Using Molecular Techniques and Next-Generation Sequencing (NGS) to Understand and Optimize Wastewater Treatment Processes

Medini Annavajhala, PhD Student

Chandran Laboratory, Columbia University

mka2136@columbia.edu

July 22, 2015

Outline

I. Traditional microscopy-focused techniques

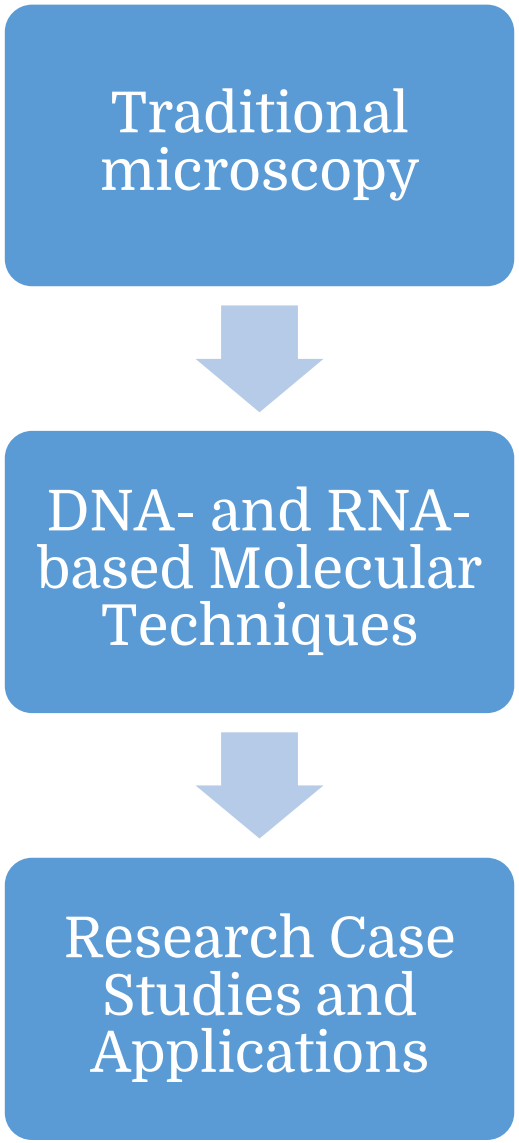
- a. Light, Fluorescence, Confocal Microscopy
- b. FISH

II. Current & Up-and-Coming DNA-/RNA-Based Molecular Techniques

- a. Stable Isotope Probing (SIP)
- b. Polymerase Chain Reaction (PCR)
  - Quantitative Real-Time PCR (qPCR/qRT-PCR)
  - Reverse Transcription PCR (RT-PCR)
- c. Next-Generation Sequencing and Bioinformatics
  - Genomics
  - Transcriptomics

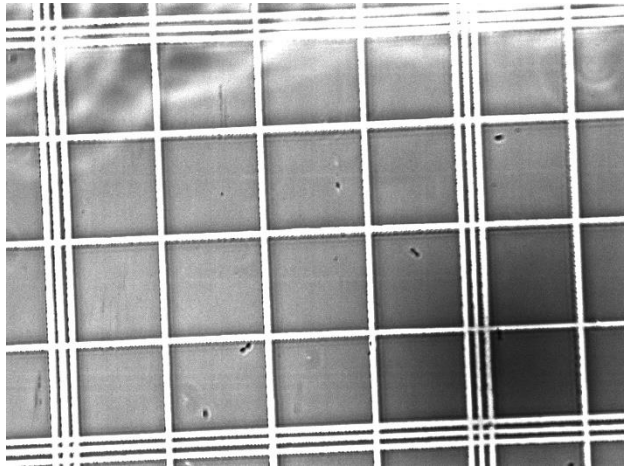
III. Research Case Studies

- a. POPs
- b. Biofuels & Alternate Endpoints
- c. GHG Emissions
- d. Annamox
- e. Full-scale WWTP monitoring

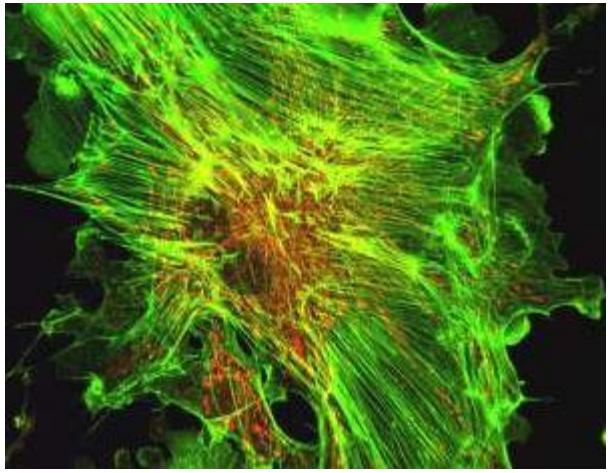


Microscopy

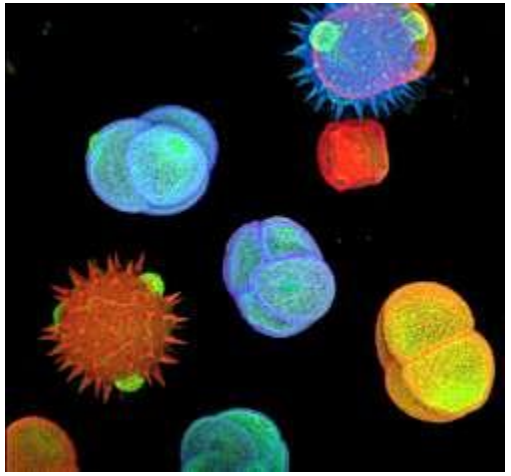
Microscopy



M. Annavajhala June 13, 2013



© Basic Science Partnership, Harvard Medical School



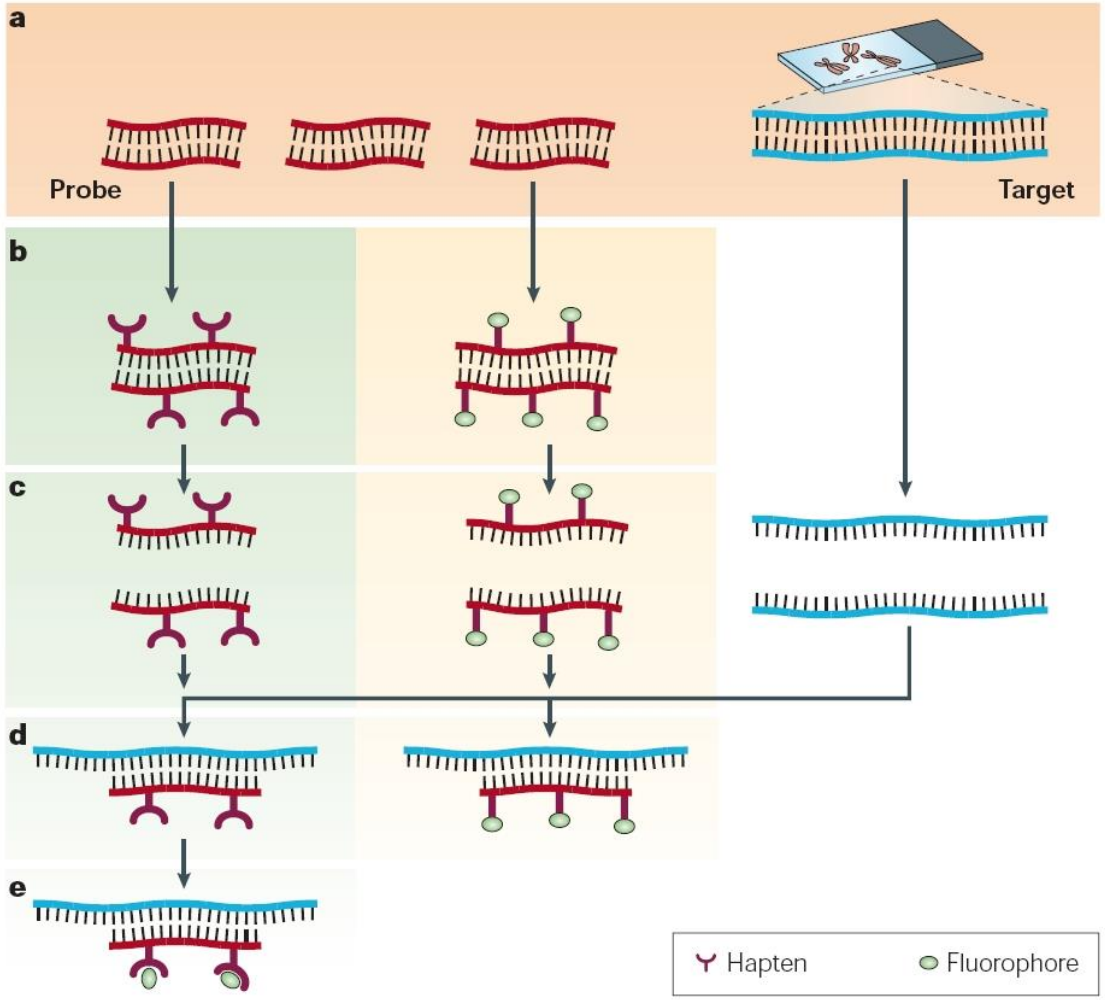
© Basic Science Partnership, Harvard Medical School

Fluorescence In-situ Hybridization (FISH)

Fluorescent probes are designed to bind to specific DNA- or RNA-sequences in the target organisms, which then can be visualized via fluorescence microscopy

