American Journal of Microbiological Research, 2016, Vol. 4, No. 4, 101-121
Available online at http://pubs.sciepub.com/ajmr/4/4/2

DOI:10.12691/ajmr-4-4-2

In Silico Prediction of a Novel Universal Multi-epitope Peptide Vaccine in the Whole Spike Glycoprotein of MERS CoV

Marwan Mustafa Badawi[1,#,*], Maryam Atif SalahEldin[2,#], Marwa Mustafa Suliman[3,#], Samah Awad AbduRahim[2,##], Alaa Abd elghafoor Mohammed[1,##], Alaa Salah Aldein SidAhmed[1], Marwa Mohamed Osman[1], Mohamed Ahmed Salih[1]

[1]Department of Biotechnology, Africa city of Technology- Khartoum, Sudan
[2]Department of medical microbiology, Faculty of Medical Laboratory Sciences, University of Khartoum-Khartoum, Sudan
[3]Sudan Armed forces hospital- Khartoum, Sudan
[#]contributed equally.
[##]contributed equally.
*Corresponding author: mmbadwi44new@gmail.com

Abstract Middle East Respiratory Syndrome (MERS) is a new viral emergent human disease caused by a novel strain of Coronavirus. First known case of MERS occurred in Jordan in April 2012, by December 2015, the disease had already struck 1,621 persons of whom 584 died. Despite of the high mortality rate of the infection, there are no clinically approved vaccines or antiviral drugs, thus, the aim of this study is to analyze Spike glycoprotein strains using *in silico* approaches looking for conservancy, which is further studied to predict all potential epitopes that can be used after *in vitro* and *in vivo* confirmation as a therapeutic peptide vaccine. Total of 255 Spike glycoprotein variants retrieved from NCBI database were aligned, to select the conserved regions for epitopes prediction. By means of IEDB analysis resource B and T cell epitopes were predicted and population coverage was calculated. Two epitopes were proposed for international therapeutic peptide vaccine for B cell (GTPPQVY and LTPRSVRSVP). Regarding T cell, FSFGVTQEY epitope was highly recommended as therapeutic peptide vaccine to interact with MHC class I along with eight other epitopes that showed good population coverage against the whole world population. Four epitopes showed high affinity to interact with MHC class II alleles (FNLTLLEPV, FAAIPFAQS, SFAAIPFAQ and FYVYKLQPL) with excellent population coverage throughout the world and Saudi Arabia. Herd immunity protocols can be conducted in countries with low population coverage to minimize the active transmission of the virus especially among people contacting camels and other groups at risk.

Keywords: *MERS CoV, Peptide vaccine, Immune Epitope Database IEDB, Epitopes, Herd immunity and Vaccine*

Cite This Article: Marwan Mustafa Badawi, Maryam Atif SalahEldin, Marwa Mustafa Suliman, Samah Awad AbduRahim, Alaa Abd elghafoor Mohammed, Alaa Salah Aldein SidAhmed, Marwa Mohamed Osman, and Mohamed Ahmed Salih, "In Silico Prediction of a Novel Universal Multi-epitope Peptide Vaccine in the Whole Spike Glycoprotein of MERS CoV." *American Journal of Microbiological Research*, vol. 4, no. 4 (2016): 101-121. doi: 10.12691/ajmr-4-4-2.

1. Introduction

Middle East respiratory syndrome (MERS) is a new viral emergent human disease caused by a novel coronavirus called Middle East Respiratory Syndrome Coronavirus (MERS-CoV) [1]. Although the disease was first reported in the Kingdom of Saudi Arabia (KSA) when the virus was isolated from a patient with fatal pneumonia and acute kidney injury in Jeddah in June 2012, however, through retrospective investigations, health officials later identified that the first known cases of MERS occurred in Jordan in April 2012 [2,3,4,5,6]. In September 2012 the World Health Organization (WHO) reported two cases of severe community-acquired pneumonia caused by MERS-CoV [7]. By December 2015, the disease had already struck 1,621 persons of whom 584 died from respiratory failure and diarrhea majority of them were reported from the Arabian Peninsula, with one large outbreak involving 186 cases in the Republic of Korea; infection is typically associated with considerable morbidity and mortality [8,9,10,11].

Most infections were geographically linked to the Middle East, i.e., Jordan, Saudi Arabia, Qatar, and United Arab Emirates (UAE), but cases also occurred in the United Kingdom, Germany, France, and Italy [12]. As known cases have been directly or indirectly related to countries in the Arabian Peninsula, all cases of MERS have been linked through travel to or residence in countries in this region [6,13]. Dromedary camels (Camelus dromedarius) were implicated for the first time as a possible source for human infection on the basis of the presence of MERS-CoV neutralizing antibodies in

dromedaries from Oman and the Canary Islands of Spain. Since then, the presence of MERS-CoV antibodies in dromedaries has been reported in Jordan, Egypt, UAE and KSA [14]. MERS affects the respiratory system (lungs and breathing tubes). Most MERS patients developed severe acute respiratory illness with symptoms of fever, cough and shortness of breath. About 3-4 out of every 10 patients reported with MERS have died [6].

As all coronaviruses, MERS-CoV has a non segmented, single stranded, positive polarity RNA genome. It is an enveloped virus with a helical nucleocapsid. There is no virion polymerase. In the electron microscope, prominent club-shaped spikes in the form of a "corona" (halo) can be seen [15]. The MERS-CoV genome contains 30,119 nucleotides and contains at least 10 predicted open reading frames, 9 of which are predicted to be expressed from a nested set of seven subgenomic mRNAs [16].

Among the structural proteins of MERS-CoV, the spike (S) glycoprotein plays the most important roles in virus infection and pathogenesis as it uses a novel coronavirus receptor for entry and is targeted by neutralizing antibodies, thus, it is considered to be a promising target for effective MERS vaccine design [17,18]. More importantly, T-cell-based cellular immunity is essential for cleaning MERS-CoV infection, yet the vaccine against the S protein mainly elicits neutralizing antibody response. Further, the high mutation rate of the S protein may result in the escape of neutralizing antibodies against MERS-CoV. Therefore, a highly conserved target that elicits both neutralizing antibody and cellular immunity against MERS-CoV is essential for an effective vaccine development [19].

Although the mortality of the infection is alarming (30-50%), as is its uncanny resemblance-at least in its clinical features to severe acute respiratory syndrome (SARS), there are no clinically approved vaccines or antiviral drugs available for either of these infections; thus, the development of effective therapeutic and preventive strategies that can be readily applied to new emergent strains is a research priority to save human lives and address the pandemic concerns [20,21,22,23]. Past efforts to develop coronavirus vaccines have used whole-inactivated virus, live-attenuated virus, recombinant protein subunit or genetic approaches [24]. Although these methods are useful for vaccine development and successful in many cases, but they are time-consuming and fail when the pathogens cannot be cultivated *in vitro*, or when the most abundant antigens are variable in sequence. Now genomic approaches allow prediction of all antigens independent of their abundance and immunogenicity during infection, without the need to grow the pathogen *in vitro* [25]. The use of peptides as therapeutics is experiencing renewed enthusiasm owing to advances in delivery, stability and design. Moreover, there is a growing emphasis on the use of peptides in vaccine design as insights into tissue-specific processing of the immunogenic epitopes of proteins and the discovery of unusually long cytotoxic T-lymphocyte epitopes broaden the range of targets and give clues to enhancing peptide immunogenicity [26].

In this study, an immunoinformatics-driven genome-wide screening strategy of vaccine targets was adopted to identify a multi-epitope vaccine candidate against whole S glycoprotein of MERS-CoV that could be suitable to trigger a significant humoral and cellular immune response.

2. Materials and Methods

2.1. Protein Sequence Retrieval

A total of 255 Spike glycoprotein of MERS CoV were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/protein/) database in April 2016, and collected from different parts of the world; 184 isolates were collected in Saudi Arabia, twenty-three in United Arab Emirates, twenty-two in South Korea, and the rest were isolated in different countries; Oman, China, USA, United Kingdom, Greece, Qatar, Egypt, Tunisia, France, Thailand and Nigeria. Retrieved strains, their accession numbers and date of collection are listed in supplementary table (S1).

2.2. Determination of Conserved Regions

The retrieved sequences were used as a platform to obtain conserved regions using multiple sequence alignment (MSA). Sequences aligned with the aid of ClustalW as implemented in the BioEdit program, version 7.0.9.0. [27]

2.3. B-cell Epitope Prediction

B cell epitope is characterized by being hydrophilic, accessible and in a beta turn region. Thus, the classical propensity scale methods and hidden Markov model programmed softwares from IEDB analysis resource (http://www.iedb.org/), were used for the following aspects: [28,29]

Prediction of linear B-cell Epitopes: BepiPred from immune epitope database (http://toolsiedb.ofg/bcell/) [30] was used as linear B-cell epitope prediction from the conserved region with a default threshold value of 0.350.

Prediction of surface accessibility: by using Emini surface accessibility prediction tool of the immune epitope database (IEDB) [31], the surface accessible epitopes were predicted from the conserved regions holding the default threshold value 1.000 or higher.

Prediction of Epitopes antigenicity sites: the kolaskar and tongaonker antigenicity method [32] was used to determine the antigenic sites with a default threshold value of 1.045.

Prediction of epitopes hydrophilicity: parker hydrophilicity prediction tool [33] was used to determine the hydrophilicity of the conserved regions; the threshold default value was 1.286.

Prediction of beta turns sites: Chou and Fasman beta turn prediction method was used with the default threshold 1.009 to determine the sites contain beta turns.

Thresholds of all tool were provided by IEDB and it is mainly calculated by the software as the average score of the tested protein for each corresponding tool.

2.4. MHC Class I Binding Predictions

Analysis of peptide binding to MHC class I molecules was assessed by the IEDB MHC I prediction tool http://tools.iedb.org/mhci/n, for MHC-I binding predication, several alleles were used including HLA-A*02, HLA-B*51, HLA-B*50, HLA-B*08, HLA-C*07 ,

HLA-C*06 and HLA-C*15 that have been reported as frequent among Saudis.[34,35,36,37,38] MHC-I peptide complex presentation to T lymphocytes undergo several steps. The attachment of cleaved peptides to MHC molecules step was predicted. prediction methods can be achieved by Artificial Neural Network (ANN), Stabilized Matrix Method (SMM), or Scoring Matrices derived from Combinatorial Peptide Libraries. Consensus method which combines ANN, SMM and comblib different methods was used [39,40,41,42,43]. Prior to prediction, all epitope lengths were set as 9mers, all internationally conserved epitopes that bind to alleles at score equal or less than 1.0 percentile rank were selected for further analysis. [44]

2.5. MHC Class II Binding Predictions

Analysis of peptide binding to MHC class II molecules was assessed by the IEDB MHC II prediction tool http://tools.immuneepitope.org/mhcii/ [45,46]. For MHC-II binding predication, the reference set of alleles were used which include HLA-DQB1*02, HLA-DQB1*03, HLA-DQB1*06, HLA-DRB1*07, HLA-DRB1*04 and HLA-DRB1*03 that are reported to be frequent among Saudis [34,35,36,37,38]. MHC class II groove has the ability to bind to peptides with different lengths, this variability in binding makes prediction as difficult as less accurate [47]. There are four prediction methods for IEDB MHC II prediction tool: ARB, SMM_align, Sturniolo's method and a consensus method. ARB predict IC 50 values through combination of searches different peptide sizes and alleles into a single global prediction based on ARB matrices. SMM-align is a matrix-based method with extensions incorporating flanking residues outside of binding grooves. It also predicts the IC50 values of peptides. The consensus approach was used which combines the outcome of the three methods. Firstly, a random scan set of Swiss-Prot proteins and achieve scores for 2,000,000 random peptides, thereafter, act as reference to rank new predictions. The consensus method uses the median rank of the three approaches as the final prediction score [48]. All internationally conserved epitopes that bind to alleles at score equal or less than 10 percentile rank were selected for further analysis.

2.6. Population Coverage Calculation

All potential MHC I and MHC II binders from Spike glycoprotein were assessed for population coverage against the whole world population and Saudi Arabia, and other populations that had been reported MERS CoV cases. Calculations achieved using the selected MHC-I and MHC-II interacted alleles by the IEDB population coverage calculation tool http://tools.iedb.org/tools/population/iedb_input [49].

2.7. Homology Modeling

The complete 3D structure of Spike glycoprotein was obtained by phyre2, (http://www.sbg.bio.ic.ac.uk/phyre2) which uses advanced remote homology detection methods to build 3D models. UCSF Chimera (version 1.8) was used to visualize the 3D structure, which is currently available within the Chimera package and available from the chimera web site (http://www.cgl.ucsf.edu/cimera). Homology modeling was achieved for further verification of the service accessibility and hydrophilicity of B lymphocyte epitopes predicted, as well as visualization of all predicted T cell epitopes in the structural level [50,51].

