

DIVERSITY MEASUREMENT IN MICROBIAL ECOLOGY

Aline Potiron, Utrecht University

<https://apotiron.github.io>



INTRODUCTION

Microbial Ecology: Microorganisms are fundamentals for our way of life.

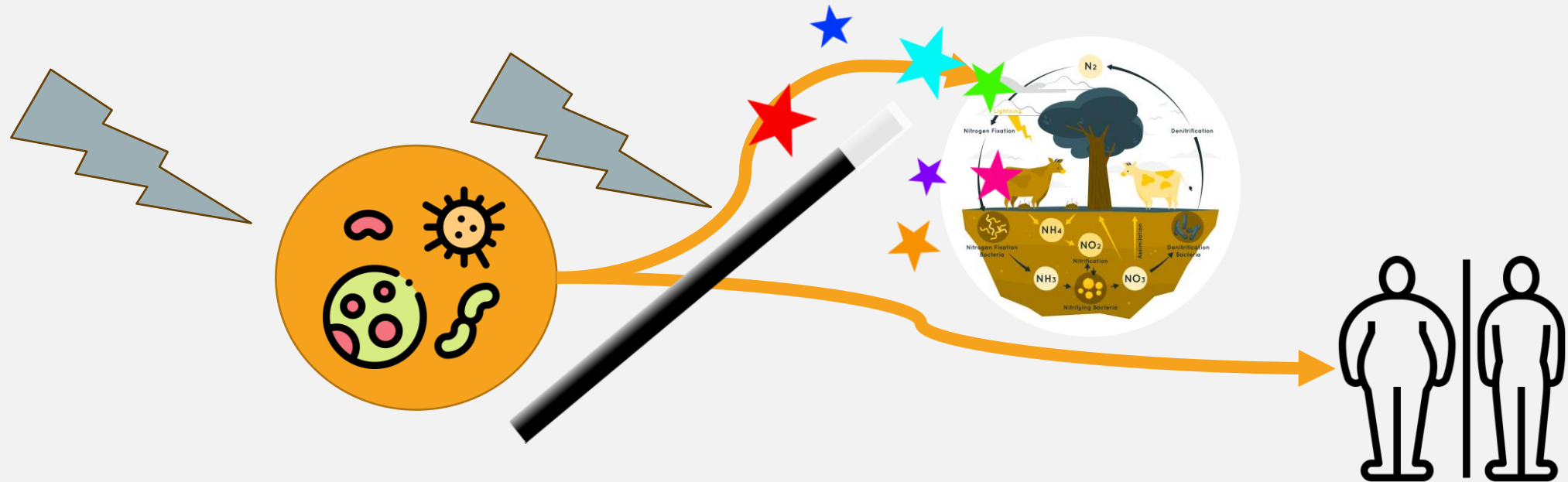
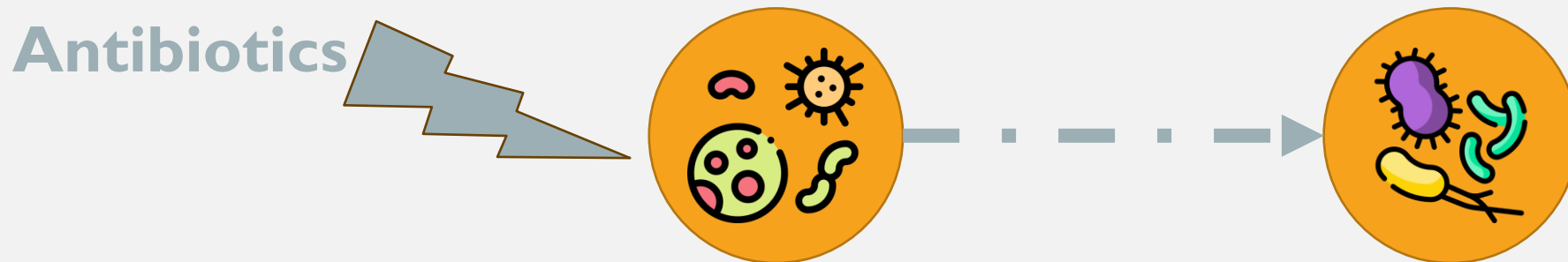


Image by [Clker-Free-Vector-Images](#) from [Pixabay](#)

INTRODUCTION

Intuitive idea of **diversity**: Disturbances (e.g., taking antibiotics) in the community will translate into changes in that community. These changes are real and mind-independent (e.g., the number and types of individuals in a community).

