# MapVEu: part II - Additional features

***Note***: *this exercise uses VectorBase.org data as an example, but the same functionality will soon become available for other VEuPathDB resources*

**Learning objectives**:
Basic/introductory features have been covered in **part I** of this tutorial
- Use date and season search
- Construct queries with 'NOT' searches & reset queries
- Interpret different tables and figures and their corresponding metadata

The map has worldwide field collected samples, some with omic/genetic related test performed
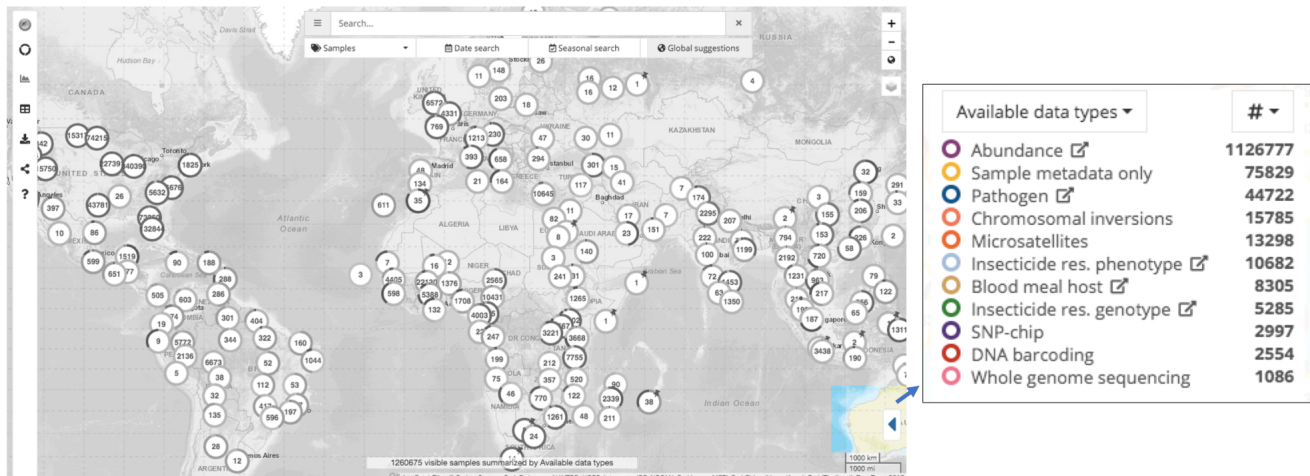


Omic/genetic tests:
- Sample metadata only
- Chromosomal inversions
- Microsatellites
- **Insecticide resistance genotypes**: identification of polymorphisms that confer resistance or susceptibility
- SNP chip
- DNA barcoding
- Whole genome sequencing

Other test or measures:
- **Population abundance**: how many arthropods were collected in a specific location?
- **Pathogen infection status**: what virus or parasite were the arthropods infected with?
- **Insecticide resistance phenotypes**: are the arthropods exposed to insecticides susceptible or resistant?
- **Blood meal host**: what was the source of the arthropod blood meal? Did it bite a human or which other animal?

In part I of the tutorial we look at 'Samples' and 'Insecticide Resistance' phenotypes, here we will look at insecticide resistance 'Genotypes' and population 'Abundance'. In the last page there are links to **bonus cases** with all map views

# Use case #1: what is the distribution of the DDT and pyrethroid insecticide susceptibility based on the knockdown resistance (kdr) mutations, for the genus *Anopheles*?

