Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 5, Issue, 5, pp.1302-1306, May, 2013

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

MOLECULAR ASPECTS & PHYLOGENETIC PROFILES OF THE SEQUENCE INFORMATICS OF MOSQUITO SPECIES

*** Sekar, T., Praveena, P., Dr. D. Sudarsanam and L. Helenmary**

Department of Advance Zoology & Biotechnology, Loyola College, Chennai-600034

ARTICLE INFO

Article History:
Received 09th February, 2013
Received in revised form
18th February, 2013
Accepted 11nd April, 2013
Published online 12th May, 2013

Key words:
Mosquito species & biodiversity,
BLAST, CLUSTAL X.

ABSTRACT

There are about more than 3000 species of mosquitoes have been recorded in the world. The four different nucleotide sequences of mosquitoes of culex, anopheles aedes aegypti in mosquito biodiversity of their sequence similarities and evolutionary relationship were analysed.

Copyright, IJCR, 2013, Academic Journals. All rights reserved.

INTRODUCTION

Vectors are the transmitters of disease-causing organisms that carry the pathogens from one host to another. Common vectors are considered to be invertebrates invariably arthropods. Technically, however, vertebrates can also act as vectors, including foxes, raccoons, and dogs, which can transmit the rabies virus to humans through a bite. Arthropods account for over 85 percent of all known animal species, and they are the most important disease vectors. Among the arthropods, mosquitoes affect human health either directly by bites, stings, or infestation of tissues, or indirectly through disease transmission therefore the most significant mode of vector-borne disease transmission is by biological transmission by blood-feeding means.

MATERIALS AND METHODS

NCBI

The National Center for Biotechnology Information (NCBI) is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health. The NCBI is located in Bethesda, Maryland and was founded in 1988. The NCBI houses genome sequencing data in GenBank and an index of biomedical research articles in PubMed Central and PubMed, as well as other information relevant to biotechnology. All these databases are available online through the Entrez search engine. The NCBI is directed by David Lipman, one of the original authors of the BLAST sequence alignment program and a widely respected figure in Bioinformatics. He also leads an intramural research program, including groups led by Stephen Altschul (another BLAST co-author), David Landsman, and Eugene Koonin (a prolific author on comparative genomics).

BLAST

In bioinformatics, Basic Local Alignment Search Tool, or BLAST, is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold. For example, following the discovery of a previously unknown gene in the mouse, a scientist will typically perform a BLAST search of the human genome to see if humans carry a similar gene; BLAST will identify sequences in the human genome that resemble the mouse gene based on similarity of sequence.

CLUSTAL-X

Clustal X is a widely used multiple sequence alignment computer program. Multiple alignments are used to find diagnostic patterns to characterize protein families; to detect or demonstrate homology between new sequences; to suggest oligonucelotide primers for PCR; as an essential prelude to molecular evolutionary analysis. The rate of appearance of new sequence data is steadily increasing and the development of multiple alignment is therefore of major importance. The majority of automatic multiple alignments are now carried out using the “progressive “approach of Feng and Doolittle.

PHYLOGENETIC TREE

Phylogram is branching diagram (tree) assumed to be estimate of phylogenetic, branch lengths is proportional to the amount of inferred evolutionary change. A cladogram is a branching diagram assumed to be an estimate of phylogeny where the branches are of equal length, thus cladograms show common ancestry, but do not indicate the amount of evolutionary “time” separating taxa. Tree distances can be shown.

SEQUENCE OF MOSQUITO SPECIES

>gi|78217044|gb|CH478182.1| Aedes aegypti strain Liverpool supercont1.998 genomic scaffold, whole genome shotgun sequence

TTTTATAAAAATCTTTTTGGGAACTTATTTCAAAATGCTCGA
AAAATCCCTCTTTTTTACTAAAAATATA
TTTTTTTATTTTTATGTTTATTAACTTTTATTACACCTTTCAA
GTCCTTTATTCACTGCCAAACCAGAAA

