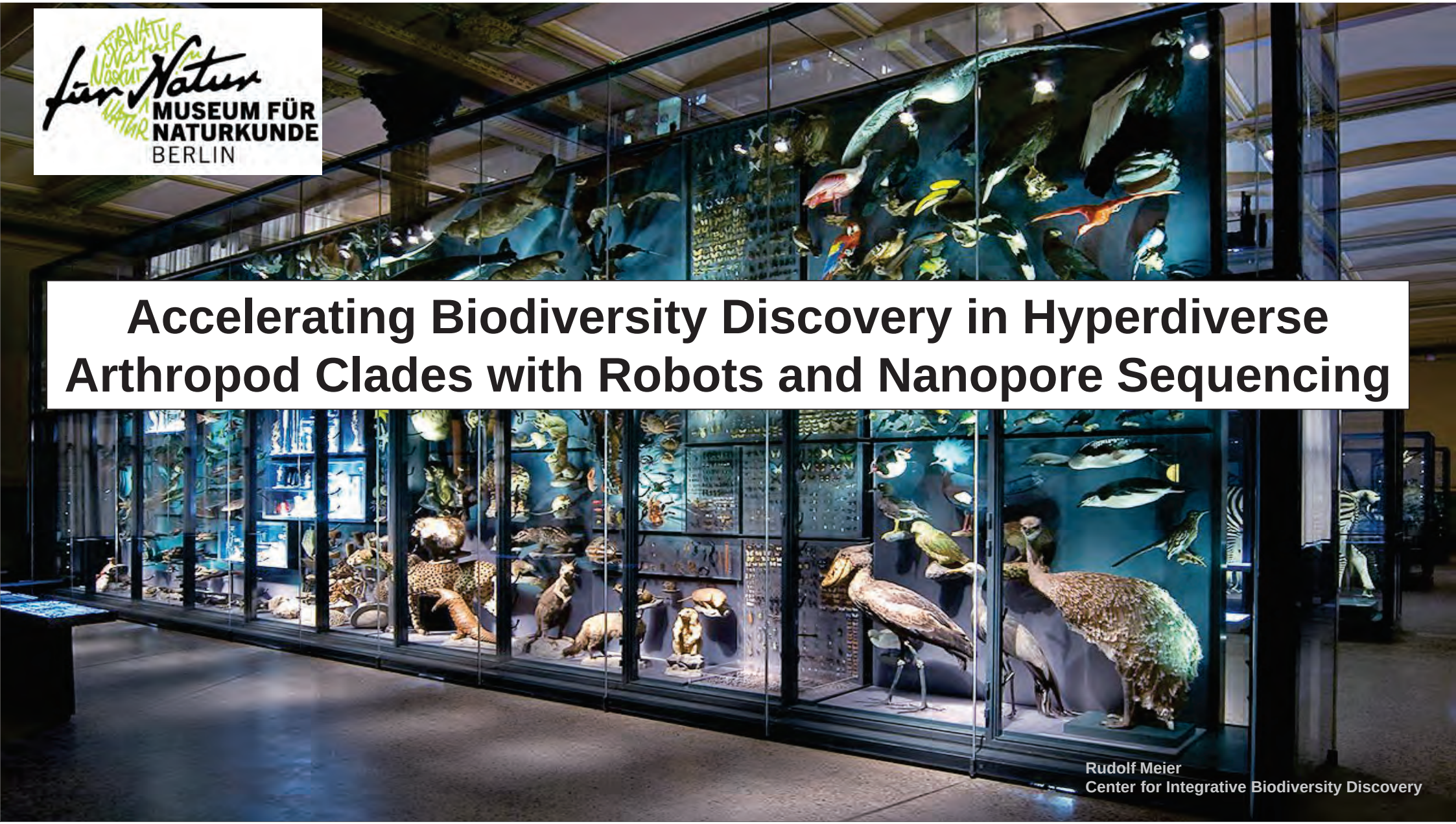


Accelerating Biodiversity Discovery in Hyperdiverse Arthropod Clades with Robots and Nanopore Sequencing

Rudolf Meier
Center for Integrative Biodiversity Discovery

Our planet has a big biodiversity problem

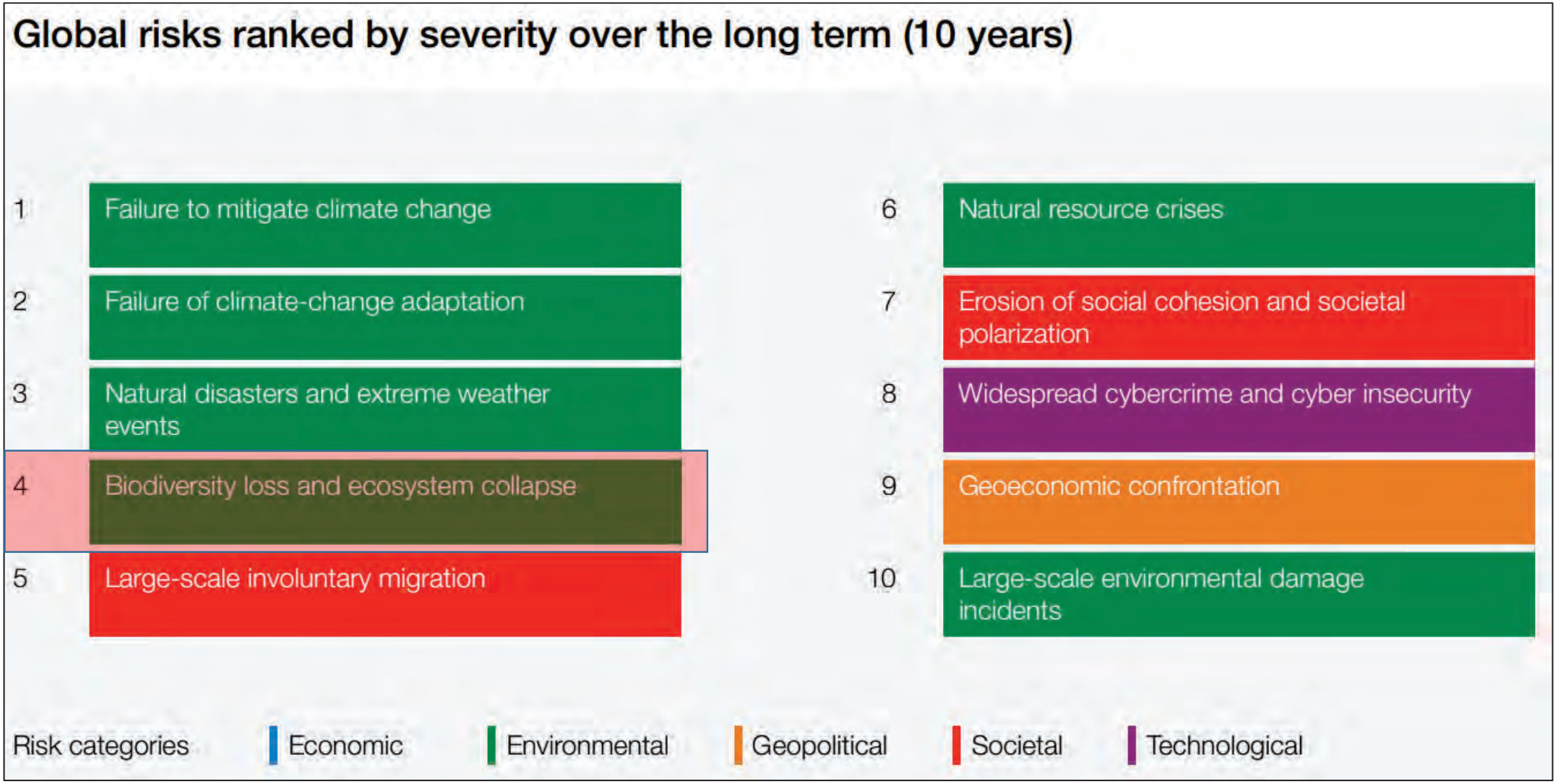
1. Biodiversity Loss: threatens the survival of human societies

