Computational Design of Ancestral and Consensus Asian Dengue Envelope Protein for Vaccine Candidate

Rahmat Azhari Kemal¹, Jeremias Ivan², Eric Bernardus Lili Sandjaja³, and Audi Putra Santosa³

¹Department of Medical Biology, Faculty of Medicine, Universitas Riau, Pekanbaru, Indonesia
²Department of BioInformatics, School of Life Sciences, Indonesia International Institute for Life Sciences. Pulomas Barat Kav. 88, Jakarta, Indonesia
³Department of BioTechnology, School of Life Sciences, Indonesia International Institute for Life Sciences. Pulomas Barat Kav. 88, Jakarta, Indonesia

Abstract
Dengue is a mosquito-borne viral disease of which incidence has rapidly increased in the last few years. Despite the recent development of a licensed dengue vaccine, safer and more efficacious dengue vaccine still needs to be developed. Dengue virus has four antigenically and genetically distinct serotypes. Ancestral sequence reconstruction (ASR) and consensus sequence (CS) might be able to overcome antigenic distinction between those four serotypes. Envelope (E) protein is responsible for a wide range of dengue virus biological activities. Domain III of the E protein (EDIII) plays a role in receptor binding for viral entry and inducing protective immunity against the dengue virus. We utilised bioinformatics software to computationally design ancestral and consensus sequences of Asian dengue E protein. E protein sequences of 987 DENV strains and 5 outgroups were retrieved from GenBank. We constructed ancestral and consensus sequences for each serotype. For ASR, ancestral sequences were gradually designed to construct ancestral sequence for all serotypes using MEGA X. For CS, all four consensus sequences were directly used to construct consensus sequence for all serotypes using UGENE 1.32. Phylogenetic tree consisting existing dengue sequences as well as ancestral and consensus sequences were visualised using FigTree 1.4.4. All ancestral and consensus sequences were analysed for conserved motifs, especially in domain III region. ASR sequences were closer to the centre of phylogenetic tree branches while consensus sequences were located among natural isolates. Further CD4 T cell immunogenicity prediction on domain III (EDIII) showed that both ASR and consensus EDIII have the two-highest combined immunogenicity scores. These sequences are potential for further in vitro and in vivo studies as dengue vaccine candidate.

Keywords: ancestral sequence, consensus, dengue, envelope protein, vaccine
1. Introduction

Dengue is a wide-spread disease caused by the dengue virus (DENV), a member of the Flaviviridae family [1]. The virus infects human, transmitted by *Aedes aegypti* and *Aedes albopictus* mosquitoes as the vector [2]. Dengue virus (DENV) comprises four different stereotypes: DENV1, DENV2, DENV3, and DENV4. All of them are in circulation in Indonesia urban areas [3 - 4]. In 2004-2010 Indonesia became the second rank country with highest dengue prevalence after Brazil; furthermore, the number kept on increasing until it reached 129,650 cases in 2015 [5 - 6]. Due to the lack of available anti-DENV drug [7], prevention is considered as the best management strategy.

Up until now, Food and Drug Administration (FDA) has only approved one vaccine against dengue fever, Dengvaxia. Although it reduces the rates of getting severe virologically confirmed dengue (VCD) among seropositive patients, it produces the opposite effect for those who have not been infected before [8 - 9]. It might be caused by Antibody-Dependent Enhancement (ADE) mechanism of the virus, which eventually increases its infectivity [10 - 11]. Such condition resulted in a controversy in Philippine, where many parents blamed the nation-wide dengue vaccination for their children's death [9, 12]. Even though those accusations have not been scientifically-confirmed, the limited utilization of the vaccine itself shows the need for better alternatives. According to McArthur et al. [13], there are several main factors that need to be assessed in developing dengue vaccine, including the viral immunology and epidemiology.

DENV is composed of a single-stranded RNA genome packed inside a core protein surrounded by scaffolds and packed by a lipid envelope. The genome encodes a polyprotein which later cleaves into three structural proteins: pre-membrane/membrane (prM/M), envelope (E), and capsid (C) alongside with two other structural (prM and C protein) and seven nonstructural (NS) proteins [14 - 15]. Envelope (E) protein is an antigenic protein that has been extensively studied. E protein is responsible for a wide range of dengue virus biological activities.