Corresponding author: sekarsolomon125@gmail.com

TGGGTCATGTTCCTGAATAAAAGCGTTTTAAATGGATATCG
AGTTATGGAGGGAAATATCCTCCCACGTG
ACATCGACTAGTGGAGACATCGAGTTATAGAAATATCGACT
TATGGAGAGCATGATGTATGGGAACTTGA
AGGGACCGCAAAATCCATCGAGTTATAGAATATATCGAGTT
ATAGAACATCGAGTTATGGAGAGTCGACT
GTAGTTATTTGTGGAATGGCGCGTCCTGAAGAATGATGGAT
TCTAGGGAATGTTTTTCTGGTTGGTGGAA
TTCCGGAGAACCATTTCCGGGGGAAAAGTTTTCTGGGGAAT
GTTACAGATTTGTTTATTGGGTATGGTCA
TAATTGTTTCAATAACACTCTTTATGAATTTTCCAGCGATTT
GTCCAAGTTGTGTAGACTAAAAATAGAT
CCCCTATCGGTCGACAATCCACACACCAAAGTATTCCTACT
GATGCAAGCTCATATGTCCCGGCTGCCTC
TTCCGAATTCCGACTACGGAACGGATACGAAATCCGTTCTG
GATCAGTCGATCCGAATCATTCAGGTAGG
TTCCAGTAGCAAAATGATCGCAACCTCGAGATAATTTAACG
TTTGATCGTTTCAGGCAATGATCGACATT
ACTGCGGAGCGTGGTTGGCTCGTAAGTACCCTTCGTATTCA
GCAACTTATGCAGTGCATTGTTCAAGCCC
GATGGATTGATGACCCAGTGGTGATGACCCTGCCAAATGTT
GAACCACATAATGCTCACGTCTTCAAACA
CGTAAAATTGGAGTAAGTATTGACAGCGAACTGTAATATGT
GAAAGCTCTATGATCATTTTTCACTTTTT
CCCAGTCATCCATACCTGACACTGCCTGCCCTGAAAGAAAA
ATGCCACCGAAAGTACGAAAACCTGGCTG

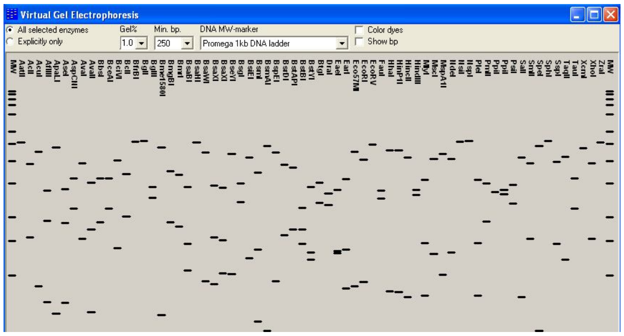

Virtual Gel Electrophoresis
● All selected enzymes
○ Explicitly only
Gel% 1.0
Min. bp. 250
DNA MW-marker Promega 1kb DNA ladder
☐ Color dyes
☐ Show bp

Fig.1 Insilco way of agarose electrophoresis of *Aedes aegypti*

>gi|442885243|dbj|AB776709.1| Anopheles gambiae VSC gene for voltage-gated sodium channel, partial cds, isolate: RTK1

CCTGTAGTGATAGGAAATTCAGTCGTAAGTAATGCAAATTA
ACATGGACCAAGATCGTTTTTACATGACA
TTGTTTTGCAGGTGCTTAACCTTTTCTTAGCCTTGCTTTTGTC
AAATTTTGGTTCAT

>gi|397739584|gb|JX260491.1| Anopheles punctipennis voucher NEONTculicid2063 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial

AACATTATACTTTATTTTTGGAGCATGATCTGGGATAGTGG
GAACTTCTCTTAGAATCCTTATTCGAGCA
GAACTAAGTCAACCAGGAGTTTTTATTGGAAATGACCAAAT
TTATAATGTTATTGTAACAGCTCATGCTT
TTATTATAATTTTTTTTTATAGTTATACCTATTATAATTGGAG
GATTTGGTAATTGATTAGTTCCTTTAAT
ATTAGGAGCCCCTGATATAGCCTTTCCTCGAATAAATAATA
TAAGATTTTGAATACTTCCTCCCTCATTA
ACTCTCCTCCTTTCTGGGGGTATAGTAGAAAGCGGGGCTGG
GACTGGATGAACTGTTTATCCCCCACTTT
CTGCTGGAACTGCTCATGCGGGGGCATCTGTAGATCTTTCT
ATTTTTTCTCTTCATTTAGCAGGAATTTC
TTCTATTTTAGGGGCAGTAAACTTTATTACTACAGTAATTAA
TATACGAACTTCTGGTATTACTTTAGAT
CGTTTACCTTTATTTGTTTGATCTGTAGTAATTACAGCGGTT
TTATTACTTTTATCCCTTCCAGTCCTTG
CAGGGGCTATTACCATACTTTTAACTGACCGAAACTTAAAT
ACTTCCTTTTTTGATCCTACAGGAGGGGG
AGACCCTATCTTATACCAACATTTATTT

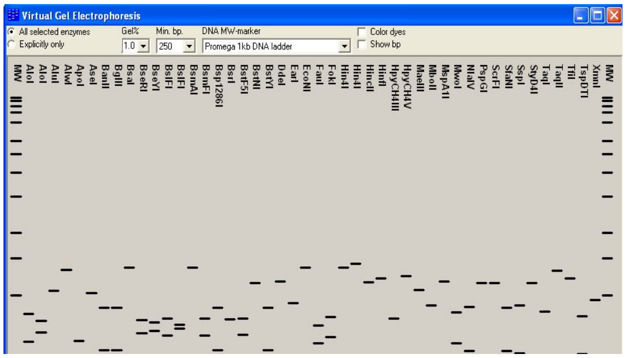

Virtual Gel Electrophoresis
● All selected enzymes
○ Explicitly only
Gel% 1.0
Min. bp. 250
DNA MW-marker Promega 1kb DNA ladder
☐ Color dyes
☐ Show bp

Fig.2 Insilico way of agarose electrophoresis of *Anopheles* species

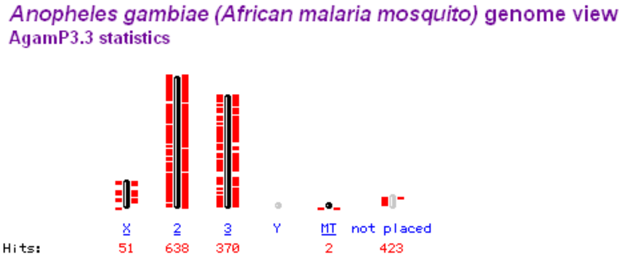


Fig.3 Map viewer of *Anopheles* species

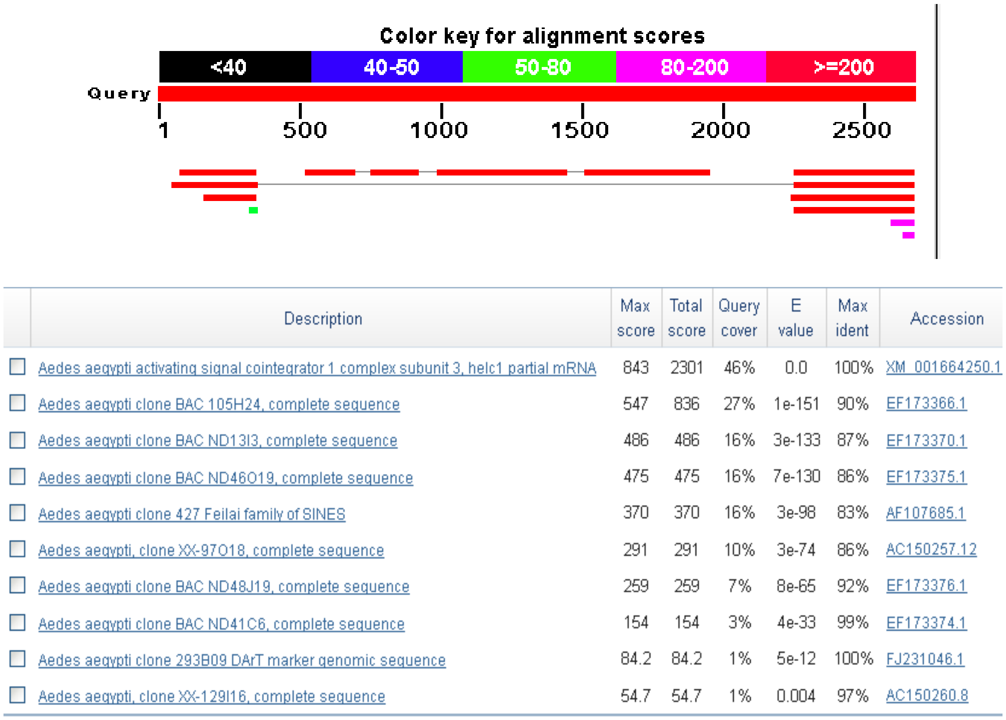
>gi|442540092|gb|JX975223.1| Plasmodium sp. MarF-2013 cytochrome b (cytb) gene, partial cds; mitochondrial

