

Towards holistic insect monitoring by accelerating species discovery, description, and identification using robots, nanopore sequencing, and Artificial Intelligence

Biodiversity research: 1990s versus 2020s

1990s: “Ivory Tower” discipline for biologists out of touch with modernity

2020s: Biodiversity loss threatens planetary health and society now cares

Who cares?

The Global Risks Report 2023 18th Edition



- 1 Failure to mitigate climate change
- 2 Failure of climate-change adaptation
- 3 Natural disasters and extreme weather events
- 4 Biodiversity loss and ecosystem collapse
- 5 Large-scale involuntary migration

Who else?

Goldman Sachs


GS SUSTAIN

Biodiversity


Assessing the Financial Links to Natural Capital

Biodiversity is the cornerstone around which most sustainability objectives sit and an area to which we think corporates and investors will likely pay increased attention in order to meet sustainability and economic

Re/insurance and Biodiversity: Insights on a Swiss Re Report




Corvus Centre for Conservation Policy

Allianz 

Allianz Trade

The new risk frontier in finance:
biodiversity loss

[DISCOVER OUR REPORT](#)




But is society interested in “Ivory Tower biodiversity”?


Forbes

FORBES > LEADERSHIP > LEADERSHIP STRATEGY


Ugly Species Deserve Biodiversity Protections, Too

World Economic Forum Contributor 

Nov 2, 2020, 04:55am EST

 Listen to article 9 minutes 

f
t
in



Hello. PHOTOHOLGIC/UNSPLASH

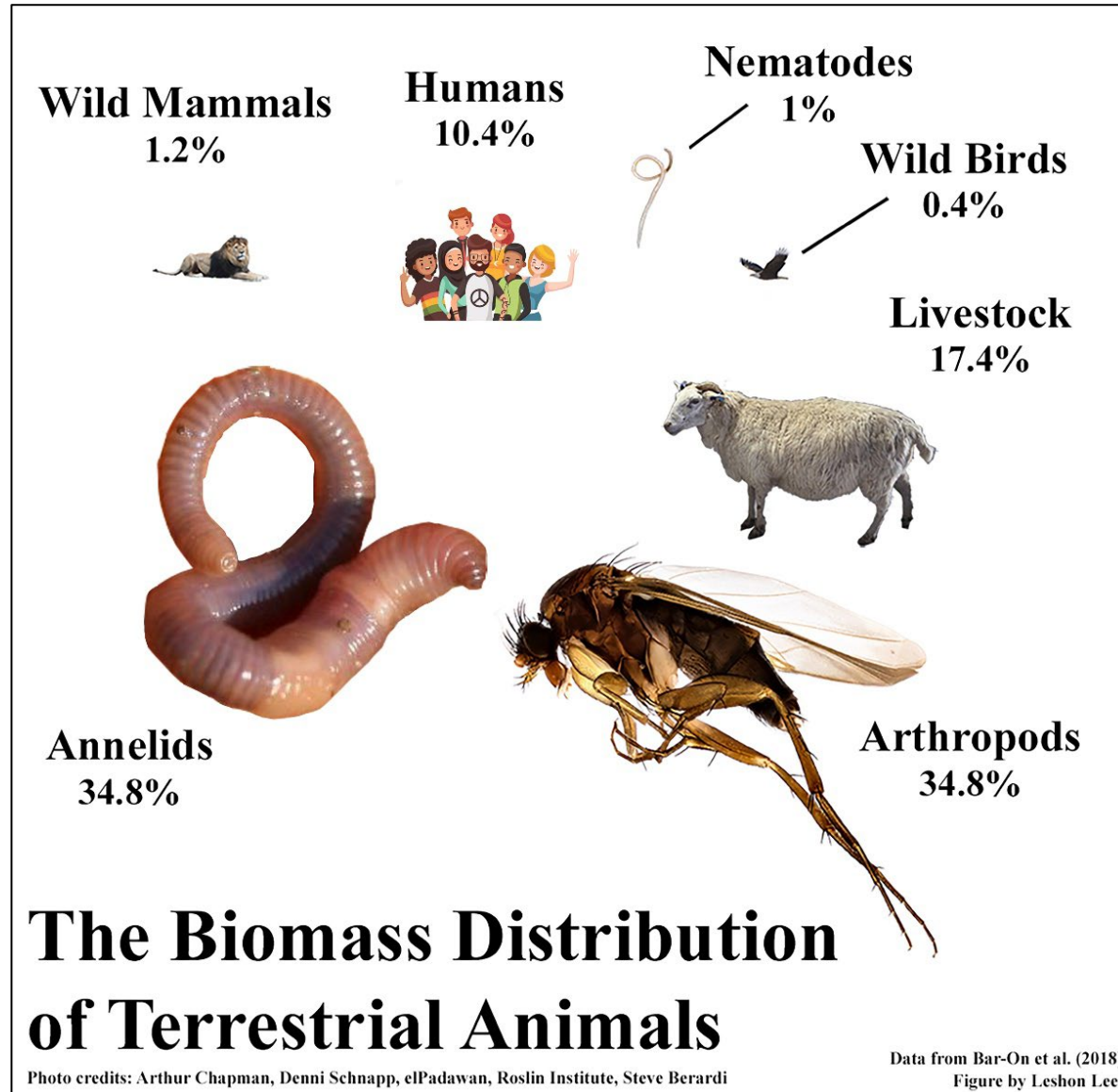
1990s: Ivory Tower taxon biases:

- pretty
- pretty large
- pretty diurnal
- pretty endangered
- pretty impressive on a tree
- pretty manageable taxonomically
- pretty unusual in many other ways...

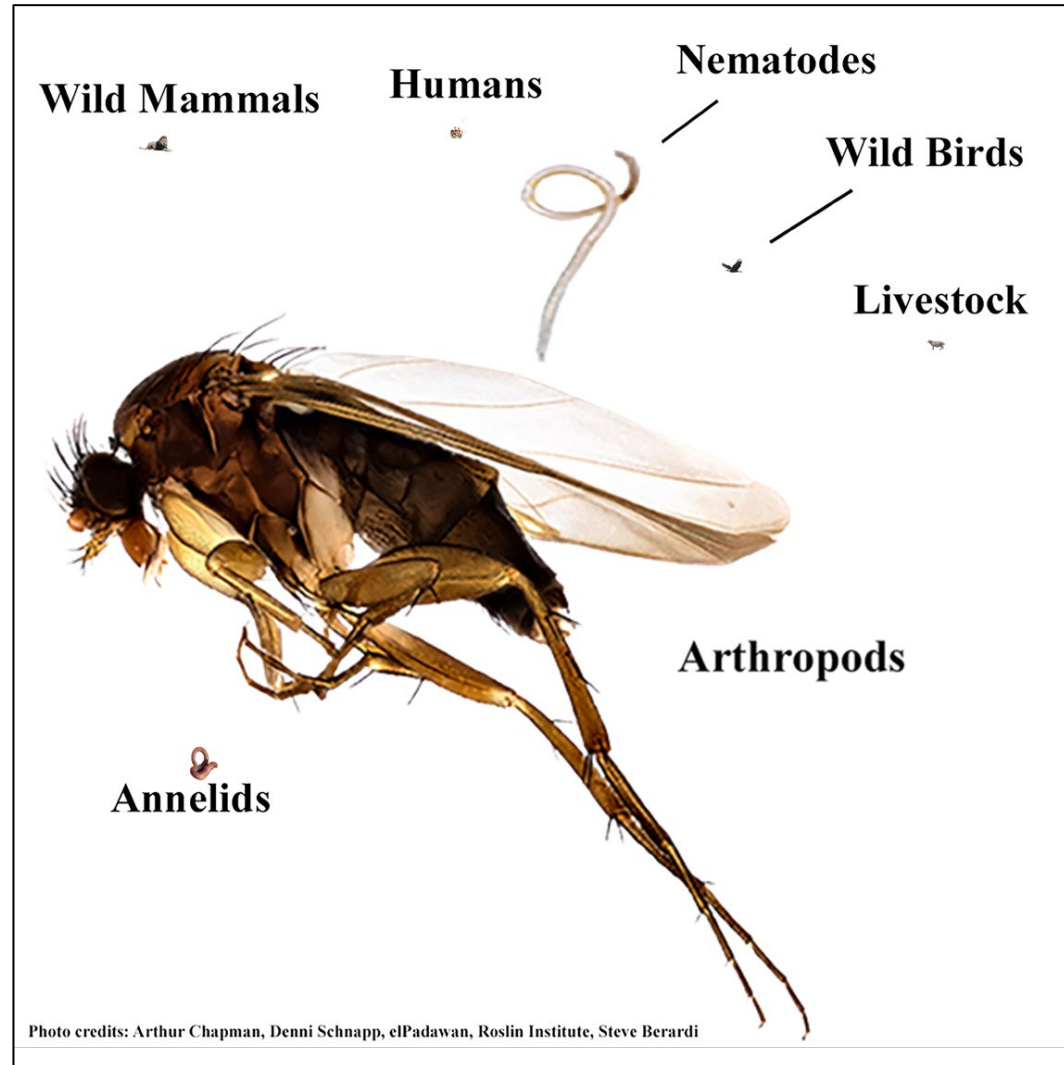
What is missing is:

- pretty important but ugly...

Terrestrial animal biodiversity: biomass



Terrestrial animal biodiversity: species diversity



Shortcomings of the old ways...

- >80% of all species undescribed
- 95% of all described species only identifiable for taxonomic experts
- very little baseline biodiversity data from standardized samples
- natural history data only available for <5% of species

Triage: which taxa need our attention?



Dr. Amrita Srivathsan



Convergence of dominance and neglect in flying insect diversity

Received: 8 August 2022

Accepted: 6 April 2023

Published online: 18 May 2023

 Check for updates

Amrita Srivathsan¹, Yuchen Ang², John M. Heraty³, Wei Song Hwang², Wan F. A. Jusoh^{2,4}, Sujatha Narayanan Kutty^{5,6}, Jayanthi Puniemoorthy⁵, Darren Yeo⁵, Tomas Roslin⁷ & Rudolf Meier^{1,5} ✉

Most of arthropod biodiversity is unknown to science. Consequently, it has been unclear whether insect communities around the world are dominated by the same or different taxa. This question can be answered through standardized sampling of biodiversity followed by estimation of species diversity and community composition with DNA barcodes. Here this approach is applied to flying insects sampled by 39 Malaise traps placed in five biogeographic regions, eight countries and numerous habitats (>225,000 specimens belonging to >25,000 species in 458 families). We find that 20 insect families (10 belonging to Diptera) account for >50% of local species diversity regardless of clade age, continent, climatic region and habitat type. Consistent differences in family-level dominance explain two-thirds of variation in community composition despite massive levels of species turnover, with most species (>97%) in the top 20 families encountered at a single site only. Alarming, the same families that dominate insect diversity are 'dark taxa' in that they suffer from extreme taxonomic neglect, with little signs of increasing activities in recent years. Taxonomic neglect tends to increase with diversity and decrease with body size. Identifying and tackling the diversity of 'dark taxa' with scalable techniques emerge as urgent priorities in biodiversity science.

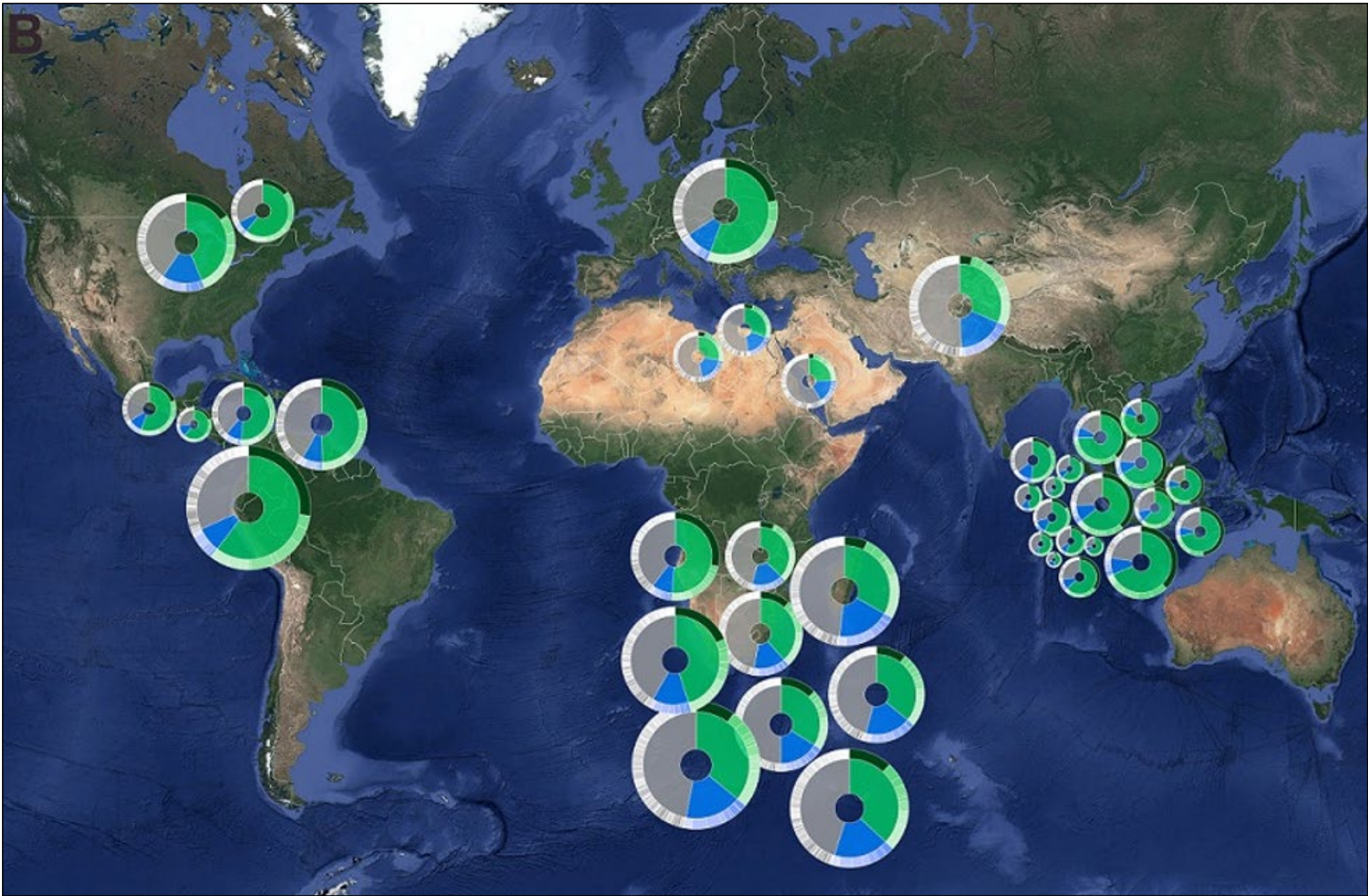


Dr. Amrita Srivathsan

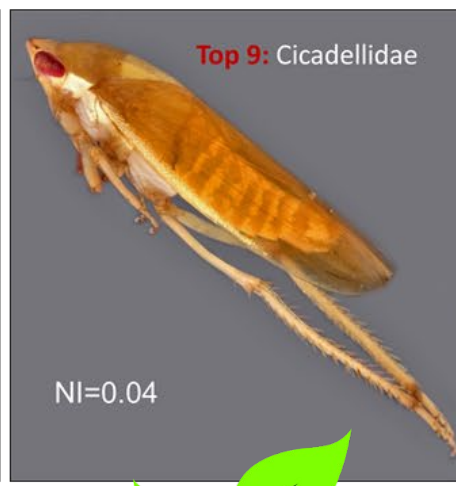
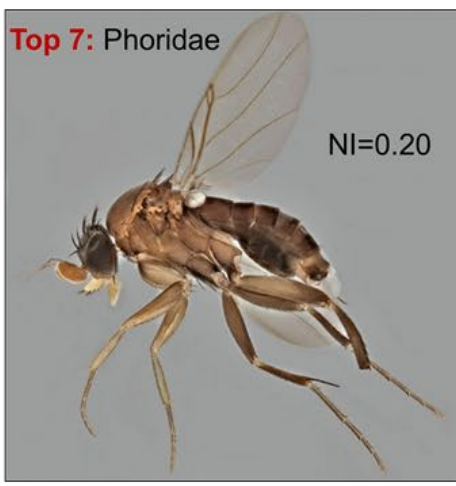
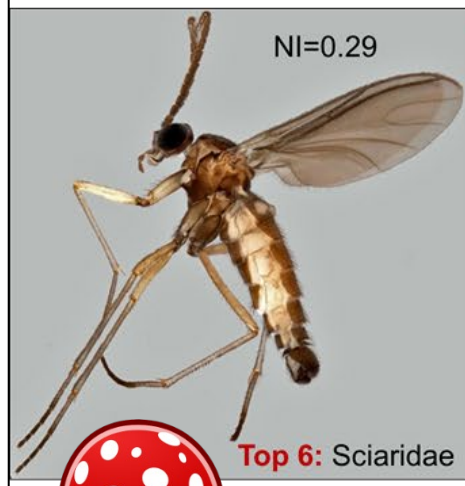
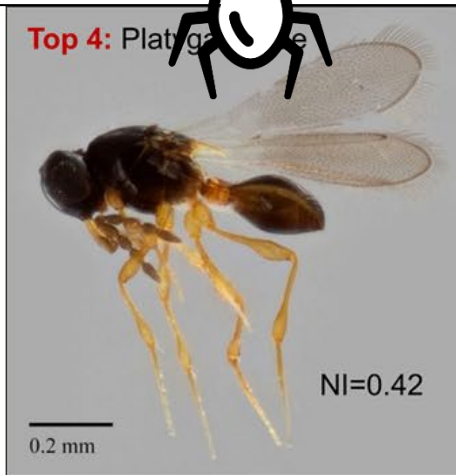
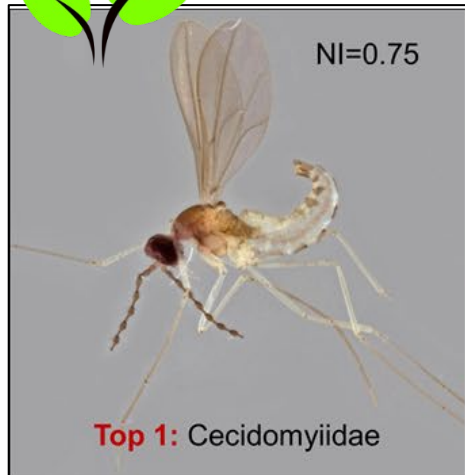
=> same 20 family-ranked insect clades contribute >50% of species diversity

