

# Phylogenetic Analysis of Some Mesophilic and Thermophilic Bacteria Based On 23S rRNA Sequence

---

Arnab Hazra<sup>1</sup>, Ankita Dutta<sup>2</sup>, Dipankar Bera<sup>1</sup>, Sudip Mahato<sup>1</sup>, Agnita Acharya<sup>1</sup>, Nandini Ghosh<sup>1\*</sup>

1) Dept. of Microbiology, Vidyasagar University, Midnapore - 721102

2) Centre for Life Sciences, Vidyasagar University, Midnapore - 721102

\* Corresponding Author: Dr. Nandini Ghosh, Assistant Professor, Department of Microbiology, Vidyasagar University, Midnapore - 721102  
Email: nandini.ghosh14@gmail.com, nandini@mail.vidyasagar.ac.in

---

*The relationship among different groups of bacteria is increasingly supported by the molecular phylogeny and forms the basic platform for understanding of their natural affiliation and subsequent classification. Comparative rRNA sequence analysis is the most widely used approach for constructing the microbial phylogeny. The more or less conserved primary structure and higher-ordered structural elements of rRNA express the evolutionary history of microbes. The prerequisite for a reliable phylogenetic conclusion is the optimal primary structural alignment and a careful data selection. In this study few mesophilic and thermophilic bacteria are selected and their evolutionary relatedness is calculated based on 23S rRNA sequence conservation. Interestingly it is found that some mesophilic bacteria are more closely related to thermophiles on the basis of 23S rRNA sequence homology. This signifies the need of reconstruction of classical microbial classification system from evolutionary point of view using conserved rRNA sequences.*

**Keywords:** 23S rRNA, Extremophiles, Multiple sequence alignment, Phylogenetic tree.

## 1. Introduction

Molecular phylogeny supports the proper understanding of relationships between different organisms and provides the platform for classification of all microorganisms. Bacterial identification and its phylogenetic position are elucidated by rRNA sequence analysis. Microbial evolution is supported by conserved primary and advanced order structural elements of 16S rRNA. ATP synthase  $\beta$  subunits, elongation factors etc. are also used as phylogenetic markers (Cowan, 1968). Now-a-days comparison of 23S rRNA sequence is also used as an efficient tool to reveal the evolutionary history and to establish phylogenetic relationship among different genera (Cliffs and Woese, 1993; Ludwig and Schleifer, 2017).

Contemporary organisms are the products of historical events during the course of evolution and their structures at all levels reflect their evolutionary history. In microorganisms, prehistoric documents of earlier evolutionary events are hard to be found at the phenotypic level. It has almost taken 100 years to finally unleash the phylogenetic relationship of prokaryotes (Cowan, 1968). The conclusive construction of the bacterial phylogenies is allowed by comparatively studying the phylogenetic marker molecules and by comparing the phenotypic level of evolution, a higher number of definable characters at the molecular and genotypic level can be obtained. The underlying genetic information for every definable phenotypic function comprises tens to thousands of evolutionary independent sequence residues. Size is an important factor regarding the usefulness of the molecules for the reconstruction of major phylogenies. Since every individual sequence position can only carry the information on a rather narrow range of evolutionary time, an increasing number of independently evolving positions or regions augments the number of phylogenetic levels which probably can be detected. Another most important advantage of large sized molecules like 16S and 23S rRNA over smaller molecules like 5S rRNA is that the localized non-random rearrangement does not tend to disturb the complete picture that is derived from comparative analysis (Trifiper and Achleifer, 1991; Zuckerandl and Pailing, 2005).

Selected genera for this study belong to Mesophiles (according to their optimum temperature of growth)

and extremophiles (according to their tolerance of extreme environmental conditions such as - extreme temperature, acidity, alkalinity or chemical concentration). RNA molecules conferring thermal adaptations are also known. The G+C content of rRNA and tRNA of thermophilic bacteria is higher than mesophiles, the base pairs of G+C forms more hydrogen bonds than that of A+T base pairs and higher the G+C contents in double-stranded stem region more is the thermostability of the RNA molecules. In this study the phylogenetic relationship between some selected mesophiles and thermophiles are established by comparing their 23S rRNA sequence.