CAACAGGAGCTTCATTTGTATTTATTTTAACTTATTTACATA
TTTTAAGAGGATTAAATTATTCATACTC
ATATTTACCTTTATCATGGATATCAGGATTAATGATATTTTT
AATATCAATAGTTACAGCTTTTATGGGT
TATGTATTACCTTGGGGTCAAATGAGTTTCTGGGGTGCAAC
TGTTATCACTAATTTATTATATTTTATCC
CTGGACTTGTTTCATGGATTTGTGGTGGATATCTTGTAAGTG
ACCCAACTTTAAAAAGATTCTTTGTATT
ACATTTTATATTTCCATTCATTGCTTTATGTATTGTATTCATT
CATATATTCTTTTTACATTTACAAGGT
AGCACTAATCCTTTAGGGTATGATACAGCTTTAAAAATACC
CTTCTATCCAAATCTATTAAGTCTCGATA
TAAAAGGATTTAATAATGTATTAGTCTTATTCTTAGCACAA
AGCTTATTTGGAATTTT

>gi|157138707|ref|XM_001664250.1|:792-1235 Aedes aegypti activating signal cointegrator 1 complex subunit 3, helc1 partial mRNA

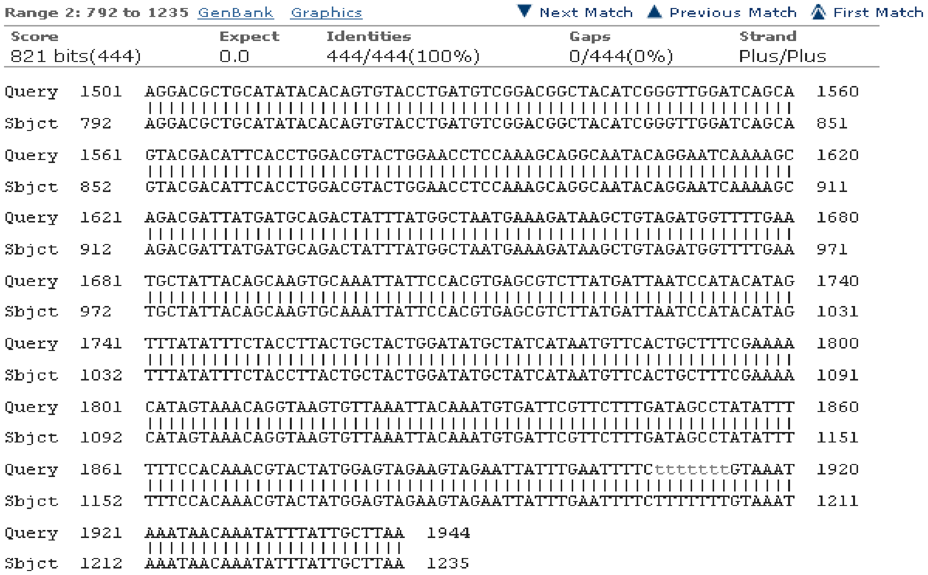
AGGACGCTGCATATACACAGTGTACCTGATGTCGGACGGCT
ACATCGGGTTGGATCAGCAGTACGACATTCACCTGGACG
TACTGGAACCTCCAAAGCAGGCAATACAGGAATCAAAAGC
AGACGATTATGATGCAGACTATTTATGGCTAATGAAAGAT
AAGCTGTAGATGGTTTTGAATGCTATTACAGCAAGTGCAAA
TTATTCCACGTGAGCGTCTTATGATTAATCCATACATAG
TTTATATTTCTACCTTACTGCTACTGGATATGCTATCATAAT
GTTCACTGCTTTCGAAAACATAGTAAACAGGTAAGTGT
TAAATTACAAATGTGATTCGTTCTTTGATAGCCTATATTTTT
TCCACAAACGTACTATGGAGTAGAAGTAGAATTATTTG
AATTTTCTTTTTTTGTAAATAAATAACAAATATTTATTGCTT
AA