- **Diversity needs to be measured**



INTRODUCTION

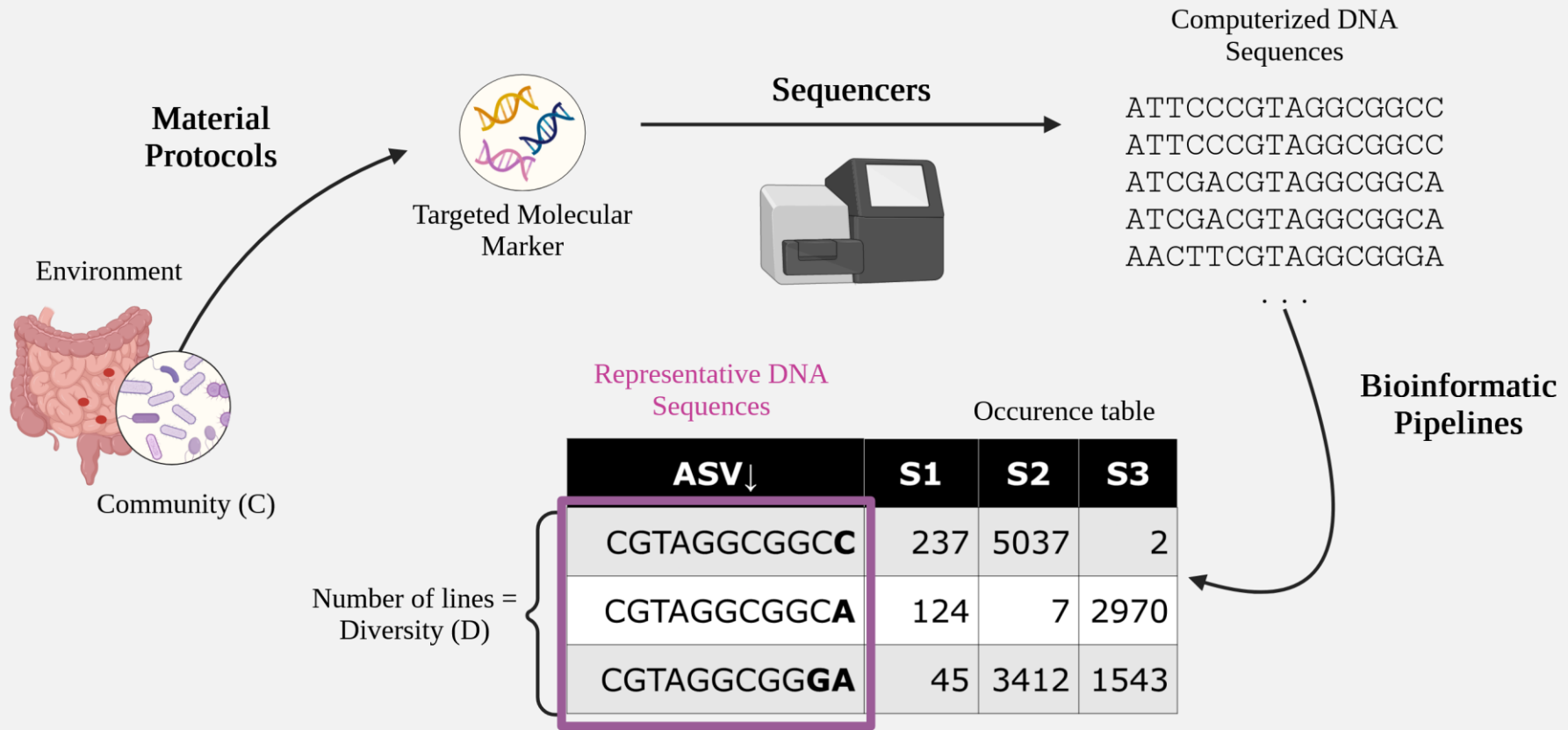
- Diversity is a **relative** notion: relative to the spatial area and the community chosen.
- The **concept** of diversity is **ambiguous**: taxonomic, genetic, functional, and phylogenetic diversity. It is associated with societal and ethical values.
- Biological diversity is measured: turn to **the practice of measurement**



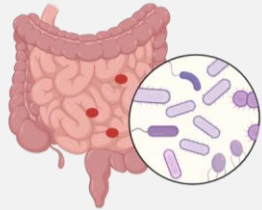
RESEARCH QUESTION

- How upstream choices (e.g., species definition) influence the value of diversity measured?
 1. Propose an analysis of diversity in microorganismal communities as a measurement process using the model-based account (MBA) of measurement.
 2. I review challenges that the MBA does not help settle.
 3. I outline the role of the species concept.