1. Let's go to the Genotypes view. It shows the prevalence of locus and their corresponding alleles. from genotype assays. Currently mostly for insecticides but it can display any other type of genotype
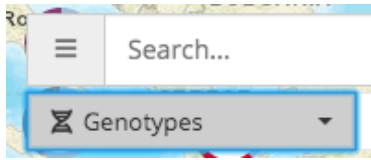


2. By default it opens in Allele, change to Locus. In the Locus let's select kdr L1014, this action adds the term as query in the search box. Go back to Allele, this time it onlys shows the alleles for this locus
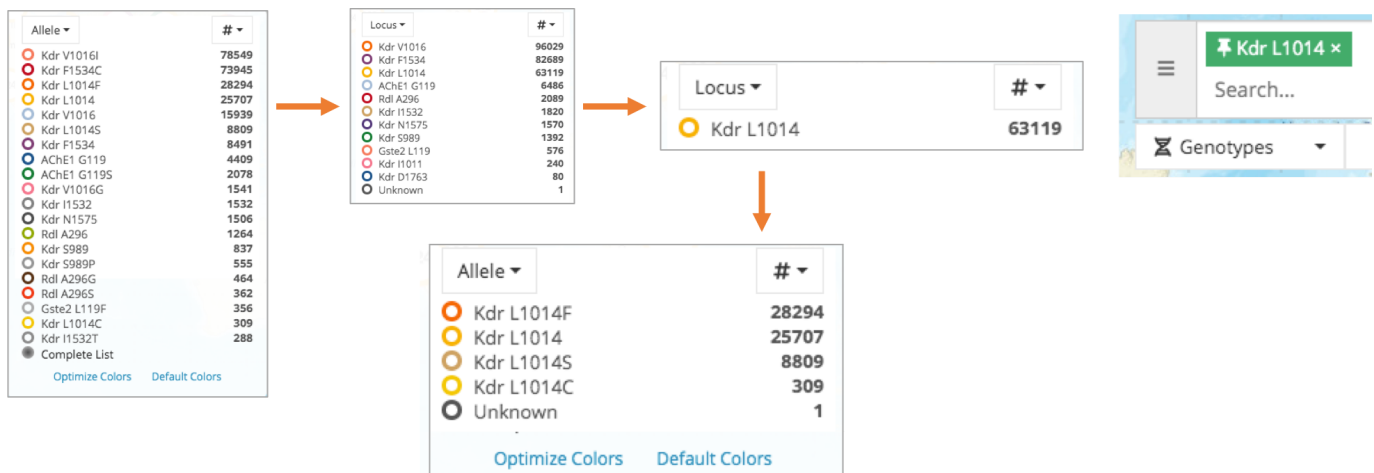


3. Click on 'optimize colors' for a better color representation. The purple Allele (Kdr L1014) is the wild type (with L-leucine), the others are the mutants (with F-phenylalanine, S-serine and C-cysteine)



| Allele ▾ | # ▾ |
|---|---|
| ○ Kdr L1014F | 28294 |
| ● Kdr L1014 | 25707 |
| ○ Kdr L1014S | 8809 |
| ○ Kdr L1014C | 309 |
| ○ Unknown | 1 |

4. Filter only for the genus *Anopheles*



2
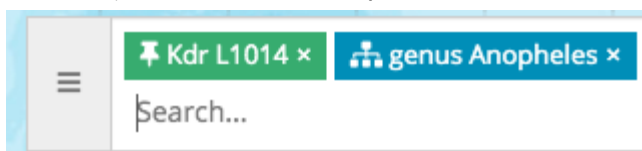
5. Filter for the collection protocol, including all except 'pyrethrum spray catch' with control (Windows) / command (Mac) + click. We perform this filter because we are looking at insecticide resistance genotypes and that collection protocol, using pyrethrum, creates a bias in our results
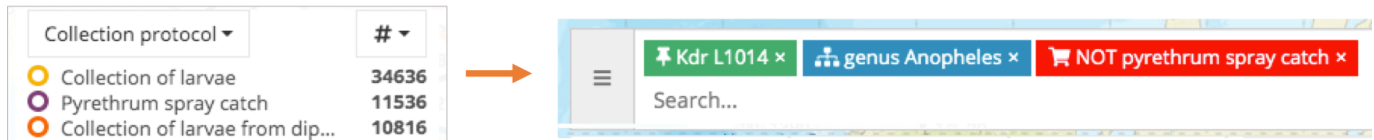