>gi|157138707|ref|XM_001664250.1|:1-172 Aedes aegypti activating signal cointegrator 1 complex subunit 3, helc1 partial mRNA



	Description	Max score	Total score	Query cover	E value	Max ident	Accession
☐	Aedes aegypti activating signal cointegrator 1 complex subunit 3, helc1 partial mRNA	843	2301	46%	0.0	100%	XM_001664250.1
☐	Aedes aegypti clone BAC 105H24, complete sequence	547	836	27%	1e-151	90%	EF173366.1
☐	Aedes aegypti clone BAC ND13I3, complete sequence	486	486	16%	3e-133	87%	EF173370.1
☐	Aedes aegypti clone BAC ND46O19, complete sequence	475	475	16%	7e-130	86%	EF173375.1
☐	Aedes aegypti clone 427 Feilai family of SINES	370	370	16%	3e-98	83%	AF107685.1
☐	Aedes aegypti, clone XX-97O18, complete sequence	291	291	10%	3e-74	86%	AC150257.12
☐	Aedes aegypti clone BAC ND48J19, complete sequence	259	259	7%	8e-65	92%	EF173376.1
☐	Aedes aegypti clone BAC ND41C6, complete sequence	154	154	3%	4e-33	99%	EF173374.1
☐	Aedes aegypti clone 293B09 DArT marker genomic sequence	84.2	84.2	1%	5e-12	100%	FJ231046.1
☐	Aedes aegypti, clone XX-129I16, complete sequence	54.7	54.7	1%	0.004	97%	AC150260.8

Fig.4 BLAST-N



```
Range 2: 792 to 1235

Score           Expect    Identities          Gaps          Strand
821 bits(444)   0.0       444/444(100%)       0/444(0%)     Plus/Plus

Query  1501  AGGACGCTGCATATACACAGTGTACCTGATGTCGGACGGCTACATCGGGTTGGATCAGCA  1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  792   AGGACGCTGCATATACACAGTGTACCTGATGTCGGACGGCTACATCGGGTTGGATCAGCA  851

Query  1561  GTACGACATTCACCTGGACGTACTGGAACCTCCAAAGCAGGCAATACAGGAATCAAAAGC  1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  852   GTACGACATTCACCTGGACGTACTGGAACCTCCAAAGCAGGCAATACAGGAATCAAAAGC  911

Query  1621  AGACGATTATGATGCAGACTATTTATGGCTAATGAAAGATAAGCTGTAGATGGTTTTGAA  1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  912   AGACGATTATGATGCAGACTATTTATGGCTAATGAAAGATAAGCTGTAGATGGTTTTGAA  971

Query  1681  TGCTATTACAGCAAGTGCAAATTATTCCACGTGAGCGTCTTATGATTAATCCATACATAG  1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  972   TGCTATTACAGCAAGTGCAAATTATTCCACGTGAGCGTCTTATGATTAATCCATACATAG  1031

Query  1741  TTTATATTTCTACCTTACTGCTACTGGATATGCTATCATAATGTTCACTGCTTTCGAAAA  1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1032  TTTATATTTCTACCTTACTGCTACTGGATATGCTATCATAATGTTCACTGCTTTCGAAAA  1091

Query  1801  CATAGTAAACAGGTAAGTGTTAAATTACAAATGTGATTCGTTCTTTGATAGCCTATATTT  1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1092  CATAGTAAACAGGTAAGTGTTAAATTACAAATGTGATTCGTTCTTTGATAGCCTATATTT  1151

Query  1861  TTTCCACAAACGTACTATGGAGTAGAAGTAGAATTATTTGAATTTTCtttttttGTAAAT  1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1152  TTTCCACAAACGTACTATGGAGTAGAAGTAGAATTATTTGAATTTTCTTTTTTTGTAAAT  1211

Query  1921  AAATAACAAATATTTATTGCTTAA  1944
             ||||||||||||||||||||||||
Sbjct  1212  AAATAACAAATATTTATTGCTTAA  1235
```

