JMID/
Journal of Microbiology and Infectious Diseases

2020; 10 (1):31-46
doi: 10.5799/jmid.700510

RESEARCH ARTICLE

Identification of Novel Vaccine Candidates against Yellow Fever Virus from the Envelope Protein: An Insilico Approach

Hind Abdelrahman Hassan[1,2], Khoubieb Ali Abdelrahman[3], Nasr Mohammed Nasr[1,2], Yassir A. Almofti[1]

[1]*Department of Molecular Biology and Bioinformatics, College of Veterinary Medicine, University of Bahri, Khartoum, Sudan*
[2]*Department of Microbiology, College of Medical laboratory Science, Omdurman Ahlia University, Omdurman, Sudan*
[3]*Department of Pharmaceutical Technology, College of Pharmacy, University of Medical Science and Technology, Khartoum- Sudan*

ABSTRACT

Objectives: Yellow fever virus (YFV) is an enveloped positive sense RNA virus. It is the causative agent of the mosquito-borne disease yellow fever. The aim of this study was to design multi epitopes vaccine for YFV from envelope protein eliciting humoral and cellular immunity.

Methods: Twenty six YFV strains envelope proteins were retrieved from NCBI. The immune epitope database analysis resources (IEDB) were used for epitopes prediction.

Results: Eleven epitopes successfully passed all B cell prediction tools, among them four epitopes 33VMAPDKPSL41, 72DKCP77, 236PPHA239 and 385LTYQ388 demonstrated higher score in Emini and Kolaskar and tongaonker software. Thus were proposed as B cells epitopes. For T cells; 28 epitopes interacted with MHC-I and the best recognizable epitopes were 471MTMSMSMIL479, 363VLIEVNPPF371, 33VMAPDKPSL41 and 226REMHHLVEF234. For MHC-II ninety epitopes were predicted and the best epitopes were 284RVKLSALTL292, 363VLIEVNPPF371, 479LVGVIMMFL487 and 226REMHHLVEF234. Strikingly the epitope 33VMAPDKPSL41 successfully interacted with both B and T cells. Also 363VLIEVNPPF371 and 226REMHHLVEF234 demonstrated successful interaction with T cells. The population coverage was 84.66% and 99.91% for MHC-I and MHC-II epitopes, respectively, and 99.99% for all T cells epitopes.

Conclusions: Taken together nine epitopes successfully proposed as vaccine candidate against YFV. In vivo and in vitro clinical trials studies are required to elucidate the effectiveness of these epitopes as vaccine. *J Microbiol Infect Dis 2020; 10(1):31-46.*

Keywords: *Yellow fever virus; NCBI; IEDB; Immunoinformatics; Insilico vaccine*

INTRODUCTION

Yellow fever virus (YFV) is an enveloped positive sense single stranded RNA virus. It belongs to flavivirus genus and member of Flaviviridae family [1]. It transmitted to human and other primates by infected female mosquito mainly by *Aedes aegypti* in tropical area of South America and Africa [2]. The yellow fever virus is the causative agent of the yellow fever which is a short duration disease generally characterized by fever, chills, loss of appetite, nausea, muscle pain and headache [2]. However, in 15% of infected people yellow skin occurred due to liver damage as well as bleeding and kidney problems [3].

Worldwide about 600 million people live in endemic areas of the disease and 90% of the infection occurred in the African continent [4]. In 2016 a large outbreak occurred in Angola and spread to neighboring countries and 11 cases

Correspondence: Yassir A. Almofti, University of Bahri, College of Veterinary Medicine, Department of Molecular Biology and Bioinformatics, Khartoum, Sudan
Email: yamofti99@gmail.com
Received: 31 May 2019 Accepted: 26 January 2020

were reported in China which considered as the first incidence of the disease in Asia [2, 4]. In Sudan the two recent outbreaks were in 2003 and 2005 [4]. In late 2016 a large outbreak began in Brazil [5]. In 2017 the sylvan outbreaks spread into the Brazilian coast [6,7]. Phylogenetic analysis has identified seven genotypes of yellow fever viruses. They assumed to be differently adapted to humans and to the vector *A. aegypti* [2,8]. Five genotypes occurred only in Africa and distributed in Angola, Central/East Africa, East Africa, West Africa I, and West Africa II [9]. West Africa genotype I found in Nigeria and the surrounding areas [4,9]. This genotype seems to be more virulent or infectious as it often associated with major outbreaks. While the three genotypes in East and Central Africa were rarely occur. The East African genotype has occurred in Kenya (1992–1993) and Sudan (2003 and 2005) [5,6].

In South America, two genotypes have been identified (South American genotypes I and II) [10]. Based on phylogenetic analysis these two genotypes appear to have originated in West Africa and later introduced into Brazil [9,10]. Genotype I have been divided into five subclasses. Bayesian analysis of genotype I and II showed that genotype I accounted for virtually all the current infections in Brazil, Colombia, Venezuela, and Trinidad and Tobago [11,12]. While genotype II accounted for all cases in Peru and other parts of the world [11,12].

Vaccination against yellow fever virus was developed in 1930s by using live attenuated YF-17D virus [13]. The vaccine conferred protection in more than 95% of the vaccinated population for up to 40 years [14,15]. The vaccine was successful and conferred protection for many years against YFV infection. However recently rare cases of fatal vaccine-associated adverse events were reported [16]. These adverse events include neurotropic diseases characterized by post-vaccinal encephalitis [16,17]. Moreover, vaccine viscerotropic diseases characterized by pan-systemic infections with liver damage similar to infection by wild type were also reported [16-18]. Furthermore, the problems of contraindication risk of the live attenuated vaccine in groups such as pregnant women, infants, elderly people, immunosuppressed and people who are sensitive to eggs were also reported [19]. Therefore, new vaccination strategies highly required to solve problems associated with live or attenuated vaccines.

YFV is a 40-50 nm enveloped virus, single stranded RNA around 11.000 long with a positive-sense [1]. The virus has single open reading frame encoding a poly protein which is cut by host proteases into structural proteins C, prM, E and non-structural proteins (NS2A, NS2B, NS3, NS4A, NS4B and NS5) [19]. The Envelope protein is responsible for virus attachment to specific receptors on target cells [20]. In addition to that many studies showed that the envelope protein is the best candidate to elicit immune system and important for recombinant vaccine production [21-23].

Insilico modeling of epitopes protein would help in manufacture of peptide vaccine, which is highly immunogenic and with minimal allergenic effects. Therefore, this study was conducted using immunoinformatic approaches to design a multi epitopes vaccine against yellow fever virus using envelope protein as an immunogen.

METHODS

Protein sequence retrieval

A total of 26 envelope proteins of yellow fever virus strains were retrieved from NCBI database (NCBI, RRID: SCR_006472) URL: http://www.ncbi.nlm.nih.gov in March 2018 from (https://www.ncbi.nlm.nih.gov/protein/?term=envelope + protein + YFV). These 26 envelop protein sequences were retrieved from different parts of the world. The accession numbers of the retrieved strains were (NP_740305 as a reference sequence, AAX47570, AAX47569, AAX47568, AAA92706, AAA92705, AAA92704, AAA92703, AAA92702, AAA92701, AAA92700, AAA92699, AAA92698, AAA92697, AAA92696, AAA92695, AAA92694, AAA92693, AAA92692, AAA92691, AAY68350, AAR86693, AAY68346, AAY68347, AAY68348 and AAY68349)

Phylogenetic relationship

The retrieved sequences were submitted to phylogeny.fr server (Phylogeny.fr, RRID: SCR_010266) URL: (http://www.phylogeny.fr/) to determine the common ancestor of each strain and the genetic relationships among strains.

