

# Key & frequently overlooked points in amplicon sequencing analysis of environmental microbiomes

by

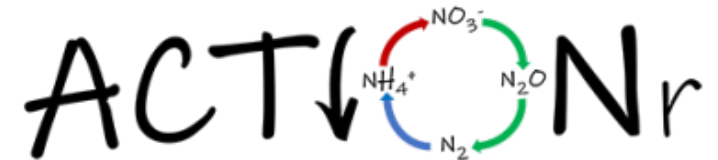
**Sotirios Vasileiadis**

Assistant Professor at Grp of Plant and Environmental Biotechnology  
Department of Biochemistry and Biotechnology

From: Bioinformatics for Microbiomes  
*Microbial diversity analysis with amplicon sequencing*

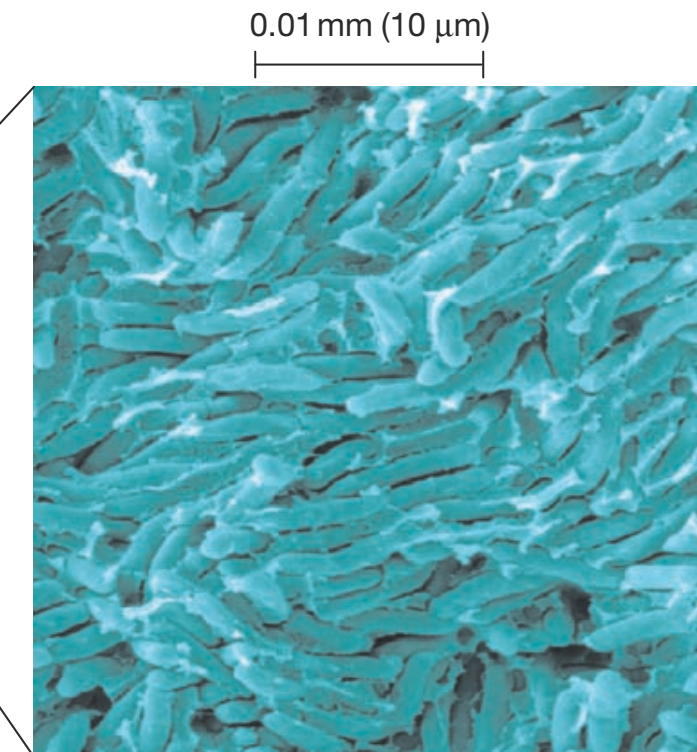
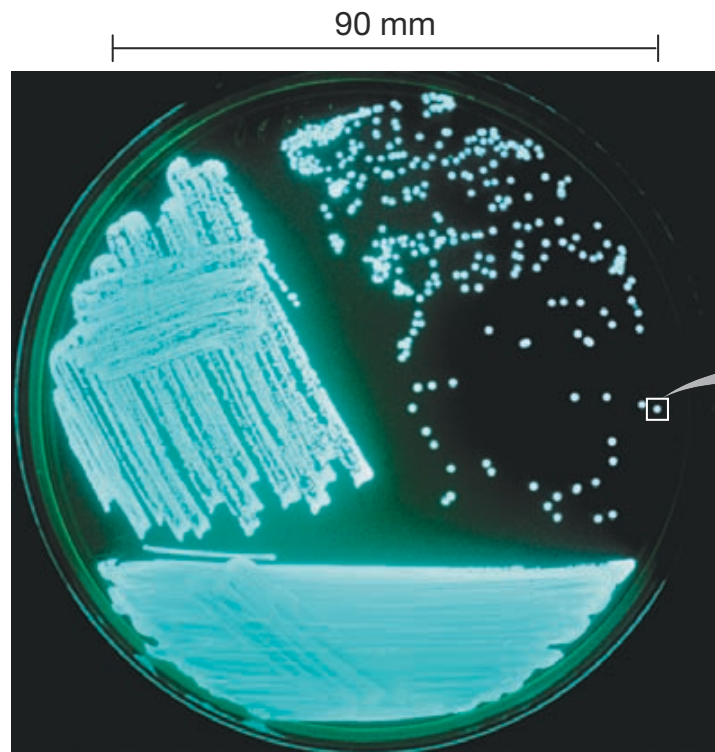
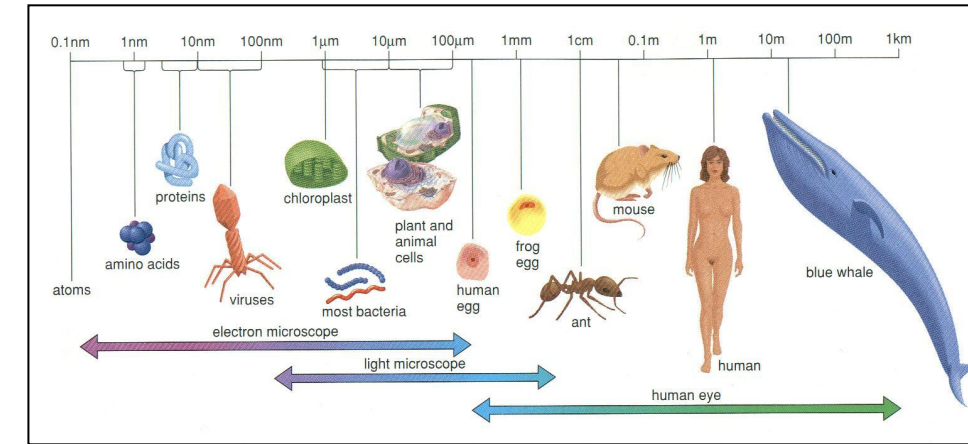
[HosMic MSc programme!!!](#)

MSc contact person: Prof K. Kormas (kkormas@uth.gr)





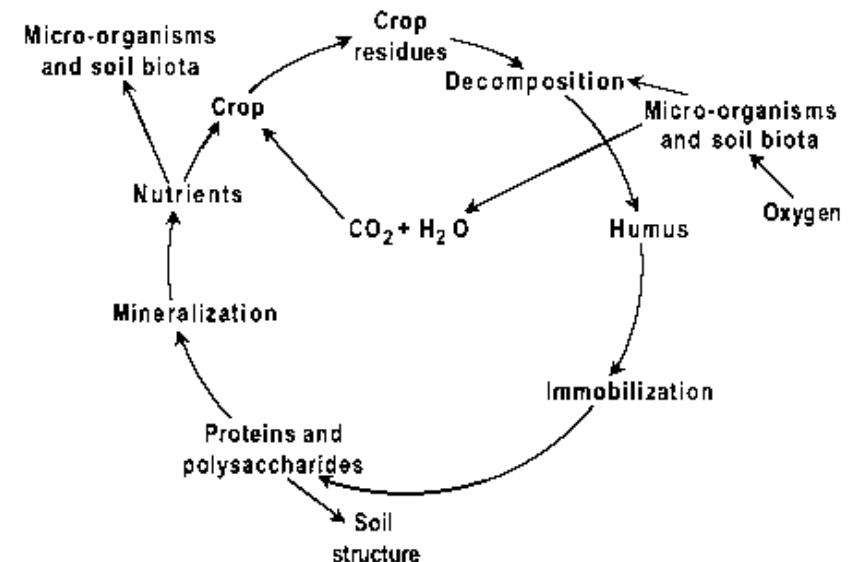
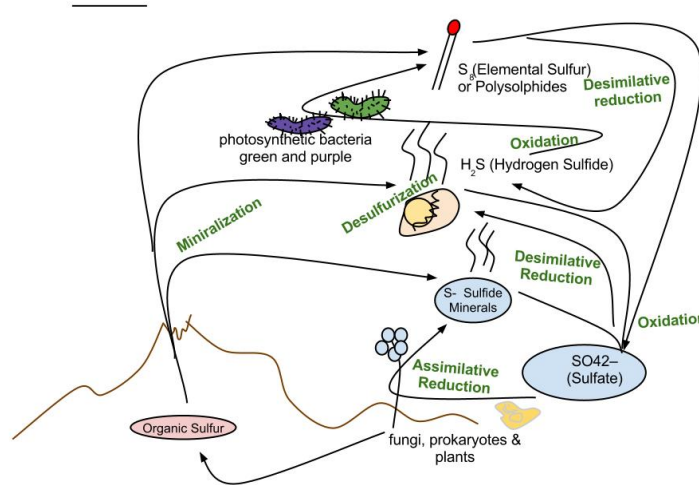
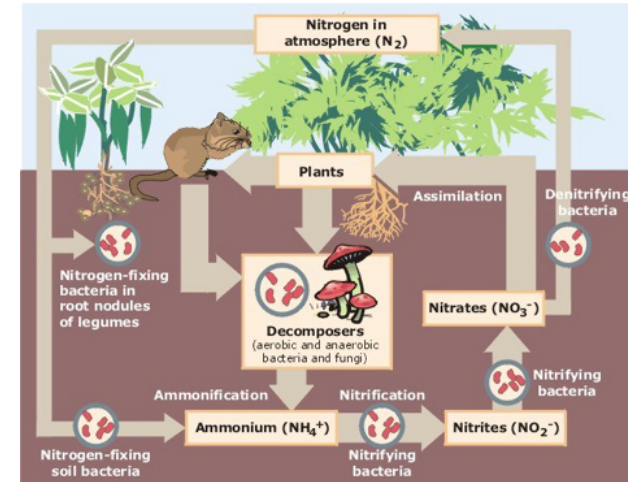
# Microorganisms





# Microbial ecosystem services

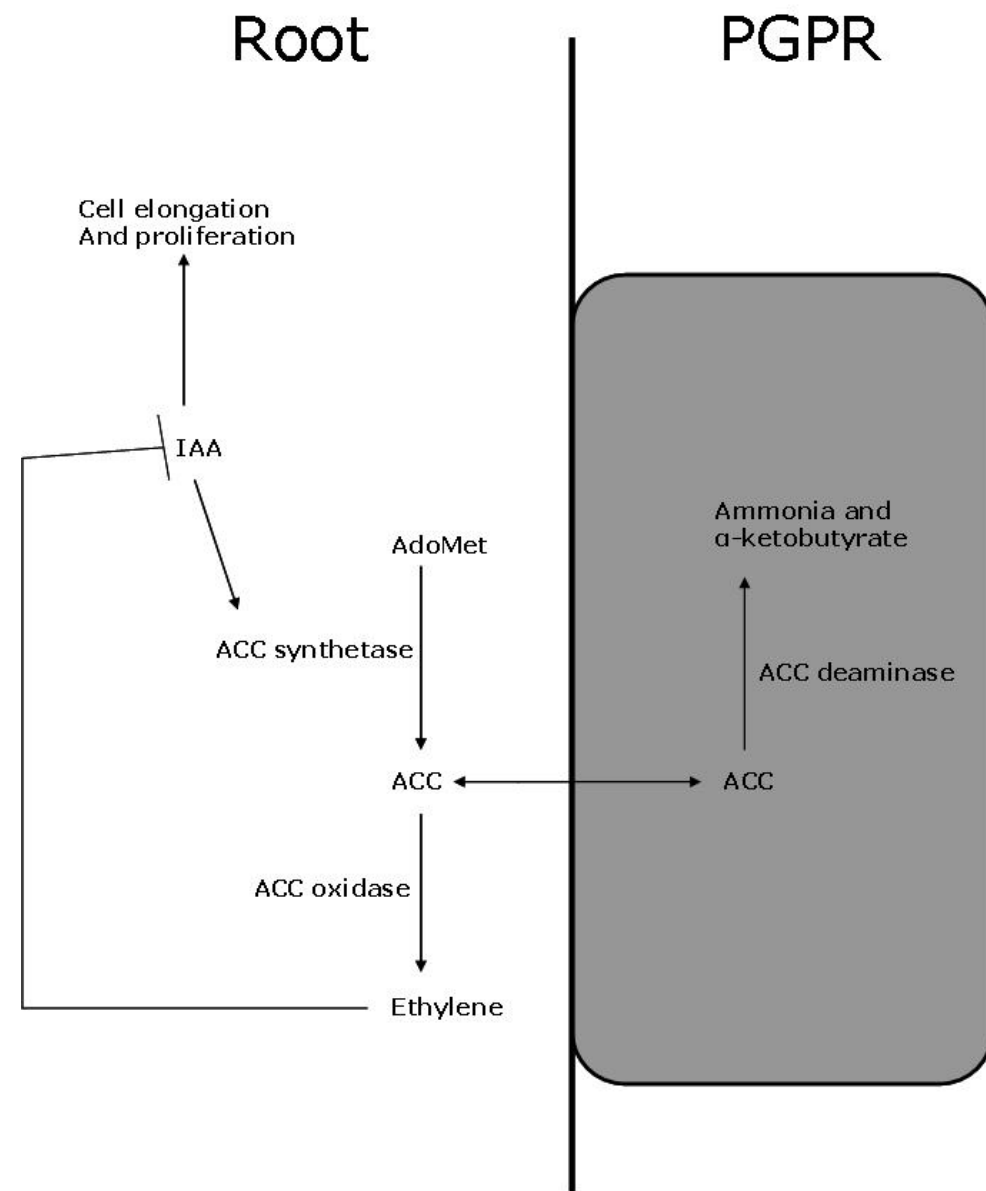
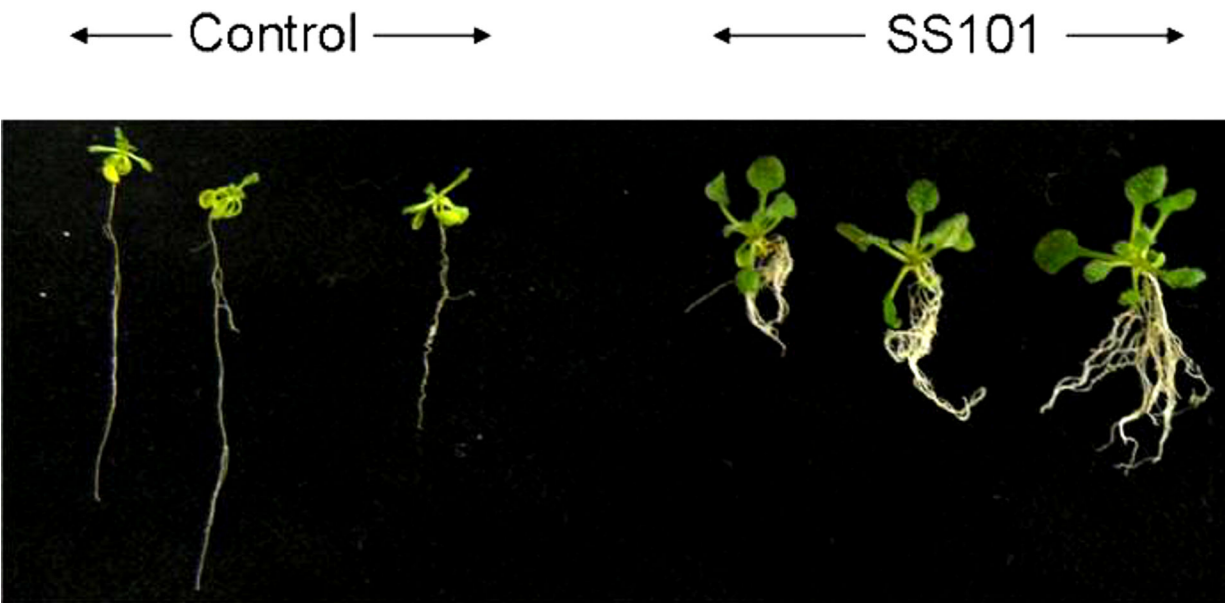
- Nutrient cycles
- Effect on animal/plant growth
- Bioremediation
- Pathogen biocontrol, source of novel antibiotics & enzymes
- ...