=> same 20 clades also account for >50% of specimen abundance

Habitat, climate, etc. don't matter



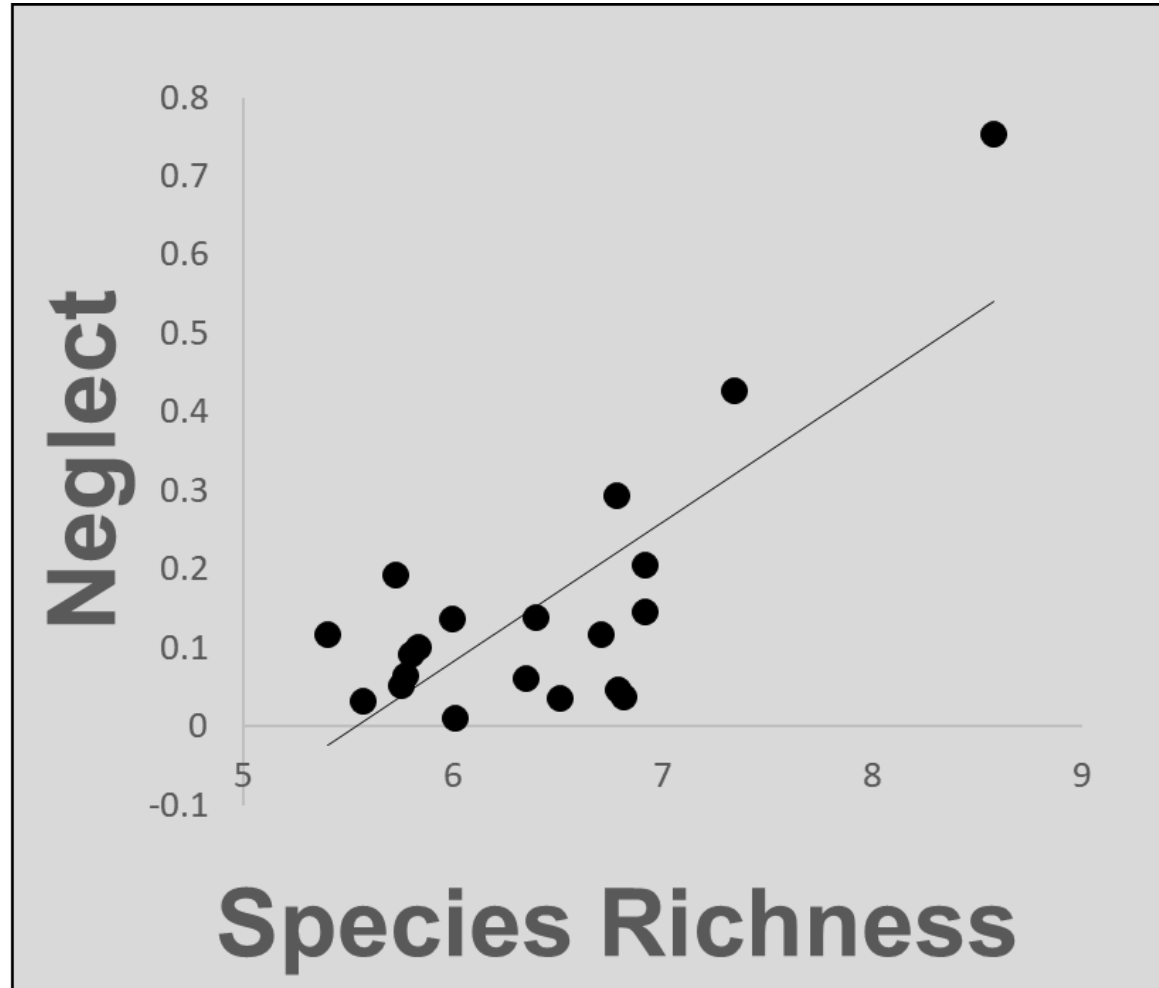
The top 10



Most of these top taxa suffer from taxonomic neglect



Dr. Amrita Srivathsan



‘Neglect Index’ (NI): ratio between the number of mOTUs found across Malaise traps and number of species described

1. Neglect increases with species richness
2. Neglect is not improving over time
3. Small body size leads to neglect

**Clearly, we need “Ugly Zoology”!
But we also need rebranding...**

“Dark Taxon Zoology”: resolving biology’s dark matter

Roderik Page in 2011: “dark taxa”

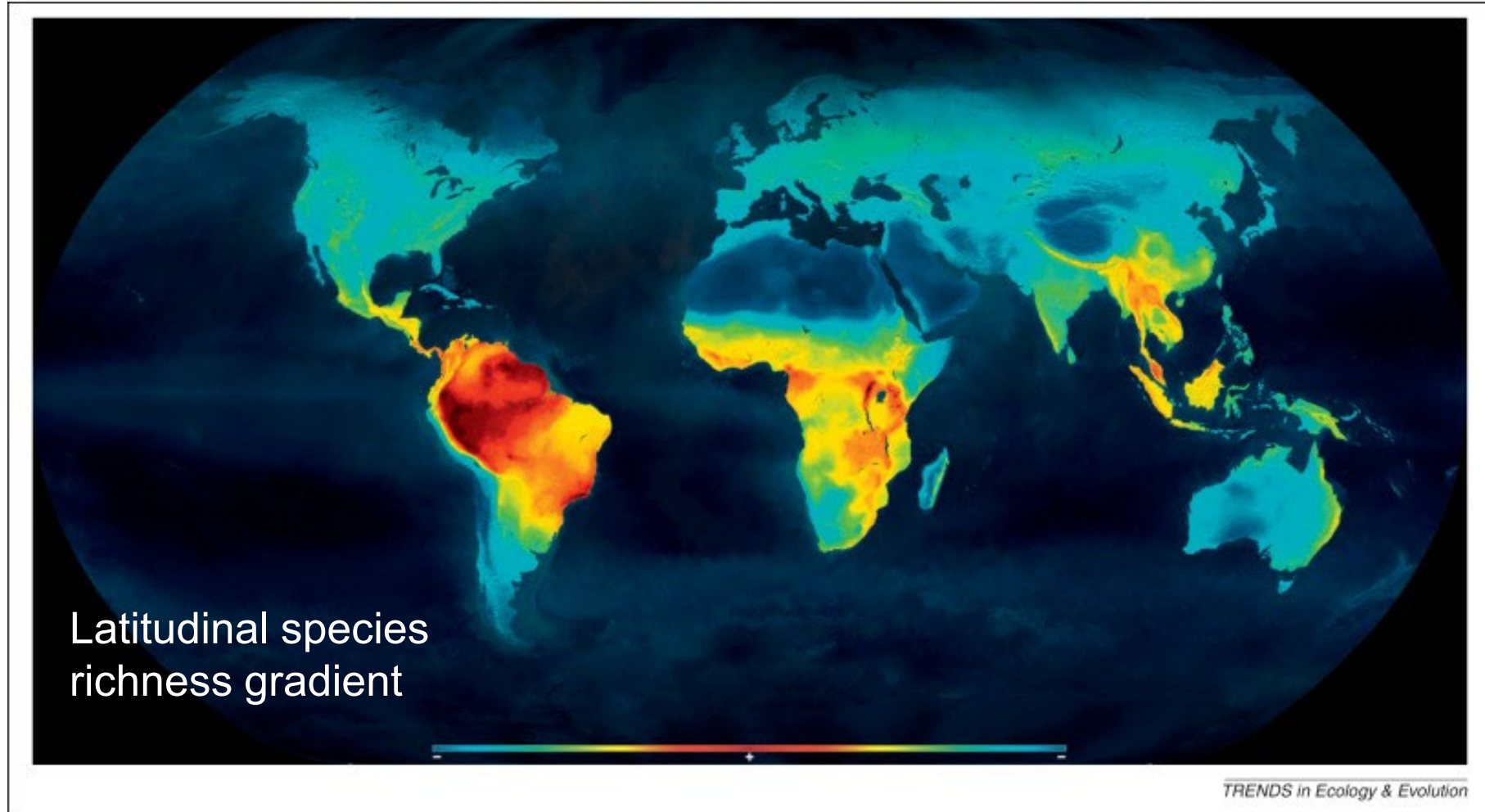
1. groups of organisms present in genetic databases
2. lacking formal scientific names

Hartop et al. 2022: “dark taxon”

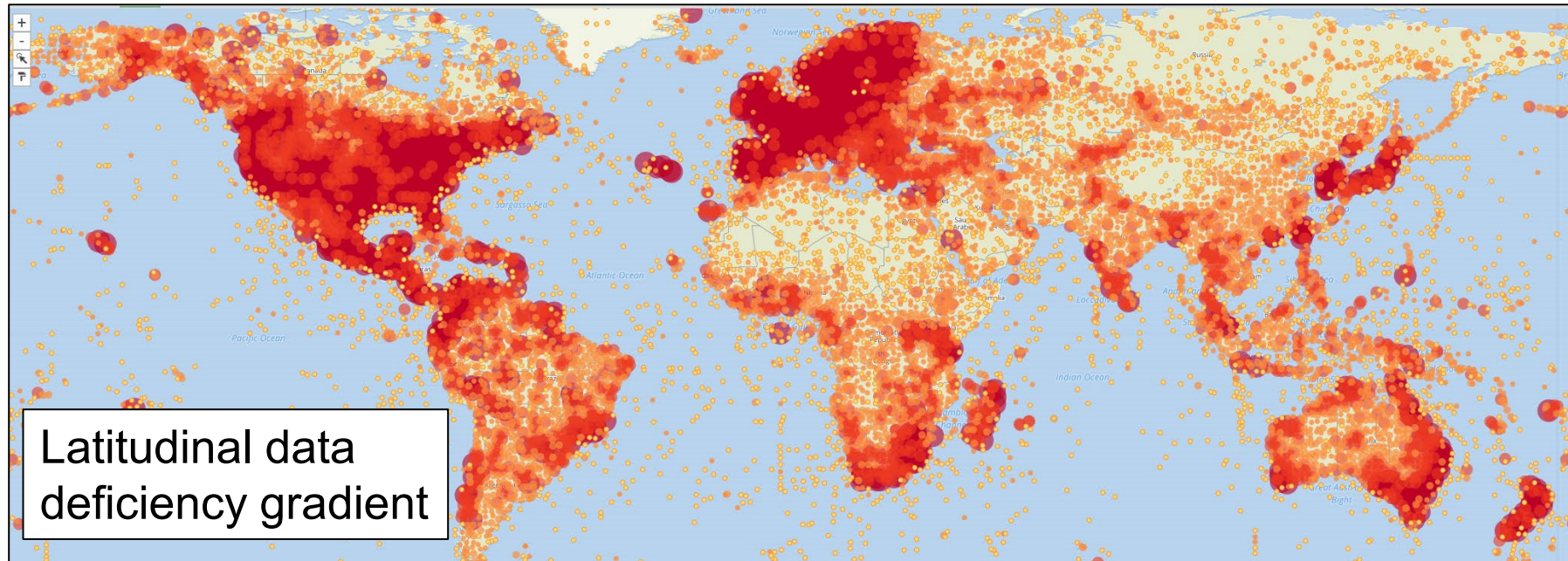
1. clade whose diversity exceeds 1,000 species
2. of which fewer than 10% of species described

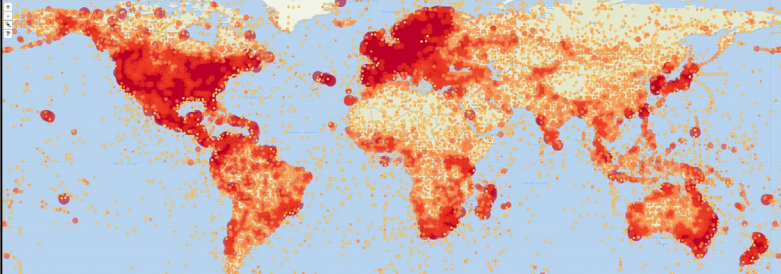
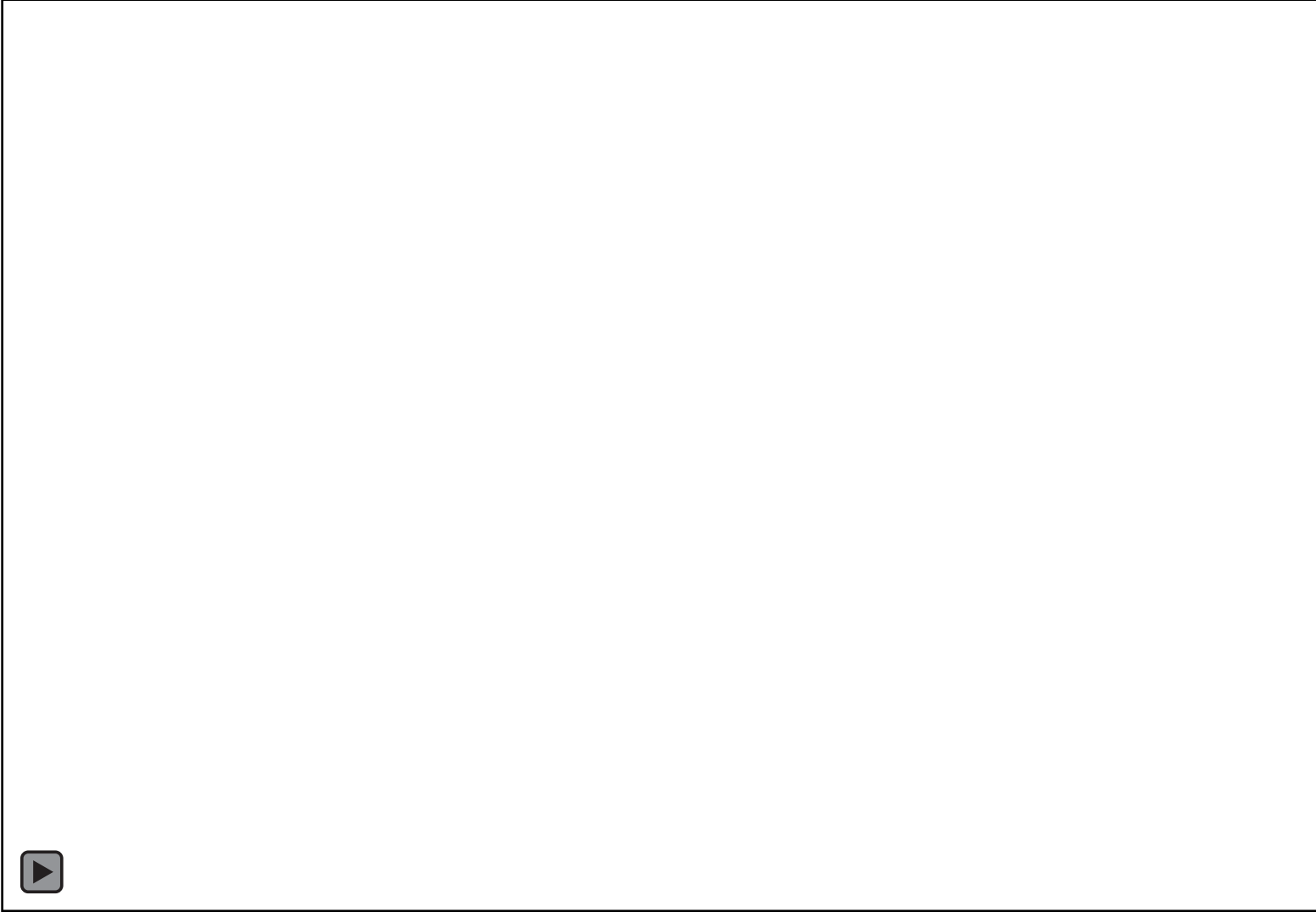
Biodiversity data crisis: more problems...

