

# MSA and Phylogenetic Analysis of ABC Transporter from Bacterial Species Causing Apical Periodontitis

**Sanu Narayanan**

*Postgraduate Student, Department of Conservative Dentistry and Endodontics  
Saveetha Dental College, Saveetha University, Chennai, India.*

**Dr. Sindhu Ramesh**

*Professor and Head (Admin), Department of Conservative Dentistry and Endodontics  
Saveetha Dental College, Saveetha University, Chennai, India.*

**Dr. Vishnu Priya V**

*Professor, Department of Biochemistry,  
Saveetha Dental College, Saveetha University, Chennai, India.*

**Dr. Gayathri R**

*Assistant Professor, Department of Biochemistry,  
Saveetha Dental College, Saveetha University, Chennai, India.*

---

## Abstract

Phylogenetic analysis is an integral part of bioinformatics as it determines evolutionary relationship between molecules or organisms. The aim of this study is to determine phylogenetic analysis of 5 bacteria ( *Campylobacter jejuni* , *Fusobacteriumnucleatum* , *Streptococcus mutants* , *Treponema denticola* and *Porphyromonasgingivalis* ) which are found in apical periodontitis using sequences of their ABC transporters. After obtaining the sequences from the database , multiple sequence analysis was done using Clustal Omega program and this is used to generate a phylogenetic tree using Mega 5 software. The phylogenetic tree shows that *Campylobacterjejuni* and *Fusobacteriumnucleatum*, and *Streptococcus mutans* and *Treponema denticola* evolved from a common ancestor where as *Porphyromonasgingivalis* underwent no changes during the course of evolution.

**Keywords:** Phylogenetic analysis, Apical periodontitis, ABC transporters, bacterial species

---

## INTRODUCTION

Apical periodontitis is the inflammatory condition of periodontal tissues apical to a tooth caused by microorganism (mainly bacteria) invading the periapical area from a necrotic infected root canal<sup>1,2</sup>. Necrosis of pulp can occur as a sequelae to caries , trauma or due to iatrogenic causes<sup>3</sup>. The microbiota of endodontic infection is mixed and as the pulp becomes necrotic, the ecology shifts and are dominated by anaerobic microorganism<sup>4</sup>. The pathogenesis of apical periodontitis is complex and involves host-tooth factors and their interrelationship.

ATP-binding cassette transporters (ABC transporters) are members of a transport system present in prokaryotes and eukaryotes and often consist of multiple subunits, one or two of which are transmembrane proteins and one or two of which are membrane-associated ATPases<sup>5,6,7</sup>. The energy released during ATP binding and hydrolysis is used by ATPase for translocation of various substrates across membranes, either for uptake or for export of the substrate<sup>8</sup>. ABC transporters are important for the pathogenic bacteria to maintain their cell viability, virulence, and pathogenicity<sup>7</sup>. The ABC transporters helps the bacteria to overcome any undesirable change occurring in the cell and plays an important role in maintaining cell viability<sup>9</sup>. Other than functioning in transport, some

bacterial ABC proteins are also involved in the regulation of several physiological processes<sup>7</sup>.

Phylogenetic analysis is a way of inferring or estimating the evolutionary relationships among molecules (“gene trees”), organisms, or both using branching, treelike diagrams that represent an estimated pedigree of the inherited relationships<sup>10</sup>.

The aim of this study is to determine the phylogenetic analysis of 5 bacterial species seen in apical periodontitis –*Campylobacter jejuni* , *Fusobacteriumnucleatum* , *Streptococcus mutants* , *Treponema denticola* and *Porphyromonasgingivalis*.

