DOI : 10.5958/2277-8934.2022.00016.9 *Vol 29 No 2, p 117-125*

REVERSE DIAGNOSTIC WORKFLOW TO SHORTLIST MERS-COV SPIKE ANTIGENIC EPITOPES IN DROMEDARY CAMELS

Sabry M. El-Bahr[1,2], Mahmoud Kandeel[1,3], I.F. Albokhadaim[1], M. Al-Nazawi[1] and Hafiz Ahmed[4]

[1]Department of Biomedical Sciences, College of Veterinary Medicine, King Faisal University, Al-Ahsa 31982, P.O. 400, Saudi Arabia

[2]Department of Biochemistry, Faculty of Veterinary Medicine, Alexandria University, Alexandria 21523, Egypt

[3]Department of Pharmacology, Faculty of Veterinary Medicine, Kafr Elsheikh University, Egypt

[4]GlycoMentra Inc., 1450 South Rolling Road, Baltimore, MD 21227, USA

ABSTRACT

In this study, we provide a comprehensive reverse diagnostic workflow, in which, the *in-silico* amino acid composition was employed retrogradely to shortlist the viral spike antigenic epitopes to infer diagnostics efficiency. This study was aimed to explore the analytical estimates of changes in epitope composition among MERS-CoV lineages. This can be used as a predictor of the effectiveness of rapid diagnostic testing. Therefore, MERS-CoV lineage-specific spike protein sequences were extracted, aligned and compared. In addition, the degree of sequence similarity, as well as pairwise comparison, phylogenetic relations and antigenic epitopes analysis, have been conducted. The current findings indicated that no differences were observed in length and range of epitopes for each virus among all studied lineages. Most of epitopes sequences were conserved. However, few sequences showed few single amino acid mutations. About 20% of epitopes were located at the receptor-binding domain (RBD) and 80% of these were located in the other Spike`s domains. These mutations were related to lineage 2 and 3 and not for lineage 1, 4 and 5. The number of difference among viruses accession no (ALA49781, ALA49594, ALA49440, ALA50001, ALA49935, ALA49847, ALA49825, ALA49561, ALA49374, ALA49803, ALA49660, ALA49352, ALA49671, ALA49341, AHX00711, AHY22565, AJG44124, AJG44091, ALA49429, ALA49418, ALA49902, AHY22525, AHX71946, AHE78108.1, AHI48672.1, AHI48550) was maximum of 8. No gaps were observed in the epitopes alignment. The identity of spike protein among the lineages ranged from 99.5 -100%. The study concluded that any of studied epitopes are suitable for production of rapid tests of MERS CoV in dromedary camels, particularly that produced from lineages 1, 4 and 5.

Key words: MERS CoV, camels, epitopes, viral lineage, rapid test

The Middle East respiratory syndrome coronavirus (MERS-CoV) was initially discovered in a patient with acute respiratory distress syndrome in the Kingdom of Saudi Arabia in October 2012 (Zaki *et al,* 2012). MERS-CoV is diagnosed primarily by molecular methods, i.e. Real-time reverse transcriptase-polymerase amplification (RT-PCR) (Corman *et al,* 2012a; 2012b), reverse transcription–loop-mediated isothermal amplification (RTLAMP) and reverse transcription–recombinase polymerase amplification (RT-RTPA) (Shirato *et al,* 2014; Abd El Wahed *et al,* 2013). MERS-CoV or closely related viruses have also been detected in seropositive camels using a variety of serological techniques. Protein microarrays (Reusken *et al,* 2013a, 2013b; Meyer *et al,* 2013), a recombinant spike immunofluorescent assay (Buchholz *et al,* 2013; Annan *et al,* 2013), an indirect enzyme-linked immunosorbent assay (ELISA) (Alexandersen *et al,* 2014), microneutralisation, and spike pseudoparticle neutralisation are some of the techniques used (Perera *et al,* 2013). Molecular tests are relatively expensive and considered problematic for screening large numbers of animals in a short period of time; therefore, a rapid, inexpensive, sensitive, and specific test for the diagnosis of MERS-CoV in camels is required. MERS-CoV antigen was detected in the nasal swabs of dromedary camels using a fast immunochromatographic technique (Song *et al,* 2015). The identification of MERS-CoV nucleocapsid protein in a short time period utilising highly specific monoclonal antibodies at room temperature is the basis of this assay. The viral spike protein of MERS-CoV, which has both conserved and highly mutable or variable regions in its sequence, is the focus of most contemporary fast diagnostic assays. The accuracy of such testing is then jeopardised by mutations,

SEND REPRINT REQUEST TO SABRY M. EL-BAHR email: selbahar@kfu.edu.sa; sabryelbahr@hotmail.com

particularly in the virus's antigenic epitopes. This study presents analytical estimates of changes in epitope composition among MERS-CoV lineages.

Materials and Methods

Retrieval of input protein data and analytical programs

The sequences of the input spike protein were retrieved from the previously analysed MERS-CoV lineages (Chu *et al,* 2018; Sabir *et al,* 2016). Table 1 provides an overview of the genomes and proteins utilised in the study. CLC Genomics Workbench 12.0 (QIAGEN, Aarhus, Denmark) and Geneious prime were used to process the sequences (Kearse *et al,* 2012).

Sequence alignment and retrieval of the spike protein

The CLC genomics program's protein sequence capabilities were used to handle the spike protein sequence. With 10 gaps open and one gap extension cost, a very accurate alignment was achieved using the protein alignment wizard's alignment parameters. To compare the sequences, a pairwise comparison matrix was created. The identity matrix was created after calculating the differences, identity percentage, gaps and mutations.

Phylogenetics

The phylogenetic tree was built using the neighbour-joining (NJ) technique and then evaluated for evolutionary links. Using the default settings, the CLC genomics program was utilised. Distances were calculated using the JTT substitution model. The neighbour-joining method was tested to 100 Bootstrap resampling repeats. (Romesburg, 2004).