# Results interpretation case #1: In India for example, results show that the wild

type alleles are mostly located towards the center and South of the country, with a higher prevalence compared with the mutants in the North. In contrast, in continental Africa, the difference seems to be East to West, with a lower proportion of the wild type in the East. In Madagascar, there are only wild type alleles.
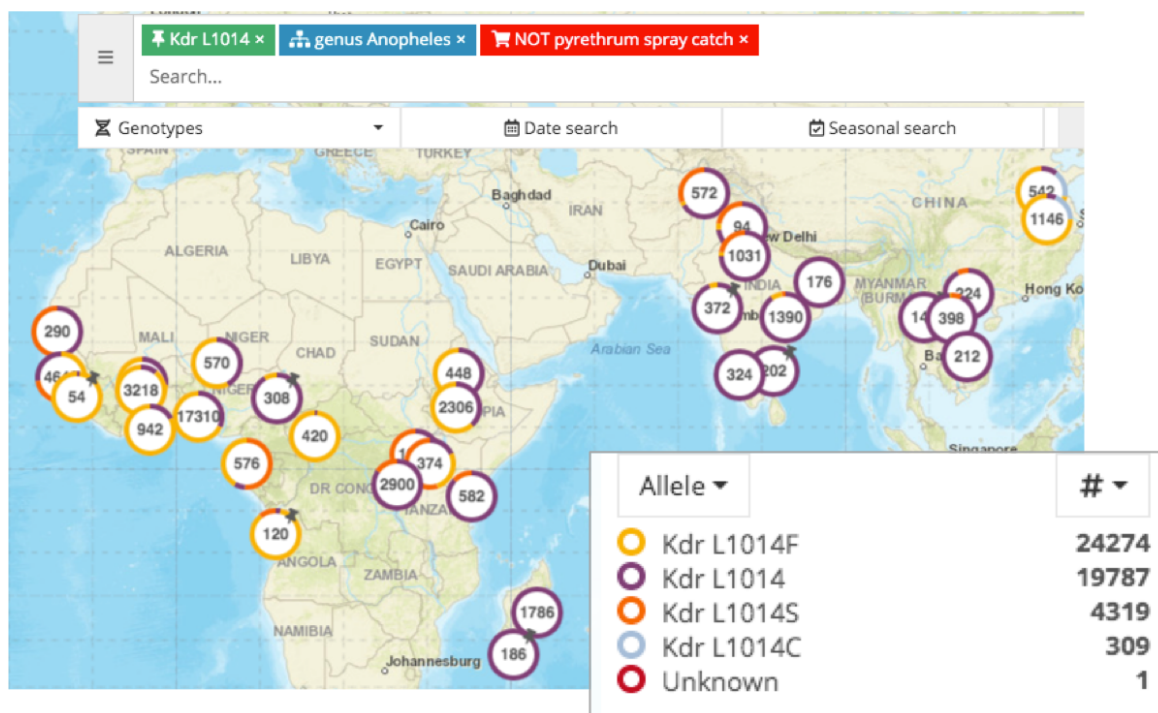


6. Clear or delete all queries and filters with a click in the search box right hand 'X'
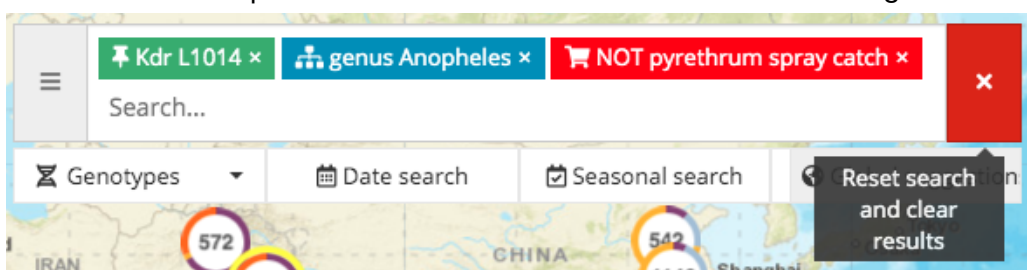
**Use case #2**: Currently most of the population abundance data come from the USA mosquito control abatement districts, and we are expecting soon to increase the mosquito data from Africa too. Most of the non-mosquito data are the *Culicoides* data coming from the UK. The map can handle any arthropod vector or related invertebrate species of interest. Let's explore some trends of the mosquito population abundance data in the USA for the state of Iowa

