Microbial processes and monitoring

Lecture 2

Learning outcomes

- Microbial kinetics
- Metabolism
- Microbial monitoring

Microbial kinetics

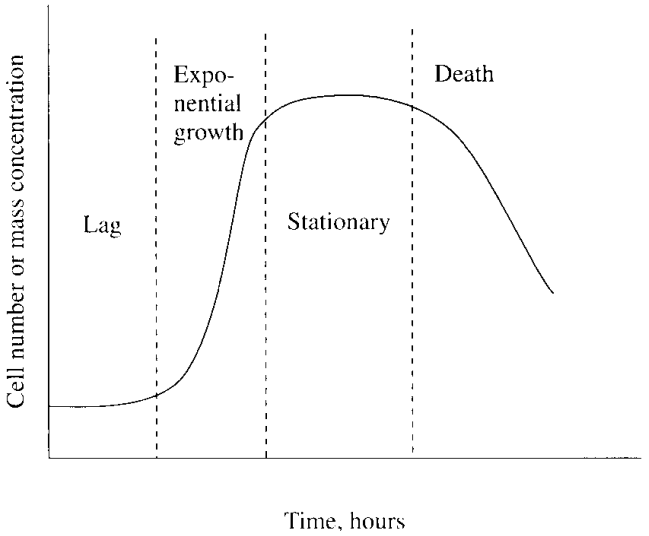
Biomass production

Bacterial biomass production is often accounted for with empirical cell ratios:

$C_5H_7O_2N$

or (more complex)

$C_{60}H_{87}O_{23}N_{12}P$



$N_t=N_0e^{kt}$ (first order kinetics)

N_t= number of microorganisms at time t

N_0= number of microorganisms at time 0

k=growth constant

for one doubling time: $N=2N_0 \Rightarrow t_d=\frac{\ln(2)}{k}$

Maximum recorded growth rates for some bacteria, measured at or near their respective optimal temperature, in complex media

Organism	Temperature, °C	Doubling time, h
Vibrio natriegens	37	0.16
Bacillus stearothermophilus	60	0.14
Escherichia coli	40	0.38
Bacillus subtilis	40	0.43
Pseudomonas putida	30	0.75*
Vibrio marinus	15	1.35
Rhodobacter sphaeroides	30	2.2
Mycobacterium tuberculosis	37	≈ 6
Nitrobacter agilis	27	≈ 20*

Source: Stanier et al., 1986.
*Grown in synthetic media.

Specific growth rate

• More generally, the **specific growth rate** (μ) defined as rate at which cells divide (generations per unit time or inverse of doubling time)

$$\frac{dX}{dt} = \mu X \qquad \text{(eq. 1)}$$

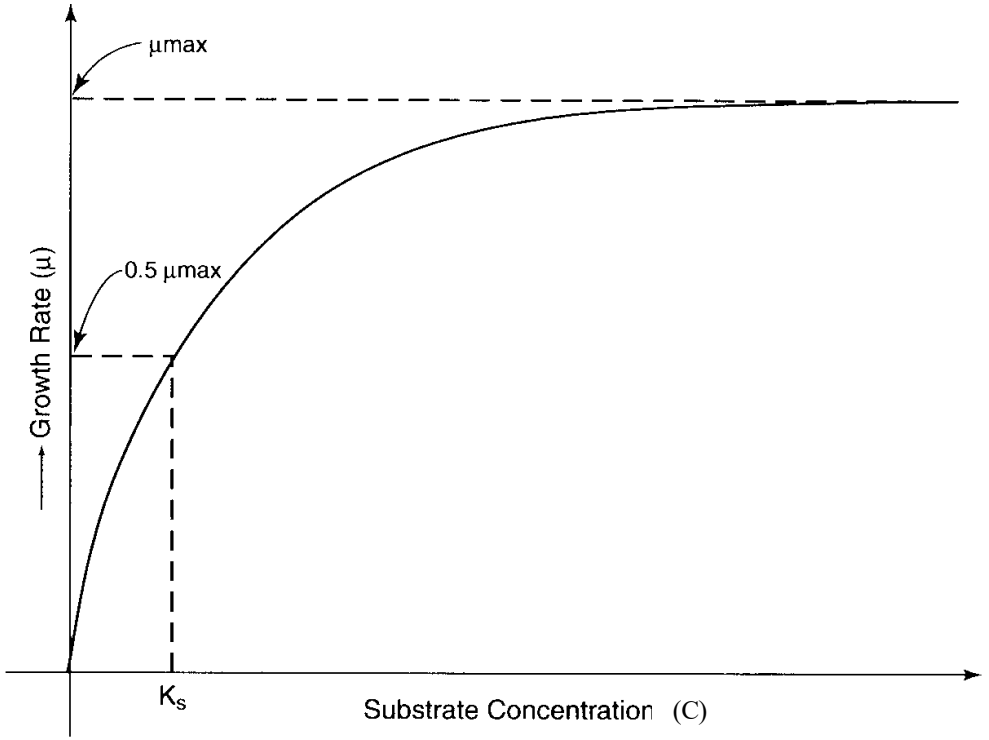
Where:

μ= specific growth rate (time^{-1})
t= time
X= biomass concentration (mass/unit volume)

• A classic relationship exists between bacterial growth rate and substrate concentration (Monod model)

Monod model

$$\mu = \mu_{max} \frac{C}{K_S + C} \qquad \text{(eq. 2)}$$



Where:

μ= specific growth rate (time^{-1})

μ_{max}= maximum specific growth rate (time^{-1})

C= concentration of substrate in solution (mass/unit volume)

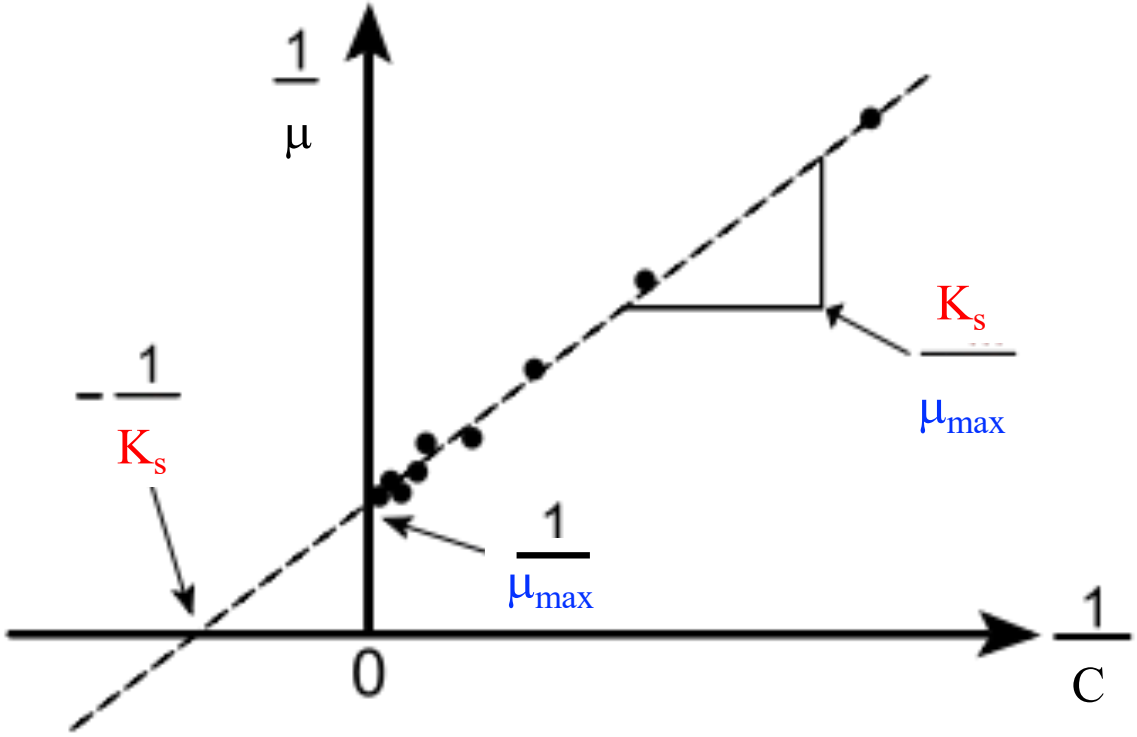
K_S= half-velocity constant (substrate concentration at which μ is one half of μ_{max}) (mass/unit volume)

Lineweaver-Burk plot

$$\mu = \frac{\mu_{max}\ C}{K_S + C}$$

$$\frac{1}{\mu} = \frac{K_s}{\mu_{max} C} + \frac{1}{\mu_{max}}$$

$$\boxed{\frac{1}{\mu} = \frac{K_s}{\mu_{max}} \times \frac{1}{C} + \frac{1}{\mu_{max}}}$$



Modeling growth

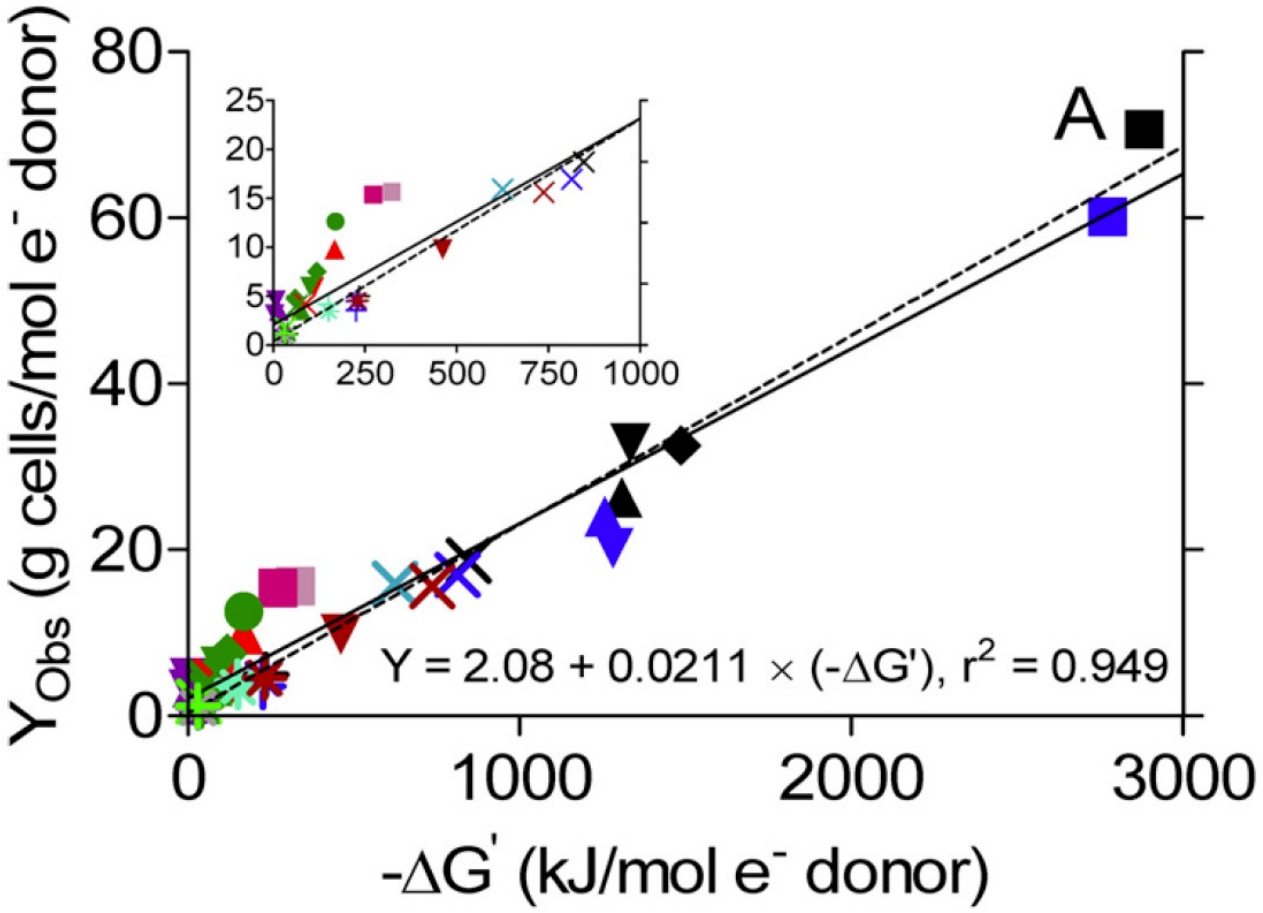
• For organic contaminants that serve as the primary substrate (the sole source of carbon and energy), biomass production is linked to substrate degradation (combine eqs. 1 & 2):

$$\frac{dX}{dt} = \left(\frac{\mu_{max} C}{K_S + C}\right) X \qquad \text{(eq. 3)}$$

• The mass of new cells synthesized per unit mass of substrate removed is constant for a given bacterium and a given substrate. Define the growth yield coefficient (Y= mass biomass/unit mass of substrate)

$$Y = \frac{dX/dt}{dC/dt} \qquad \text{(eq. 4)}$$

Growth yield coefficient (Y)



Roden and Jin, 2011

Y => related to the Gibbs free energy of the reaction

Y => Can be estimated

Modeling growth

Combining eqs. 3 & 4

$$\frac{dC}{dt} = \left(\frac{\mu_{max} C}{Y * (K_S + C)} \right) X \qquad \text{(eq. 5)}$$

The term $\boldsymbol{\mu_{max}/Y}$ can be replaced by the term **k**: degradation rate constant (or maximum rate of substrate removal per unit weight of biomass) (unit mass substrate/ (unit mass biomass . time))

$$\boxed{\frac{\mu_{max}}{Y} = \text{cst} = k}$$

[mass substrate/(mass biomass.time)]

$$\frac{dC}{dt} = \left(\frac{kC}{K_S + C} \right) X \qquad \text{(eq. 6)}$$

and

$$\frac{dX}{dt} = \left(\frac{YkC}{K_S + C} \right) X \qquad \text{(eq. 7)}$$

Modeling growth

In any biological treatment, a portion of the cells are not proliferating but rather in maintenance or dead phase. Need to take death of cells into account

$$\frac{dX}{dt} = \left(\frac{YkC}{K_s + C} - b \right) X \qquad \text{(eq. 8)}$$

Where b= endogenous decay constant (time^{-1})

Minimum substrate concentration

In many cases, the concentration of contaminants is low and does not support viable biomass. At low concentrations, a threshold will be reached where energy needs are not met and biomass loss is observed
This minimum substrate concentration occurs when

$$bX = \left(\frac{YkC_{min}}{K_S + C_{min}} \right) X \qquad \text{(eq. 9)}$$

Where
C_{min} = minimum concentration of substrate to support growth (mass/unit volume)

$$C_{min} = \frac{bK_S}{Yk - b} \qquad \text{(eq. 10)}$$

Typically, C_{min} is in the range of 0.1 to 1.0 mg/L. Some contaminants need to be removed to below 0.01 mg/L. In such cases, it may be necessary to induce cometabolism by adding a primary substrate.