Spike`s antigenic epitopes

The potential antigenic epitopes in studied MERS-CoV lineages were searched at EMBOSS

Table 1. The accession numbers (no.) and description of the genomes and proteins used in this study.

Lineage	Accession no. of full genome	Accession no. of Spike	Description
5	Camel/Jeddah/Jd87/2015	ALA49781	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Jeddah/N62(b)/2014	ALA49594	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Jeddah/D38/2014	ALA49440	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Taif/T68/2015	ALA50001	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Riyadh/Ry79/2015	ALA49935	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Riyadh/Ry159/2015	ALA49847	Spike protein [Middle East respiratory syndrome-related coronavirus].
5	Camel/Riyadh/Ry136/2015	ALA49825	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/D90/2014	ALA49561	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/401/2014	ALA49374	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/Jd175/2015	ALA49803	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/S100/2014	ALA49660	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Riyadh/Ry23N/2014	ALA49352	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/S73/2014	ALA49671	Spike protein [Middle East respiratory syndrome-related coronavirus].
3	Camel/Jeddah/F13A/2014	ALA49341	Spike protein [Middle East respiratory syndrome-related coronavirus].
2	KFU-HKU/13/2013	AHX00711	Spike protein [Middle East respiratory syndrome-related coronavirus].
2	Camel/KSA/376/2013	AHY22565	Spike protein [Middle East respiratory syndrome-related coronavirus].
2	Camel/UAE/D1209/2014	AJG44124	S [Middle East respiratory syndrome-related coronavirus].
2	Camel/UAE/D1164.14/2014	AJG44091	S [Middle East respiratory syndrome-related coronavirus].
4	Camel/Jeddah/D36/2014	ALA49429	Spike protein [Middle East respiratory syndrome-related coronavirus].
4	Camel/Jeddah/D35/2014	ALA49418	Spike protein [Middle East respiratory syndrome-related coronavirus].
4	Camel/Riyadh/Ry179/2015	ALA49902	Spike protein [Middle East respiratory syndrome-related coronavirus].
4	Camel/KSA/505/2014	AHY22525	Spike protein [Middle East respiratory syndrome-related coronavirus].
4	Camel/Qatar/2/2014	AHX71946	Spike protein [Middle East respiratory syndrome-related coronavirus].
1	Camel/Jeddah-Camel-1/2013	AHE78108.1	S [Middle East respiratory syndrome-related coronavirus]
1	Taif/1/2013	AHI48672.1	S protein [Middle East respiratory syndrome-related coronavirus]
1	Wadi-Ad-Dawasir_1/2013	AHI48550	S protein [Middle East respiratory syndrome-related coronavirus]

antigenic prediction tool (Kolaskar and Tongaonkar, 1990). The antigenic determination is based on a semi-empirical approach based on physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes. The minimal length of antigenic region was set to six. The output format was set to EMBOSS motif.

Results

Antigenic epitopes

Antigenic epitope analysis requests were sent to the EMBOSS antigenic prediction tool. Based on the output epitopes, the 15 epitopes with the highest epitope score were extracted and compared among MERS-CoV lineages. Table 2 summarises the

Table 2. The antigenic epitopes predicted by EMBOSS antigenic detection in MERS-CoV lineages. The epitopes are in descending order according to the predicted score. Top 15 epitopes for each virus were selected.

Virus	#	In RBD (YES/ No)	Sequence	length	range	Conserved/ mutations no.	score
AHE78108.1 AHI48550.1 AHX71946.1 AHY22525 AHY22565.1 AJG44124.1 ALA49341.1 ALA49352.1 ALA49418.1 ALA49429.1 ALA49440.1 ALA49561.1 ALA49594.1 ALA49660.1 ALA49671.1 ALA49781.1 ALA49803.1 ALA49825.1 ALA49847.1 ALA49902.1 ALA49935.1 ALA50001.1	1	No	YIWLGFIAGLVALALCVFFILCCTGCGTN	29	1298->1326	Conserved	1.258
	2	No	NYYCLRACVSVPVSVIYD	18	647->664	Conserved	1.243
	3	No	SGFCGQGTHIVSFVVNAP	18	1114->1131	Conserved	1.216
	4	Yes	YSPCVSIVPST	11	523->533	Conserved	1.210
	5	No	ARDLICAQYVAGYKVLPPLM	20	920->939	Conserved	1.205
	6	No	YGPLQTPVGCVLGLVNSSLFVEDCKLPL GQSLCALPDTP	39	704->742	Conserved	1.204
	7	Yes	NPTCLILATVPHNLT	15	475->489	Conserved	1.203
	8	No	TTLLDLTYEMLSLQQVVKALNESYIDLK	28	1257->1284	Conserved	1.187
	9	Yes	NYNLTKLLSLFSVNDFTCSQISPAAIASN CYSSLILDYFSYPLS	44	408->451	T424I (ALA49341.1) L411F (ALA49374.1)	1.178
	10	No	SVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQT	35	4->38	Conserved	1.178
	11	No	ASQLGNCVEYSLYGVSGRG	19	597->615	Conserved	1.176
	12	No	NHTLVLLPDGCGTLLRAFYCILEP	24	166->189	Conserved	1.173
	13	No	TLNAFVAQQLVRSESAALSAQLAKD	25	1077->1101	Conserved	1.172
	14	No	SFGVTQEYIQTTIQKVTVDCKQYVCNGF	28	787->814	V810I (AJG44091.1) V810I (AHX00711.1)	1.168
	15	No	GLYFMHVGYYPSNHIEVVSAYGLCDAA	27	1133->1159	A1159S (AHY22565.1)	1.165
AHX00711.1 AJG44091.1	1	No	YIWLGFIAGLVALALCVFFILCCTGCGTN	29	1298->1326	Conserved	1.258
	2	No	NYYCLRACVSVPVSVIYD	18	647->664	Conserved	1.243
	3	No	SGFCGQGTHIVSFVVNAP	18	1114->1131	Conserved	1.216
	4	Yes	YSPCVSIVPST	11	523->533	Conserved	1.210
	5	No	ARDLICAQYVAGYKVLPPLM	20	920->939	Conserved	1.205
	6	No	YGPLQTPVGCVLGLVNSSLFVEDCKLPLGQSLCAL PDTP	39	704->742	Conserved	1.204
	7	Yes	NPTCLILATVPHNLT	15	475->489	Conserved	1.203
	8	No	TTLLDLTYEMLSLQQVVKALNESYIDLK	28	1257->1284	Conserved	1.187
	9	Yes	NYNLTKLLSLFSVNDFTCSQISPAAIASN CYSSLILDYFSYPLS	44	408->451	T424I (ALA49341.1) L411F (ALA49374.1)	1.178
	10	No	SVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQT	35	4->38	Conserved	1.178
	11	No	ASQLGNCVEYSLYGVSGRG	19	597->615	Conserved	1.176
	12	No	NHTLVLLPDGCGTLLRAFYCILEP	24	166->189	Conserved	1.173
	13	No	TLNAFVAQQLVRSESAALSAQLAKD	25	1077->1101	Conserved	1.172
	14	No	GLYFMHVGYYPSNHIEVVSAYGLCDAA	27	1133->1159	A1159S (AHY22565.1)	1.165
	15	No	HATLFGSVACEHI	13	670->682	Conserved	1.155

