Nomenclature, Classification and Identification

Anil Shrestha
Msc Medical Microbiology

Terminologies

- Nomenclature :- concerned with assignment of name to taxonomic groups in agreement of published rules.
- Classification :- ordering organism on the basis of shared properties in particular taxon.
- Identification :- process of obtaining data on the properties of organism i.e, characterization and determination which species is belong to based on direct comparison to known taxonomic group.
- Systemics:- any study of nature of organism, when the knowledge gained is used in taxonomy.

  Study of morphology, ecology, epidemiology, biochemistry, genetics and physiology.

Classification of bacteria

- The main aim of bacterial taxonomy is to setup classification, an ordering of bacteria into group based on common properties. This classification can then be used to identify individual bacterial species or strains (identification) and name them (Nomenclature).
- Classification is used to study the relationship of different bacterial groups based on shared phenotypic properties and probably a common evolutionary history.
- A wide range of morphologies, biochemical, molecular characters are measured and then information derived. From this is used to create a hierarchical classification system into which individual bacteria can be placed.

- Caroleus Linnaeus :- The eighteen century Swedish botanist Carlous Linnaeus is credited with foundly the science of taxonomy and he originated the binomial nomenclature.

Binomial nomenclature

Latinized name composed of two words;

--1st word is a genus name (always capitalized) Always a noun

--2nd word is species name (not capitalized)

Usually an adjective

Species name must be continued with genus name.

Meaning of some microorganisms

1. *Escherichia coli*- Named after Theoder escherich in 1888; found in colon.
2. *Haemophilus ducreyii*– Haemo-blood, phil-love. Named after Auguston Ducrey in 1889.
3. *Neisseria gonorrhea*- Named after Albert Neisser in 1879; cause the disease gonorrh oea.
4. *Saccharomyces cerevisiae*- Saccaro=sugar, myco=mold, cerevisiae=beer or ale.
5. *Staphylococccus aureus*- aureus= golden, staphylo=clustrer, Kokkus= berry
6. *Streptococcus lactis*- strepto=twisted chain, kokkus=berry, lacto=milk

7. *Rhizopus nigricans*- Rhizo= rod like, nigricans=black colour

8. *Penicillum notatum*- Penicil=tutt like (pencil) apperance, notatum= spores easily spread in air.

9. *Streptococcus pyogens*- strepto=twisted chain, kokkus=berry, pyogens= from pus

10. *Klebsiella pneumoniae*- kleb= honours of discovered, pneunmoniae= disease it cause

11. *Salmonella typhimurium*- salmon= honours of Daniel Salmon, Typhimurium=causes stupor(typhi) in mice.

12. *Trypanosoma cruzi*- Trypane=borer, soma=body, cruzi=honour of epidemiologist oswaldocruz

Classification of organism

- Kingdom
- Domain
- Phylum or Division
- Class
- Order
- Family
- Genus
- Species

Polyphasic taxonomy

- This approach includes phenotypic, phylogenetic, genotypic feature and Numerical Taxonomy.
- To understand how all of these data are incorporated into a coherent profile of taxonomic criteria, we must first consider the individual components and determine how they are assessed quantitatively through numerical taxonomy.

1) Phenetic Classification

- **Phenetic system -**which groups organisms together based on the mutual similarity of their phenotypic characteristics.
- This classification system succeeded in bringing order to biological diversity and clarified the function of morphological structures. For example, because motility and flagella are always associated in particular microorganisms, it is reasonable to suppose that flagella are involved in at least some types of motility.
- Although phenetic studies can reveal possible evolutionary relationships, they are not dependent on phylogenetic analysis.
- The best phenetic classification is one constructed by comparing as many attributes as possible. Organisms sharing many characteristics make up a single group or taxon.

2) Phylogenetic Classification

- Phylogeny (Greek: Phylum-tribe or race and gene is genera or origin) is refers to evolutionary development of a species.
- Phylogenetic classification : classified organism on the basis of evolutionary relationship.
- At the end of 20^{th} century, woose and fox were able to assess evolutionary relationship between microbes causing nucleotide sequence.
- Currently over 2,00,000 different 16S and 18S rRNA sequence are available.

3) Genotypic Classification

- Organism are classified on the basis of genetic similarities.
- Individual gene or whole gene can be compared.

4) Numerical Taxonomy

- Peter H. A. Sneath and Robert Sokal have defined numerical taxonomy as “the grouping by numerical methods of taxonomic units into taxa on the basis of their character states.”
- Information about the properties of organism is converted into a form suitable for numerical analysis and then compared by means of computer.
- Process begins with determination of presence or absence of selected character in a group of organism under study.

- Many characters at least 50, preferably >100 should be compared for an accurate and reliable classification of organism. Characters used are:
  - Morphological character
  - Chemical character
  - Cultural character
  - Metabolic character
  - Antigenic character
  - Biochemical character
  - Physiological character
  - Ecological character

- After character analysis, association coefficient is calculated.
- Association coefficient is a function that measures the agreement between characters possessed by two organism.
- Simple matching coefficient and Jaccard coefficient is calculated.
- Simple matching coefficient (SSM) is the proportion of characters that match regardless of whether the characters is present or absent.
- Jaccard coefficient is calculated by ignoring any characters that both organism lack.
- Both coefficient increase linearly in value from 0.0 (no matches) to 1.0 (100% matches)

Table 19.2 The Calculation of Association Coefficients for Two Organisms

In this example, organisms A and B are compared in terms of the characters they do and do not share. The terms in the association coefficient equations are defined as follows:

		Organism B: 1	Organism B: 0
Organism A	1	a	b
	0	c	d

a = number of characters coded as present (1) for both organisms

b and c = numbers of characters differing (1,0 or 0,1) between the two organisms

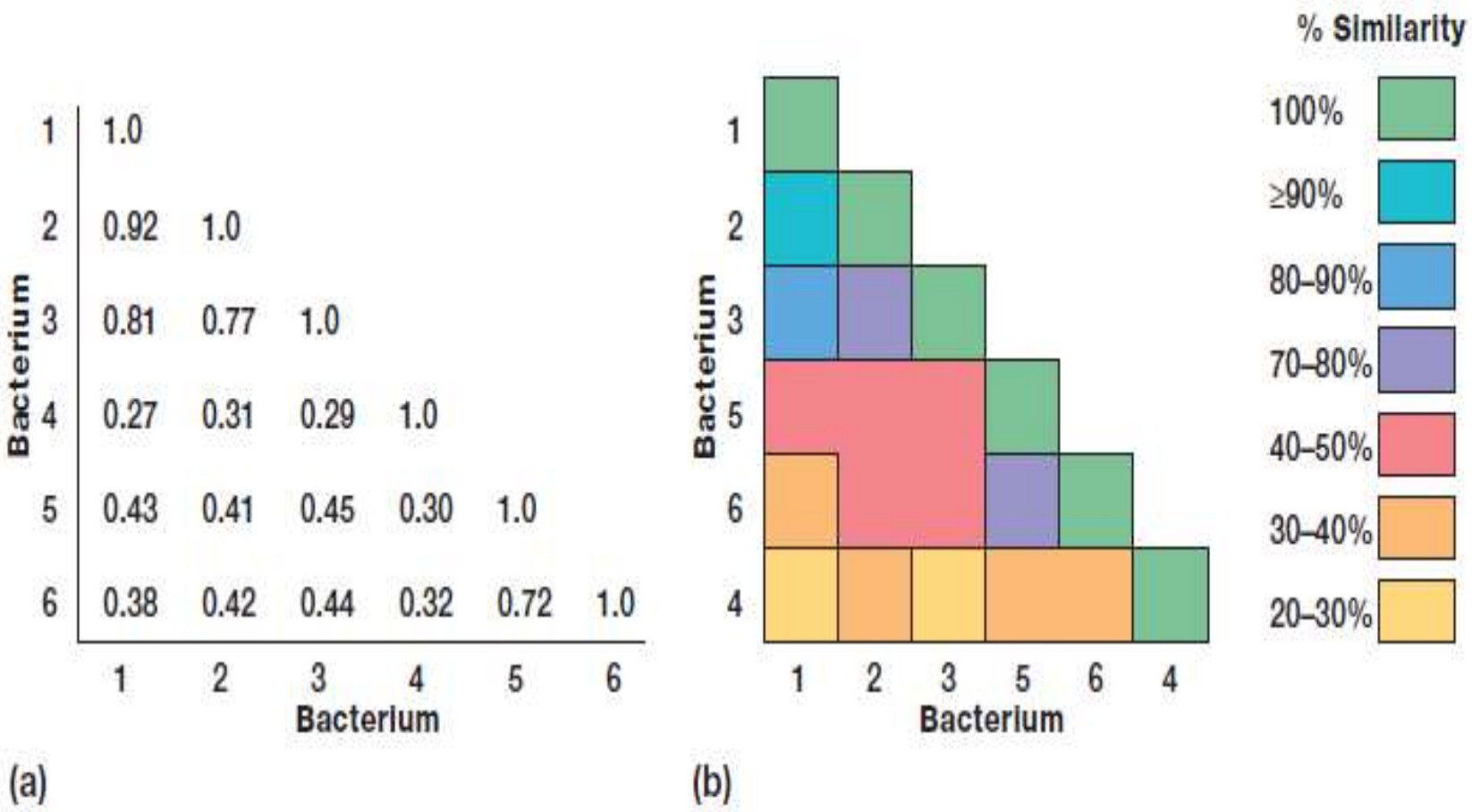
d = number of characters absent (0) in both organisms