E protein is composed of three domains (domain I, II, and II). Domain III of the E protein (EDIII) plays a role in receptor binding for viral entry and inducing protective immunity against the virus [1]. Domain III is able to trigger receptor binding that will lead to the virus entry and replication of virus in the host [14]. EDIII-specific antibodies can give protection to DENV infection and considered as protection correlated for anti-DENV vaccines [16 - 17]. Depending on the specificity, concentration and affinity of antibodies targeting the E protein, the antibodies may block virus infection or promote enhancement of infection.
via cell entry mediated by Fc-γ receptors [18 - 19]. EDIII domain is also used to develop recombinant protein vaccines [1].

Due to the co-circulation of all four DENV serotypes in Indonesia, a vaccine capable to induce immunity against all serotypes is highly needed. One consideration in dengue vaccine is antibody-dependent enhancement. The vaccine needs to induce strong immunity against all four serotypes. The presence of low affinity suboptimal neutralising antibodies against heterotypic DENV can cause ADE, which is enhanced viral uptake and replication in macrophages and monocytes [20].

Ancestral and consensus sequences can be designed to minimise the genetic differences between vaccine strains and naturally-occurring strains [21]. Ancestral sequence is computationally derived to represent the most recent common ancestor of a pool of viruses [22]. It maximises the likelihood that selected amino acid represents universally shared evolutionary forces. As for the consensus sequence, it incorporates the most common amino acid at each position [21]. Construction of ancestral and consensus sequences as vaccine candidate has been utilised for other highly divergent viruses, such as hepatitis C virus HCV [21], avian influenza H5N1 [22 - 23], and HIV [24]. Considering these, we have utilised available bioinformatics software to design ancestral and consensus E protein of DENV.

2. Method

2.1. Sequence collection

We retrieved 987 E protein sequences of DENV and 5 of Japanese Encephalitis Virus (JEV) from National Centre for Biotechnology Information (NCBI). All viruses were isolated from human patients in Asia. In order to avoid certain country representation, only 10 isolates were retrieved from each country every year. The number of isolates per sampling location and time were shown in Table 1 and 2, respectively.

2.2. Initial phylogenetic tree construction

Multiple Sequence Alignment (MSA) of the initial sequences was done by using MUSCLE [25], followed by phylogenetic tree construction by using Maximum-Likelihood (ML) algorithm with 250 bootstrapping in MEGA X [26].
2.3. Ancestral and consensus sequence reconstruction

The ancestral sequence of the virus was constructed by using ML algorithm based on the generated phylogenetic tree in MEGA X. On the other hand, MSAs were done for each serotype by using MUSCLE, followed by strict serotype-based consensus sequence
reconstruction by using UniPro UGENE 1.32 [27] with 50% threshold. The final consensus sequence was inferred from the consensus of the four serotypes by using the same parameter.

2.4. Sequence analysis

Multiple Sequence Alignment (MSA) of the initial, ancestral, and consensus sequences was done by using MUSCLE, followed by phylogenetic tree construction by using ML algorithm with 250 bootstrapping in MEGA X. The tree was then visualized by using FigTree 1.4.4 [28].

Analysis of EDIII of ancestral and consensus sequences were conducted by MSA in ClustalW. Next, CD4 T cell immunogenicity predictions on EDIII of ancestral and consensus sequences were done by using IEDB T Cell Epitope Prediction Tools [29 - 30], which was available at http://tools.iedb.org/CD4episcore/.

3. Results and Discussion

3.1. Sequence collection

In this study, we retrieved 987 sequences of DENV E protein comprised of four serotypes: DENV1 (324), DENV2 (365), DENV3 (233), and DENV4 (65). All of the sequences came from Asia with isolation time ranged from 1956 to 2017. In addition, we also retrieved five E protein of JEV to be used as an outgroup in the phylogenetic tree construction.

3.2. Phylogenetic position of constructed sequences

Initial phylogenetic tree showed that DENV is well-clustered between serotypes (Figure 1). DENV1 had the closest relationship with DENV3, followed by DENV2 and DENV4, respectively. It was also observable that the branching within-serotype was very shallow, whereas the ones between-serotypes were deep. This might infer that each DENV serotype had majorly evolved since their divergences. Incorporation of ancestral and consensus sequences also supported this structure (Figure 2).