AMPLICON SEQUENCING



Environment



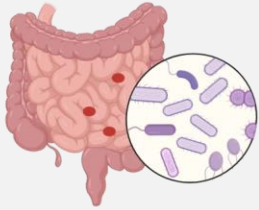
Community (C)

System under
measurement

MODEL OF AMPLICON SEQUENCING

A.

Environment



Community (C)

The measurand is the
diversity (D)

B.

Idealization



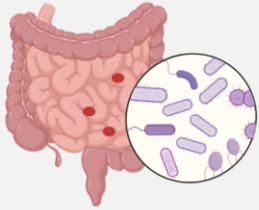
Targeted Molecular
Marker (C*)

The measurand is
modeled as the
species richness (D*)



A.

Environment



Community (C)

B.

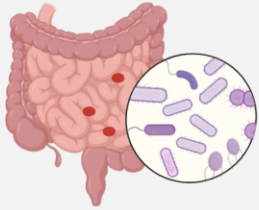
Idealization



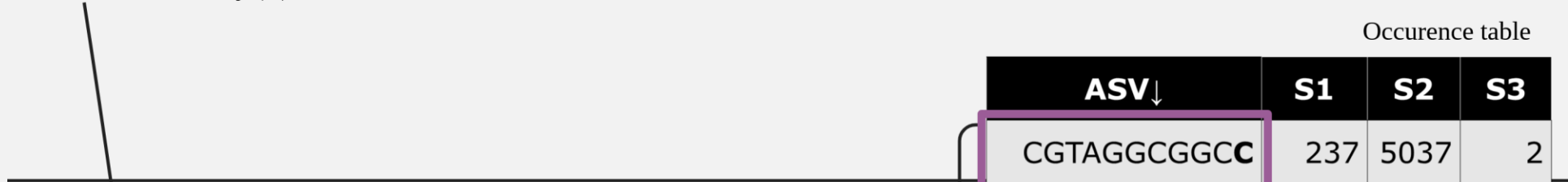
Targeted Molecular
Marker (C*)

A.

Environment



Community (C)



B.

Idealization



Targeted Molecular
Marker (C*)

Number of Z* =
Diversity (D*)

Occurrence table

ASV↓	S1	S2	S3
CGTAGGCGGCC	237	5037	2
CGTAGGCGGCA	124	7	2970
CGTAGGCGGGA	45	3412	1543

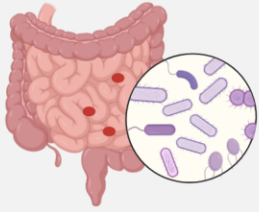
Representative DNA
Sequences (Z*)

ATTCCCGTAGGCGGCC
ATTCCCGTAGGCGGCC
ATCGACGTAGGCGGCA
ATCGACGTAGGCGGCA
AACTTCGTAGGCGGGA

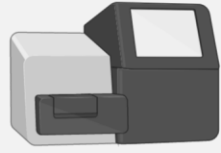
...
Computerized DNA
Sequences (Y*)

A.

Environment



Community (C)



Sequencers (mS)

Bioinformatic Pipelines (mB)

Occurrence table

ASV↓	S1	S2	S3
CGTAGGCGGCC	237	5037	2
CGTAGGCGGCA	124	7	2970
CGTAGGCGGGA	45	3412	1543

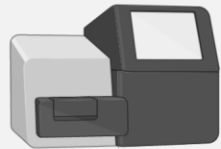
B.

Idealization



Targeted Molecular Marker (C*)

Ideal Sequencers (mS*)



Number of Z* = Diversity (D*)

ATTCCCGTAGGCGGCC
 ATTCCCGTAGGCGGCC
 ATCGACGTAGGCGGCA
 ATCGACGTAGGCGGCA
 AACTTCGTAGGCGGGA
 . . .

