Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Disease

journal homepage: www.elsevier.com/locate/apjtd

Document heading doi: 10.1016/S2222-1808(14)60681-2

Homology and conservation of amino acids in E-protein sequences of dengue serotypes

Ramesh Venkatachalam, Vijayakumar Subramaniyan*

P.G. and Research Department of Botany and Microbiology, A. Veeriya Vandayar Memorial Sri Pushpam College, Poondi-613 503, Thanjavur District, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 14 Jul 2014
Received in revised form 14 Aug 2014
Accepted 24 Aug 2014
Available online 28 Aug 2014

Keywords:
Conservation
Dengue serotypes
E-proteins
Homology
In silico analysis

ABSTRACT

Objective: To identify the homology and phylogenetic relationship among the four dengue virus (DENV) serotypes, and conservation of amino acid in E-proteins and to find out the phylogenetic relationship among the strains of four DENV serotypes.
Methods: Clustal W analysis for homology and phylogram, European molecular biology open software suite for pairwise alignment of amino acid sequences and BLAST-P analysis for various strains of four DENV serotypes were carried out.
Results: Homology of E-protein sequences of four DENV serotypes indicated a close relationship of DENV-1 with DENV-3. DENV-2 showed close relationship with DENV-1 and -3 forming a single cluster whereas DENV-4 alone formed group with a single serotype. In the multiple sequence alignment, 19 amino acid conserved groups were observed. BLAST-P analysis showed more number of 100% similarity among DENV-1 and -3 strains whereas only few strains showed 100% similarity in DENV-4. However, 100% similarity was absent among the DENV-3 strains.
Conclusions: From the present study, phylogenetically all the four DENV serotypes were related but DENV-1, -2 and -3 were very closely related whereas DENV-4 was somewhat distant from the other three serotypes.

1. Introduction

Dengue fever is caused by four dengue virus (DENV) serotypes, DENV-1, DENV-2, DENV-3 and DENV-4[1]. The four serotypes are recognized based on their antigenicity and immunogenic property in the human body. Dengue virus genome is a single stranded RNA which is translated into a single polyprotein, proteome containing three structural proteins and seven non-structural proteins. The structural proteins are capsid protein (C), membrane protein (M) and envelope protein (E)[2]. Among them, E-protein is considered more important than others. Therefore in the present study E-proteins of four dengue serotypes have been analyzed and discussed for their homology and interrelationships within the serotypes and within the strains of each serotype.

E-protein is basically formed by 495 amino acids. This number is not constant for all four serotypes. DENV-1 and DENV-4 showed 495 amino acids in their sequences. DENV-3 showed 493 whereas only 480 were found in the case of DENV-2. This variation in the number is due to the deletion of some of the amino acids in their sequences in the middle or at the tail end. All the amino acids in the sequences of four serotypes are not similar and show variations. This variation is due to the mutations occurred in the position of some of the amino acids in the sequences. The replaced amino acids are qualitatively similar in their properties and characters of replaced amino acids.

2. Materials and methods

Amino acid sequences of E-proteins of dengue serotypes DENV-1, DENV-2, DENV-3 and DENV-4 were retrieved from the National Center for Biotechnology Information[3]. The sequences were converted into FASTA format. The retrieved sequences were fed with ClustalW for multiple sequence alignment to observe sequence homology and the phylogram was drawn for all proteins[4]. Pairwise alignment of amino acid sequences was carried out using European molecular biology open software suite[5]. Homology analysis of E-proteins was carried out using BLAST-P[6], to find out the similarity within the respective strains of the serotypes[7].

*Corresponding author: Vijayakumar Subramaniyan, P.G. and Research Department of Botany and Microbiology, A. Veeriya Vandayar Memorial Sri Pushpam College, Poondi-613 503, Thanjavur District, Tamil Nadu, India.
Tel: 04374-239523
Fax: 04374-239477
E-mail: svijaya_kumar2579@rediff.com
Foundation Project: Supported by University Grants Commission New Delhi, India for the Major Research Project [MRP R. No: 41-472/2012(SR)].