7. Let's go to the population "Abundance" view.
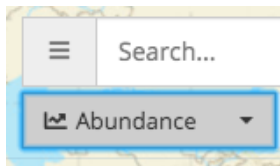


8. Query the data for the state of Iowa. Notice most of the data is from *Aedes vexans*.
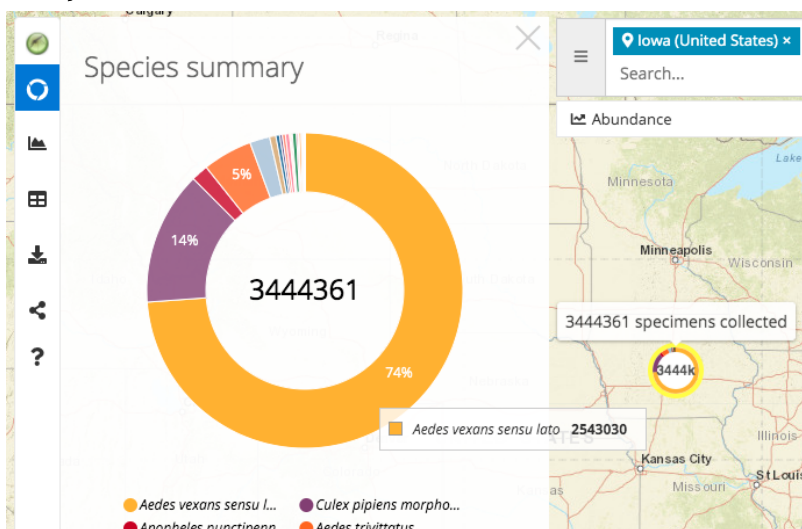


9. Now let's explore the species prevalence in time. On the left of the map, open the timeline graph. Using the double lines, expand it to show a yearly view. There is data from 1969 to 2017, with an unusual data peak in 1975 and data missing from 1971 to 74. This is 44 years of data!
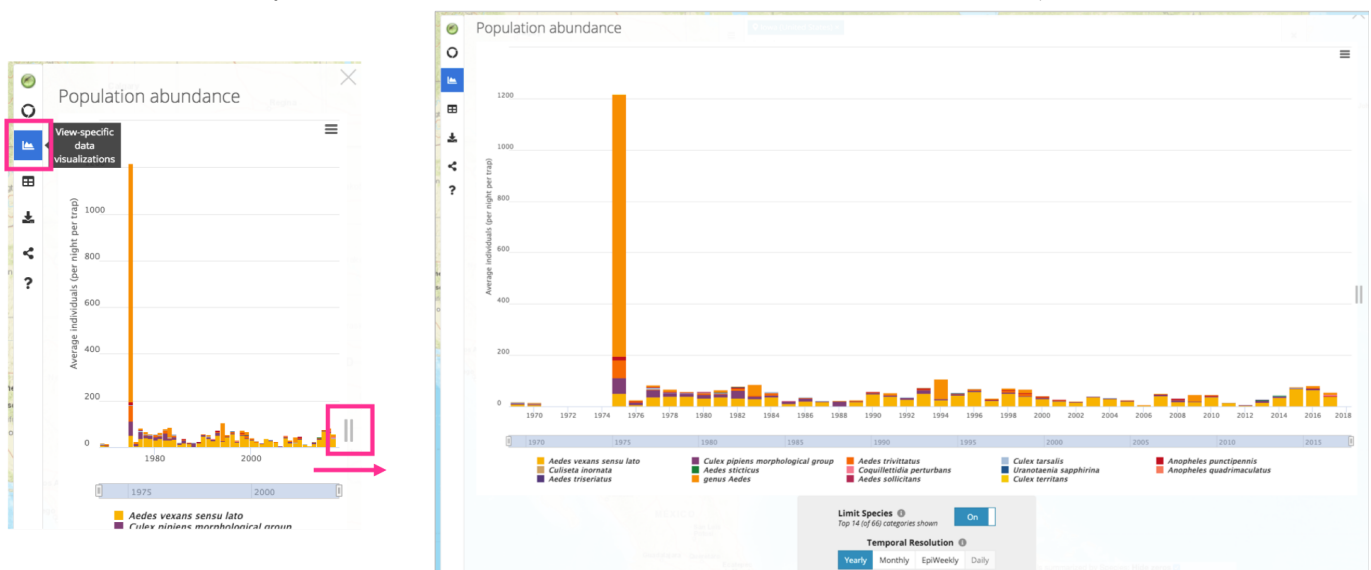
10. What are the most prevalent species in each date range? Select one range at the time, open and expand the timeline graph, and change the temporal resolution to year. Hover with the mouse over the bars, to identify the most prevalent species.