Based on the Figure, the ASR sequences were located near to the divergence tips of each serotype [31 - 32], while the consensus sequences were found among natural isolates. This result was similar to the previous studies [33 - 34]. Specifically, it might
correlate with the consensus threshold and genetic evolution of each serotype. As mentioned before, the short within-serotype branching might infer low genetic diversity of E protein between DENV isolates with the same serotype. In comparison, highly divergent viruses would have longer branches, indicating a greater genetic distance between isolates [22 - 24]. However, as we set the threshold to be 50% with a strict algorithm, this produced gaps that might reduce the representation for all isolates [35]. Nevertheless, both Figures showed that the E protein of DENV had clear clustering between serotypes, where ancestral and consensus sequences summarized the genetic variation between the isolates.

![Figure 1](Initial phylogenetic tree of DENV based on E protein. Coloring was based on serotypes: red (DENV1), yellow (DENV2), green (DENV3), blue (DENV4), black (JEV; outgroup).)
Figure 2: Phylogenetic tree of DENV based on E protein. Coloring was based on serotypes: red (DENV1), yellow (DENV2), green (DENV3), blue (DENV4), black (JEV; outgroup), pink (ancestral and consensus sequences).

3.3. Analysis of EDIII in constructed sequences

EDIII of ASR and consensus sequences are 103-amino acids long, with conserved N-terminal of KGMSY and C-terminal of KGSS. All sequences contain cysteine residue in positions of 8 and 39 [1]. CD4 T cell immunogenicity prediction on EDIII (Table 3) showed that both ASR and consensus sequences for all 4 serotypes had the two-highest combined immunogenicity scores. ASR and consensus sequences for DENV1 and for DENV2 were also the ten peptides with highest combined immunogenicity.

Sequences have conserved linear B cell neutralising epitope 16KEVAETQHGT25. Substitution in 18V and 19A residues are putative antigenically silent [36 -37]. ETQH is also an epitope recognised by 2H12, cross-reactive antibody capable to neutralise all
TABLE 3: Peptides with the ten highest combined immunogenicity score.

| Sequence          | Peptide              | Score     |
|-------------------|----------------------|-----------|
| Consensus -- All  | RLITANPIVTNKESP      | 49,67564  |
| ASR -- All        | RDVNKKKVVGRLITV      | 49,48832  |
| Consensus -- All  | RDVNKKKVNGRLITA      | 49,07948  |
| ASR -- DENV 1     | RLITANPIVTDEKEP      | 49,05320  |
| ASR -- DENV 1 & 3 | RLITANPIVTDEKEP      | 49,05320  |
| Consensus -- DENV 1| RLITANPIVTDEKEP      | 49,05320  |
| Consensus -- DENV 2| KGMSYSMCTGFKV      | 48,97028  |
| ASR -- DENV 2     | MDLEKRHVLGRLITV      | 48,20512  |
| Cons -- DENV 2    | MDLEKRHVLGRLITV      | 48,20512  |
| ASR -- DENV 2     | RLITVPNIVTEKDSP      | 48,14608  |

four DENV serotypes [38]. These sequences are potential for further in vitro and in vivo studies as dengue vaccine candidate.

Construction of ancestral and consensus sequences has been utilised to design vaccine for other highly divergent viruses. Burke et al. [21] compared the capacity of ancestral, consensus, and natural strains of HCV to expand CD8 T cell populations in vitro. Ancestral sequence expanded CD8+ T cells into broader and more robust recognition towards highly diverse circulating HCV strains. On the contrary in HIV vaccine design, Kothe et al. [24] observed that while pseudovirions containing ancestral and consensus HIV envelope proteins induced comparable immune responses, the consensus was significantly more infectious than the ancestral.

4. Conclusion

We have designed ancestral and consensus sequence of DENV EDIII. However, ancestral and consensus sequences reconstruction techniques are highly dependent on several aspects, including the number of sequences, software, and parameters that were used. Nevertheless, the results of this study warrant further in vitro and in vivo studies to analyse immunogenic properties of these computationally designed dengue vaccine candidate.