Determination of conserved regions

The retrieved protein sequences were aligned for finding the conserved regions among yellow fever envelope protein variants using Clustal W multiple alignment program which implemented in the offline Bioedit software (version 7.2.5.0) [24]. The conserved epitopes were then analyzed by different prediction tool in immune epitope database IEDB (Immune Epitope Database and Analysis Resource, RRID: SCR_006604) URL: (http://www.immuneepitope.org/)

B cell epitope prediction

Tools from IEDB were used to identify the B cell epitope prediction, including Bepipred linear epitopes analysis, Emini surface accessibility and Kolaskar and Tongaonkar for antigenicity scale.

Prediction of linear epitopes

Bepipred linear epitope prediction tool in IEDB (http://toolsiedb.ofg/bcell/) were used to identify linear epitopes from the envelope protein. The epitopes were predicted from conserved regions with a default threshold value -0.043.

Prediction of surface accessibility

Epitopes surface accessibility was determined using Emini surface accessibility prediction tool in IEDB at (http://tools.iedb.org/bcell/). The surface accessible epitopes were predicted from the conserved region holding the default threshold value 1.000.

Prediction of epitopes antigenicity sites

The kolaskar and tongaonkar antigenicity tool in IEDB was used to determine the antigenic epitopes with a default threshold value of 1.028 (http://tools.iedb.org/bcell/)

MHC epitopes prediction

IEDB server (http://www.iedb.org) was used through specific tools to determine MHC-1 and MHC-II binding epitopes. This server uses specific scoring IC50 (inhibitory concentration 50) to predict epitopes that bind to different MHC class I and MHC class II alleles.

MHC class I epitope prediction

The method for prediction of MHC class I affinity was tested on large set of quantitative peptides of MHC class I measurement affinity on the IEDB (http://tools.iedb.org/mhci/). The Prediction methods achieved by Artificial Neural Network (ANN), Stabilized Matrix Method (SMM) or Scoring Matrices derived from Combinatorial Peptide Libraries [25]. By using artificial neutral network (ANN) method the length of epitopes was set as 9mers. All conserved epitopes bound with score equal to or less than 300 IC50 were chosen for further analysis.

MHC class II epitope prediction

MHC-II binding tool from IEDB was used by applying NN align as prediction method (http://tools.iedb.org/mhcii/). IC50 prediction value equal to or less than 3000 was used to predict epitopes for MHC-II [26]. All conserved epitopes were chosen for more analysis.

Population coverage

All potential MHC-I and MHC-II binders of yellow fever virus envelope protein were assessed for population coverage against the whole world population with the selected MHC-I and MHC-II interacted alleles by the IEDB population coverage calculation tool at (http://tools.iedb.org/tools/population/iedb_input)

Homology modeling

Raptor X structure prediction server was used to predict the 3D structure of the envelope protein (http://raptorx.uchicago.edu/). The structures of proposed B and T cells epitopes that would be utilized as vaccine candidates were demonstrated by Chimera1.8 (http://www.cgl.ucsf.edu/cimera).

RESULTS

Phylogenetic tree

Figure 1 showed the phylogenetic analysis of the retrieved strains. The phylogenetic tree showed that the South America strains were closely related to each other, although Brazil 1979 strains were closely related to strain Trinidad1979. Also, the strains from Sudan were closely related to strains from Ethiopia and Central Africa Republic. Generally, all strains of African countries were shown to be clustered together. Moreover, strains from South American were also clustered together. The strain from USA was closely related to the strains from Trinidad and Senegal. This may indicate the transmission of the virus from Africa or South America to USA.

Alignment

Sequence alignment of all retrieved strains was performed using ClustalW that presented by Bioedit software. The retrieved sequences demonstrated areas of conservancy and non-conservancy when sequences were aligned. The conserved regions were recognized by identity of amino acid sequences among the retrieved sequences. All the predicted epitopes that showed 100% conservancy in the tools of B and T lymphocytes were used for the further analysis.

Prediction of B cell epitope

The reference envelope protein sequence was subjected to Bepipred linear epitope, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity methods in IEDB. These methods predicted specific peptides in the protein that were linear, at surface and immunogenic, respectively, and can bind to B cell receptors. As shown in Figure (2), for Bepipred linear epitope prediction method the average binders score of the envelope protein to B cell was -0.043 with a maximum of 1.909 and a minimum of -2.756. Twenty-eight epitopes were predicted eliciting the B cell from the conserved regions and all values of the predicted linear epitopes were equal to or greater than the default threshold -0.043. In Emini surface accessibility prediction the average surface accessibility area of the protein was 1.000, with a maximum of 6.001 and a minimum of 0.064. Nineteen epitopes were potentially in the surface by passing the default threshold 1.000. In Kolaskar and Tongaonkar antigenicity the average of antigenicity was 1.028 with a maximum of 1.204 and minimum of 0.835. Fourteen epitopes gave score above the default threshold 1.028. Eleven epitopes successfully overlapped the three tools. Among them four epitopes namely 33VMAPDKPSL41, 72DKCP77, 236PPHA239 and 385LTYQ388 were proposed as B cell epitopes. The result of the all predicted epitopes that interacted with B cell was illustrated in Table (1) and the positions of the four proposed epitopes in the 3D structural level of envelope protein were shown in Figure 3.

Prediction of cytotoxic T-lymphocyte epitopes and interaction with MHC class I

The reference sequence of envelope protein was analyzed using IEDB MHC-1 binding prediction tools to predict T cell epitopes interacting with MHC Class I alleles. Based on Artificial Neural Network (ANN) with half-maximal inhibitory concentration (IC50) ≤300, 28 epitopes were predicted to interact with different MHC-1 alleles. The epitopes and their corresponding MHC-1 alleles were shown in Table 2. Four epitopes namely 471MTMSMSMIL479, 363VLIEVNPPF371, 33VMAPDKPSL41 and 226REMHHLVEF234 demonstrated higher interaction with MHC-1. Therefore, they were predicted as T cytotoxic cells epitopes. The position of these predicted epitopes in the 3D structural level in the envelop protein was illustrated in Figure 4.

Prediction of T helper cell epitopes and their interaction with MHC class II

The reference sequence of the envelope protein was analyzed using IEDB MHC-II binding prediction tools. Based on NN-align with half-maximal inhibitory concentration (IC50) ≤ 3000 there were ninety predicted epitopes found to interact with MHC-II alleles. Four epitopes namely 284RVKLSALTL292, 363VLIEVNPPF371, 479LVGVIMMFL487 and 226REMHHLVEF234 were interacted with most frequent MHC class II alleles. Therefore they were predicted as T helper cells epitopes. These four epitopes and their corresponding MHC- II alleles were shown in Table (3). The position of these predicted epitopes in the 3D structural level in the envelop protein was illustrated in Figure 4.