a. 1985 - 1989: this shows that the two most prevalent records are from two species complexes, *Aedes vexans* and *Culex pipiens*.
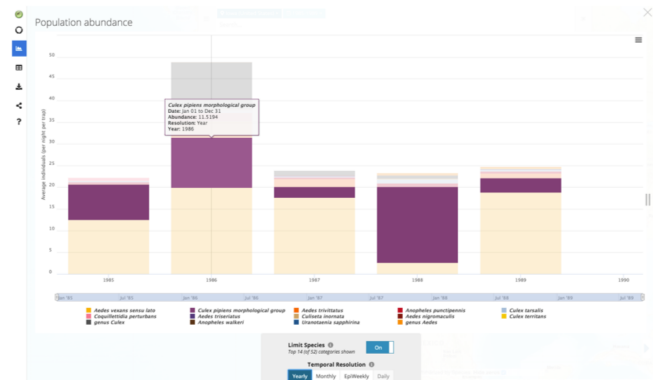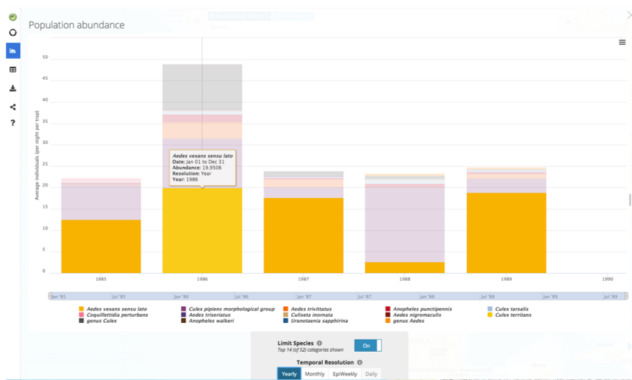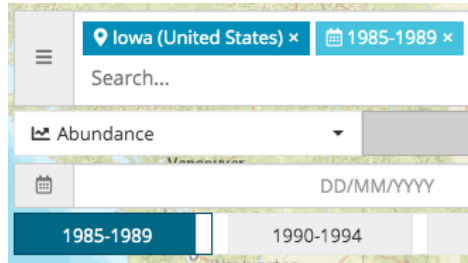




