MapVEu

Note: this exercise uses VectorBase.org. MapVEu is limited to only certain VEuPathDB sites at this time.

What is MapVEu?

MapVEu is a web application and exploratory data analysis platform that facilitates access to and exploration of geospatial data. It integrates genomic, phenotypic and population data for traits such as insecticide resistance genotypes and phenotypes, genetic variation with microsatellites, chromosomal inversions and SNPs, population abundance, pathogen infection status and blood meal identification.

It is a powerful tool, but has many features that are not relevant to any particular dataset.

For today's analysis, we will be looking at tick data from the State of Maine. Citizens can collect ticks they find on themselves and send them for testing for pathogens. If interested, learn more about this service at https://extension.umaine.edu/ticks/management/tick-surveillance/.

Part 1: State of Maine, USA, Tick Surveillance:

https://vectorbase.org/vectorbase/app/workspace/maps/DS_e18287e335/

(click "New Analysis" to begin)

Things to try:

1. From Configure Map > Donuts , change the variable from species to a different field.

2. From Configure Map, try (explore) adding some filters (lets save Pathogen detection assays for later). Perhaps you will filter out a species, property type, data collected and/or feeding status. Note some filters will preclude using other filters (e.g. if this was a mosquito study, and you choose "Anopheles," there would be no samples that have "dengue" You can remove filters (at any time by clicking the X's shown, below)
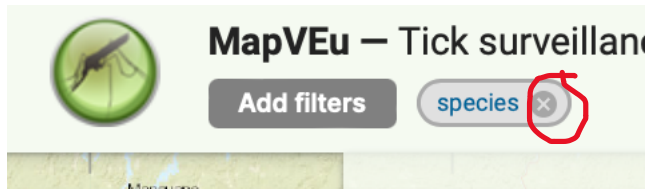
3. Please remove all filters (clicking the X's shown, above)

4. Let's look at infection data:

   a. Filter for ticks infected by Babesia – what kinds of ticks spread this disease? (hint: filter for Babesia positive ticks to be safe, then set doughnuts to tick species)
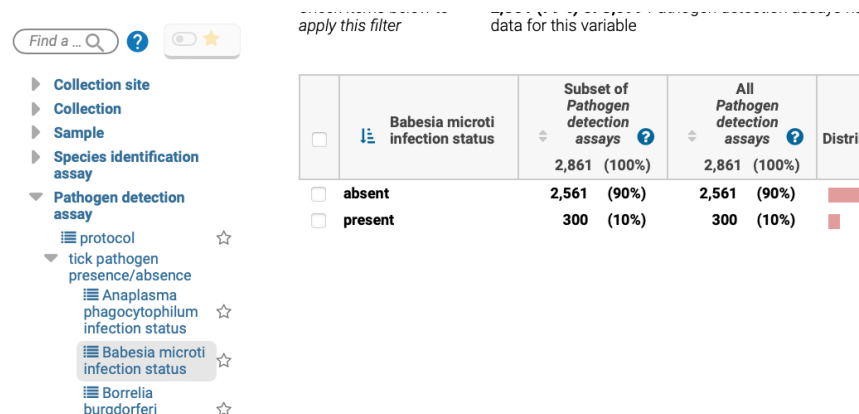
   b. Try Ehrlichia species to see what tick species is infected by this pathogen (hint: the health authorities never found it, but you can see what tick species they were testing)

   c. Clear filter and look for ticks doubly infected with Babesia and Lyme disease (*Borrelia burgdorferi*)

   d. Thinking about b and c above, why can you not look for ticks doubly infected by both Ehrlichia and Borrelia, in this dataset?