## 2. Materials and Method

### 2.1. Target Organisms:

Five mesophilic and five thermophilic bacteria were selected. Among the selected mesophiles most of the bacteria belong to Class III  $\gamma$  proteobacteria and most of the thermophiles belong to high G+C content in Bergey's Manual of Systematic Bacteriology; selected mesophiles and thermophiles are mentioned in Table 1 and 2.

### 2.2. Search of 23S rRNA sequences of the selected mesophiles and thermophiles:

23S rRNA sequence data of the selected mesophiles and thermophiles or extremophiles were collected from NCBI (National Centre for Biotechnology Information) and their accession numbers were mentioned in Table 1 and 2.

**Table 1: Target mesophilic bacteria and respective 23s rRNA accession number**

Sl. No.	Name of the bacteria	23S rRNA accession no.
1	<i>Escherichia coli</i>	>NR_076322.1
2	<i>Staphylococcus aureus</i> subsp. aureus strain N315	>NR_076325.1
3	<i>Salmonella enterica</i> subsp. Entericasero var Typhi strain Ty2	>NR_076191.1
4	<i>Vibrio cholerae</i> O1 biovar El Tor strain N16961	>NR_076165.1
5	<i>Pseudomonas aeruginosa</i> strain PAO1	>NR_076166.1

**Table 2: Target thermophilic bacteria and respective 23S rRNA accession number**

Sl. No.	Name of the bacteria	23S rRNA accession no.
1	<i>Deinococcus radiodurans</i> strain R1	>NR_076078.2
2	<i>Thermus thermophilus</i> strain HB27	>NR_076211.1
3	<i>Thermus scotoductus</i> strain SA-01	>NR_076136.1
4	<i>Chloroflexus aurantiacus</i> strain J-10-fl	>NR_076553.1
5	<i>Spirochaeta africana</i> strain Z-7692	>NR_103194.1

### 2.3 Prediction of secondary structure of 23S rRNA:

Secondary structures of selected 23S rRNA sequences were predicted using RNAfold webserver (<http://rna.tbi.univie.ac.at/>). Energy minimization was done. The centroid structure was built (Lorenz et al., 2011).

### 2.4 Multiple Sequence Alignment (MSA):

To understand the sequence similarities between different strains or different species multiple sequence alignment is used. In this study the 23S rRNA sequences of the selected mesophiles and thermophiles were aligned in Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo>).

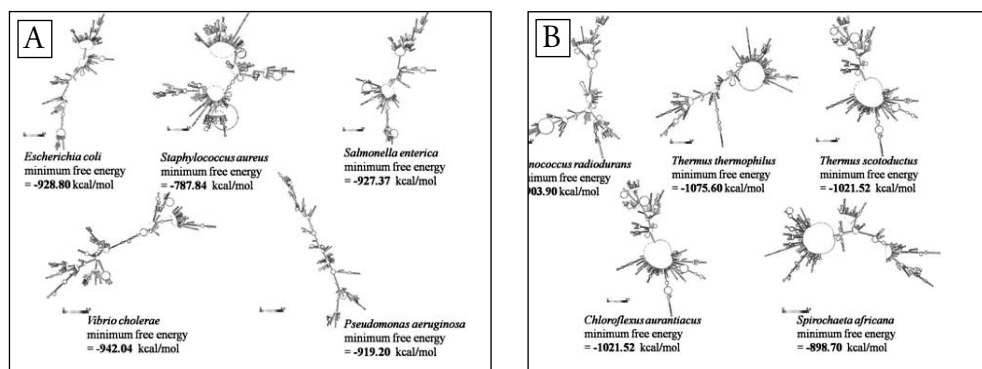
### 2.5 Phylogenetic tree construction:

Alignments were also used to establish evolutionary relationship and construction of the phylogenetic tree. The sequences were used for constructing the tree by using MUSCLE alignment software. Phylogenetic tree was constructed by maximum parsimony method using 1000 bootstrap values. *Acorus calamus* plastid genome (>NC\_007407.1) was taken as outlayer.