## METHODOLOGY

- 1) Collection of ABC transporter sequence  
Amino acid sequence of ABC transporter of the 5 bacterial species (*Campylobacter jejuni* , *Fusobacteriumnucleatum* , *Streptococcus mutants* , *Treponema denticola* and *Porphyromonasgingivalis*) was obtained from NCBI (National Centre for Biotechnology Information) database.
- 2) Multiple sequence alignment  
After obtaining the ABC transporter sequence of these 5 bacterial species , the sequences are

subjected to alignment. Since alignment has to be done for 5 bacterial species, multiple sequence alignment was done using Clustal Omega program.

### 3) Phylogenetic analysis

The phylogenetic analysis was done using the software MEGA5. The phylogenetic analysis as a whole can be classified into distance based and character based analysis. However, MSA is the raw material for both. The various types of grouping under distance based are UPGMA, neighbor – joining method and Fisch Margoliash method. In this study, a phylogenetic tree was constructed and accordingly the phylogenetic relationship was analysed.

## RESULTS

The phylogenetic tree generated through the MEGAS software has been studied. Figure 1 shows the phylogenetic tree representing the evolutionary relationship between the various species such as *Campylobacter jejuni*, *Fusobacterium nucleatum*, *Streptococcus mutans*, *Trypanema denticola* and *Porphyromonas gingivalis* using the ABC transporter sequence.

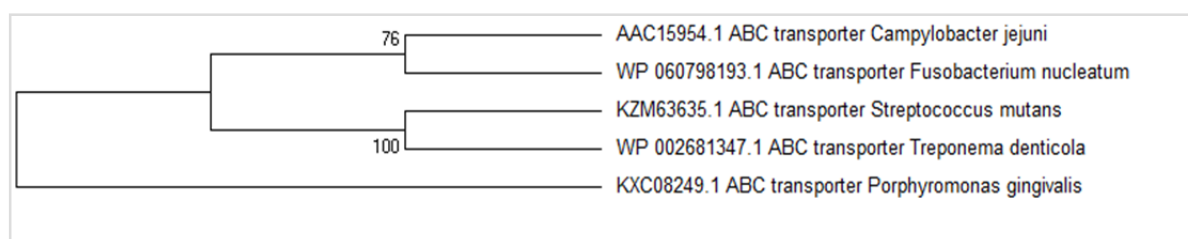
## DISCUSSION

Multiple sequence alignments (MSA) are an essential and widely used computational procedure for biological sequence analysis in molecular biology, computational biology, and bioinformatics. MSA are completed where homologous sequences are compared in order to perform phylogenetic reconstruction, protein secondary and tertiary structure analysis, and protein function prediction analysis<sup>12</sup>. Different methods for producing multiple sequence alignment exist, and their use depends on user preferences and sequence length and type<sup>13</sup>.

The most popular heuristic used from which the majority of multiple sequence alignments are generated is that developed by Feng and Doolittle, which they referred

to as “progressive alignment”<sup>14,15</sup>. Progressive alignment works by building the full alignment progressively, firstly completing pairwise alignments using methods such as the Needleman-Wunsch algorithm, Smith-Waterman algorithm, k-tuple algorithm, or k-mer algorithm, and then the sequences are clustered together to show the relationship between them using methods such as mBed and k-means<sup>16,17,18</sup>. Similarity scores are normally converted to distance scores and guide trees are constructed using these scores by guide tree building methods such as Neighbour-Joining (NJ) and Unweighted Pair Group Method with Arithmetic Mean UPGMA<sup>19,20</sup>. Once the guide tree is built, the multiple sequence alignment is assembled by adding sequences to the alignment one by one according to the guide tree, that is, the most similar sequences added first and then gradually adding more distant sequences. In this study, progressive alignment was done using Clustal Omega to produce multiple sequence alignment<sup>13</sup>.