Conflict of Interest

The authors declare no conflict of interest.
References

[1] H. Fahimi, M. Mohammadipour, H. Kashani, F. Parvini, & M. Sadeghizadeh. “Dengue viruses and promising envelope protein domain III-based vaccines,” Appl. Microbiol. Biotechnol., vol. 102, no. 7, pp. 2977-2996, 2018.

[2] S. Hasan, S. F. Jamdar, M. Alalowi, & S. M. Al Ageel Al Beaiji. “Dengue virus: A global human threat: Review of literature,” J. Int. Soc. Prev. Community. Dent., vol. 6, no. 1, p. 1, 2016.

[3] R. T. Sasmono, A. F. Taurel, A. Prayitno, B. Sitompul, B. Yohan, R. F. Hayati, A. Bouckenooghe, S. R. Hadinegoro, J. Nealon. “Dengue virus serotype distribution based on serological evidence in pediatric urban population in Indonesia,” PLoS. Negl. Trop. Dis, vol. 12, no. 6, 2018.

[4] World Health Organization. “Dengue,” Retrieved from http://www.searo.who.int/entity/vector_borne_tropical_diseases/data/data_factsheet/en/index1.html, 2019.

[5] A. Maula, A. Fuad, & A. Utarini. “Ten-years trend of dengue research in Indonesia and South-east Asian countries: a bibliometric analysis,” Global Health Action, vol. 11, no. 1, p. 1504398, 2018.

[6] Departemen Kesehatan Republik Indonesia, “Profil Kesehatan Indonesia 2014,” Retrieved from http://www.depkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/profil-kesehatan-indonesia-2014.pdf, 2014.

[7] Y. S. Tian,, Y. Zhou, T. Takagi,, M. Kameoka, & N. Kawashita. “Dengue virus and its inhibitors: A brief review,” Chemical and Pharmaceutical Bulletin, vol. 66, no. 3, pp. 191-206. 2018.

[8] S. Sridhar, A. Luedtke, E. Langevin, M. Zhu, M. Bonaparte, & T. Machabert, S. Savarino, B. Zambrano, A. Moureau, A. Khromava, Z. Moodie, T. Westling, C. Mascareñas, C. Frago, M. Cortés, D. Chansinghakul, F. Noriega, A. Bouckenooghe, J. Chen, S. P. Ng, P. B. Gilbert, S. Gurunathan, C. A. DiazGranados. “Effect of Dengue Serostatus on Dengue Vaccine Safety and Efficacy,” N. Engl. J. Med., vol. 379, no. 4, pp. 327-340, 2018.

[9] R. Voelker. “Dengue Vaccine Gets the Nod,” JAMA, vol. 321, no. 21, p. 2066, 2019

[10] J. Flipse, M. A. Diosa-Toro, T. E. Hoornweg, D. P. I. van de Pol, S. Urcuqui-Inchima, & J. M. Smit. “Antibody-Dependent Enhancement of Dengue Virus Infection in Primary Human Macrophages; Balancing Higher Fusion against Antiviral Responses,” Scientific Reports, vol. 6, no. 1, 2016.

[11] S. Yasmin & M. Mukerjee. “The Dengue Debacle,” Scientific American, vol. 320, no. 4, pp. 38-47, 2019.
[12] K. Fatima & N. I. Syed. "Dengvaxia controversy: impact on vaccine hesitancy," J. Glob. Health, vol. 8, no. 2, p. 020312, 2018.

[13] M. McArthur, M. Sztein, & R. Edelman. "Dengue vaccines: recent developments, ongoing challenges and current candidates," Expert Rev. Vaccines, vol. 12, no. 8, pp. 933-953, 2013.

[14] D. L. N. F. Maeda, M. T. Batista, L. R. Pereira, M. de Jesus Cintra, J. H. Amorim, C. Mathias-Santos, S. A. Pereira, S. B. Boscardin, S. dos Ramos Silva, E. L. Faquim-Mauro, V. B. Silveira, D. B. L. Oliveira, S. A. Johnston, L. C. de Souza Ferreira, & J. F. Rodriguez. "Adjuvant-mediated epitope specificity and enhanced neutralizing activity of antibodies targeting dengue virus envelope protein," Front. Immunol., vol. 8, p. 1175, 2017.