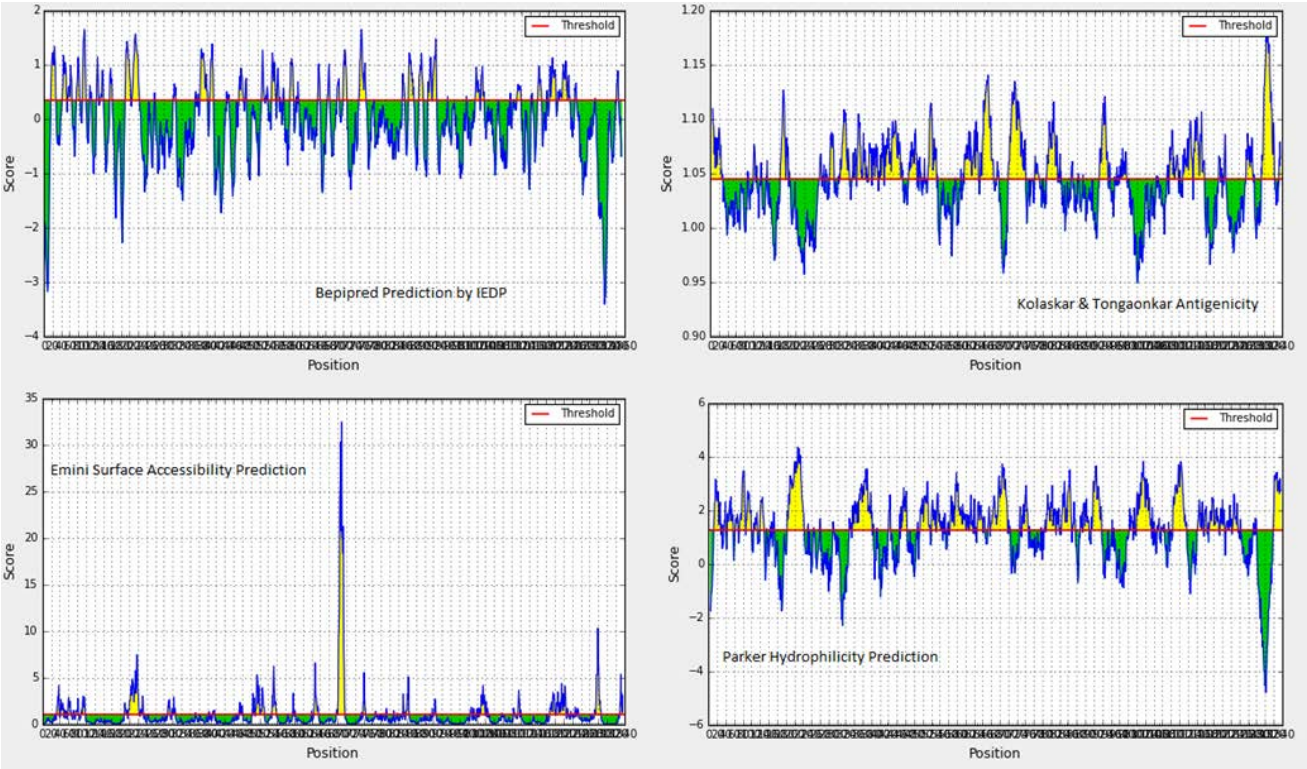

Figure 1. Prediction of B-cell epitopes by different scales

Yellow areas above threshold (red line) are proposed to be a part of B cell epitope. While green areas are not.

3. Results

3.1. Prediction of B-cell Epitopes

Spike glycoprotein was subjected to Bepipred linear epitope prediction, Emini surface accessibility, Kolaskar and Tongaonkar antigenicity, Parker hydrophobicity and Chou and Fasman beta turn prediction methods in IEDB, Figure 1.

In Bepipred Linear Epitope Prediction method; the average binders score of Spike glycoprotein to B cell was 0.35, all values equal or greater than the default threshold 0.35 were predicted to be potential B cell binders.

In Emini surface accessibility prediction; the average surface accessibility areas of the protein was scored as 1.000, all values equal or greater than the default threshold 1.0 were regarded potentially in the surface.

The default threshold of antigenicity of the protein was 1.045; all values greater than 1.045 were considered as potential antigenic determinants.

In Parker hydrophilicity prediction; the average hydrophilicity score of the protein was 1.286, all values equal or greater than the default threshold 1.286 were potentially hydrophilic.

Two internationally conserved epitopes had succeeded all prediction methods, epitope **GTPPQVY** from 391 to 397 was found to have the highest score, followed by **LTPRSVRSVP** from 745 to 754. The result is summarized in Table 1 and proposed epitopes are shown in Figure 2 at the structural level of the spike protein.

Table 1. list of B- cell epitopes predicted by different scales

Start	End	Peptide	Length	Surface Score[1]	Antigenicity Score[2]	Hydrophilicity Score[3]	Beta-Turn Score[4]
16	22	ESYVDVG	7	0.714	1.076	2.957	1.047
42	50	KTWPRPIDV	9	1.738	1.015	0.844	1.039
55	57	GII	3	0.25	1.059	-3.433	0.833
59	64	PQGRTY	6	3.13	0.983	3.55	1.185
76	83	PYQGDHGD	8	2.722	0.978	4.963	1.329
92	96	ATGTT	5	0.995	0.933	4.68	1.02
109	111	VKQ	3	1.322	1.109	2.667	0.83
154	157	GNFS	4	0.763	0.938	2.5	1.288
190	193	RSGN	4	1.725	0.884	5.85	1.375
195	197	CPA	3	0.431	1.18	1.867	1.123
214	223	CSDGNYNRNA	10	1.743	0.959	4.9	1.307
305	308	SDRK	4	3.621	0.92	6.6	1.212
354	356	FDV	3	0.552	1.113	-0.967	0.853
365	368	SFEA	4	0.839	1.004	1.8	0.857
370	373	PSGS	4	1.135	0.991	5.2	1.485
375	383	VEQAEGVEC	9	0.374	1.076	3.467	0.846
391	397	GTPPQVY	7	1.482	1.067	2.214	1.169
462	466	GPISQ	5	0.824	1.023	2.46	1.192
538	543	GDYYRK	6	4.239	0.977	3.633	1.21
546	550	SPLEG	5	0.97	1.01	2.58	1.168
576	581	QYGTDT	6	2.491	0.956	5.033	1.177
641	646	YYSDDG	6	2.422	0.99	4.733	1.365
688	692	QYSRS	5	3.16	1.015	4.26	1.186
701	709	DSTYGPLQT	9	2.246	1.007	3.289	1.178
738	741	LPDT	4	1.269	1.022	2.025	1.132
745	754	LTPRSVRSVP	10	1.23	1.082	1.42	1.035
840	845	LRQDDS	6	2.788	0.98	4.583	1.145
857	866	QSSPIIPGFG	10	0.353	1.031	0.94	1.154
878	885	SISTGSRS	8	1.104	0.982	4.137	1.207
900	908	IADPGYMQG	9	0.677	0.988	1.944	1.106
911	913	DCM	3	0.456	1.035	2.4	1.083
915	917	QGP	3	1.363	0.984	4.6	1.353
1010	1019	TGFTTTNEAF	10	1.031	0.938	2.5	0.956
1024	1031	DAVNNNAQ	8	1.59	0.965	4.688	1.118
1108	1118	KAQSKRSGFCG	11	1.031	1.008	3.664	1.125
1160	1165	NPTNCI	6	0.578	1.015	2.45	1.21
1185	1192	SYTGSSFY	8	1.276	1.029	2.175	1.211
1196	1198	PIT	3	0.804	1.042	-0.233	0.983
1205	1207	VAP	3	0.596	1.17	0.167	0.893
1210	1223	TYQNISTNLPPPLL	14	1.396	1.047	0.407	1.099
1228	1232	GIDFQ	5	0.575	1	0.9	1.014
1244	1247	TSIP	4	0.866	1.034	1.45	1.095
1342	1345	EYDL	4	1.544	1.032	1.675	0.982

[1] threshold :1.000 [2] threshold : 1.045 [3] threshold: 1.286 [4] threshold : 1.009

Position of peptides is according to position of amino acids in the Spike glycoprotein.

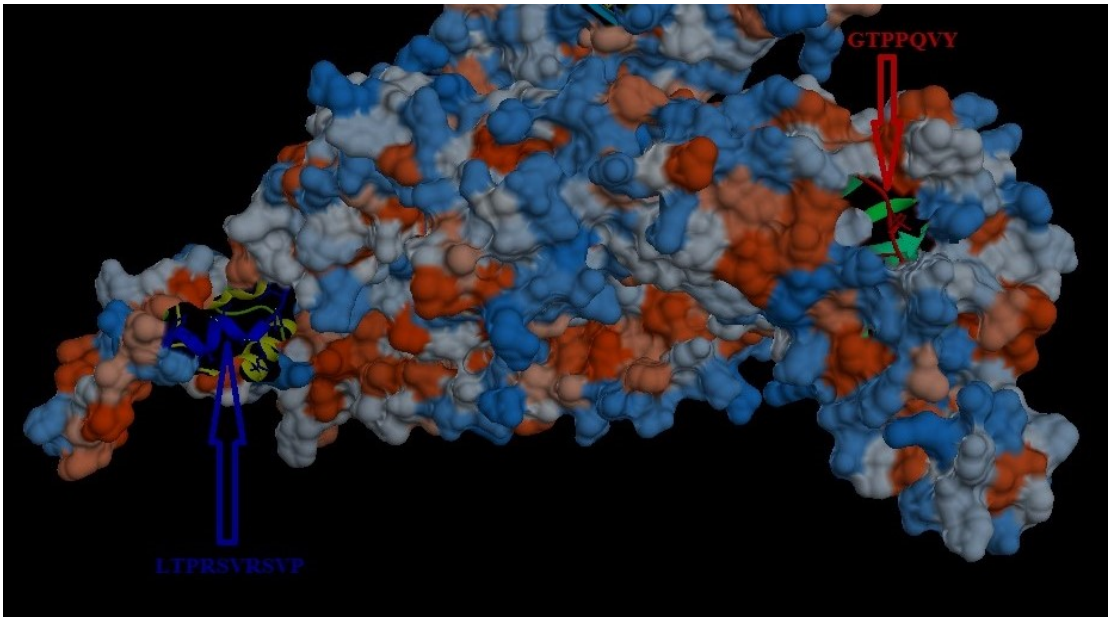

Figure 2. proposed B-Cell Epitopes

Proposed epitopes of B cell that are conserved in all retrieved Spike glycoprotein in the structural level.

The Chou and Fasman beta turn prediction method was used with the default threshold 1.009, for more confirmation of the prediction.

3.2. Prediction of Cytotoxic T-lymphocyte Epitopes and Interaction with MHC Class I

Based on Consensus (ann/smm/comblib_sidney2008) with percentile rank ≤1, 123 epitopes were predicted to interact with different MHC I alleles. The epitope FSFGVTQEY showed high affinity to interact with 9 alleles. All epitopes and their corresponding MHC-1 alleles are shown in Table 2. Figure 3 displays epitopes in the structural level.

Table 2. list of epitopes that have binding affinity with the MHC Class I alleles

Epitope	Start	End	Allele	Percentile Rank	Smm - ic50
YYSIIPHSI	292	300	HLA-A*24:02	0.15	53.04
			HLA-A*23:01	0.3	43.8
VVNAPNGLY	1127	1135	HLA-A*30:02	0.2	40.59
			HLA-A*29:02	0.4	48.61
LYGGNMFQF	273	281	HLA-A*23:01	0.2	22.26
			HLA-A*24:02	0.25	35.21
ITYQGLFPY	69	77	HLA-A*29:02	0.2	15.59
			HLA-B*46:01	0.25	1000.64
			HLA-A*30:02	0.45	68.62
			HLA-A*01:01	0.55	592.31
			HLA-A*11:01	0.6	80.97
			HLA-A*26:01	0.75	557.97
			HLA-B*15:01	0.8	123.88
			HLA-B*58:01	0.8	118.39
			HLA-A*25:01	0.9	1014.1
FSFGVTQEY	786	794	HLA-B*46:01	0.2	53.99
			HLA-A*01:01	0.3	304.47
			HLA-B*35:01	0.4	19.26
			HLA-A*29:02	0.55	55.43
			HLA-A*26:01	0.75	978.61
			HLA-C*06:02	0.9	42.93
			HLA-A*25:01	0.95	1644.67
			HLA-B*15:01	1	152.41
			HLA-B*58:01	1	158.23
FSSRYVDLY	266	274	HLA-A*01:01	0.2	57.09
			HLA-A*29:02	0.85	125.24
NTTLLDLTY	1256	1264	HLA-A*01:01	0.2	149.12
			HLA-A*29:02	0.55	58.31
YEMLSLQQV	1264	1272	HLA-B*44:02	0.2	144.41
			HLA-B*40:02	0.35	43.94
IPFAQSIFY	970	978	HLA-B*35:01	0.2	3.33
			HLA-B*53:01	0.4	90.34

