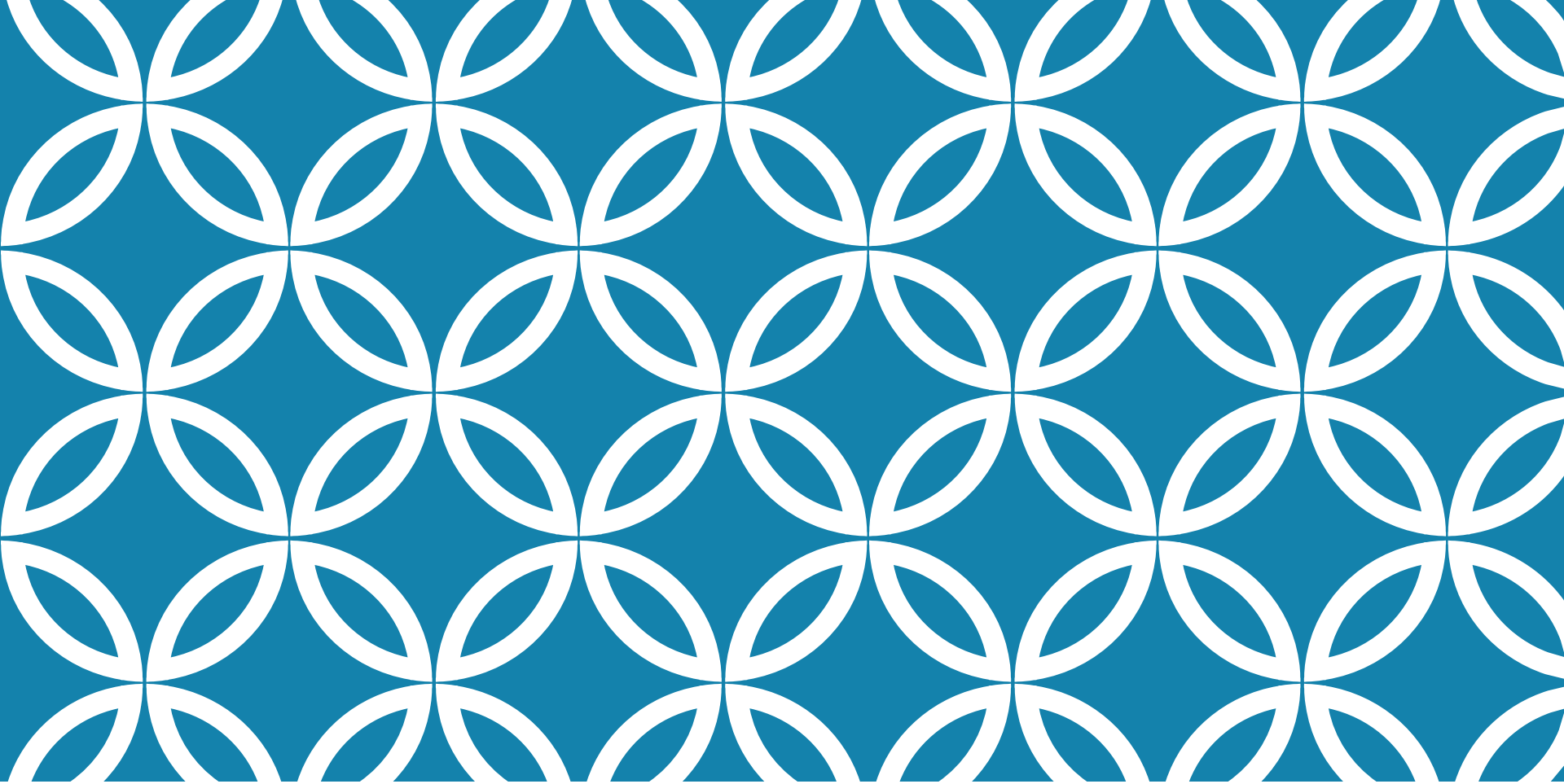


METHODS OF STUDYING MICROBIAL DIVERSITY

INTRODUCTION

Soil bacteria & fungi play important roles in biogeochemical cycles

- Influence above-ground ecosystems

Diversity of soil microbes are limited

- Due to our inability to study soil microbes

Estimated in 1 g of soil = 4 000 – 5 000 different bacterial "genomic units" based on DNA-DNA reassociation.

About 5000 bacteral species have been described.

Only 1% of soil microbes are culturable by standard lab practices

It is not known whether the 1% represents the actual population

- An estimated 1,500,000 species of fungi exist in the world
- Many fungi cannot be cultured by current standard lab methods
- Very little research has been undertaken for soil fungi unlike bacteria
- All organisms in the biosphere depend on microbial activity
- Soil microbes are vital for the continuing cycling of nutrients and for driving above-ground ecosystem

GENERAL LIMITATIONS IN STUDYING MICROBIAL DIVERSITY

Limitations in methodology & lack of taxonomical knowledge

Difficult to study diversity when you can't even identify or categorize the species present

1. SPATIAL HETEROGENEITY

Typically 1 – 5 g of soil taken as representative of microbial community

Limitation:

Innate heterogeneity of soil and spatial distribution of the microbes

Dependent on different soil properties

Bias results based on 1 – 5 g – Could favour detection of dominant populations

2. INABILITY TO CULTURE SOIL MICROBES

Soil bacteria and fungi have immense phenotypic and genotypic diversity

Limitation:

99% of soil microbes can't be cultured

3. LIMITATIONS OF MOLECULAR METHODS

Molecular techniques based on PCR have been used to overcome limitations of culture-based methods

Limitation:

Have there own limitations depending on technique

- Cell lysis efficiency
- DNA or RNA extraction methods – Inhibitory substances such as humic acids?
- Differential amplification of target genes – 16S rRNA, 18S rRNA or ITS region
  - Different affinities of primers to templates
  - Different copy numbers of target genes
  - Hybridization efficiency
  - Primer specificity
  - Sequences with lower G+C content separate more efficiently in the denaturing step of PCR and could be preferencially amplified

4. TAXONOMIC AMBIGUITY OF MICROBES

- Problem of defining microbial species
- No single definition of the term “species”
- Traditional species definition was based on higher plants & animals
- Genetic plasticity of bacteria allowing DNA transfer through plasmids, bacteriophages and transposons complicates the concept of bacterial species
- Most of the fungal taxonomy is based on the sexual states

METHODS OF STUDYING MICROBIAL DIVERSITY

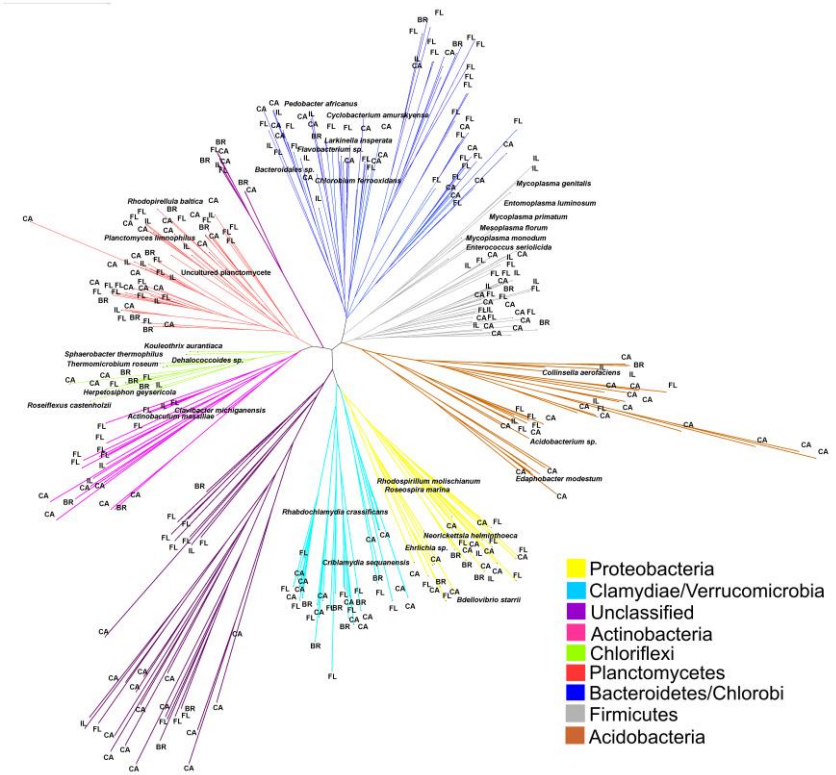
Species diversity consists of:

1. Species richness
2. Total no. of species
3. Species evenness
4. Distribution of species

2 Major Methods:

A. Biochemical

B. Molecular



BIOCHEMICAL METHODS

1. Plate Count

   Selective plating

   Direct viable counts

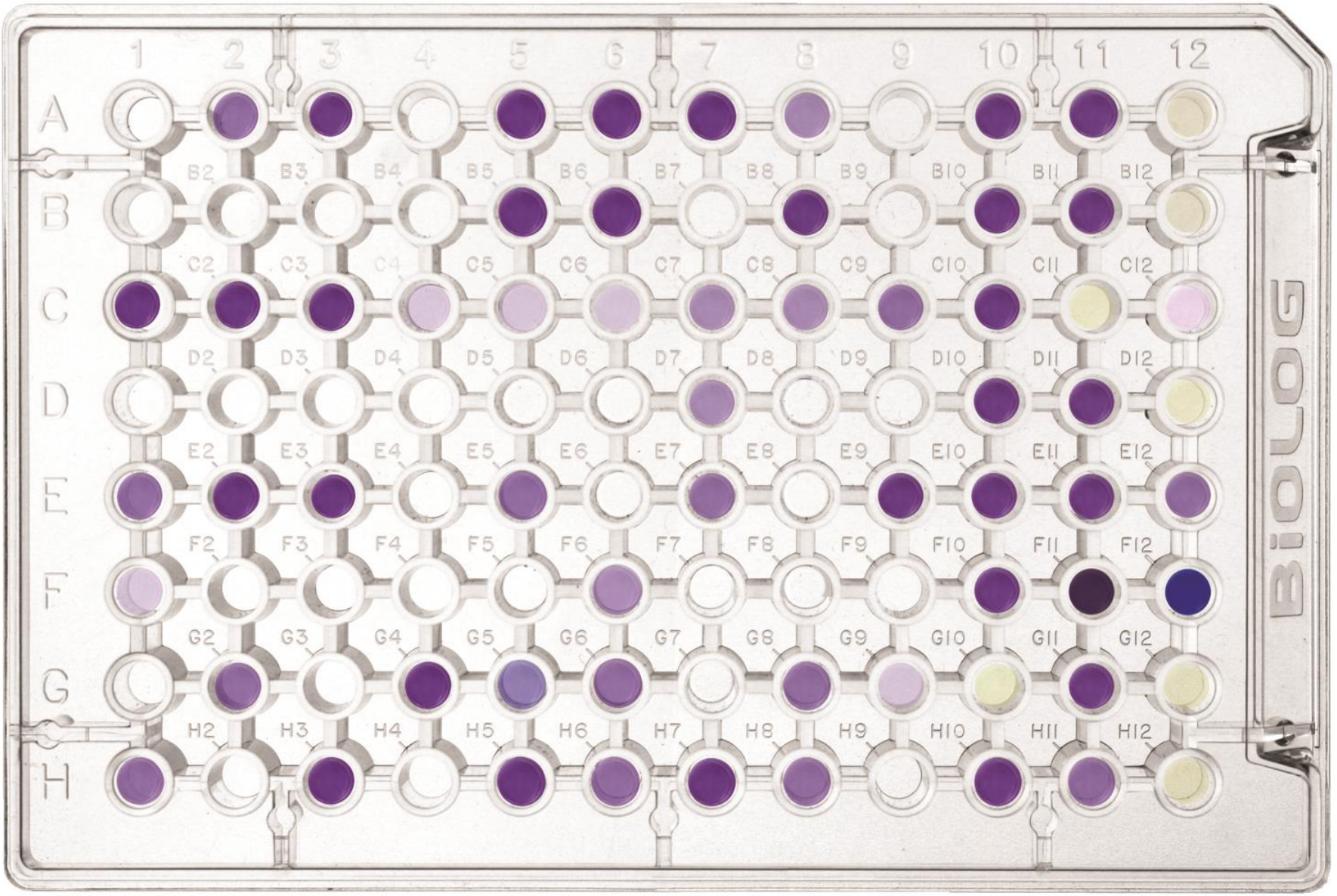
2. Community physiological profiling (CLPP)

   Uses 96-well microtitre plates

   Produces sole carbon source utilisation pattern (SSCUP)

   Different plates for Gram-positive & Gram-negative organisms

   High-throughput


1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12
B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12
D1 D2 D3 D4 D5 D6 D7 D8 D9 D10 D11 D12
E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11 E12
F1 F2 F3 F4 F5 F6 F7 F8 F9 F10 F11 F12
G1 G2 G3 G4 G5 G6 G7 G8 G9 G10 G11 G12
H1 H2 H3 H4 H5 H6 H7 H8 H9 H10 H11 H12
BIOLOG

BIOCHEMICAL METHODS

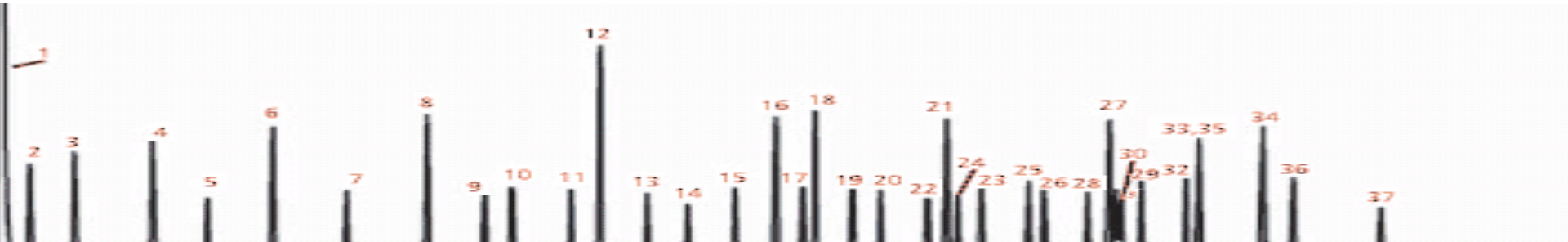
3. Fatty acid methyl ester analysis (FAME)

   Determined based on grouping of fatty acids

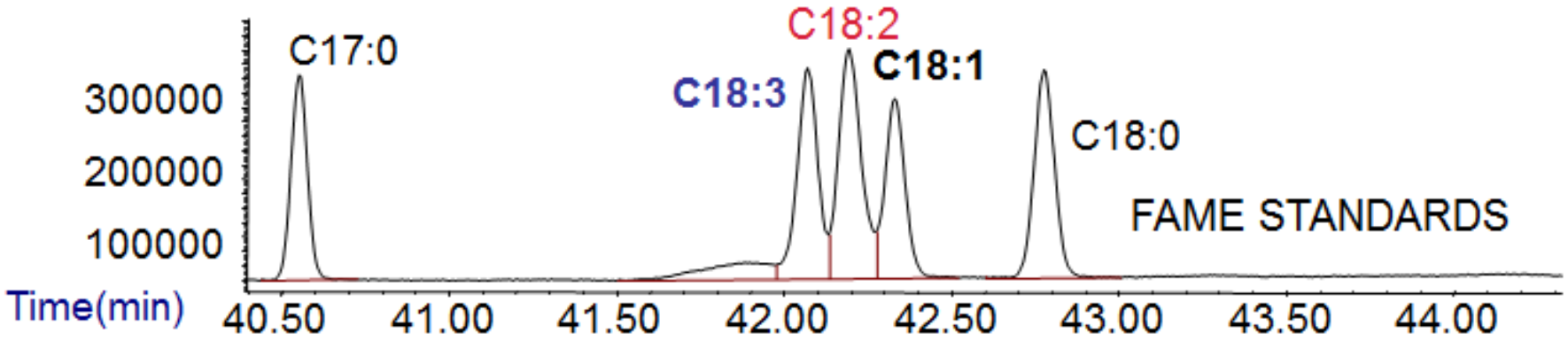
   Fatty acids make up constant proportion of cell biomass