2. Lack of unbiased data to address the problem

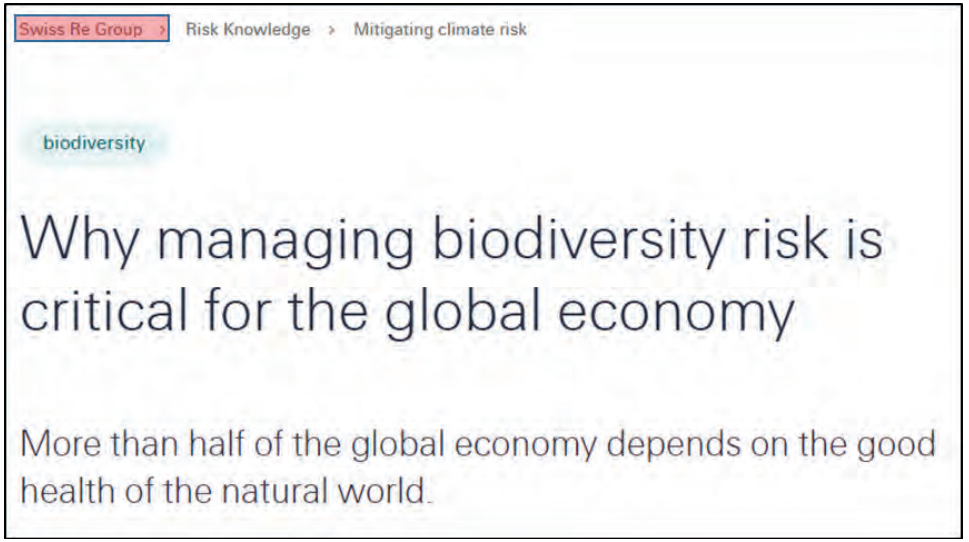
The Global Risks Report 2023
18th Edition

WORLD ECONOMIC FORUM

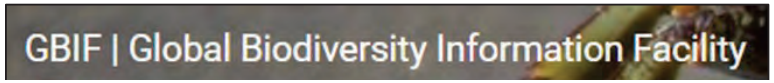
Global risks ranked by severity over the long term (10 years)



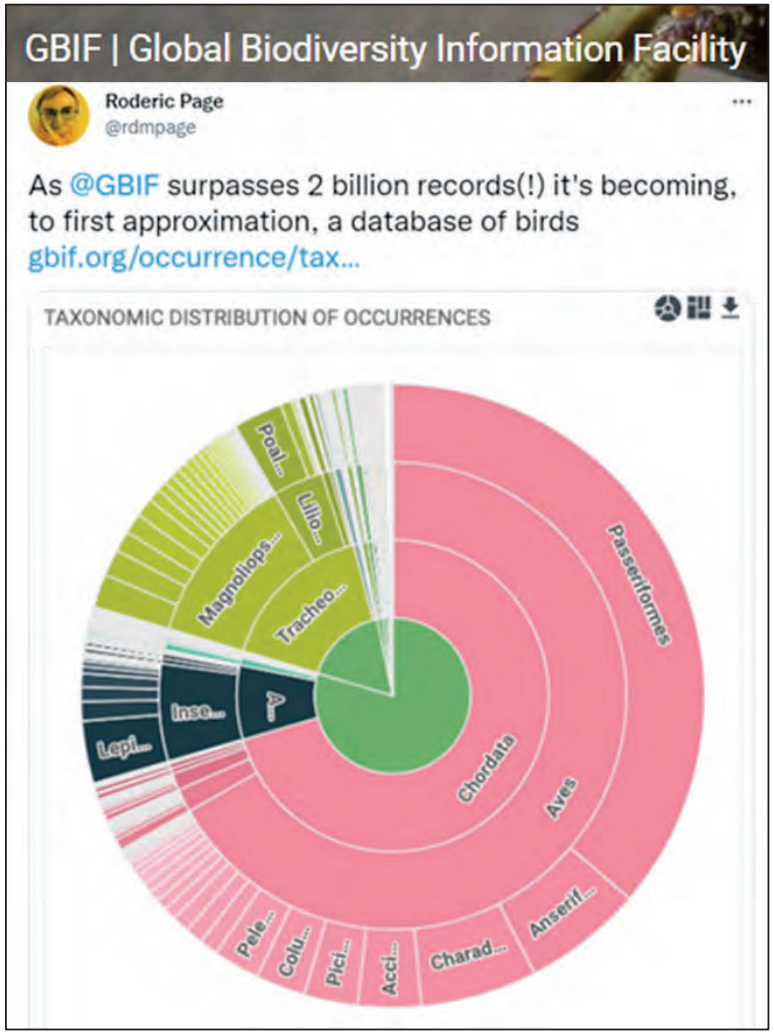
Not only the World Economic Forum…



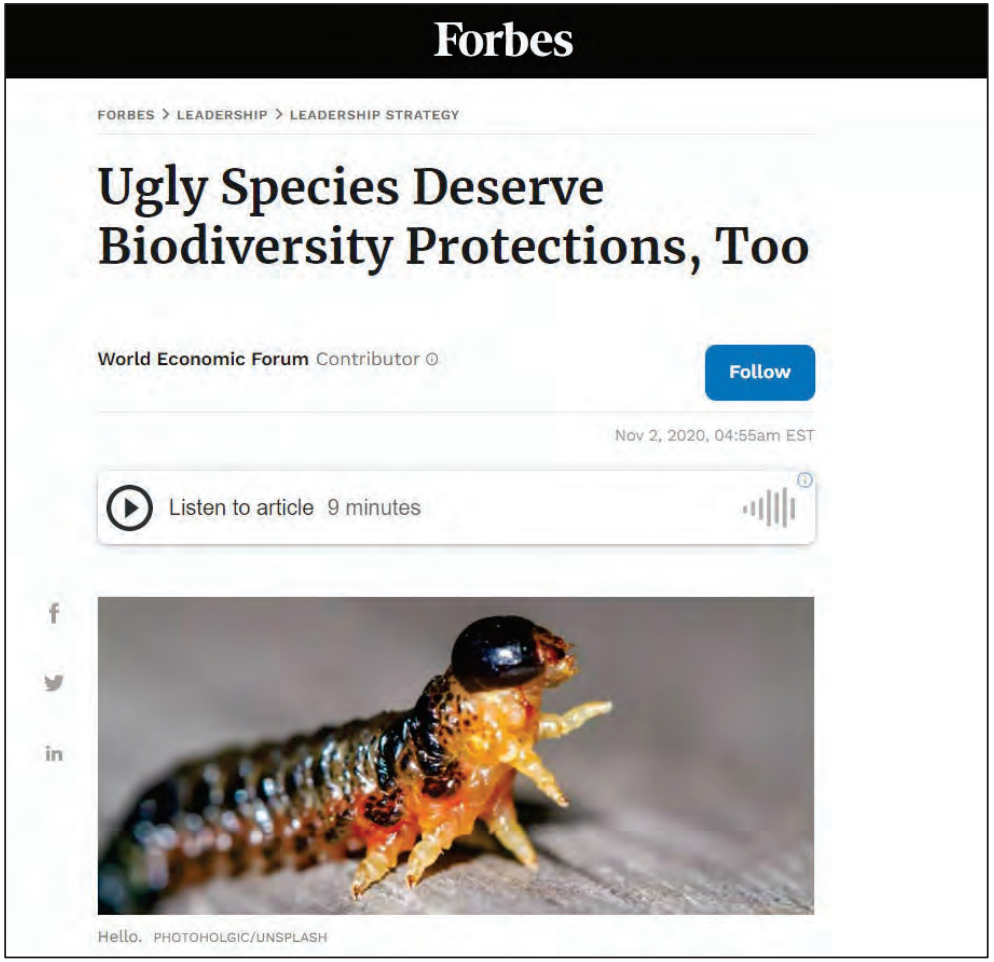
2. Lack of unbiased data to address the problem