3. Results

3.1. Homology of E–proteins

Multiple sequence alignment of all the four E–protein sequences was carried out to understand the homology and conservation of amino acids within the four serotypes (Figure 1). The phylogram deduced from the E–protein sequences of four dengue serotypes showed similarity and differences among the serotypes. The analysis indicated a close relationship of DENV–1 with DENV–3. DENV–2 showed close relationship with DENV–1 and –3 forming a single cluster whereas DENV–4 alone formed a single serotype which showed a distant relationship with other serotypes in Figure 2.

```
Denguevirus1    MRCVGIGSRDFVEGLSGATWVDVVLEHGSCVTTMAKDKPTLDIELLKTEVTNPAVLRKLC
Denguevirus2    MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEAKQPATLRKYC
Denguevirus3    MRCVGVGNRDFVEGLSGATWVDVVLEHGGCVTTMAKNKPTLDIELQKTEATQLATLRKLC
Denguevirus4    MRCVGVGNRDFVEGVSGGAWVDLVLEHGGCVTTMAQGKPTLDFELIKTTAKEVALLRTYC
                ***:*:.:******:** .:***:*****.******:.*****:** ** ..: * **. *

Denguevirus1    IEAKISNTTTDSRCPTQGEATLVEEQDANFVCRRTFVDRGWGNGCGLFGKGSLLTCAKFK
Denguevirus2    IEAKLTNTTTESRCPTQGEPSLNEEQDKRFVCKHSMVDRGWGNGCGLFGKGGIVTCAMFT
Denguevirus3    IEGKITNITTDSRCPTQGEAVLPEEQDQNYVCKHTYVDRGWGNGCGLFGKGSLVTCAKFQ
Denguevirus4    IEASISNITTATRCPTQGEPYLKEEQDQQYICRRDVVDRGWGNGCGLFGKGGVVTCAKFS
                **.:.:* ** :*******. * **** .:.*:: ***************.:.*** *

Denguevirus1    CVTKLEGKIVQYENLKYSVIVTVHTGDQHQVGNESTEHGTTATITPQAPTTEIQLTDYGA
Denguevirus2    CKKNMKGKIVQPENLEYTIVITPHSGEEHAVGNDTGKHGKEIKITPQSSITEAELTGYGT
Denguevirus3    CLEPIEGKVVQYENLKYTVIITVHTGDQHQVGN--ETQGVTAEITPQASTTEAILPEYGT
Denguevirus4    CSGKITGNLVQIENLEYTVVVTVHNGDTHAVGNDTSNHGVTATITPRSPSVEVKLPDYGE
                *  : *::** ***:*:...* *.*:*.* ***   :*   ***:.. .*  *. **

Denguevirus1    LTLDCSPRTGLDFNEMVLLTMKEKSWLVHKQWFLDLPLPWTSGASTSQETWNRQDLLVTF
Denguevirus2    VTMECSPRTGLDFNEMVLLQMEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTF
Denguevirus3    LGLECSPRTGLDFNEMILLTMKNKAWMVHRQWFFDLPLPWTSGATTETPTWNRKELLVTF
Denguevirus4    LTLDCEPRSGIDFNEMILMKMKKKTWLVHKQWFLDLPLPWAAGADTSEVHWNYKERMVTF
                :.:*.**:*:*****.*: *:.*:*.**.***:****** .** *.  * :.:.***

Denguevirus1    KTAHAKKQEVVVLGSQEGAMHTALTGATEIQTSGTTTIFAGHLKCRLKMDKLTLKGMSYV
Denguevirus2    KNPHAKKQDVVVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCRLRMDKLQLKGMSYS
Denguevirus3    KNAHAKKQEVVVLGSQEGAMHTALTGATEIQNSGGTSIFAGHLKCRLKMDKLELKGMSYA
Denguevirus4    KVPHAKRQDVTVLGSQEGAMHSALTGATEVDSGDGNHMFAGHLKCKVRMEKLRIKGMSYT
                * .***:*:*.**********:*******:. . . .:*:*****:..*:** .*****

Denguevirus1    MCTGSFKLEKEVAETQHGTVLVQIKYEGTDAPCKIPFSTQDEKGVTQNGRLITANPIVTD
Denguevirus2    MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEIMDLEKRHVLGRLITVNPIVTE
Denguevirus3    MCTNTFVLKKEVSETQHGTILIKVEYKGEDAPCKIPFSTEDEQGKAHNGRLITANPVVTK
Denguevirus4    MCSGKFSIDKEMAETQHGTTVIKVKYEGAGAPCKVPIEIRDVNKEKVVGRVISSTPFAEN
                **:.* : **::****** .....*:* .:***:*:. *:.  * :    **:*:.*:. .

Denguevirus1    KEKPVNIEAEPPFGESYIVIGAGEKALKLSWFKKGSSIGKMFEATARGARRMAILGDTAW
Denguevirus2    KDSPVNIEAEPPFGDSYIIIGVEPGQLKLNWFKKGSSIGQMFETTMRGAKRMAILGDTAW
Denguevirus3    KEEPVNIEAEPPFGESNIVIGIGDNALKINWYKKGSSIGKMFEATARGARRMAILGDTAW
Denguevirus4    TNSVTNIELEPPFGDSYIIIGVGDSALTLHWFRKGSSIGKMFESTYRGAKRMAILGETAW
                .:. .*** *****:* *:**    *.: *::******:***:* ***:*******:***

Denguevirus1    DFGSIGGVFTSVGKLVHQIFGTAYGVLFSGVSWTMKIGIGVLLTWLGLNSRSTSLSMTCI
Denguevirus2    DFGSLGGVFTSIGKALHQVFGAIYGAAFSGVSWTMKILIGVIITWIGMNSRSTSLSVSLV
Denguevirus3    DFGSVGGVLNSLGKMVHQIFGSAYTALFSGVSWVMKIGIGVLLTWIGLNSKNTSMSFSCI
Denguevirus4    DFGSVGGLLTSLGKAVHQVFGSVYTTMFGGVSWMVRILIGFLVLWIGTNSRNTSMAMTCI
                ****.**:..*.**.:**:**: * . *.**** .:* **.:: *:* **:.**:.::: :

Denguevirus1    AVGLVTLYLGVMVQA
Denguevirus2    ---------------
Denguevirus3    AIGIITLYLGAVVQA
Denguevirus4    AVGGITLFLGFTVQA
```