5. Please clear all filters

6. Let us make a supporting plot showing species variation over time:
   a. As shown below click Donuts > Supporting plots > Histogram
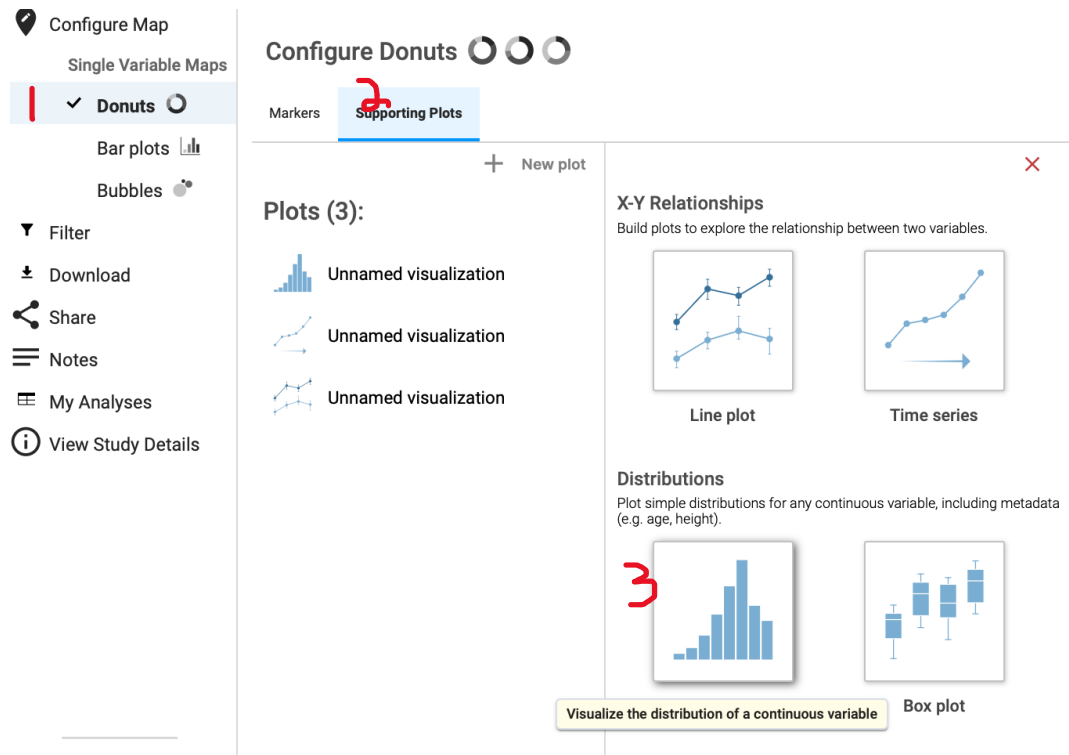
b.

A new window will open. Under axis variables, toggle on "Only show compatible variables" (as shown) then select "specimen collection start date"
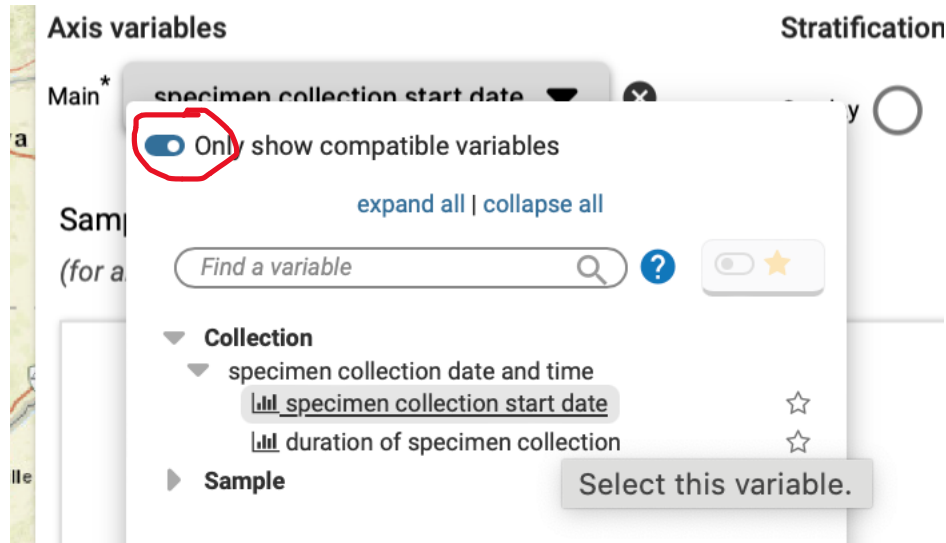
c. Select Species as your stratification variable
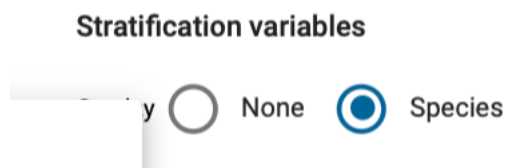
d. You should now have a graph of seasonal changes in species composition over time. Scroll down in that graph window to see x-axis range and try remove dates prior to the 2022 season.