Table 2 continued....

ALA49374.1	1	No	YIWLGFIAGLVALALCVFFILCCTGCGTN	29	1298->1326	Conserved	1.258
	2	No	NYYCLRACVSVPVSVIYD	18	647->664	Conserved	1.243
	3	No	SGFCGQGTHIVSFVVNAP	18	1114->1131	Conserved	1.216
	4	Yes	YSPCVSIVPST	11	523->533	Conserved	1.210
	5	No	ARDLICAQYVAGYKVLPPLM	20	920->939	Conserved	1.205
	6	No	YGPLQTPVGCVLGLVNSSLFVEDCKLPLGQSLCA LPDTP	39	704->742	Conserved	1.204
	7	Yes	NPTCLILATVPHNLT	15	475->489	Conserved	1.203
	8	No	TTLLDLTYEMLSLQQVVKALNESYIDLK	28	1257->1284	Conserved	1.187
	9	Yes	TKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLS	40	412->451	T424I (ALA49341.1)	1.178
	10	No	SVFLLMFLLTPTESYVDVGPDSVKSACIEVDIQQT	35	4->38	Conserved	1.178
	11	No	ASQLGNCVEYSLYGVSGRG	19	597->615	Conserved	1.176
	12	No	NHTLVLLPDGCGTLLRAFYCILEP	24	166->189	Conserved	1.173
	13	No	TLNAFVAQQLVRSESAALSAQLAKD	25	1077->1101	Conserved	1.172
	14	No	SFGVTQEYIQTTIQKVTVDCKQYVCNGF	28	787->814	V810I (AJG44091.1) V810I (AHX00711.1)	1.168
	15	No	GLYFMHVGYYPSNHIEVVSAYGLCDAA	27	1133->1159	A1159S (AHY22565.1)	1.165

epitopes sequences, their respective location on either RBD or other S protein domains, the range of amino acid, conservation among lineages and the predicted mutations. The epitopes were arranged in descending order, according to their antigenic scores.

Spike alignment

The spike protein sequences of various MERS-CoV lineages were aligned (Fig 1). The alignment showed a mostly conserved amino acid sequence with few mutations, which are summarised in table 3. These mutations were related to lineage 2 and 3. However, these mutations were few and involved only single amino acid. The current findings indicated that there were no differences observed in length and range of epitopes for each virus among all studied lineages (Table 2). Most of epitopes sequences were conserved. However, few sequences underwent little mutation involving single amino acids (Table 2). Except for AHX00711.1 and AJG44091.1, single amino acid mutations were observed at epitope no. 9, 14 and 15 (Table 2). At epitope 9, these mutations were T424I (ALA49341.1) and L411F (ALA49374.1) (Table 2). At epitope 14, these mutations were V810I (AJG44091.1) and V810I (AHX00711.1) (Table 2). At epitope 15, this mutation was A1159S (AHY22565.1) (Table 2). For AHX00711.1 and AJG44091.1, the single amino acid mutations were at epitopes no. 9 and 14 (Table 2). At epitope 9, these mutations were T424I (ALA49341.1) and L411F (ALA49374.1). At epitope 14, this mutation was A1159S (AHY22565.1) (Table 2). Data summarised in table 3 indicated that 20% of epitopes sequences are located at the receptor-binding domain (RBD) and 80% of these sequences are located away from the RBD. This finding suggests the reliability of using a rapid test based on spike protein epitopes in dromedary camels by targeting multiple domains epitopes on the surface of MERS-CoV spike.

Table 3. List of observed mutations in virus spike in the studied lineages.

Mutation	Present in all lineages (Yes/No)	lineage	Accession number of mutants
T424I	No	3	ALA49341.1
L411F	Yes	3	ALA49374.1
V810I	Yes	2	AJG44091.1
V810I	Yes	2	AHX00711.1
A1159S	Yes	2	AHY22565.1

Pairwise comparison matrix

Pairwise comparison matrix revealed the absence of gaps in the alignment that indicates the lack of amino acid insertions or deletions in the spike protein among the lineages (Fig 2). In addition, the number of differences among viruses was a maximum of 8. The highest difference was 8 and it was among virus no 25 when matched with virus no 23 and 22 (Fig 2). Fortunately, most of these mutations are not involved in the antigenic epitopes. Pairwise comparison matrix revealed a higher identity of the spike protein among the lineages which ranged from 99.5 -100% (Fig 3).