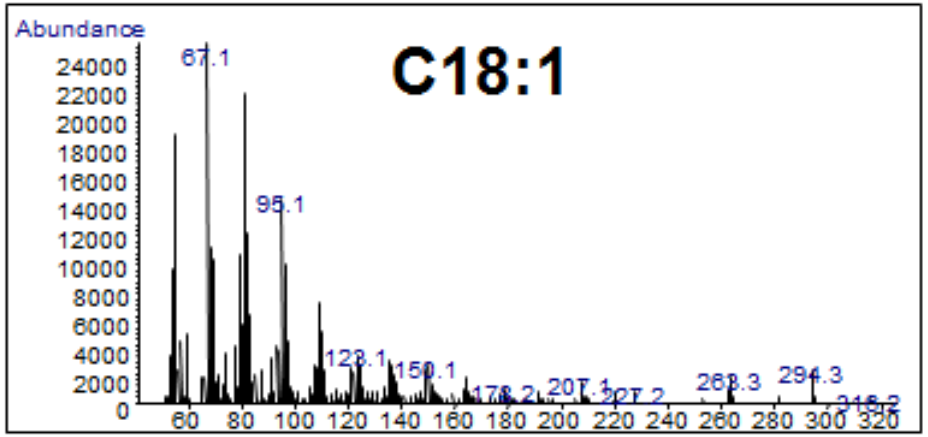
   There are signature fatty acids to differentiate between different taxonomic groups

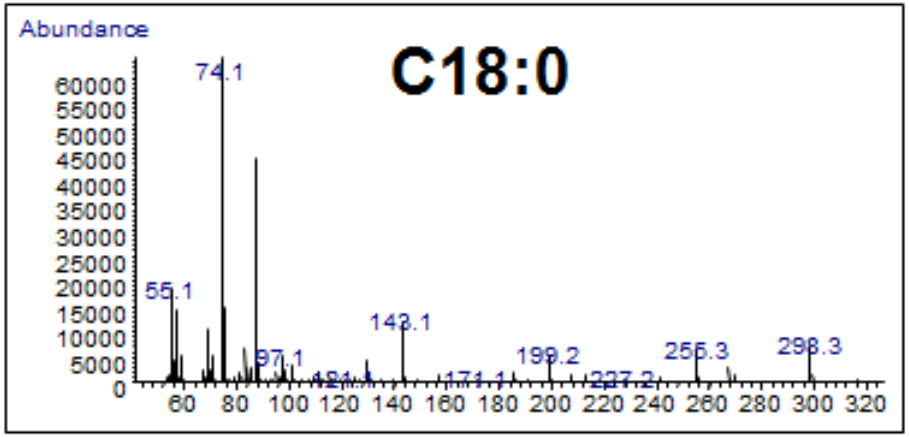
   Change in fatty acid profile indicates changes in microbial community

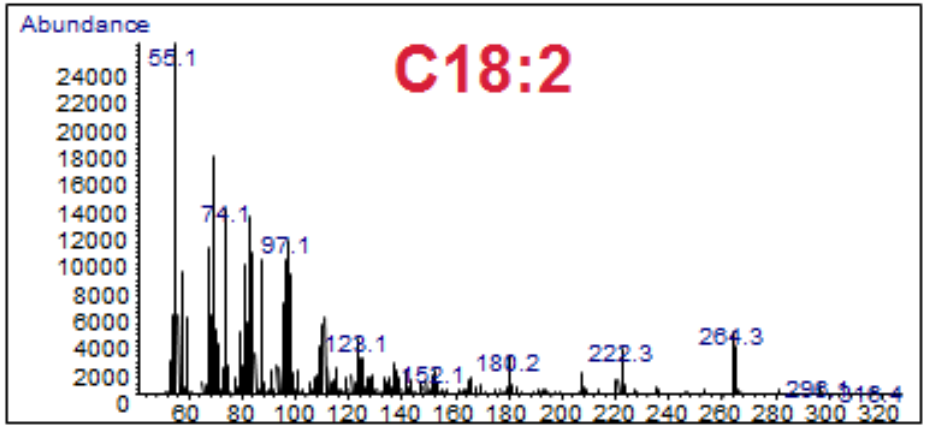


FATTY ACID METYL ESTERS (FAME) ANALYSIS BY GC-MS









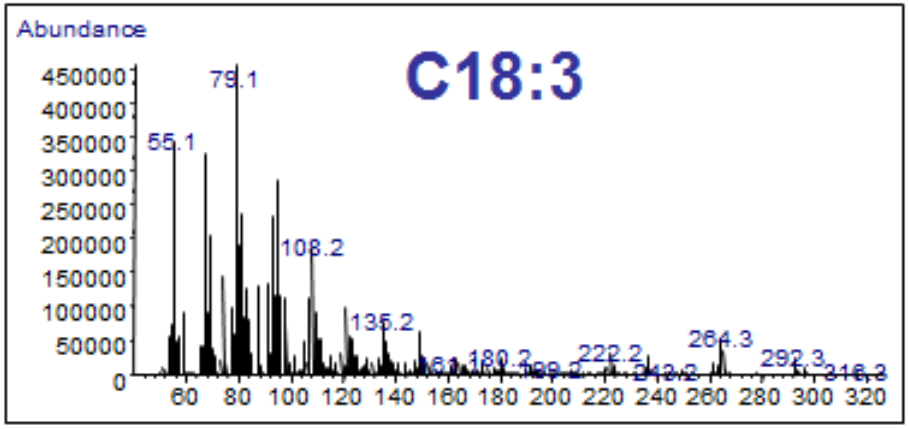


TABLE 1: ADVANTAGES & DISADVANTAGES OF BIOCHEMICAL METHODS

Technique	Advantages	Disadvantages
Plate Counts	• Fast • Cost effective	• Unculturable microbes not detected • Bias towards fast growing organisms • Bias towards fungal species which produces large amount of spores
Community level physiological profiling (CLPP)	• Fast • Highly reproducible • Relatively inexpensive • Differentiate between microbial communities • Generates large amount of data	• Only represents culturable fraction of the community • Favours fast growing organisms • Only represents those organisms capable of utilizing available carbon sources • Potential metabolic diversity • Sensitive to inoculum density
Fatty acid methyl ester analysis (FAME)	• No culturing needed • Direct extraction from soil	• Can be influenced by external factors

MOLECULAR BASED METHODS

Based on the basic unit of life: nucleic acids

Methods incl.:

1. G+C content
2. Nucleic acid re-association and hybridization
3. DNA Microarrays
4. DNA Cloning & Sequencing
5. PCR-based methods
   - **DGGE/TGGE**
   - Single strand conformation polymorphism (SSCP)
   - Restriction fragment length polylymorphism (RFLP)/Amplified ribosomal DNA restriction analysis (ARISA)
   - **Terminal restriction fragment length polymorphism (T-RFLP)**
   - Ribosomal intergenic spacer analysis IRISA)/Automated ribosomal intergenic spacer analysis (ARISA)
   - Highly repeated sequence characterization or microsatellite regions

