



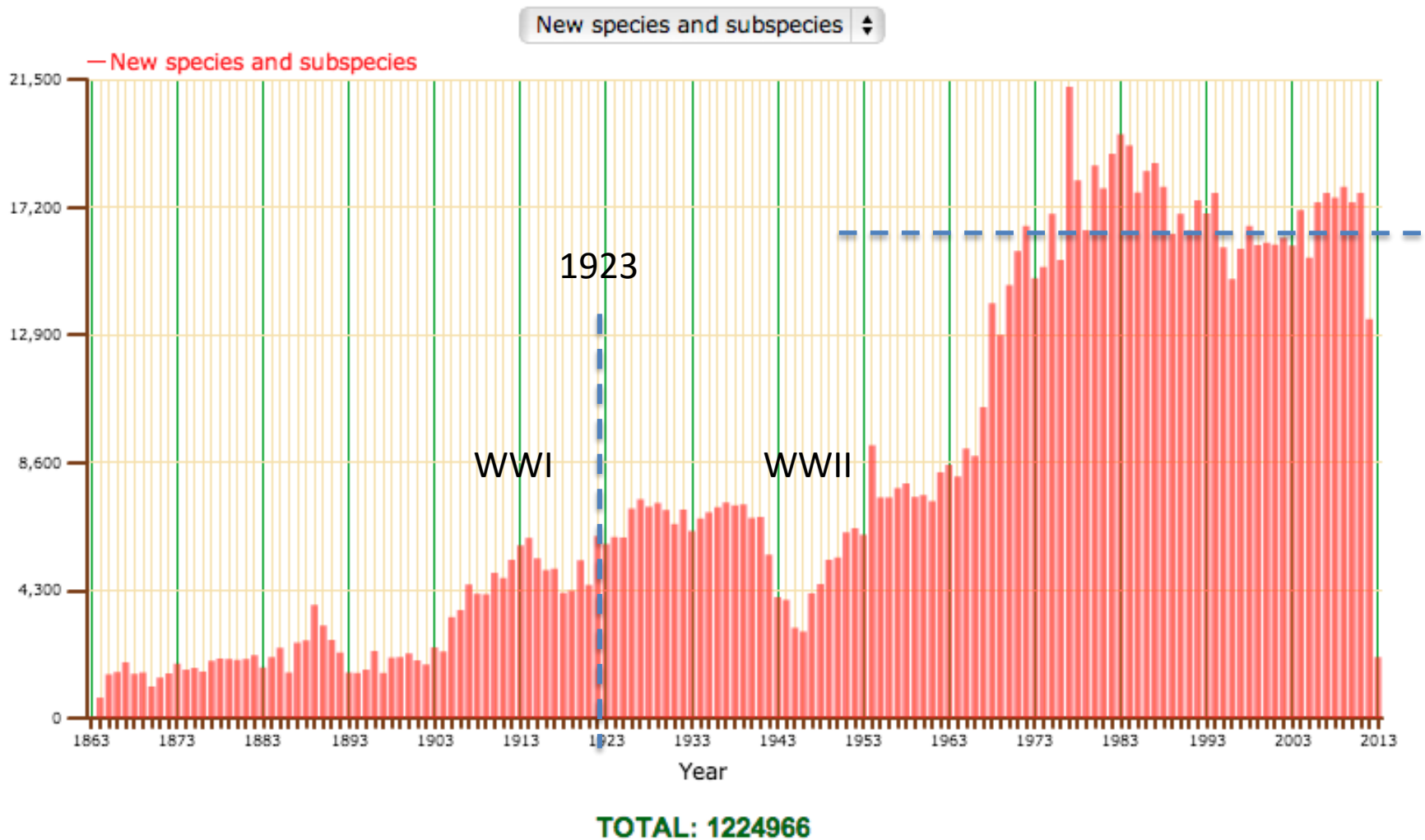
@rdmpage

<http://iphylo.blogspot.com>

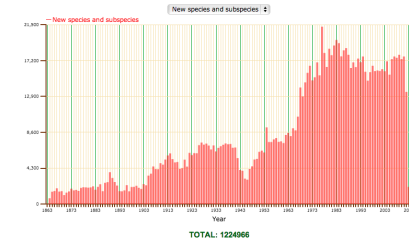
# Two graphs, three responses



# Numbers of new animal names

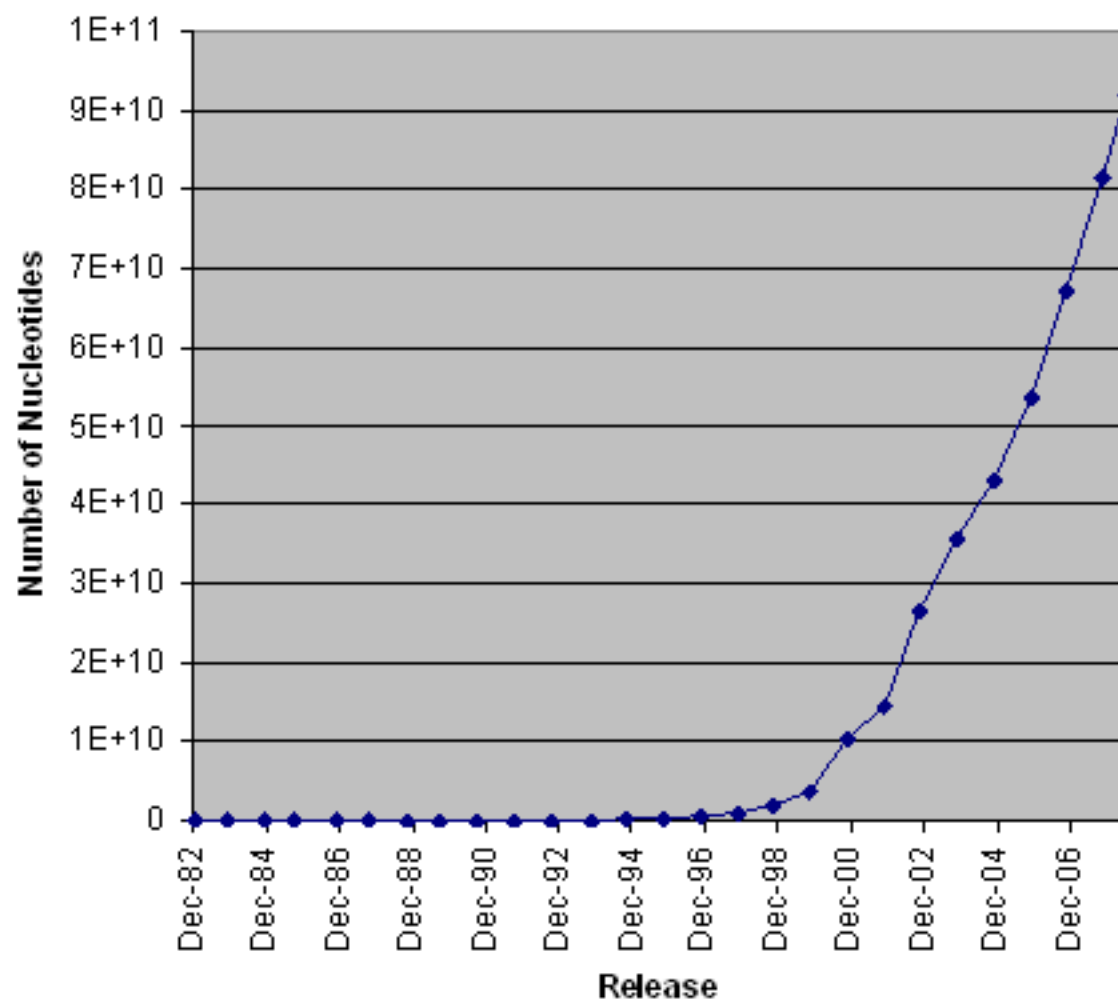


# Implications

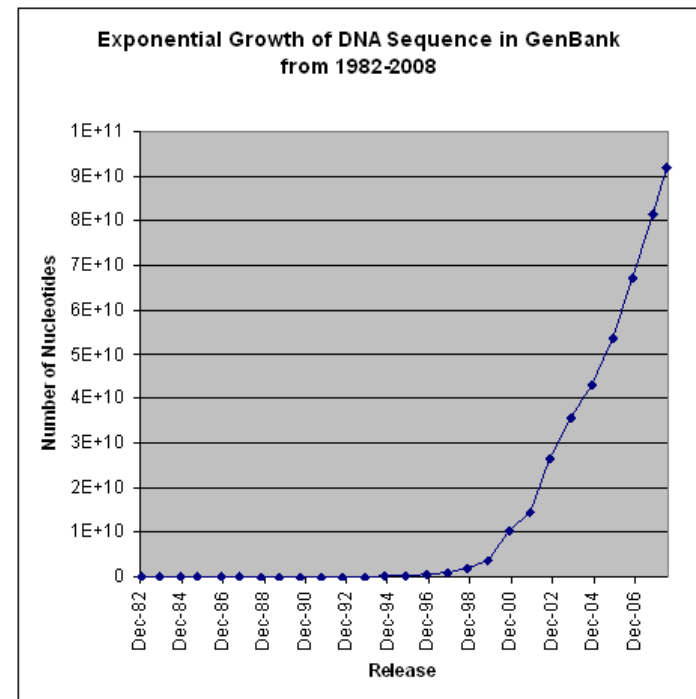
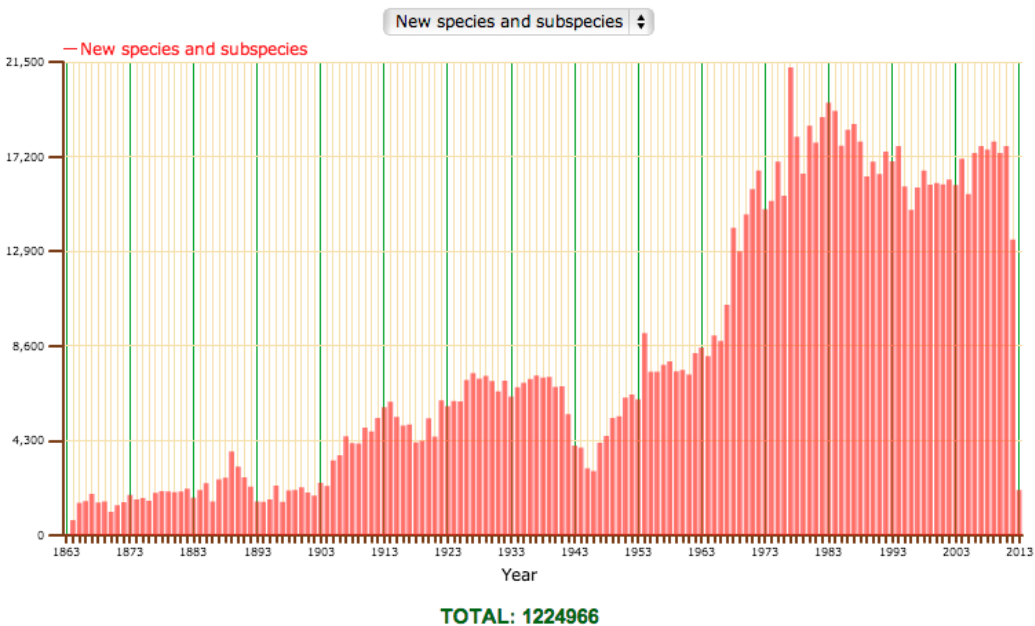


- Taxonomists are working at capacity