Modeling growth

It is difficult to measure growth in subsurface, so make approximations:

- $C << K_S$

$$\frac{dC}{dt} = \frac{kC}{K_S} X \qquad \text{(eq. 11)}$$

- There is little growth, so X is constant

$$\ln \frac{C}{C_0} = - \frac{k}{K_S} Xt \qquad \text{(eq. 12)}$$

Where:

t= duration of treatment (time)

C_0= initial substrate concentration (mass/unit volume)

$$k' = \frac{k}{K_S} X \qquad \text{(eq. 13)}$$

Where:

k' = first-order degradation rate constant ($time^{-1}$)

Modeling growth

It is difficult to measure growth in subsurface, so make approximations

$$C = C_0 e^{-k' t} \qquad \text{(eq. 14)}$$

- Typically, half-lives of contaminants is reported:

$$k' = \frac{0.693}{t_{1/2}} \qquad \text{(eq. 15)}$$

Where:

$t_{1/2}$ = half-life (time)

In the subsurface

Use the solid phase concentration of cells (attached to soil particles) rather than solution phase cell conc.

(from eq. 3)

$$\frac{dX}{dt} = \left(\frac{\mu_{max} C}{K_S + C}\right) B \qquad \text{(eq. 16)}$$

Where B= solid phase concentration of cells mg cells/g soil

and $B = B_0 + Y * (C_0 - C)$ (eq. 17)

Example: Y = 0.5 g/g for aerobic bacteria or <0.1 g/g in some soils

Full equation

Putting the three equations together : eq. 5, 21, 22:

$$\frac{dC}{dt} = \frac{\mu_{max}}{Y} \frac{C\,[B_0 + Y\,(C_0\text{-}C)]}{K_s + C} \qquad \text{(eq. 18)}$$

In practice, (eq. 18) is often simplified:

Model	Necessary conditions	Differential form	Integral form
Zeroth-order	$C_0 >> K_S$ $B_0 >> YC_0$	$\frac{dC}{dt} = \frac{\mu_{max}}{Y} B_0$	$C = C_0 + \frac{\mu_{max} B_0}{Y} t$
Monod-no growth	B= constant	$\frac{dC}{dt} = \frac{\mu_{max} B_0 C}{Y(K_S + C)}$	$K_S \ln \frac{C}{C_0} + C - C_0 = \frac{B_0}{Y} \mu_{max} t$
First-order-no growth	$C_0 << K_S$ B_0= constant	$\frac{dC}{dt} = \frac{\mu_{max} B_0 C}{YK_S}$	$C = C_0 \exp\left(\frac{\mu_{max} B_0}{YK_S} t\right)$

Important kinetic parameters

Term	Definition	Significance
k	Degradation rate constant ($time^{-1}$)	Indication of degradability of an organic compound; for biological treatment to be effective high values are preferred
K_s	Half-velocity constant (mass/unit volume)	Indication of efficiency of the degradation; to degrade organics to low concentrations, low values are needed.
Y	Yield of biomass generated per unit mass of substrate removed (mass/unit mass)	Indication of utilization of substrate for the production of biomass
b	Endogenous decay constant ($time^{-1}$)	Indication of the rate of loss of biomass due to natural death and decay

Microbially-catalyzed degradation

Microbial growth requirements

- Electron acceptors: O_2 (aerobic) or NO_3^-, Fe(III), SO_4^{2-}, etc..
- Moisture: absolute minimum needed for biological treatment is 40% saturation of soil /sediment
- pH: most bacteria grow best at pH 6-8. Most bacteria die off at pH<4 and pH>9.5
- Inorganic nutrients: empirical equation for cellular material is $C_{60}H_{87}O_{23}N_{12}P$ (or simply $C_5H_7O_2N$). Rule of thumb TOC:N:P mass ratio of 20:5:1.
- Micronutrients: S, K, Ca, Mg, Fe, Ni, Cu, Zn, vitamins

Energetics

- Microbial metabolism is based on thermodynamics: If there is energy available in a reaction, a microbe may catalyze that reaction
- Most biologically-mediated transformations rely on redox transformations
- The relationship between energy and redox potential is captured in the Nernst equation:

$\Delta G^0 = -nF\ \Delta E_0$

where n= number of electrons transferred

F=faraday's constant 96,630 J/V

$\Delta E_0 = E_0$ (e- accepting couple)- E_0(e- donating couple)

TABLE 5.3

The electron tower: standard reduction potentials at 25°C and pH 7 for selected environmentally important redox couples

Half reaction	E_0, V
$6CO_2 + 24H^+ + 24e^- = C_6H_{12}O_6 + 6H_2O$	–0.43
$2H^- + 2e^- = H_2$	–0.41
$CO_2 + 6H^- + 6e^- = CH_3OH + H_2O$	–0.38
$NAD^+ + 2H^+ + 2e^- = NADH + H^+$	–0.32
$CO_2 + HCO_3^- + 8H^- + 8e^- = CH_3COO^- + 3H_2O$	–0.29
$CO_{2(g)} + 8H^+ + 8e^- = CH_{4(g)} + 2H_2O$	–0.25
$S_{(s)} + 2H^- + 2e^- = H_2S_{(g)}$	–0.24
$SO_4^{-2} + 9H^- + 8e^- = HS^- + 4H_2O$	–0.22
Pyruvate + $2H^+ + 2e^-$ = lactate	–0.19
$FeOOH_{(s)} + HCO_3^- + 2H^+ + e^- = FeCO_{3(s)} + 2H_2O$	–0.05*
$CH_3SOCH_3 + 2H^- + 2e^- = CH_3SCH_3 + H_2O$	0.16
$NO_3^- + 10H^+ + 8e^- = NH_4^+ + 3H_2O$	0.36
$NO_3^- + 2H^- + 2e^- = NO_2^- + H_2O$	0.42
$MnO_{2(s)} + HCO_3^- + 3H^- + 2e^- = MnCO_{3(s)} + 2H_2O$	0.52*
$CHCl_3 + H^- + 2e^- = CH_2Cl_2 + Cl^-$	0.56
$CCl_4 + H^+ + 2e^- = CHCl_3 + Cl^-$	0.67
$2NO_3^- + 12H^+ + 10e^- = N_2 + 6H_2O$	0.74
$Fe^{3+} + e^- = Fe^{2-}$	0.76
$O_{2(g)} + 4H^+ + 4e^- = 2H_2O$	0.82
$CCl_3CCl_3 + 2e^- = CCl_2CCl_2 + 2Cl$	1.13
$2HOCL + 2H^- + 2e^- = Cl_2 + 2H_2O$	1.18

*Based on $[HCO_3^-] = 10^{-3}$ M.

Problem:

A groundwater is contaminated with hexachloroethane Cl_3CCCl_3.

An environmental engineer is concerned that this compound will be transformed to tetrachloroethylene (Cl_2CCCl_2), an even more toxic compound, by reacting with nitrite in the groundwater.

Do you think this is likely to happen? Why or why not?

Degradation strategies vary depending on the compound

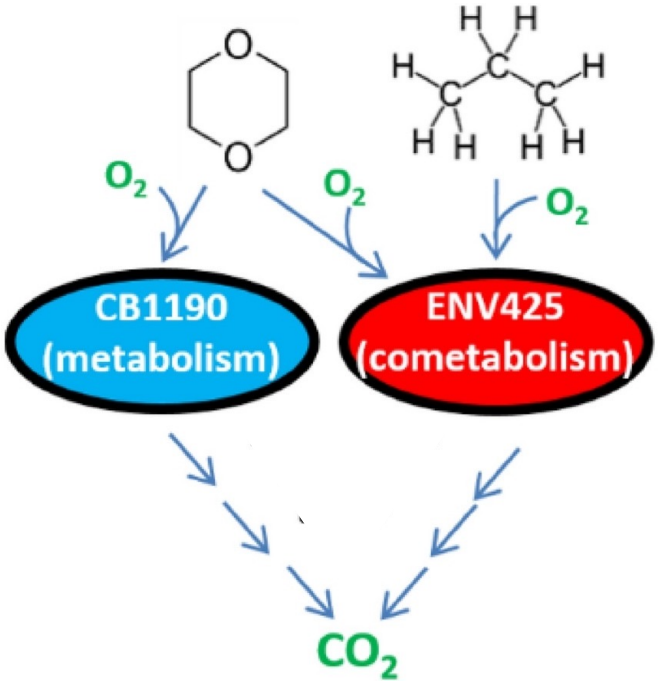
- Aerobic conditions:
  - Contaminant as an e^- donor and C source and O_2 as e^- acceptor- couple **growth** to the oxidation of contaminant
  - **Cometabolism**: organism grows on another primary substrate and also oxidizes the contaminant but obtains no benefit-O_2 as e^- acceptor

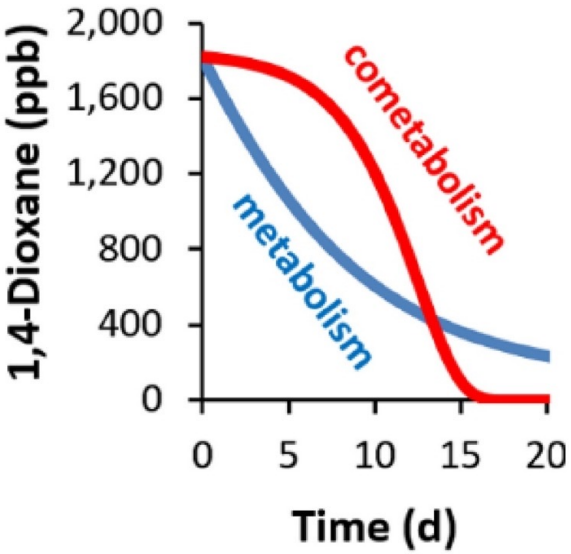
- Anaerobic conditions:

  - Contaminant as an e^- donor and C source and NO_3^-, SO_4^{2-} or Fe(III) as e^- acceptor- couple **growth** to the oxidation of contaminant
  - Contaminant as an e^- acceptor and e.g., H_2 as e^- donor: **Halorespiration**
  - **Cometabolism**: organism grows on another primary substrate and also oxidizes or reduces the contaminant (serves as e^- donor or acceptor but not coupled to growth)

Metabolism vs. cometabolism

- Organic compounds used as source of energy are most likely to be degraded
- Some compounds not used as energy source can be transformed through cometabolism
- Cometabolism: fortuitous transformation of a compound by enzymes designed for other purposes





1,4-dioxane metabolism by *Pseudonocardia dioxanivorans* (CB1190) compared to cometabolism by the propanotroph *Rhodococcus ruber (*ENV425)

Barajas-Rodriguez et al., 2018, JHM

Cometabolism

Transformation of a substrate by a microorganism that is unable to use the substrate as an energy source

Many enzymes can carry out cometabolism:

oxygenases

- Methane monooxygenase: methylotrophs can transform DCE
- Toluene dioxygenase: incorporates both atoms of molecular oxygen into toluene. Also transforms TCE
- Ammonia monoxygenase (*Nitrosomonas europaea*): chemoautotroph that cometabolizes TCE, DCE

CO dehydrogenase

- CO dehydrogenase of acetoclastic methanogens
- CO dehydrogenase pathway (oxidative) sulfate reducers
- CO dehydrogenase pathway (reductive) in homoacetogens

Co-factors

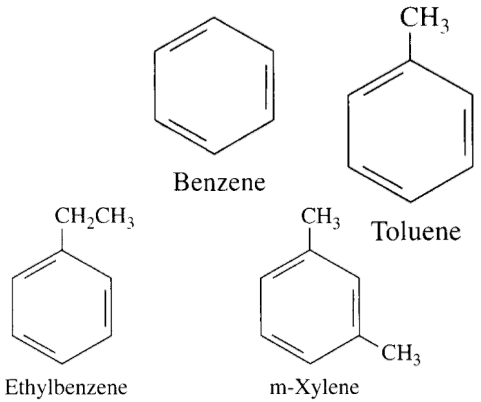
- Cometabolic reductive dehalogenation catalyzed by coenzyme B12 that is present in many anaerobic bacteria such as Clostridia, propionic acid bacteria
- Cytochrome P_{450} from *Pseudomonas* spp.
- F_{430} in methyl-CoM reductase in methanogens