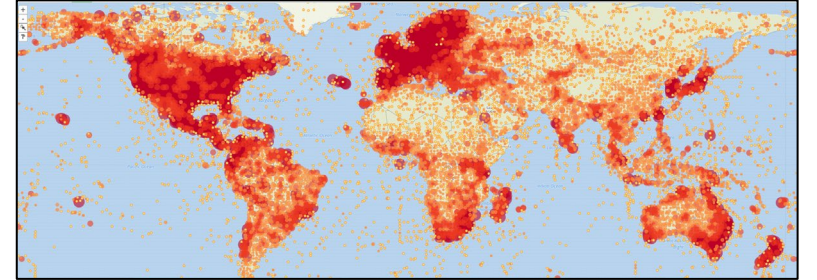
Geographic biases in biodiversity data



Insect data in GBIF

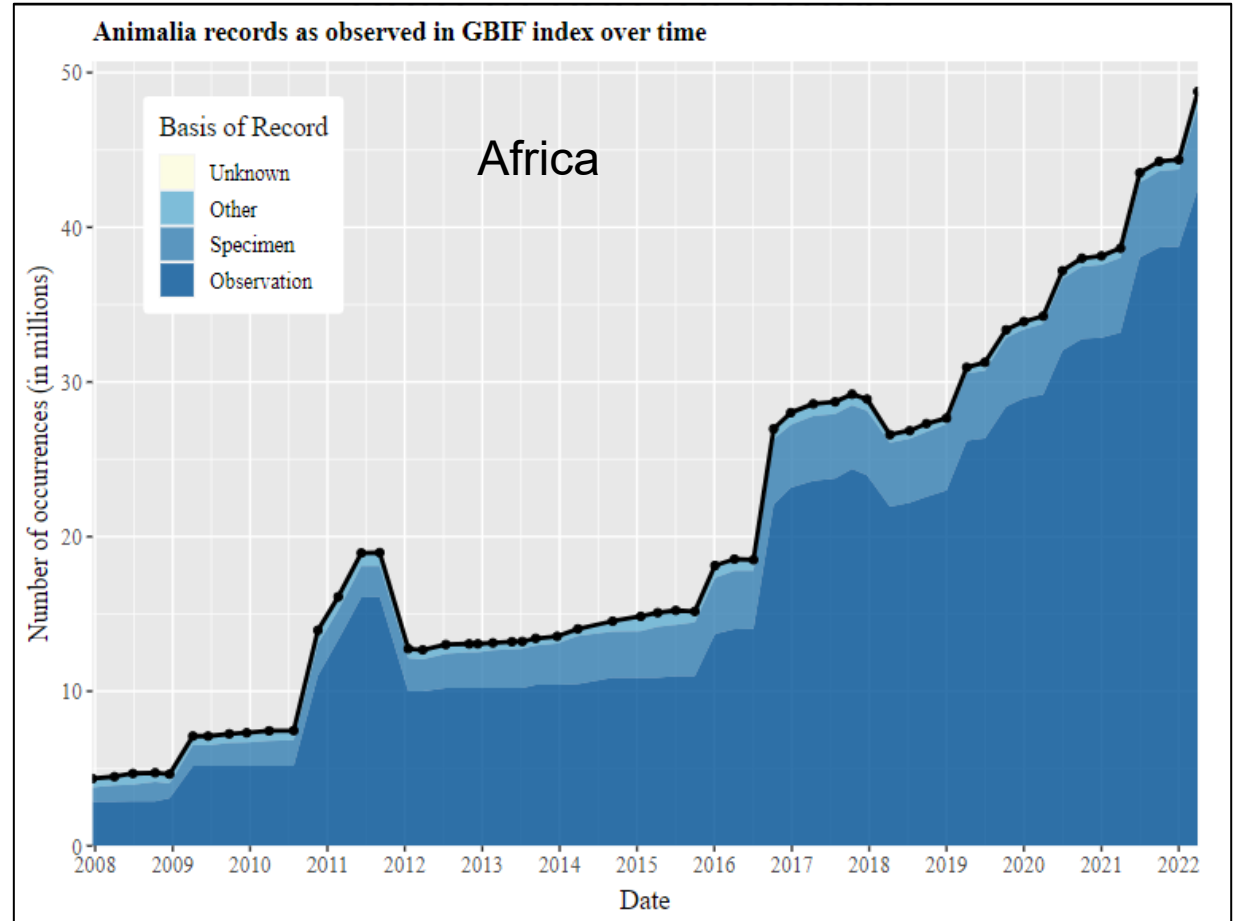
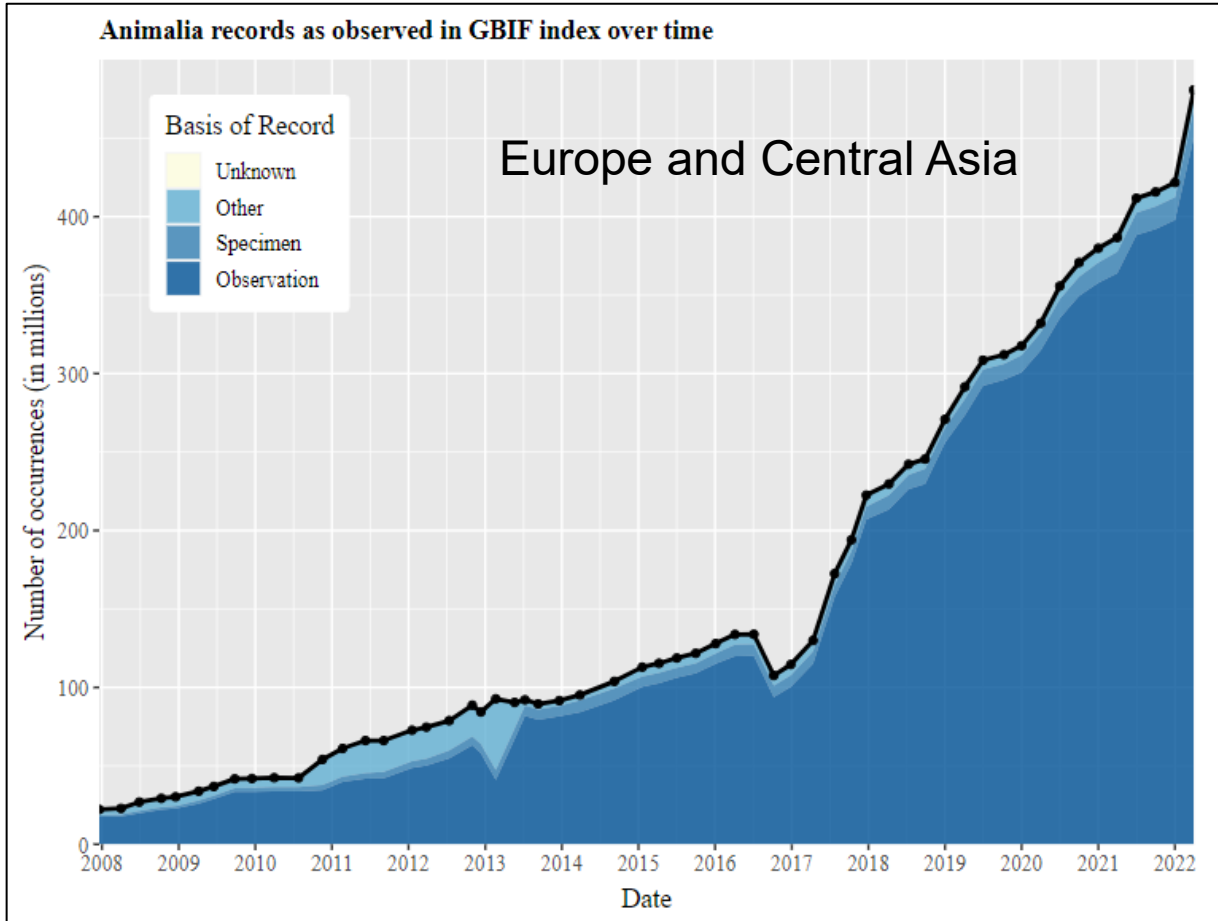






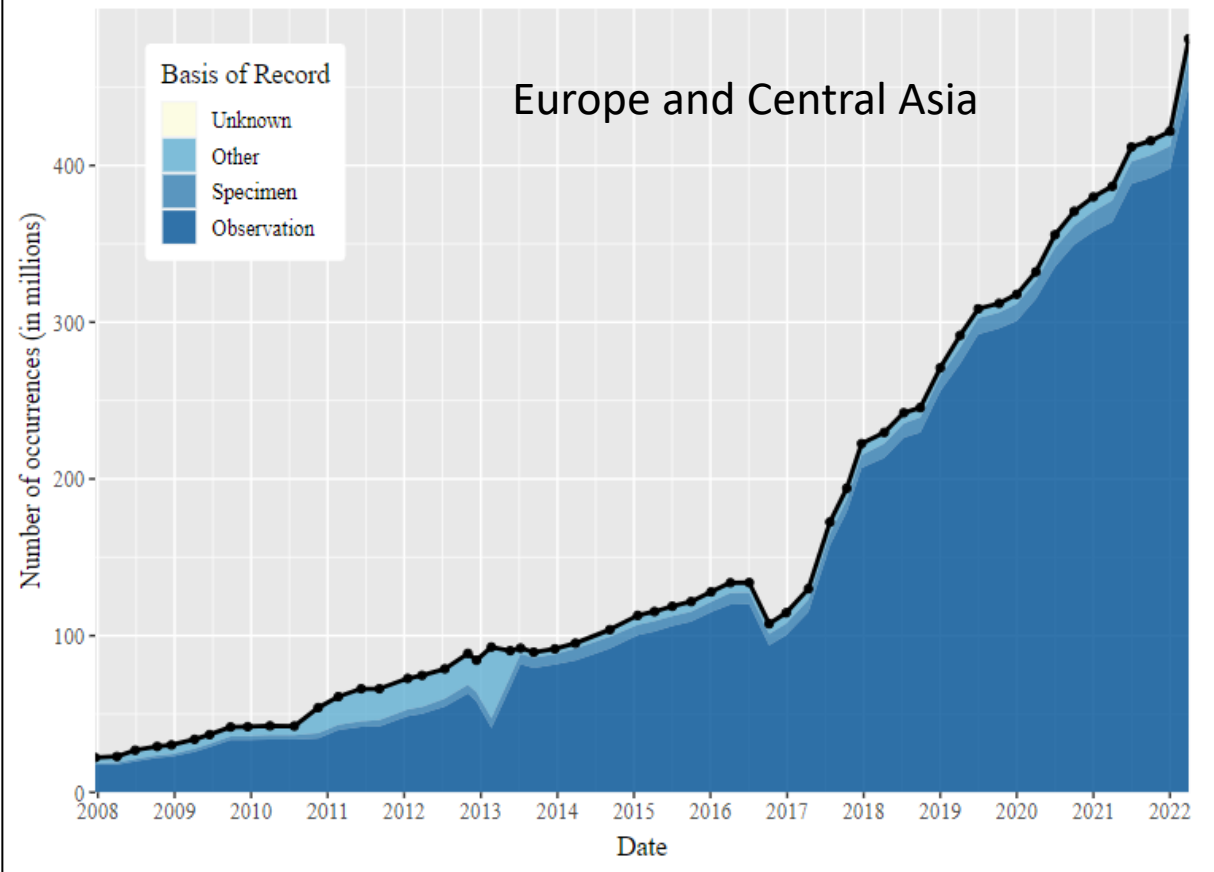
20% of insect data are for the UK

Is the situation improving?

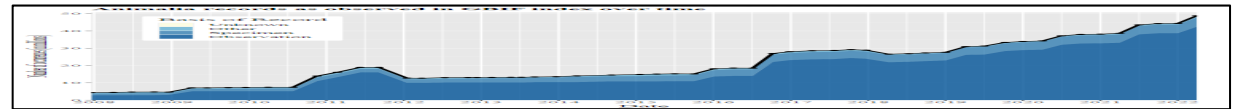


Is the situation improving?

Animalia records as observed in GBIF index over time



Africa

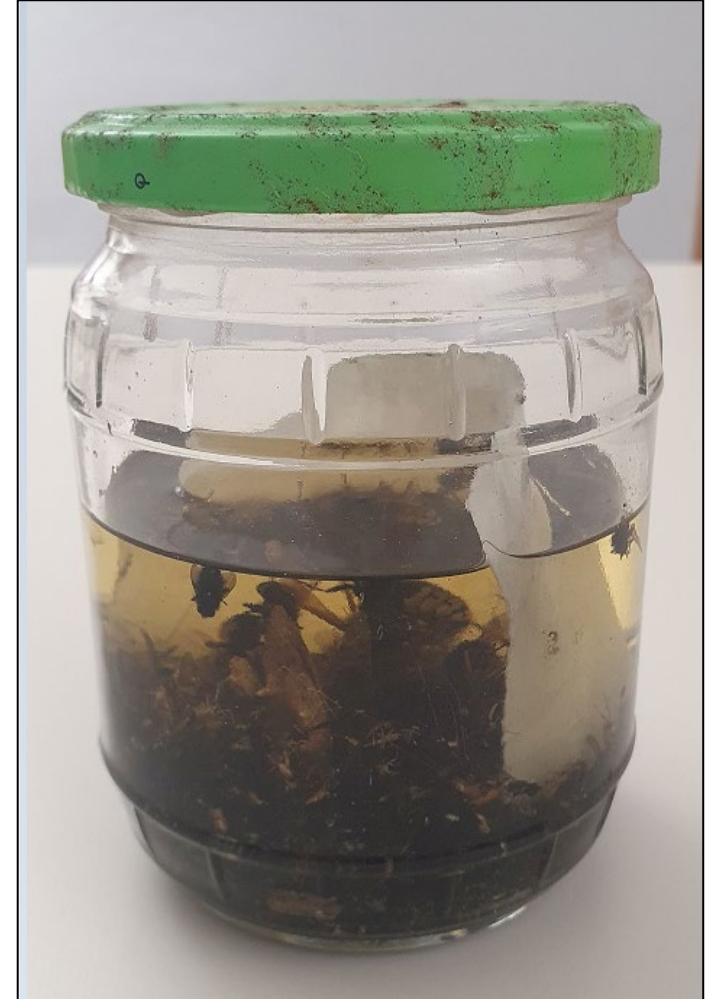
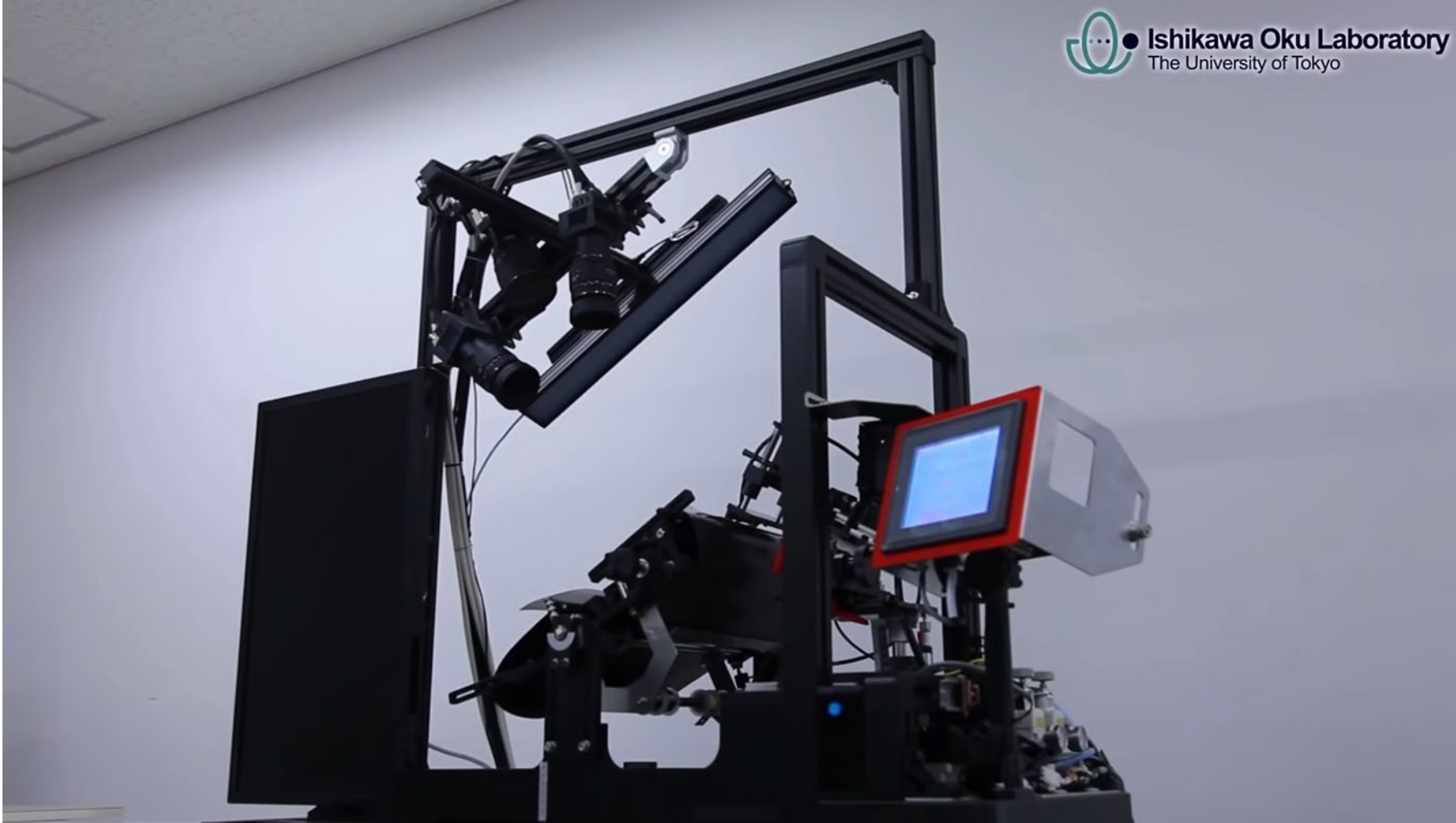


Biodiversity data crisis

- we need to generate more data fast
- data should be for standardized samples

Let's analyze biodiversity samples like libraries

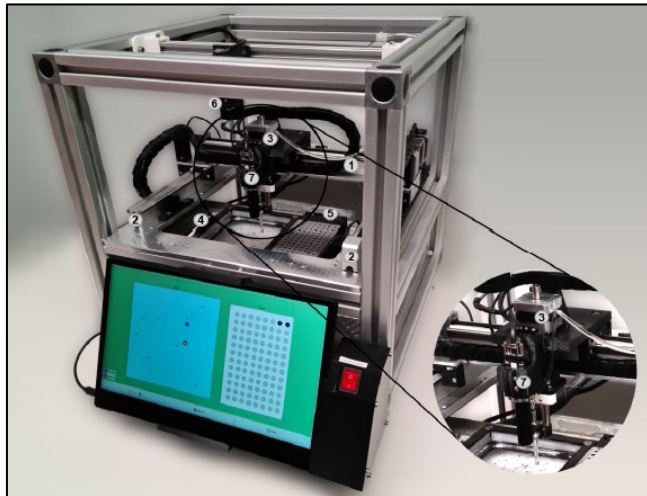
 Ishikawa Oku Laboratory
The University of Tokyo



Species Discovery Factory



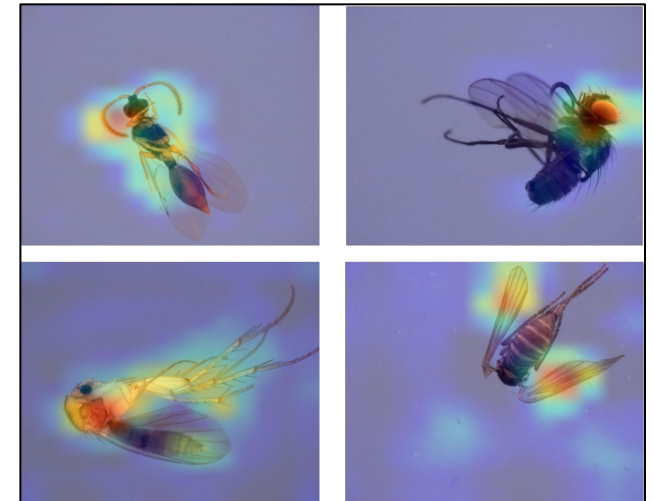
1. DiversityScanner:
Digitization with robots

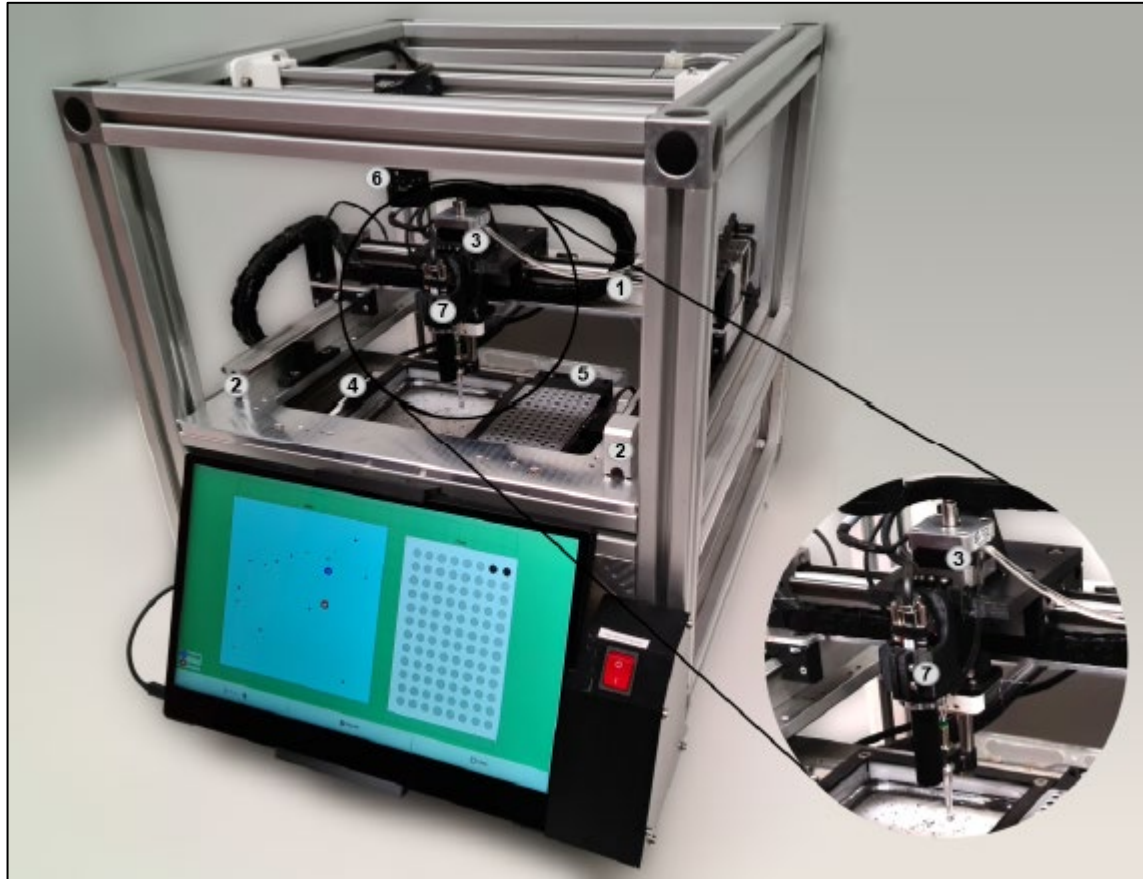


2. Nanopore Sequencing:
Sorting with DNA barcodes

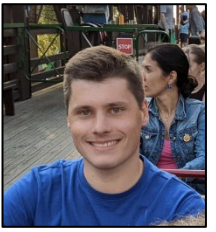


3. Identified images:
biomonitoring with AI





- robot finds insects in a sample
- images each specimen
- classifies the specimen
- measures specimen
- prepares it for DNA sequencing



