Topicos Lab

Carlos Zambrana-Torrelio

2022-04-22

Let's map the distribution of Mammarenavirus hosts. First, let's make a table with species and number of viruses.

```
virion = vroom(here("data/Virion.csv.gz"))
```

```
## Rows: 484464 Columns: 34
## -- Column specification -----------------------------------------------------
## Delimiter: "\t"
## chr (19): Host, Virus, HostGenus, HostFamily, HostOrder, HostClass, HostOrig...
## dbl (11): HostTaxID, VirusTaxID, PublicationYear, PMID, ReleaseYear, Release...
## lgl  (4): HostNCBIResolved, VirusNCBIResolved, ICTVRatified, HostFlagID
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
virion %>%
  filter(Host == 'homo sapiens') %>%
  select(Virus) %>%
  distinct() %>%
  pull(Virus) -> zoonoses

virion %>%
  filter(VirusGenus == "mammarenavirus",
         HostOrder == "rodentia",
         DetectionMethod == "Isolation/Observation") %>%
  filter(Virus %in% zoonoses) %>%
  select(Host, Virus) %>%
  distinct() %>%
  group_by(Host) %>%
  summarize(NVirus = n_distinct(Virus)) -> nvir
```

Now let's read spatial data obtained from the IUCN.

```
## Reading layer 'rodents_arena' from data source
##   '/home/rstudio/repos/topicos2022/data/rodents_arena/rodents_arena.shp'
##   using driver 'ESRI Shapefile'
## Simple feature collection with 31 features and 29 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -161.9321 ymin: -54.77828 xmax: 179.999 ymax: 73.53033
## Geodetic CRS:  WGS 84
```

Let's map the results. First, we need to convert the IUCN polygons to rasters. We are using the `elevation` at a spatial resolution of 10' (check `prep_files.R`).

```r
r = raster::raster(here("data/wc10/alt.bil"))
# Count the number of overlapping hosts
map.num.m = fasterize(iucn, r, field = NULL, fun = 'count')
# Sum the number of viruses per host
map.sum.m <- fasterize(iucn, r, field = "NVirus", fun = 'sum')
```