Analysis of the population coverage

The predicted epitopes from the envelope protein that interacting with MHC Class I and II alleles were subjected to population coverage analysis. As shown in Table (4), the MHC class I demonstrated four epitopes that highly interacted with most frequent MHC class I alleles. For instance, the epitope 471MTMSMSMIL479 demonstrated highest percentage (72.51%), followed by 363VLIEVNPPF371 (57.89%), 33VMAPDKPSL41 (45.80%) and 226REMHHLVEF234 (40.22%). The epitope set of these four epitopes against MHC-1 gave high percentage (84.66%) against the whole world population using IEDB population coverage tool.

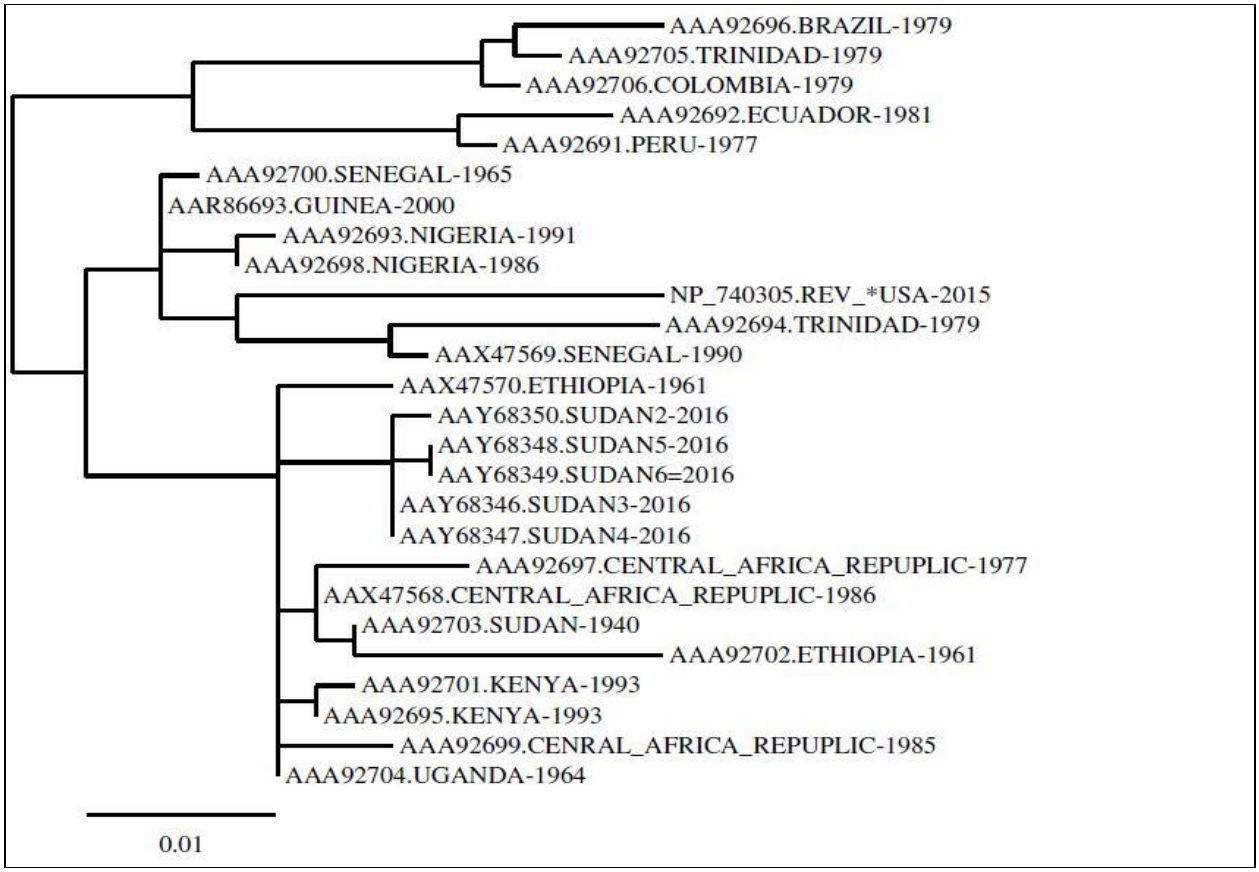

Figure 1. Phylogenetic tree of the envelop proteins of the retrieved strains. The retrieved strains demonstrated divergence in their common ancestors. *Reference sequence

Table1. B-cell predicted epitopes. The position of epitopes was according to the position of amino acids in the envelope protein of YFV.

Epitopes	Start	End	Length	Surface accessibility (1.000)	Antigenicity (1.028)
RDFIEGVHGGTW	9	20	12	0.437	0.979
SATLEQ	22	27	6	1.393	1.017
VMAPDKPSLDI	33	43	11	0.804	1.043
VMAPDKPSL*	33	41	9	1.037	1.051
KPSL*	38	41	4	1.493	1.064
DKCPSTGEAHL	72	82	11	1.091	1.031
DKCPST*	72	77	6	1.543	1.032
DKCP*	72	75	4	1.209	1.068
RTYSDR	94	99	6	5.597	0.949
WGNG	101	104	4	0.723	0.854
IAEME	195	199	5	0.746	0.949
TLPWQSGS	214	221	8	1.149	1.004
LPWQ*	215	218	4	1.014	1.055
FEPPHAA	234	240	7	1.168	1.043
PPHA*	236	239	4	1.436	1.074

QEGSLKTA	250	257	8	1.845	0.988
SLKTA*	253	257	5	1.144	1.033
RVTKD	263	267	5	2.487	0.992
KNPTD	308	312	5	4.254	0.909
GHGT	314	317	4	0.84	0.941
KGAPC	326	330	5	0.588	1.069
PIASTN	354	359	6	0.982	0.996
DEVLIE	361	366	6	0.62	1.059
NPPFGD	368	373	6	1.586	0.956
GDSRLTYQWHKEGSSI	381	396	16	3.298	0.987
LTYQWH*	385	391	7	2.168	1.038
LTYQ*	385	388	4	1.411	1.084
SAGG	422	425	4	0.579	0.956

*peptides revealed higher score if they were shorten in all tools (Bepipred linear epitope, Emini surface accessibility and Kolaskar and Tongaonkar antigenicity methods).
The proposed B cell epitopes in this study were underlined.

Table 2. List of epitopes that had binding affinity with MHC-I alleles. The position of peptides is according to position of amino acids in the envelope protein of YFV. The proposed T cytotoxic cell epitopes in this study were underlined.

