

JOURNAL OF ADVANCEMENT IN

MEDICAL AND LIFE SCIENCES

Journal homepage: http://scienceq.org/Journals/JALS.php

Open Access

Plotting Phylogenetic Tree Using MOLEBLAST; Technical Guide Adebule, Adenivi Philip^{1*}

^{1*}Department of Microbiology, Ekiti State University, P.M.B. 5363, Ado Ekiti, Nigeria.

*Corresponding author: Adebule, Adeniyi Philip, Email: philbule@yahoo.com

Received: January 28, 2018, Accepted: March 15, 2018, Published: March 15, 2018.

ABSTRACT:

Article

The quest for neighbour sequence of the query for plotting the phylogenetic tree above led to the use of MOLE-BLAST. The construction of phylogenetic tree was demonstrated using three (3) organisms with the accession numbers; NR_074540.1, NR_025685.1 and NR_117706.1 obtained from food waste contaminated soil in Ekiti State, Nigeria and identified as *Bacillus cereus* ATCC strain, *Achromobacter insotitus* strain and *Achromobacter insuavis* strain respectively. These organisms were seen to show some level of relatedness based on the tree diagram. The study guide for the construction phylogenetic tree exposes the easy approach that is needed to utilize the MOLE-BLAST for neighbour search for submitted sequence queries.

Keyword: Phylogenetic Tree, MOLEBLAST, BLAST, Sequence Analysis, Genetics

INTRODUCTION

Mole blast as a neighbor search tool helps taxonomists find the closest database neighbor of submitted query sequences. It can also be termed as a tool that provides taxonomic context for queries by computing the multiple queries sequence alignment [1] along with their top BLAST database hit [2]; discovering their relationship to each other and representing neighborhood search result in a phylogenetic tree. Mole BLAST is a subset of BLAST (Basic local Alignment Search tool), a tool which finds regions of similarity between biological (nucleotide and protein) sequences, compare them to the sequences database and later calculated their statistical significance of matches. BLAST can be used to identify members of gene families by taking into use functional and evolutionary relation between sequences already documented at the database (GENE-BANK) [3].

Mole-BLAST differs from other BLAST tools such as smart BLAST (which finds proteins highly similar to your query); primer-BLAST (Designs primers specific to your PCR templates); Global Align (compare two sequences across their entire span); CD-search which find conserved domains in your sequence, GED (find matches to gene expression profites); igBLAST (search immunoglobutins and t-cell receptor sequences); CDART which find sequences with similar conserved domain architecture; and Bio Assay (which search protein or nucleotide targets in PubChem Bio Assay). MOLE BLAST is a useful tool for Taxonomist, Ecologists, and other scientist to submit sequence to NCBI (National centre for Biotechnology information) to confirm the correct taxonomic annotation of sequences. Also with MOLE-BLAST, assessment of sequence membership in taxonomy groups, discovery of neighbor sequences, visualization of relatedness of type reference specimens to sequences and separation of a large set of sequence into different genes or loci are possible.

MATERIAL AND METHODS

Source and Identity of Organisms

Three (3) organisms with the accession numbers; NR_074540.1, NR_025685.1 and NR_117706.1 were identified as *Bacillus cereus* ATCC strain, *Achromobacter insotitus* strain and *Achromobacter insuavis* strain respectively, were obtained from food waste contaminated soil in Ekiti State, Nigeria.

Open the MOLE-BLAST Interface

In other to access the MOLE-BLAST interface, firstly visit the NCBT website (www.ncbi.nim.nhl.gov), and then click on the BLAST icon situated at the popular resources section (Fig. 1). On the BLAST page, locate the MOLE-BLAST icon and click

to the MOLE-BLAST interface or visit *http://blast.ncbi.nlm.nih.gov/blast/moleblast.cgi* as showcased in Fig. 2.