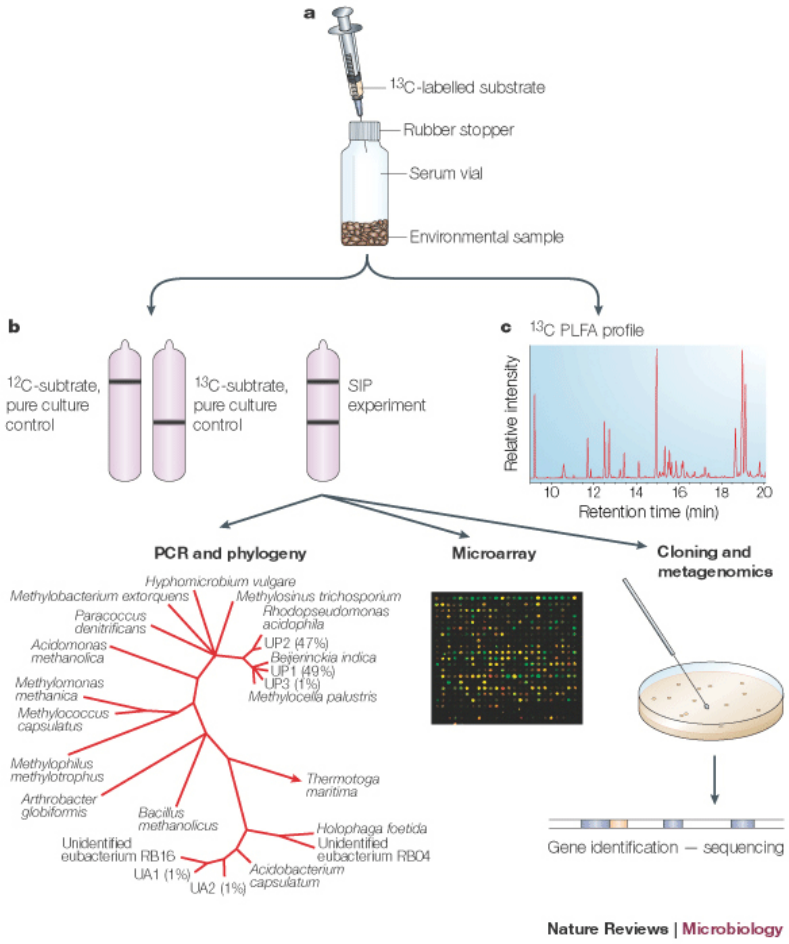
Applications

- Detecting pathogenic bacteria
- Visualization of specific species in mixed cultures
- Localize DNA and RNA targets in a variety of sample types



Current Practices:
Stable Isotope Probing
PCR Basics & Tools
Next-Gen Sequencing

Stable Isotope Probing (SIP)



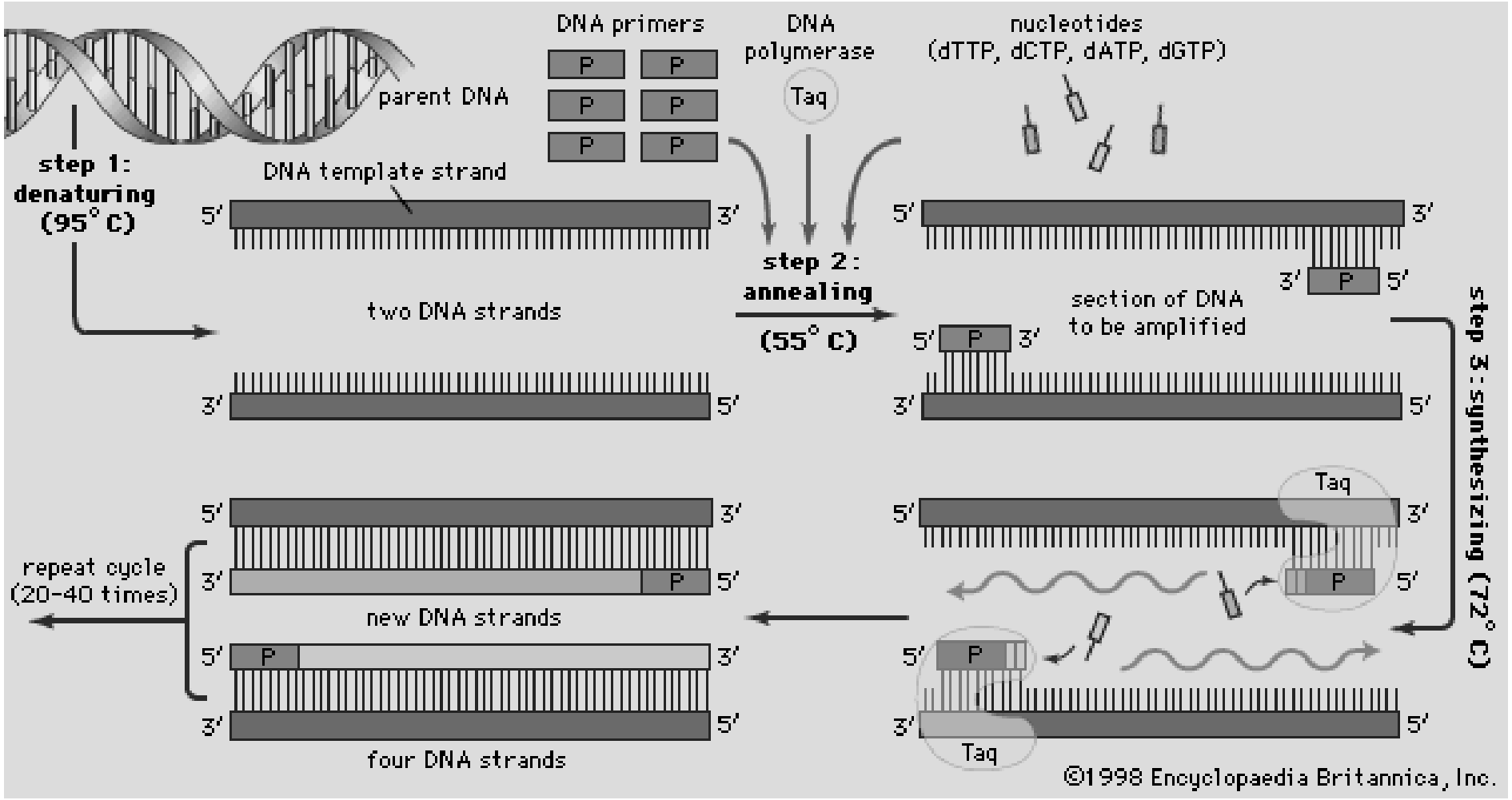
Dumont, MG; Murrell, JC (June 2005). Stable Isotope Probing- linking microbial identity to function. *Nature Reviews Microbiology* **3**, 499-504.

Compounds containing stable isotopes are used to treat an environmental sample; the ability or inability to incorporate the isotope into DNA, RNA, or PLFAs indicates an organism's ability to perform a specific reaction.

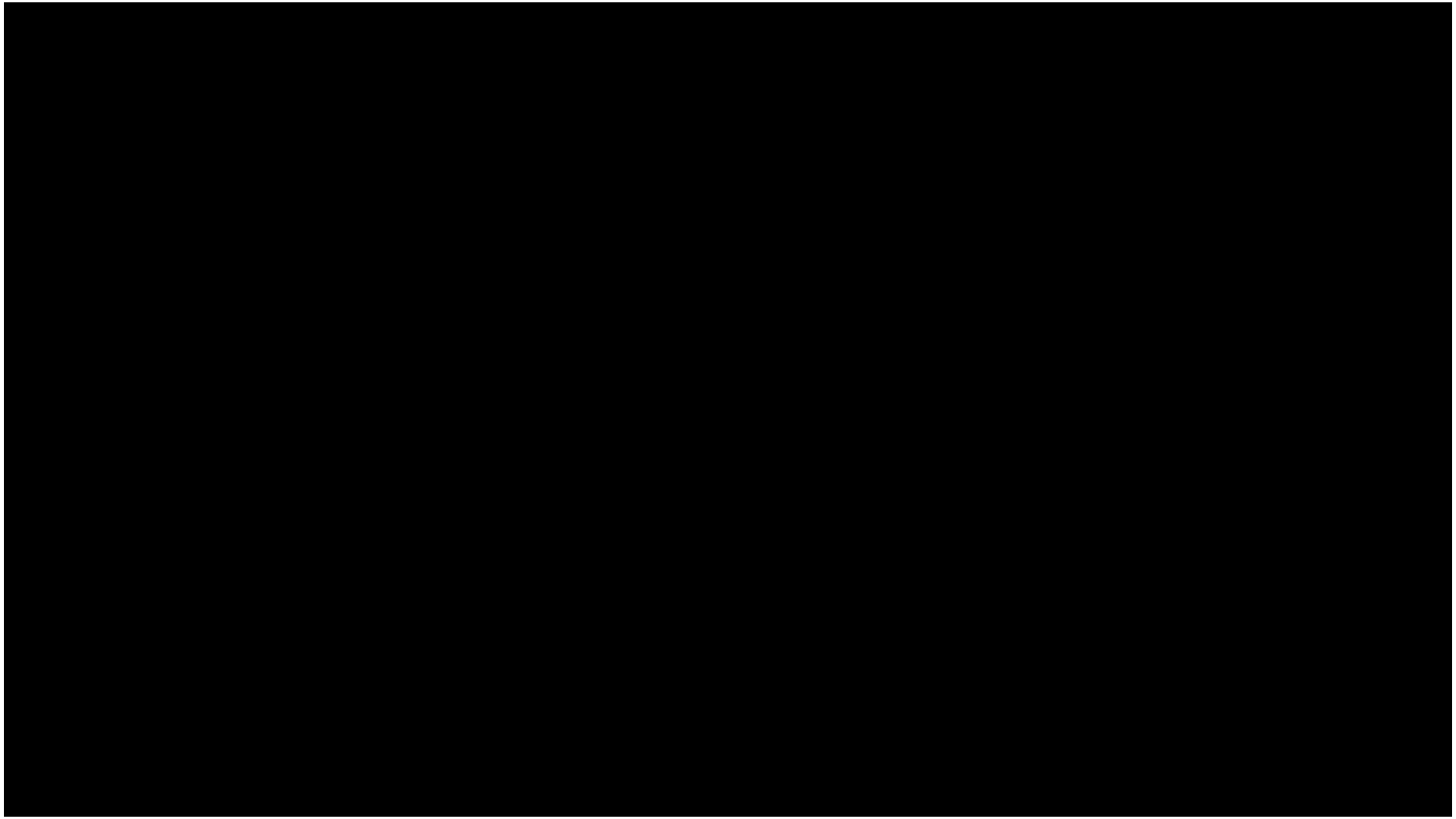
Applications

- Determine which species are involved in a specific reaction pathway
- Assess whether the organisms in your sample are capable of/actively performing a specific reaction

Polymerase Chain Reaction (PCR)



Polymerase Chain Reaction (PCR)



Polymerase Chain Reaction (PCR)

Things to keep in mind:

-Primers can be designed and optimized for specific goals

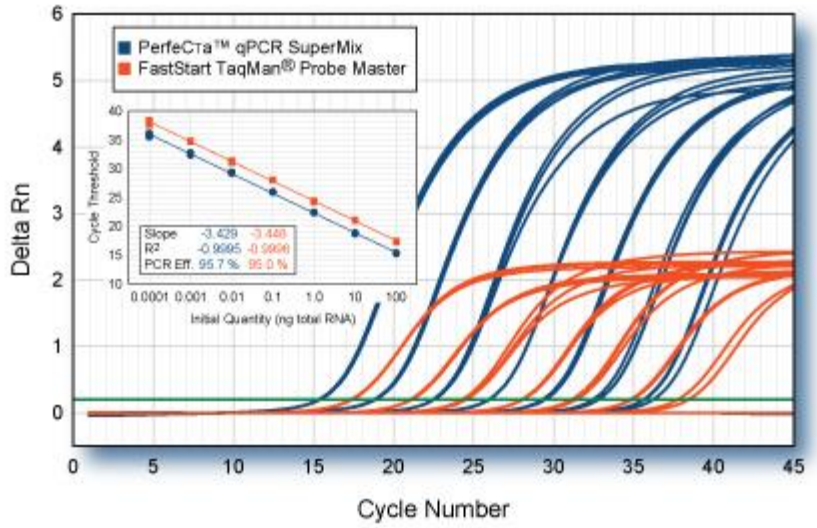
-Multiplex PCR can be used to amplify several regions at once

-Relatively low-cost

-Intermediate skill level and optimization of reaction cycles required

Quantitative PCR (qPCR)

By using a set of standards of known concentration, qPCR can help quantify the amount of starting template material



Quanta Biosciences

Applications

- Quantifying relative abundance of microbial species in a mixed community
- Quantifying relative abundance of genes of interest across samples

Reverse-Transcription PCR (RT-PCR)

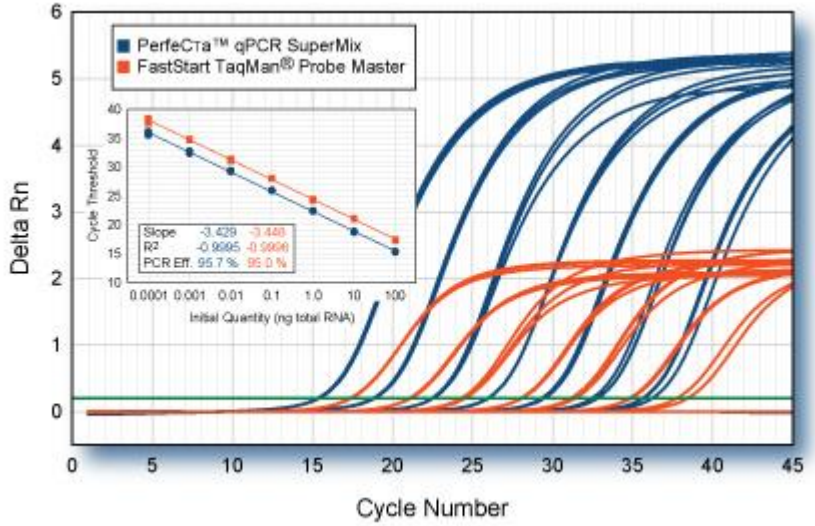
By first converting RNA samples into cDNA, even RNA can be analyzed/quantified with PCR

RNA

↓ Reverse Transcription

Complementary DNA (cDNA)

↓ qPCR/PCR



Applications

- Determining presence/absence of specific genes at the RNA level
- Differential gene expression across samples (assess your community's activity in different conditions)

Bioinformatics

Using NGS DNA-, RNA-, and protein-level information to connect microbial genetics and expression to wastewater treatment processes

Which microorganisms are present in the sample?

16S Amplicon Sequencing (Metagenomics)

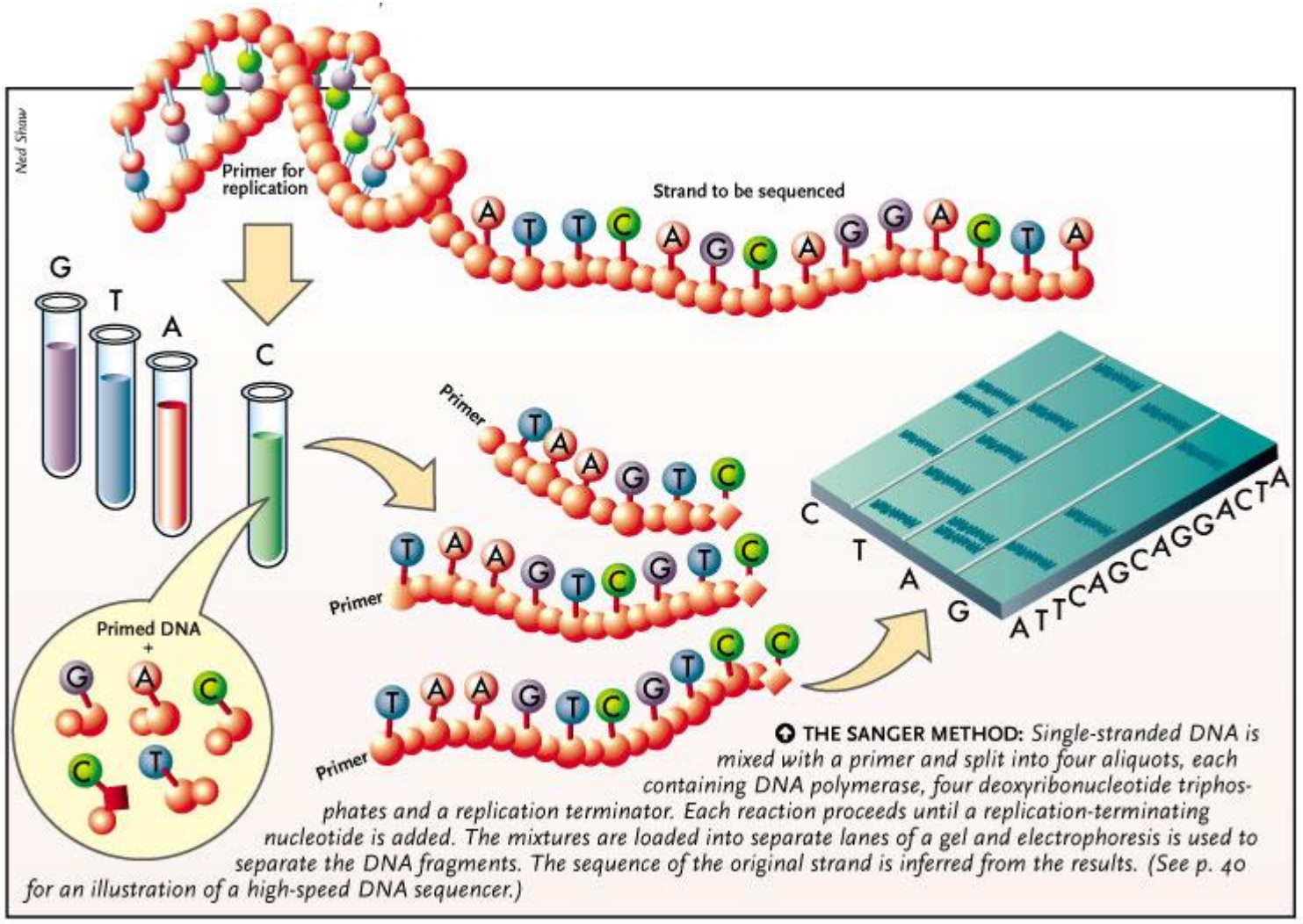
Which metabolic & functional pathways are active in the sample?

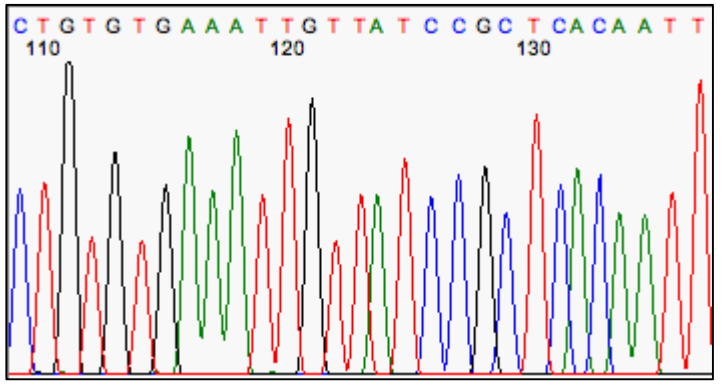
Shotgun DNA Sequencing (Metagenomics), RNA-Seq (Metatranscriptomics), Metaproteomics

How are genes being differentially expressed or gene products being differentially produced across samples?