Epitope	Start	End	Allele	ANN-ic50	Percentile
AKFTCAKSM	117	125	HLA-C*14:02	211.79	0.2
ALTLKGTSY	289	297	HLA-B*15:01	287.68	0.2
APDKPSLDI	35	43	HLA-B*07:02	152.86	0.2
CPSTGEAHL	74	82	HLA-B*53:01	208.8	0.1
FEPPHAATI	234	242	HLA-C*12:03	40.63	0.2
			HLA-C*14:02	283.4	0.2
GVIMMFLSL	481	489	HLA-A*02:06	56.25	0.2
			HLA-B*08:01	294.46	0.1
			HLA-B*15:01	213.05	0.2
IEGVHGGTW	12	20	HLA-B*44:03	208.47	1.3
ILVGVIMMF	478	486	HLA-A*23:01	113.23	0.2
			HLA-B*15:01	201.37	0.2
IMMFLSLGV	483	491	HLA-A*02:01	4.22	0.1
			HLA-A*02:06	19.67	0.2
			HLA-A*32:01	297.72	0.2
KEGSSIGKL	391	399	HLA-B*40:01	180.18	0.1
			HLA-B*40:02	224.46	0.1
KTALTGAMR	255	262	HLA-A*11:01	299.56	0.2
			HLA-A*31:01	34.93	0.2
			HLA-A*68:01	138.1	4.9
LTGAMRVTK	258	266	HLA-A*03:01	273.03	1.3
			HLA-A*11:01	50.49	0.2
			HLA-A*30:01	187.06	0.2
			HLA-A*68:01	149.24	5.2

LTLKGTSYK	290	298	HLA-A*03:01	17.73	0.2
			HLA-A*11:01	14.39	0.2
			HLA-A*30:01	55.31	0.2
			HLA-A*31:01	247.26	0.2
			HLA-A*68:01	33.57	4.9
MILVGVIMM	477	485	HLA-B*35:01	174.92	0.3
MSMILVGVI	475	483	HLA-A*68:02	92.25	0.2
			HLA-B*58:01	111.43	0.2
MSMSMILVG	473	481	HLA-B*15:01	202.53	0.2
MTMSMSMIL	471	479	HLA-A*02:01	44.95	0.2
			HLA-A*02:06	60.89	0.2
			HLA-A*30:01	112.05	0.2
			HLA-A*32:01	16.63	0.2
			HLA-A*68:02	5.5	0.2
			HLA-B*08:01	162.77	0.1
			HLA-B*15:01	162.77	0.2
			HLA-B*35:01	192.11	0.3
			HLA-B*39:01	13.03	0.2
			HLA-B*53:01	277.77	0.1
			HLA-B*57:01	196.85	0.2
			HLA-B*58:01	52.23	0.2
			HLA-C*03:03	261.08	0.2
			HLA-C*14:02	88.5	0.2
			HLA-C*15:02	49.05	0.1
QEGSLKTAL	250	258	HLA-B*40:01	71.01	0.1
			HLA-B*40:02	123.87	0.1
REMHHLVEF	226	234	HLA-A*32:01	33.38	0.2
			HLA-B*15:01	133.2	0.2
			HLA-B*18:01	103.67	0.1
			HLA-B*40:01	14.01	0.1
			HLA-B*40:02	18.42	0.1
			HLA-B*44:02	24.9	0.1
			HLA-B*44:03	34.33	0.1
RNMTMSMSM	469	477	HLA-A*32:01	72.14	0.2
			HLA-B*08:01	201.2	0.2
			HLA-B*15:01	126.54	0.2
			HLA-C*14:02	75.74	0.2
			HLA-C*15:02	57.41	0.1
RVKLSALTL	284	292	HLA-A*30:01	11.28	0.2
			HLA-B*07:02	222.64	0.2
SMILVGVIM	476	484	HLA-B*15:01	122.08	0.2
SMSMILVGV	474	482	HLA-A*02:01	11.64	0.2
			HLA-A*02:06	22.78	0.2
			HLA-A*68:02	47.79	0.2

SRLTYQWHK	383	391	HLA-B*27:05	35.77	0.2
TMSMSMILV	472	480	HLA-A*02:01	23.48	0.2
			HLA-A*02:06	44.8	0.2
			HLA-A*68:02	121.79	0.2
VEFEPPHAA	232	240	HLA-B*18:01	154.84	0.1
			HLA-B*40:02	152.41	0.1
VLIEVNPPF	363	371	HLA-A*02:01	105.4	0.2
			HLA-A*02:06	17.81	0.2
			HLA-A*23:01	162.63	0.2
			HLA-A*32:01	73.35	0.2
			HLA-B*15:01	17.79	0.1
			HLA-B*15:02	248.72	0.1
			HLA-B*35:01	298.06	0.3
VMAPDKPSL	33	41	HLA-A*02:01	294.47	0.2
			HLA-C*03:03	111.1	0.2
			HLA-C*14:02	161.39	0.2

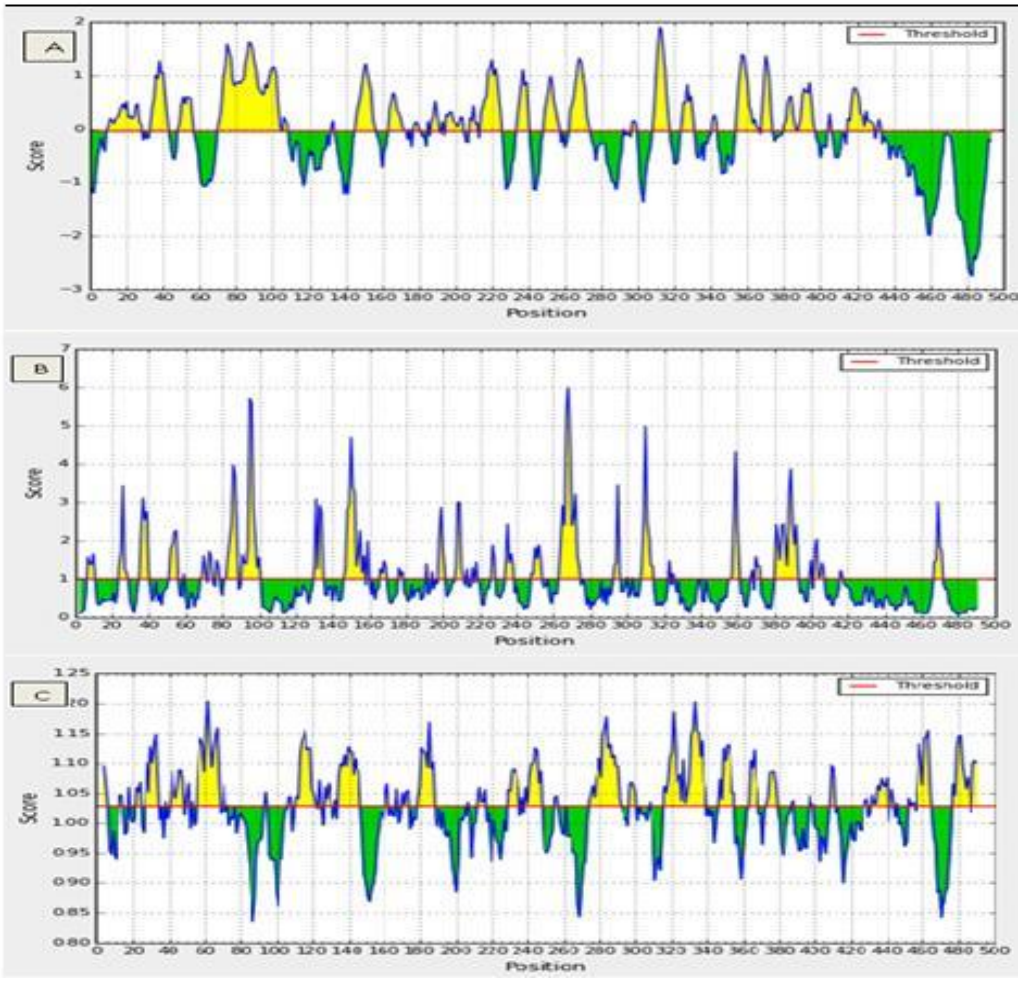

Figure 2. Prediction of B-cell epitopes by different IEDB scales (A- Bepipred linear epitope prediction, B- Emini surface accessibility, C- Kolaskar and Tongaonkar antigenicity prediction). Regions above threshold (red line) are proposed to be a part of B cell epitope while regions below the threshold (red line) are not.