## 3. Results and Discussion

### 3.1 Variation in the secondary structure of 23S rRNA:

The secondary structures of 23S rRNA of the selected mesophiles and thermophiles were predicted. There was ample difference between the folding patterns so as to overall structure of the rRNA molecules of selected micro-organisms when viewed from the same angle. The predicted secondary structures of the rRNA molecules were depicted in figure 1A & B.



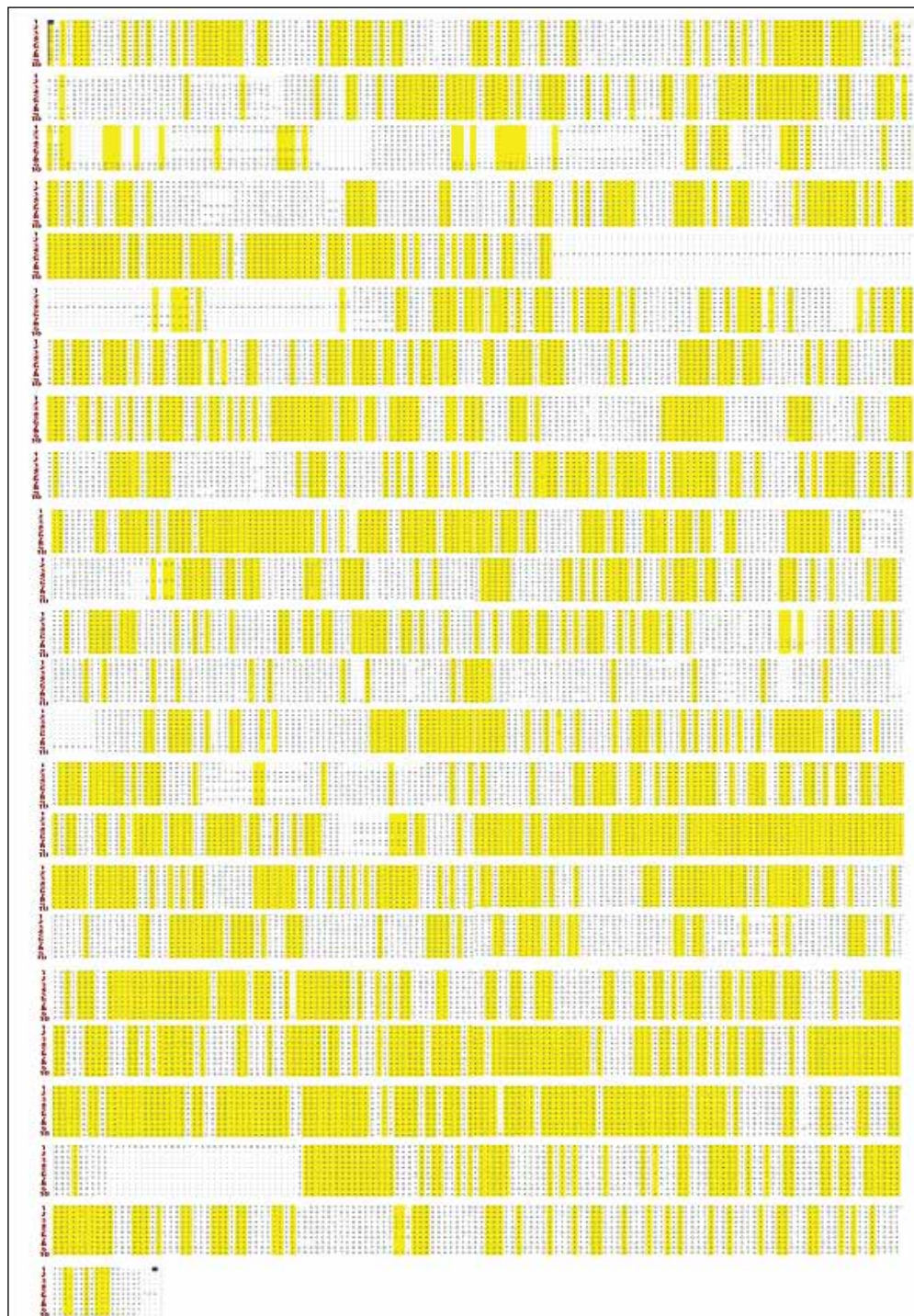
**Figure 1:** Predicted centroid secondary structures of selected mesophilic (A) and thermophilic bacteria (B). The scale represents the base pair probabilities.