Fig 5. Similar sequences

AGCGATTTGTCCAAGTTGTGTAGACTAAAAATAGATCCCCT
ATCGGTCGACAATCCACACACCAAAGTATTCCTACTGAT
GCAAGCTCATATGTCCCGGCTGCCTCTTCCGAATTCCGACT
ACGGAACGGATACGAAATCCGTTCTGGATCAGTCGATCC
GAATCATTCAGG

>gi|157138707|ref|XM_001664250.1|:167-338 Aedes aegypti activating signal cointegrator 1 complex subunit 3, helc1 partial mRNA

TTCAGGCAATGATCGACATTACTGCGGAGCGTGGTTGGCTC
GTAAGTACCCTTCGTATTCAGCAACTTATGCAGTGCATT
GTTCAAGCCCGATGGATTGATGACCCAGTGGTGATGACCCT
GCCAAATGTTGAACCACATAATGCTCACGTCTTCAAACA
CGTAAAATTGGA

>gi|78126903|gb|AC150257.12|:58231-58510 Aedes aegypti, clone XX-97O18, complete sequence

TTTTTAATGTTTTTTAACATTTATTACAACTTGAAAGTATTT
ATTCGCTTTCAAACAAGAATTAGTCCGAGTTTCCTCAT
ACAAGAAAGATTTTAAAATGGATATCGAGTTATGGAGGGA
AATTTTCCTCTCACGTGACATCGACTAATGGAGAGATCGA
GTTATAGAAATATCGACTTATGGAGAGCACGATGTATGGGA
ATTTGAAGGGACCGATAAATCCATCGAGTTATATAGTAT
ATCGAGTTATAGAATATCGAGTTGTGGAGAGTCGACTGTA

>gi|122937776|gb|EF173376.1|:59810-59998 Aedes aegypti clone BAC ND48J19, complete sequence

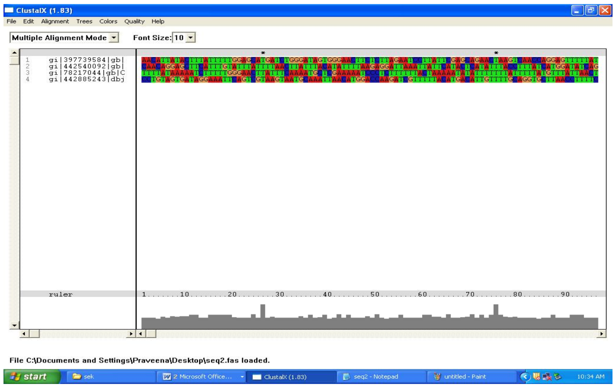
TTTTAAATATATATCGAGTTATGGAGAGAAAATTACCTCCC
ACGTGACATCGACTAGTGGAGATATCGAGTTATAGAAAT
ATCGAAATATGGAGAGCACGATGTATGGAAATTTGAAGGG
ACCGAAAAATCCATCGAGTTATAGAGTATATCGAGTTATA
GAATATCGAGTTGTGGAGAGTCGACTGTA
>gi|122937769|gb|EF173374.1|:40108-40193 Aedes aegypti clone BAC ND41C6, complete sequence

GCAGGCAACAGTGGTGCTCCTGGCGGGAAGAATAGATCAA
AGTAATCCAGGCTACTATGTTCTACAGCAAAAAAGCAGTG
TTGCTC
>gi|111038199|gb|AC150260.8|:21734-21765 Aedes aegypti, clone XX-129I16, complete sequence

