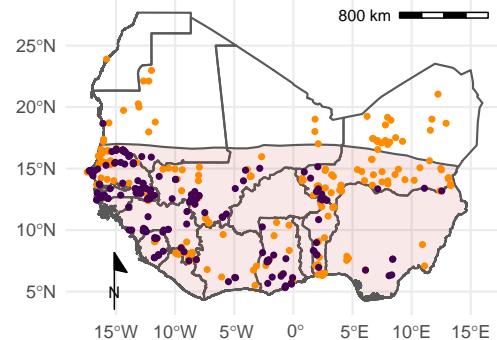
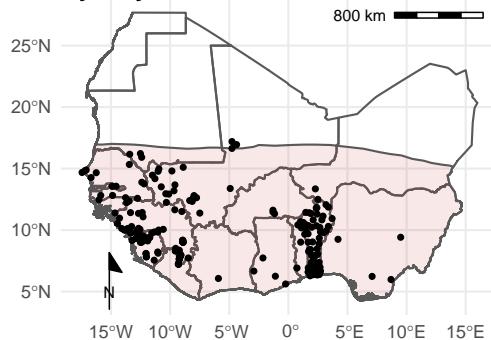


B

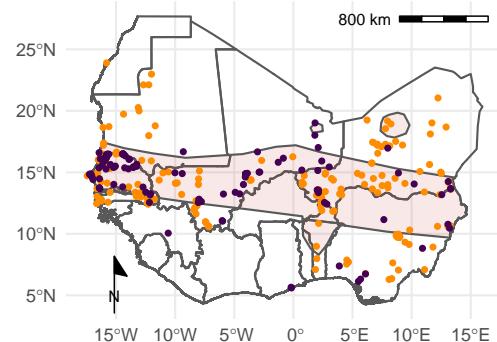
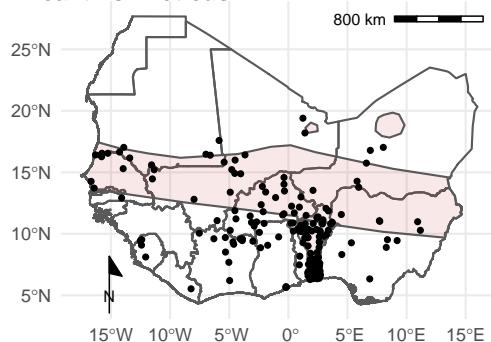


Supplementary Fig 4

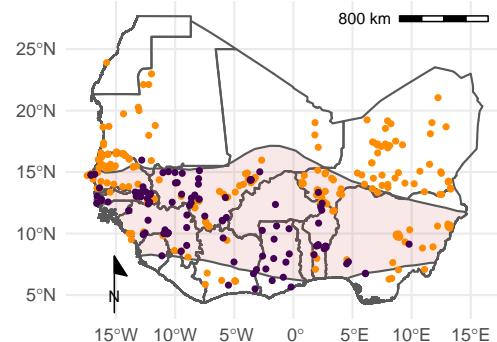
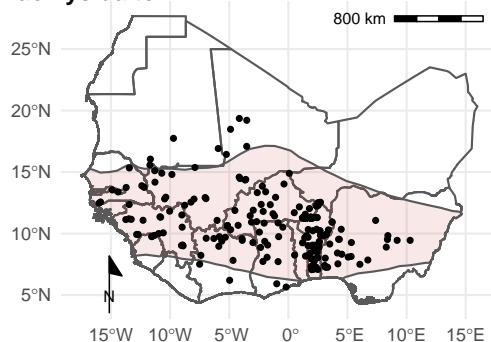
Mastomys erythroleucus



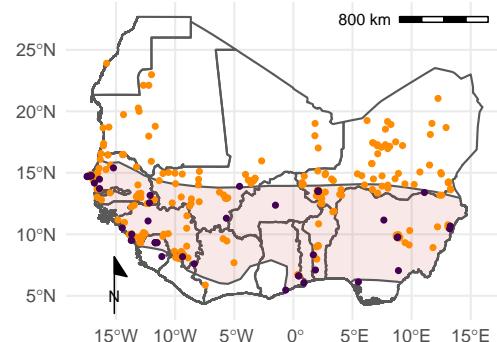
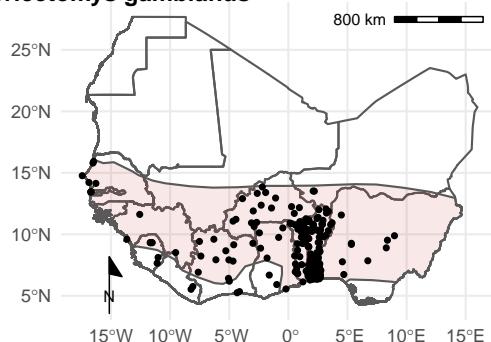
Arvicantis niloticus



Praomys daltoni



Cricetomys gambianus



Detection/Non-detection • Detection • Non-detection



Figure 9: Supplementary Fig 4. A) 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. B) 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.