### 3.2 Sequence conservation between different bacterial groups:

Among the selected microorganisms 66%-98% sequence conservation was observed. Identity matrix was represented in table 3. Maximum sequence similarity was observed between *E. coli* and *Salmonella enterica*, whereas minimum sequence homology was found in between *Chloroflexus aurantiacus* and *Spirochaeta africana* although both of them were thermophiles. Interestingly 95.75% sequence homology was present within two different species of *Thermus*, whether 97.93% sequence homology was present between *E. coli* and *Salmonella enterica*. About 80%-90% sequence conservation was observed in the multiple sequence alignment between 2500-2700 nucleotide of the selected sequences. Sequence alignment was represented as fig. 2.

**Table 3:** Sequence identity matrix (in terms of % of sequence similarity) of selected mesophilic and thermophilic bacteria

	<i>E. coli</i>	<i>S. aureus</i>	<i>S. enterica</i>	<i>V. cholerae</i>	<i>P. aeruginosa</i>	<i>D. radiodurans</i>	<i>T. thermophilus</i>	<i>T. scotoductus</i>	<i>C. aurantiacus</i>	<i>S. africana</i>
<i>Escherichia coli</i>	100									
<i>Staphylococcus aureus</i>	72.83	100								
<i>Salmonella enterica</i>	97.93	72.55	100							
<i>Vibrio cholerae</i>	88.17	73.13	87.77	100						
<i>Pseudomonas aeruginosa</i>	84.91	74.32	84.79	84.98	100					
<i>Deinococcus radiodurans</i>	70.19	73.18	70.27	70.27	71.65	100				
<i>Thermus thermophilus</i>	70.43	73.03	70.50	70.50	72.33	77.94	100			
<i>Thermus scotoductus</i>	70.57	73.39	70.79	70.79	72.40	77.81	95.75	100		
<i>Chloroflexus aurantiacus</i>	67.92	68.78	67.93	67.57	68.22	71.07	71.23	70.60	100	
<i>Spirochaeta africana</i>	73.76	74.81	73.55	73.73	74.91	70.47	71.72	71.54	66.54	100



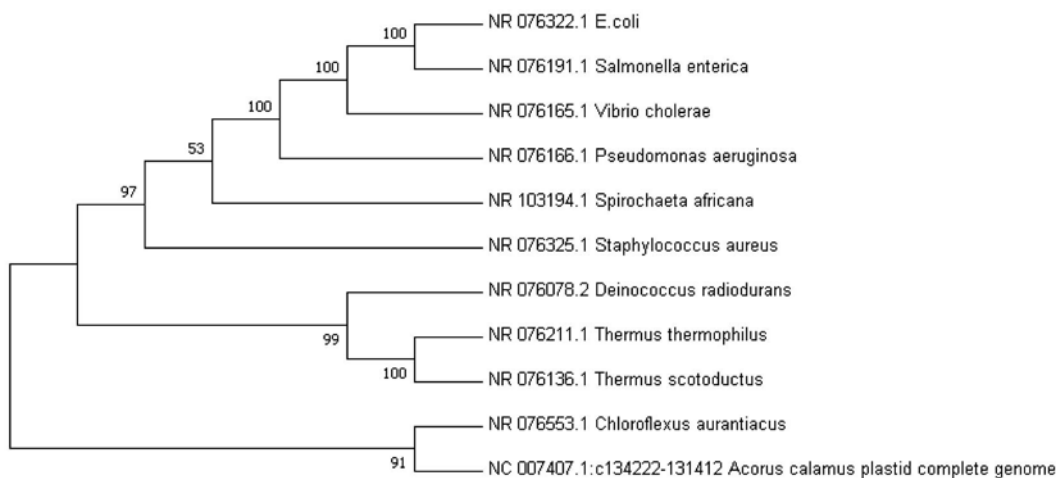
**Figure 2:** Multiple sequence alignment of the 23S rRNA sequence of selected bacteria. Conserved regions are highlighted in yellow colour. In this alignment the microorganisms were represented by numbers that are listed below -1: *Deinococcus radiodurans*, 2: *Thermus scotundus*, 3: *Vibrio cholerae*, 4: *Pseudomonas aeruginosa*, 5: *Salmonella enterica*, 6: *Thermus thermophilus*, 7: *E. coli*, 8: *Staphylococcus aureus*, 9: *Chloroflexus aurantiacus*, 10: *Spirochaeta africana*