			HLA-A*29:02	0.9	109.08
VYKLQPLTF	315	323	HLA-A*24:02	0.2	36.19
			HLA-A*23:01	0.3	23.14
ESFDVESGV	352	360	HLA-A*68:02	0.2	7.37
IEVVSAYGL	1147	1155	HLA-B*40:01	0.25	34.43
KEYFNLRNC	229	237	HLA-B*40:02	0.25	25.28
FQDELDEFF	1231	1239	HLA-B*38:01	0.25	95.29
			HLA-A*02:06	0.4	7.07
SFKEYFNLR	227	235	HLA-A*31:01	0.25	9.45
DTIKYYSII	288	296	HLA-A*25:01	0.25	55.09
			HLA-A*26:01	0.4	197.52
FQNCTAVGV	617	625	HLA-A*02:06	0.25	3.36
			HLA-B*39:01	0.55	75.93
ILDYFSYPL	442	450	HLA-A*02:01	0.3	10.72
			HLA-A*32:01	0.5	55.58
SLILDYFSY	440	448	HLA-A*29:02	0.3	38.97
			HLA-A*30:02	0.7	163.48
			HLA-A*25:01	1	875.14
GLYFMHVGY	1133	1141	HLA-A*29:02	0.3	38.88
			HLA-A*03:01	0.75	176.93
LLDFSVDGY	324	332	HLA-A*01:01	0.3	209.19
WSYTGSSFY	1184	1192	HLA-A*01:01	0.3	217.55
			HLA-A*30:02	0.4	109.76
			HLA-A*29:02	0.5	83.13
			HLA-B*15:01	0.6	86.5
			HLA-B*35:01	1	43.11
AFYCILEPR	182	190	HLA-A*31:01	0.3	12.78
LSIPTNFSF	780	788	HLA-B*15:01	0.3	44.16
			HLA-B*58:01	0.3	19.24
			HLA-B*57:01	0.4	236.49
			HLA-A*23:01	0.55	194.73
MEAAYTSSL	943	951	HLA-B*44:02	0.3	
			HLA-B*40:01	0.4	63.66
			HLA-B*39:01	0.4	49.14
			HLA-B*40:02	0.45	88.28
			HLA-B*48:01	0.65	738.92
			HLA-B*18:01	0.65	357.31
LASELSNTF	1036	1044	HLA-B*35:01	0.3	11.71
			HLA-B*53:01	0.6	145.51
			HLA-B*46:01	0.85	4337.9
			HLA-B*15:01	1	150.31
SPLEGGGWL	546	554	HLA-B*07:02	0.3	28.63
KQFANGFVV	110	118	HLA-A*02:06	0.3	5.77
			HLA-B*48:01	0.6	1070.53
YYSDDGNYY	641	649	HLA-A*29:02	0.3	21.57
GYYSDDGNY	640	648	HLA-A*30:02	0.35	85.01
			HLA-A*29:02	0.55	46.53
TTITKPLKY	489	497	HLA-A*29:02	0.35	34.65
			HLA-A*01:01	0.4	323.26
			HLA-A*25:01	0.45	167.91
			HLA-A*26:01	1	565.73
LYFMHVGYY	1134	1142	HLA-A*29:02	0.35	26.47
			HLA-A*30:02	0.45	136.29
LSMKSDLSV	450	458	HLA-C*15:02	0.35	11.39
ITEDEILEW	245	253	HLA-B*57:01	0.35	114.77
			HLA-B*58:01	0.9	88.57
KLQPLTFLL	317	325	HLA-A*02:01	0.4	18.8
			HLA-A*32:01	0.5	50.23
TLLDLTYEM	0.4	27.93	HLA-A*02:01	1258	1266
CYSSLILDY	437	445	HLA-A*29:02	0.4	30.46
YSDDGNYYC	642	650	HLA-A*01:01	0.4	206.8
GQGTHIVSF	1118	1126	HLA-B*15:01	0.4	48.53
YPLSMKSDL	448	456	HLA-B*07:02	0.4	47.3
			HLA-B*53:01	0.9	264.78
WTAGLSSFA	960	968	HLA-A*68:02	0.4	22.94

SSFAAIPFA	965	973	HLA-A*30:01	0.4	25.47
LLRAFYCIL	179	187	HLA-B*08:01	0.4	75.43
QTAQGVHLF	258	266	HLA-A*26:01	0.45	393.2
			HLA-B*58:01	0.7	91.26
			HLA-A*25:01	0.75	1071.72
			HLA-A*32:01	0.9	111.67
EVVSAYGLC	1148	1156	HLA-A*26:01	0.45	213.6
MLSLQQVVK	1266	1274	HLA-A*03:01	0.45	58.99
KVTIADPGY	897	905	HLA-A*30:02	0.45	92.57
TEDEILEWF	246	254	HLA-B*18:01	0.45	203.25
ITKPLKYSY	491	499	HLA-A*30:02	0.5	105.79
			HLA-B*57:01	0.55	358.77
			HLA-A*29:02	0.75	107.59
KELGNYTYY	1284	1292	HLA-B*18:01	0.5	255.88
			HLA-B*44:02	0.95	626.04
LSSFAAIPF	964	972	HLA-B*15:01	0.5	77.27
LTFLLDFSV	321	329	HLA-A*68:02	0.5	28.09
ESAALSAQL	1090	1098	HLA-A*68:02	0.5	38.07
AAIPFAQSI	968	976	HLA-C*15:02	0.5	15.97
FANGFVVRI	112	120	HLA-B*53:01	0.7	103.01
			HLA-B*53:01	0.7	103.01
RQDDSVRNL	841	849	HLA-B*48:01	0.5	499.57
FNLRNCTFM	232	240	HLA-B*14:02	0.55	250.06
ITITYQGLF	67	75	HLA-B*57:01	0.55	225.33
			HLA-A*26:01	0.65	313.77
			HLA-B*58:01	0.8	102.16
MDVNMEAAY	939	947	HLA-B*35:01	0.6	11.66
LPDGCGTLL	172	180	HLA-B*07:02	0.6	48.85
IAFNHPIQV	762	770	HLA-B*51:01	0.6	1234.89
			HLA-C*06:02	1	67.42
ASIAFNHPI	760	768	HLA-C*15:02	0.6	20.02
			HLA-A*32:01	0.8	101.38
MIHSVFLLM	1	9	HLA-A*26:01	0.65	547.78
			HLA-A*29:02	0.85	138.6
FLLTPTESY	10	18	HLA-A*29:02	0.65	84.09
			HLA-B*46:01	0.95	2255.69
			HLA-B*15:01	1	155.24
TYQNISTNL	1210	1218	HLA-A*23:01	0.65	248.56
			HLA-A*24:02	0.7	418.44
MTEQLQMGF	563	571	HLA-A*01:01	0.65	752.58
SQLGNCVEY	598	606	HLA-A*30:02	0.7	129.86
VECDFSPLL	381	389	HLA-B*40:01	0.7	101.13
RRDSTYGPL	699	707	HLA-B*39:01	0.7	80.43
			HLA-B*27:05	0.8	173.61
			HLA-B*38:01	1	1054.56
FSVDGYIRR	327	335	HLA-A*68:01	0.7	39.76
MLGSSVGNF	148	156	HLA-B*15:01	0.7	96.83
QQRFVYDAY	627	635	HLA-B*15:01	0.7	59.7
SVFLLMFLL	4	12	HLA-A*68:02	0.7	55.03
FIAGLVALA	1303	1311	HLA-A*68:02	0.7	49.16
LSPLEGGGW	545	553	HLA-B*57:01	0.7	177.75
LINGRLTTL	1070	1078	HLA-B*14:02	0.7	483.11
SFDVESGVY	353	361	HLA-A*30:02	0.75	100.57
DEWSYTGSS	1182	1190	HLA-B*18:01	0.75	304.82
GNYNRNASL	217	225	HLA-B*14:02	0.75	240.46
TQEYIQTTI	791	799	HLA-B*38:01	0.75	680.88
GVHLFSSRY	262	270	HLA-A*29:02	0.75	105.87
YFNLRNCTF	231	239	HLA-A*24:02	0.75	385.15
QPLTFLLDF	319	327	HLA-B*53:01	0.8	328
FAQSIFYRL	972	980	HLA-B*53:01	0.8	190.06
			HLA-B*35:03	0.95	486.62
FYVYKLQPL	313	321	HLA-A*23:01	0.8	417.28
LLMFLLTPT	7	15	HLA-A*02:01	0.8	45.2
LLFDKVTIA	893	901	HLA-A*02:01	0.8	48.88
TYSNITITY	63	71	HLA-A*29:02	0.8	77.58

NLRNCTFMY	233	241	HLA-A*29:02	0.8	79.94
YGVSGRGVF	609	617	HLA-B*46:01	0.8	2340.35
AALSAQLAK	1092	1100	HLA-A*11:01	0.8	62.56
YSSLILDYF	438	446	HLA-A*01:01	0.8	631.76
GRLTTLNAF	1073	1081	HLA-B*27:05	0.8	163.9
LVRSESAAL	1086	1094	HLA-B*07:02	0.8	134.24
TSIPNFGSL	1244	1252	HLA-A*25:01	0.8	1358.56
FVVRIGAAA	116	124	HLA-A*68:02	0.8	65.55
QSSPIIPGF	857	865	HLA-B*58:01	0.8	100.53
RLASIAFNH	758	766	HLA-A*03:01	0.8	248.77
IPTNFSFGV	782	790	HLA-B*51:01	0.8	1540.39
DEILEWFGI	248	256	HLA-B*44:02	0.85	386.9
DELDEFFKN	1233	1241	HLA-B*18:01	0.85	177.44
LEFANDTKI	588	596	HLA-B*40:01	0.85	99.29
			HLA-B*40:02	0.85	134.85
YYRKQLSPL	540	548	HLA-A*23:01	0.85	254.94
			HLA-A*24:02	1	691.24
GTNCMGKLK	1324	1332	HLA-A*11:01	0.85	105.27
QYVAGYKVL	927	935	HLA-A*23:01	0.9	401.26
NHIEVVSAY	1145	1153	HLA-B*18:01	0.9	566.29
			HLA-A*26:01	1	457.73
RLTTLNAFV	1074	1082	HLA-A*02:01	0.9	47
YSLYGVSGR	606	614	HLA-A*68:01	0.9	60.88
VAMTEQLQM	561	569	HLA-B*46:01	0.9	847.77
IHSVFLLMF	2	10	HLA-A*24:02	0.9	491.62
HSVFLLMFL	3	11	HLA-C*15:02	0.9	21.45
SVRNLFASV	845	853	HLA-A*30:01	0.9	55.33
			HLA-B*07:02	1	252.87
			HLA-B*08:01	1	345.58
TLLRAFYCI	178	186	HLA-A*23:01	0.95	328.42
ESYIDLKEL	1278	1286	HLA-C*15:02	0.95	7.73
KVTVDCKQY	801	809	HLA-A*30:02	1	214.02
ICAQYVAGY	924	932	HLA-A*30:02	1	256.13
GWTAGLSSF	959	967	HLA-A*23:01	1	269.42
TQTAQGVHL	257	265	HLA-B*39:01	1	182.56
AIEDLLFDK	889	897	HLA-A*11:01	1	96.45
TFGAISASI	1043	1051	HLA-A*24:02	1	643.62
YYPSNHIEV	1141	1149	HLA-A*24:02	1	585.64
SIFYRLNGV	975	983	HLA-A*68:02	1	79.91
LPPLMDVNM	935	943	HLA-B*51:01	1	1998.16

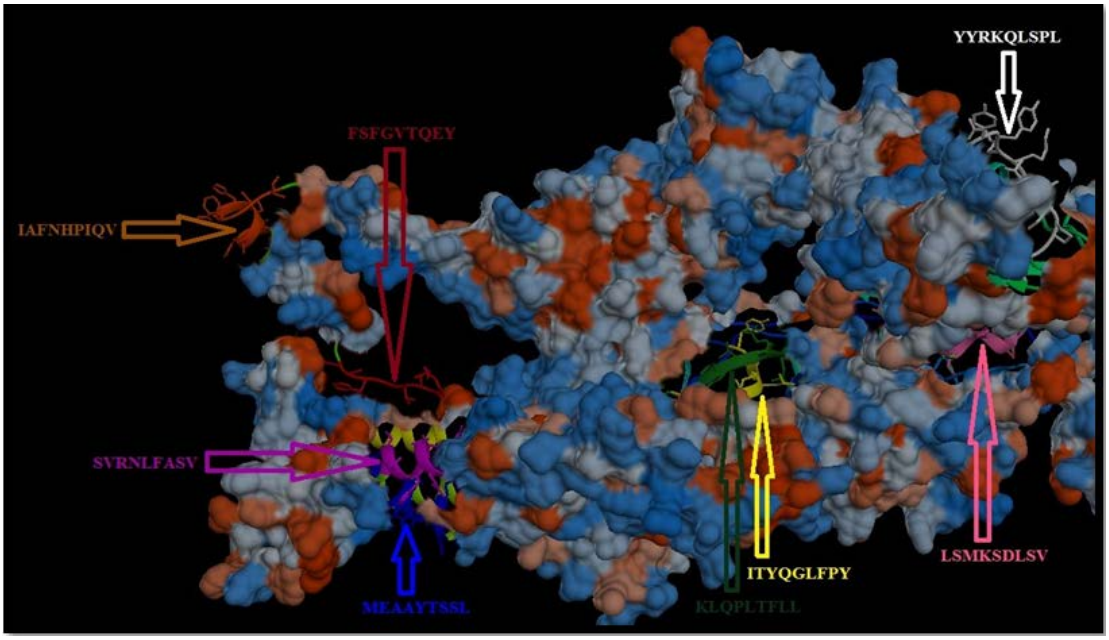

Figure 3. proposed T-Cell epitopes that interact with MHC Class I

Proposed epitopes of MHC I that are conserved in all Spike glycoprotein strains in the structural level

3.3. Prediction of T Helper Cell Epitopes and Interaction with MHC Class II

Based on Consensus (smm/nn/sturniolo) with percentile rank ≤10, there were 374 predicted epitopes found to interact with MHC-II alleles [data not shown] from which the peptide (core) **FNLTLLEPV** shows very high binding affinity to 26 alleles. The promiscuous epitopes and those which bind more than twelve different alleles are summarized in Table 3. Figure 4 displays epitopes in the structural level.