Types of substrate

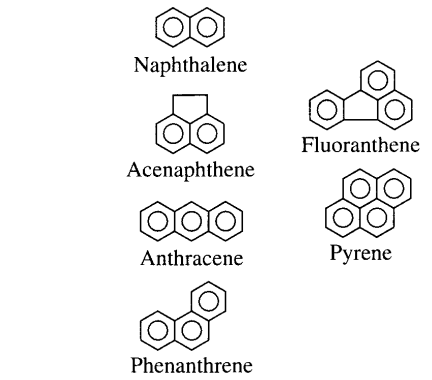
- Primary substrate: sole source of C and energy
- Secondary substrate: for cometabolism the substrate being degraded is not the same as that supporting growth
- Limiting substrate: one substrate (often O_2) can be limiting

Petroleum hydrocarbons

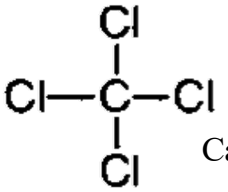
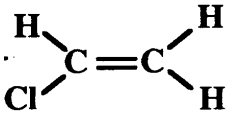
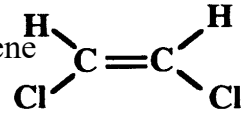
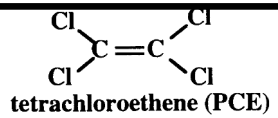
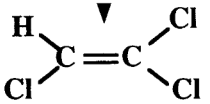
Benzene- Toluene- Ethylbenzene- Xylene (BTEX)

compound	aerobic		anaerobic		
	e- donor	process	e^- donor	e^- acceptor	process
			BTEX	nitrate	Anaerobic addition of O (from fumarate)
	Methane, toluene, NH_4^+	Cometabolism (oxygenase)			
	BTEX	Respiration (O_2 e^- acceptor)			

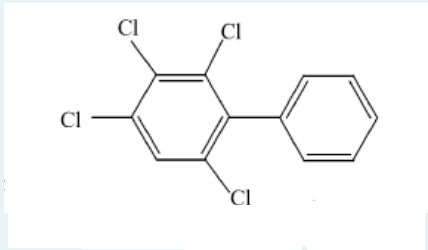
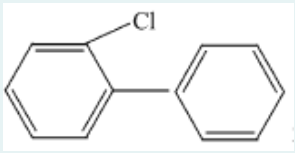
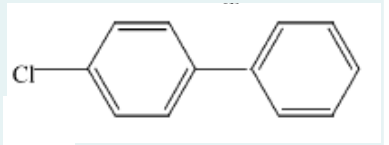
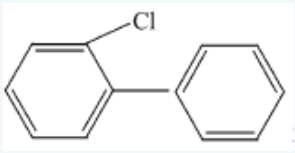
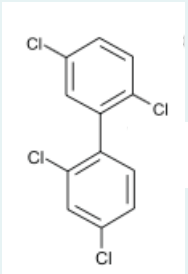
Polycyclic aromatic hydrocarbon (PAH)

compound	aerobic		anaerobic		
	e- donor	process	e^- donor	e^- acceptor	process
			PAH	nitrate	Denitrification
	Biphenyl or m-xylene	Cometabolism (dioxygenase)			
	Naphthalene	Respiration (O_2 e^- acceptor)			

Halogenated hydrocarbons

compound	aerobic		anaerobic		
	e- donor	process	e$^-$ donor	e$^-$ acceptor	process
 Carbon tetrachloride			H_2	CO_2 to CH_4	cometabolism
			H_2	CO_2 to acet.	cometabolism
			Acetate	SO_4^{2-}	cometabolism
			Lactate	NO_3^-	PDTC secretion
 Vinyl chloride	VC or DCE	respiration			
	Methane, toluene, NH_4^+	cometabolism			
dichloroethylene 			H_2	VC, DCE	halorespiration
			H_2	CO_2	cometabolism
 tetrachloroethene (PCE)			Methanol	Methanol	cometabolism
			H_2	PCE, TCE	halorespiration
 trichloroethene (TCE)	Methane, toluene, NH_4^+	Cometabolism of TCE, PCE			

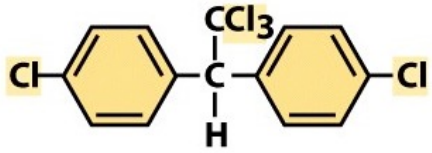
Polychlorinated biphenyls

Compound	aerobic	anaerobic
 2,3,4,6-tetrachlorobiphenyl		Dechlorination to: 4-chlorobiphenyl  2-chlorobiphenyl 
	Carbon source	
 2,2',4,5'-tetrachlorobiphenyl	2,3-dioxygenase attack *Pseudomonas* sp. 2	

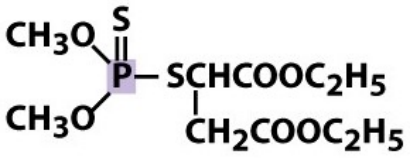
Pesticides

Persistence of herbicides and insecticides in soil

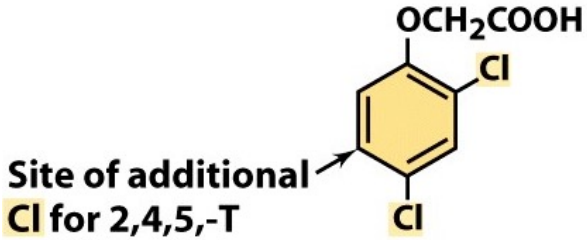
Substance	Time for 75–100% disappearance
Chlorinated insecticides	
DDT [1,1,1-trichloro-2,2-bis-(ρ-chlorophenyl)ethane][a]	4 years
Aldrin	3 years
Chlordane	5 years
Heptachlor	2 years
Lindane (hexachlorocyclohexane)	3 years
Organophosphate insecticides	
Diazinon	12 weeks
Malathion[a]	1 week
Parathion	1 week
Herbicides	
2,4-D (2,4-dichlorophenoxyacetic acid)	4 weeks
2,4,5-T (2,4,5-trichlorophenoxyacetic acid)	20 weeks
Dalapon	8 weeks
Atrazine[a]	40 weeks
Simazine	48 weeks
Propazine	1.5 years



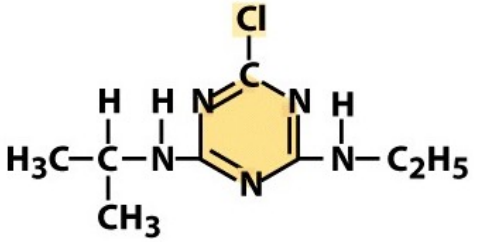
DDT; dichlorodiphenyltrichloro-ethane (an organochlorine)



Malathion; mercaptosuccinic acid diethyl ester (an organophosphate)



2,4-D; 2,4-dichlorophenoxy acetic acid (a chlorophenoxy acetic acid derivative)



Atrazine, 2-chloro-4-ethylamino -6-isopropylaminotriazine (a triazine derivative)

Biological monitoring

Characterization of microbial community and activity

- Culture-dependent approaches
- DNA-based approaches
- RNA-based approaches

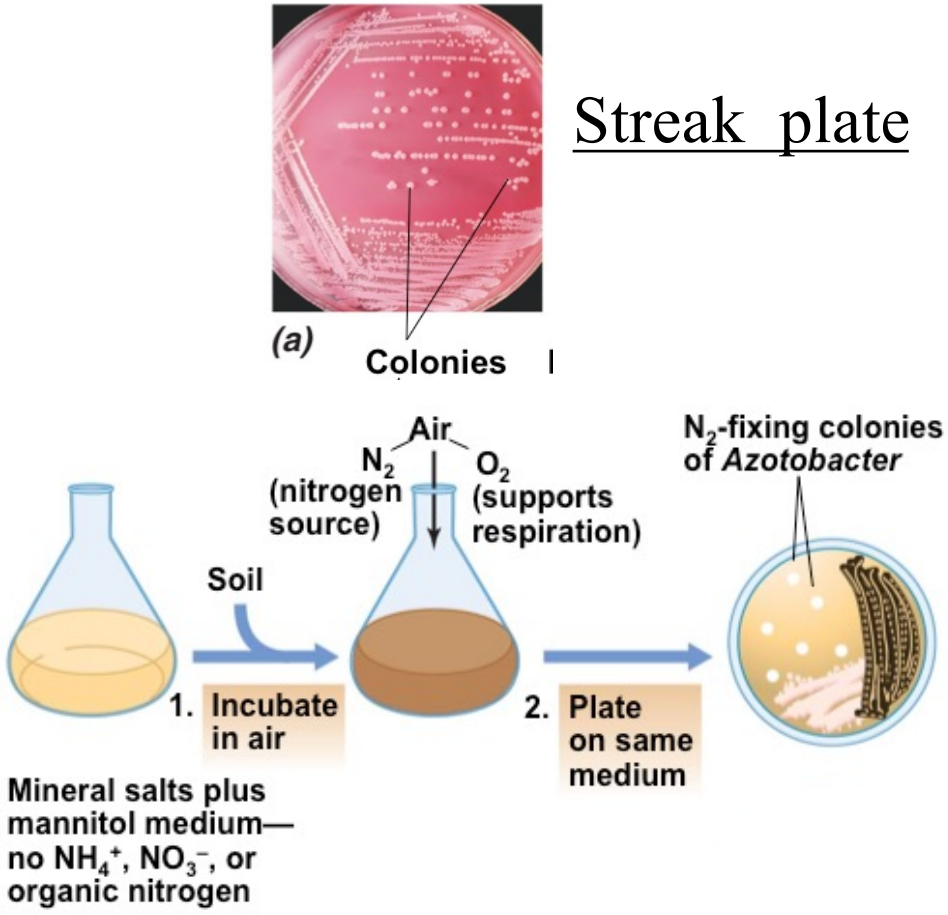
Culture-dependent characterization

- All culture-dependent methods have a major limitation: < 0.5% of soil microorganisms are cultivable.
- **Enrichment culture**: a medium and set of incubation conditions are established that are selective for the desired organism
- Limitations:
  - Negative result does not guarantee absence of organism/ Can prove a positive but not a negative
  - No information on ecological importance or abundance of organism
  - Bias: liquid enrichment results different from plating results
  - Bias: rapidly growing ('weed') species appear quantitatively important (combat with dilution of inoculum)

Dark-chemolithotrophic bacteria: main C source, CO_2 (medium must lack organic C)			
Incubation in air: aerobic respiration			
Electron donor	***Electron acceptor***	***Organisms enriched***	***Inoculum***
NH_4^+	O_2	Ammonia-oxidizing bacteria (*Nitrosomonas*)	Soil, mud; sewage effluent
NO_2^-	O_2	Nitrite-oxidizing bacteria (*Nitrobacter, Nitrospira*)	
H_2	O_2	Hydrogen bacteria (various genera)	
H_2S, S^0, $S_2O_3^{2-}$	O_2	*Thiobacillus* spp.	
Fe^{2+}, low pH	O_2	*Acidithiobacillus ferrooxidans*	
Anoxic incubation			**Inoculum**
S^0, $S_2O_3^{2-}$	NO_3^-	*Thiobacillus denitrificans*	Mud, lake sediments, soil
H_2	NO_3^- + yeast extract	*Paracoccus denitrificans*	

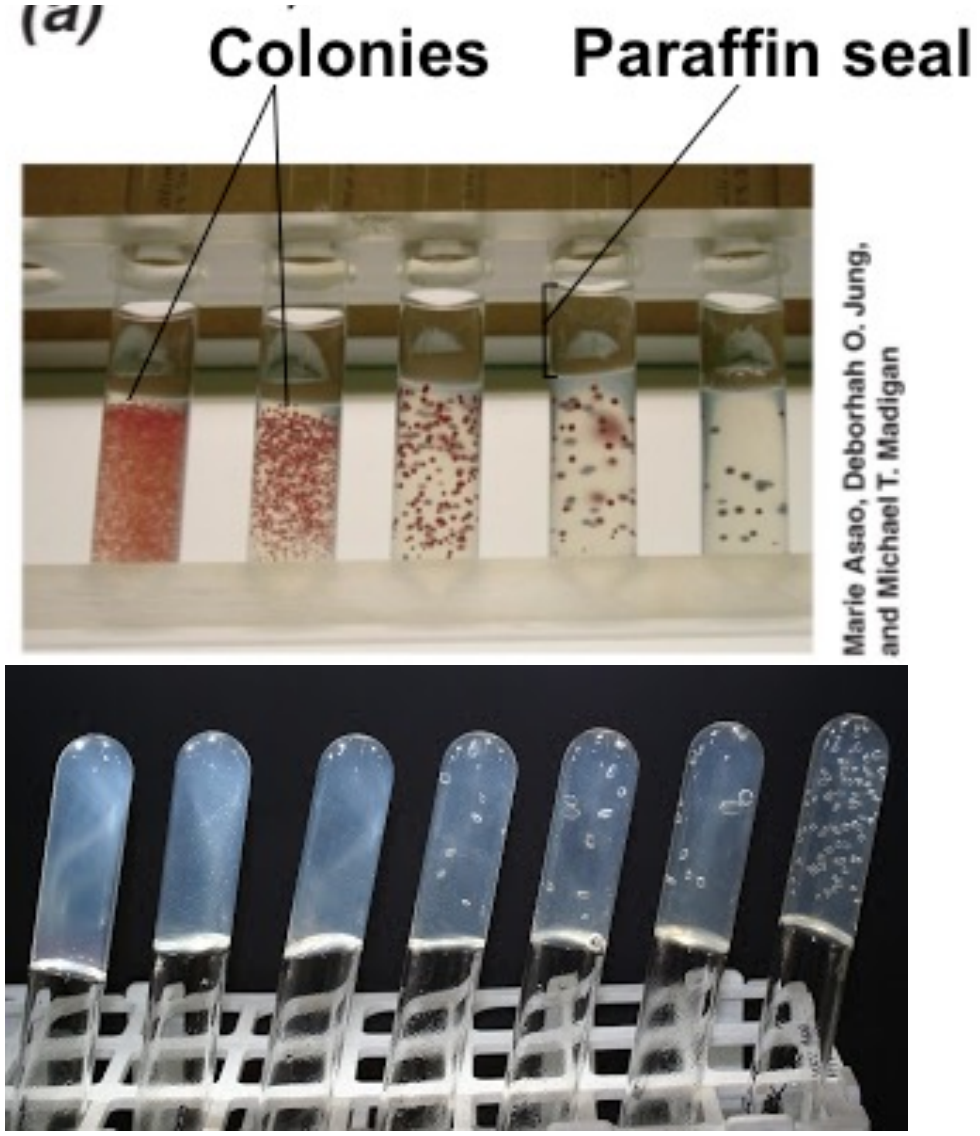
Culture-dependent characterization

- **Isolation**: obtaining a pure culture (a single kind of microorganism)
  - Streak plate: for organisms that form colonies on agar
  - Agar shake: mixed culture diluted in tubes of molten agar (colonies embedded in agar). Useful to isolate anaerobic microbes. Successive dilutions into molten agar medium.
  - Liquid dilution: serially dilute an inoculum into a liquid medium. Most probable number (MPN) method.