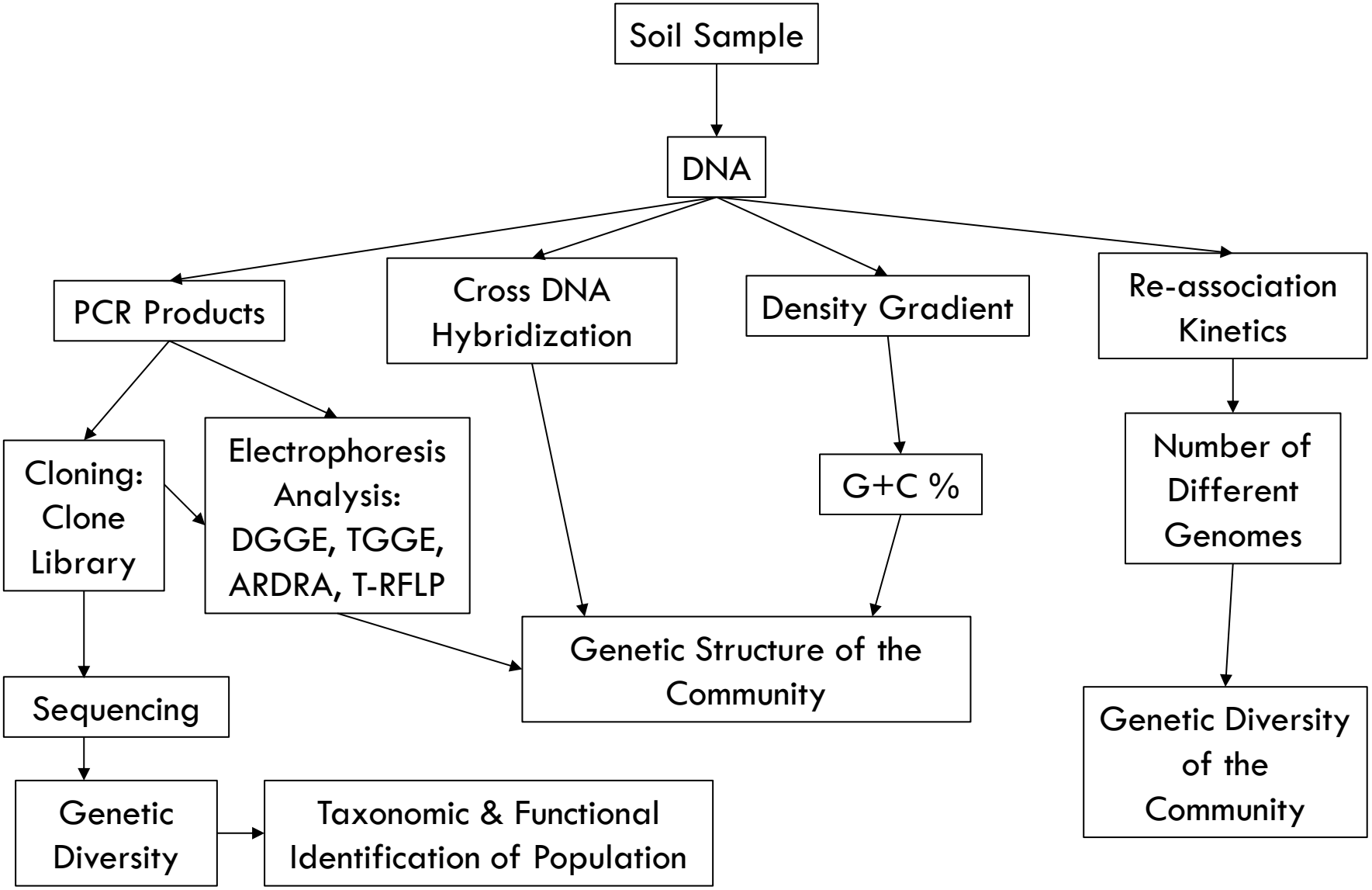
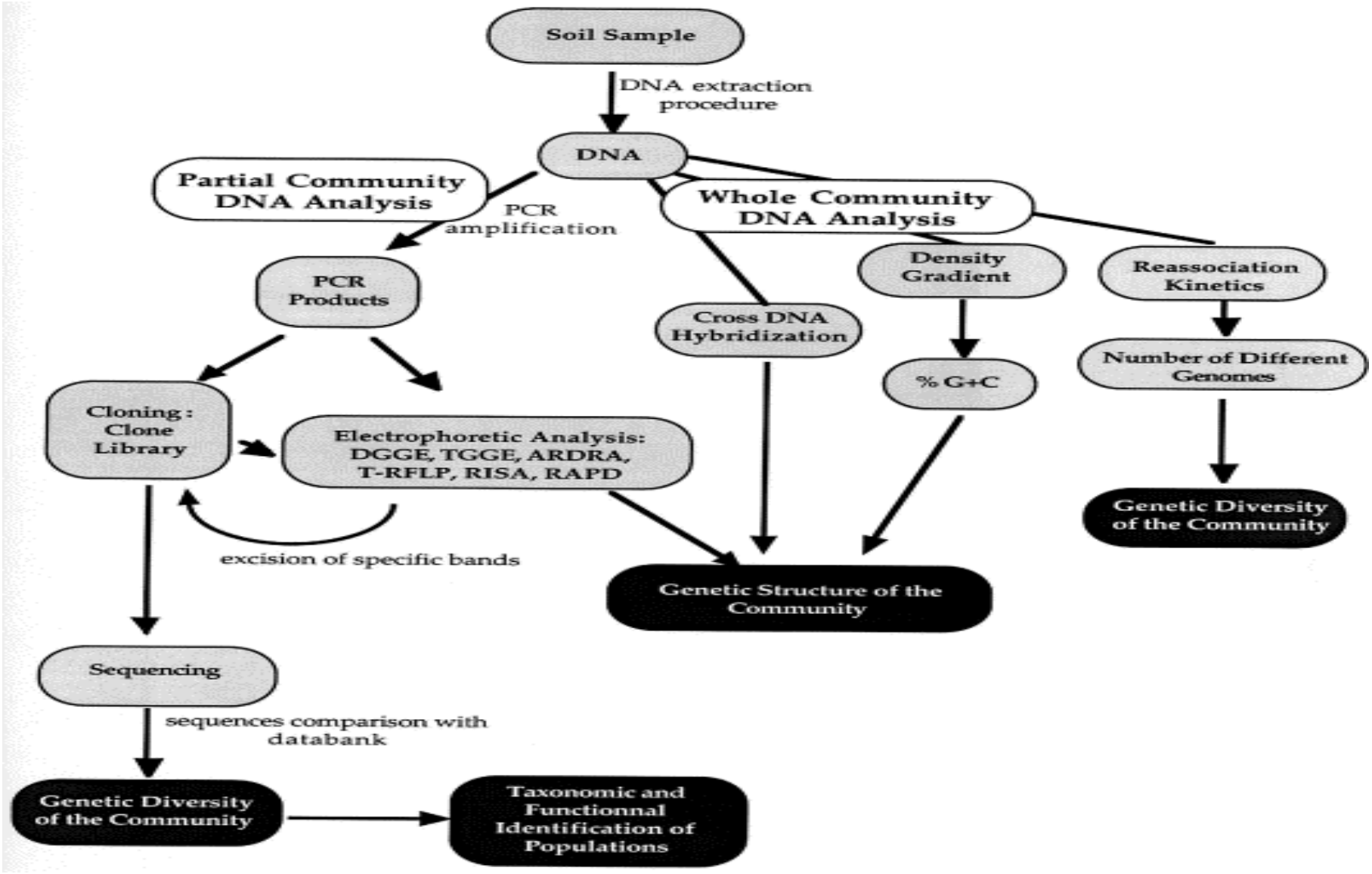
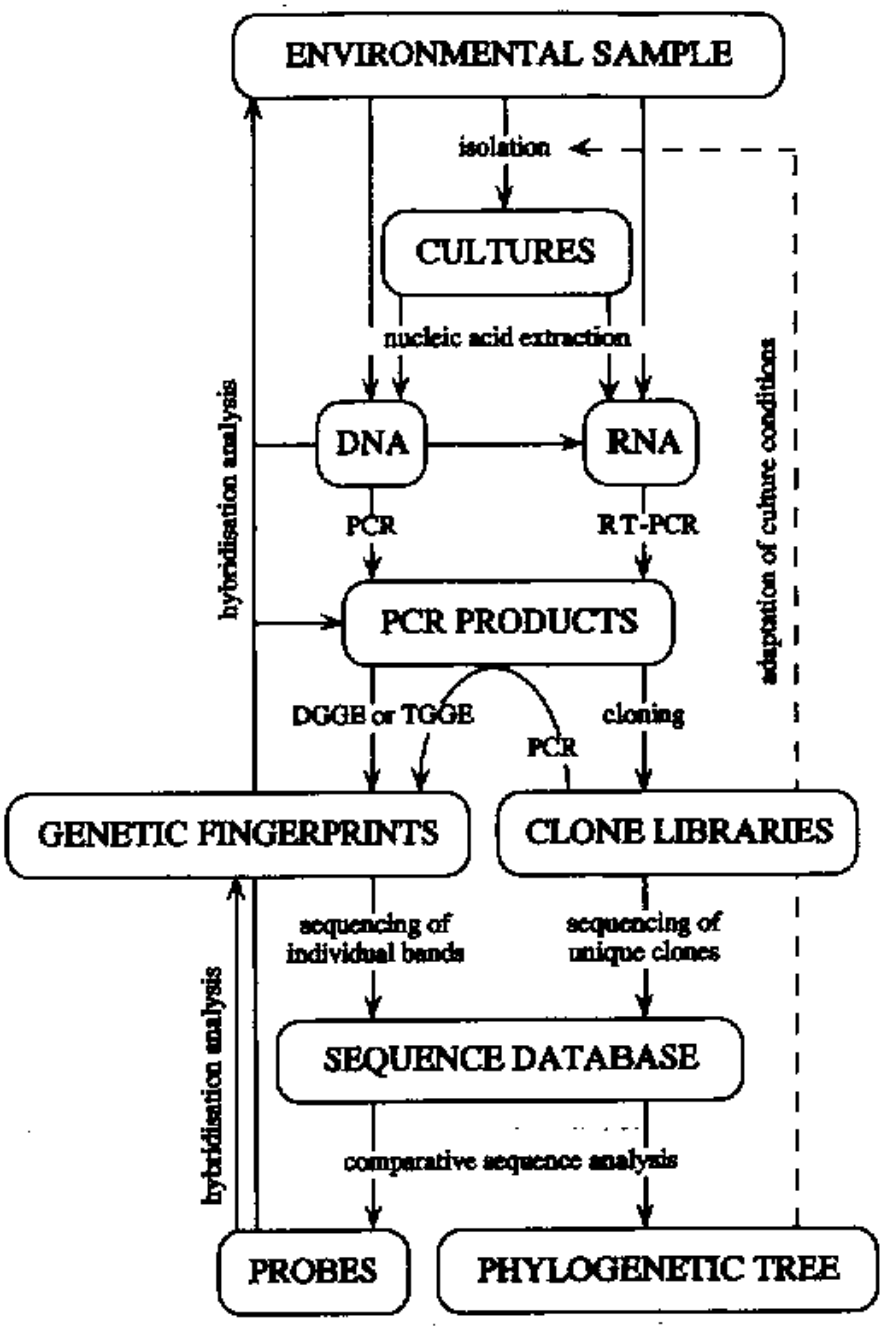


Figure 1: Scheme illustrating the possibilities for application of molecular techniques to study bacterial communities.