Table 3. list of epitopes that have binding affinity with ≥12 Class II alleles

Epitope	Allele	Percentile rank	Peptide	Start	End	Smm-ic50
FKLSIPTNF	HLA-DRB3*01:01	0.01	SYFKLSIPTNFSFGV	776	790	316
	HLA-DRB1*09:01	0.04	SYFKLSIPTNFSFGV	776	776	48
	HLA-DRB1*07:01	0.52	SYFKLSIPTNFSFGV	776	776	13
	HLA-DRB1*04:05	1.28	SYFKLSIPTNFSFGV	776	776	99
	HLA-DRB5*01:01	1.27	SYFKLSIPTNFSFGV	776	776	65
	HLA-DRB1*04:01	2.1	SYFKLSIPTNFSFGV	776	776	139
	HLA-DRB1*03:01	2.66	SYFKLSIPTNFSFGV	776	776	371
	HLA-DRB1*11:01	4	SYFKLSIPTNFSFGV	776	776	378
	HLA-DRB1*13:02	6.41	SYFKLSIPTNFSFGV	776	776	229
	HLA-DRB1*08:02	9.17	NSSYFKLSIPTNFSF	774	788	1056
	HLA-DRB1*01:01	10	NSSYFKLSIPTNFSF	774	788	51
VFLLMFLLT	HLA-DPA1*01:03	0.01	IHSVFLLMFLLTPTE	2	16	32
	HLA-DPB1*02:01	0.01	IHSVFLLMFLLTPTE	2	16	32
	HLA-DPA1*01	0.02	IHSVFLLMFLLTPTE	2	16	71
	HLA-DPB1*04:01	0.02	IHSVFLLMFLLTPTE	2	16	71
	HLA-DRB1*11:01	0.38	IHSVFLLMFLLTPTE	2	16	203
	HLA-DPA1*02:01	2.47	IHSVFLLMFLLTPTE	2	16	254
	HLA-DPB1*05:01	2.47	IHSVFLLMFLLTPTE	2	16	254
	HLA-DQA1*04:01	4.93	IHSVFLLMFLLTPTE	2	16	2387
	HLA-DQB1*04:02	4.93	IHSVFLLMFLLTPTE	2	16	2387
	HLA-DPB1*01:01	7.64	IHSVFLLMFLLTPTE	2	16	18
	HLA-DPA1*02:01	7.64	IHSVFLLMFLLTPTE	2	16	18
	HLA-DRB4*01:01	7.32	MIHSVFLLMFLLTPT	2	15	222
IAGLVALAL	HLA-DPA1*01/DPB1*04:01	9.11	LGFIAGLVALALCVF	1301	1315	952
	HLA-DPA1*03:01/DPB1*04:02	8.31	LGFIAGLVALALCVF	1301	1315	206
	HLA-DRB1*01:01	0.77	LGFIAGLVALALCVF	1301	1315	7
	HLA-DQA1*01:02/DQB1*06:02	1.43	LGFIAGLVALALCVF	1301	1315	107
	HLA-DRB1*11:01	1.82	LGFIAGLVALALCVF	1301	1315	295
	HLA-DQA1*03:01/DQB1*03:02	6.66	LGFIAGLVALALCVF	1301	1315	1043
	HLA-DQA1*01:01/DQB1*05:01	6.15	LGFIAGLVALALCVF	1301	1315	3602
	HLA-DRB1*15:01	3.26	LGFIAGLVALALCVF	1301	1315	196
	HLA-DPA1*02:01/DPB1*01:01	2.35	GFIAGLVALALCVFF	1302	1316	492
	HLA-DRB1*08:02	2.73	GFIAGLVALALCVFF	1302	1316	763
	HLA-DQA1*05:01/DQB1*02:01	4.6	GFIAGLVALALCVFF	1302	1316	774
	HLA-DRB1*04:01	7.08	GFIAGLVALALCVFF	1302	1316	204
	HLA-DQA1*04:01/DQB1*04:02	7.43	IWLGFIAGLVALALC	1299	1313	889
	HLA-DRB1*07:01	6.4	IWLGFIAGLVALALC	1299	1313	124
FLLMFLLTP	HLA-DPA1*01/DPB1*04:01	0.06	SVFLLMFLLTPTESY	4	18	112
	HLA-DRB1*04:01	4.16	MIHSVFLLMFLLTPT	1	15	142
	HLA-DRB1*11:01	1.14	MIHSVFLLMFLLTPT	1	15	454
	HLA-DPA1*03:01/DPB1*04:02	1.17	MIHSVFLLMFLLTPT	1	15	19
	HLA-DRB3*01:01	1.58	SVFLLMFLLTPTESY	4	18	817
	HLA-DQA1*01:01/DQB1*05:01	2.06	MIHSVFLLMFLLTPT	1	15	564
	HLA-DRB1*04:05	3.04	MIHSVFLLMFLLTPT	1	15	58
	HLA-DRB1*12:01	4.06	MIHSVFLLMFLLTPT	1	15	427
	HLA-DRB5*01:01	5.9	MIHSVFLLMFLLTPT	1	15	701
	HLA-DQA1*05:01/DQB1*02:01	6.49	MIHSVFLLMFLLTPT	1	15	940
	HLA-DPA1*01:03/DPB1*02:01	7.23	FLLMFLLTPTESYVD	6	20	366
FYVYKLQPL	HLA-DPA1*02:01/DPB1*01:01	0.11	WAAFYVYKLQPLTFL	310	324	97
	HLA-DPA1*03:01/DPB1*04:02	0.27	WAAFYVYKLQPLTFL	310	324	78
	HLA-DPA1*01:03/DPB1*02:01	0.28	WAAFYVYKLQPLTFL	310	324	142
	HLA-DPA1*01/DPB1*04:01	1.66	WAAFYVYKLQPLTFL	310	324	268
	HLA-DQA1*01:01/DQB1*05:01	1.91	WAAFYVYKLQPLTFL	310	324	540
	HLA-DRB4*01:01	4.06	WAAFYVYKLQPLTFL	310	324	159

	HLA-DRB1*15:01	2.61	WAAFYVYKLQPLTFL	310	324	165
	HLA-DRB1*04:05	3.3	WAAFYVYKLQPLTFL	310	324	152
	HLA-DRB1*07:01	2.94	AWAAFYVYKLQPLTF	309	323	205
	HLA-DRB1*11:01	5.05	KAWAAFYVYKLQPLT	308	322	242
	HLA-DRB3*01:01	5.17	WAAFYVYKLQPLTFL	310	324	790
	HLA-DRB1*08:02	6.32	KAWAAFYVYKLQPLT	308	322	2141
YKLQPLTFL	HLA-DPA1*02:01/DPB1*01:01	0.1	AFYVYKLQPLTFLLD	312	326	50
	HLA-DPA1*01:03/DPB1*02:01	0.43	AFYVYKLQPLTFLLD	312	326	64
	HLA-DRB1*09:01	1.91	AAFYVYKLQPLTFLL	311	325	280
	HLA-DPA1*02:01/DPB1*05:01	2.1	AFYVYKLQPLTFLLD	312	326	445
	HLA-DRB5*01:01	2.68	FYVYKLQPLTFLLDF	313	327	215
	HLA-DRB1*04:01	2.97	AAFYVYKLQPLTFLL	311	325	253
	HLA-DRB4*01:01	3.84	YVYKLQPLTFLLDFS	314	328	314
	HLA-DRB1*12:01	4.94	AFYVYKLQPLTFLLD	312	326	491
	HLA-DRB3*01:01	5.29	AAFYVYKLQPLTFLL	311	325	768
	HLA-DRB1*01:01	5.4	AAFYVYKLQPLTFLL	311	325	28
	HLA-DRB1*15:01	7.64	YVYKLQPLTFLLDFS	314	328	826
	HLA-DQA1*05:01/DQB1*02:01	8.64	AAFYVYKLQPLTFLL	311	325	1116
LQPLTFLLD	HLA-DPA1*02:01/DPB1*01:01	0.58	VYKLQPLTFLLDFSV	315	329	69
	HLA-DPA1*03:01/DPB1*04:02	2.99	VYKLQPLTFLLDFSV	315	329	158
	HLA-DPA1*01:03/DPB1*02:01	2.3	VYKLQPLTFLLDFSV	315	329	67
	HLA-DPA1*01/DPB1*04:01	2.35	VYKLQPLTFLLDFSV	315	329	94
	HLA-DPA1*02:01/DPB1*05:01	2.32	VYKLQPLTFLLDFSV	315	329	622
	HLA-DRB1*04:05	5.01	VYKLQPLTFLLDFSV	315	329	340
	HLA-DRB4*01:01	6.78	YKLQPLTFLLDFSVD	316	330	589
FAQSIFYRL	HLA-DPA1*01/DPB1*04:01	0.17	FAAIPFAQSIFYRLN	967	981	19
	HLA-DPA1*02:01/DPB1*01:01	0.56	FAAIPFAQSIFYRLN	967	981	65
	HLA-DPA1*01:03/DPB1*02:01	0.74	FAAIPFAQSIFYRLN	967	981	18
	HLA-DRB1*01:01	1.36	PFAQSIFYRLNGVGI	971	985	10
	HLA-DRB1*09:01	1.44	PFAQSIFYRLNGVGI	971	985	255
	HLA-DRB1*07:01	1.78	PFAQSIFYRLNGVGI	971	985	98
	HLA-DPA1*03:01/DPB1*04:02	2.62	FAAIPFAQSIFYRLN	967	981	80
	HLA-DRB3*01:01	4.63	FAQSIFYRLNGVGIT	972	986	267
	HLA-DRB1*03:01	5.88	FAQSIFYRLNGVGIT	972	986	3237
	HLA-DRB5*01:01	6.27	FAAIPFAQSIFYRLN	967	981	394
	HLA-DRB1*12:01	7.23	IPFAQSIFYRLNGVG	970	984	657
	HLA-DRB1*11:01	7.59	FAAIPFAQSIFYRLN	967	981	1123
	HLA-DRB1*04:05	9.35	PFAQSIFYRLNGVGI	971	985	655
FIAGLVALA	HLA-DRB1*09:01	0.28	YIWLGFIAGLVALAL	1298	1312	61
	HLA-DPA1*03:01/DPB1*04:02	0.29	YIWLGFIAGLVALAL	1298	1312	206
	HLA-DQA1*05:01/DQB1*03:01	0.55	YIWLGFIAGLVALAL	1298	1312	9
	HLA-DRB1*12:01	0.57	YIWLGFIAGLVALAL	1298	1312	140
	HLA-DRB1*01:01	0.77	YIWLGFIAGLVALAL	1298	1312	7
	HLA-DPA1*02:01/DPB1*01:01	1.19	LGFIAGLVALALCVF	1301	1315	341
	HLA-DQA1*01:02/DQB1*06:02	1.66	YIWLGFIAGLVALAL	1298	1312	117
	HLA-DQA1*05:01/DQB1*02:01	1.78	YIWLGFIAGLVALAL	1298	1312	487
	HLA-DRB1*03:01	6.66	YIWLGFIAGLVALAL	1298	1312	1448
	HLA-DRB1*08:02	2.73	YIWLGFIAGLVALAL	1298	1312	393
	HLA-DRB1*04:05	5.13	YIWLGFIAGLVALAL	1298	1312	81
	HLA-DRB1*04:01	7.08	YIWLGFIAGLVALAL	1298	1312	121
FNLTLLEPV	HLA-DPA1*03:01/DPB1*04:02	0.33	FGGDFNLTLLEPVSI	865	879	23
	HLA-DPA1*02:01/DPB1*01:01	0.36	FGGDFNLTLLEPVSI	865	879	93
	HLA-DRB1*04:05	0.66	FGGDFNLTLLEPVSI	865	879	42
	HLA-DRB1*09:01	1.7	FGGDFNLTLLEPVSI	865	879	119
	HLA-DQA1*01:01/DQB1*05:01	2.22	FGGDFNLTLLEPVSI	865	879	1215
	HLA-DPA1*01:03/DPB1*02:01	2.76	FGGDFNLTLLEPVSI	865	879	54
	HLA-DQA1*03:01/DQB1*03:02	3.57	FGGDFNLTLLEPVSI	865	879	1142
	HLA-DRB5*01:01	5.05	FGGDFNLTLLEPVSI	865	879	567
	HLA-DRB1*03:01	5.24	GGDFNLTLLEPVSIS	866	880	2459
	HLA-DRB1*11:01	5.43	FGGDFNLTLLEPVSI	865	879	453
	HLA-DRB1*01:01	5.59	FGGDFNLTLLEPVSI	865	879	29
	HLA-DRB1*12:01	5.74	GDFNLTLLEPVSIST	867	881	550
	HLA-DPA1*01/DPB1*04:01	6.85	FGGDFNLTLLEPVSI	865	879	134
	HLA-DQA1*05:01/DQB1*02:01	7.12	GGDFNLTLLEPVSIS	866	880	992
	HLA-DRB1*07:01	8.77	FGGDFNLTLLEPVSI	865	879	46