Total number of characters compared = $a + b + c + d$

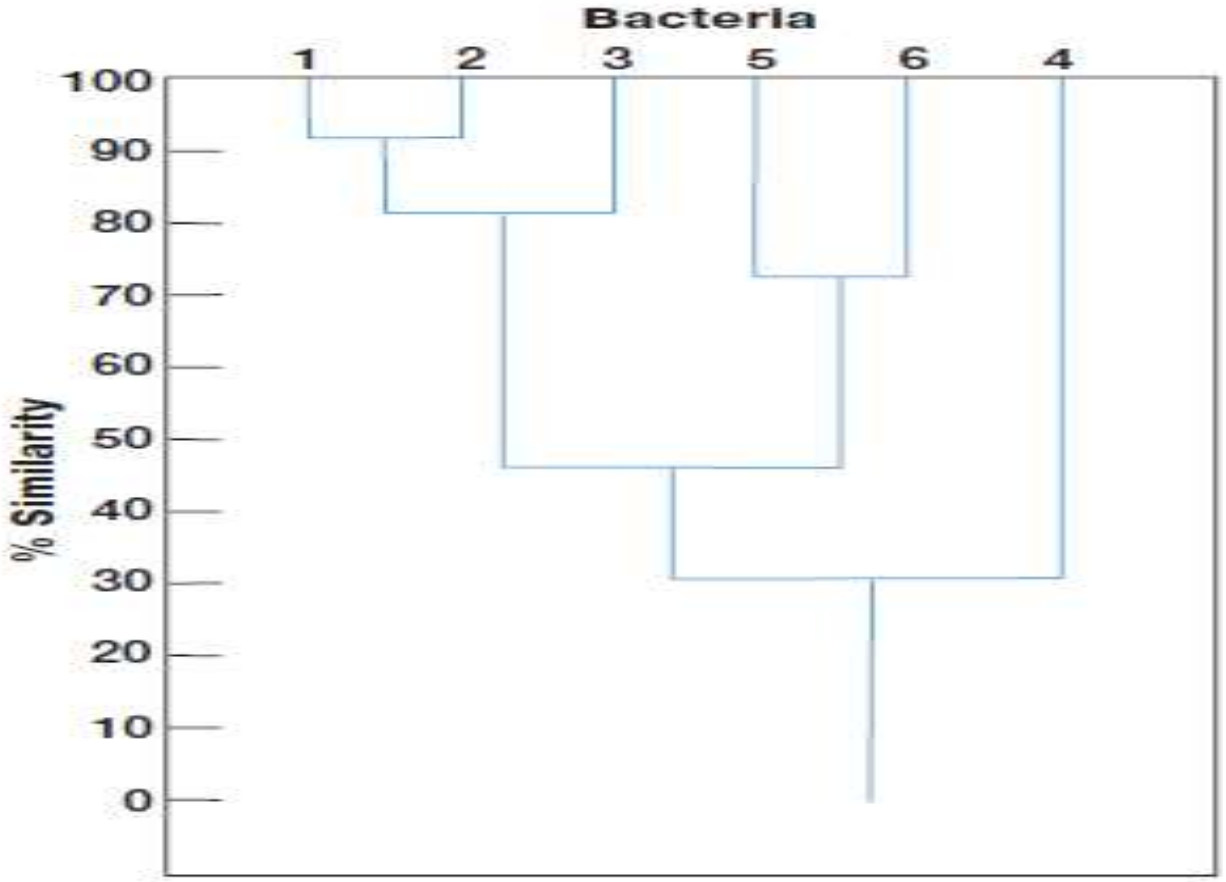
The simple matching coefficient $(S_{SM}) = \dfrac{a + d}{a + b + c + d}$

The Jaccard coefficient $(S_j) = \dfrac{a}{a + b + c}$

- The simple matching coefficients, or other association coefficients, are then arranged to form a **similarity matrix.** This is a matrix in which the rows and columns represent organisms, and each value is an association coefficient measuring the similarity of two different organisms so that each organism is compared to every other one in the table (***α***). Organisms with great similarity are grouped together and separated from dissimilar organisms (figure *b*); such groups of organisms are called **phenons** (sometimes called phenoms).


Bacterium
1 1.0
2 0.92 1.0
3 0.81 0.77 1.0
4 0.27 0.31 0.29 1.0
5 0.43 0.41 0.45 0.30 1.0
6 0.38 0.42 0.44 0.32 0.72 1.0
1 2 3 4 5 6
Bacterium
(a)
Bacterium
1 2 3 5 6 4
Bacterium
(b)
% Similarity
100%
≥90%
80–90%
70–80%
40–50%
30–40%
20–30%

The results of numerical taxonomic analysis are often summarized with a treelike diagram called a **dendrogram.**



Technique used for determining microbial taxonomy and phylogeny

- Classical methods
- Molecular methods

Classical methods

1. Morphological characteristics
2. Physiological characteristics
3. Biochemical characteristics
4. Ecological characteristics
5. Genetic characteristics

1) Morphological characteristics

- Important in microbial taxonomy.
- Easy to study and analyse particularly in eukaryotic microorganism and more complex prokaryotes.
- Morphological similarity is often a good indication of phylogenetic relatedness.
- Morphological characters are used in the classification and identification of microorganism
- Light microscopy is an important tool.
- The transmission and scanning electron has greater importance to visualize all microbes.

Table 19.4 Some Morphological Features Used in Classification and Identification

Feature	Microbial Groups
Cell shape	All major groups[a]
Cell size	All major groups
Colonial morphology	All major groups
Ultrastructural characteristics	All major groups
Staining behavior	Bacteria, some fungi
Cilia and flagella	All major groups
Mechanism of motility	Gliding bacteria, spirochetes
Endospore shape and location	Endospore-forming bacteria
Spore morphology and location	Bacteria, protists, fungi
Cellular inclusions	All major groups
Color	All major groups

[a]Used in classifying and identifying at least some bacteria, fungi, and protists.

2) Physiological and metabolic character

- Are very useful because they are related to the nature and activity of microbial enzymes and transport proteins.
- Proteins are gene product, analysis of these characteristics provides an indirect comparison of microbial genomes.

Table 19.5	Some Physiological and Metabolic Characteristics Used in Classification and Identification
Carbon and nitrogen sources	
Cell wall constituents	
Energy sources	
Fermentation products	
General nutritional type	
Growth temperature optimum and range	
Luminescence	
Mechanisms of energy conversion	
Motility	
Osmotic tolerance	
Oxygen relationships	
pH optimum and growth range	
Photosynthetic pigments	
Salt requirements and tolerance	
Secondary metabolites formed	
Sensitivity to metabolic inhibitors and antibiotics	
Storage inclusions	

3) Ecological characteristics

- Habitat and distribution of organism in nature and interaction between and among species in natural environment.
- Microorganism found in marine environment differ from those found in fresh water.
- The microbial population of oral cavity differ from the intestinal tract various characters are included;
  - Life cycle pattern
  - Nature of symbiotic ralationship
  - Ability to cause in particular host
  - Habitat preferences such as requirement of temperature, pH, oxygen and osmotic concentration.