2. Lack of unbiased data to address the problem



2. Lack of unbiased data to address the problem



Forbes

FORBES > LEADERSHIP > LEADERSHIP STRATEGY

Ugly Species Deserve Biodiversity Protections, Too

World Economic Forum Contributor ⓘ

Follow

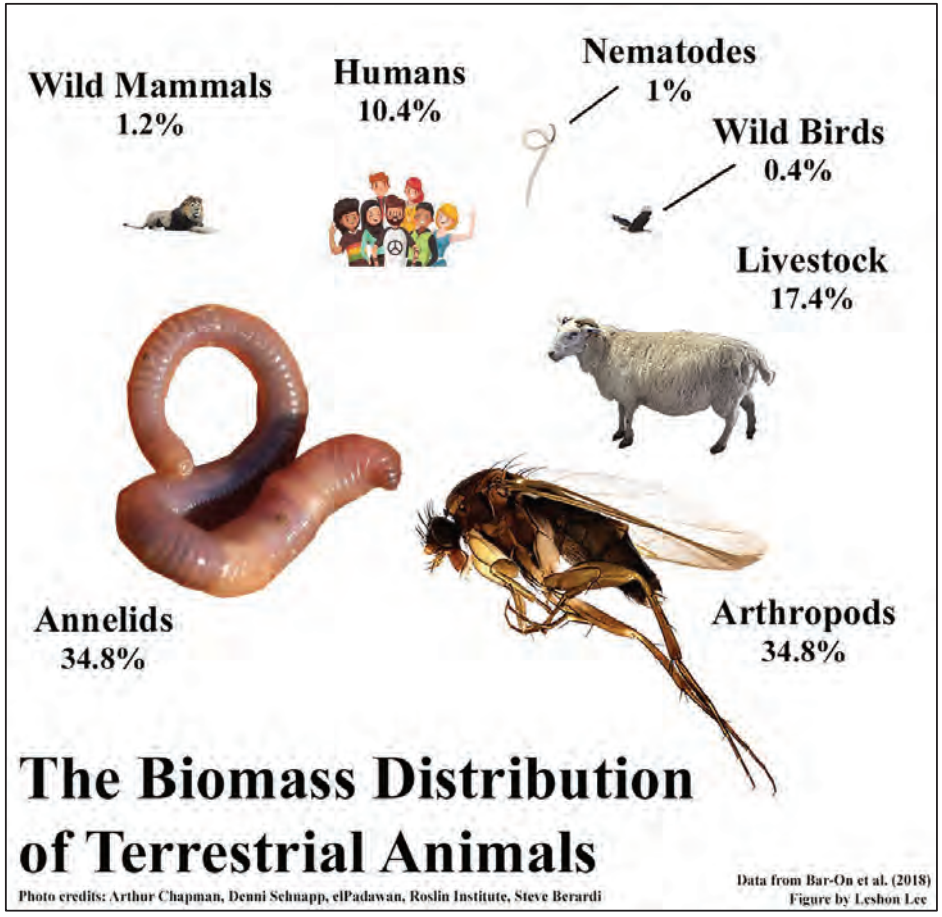
Nov 2, 2020, 04:55am EST

Listen to article 9 minutes

Hello. PHOTOHOLGIC/UNSPLASH

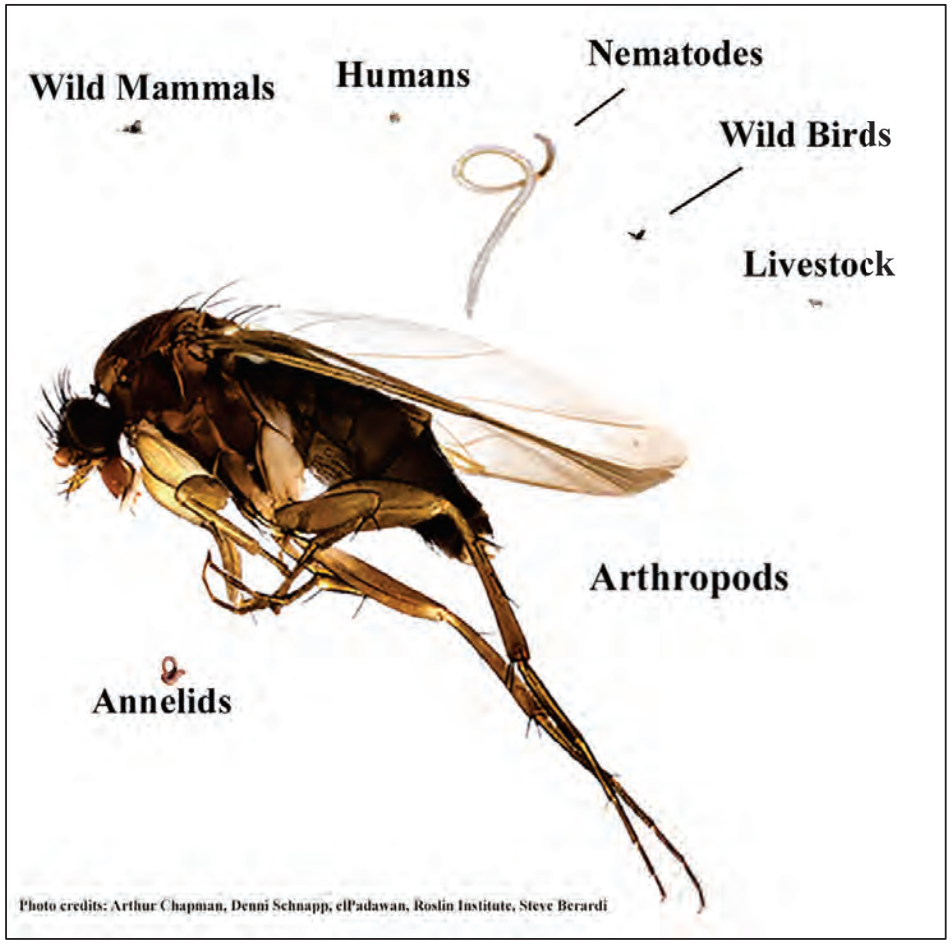
“Ugly” species matter

Taxon bias:
terrestrial animal biomass



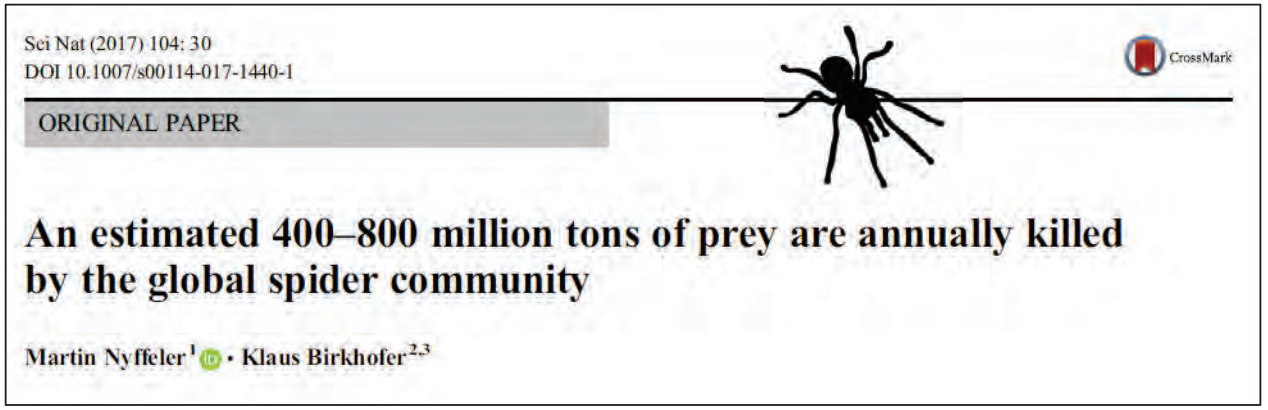
“Ugly” species matter

Taxon bias:
animal species diversity



“Ugly” species matter

Taxon bias:
Ecosystem impact


Sci Nat (2017) 104: 30
DOI 10.1007/s00114-017-1440-1

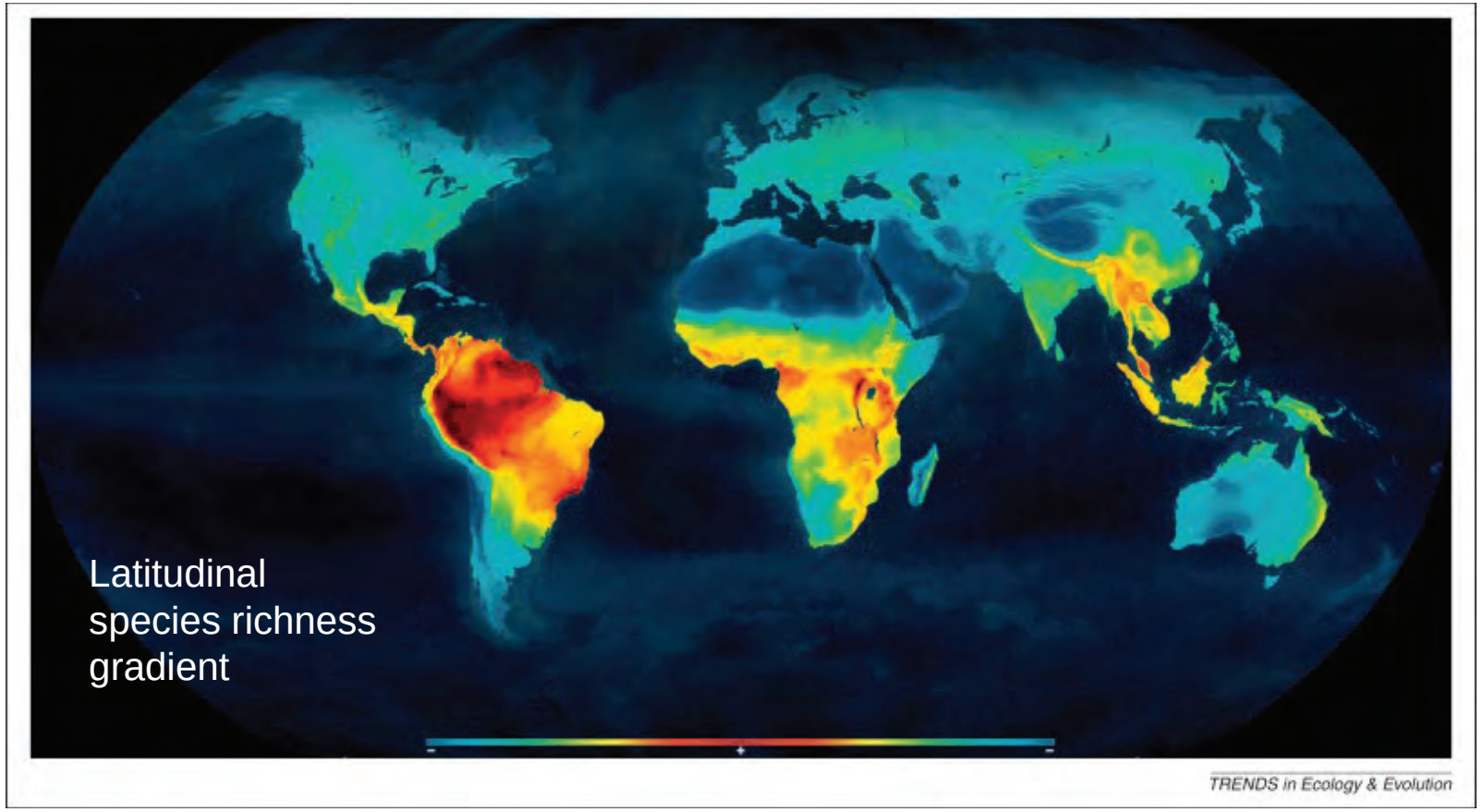
ORIGINAL PAPER

An estimated 400–800 million tons of prey are annually killed by the global spider community