RNA-Seq (Metatranscriptomics), Metaproteomics

Traditional Sanger Sequencing

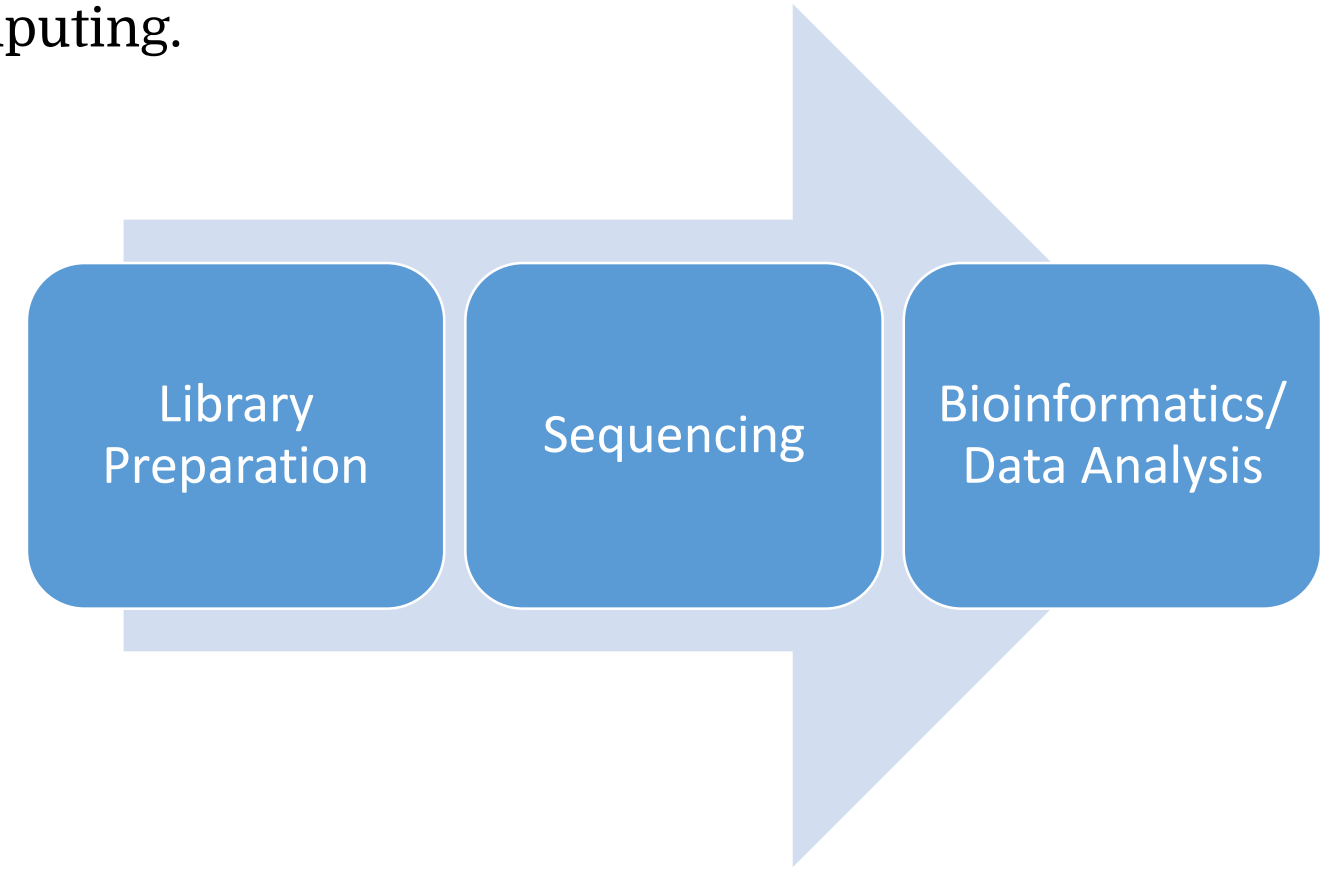




The University of Michigan,
DNA Sequencing Core

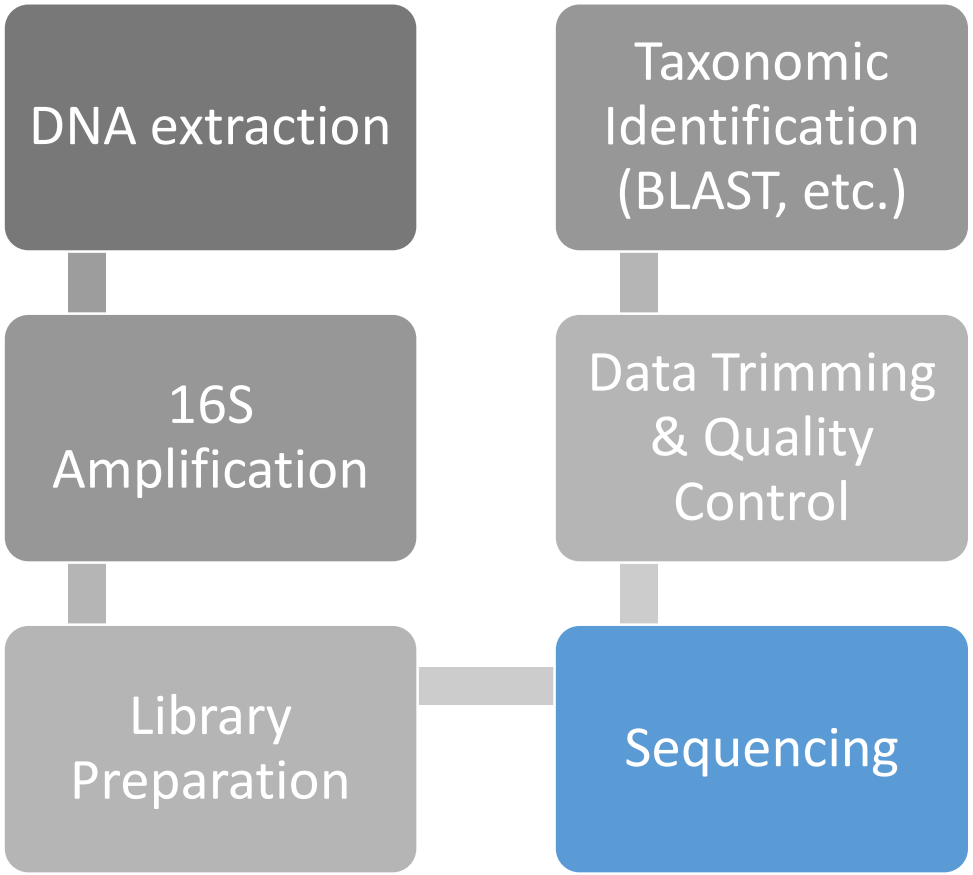
Next-Generation Sequencing (NGS)

Sequencing refers to the process of determining the order (sequence) of base pairs (A,T,C,G) which encode genetic information. Next-Generation Sequencing uses semiconductor chips to convert the biochemical information held in a DNA or cDNA sequence to electronic data which can be processed and analyzed with high-throughput computing.

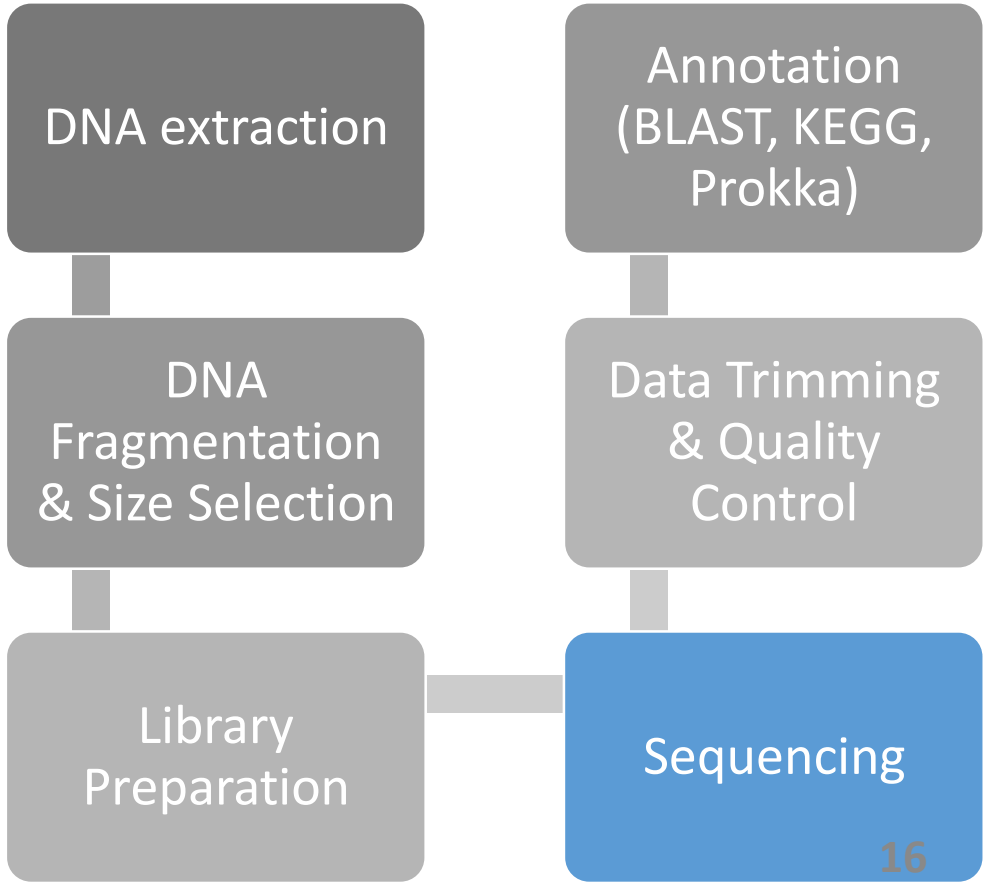


(Meta)genomics

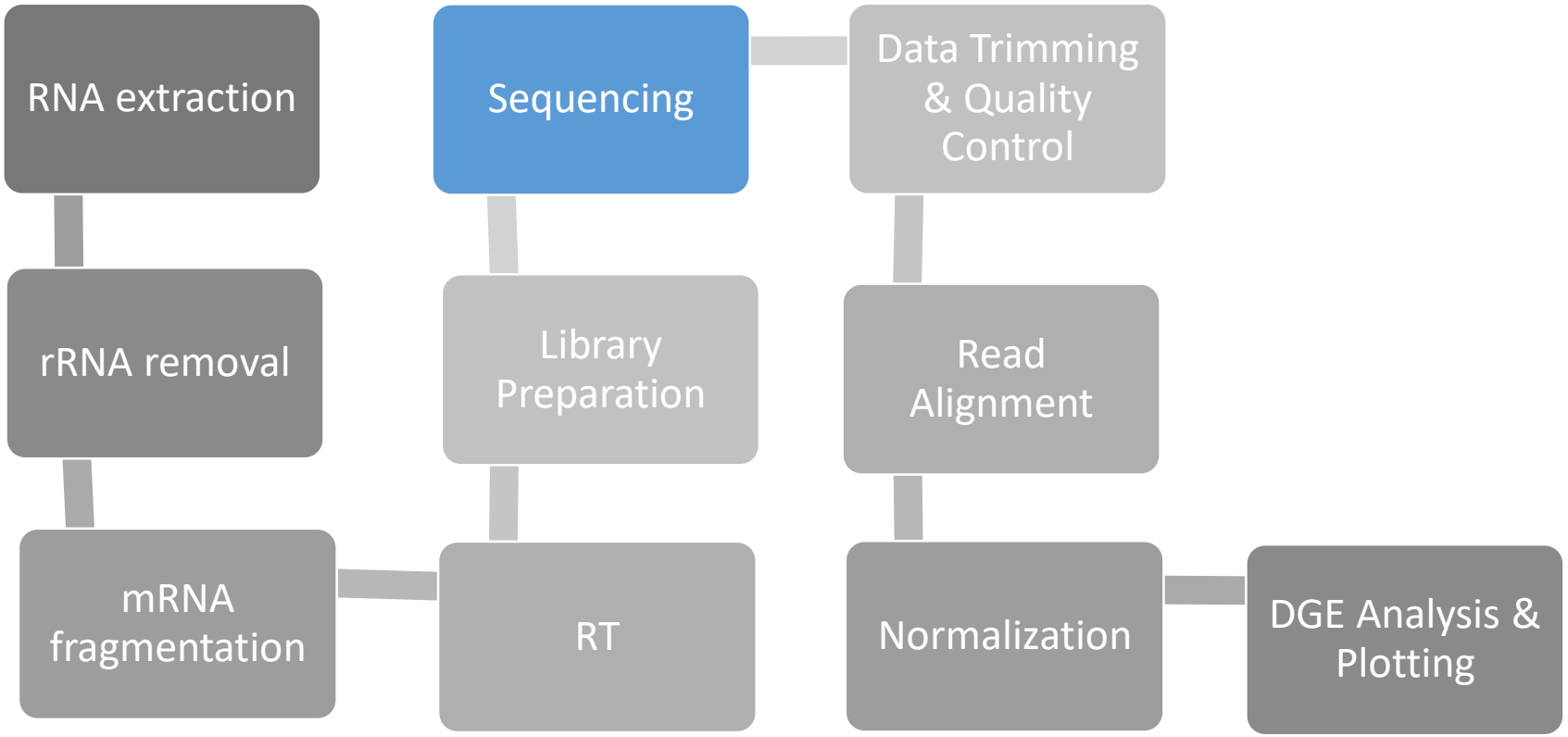
Identifying the Community (16S Amplicon Sequencing)