Table 3. List of top four epitopes that had high binding affinity with MHC-II alleles. The position of peptides is according to position of amino acids in the envelope protein.

core sequence	start	end	peptide sequence	Allele	ic-50
LVGVIMMFL	473	487	MSMSMILVGVIMMFL	HLA-DPA1*01/DPB1*04:01	1245.4
				HLA-DPA1*01:03/DPB1*02:01	634.9
				HLA-DPA1*02:01/DPB1*01:01	544.6
				HLA-DPA1*03:01/DPB1*04:02	295.6
				HLA-DQA1*05:01/DQB1*02:01	1587.8
				HLA-DRB1*04:05	1373.3
				HLA-DRB1*07:01	494.4
				HLA-DRB1*15:01	560.2
	474	488	SMSMILVGVIMMFLS	HLA-DPA1*01/DPB1*04:01	1167.8
				HLA-DPA1*01:03/DPB1*02:01	446.7
				HLA-DPA1*02:01/DPB1*01:01	550.9
				HLA-DPA1*03:01/DPB1*04:02	223
				HLA-DRB1*04:05	1335.5
				HLA-DRB1*07:01	732.3
				HLA-DRB1*15:01	540.4
	475	489	MSMILVGVIMMFLSL	HLA-DPA1*01/DPB1*04:01	968.9
				HLA-DPA1*01:03/DPB1*02:01	329.6
				HLA-DPA1*02:01/DPB1*01:01	495.8
				HLA-DPA1*03:01/DPB1*04:02	165.9
				HLA-DRB1*04:05	1189.7
				HLA-DRB1*07:01	873.4
				HLA-DRB1*15:01	425.1
	476	490	SMILVGVIMMFLSLG	HLA-DPA1*01/DPB1*04:01	731.2
				HLA-DPA1*01:03/DPB1*02:01	356.9
				HLA-DPA1*02:01/DPB1*01:01	591.3
				HLA-DRB1*07:01	1910.5
				HLA-DRB1*15:01	477.4
	477	491	MILVGVIMMFLSLGV	HLA-DPA1*01/DPB1*04:01	827.4
				HLA-DPA1*01:03/DPB1*02:01	369.9
				HLA-DPA1*02:01/DPB1*01:01	684.7
				HLA-DQA1*05:01/DQB1*02:01	1834.3
	478	492	ILVGVIMMFLSLGVG	HLA-DPA1*01/DPB1*04:01	942.2
				HLA-DPA1*01:03/DPB1*02:01	443
				HLA-DPA1*02:01/DPB1*01:01	847.6
	479	493	LVGVIMMFLSLGVGA	HLA-DPA1*01/DPB1*04:01	1099.8
				HLA-DPA1*01:03/DPB1*02:01	543.2
				HLA-DPA1*02:01/DPB1*01:01	1052.1
				HLA-DQA1*05:01/DQB1*02:01	2965.6
RVKLSALTL	278	292	GGHVSCRVKLSALTL	HLA-DPA1*02:01/DPB1*01:01	1762.8
				HLA-DPA1*02:01/DPB1*05:01	2800.5

			HLA-DPA1*03:01/DPB1*04:02	1804.5
			HLA-DRB1*01:01	44.2
			HLA-DRB1*04:01	1148.1
			HLA-DRB1*04:04	974
			HLA-DRB1*07:01	74.1
			HLA-DRB1*09:01	488.6
			HLA-DRB1*13:02	1838.5
			HLA-DRB4*01:01	144.8
279	293	GHVSCRVKLSALTLK	HLA-DPA1*01:03/DPB1*02:01	2225.6
			HLA-DRB1*01:01	31.2
			HLA-DRB1*04:01	755.1
			HLA-DRB1*04:04	471.8
			HLA-DRB1*07:01	143.3
			HLA-DRB1*09:01	408.6
			HLA-DRB1*13:02	1150
			HLA-DRB4*01:01	110.9
280	294	HVSCRVKLSALTLKG	HLA-DPA1*01:03/DPB1*02:01	1720.1
			HLA-DRB1*01:01	25.6
			HLA-DRB1*04:04	422.7
			HLA-DRB1*07:01	321.5
			HLA-DRB1*09:01	431
			HLA-DRB1*13:02	1101.7
			HLA-DRB4*01:01	86.9
281	295	VSCRVKLSALTLKGT	HLA-DPA1*01:03/DPB1*02:01	1549
			HLA-DRB1*01:01	23.6
			HLA-DRB1*07:01	558.5
			HLA-DRB1*08:02	370
			HLA-DRB1*09:01	505.3
			HLA-DRB1*13:02	954.7
			HLA-DRB4*01:01	62.4
282	296	SCRVKLSALTLKGTS	HLA-DPA1*01:03/DPB1*02:01	1695.1
			HLA-DRB1*01:01	36
			HLA-DRB1*07:01	845.8
			HLA-DRB1*09:01	711.5
			HLA-DRB1*13:02	1063.9
			HLA-DRB1*15:01	110.8
			HLA-DRB4*01:01	55.1
283	297	CRVKLSALTLKGTSY	HLA-DPA1*01/DPB1*04:01	2471.9
			HLA-DRB1*08:02	554.8
			HLA-DRB1*09:01	1013.8
			HLA-DRB1*13:02	1515.3
			HLA-DRB4*01:01	82.3
284	298	RVKLSALTLKGTSYK	HLA-DRB1*13:02	2095.5

				HLA-DRB4*01:01	90.4
VLIEVNPPF	357	371	STNDDEVLIEVNPPF	HLA-DQA1*05:01/DQB1*02:01	162.4
				HLA-DRB1*01:01	2207.4
				HLA-DRB1*04:01	120.3
				HLA-DRB1*04:04	645.4
				HLA-DRB1*04:05	563.2
				HLA-DRB1*07:01	2790.6
				HLA-DRB1*08:02	2573.3
				HLA-DRB1*13:02	172
				HLA-DRB1*15:01	584.3
				HLA-DRB3*01:01	28.3
				HLA-DRB4*01:01	1611.4
	358	372	TNDDEVLIEVNPPFG	HLA-DPA1*02:01/DPB1*01:01	2202.3
				HLA-DQA1*05:01/DQB1*02:01	190.4
				HLA-DRB1*01:01	1331.4
				HLA-DRB1*03:01	2280.1
				HLA-DRB1*04:01	87.5
				HLA-DRB1*04:04	374.8
				HLA-DRB1*04:05	495.3
				HLA-DRB1*08:02	1142.8
				HLA-DRB1*13:02	61.3
				HLA-DRB1*15:01	357.7
				HLA-DRB3*01:01	30.2
				HLA-DRB4*01:01	1273.2
	359	373	NDDEVLIEVNPPFGD	HLA-DPA1*02:01/DPB1*01:01	2219.8
				HLA-DQA1*05:01/DQB1*02:01	173.5
				HLA-DRB1*01:01	1364.4
				HLA-DRB1*03:01	1408.3
				HLA-DRB1*04:01	78.8
				HLA-DRB1*04:04	292.8
				HLA-DRB1*04:05	447.5
				HLA-DRB1*08:02	805.8
				HLA-DRB1*13:02	39.3
				HLA-DRB1*15:01	279.1
				HLA-DRB3*01:01	27.3
				HLA-DRB4*01:01	1019.9
	360	374	DDEVLIEVNPPFGDS	HLA-DPA1*01:03/DPB1*02:01	2555.1
				HLA-DPA1*02:01/DPB1*01:01	2348.3
				HLA-DQA1*01:01/DQB1*05:01	2996.7
				HLA-DQA1*05:01/DQB1*02:01	220.6
				HLA-DRB1*01:01	627.6
				HLA-DRB1*03:01	816.1
				HLA-DRB1*04:01	61.4