Lorenz Wühl, KIT



MOLECULAR ECOLOGY RESOURCES

RESOURCE ARTICLE | [Open Access](#)

DiversityScanner: Robotic handling of small invertebrates with machine learning methods

Lorenz Wühl, Christian Pylatiuk , Matthias Giersch, Florian Lapp, Thomas von Rintelen, Michael Balke, Stefan Schmidt, Pierfilippo Cerretti, Rudolf Meier 

First published: 04 December 2021 | <https://doi.org/10.1111/1755-0998.13567>

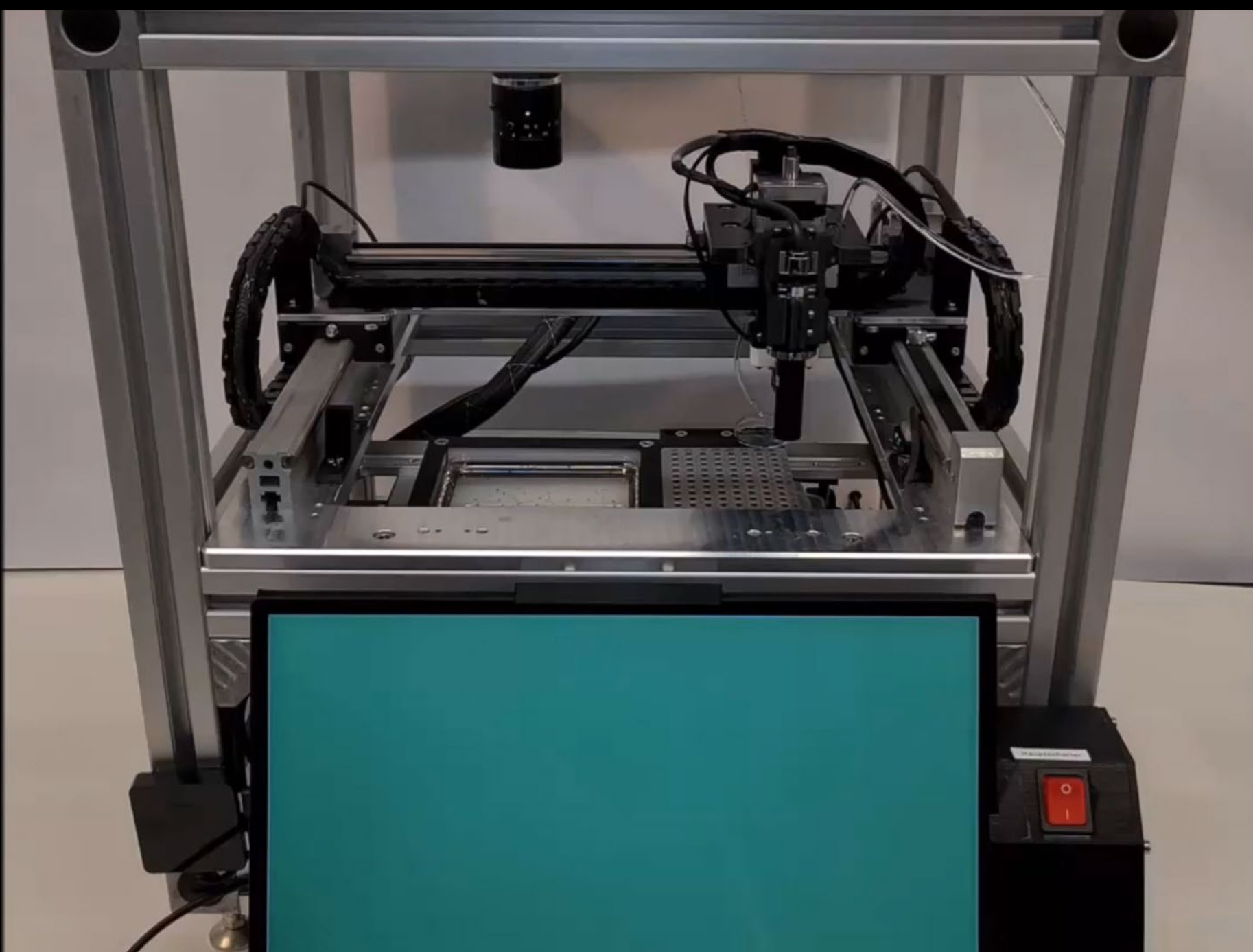
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:10.1111/1755-0998.13567



Prof. C. Pylatiuk



Lorenz Wührl



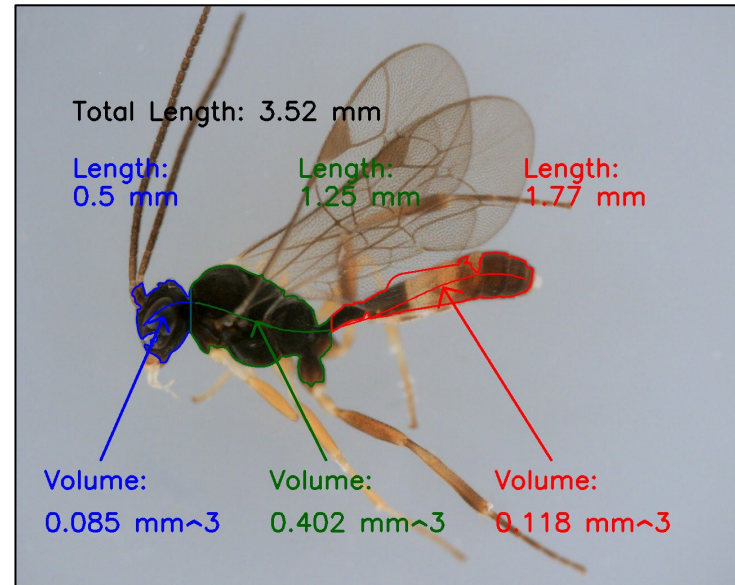
Detailed images for:

- taxonomic work
- training of AI models

=> image-based biomonitoring

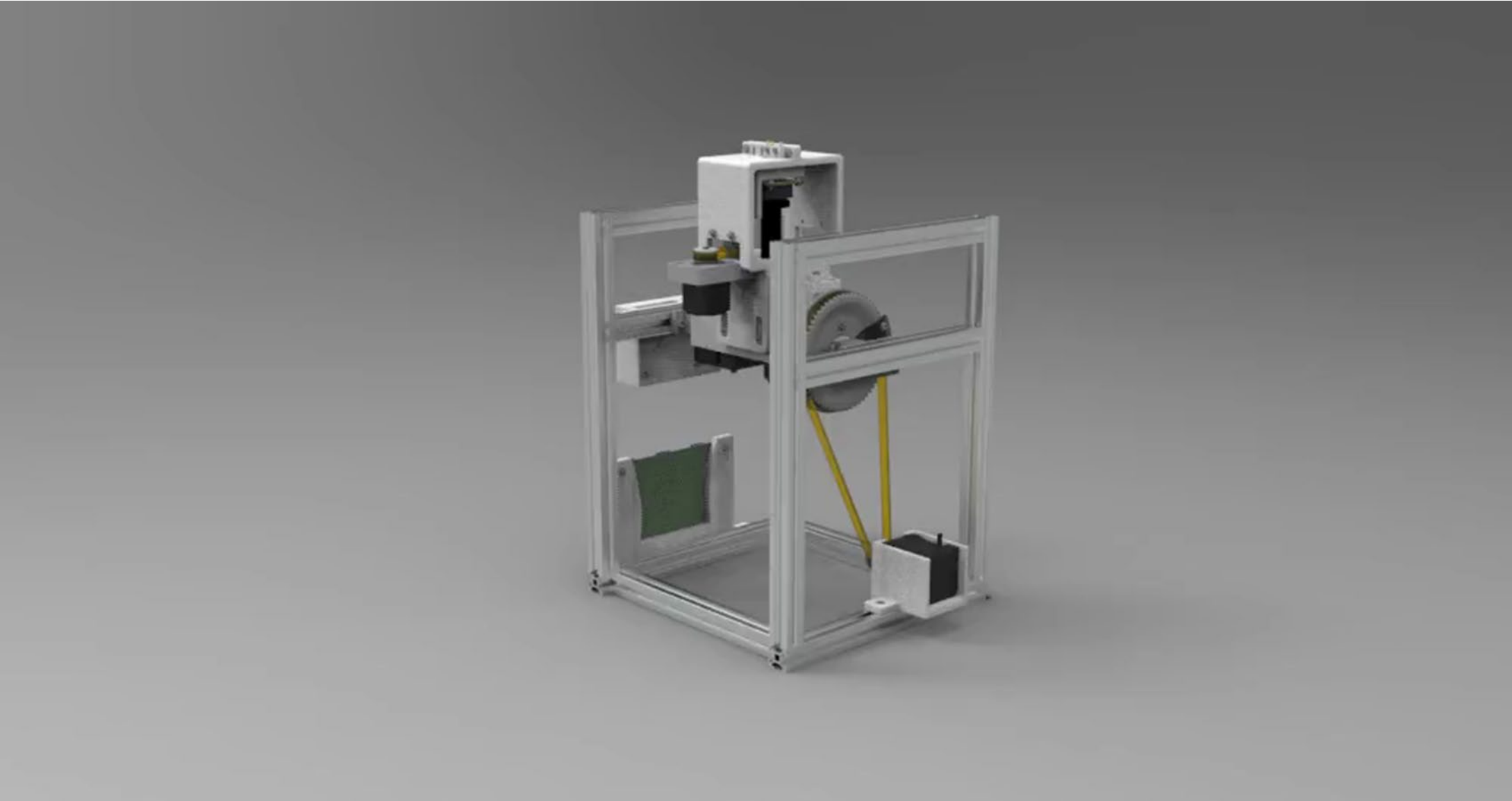


DiversityScanner
Biomass estimation





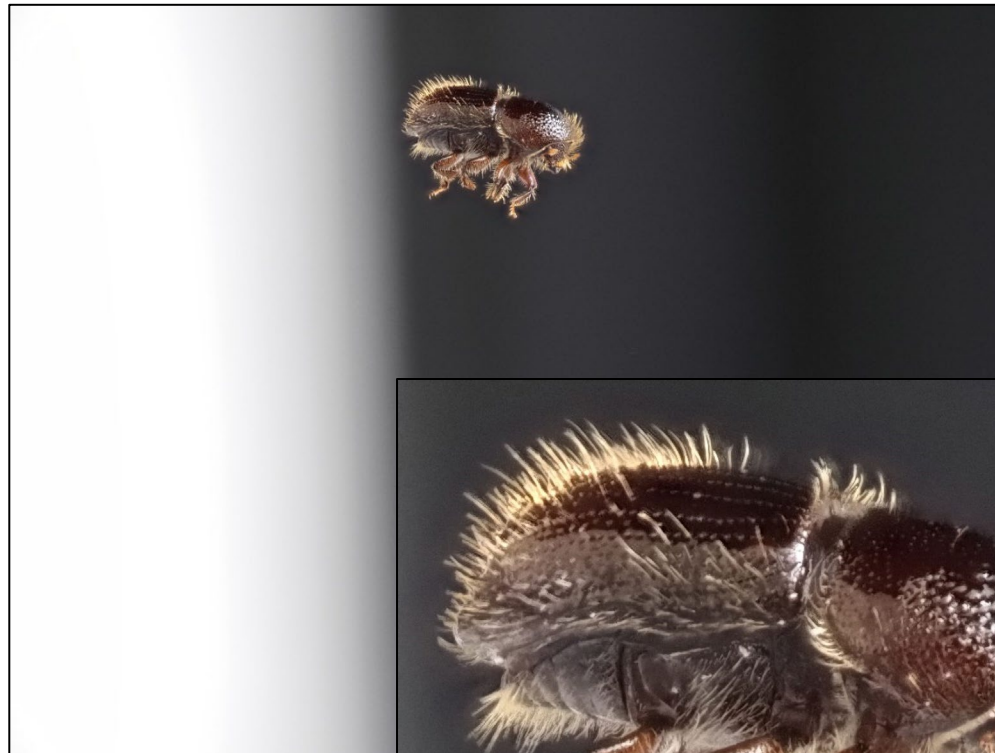
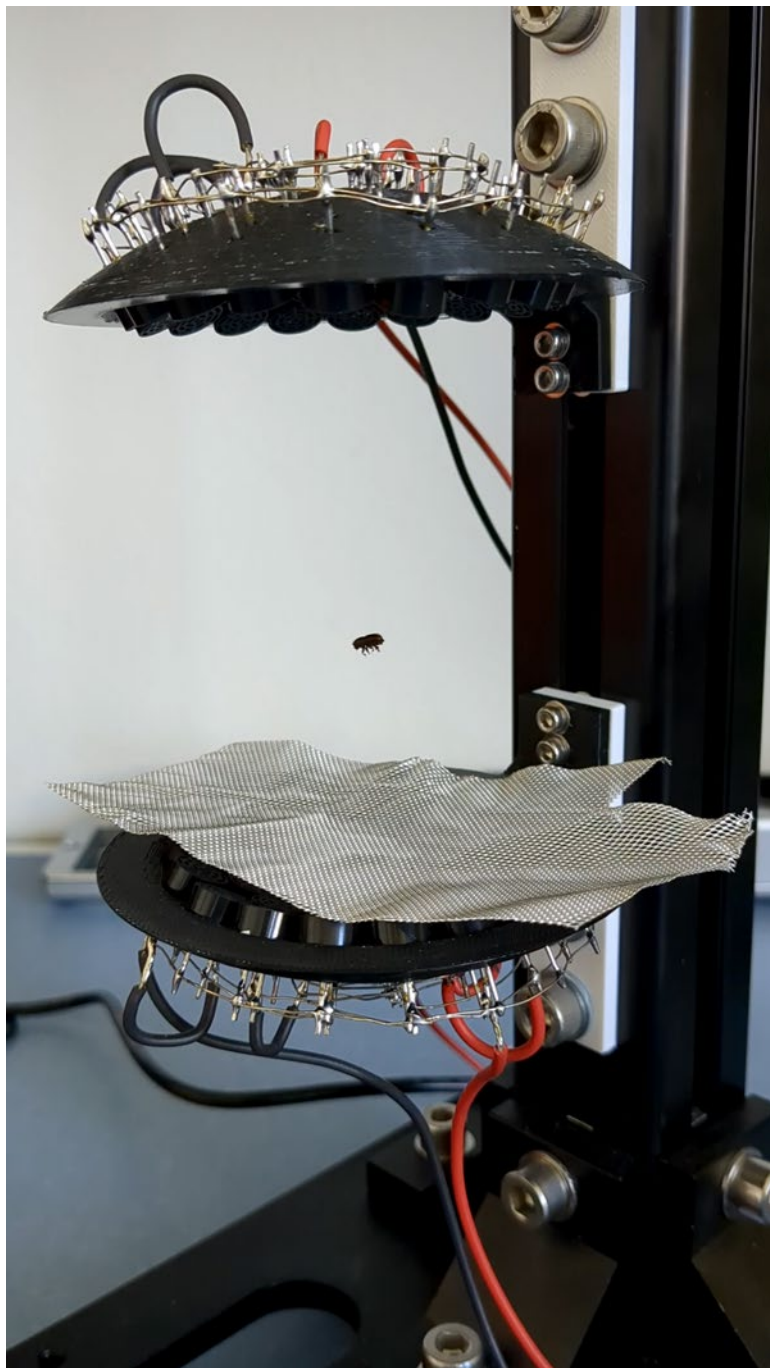
Lorenz Wüthli, KIT





Lorenz Wühl, KIT





Nathalie Klug, KIT

Imaging from multiple perspectives also important for training AI identification models

Different “DiversityScanners” for different purposes



Lorenz Wüthrl, KIT

Entomoscope: An Open-Source Photomicroscope for Biodiversity Discovery

LORENZ WÜHRL¹, LUCA RETTENBERGER¹, RUDOLF MEIER², EMILY HARTOP², JULIEN GRAF², and CHRISTIAN PYLATIUK¹,

¹Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany (lorenz.wuehrl, luca.retttenberger, pylatiuk}@kit.edu
²Center for Integrative Biodiversity Discovery, Leibniz Institute for Evolution and Biodiversity Science, Museum für Naturkunde Berlin, Germany (rudolf.meier, emily.hartop, julien.graf}@mfn.berlin

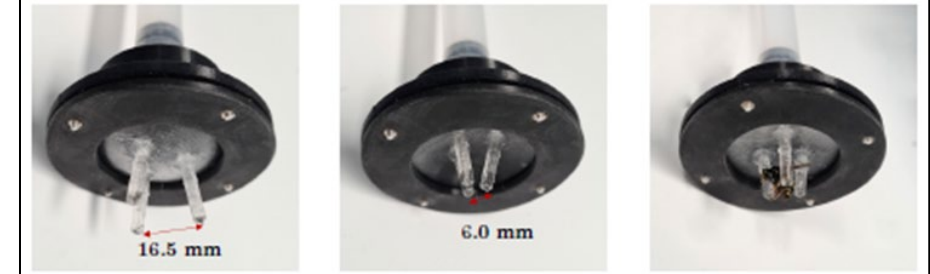
Corresponding author: Lorenz Wüthrl (e-mail: lorenz.wuehrl@kit.edu).

This research was supported by the Center for Integrative Biodiversity Discovery at the Museum für Naturkunde Berlin and the program Natural, Artificial and Cognitive Information Processing (NACIP) of the Helmholtz-Association.