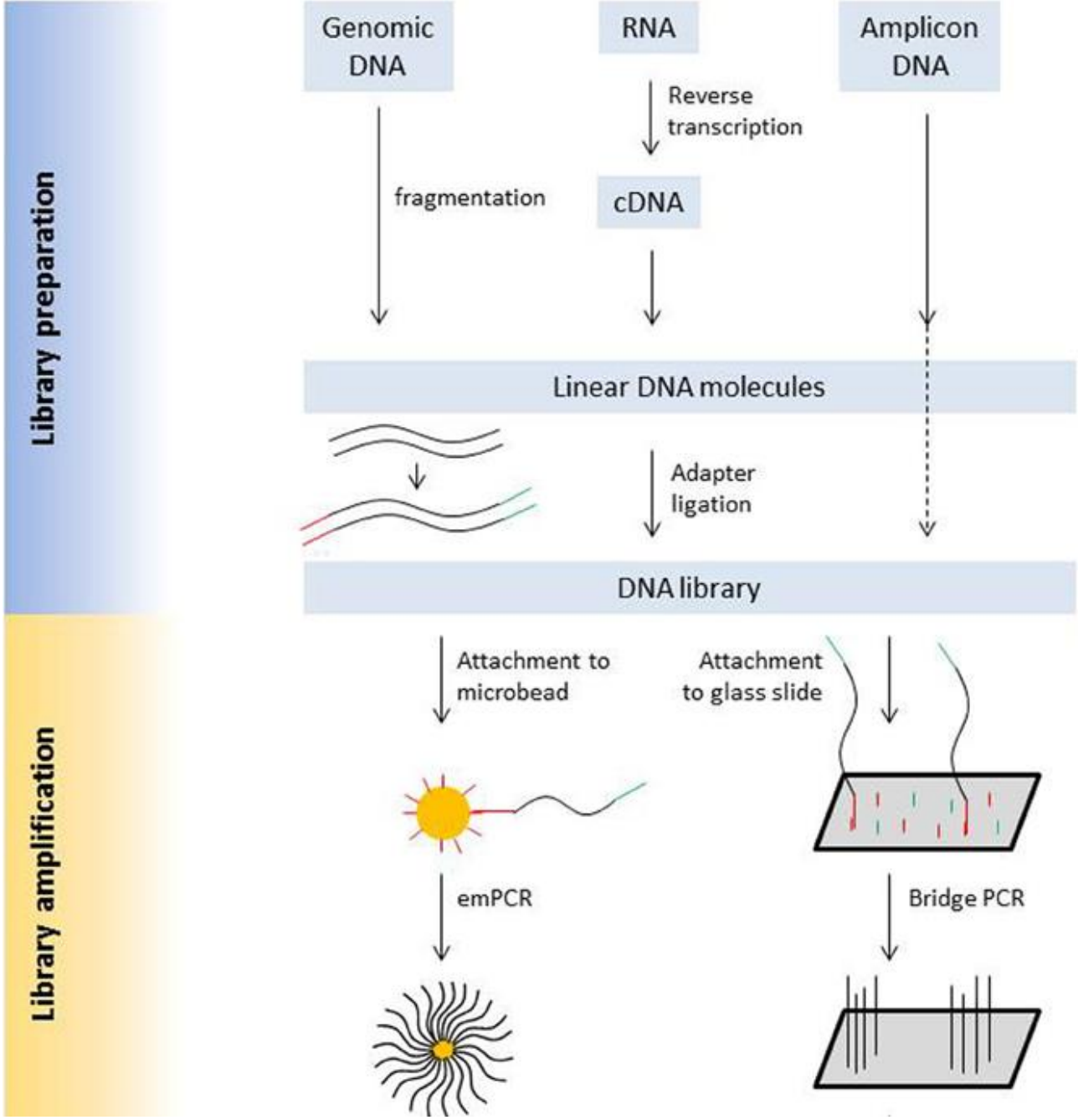


Identifying Functional Pathways (Shotgun Sequencing)



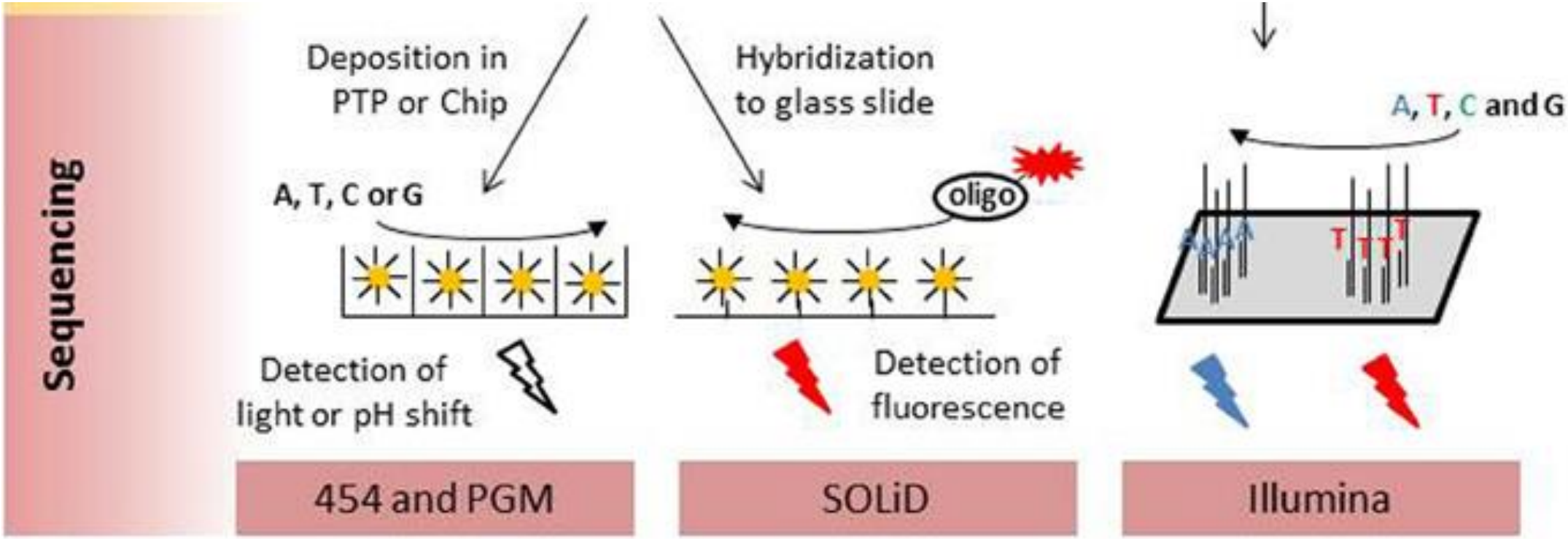
(Meta)transcriptomics (RNA-Seq)





Knief, Claudia (2014). Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front. Plant Sci.* **5**

NGS Mechanism



Knief, Claudia (2014). Analysis of plant microbe interactions in the era of next generation sequencing technologies. *Front. Plant Sci.* **5**

Bioinformatics

Using NGS DNA-, RNA-, and protein-level information to connect microbial genetics and expression to wastewater treatment processes

Which microorganisms are present in the sample?

16S Amplicon Sequencing (Metagenomics)

Which metabolic & functional pathways are active in the sample?

Shotgun DNA Sequencing (Metagenomics), RNA-Seq (Metatranscriptomics), Metaproteomics

How are genes being differentially expressed or gene products being differentially produced across samples?

RNA-Seq (Metatranscriptomics), Metaproteomics

Bioinformatics

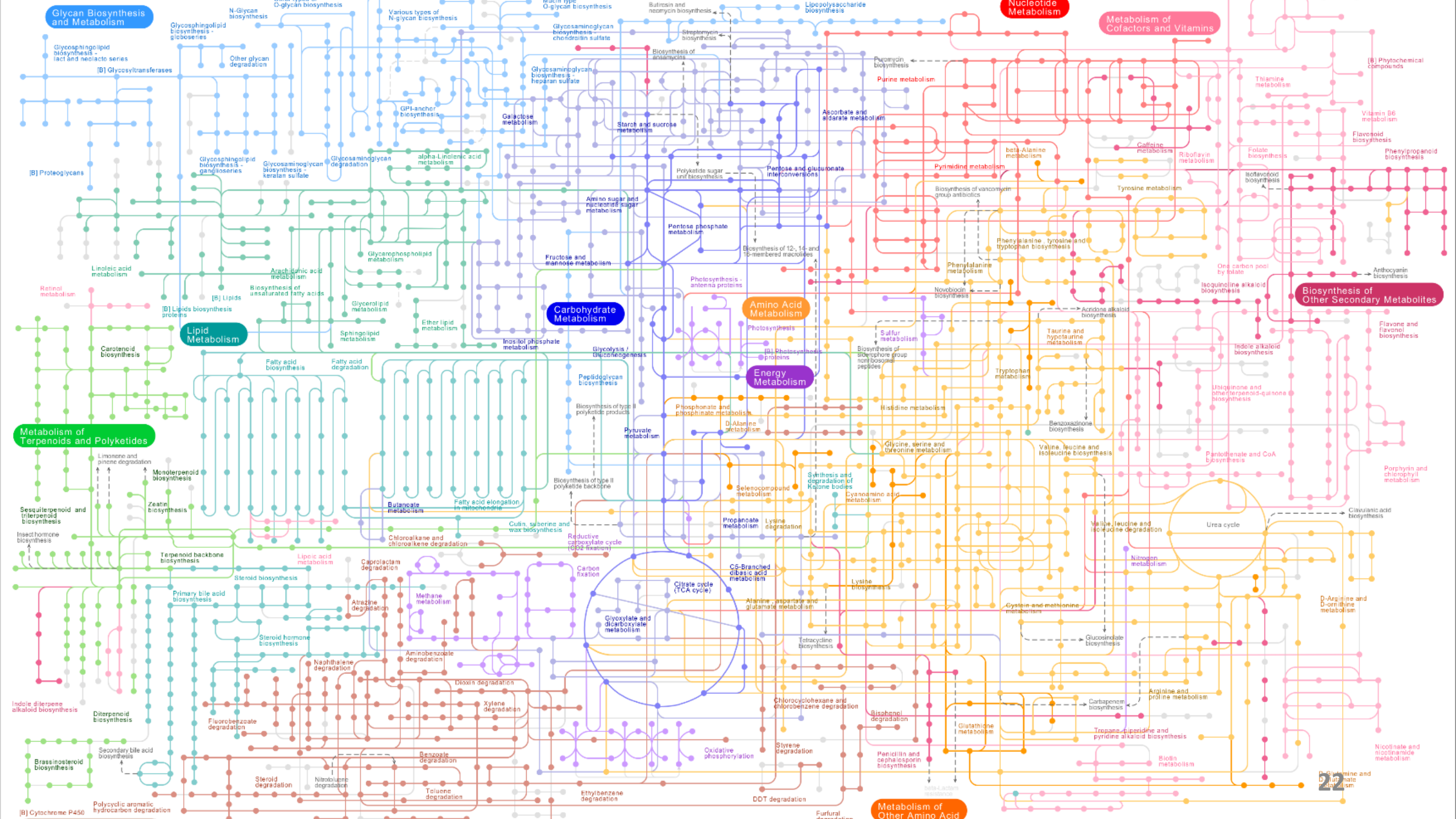
Things to keep in mind:

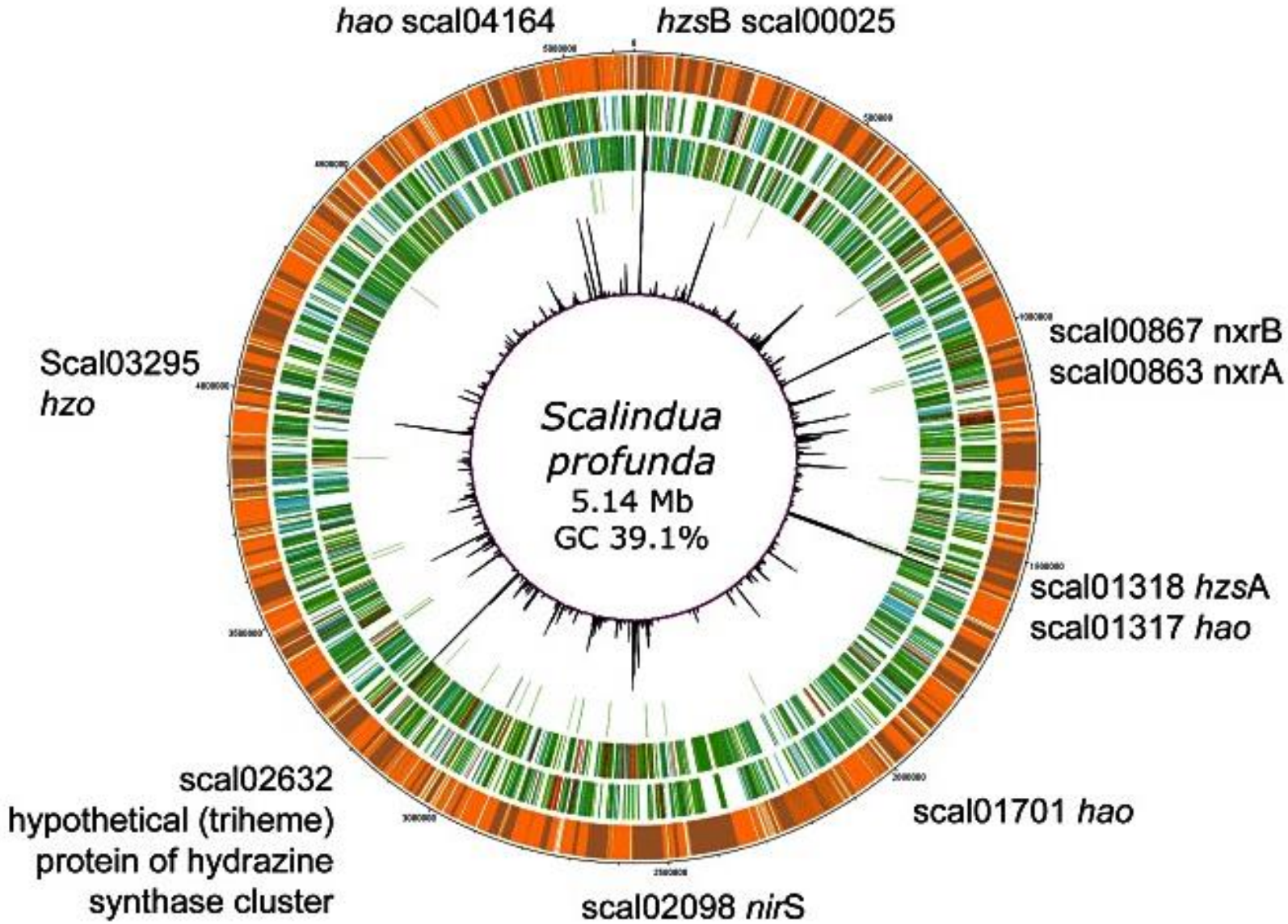
-Still fairly costly

-Difficult techniques → most send samples to a core facility rather than sequencing on-site

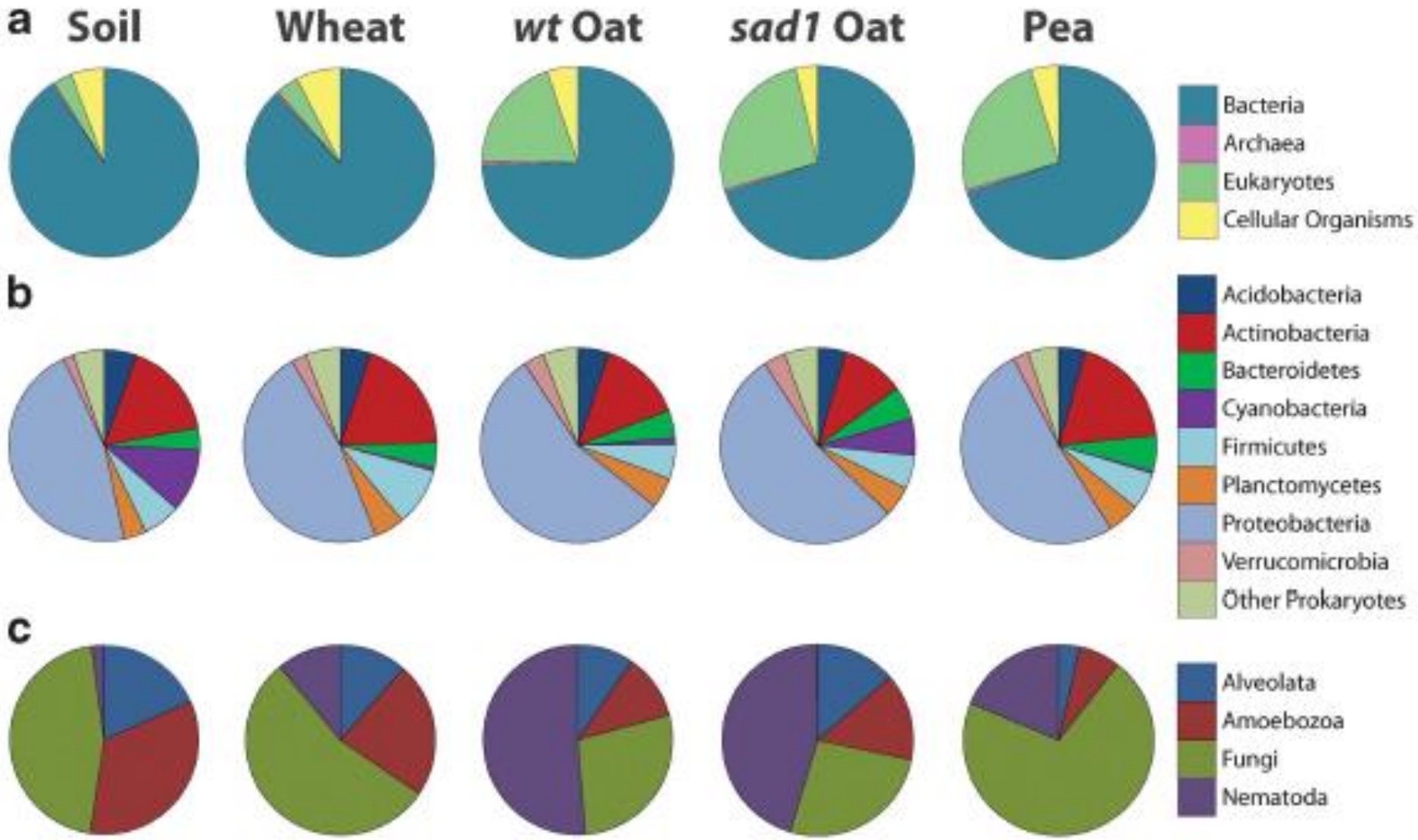
-Huge amounts of data to sift through and analyze → know what final information you want BEFORE you sequence