4) Biochemical characteristics

- Presence or absence of specific enzymes in microorganisms.
- Constituents of the cells such of cell wall lipid or amino acid content, membrane protein content or presence of specific pigments etc.

5) Genetic characteristics

- Characteristics of hereditary material of cell (DNA and RNA).
- The sequence of nucleotide bases in the DNA etc.
- DNA base composition
- Occurrence and function of other kinds of extrachromosal genetic material such as plasmid.
- Chromosomal gene exchange through transformation, conjugation and transduction.
- Transformation is useful for study of *Bacillus, Micrococcus, Haemophilus, Rhizobium.*
- Conjugation studies also produce taxonomically useful data, Particularly with enteric bacteria.eg : *E. coli* can conjugate with *Salmonella* but not with *Proteus* and *Enterobacter*.
- Plasmid are useful for analysis of phenotypic trait.

Molecular methods

- Analysis of nucleic acid (DNA and RNA) and protein.
- Phylogenetic inference based on molecular characteristics provide the most strong and reliable analysis of microbial evolution.
- Molecular method used in classification and identification of microbes are;
  a) Nucleic-acid base composition
  b) Nucleic-acid hybridization
  c) Nucleic-acid sequencing
  d) Genome fingerprinting
  e) Amino acid sequencing

a) Nucleic acid base composition

- DNA base composition is widely used by taxonomist for classification.
- DNA contains four purine and pyrimidine bases; Adenine (A), Guanine (G), Cytosine (C), Thymine (T).
- In double standard DNA, A pairs with T and G pairs with C.

$$\text{Mol\% G + C} = \frac{\text{G + C}}{\text{G + C + A + T}} \times 100$$

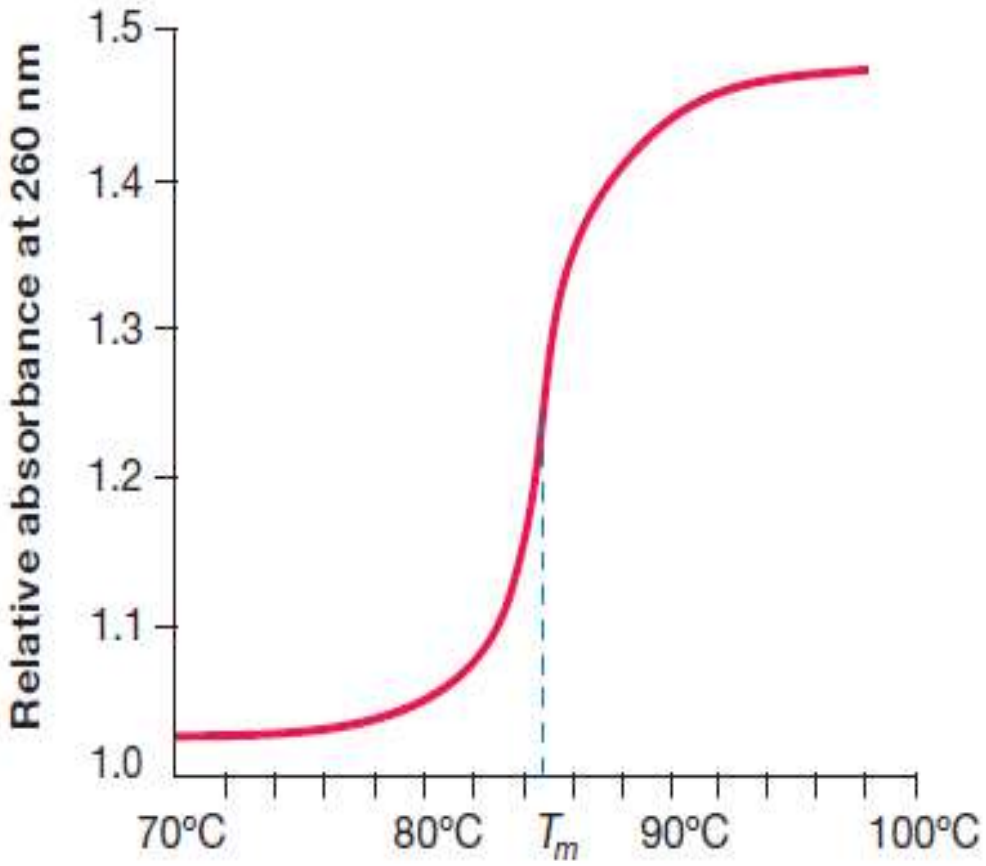
- G+C content can also can be determined by hydrolysis of DNA and analysis of its bases with high performance liquid chromatography.
- G + C content can also be determined from melting temperature (Tm) of DNA. In double standard DNA, three hydrogen bonds joins GC base pair and two bonds joins AT base pairs.

- DNA with greater G+C content will have more hydrogen bond and its strands will separate only at higher temperature. i.e, it will have a higher melting point.
- DNA melting point can be easily followed spectrophotometrically because the absorbance of 260nm UV light by DNA increases during strand separation.
- When a DNA sample is slowly heated, absorbance increases, as hydrogen bonds are broken and reaches a steady state when all DNA has become single standard.