Phylogenetic analysis

The spike protein of MERS-CoV was clustered in 3 main lineages (Fig 4). This indicates that the virus diversity is less common in virus spike, compared with the full genome.

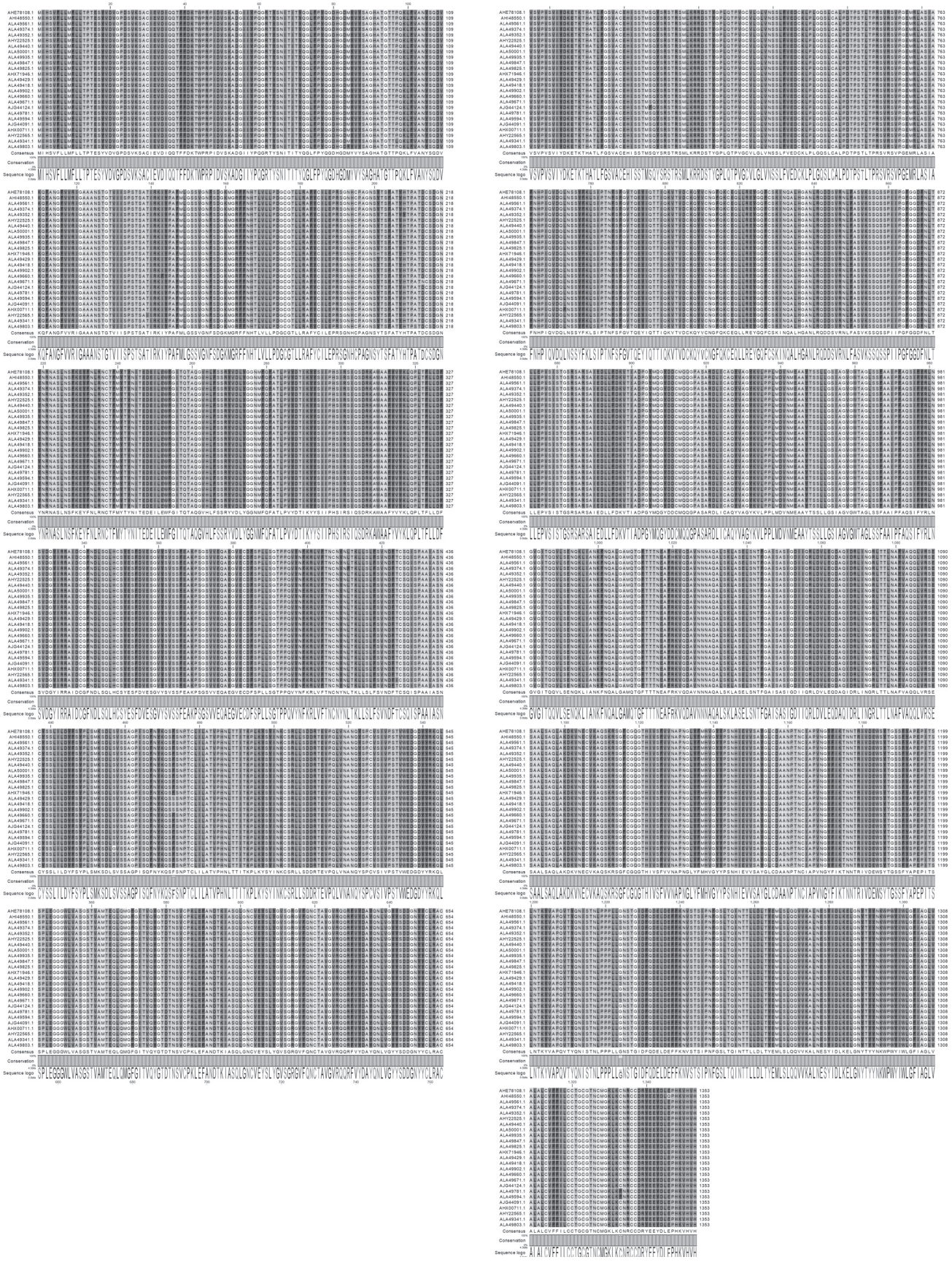

Fig 1. Sequence alignment of spike protein from different MERS-CoV lineages.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AHE78108.1	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AHI48550.1	2	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49561.1	3	2	3		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49374.1	4	3	4	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49352.1	5	2	3	2	3		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AHY22525.1	6	1	2	1	2	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49440.1	7	1	2	1	2	1	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA50001.1	8	1	2	1	2	1	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49935.1	9	1	2	1	2	1	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49847.1	10	1	2	1	2	1	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49825.1	11	1	2	1	2	1	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0
AHX71946.1	12	2	3	2	3	2	1	1	1	1	1	1		0	0	0	0	0	0	0	0	0	0	0	0	0
ALA49429.1	13	3	4	1	2	3	2	2	2	2	2	2	3		0	0	0	0	0	0	0	0	0	0	0	0
ALA49418.1	14	3	4	1	2	3	2	2	2	2	2	2	3	0		0	0	0	0	0	0	0	0	0	0	0
ALA49902.1	15	2	3	2	3	2	1	1	1	1	1	1	2	1	1		0	0	0	0	0	0	0	0	0	0
ALA49660.1	16	4	5	2	3	4	3	3	3	3	3	3	4	3	3	4		0	0	0	0	0	0	0	0	0
ALA49671.1	17	4	5	2	3	4	3	3	3	3	3	3	4	3	3	4	2		0	0	0	0	0	0	0	0
AJG44124.1	18	3	4	3	4	3	2	2	2	2	2	2	3	4	4	3	5	5		0	0	0	0	0	0	0
ALA49781.1	19	2	3	2	3	2	1	1	1	1	1	1	2	3	3	2	4	4	3		0	0	0	0	0	0
ALA49594.1	20	2	3	2	3	2	1	1	1	1	1	1	2	3	3	2	4	4	3	0		0	0	0	0	0
AJG44091.1	21	4	5	4	5	4	3	3	3	3	3	3	4	5	5	4	6	6	3	4	4		0	0	0	0
AHX00711.1	22	5	6	5	6	5	4	4	4	4	4	4	5	6	6	5	7	7	4	5	5	3		0	0	0
AHY22565.1	23	5	6	5	6	5	4	4	4	4	4	4	5	6	6	5	7	7	4	5	5	5	6		0	0
ALA49341.1	24	3	4	3	4	3	2	2	2	2	2	2	3	4	4	3	5	5	4	3	3	5	6	6		0
ALA49803.1	25	5	6	3	4	5	4	4	4	4	4	4	5	4	4	5	5	5	6	5	5	7	8	8	6	

