

Reenacting of Phylogenetic Tree for Cyclooxygenase DNA Sequences by using MUSCLE

Chukka Santhaiah* and A. Rama Mohan Reddy

C. S. E Department, S. V. University College of Engineering, S. V. University, Tirupati - 517502, Andhra Pradesh, India;
chukka.santhsvu006@gmail.com, ramamohansvu@yahoo.com

Abstract

Background/Objectives: Multiple Sequence alignment based on Log Expectation (MUSCLE) is a new approach for multiple sequence alignments of DNA or Protein sequences. The MUSCLE shows the high accuracy than the Clustal W and MAFFT and T-Coffee approaches. Based on MUSCLE approach, we aim to construct the Phylogenetic tree for Cyclooxygenase sequences. **Method/ Statistical Analysis:** In MUSCLE there is a tree steps of process to find the accuracy of alignment i.e. are draft progressive, improved progressive and refinement. In each step it has the another sub steps to complete the alignment. **Findings:** MUSCLE is the faster algorithm than the ClustalW by adding the additional options for input sequences. It executes more sequences than the ClustalW and MAFFT. The MUSCLE displays the results with accuracy values of every input sequence. **Application/Improvement:** MUSCLE gives the Phylogenetic tree with very closeness of the sequences of input sequences alignment.

Keywords: DNA, Hippocampus, MUSCLE, Phylogenetic Tree

1. Introduction

Phylogenetic is the cram of development affairs linking organisms. Information has allowed spectacular evolution in phylogenetics. One part of this within the biology itself. The innovation of DNA and the capability for biologists to sequences of DNA has the slightest revolutionized the sequences¹. Phylogenetics passes directly to the evolutionary circle of relatives together with numerous organisms. A few confirmations illustrate that several organisms distribute analogous characteristics, on the equal time as starting the same antecedent and other characteristics are separate to help organisms to become acquainted with their specific surroundings groups. So the analyzing of phylogeny in conjunction with one of kind organisms locate the similarities and differences amongst exclusive species^{2,3}.

The phylogenetic tree is also termed as evolutionary tree symbolizes an evolutionary relations or phylogeny amongst species, and plays essential roles in organic research. The leaf nodes stand for species and inner

nodes are revealed as the common ancestor of its child nodes⁴. The extent of a DNA connecting one node to any other indicates evolution time. The building of evolutionary trees considers genomic sequences. The genomic sequences are generally long; therefore, examining genomic sequences from corner to corner species for building phylogenetic trees is computationally high cost. The hippocampus is a mind association that performs a big obligation in advantage knowledge and memory in people and animals. Huge electrophysiological research has validated the existence of activity-established synaptic plasticity including Long Term Potentiation/Long-Time period Despair (LTP/LTD) in hippocampal pathways^{5,6}. Recently, genetic analyses have certainly confirmed the crucial role of the hippocampal synaptic plasticity in memories formation^{7,8}. Despite the significance of the hippocampus in getting to know and memories, our knowledge of the genetic applications underlying the developing hippocampus is quite limited.

Prostaglandin H synthase, or else called Cyclooxygenase (COX), catalyzes the primary metabolic step in the transfor-

*Author for correspondence

mation of Arachidonic Acid (AA) to the bioactive produce prostaglandins and thromboxanes⁸. The continuation of isoforms of prostaglandin H synthase, namely COX-1 and COX-2, has been incorrigible in numerous organs, collectively with brain^{9,10}. Mice lacking in either COX-1 (COX-1^{-/-}) or COX-2 (COX-2^{-/-}) are handy and has been use to transport ahead our be privy to the physiologic and pathologic roles of the person COX isoforms¹¹.

The multiple sequence alignment is helpful in protein, DNA and RNA for biological sequences. In more than one sequences, alignment represents the evolutionary relation among organs. It implies to the quantity of sequences distribute an evolutionary courting, which means that they are all descendents from a common ancestor. This constituency might also talk to useful, structural, or evolutionary relationships among the sequences. Alignments can reproduce a degree of evolutionary vary among sequences that are descendants from a commonplace ancestor. There are different strategies available to generate phylogenetic trees and the phylogentic tree is based on genetic similarity of genes. On these images the MUSCLE is used for building a phylogenetic tree.

2. Methodology

The methodology purpose is to generate the Phylogenetic tree for Hippocampus DNA sequences based on the its accession numbers. The gene accession numbers are paste in the NCBI database and it gives the DNA sequences. Based on these sequences, Phylogenetic tree is reconstructed. The methodology follows different operations.

- a. Get the gene accession numbers from the genome table.
- b. Paste the accession numbers in the NCBI for DNA sequences.
- c. Then select the DNA sequences of each accession number of genes FASTA format.
- d. Save the FASTA format sequences in Ms-Word format for input to the MUSCLE.
- e. Select the MUSCLE from google.
- f. Paste all nucleotide sequences in MUSCLE.
- g. Then adjust the options for execution of MUSCLE.
- h. To opt the Phylogenetic tree in second iteration of the MUSCLE process.
- i. Then submit the job and finally get the Phylogenetic tree with accuracy values of the input sequences.