- Maybe we are running out of species to discover...(!)
- Most taxonomic work is in the past (“legacy”)

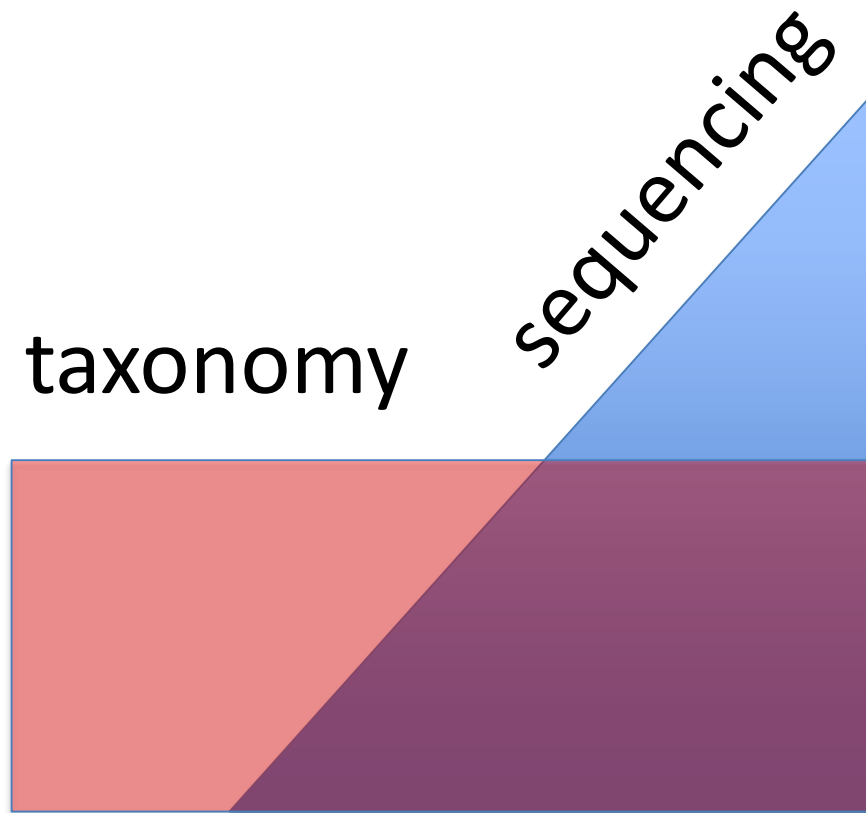
### Exponential Growth of DNA Sequence in GenBank from 1982-2008



# Two graphs



# Two graphs

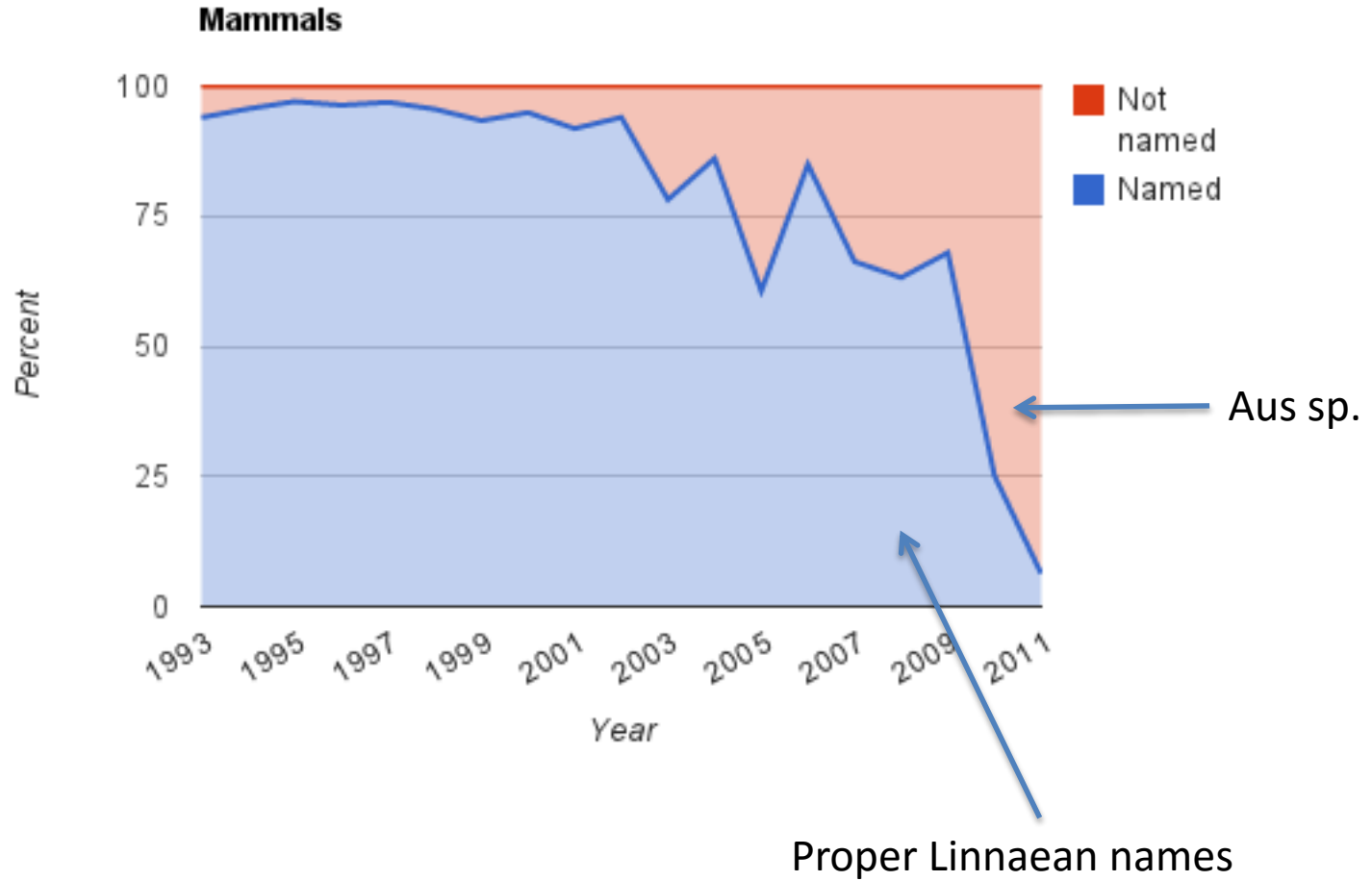


# Dark taxa

(taxa with sequences but no proper  
scientific name)