				HLA-DRB1*04:04	250.6
				HLA-DRB1*04:05	444.4
				HLA-DRB1*08:02	469.7
				HLA-DRB1*13:02	36.5
				HLA-DRB1*15:01	212.5
				HLA-DRB3*01:01	29.6
				HLA-DRB4*01:01	801.6
	361	375	DEVLIEVNPPFGDSY	HLA-DPA1*01:03/DPB1*02:01	2297.1
				HLA-DPA1*02:01/DPB1*01:01	2516.8
				HLA-DQA1*05:01/DQB1*02:01	262.6
				HLA-DRB1*03:01	1087
				HLA-DRB1*04:01	74.8
				HLA-DRB1*04:04	217.6
				HLA-DRB1*04:05	577
				HLA-DRB1*08:02	428.1
				HLA-DRB1*13:02	38.7
				HLA-DRB1*15:01	206.9
				HLA-DRB3*01:01	55.4
				HLA-DRB4*01:01	751.3
	362	376	EVLIEVNPPFGDSYI	HLA-DPA1*01:03/DPB1*02:01	2153.9
				HLA-DPA1*02:01/DPB1*01:01	2747.9
				HLA-DQA1*05:01/DQB1*02:01	318.3
				HLA-DRB1*03:01	1535.5
				HLA-DRB1*04:01	101.3
				HLA-DRB1*04:04	219.3
				HLA-DRB1*04:05	1023.7
				HLA-DRB1*08:02	477.1
				HLA-DRB1*15:01	179.3
				HLA-DRB3*01:01	114.5
				HLA-DRB4*01:01	722.2
	363	377	VLIEVNPPFGDSYII	HLA-DPA1*01:03/DPB1*02:01	2314.4
				HLA-DQA1*05:01/DQB1*02:01	532.7
				HLA-DRB1*04:01	195.4
				HLA-DRB1*04:04	463.5
				HLA-DRB1*04:05	1591.7
				HLA-DRB1*08:02	638.2
				HLA-DRB1*15:01	199
				HLA-DRB3*01:01	268.6
				HLA-DRB4*01:01	1225.1
REMHHLVEF	220	234	GSGGVWREMHHLVEF	HLA-DQA1*03:01/DQB1*03:02	2849
				HLA-DQA1*05:01/DQB1*02:01	2188.1
				HLA-DRB1*07:01	1467.4
	223	237	GVWREMHHLVEFEPP	HLA-DPA1*01:03/DPB1*02:01	500.4

224	238	VWREMHHLVEFEPPH	HLA-DPA1*01:03/DPB1*02:01	394
			HLA-DRB1*09:01	2640
225	239	WREMHHLVEFEPPHA	HLA-DPA1*01/DPB1*04:01	1419.2
226	240	REMHHLVEFEPPHAA	HLA-DPA1*01/DPB1*04:01	1427.6

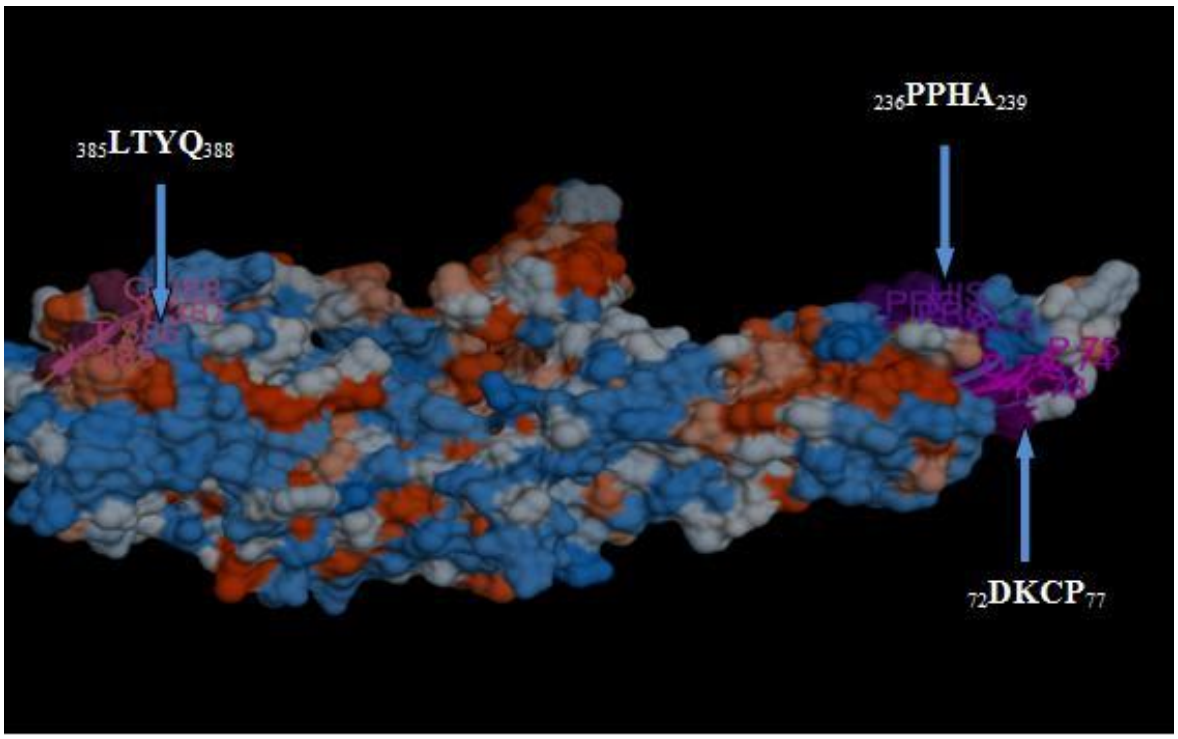

Figure 3. Position of proposed conserved B cell epitopes in structural level of envelop protein. The epitope VMAPDKPSL) was shown in figure (4). The epitopes showed conservancy, surface accessibility and antigenicity using IEDB software.

Table 4. The population coverage (PC) of MHC-I and MHC-II for the proposed epitopes. The population coverage of MHC-I and MHC-II combined alleles was calculated for all proposed epitopes.