Culture-dependent characterization

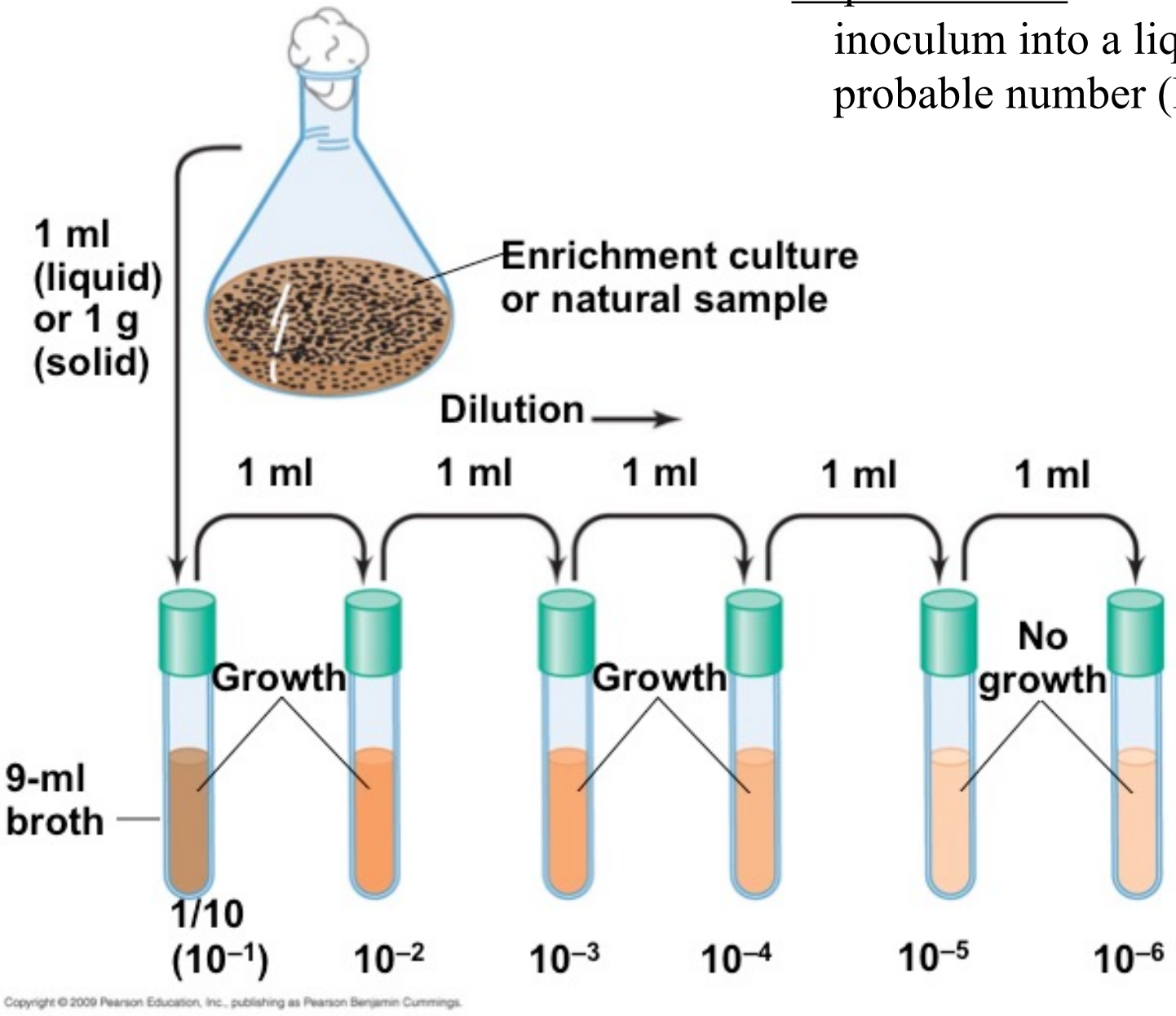
Agar shake: mixed culture diluted in tubes of molten agar (colonies embedded in agar). Useful to isolate anaerobic microbes. Successive dilutions into molten agar medium.



Agar shakes

Culture-dependent characterization

Liquid dilution: serially dilute an inoculum into a liquid medium. Most probable number (MPN) method.

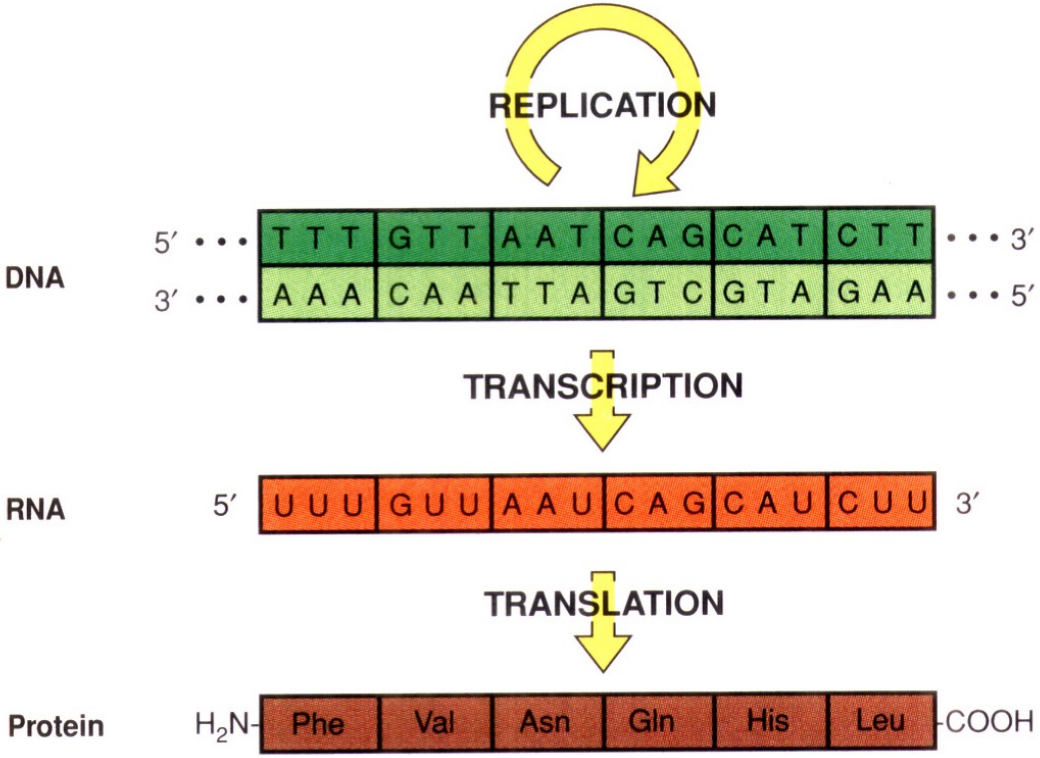


Culture-independent approaches

Culture-independent genetic analysis of microbiomes

- 16S rRNA amplicon sequencing
- Metagenomics

Molecular biology primer

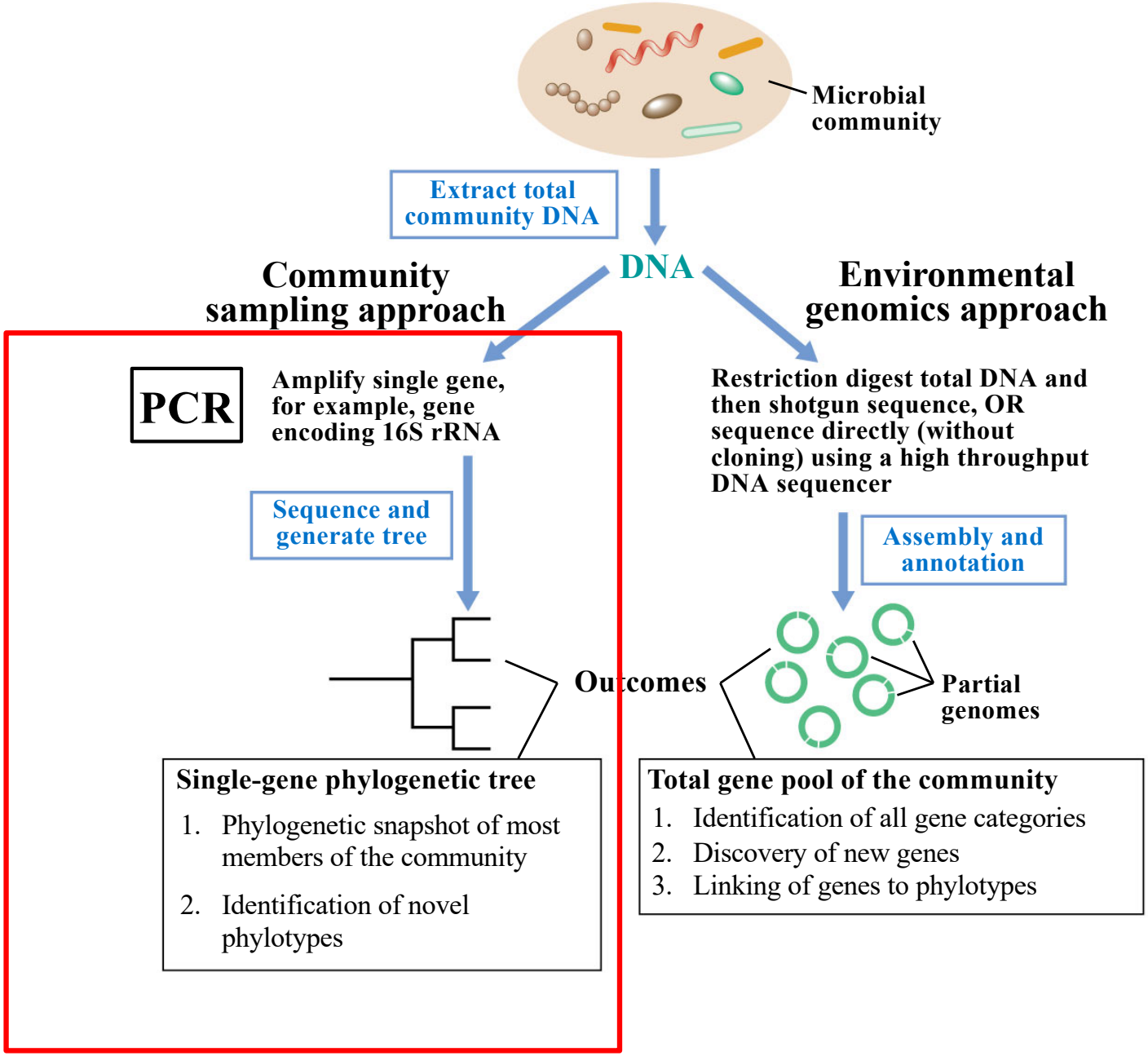


- DNA is somewhat stable even if microbe is dead
- DNA-based methods identify the presence of a microorganism but not necessarily activity