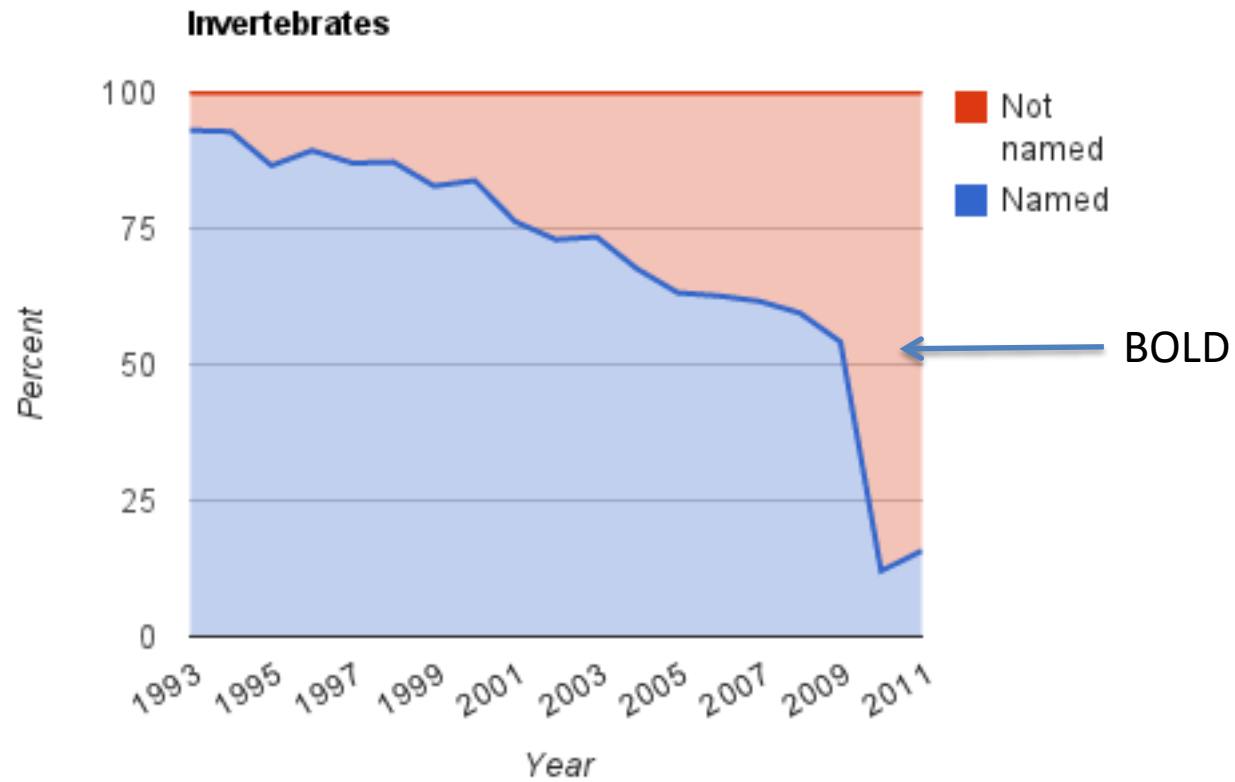
<http://iphylo.blogspot.ca/2011/04/dark-taxa-genbank-in-post-taxonomic.html>

# Mammals

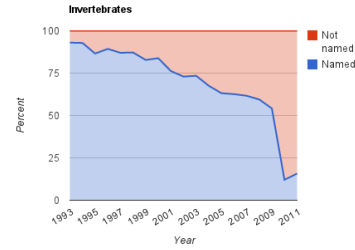




# “Invertebrates”

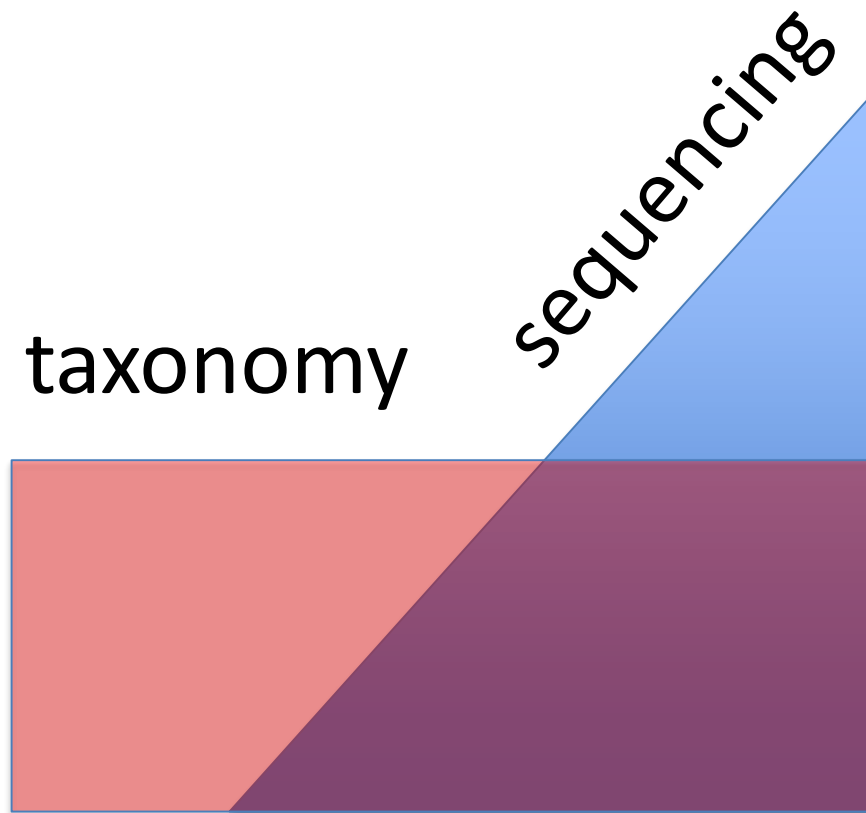


# Dark taxa



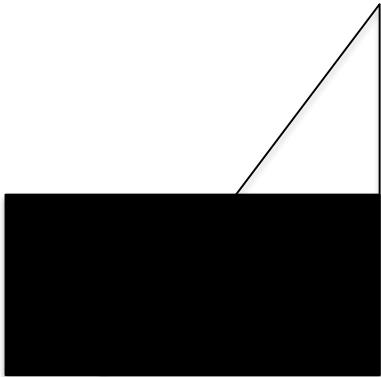
- Disconnect between taxonomy and genomics
- Are “dark taxa” species we already know about, or do they represent new diversity?