# Microbial ecosystem services

- Nutrient cycles
- Effect on animal/plant growth
- Bioremediation
- Pathogen biocontrol, source of novel antibiotics & enzymes
- ...

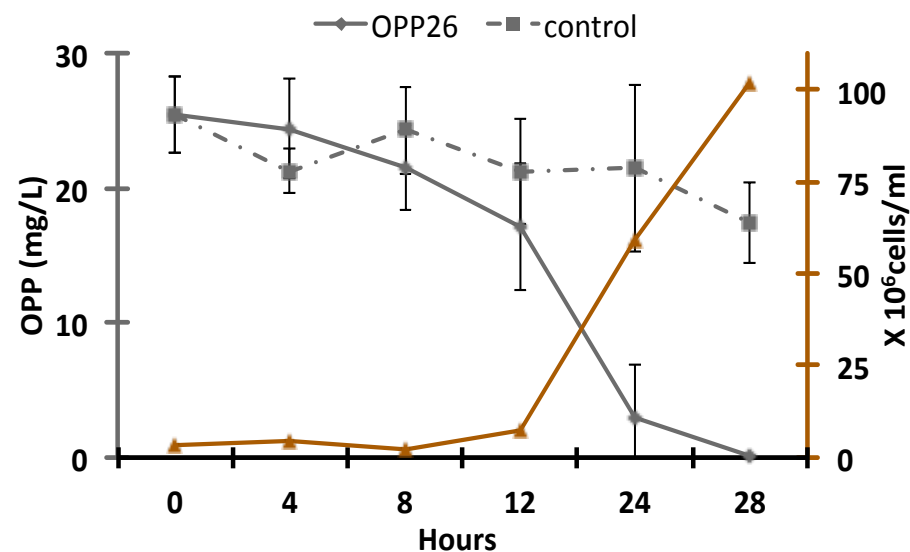
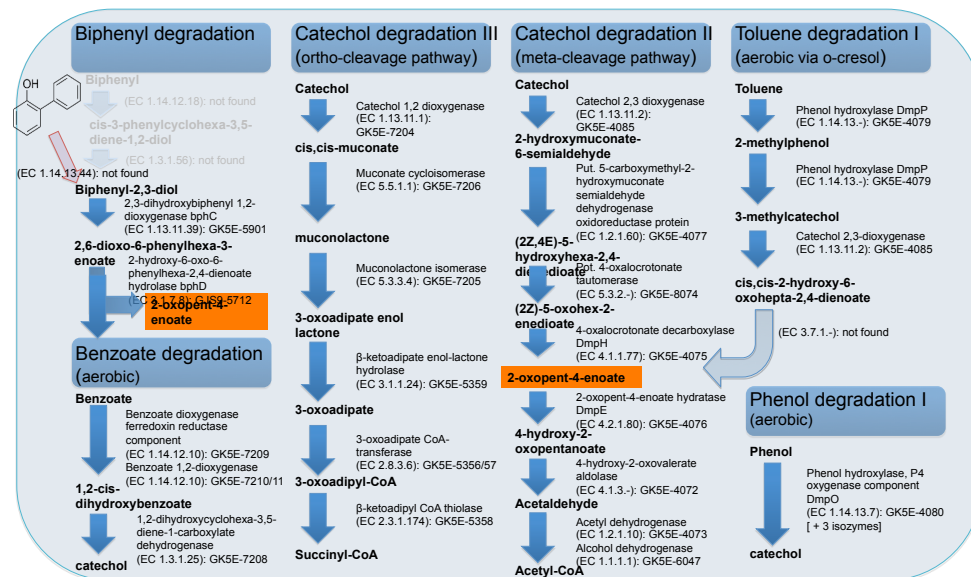






# Microbial ecosystem services

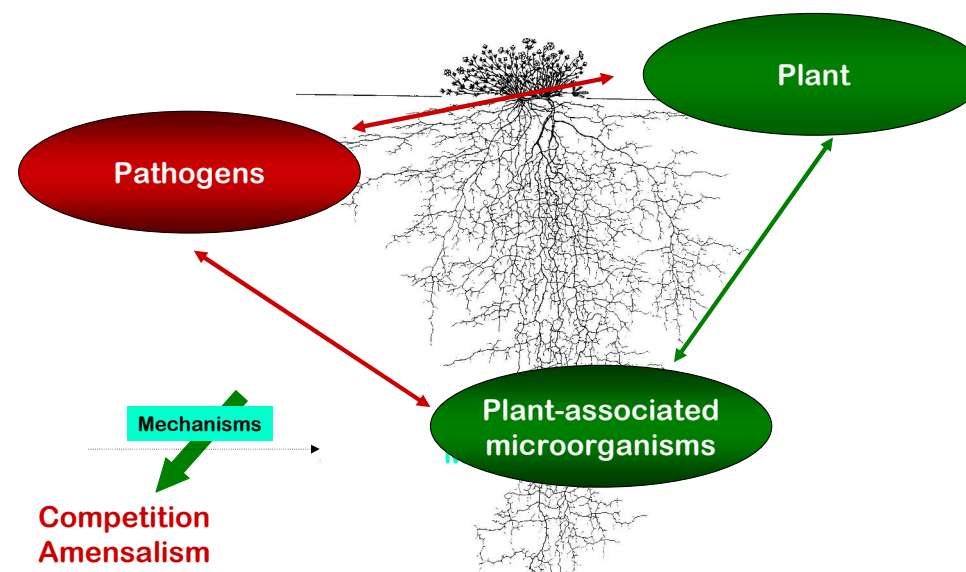
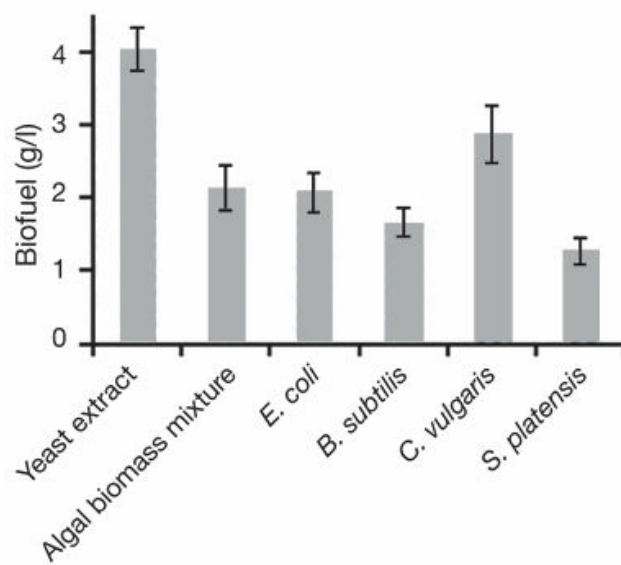
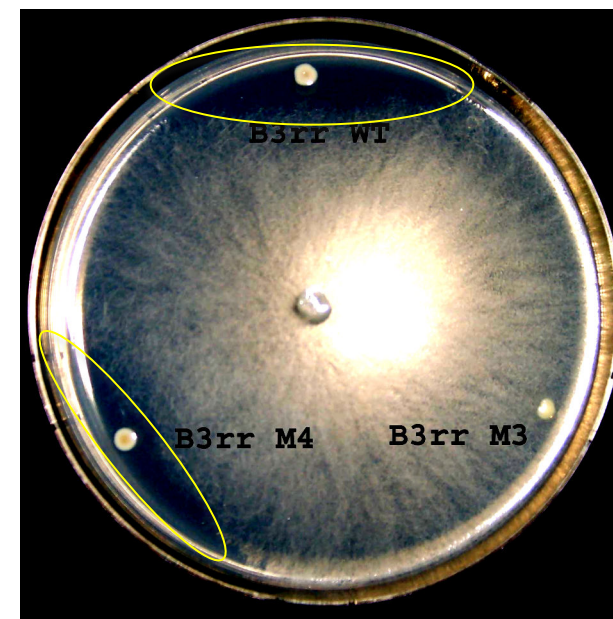
- Nutrient cycles
- Effect on animal/plant growth
- Bioremediation
- Pathogen biocontrol, source of novel antibiotics & enzymes
- ...





# Microbial ecosystem services

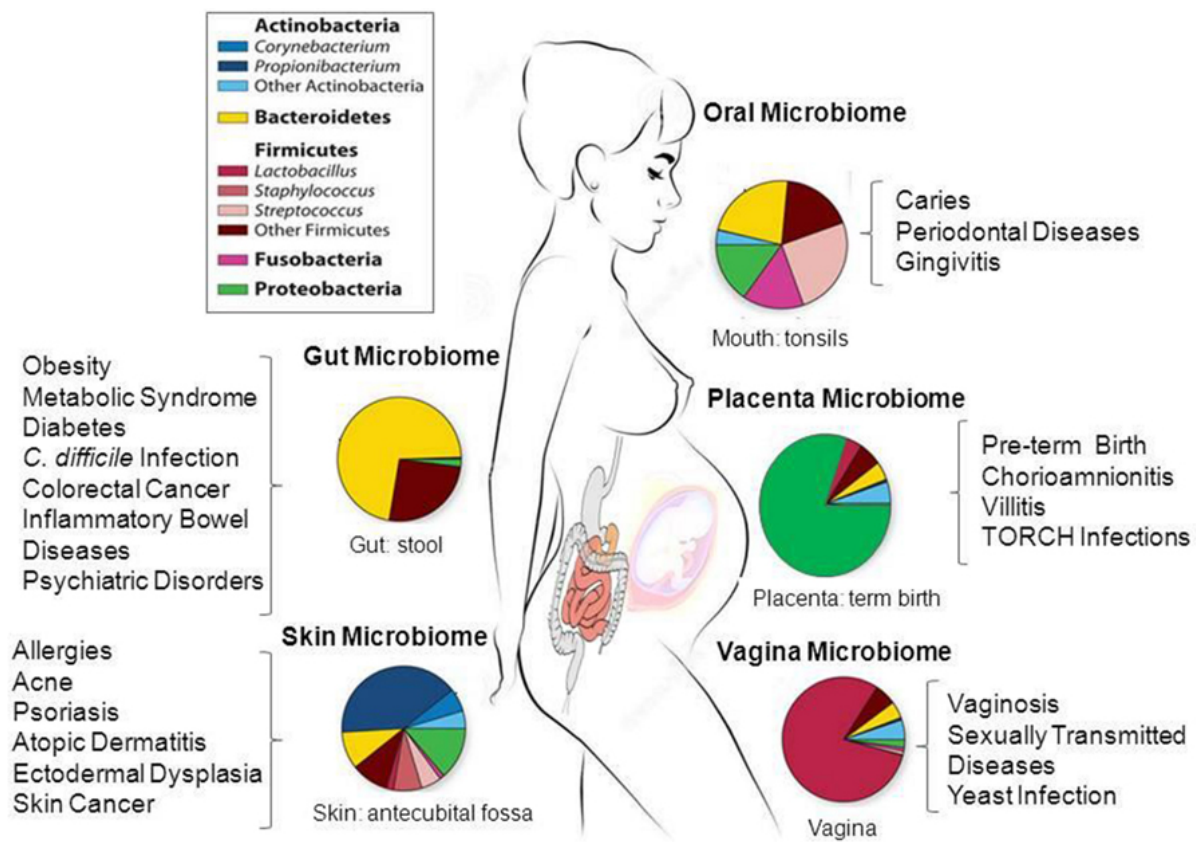
- Nutrient cycles
- Effect on animal/plant growth
- Bioremediation
- Pathogen biocontrol, source of novel antibiotics & enzymes
- ...