The framework of the MUSCLE and its procedure is represented in the Figure 7 and the required gene accession numbers is shows in the Table 1.

3. Implementation

The basic strategy used by MUSCLE is similar to that used by PRRP [13] and MAFFT [14]. A progressive alignment is built, to which horizontal refinement is then applied.

The MUSCLE is implemented by using basic approaches like ClustalW and MAFFT and then it applies the refinement process.

This approach has the three stages for progressive alignment.

1. Draft Progressive
2. Improved Progressive
3. Refinement

And each stage has the four sub-stages for complete MUSCLE approach.

The draft progressive is the first stage of progressive alignment and it contains the four steps. The draft progressive stage is completed in four steps which are as follows and is shown in Figure 1:

- i. Apply similarity measure for the given Hippocampus DNA sequences

Table 1. The hippocampus accession numbers of COX

Gene Accession Number	cox-1	cox-2
BG064367	-1.84	-1.94
BG064631	1.95	1.81
BG065181	1.52	2.09
BG065688	3.21	2.62
BG066667	1.6	2.07
BG067592	2.13	1.54
BG067766	1.65	1.51
BG069493	-1.84	-1.52
BG078482	1.67	1.84
BG079093	2.26	2.02
BG081977	-1.85	-1.94
BG083840	2.06	2.23

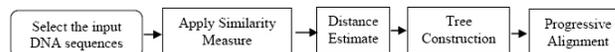


Figure 1 The Process of Draft progressive technique

- ii. Estimate the distance among the DNA sequences
- iii. Construct the tree based on the results obtained from step(ii)
- iv. Perform the progressive alignment for the generated sequences

The improved progressive is second stage of the refinement process and it has four steps for sequence alignment. The improved progressive stage also has four steps for completing its process with one different step (tree comparison) than the draft progressive stage as shown in Figure 2.

- i. Apply similarity measure for the given Hippocampus DNA sequences
- ii. Construct the tree based on the results obtained from step(i)
- iii. Compare the tree with the tree generated from draft progressive technique
- iv. Perform the progressive alignment for the generated sequences

The refinement is the last stage of the progressive alignment and it has the four steps for the sequence alignment which are as follows and is shown in Figure 3.

- i. Perform the choice of bipartition
- ii. Extract the profile from the results of step(i)
- iii. Realignment the sequences
- iv. Accept the input sequences if DNA sequences are similar gene otherwise reject.

4. Results

In the MUSCLE approach the phylogenetic tree is constructed based on Hippocampus DNA sequences, its DNA sequences are collected National Centre for Biotechnology Information (NCBI) databank. The similarity is very close with those gene accession numbers formed as a neighbor and less closeness genes also formed

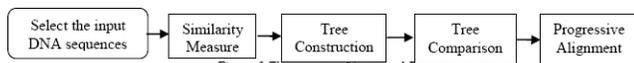


Figure 2 The process of Improved Progressive steps

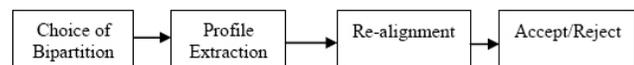


Figure 3. The process of refinement steps.

group but more distance. The percent identity matrix shows the diagonal value is zero for the corresponding genes and shown in the Figure 4.

The phylogenetic tree is constructed with MUSCLE based on DNA sequences and it is Phylogram as Cladogram in Figure 5 and Phylogram as Real in Figures 6, 7.

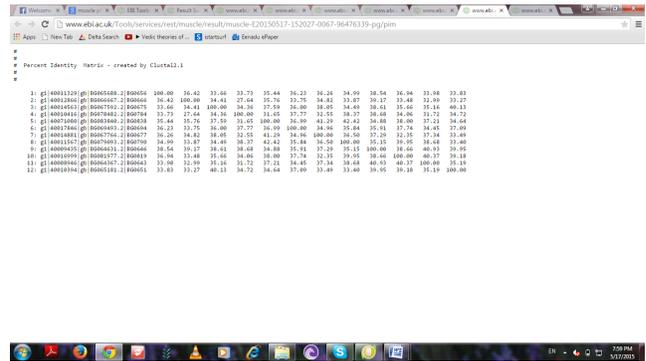


Figure 4. The identity matrix for input sequences.

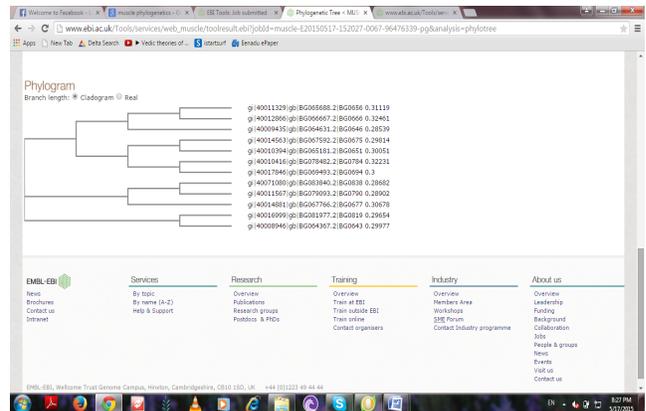


Figure 5. The phylogenetic tree of input sequences cladogram.