# Two graphs

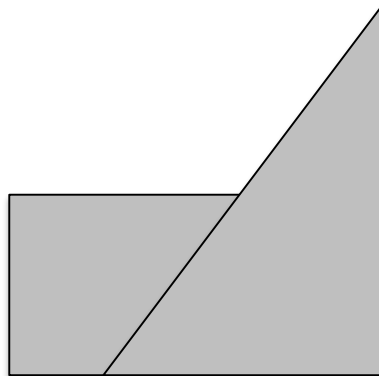


# Three responses

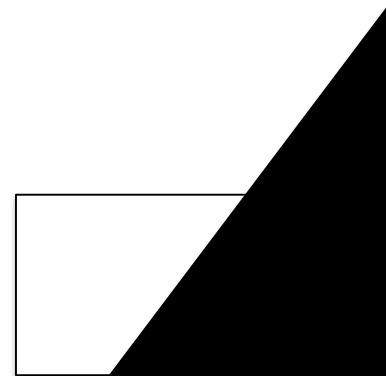
Make  
taxonomy  
digital



Integrate  
taxonomy &  
sequences



Sequences  
eat the  
world







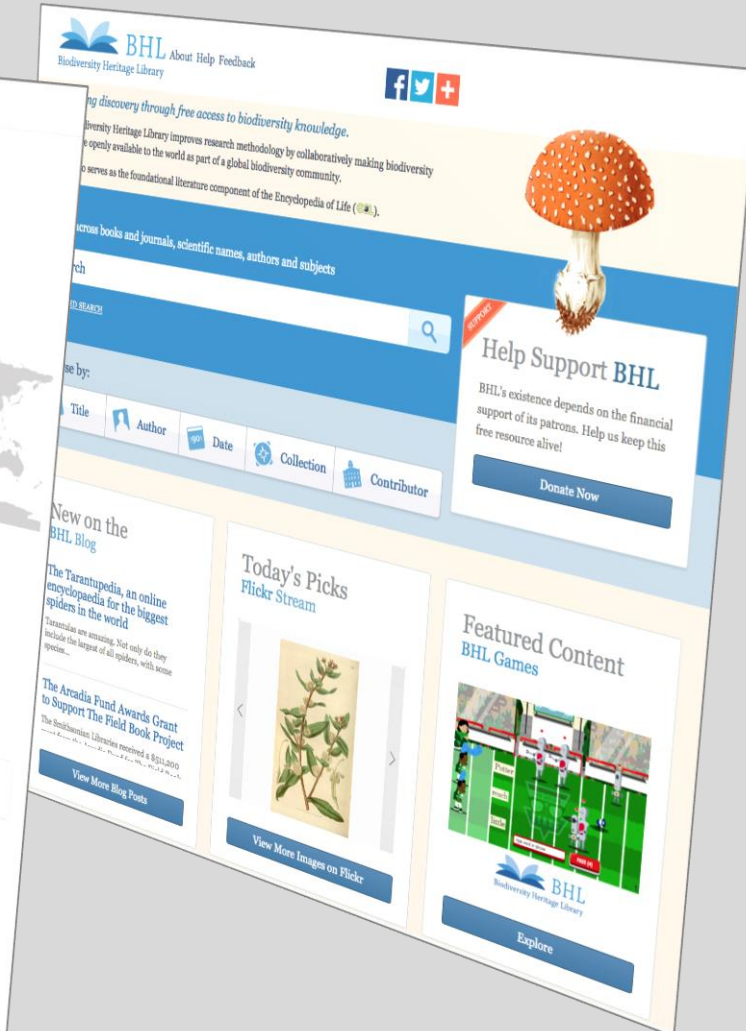
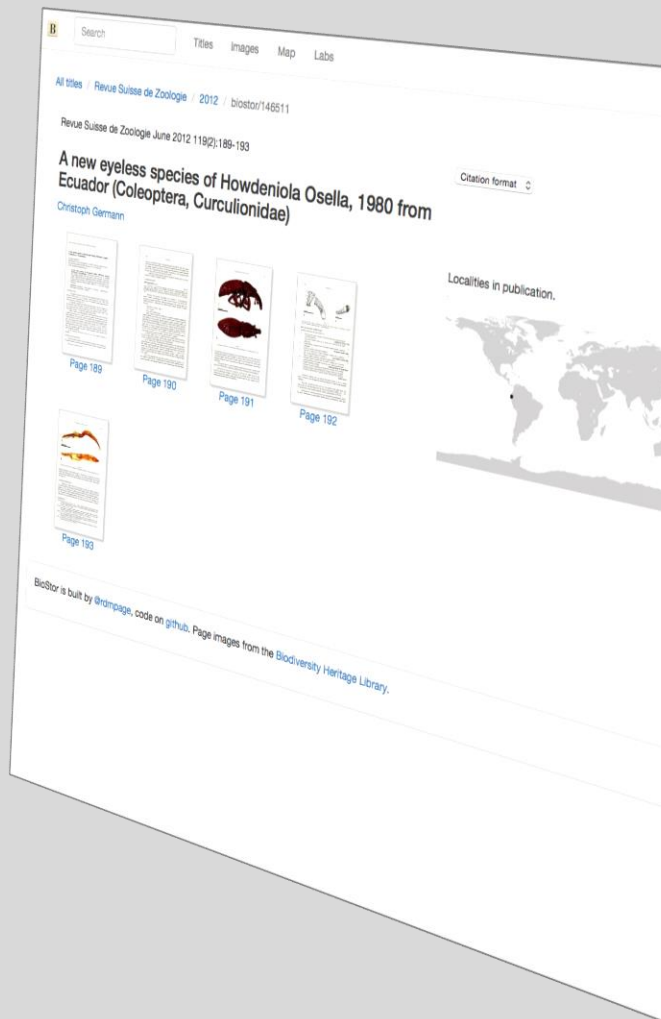
[https://www.flickr.com/photos/nhm\\_beetle\\_id/15930177695](https://www.flickr.com/photos/nhm_beetle_id/15930177695)



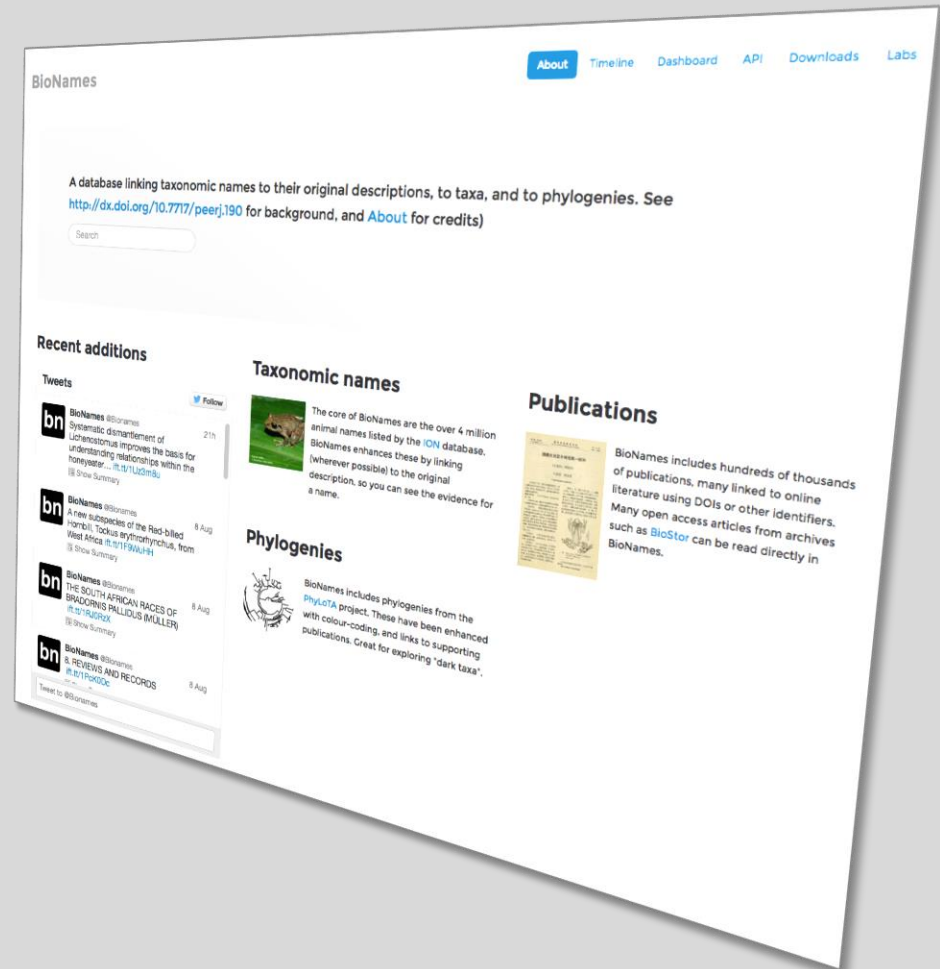


GBIF.org 500 million records

# Digitise the literature



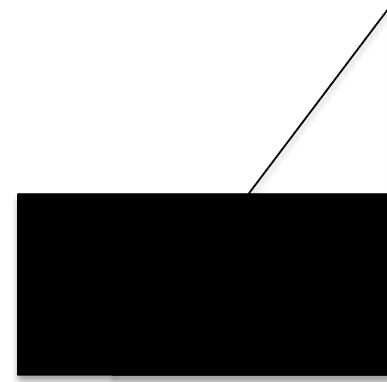
- Linking taxonomic names to published literature
- DOIs, BHL, BioStor, PDFs, Gallica, etc.
- 4 million names
- 400K publications





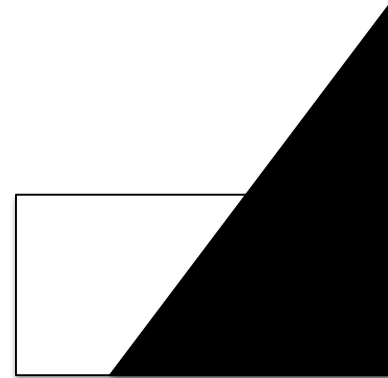
# Response 1: Modernise taxonomy

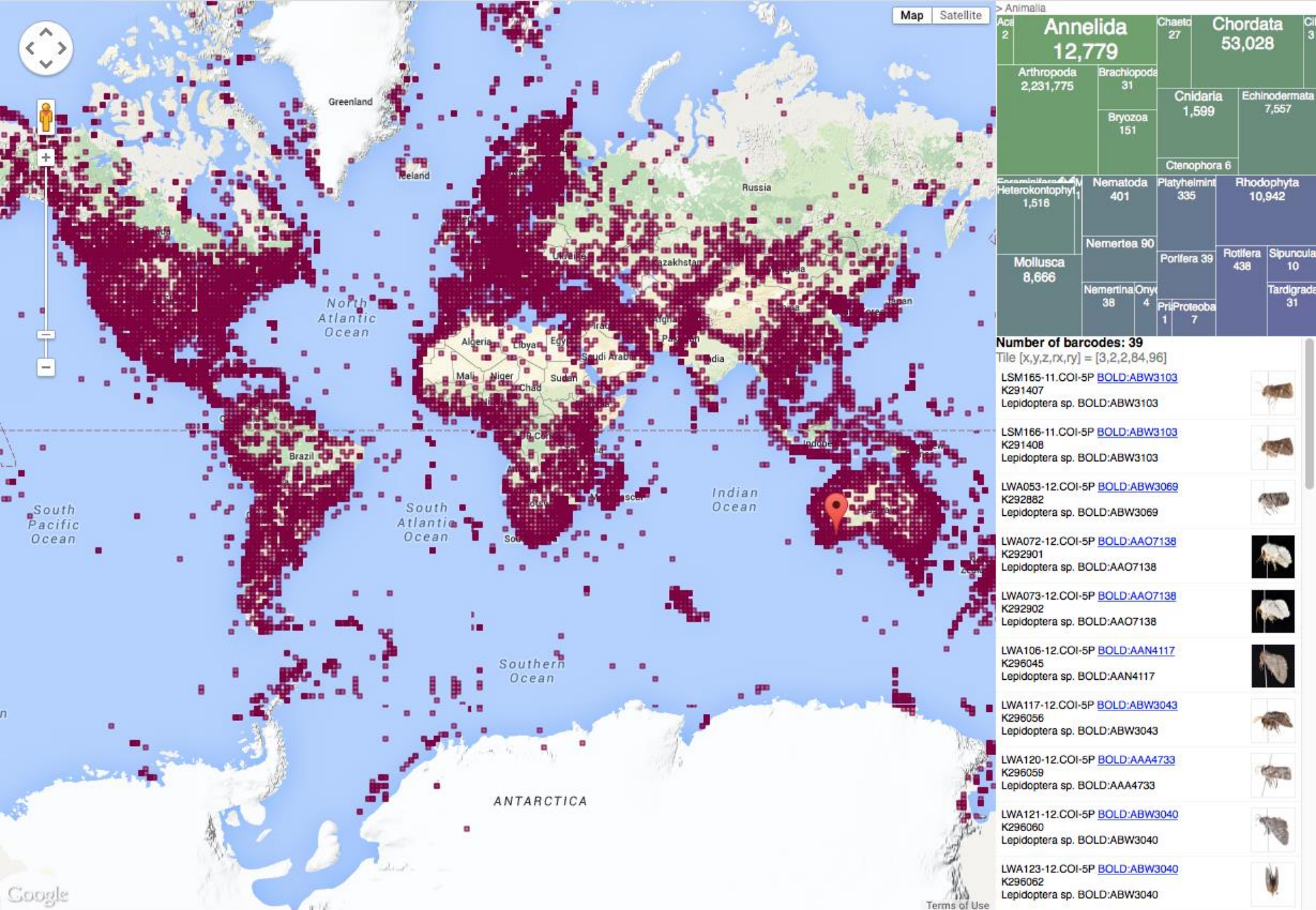
- Barcoding is digitising life (organism to string of letters)
- Mass digitisation of specimens and literature, etc.
- Cross-links between data sources