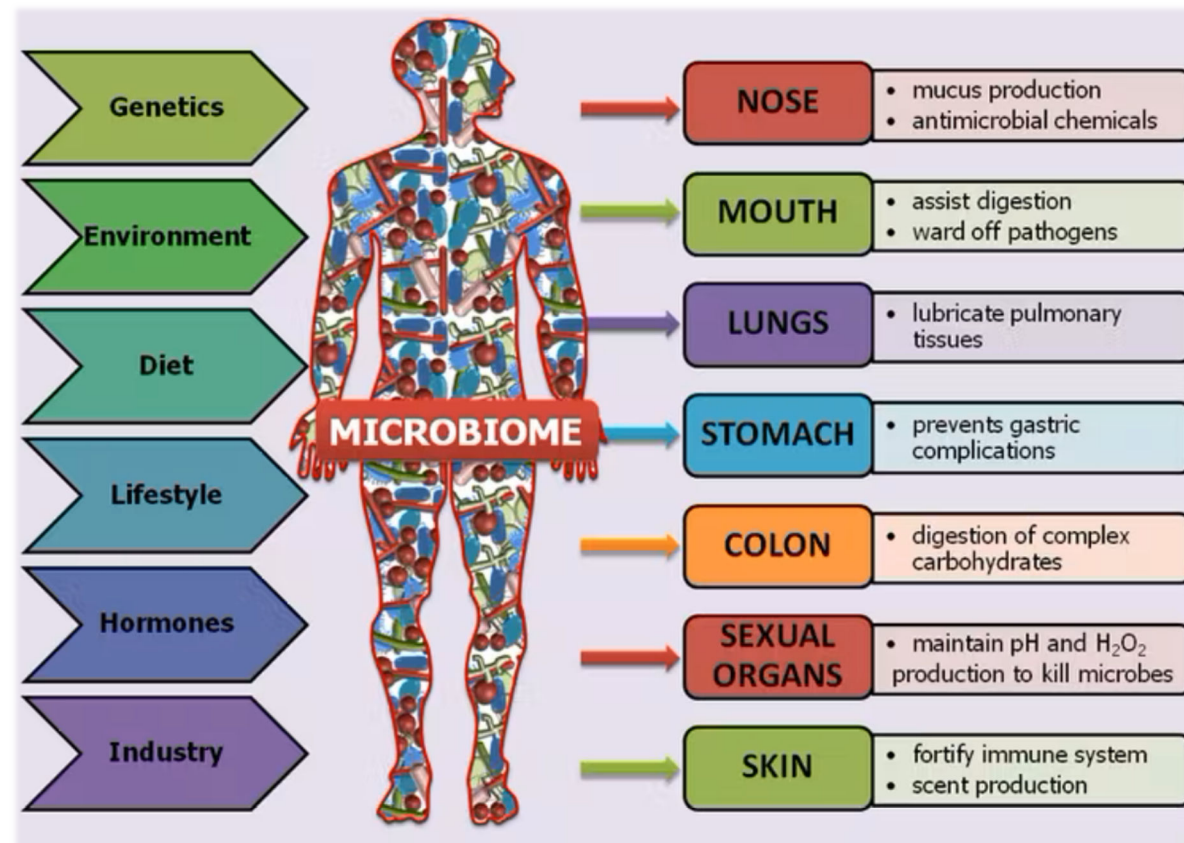


# The human microbiome: pathogenesis and control

## Pathogenicity



## Beneficial microbes

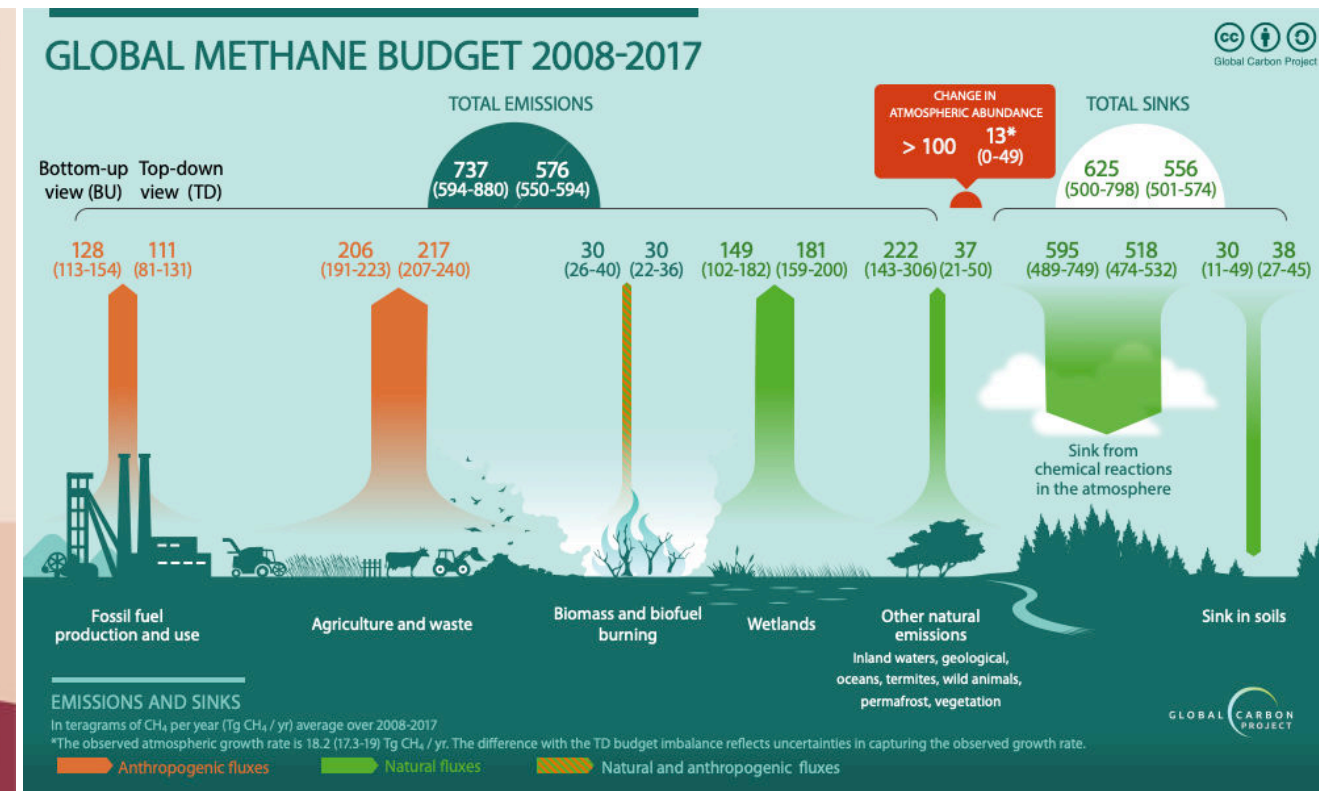
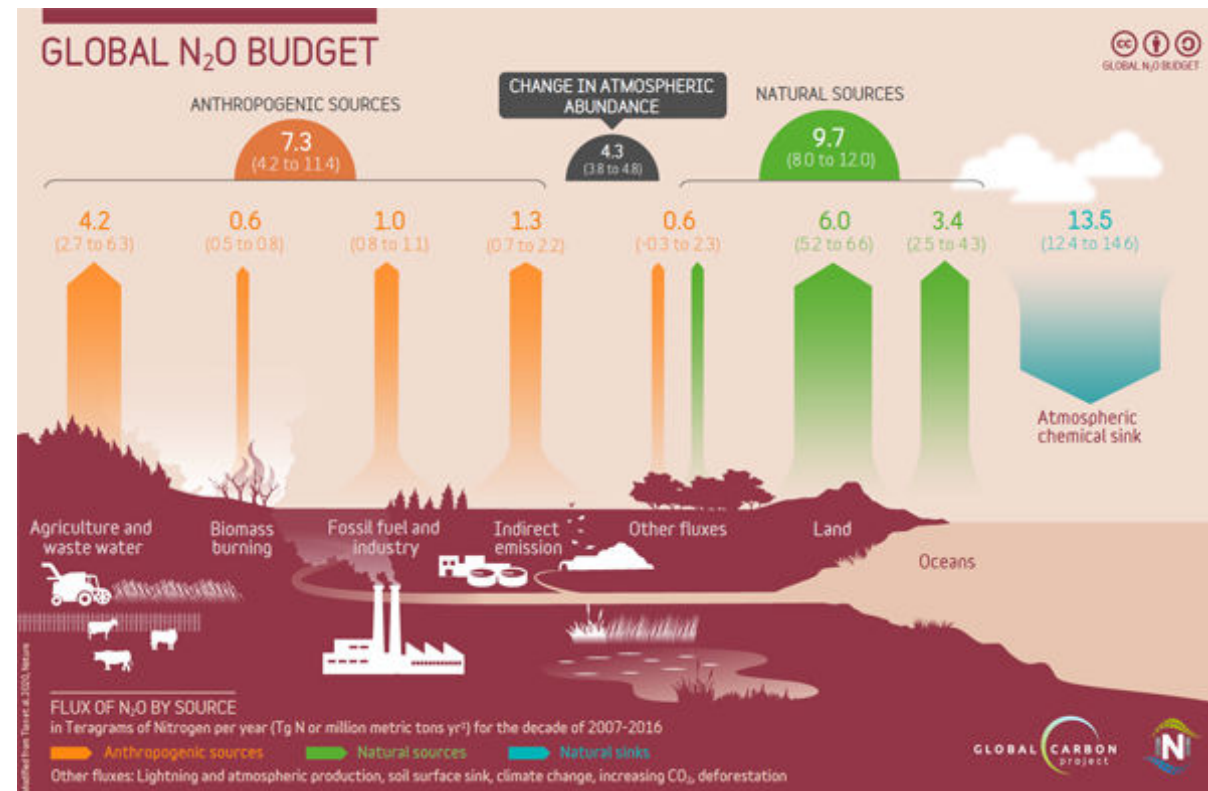






# Microbes shape conditions for life...

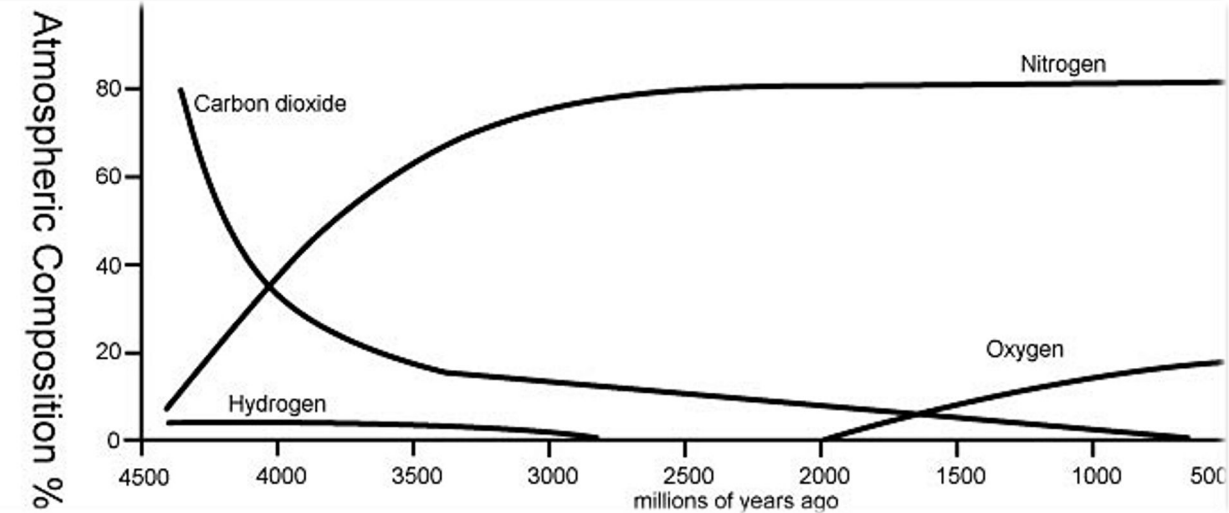
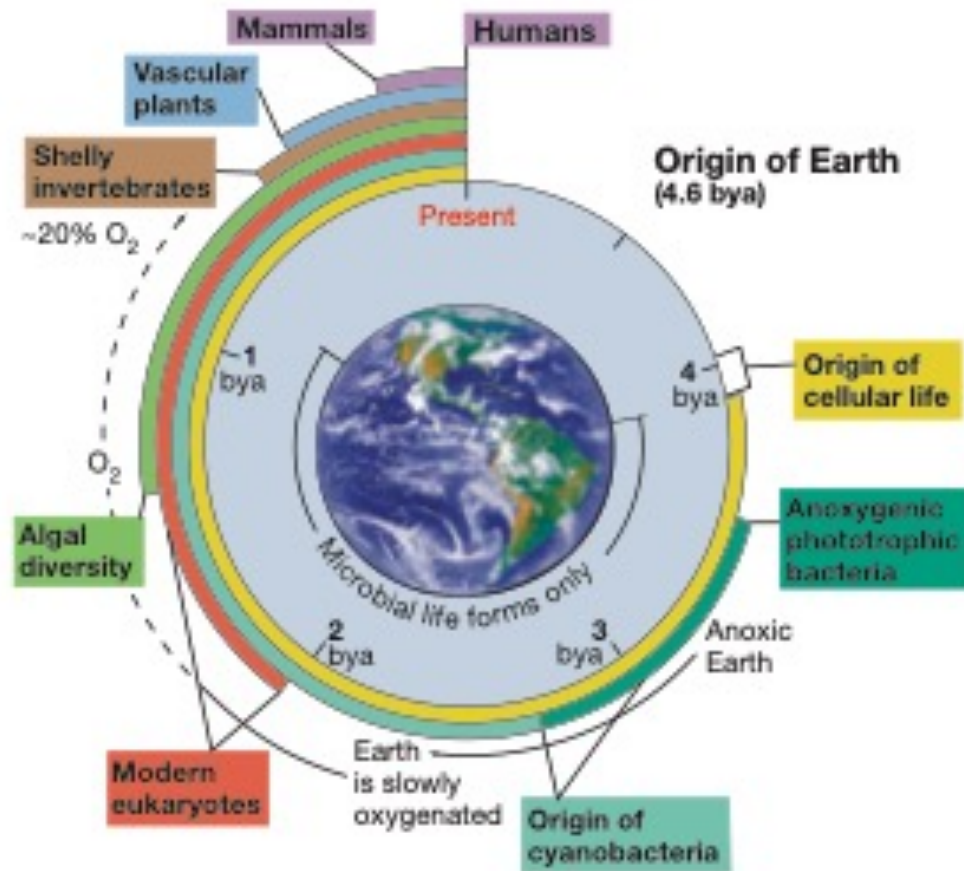
- Greenhouse gasses of great impact...