Martin Nyffeler[1] · Klaus Birkhofer[2,3]

Humans: annual meat/fish consumption: 400 mio. tons

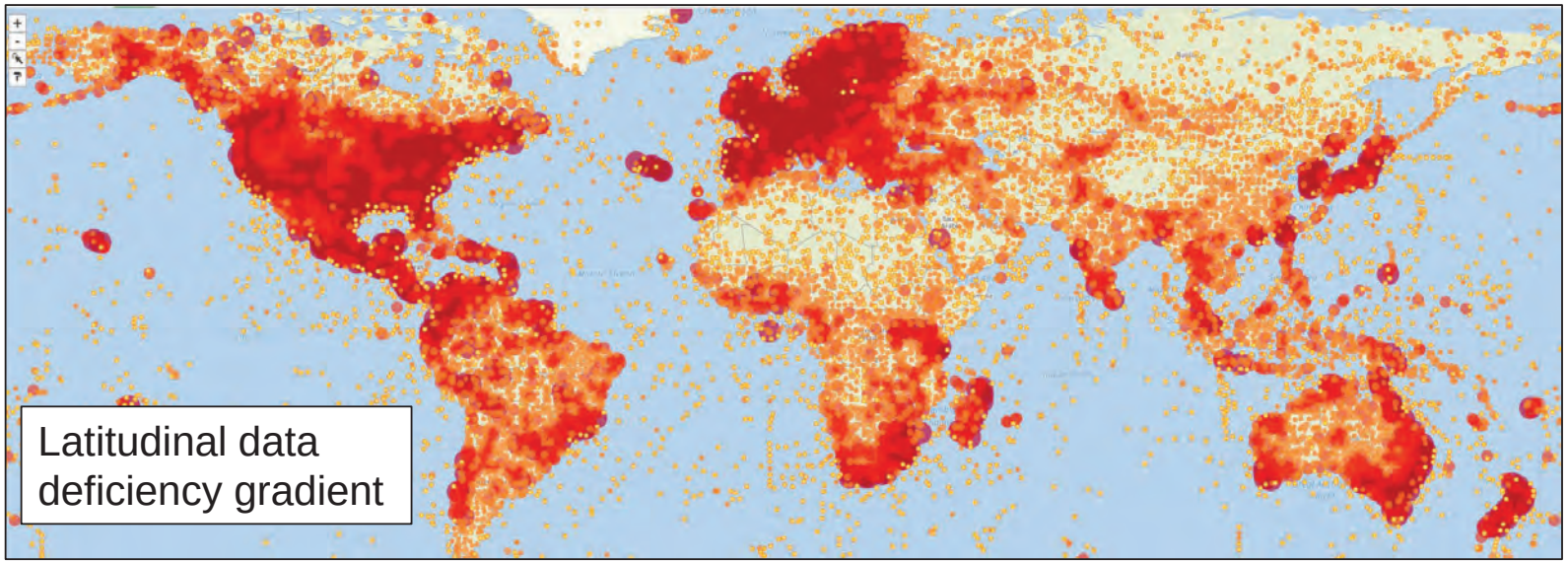
Biodiversity data problem is worse at the global scale

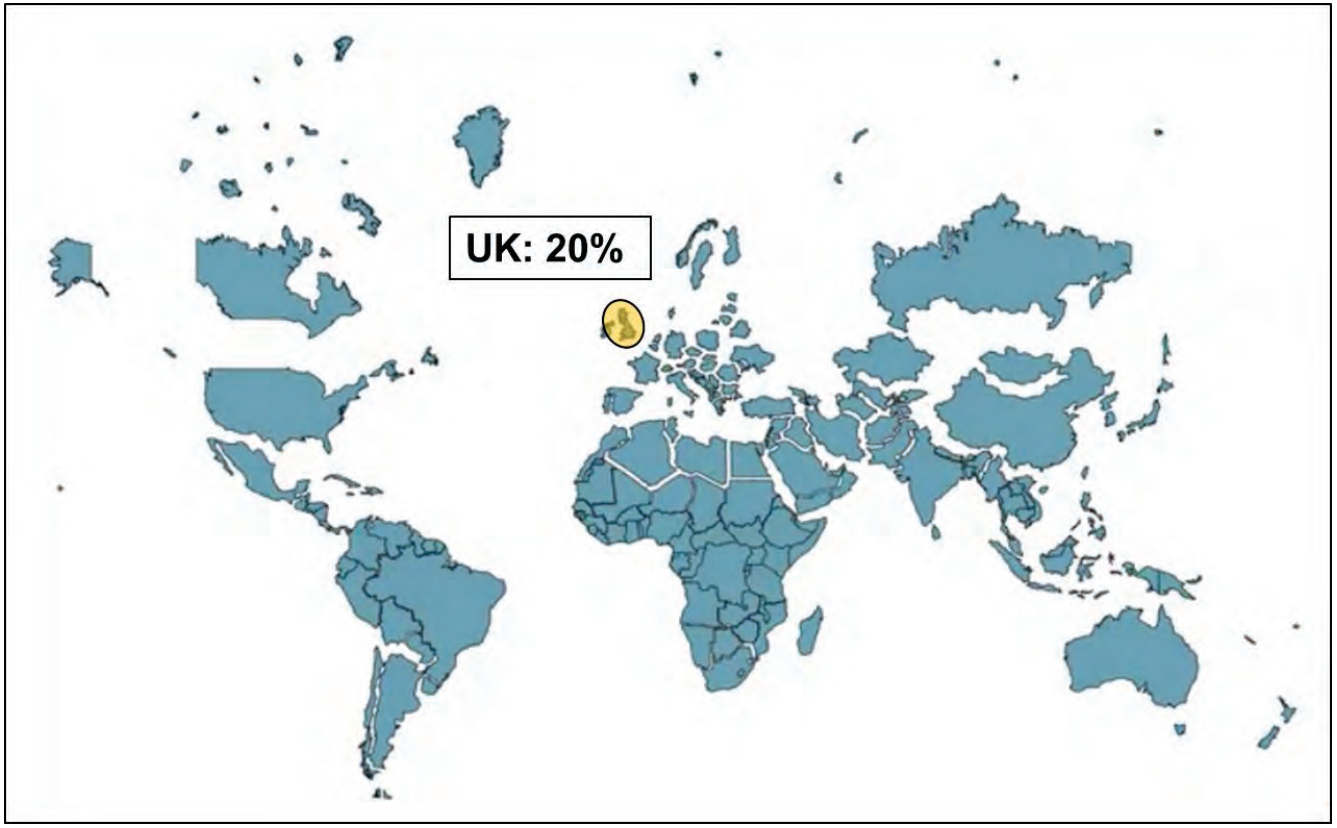


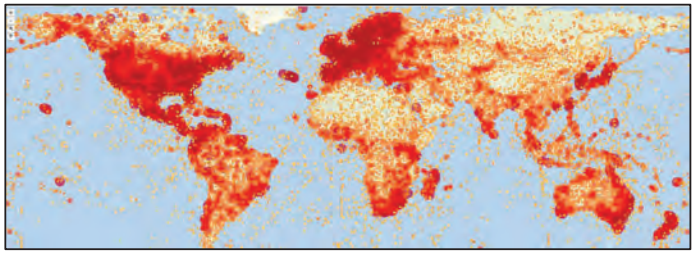
GBIF insect data

Insect data in GBIF:

- 169 million records
- 47 million specimens





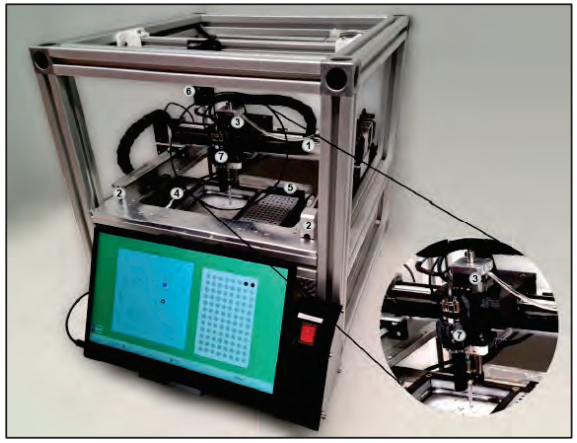


Biodiversity: overcome the taxon and geographic biases

- we need cost-effective methods for collecting data
- methods have to be suitable for machines and humans alike

Species Discovery Factory

1. DiversityScanner:
Digitization with robots

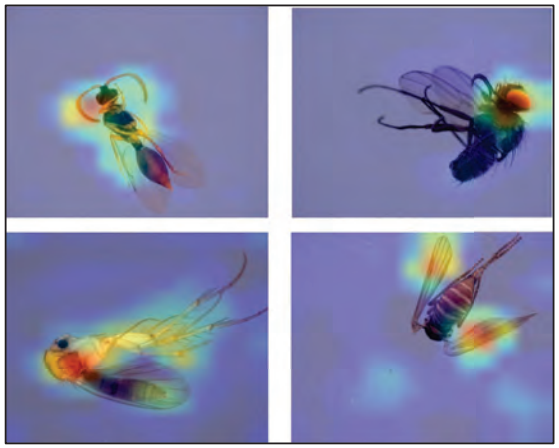


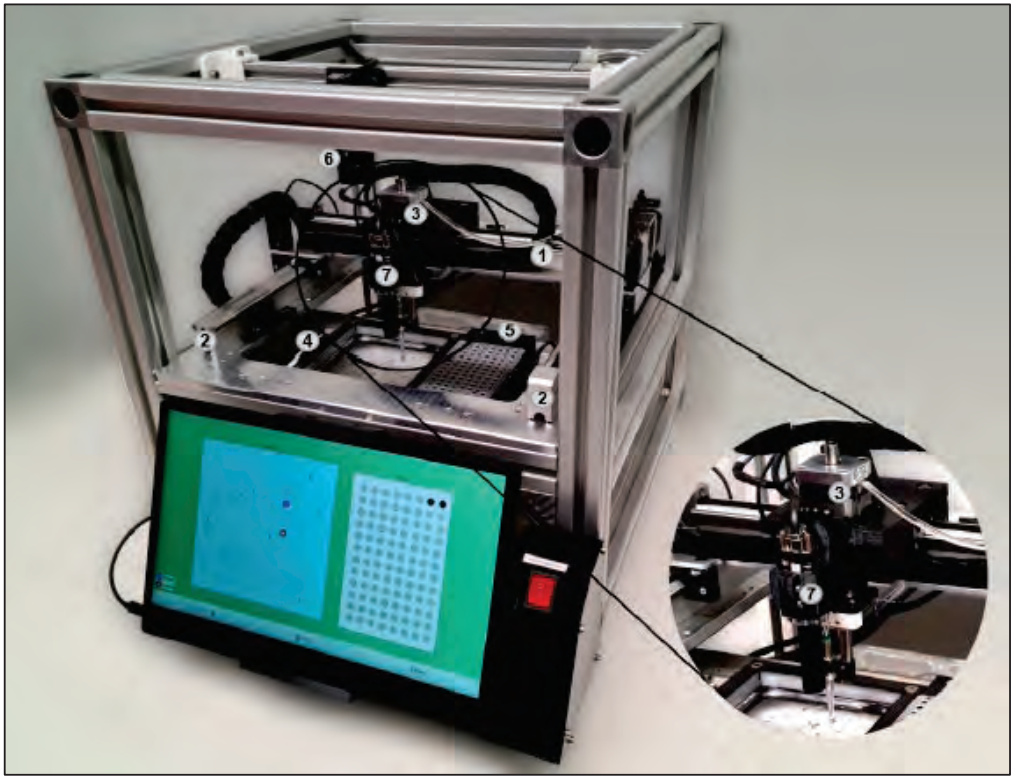
+

2. Nanopore Sequencing:
Sorting with DNA barcodes

=

3. Identified images:
biomonitoring with AI





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

MOLECULAR ECOLOGY
RESOURCES

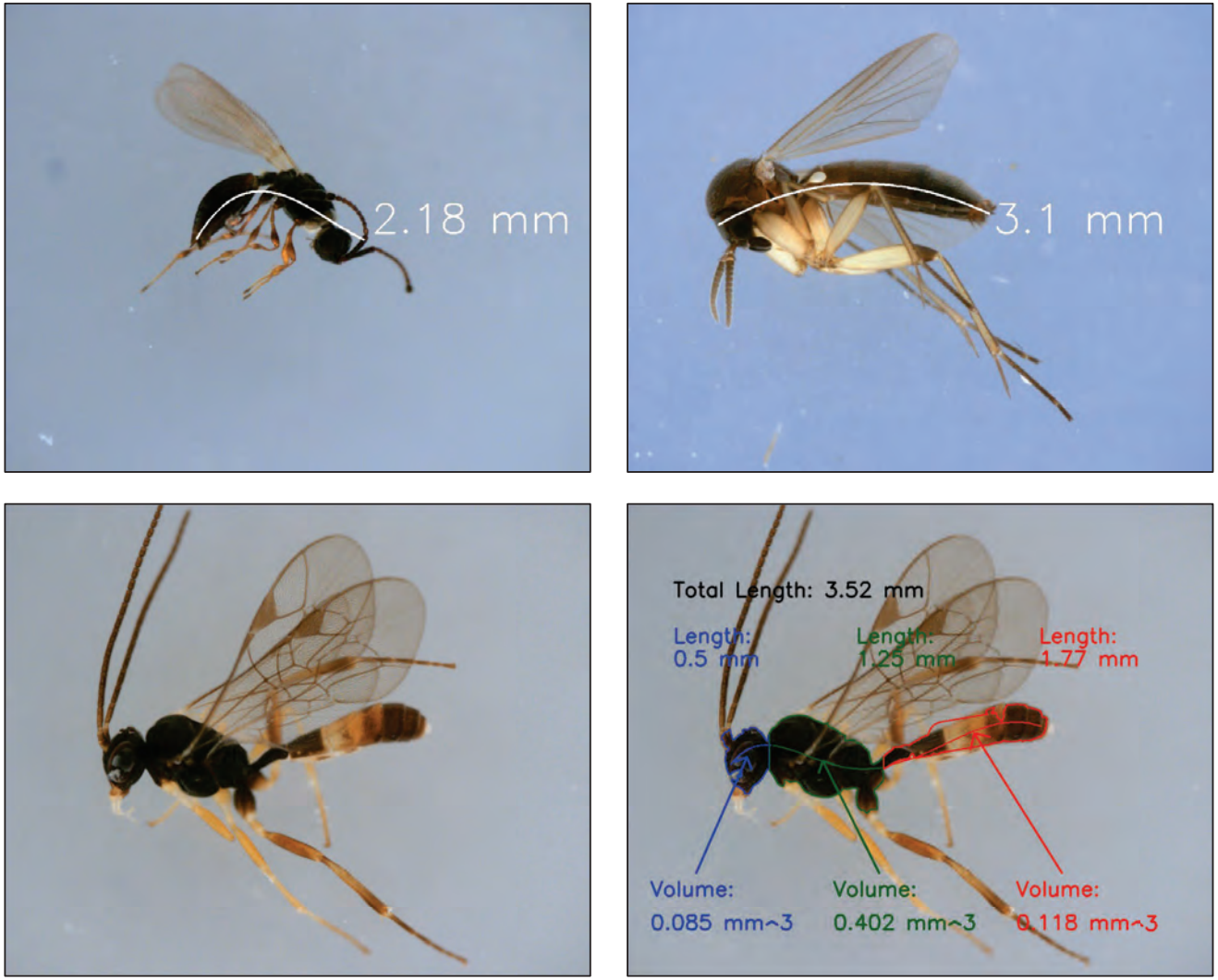
RESOURCE ARTICLE | Open Access

DiversityScanner: Robotic handling of small invertebrates with machine learning methods