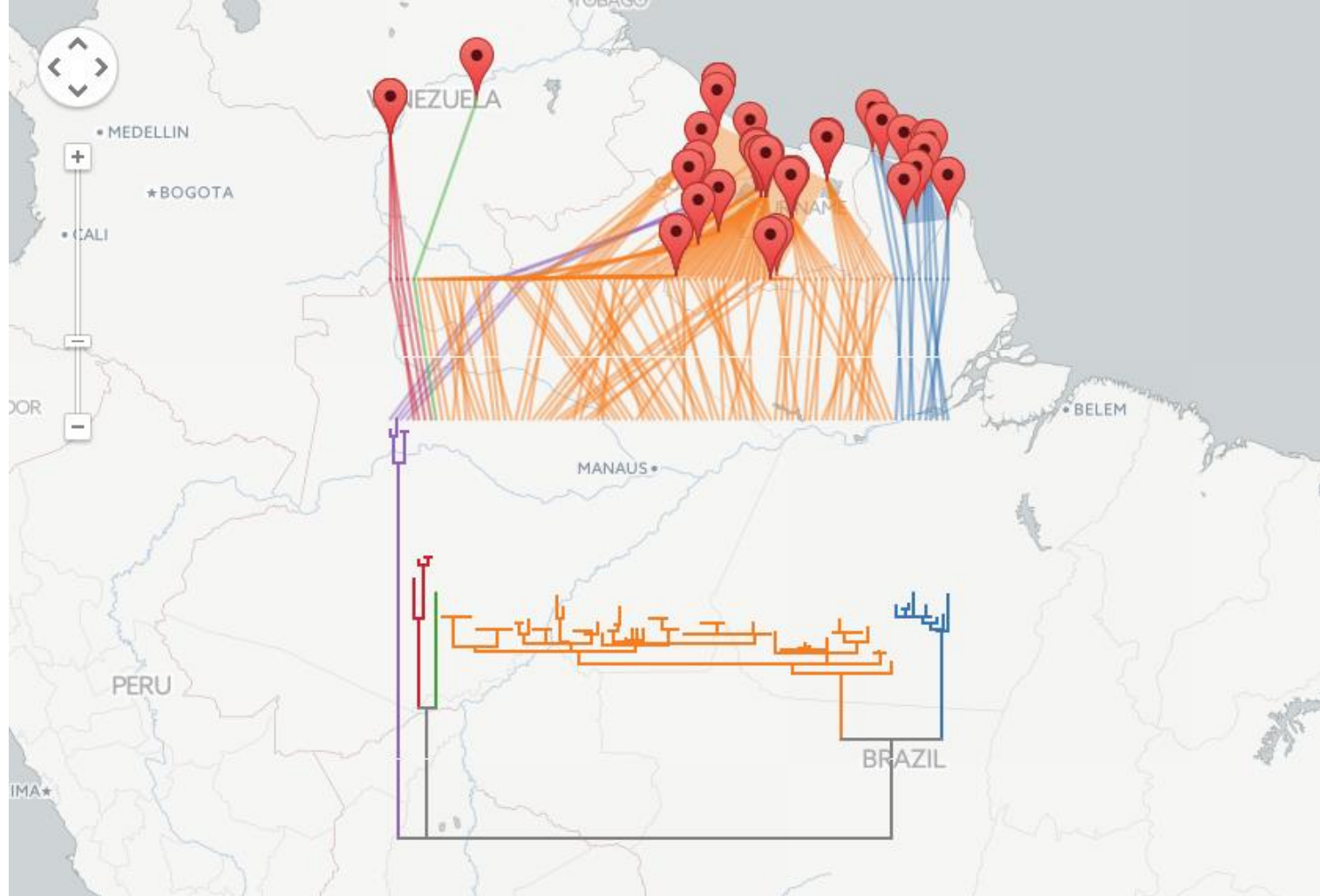
# Response 2: Barcode-only world

- Imagine barcodes were all we had (i.e., we're microbiologists)
- How would we visualise biodiversity?





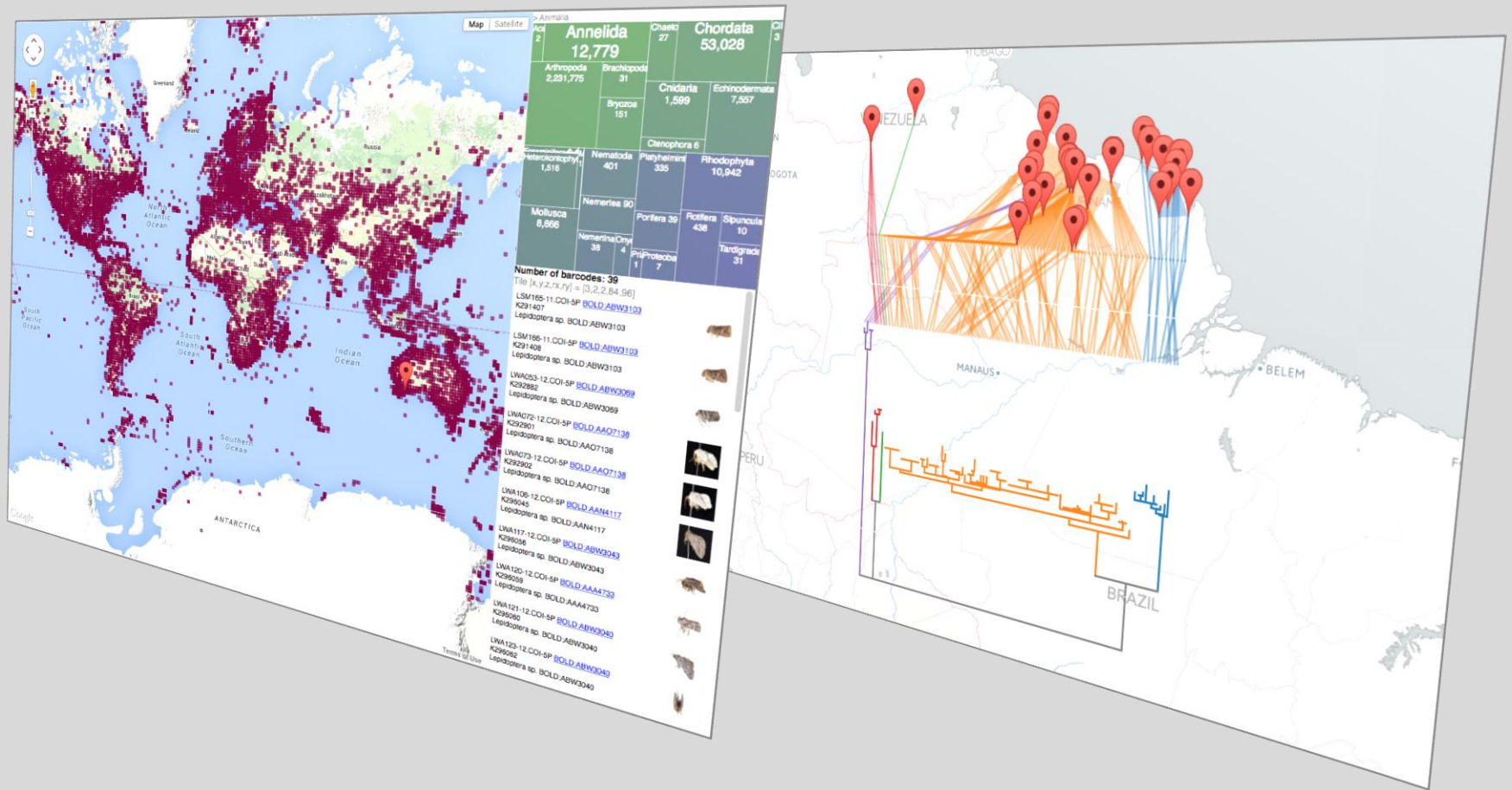
<http://iphylo.org/~rpage/bold-map>



Page R. **Visualising Geophylogenies in Web Maps Using GeoJSON.** PLOS Currents Tree of Life. 2015 Jun 23 . Edition 1. doi: 10.1371/currents.tol.8f3c6526c49b136b98ec28e00b570a1e



# These are (too) simple



# Challenges

- Can we build a complete tree of barcodes?
- Can we (quickly) compute sequence diversity for an area?
- Can we visualise ecological communities?