b. 1985 - 1989: To know the exact species distribution for this date range, go back to the pie chart.
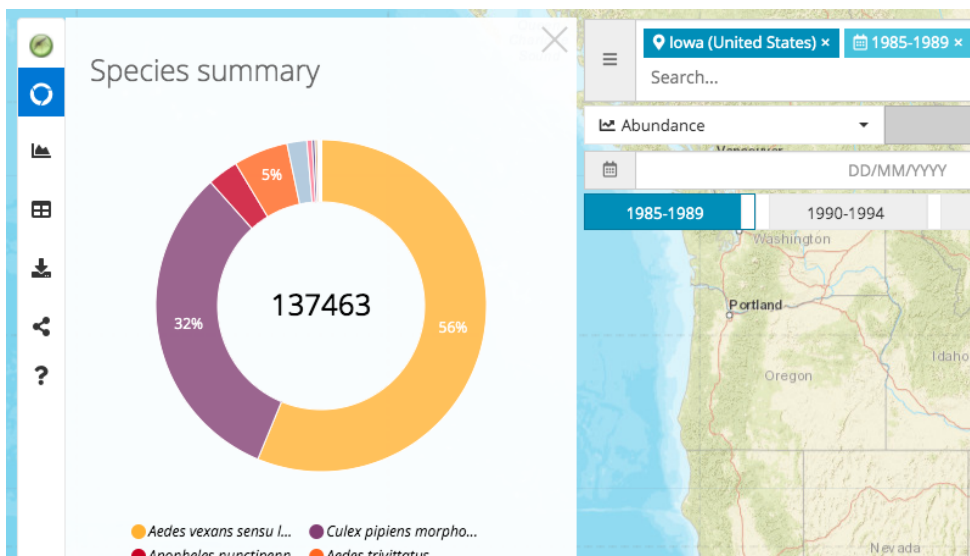


c. We could repeat these steps for each date range, to identify any possible species patterns.

11. Data can also be explored at lower geographic resolution, looking for patterns from north to south and east to west in the State

12. To explore species abundance patterns of increase and decrease monthly, you could look at the data one month at the time
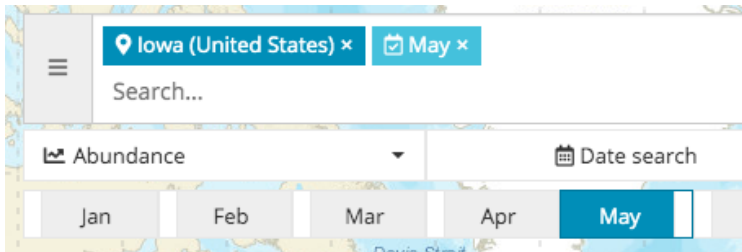


13. Metadata for each record is shown in the sample details, where the dotted lines indicate filters that can be applied in the top search box or the right hand side filter.
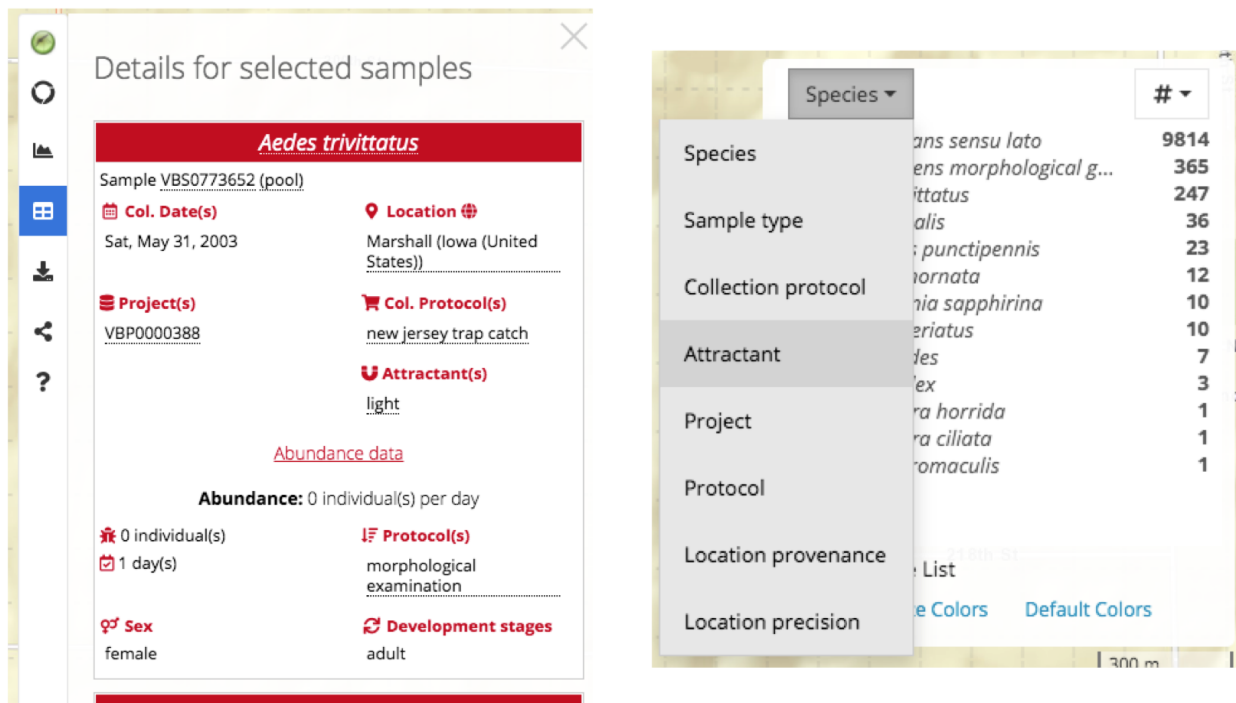