MHC-I	PC	MHC-II	PC	MHC-I&MHC-II	PC
MTMSMSMIL	72.51%	RVKLSALTL	99.13%	RVKLSALTL	99.27%
VLIEVNPPF	57.89%	VLIEVNPPF	99.08%	VLIEVNPPF	99.61%
VMAPDKPSL	45.80%	LVGVIMMFL	98.80%	LVGVIMMFL	98.80%
REMHHLVEF	40.22%	REMHHLVEF	96.80%	MTMSMSMIL	97.20%
Epitopes Set	84.66%	Epitopes Set	99.91%	VMAPDKPSL	47.05%
				REMHHLVEF	98.09%
				Epitopes Set	99.99%

PC; Population Coverage

Also four epitopes highly interacted with most frequent MHC class II alleles. The epitope 284RVKLSALTL292 demonstrated highest percentage (99.13%) followed by 363VLIEVNPPF371 (99.08%), 479LVGVIMMFL487 (98.80%) and 226REMHHLVEF234 (96.80%) (Table 4).The epitope set of these four epitopes against MHC-II alleles gave high percentage (99.91%) against

the whole world population using IEDB population coverage tool.

It is noteworthy that the epitopes 363VLIEVNPPF371 and 226REMHHLVEF234 interacted with both MHC-I and II alleles with high affinity and high population coverage for each class. All proposed epitopes were subjected to population coverage tools to assess population coverage of their MHC-I and MHC-II combined alleles. The population coverage of the proposed epitopes against the combined alleles was 99.99%. This result strengthens the impact of these elected epitopes as vaccine candidates against YFV.

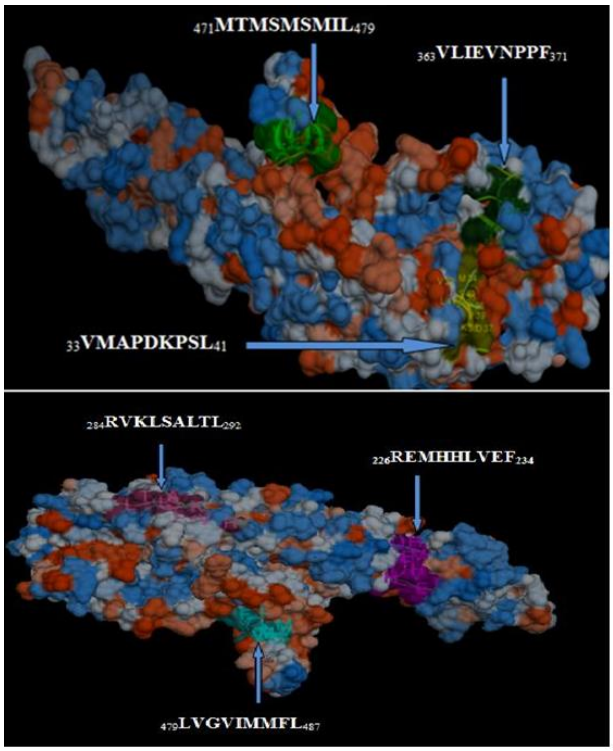

Figure 4. Position of proposed conserved T cytotoxic cell (MHC class I) and T helper cell (MHC class II) epitopes in structural level of envelop protein

DISCUSSION

In this study, immunoinformatic tools at IEDB were used to determine the 100% conserved regions that could be predicted as highly potential immunogenic epitopes that elicit both B and T cells. Many studies concluded that the envelope protein was critical to induce strong humoral and cellular immune response against yellow fever virus [27,28]. Therefore in this study the envelope protein was subjected to B cell epitope prediction tool in IEDB to search for conserved, linear and antigenic epitopes. Eleven epitopes successfully overlapped the three tools used to predict epitopes that interacted with the humoral immunity. Four epitopes namely 72DKCP77, 236PPHA239, 385LTYQ388 and 33VMAPDKPSL41 were selected as B cell epitopes. The epitope 385LTYQ388 showed surface accessibility antigenicity score 1.411 and 1.084, respectively, and was found the most satisfactory peptide for eliciting B cell. Moreover this epitope was overlapped with 383SRLTYQWHK391 which has affinity to MHC-I and MHC-II alleles and has 50.52% in the combined population coverage for MHC-I and MHC-II. Furthermore the epitope 236PPHA239 gave 1.436 as surface accessibility score and 1.074 in antigenicity score overlapped with the epitope 232VEFEPPHAA240 that has affinity to MHC-I and MHC-II alleles and has 79.73% in the combined population coverage for MHC-I and MHC-II. This indicated the importance of these regions in induction of B and T cells against yellow fever virus. The epitope 72DKCP77 gave 1.209 as surface accessibility score and 1.068 as antigenicity score. Strikingly the epitope 33VMAPDKPSL41 demonstrated favorable interaction with B cells and interacted with both MHC-I and MHC-II alleles of T cells. In B cell it was found to be linear, conserved, surface accessible and antigenic (Table 1). In the T cells it interacted with high affinity with MHC-I alleles with better population coverage but it revealed low affinity and less population coverage of MHC-II. Therefore this epitope was chosen as a vaccine candidate base on its promising results against B cells and MHC-I of the T cells.

Concerning the T cell prediction tools; twenty eight and ninety epitopes were predicted from the envelope protein interacting with most frequent alleles of MHC-I and MHC-II, respectively. Beside this epitope 33VMAPDKPSL41 three epitopes namely 471MTMSMSMIL479, 363VLIEVNPPF371 and 226REMHHLVEF234 were interacted with most frequent alleles of MHC-I. Also they got higher percentages in the population coverage. Therefore were proposed as a vaccine candidate against cytotoxic T cell. For MHC-II, the epitopes 363VLIEVNPPF371 and 226REMHHLVEF234 that interacted with MHC-I were also found interacting with MHC-II frequent

alleles with high population coverage percentages. In addition to that the epitopes 284RVKLSALTL292, 479LVGVIMMFL487 interacted only with MHC-II alleles with high population coverage. Therefore were chosen as the promising epitopes for T helper cell. All these epitopes were tested for the population coverage against the whole world that had the potential to develop immune response against these epitopes. All the proposed (combined epitopes against MHC-I and MHC-II) demonstrated epitope set of 99.99%. This result potentiated their ability to act as a vaccine candidate against T cells.

One report by de Melo et al (2013) provided six epitopes from the YFV envelope protein that elicited both CD4+ and CD8+ T cells [29]. These epitopes were E57-71, E65-79, E72-87, E337-351, E345-359 and E361-375. In this study two epitopes 72DKCP77 and 363VLIEVNPPF371 overlapped with their epitope E72-87 and E361-375, respectively. The former was found eliciting the CD8+ T cells and the latter eliciting both CD4+ and CD8+ T cells in this study. Also another study conducted by Milton et al (2015) using ELISPOT showed that the peptide from yellow fever virus envelope protein E 57-71 and E 329-343 produce the highest CD8+ T cell responses and peptides E57-71, E61-75 , E129-145 and E135-147 were able to induce a high CD4+ T cell response in murine [30]. None of these peptides agreed with our predicted peptides.