### 3.3 Evolutionary relationship between selected taxa:

The phylogenetic tree was constructed based on 23S rRNA sequence of selected taxa. The 23S rRNA tree has two clusters or the tip species, mesophiles were present in cluster 1 and the thermophiles were present in cluster 2. The phylogenetic tree represents the evolutionary relationship of the extremophiles with a common or shared ancestry with the mesophiles. The closely related strains were placed closer to each other compared to the distantly related ones. The branches were also spaced proportional to the accounted changes or the evolutionary time; this helps us to draw the inference regarding the species those had changed over time. Grouping of species were also accomplished based on the similarities in their genetic set up. In the above mentioned phylogenetic tree, Cluster 1 and 2 or more specifically the two separate Clades represent their ancestor and the descendants from the ancestor. From the tree thus can be concluded that the mesophiles and the thermophiles originated and then diverged from a common ancestor which is known as LUCA, Last Universal Common Ancestor.

From the tree it was evident that *Spirochaeta africana* had diverged from the same ancestor with that of mesophiles rather than thermophiles. Again, *Chloroflexus aurantiacus* formed a completely separate clade from the others and was much more closely related with the plastid sequence of a green plant *Acorus calamus*. It can be concluded that rapid speciation has taken place in the evolutionary history of the selected taxa.



**Figure 3:** Phylogenetic tree constructed based on 23S rRNA showing the relationships of the mesophilic and thermophilic bacteria.

### 4. Conclusion

Rapid evolution of methods of classification of organisms based on their molecular phylogeny has formed the basis of this study. 16S rRNA is now well accepted as a tool to establish phylogenetic relationship among different taxa. Now-a-days along with this, researchers are interested about some other molecular markers for strengthening the phylogenetic relationship. Different methods of sequence alignment, phylogeny reconstruction, error check and signature analysis are currently being developed. From this study it can be concluded that mesophiles and thermophiles share common ancestor and they have diverged during the course of evolution. Some of the thermophiles are even more closely related to mesophiles than other thermophiles. So there is a need to revise the classical system of microbial classification in light of their evolutionary relationship using modern day tools.

### **Acknowledgement:**

All the authors are grateful to the Vice Chancellor of Vidyasagar University for providing infrastructural facility to carry out this study.

### **References**

- Cowan, S.T. (1968). A dictionary of microbial taxonomic usage. Oliver & Boyd, Edinburgh.
- Cliffs, G.J. and Woese, C.R. (1993). Ribosomal RNA: a key to phylogeny. In: Stanier, R.Y., Ingraham, M.L., Wheelis, M.L. and Pains, P.R. *The Microbial World*, FASEB J. 7, 113-123.
- Ludwig, W. and Schleifer, K.H. (2017). Bacterial phylogeny based on 16S and sequence analysis 23S rRNA. *Federation of European Microbiological Societies* 21, D-80290.
- Trifper, H.G. and Schleifer, K.H., 1991. Prokaryote characterization and identification. In: Balows, A. Triiper, H.G., Dworkin, M., Harder, W., Schleifer, K.H. (Eds.), *The Prokaryotes*, second ed., Springer New York, 1, 126-148.
- Zuckerandl, E. and Pauling, L. (2005). Molecules as documents of evolutionary history. *J. Theor. Biol.* 8, 357-366.
- Lorenz, R., Bernhart, S. H., Zu Siederdisen, C. H., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). Vienna RNA Package 2.0. *Algorithms for molecular biology*, 6(1), 26.