Entry of sequences queries

Queries can be entered as nucleotide accessions, gis, or FASTA sequences with up to 300 input sequences with up to 5000 bases each. The query sequences to be aligned should be posted in the text area (Fig. 3).



Fig. 2: MOLEBLAST interphase

IOLE-BLAST	Neighbor Search Tool
acleotide	
Enter Query Sequence	MOLE BLAST searches for closest neighbors 🔮
Fatas anala alida a sasanla	ns. ols. or FASTA sequences(up to 200 input sequences with up to 5000 bases each) Q
NR_074540.1 NR_025685.1 NR_117706.1	land Bast as the new model and an endowing a discussion with the new model and the second second second second
ner nucleonde accession m_074540.1 MP_025485.1 NP_117706.1 Or, upload FASTA file	Choose File. No file chosen
nter nucleotos accesso m2º1540.1 m2º1545.1 m2º117706.1 Or, upload FASTA file Job Title	Choose File. No file chosen
The mucleone accesso mp.ortsto mp.ortsto mp.ortsto mp.ortsto mp.introne.i po upload FASTA file Job Title Choose Search Set	Choose File. No file chosen

Fig. 3: Sequence query entry page

The format of the input is automatically determined. FASTA File is chosen and uploaded in the case of FASTA while accession number can also be imputed. In either case, the same result will be gotten. A title is given to the queries while the search set will be chosen (i.e. nucleotide collection [nr/nt]) as seen in Fig. 3. After adequate entries of queries are made, the ALIGN icon is clicked to query sequences that belong to a cluster along with their top BLAST search whose results are afterwards submitted to Muscle [4] for multiple alignment. This carries out the discovery of neighbour sequence and compares it with the sequences in the data base to draw the phylogenetic tree. **RESULT AND DISCUSSION**

The quest for neighbour sequence of the query for plotting the phylogenetic tree above led to the use of MOLE-BLAST. In this phylogenetic tree construction guide, Fig. 4 shows the neighbour joining search tools calculating the relatedness of the sequence entry. The result from the calculation is represented in Fig. 5 with phylogenetic tree view either as neighbour joining [5] (5A) or fast minimum evolution [6] (5B). The tree view can be modified using maximum sequence differences (ranging from 0.05 to 0.75) and sequence label [sequence IDo, Taxonomic Name (if available), sequence tittle (if available), Blast name (if available), sequence ID (Blast name), Taxonomic Name (Sequence ID)]. In case of multiple sequences (mostly from difference family), the locus could be more than one (such as locus 1, 2 e. t. c.) (Fig. 6A and B)

The alignment of the neighbour sequences (Fig. 7) discovered at the Gene Bank can be accessed by clicking the "see Alignment" on the phylogenetic tree view page. Better view and modification of the phylogenetic tree view can be accessed by utilizing the different page tool such as the arrow (up, down, right and left), the zoom and refresh key. By clicking the tools icon, other sub-icon is displayed such as download, layout, sort, expand, edit label e. t. c. (Fig. 8). Further click on the layout icon, further give the opportunity to restructure the phylogenetic tree views (Rectangular cladogram, Slant cladogram, Radial tree, circular tree, use distance or mid-point).

/	
MOLE-BLAST	Neighbor Search Tool
Please, do not close the browser before the processing	g is completed
Vale BLAST Request ID	SUZX6R9W413
Status	Calculating
Jubmitted at	Fit Aug 11 05:57:56 2017
Surrent time	Fri Aug 11 05:58:21 2017
lime since submission	25 seconds
his page will be automatically updated in 10 seconds	
III JAT is a residueat trademark of the National Deans of Maderice	

Fig. 4: Neighbor Search Tool

The download icon give room for the easy access of the phylogenetic tree either as ASN text file, ASN binary file, Newick file, NEXUS file or PDF file. For the purpose of this guide, the tree is downloaded as PDF file as shown in Fig.9.