Lorenz Wührl, 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

DiversityScanner
Biomass estimation

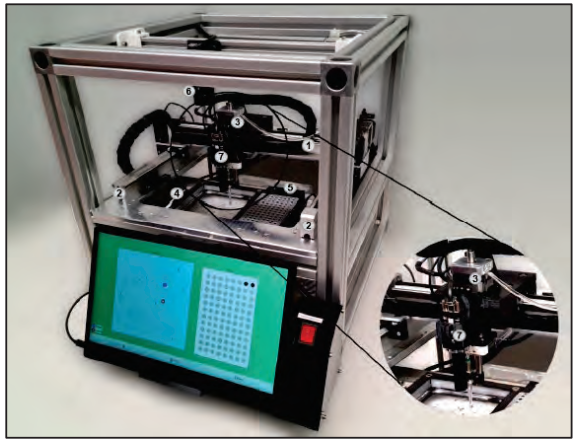







Species Discovery Factory

1. DiversityScanner:
Digitization with robots

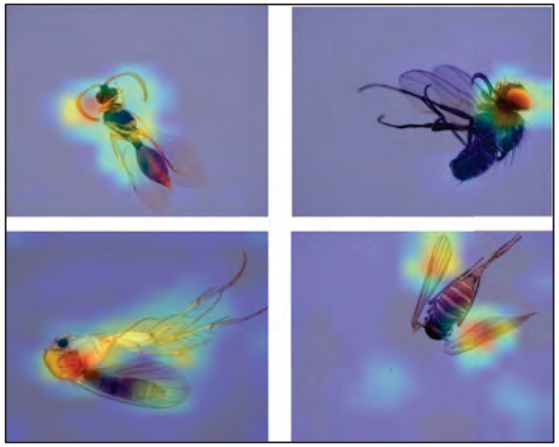


+

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

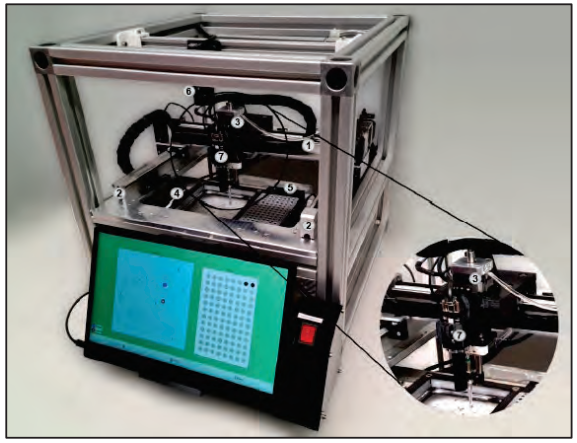
ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone

Amrita Srivathsan[1], Leshon Lee[1], Kazutaka Katoh[2,3], Emily Hartop[4,5], Sujatha Narayanan Kutty[1,6], Johnathan Wong[1], Darren Yeo[1] and Rudolf Meier[1,7*]

Mobile labs for biodiversity discovery:
- simplified, fast, and robust techniques for specimen-based barcoding (10 cents/specimen)
- MinION: pocket-size sequencer

Species Discovery Factory

1. DiversityScanner:
Digitization with robots

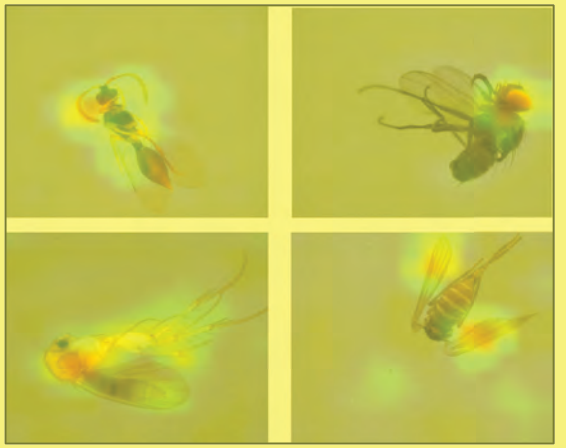


+

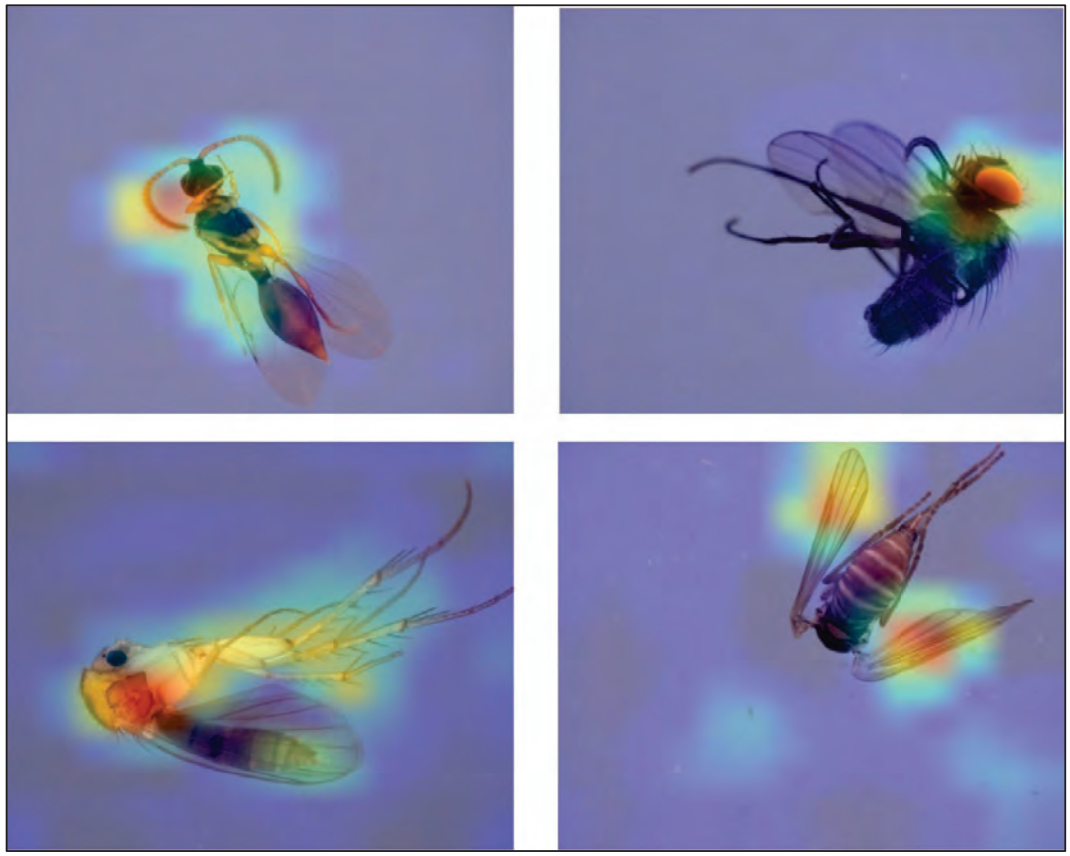
2. Nanopore Sequencing:
Sorting with DNA barcodes

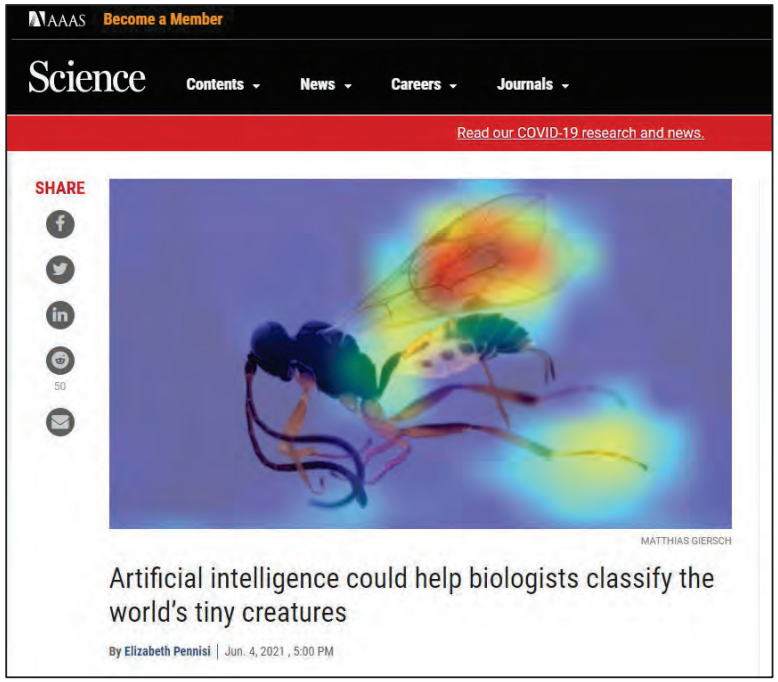
=

3. Identified images:
biomonitoring with AI



3. Identification with AI: training of Convolutional Neural Networks





3. Identification with AI: training of Convolutional Neural Networks

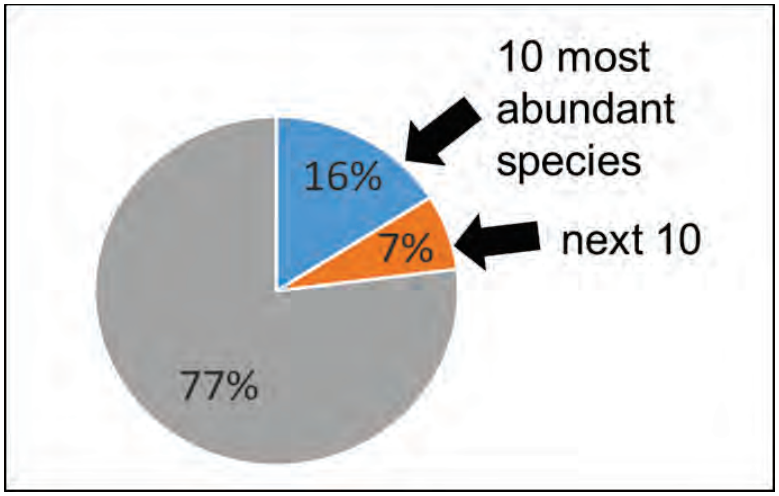
Class (Taxon)	Result	Class (Taxon)	Result
Diptera Acalyptratae	91%	Diptera Psychodidae	89%
Diptera Calyptratae	83%	Diptera Sciaridae	92%
Diptera Cecidomyiidae	91%	Hemiptera Cicadellidae	100%
Diptera Chironomidae	97%	Hymenoptera Braconidae	82%
Diptera Dolichopodidae	86%	Hymenoptera Diapriidae	100%
Diptera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