# A scientific paper is a static view of the data, in contrast when you submit to MapVEu (VectorBase.org):

- the data on the map gets updated every two months
- the map allows data visualization with different filters
- the data can be analysed at different geographic levels
- the data from both scientific research and control/monitor agencies can be compared
- your papers gets more exposure and users can find it both with a literature search or with a query in MapVEu

- if a reader from your papers wishes to have your raw data, they do not need to contact you, it can be download from MapVEu

# **Bonus cases:** Try these queries using MapVEu

Samples:
- PCR-based species identification
- animal baited net trap catch, genus *Anopheles*
- genotyping by high throughput sequencing, pyrethrum spray catch

Insecticide Resistance:
Insecticide resistance assay results can be filtered by species, sample type, collection protocol, project, assay protocol and insecticide.
- *Anopheles gambiae*, West Africa, malathion
- WHO larvicide dose response test, temephos, collection of larvae
- Africa + sample type (individual or pool)
- Pyrethroid + genus *Aedes* + collection protocol (ovitrap, CDC light trap, …)
- CDC bottle bioassay + permethrin + collection of larvae + project (VBP*)

Genotypes:
Shows the distribution and proportion of insecticide resistance alleles. It can be filtered by a single locus or allele, species, sample type, collection protocol, project or assay protocol.
- Madagascar, Kdr L1014, *Anopheles arabiensis*
- *Anopheles gambiae, Anopheles gambia*e sensu lato, *Anopheles gambiae* x *Anopheles coluzzii*
- AChE1 G119, Rdl A296
- India (displaying the markers as 'Allele')

Abundance:
Temporal abundance that can be filtered by species, sample type, collection protocol, project and assay protocol. We have large multi year datasets from regular USA mosquito control district surveillance. We also have records from other countries.
- State of Iowa, genus *Culex*, CDC gravid trap, Lyric Bartholomay
- VBP0000213, morphological examination, RT-PCR assay
- pool, house resting-daytime catch
- *Aedes albopictus*
- Location + sex + identification method + taxon + date range + trap + attractant + GPS provenance and precision
- Florida +  BG-Counter trap[1]

---

[1] Traps that automatically identify the collected material (to Culicidae). We obtain the data directly from the manufactures API

Pathogen:

The pathogen status view provides vector pathogen-status testing, which can be filtered by pathogen, infection status, vector species, sample type, collection protocol, project and protocol assay.

- West Nile virus, *Culex pipiens* group
- Query with no keywords: directly zoom in and out and move around the map
- *Plasmodium*