	HLA-DRB3*01:01	8.79	GGDFNLTLLEPVSIS	866	880	1971
	HLA-DRB1*13:02	9.28	FGGDFNLTLLEPVSI	865	879	337
SFAAIPFAQ	HLA-DPA1*01/DPB1*04:01	1.45	TAGLSSFAAIPFAQS	961	975	
	HLA-DQA1*03:01/DQB1*03:02	4.19	TAGLSSFAAIPFAQS	961	975	1256
	HLA-DPA1*01:03/DPB1*02:01	1.81	TAGLSSFAAIPFAQS	961	975	90
	HLA-DQA1*04:01/DQB1*04:02	2.29	TAGLSSFAAIPFAQS	961	975	567
	HLA-DQA1*05:01/DQB1*03:01	4.04	TAGLSSFAAIPFAQS	961	975	62
	HLA-DPA1*02:01/DPB1*01:01	8.18	TAGLSSFAAIPFAQS	961	975	178
FAAIPFAQS	HLA-DRB1*08:02	0.37	LSSFAAIPFAQSIFY	964	978	316
	HLA-DPA1*01/DPB1*04:01	1.08	LSSFAAIPFAQSIFY	964	978	48
	HLA-DPA1*01:03/DPB1*02:01	1.1	LSSFAAIPFAQSIFY	964	978	83
	HLA-DRB1*09:01	2	LSSFAAIPFAQSIFY	964	978	289
	HLA-DPA1*03:01/DPB1*04:02	9.48	LSSFAAIPFAQSIFY	964	978	160
	HLA-DRB1*11:01	4.24	LSSFAAIPFAQSIFY	964	978	623
	HLA-DPA1*02:01/DPB1*01:01	5.41	LSSFAAIPFAQSIFY	964	978	141
	HLA-DRB3*01:01	6.41	LSSFAAIPFAQSIFY	964	978	4180
	HLA-DRB1*04:01	7.44	LSSFAAIPFAQSIFY	964	978	619
	HLA-DPA1*02:01/DPB1*05:01	7.8	SSFAAIPFAQSIFYR	965	979	115
	HLA-DQA1*01:02/DQB1*06:02	9.26	LSSFAAIPFAQSIFY	964	978	310
	HLA-DRB1*04:05	9.35	SSFAAIPFAQSIFYR	965	979	655
	HLA-DRB5*01:01	9.95	GLSSFAAIPFAQSIF	963	977	913
	HLA-DQA1*04:01/DQB1*04:02	3.63	SSFAAIPFAQSIFYR	965	979	997
FLLTPTESY	HLA-DRB1*12:01	0.39	VFLLMFLLTPTESYV	5	19	118
	HLA-DRB3*01:01	0.96	LLMFLLTPTESYVDV	7	21	765
	HLA-DRB1*04:01	0.99	LLMFLLTPTESYVDV	7	21	146
	HLA-DRB1*03:01	1.99	LLMFLLTPTESYVDV	7	21	268
	HLA-DRB1*11:01	2.21	LLMFLLTPTESYVDV	7	21	397
	HLA-DRB1*04:05	2.8	MFLLTPTESYVDVGP	9	23	823
	HLA-DRB1*07:01	3.45	LLMFLLTPTESYVDV	7	21	148
	HLA-DPA1*03:01/DPB1*04:02	5.46	LLMFLLTPTESYVDV	7	21	245
	HLA-DRB5*01:01	6.18	VFLLMFLLTPTESYV	5	19	731
	HLA-DRB1*09:01	8.72	LLMFLLTPTESYVDV	7	21	888
	HLA-DQA1*05:01/DQB1*02:01	7.95	VFLLMFLLTPTESYV	5	19	1060
	HLA-DPA1*01:03/DPB1*02:01	7.42	LLMFLLTPTESYVDV	7	21	527
YEMLSLQQV	HLA-DRB4*01:01	0.59	DLTYEMLSLQQVVKA	1261	1275	59
	HLA-DPA1*02:01/DPB1*01:01	0.74	DLTYEMLSLQQVVKA	1261	1275	267
	HLA-DPA1*03:01/DPB1*04:02	2.04	DLTYEMLSLQQVVKA	1261	1275	430
	HLA-DPA1*01:03/DPB1*02:01	5.86	DLTYEMLSLQQVVKA	1261	1275	565
	HLA-DQA1*04:01/DQB1*04:02	4.2	LLDLTYEMLSLQQVV	1259	1273	1825
	HLA-DRB1*11:01	4.33	DLTYEMLSLQQVVKA	1261	1275	583
	HLA-DRB1*09:01	4.7	DLTYEMLSLQQVVKA	1261	1275	225
	HLA-DQA1*03:01/DQB1*03:02	9.18	DLTYEMLSLQQVVKA	1261	1275	1318
	HLA-DPA1*01/DPB1*04:01	8.25	DLTYEMLSLQQVVKA	1261	1275	1276
	HLA-DRB1*08:02	6.37	LDLTYEMLSLQQVVK	1260	1274	2155
	HLA-DRB1*01:01	6.87	DLTYEMLSLQQVVKA	1261	1275	35
	HLA-DQA1*01:02/DQB1*06:02	7.06	DLTYEMLSLQQVVKA	1261	1275	316
FSYPLSMKS	HLA-DRB1*04:01	3.37	ILDYFSYPLSMKSDL	442	456	223
	HLA-DPA1*03:01/DPB1*04:02	2.32	ILDYFSYPLSMKSDL	442	456	481
	HLA-DRB1*09:01	3.87	ILDYFSYPLSMKSDL	442	456	462
	HLA-DRB5*01:01	4.34	ILDYFSYPLSMKSDL	442	456	103
	HLA-DRB1*04:05	4.45	ILDYFSYPLSMKSDL	442	456	285
	HLA-DRB1*12:01	7	ILDYFSYPLSMKSDL	442	456	641
	HLA-DRB1*11:01	5.43	ILDYFSYPLSMKSDL	442	456	194
	HLA-DPA1*01/DPB1*04:01	8.92	ILDYFSYPLSMKSDL	442	456	1343
	HLA-DRB1*03:01	9.58	ILDYFSYPLSMKSDL	442	456	3168
	HLA-DRB1*08:02	9.61	DYFSYPLSMKSDLSV	444	458	
LILDYFSYP	HLA-DPA1*02:01/DPB1*01:01	0.56	SSLILDYFSYPLSMK	439	453	80
	HLA-DRB1*15:01	0.71	NCYSSLILDYFSYPL	436	450	41
	HLA-DPA1*01:03/DPB1*02:01	1.39	NCYSSLILDYFSYPL	436	450	132
	HLA-DPA1*01/DPB1*04:01	2.03	SSLILDYFSYPLSMK	439	453	333
	HLA-DQA1*01:01/DQB1*05:01	2.3	NCYSSLILDYFSYPL	436	450	1065
	HLA-DPA1*03:01/DPB1*04:02	4.43	SSLILDYFSYPLSMK	439	453	233
	HLA-DRB1*03:01	3.6	NCYSSLILDYFSYPL	436	450	1252
	HLA-DPA1*02:01/DPB1*05:01	4.32	SSLILDYFSYPLSMK	439	453	496
	HLA-DRB3*01:01	4.95	NCYSSLILDYFSYPL	436	450	497

	HLA-DRB1*12:01	9.22	SSLILDYFSYPLSMK	439	453	802
LLRAFYCIL	HLA-DPA1*02:01/DPB1*01:01	3.87	DGCGTLLRAFYCILE	174	188	676
	HLA-DPA1*01:03/DPB1*02:01	3.23	DGCGTLLRAFYCILE	174	188	196
	HLA-DQA1*01:01/DQB1*05:01	6.39	DGCGTLLRAFYCILE	174	188	1605
	HLA-DPA1*01/DPB1*04:01	6.81	DGCGTLLRAFYCILE	174	188	372
	HLA-DRB3*01:01	7.78	DGCGTLLRAFYCILE	174	188	9351
	HLA-DRB5*01:01	8.79	GTLLRAFYCILEPRS	177	191	1047
	HLA-DPA1*02:01/DPB1*05:01	9.83	DGCGTLLRAFYCILE	174	188	1320
YIQTTIQKV	HLA-DRB1*07:01	0.62	TQEYIQTTIQKVTVD	791	805	73
	HLA-DPA1*02:01/DPB1*01:01	2.14	TQEYIQTTIQKVTVD	791	805	281
	HLA-DPA1*01:03/DPB1*02:01	3.49	TQEYIQTTIQKVTVD	791	805	594
	HLA-DRB4*01:01	4.81	TQEYIQTTIQKVTVD	791	805	186
	HLA-DPA1*03:01/DPB1*04:02	5.45	TQEYIQTTIQKVTVD	791	805	342
	HLA-DPA1*01/DPB1*04:01	5.73	TQEYIQTTIQKVTVD	791	805	954
	HLA-DRB1*09:01	8.31	TQEYIQTTIQKVTVD	791	805	513
	HLA-DRB1*04:05	9.81	TQEYIQTTIQKVTVD	791	805	490
LSSFAAIPF	HLA-DRB1*09:01	1.24	TAGLSSFAAIPFAQS	961	975	210
	HLA-DRB1*15:01	3.93	TAGLSSFAAIPFAQS	961	975	227
	HLA-DRB1*08:02	3.26	TAGLSSFAAIPFAQS	961	975	295
	HLA-DRB1*07:01	8.74	TAGLSSFAAIPFAQS	961	975	49
	HLA-DQA1*01:01/DQB1*05:01	4.39	TAGLSSFAAIPFAQS	961	975	1464
	HLA-DQA1*03:01/DQB1*03:02	5.23	WTAGLSSFAAIPFAQ	960	974	1433
	HLA-DRB3*01:01	6.41	TAGLSSFAAIPFAQS	961	975	3785
	HLA-DRB5*01:01	9.95	TAGLSSFAAIPFAQS	961	975	429
	HLA-DRB1*04:05	7.58	TAGLSSFAAIPFAQS	961	975	159
	HLA-DRB1*11:01	7.59	TAGLSSFAAIPFAQS	961	975	574
	HLA-DRB4*01:01	8.44	TAGLSSFAAIPFAQS	961	975	966
	HLA-DPA1*02:01/DPB1*01:01	8.6	WTAGLSSFAAIPFAQ	960	974	1167
	HLA-DRB1*04:01	7.44	TAGLSSFAAIPFAQS	961	975	306

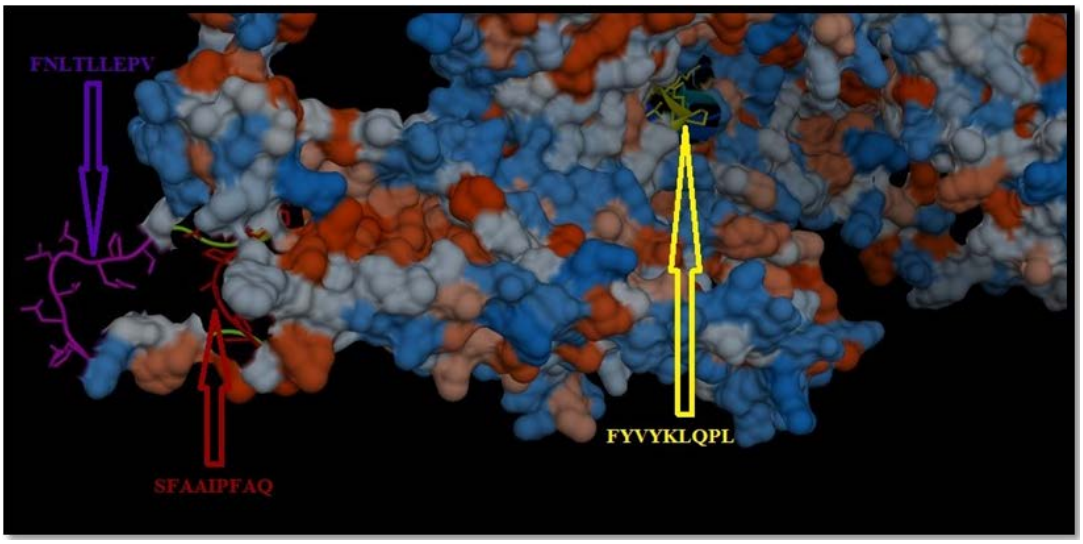

Figure 4. proposed T-Cell epitopes that interact with MHC Class II

Epitopes of MHC II that are conserved in all international strains are in the structural level.

3.4. Analysis of the Population Coverage

Epitopes that are suggested interacting with MHC-I and II alleles especially high affinity binding epitopes and that can bind to different set of alleles were selected for population coverage analysis. The results of population coverage of proposed epitopes are listed in Table 4.

In MHC class I, three epitopes that interact with most frequent MHC class I alleles **(**ITYQGLFPY, FSFGVTQEY and KLQPLTFLL) gave high percentage against the whole world population by IEDB population coverage tool. The population coverage of FSFGVTQEY and IAFNHPIQV show good coverage among Saudis (60.35% and 54.70%) respectively, Table 5 represents the population coverage for the proposed epitopes and shows their corresponding coverage percentage.

Also in MHC class II, four epitopes that interact with most frequent MHC class II alleles (FNLTLLEPV, FAAIPFAQS, SFAAIPFAQ and FYVYKLQPL) gave high percentage against the whole world population by IEDB population coverage tool. The population coverage for these proposed epitopes is excellent internationally as (99.99%) as well as Saudi Arabia (99.62%) and (100%) in England. Table 5 represents the proposed epitopes. FNLTLLEPV which binds to DQB1*02:01 and DQB1*03:02 that is reported in high frequency in Saudi

Arabia was found to definitely increase the population coverage in Saudi Arabia as seen in tables (5A, B and C).