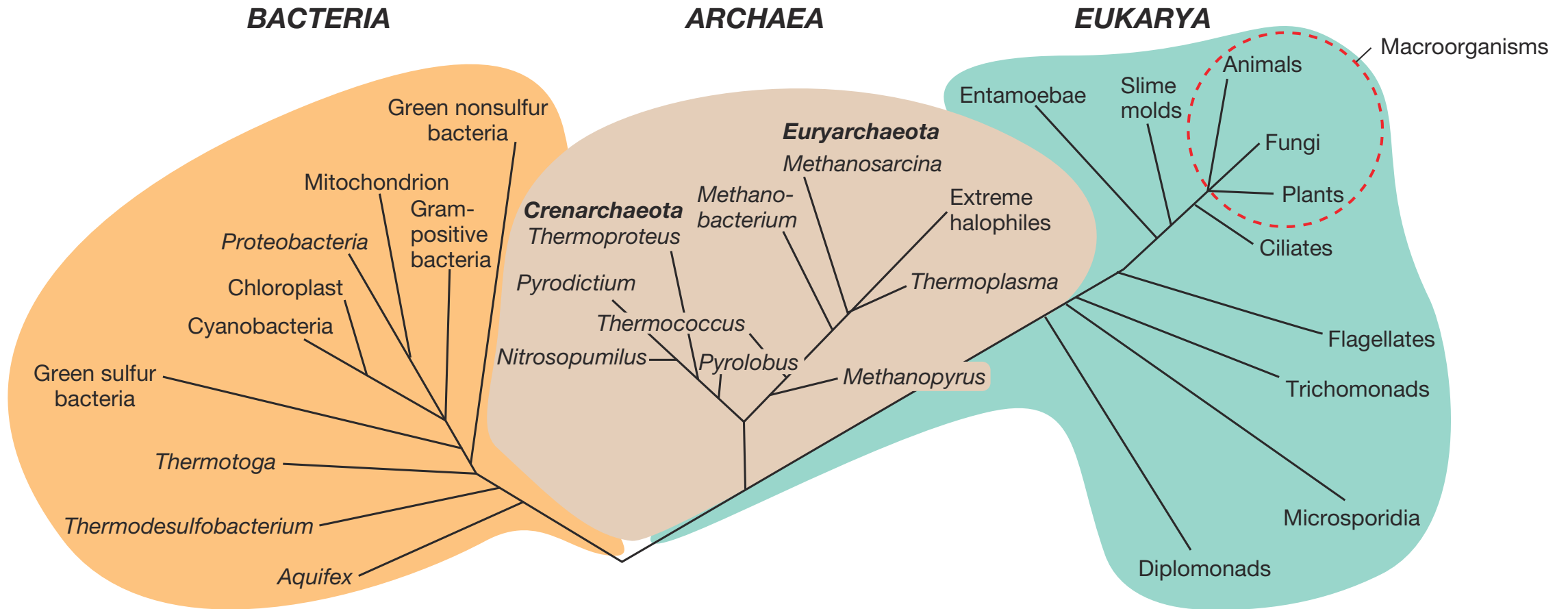


# Microbes shape conditions for life...



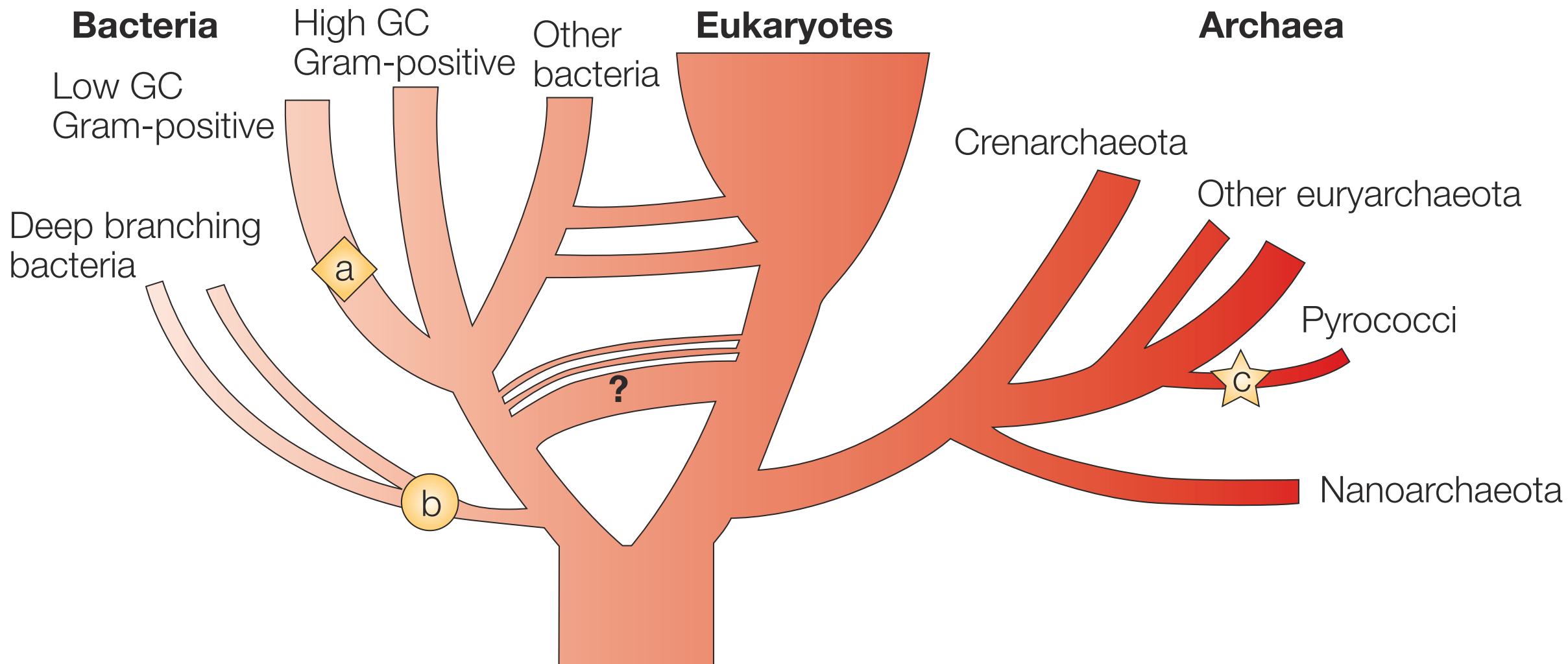


# they dominate life...





# ... & they mess with it!!!







# They outnumber all of us...

## Microbiome

IN NUMBERS

**100 Trillion**

sybiotic microbes live in and on every person and make up the human microbiota

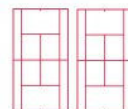
The human body has more microbes than there are stars in the milky way

**95%**

of our microbiota is located in the GI tract

**150:1**

The genes in your microbiome outnumber the genes in our genome by about 150 to one



The surface area of the GI tract is the same size as 2 tennis courts

You have

**1.3X**

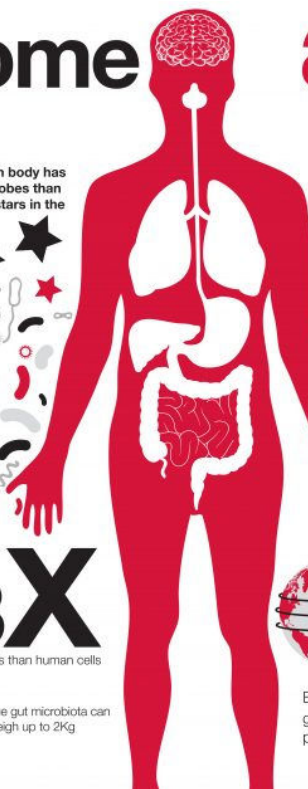
more microbes than human cells

**>10,000**

Number of different microbial species that researchers have identified living in and on the human body

**2kg**

The gut microbiota can weigh up to 2Kg



**aps**  
Microbiome  
Ireland

Intertacing Food & Medicine

The microbiome is more medically accessible and manipulable than the human genome

It is thought that

**90%**

of disease can be linked in some way back to the gut and health of the microbiome

**5:1**

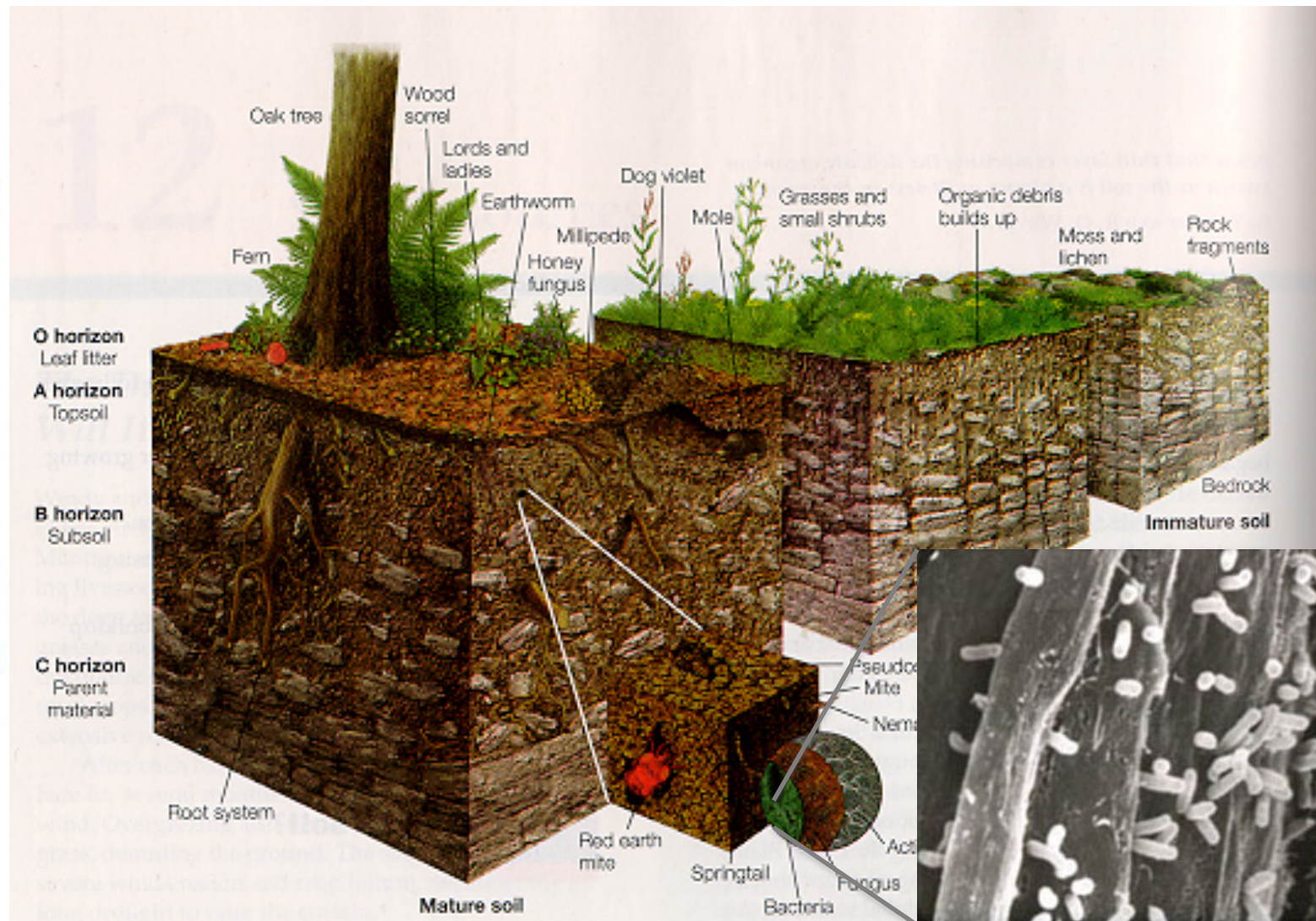
Viruses:Bacteria in the gut microbiota

**2.5**

The number of times your body's microbes would circle the earth if positioned end to end



Each individual has a unique gut microbiota, as personal as a fingerprint







# So... who is the “puppet-master”?

- $10^9$  cells &  $1-5 \times 10^4$  strains per temperate soil gram
- $10^4$  strains in the human body (prokaryotic cells comprise 70% of the total cells of the human holobiome)
- 100 M more microorganisms in the oceans than the stars of the known universe
- 400 g of *Clostridium botulinum* neurotoxin can eliminate mankind

## EDITORIAL

### Microbiology by numbers

The scale of life in the microbial world is such that amazing numbers become commonplace. These numbers can be sources of inspiration for those in the field and used to inspire awe in the next generation of microbiologists.

**Nature E.** 2011. Microbiology by numbers. *Nature* **9**:628-628.

**Dance A.** 2008. Soil ecology: What lies beneath. *Nature* **455**:724-725.  
And citations therein...

### What lies beneath

More creatures live in soil than any other environment on Earth. But what are they all doing there? **Amber Dance** reports on the world's widest biodiversity.

terial cells each day. Moving onto dry land, the number of microorganisms in a teaspoon of soil ( $1 \times 10^9$ ) is the same as the number of humans currently living in Africa. Even more amazingly, dental plaque is so densely

...The first DNA-based estimate of soil microbial biodiversity, published in 1990, counted about 4,000 different bacterial genomes per gram of soil. Since then, various studies and models have pushed the number up as high as 830,000 species per gram, down to 2,000, and back up again. Most recently, Triplett and his colleagues ran 139,000 individual sequences — more than other studies have used — and came up with an estimate of 10,000 to 50,000 species per gram of soil...



# So... to summarize

## Microorganisms

- Shape our home, Earth
- Important for our progress/health

## Important traits

- Highly diverse
- High numbers

-----

How do we learn more about them?



# Microbiome analysis methods

- Culture-based
- Culture independent



# Microbiome analysis methods

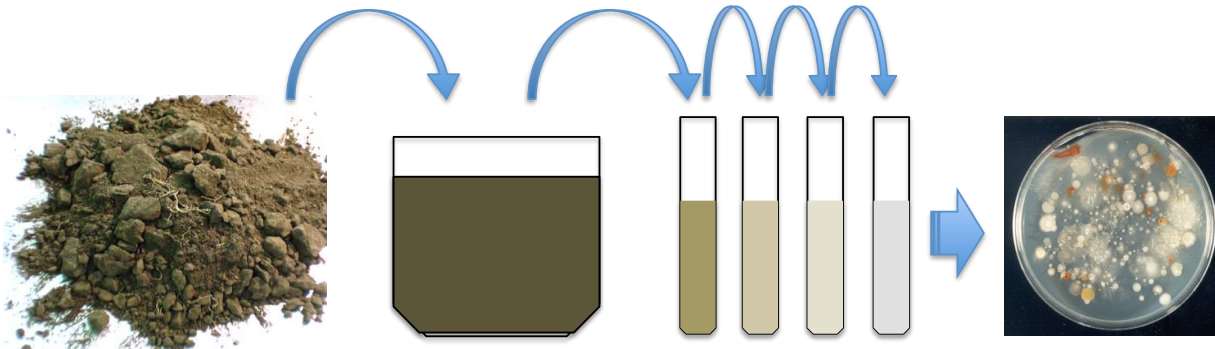
- Culture-based
- Culture independent





# Culture-based

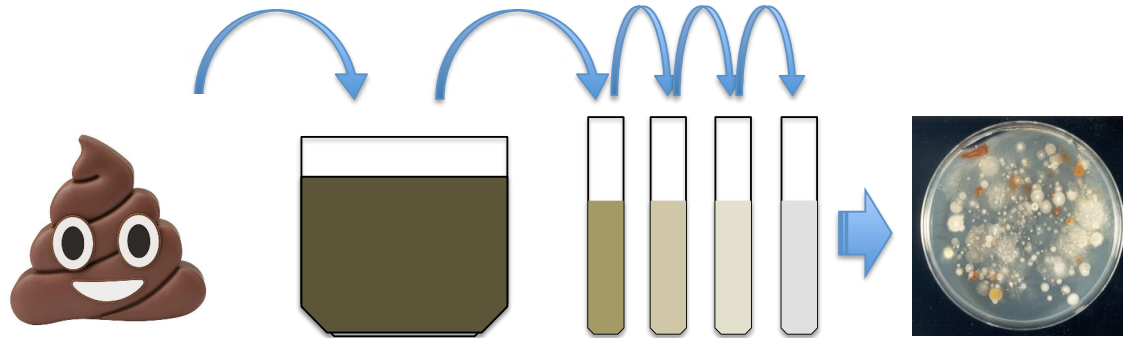
- Traditional





# Culture-based

- Traditional



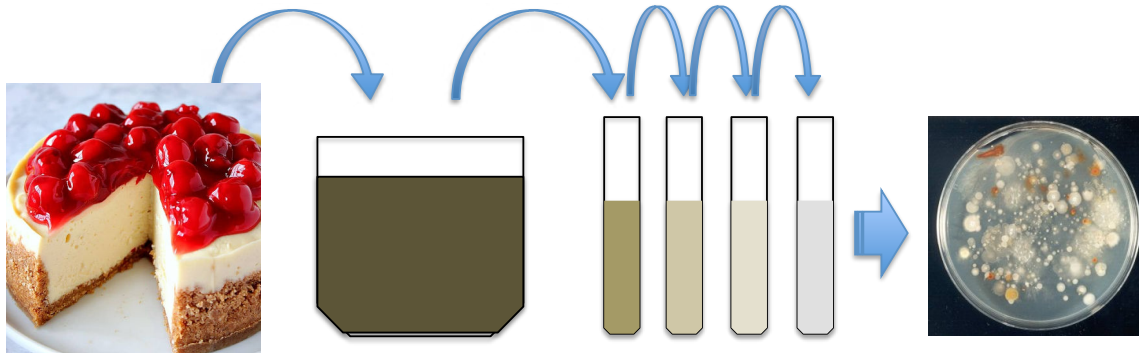


ΤΜΗΜΑ

Βιοχημείας &  
Βιοτεχνολογίας  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

# Culture-based

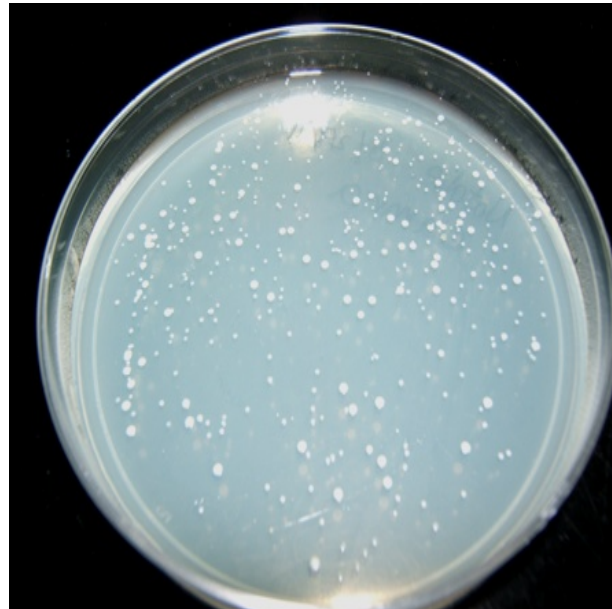
- Traditional





# Culture-based

- Traditional



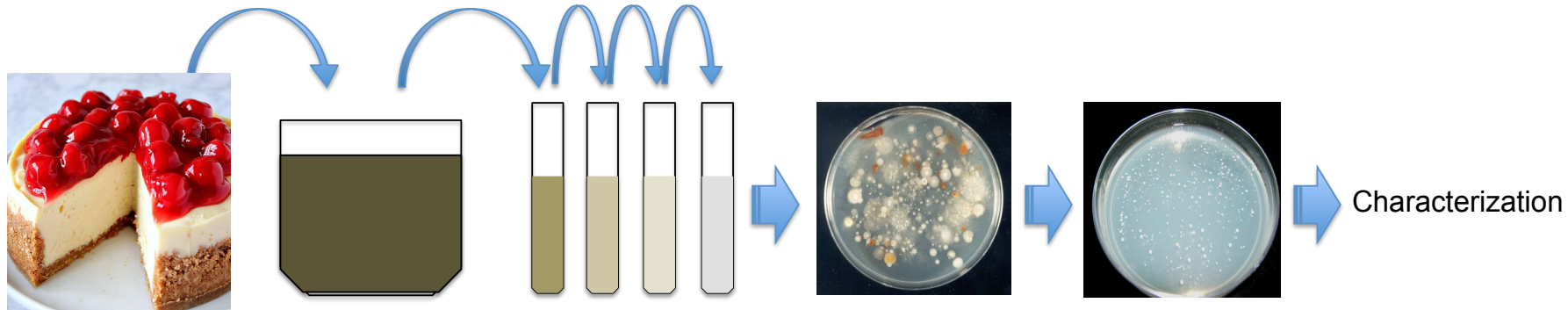
Characterization





# Culture-based

- Traditional
- High throughput (robotic colony pickers)