Schematic representation of the different molecular approaches for assessing the genetic diversity and structure of soil bacterial communities



- Flow diagram of the different steps in the study of the structure and function of microbial communities

- DGGE or TGGE strategy study the presence (DNA) and activity (rRNA or mRNA) of bacterial populations in complex mixtures

- To obtain information about particular bacterial population within a community;
  - Hybridization analysis with taxon-specific probes
  - Excision of individual bands from gel and sequence

MOLECULAR METHODS

1. GUANINE & CYTOSINE (G+C) CONTENT

Microbial diversity analysis based on G+C content of DNA of microorganisms

G+C content of taxonomical related groups only differ by 3% - 5%

- Not influenced by PCR biases
- Includes all DNA extracted
- Quantitative and can uncover rare members in the microbial population
- Requires large amount of DNA – up to 50 µg

A complementary test is recommended:

- G+C content
- rDNA restriction analysis (ARDRA)
- rDNA sequence analysis

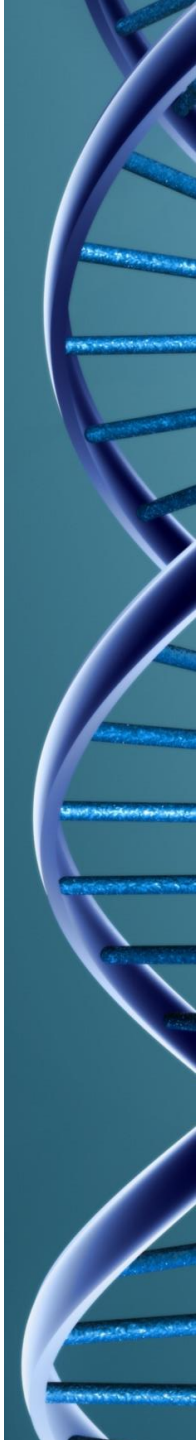
MOLECULAR METHODS

2. NUCLEIC ACID RE-ASSOCIATION & HYBRIDIZATION

DNA re-association is a measure of genetic complexity of the microbial community

Requires: total DNA

- Extraction and purification of total DNA from environmental samples
- Denaturation and re-annealing of DNA
- Rate of hybridization or reassociation depends on sequence similarity

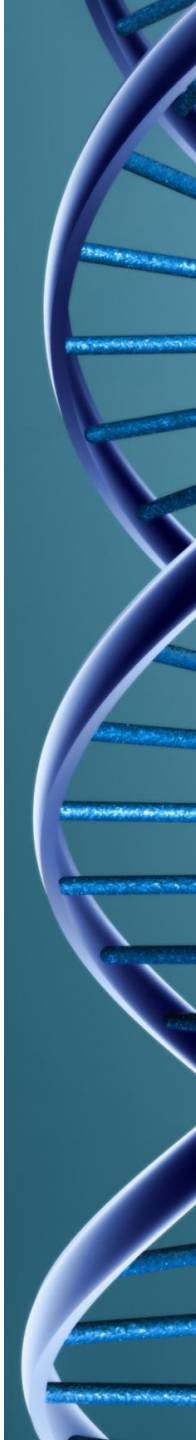


- Rate of hybridization is depended on the similarity of sequences

- Increase in complexity or diversity of DNA sequences = decrease in the rate of DNA re-association

- Time needed for half of the DNA to re-associate is used as a diversity index
  - This takes into account : the amount & distribution of DNA re-association

- Similarity between communities of two different samples can be studied by measuring the degree of DNA similarity through hybridization kinetics

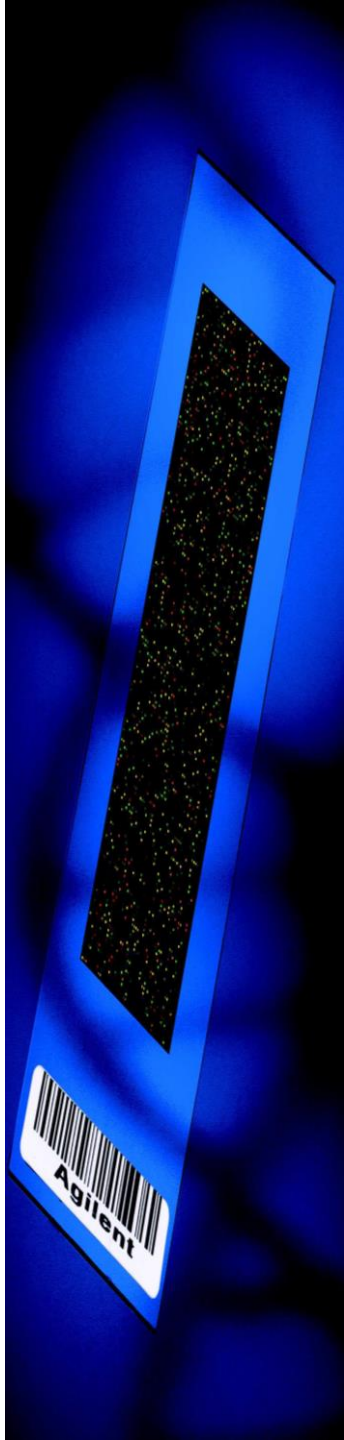


- Nucleic acid hybridization can be done on extracted DNA or RNA or in situ.
- Requires: target probes
  - Oligonucleotide or polynucleotide designed from known sequences
  - Ranges in specificity from domain to species
  - Incl. fluorescent markers (fluorescein, rhodamine)
- Quantitative dot blot can be used to measure relative abundance
- Traditionally, radioactive isotopes were used to label the probes but fluorescent probes are often preferred e.g. for fluorescent in situ hybridization (FISH) techniques.
- Lacks sensitivity – May not detect sequences present in low copy number

MOLECULAR METHODS

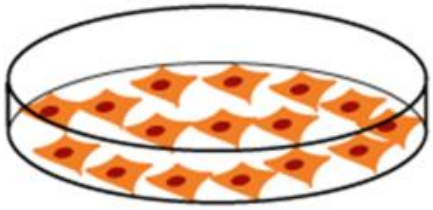
3. DNA MICROARRAYS

- DNA-DNA hybridization may be used together with DNA microarrays to detect and identify bacterial species
- Valuable tool in microbial diversity studies
- Single array can contain thousands of DNA sequences with high specificity
  - **Functional diversity** – using specific target genes e.g. nitrate reductase
  - **Species diversity** – contain sample of environmental "standards" (DNA fragments with less than 70% hybridization representing different species found in environmental sample.
- High-throughput



Oligonucleotide microarray

Cells of person 1/condition 1



RNA isolation

mRNA

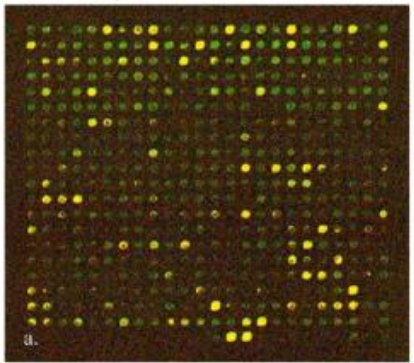
Reverse transcriptase labeling

cDNA

"Green Fluorescent" Targets

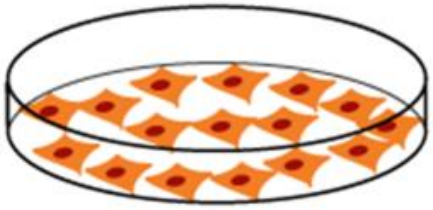
Hybridize to microarray

Microarray with short ssDNA spanning the entire genome

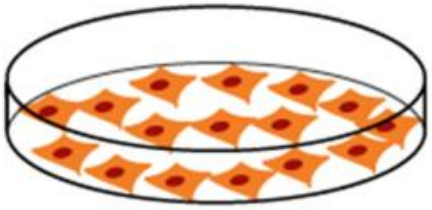


cDNA microarray

Cancer cells



Normal cells



RNA isolation

mRNA

mRNA

Reverse transcriptase labeling

cDNA

"Red Fluorescent" Targets

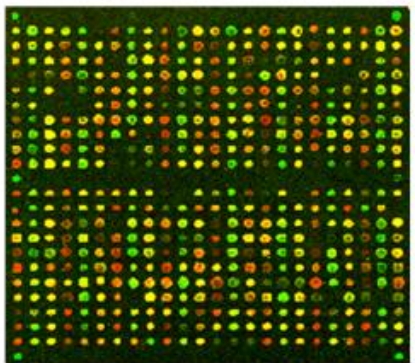
cDNA

"Green Fluorescent" Targets

Combine targets

Hybridize to microarray

Microarray with long cDNA covering the transcriptional activity of the cell type



REVERSE SAMPLE GENOME PROBING (RSGP)