Table 4. population coverage of all epitopes in both MHC class I and II in the world

EPITOPE	COVERAGE CLASS 1	NO. OF ALLELS	EPITOPE2	COVERAGE CLASS II	NO. OF ALLELS
ITYQGLFPY	51.96%	9	FYVYKLQPL	98.63%	17
FSFGVTQEY	53.37%	9	FNLTLLEPV	99.79%	24
MEAAYTSSL	28.72%	6	SFAAIPFAQ	98.79%	12
SVRNLFASV	25.62%	3	FAAIPFAQS	99.09%	20
KLQPLTFLL	42.66%	2	**Epitope set**	**99.99%**	
YYRKQLSPL	26.18%	2			
IAFNHPIQV	21.80%	2			
LSMKSDLSV	4.41%	1			
GLYFMHVGY	20.35%	2			
epitope set	**98.27%**				

Table (5A). population coverage for Class I proposed epitopes for selected regions

Epitope/population	World	England	Korea, South	Philippines
ITYQGLFPY	51.96%	64.25%	57.44%	13.51%
FSFGVTQEY	53.37%	65.95%	56.82%	0%
MEAAYTSSL	28.72%	39.16%	30.70%	73.99%
SVRNLFASV	25.62%	52.64%	13.58%	0%
KLQPLTFLL	42.66%	53.12%	30.75%	51%
YYRKQLSPL	26.18%	16.61%	38.66%	53.76%
IAFNHPIQV	21.80%	26.71%	22.93%	0%
LSMKSDLSV	4.41%	4.55%	4.88%	17.19%
GLYFMHVGY	20.35%	31.05%	3.56%	0%
epitope set	**98.27%**	**99.93%**	**97.36%**	**97.93%**

Table (5B). population coverage for Class II proposed epitopes for selected regions

Epitope/population	Oman	UAE	Saudia Arabia	Jordan
ITYQGLFPY	61.93%	2.19%	41.03%	36.69%
FSFGVTQEY	50.97%	2.19%	60.35%	36.22%
MEAAYTSSL	9.18%	8.99%	5.62%	9.94%
SVRNLFASV	23.79%	0%	30.14%	23.87%
KLQPLTFLL	55.11%	0%	38.22%	30.28%
YYRKQLSPL	15.36%	0%	22.03%	14.07%
IAFNHPIQV	32.43%	0%	54.70%	26.64%
LSMKSDLSV	0%	0%	16.83%	8.80%
GLYFMHVGY	11.45%	0%	14.44%	10.13%
epitope set	**98.97%**	**11.08%**	**97.29%**	**90.63%**

Table (5C). population coverage for Class I proposed epitopes for selected regions

epitope/population	Iran	Lebanon	Israel
ITYQGLFPY	46.21%	0%	46.92%
FSFGVTQEY	43.84%	22.91%	53.34%
MEAAYTSSL	13.7%	0%	27.25%
SVRNLFASV	20.97%	0%	10.05%
KLQPLTFLL	46.56%	0%	25.85%
YYRKQLSPL	8.80%	0%	18.81%
IAFNHPIQV	30.23%	22.91%	25.63%
LSMKSDLSV	17.88%	9.75%	5.71%
GLYFMHVGY	10.89%	0%	10.04%
epitope set	**96.28%**	**31.44%**	**92.08%**

Table 6. population coverage for Class II proposed epitopes for selected regions

Population area	FYVYKLQPL	FNLTLLEPV	SFAAIPFAQ	FAAIPFAQS	Epitope set
World	98.63%	99.79%	98.79%	99.09%	**99.99%**
Saudi Arabia	96.41%	98.92%	90.07%	94.19%	**99.62%**
Jordan	44.93%	93.99%	82.04%	59.54%	**99.19%**
England	97.70%	99.82%	96.65%	97.63%	**100%**
Korea; South	76.79%	93.71%	70.39%	88.24%	**99.29%**
Philippines	21.36%	27.55%	0.00%	23.42%	**28.56%**
Iran	67.16%	95.50%	77.59%	50.10%	**99.38%**
Israel	66.78%	96.17%	81.43%	45.83%	**99.58%**
Lebanon	97.28%	99.05%	93.24%	96.25%	**99.90%**

4. Discussion

Vaccination has proven to be the mainstay in prevention of various deadly infectious diseases. Historically, live-attenuated or inactivated forms of microbial pathogens (viruses, bacteria, etc.) have been used for induction of immune responses that protect the host against the subsequent infections, such vaccine might contain unnecessary proteins for the induction of protective immunity and may lead to allergenic responses, This has created an interest in peptide vaccines containing only epitopes capable of inducing positive, desirable T cell and B cell mediated immune responses. There are many peptide vaccines under development, such as vaccine for human immunodeficiency virus (HIV), hepatitis C virus (HCV), malaria, foot and mouth disease, swine fever, influenza, anthrax, human papilloma virus (HPV), therapeutic anti-cancer vaccines pancreatic cancer, melanoma, non-small cell lung cancer, advanced hepatocellular carcinoma cutaneous T-cell lymphoma and B-Cell chronic lymphocytic leukemia. [52-69]

In this study, we aimed to determine the highly potential immunogenic epitopes for B and T cells - the prime molecules of cell mediated and humoral immunity - as vaccine candidates for the highly lethal MERS coronaviral infection using the whole Spike glycoprotein as a target. Unlike previous studies and for better determination of the best candidate epitopes; all S protein was under investigation not only the receptor binding domain (RBD). Both complete S protein and RBD were *in vitro* and *in vivo* proven to elicit B lymphocyte to produce antibodies [12,70-75], therefore, we conducted this study to propose epitopes that would be enough to select out of these large non conserved domains to hopefully achieve the same result.

Conservancy in S protein in MERS CoV was found promising for peptide vaccine design, however, *Tuhin ali et al* found that a peptide region of 367-606 (which is the receptor binding domain) remained conserved, unlike our findings which showed that this region is no longer conserved, this region is no longer conserved, Tuhin ali et al,.2014 also found epitope CYSSLILDY interacting with 11 different MHC I alleles at threshold IC50 $\leq$ 100 and percentile rank $\leq$ 1, while we found this epitope only interacting with one allele as illustrated in Table 1, furthermore, this epitope was not predicted as B lymphocyte in our study [76].

To determine a potential and effective peptide antigen for B cell, epitopes should get above threshold scores in Bepipred linear epitope prediction, Emini surface accessibility, Parker hydrophobicity, Kolaskar and Tongaonkar antigenicity and Chou and Fasman beta turn prediction methods in IEDB. Epitopes illustrated in Table 1, are the only conserved regions from all retrieved strains of MERS coronavirus Spike glycoprotein that are available in NCBI database until 15th April 2016 and have high probability of activating humoral immune response. In the RBD, the only conserved epitope that was found activating B lymphocyte immune response was the 7mer epitope 391 GTPPQVY 397. LTPRSVRSVP epitope is located in S2 domain, and was the only conserved epitope that was found passing the thresholds in S2 domain of Spike glycoprotein.

Wang, L. et al had found important residues in RBD, that if mutated will affect the binding of RBD with the neutralizing antibody; 535W and 536E. 535W is conserved residue, the following 536E residue is not conserved as only one Spike glycoprotein retrieved in this study (gb/ALJ76286), collected from Taif - Saudi Arabia in 2014, is showing novel mutation in this residue. Although, these residues are no longer located in conserved peptide, a focus needs to be targeted to this region as they show high score in bepipred prediction tool [77].

Since the immune response of T cell is long lasting response comparing with B cell, where the antigen can easily escape the antibody memory response [54]. Additionally, CD8+ T and CD4+ T cell responses play a major role in antiviral immunity [55], designing of vaccine against T cell epitope is much more promising. For MHC Class I Alleles prediction, we chose the most common HLA-A and HLA-B alleles [56], and according to several studies in the allele frequencies among Saudis (the most endemic area) the highest frequency of HLA-A was observed for A*02, followed by A*68, A*24, and A*26, for HLA-B locus, B*51, B*50, and B*08 represent very high frequency, the highest observed frequencies for HLA-C were C*07, C*06 , and C*15, whereas for MHC II, DQB1*02, DQB1*03, and DQB1*06 were representing the highly frequent alleles. The binding affinity of those alleles with all of the conserved epitopes of S protein was examined [34,35,36,37,38].

Among 123 internationally conserved T cell epitopes predicted to interact with MHC Class I as shown in Table 2, ITYQGLFPY and FSFGVTQEY were found to interact with 9 alleles (HLA-A*29:02, HLA-B*46:01, HLA-A*30:02, HLA-A*01:01, HLA-A*11:01, HLA-A*26:01, HLA-B*15:01, HLA-B*58:01, HLA-A*25:01) and (HLA-B*46:01, HLA-A*01:01, HLA-B*35:01, HLA-A*29:02, HLA-A*26:01, HLA-C*06:02, HLA-A*25:01, HLA-B*15:01, HLA-B*58:01) respectively. IAFNHPIQV was found to bind with the HLA-B*51:01, HLA-C*06:02 that are highly frequent among Saudis as well as AAIPFAQSI which interact with the HLA-C*15:02. (FSFGVTQEY, MEAAYTSSL, SVRNLFASV, KLQPLTFLL, IAFNHPIQV and GLYFMHVGY) among the proposed MHCI binders are located in S2 domain, and only LSMKSDLSV is found to be located in RBD.

We found that 374 internationally conserved epitopes in S glycoprotein interact with MHC-II alleles as represented in Table 3. Epitope FNLTLLEPV is showing very high binding affinity to 26 MHC II alleles (HLA-DPA1*03:01, HLA-DPB1*04:02, HLA-DPA1*02:01, HLA-DPB1*01:01, HLA-DRB1*04:05, HLA-DRB1*09:01, HLA-DQA1*01:01, HLA-DQB1*05:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DQA1*03:01, HLA-DQB1*03:02, HLA-DRB5*01:01, HLA-DRB1*03:01, HLA-DRB1*11:01, HLA-DRB1*01:01, HLA-DRB1*12:01, HLA-DPA1*01:03, HLA-DPB1*02:01, HLA-DPA1*01, HLA-DPB1*04:01, HLA-DQA1*05:01, HLA-DQB1*02:01, HLA-DRB1*07:01, HLA-DRB3*01:01, HLA-DRB1*13:02) among which highly frequent alleles in Saudis, this epitope alone showed 99.79% coverage in the world, 98.92% in Saudi Arabia and 99.82% in England. FYVYKLQPL, FAAIPFAQS and SFAAIPFAQ also showed very high binding affinity. FNLTLLEPV,

FAAIPFAQS and SFAAIPFAQ are all located in S2 domain. All proposed MHC I and MHC II epitopes as illustrated in Table 5, Table 6 were better chosen to serve the best population coverage percentage as well as the lowest number of peptides to be used as multi epitope vaccines against Middle East Respiratory Syndrome.

5. Conclusion

Conventional peptide vaccine development methods are costly, and time consuming, *insilico* prediction is highly appreciated as it selects specific peptides in protein, which then tested *in vitro* and *in vivo* to verify and prove the effectiveness of the proposed epitopes to induce an immune response, as well as to be used as a diagnostic screening test. Herd immunity protocols can be conducted in countries with low population coverage to minimize the active transmission of the virus, especially among people contacting camels and other groups at risk. Ultimately, a vaccine needs to target a broad set of zoonotic CoVs, because the emergence of two zoonotic CoV infections in 10 years suggests that this will happen repeatedly.

Acknowledgments

Authors would like to thank African City of Technology members for their cooperation.

Competing Interest

The authors declare that they have no competing interests.

References

[1] World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV). http://www.who.int/mediacentre/factsheets/mers-cov/en/.

[2] Memish ZA, Cotten M, Watson SJ, Kellam P, Zumla A, Alhakeem RF, *etal*. Community Case Clusters of Middle East Respiratory Syndrome Coronavirus in Hafr Al-Batin, Kingdom of Saudi Arabia: A Descriptive Genomic study. Int J Infect Dis. 2014; 23: 63-68.

[3] Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palser AL, Ong SH, *etal*. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. MBio. 2014; 5(1):e01062-13.

[4] Chastel C. Middle East respiratory syndrome (MERS): bats or dromedary, which of them is responsible?. Bull Soc Pathol Exot. 2014; 107(2):69-73.

[5] Al-Dorzi HM, Alsolamy S, Arabi YM. Critically ill patients with Middle East respiratory syndrome coronavirus infection. Crit Care. 2016; 20(1): 65.

[6] Centers for Disease Control and Prevention. Middle East Respiratory Syndrome (MERS). http://www.cdc.gov/coronavirus/mers/about/index.htm.

[7] World Health Organization Regional Office for the Eastern Mediterranean. New coronavirus identified in two patients in the EMR. Weekly Epidemiologi-cal Monitor. 2012; 5(39) http://www.emro.who.int/images/stories/csr/documents/epi___issue_no__39.coronavirus.pdf.

[8] World Health Organization.Middle East respiratory syndrome coronavirus (MERS-CoV) – Saudi Arabia. http://www.who.int/csr/don/4-december-2015-mers-saudi-arabia/en/.

[9] Malczyk AH, Kupke A, Prüfer S, Scheuplein VA, Hutzler S, Kreuz D *etal*. A Highly Immunogenic and Protective Middle East Respiratory Syndrome Coronavirus Vaccine Based on a Recombinant Measles Virus Vaccine Platform. J Virol. 2015; 89(22):11654-67.

[10] Scobey T, Yount BL, Sims AC, Donaldson EF, Agnihothram SS, Menachery VD, etal. Reverse genetics with a full-length infectious cDNA of the Middle East respiratory syndrome coronavirus. Proc Natl Acad Sci. 2013; 110(40):16157-62.

[11] Public Health England. Risk assessment of Middle East respiratory syndrome coronavirus (MERS-CoV). https://www.gov.uk/government/uploads/system/uploads/attachment_data 505701/MERS-COV_RA_Mar2016_240216_RP__3_.pdf.

[12] Song F, Fux R, Provacia LB, Volz A, Eickmann M, Becker S, *etal*. Middle East respiratory syndrome coronavirus spike protein delivered by modified vaccinia virus Ankara efficiently induces virus-neutralizing antibodies. J Virol. 2013;87(21):11950-4.

[13] Memish ZA, Cotten M, Meyer B, Watson SJ, Alsahafi AJ, Al Rabeeah AA, *etal*. Human infection with MERS coronavirus after exposure to infected camels, Saudi Arabia, 2013. Emerg Infect Dis. 2014 Jun; 20(6):1012-5.

[14] Reusken C, Messadi L, Feyisa A, Ularamu H, Godeke G, Danmarwa A, *etal*. Geographic distribution of MERS coronavirus among dromedary camels, Africa .Emerging Infectious Diseases. 2014; 20(8):1370-1374.

[15] Levinson W. Medical Microbiology and Immunology (examination and board review). 8th edition. McGraw-Hill, San Francisco, 2004; p268.

[16] van Boheemen S, de Graaf M, Lauber C, Bestebroer TM, Raj VS, Zaki AM, etal. Genomic characterization of a newly discovered coronavirus associated with acute respiratory distress syndrome in humans. mBio. 2012; 3(6):e00473-12.