GAATATCGAGTTATGGAGAGTCGACTGTAGTT

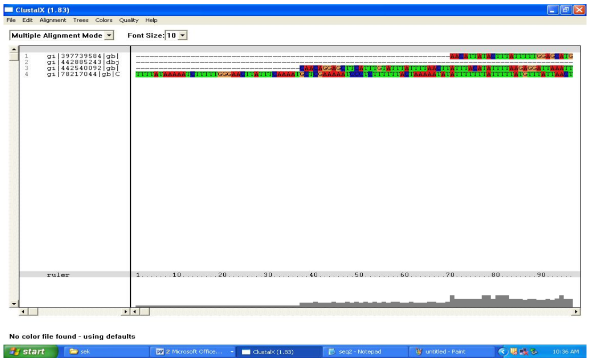
PROT DIST

```
  5
gi|1471544    0.00000 0.05055 0.37661 0.72457 0.83569
gi|1215926    0.05055 0.00000 0.39864 0.70187 0.83576
gi|1691803    0.37661 0.39864 0.00000 0.63743 0.82960
gi|1166944    0.72457 0.70187 0.63743 0.00000 0.79460
gi|1776769    0.83569 0.83576 0.82960 0.79460 0.00000
```



File C:\Documents and Settings\Praveena\Desktop\seq2.fas loaded.

MULTIPLE SEQUENCE ALIGNMENT WINDOWS



No color file found - using defaults

Fig 6&7CLUSTAL X

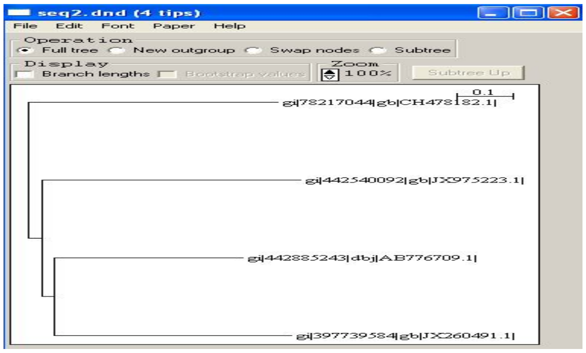


Fig. 8. NJ –PLOT

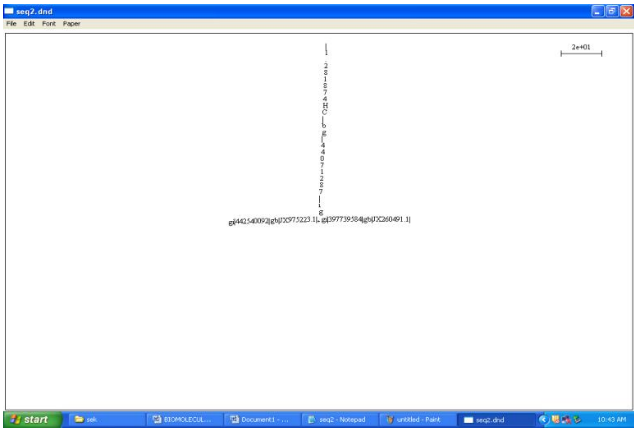
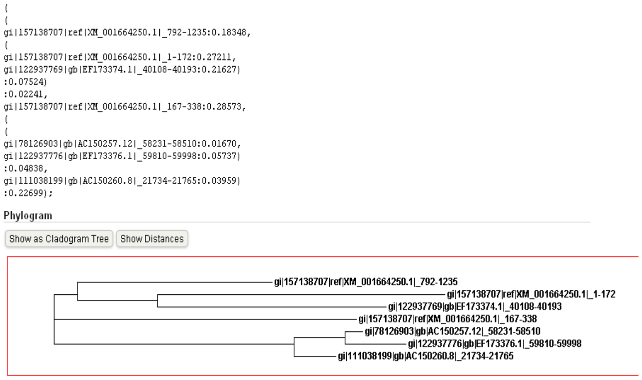