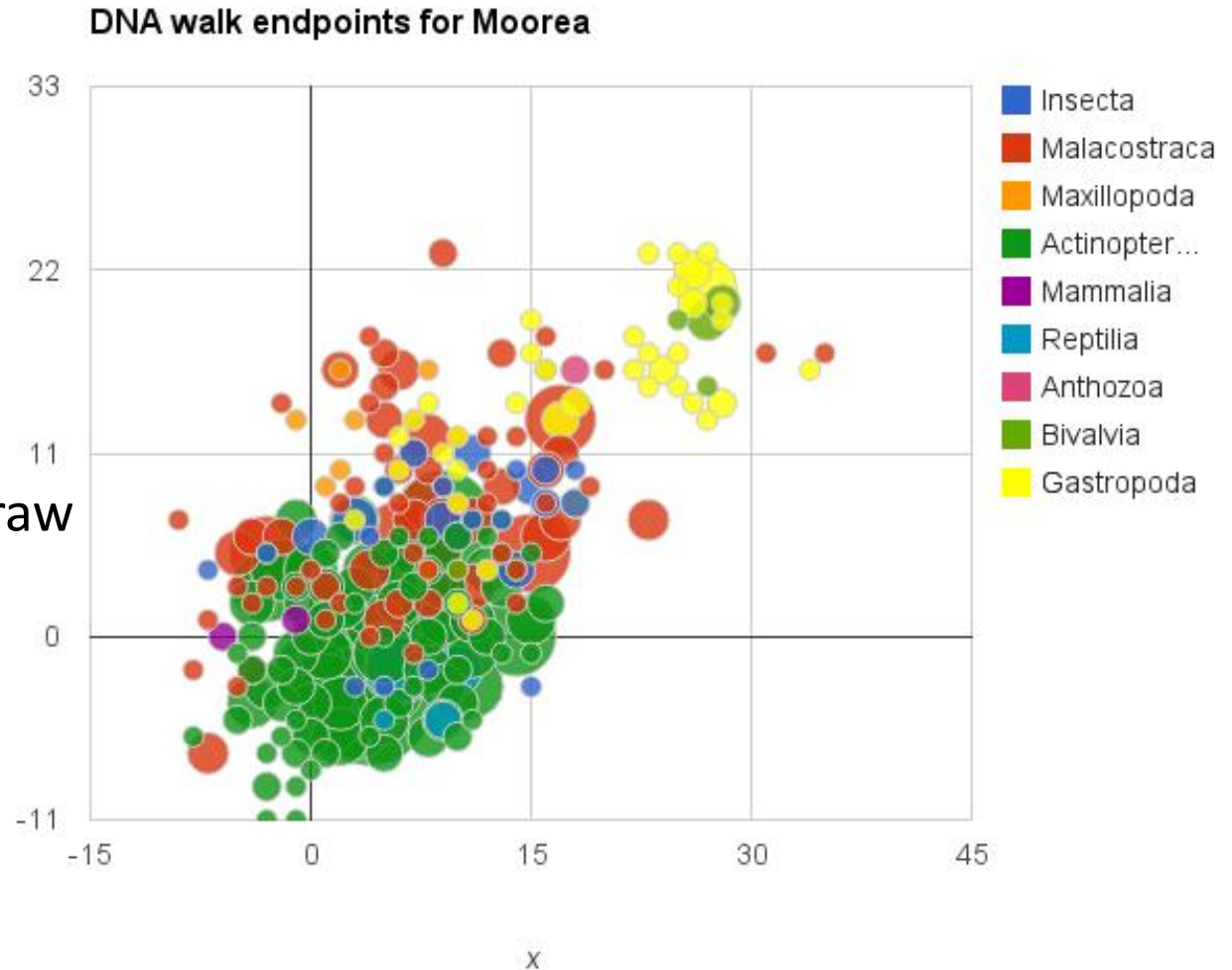
# DNA walks

Start at (0,0) and first position in sequence

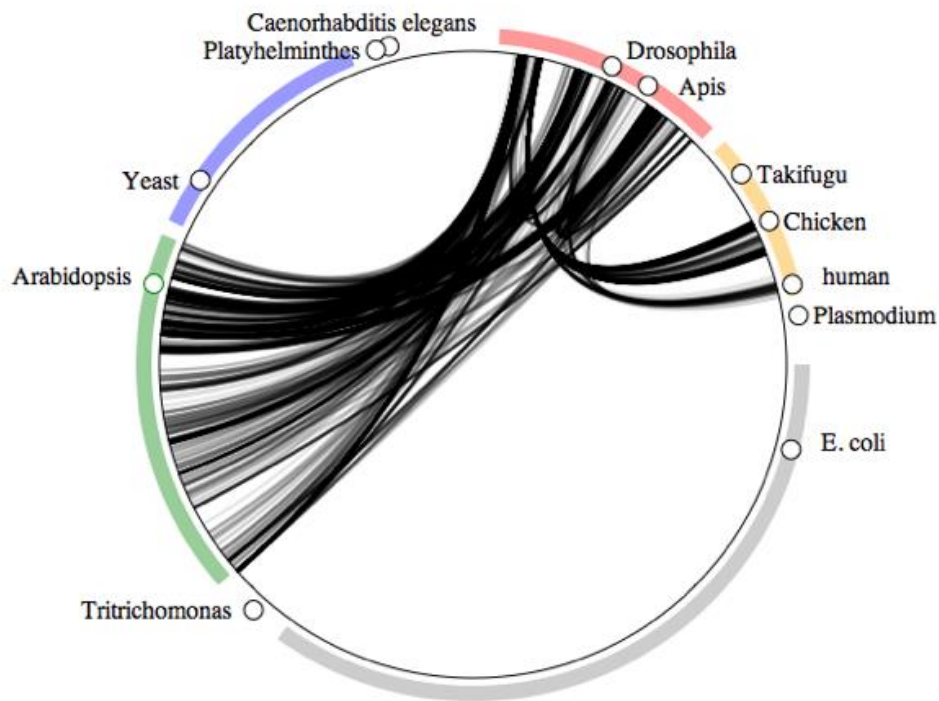
Move in x,y plane according to whether next base is different

At end of sequence draw point

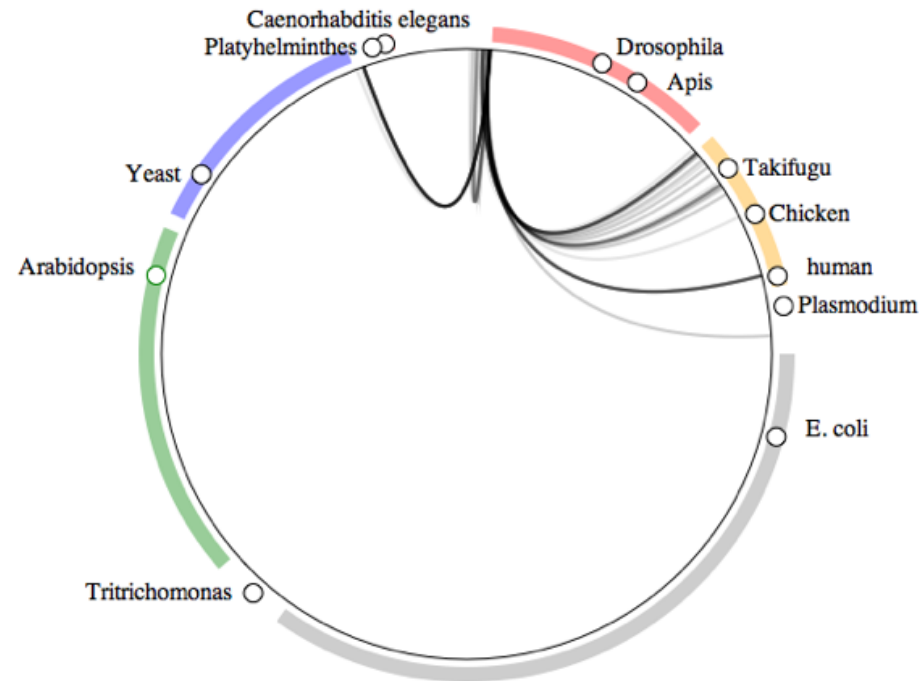
Plot shows  
DNA mini-barcodes  
(127 bp)