[15] S. Noisakran, N. Onlamoon, P. Songprakhon, H. Hsiao, K. Chokephaibulkit, & G. Perng. "Cells in Dengue Virus Infection In Vivo," Advances In Virology, vol. 2010, pp. 1-15, 2010.

[16] C. Y. Chiang, M. H. Huang, C. H. Hsieh, M. Y. Chen, H. H. Liu, J. P. Tsai, Y. S. Li, C. Y. Chang, S. J. Liu, P. Chong, C. H. Leng, & H. W. Chen. "Dengue-1 envelope protein domain III along with PELC and CpG oligodeoxynucleotides synergistically enhances immune responses," PLoS Negl. Trop. Dis., vol. 6, no. 5, p. e1645, 2012.

[17] S. Sukupolvi-Petty, S. K. Austin, W. E. Purtha, T. Oliphant, G. E. Nybakken, J. J. Schlesinger, J. T. Roehrig, G. D. Gromowski, A. D. Barrett, D. H. Fremont, M. S. Diamond. "Type-and subcomplex-specific neutralizing antibodies against domain III of dengue virus type 2 envelope protein recognize adjacent epitopes," J. Virol., vol. 81, no. 23, pp. 12816-12826, 2007.

[18] S. M. Lok, V. Kostyuchenko, G. E. Nybakken, H. A. Holdaway, A. J. Battisti, S. Sukupolvi-Petty, D. Sedlak, D. H. Fremont, P. R. Chipman, J. T. Roehrig, M. S. Diamond, R. J. Kuhn, & M. G. Rossmann. "Binding of a neutralizing antibody to dengue virus alters the arrangement of surface glycoproteins," Nat. Struct. Mol. Biol., vol. 15, no. 3, p. 312, 2008.

[19] M. Poggianella, J. L. S. Campos, K. R. Chan, H. C. Tan, M. Bestagno, E. E. Ooi, & O. R. Burrone. "Dengue E protein domain III-based DNA immunisation induces strong antibody responses to all four viral serotypes," PLoS Negl. Trop. Dis., vol. 9, no. 7, 2015.

[20] J. Torresi, G. Ebert, & M. Pellegrini. "Vaccines licensed and in clinical trials for the prevention of dengue," Hum. Vaccin. Immunother., vol. 13, no. 5, pp. 1059-1072, 2017.

[21] K. P. Burke, S. Munshaw, W. O. Osburn, J. Levine, L. Liu, J. Sidney, A. Sette, S. C. Ray, & A. L. Cox. "Immunogenicity and cross-reactivity of a representative ancestral
sequence in hepatitis C virus infection," J. Immunol., vol. 188, no. 10, pp. 5177-5188, 2012.

[22] M. F. Ducatez, J. Bahl, Y. Griffin, E. Stigger-Rosser, J. Franks, S. Barman, D. Vijaykrishna, A. Webb, Y. Guan, R. G. Webster, G. J. D. Smith, R. J. Webby. "Feasibility of reconstructed ancestral H5N1 influenza viruses for cross-clade protective vaccine development," Proc. Nat. Acad. Sci., vol. 108, no. 1, pp. 349-354, 2010.

[23] M. W. Chen, T. J. R. Cheng, Y. Huang, J. T. Jan, S. H. Ma, L. Y. Alice, W. Chi-Huey, & D. D. Ho. "A consensus–hemagglutinin-based DNA vaccine that protects mice against divergent H5N1 influenza viruses," Proc. Natl. Acad. Sci., vol. 105, no. 36, pp. 13538-13543, 2008.

[24] D. L. Kothe, Y. Li, J. M. Decker, F. Bibollet-Ruche, K. P. Zammit, M. G. Salazar, Y. Chen, Z. Weng, E. A. Weaver, F. Gao, B. F. Haynes, G. M. Shaw, B. T. Korber, B. H. Hahn. "Ancestral and consensus envelope immunogens for HIV-1 subtype C," Virology, vol. 352, no. 2, pp. 438-449, 2016.

[25] R. Edgar. "MUSCLE: multiple sequence alignment with high accuracy and high throughput," Nucleic Acids Res., vol. 32, no. 5, pp. 1792-1797, 2004.