# Culture-based

- Traditional
- High throughput (robotic colony pickers)

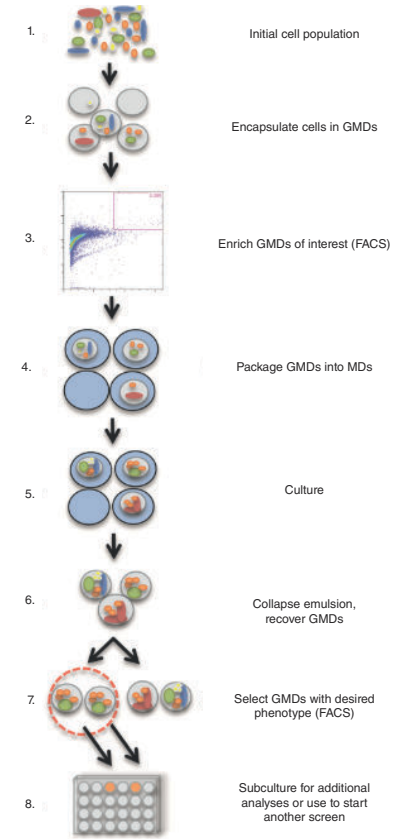
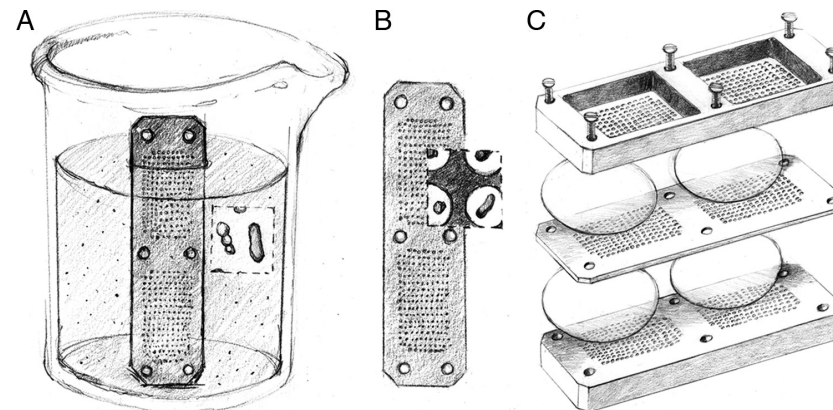
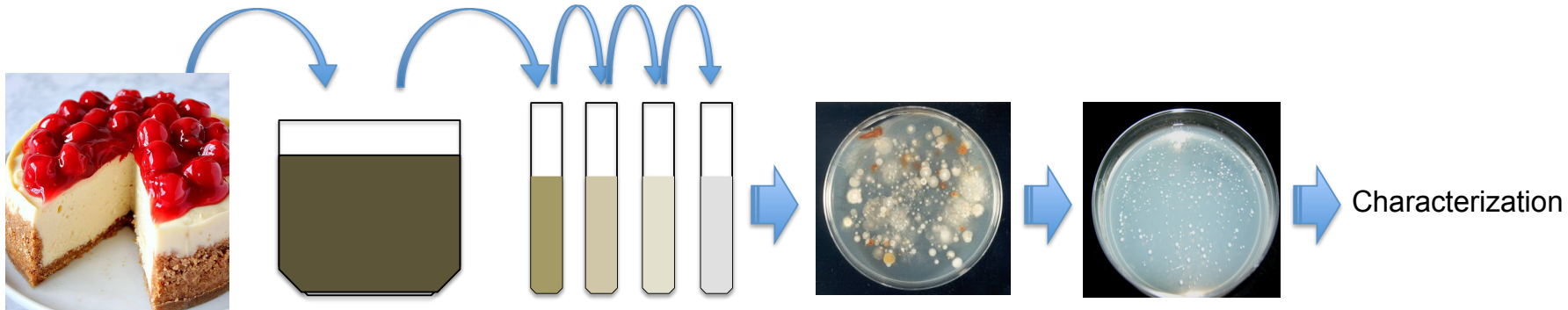
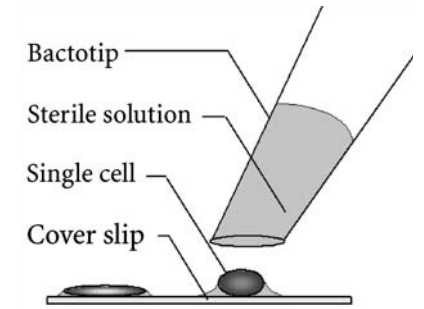
Isolation is hampered by:

- *Common nutritional needs*
- *Common antibiotic resistances*
- *Auxotrophies (e.g. B12)*
- *Parasitism/predation phenomena (e.g. Saccharibacteria, BALOs)*
- *Competition-free single cell methods are usually laborious/expensive*
- *Yet undiscovered nutrient combinations (culturomics)*



# Culture-based

- Traditional
- High throughput (robotic colony pickers)
- Competition removal (i-Chip, micromanipulation, microbial gel droplets)





# Culture-based

- Traditional
- High throughput (robotic colony pickers)
- Competition removal (i-Chip, micromanipulation, microbial gel droplets)

***Despite the progress, several obstacles still exist***



# Microbiome analysis methods

- Culture-based
- Culture independent



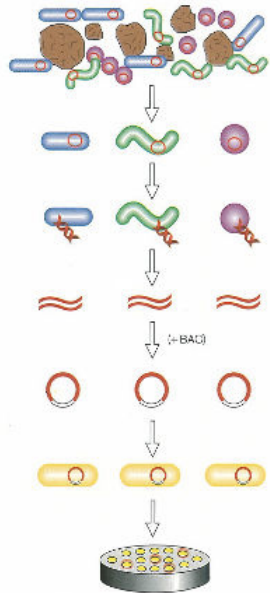
DEPARTMENT OF

**Biochemistry &  
Biotechnology**

UNIVERSITY OF THESSALY

# Culture-independent

1998



## Metagenomics

“The functional and sequence-based analysis of the collective environmental genetic content”

## Aims

- 1) Diversity analysis
- 2) Functional characterization

Handelsman J *et al.* 1998 *Chem. & Biolog.* R245-R249





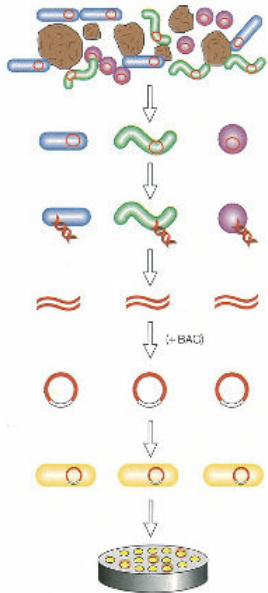
DEPARTMENT OF

**Biochemistry &  
Biotechnology**

UNIVERSITY OF THESSALY

# Culture-independent

1998



## Metagenomics

“The functional and sequence-based analysis of the collective environmental genetic content”

## Aims

- 1) Diversity analysis
- 2) Functional characterization

Handelsman J *et al.* 1998 *Chem. & Biolog.* R245-R249



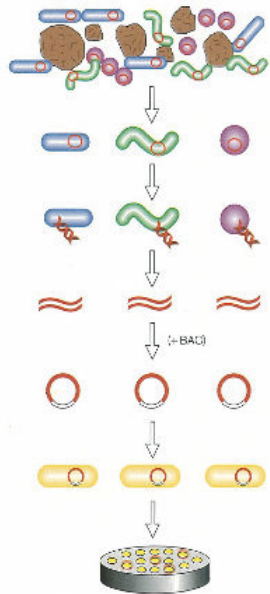
DEPARTMENT OF

**Biochemistry &  
Biotechnology**

UNIVERSITY OF THESSALY

# Culture-independent

1998



x 32 500 freezers for storing the DNA  
derived from microorganisms of 1 g of soil

## Metagenomics

“The functional and sequence-based  
analysis of the collective environmental  
genetic content”

## Aims

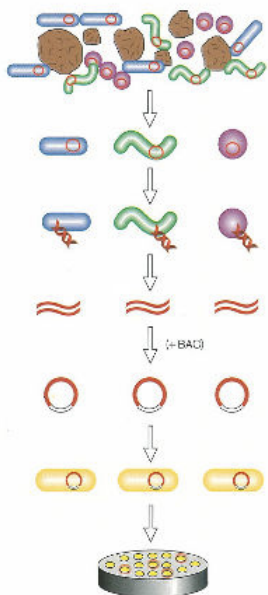
- 1) Diversity analysis
- 2) Functional characterization

Handelsman J *et al.* 1998 *Chem. & Biolog.* R245-R249

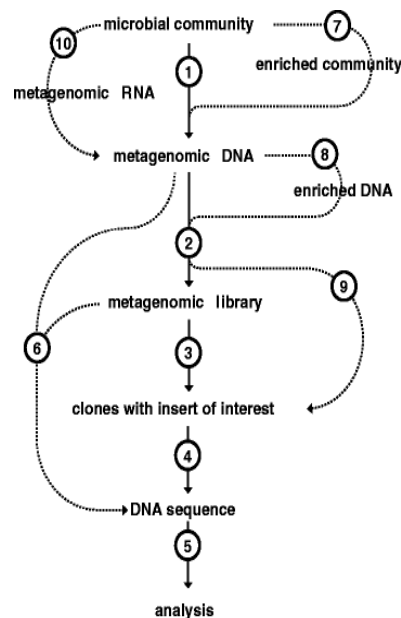


# Culture-independent

1998

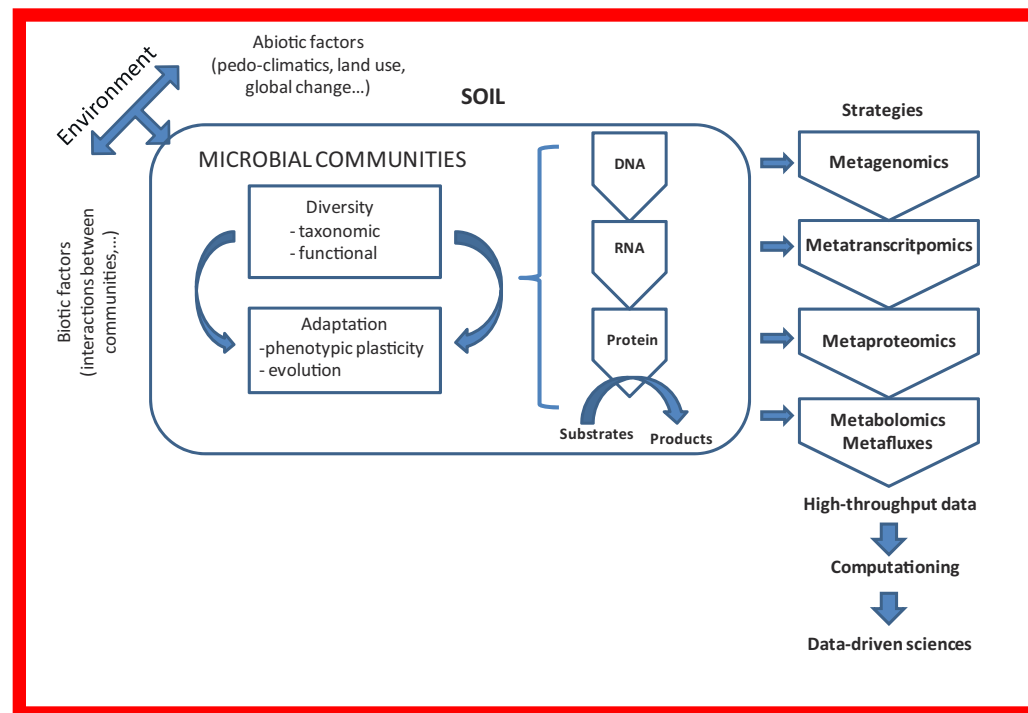


2007



2011

Omics



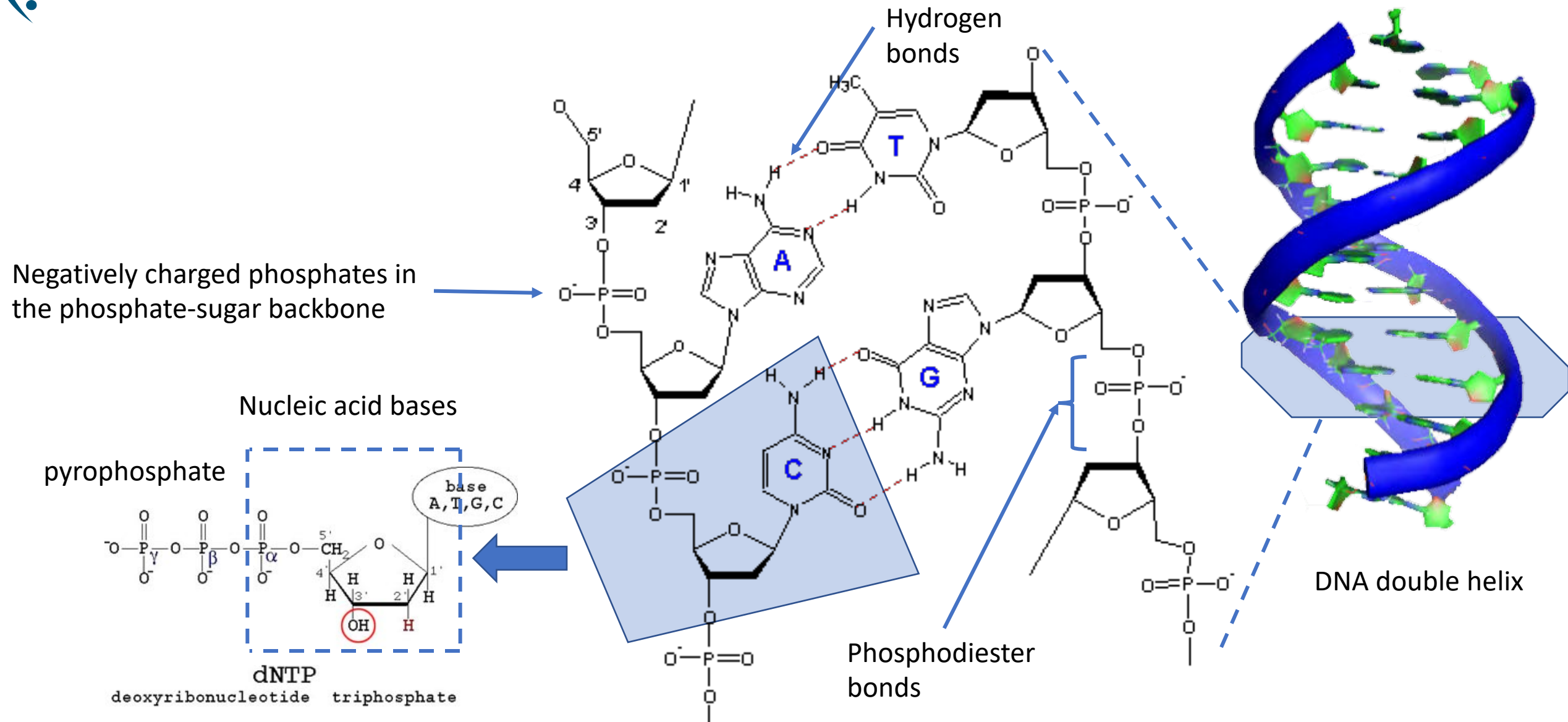
## Metagenomics

“The functional and sequence-based analysis of the collective environmental genetic content”