[You may normally access the map from the VectorBase menu bar. Select tools > MapVEu — Geolocation visualizations of field-based studies. For bandwidth reasons, please do not do this during the session]

Part 2: Upload your own dataset ("EDA My Studies")

Our system allows you to upload and explore your own datasets. Since you probably don't have a dataset with you right now that is ready, we have prepared one for you. This dataset has some genomic data on mosquito insecticide resistance form a few merged datasets.

 You can download it [here](#).

We will use our sister site, www.ClinEpiDB.org, as we have not fully moved this functionality to VectorBase.org yet. You will use same login credentials.

1. Open your datafile in Excel and look at column names:

    Note dates are in YYYY-MM-DD format
    GPS fields are labeled, exactly, "Latitude" and "Longitude"

    No special characters

2. Go to ClinEpiDB, Go to Workspace > My Studies
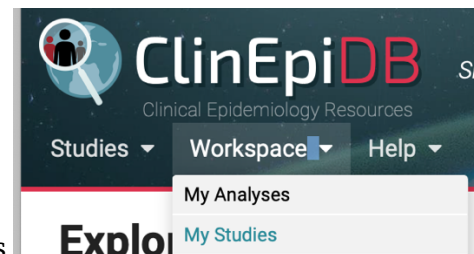
3. Upload your datafile (I've provided some suggested titles/descriptions)

    Suggested Name: IR data (your name)
    Suggested Title: Some IR data on KDR N1575

4. And now you wait! Every 3 min please push the 'refresh' button on your browser

My Study: *IR data* ✏️

Status: 🕐 This study is queued. Please check again soon (reload the page).

Owner: Me

5. Once your data is loaded, click to "Explore in MapVEu"!

My Study: *IR data* ✏️

Status: ✅ This study is installed and ready to use in ClinEpiDB. ✓

Owner: Me
Description: *No Description* ✏️
ID: 9XPjzBKK2t4
Data type: ISA Study (isasimple 1.0)
Summary: Some sample IR data on KDR N1575 ✏️
Created: 6/24/2024, 6:15:26 AM
Data set size: 1.14 KB
Quota usage: 0% of 10.74 GB

🏛️ **Explore in ClinEpiDB**

🏛️ **Explore in MapVEu**

6. Explore in MapVEu

   a. Make a map showing distribution of data by species
   b. Make a map showing of data by KDR genotype
   c. Try making a contingency table supporting plot of genotype by species (as shown below)