Table 19.6 Representative G + C Content of Microorganisms

Organism	Percent G + C	Organism	Percent G + C
Bacteria		*Rhodospirillum*	62–66
Actinomyces	59–73	*Rickettsia*	29–33
Anabaena	39–44	*Salmonella*	50–53
Bacillus	32–62	*Spirillum*	38
Bacteroides	28–61	*Spirochaeta*	51–65
Bdellovibrio	49.5–51	*Staphylococcus*	30–38
Caulobacter	62–65	*Streptococcus*	33–44
Chlamydia	41–44	*Streptomyces*	69–73
Chlorobium	49–58	*Sulfolobus*	31–37
Chromatium	48–70	*Thermoplasma*	46
Clostridium	21–54	*Thiobacillus*	52–68
Cytophaga	33–42	*Treponema*	25–53

A DNA Melting Curve.

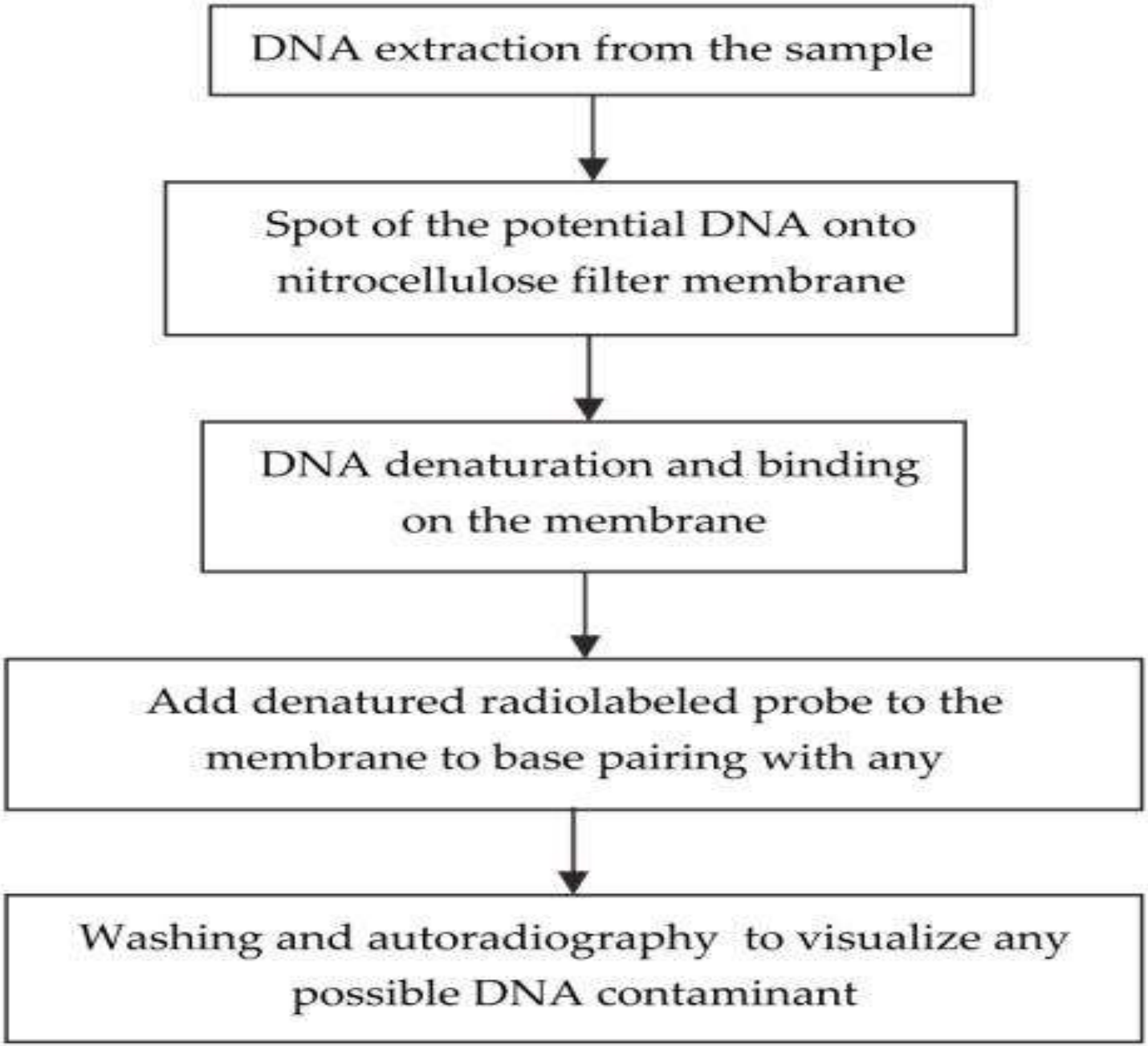


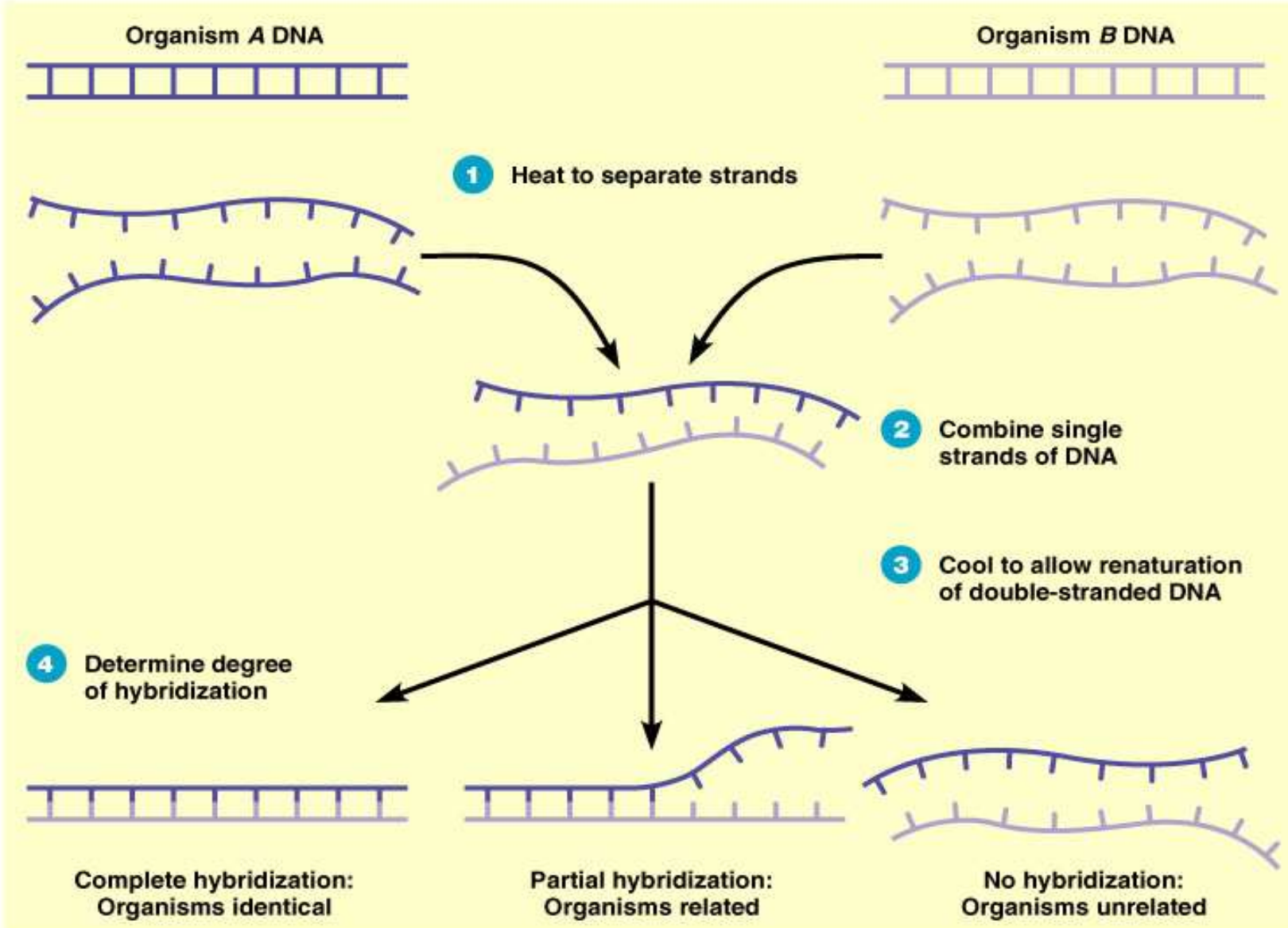
The T_m is indicated.

- G+C content of DNA from animals and higher plants average around 40% and ranges between 30-50%.
- DNA of both eukaryotic and prokaryotic microorganism varies greatly in G+C content.
- Prokaryotes G+C content is the most variable, ranging between 25-80%.
- G+C content of strains within a particular species is constant.
- If two organism differ in their G+C content by more than about 10%, their genome have quite different base sequence.
- But alike phenotyping, similar G+C content suggest dose relatedness.
- G+C content data are important for:
  - To confirm a taxonomic scheme developed by using other data.
  - Useful in characterization of prokaryotic genera.

b) Nucleic acid hybridization

- A G+C ratio measures the percentage of each nucleotide present in an organisms DNA but not measures the information on the nucleotide sequence.
- DNA homology measures the order of nucleotides as well as the composition of genome in an organisms DNA.
- DNA homology provides a tool for classifying microorganisms especially for establishing the affinities of the strain within species and species within genus.