- Used to analyze microbial community composition of the most dominant culturable species, using genome microarrays
- 4 Steps involved:

1. Isolation of genomic DNA from pure cultures

2. Cross hybridization testing to obtain DNA fragments with less than 70% cross-hybridization – DNA fragments with >70% = same species

3. Preparation of genome arrays onto solid support

4. Random labelling of a defined mixture of total community DNA and internal standard

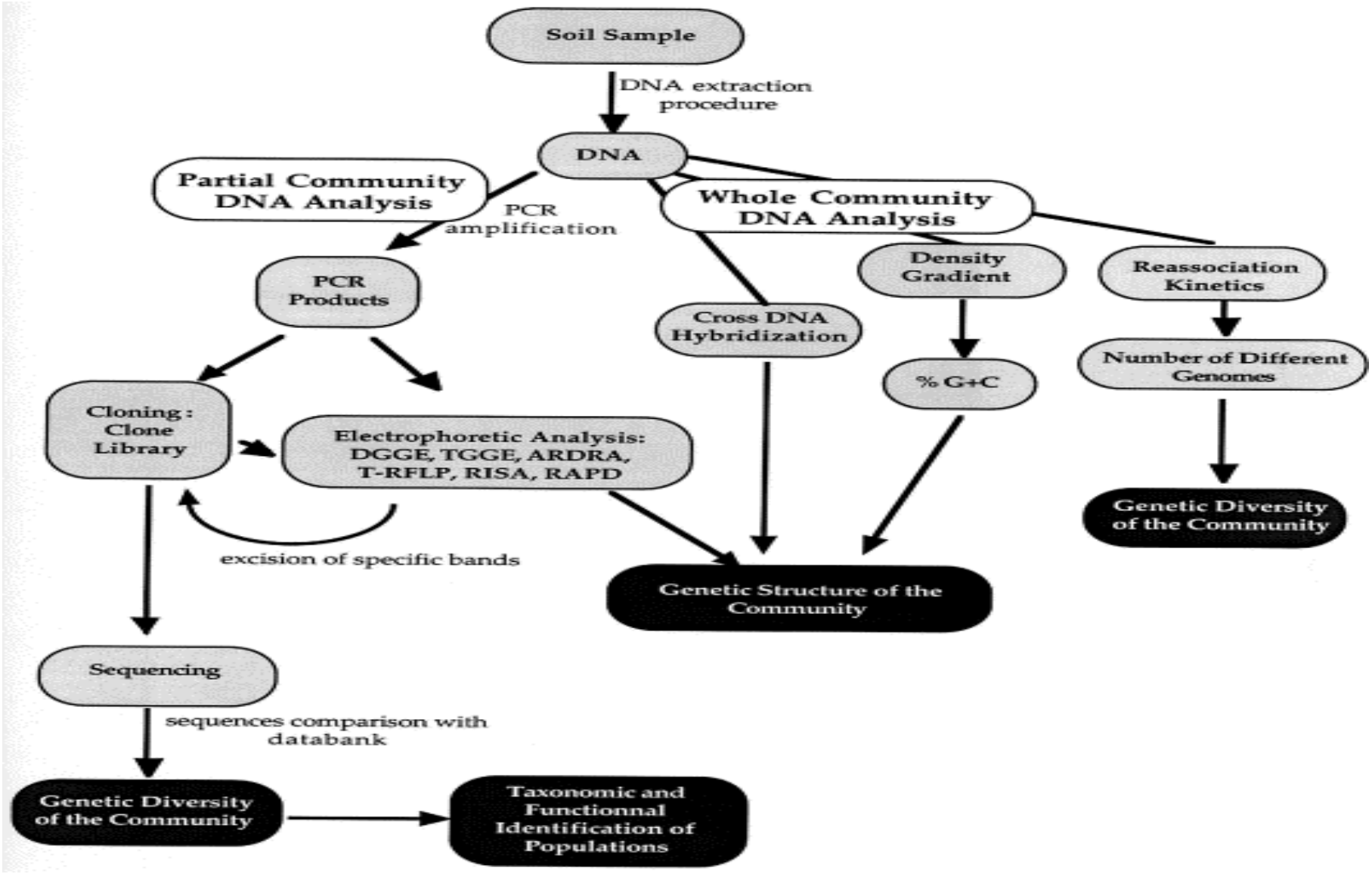
- Used to study the microbial population in produced waters and on corrosion coupons of a select number of Western Canadian oil fields using 16 distinct genomes of sulfate-reducing bacteria (SRB) and 4 of heterotrophs.

PCR-BASED MOLECULAR METHODS

- PCR targeting the 16S rDNA
  - – for prokaryote diversity
  - - allows identification of prokaryotes and prediction of phylogenetic relationship

- 18S rDNA and internal transcribed spacer (ITS) region
  - – used to study fungal communities
  - – Available databases not as extensive as for prokaryotes

- Initially, molecular-based methods for ecological studies relied on cloning of target genes isolated from environmental samples
- Although sequencing has become routine, sequencing thousands of clones is cumbersome
- DNA is extracted from the environmental sample and purified. Target DNA (16S, 18S or ITS) is amplified using universal or specific primers and the resulting products are separated in different ways.



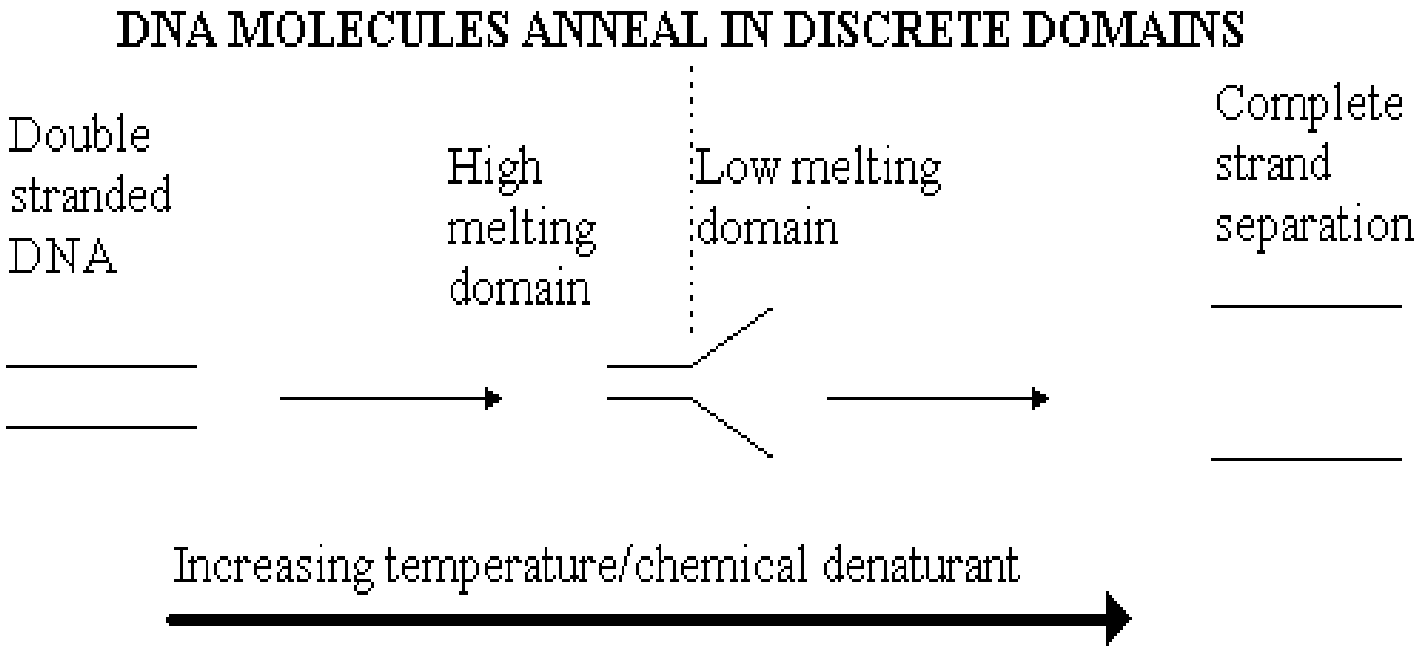
Schematic representation of the different molecular approaches for assessing the genetic diversity and structure of soil bacterial communities