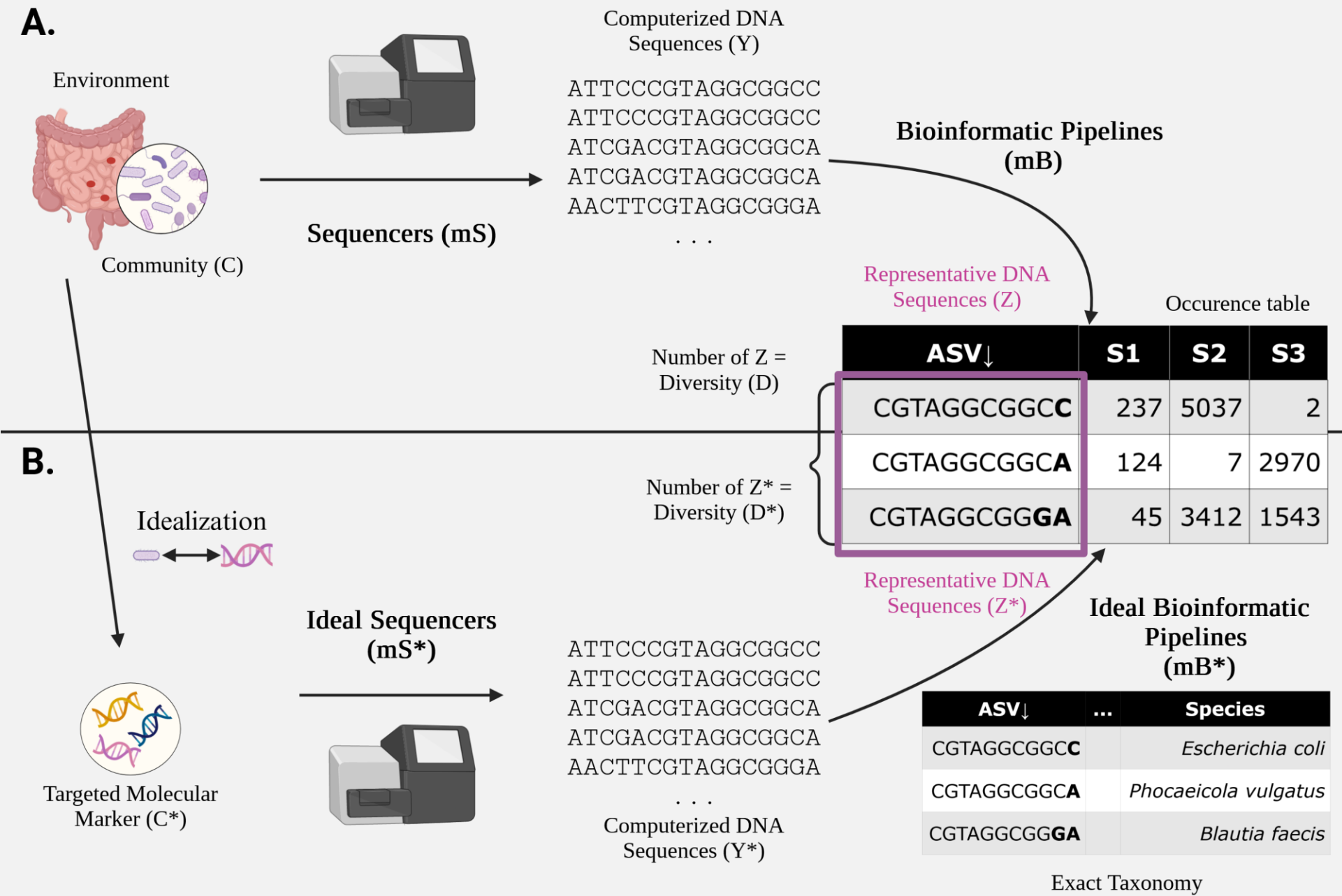
Computerized DNA Sequences (Y*)

Representative DNA Sequences (Z*)

Ideal Bioinformatic Pipelines (mB*)

ASV↓	...	Species
CGTAGGCGGCC		<i>Escherichia coli</i>
CGTAGGCGGCA		<i>Phocaeicola vulgatus</i>
CGTAGGCGGGA		<i>Blautia faecis</i>

Exact Taxonomy



CONSEQUENCES OF/FOR THE MBA

ADVANTAGES

- **Identify** and **localize** the different **choices needed in the modelling process** of the measurement.
- Make space for the notion of **purpose** of the measurement.
- Emphasize the role of **calibration**.