-Many different software packages


Glycan Biosynthesis and Metabolism
Carbohydrate Metabolism
Lipid Metabolism
Amino Acid Metabolism
Energy Metabolism
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Biosynthesis of Other Secondary Metabolites
Metabolism of Terpenoids and Polyketides
Metabolism of Other Amino Acid
Glycosphingolipid biosynthesis - lacto and neolacto series
Glycosphingolipid biosynthesis - globoseries
N-Glycan biosynthesis
Various types of N-glycan biosynthesis
Other glycan degradation
[B] Glycosyltransferases
[B] Proteoglycans
GPI-anchor biosynthesis
Glycosphingolipid biosynthesis - ganglioseries
Glycosaminoglycan biosynthesis - keratan sulfate
Glycosaminoglycan degradation
Glycosaminoglycan biosynthesis - chondroitin sulfate
Glycosaminoglycan biosynthesis - heparan sulfate
Galactose metabolism
Starch and sucrose metabolism
alpha-Linolenic acid metabolism
Amino sugar and nucleotide sugar metabolism
Pentose phosphate metabolism
Fructose and mannose metabolism
Glycerophospholipid metabolism
Linoleic acid metabolism
Arachidonic acid metabolism
Biosynthesis of unsaturated fatty acids
[B] Lipids
[B] Lipids biosynthesis proteins
Glycerolipid metabolism
Ether lipid metabolism
Sphingolipid metabolism
Inositol phosphate metabolism
Retinol metabolism
Carotenoid biosynthesis
Fatty acid biosynthesis
Fatty acid degradation
Peptidoglycan biosynthesis
Glycolysis / Gluconeogenesis
Pyruvate metabolism
Biosynthesis of type II polyketide products
Biosynthesis of type II polyketide backbone
Butanoate metabolism
Fatty acid elongation in mitochondria
Propanoate metabolism
Lysine degradation
Reductive carboxylate cycle (CO2 fixation)
Carbon fixation
Citrate cycle (TCA cycle)
Glyoxylate and dicarboxylate metabolism
C5-Branched dibasic acid metabolism
Alanine, aspartate and glutamate metabolism
Lysine biosynthesis
Cutin, suberine and wax biosynthesis
Chloroalkane and chloroalkene degradation
Caprolactam degradation
Lipoic acid metabolism
Methane metabolism
Atrazine degradation
Aminobenzoate degradation
Naphthalene degradation
Dioxin degradation
Xylene degradation
Fluorobenzoate degradation
Benzoate degradation
Toluene degradation
Nitrotoluene degradation
Steroid degradation
Ethylbenzene degradation
Styrene degradation
Oxidative phosphorylation
Chlorocyclohexane and chlorobenzene degradation
DDT degradation
Tetracycline biosynthesis
Penicillin and cephalosporin biosynthesis
beta-Lactam resistance
Bisphenol degradation
Glutathione metabolism
Carbapenem biosynthesis
Arginine and proline metabolism
Tropane, piperidine and pyridine alkaloid biosynthesis
Biotin metabolism
Nicotinate and nicotinamide metabolism
D-Glutamine and D-glutamate metabolism
D-Arginine and D-ornithine metabolism
Urea cycle
Nitrogen metabolism
Valine, leucine and isoleucine degradation
Valine, leucine and isoleucine biosynthesis
Glycine, serine and threonine metabolism
Cysteine and methionine metabolism
Glucosinolate biosynthesis
Clavulanic acid biosynthesis
Pantothenate and CoA biosynthesis
Porphyrin and chlorophyll metabolism
Histidine metabolism
Benzoxazinone biosynthesis
Tryptophan metabolism
Indole alkaloid biosynthesis
Ubiquinone and other terpenoid-quinone biosynthesis
Taurine and hypotaurine metabolism
Sulfur metabolism
Acridone alkaloid biosynthesis
Isoquinoline alkaloid biosynthesis
Tyrosine metabolism
Phenylalanine metabolism
Phenylalanine, tyrosine and tryptophan biosynthesis
Novobiocin biosynthesis
Biosynthesis of vancomycin group antibiotics
Pyrimidine metabolism
beta-Alanine metabolism
Caffeine metabolism
Riboflavin metabolism
Folate biosynthesis
One carbon pool by folate
Isoflavonoid biosynthesis
Flavone and flavonol biosynthesis
Anthocyanin biosynthesis
Phenylpropanoid biosynthesis
Flavonoid biosynthesis
Vitamin B6 metabolism
Thiamine metabolism
[B] Phytochemical compounds
Ascorbate and aldarate metabolism
Pentose and glucuronate interconversions
Purine metabolism
Puromycin biosynthesis
Streptomycin biosynthesis
Butirosin and neomycin biosynthesis
Lipopolysaccharide biosynthesis
Biosynthesis of ansamycins
Biosynthesis of 12-, 14- and 16-membered macrolides
Photosynthesis - antenna proteins
Photosynthesis
[B] Photosynthetic proteins
Phosphonate and phosphinate metabolism
D-Alanine metabolism
Selenocompound metabolism
Synthesis and degradation of ketone bodies
Cyanoamino acid metabolism
Biosynthesis of siderophore group nonribosomal peptides
Limonene and pinene degradation
Monoterpenoid biosynthesis
Zeatin biosynthesis
Sesquiterpenoid and triterpenoid biosynthesis
Insect hormone biosynthesis
Terpenoid backbone biosynthesis
Steroid biosynthesis
Primary bile acid biosynthesis
Steroid hormone biosynthesis
Secondary bile acid biosynthesis
Brassinosteroid biosynthesis
[B] Cytochrome P450
Diterpenoid biosynthesis
Indole diterpene alkaloid biosynthesis
Polycyclic aromatic hydrocarbon degradation
Furfural degradation
Metabolism of Other Amino Acid
22



Vossenberg, J, et al. (2013). *Environmental Microbiology* 15(5):1275-1289

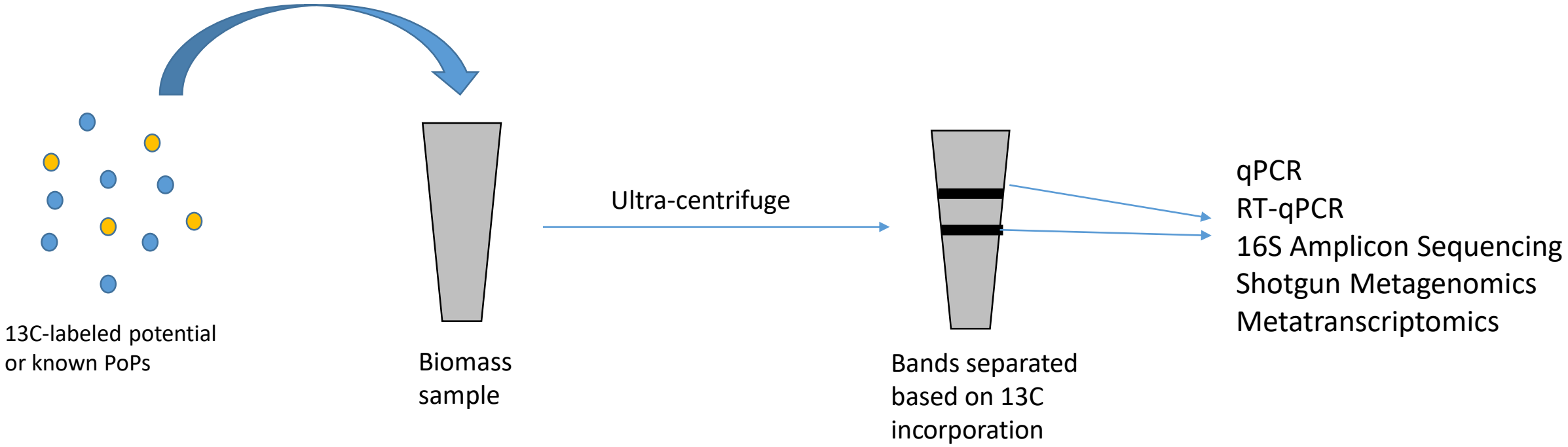


Turner, T, et al. (2013). *The ISME Journal* **7**, 2248-2258

Research Studies

Persistent Organic Pollutants

PoPs (SIP, PCR, NGS)



Who can incorporate these compounds?
Can microbes be used to treat areas polluted with PoPs and then be more safely discarded or used?

Biofuels & Alternate Endpoints

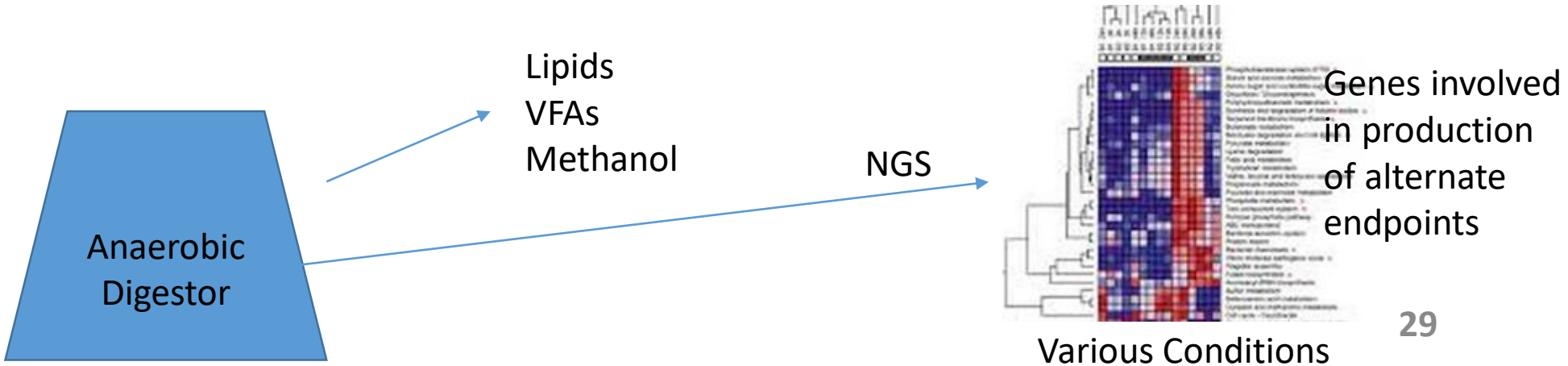
Biofuels (NGS)

Current research is more focused on **wastewater reuse** than treatment

How can NGS and other molecular techniques be applied?