Fig. 9. UN ROOTED TREE



```
(
(
gi|157138707|ref|XM_001664250.1|_792-1235:0.18348,
(
gi|157138707|ref|XM_001664250.1|_1-172:0.27211,
gi|122937769|gb|EF173374.1|_40108-40193:0.21627)
:0.07524)
:0.02241,
gi|157138707|ref|XM_001664250.1|_167-338:0.28573,
(
(
gi|78126903|gb|AC150257.12|_58231-58510:0.01670,
gi|122937776|gb|EF173376.1|_59810-59998:0.05737)
:0.04838,
gi|111038199|gb|AC150260.8|_21734-21765:0.03959)
:0.22699);
```

Fig.10 Clustal W

DISCUSSION

Globally there are about more than 3500 species of mosquito which show variety and variability in their characteristic features including habit, feeding habits, geographical distribution, larval habitat, external features in relation to other species, diseases caused, proteomics, genomics, metabolomics etc.,

The sequence of anopheles, *culex, and aedes aegypti* mosquito were blasted to find the similarity and homology between the sequence and different species according to the E-value and score obtained. Clsutal X was performed for multiple sequence alignment to find the evolutionary relationships among various organisms are performed. Cladogram and phylogram to find the evolutionary distance & the relationship between each species and they are evaluated for nucleotide sequence. Isolation and characterization of genes that are involved in critical behaviours of economically important insects such as disease vectors and agricultural pests. One example of such a behavior that we are investigating is host (i.e. blood-meal source) seeking/selection in several Anopheline mosquitoes comprising the principal mosquito vectors for malaria and other diseases that are responsible for over 3 million deaths annually. Studying the genes encodijg the molecular components of olfaction (smell) in the vector mosquito Anopheles Gambians. Because olfaction is the dominant sensory input for host preference behaviours in mosquitoes and other insects. Since these behaviors make a critical contribution to the vectorial capacity of this mosquito as well as playing similar roles in the overall impact of many other insects of economic and medical importance, an analysis of the molecular basis for olfactory discrimination in this system is of fundamental significance. In future, more information on their genetical features of these mosquitoes with respect to the pathogen they carry would be introduced to have a complete data about them for better identification and potential compatability to treat the mosquito borne diseases.

REFERENCES

1. Adele Ahmed., Rama Rao, V., Radhakrishna P., 1999, A model based on physical and physiological characteristics for the generation of high wing beat frequency in mosquito flight. *Bull Pure & Appl. Sci.* 18, 49-54.
2. Bakonyi T., Hubálek Z., Rudolf I., Nowotny N. 2005. Novel flavivirus or new lineage of West Nile virus, Central Europe. *Emerg. Infect. Dis.* 11: 225-231.
3. Cancrini, G., Frangipane di Regalbono, A., Ricci, I., Tessarin, C., Grabrielli, S. & Pietrobelli, M. (2003a) Aedes albopictus is a natural vector of Dirofilaria immitis in Italy. *Veterinary Parasitology,* 118(3-4) pp195-202
4. Daniel Lawson, Peter Arensburger, Peter Atkinson, Nora J. Besansky, 2008 VectorBase: a data resource for invertebrate vector genomics Nucleic Acids Research, 2009, Vol. 37, Database issue D583–D587 doi:10.1093/nar/gkn857
5. Elizondo Quiroga, 2002. Taxonomia y distribución de los mosquitos (Diptera: Culcidae) de las regiones fisiograficas Llanura costera Del Golfo y Sierra Madre Oriental, del estado de Nuevo Leon, Mexico *[Tesis de Maestria]* Universidad Autónoma de Nuevo León, San Nicolás de los Garza, México.
6. Fasulo TR, Dame DA. (2002). Mosquitoes 1 and 2. Bug Tutorials. University of Florida/IFAS. CD-ROM. SW-160.
7. Garner P, Graves PM. The benefits of artemisinin combination therapy for malaria extend beyond the individual patient. *PLoS Med. 2005 Apr; 2(4):e105. Epub 2005 Apr 26.*
8. Halstead S B, (1997) in Dengue and Dengue Hemorrhagic Fever, Eds GublerD J, Kuno G (CAB International, Wallingford, U.K.) pp 23–44.

****