Fig 2. Comparative pairwise matrix for sequences of spike protein from different MERS-CoV lineages. The upper diagonal panel is the number of gaps. The lower diagonal panel is the number of amino acid differences.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
AHE78108.1	1		7.3939E-4	0.0015	0.0022	0.0015	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	0.0015	0.0022	0.0022	0.0015	0.0030	0.0030	0.0022	0.0015	0.0015	0.0030	0.0037	0.0037	0.0022	0.0037
AHI48550.1	2	99.9261		0.0022	0.0030	0.0022	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.0022	0.0030	0.0030	0.0022	0.0037	0.0037	0.0030	0.0022	0.0022	0.0037	0.0044	0.0044	0.0030	0.0044
ALA49561.1	3	99.8522	99.7783		7.3939E-4	0.0015	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	0.0015	7.3939E-4	7.3939E-4	0.0015	0.0015	0.0015	0.0022	0.0015	0.0015	0.0030	0.0037	0.0037	0.0022	0.0022
ALA49374.1	4	99.7783	99.7044	99.9261		0.0022	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.0022	0.0015	0.0015	0.0022	0.0022	0.0022	0.0030	0.0022	0.0022	0.0037	0.0044	0.0044	0.0030	0.0030
ALA49352.1	5	99.8522	99.7783	99.8522	99.7783		7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	7.3939E-4	0.0015	0.0022	0.0022	0.0015	0.0030	0.0030	0.0022	0.0015	0.0015	0.0030	0.0037	0.0037	0.0022	0.0037
AHY22525.1	6	99.9261	99.8522	99.9261	99.8522	99.9261		1.0547E-16	1.0547E-16	1.0547E-16	1.0547E-16	1.0547E-16	7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
ALA49440.1	7	99.9261	99.8522	99.9261	99.8522	99.9261	100.0000		1.0547E-16	1.0547E-16	1.0547E-16	1.0547E-16	7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
ALA50001.1	8	99.9261	99.8522	99.9261	99.8522	99.9261	100.0000	100.0000		1.0547E-16	1.0547E-16	1.0547E-16	7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
ALA49935.1	9	99.9261	99.8522	99.9261	99.8522	99.9261	100.0000	100.0000	100.0000		1.0547E-16	1.0547E-16	7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
ALA49847.1	10	99.9261	99.8522	99.9261	99.8522	99.9261	100.0000	100.0000	100.0000	100.0000		1.0547E-16	7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
ALA49825.1	11	99.9261	99.8522	99.9261	99.8522	99.9261	100.0000	100.0000	100.0000	100.0000	100.0000		7.3939E-4	0.0015	0.0015	7.3939E-4	0.0022	0.0022	0.0015	7.3939E-4	7.3939E-4	0.0022	0.0030	0.0030	0.0015	0.0030
AHX71946.1	12	99.8522	99.7783	99.8522	99.7783	99.8522	99.9261	99.9261	99.9261	99.9261	99.9261	99.9261		0.0022	0.0022	0.0015	0.0030	0.0030	0.0022	0.0015	0.0015	0.0030	0.0037	0.0037	0.0022	0.0037
ALA49429.1	13	99.7783	99.7044	99.9261	99.8522	99.7783	99.8522	99.8522	99.8522	99.8522	99.8522	99.8522	99.7783		1.0547E-16	7.3939E-4	0.0022	0.0022	0.0030	0.0022	0.0022	0.0037	0.0044	0.0044	0.0030	0.0030
ALA49418.1	14	99.7783	99.7044	99.9261	99.8522	99.7783	99.8522	99.8522	99.8522	99.8522	99.8522	99.8522	99.7783	100.0000		7.3939E-4	0.0022	0.0022	0.0030	0.0022	0.0022	0.0037	0.0044	0.0044	0.0030	0.0030
ALA49902.1	15	99.8522	99.7783	99.8522	99.7783	99.8522	99.9261	99.9261	99.9261	99.9261	99.9261	99.9261	99.8522	99.9261	99.9261		0.0030	0.0030	0.0022	0.0015	0.0015	0.0030	0.0037	0.0037	0.0022	0.0037
ALA49660.1	16	99.7044	99.6305	99.8522	99.7783	99.7044	99.7783	99.7783	99.7783	99.7783	99.7783	99.7783	99.7044	99.7783	99.7783	99.7044		0.0015	0.0037	0.0030	0.0030	0.0044	0.0052	0.0052	0.0037	0.0037
ALA49671.1	17	99.7044	99.6305	99.8522	99.7783	99.7044	99.7783	99.7783	99.7783	99.7783	99.7783	99.7783	99.7044	99.7783	99.7783	99.7044	99.8522		0.0037	0.0030	0.0030	0.0044	0.0052	0.0052	0.0037	0.0037
AJG44124.1	18	99.7783	99.7044	99.7783	99.7044	99.7783	99.8522	99.8522	99.8522	99.8522	99.8522	99.8522	99.7783	99.7044	99.7044	99.7783	99.6305	99.6305		0.0022	0.0022	0.0022	0.0030	0.0030	0.0030	0.0044
ALA49781.1	19	99.8522	99.7783	99.8522	99.7783	99.8522	99.9261	99.9261	99.9261	99.9261	99.9261	99.9261	99.8522	99.7783	99.7783	99.8522	99.7044	99.7044	99.7783		1.0547E-16	0.0030	0.0037	0.0037	0.0022	0.0037
ALA49594.1	20	99.8522	99.7783	99.8522	99.7783	99.8522	99.9261	99.9261	99.9261	99.9261	99.9261	99.9261	99.8522	99.7783	99.7783	99.8522	99.7044	99.7044	99.7783	100.0000		0.0030	0.0037	0.0037	0.0022	0.0037
AJG44091.1	21	99.7044	99.6305	99.7044	99.6305	99.7044	99.7783	99.7783	99.7783	99.7783	99.7783	99.7783	99.7044	99.6305	99.6305	99.7044	99.5565	99.5565	99.7783	99.7044	99.7044		0.0022	0.0037	0.0037	0.0052
AHX00711.1	22	99.6305	99.5565	99.6305	99.5565	99.6305	99.7044	99.7044	99.7044	99.7044	99.7044	99.7044	99.6305	99.5565	99.5565	99.6305	99.4826	99.4826	99.7044	99.6305	99.6305	99.7783		0.0044	0.0044	0.0059
AHY22565.1	23	99.6305	99.5565	99.6305	99.5565	99.6305	99.7044	99.7044	99.7044	99.7044	99.7044	99.7044	99.6305	99.5565	99.5565	99.6305	99.4826	99.4826	99.7044	99.6305	99.6305	99.6305	99.5565		0.0044	0.0059
ALA49341.1	24	99.7783	99.7044	99.7783	99.7044	99.7783	99.8522	99.8522	99.8522	99.8522	99.8522	99.8522	99.7783	99.7044	99.7044	99.7783	99.6305	99.6305	99.7044	99.7783	99.7783	99.6305	99.5565	99.5565		0.0044
ALA49803.1	25	99.6305	99.5565	99.7783	99.7044	99.6305	99.7044	99.7044	99.7044	99.7044	99.7044	99.7044	99.6305	99.7044	99.7044	99.6305	99.6305	99.6305	99.5565	99.6305	99.6305	99.4826	99.4087	99.4087	99.5565	