DNA extraction from the sample
Spot of the potential DNA onto nitrocellulose filter membrane
DNA denaturation and binding on the membrane
Add denatured radiolabeled probe to the membrane to base pairing with any
Washing and autoradiography to visualize any possible DNA contaminant



- This includes DNA-DNA homology or DNA base sequence and RNA sequence.
- If a mixture of SSDNA (formed by heating dsDNA) is cooled at hold at a temperature of about 25°C below the Tm, strands with complementary base sequence will reassociate to form stable ds DNA and non complementary strands will remain single.
- Incubation at 10-15°C below the Tm permits hybrid formation with almost identical strands.

DNA-DNA hybridization

- There is no fixed convention as to how much hybridization between two DNA is necessary to assign two organism to the same taxonomic rank.
- Two strains whose DNA show at least 70% relatedness under optimal hybridization conditions and less than 5% difference in Tm often are considered members of same species.
- If DNA molecules are very different in sequence, they will not form a stable detectable hybrid. Therefore DNA-DNA hybridization is used to study only closely related microorganism.

c) Nucleic acid sequencing

- It is the process of determining sequenceof nucleotide in a strand of DNA or RNA.
- rRNA from small ribosomal subunit (16s and 18s rRNA from prokaryotes and eukaryotes respectively) have become the molecule of choice for inferring microbial phylogenies and making taxonomic assignment at genus level.
- Small subunit rRNA (SSU rRNA) are used because
  - They play same role in all microorganism.
  - Ribosome is absolutely necessary for survival and SSU r RNA are part of complex ribosomal structure.