Clustal Omega is a package for performing fast and accurate multiple sequence alignments (MSAs) of potentially large numbers of protein or DNA/RNA sequences<sup>21</sup>. It is the latest version of the popular and widely used Clustal MSA package<sup>22,23</sup>. Clustal Omega retains the basic progressive alignment MSA approach of the older ClustalX and ClustalW implementations, where the order of alignments is determined by a so called guide-tree, which in turn is constructed from pairwise distances amongst the sequences. The main improvements over ClustalW2 are use of the mBed algorithm for creating guide trees of any size and a very accurate profile-profile aligner, based on the HAlign package<sup>24,25</sup>. Sequence input to Clustal Omega is handled by the Squid routines (<http://selab.janelia.org/software.html>), and permissible input formats are a2m (fasta/vienna), clustal, msf, phylip, selex and stockholm. Output can be in the same formats.



**Phylogenetic tree**

TABLE I: Types of multiple sequence alignment and corresponding algorithms.

Types of MSA alignment	MSA algorithms
Pairwise alignment	Needleman-Wunsch, k-mer, k-tuple, and Smith-Waterman algorithms.
Progressive alignment	Clustal Omega, ClustalW, MAFFT, Kalign, Proalign, MUSCLE, Dialign, ProbCons, and MSAProbs.
Iterative progressive alignment	PRRP, MUSCLE, DIALIGN, SAGA, and T-COFFEE.
Homology search tools	BLAST, PSI-BLAST, and FASTA.
Structure incorporating alignment	3D-COFFEE, EXPRESSO, and MICAlign.
Motif alignment	PHI-Blast, GLAM2.
Short-read alignment	Bowtie, Maq, and SOAP.

Obtaining the sequence alignment is the first step of phylogenetic analysis which is the inferring of evolutionary relationship between molecules, organisms or both. Tree building from these sequences may be either distance based or character based. In this study character based method is been used. The character based methods allows the assessment of the reliability of each base position in an alignment on the basis of all other base positions<sup>10</sup>. The most common character-based methods include maximum parsimony and maximum likelihood. Maximum parsimony tree is the shortest—the one with the fewest changes—which, by definition, is also the one with the fewest parallel changes.

The Molecular Evolutionary Genetics Analysis (MEGA) software was developed with the goal of providing a biologist centric, integrated suite of tools for statistical analyses of DNA and protein sequence data from an evolutionary standpoint. Over the years, it has grown to include tools for sequence alignment, phylogenetic tree reconstruction and visualization, testing an array of evolutionary hypotheses, estimating sequence divergences, web-based acquisition of sequence data, and expert systems to generate natural language descriptions of the analysis methods and data chosen by the user<sup>26</sup>. In this study, MEGA5 software is used to generate the phylogenetic tree.

#### CLINICAL IMPLICATIONS

Closely evolved microorganisms may exhibit similar characters like drug resistance, and mode of inhibition.

#### CONCLUSION

*Campylobacter jejuni*, *Fusobacterium nucleatum*, *Streptococcus mutans*, *Treponema denticola* and *Porphyromonas gingivalis* are commonly seen in apical periodontitis and as ABC transporters are important in survival, virulence and drug resistance of these microorganisms, these sequences are compared to determine the phylogeny.

There is strong evidence that *Treponema denticola* and *Streptococcus mutans* are closely related. *Campylobacter jejuni* and *Fusobacterium nucleatum* also evolved from a common ancestor. *Porphyromonas gingivalis* evolved with least changes when compared to other species.