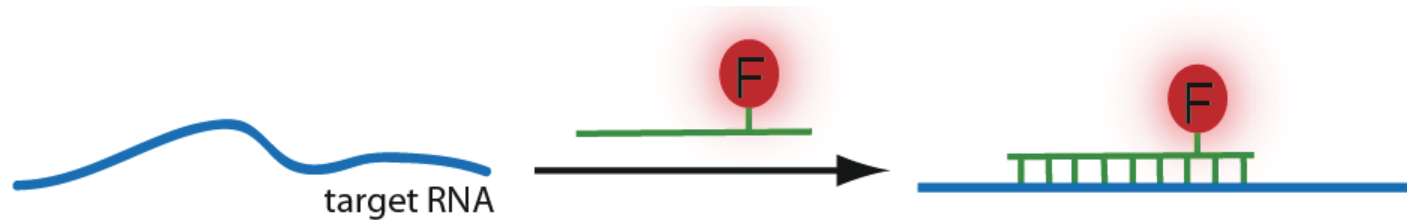
## Aims

- 1) Diversity analysis
- 2) Functional characterization

# The very basics: DNA (summary)



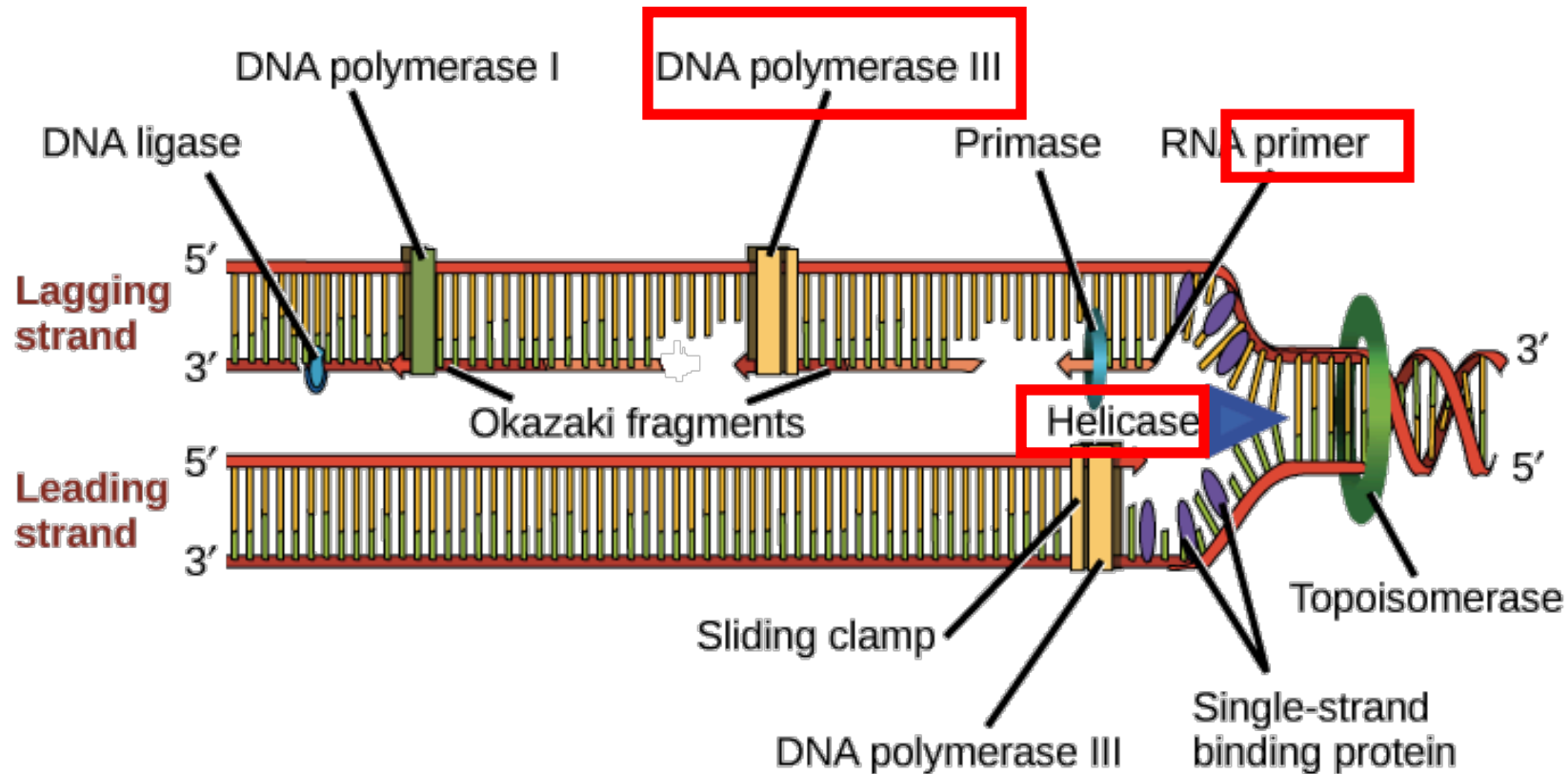
# The basics: oligo probe hybridization







# The very basics: DNA replication *in vivo*



# The basics: DNA amplification *in vitro* (PCR)

**PCR Demo**

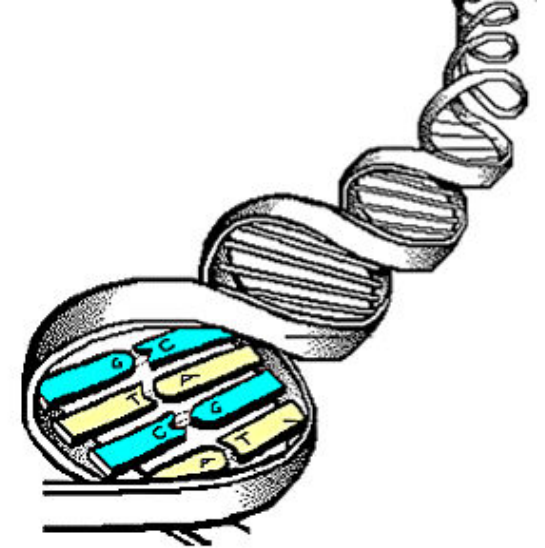
# The basics: DNA amplification *in vitro* (PCR)

## PCR Reaction Components

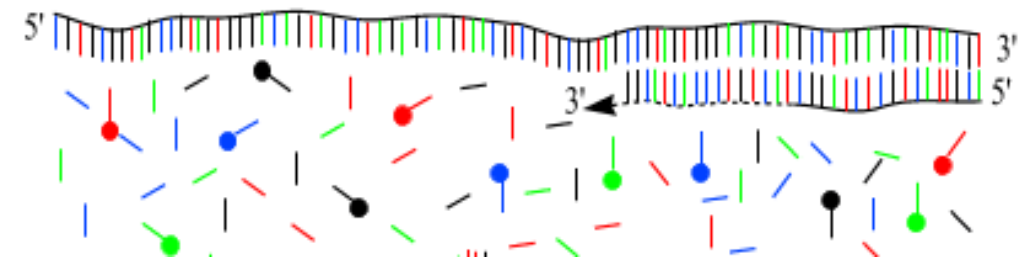
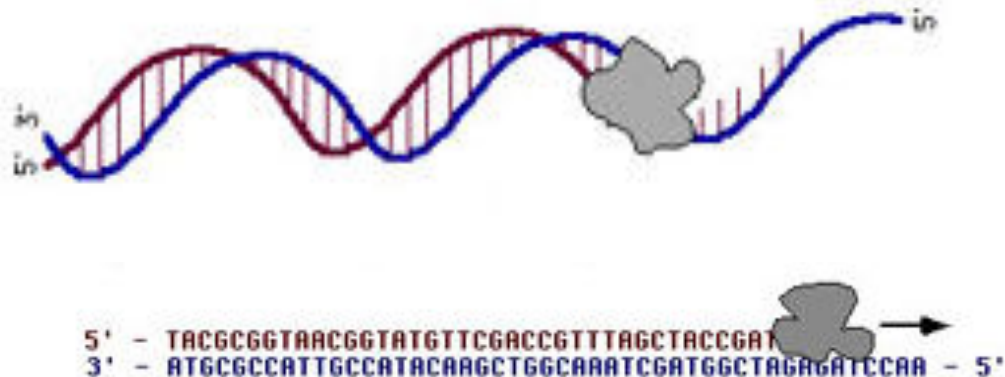
## Review

### Summary:

- Water
- Buffer
- DNA template
- Primers
- Nucleotides
- Mg<sup>++</sup> ions
- DNA Polymerase

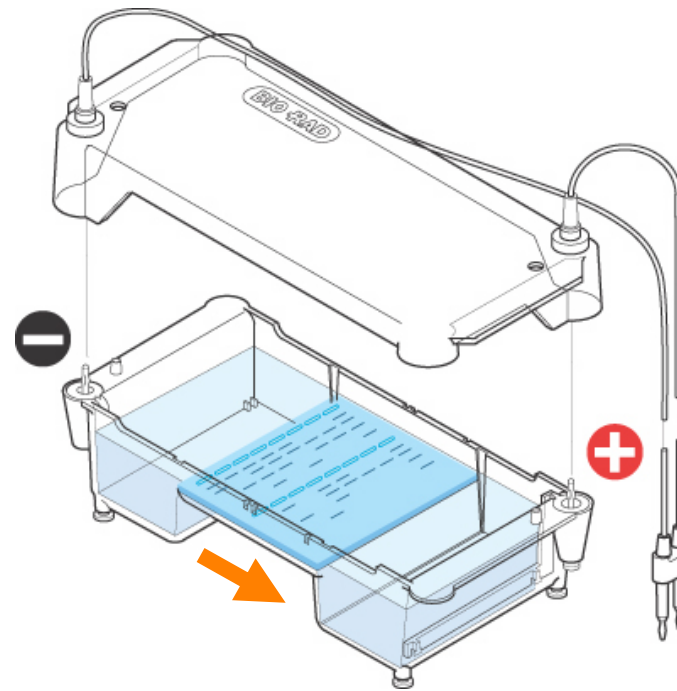
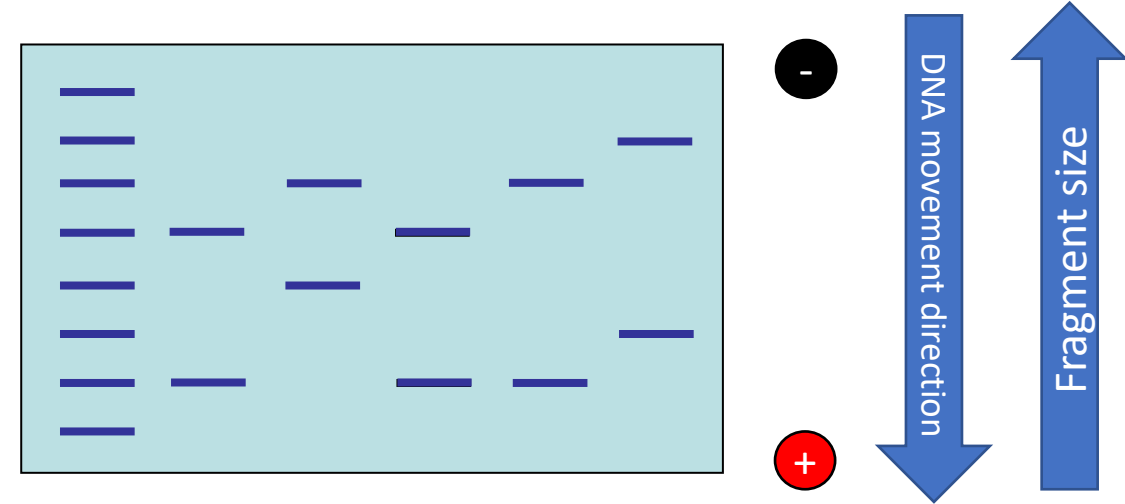


TACGCGGTACGGTATGTTTCGACCGTTTAGCTACCGAT•  
 TACGCGGTACGGTATGTTTCGACCGTTTAGCT•  
 TACGCGGTACGGTATGTTTCGACCGTTT•  
 TACGCGGTACGGTATGTTTCGACCGTT•  
 TACGCGGTACGGTATGTTTCGACCGT•  
 TACGCGGTACGGTATGTT•  
 TACGCGGTACGGTATGT•  
 TACGCGGTACGGTAT•  
 TACGCGGTACGGT•  
 TACGCGGT•