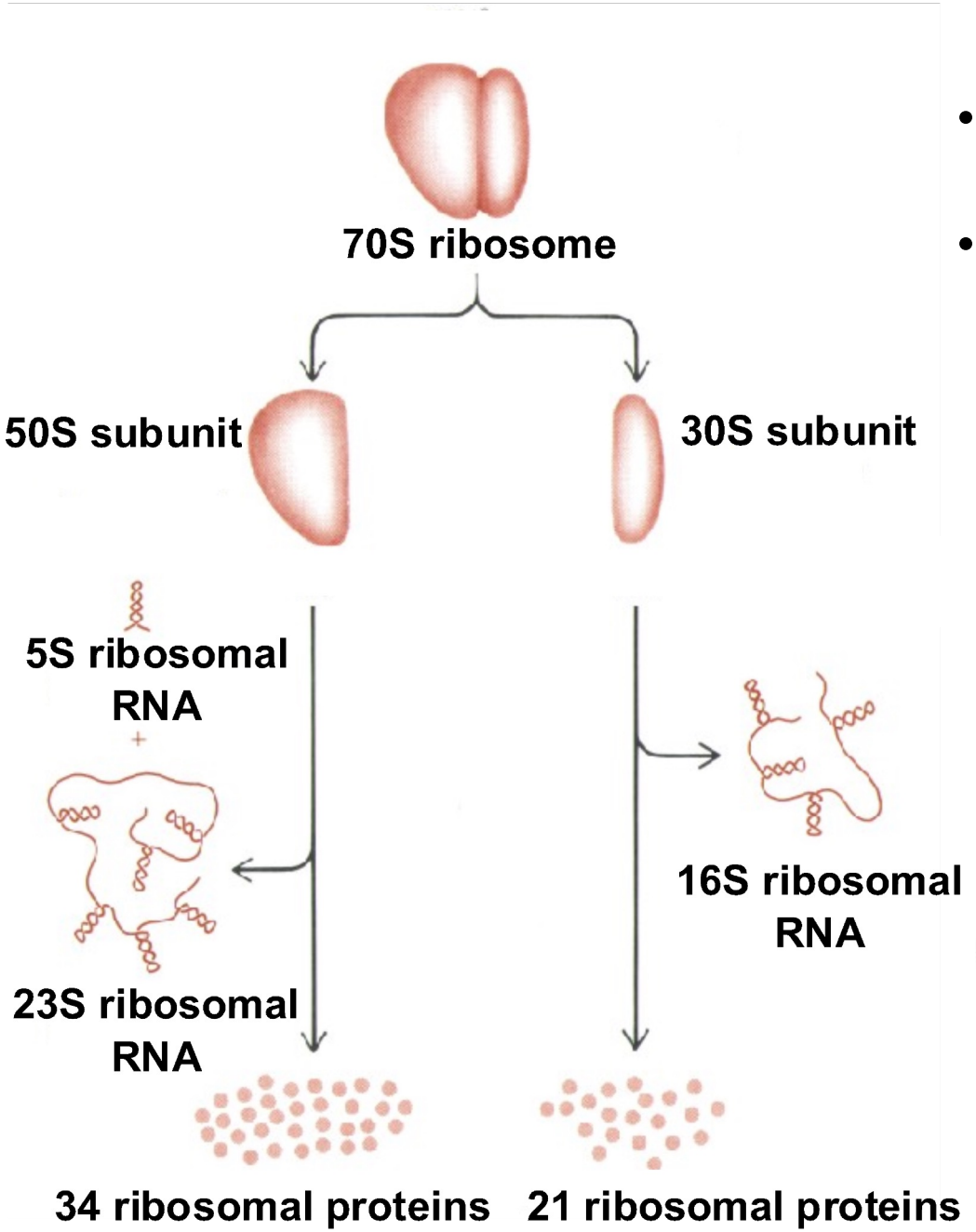
- RNA only transcribed when protein is needed
- messenger RNA (mRNA) is degraded quickly
- mRNA-based methods probe active population

- proteins only present when need for specific activity
- proteins fairly stable
- few protein-based techniques because difficult to obtain protein from soil /sediment

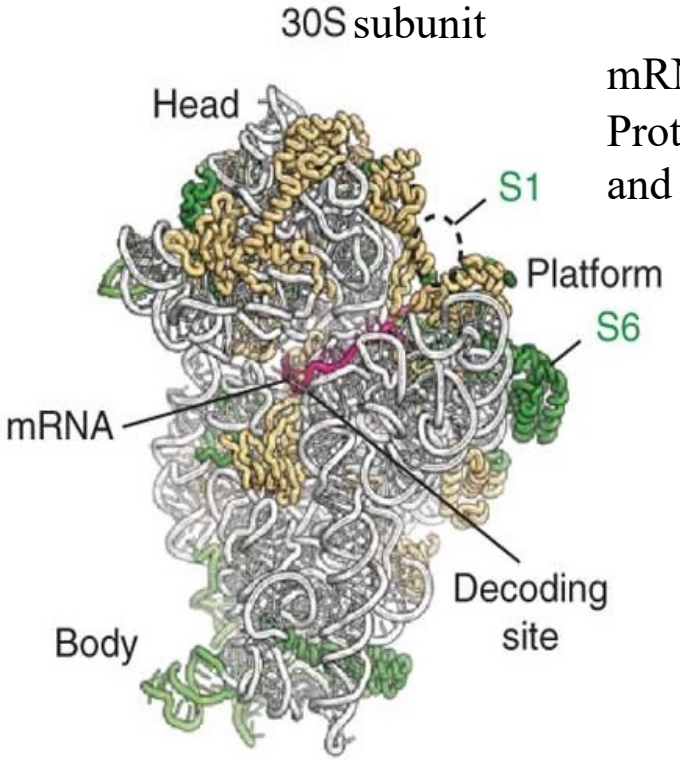
Single-gene *vs*. genomic approaches



Ribosomes and ribosomal RNA

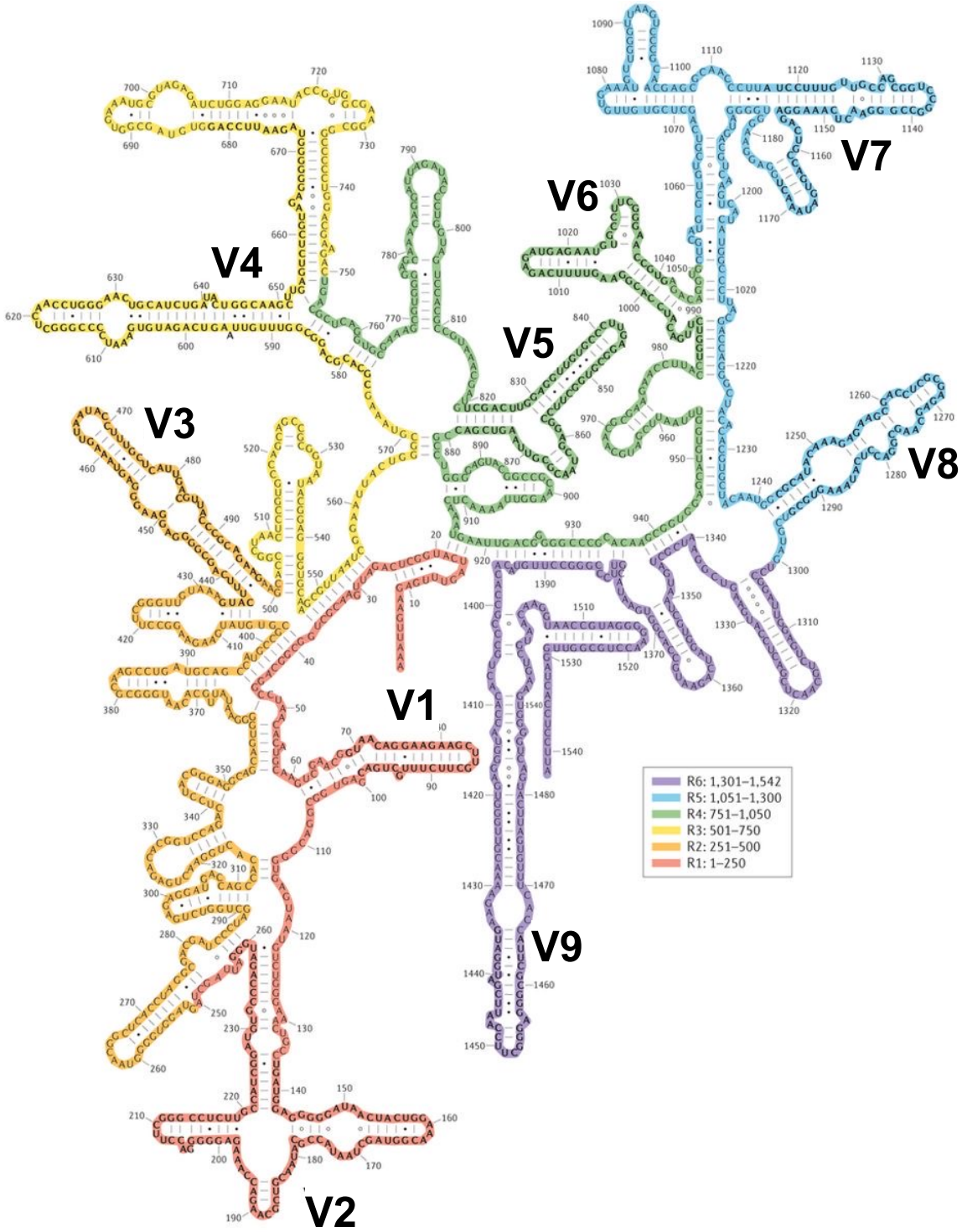


- Ribosomes are responsible for the translation of mRNA to protein
- They themselves consist of ribosomal RNA and proteins
- Bacteria/Archaea
  5S rRNA: 120 bp
  16S rRNA: 1500 bp
  23S rRNA: 3000 bp
  + 55 Proteins



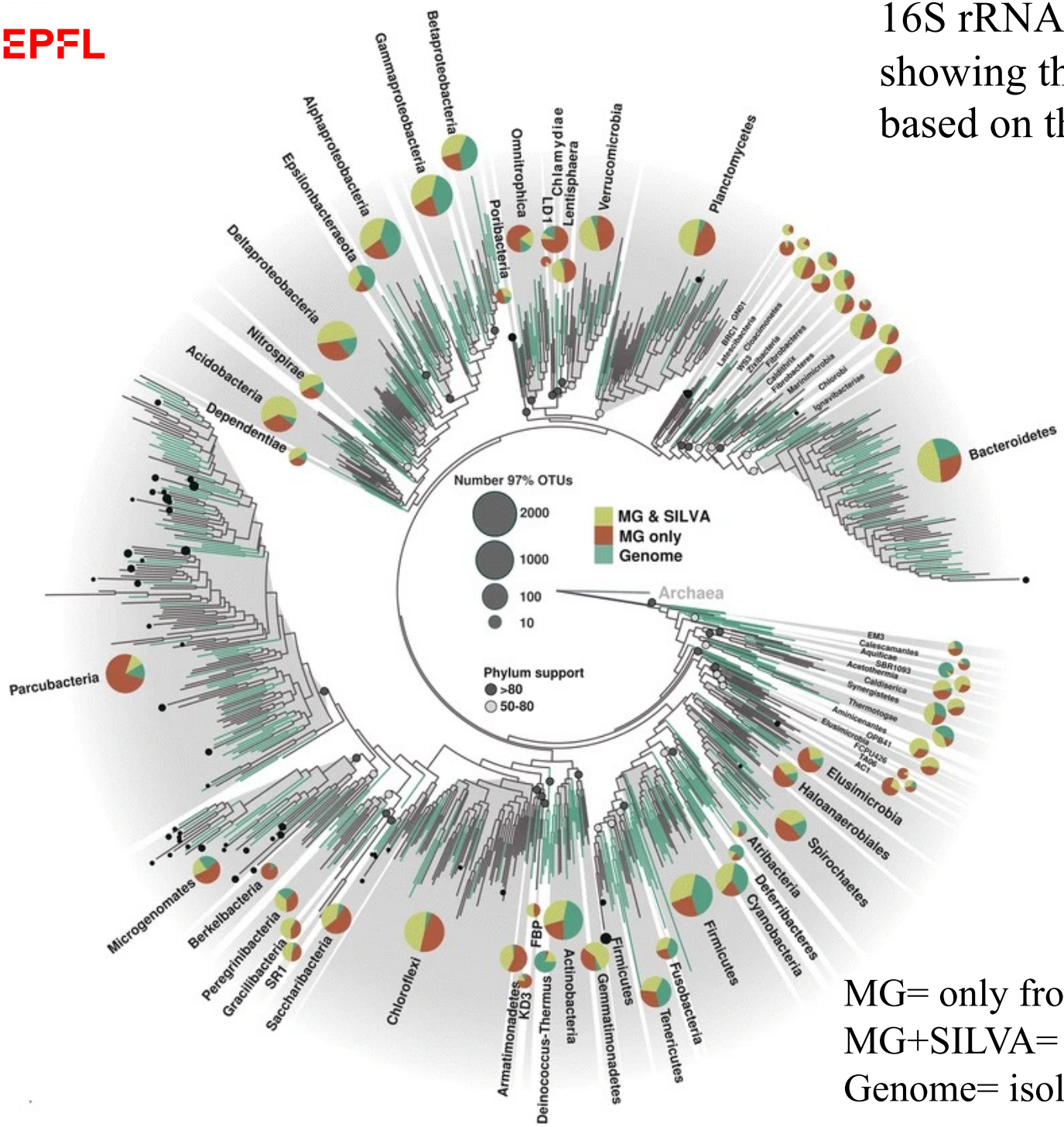
mRNA in white
Protein in orange and green

16S rRNA – variable regions



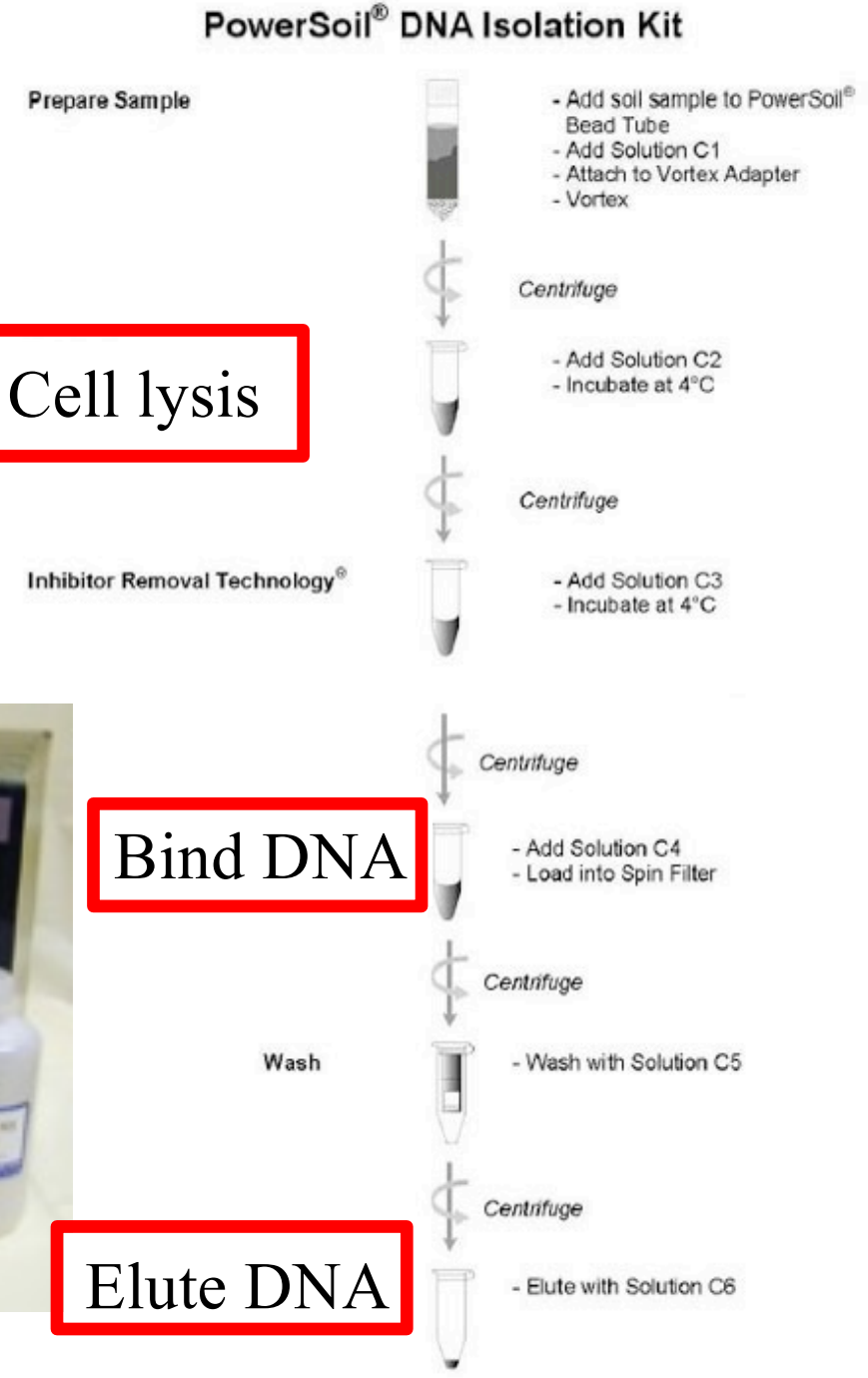
Nature Reviews | Microbiology

16S rRNA-based phylogenetic tree showing the distance between taxa based on this gene.