Lorenz Wüthrl*, Leonard Keller, Nathalie Klug, Hossein Shirali, Rudolf Meier, and Christian Pylatiuk

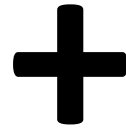
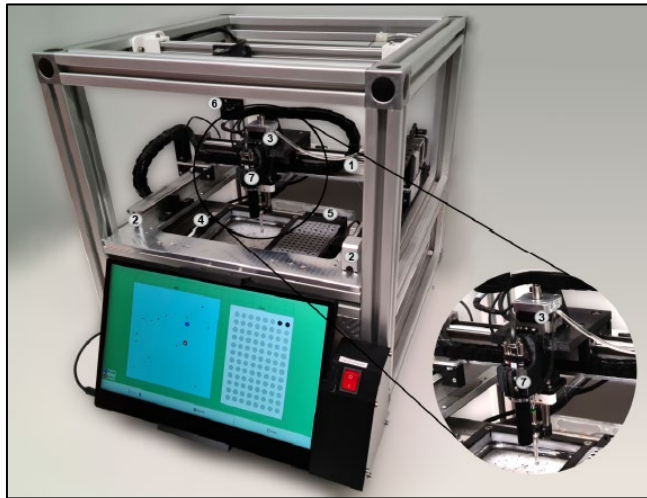
Automated Handling of Biological Objects with a Flexible Gripper for Biodiversity Research



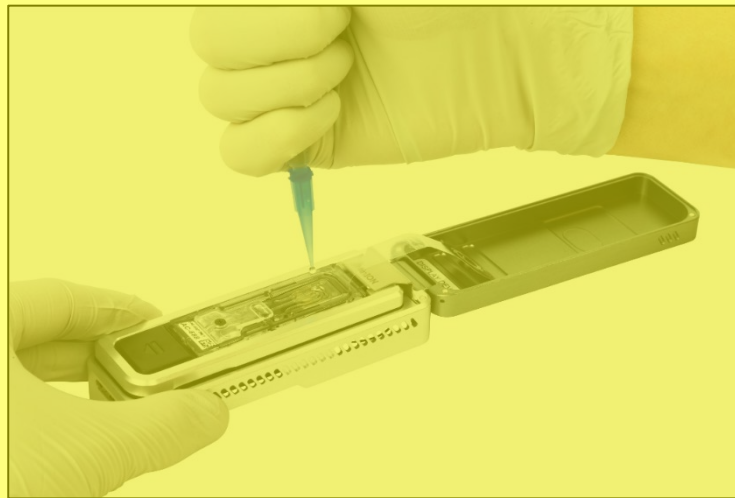
Species Discovery Factory



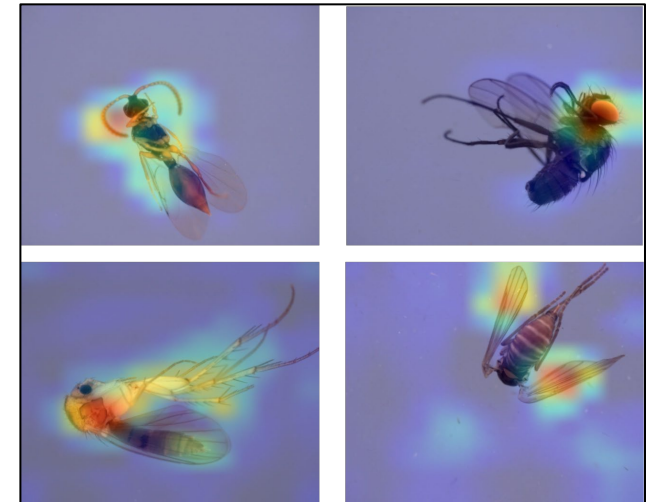
1. DiversityScanner: Digitization with robots



2. Nanopore Sequencing: Sorting with DNA barcodes



3. Identified images: biomonitoring with AI



2. Nanopore Sequencing

Srivathsan et al. *BMC Biology* (2021) 19:217
<https://doi.org/10.1186/s12915-021-01141-x>

BMC Biology

METHODOLOGY ARTICLE Open Access

ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone

Amrita Srivathsan¹, Leshon Lee¹, Kazutaka Katoh^{2,3}, Emily Hartop^{4,5}, Sujatha Narayanan Kutty^{1,6}, Johnathan Wong¹, Darren Yeo¹ and Rudolf Meier^{1,7*}

Check for updates

“Megabarcoding”:

- simplified, fast, and robust techniques for specimen-based barcoding
- 10 cents/specimen







Matching real-time sequencing with real-time DNA barcoding



Dr. Amrita Srivathsan

Cladistics
The Willi Hennig Society

Research Article | Open Access |

ONTbarcoder 2.0: rapid species discovery and identification with real-time barcoding facilitated by Oxford Nanopore R10.4

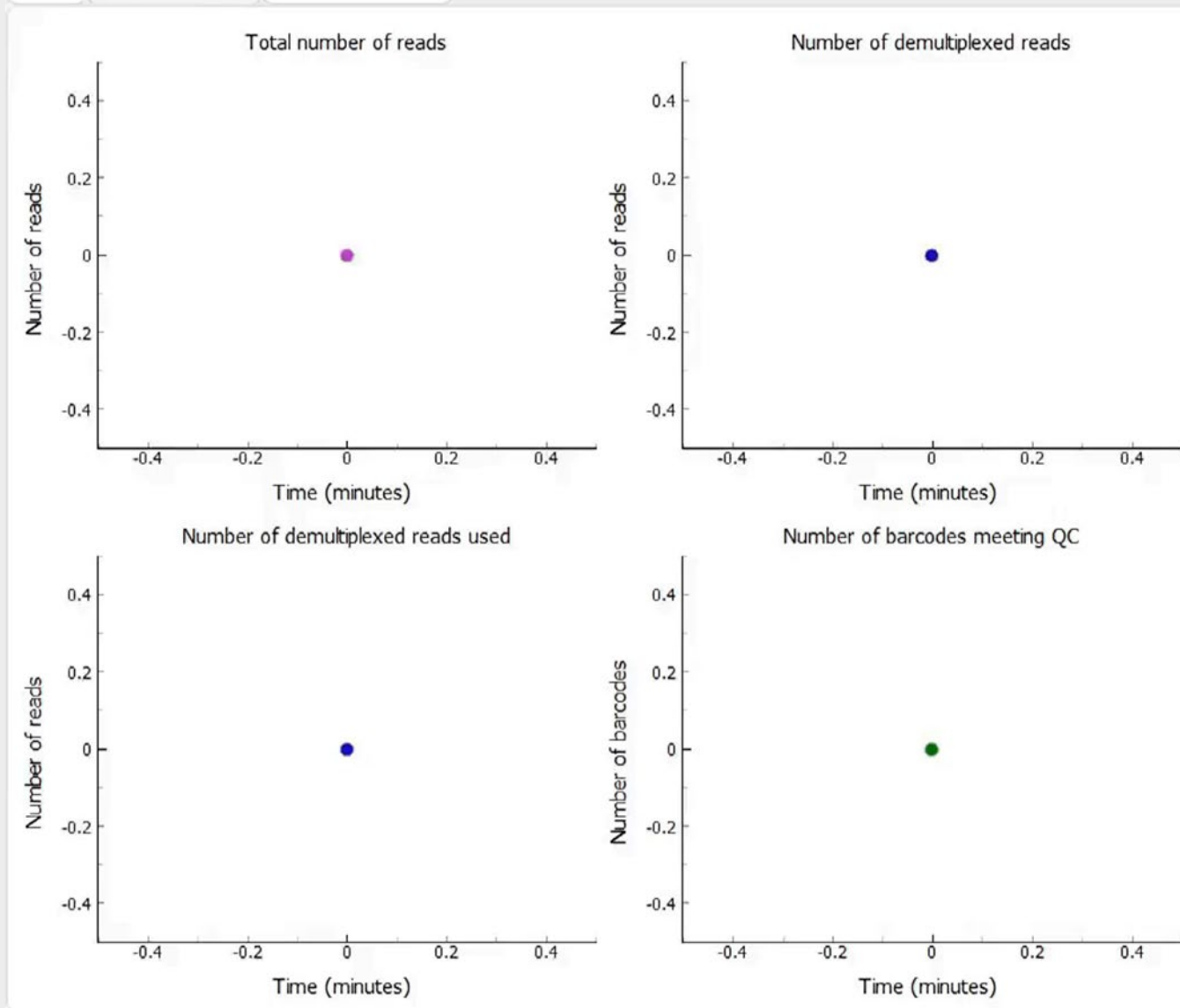
Amrita Srivathsan , Vivian Feng, Daniel Suárez, Brent Emerson, Rudolf Meier

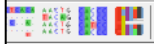
First published: 02 December 2023 | <https://doi.org/10.1111/cla.12566>



Flongle R10.4

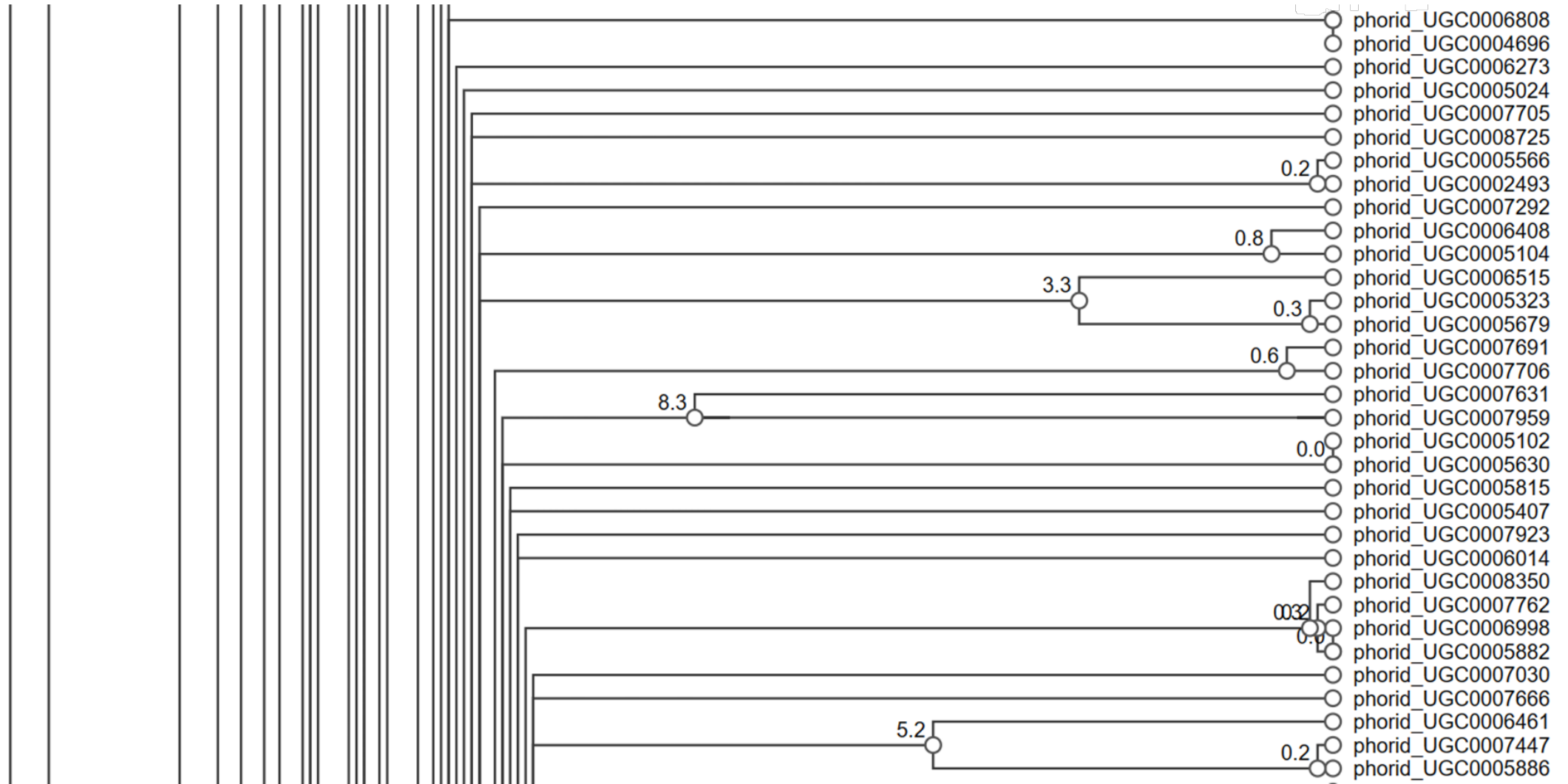
Setup Real-time barcoding Compare barcode sets

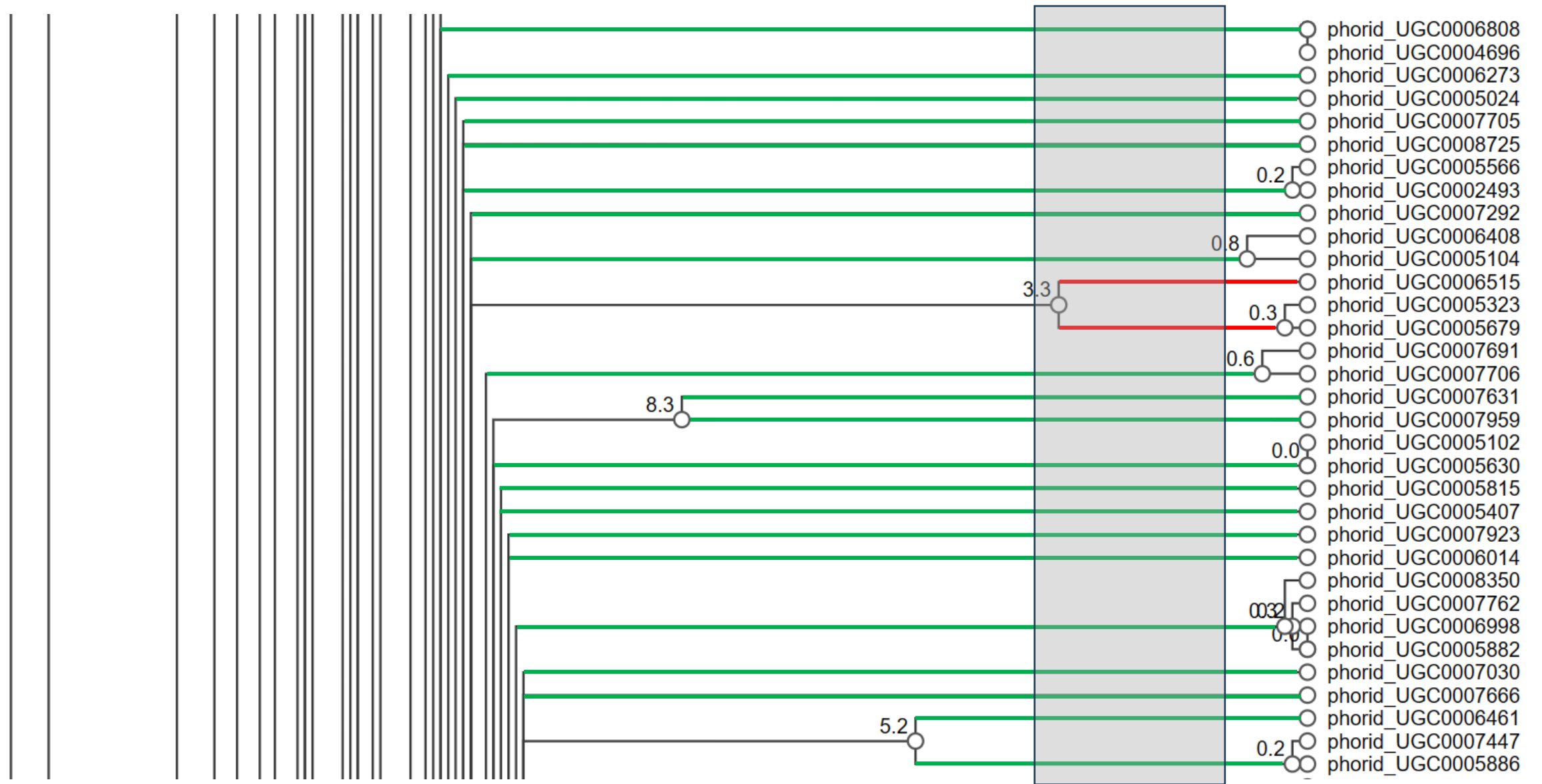




	10	20	30	40	50	60	70	80	90	100	110	120	130	140
Wenya_CN2281_CNMyco25_22AUG2	A	A	C	T	T	A	T	T	T	A	T	T	T	T
Wenya_CN5195_CNMyco56_26JUN21	G	G	C	T	T	A	T	T	T	G	G	C	T	T
Wenya_CN1952_CNMyco21_15AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2548_CNMyco28_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0300_CNMyco04_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0689_CNMyco08_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3353_CNMyco036_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0785_CNMyco09_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2598_CNMyco28_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3590_CNMyco39_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4254_CNMyco46_12SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5876_CNMyco63_redo1_2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3472_CNMyco37_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4701_CNMyco51_redo1_0	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4560_CNMyco49_26SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2511_CNMyco27_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2554_CNMyco28_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5958_CNMyco64_12OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Davenport_maggot3_maggot3_NA_1	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4754_CNMyco51_redo1_0	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1256_CNMyco014_8AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2017_CNMyco22_redo2_2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5860_CNMyco63_09SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0720_CNMyco08_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0324_CNMyco04_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1196_CNMyco13_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1066_CNMyco12_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4818_CNMyco52_10OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3277_CNMyco35_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3767_CNMyco041_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1457_CNMyco16_15AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0713_CNMyco08_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3702_CNMyco040_5septe	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3437_CNMyco37_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0600_CNMyco07_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3374_CNMyco036_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2307_CNMyco25_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4834_CNMyco52_10OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN2002_CNMyco22_redo2_2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1378_CNMyco15_15AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5532_CNMyco60_24JUL20	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1221_CNMyco13_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5021_CNMyco54_08MAY2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4884_CNMyco53_redo1_2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4740_CNMyco51_03OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4976_CNMyco54_redo1_0	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0234_CNMyco03_18JULY2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1978_CNMyco022_22AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4697_CNMyco50_03OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1075_CNMyco12_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4581_CNMyco49_redo1_2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5043_CNMyco54_redo1_3	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5022_CNMyco54_redo1_0	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0321_CNMyco04_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3138_CNMyco034_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Davenport_SK00057_SK1_NA_NASK	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5357_CNMyco58_10JUL20	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1585_CNMyco17_15AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3591_CNMyco39_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3762_CNMyco041_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4151_CNMyco045_5septe	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3639_CNMyco39_05SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3663_CNMyco39_05SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3216_CNMyco35_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0273_CNMyco03_25JULY2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3757_CNMyco040_5septe	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4266_CNMyco46_12SEP22	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0605_CNMyco07_01AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0778_CNMyco09_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN3323_CNMyco036_29AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN5942_CNMyco64_12OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN1265_CNMyco014_8AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN0943_CNMyco11_08AUG2	G	G	A	T	T	A	T	T	T	G	G	A	T	T
Wenya_CN4713_CNMyco51_03OCT2	G	G	A	T	T	A	T	T	T	G	G	A	T	T