AI identification tools evolve

- common taxa benefit first
- evolve from family => genus => species

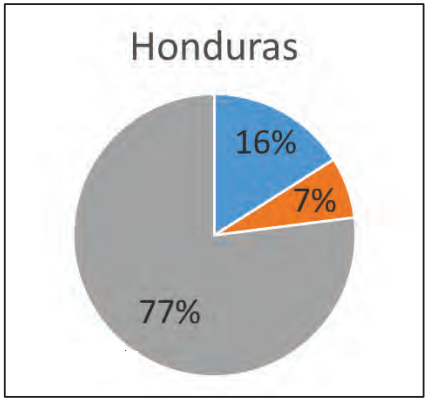
For how many species do we need AI identification tools?

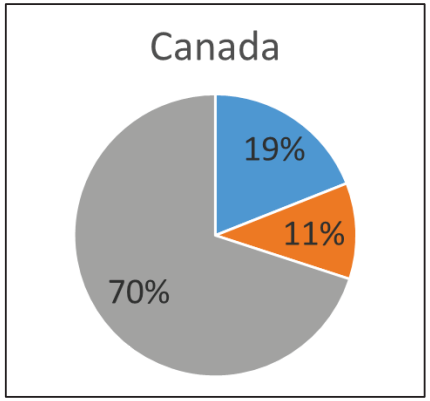
Honduras:
7229 specimens in >1700 species

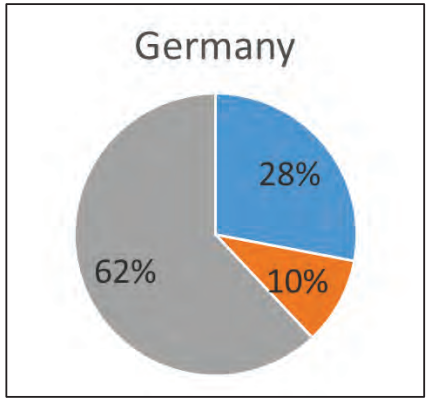


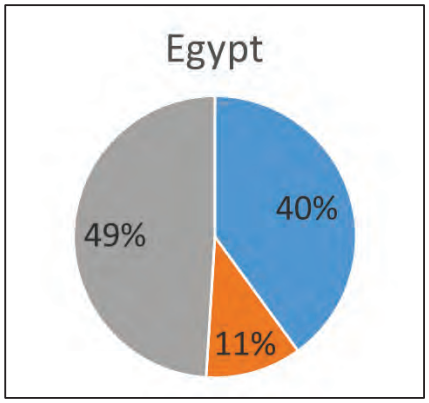
=> 23% of all specimens can be covered by AI tools for 20 species

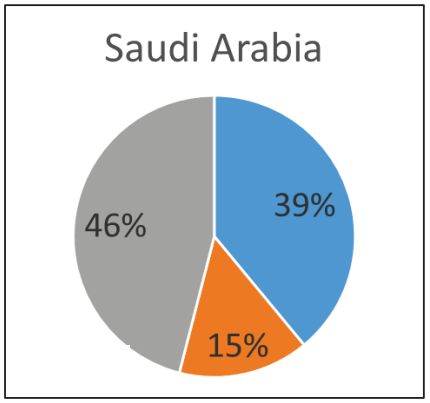
For how many species do we need AI identification tools?

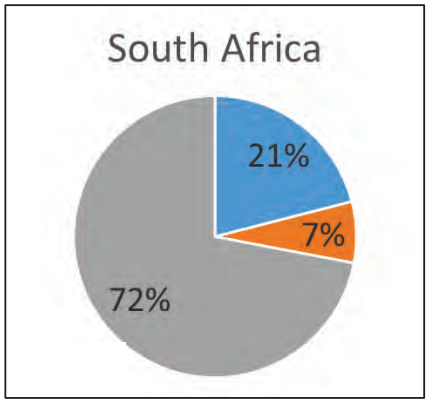


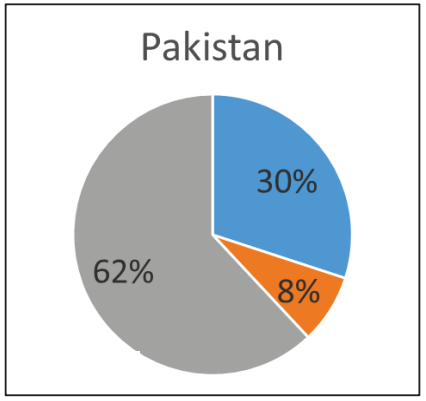


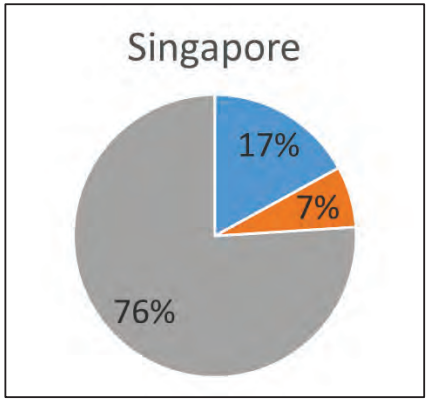












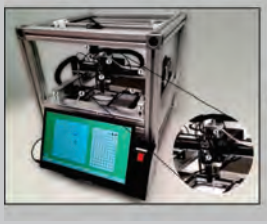
How to get started?

Natural environments:
Step 1: process 1000 specimens for 500 sites/sample (half a million specimens)
=> discovery of 25,000 species
Step 2: image numerically important species and train AI tools
=> AI tools for thousands of common species

Agricultural or forestry environments:
fewer important species => fewer specimens need to be processed