MG= only from amplicons, no isolates
MG+SILVA= at least one isolate
Genome= isolated and sequenced genome

DNA isolation from environmental sample

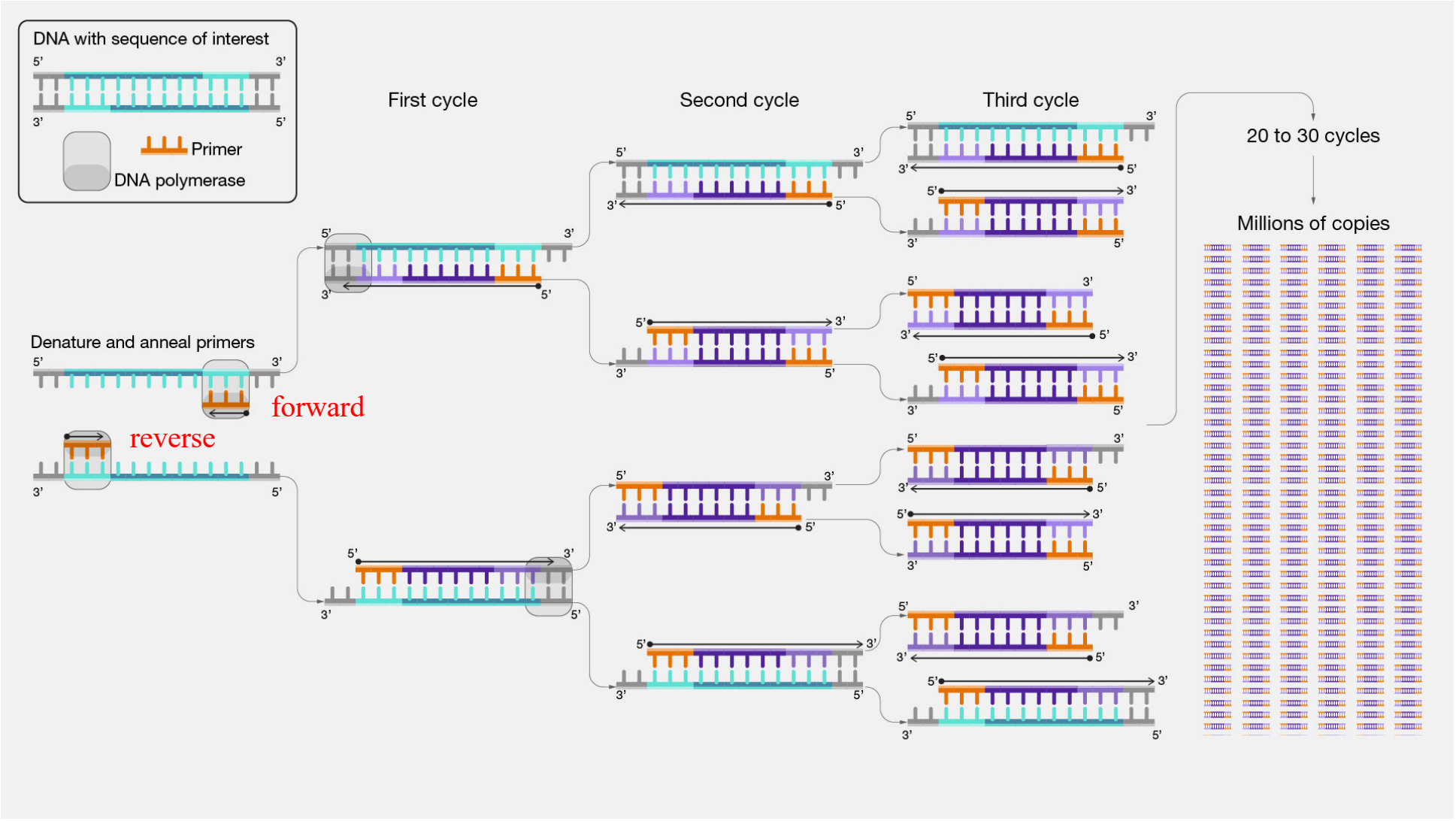


Cell lysis

Bind DNA

Elute DNA

Polymerase-chain reaction (**PCR**)