	National Center for Biotechnology Information	Sign in to NC
MOLE-BLAST	Neighbor Search Tool	Home Recent Results He
	Phylogenetic Tree View	
	This tree is based on NUSCLE multiple alignment computed for Mole-BLAST	
et Tree See alignment Moleblast RID SUZX6R9W413	Database 16	NA_typestr
ee method Max Seq Difference	Sequence Label Collapse Mode Locus	
leighber Joining 💽 😏 🛛 0.75 💽 🖌	😥 Sequence Title (if availa 💌 🤢 Show All 💌 😏 Locus 1 💌 🚱	
	Mouse over an internal node for a subtree or alignment. Click on tree label to select an	equence to download Hide leasen
	⁴⁹ Bacillus Systemsis stram BCT-3112 185 ribosomal RNA (emA) gene, complete sequence ⁴⁰ Bacillus cercus ATCC 14579 165 ribosomal RNA (emA) gene, complete sequence	Blast sames color map





Fig. 5B: Phylogenetic tree view (Fast Minimum Evolution)

	Sequence Label	Collapse Mode Locus
	Sequence Title (if availa 📉 😡	Show All 💌 🕑 Locus
	Sequence ID	over an internal node for a subtree or align
	Taxonomic Name (if available)	
	Sequence Title (if available)	illus toyonensis strain BCT-7112 165 ribos
1	Blast Name(if available)	cillus cereus ATCC 14579 16S ribosomal R
	Sequence ID(Blast Name)	cillus cereus strain CCM 2010 16S riboson
	Taxonomic Name/Samanas ID	cillus cereus strain IAM 12605 16S riboso
	Taxonomic Name(Sequence ID	Bacillus cereus strain ATCC 14579 16S rib
	3	Bacillus cereus strain NBRC 15305 16S rib
	L.S.	chromobacter aegrifaciens strain LMG 268
	5	Achromobacter spiritinus strain R-46660
		Achromobacter insuavis strain 1 MG 26845
		A A A A A A A A A A A A A A A A A A A