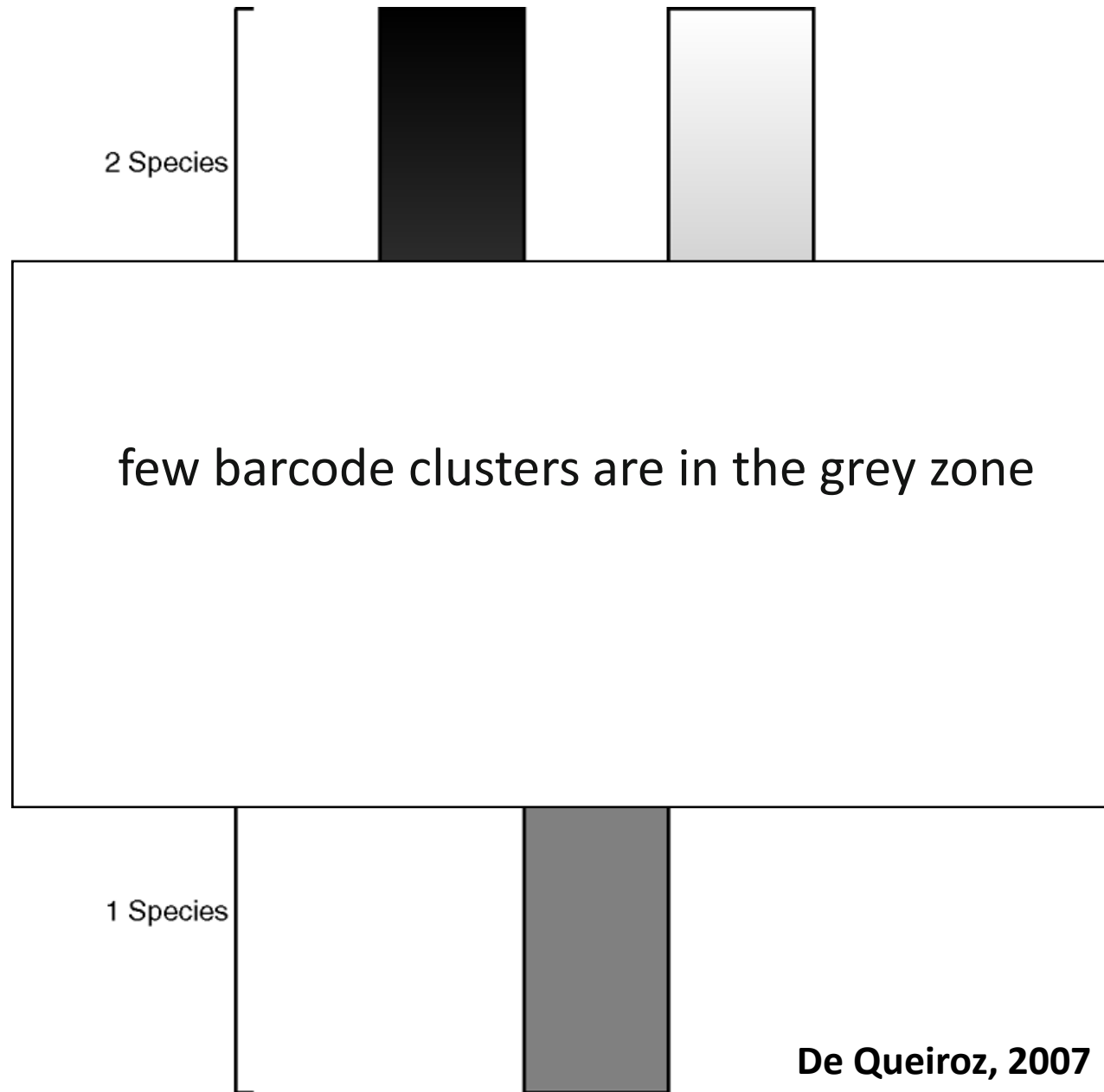
“Cluster Fusion Diagram”





1-3.5%

=> most haplotype distinct and rare => almost certainly species



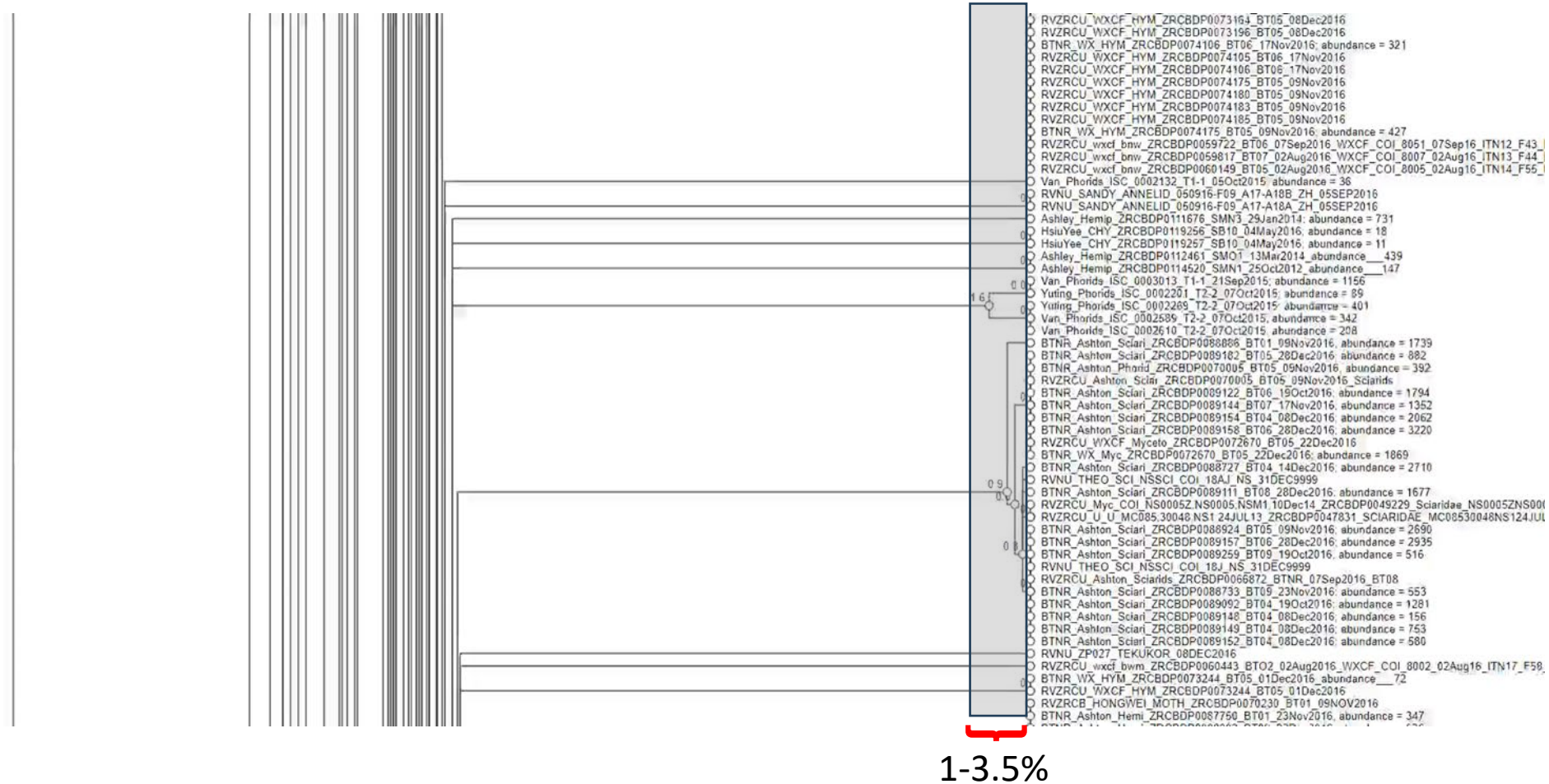
De Queiroz, 2007

Singapore dataset: 350,000 specimens

=> ca. 9,000 species with distribution and abundance information

=> 45 minute-movie

Few species are in the “grey zone” where species limits are unclear



DNA barcoding or Integrative Taxonomy?



Dr. Emily Hartop

Step 1. DNA barcoding for presorting specimens to putative species

Step 2. Morphological checks using a subset of specimens representing

- main haplotypes for distinct barcode clusters
- main and divergent haplotypes for “grey-zone clusters”

Step 3: Resolving incongruence

Step 4: Identification or Description

JOURNAL ARTICLE

Towards Large-Scale Integrative Taxonomy (LIT): Resolving the Data Conundrum for Dark Taxa

Emily Hartop, Amrita Srivathsan, Fredrik Ronquist, Rudolf Meier 

Systematic Biology, Volume 71, Issue 6, November 2022, Pages 1404–1422,

<https://doi.org/10.1093/sysbio/syac033>

Published: 13 May 2022 **Article history** ▼



Managing specimens: voucher handling



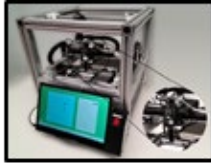
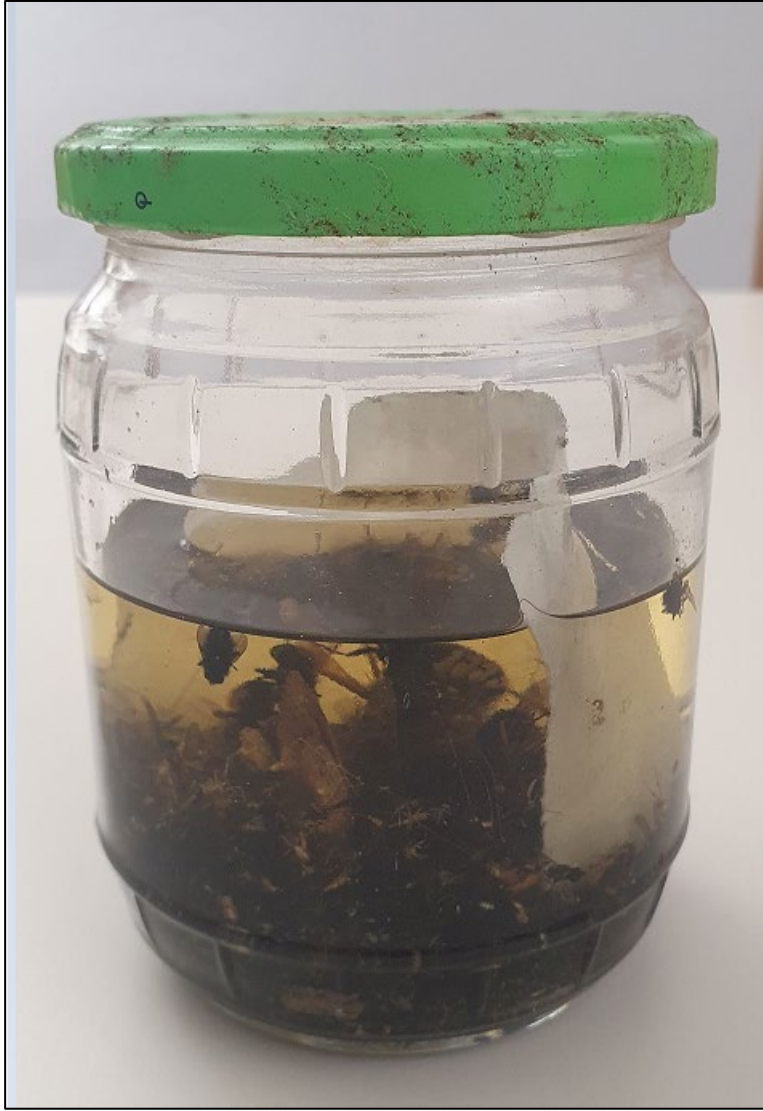
Vending machine for specimens

Carousel system for high density, low footprint storage

Pneumatic technology for vial storage and retrieval

function	comPOUND
temperature	-20°C, +4°C, ambient
dimension (w x d x h)	1.2 x 1.65 x 2.4 m (48 x 65 x 93 inches)
capacity	100,000 x 1.4 mL vials, 200,000 x 0.5 mL vials
throughput	600 vials/hour (1 sample/6 secs/unit)

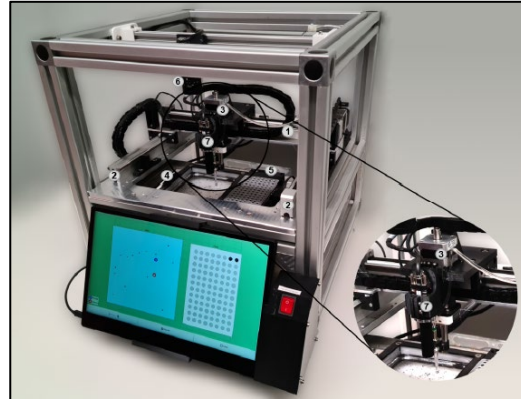




Managing data: from Entomoscope to Zenodo and GBIF

Automatic data upload to CERN's Zenodo

- Entomoscope and DiversityScanner software has Zenodo upload feature
- each specimen gets Zenodo entry with DOI (images, barcodes)



Dr. Amrita Srivathsan



Dr. Donat Agosti, Plazi

**CERN's 70th Anniversary and its contributions to Open Science:
From Particle Physics to Biodiversity Monitoring**

**18 September 2024, 6 pm
at the Embassy of Switzerland in Germany**

Automatic data upload to Zenodo/GBIF

ENIMAS - Entomoscope Imaging Software

- > AMRTEST
- > BMT
- > IDPSBG
- > PSC
- > SBG
- > SBK
- > test

Structure: Single
Code-prefix: SBK
Plate-ID: SBK001
Sample-ID: BWM001
Code-number: 150


Zenodo Upload

Current Lens: Lensagon TCSM-03-110 - Lensation - 135mm

Change... Settings...

Stack speed: 10 %
N. Stacks: 10
Step size: 0,80 mm

Reference Axis
STOP



SpeciesID:
Prediction: *Vespula_vulgaris*
Probability: 0.9302
[Open in Wikipedia](#)



Dr. Amrita Srivathsan



Dr. Donat Agosti, Plazi

How quickly can we generate good biodiversity data?

How quickly can we generate good biodiversity data?

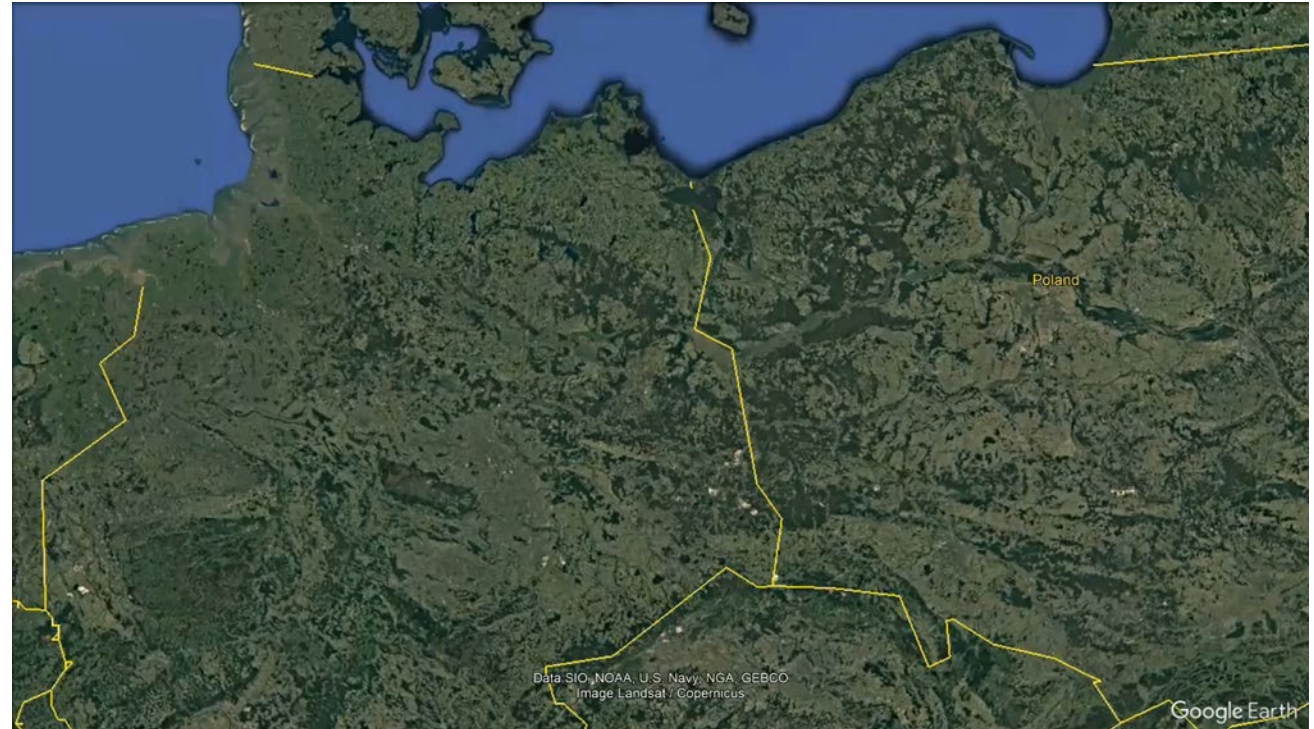


Cem Kir

5000 new data points for Berlin Kreuzberg

Bachelor's thesis

- 1 Malaise trap: 2-week July sample
- 5951 individuals
- 750 species



Let's imagine: 1000 BSc theses per year

- => 5 million new data points
- distribution of thousands of species known

imagine doing this for several years and in many places...

My challenge to you: be the first country that knows its arthropods

Step 1: barcode 1,000 specimens for 500 sites/sample (500,000 specimens)

=> species discovery rate can be 5-10% (i.e., 25,000 species)

Step 2: image numerically important species to train AI models

=> AI tools for common species

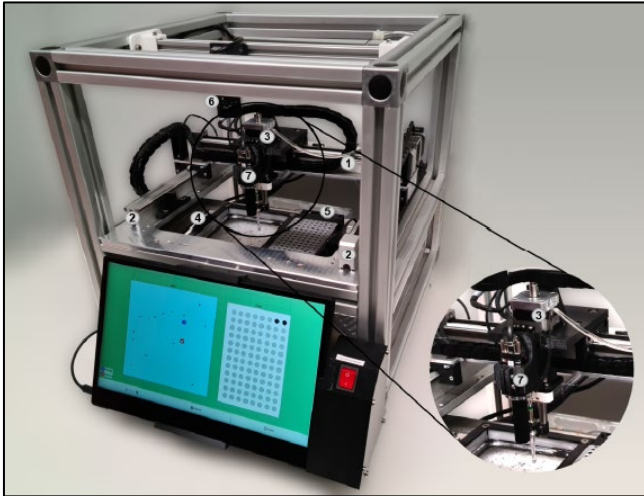
This would yield:

- unbiased abundance and diversity data for 500 sites
- barcodes for 25,000 species to interpret metabarcoding data
- vouchers, barcodes, and AI models for the most abundant species
- critical information on distributions and ecology of these species

Species Discovery Factory



1. DiversityScanner:
Digitization with robots



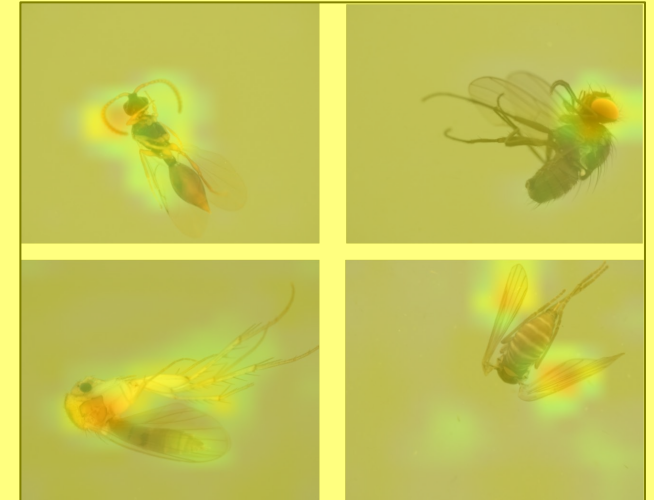
+

2. Nanopore Sequencing:
Sorting with DNA barcodes

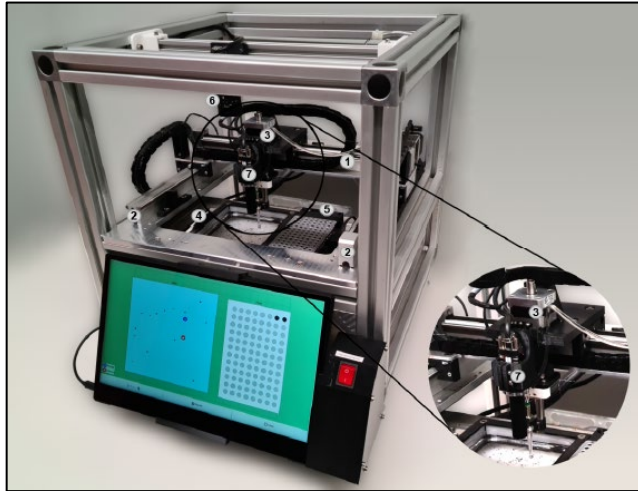


=

3. Identified images:
biomonitoring with AI



3. Identification with AI



Class (Taxon)	Result	Class (Taxon)	Result
Diptera Acalyptratae	91%	Diptera Psychodidae	89%
Diptera Calyptratae	83%	Diptera Sciaridae	92%
Diptera Cecidomyiidae	91%	Hemiptera Cicadellidae	100%
Diptera Chironomidae	97%	Hymenoptera Braconidae	82%
Diptera Dolichopodidae	86%	Hymenoptera Diapriidae	100%
Diptera Empididae & Hybotidae	87%	Hymenoptera Ichneumonidae	75%
Diptera Keroplatidae & Mycetophilidae	99%	Other	81%
Diptera Phoridae	97%		

AI identification tools evolve

- common taxa benefit first
- tools evolve from family => genus => species
- latest version of DiversityScanner/Entomoscope has an AI model for common species in Berlin



Hossein Shirali, KIT

ENIMAS - Entomoscope Imaging Software

- > AMRTEST
- > BMT
- > IDPSBG
- > PSC
- > SBG
- > SBK
- > test

Structure: Single

Code-prefix: SBK

Plate-ID: SBK001

Sample-ID: BWM001

Code-number: 150

Zenodo Upload

Current Lens: Lensagon TCSM-03-110 - Lensation - 135mm

Change... Settings...