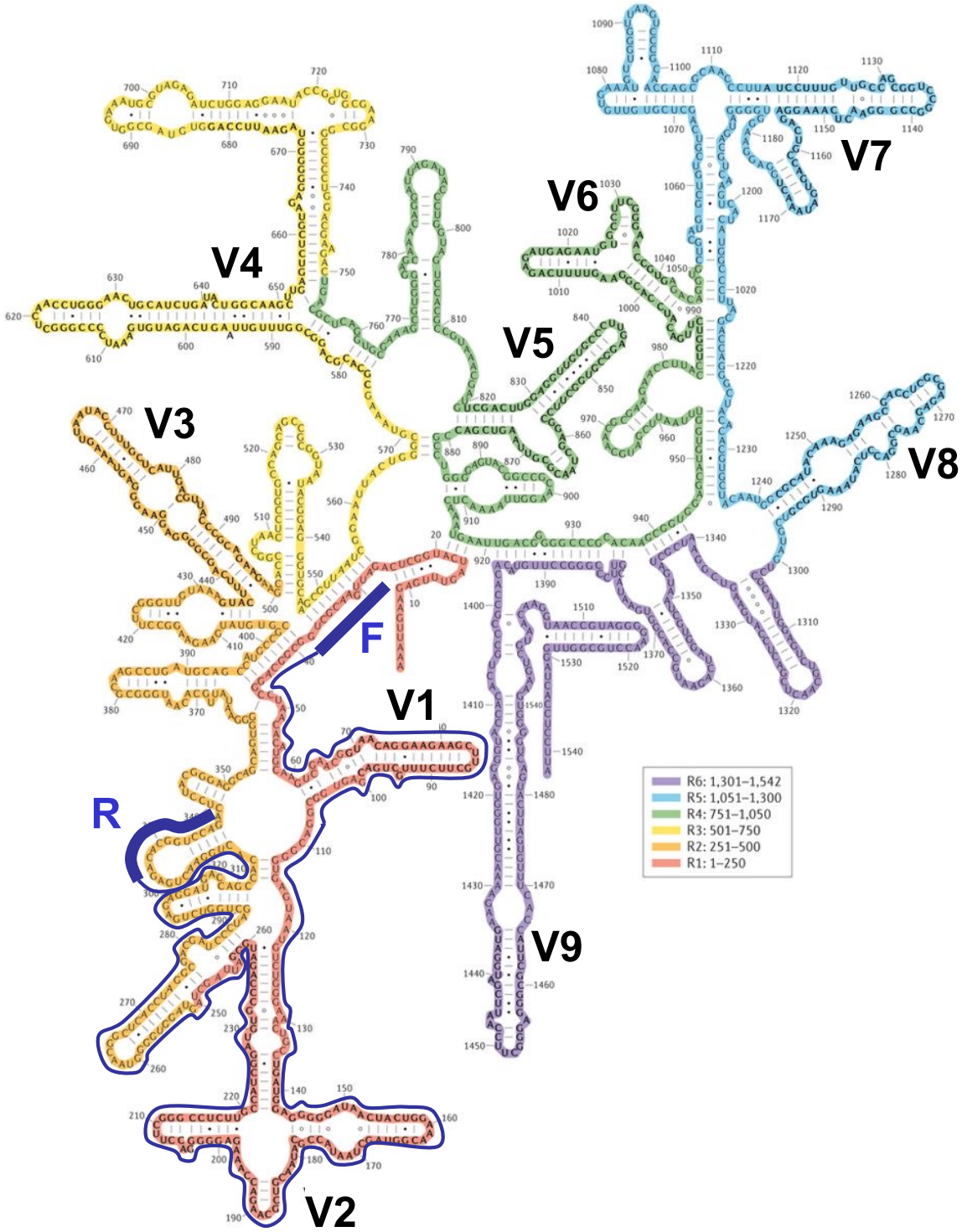


Repetitive amplification of target genes from DNA using specific primers

→ 2^n copies after n cycles

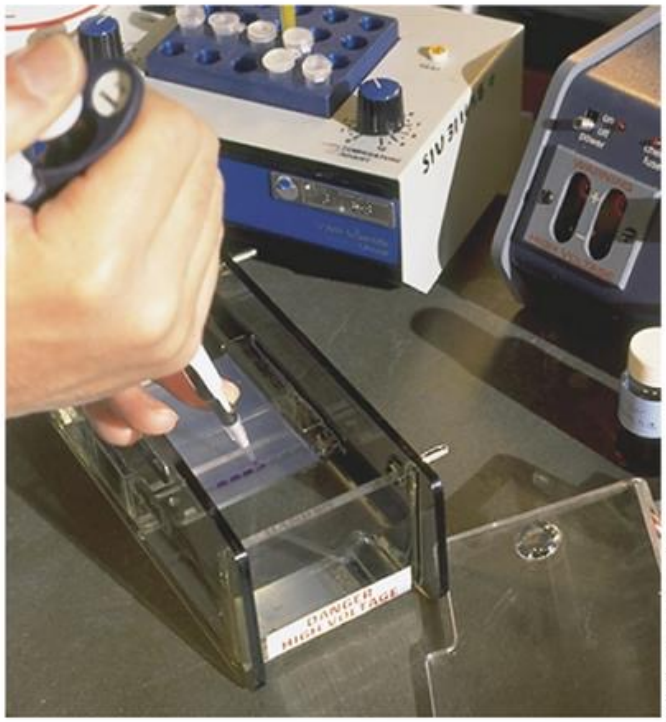
EPFL

16S rRNA – variable regions

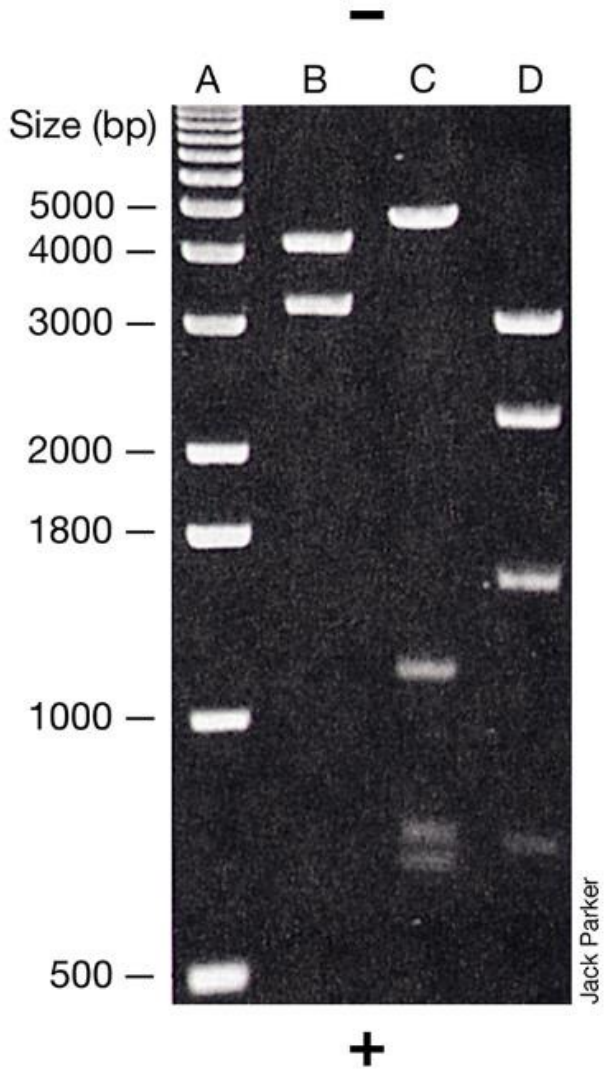


Nature Reviews | Microbiology

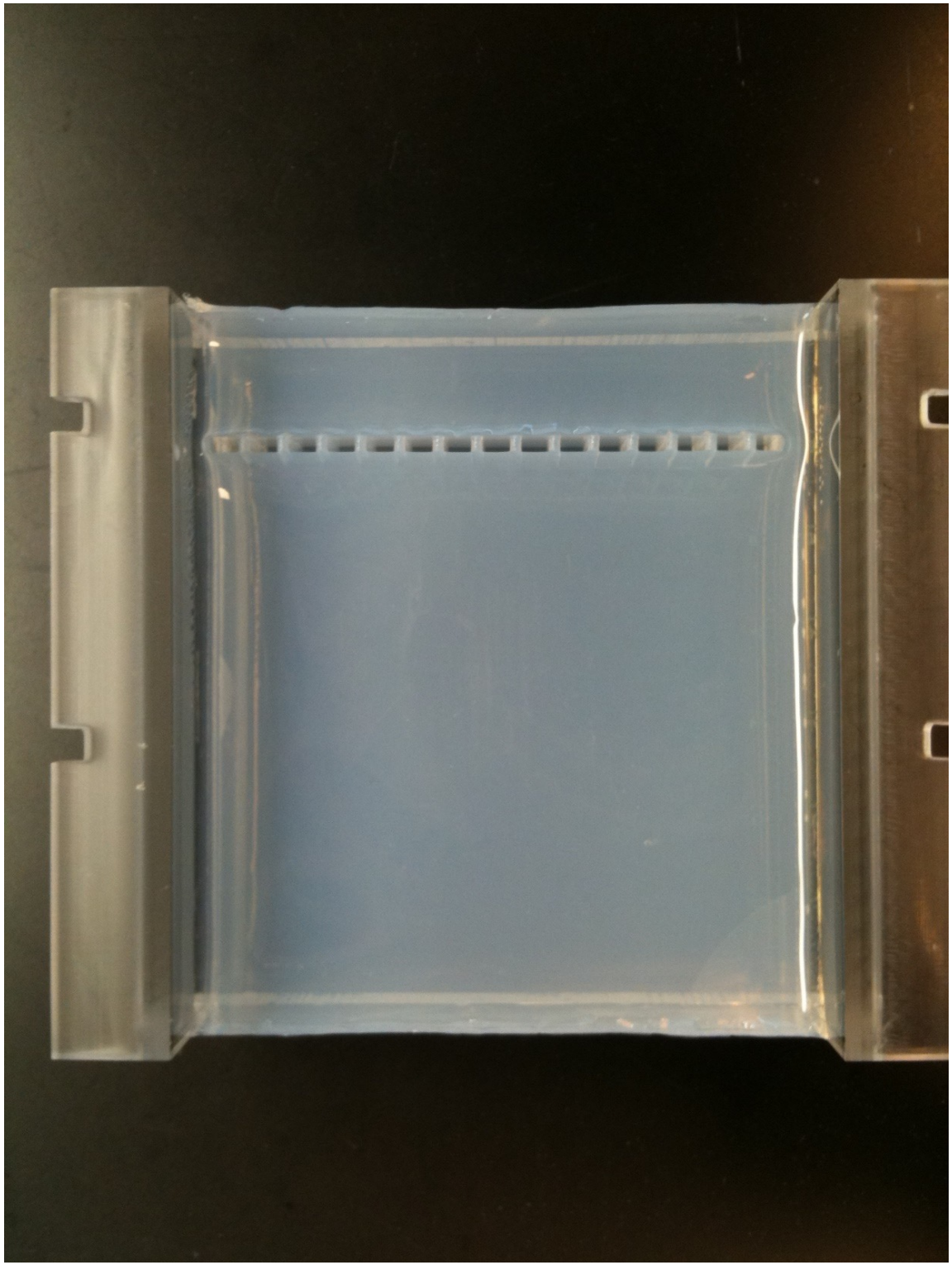
Agarose gel electrophoresis

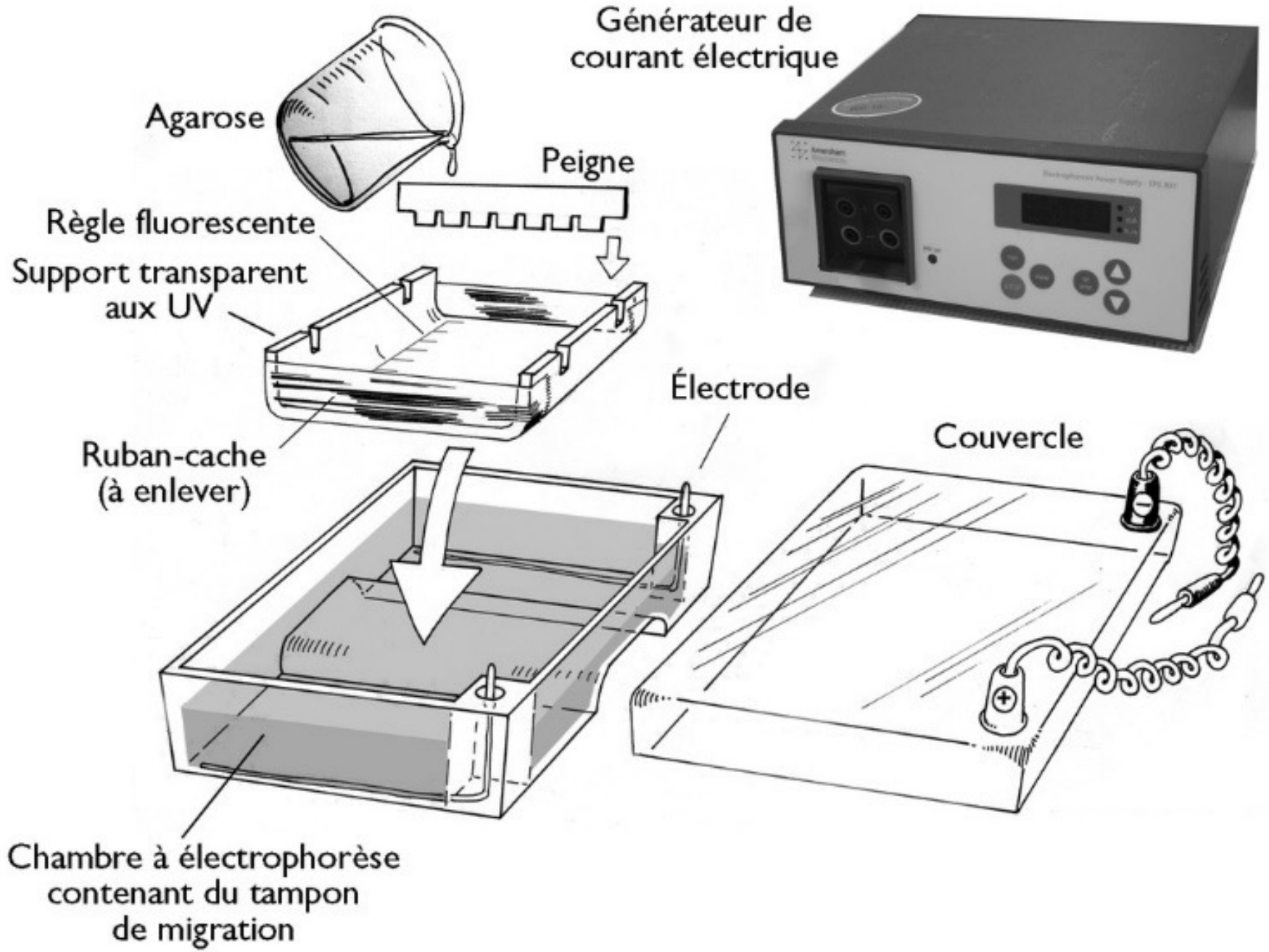


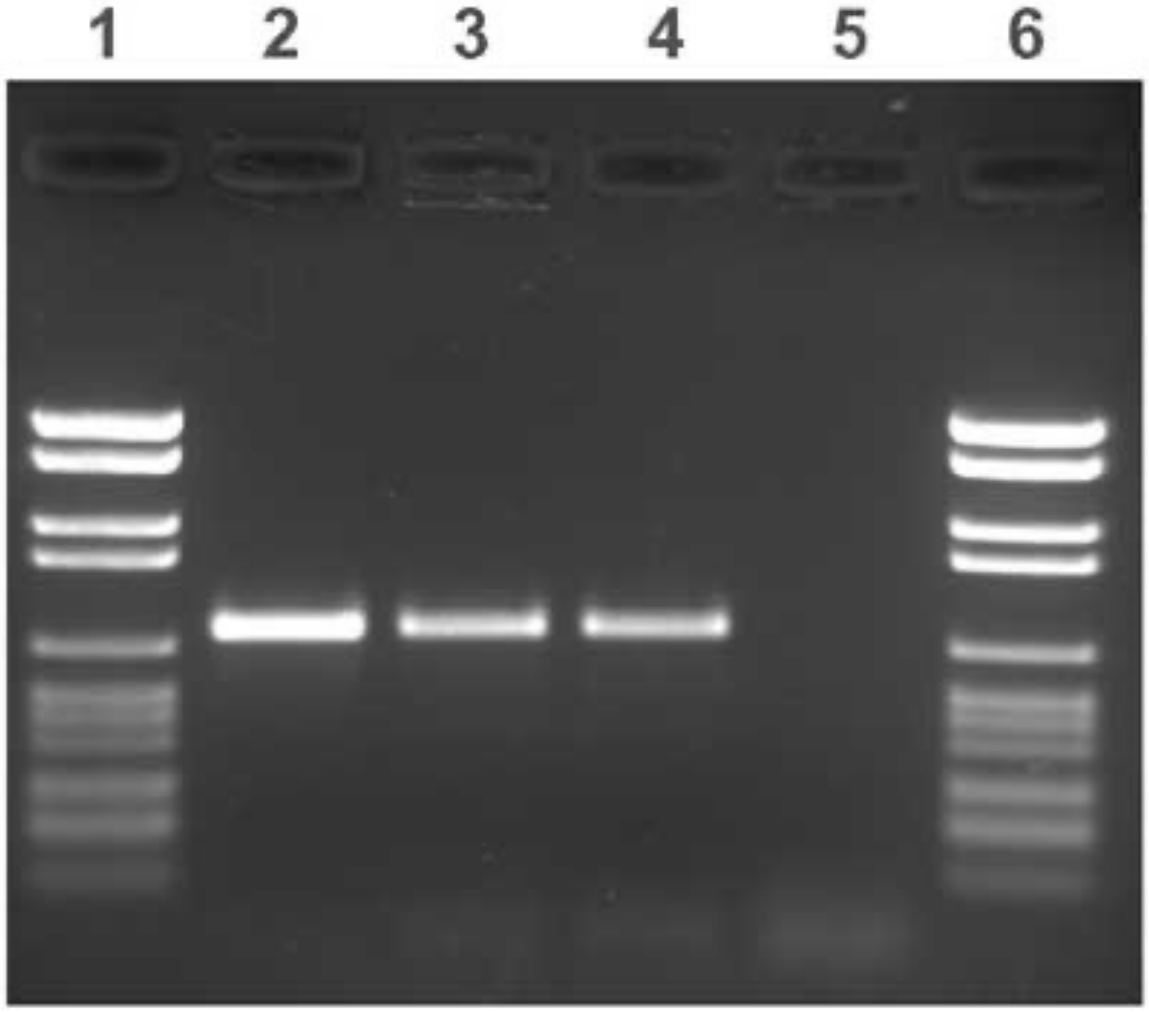
(a)



(b)

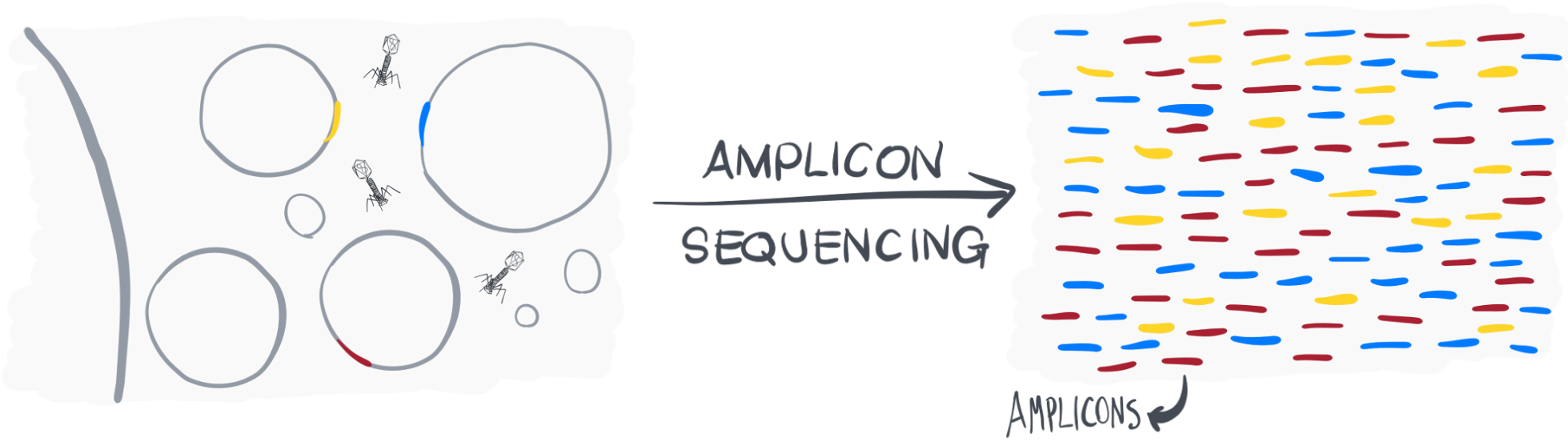



Générateur de
courant électrique
Agarose
Peigne
Règle fluorescente
Support transparent
aux UV
Électrode
Couvercle
Ruban-cache
(à enlever)
Chambre à électrophorèse
contenant du tampon
de migration