- The genes encoding SSU rRNAs cannot tolerate large mutations. Thus these genes change very slowly with time and do not appear to be subject to horizontal gene transfer, an important factor in comparing sequences from different phyla.
- Presence of certain sequence in SSU rRNAs that are variable among organism and other regions that are quite stable. The variable regions enable comparison of distantly related microorganism.
- Cells do not have to be culture in laboratory which is benefit of this technique.
- PCR is used to amplify genes encoding 16 SSU rRNA from genomic DNA and then sequence that PCR product by standard dideoxy DNA sequencing method. i.e, Sanger sequencing method.

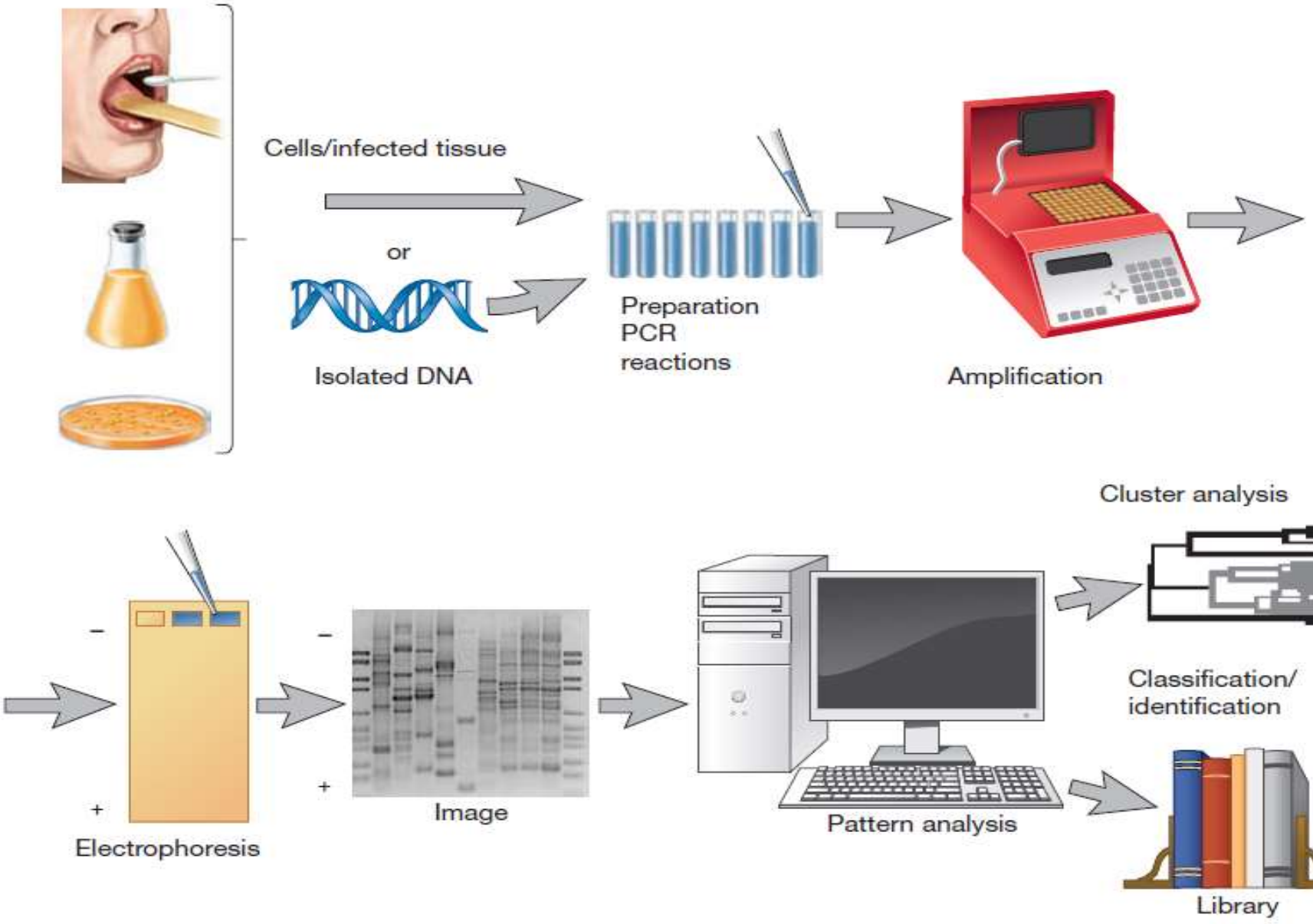
- Procedure is rapid and specific and usually using synthetically produced oligonucleotide primers complementary to conserved sequence in the SSU rRNAs as PCR primers, a tiny amount of DNA from microbial culture can produce a huge amount of PCR product for sequencing reactions.
- Once the sequencing is done, either manually or by automated sequencers, the data are ready for computer analysis.
- Newly generated sequence are compared with sequences in the ribosome database project (RDP). To date RDP has sequenced over 2,00,000 microbes.
- When comparing rRNA sequence between two microorganisms, their relatedness can be represented by an association coefficient or S*ab* value. The higher the S*ab* values, the more closely the organisms are related to each other. If the sequences of the 16SrRNAs of two organisms are identical, the S*ab* value is 1.0.

Steps in Nucleic acid sequencing

1) Isolate the bacterial DNA from culture (16SrRNA genes).
2) Heat to separate strands, add specific primer i.e, rRNA primer.
3) Primer extension with DNA polymerase enzyme.
4) Repeat above steps to obtain many copies of 16SrRNA genes.
5) Run agarose gel electrophoresis and check for correct sized product.
6) Purify and sequence the product.

d) Genomic finger printing

- Fragments of DNA generated by restriction endonuclease digestion will form pattern when subjected to electrophoresis. These pattern are called DNA finger prints or RFLP (restriction fragment length polymorphism pattern).
- It can also be used to classify microbes and help to determine phylogenetic infection.
- DNA fingerprinting is used to determine the source of hospital infection.