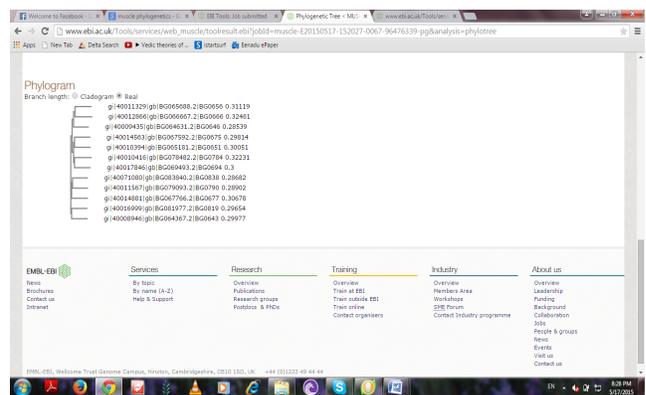


Figure 6. The phylogenetic tree of input sequence.

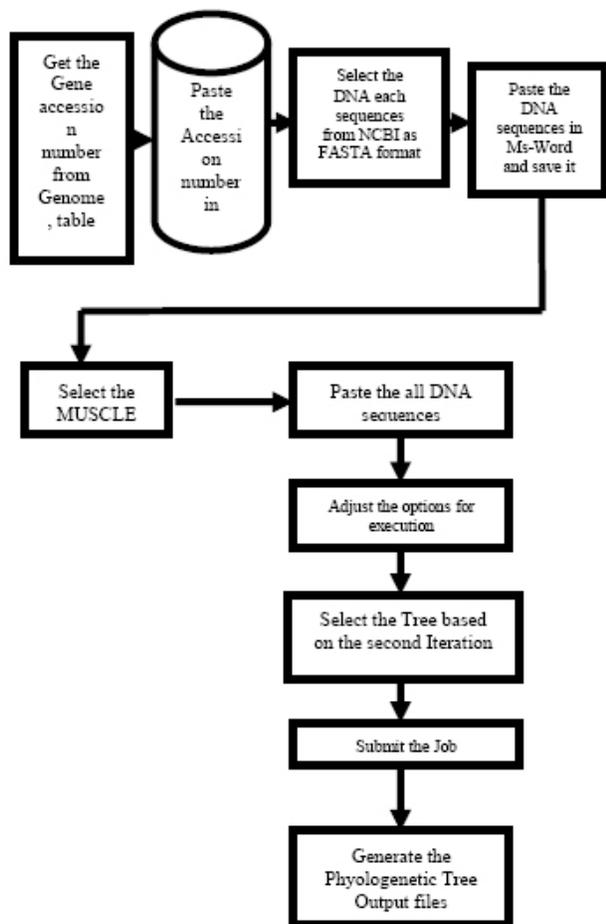


Figure 7. The framework of the MUSCLE approach.

5. Conclusion

To discover the evolutionary relationships of Hippocampus. Right here the closely sequences are formed as an institution of genes. The above outcomes are following the unweighted pair group arithmetic suggested (UPMA) and neighbor is part of techniques and indicates the rooted and unrooted evolutionary trees. The rooted trees are indicating the clear

closeness of the genes and unrooted trees are also closeness however, it's far in some way no longer clear.

6. Acknowledgement

I thank my supervisor Dr. A. Rama Mohan Reddy for giving his support in my work and UGC for providing the fellowship to do research work.

7. References

1. Garavito RM, Malkowski MG, DeWitt DL. The structures of prostaglandin endoperoxide H synthases. *Prostaglandins Other Lipid Mediat.* 2002; 1(68-69):129-52.
2. Tree of Life Project. What is Phylogeny? Tree of life web project [Internet]. [Cited 2008]. Available from: tolweb.org/tree/learn/concepts/whatisphylogeny.htm.
3. Potter R. Constructing phylogenetic tree using multiple sequence alignment, Master Thesis, University of Washington; 2008.
4. Santhaiah C. *International Journal of Advanced Research in Computer Science.* 2014; 5(5):115-17.
5. Bear MF, Malenka RL. *Current Opinion in Neurobiology.* 1994; 4(1):389-99.
6. Malenka RC, Nicoll RA. *Science.* 1999; 285(5):1870-4.
7. Tsien JZ. *Current Opinion in Neurobiology.* 2000; 10(2):266-73.
8. Smith WL, DeWitt DL, Garavito RM. Cyclooxygenases: structural, cellular, and molecular biology. *Annual Review of Biochemistry.* 2000; 69(4):145-82.
9. Kam PC, See S. Cyclo-oxygenase isoenzymes: Physiological and pharmacological role. *Anaesthesia.* 2000; 55(5):442-49.
10. Katori M, Majima M. Cyclooxygenase-2: Its rich diversity of roles and possible application of its selective inhibitors. *Inflammation Research.* 2000; 49(8):367-92.
11. Santhaiah C, Rama A, Reddy M. Role of computers in bioinformatics by using different biological datasets. *Journal of Computer Engineering.* 2014; 16(2):17-20.