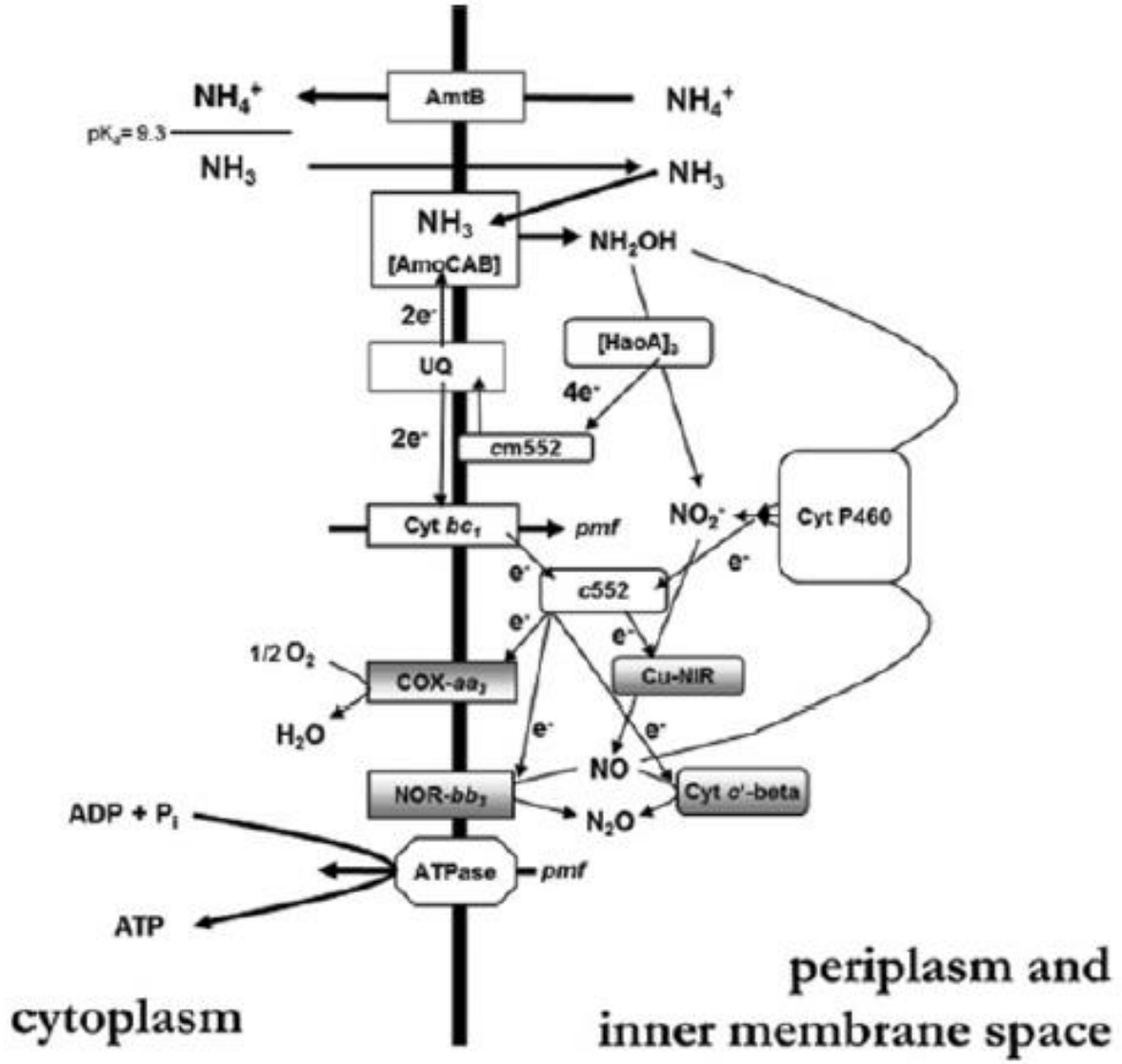
- What biochemical processes are possible in your samples?
- Perhaps side reactions are already occurring which you can identify
- Under which conditions is the production of alternative endpoints optimized

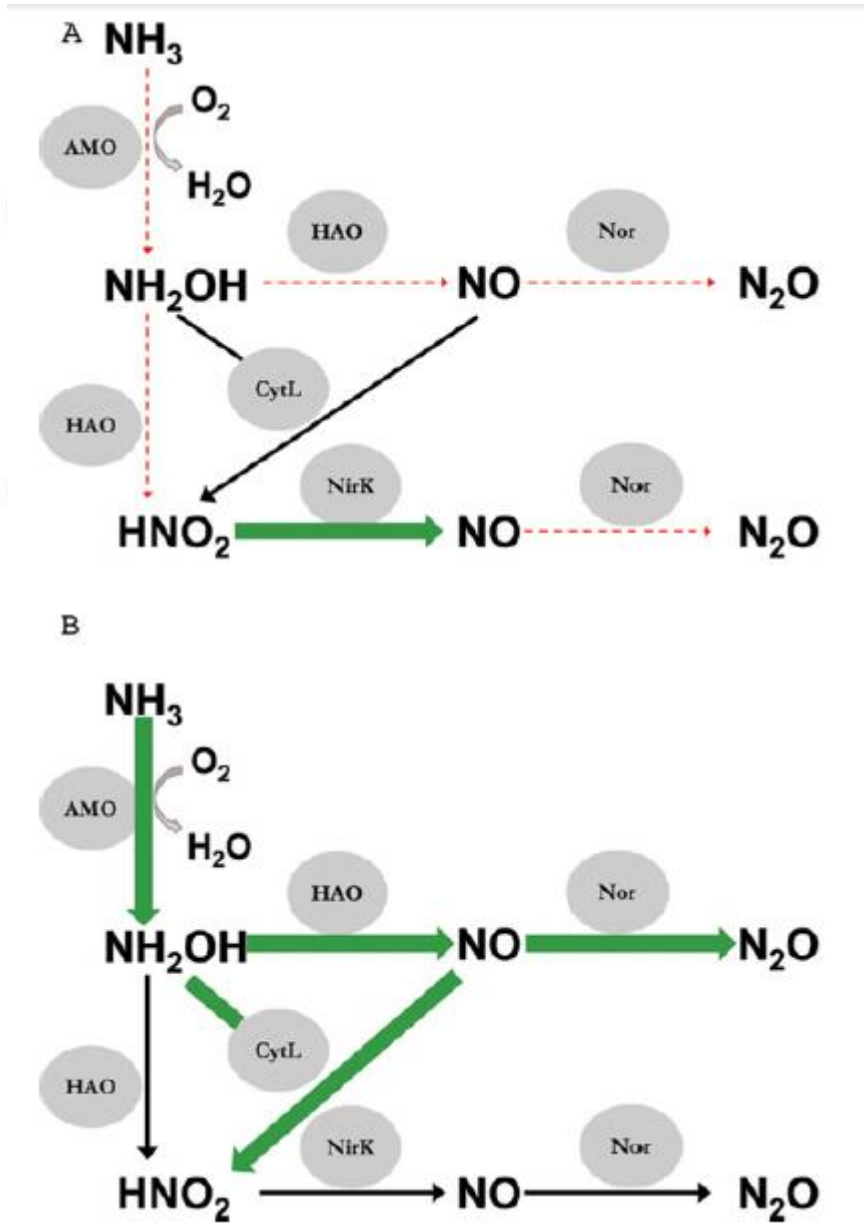
Example:



Greenhouse Gas Emissions

Who is producing these gases and under which conditions?

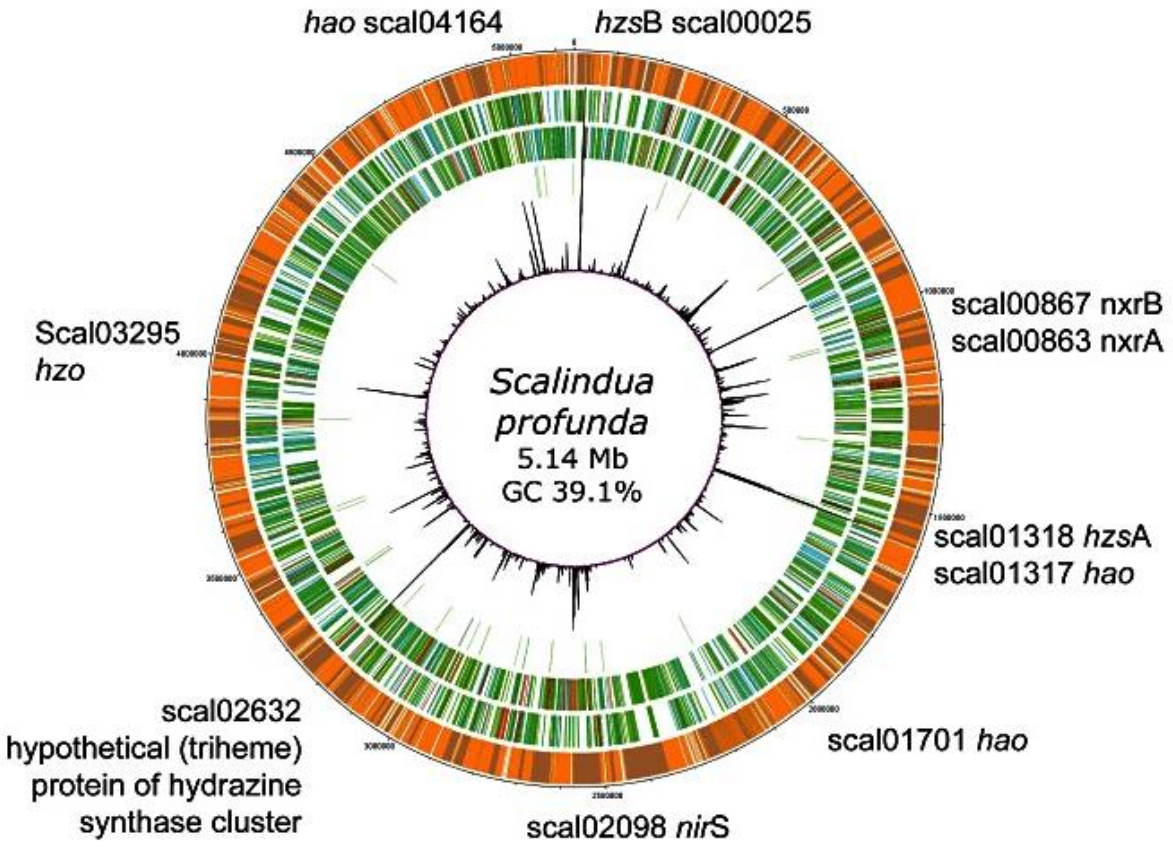




Chandran, K (2011). *Biochem. Soc. Trans.* **39**, 1832-1837

Anaerobic Ammonia Oxidation (Anammox)

Finding & Understanding Novel Species



Vossenberg, J, et al. (2013). *Environmental Microbiology* 15(5):1275-1289

Full-Scale WWTP Monitoring

Full-Scale Monitoring/Research

Which organisms are currently present in my treatment plant?

What are these organisms capable of doing?

What is the activity of these organisms?

Linking chemical & gaseous profiles to molecular data

Molecular “fingerprinting”

Conclusions

- Molecular techniques now include a wider range of possibilities than ever
- In R&D situations, techniques such as NGS can be useful for exploratory understanding of a system
- For more practical applications, SIP, PCR-based techniques, NGS can all be used to answer very specific questions
- These techniques vary greatly in cost, expertise required, input materials, but determining which can be used to answer your question of interest is the first step

Thank You!