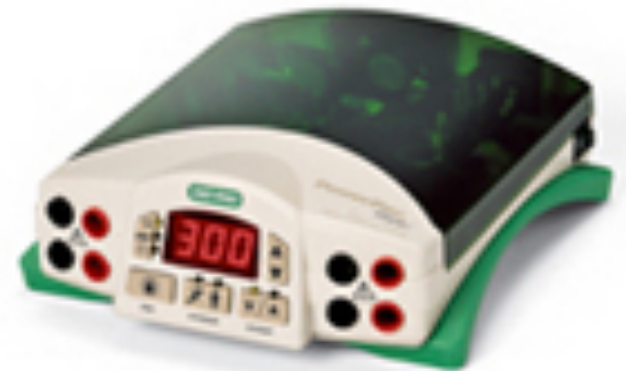


# The basics: DNA amplification *in vitro* (PCR)

## PCR product analysis in agarose gels

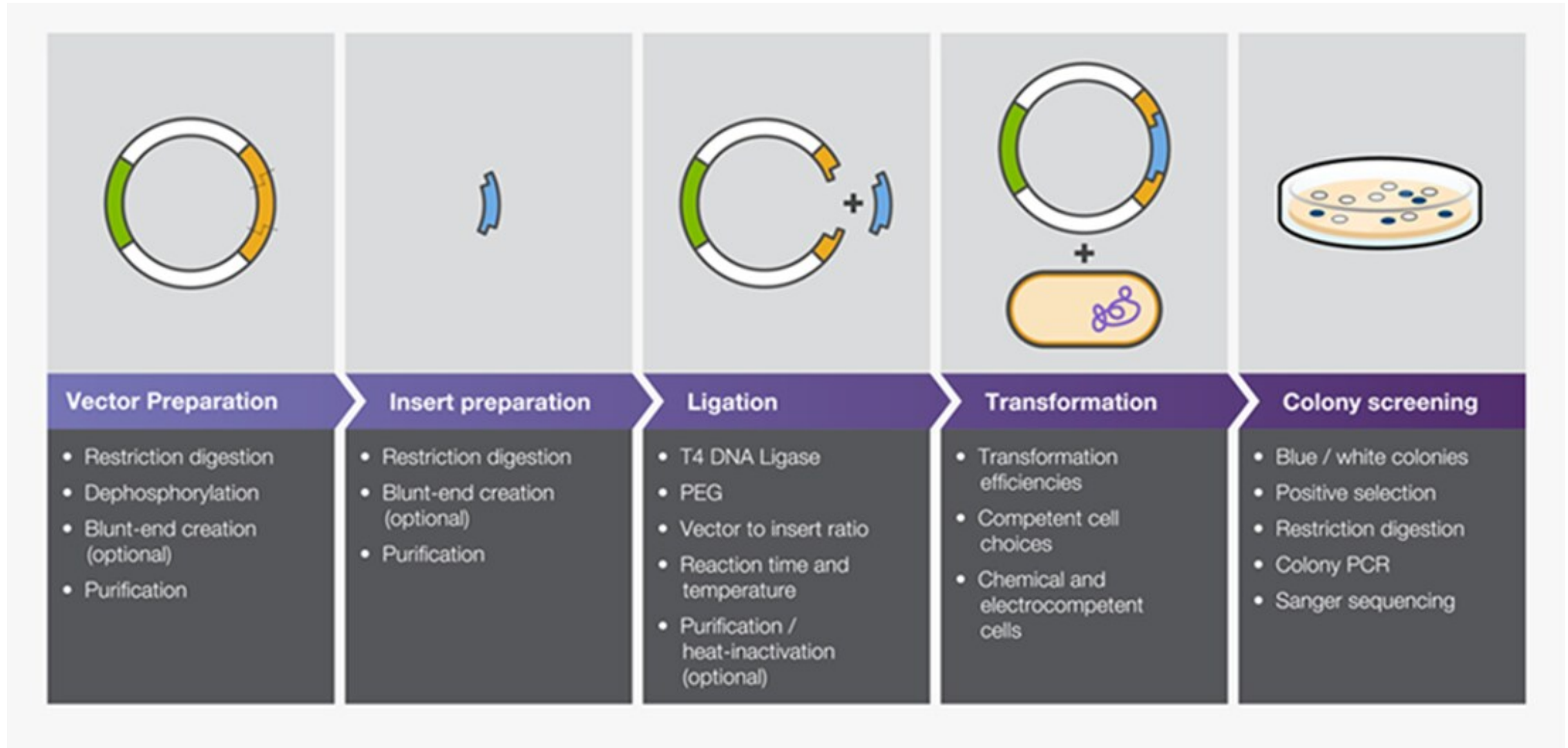


**Gel running**





# The basics: cloning














# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches

Category	Hybridization based			PCR based			
Method	FISH, CARD FISH microscopy Flow cytometry		Microarrays	DGGE	T-RFLP, ARISA	qPCR	HTS
							
Traits							
Analyzed samples per day or run	Tens per day	Tens per day	One per run	Tens per day	Hundreds per day	Hundreds per day	Hundreds per run
Signatures simultaneously screened per sample	Less than 10	Less than 10	Thousands, to hundreds of thousands	Tens	hundreds	One	Hundreds, to hundreds of thousands, to millions
Qualitative / Quantitative	Quantitative	Quantitative	Quantitative **	Semi-quantitative	Quantitative *	Quantitative	Quantitative *
Automated	No	Yes	Yes	No	Yes	Yes	Yes
Confirmation tests or suggested experimental validation	Double probing	Double probing	qPCR	Cloning and sequencing	Cloning, clone to polymorphisms match, clone sequencing	No	No

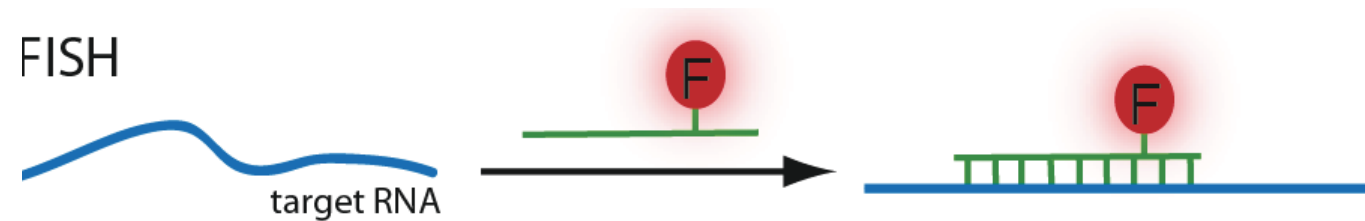
\*: careful preparation (e.g. low number of PCR cycles) is necessary for reducing intensity of the PCR plateau effect on quantification abilities

\*\* : the PCR plateau effect introduced bias is applicable in case a single marker gene like the SSU is screened through multiple taxa after PCR



# Culture independent approaches (diversity)

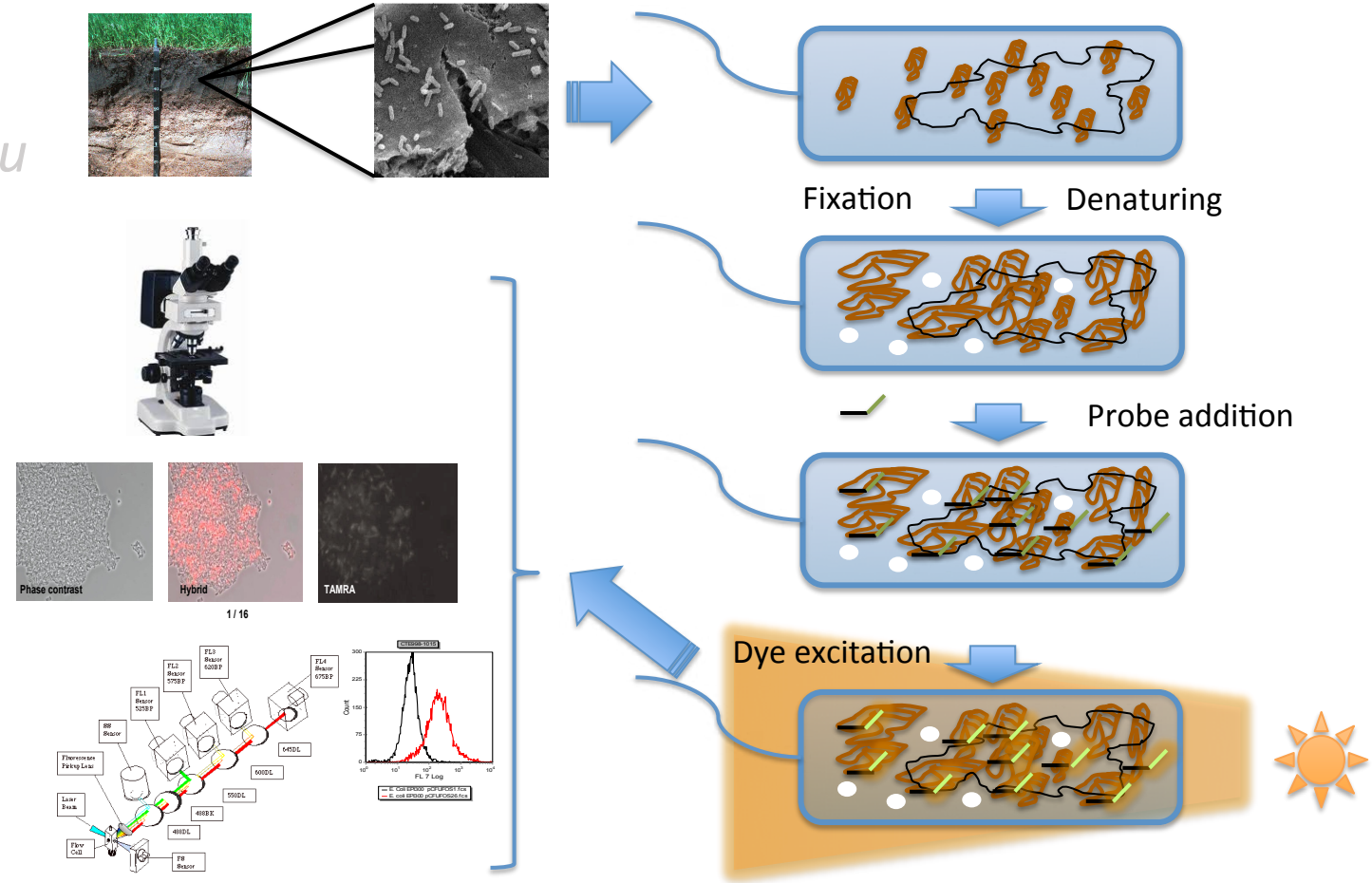
- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

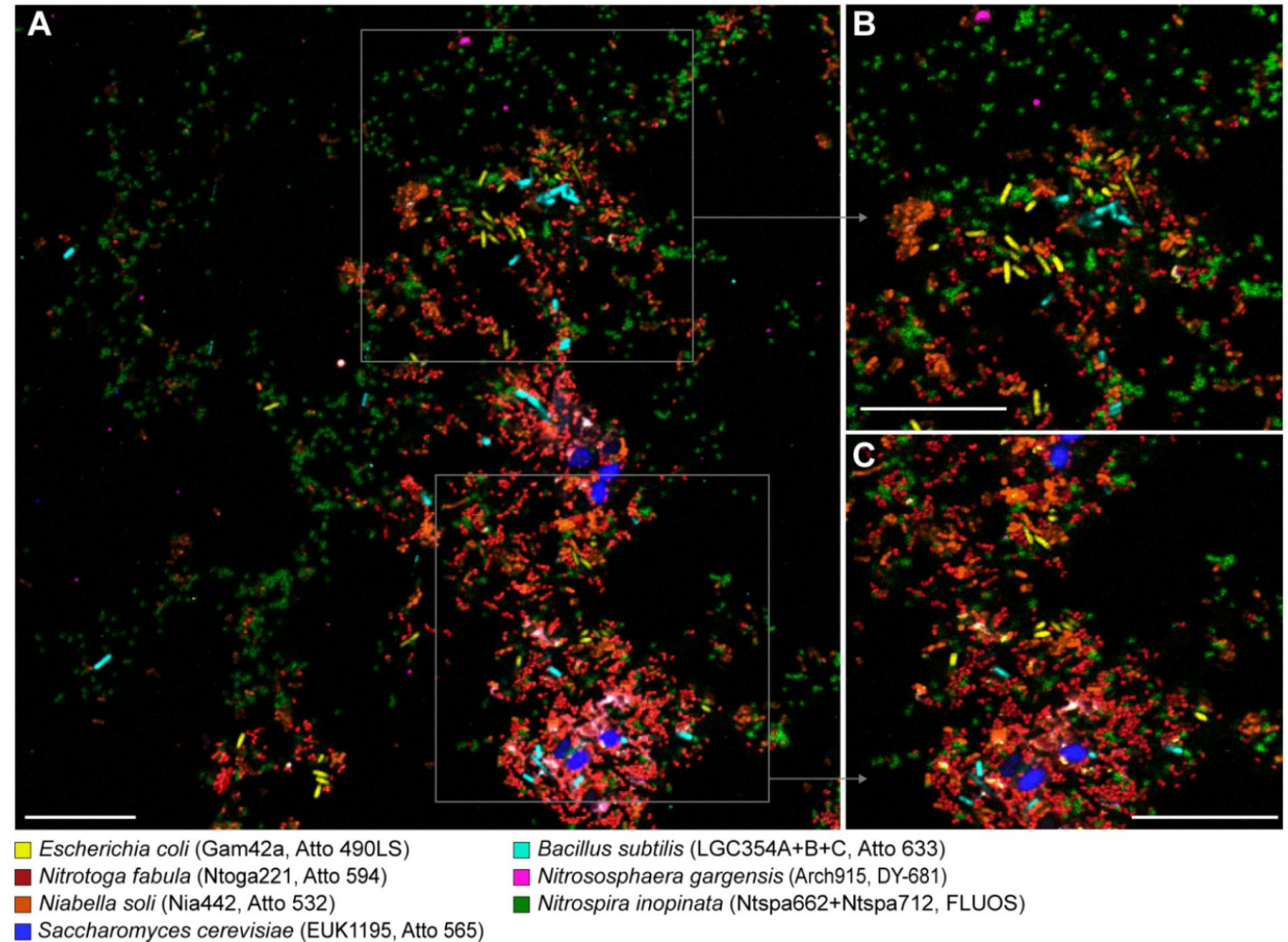
- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

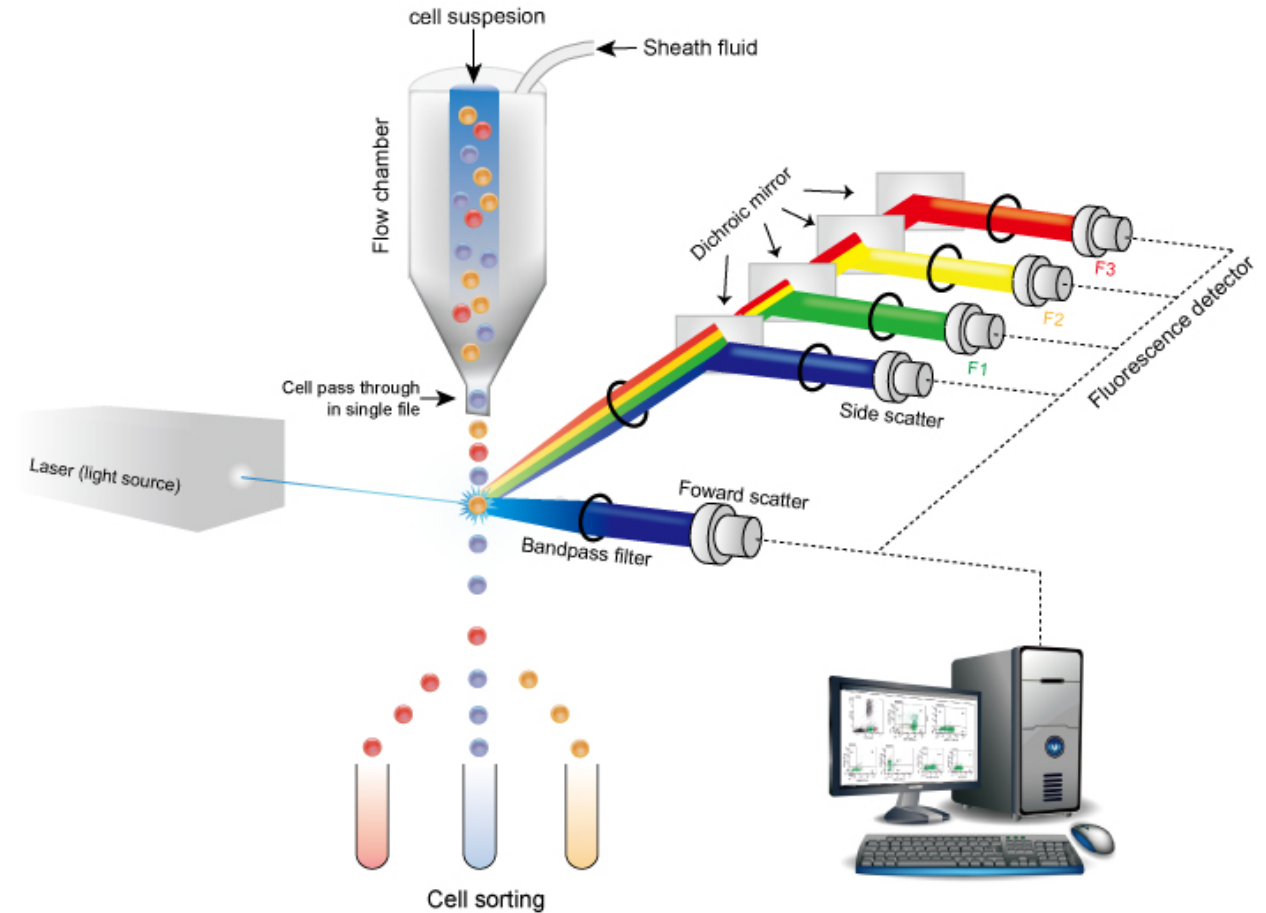
- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