Conclusion

Peptide–based vaccine is a one of immunoinformatics applications that based on identification and chemical synthesis of B and T-cell epitopes that induce humoral and cellular response. Peptides have become more desirable vaccine candidates owing to their relatively easy production and construction, chemical stability, and absence of infectious potential, which lessens the time and reduce cost. In this study nine epitopes were predicted to act as a vaccine candidate against YFV. The efficacy and safety of the predicted epitopes by this computational analysis are needed to be evaluated in an animal model to confirm their efficacy in inducing protective immune response.

ACKNOWLEDGEMENT

Authors would like to thank the staff members of College of Veterinary Medicine, University of Bahri, Sudan for their cooperation and support.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Lindenbach BD, Thiel HJ, Rice CM. Flaviviridae: The Viruses and Their Replication. In: Knipe, D.M. and Howley, O.M., Eds., Fields Virology, 5th Edition, Lippincot William & Wilkins, Philadelphia 2011; 1101-1151.
2. Tolle MA. Mosquito-borne diseases. Curr Probl Pediatr Adolesc Health Care 2009; 39(4):97-140.
3. Wilder-Smith A, W. Leong Y. Importation of yellow fever into China: assessing travel patterns. J Travel Med 2017; 24 (4). doi: 10.1093/jtm/tax008
4. Woodall JP and Yuill TM. Why is the yellow fever outbreak in Angola a 'threat to the entire world'?. Intern J Infect Dis 2016; 48:96-97.
5. Ellis BR and B. AD. The enigma of yellow fever in East Africa". Rev Med Virol 2008,18: p. 331-346.
6. Patricia Najera Hamrick, Sylvain Aldighieri, Gustavo Machado, et al. Geographic patterns and environmental factors associated with human yellow fever presence in the Americas. PLoS Negl Trop Dis 2017; 11(9): e0005897.
7. Mutebi JP, Barrett AD. The epidemiology of yellow fever in Africa. Microbes Infect 2002; 4 (14): 1459–1468. doi:10.1016/S1286-4579(02)00028-X.
8. Barrett AD, Higgs S. Yellow fever: a disease that has yet to be conquered. Annu Rev Entomol 2007; 52: 209–29. doi:10.1146/annurev.ento.52.110405.091454
9. Mutebi JP, Rijnbrand RC, Wang H, et al. Genetic relationships and evolution of genotypes of yellow fever virus and other members of the yellow fever virus group within the genus Flavivirus based on the 3' noncoding region". J Virol 2004; 78 (18): 9652–9665. doi:10.1128/JVI.78.18.9652-9665.2004.
10. Auguste AJ, Lemey P, Pybus OG, et al. Yellow fever virus maintenance in Trinidad and its dispersal throughout the Americas. J Virol 2010; 84 (19): 9967–9977. doi:10.1128/JVI.00588-10
11. de Souza RP, Foster PG, Sallum MA, et al. Detection of a new yellow fever virus lineage within the South American genotype I in BrazilJ". Med Virol 2010; 82 (1): 175–185. doi:10.1002/jmv.21606
12. Mir D, Delatorre E, Bonaldo M, Lourenço-de-Oliveira R, Vicente AC, Bello G. Phylodynamics of Yellow Fever Virus in the Americas: new insights into the origin of the 2017 Brazilian outbreak. Sci

Rep 2017. 7 (1): 7385. doi:10.1038/s41598-017-07873-7

13. Vainio J, F Cutts, eds. Yellow Fever. WHO Division of Emerging and other Communicable Diseases Surveillance and Control 1998.

14. Monath TP. The absence of yellow fever in Asia: hypotheses. A cause for concern? Virus Inf Exch Newslett 1989; 106–107.

15. Cathey JT, Marr JS. Yellow fever, Asia and the East African slave trade. Trans R Soc Trop Med Hyg 2014; 108(5): 252–257. doi:10.1093/trstmh/tru043

16. Reinhardt B, Jaspert R, Niedrig M, Kostner C, L'age-Stehr J. Development of viremia and humoral and cellular parameters of immune activation after vaccination with yellow fever virus strain 17D: a model of human flavivirus infection. J Med Virol 1998; 56:159-167.

17. Jennings AD, Gibson CA, Miller BR, et al. Analysis of a yellow fever virus isolated from a fatal case of vaccine-associated human encephalitis. J Infect 1994; 169: p. 512-518.

18. Chan RC, Penney DJ, Little D, Carter IW, Roberts JA, Rawlinson WD. Hepatitis and death following vaccination with 17D-204 yellow fever vaccine. Lancet 2001; 358: p. 121-122.

19. Vasconcelos PF, Luna EJ, Galler R, Silva LJ, et al; Brazilian Yellow Fever Vaccine Evaluation Group. Serious adverse events associated with yellow fever 17DD vaccine in Brazil: a report of two cases. Lancet 2001; 358: p. 91-97.

20. Yang H, Yang H, Li Z, et al. Japanese encephalitis virus/yellow fever virus chimera is safe and confers full protection against yellow fever virus in intracerebrally challenged mice. Vaccine 2018; 36(18): p. 2450-2455

21. Bassi MR, Larsen MA, Kongsgaard M, et al. Vaccination with Replication Deficient Adenovectors Encoding YF-17D Antigens Induces Long-Lasting Protection from Severe Yellow Fever Virus Infection in Mice. PLOS.org 2016; 10(2):e0004464. doi: 10.1371/journal.pntd.0004464

22. Ying-Tsang Lo, Tun-Wen Pai, Wei-Kuo Wu, Hao-Teng C. Prediction of conformational epitopes with the use of a knowledge-based energy function and geometrically related neighboring residue characteristics. BMC Bioinformatics 2013; 14: S3.

23. Pappalardo F, Pennisi M, Castiglione F, Motta S. Vaccine protocols optimization: in silico experiences, Biotech Adv 2010; 28: 82-93.

24. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 1999; 41: p. 95-98.

25. Peters B and S. A. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC Bioinformatics 2005; 6: p. 132.

26. Zhang Q, Wang P, Kim Y, et al. Immune epitope database analysis resource (IEDB-AR). Nucleic Acids Res 2008; 36: W513–W518.

27. Maciel M Jr, Kellathur SN, Chikhlikar P, et al. Comprehensive analysis of T cell epitope discovery strategies using 17DD yellow fever virus structural proteins and BALB/c (H2d) mice model. Virology 2008; 378: 105–117.

28. Tottey S, Shoji Y, Jones RM, et al. Plant-Produced Subunit Vaccine Candidates against Yellow Fever Induce Virus Neutralizing Antibodies and Confer Protection against Viral Challenge in Animal Models. Ameri Soci Trop Med Hyg 2018; 98(2): 420-431.

29. de Melo AB, Nascimento EJ, Braga-Neto U, et al. T-cell Memory Responses Elicited by Yellow Fever Vaccine are Targeted to Overlapping Epitopes Containing Multiple HLA-I and-II Binding Motifs. 2013; PloS Negl Trop Dis 7(1):e1938. https://doi.org/10.1371/journal.pntd.0001938.

30. Maciel M Jr, Cruz Fda S, Cordeiro MT, et al. A DNA Vaccine against Yellow Fever Virus: Development and Evaluation. plos.org 2015; 13;9(4):e0003693. doi: 10.1371/journal.pntd.0003693. eCollection.