Figure 1. Multiple sequence alignment of E–proteins of DENV serotypes showing homology and conserved regions using ClustalW.

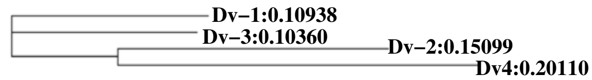

Figure 2. Phylogram for DENV–1 to –4.

A highest percentage of 78.4% identical of amino acids have been shown in DENV–1 and DENV–3, and 66.1% homology was observed between DENV–1 and DENV–2, whereas DENV–3 and DENV–4 showed only 63.4% alone indicating its distant relationship with other three serotypes. In the similarity study of the sequences, a highest percentage of 89.5% was observed between DENV–1 and DENV–3, 80.4% between DENV–1 and DENV–2, and DENV–2 and DENV–3, but DENV–4 showed only 79.9% and less than 79% with DENV–2 and DENV–3, respectively (Table 1). There are six figures of pair wise alignment (identity and similarity) sequences in DENV–1 to –4, but only one of them is presented here as a representative (Figure 3).

Table 1
Identity and similarity percentage of E–protein amino acids from DENV serotypes.

Dengue serotypes	Identity	Similarity
DENV–1 with DENV–2	66.1	80.4
DENV–1 with DENV–3	78.4	89.5
DENV–1 with DENV–4	63.3	79.9
DENV–2 with DENV–3	66.3	80.4
DENV–2 with DENV–4	62.8	78.8
DENV–3 with DENV–4	63.4	78.2

3.2. Conserved amino acids in E–proteins

Conserved amino acids, either single or more groups in the sequences, were identified on the basis of multiple sequence alignment of E–proteins. In the multiple sequence alignment, 19 amino acid conserved groups were observed. Number of amino acids in each group was ranging from 4 to 15. A maximum of 15 amino acids in a single group formed a highly significant conserved group. The identical conserved amino acids among the four E–proteins of serotypes were ranging from 62.8% to 78.4%. On the other hand, the similarity percentage of amino acids, after natural mutations that have taken place in the amino acids which had replaced by functionally similar amino acids, was ranging from 78.8% to 89.5% (Table 2, Figure 1).