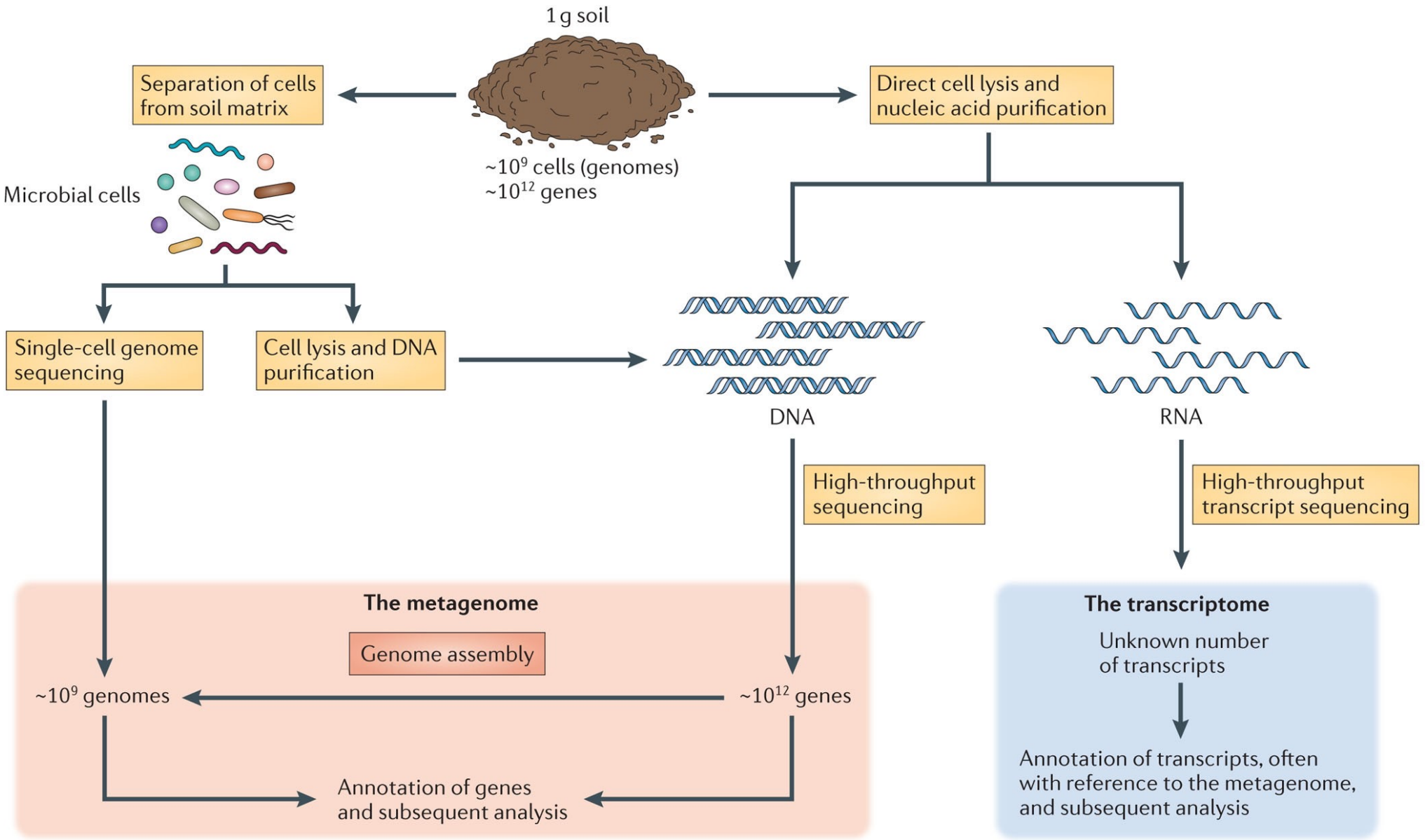

1
2
3
4
5
6

Ribosomal RNA as Evolutionary marker

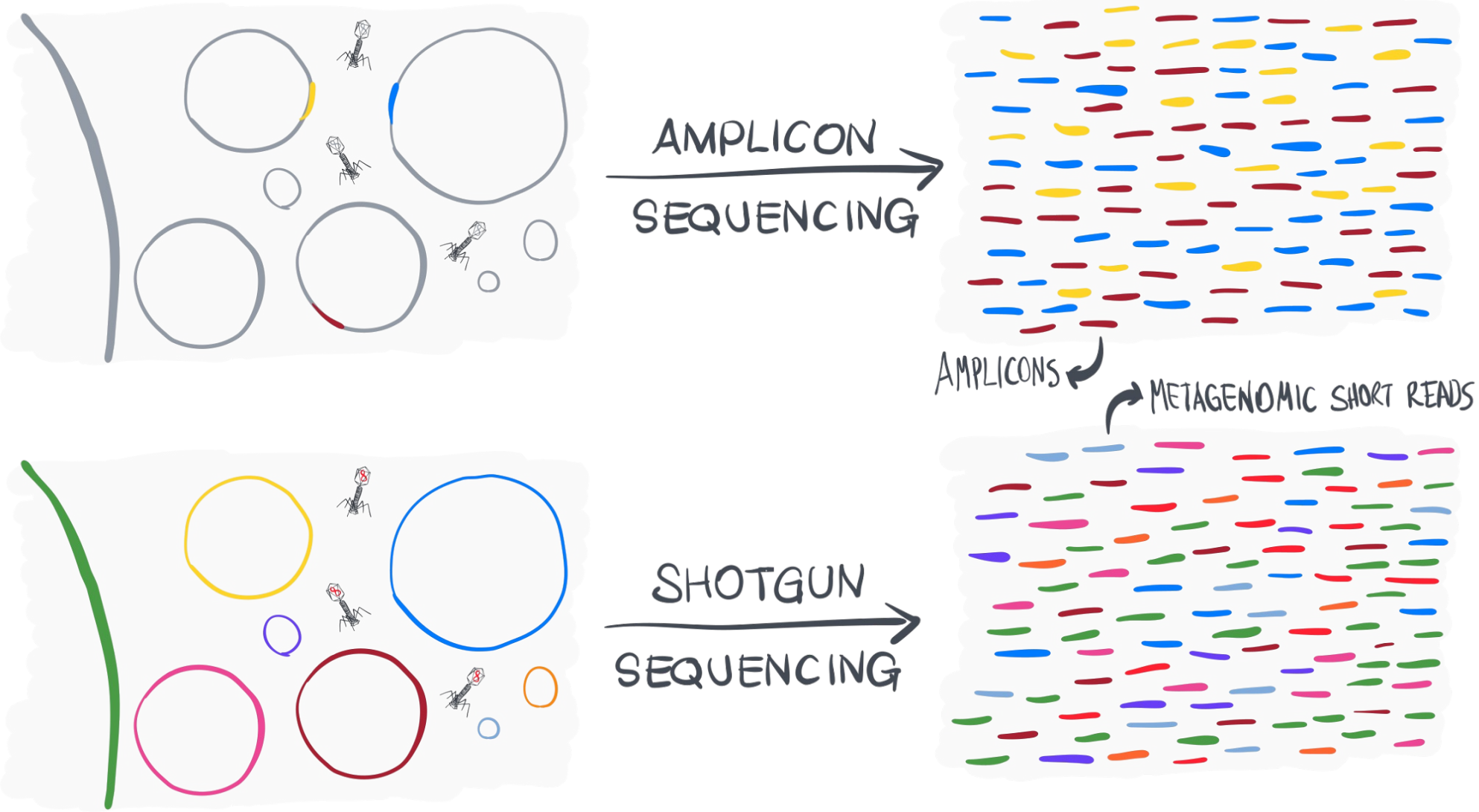
E. coli	**GCGGTAAT** ACGGAGGGTG**CAAGCGTT**AATCGGAATTACTGGGCG **TAAAGCG**
A. nidulans	**GCGGTAAT** ACGGGAGAGG**CAAGCGTT**ATCCGGAATTATTGGGCG **TAAAGCG**
T. maratima	**GCGGTAAT** ACGTAGGGGG**CAAGCGTT**ACCCGGATTTACTGGGCG **TAAAGGG**
M. vannielii	**GCGGTAAT** ACCGACGGCC**CGAGTGGT**AGCCACTCTTATTGGGCC **TAAAGCG**
T. celer	**GCGGTAAT** ACCGGCGGCC**CGAGTGGT**GGCCGCTATTATTGGGCC **TAAAGCG**
S. sulfotaricus	**GCGGTAAT** ACCAGCTCCG**CGAGTGGT**CGGGGTGATTACTGGGCC **TAAGCGG**


AMPLICON
SEQUENCING
AMPLICONS

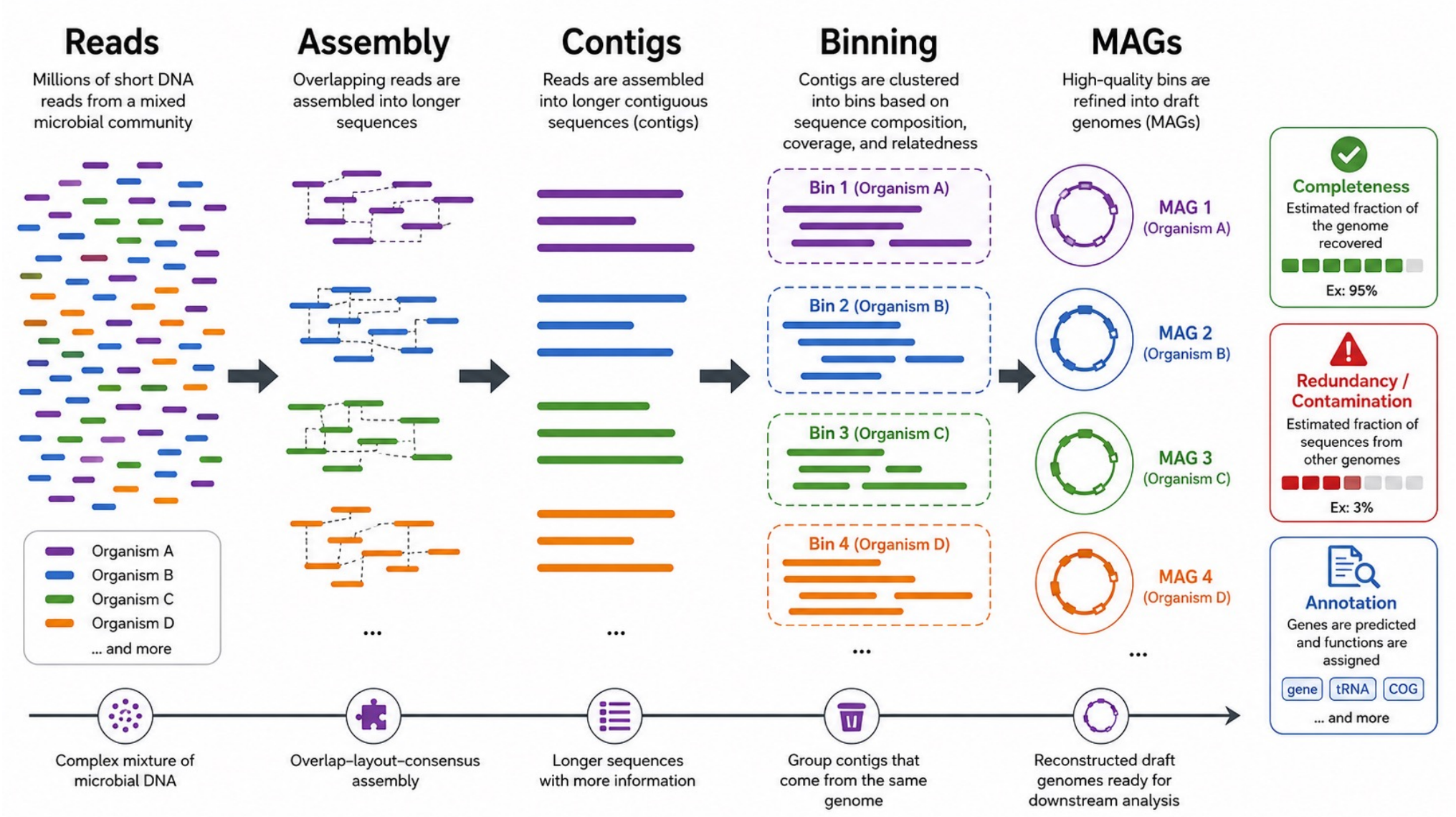
Microbial characterization



Nature Reviews | Microbiology


AMPLICON
SEQUENCING
AMPLICONS
METAGENOMIC SHORT READS
SHOTGUN
SEQUENCING

Recap


Reads
Millions of short DNA reads from a mixed microbial community
Assembly
Overlapping reads are assembled into longer sequences
Contigs
Reads are assembled into longer contiguous sequences (contigs)
Binning
Contigs are clustered into bins based on sequence composition, coverage, and relatedness
MAGs
High-quality bins are refined into draft genomes (MAGs)
Organism A
Organism B
Organism C
Organism D
... and more
Bin 1 (Organism A)
Bin 2 (Organism B)
Bin 3 (Organism C)
Bin 4 (Organism D)
MAG 1 (Organism A)
MAG 2 (Organism B)
MAG 3 (Organism C)
MAG 4 (Organism D)
Complex mixture of microbial DNA
Overlap–layout–consensus assembly
Longer sequences with more information
Group contigs that come from the same genome
Reconstructed draft genomes ready for downstream analysis
Completeness
Estimated fraction of the genome recovered
Ex: 95%
Redundancy / Contamination
Estimated fraction of sequences from other genomes
Ex: 3%
Annotation
Genes are predicted and functions are assigned
gene
tRNA
COG
... and more