# “Symbiome”



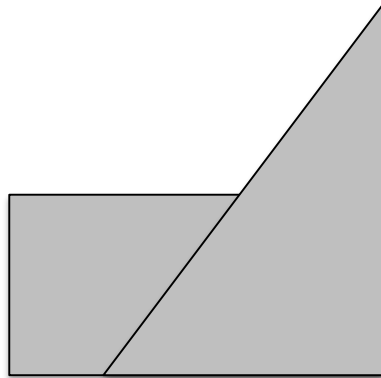
Insects



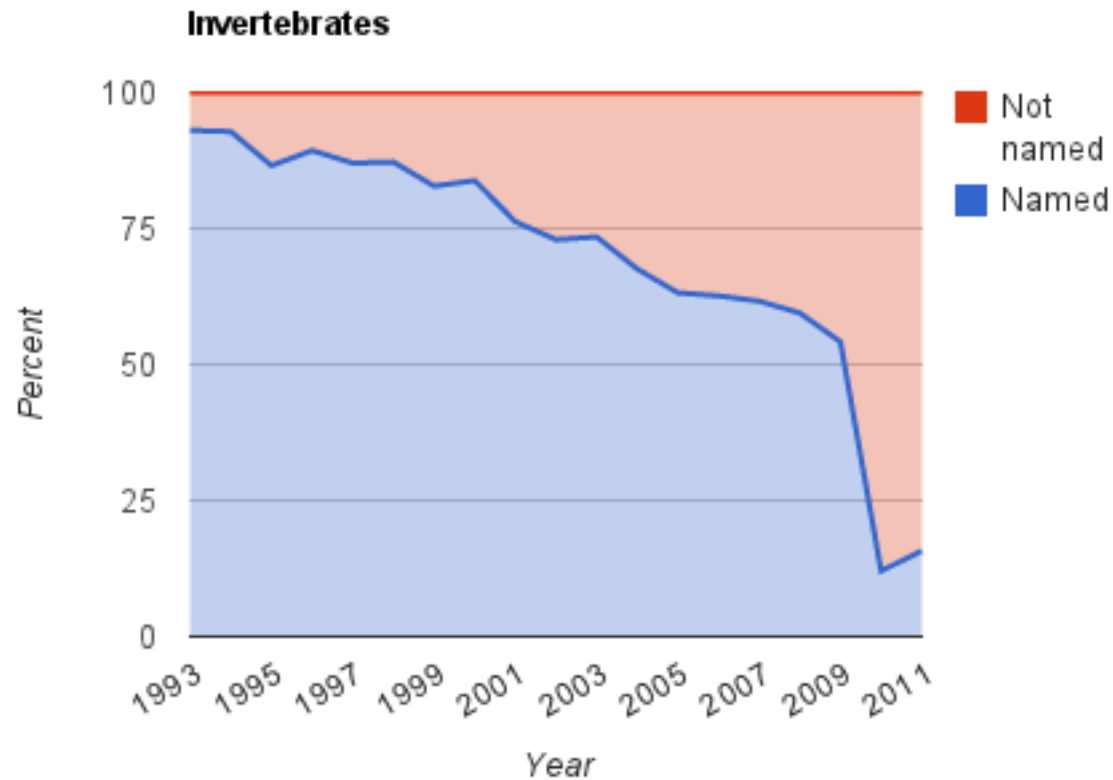
Crustacea

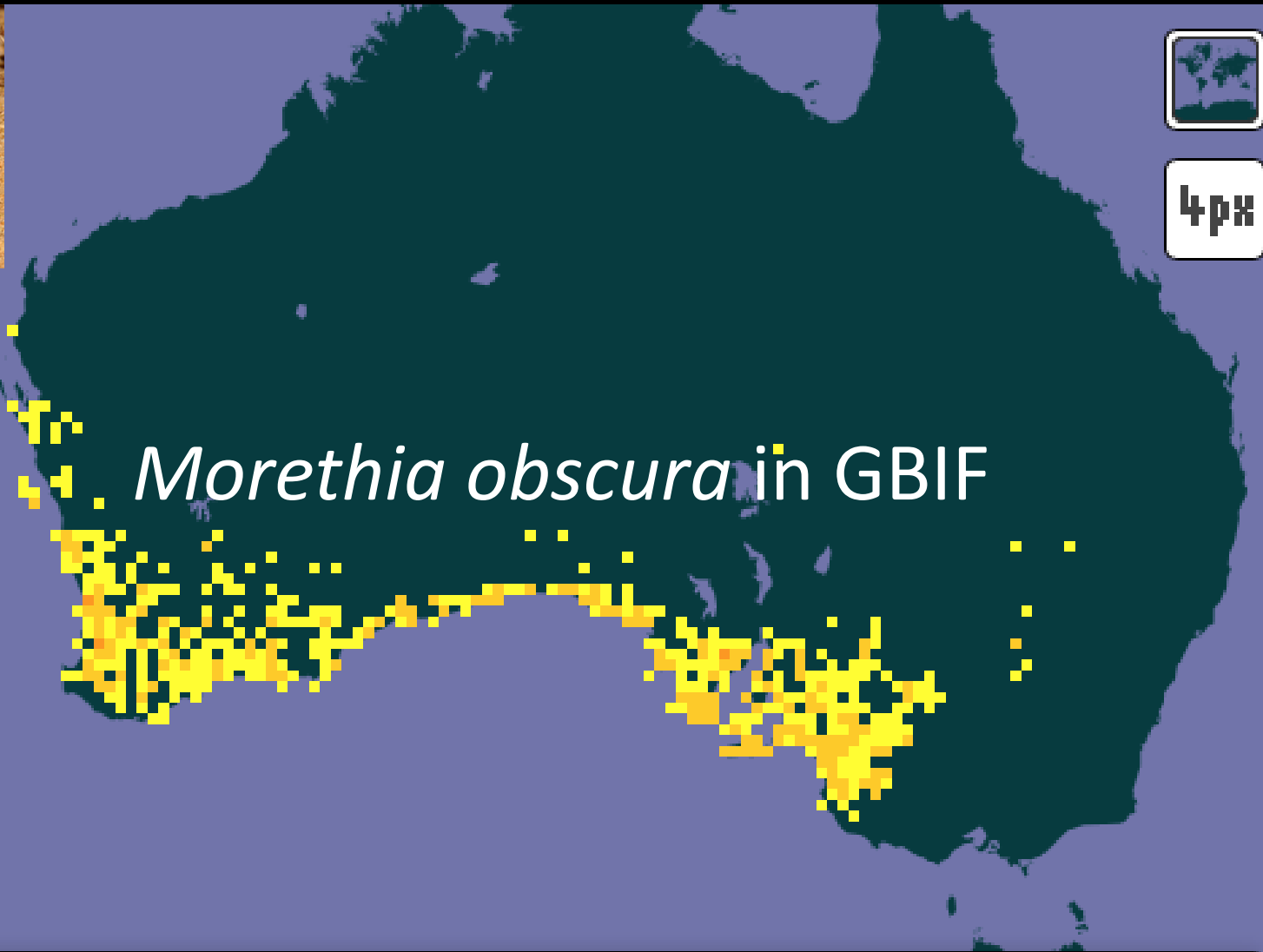


# Response 3: Link taxonomy and barcodes



# Integrate using taxonomic names?



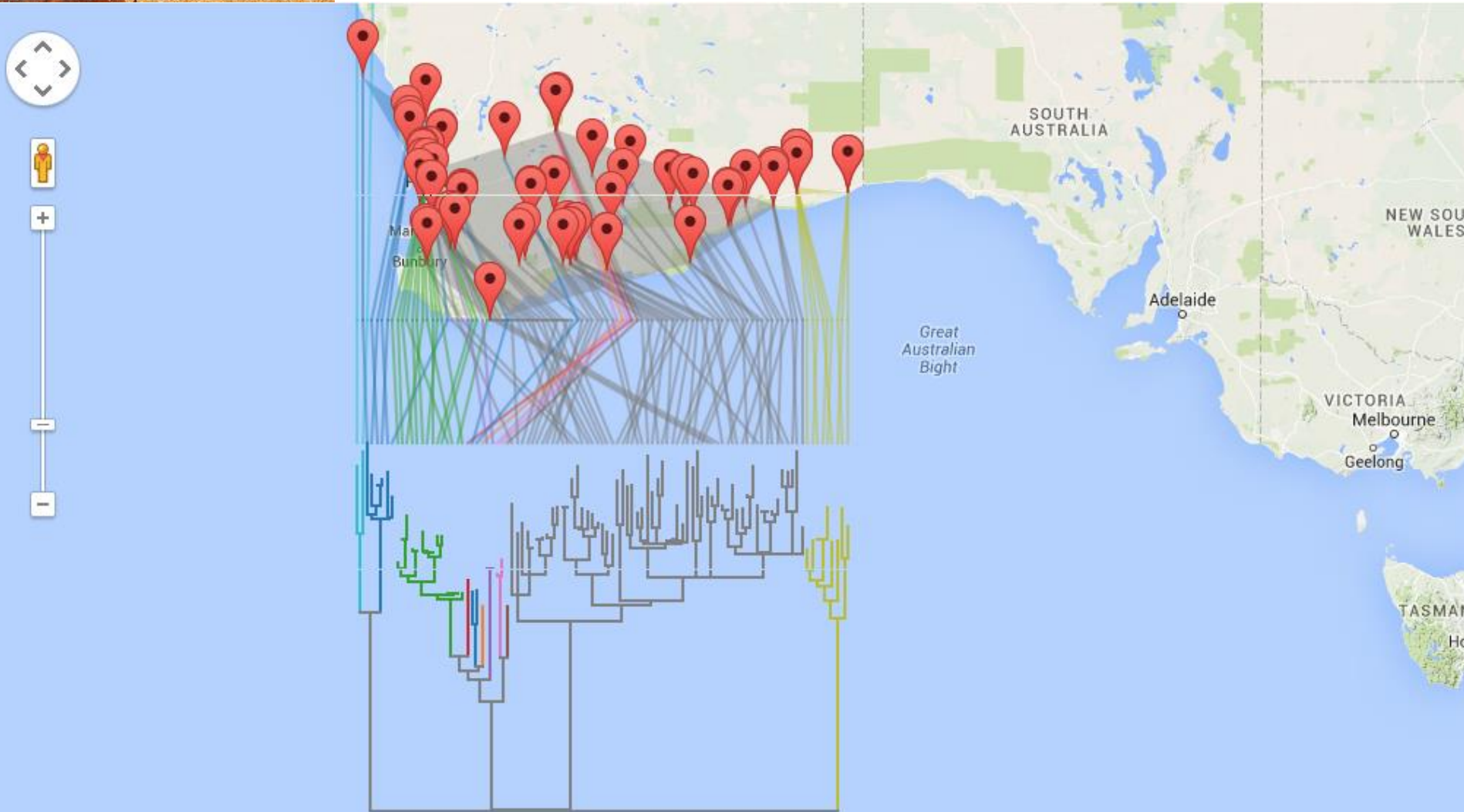


*Morethia obscura* in GBIF





# BINs for *Morethia obscura* in BOLD



# Integrate using specimens?

- Barcodes have vouchers
- GBIF has specimens
- Why isn't BOLD in GBIF?
- Actually, parts of it already are...

## Summary

### FULL TITLE

Zoologische Staatssammlung Muenchen - International Barcode of Life (iBOL) - Barcode of Life Project Specimen Data

### LANGUAGE OF METADATA

ENGLISH

**DOI** [doi:10.15468/tfpnkp](https://doi.org/10.15468/tfpnkp)

### PUBLISHED BY

[Zoologische Staatssammlung  
München/Staatliche  
Naturwissenschaftliche  
Sammlungen Bayerns](#)

### REGISTRATION DATE

Dec 17, 2013

### HOSTED BY

[Museum für Naturkunde Berlin](#)

### SERVED BY

[BioCASE Installation Museum für  
Naturkunde Berlin](#)

<http://doi.org/10.15468/tfpnkp>

# BOLD samples in GBIF (twice)

(but provided by ZSM and EMBL, **not** BOLD)

## GBIF

BC ZSM Lep 10234  
GBIF 883514761  
*Casbia rectaria* Walker, 1866

## BOLD

GWORH520-09  