Fig 3. Comparative pairwise matrix for sequences of spike protein from different MERS-CoV lineages. The upper diagonal panel is the distance. The lower diagonal panel is the identity per cent.

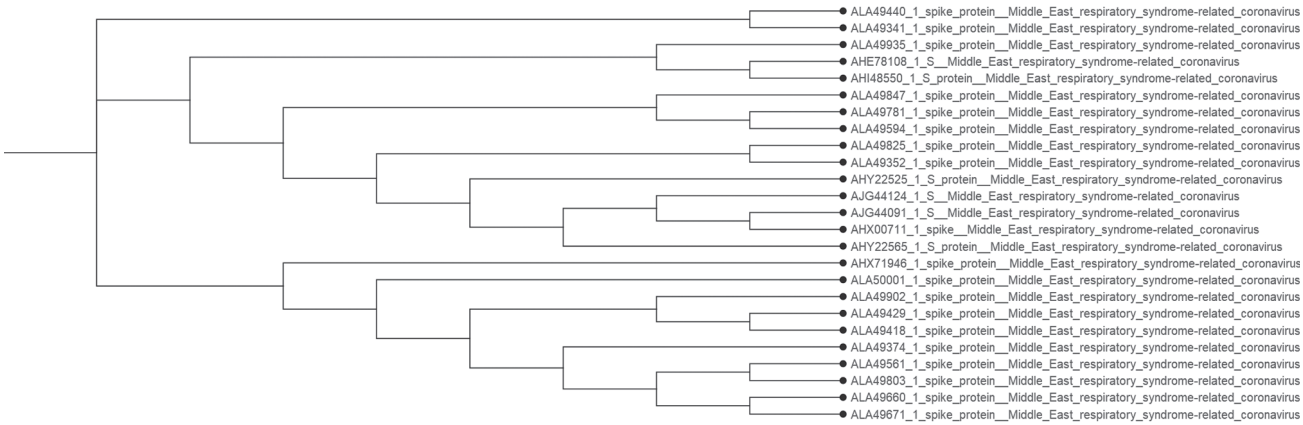

Fig 4. Cladogram of spike protein from different MERS-CoV lineages.

Discussion

MERS-CoVs have been sequenced and classified in both humans and camels since the first human MERS-CoV was discovered (Cotten *et al,* 2014; Lau *et al,* 2016; Sabir *et al,* 2016). Sabir *et al* (2016) sequenced and analysed typical MERS-CoV genomes from Saudi Arabia, including 67 from camels and discovered different clade B MERS-CoV lineages. Co-

infections with MERS-CoV and other coronaviruses were prevalent, as were additional non-MERS-CoVs. Another study sequenced 10 different camel MERS-CoVs from Abu Dhabi and discovered viruses in many clade B lineages, including a 6th lineage and a camel MERS-CoV within clade A. The clade A MERS-CoVs are assumed to be older and not circulating today, therefore, this study provides valuable insight into camel MERS-CoV variety in nature (Lau *et al,* 2016). Additional efforts to sequence MERS-CoV from camels have focused on nucleocapsid and spike genes as possible locations for finding genetic diversity (van Doremalen *et al,* 2017). Several human coronaviruses are known to have originated from zoonotic sources (Millet *et al,* 2016; Muhairi *et al,* 2016). At a live animal market in the Emirate of Abu Dhabi's eastern region, 376 camels were checked for MERS-Cov (Yusof *et al,* 2017). From 139 samples, 126 whole genomes and 3 nearly complete genomes were recovered. Five of the remaining 10 samples yielded spike gene sequences. Within clade B, the genomes of camel MERS-CoV represented 3 recognised and 2 potentially novel lineages. Camel and human MERS-CoV sequences are jumbled together within lineages. In the camel samples, the researchers discovered 10 recombination events. The junctions between ORF1b and S were the most common recombination breakpoints. MERS-CoV infection in humans may have resulted from the continual transfer of various MERS-CoV lineages from camels, according to evidence. The camel MERS-CoV genomes sequenced in this investigation support this idea (Yusof *et al,* 2017).