Table 2
Showing conserved groups of amino acids in E–proteins of DENV serotypes.

Serotype number	Number of amino acids in a group	Number of groups
1	4	2
2	5	6
3	6	6
4	7	3
5	10	1
6	15	1
Total number of groups		19

```
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 495
# Identity:     327/495 (66.1%)
# Similarity:   398/495 (80.4%)
# Gaps:          15/495 ( 3.0%)
# Score: 1841.0

EMBOSS_001         1 MRCVGIGSRDFVEGLSGATWVDVVLEHGSCVTTMAKDKPTLDIELLKTEV     50
                     |||:||.:||||||:||.:|||:|||||||||||||:|||||.||:|||.
EMBOSS_001         1 MRCIGISNRDFVEGVSGGSWVDIVLEHGSCVTTMAKNKPTLDFELIKTEA     50

EMBOSS_001        51 TNPAVLRKLCIEAKISNITTDSRCPTQGEATLVEEQDANFVCRRTFVDRG    100
                     ..||.|||.|||||::||||:||||||||.:|.|||||..|||:.:.||||
EMBOSS_001        51 KQPATLRKYCIEAKLTNTTTESRCPTQGEPSLNEEQDKRFVCKHSMVDRG    100

EMBOSS_001       101 WGNGCGLFGKGSLLTCAKFKCVTKLEGKIVQYENLKYSVIVTVHTGDQHQ    150
                     ||||||||||.::|||.|.|...::|||||.|||:|::::|.|:|::|.
EMBOSS_001       101 WGNGCGLFGKGGIVTCAMFTCKKNMKGKIVQPENLEYTIVITPHSGEEHA    150

EMBOSS_001       151 VGNESTEHGTTATITPQAPTTEIQLTDYGALTLDCSPRTGLDFNEMVLLT    200
                     |||::.:||....||||:..||.:||.||.:|::|||||||||||||||.
EMBOSS_001       151 VGNDTGKHGKEIKITPQSSITEAELTGYGTVTMECSPRTGLDFNEMVLLQ    200

EMBOSS_001       201 MKEKSWLVHKQWFLDLPLPWTSGASTSQETWNRQDLLVTFKTAHAKKQEV    250
                     |::|:||||:||||||||||..||.|....|.:::.|||||..||||||:|
EMBOSS_001       201 MEDKAWLVHRQWFLDLPLPWLPGADTQGSNWIQKETLVTFKNPHAKKQDV    250

EMBOSS_001       251 VVLGSQEGAMHTALTGATEIQTSGTTTIFAGHLKCRLKMDKLTLKGMSYV    300
                     ||||||||||||||||||||.|....:|.|||||||:||||.||||||.
EMBOSS_001       251 VVLGSQEGAMHTALTGATEIQMSSGNLLFTGHLKCRLRMDKLQLKGMSYS    300

EMBOSS_001       301 MCTGSFKLEKEVAETQHGTVLVQIKYEGTDAPCKIPFSTQDEKGVTQNGR    350
                     ||||.||:.||:|||||||::::::|||..:||||||...|.:.....||
EMBOSS_001       301 MCTGKFKIVKEIAETQHGTIVIRVQYEGDGSPCKIPFEIMDLEKRHVLGR    350

EMBOSS_001       351 LITANPIVTDKEKPVNIEAEPPFGESYIVIGAGEKALKLSWFKKGSSIGK    400
                     |||.|||||:|:.||||||||||:|||:||.....|||:||||||||:
EMBOSS_001       351 LITVNPIVTEKDSPVNIEAEPPFGDSYIIIGVEPGQLKLNWFKKGSSIGQ    400

EMBOSS_001       401 MFEATARGARRMAILGDTAWDFGSIGGVFTSVGKLVHQIFGTAYGVLFSG    450
                     |||.|.|||:||||||||||||:||||||:||.:||:||..||..|||
EMBOSS_001       401 MFETTMRGAKRMAILGDTAWDFGSLGGVFTSIGKALHQVFGAIYGAAFSG    450

EMBOSS_001       451 VSWTMKIGIGVLLTWLGLNSRSTSLSMTCIAVGLVTLYLGVMVQA    495
                     |||||||.|||::||:|:||||||||::.:
EMBOSS_001       451 VSWTMKILIGVIITWIGMNSRSTSLSVSLV----------------    480
```

Figure 3. Pairwise alignment of DENV-1 and -2.