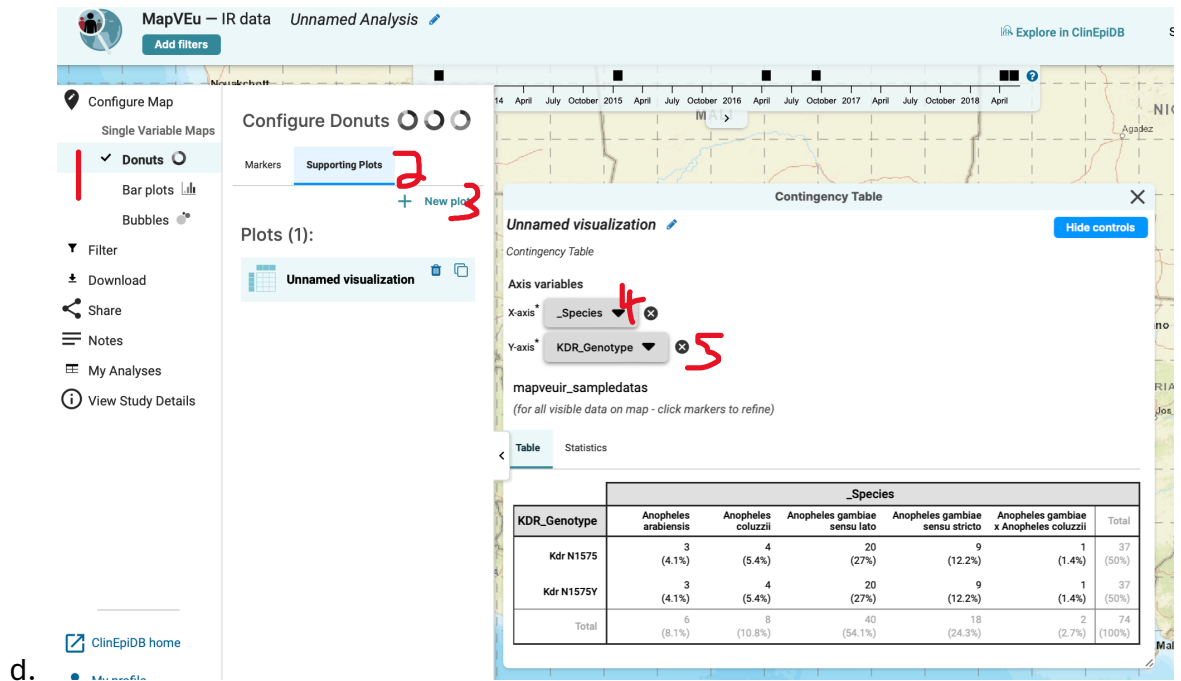MapVEu — IR data *Unnamed Analysis* ✏

Add filters   🔍 Explore in ClinEpiDB

📍 Configure Map

    Single Variable Maps

✓ Donuts ◯

    Bar plots 📊

    Bubbles 💧

▼ Filter

⬇ Download

◁ Share

≡ Notes

⊞ My Analyses

ⓘ View Study Details

🔗 ClinEpiDB home

d.

Configure Donuts ◯ ◯ ◯

Markers | **Supporting Plots**

＋ New plot

Plots (1):

| Unnamed visualization | 🗑 📋 |

Contingency Table   ✕

Unnamed visualization ✏   **Hide controls**

Contingency Table

Axis variables

X-axis* | _Species ▼ | ✕

Y-axis* | KDR_Genotype ▼ | ✕

mapveuir_sampledatas
(for all visible data on map - click markers to refine)

Table | Statistics

	_Species					
KDR_Genotype	Anopheles arabiensis	Anopheles coluzzii	Anopheles gambiae sensu lato	Anopheles gambiae sensu stricto	Anopheles gambiae x Anopheles coluzzii	Total
Kdr N1575	3 (4.1%)	4 (5.4%)	20 (27%)	9 (12.2%)	1 (1.4%)	37 (50%)
Kdr N1575Y	3 (4.1%)	4 (5.4%)	20 (27%)	9 (12.2%)	1 (1.4%)	37 (50%)
Total	6 (8.1%)	8 (10.8%)	40 (54.1%)	18 (24.3%)	2 (2.7%)	74 (100%)

Part 3: Try a dataset with numerical values and filters

https://clinepidb.org/analysis/KJOec5h

Try filtering by Sample > Lab test > blood test > Hemoglobin

Part 4: Think and discuss at your table about ways you could use MapVEu in your own research