# Population Biology (PopBio)

Advanced Search

Answer key

We've already introduced the basics of search and learnt how to query the map interface, now let's see how to find more specific information and data.

1. Click the "Advanced Search" link below the search box in the top-right corner of any page in VectorBase. You will land on a page showing every single search result available in VectorBase.
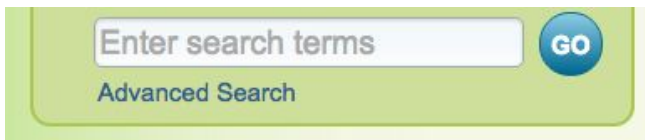
As you may remember from the Search tutorial, there is a "Filter Results" section on the left hand side. You can examine in detail the subsections of VectorBase data by clicking on the links in the filter section, and "back out" again by clicking the "Reset Filter" links.

2. Click on "Population Biology". How many samples are in the database?

Sample genotype	306,465
Sample	102,514
Sample phenotype	58,442

Sample search

3. Click on Sample and scroll down the filter section to see links for restricting further by species and project. Which project has the most samples?

VBP0000005

4. Click on that project[1] and see which species it contains data for.

> *Anopheles arabiensis, An. coluzzii, An. gambiae* and *An. gambiae sensu lato*

Reset the project filter, …

and open up the Advanced Search panel so that you can see the "Add field" link.

Click on "Add field" and check the boxes for "Collection date" and "Collection site". Click anywhere outside the popup menu to dismiss it.

---

[1] Notice that after you filter you may still see more than one project. This is because sometimes a sample or assay might be re-used in different project.

We're going to search for samples collected during the year 2000 in different regions of Africa. First enter the date range[2] 2000-01 to 2000-12, then collection site "West Africa" (with or without quotes). Click "Go" or just "Enter/Return" in your keyboard.

5. Write down the number of results for search for both West and East Africa

West Africa	East Africa
194	208

Practice adding and removing more advanced search filters. For example you could restrict results to "population" in "Sample type" (other common values for Sample type are "pool" and "individual").

6. How many "individual" type samples were collected in Mali? (no date restriction).

14,242

7. Add a "Search by circular geo-location" to restrict results to those collected within 100km of coordinates11.314, -5.695 (Sikasso, Mali). How many results are there now?

993

The main search box

Note that while using advanced search it is normal to ignore the main search box (labelled "Enter terms") just above the advanced search fields. Just leave the asterisk >*< there and use the advanced search fields only. You will get more results this way.

Sample genotype search

Start a completely new advanced search (link in top-right of any VectorBase page). In the advanced search menu, select "Population Biology" and "Sample Genotypes". If you click "GO" now, it's the same as if you had click using the search filters into PopBio Sample genotype, but with fewer clicks.

Now add an advanced search field for "Collection/catch method" and enter "aspiration" and click "GO" or "Enter/Return" in your keyboard. Inspect the results.

---

[2] In some browsers a graphical "date picker" will appear, in others you may have to type the YYYY-MM date.

You may notice that the collection protocol is actually called "catch by aspiration". You can alter your search to "catch by aspiration" or even "MIRO:30000043"[3] which is the accession code/formal ID for the ontology term that is stored in the PopBio database.

How would you find out this ontology term accession? An easy way, is to click on the green links in the PopBio projects of an item you are already interested in.

To find the green link for a "catch by aspiration" ontology term,
- open one of the search results in a new tab,
- navigate to the sample page,
- and then to the collection page.

Click on the green text (example below) and you will see details about the ontology term and its accession code.

Field collection

VectorBase stable ID	VBA0182187
Type / Name	Field collection / G07-00021.c1
Description	A. coluzzii field isolate, Goundry, Burkina Faso, individual insect
Performed by	n/a
start date	2007-06
end date	2007-11

Protocols

Name	VBA0182187:ADULT
Description	n/a
Type	house resting - daytime catch
External link	

Collection site

Name	Loumbila Department
Coordinates	12.5003,-1.3336 (WGS 84)
Altitude	
collection site	Loumbila Department
village	Goundry

---

[3] Mosquito Insecticide Resistance Ontology, MIRO. For ontologies and acronyms please visit The OBO Foundry site http://www.obofoundry.org/ and the NCBO's BioPortal site http://bioportal.bioontology.org/ontologies

When you click on "house resting - daytime catch" you see that this term is a child (more specific sub-term) of the term "catch by aspiration" (see below). When you search with a more general term in VectorBase, you get results for items that are annotated with its more specific child terms, as we have seen here.

Ontology term: house resting - daytime catch   [close]
Accession: MIRO:30000007
Definition: A field population catch method designed to catch specimens that are resting in a house during daytime.
Synonyms:
Parent term: catch by aspiration

Genotype assay subtypes

Using the "Add field" menu, remove the "Collection/catch method" field and add the "**Genotype class**" field. The values you can enter here are
- chromosomal inversion
- microsatellite
- mutated protein

Let's search for "chromosomal inversion" genotype assays…

7. How many are there in total?

90,550

8. How many chromosomal inversion assays used protocol "cytological chromosome examination"" (MIRO:30000037)? (Hint, use the "Protocols/Method" field.)

88,257

Now let's add the "**Genotype name**" field and search for a specific chromosomal inversion genotype.

There are a small number of well documented chromosomal inversions. 2Rc is one of them.
Type '2Rc' into the box (*case sensitive*) and press "Go". How many hits you obtained?

14,533

9. You can use the "Inversion count" field to search for homozygous inverted (2), heterozygous inverted (1), and homozygous non-inverted/wild type (0). How many samples have the "2Rc heterozygous inverted" genotype?

2,430

Inversion count

The values can only be 0, 1 or 2

`2` `≤ ⇕` *Inversion count*

Downloading results

Download a text file of comma-separated results using the "Export results" button. This feature is recent, please let us know about other fields/columns that you think we should include in the file.

Insecticide resistance search

11. Switch the Domain/sub-domain to "Sample phenotypes" and clear/cancel the Protocol filter. How many results you obtain?

2,430

.
Now find out how many phenotypes we have for the insecticide "deltamethrin"?

1,699

Searches on resistance values (the assay results) will be coming soon, but for now you can view and compare resistance data on the PopBio Map!

12. Higher level ontology term search:

Change the "Insecticide" field query to "pyrethroid" - you should get more results than you got with "deltamethrin". Try also "carbamate".

pyrethroid	carbamate
4,011	1,260

Which class of insecticide has more results in VectorBase at the moment?

pyrethroid

Remember you can browse through most of the ontologies used in VectorBase with our Ontology Browser (tools menu).

Project search

Finally we will show you a useful way to get a list of all insecticide resistance projects in the PopBio database.Change the Domain/sub-domain to "Population Biology - Projects", add the "Study design" field and enter "compound treatment design" into the box (and click "GO"). How many results you obtain?

159

Wrap up

Now you have had a comprehensive introduction to Advanced Search of PopBio data in VectorBase. Similar principles apply to other data domains, such as expression and genome annotation data. *Remember to look out for the improved and hopefully simplified PopBio search coming soon.*

Optional exercises

1. Use Advanced Search and/or the main PopBio Map Search to compare the species collected using animal vs. human/man baited catch methods. Choose one geographical area for this analysis. It may be useful to filter on "population"-type Samples. In the PopBio map, you can select each of the animal-related catch protocols or you can type "animal" and select the "~7721 *in Collection protocols*" from More suggestions.

2. As above, use Advanced Search and/or the main PopBio Map to compare species distributions between two neighboring geographical regions (where some species will be shared). Filter on sample type = "population" or "individual" (but not both). If using the map, you may want to zoom out fully.