[17] Qian Z, Dominguez SR, Holmes KV. Role of the Spike Glycoprotein of Human Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in Virus Entry and Syncytia Formation. PLoS One. 2013; 8(10): e76469.

[18] Gierer S, Bertram S, Kaup F, Wrensch F, Heurich A, Krämer-Kühl A, *etal*. The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. J Virol. 2013 May; 87(10):5502-11.

[19] Shi J, Zhang J, Li S, Sun J, Teng Y, Wu M, *etal*. Epitope-Based Vaccine Target Screening against Highly Pathogenic MERS-CoV: An *In Silico* Approach Applied to Emerging Infectious Diseases. PLoS One. 2015; 10(12): e0144475.

[20] de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, etal. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. J Virol. 2013;87(14):7790-2.

[21] World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV) .http:// www.who.int/mediacentre/factsheets/mers-cov/en/.

[22] Graham RL, Donaldson EF, Baric RS. A decade after SARS: strategies for controlling emerging coronaviruses. Nat Rev Microbiol. 2013 D; 11(12):836-48.

[23] Ying T, Du L, Ju TW, Prabakaran P, Lau CC, Lu L, *etal*. Exceptionally potent neutralization of Middle East respiratory syndrome coronavirus by human monoclonal antibodies. J Virol. 2014; 88(14):7796-805.

[24] Wang L, Shi W, Joyce MG, Modjarrad K, Zhang Y, Leung K, *etal*. Evaluation of candidate vaccine approaches for MERS-CoV. Nat Commun. 2015; 6:7712.

[25] Rappuoli R. Reverse vaccinology. Curr Opin Microbiol. 2000; 3 (5):445-50.

[26] Purcell AW, McCluskey J, Rossjohn J. More than one reason to rethink the use of peptides in vaccine design. Nat Rev Drug Discov. 2007; 6(5):404-14.

[27] Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98.

[28] Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, Wheeler DK, Gabbard JL, Hix D, Sette A, Peters B. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2014 Oct 9. pii: gku938. [Epub ahead of print] PubMed PMID: 25300482.

[29] Anayet Hasan, Mehjabeen Hossain and Md. Jibran Alam. A Computational Assay to Design an Epitope-Based Peptide Vaccine Against Saint Louis Encephalitis Virus. *Bioinformatics and Biology Insights* 2013:7 347-355.

[30] Jens Erik Pontoppidan Larsen, Ole Lund and Morten Nielsen. Improved method for predicting linear B-cell epitopes. Immunome Res. 2006; 2: 2.

[31] Emini EA, Hughes JV, Perlow DS, Boger J. 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 55:836-839.

[32] Kolaskar AS, Tongaonkar PC. 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*276:172-174.

[33] Parker JM, Guo D, Hodges RS. 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry* 25:5425-5432.PMID: 2430611.

[34] Osman et al. HLA-A, -B, -C, -DRB1, and -DQB1 allele Lineages and Haplotype Frequencies among Saudis. Immunology and Immunogenetics Insights 2014:6 1-6.

[35] Hajeer AH, Sawidan FA, Bohlega S, et al. HLA class I and class II polymorphisms in Saudi patients with myasthenia gravis. Int J Immunogenet. 2009;36: 169-172.

[36] Hajeer AH, Al Balwi MA, Aytül Uyar F, et al. HLA-A, -B, -C, -DBR1 and -DQB1 allele and haplotype frequencies in Saudis using next generation sequencing technique. Tissue Antigens. 2013;82:2520258.

[37] Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. Nucleic Acids Res. 2011;39:D9130D919.

[38] Valluri V, Mustafa M, Santhosh A, et al. Frequencies of HLA-A, HLA-B, HLA-DR, and HLA-DQ phenotypes in the United Arab Emirates population. Tissue Antigens. 2005;66:107-113.

[39] Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B. 2012. Immune epitope database analysis resource. NAR.

[40] Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, Brunak S, Lund O. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12:1007-1017.

[41] Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, and Nielsen M. 2008. NetMHC-3.0: Accurate web accessible predictions of Human, Mouse, and Monkey MHC class I affinities for peptides of length 8-11. NAR 36:W509-512.

[42] Peters B, Sette A. 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. BMC Bioinformatics 6:132.

[43] Sidney J, Assarsson E, Moore C, Ngo S, Pinilla C, Sette A, Peters B. 2008. Quantitative peptide binding motifs for 19 human and mouse MHC class I molecules derived using positional scanning combinatorial peptide libraries. Immunome Res 4:2.

[44] Kim Y, Ponomarenko J, Zhu Z, et al. Immune epitope database analysis resource. Nucleic Acids Research. 2012;40(Web Server issue):W525-W530.

[45] Wang P, Sidney J, Dow C, Mothé B, Sette A, Peters B. 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol. 4(4):e1000048.

[46] Wang P, Sidney J, Kim Y, Sette A, Lund O, Nielsen M, Peters B. 2010. Peptide binding predictions for HLA DR, DP and DQ molecules. BMC Bioinformatics. 11:568.

[47] Pratik Narain Srivastava, Richa Jain, Shyam Dhar Dubey, Sharad Bhatnagar, Nabeel Ahmad. Prediction of Epitope-Based Peptides for Vaccine Development from Coat Proteins GP2 and VP24 of Ebola Virus Using Immunoinformatics, International Journal of Peptide Research and Therapeutics (2016) 22:119-133.

[48] Zhang, Q., Wang, P., Kim, Y., Haste-Andersen, P., Beaver, J., Bourne, P. E. *et al*. (2008). Immune epitope database analysis resource (IEDB-AR).*Nucleic Acids Research*, *36*(Web Server issue), W513-W518.

[49] Bui HH,Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006 Mar 17;7:153.

[50] The Phyre2 web portal for protein modeling, prediction and analysis Kelley LA et al. Nature Protocols 10, 845-858 (2015).

[51] UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem.* 2004 Oct;25(13):1605-12.

[52] Bachler, B.C.; Humbert, M.; Palikuqi, B.; Siddappa, N.B.; Lakhashe, S.K.; Rasmussen, R.A.; Ruprecht, R.M. Novel biopanning strategy to identify epitopes associated with vaccine protection. J. Virol. 2013, 87, 4403-4416.

[53] Perrie, Y.; Kirby, D.; Bramwell, V.W.; Mohammed, A.R. Recent developments in particulate-based vaccines. Recent Pat. Drug Deliv. Formul. 2007, 1, 117-129.

[54] Black, M.; Trent, A.; Tirrell, M.; Olive, C. Advances in the design and delivery of peptide subunit vaccines with a focus on toll-like receptor agonists. Expert Rev. Vaccines 2010, 9, 157-173.

[55] Sesardic, D. Synthetic peptide vaccines. J. Med. Microbiol. 1993, 39, 241-242.

[56] Liu, Y.; McNevin, J.; Zhao, H.; Tebit, D.M.; Troyer, R.M.; McSweyn, M.; Ghosh, A.K.; Shriner, D.; Arts, E.J.; McElrath, M.J.; et al. Evolution of human immunodeficiency virus type 1 cytotoxic T-lymphocyte epitopes: Fitness-balanced escape. J. Virol. 2007, 81, 12179-12188.

[57] Kolesanova, E.F.; Sanzhakov, M.A.; Kharybin, O.N. Development of the schedule for multiple parallel ―difficult Peptide synthesis on pins. Int. J. Pept. 2013.

[58] Epstein, J.E.; Giersing, B.; Mullen, G.; Moorthy, V.; Richie, T.L. Malaria vaccines: Are we getting closer? Curr. Opin. Mol. Ther. 2007, 9, 12-24.

[59] Volpina, O.M.; Gelfanov, V.M.; Yarov, A.V.; Surovoy, A.Y.; Chepurkin, A.V.; Ivanov, V.T. New virus-specific T-helper epitopes of foot-and-mouth disease viral VP1 protein. FEBS Lett. 1993, 333, 175-178.

[60] Tarradas, J.; Monso, M.; Munoz, M.; Rosell, R.; Fraile, L.; Frías, M.T.; Domingo, M.; Andreu, D.; Sobrino, F.; Ganges, L. Partial protection against classical swine fever virus elicited by dendrimeric vaccine-candidate peptides in domestic pigs. Vaccine 2011, 29, 4422-4429.

[61] Stanekova, Z.; Kiraly, J.; Stropkovska, A.; Mikušková, T.; Mucha, V.; Kostolanský, F.; Varečková, E. Heterosubtypic protective immunity against influenza a virus induced by fusion peptide of the hemagglutinin in comparison to ectodomain of M2 protein. Acta Virol. 2011, 55, 61-67.

[62] Oscherwitz, J.; Yu, F.; Cease, K.B. A synthetic peptide vaccine directed against the 2ss2–2ss3 loop of domain 2 of protective antigen protects rabbits from inhalation anthrax. J. Immunol. 2010, 185, 3661-3668.

[63] Solares, A.M.; Baladron, I.; Ramos, T.; Valenzuela, C.; Borbon, Z.; Fanjull, S.; Gonzalez, L.; Castillo, D.; Esmir, J.; Granadillo, M.; et al. Safety and immunogenicity of a human papillomavirus peptide vaccine (CIGB-228) in women with high-grade cervical intraepithelial neoplasia: first-in-human, proof-of-concept trial. ISRN Obstet. Gynecol. 2011.

[64] Bernhardt, S.L.; Gjertsen, M.K.; Trachsel, S.; Møller, M.; Eriksen, J.A.; Meo, M.; Buanes, T.; Gaudernack, G. Telomerase peptide vaccination of patients with non-resectable pancreatic cancer: A dose escalating phase I/II study. Br. J. Cancer 2006, 95, 1474-1482.

[65] Brunsvig, P.F.; Aamdal, S.; Gjertsen, M.K.; Kvalheim, G.; Markowski-Grimsrud, C.J.; Sve, I.; Dyrhaug, M.; Trachsel, S.; Møller, M.; Eriksen, J.A.; et al. Telomerase peptide vaccination: A phase I/II study in patients with non-small cell lung cancer. Cancer Immunol. Immunother. 2006, 55, 1553-1564.

[66] Brunsvig, P.F.; Kyte, J.A.; Kersten, C.; Sundstrøm, S.; Møller, M.; Nyakas, M.; Hansen, G.L.; Gaudernack, G.; Aamdal, S. Telomerase peptide vaccination in NSCLC: A phase II trial in stage III patients vaccinated after chemoradiotherapy and an 8-year update on a phase I/II trial. Clin. Cancer Res. 2011, 17, 6847-6857.

[67] Kyte, J.A.; Gaudernack, G.; Dueland, S.; Trachsel, S.; Julsrud, L.; Aamdal, S. Telomerase peptide vaccination combined with temozolomide: A clinical trial in stage IV melanoma patients. Clin. Cancer Res. 2011, 17, 4568-4580.

[68] Greten, T.F.; Forner, A.; Korangy, F.; N'Kontchou, G.; Barget, N.; Ayuso, C.; Ormandy, L.A.; Manns, M.P.; Beaugrand, M.; Bruix, J. A phase II open label trial evaluating safety and efficacy of a telomerase peptide vaccination in patients with advanced hepatocellular carcinoma. BMC Cancer 2010, 10, e209.

[69] Kyte, J.A.; Trachsel, S.; Risberg, B.; Thor, S.P.; Lislerud, K.; Gaudernack, G. Unconventional cytokine profiles and development of T cell memory in long-term survivors after cancer vaccination. Cancer Immunol. Immunother. 2009, 58, 1609-1626. 25.

[70] Du L, Zhao G, Yang Y, Qiu H, Wang L, Kou Z, et al. A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein. J Virol2014;88:7045-53.

[71] Jiang L, Wang N, Zuo T, Shi X, Poon KM, Wu Y, et al. Potent neutralizationof MERS-CoV by human neutralizing monoclonal antibodies to the viral spikeglycoprotein. Sci Transl Med 2014;6:234ra59.

[72] Ying T, Du L, Ju TW, Prabakaran P, Lau CC, Lu L, et al. Exceptionally potent neutralization of middle East respiratory syndrome coronavirus by human monoclonal antibodies. J Virol 2014;88:7796-805.

[73] Zhang N, Jiang S, Du L. Current advancements and potential strategies in thedevelopment of MERS-CoV vaccines. Expert Rev Vaccines 2014;13:761-74.

[74] Mou H, Raj VS, van Kuppeveld FJ, Rottier PJ, Haagmans BL, Bosch BJ. The receptor binding domain of the new Middle East respiratory syndrome coronavirus maps to a 231-residue region in the spike protein that efficiently elicits neutralizing antibodies. J Virol 2013;87:9379-83.

[75] Du L, Zhao G, Kou Z, Ma C, Sun S, Poon VK, et al. Identification of a receptor-binding domain in the S protein of the novel human coronavirus Middle Eastrespiratory syndrome coronavirus as an essential target for vaccine develop-ment. J Virol 2013;87:9939-42.

[76] Tuhin ali *et al*, Bioinformation 10(8): 533-538 (2014).

[77] Wang, L. et al. Evaluation of candidate vaccine approaches for MERS-CoV. Nat. Commun. 6:7712.