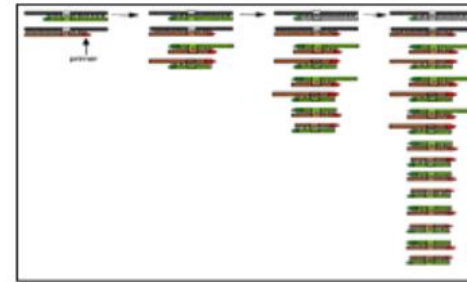


# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches

## PhyloChip analysis

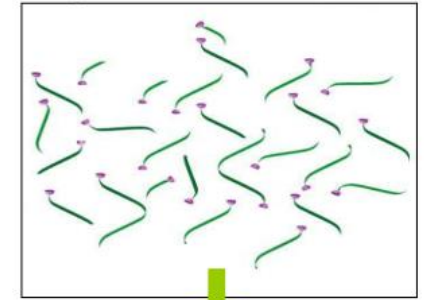
PCR amplification of  
community 16S rRNA genes



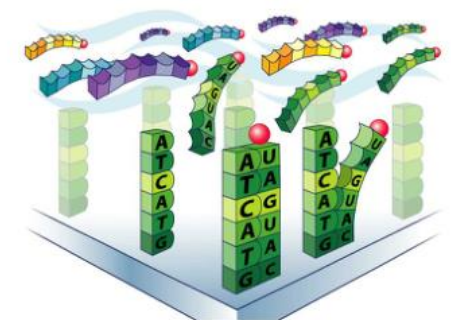
Extract DNA/RNA  
Environmental  
sample



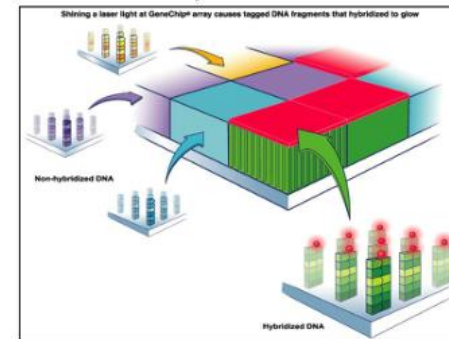
Fragment and biotin label



Hybridize to array



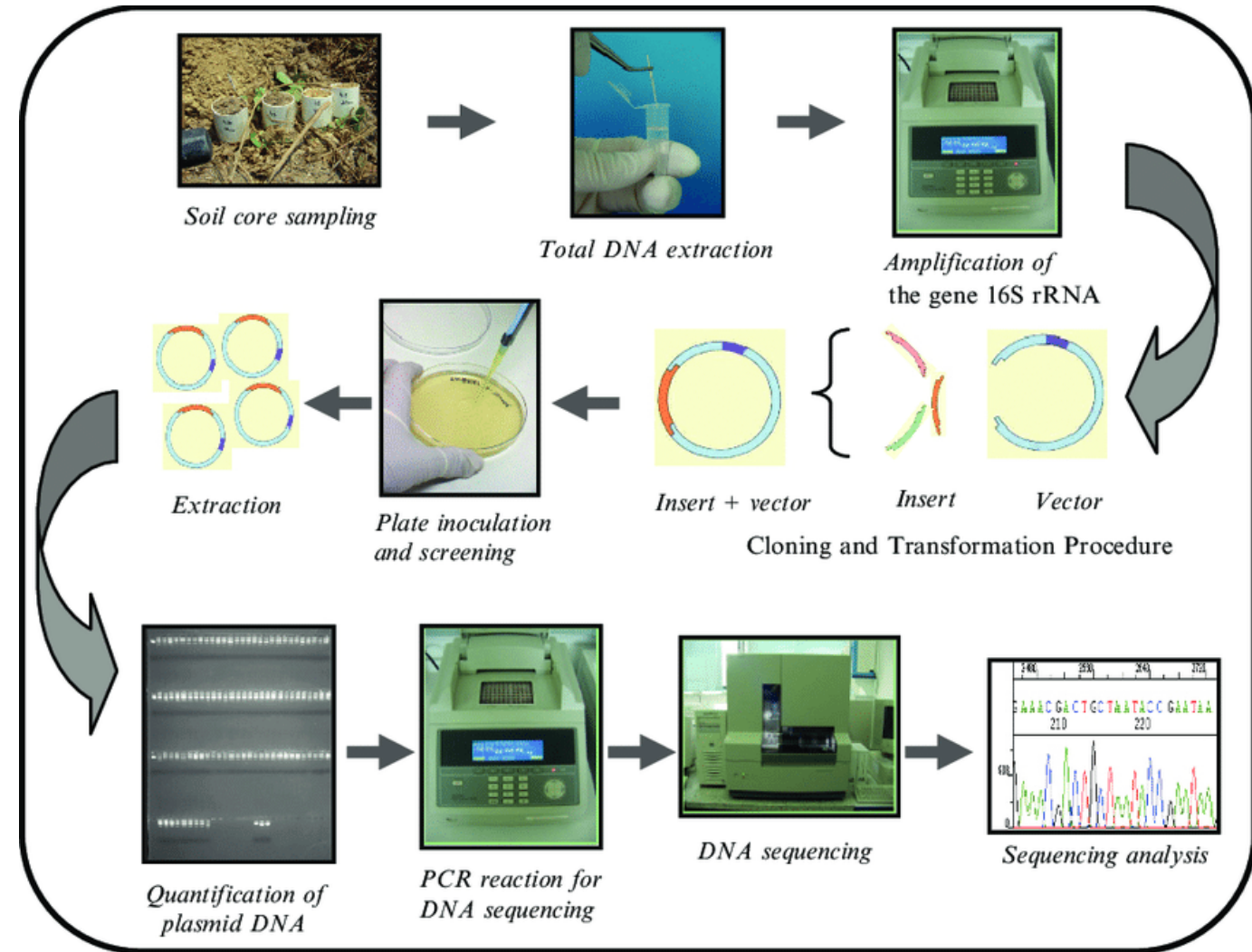
Wash, stain and scan





# Culture independent approaches (diversity)

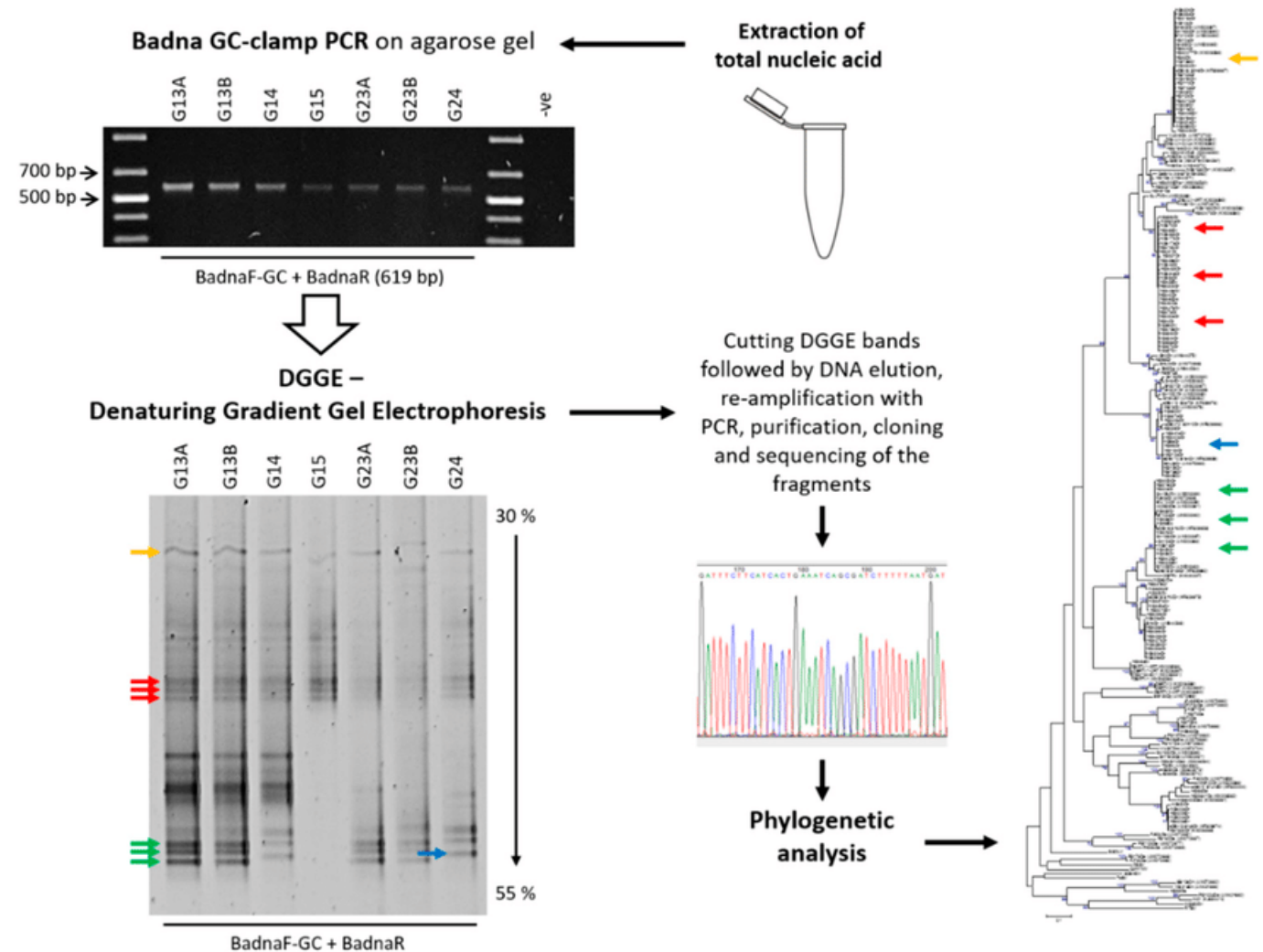
- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

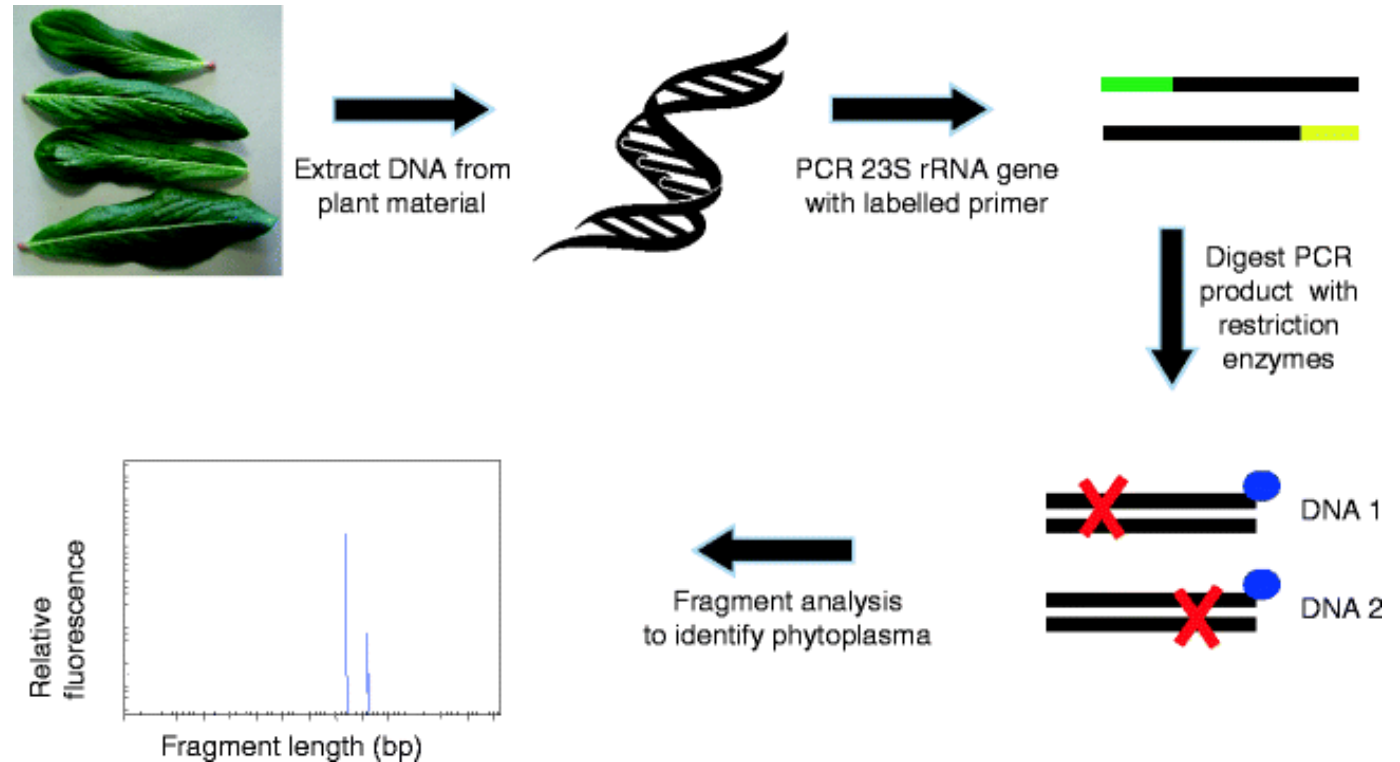
- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches





# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches
- Great... however
  - Probe-based:
    - FISH: few targets at the time, low/intermediate throughput
    - Microarrays: plenty of targets, low throughput and cost ineffective
  - PCR based approaches examined thus far: higher throughput, and more targets, but:
    - for achieving near ambient resolution they can become super expensive





# Culture independent approaches (diversity)

- Hybridization (probes) e.g.:
  - Microscopy (fluorescent *in situ* hybridization - FISH)
  - Flow cytometry
  - Microarrays (Phylochip)
- PCR amplification e.g.:
  - Cloning/clone-sequencing
  - DGGE
  - T-RFLP, ARISA
  - Next generation sequencing approaches

**Solution?**

In cases of highly diverse environments, we resort to

Next generation sequencing (NGS) of  
phylogenetic marker gene amplicons (PMGA)