DGGE & TGGE

- In DGGE & TGGE, PCR-amplified DNA fragments of the same length but with different DNA sequences can be separated
- Separation is based on the decreased electrophoretic mobility of a partially melted DS DNA molecule in polyacrylamide gels containing a linear gradient of DNA denaturants (a mixture of urea and formamide) or a linear temperature gradient.
- The melting of DNA fragments proceeds in discrete so-called *melting domains*: stretches of base-pairs with an identical melting temperature (T_m).
- Once a domain with the lowest melting temperature reaches its T_m at a particular position in the denaturing or temperature gradient gel, a transition of a helical to a partially melted molecule occurs, and the migration of the molecule will practically halt.

- Sequence variation within such domains causes the melting temperatures to differ, and molecules with different sequences will stop migrating at different positions in the gel.
- The temperature at which a DNA duplex melts is influenced by two factors:
  - **The hydrogen bonds formed between complimentary base pairs, GC rich regions melt at higher temperatures than regions that are AT rich.**
  - **The attraction between neighbouring bases of the same strand or "stacking"**
- A GC-rich sequence (between 30 and 50 nucleotides), a so-called GC-clamp is normally attached to the 5'-end of one of the PCR-primers, co-amplified and thus introduced into the amplified DNA fragments.
- The GC-rich sequence acts as a **high melting domain:**
  - Prevents the two DNA strands from complete dissociation into single strands

DIAGRAMATIC ILLUSTRATION OF DGGE PRINCIPLE



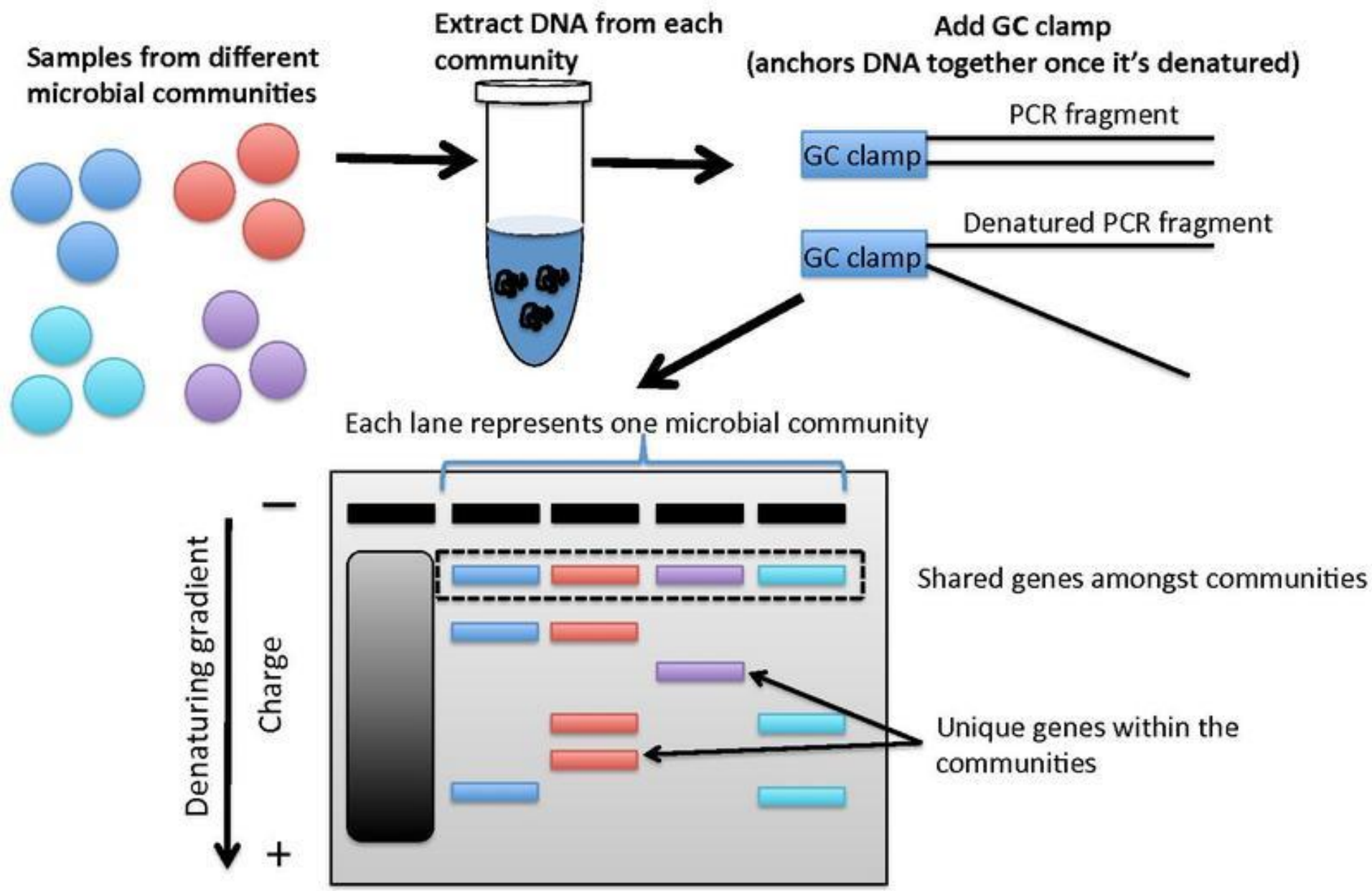
•DS DNA is subjected to denaturant environment in a denaturing gradient polyacrylamide gel and will melt in discrete segments called "melting domains"

• When the Tm of the lowest melting domain is reached, the DNA will become partially melted, creating branched molecules

• Partial melting reduces the DNA's mobility in the polyacrylamide gel

• The branched structure of the SS moiety of the molecules becomes entangled in the gel matrix and no further movement occurs

DGGE: Denaturing Gradient Gel Electrophoresis



APPLICATIONS

Studying community complexity

Studying community changes

Monitoring the enrichment and isolation of bacteria

Comparison of different DNA extraction protocols

Screening of clone libraries

Determining PCR and cloning biases

LIMITATIONS

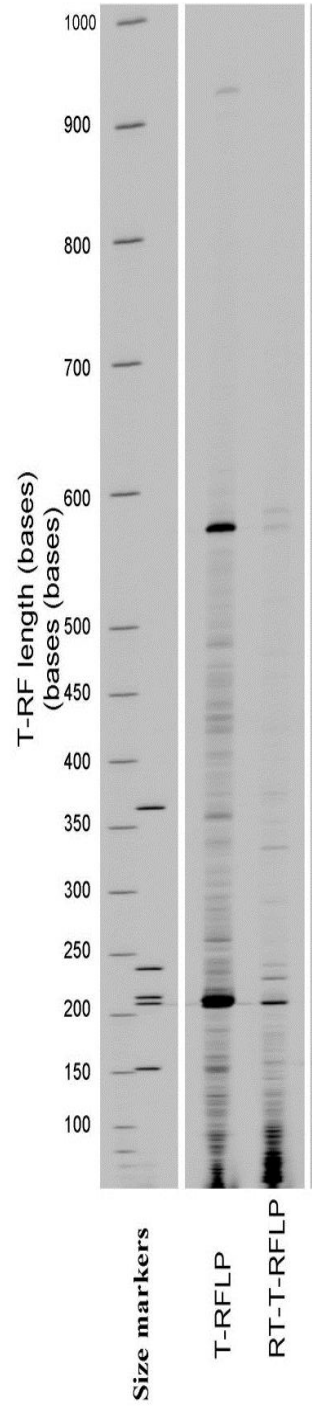
- **Separation of only relatively small fragments**, up to 500 bp, thus limiting the amount of sequence information for phylogenetic inferences and probe design.
- Difficulty in separating DNA fragments which have certain amount of sequence variation.
- Use of different regions of the 16S rRNA and different DGGE conditions might result in different resolutions of separation.
- Limitation in the maximum number of the different DNA fragments which can be separated by DGGE.
- Several studies revealed that only bacterial populations that make up 1% or more of the total community can be detected by PCR-DGGE.
- **Co-migration of DNA fragments.**
- **Multiple heterogeneous rRNA operons -** One organism may produce more than one DGGE band.
- **Primer degeneracies** and **greatly varying G+C contents of amplified DNA molecules** have also been suspected to cause differential amplification efficiencies.

RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)/ AMPLIFIED RIBOSOMAL DNA RESTRICTION ANALYSIS (ARDRA)

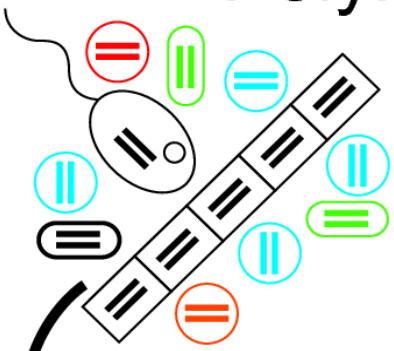
- PCR-amplified rDNA is digested with a 4-base pair cutting restriction enzyme.
- Different fragment lengths are detected using agarose or non-denaturing polyacrylamide gel electrophoresis in the case of community analysis.
- RFLP banding patterns can be used to screen clones or used to measure bacterial community structure.
- This method is useful for detecting structural changes in microbial communities but not as a measure of diversity or detection of specific phylogenetic groups.
- Banding patterns in diverse communities become too complex to analyze using RFLP since a single species could have four to six restriction fragments.
- Using a six-base cutting enzyme, the number of restriction fragments per species could be reduced, thereby increasing the resolution of this method.

TERMINAL RESTRICTION FRAGMENT LENGTH POLYMORPHISM (T-RFLP)

- Follows the same principle as RFLP except that one PCR primer is labelled with a fluorescent dye, such as TET (4,7,2V,7V-tetrachloro-6-carboxyfluorescein) or 6-FAM (phosphoramidite fluorochrome 5-carboxyfluorescein).
- It allows for the detection of only the labelled terminal restriction fragment.
- This simplifies the banding pattern, thus allowing the analysis of complex communities.
- It also provides information on diversity as each visible band represents a single operational taxonomic unit or ribotype.
- The banding pattern can be used to measure species richness and evenness as well as similarities between samples.
- Technique can be automated

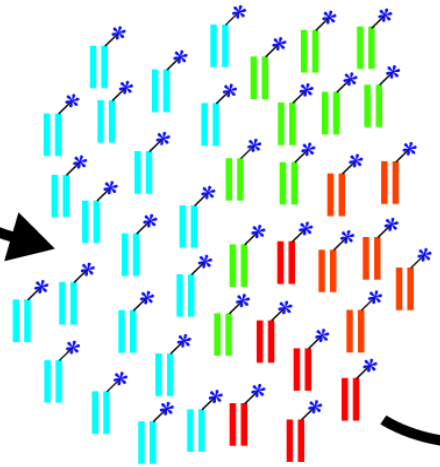


Terminal Restriction Fragment Length Polymorphism Analysis (TRFLP)

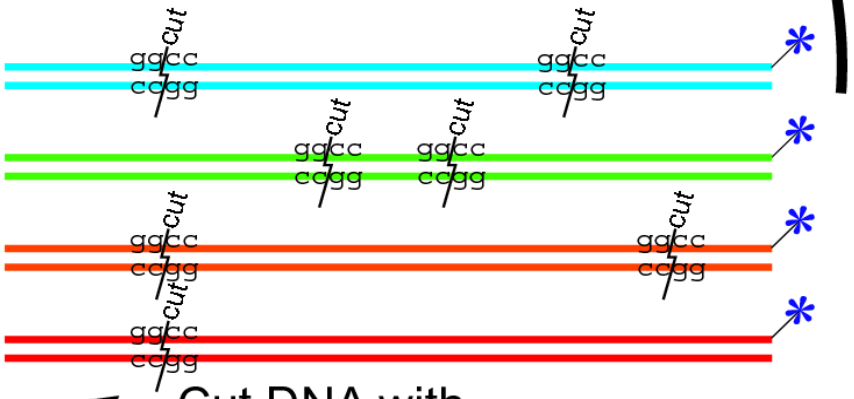


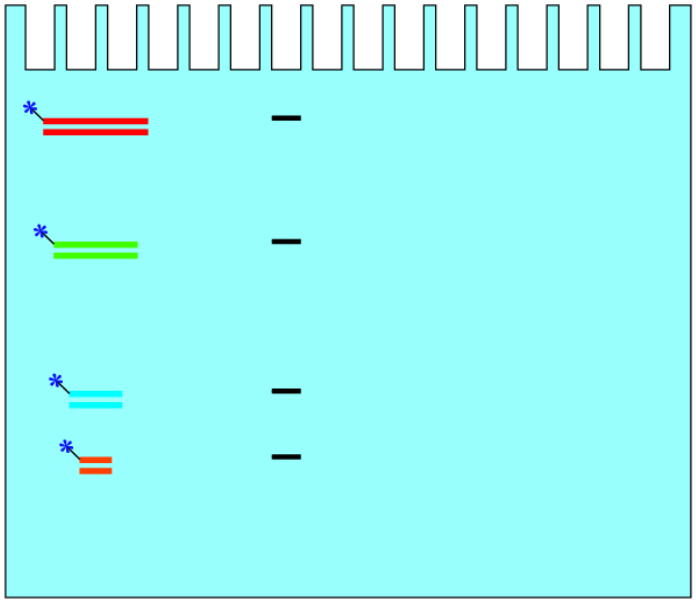
Extract DNA

PCR 16S rRNA genes with labelled primer



Cut DNA with Restriction Enzymes





Gel Electrophoresis of terminal restriction fragments

TABLE 2: ADVANTAGES & DISADVANTAGES OF MOLECULAR METHODS

Technique	Advantages	Disadvantages
G+C	• Not influenced by PCR biases • Incl. all DNA extracted • Quantitative • Incl. rare members of the community	• Requires large quantities of DNA • Dependent on lysing and extraction efficiency • Coarse level of resolution
Nucleic acid re-association & hybridization	• Total DNA extracted • Not influenced by PCR biases • Study DNA or RNA • Can be studies *in situ*	• Lack of sensitivity • Sequences need to be in high copy no. to be detected • Dependent on lysing & extraction efficiency
DNA Microarrays & DNA Hybridization	• Same as nucleic acid hybridization • Thousands of genes can be analysed	• Only detect most abundant species • Need to be able to culture organisms • Only accurate in low diversity systems
DGGE & TGGE	• Large no. of samples can be analysed simultaneously • Reliable, reproducible and rapid	• PCR biases • Dependent on lysing & extraction efficiency • Sample handing can influence community • Co-migration • Only detects dominant species
T-RFLP	• Simpler banding patterns than RFLP • Can be automated • High-throughput • Compare differences in microbial communities	• Dependent on extraction & lysing efficiency • PCR biases • Type of *Taq* can increase variability • Choice of universal primers • Choice of restriction enzymes will influence community fingerprint

RIBOSOMAL INTERGENIC SPACER ANALYSIS (RISA)/AUTOMATED RIBOSOMAL INTERGENIC SPACER ANALYSIS (ARISA)

- In RISA and ARISA, the intergenic spacer (IGS) region between the 16S and 23S ribosomal subunits is amplified by PCR, denatured and separated on a polyacrlyamide gel under denaturing conditions.

- This region may encode tRNAs and is useful for differentiating between bacterial strains and closely related species because of heterogeneity of the IGS length and sequence.

- In RISA, the sequence polymorphisms are detected using silver stain
- In ARISA the forward primer is fluorescently labelled and is automatically detected.

HIGHLY REPEATED SEQUENCE CHARACTERIZATION OR MICROSATELLITE REGIONS

- Many organisms, both prokaryotic and eukaryotic, contain highly repetitive short DNA sequences that are 1–10 base pairs long repeated throughout their genomes.
- Depending on the rate of evolution, these sequences may be diagnostic and allow differentiation down to the species or strain level.
- The method, also termed rep-PCR, has been used for identification of bacteria since it provides a genomic fingerprint of chromosome structure, and chromosome structure is considered to be variable between strains.
- Highly repeated sequences are also referred to as microsatellite regions and have been used for identification of mycorrhizal fungi.
- Fingerprinting of PCR-amplified microsatellites can be compared using similarity indices to investigate difference at the inter- and intraspecific level.