Stack picture

Stack speed: 10 %

N. Stacks: 10

Step size: 0,80 mm

STOP



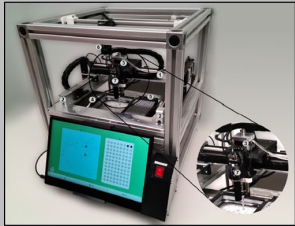
Hossein Shirali, KIT

Does the model really work? some species are easy some are hard

Species Discovery Factory



1. DiversityScanner:
Digitization with robots



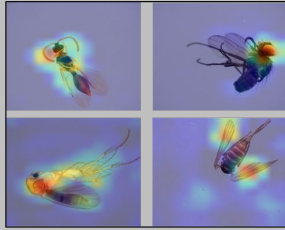
+

2. Nanopore Sequencing:
Sorting with DNA barcodes



=

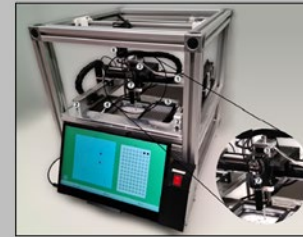
3. Identified images:
biomonitoring with AI



Species Biomonitoring Factory

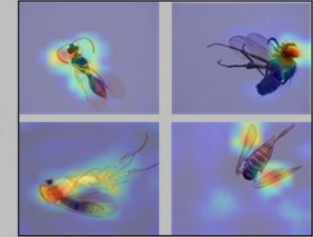


1. DiversityScanner:
Digitization with robots



=

2. Identified images:
biomonitoring with AI






Could generate large amounts of data quickly

Shortcomings of the old ways...

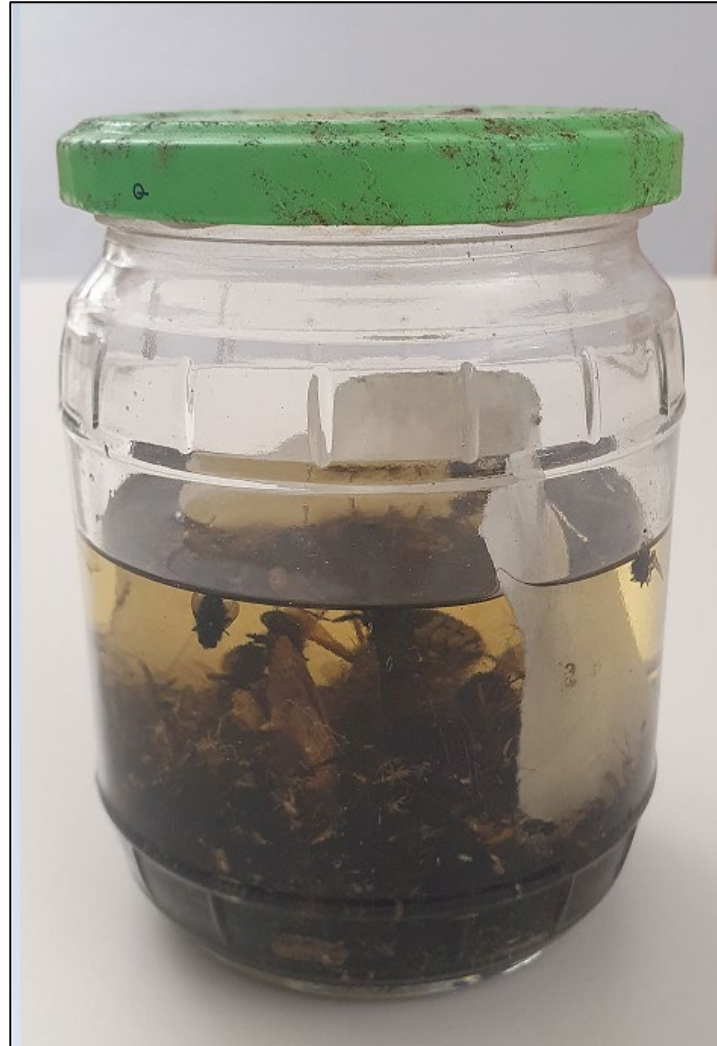
- >80% of all species undescribed
- 95% of all described species only identifiable for taxonomic experts
- very little baseline biodiversity data from standardized samples
- natural history data only available for <5% of species



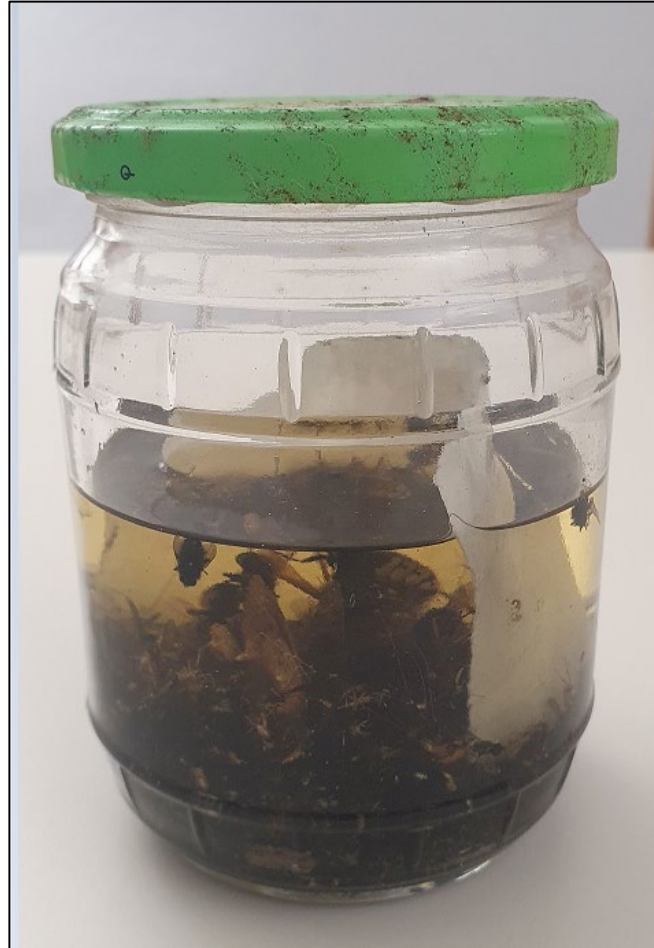
Shortcomings of the old ways...

- >80% of all species undescribed 
- 95% of all described species only identifiable for taxonomic experts 
- very little baseline biodiversity data from standardized samples 
- natural history data only available for <5% of species

Too many species... isn't it hopeless?



**”Dark Taxonomy“: focus taxonomic work on species
important for biomonitoring**



”Dark Taxonomy“ with biomonitoring focus

Step 1: Get fresh material for a species-rich taxon using traps also used for biomonitoring



Taxonomic
sample

“Dark Taxonomy”

Step 1: Get fresh material from many traps used in biomonitoring

Step 2: Pre-sort specimens with DNA barcodes



”Dark Taxonomy“: phase 1 of taxonomic exploration

Step 1: Get fresh material from many traps used in biomonitoring

Step 2: Pre-sort specimens with DNA barcodes

Step 3: Integrative taxonomy with two data sources

“Dark Taxonomy”

Step 1: Get fresh material from many traps used in biomonitoring


Step 2: Pre-sort specimens with DNA barcodes

Step 3: Integrative taxonomy with two data sources


Step 4: Identify/describe species

120 species of fungus gnats (Diptera: Mycetophilidae) discovered in Singapore:

- 5 described species
- 115 new species



Cold Spring Harbor Laboratory



bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

New Results [Follow this preprint](#)

An integrative taxonomic treatment of the Mycetophilidae (Diptera: Bibionomorpha) from Singapore reveals 115 new species on 730km²

[ID](#) Dalton de Souza Amorim, [ID](#) Sarah Siqueira Oliveira, Maria Isabel P.A. Balbi, [ID](#) Yuchen Ang, [ID](#) Darren Yeo, [ID](#) Amrita Srivathsan, [ID](#) Rudolf Meier

doi: <https://doi.org/10.1101/2023.09.02.555672>



“Dark Taxonomy”

Step 1: Get fresh material from many traps used in biomonitoring

Step 2: Pre-sort specimens with DNA barcodes

Step 3: Integrative taxonomy with two data sources

Step 4: Identify or describe species

Step 5: Test whether all species important for biomonitoring are covered



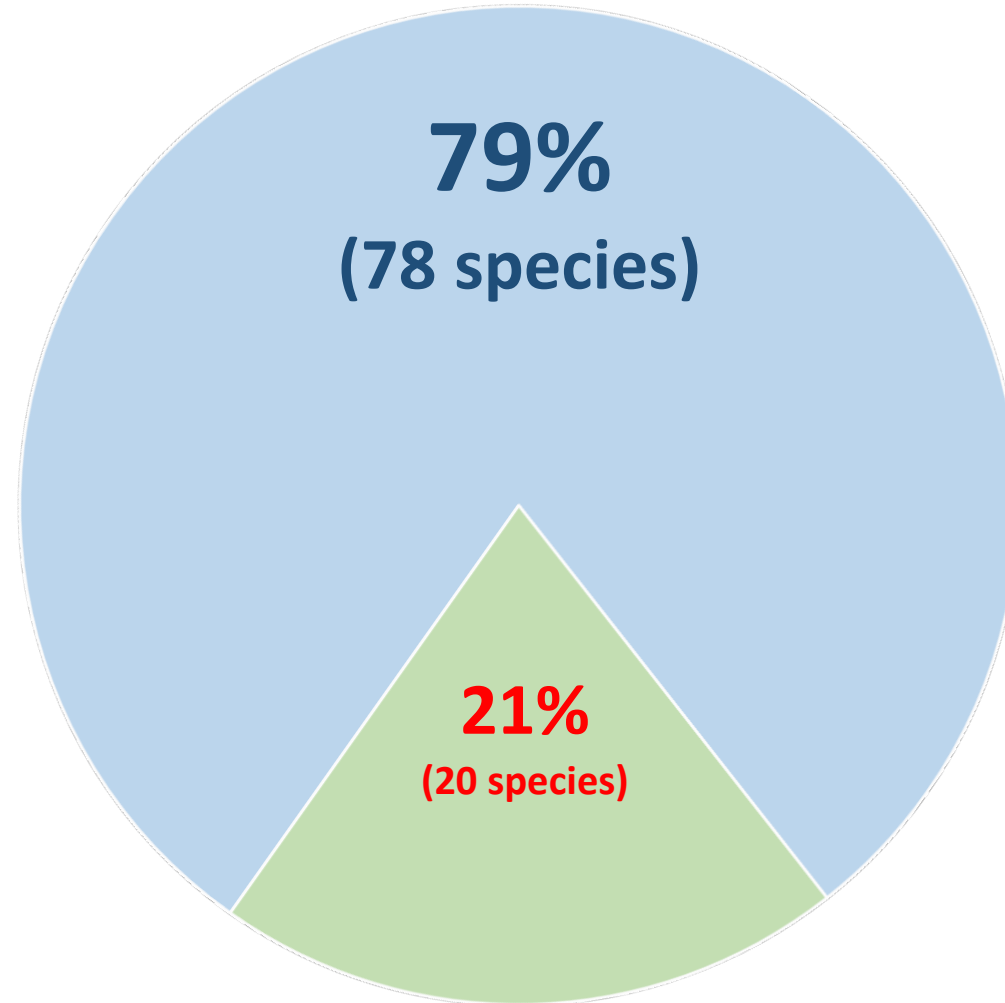
Taxonomic
sample



Test
sample

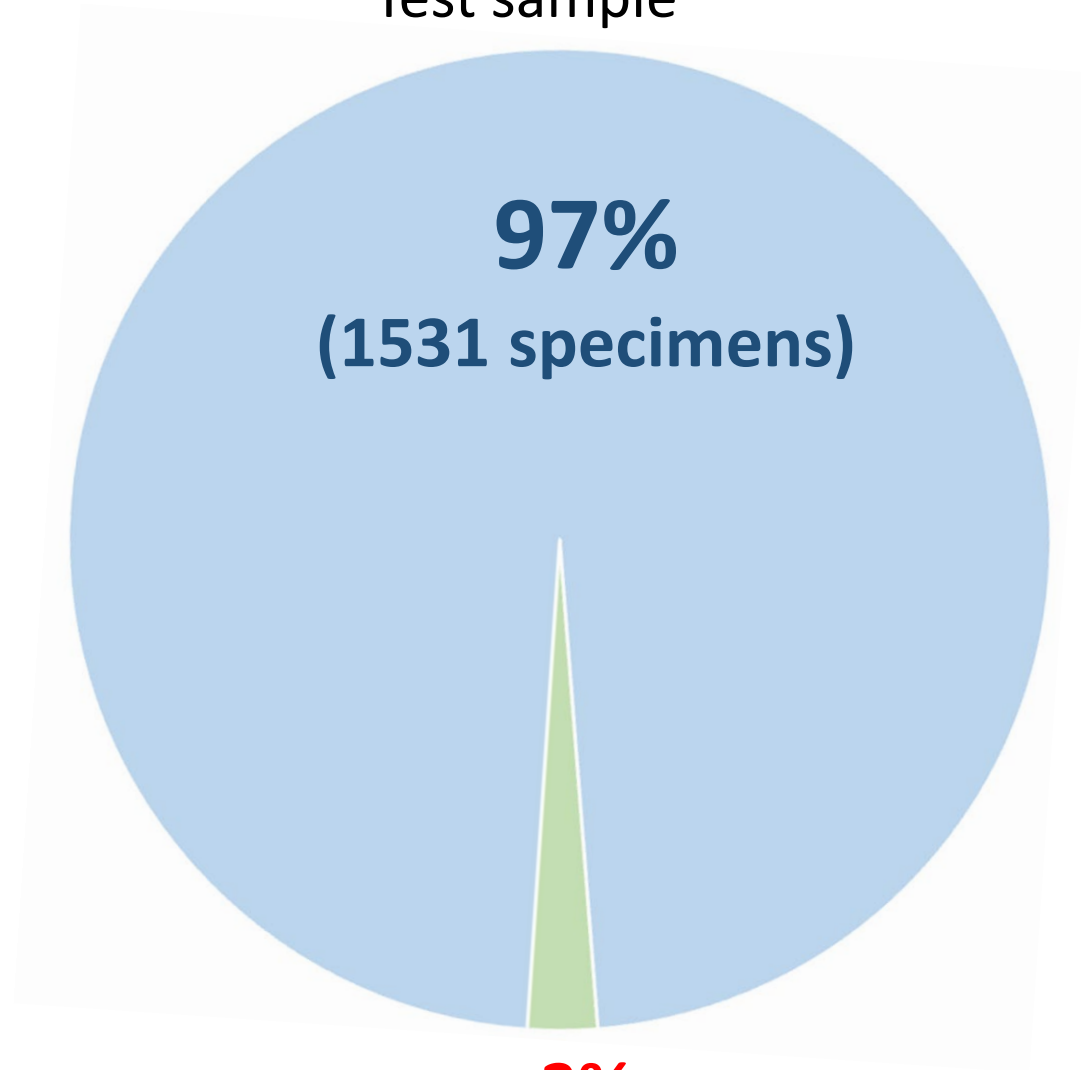
Mycetophilidae would be “biomonitoring-enabled” if test sample consists mostly of described species also in taxonomic sample

Test sample



What about specimens?

Test sample



97%
(1531 specimens)

3%
(36 specimens)

Dark Taxonomy:

from unknown to ready for biomonitoring in 5 steps



Cold Spring Harbor Laboratory

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

New Results [Follow this preprint](#)

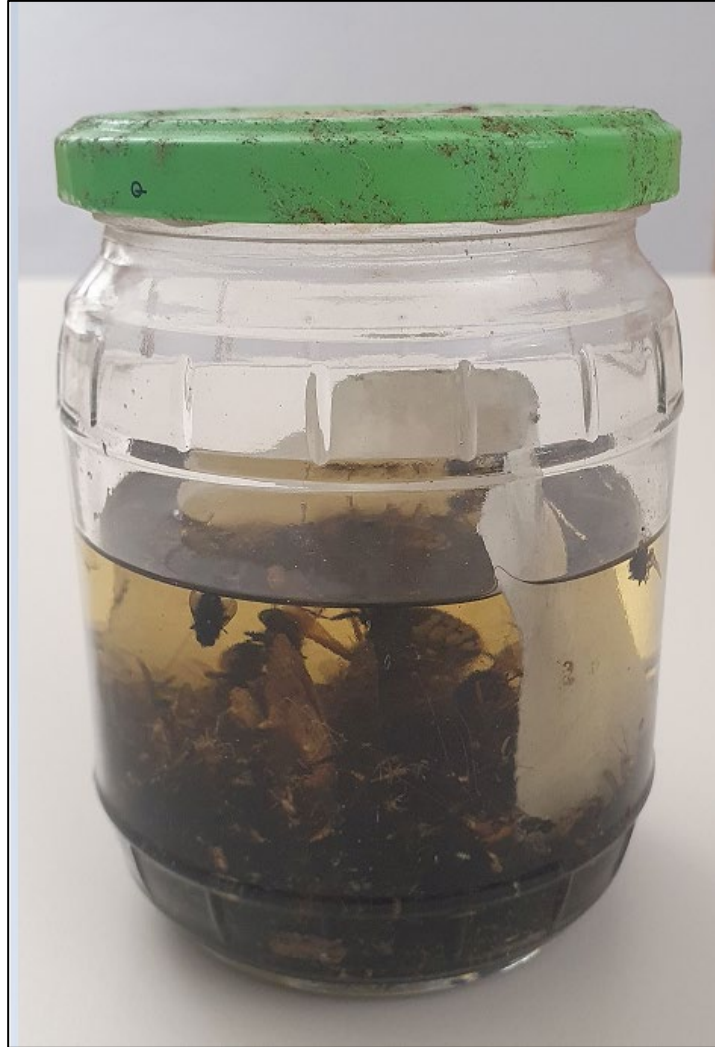
"Dark taxonomy": a new protocol for overcoming the taxonomic impediments for dark taxa and broadening the taxon base for biodiversity assessment

 Rudolf Meier,  Amrita Srivathsan,  Sarah Siqueira Oliveira, Maria Isabel P.A. Balbi,  Yuchen Ang,  Darren Yeo,  Dalton de Souza Amorim





doi: <https://doi.org/10.1101/2023.08.31.555664>



=> insect samples and faunas aren't hopeless



Shortcomings of the old ways...