3.3. BLAST-P analysis of dengue strains

BLAST-P analysis indicated the relationship of large number of dengue strains so far reported within the dengue serotype. In DENV-1, 100% identity was observed in 15 strains and other strains showed 99% identity. Thus genotypically two groups have formed. In DENV-3, two groups were formed as well but with 100% identity in few strains and 99% in most of the cases. In DENV-2, all

the strains showed only 99% identity. In DENV-4, three genotypic groups were formed with very few strains showing 100% and 99% whereas most of the others showed 98% identity (Tables 3, 4, 5 and 6).

Table 3
BLAST-P analysis for DENV-1 strains.

Serotype number	Accession number	Identity percentage
1	ACN42675.1	100
2	AGT63074.1	100
3	BAN63085.1	100
4	AAW64439.1	100
5	ACO06148.1	100
6	ACN42673.1	100
7	ACA49078.1	100
8	AEF01522.1	100
9	AEB98757.1	100
10	AEB98756.1	100
11	ADI80636.1	100
12	ACY70579.1	100
13	ACQ44314.1	100
14	ACL99154.1	100
15	ACJ04262.1	100
16	AAW28143.1	99
17	ACI15826.1	99
18	AAW23164.1	99
19	AGT63075.1	99
20	AAW28142.1	99

Table 4
BLAST-P analysis for DENV-2 strains.

Serotype number	Accession number	Identity percentage
1	CAR65175.1	99
2	CAR65156.1	99
3	CAR65155.1	99
4	CAR65152.1	99
5	CAR65149.1	99
6	ACA58343.1	99
7	ACA48988.1	99
8	ACA48933.1	99
9	ACA48929.1	99
10	ACA48915.1	99
11	AGO67248.1	99
12	BAH86603.1	99
13	CAR65176.2	99
14	CAR65173.2	99
15	CAR65171.2	99
16	CAR65170.2	99
17	CAR65157.1	99
18	CAR65154.1	99
19	CAR65148.1	99
20	CAR65144.1	99

Table 5
BLAST-P analysis for DENV-3 strains.

Serotype number	Accession number	Identity percentage
1	ACN54322.1	100
2	AEZ01357.1	100
3	ADC35596.1	100
4	ACQ44483.1	100
5	ACQ83328.1	100
6	ADG85695.1	99

Table 5 continued
BLAST-P analysis for DENV-3 strains.

Serotype number	Accession number	Identity percentage
7	ACL99103.1	99
8	AGN94907.1	99
9	AGN94904.1	99
10	AGN94903.1	99
11	AGN94899.1	99
12	AGN94897.1	99
13	AGN94896.1	99
14	ADM63678.1	99
15	ACV04798.1	99
16	ACQ44503.1	99
17	ACQ44502.1	99
18	ACQ44501.1	99
19	ACQ44498.1	99
20	ACQ44380.1	99

Table 6
BLAST-P analysis for DENV-4 strains.

Serotype number	Accession number	Identity percentage
1	AAG30148.1	100
2	AEG28980.1	100
3	ACN54389.1	99
4	ACC68749.1	99
5	ACC68750.1	99
6	ACJ65015.1	99
7	AAU89378.1	99
8	AAB70685.1	99
9	BAC77234.1	99
10	AGE13481.1	98
11	AGE51420.1	98
12	ACS32013.1	98
13	ACH61688.1	98
14	ACS37018.1	98
15	ACS32016.1	98
16	ACS32014.1	98
17	ACQ44408.1	98
18	ACQ44405.1	98
19	ACQ44403.1	98
20	ACQ44402.1	98