The current findings indicated that there were no differences observed in length and range of epitopes for each virus among all studied lineages. Most of epitopes sequences were conserved. However, few sequences underwent little mutation involving single amino acids. The top 15 epitopes in dromedary camels have a conserved sequence, indicating that a quick test based on spike protein epitopes is reliable. According to the present data, 20% of epitope sequences are found near the RBD, whereas the remaining 80% are found distant from the RBD. In dromedary camels, this research proved the reliability of the previously mentioned rapid test based on spike protein epitopes. In the current study, the alignment of spike protein sequences from several MERS-CoV lineages indicated a small number of alterations. Lineages 2 and 3 were affected by these changes. Using rapid testing for MERS-CoV detection based on epitopes from lineages 1, 4, and 5 seems to be effective, based on the current findings. Furthermore, given the changes in lineages 2 and 3 were rare and involved just a single amino acid, these quick tests may be successful when employing any epitope from these lineages.

The current findings revealed the absence of gaps in the alignment that indicates the lack of amino acid insertions or deletions in the spike protein among the lineages. This finding confirmed the effectiveness of any of studied epitopes for rapid test designing in dromedary camels. The current finding indicated that, the number of difference among viruses accession number was a maximum of 8, most of them were not involved in the antigenic epitopes. This finding confirmed the suitability of any of the studied epitopes for production of rapid tests of MERS CoV in dromedary camels. The current findings indicated a higher identity of the spike protein among the lineages which ranged from 99.5 -100%. This finding provides new evidence of the effectiveness of any of the studied epitopes for the production of rapid tests of MERS CoV in dromedary camels as illustrated above.

Because identifying antigens and epitopes using an experimental method may be difficult, expensive, and time-consuming, using an *in-silico* strategy to uncover novel epitopes has become the preferred alternative. This approach is recognised as one of the most effective in identifying antigens because it screens the whole microbial proteome using a variety of prediction algorithms. In conclusion, our comprehensive technique encompasses antigenic epitope screening both horizontally (through the whole coding areas of the Spike) and vertically (across various MERS-CoV lineages). Mutations in the targeted epitopes might have an influence on the diagnostics' population coverage and efficiency. After examining current MERS-CoV mutants across lineages, we evaluated for immunogenicity conservation in the chosen epitopes to provide more diagnostic options. The efficiency of diagnostics based on the provided configurations is expected to be high due to the observed low mutagenicity rate of MERS-CoV. Combining the present findings with experimental confirmation is a pre requisite for successful diagnostics.

Acknowledgements

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia (Project # 171002; Grant 271).

References

Abd El Wahed A, Patel P, Heidenreich D, Hufert FT and Weidmann M. Reverse transcription recombinase polymerase amplification assay for the detection of Middle East respiratory syndrome coronavirus. PLOS Currents. 2013; 12;5:ecurrents.DOI:10.1371/currents.outbreaks.

Alexandersen S1, Kobinger GP, Soule G and Wernery U. Middle East respiratory syndrome coronavirus antibody reactors among camels in Dubai, United Arab Emirates, in 2005. Transboundary and Emerging Diseases. 2014; 61:105-108.

Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE, Badu EK, Anti P, Agbenyega O, Meyer B, Oppong S, Sarkodie YA, Kalko EK, Lina PH, Godlevska EV, Reusken C, Seebens A, Gloza-Rausch F, Vallo P, Tschapka M, Drosten C and Drexler JF. Human betacoronavirus 2c EMC/2012-related viruses in bats, Ghana and Europe. Emerging Infectious Diseases. 2013; 19:456-459.

Buchholz U, Muller MA, Nitsche A, Sanewski A, Wevering N, Bauer- Balci T, Bonin F, Drosten C, Schweiger B, Wolff T, Muth D, Meyer B, Buda S, Krause G, Schaade L and Haas W. Contact investigation of a case of human novel coronavirus infection treated in a German hospital, October-November 2012. Eurosurveillance. 2013; 18(8):20406.

Chu DKW, Hui KPY, Perera R, Miguel E, Niemeyer D, Zhao J, Channappanavar R, Dudas G, Oladipo JO, Traoré A, Fassi-Fihri O, Ali A, Demissié GF, Muth D, Chan MCW, Nicholls JM, Meyerholz DK, Kuranga SA, Mamo G, Zhou Z, So RTY, Hemida MG, Webby RJ, Roger F, Rambaut A, Poon LLM, Perlman S, Drosten C, Chevalier V and Peiris M. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. Proceedings of the National Academy of Sciences. 2018; 115:3144-3149.

Corman V, Eckerle I, Bleicker T, Zaki A, Landt O, Eschbach-Bludau M, van Boheemen S, Gopal R, Ballhause M, Bestebroer T, Muth D, Muller M, Drexler J, Zambon M, Osterhaus A, Fouchier R and Drosten C. Detection of a novel human coronavirus by real-time reversetranscription polymerase chain reaction. Eurosurveillance. 2012a; 17:pii20285.

Corman VM, Muller MA, Costabel U, Timm J, Binger T, Meyer B, Kreher P, Lattwein E, Eschbach-Bludau M, Nitsche A, Bleicker T, Landt O, Schweiger B, Drexler JF, Osterhaus AD, Haagmans BL, Dittmer U, Bonin F, Wolff T and Drosten C. Assays for laboratory confirmation of novel human coronavirus (hCoV-EMC) infections. Eurosurveillance. 2012b; 17(49):20334.

Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palser AL, Ong SH, Al Rabeeah AA, Alhakeem RF, Assiri A, Al-Tawfiq JA, Albarrak A, Barry M, Shibl A, Alrabiah FA, Hajjar S, Balkhy HH, Flemban H, Rambaut A, Kellam P and Memish ZA. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. mBio. 2014; 5(1):e01062-13. DOI:10.1128/mBio.01062-13.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P and Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organisation and analysis of sequence data. Bioinformatics. 2012; 28:1647-1649.

Kolaskar AS and Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Letters. 1990; 276:172-174.

Lau SK, Wernery R, Wong EY, Joseph S, Tsang AK, Patteril NA, Elizabeth SK, Chan KH, Muhammed R, Kinne J, Yuen KY, Wernery U and Woo PC. Polyphyletic origin of MERS coronaviruses and isolation of a novel clade A strain from dromedary camels in the United Arab Emirates. Emerging Microbes and Infections. 2016; 5(1):1-9.

Meyer B, Müller MA, Corman VM, Reusken CBEM, Ritz D, Godeke GJ, Lattwein E, Kallies S, Siemens A, van Beek J, Drexler JF, Muth D, Bosch B-J, Wernery U, Koopmans MPG, Wernery R and Drosten C. Antibodies against MERS coronavirus in dromedary camels, United Arab Emirates, 2003 and 2013. Emerging Infectious Diseases. 2003; 20:552-559. DOI:10.3201/eid2004.131746.

Millet JK, Goldstein ME, Labitt RN, Hsu HL, Daniel S and Whittaker GR. A camel-derived MERS-CoV with a variant spike protein cleavage site and distinct fusion activation properties. Emerging Microbes and Infections. 2016; 5(1):1-9.

Muhairi SA, Hosani FA, Eltahir YM, Mulla MA, Yusof MF, Serhan WS, Hashem FM, Elsayed EA, Marzoug BA and Abdelazim AS. Epidemiological investigation of Middle East respiratory syndrome coronavirus in dromedary camel farms linked with human infection in Abu Dhabi Emirate, United Arab Emirates. Virus Genes. 2016; 52: 848-854. DOI:10.1007/s11262-016-1367-1.

Perera RA, Wang P, Gomaa MR, El-Shesheny R, Kandeil A, Bagato O, Siu LY, Shehata MM, Kayed AS, Moatasim Y, Li M, Poon LL, Guan Y, Webby RJ, Ali MA, Peiris JS and Kayali G. Seroepidemiology for MERS coronavirus using microneutralisation and pseudoparticle virus neutralisation assays reveal a high prevalence of antibody in dromedary camels in Egypt. Eurosurveillance. 2013; 5;18(36):pii20574. DOI:10.2807/1560-7917.es2013.18.36.20574.

Reusken C, Mou H, Godeke GJ, van der Hoek L, Meyer B, Muller MA, Haagmans B, de Sousa R, Schuurman N, Dittmer U, Rottier P, Osterhaus A, Drosten C, Bosch BJ and Koopmans M. Specific serology for emerging human coronaviruses by protein microarray. Eurosurveillance. 2013b; 18(14):20441. DOI: 10.2807/1560-7917.es2013.18.14.20441.

Reusken CB, Haagmans BL, Muller MA, Gutierrez C, Godeke GJ, Meyer B, Muth D, Raj VS, Smits-De Vries L, Corman VM, Drexler J-F, Smits SL, El Tahir YE, De Sousa R, van Beek J, Nowotny N, van Maanen K, Hidalgo-Hermoso E, Bosch B-J, Rottier P, Osterhaus A, Gortazar-Schmidt C, Drosten C and Koopmans MPG. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. The Lancet Infectious Diseases. 2013a; 13:859-866. DOI:10.1016/S1473-3099(13)70164-6.

Romesburg C. Cluster analysis for researchers. Lulu. com. (2004).

Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, Abo-Aba SE, Qureshi MI, Abu-Zeid M, Zhang Y, Khiyami MA, Alharbi NS, Hajrah NH, Sabir MJ, Mutwakil MH, Kabli SA, Alsulaimany FA, Obaid AY, Zhou B, Smith DK, Holmes EC, Zhu H and Guan Y. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. Science. 2016; 351:81-84.

Shirato K, Yano T, Senba S, Akachi S, Kobayashi T, Nishinaka T, Notomi T and Matsuyama S. Detection of Middle East respiratory syndrome coronavirus using reverse transcription loop-mediated isothermal amplification (RT-LAMP). Virology Journal. 2014; 11(1):1-11.

Song D, Ha G, Serhan W, Eltahir Y, Yusof M, Hashem F, Elsayed E, Marzoug B, Abdelazim A and Al Muhairi S. Development and validation of a rapid immunochromatographic assay for detection of Middle East respiratory syndrome coronavirus antigen in dromedary camels. Journal of Clinical Microbiology. 2015; (4):1178-82.

van Doremalen N, Hijazeen ZS, Holloway P, Al Omari B, McDowell C, Adney D, Talafha HA, Guitian J, Steel J, Amarin N, Tibbo M, Abu-Basha E, Al-Majali AM, Munster VJ and Richt JA. High prevalence of middle east respiratory coronavirus in young dromedary camels in Jordan. Vector Borne Zoonotic Diseases. 2017; 17:155-159.

Yusof MF, Queen K, Eltahir YM, Paden CR, Al Hammadi ZMAH, Tao Y, Li Y, Khalafalla AI, Shi M, Zhang J, Mohamed MSAE, Abd Elaal Ahmed MH, Azeez IA, Bensalah OK, Eldahab ZS, Al Hosani FI, Gerber SI, Hall AJ, Tong S and Al Muhairi SS. Diversity of Middle East respiratory syndrome coronaviruses in 109 dromedary camels based on full-genome sequencing, Abu Dhabi, United Arab Emirates. Emerging Microbes and Infections. 2017; 6(11): 101-111.

Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD and Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. The New England Journal of Medicine. 2012; 367:1814-1820.