BC ZSM Lep 10234  
BOLD:AAA4623  
*Casbia rectaria*



## GBIF

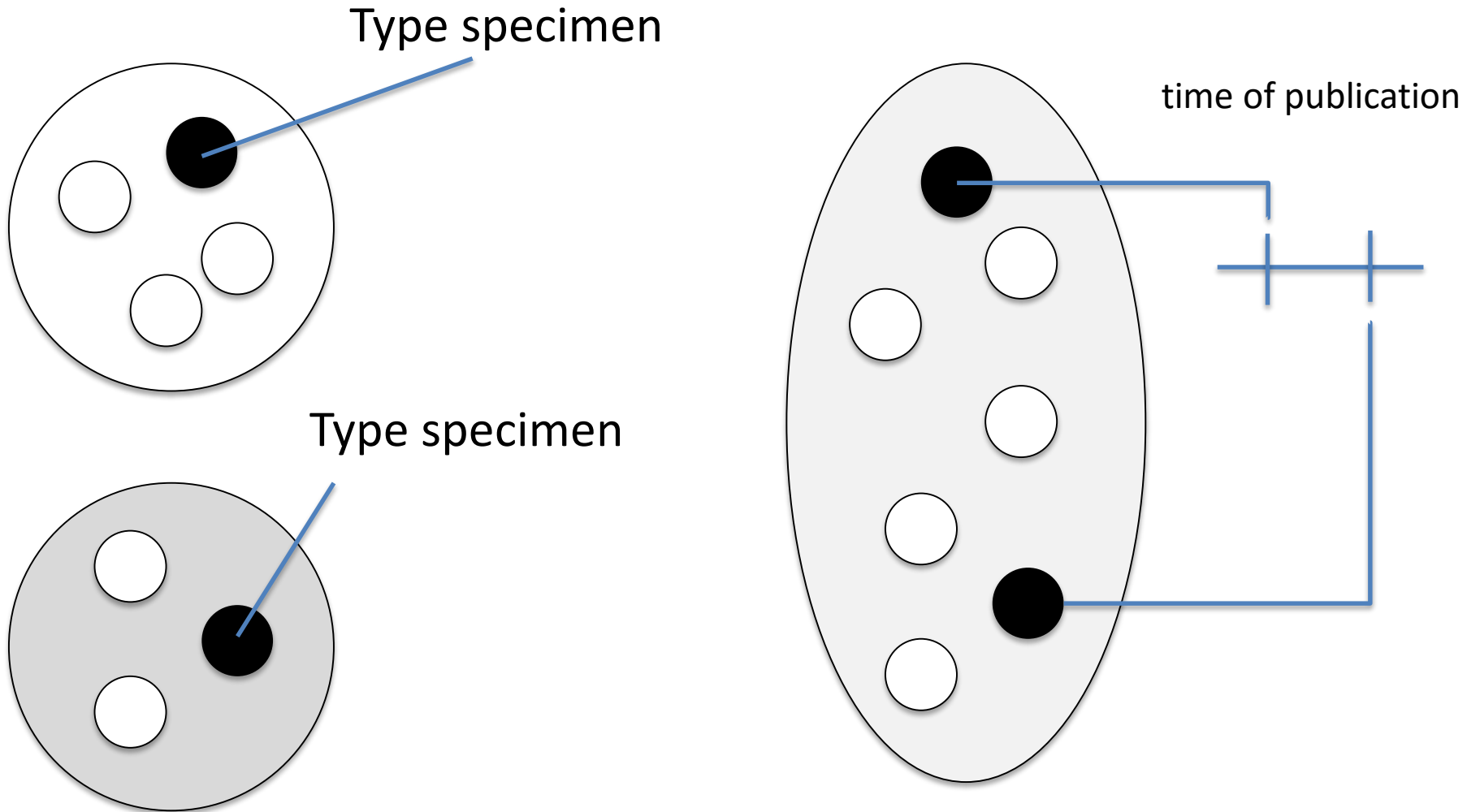
GU655831  
GBIF 1080492017  
**Lepidoptera**

7bcace52-85a0-4c5c-991e-555d4479e42c

## EMBL

GU655831  
BC ZSM Lep 10234  
**Lepidoptera sp. BOLD:AAA4623**

# The role of type specimens







# Search occurrences

Use the filters to customize search results

6,648

Occurrences

Download

6,648 results

11,664 bird holotype specimens:

For 6,648 GBIF doesn't understand the name

Type status: Holotype

## Aves

Published in Natural History Museum (London)  
Collection Specimens

Myanmar

N/A

Specimen

1 / 2012

1057526476

Type status: Holotype

## Cinnyris cupreus

Published in Natural History Museum (London)  
Collection Specimens

Uganda

N/A

Specimen

6 / 2012

1056117638 · Cat. 1936.4.12.3039

Type status: Holotype

## Bhringa Hodgson, 1837

Published in Natural History Museum (London)  
Collection Specimens

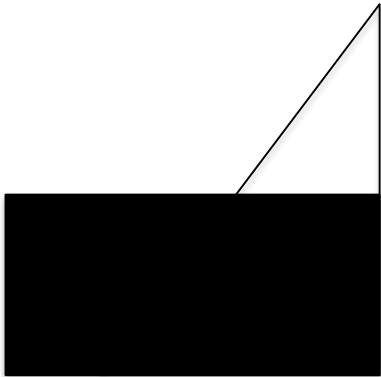
N/A

Specimen

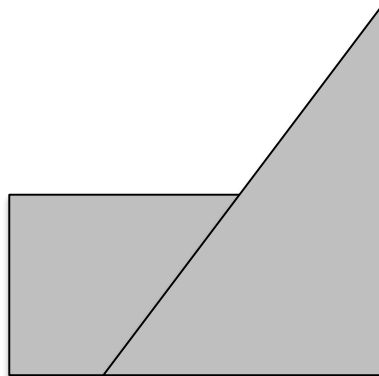
10 / 2009

# Three responses

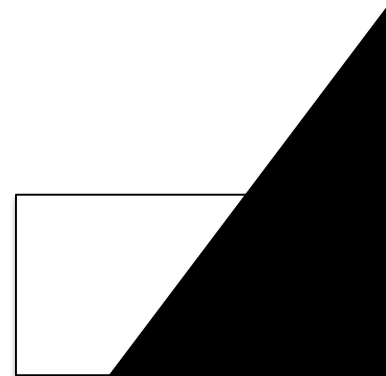
Make  
taxonomy  
digital



Integrate  
taxonomy &  
sequences



Sequences eat  
the world



Number of names (5 year intervals)