#### REFERENCES

- 1) Pashley DH. Dynamics of the pulpo-dentin complex. Review. Crit Rev Oral Biol Med 1996; 7: 104–133.
- 2) Kakehashi S, Stanley HR, Fitzgerald RJ. The effects of surgical exposures of dental pulps in germfree and conventional laboratory rats. J Southern Calif Dent Assoc 1966; 34: 449–451.
- 3) P. V. Abbott and C. Yu, "A clinical classification of the status of the pulp and the root canal system," Australian Dental Journal, vol. 52, supplement 1, pp. S17–S31, 2007.
- 4) Siqueira JF, Roca IN: Diversity of endodontic microbiota revisited. J Dent Res. 2009, 88 (11): 969-981.
- 5) Jones PM, George AM. "The ABC transporter structure and mechanism: perspectives on recent research". Cellular and Molecular Life Sciences. 2004,61 (6): 682–99.
- 6) Ponte-Sucre A, ed.ABC Transporters in Microorganisms. Caister Academic Press. 2009.
- 7) Davidson AL, Dassa E, Orelle C, Chen J. "Structure, function, and evolution of bacterial ATP-binding cassette systems". Microbiology and Molecular Biology Reviews.2008, 72 (2): 317–64.
- 8) Goffeau A, de Hertogh B, Baret PV. "ABC Transporters". In Lane WJ, Lennarz MD. Encyclopedia of Biological Chemistry (Second ed.) 2003.
- 9) Poolman B, Spitzer JJ, Wood JM. "Bacterial osmosensing: roles of membrane structure and electrostatics in lipid-protein and protein-protein interactions". Biochimica et BiophysicaActa. 2004, 1666 (1-2): 88–10.
- 10) Brinkman, F. S., and D. D. Leipe. Phylogenetic analysis. Methods Biochem. Anal. 2001, 43:323–358.
- 11) Daugelaite J, O'Driscoll A, Sleator RD. An overview of multiple sequence alignments and cloud computing in bioinformatics, ISRN Biomathematics; 2013. p. 14.
- 12) C. Kemena and C. Notredame, "Upcoming challenges for multiple sequence alignment methods in the high-throughput era," Bioinformatics, 2009 vol. 25, no. 19, pp. 2455–2465.
- 13) J. Daugelaite, A. O'Driscoll, and R. D. Sleator, "An overview of multiple sequence alignments and cloud computing in bioinformatics," ISRN Biomathematics, vol. 2013.
- 14) I. M. Wallace, G. Blackshields, and D. G. Higgins, "Multiple sequence alignments," Current Opinion in Structural Biology, 2005, vol. 15, no. 3, pp. 261–266.
- 15) D.-F. Feng and R. F. Doolittle, "Progressive sequence alignment as a prerequisite to correct phylogenetic trees," Journal of Molecular Evolution, 1987, vol. 25, no. 4, pp. 351–360.
- 16) W. J. Wilbur and D. J. Lipman, "Rapid similarity searches of nucleic acid and protein data banks," Proceedings of the National Academy of Sciences of the United States of America, 1983, vol. 80, no. 3, pp. 726–730.
- 17) R.C.Edgar, "MUSCLE:multiplesquencealignmentwithhigh accuracy and high throughput," Nucleic Acids Research, 2004, vol. 32, no. 5, pp. 1792–1797.
- 18) F. Sievers, A. Wilm, D. Dineen, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," Molecular Systems Biology, 2011, vol. 7, article 539.
- 19) N. Saitou and M. Nei, "neighbor-joining method: a new method for reconstructing phylogenetic trees," Molecular Biology and Evolution, 1987, vol. 4, no. 4, pp. 406–425.
- 20) I. Gronau and S. Moran. "Optimal implementations of UPGMA and other common clustering algorithms," Information Processing Letters, 2007, vol. 104, no. 6, pp. 205–210, 2007.
- 21) Sievers F, Wilm A, Dineen D. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011, 7:539.
- 22) Higgins DG, Sharp PM. Clustal : a package for performing multiple sequence alignment on a microcomputer. Gene 1988, 73(1): 237–244
- 23) Larkin MA, Blackshields G, Brown NP. Clustal W and Clustal X version 2.0. Bioinformatics 2007, 23:2947–2948.
- 24) Blackshields G, Sievers F, Shi W. Sequence embedding for fast construction of guide trees for multiple sequence alignment. Algorithms Mol Biol 2011, 5:21.
- 25) So ding J. Protein homology detection by HMM-HMM comparison. Bioinformatics. 2005, 21:951–960.
- 26) Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011, 28:2731–2739.