Species Discovery Factory

1. DiversityScanner:
Digitization with robots

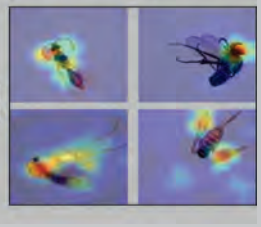


+

2. Nanopore Sequencing:
Sorting with DNA barcodes

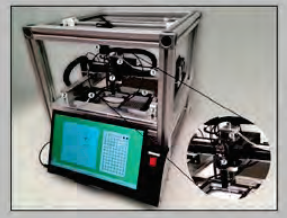
=

3. Identified images:
biomonitoring with AI



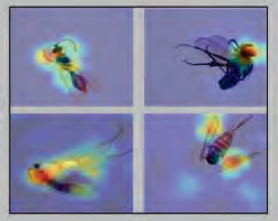
Species Biomonitoring Factory

1. DiversityScanner:
Digitization with robots



=

2. Identified images:
biomonitoring with AI



Environmental impact of species: Molecular Ecology

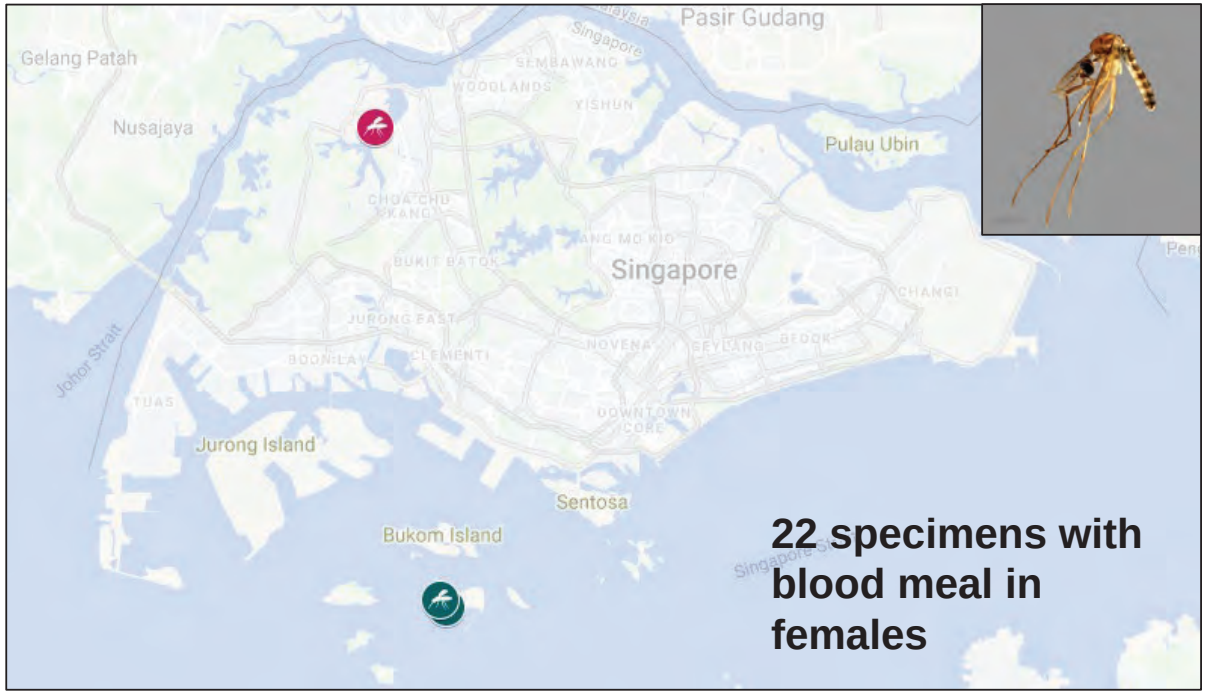
- 4000 mosquitoes from Singapore processed

=> ca. 120 species detected

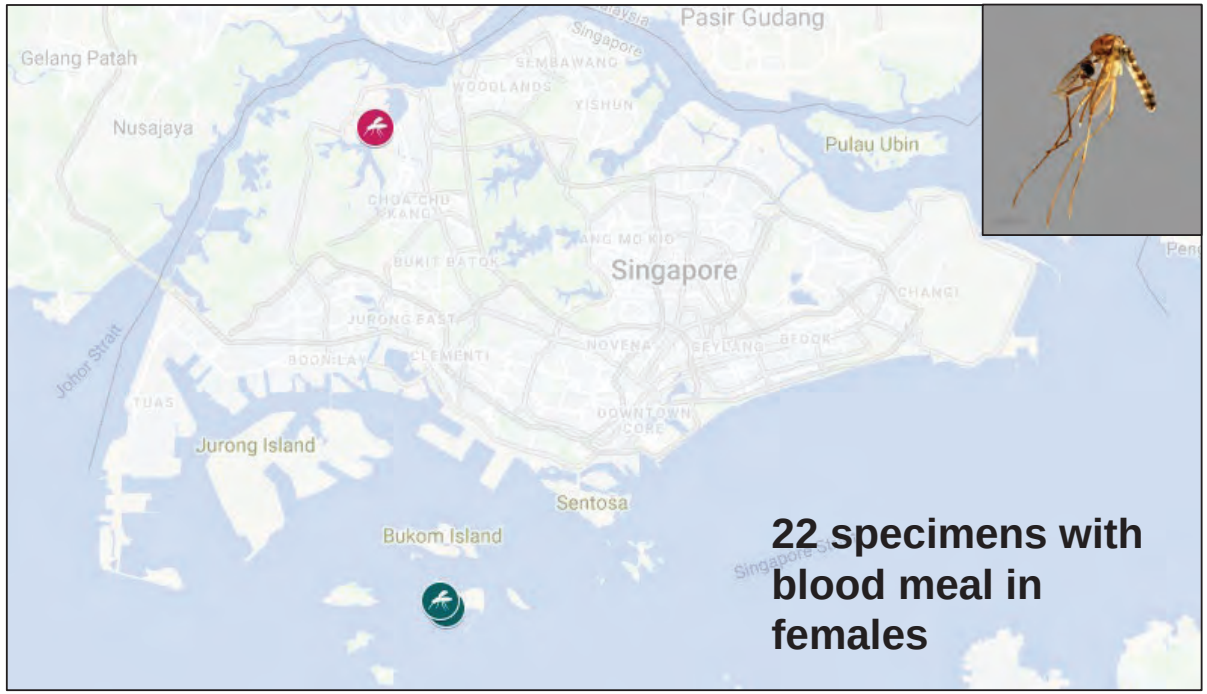
=> females for ca. 10 species had visible blood meals

=> metabarcoding of gut content

Aedes sp. ZRCBDP00009685

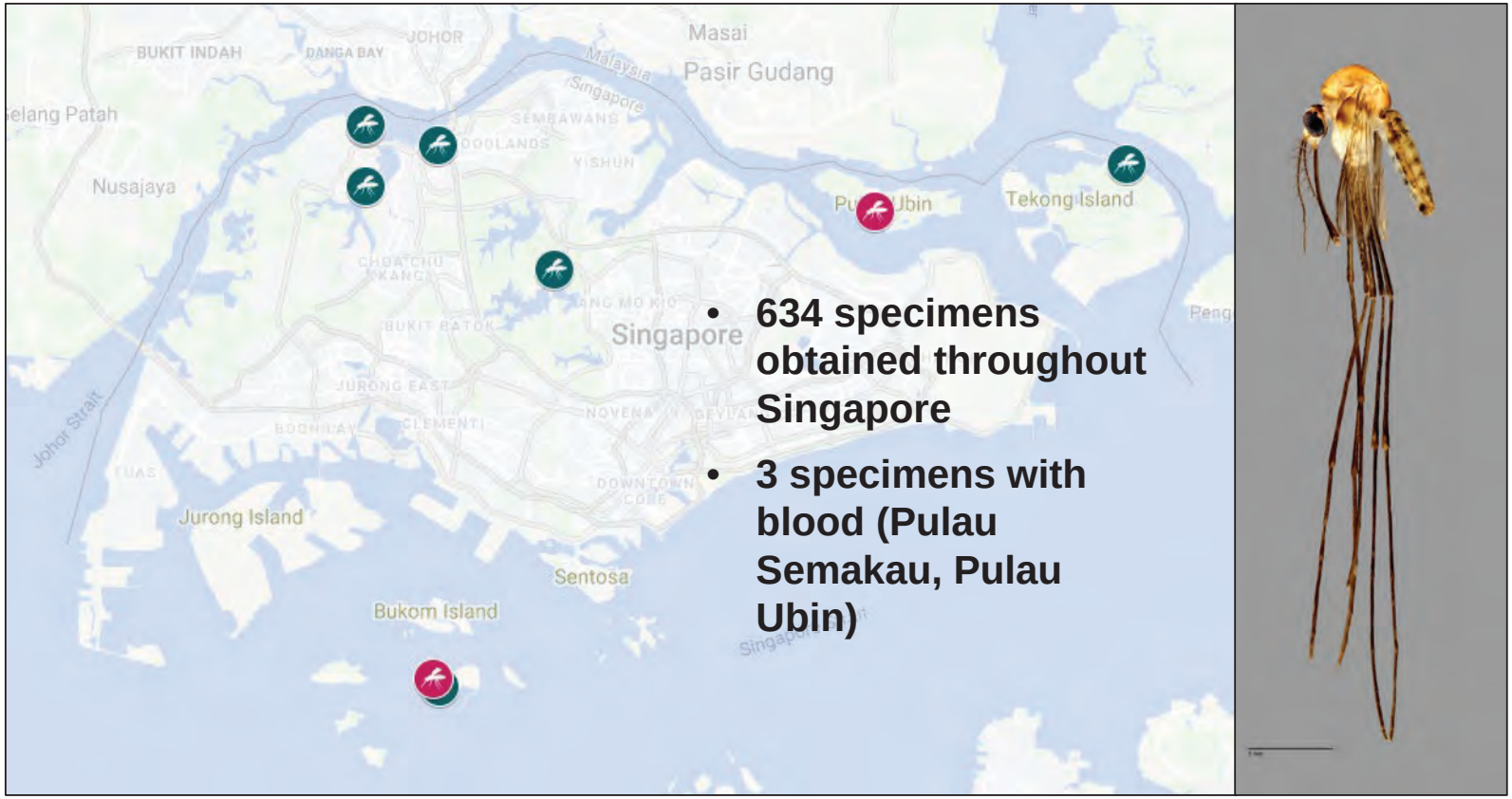


Aedes sp. ZRCBDP00009685



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

Uranotaenia sp. ZRCBDP0009257




Leong
Tzi Ming

Conclusions

- The world expects us to monitor biodiversity holistically
- A major challenge is not enough and biased data
- Particularly data deficient are insects although they have high abundance and species diversity
- Fortunately, automation and machine learning can overcome this challenge

=> cost-effective data generation tools for machines and humans are feasible