4. Discussion

The study of homology is very important from the phylogenetic point of view. Homology indicates the evolution of polymorphic genotypes. Different serotypes evolve from a single basic genotype through the mutation occurred in a single or a group of amino acids in the sequences. This may occur in the middle or at the end of the sequence. This phenotypic change of amino acids may due to the need based and environmentally related process. This development of diversity is found in various kinds of living organisms in the world. Similar evolution also has occurred in the genome and proteome of *Flavivirus* and resulted in the formation of four serotypes

like DENV-1, DENV-2, DENV-3 and DENV-4. The E-protein of the dengue virus is basically formed by 495 amino acid residues. However, the number of amino acids in all the four types is not similar. All 495 amino acids are found in DENV-1 and DENV-4. DENV-3 showed 493 after the deletion of two amino acids in the sequence and 480 amino acids were found in the case of DENV-2 due to the deletion of 15 amino acids either in the middle or at the end of the sequence. Such a kind of analysis and in depth study in this area has not carried out in the recent past. Fu *et al.*[8], have studied the homology of E-proteins dengue serotypes and reported that the four serotypes of DENV (DENV-1 to -4) shared approximately 65%-75% homology at the amino acid level. In the present study, the homology among the dengue serotypes was found to be 63.0% to 78.7%.

In the present study from the phylogram, it has been concluded that dengue serotypes 1, 2 and 3 formed a single group and the fourth group was distant from the above three serotypes, forming two clusters. Based on the BLAST-P analysis of various strains of four serotypes of DENV, the strains of DENV-1, DENV-2 and DENV-3 showed 99% homology. On the other hand, DENV-4 showed only 98% homology.

Conservation of amino acids in a group of serotypes or species is an indication of phylogenetic closeness among them. Conservation of amino acid sequences is indicated by the presence of identical amino acid residues. Sequence similarities serve as evidence for structural and functional conservation as well as the presence of evolutionary relationships between the sequences[8,9]. In the present study, through the multiple sequence alignment, 19 amino acid conserved groups were observed. Number of amino acids in each group was ranging from 4 to 15. A maximum of 15 amino acids in a single group formed a highly significant conserved group.

E-proteins of dengue virus serotypes were not similar but homologous with each other. Homology of identical amino acid sequences in the four dengue serotypes varied from 62.8% to 78.4% in their amino acid sequences. However, the similarity percentage of amino acids, after natural mutations that have taken place in the amino acids after replacement by functionally similar amino acids, was ranging from 78.8% to 89.5%. Among the four serotypes of dengue virus, DENV-1, DENV-2 and DENV-3 formed a single group whereas DENV-4 alone formed a single cluster which showed a distant relationship with other serotypes, based on the multiple sequence alignment of phylogram. Based on the BLAST-P analysis of various strains of four serotypes, most of the DENV-1, DENV-2 and DENV-3 strains showed 99% homology whereas DENV-4 showed only 98% homology. In the multiple sequence alignment, 19 conserved amino acid groups were observed. Number of amino acids in each group was ranging from 4 to 15. A maximum of 15 amino acid sequence in a single group formed a highly significant conserved group.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgements

The authors are grateful to the University Grants Commission New Delhi, India for the financial assistance given through the Major Research Project [MRP R. No: 41-472/2012(SR)]. The authors specially express their thanks to the management of A. Veeriya Vandayar Memorial Sri Pushpam College (Autonomous), Poondi, for providing necessary facilities and support to carry out this work.

References

[1] Whitehead SS, Blaney JE, Durbin AP, Murphy BR. Prospects for a dengue virus vaccine. *Nat Rev Microbiol* 2007; **5**: 518-528.

[2] Lindenbach BD, Thiel HJ, Rice CM. Flaviviridae: the viruses and their replication. In: Knipe DM, Howley PM, editors. *Fields virology*. Vol 1. 5th ed. Philadelphia: Lippincott-Raven Publishers; 2007, p. 1101-1152.

[3] Federhen S. The NCBI taxonomy database. *Nucleic Acids Res* 2012; **40**(D1): D136-D143.

[4] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389-3402.

[5] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet* 2000; **16**(6): 276-277.

[6] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; **215**(3): 403-410.

[7] Zhang Z, Miller W, Schäffer AA, Madden TL, Lipman DJ, Koonin EV, et al. Protein sequence similarity searches using patterns as seeds. *Nucl Acids Res* 1998; **26**: 3986-3990.

[8] Fu J, Tan BH, Yap EH, Chan YC, Tan YH. Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90). *Virology* 1992; **188**(2): 953-958.

[9] Gross L. Are "ultraconserved" genetic elements really indispensable. *PLoS Biol* 2007; doi: 10.1371/journal.pbio.0050253.