Fig. 6A: Sequence label of trees



Fig. 6B: Collapse Mode

Acce	antes .	Description	
ER CHARLES ADDRESS		A ferrendrative and an entry of the Additional 2017 Statements and an	
TH AND INTROD		Accessed where we can be accessed and the SAMA SAMA SAMA (Response) and the company of the second same	
ER BITT / BEAT		Activities and an and a second s	
C# CE017#31	4	Accessed as the ruph and Brooks ACH \$5,1937 parameter	
CP012044	1	Automotivation ter sphericalitation atract MWOOT, complete generation	
		Activitieshanted as: 06.0.1 103 resonanted RNA game, partial resonances	
C2 .00.031993.		Acheomobacter meditas strain G3 161 ribonimal Mah gene, partial sequence	
E3 PVP. 022641	1.1	Astronomication tradition attain LMG 0003 103 (Revenue RNA gene, and tal opportuni	
E1000000000000000000000000000000000000	b	Achromobactel sp. Ko 105 received this gene, partial sequence	
NR.074140	1.3	Bacillus satistis ATCC 14579 108 (Bosternal RHA (HrsA) gains, introduite projection	
STALLARD,	3	Bacilius cereius atrain M1.5, complete genoeme	
CT016505.		Bardhar termen stadt R., complete segments	
50017224		and the second state of th	
CA PROPERTY.		To existing a second se	
TR N.ITELAST	1	A substitution of a substantial and an A 17.7 A 8.8 a Research of WMA instant, successful even advance	
E .martient		A supplication on VEC bits 100 microsoval Phile series (series) and a series	
The name is named		Basisfield are 1910 B 40032 15/2 effective 1934 limited and a second rate	
C		Length And	
CA LANDOTTAN		One officer and here is an end of the electronic of the particular and an electronic of the particular and t	
	9	Bastillus thatinglenate strain c25, complete percenta	
515132323.	1	Astronomicanter encolinia strain DDM 22007, complete generation	
CTRALAMENT.		Annomabarter sylanomises etran FEAARGOS, 162, complete genome	
	******	annota vannt Valova Vana Aast	
S MANAGE		A TELEVISION AND A TELEVIS	
B HEALTH			
	******		******
	******		******
	040000 1 1 1 1 1 1 1 1 1 1 1 1 1 1		******
	1 1 1 1 1 1 1 1 1 1 1 1 1 1		******
	****** * * * * * * * * * * *		******
	******* * * * * * * * * * * * * * * *		******
	******* ******* **********************		****** #* #* #* #* #* #* #* #* #* #* #*
	· · · · · · · · · · · · · · · · · · ·		448484 249444 44 48444 48444 48444 48 48 48 48 48
	00000000000000000000000000000000000000		440380 Noreston 44 44 44 40 40 40 40 40 40 40 40 40 40
	00000000 1 1 1 1 1 1 1 1 1 1 1 1 1		440384 Notestea 44 84 84 84 84 84 84 84 84 84 84 84 84
	0 4 0 4 4 4 0 4 0 4 4 4 4 1 4 1 4 1 4 1 4 1 4 1 4		****** N===== A A A A A A A A A A A A A A A A
	0 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		***** ** ** ** ** ** ** ** ** ** ** **
	0 4 0 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		**************************************
	0 4 0 4 4 4 1 0 4 1 4 1 4 1 1 1 1 1 1 1 1 1 1 1 1		14.0.4.6. 10.0.2.6.00 14. 14. 14. 14. 14. 14. 14. 14.
	0		****** *** *** *** *** *** *** *** ***
	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		440444 440 44 44 44 44 44 44 44
	U		****** ** ** ** ** ** ** ** ** ** ** **
	0.000 0.0000000 0.0000000 0.0000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.00000000		****** ** ** ** ** ** ** ** ** ** ** **
	V 400 4 4 4 4 4 5 4 4 5 5 6 6 6 7 6 7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8		****** ** ** ** ** ** ** ** ** ** ** **
Image: A and a second	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		****** ** ** ** ** ** ** ** ** ** ** **
	Norman Norman 2 2 2 2 2 2 2 2 2 2 2 2 2		****** *** *** *** *** *** *** *** ***

Fig. 7: Alignment of neighbour sequence



Fig. 9: Sample of phylogenetic tree downloaded in PDF format

CONCLUSION

The study guide for the construction phylogenetic tree exposes the easy approach needed to utilize the MOLE-BLAST for neighbour search for submitted sequence queries. This will further help taxonomist and scientist to verify the correct taxonomic annotation of sequences. In the same biotechnology and genetic student will find it easy to handle MOLE-BLAST search tool to foster better research delivery.

FUTURE FOCUS

- BLAST made easy
- Achieving BLASTING with Mega 7.0.

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. Journal of Molecular Biology 1990;215:403–410. [PubMed: 2231712]
- 2. National Center for Biotechnology Information (NCBI), 2016. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 44: 7-19.
- Wolfe BE, Button JE, Santarelli M and Dutton RJ (2014) Cheese rind communi-ties provide tractable systems for in situ and in vitro studies of microbial diversi-ty, Cell 158:422-433
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1997
- Saitou N and Nei M (1987) The neighbor-joining method: a new method for re-constructing phylogenetic trees. Mol Biol Evol. 4:406-425
- 6. Desper R and Gascuel O (2004) Theoretical foundation of the balanced mini-mum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. Mol Biol Evol. 21:587-298.

Citation: Adebule, Adeniyi Philip (2018). Plotting Phylogenetic Tree Using MOLEBLAST; Technical Guide, DR Congo. J. of Advancement in Medical and Life Sciences. V6I3-03. DOI: 10.5281/zenodo.1198928

Copyright: © 2018 Adebule, Adeniyi Philip. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.