Supplementary material

Table (S1). Spike Glycoprotein strains retrieved, accession numbers and date of collection

Accession No	Country of collection	Date of collection
AIY60588.1	United Arab Emirates	13-Apr-14
AIY60578.1	United Arab Emirates	15-Nov-13
AIY60568.1	United Arab Emirates	17-Apr-14
AIY60558.1	United Arab Emirates	7-Mar-14
AIY60548.1	United Arab Emirates	19-Apr-14
AIY60538.1	United Arab Emirates	10-Apr-14
AIY60528.1	United Arab Emirates	10-Apr-14
AIY60518.1	United Arab Emirates	7-Apr-14
AKK52612.1	Saudi Arabia	1-Mar-15
AKK52602.1	Saudi Arabia	10-Feb-15
AKK52592.1	Saudi Arabia	1-Mar-15
AKK52582.1	Saudi Arabia	10-Feb-15
ALM26400.1	China	27-May-15
AHZ58501.1	USA	30-Apr-14
AHZ20790.1	Greece	18-Apr-14
ALW82753.1	Saudi Arabia	10-Feb-15
ALW82742.1	Saudi Arabia	13-Feb-15
ALW82731.1	Saudi Arabia	2-Feb-15
ALW82720.1	Saudi Arabia	2-Feb-15
ALW82702.1	Saudi Arabia	12-Feb-15
ALW82709.1	Saudi Arabia	5-Feb-15
ALW82691.1	Saudi Arabia	15-Feb-15
ALW82680.1	Saudi Arabia	7-Feb-15
ALW82669.1	Saudi Arabia	27-Mar-15
ALW82658.1	Saudi Arabia	10-May-15
ALW82647.1	Saudi Arabia	22-Feb-15
ALW82636.1	Saudi Arabia	9-Feb-15
AKM76239.1	Oman	28-Dec-13
AKM76229.1	Oman	28-Oct-13
AIZ48769.1	United Arab Emirates	29-Oct-13
AIZ48760.1	USA	Jun-14
AHZ64057.1	USA	10-May-14
AHC74098.1	Qatar	17-Oct-13
AHC74088.1	Qatar	13-Oct-13
ALD51904.1	Thailand	17-Jun-15
AJD81440.1	United Kingdom	13-Feb-13
AGV08379.1	Saudi Arabia	23-Oct-12
AML60270.1	South Korea	8-Jun-15
ALB08322.1	South Korea	24-Jun-15
ALB08311.1	South Korea	21-Jun-15
ALB08300.1	South Korea	18-Jun-15
ALB08289.1	South Korea	3-Jun-15
ALB08278.1	South Korea	13-Jun-15
ALB08267.1	South Korea	31-May-15
ALB08257.1	South Korea	19-Jun-15

ALB08246.1	South Korea	29-Jun-15
AKJ80137.2	China	27-May-15
AGV08584.1	Saudi Arabia	30-Oct-12
AGV08573.1	Saudi Arabia	23-May-13
AGV08558.1	Saudi Arabia	15-May-13
AGV08546.1	Saudi Arabia	11-May-13
AGV08535.1	Saudi Arabia	12-May-13
AGV08524.1	Saudi Arabia	8-May-13
AGV08505.1	Saudi Arabia	3-May-13
AGV08492.1	Saudi Arabia	30-May-13
AGV08480.1	Saudi Arabia	23-May-13
AGV08467.1	Saudi Arabia	13-May-13
AGV08455.1	Saudi Arabia	4-Jun-13
AGV08444.1	Saudi Arabia	7-May-13
AGV08438.1	Saudi Arabia	1-May-13
AGV08426.1	Saudi Arabia	1-May-13
AGV08408.1	Saudi Arabia	19-Jun-12
AGV08390.1	Saudi Arabia	5-Feb-13
AGN70973.1	Saudi Arabia	22-Apr-13
AGN70962.1	Saudi Arabia	9-May-13
AGN70951.1	Saudi Arabia	21-Apr-13
AGN70929.1	Saudi Arabia	1-May-13
AHY22565.1	Saudi Arabia	Nov-13
AHY22555.1	Saudi Arabia	Nov-13
AHY22545.1	Saudi Arabia	Nov-13
AHY22535.1	Saudi Arabia	Nov-13
AHY22525.1	Saudi Arabia	Nov-13
AHI48739.1	Saudi Arabia	5-Aug-13
AHI48737.1	Saudi Arabia	26-Aug-13
AHI48711.1	Saudi Arabia	19-Jun-13
AHI48702.1	Saudi Arabia	6-Aug-13
AHI48692.1	Saudi Arabia	15-Jul-13
AHI48682.1	Saudi Arabia	8-Aug-13
AHI48672.1	Saudi Arabia	12-Jun-13
AHI48662.1	Saudi Arabia	13-Aug-13
AHI48652.1	Saudi Arabia	18-Jun-13
AHI48626.1	Saudi Arabia	17-Jul-13
AHI48616.1	Saudi Arabia	11-Sep-13
AHI48605.1	Saudi Arabia	1-Mar-13
AHI48594.1	Saudi Arabia	12-Jun-13
AHI48583.1	Saudi Arabia	2-Jul-13
AHI48572.1	Saudi Arabia	15-Aug-13
AHI48561.1	Saudi Arabia	5-Aug-13
AHI48550.1	Saudi Arabia	12-Jun-13
AHI48539.1	Saudi Arabia	28-Aug-13
AHI48528.1	Saudi Arabia	17-Jul-13
AHI48517.1	Saudi Arabia	2-May-13
AKL80615.1	China	28-May-15
AKL80604.1	China	28-May-15
AKL80593.1	China	28-May-15
AKI29284.1	Saudi Arabia	6-Jan-15
AKI29275.1	Saudi Arabia	26-Jan-15
AKI29265.1	Saudi Arabia	21-Jan-15
AKI29255.1	Saudi Arabia	21-Jan-15
AJD81451.1	United Kingdom	10-Feb-13
AID50418.1	United Kingdom	10-Feb-13
AHN10812.1	Saudi Arabia	6-Nov-13
AHB33326.1	France	7-May-13
AHE78108.1	Saudi Arabia	5-Nov-13
AHE78097.1	Saudi Arabia	8-Nov-13
AIZ74450.1	France	7-May-13
AIZ74439.1	France	7-May-13
AIZ74433.1	France	7-May-13
AIZ74417.1	France	26-Apr-13
AIZ74405.1	France	26-Apr-13

AJG44124.1	United Arab Emirates	N.A*
AJG44113.1	United Arab Emirates	Jun-14
AJG44102.1	United Arab Emirates	Jun-14
AJG44091.1	United Arab Emirates	Jun-14
AJG44080.1	United Arab Emirates	Jun-14
AJG44069.1	United Arab Emirates	Jun-14
AJG44058.1	United Arab Emirates	Jun-14
AHL18090.1	Egypt	2013
ALT66880.1	United Arab Emirates	4-Mar-14
ALT66870.1	United Arab Emirates	12-Mar-14
ALS20350.1	Nigeria	13-Jan-15
ALR69641.1	United Arab Emirates: Dubai	N.A*
AID55103.1	Saudi Arabia	2014
AID55102.1	Saudi Arabia	2014
AID55101.1	Saudi Arabia	2014
AID55100.1	Saudi Arabia	2014
AID55099.1	Saudi Arabia	2014
AID55098.1	Saudi Arabia	2014
AID55097.1	Saudi Arabia	2014
AID55096.1	Saudi Arabia	2014
AID55095.1	Saudi Arabia	2014
AID55094.1	Saudi Arabia	2014
AID55093.1	Saudi Arabia	2014
AID55092.1	Saudi Arabia	2014
AID55091.1	Saudi Arabia	2014
AID55090.1	Saudi Arabia	2014
AID55089.1	Saudi Arabia	2014
AID55088.1	Saudi Arabia	2014
AID55087.1	Saudi Arabia	2014
AID55086.1	Saudi Arabia	2014
AID55085.1	Saudi Arabia	2014
AID55084.1	Saudi Arabia	2014
AID55083.1	Saudi Arabia	2014
AID55082.1	Saudi Arabia	2014
AID55081.1	Saudi Arabia	2014
AID55080.1	Saudi Arabia	2014
AID55079.1	Saudi Arabia	2014
AID55078.1	Saudi Arabia	2014
AID55077.1	Saudi Arabia	2014
AID55076.1	Saudi Arabia	2014
AID55075.1	Saudi Arabia	2014
AID55074.1	Saudi Arabia	2014
AID55073.1	Saudi Arabia	22-Apr-14
AID55072.1	Saudi Arabia	15-Apr-14
AID55071.1	Saudi Arabia	21-Apr-14
AID55070.1	Saudi Arabia	14-Apr-14
AID55069.1	Saudi Arabia	12-Apr-14
AID55068.1	Saudi Arabia	7-Apr-14
AID55067.1	Saudi Arabia	2014
AID55066.1	Saudi Arabia	2014
ALA50067.1	Saudi Arabia	Mar-15
ALA50056.1	Saudi Arabia	Mar-15
ALA50045.1	Saudi Arabia	Apr-15
ALA50034.1	Saudi Arabia	Apr-15
ALA50023.1	Saudi Arabia	Apr-15
ALA50012.1	Saudi Arabia	Apr-15
ALA50001.1	Saudi Arabia	Apr-15
ALA49990.1	Saudi Arabia	Apr-15
ALA49979.1	Saudi Arabia	Apr-15
ALA49968.1	Saudi Arabia	Apr-15
ALA49957.1	Saudi Arabia	Apr-15
ALA49946.1	Saudi Arabia	Mar-15
ALA49935.1	Saudi Arabia	Mar-15
ALA49924.1	Saudi Arabia	Mar-15
ALA49913.1	Saudi Arabia	Mar-15

ALA49902.1	Saudi Arabia	Mar-15
ALA49891.1	Saudi Arabia	Mar-15
ALA49880.1	Saudi Arabia	Mar-15
ALA49869.1	Saudi Arabia	Mar-15
ALA49858.1	Saudi Arabia	Mar-15
ALA49847.1	Saudi Arabia	Mar-15
ALA49836.1	Saudi Arabia	Mar-15
ALA49825.1	Saudi Arabia	Mar-15
ALA49814.1	Saudi Arabia	Feb-15
ALA49803.1	Saudi Arabia	Feb-15
ALA49792.1	Saudi Arabia	Jan-15
ALA49781.1	Saudi Arabia	Jan-15
ALA49770.1	Saudi Arabia	Jan-15
ALA49759.1	Saudi Arabia	Jan-15
ALA49748.1	Saudi Arabia	Jan-15
ALA49737.1	Saudi Arabia	Jan-15
ALA49726.1	Saudi Arabia	Jan-15
ALA49715.1	Saudi Arabia	Jan-15
ALA49704.1	Saudi Arabia	Sep-14
ALA49693.1	Saudi Arabia	Sep-14
ALA49682.1	Saudi Arabia	Sep-14
ALA49671.1	Saudi Arabia	Sep-14
ALA49660.1	Saudi Arabia	Sep-14
ALA49649.1	Saudi Arabia	Oct-14
ALA49638.1	Saudi Arabia	Oct-14
ALA49627.1	Saudi Arabia	Oct-14
ALA49616.1	Saudi Arabia	Oct-14
ALA49605.1	Saudi Arabia	Nov-14
ALA49594.1	Saudi Arabia	Nov-14
ALA49583.1	Saudi Arabia	Nov-14
ALA49572.1	Saudi Arabia	Dec-14
ALA49561.1	Saudi Arabia	Dec-14
ALA49550.1	Saudi Arabia	Dec-14
ALA49539.1	Saudi Arabia	Dec-14
ALA49528.1	Saudi Arabia	Dec-14
ALA49517.1	Saudi Arabia	Dec-14
ALA49506.1	Saudi Arabia	Dec-14
ALA49495.1	Saudi Arabia	Dec-14
ALA49484.1	Saudi Arabia	Dec-14
ALA49473.1	Saudi Arabia	Dec-14
ALA49462.1	Saudi Arabia	Dec-14
ALA49451.1	Saudi Arabia	Dec-14
ALA49440.1	Saudi Arabia	Dec-14
ALA49429.1	Saudi Arabia	Dec-14
ALA49418.1	Saudi Arabia	Dec-14
ALA49407.1	Saudi Arabia	Dec-14
ALA49396.1	Saudi Arabia	Dec-14
ALA49385.1	Saudi Arabia	Dec-14
ALA49374.1	Saudi Arabia	Sep-14
ALA49363.1	Saudi Arabia	Jul-14
ALA49352.1	Saudi Arabia	Jul-14
ALA49341.1	Saudi Arabia	May-14
ALK80311.1	South Korea	22-Jun-15
ALK80301.1	South Korea	10-Jun-15
ALK80291.1	South Korea	10-Jun-15
ALK80281.1	South Korea	11-Jun-15
ALK80271.1	South Korea	10-Jun-15
ALK80261.1	South Korea	10-Jun-15
ALK80251.1	South Korea	9-Jun-15
ALK80242.1	South Korea	10-Jun-15
ALK80232.1	South Korea	11-Jun-15
ALK80222.1	South Korea	8-Jun-15
ALK80212.1	South Korea	8-Jun-15
ALK80202.1	South Korea	11-Jun-15
ALK80192.1	South Korea	11-Jun-15

AMO03401.1	Egypt	17-Dec-14
AHZ90568.1	Tunisia	8-May-13
ALJ76286.1	Saudi Arabia	22-Nov-14
ALJ76285.1	Saudi Arabia	12-Nov-14
ALJ76284.1	Saudi Arabia	27-Oct-14
ALJ76283.1	Saudi Arabia	27-Oct-14
ALJ76282.1	Saudi Arabia	24-Oct-14
ALJ76281.1	Saudi Arabia	10-Sep-14
ALJ76280.1	Saudi Arabia	27-Oct-14
ALJ76279.1	Saudi Arabia	11-Nov-14
ALJ76278.1	Saudi Arabia	4-Nov-14
ALJ76277.1	Saudi Arabia	28-Nov-14
AKL59401.1	South Korea	20-May-15
AHX71946.1	Qatar	16-Feb-14
YP_009047204.1**		13-Jun-12

* Data not available.
** Ref sequence.