- >80% of all species undescribed 
- 95% of all described species only identifiable for taxonomic experts 
- very little baseline biodiversity data from standardized samples 
- natural history data only available for <5% of species 

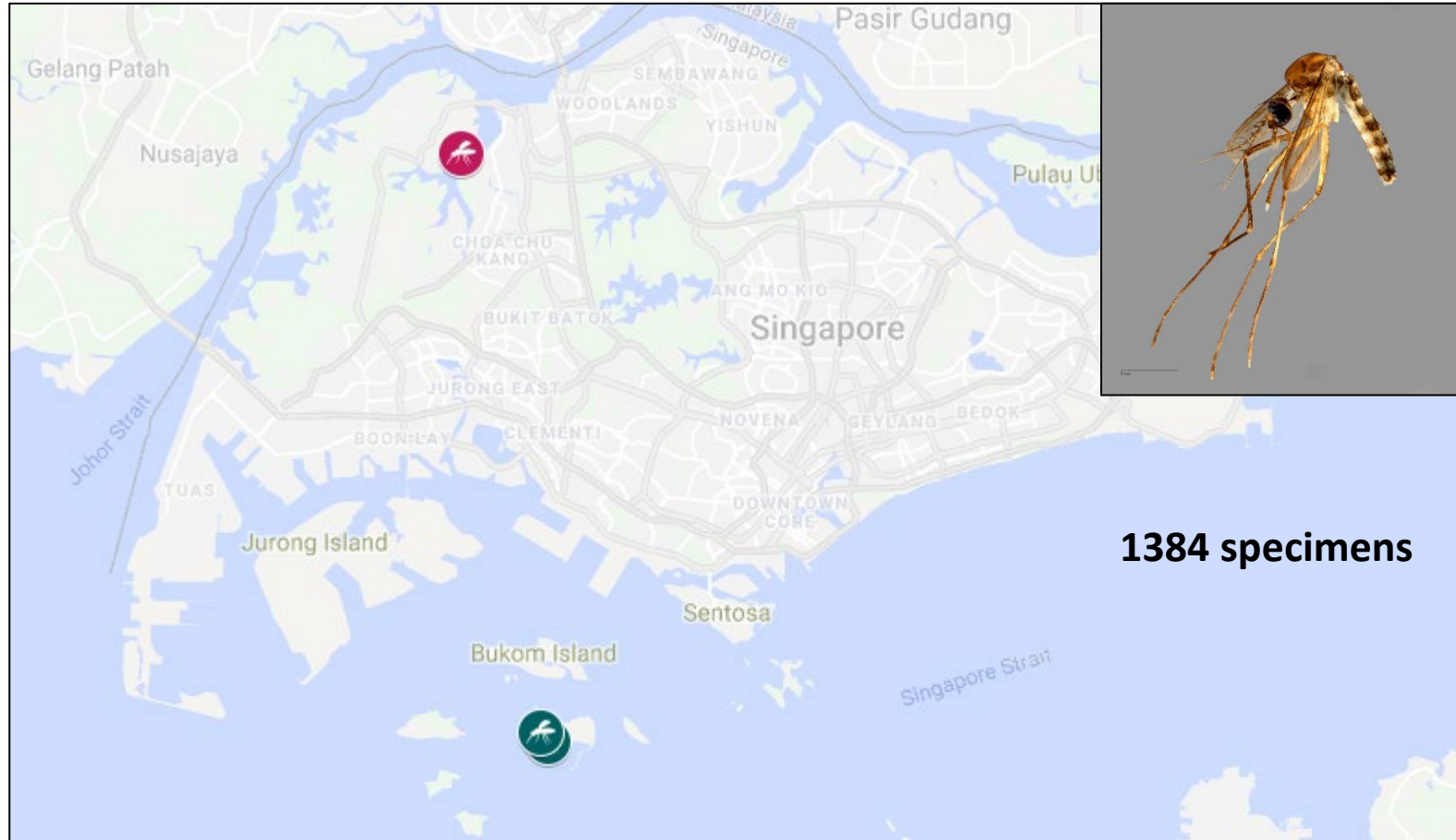
Our recommendation: resolve the natural history of common species first

Example 1: Mosquito diet

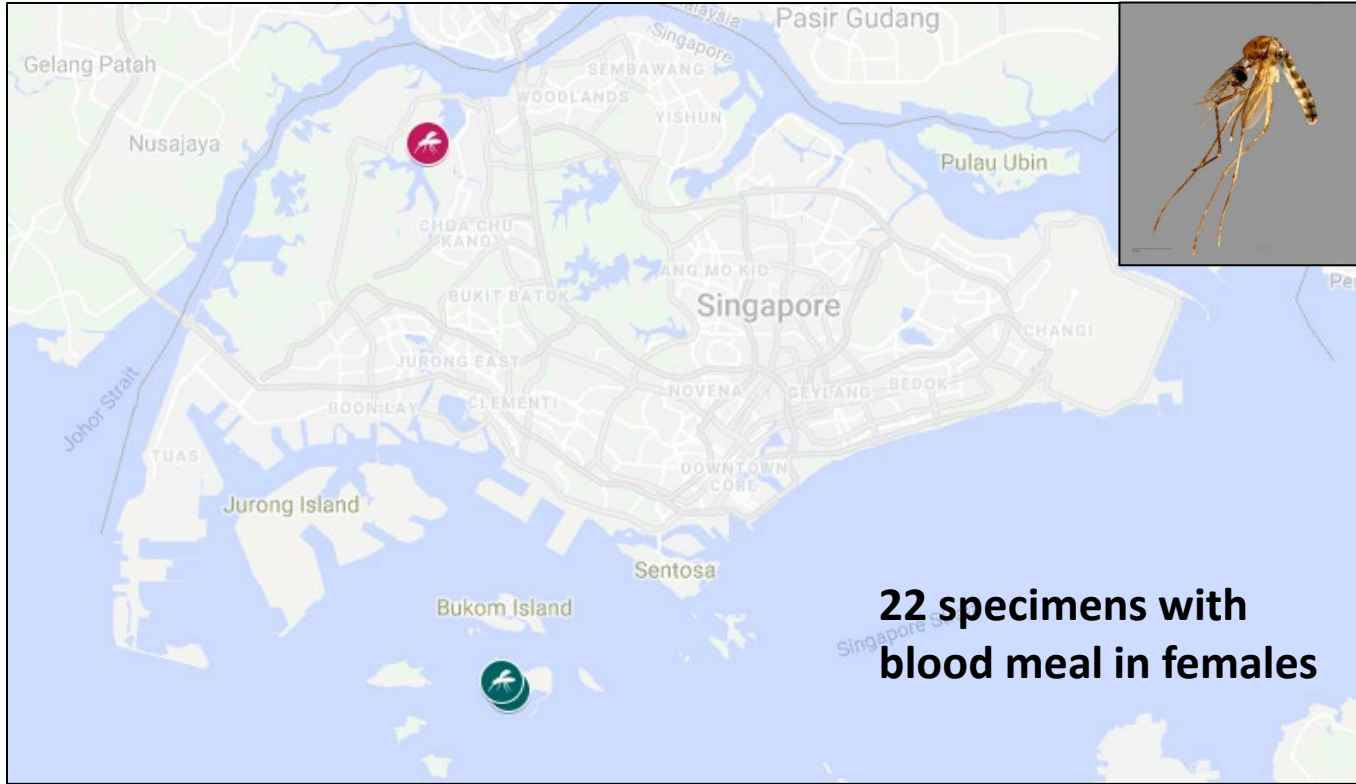
- 4000 mosquitoes barcoded
- => ca. 120 species detected (1/3rd identifiable via BLAST)
- => females for ca. 10 species had visible blood meals
- => metabarcoding of gut content



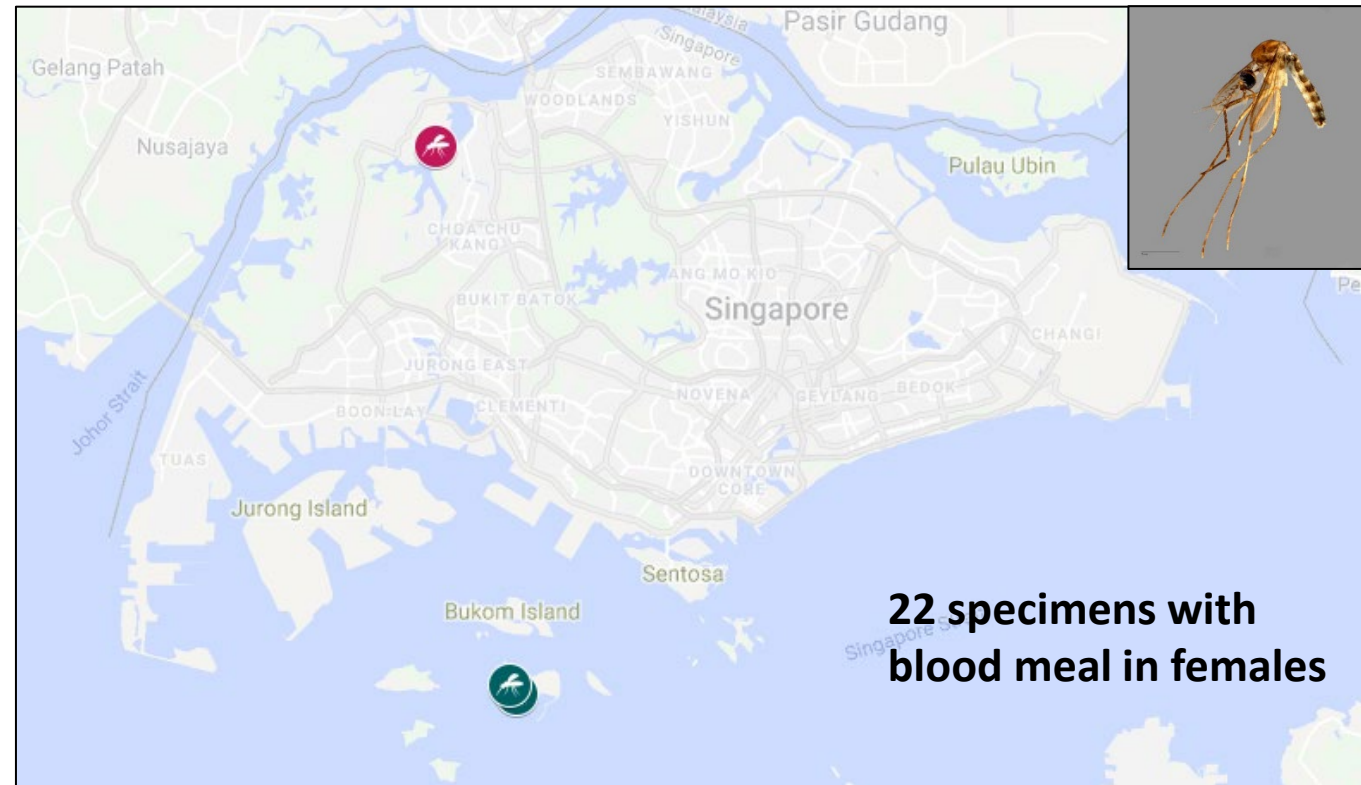
Aedes sp. ZRCBDP00009685



Aedes sp. ZRCBDP00009685

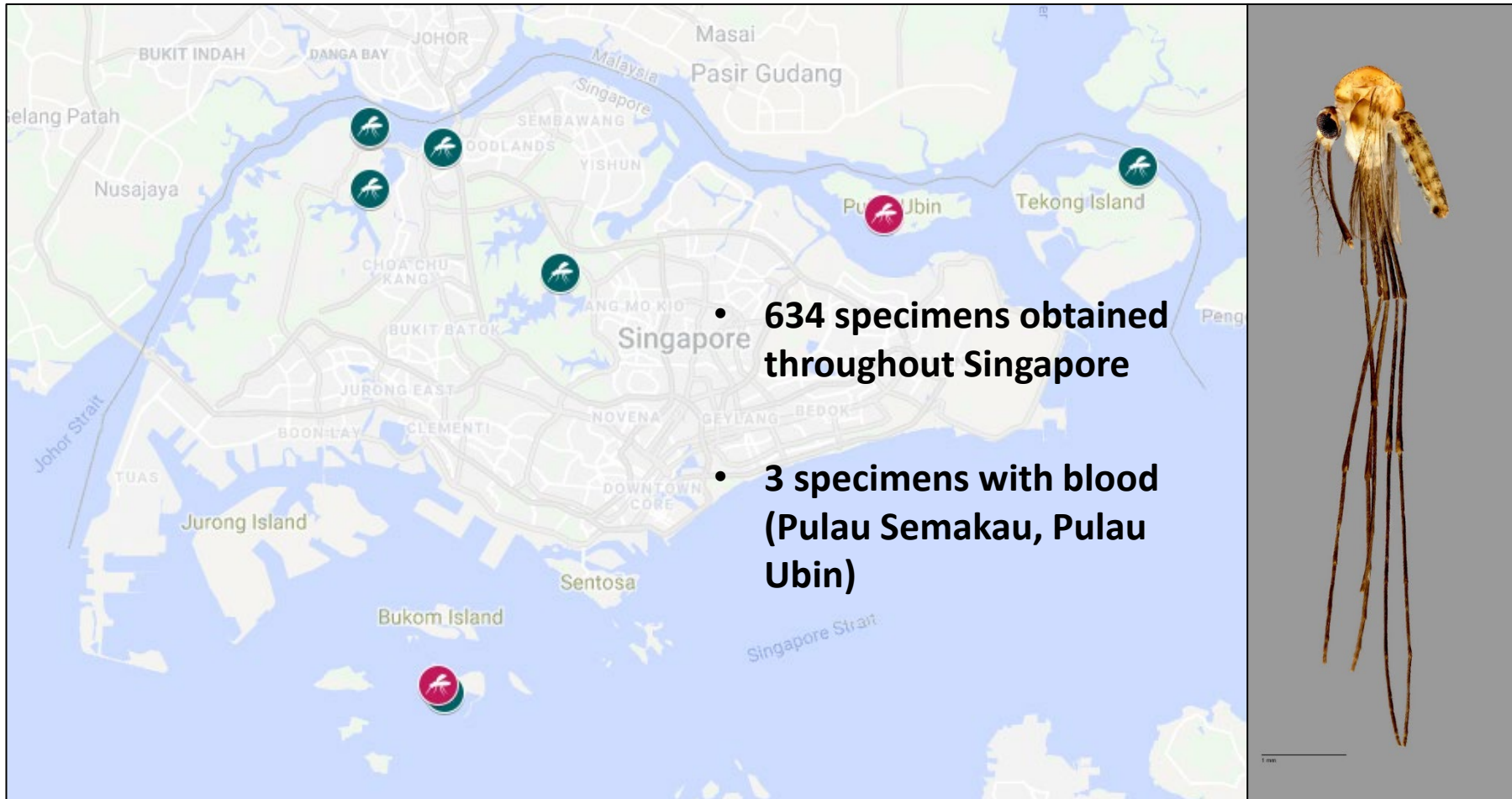


Aedes sp. ZRCBDP00009685



=> high potential as an agent transmitting viruses between distantly related hosts, but fortunately with restricted distribution

Uranotaenia sp. ZRCBDP0009257





*Leong
Tzi Ming*

Example 2: Dragonfly nymph ecology

Singapore's odonates (122 extant spp)





- 1500 specimens sequenced

=> larval and adult matches for 59 species

=> molecular work: 1-week

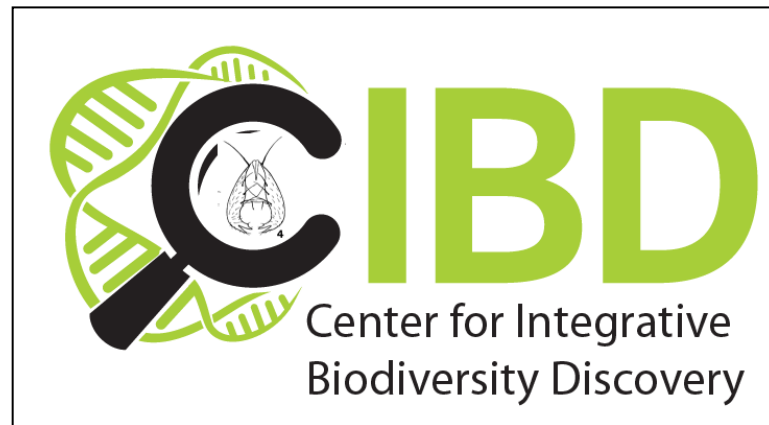


Shortcomings of the old ways...

- >80% of all species undescribed 
- 95% of all described species only identifiable for taxonomic experts 
- very little baseline biodiversity data from standardized samples 
- natural history data only available for <5% of species 

Conclusions

- Society expects us to monitor biodiversity holistically: we are failing
- We can catch up quickly as long as we use robotics, new DNA sequencers, and AI
- But we need “Dark Taxon Biology” to complement existing fields in biology
- Goal should be that biodiversity samples can be scanned like libraries
 - already realized for some samples and taxa
 - needs to be expanded to dark taxa





Amrita Srivathsan



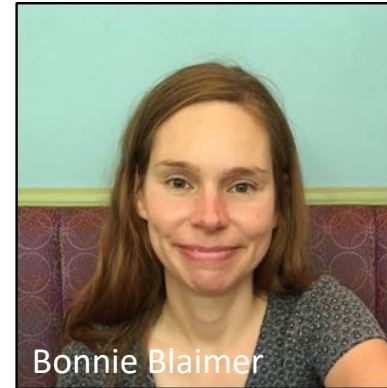
Ambrosio Torres



Bernardo Santos



Birger Neuhaus



Bonnie Blaimer



Dr. Liliia Serbina



Paco Hita-Garcia



Gregor Hagedorn



Johannes Frisch



Jörg Freyhof



Josh Penalba



Kristina von Rintelen



Michael Ohl



Theo Leger



Thomas von Rintelen



Dr. Alex van Dam