ISSUES

- The choice of model remains **difficult, in particular what to include in the measuring instrument**.
- System under measurement vs. object interacting with the measuring instrument.
- Measurement indications vs. measurement results

THE ROLE OF THE SPECIES CONCEPT

- **Taxa Delineation Issue:** Ambiguity of the concept used to delineate microorganismal taxa, in particular species of bacteria.
 - Difficult to answer diversity-related questions: **(A)** How many taxa compose this community? and **(B)** How different are these taxa?
- *Example:* OTUs/ASVs are **two different ways of delineating phylogenetic similarity**. It does not always correspond to a specific taxa, species or genus.
 - Add another layer at which communities can be compared (species or ASVs?), what is the meaning of this layer? Risk of doing “ecology of molecules”
 - Problem for comparing scientific publications among them

THE ROLE OF THE SPECIES CONCEPT

- **Unstable Microbial Classification:** Few valid taxon names are published and there are regular changes in the terminology within the taxonomy.
 - Without the possibility of attaching an individual to a category, **(B)** is difficult to answer.
 - It is difficult to compare scientific production over time.
- *Example:* Problem to link DNA sequences to a moving classification.
 - Few valid taxon names = few reliable link between DNA sequences and taxon name.
 - Need a lot of background knowledge (databases).

THE ROLE OF THE SPECIES CONCEPT

- **Microorganism Isolation Issue:** Difficult to isolate microbial individuals despite this being necessary for a stable taxonomy.
 - Additional challenge for **diversity measurement**.
- *Example:* Pooling of the microorganisms and their DNA
 - Difficulties in equating one DNA molecule to one microorganism

THE ROLE OF THE SPECIES CONCEPT

- The species concept enters measurements in the **modelling activities**, so it creates **relevant uncertainties** to the parameter diversity.
- The diversity value changes depending **on which concrete definition of species researchers choose**.



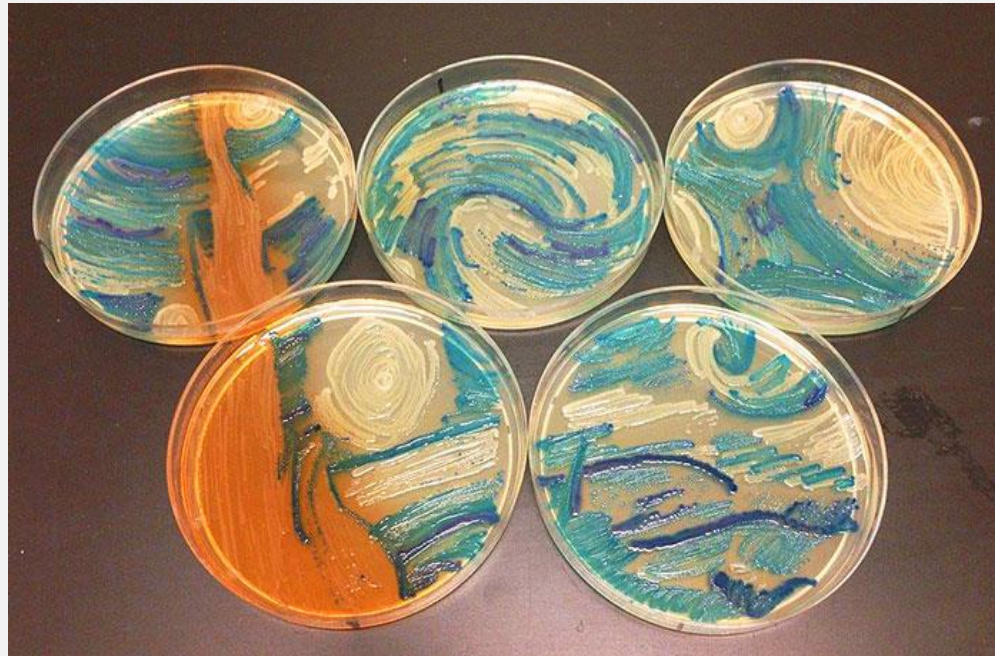
THE ROLE OF THE SPECIES CONCEPT

- There is a tension between the system under measurement that is supposed to exacerbate **different species** (diversity) and its idealization in the model of the measurement.
 - Is there a “real” value for such parameters?
 - **Solution:** choose a species definition that fits **the measurement’s purpose or is stable enough for calibration.**
- ⇒ **Concrete** and potentially local definitions are needed.
- In conservation biology, a **general** definition of species is used to measure diversity across environments.

CONCLUSION

- Diversity needs to be measured in microbial ecology.
- The model-based account gives insight into **the upstream decisions needed before the measurement process.**
- Decisions concern **the model itself**, the MBA does not help to solve these tension (e.g., what is part of the measuring instrument), and they concern the **idealization made in the model** (e.g., which concrete species concept to use)
- The species concept is **a source of uncertainty in the diversity measurement.**

THANK YOU!



Vincent van Gogh's "The Starry Night" by Melanie Sullivan of Missouri.
American Society of Microbiology.