[26] S. Kumar, G. Stecher, M. Li, C. Knyaz, & K. Tamura. "MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms," Mol. Biol. Evol., vol. 35, no. 6, pp. 1547-1549, 2018.

[27] K. Okonechnikov, O. Golosova, & M. Fursov. "Unipro UGENE: a unified bioinformatics toolkit," Bioinformatics, vol. 28, no.8, pp. 1166-1167, 2012.

[28] A. Rambaut. "FigTree: tree figure drawing tool version 1.4.4," Available at: http://tree.bio.ed.ac.uk/software/figtree, 2012.

[29] S. Dhanda, E. Karosiene, L. Edwards, A. Grifoni, S. Paul, M. Andreatta, D. Weiskopf, J. Sidney, M. Nielsen, B. Peters, & A. Sette. "Predicting HLA CD4 Immunogenicity in Human Populations," Front. Immunol., vol. 9, p. 1369, 2018.

[30] S. Paul, C. S. Lindestam Arlehamn, T. J. Scriba, M. B. Dillon, C. Oseroff, D. Hinz, D. M. McKinney, S. C. Pro, J. Sidney, B. Peters, & A. Sette. "Development and validation of a broad scheme for prediction of HLA class II restricted T cell epitopes," J. Immunol. Methods., vol. 422, pp. 28-34, 2015.

[31] V. Risso, J. Gavira, D. Mejia-Carmona, E. Gaucher, & J. Sanchez-Ruiz. "Hyper-stability and Substrate Promiscuity in Laboratory Resurrections of Precambrian β-Lactamases," J. Am. Chem. Soc., vol. 135, no. 8, pp. 2899-2902, 2013

[32] J. B. Joy, R. H. Liang, R. M. McCloskey, T. Nguyen, & A. F. Y. Poon. "Ancestral Reconstruction," PLoS. Comput. Biol., vol. 12, no. 7, p. e1004763, 2016.
[33] G. S. Kesturu, Colleton, B. A. Colleton, Y. Liu, L. Heath, O. S. Shaikh, C. R. Jr. Rinaldo, & R. Shankarappa. “Minimization of genetic distances by the consensus, ancestral, and center-of-tree (COT) sequences for HIV-1 variants within an infected individual and the design of reagents to test immune reactivity,” Virology, vol. 348, no. 2, pp. 437-448, 2006.

[34] H. Ross, D. Nickle, Y. Liu, L. Heath, M. Jensen, A. Rodrigo, & J. Mullins. “Sources of Variation in Ancestral Sequence Reconstruction for HIV-1 Envelope Genes,” Evol. Bioinformatics, vol. 2, 2007.

[35] T. Schneider. “Consensus Sequence Zen,” Appl. Bioinformatics, vol. 1, no. 3, pp. 111-119, 2002.

[36] X. Q. Li, L. W. Qiu, Y. Chen, K. Wen, J. P. Cai, J. Chen, Y. X. Pan, J. Li, D. M. Hu, Y. F. Huang, L. D. Liu, X. X. Ding, Y. H. Guo, & X. Y. Che. “Dengue virus envelope domain III immunization elicits predominantly cross-reactive, poorly neutralizing antibodies localized to the AB loop: implications for dengue vaccine design,” J. Gen. Virol., vol. 94, no. 10, pp. 2191-2201, 2013.

[37] Y. Lin, K. Wen, Y. Guo, L. Qiu, Y. Pan, L. Yu, B. Di, & Y. Chen. “Mapping of the B cell neutralizing epitopes on ED III of envelope protein from Dengue virus,” Bing Du Xue Bao (Chinese Journal of Virology), vol. 31, no. 6, pp. 665-673, 2015.

[38] C. M. Midgley, A. Flanagan, H. B. Tran, W. Dejnirattisai, K. Chawansuntati, A. Jumnainsong, W. Wongwiwat, T. Duangchinda, J. Mongkolsapaya, J. M. Grimes, G. R. Screaton. A. “Structural Analysis of a Dengue Cross-Reactive Antibody Complexed with Envelope Domain III Reveals the Molecular Basis of Cross-Reactivity,” J. Immunol. vol. 188, no. 10, pp. 4971-4979, 2012.