An Overview of the Genomic Fingerprinting Technique Based on Repetitive Nucleotide Sequences.

Amino acid sequencing

- Sequencing of proteins directly related mRNA sequence and therefore represents the genes coding for their synthesis.
- The most direct approach is to determine the amino acid sequence of proteins with the same function. The value of a given protein in taxonomic and phylogenetic studies varies. The sequences of proteins with dissimilar functions often change at different rates; some sequences change quite rapidly whereas others are very stable.
- Nevertheless, if the sequences of proteins with the same function are similar, the organisms possessing them may be closely related.

- The sequences of cytochromes and other electron transport proteins, histones and heat-shock proteins, transcription and translation proteins, and a variety of metabolic enzymes have been used in taxonomic and phylogenetic studies.
- A sequence of 20 amino acids has more information per site than a sequence of four nucleotides.
- The most direct approach is to determine the amino acid sequence of proteins with the same function. Because protein sequencing is slow and expensive, more indirect methods of comparing proteins frequently have been employed. The electrophoretic mobility of proteins is useful in studying relationships at the species and subspecies levels.
- Antibodies can discriminate between very similar proteins, and immunologic techniques are used to compare proteins from different microorganisms.

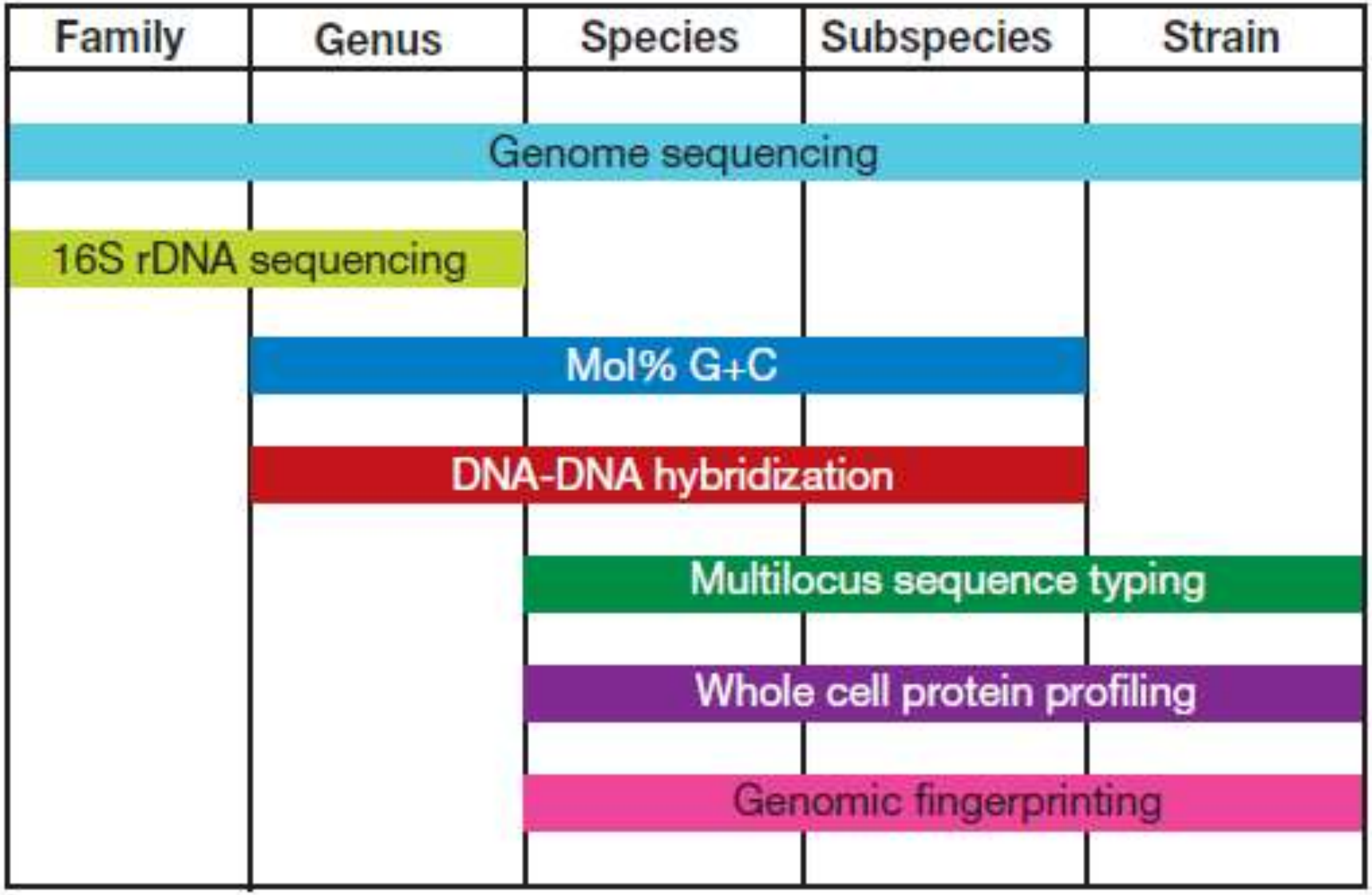


Figure 19.12 Relative Taxonomic Resolution of Various Molecular Techniques.