



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Genetic diversity of MERS-CoV spike protein gene in Saudi Arabia

Sayed S. Sohrab^{a,b,*}, Esam I. Azhar^{a,b}

^a Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

^b Medical Laboratory Technology Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

ARTICLE INFO

Article history:

Received 10 September 2019

Received in revised form 2 November 2019

Accepted 17 November 2019

Keywords:

Genetic diversity

MERS-CoV

Human

Camel

Saudi Arabia

ABSTRACT

Background: Middle East respiratory syndrome coronavirus (MERS-CoV) was primarily detected in 2012 and still causing disease in human and camel. Camel and bats have been identified as a potential source of virus for disease spread to human. Although, significant information related to MERS-CoV disease, spread, infection, epidemiology, clinical features have been published, A little information is available on the sequence diversity of Spike protein gene. The Spike protein gene plays a significant role in virus attachment to host cells. Recently, the information about recombinant MERS-CoV has been published. So, this work was designed to identify the emergence of any another recombinant virus in Jeddah, Saudi Arabia.

Methods: In this study samples were collected from both human and camels and the Spike protein gene was amplified and sequenced. The nucleotide and amino acid sequences of MERS-CoV Spike protein gene were used to analyze the recombination, genetic diversity and phylogenetic relationship with selected sequences from Saudi Arabia.

Results: The nucleotide sequence identity ranged from 65.7% to 99.8% among all the samples collected from human and camels from various locations in the Kingdom. The lowest similarity (65.7%) was observed in samples from Madinah and Dammam. The phylogenetic relationship formed different clusters with multiple isolates from various locations. The sample collected from human in Jeddah hospital formed a closed cluster with human samples collected from Buraydah, while camel sample formed a closed cluster with Hufuf isolates. The phylogenetic tree by using Aminoacid sequences formed closed cluster with Dammam, Makkah and Duba isolates. The amino acid sequences variations were observed in 28/35 samples and two unique amino acid sequences variations were observed in all samples analyzed while total 19 nucleotides sequences variations were observed in the Spike protein gene. The minor recombination events were identified in eight different sequences at various hotspots in both human and camel samples using recombination detection programme.

Conclusion: The generated information from this study is very valuable and it will be used to design and develop therapeutic compounds and vaccine to control the MERS-CoV disease spread in not only in the Kingdom but also globally.

© 2019 The Authors. Published by Elsevier Ltd on behalf of King Saud Bin Abdulaziz University for Health Sciences. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Coronaviruses (CoVs), belongs to the family *Coronaviridae*. They have single stranded positive-sense RNA genome with ~26–32 kb in size [1,2]. The genomic organization and expression pattern are similar in all coronaviruses. It is well known that multiple CoVs are found naturally and their genetic recombination hap-

pens, and novel viruses emerged with variable characteristics with their extended hosts [3]. In 2012, MERS-CoV was identified from a 60-year-old patient, died with severe pneumonia, hospitalized in Jeddah, Saudi Arabia [4]. MERS-CoV has become a deadly pathogen for the occurrence of an outbreak and causes zoonotic respiratory disease. The MERS-CoV has spread into 27 countries and has now become a major health concern for both human and animal globally. The bats and dromedary camels serve as a reservoir and potential source of MERS-CoV for human infections. Currently, total of 2458 laboratory-confirmed cases with 848 deaths have been reported [5]. The MERS-CoV mainly transmitted to humans in family clusters, healthcare workers, household contacts, through

* Corresponding author at: Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Post Box No: 80216, Jeddah, 21589, Saudi Arabia.

E-mail address: ssohrab@kau.edu.sa (S.S. Sohrab).

contact with camels, as well as community settings [6–8]. MERS-CoV is responsible for causing lower respiratory infections with fever and cough followed by shortness of breath and organ failure in severe cases with comorbidities [9–11]. Earlier it was believed that inter-human transmission is limited but the nosocomial infection was reported in healthcare facilities due to inadequate infection control which resulted in larger outbreaks with MERS-CoV confirming human to human infection [12]. MERS-CoV has been isolated and sequenced from camel and its infected owner patient as well as in air sample collected from the same barn that sheltered the infected camel and showed the identical sequences from both camel and human which confirms the direct transmission from camel to human [7,13]. MERS-CoV has been detected in upper and lower respiratory secretions at relatively high virus load and in fecal samples [14]. The identification of MERS-CoV and the neutralizing antibodies as well as a similar coronavirus has already been identified from camels and bats from multiple locations; Ghana, Europe, South Africa, Oman, the Canary Islands, UAE, Korea, and Egypt [12,14,15]. Recently, a mutant MERS-CoV has been identified from South Korea. The mutation was observed in Ribosomal Binding Domain (RBD) domain of Spike protein gene [16,17]. Additionally, the MERS-CoV neutralizing antibodies were detected recently in the young goats and sheep from Jordan and Egypt [15,18]. Recently, a large outbreak observed due to the unusual presentation of MERS-CoV from Riyadh, Saudi Arabia [19].

The CoVs have high sequence variation which favors the possible recombination, mutations, and emergence of a novel and recombinant virus with variable characters and extended hosts. The emergence of dominant and recombinant MERS-CoV which caused a human outbreak in 2015 has been reported from camel [3]. The zoonotic introduction was suspected after the MERS-CoV identification [18] and in a recent study from UAE the genetic diversity of MERS-CoV full-genome from both human and dromedary camel was analyzed, and very closed sequence similarity was observed which confirms the zoonotic introductions [20,21]. Additionally, the zoonotic introduction time and seasons play an important role in the disease spread. Based on the analysis of the distribution of human outbreak cluster size, it has also been demonstrated that the time of zoonotic introduction and season plays a significant role in human outbreak driven by MERS-CoV in Arabian Peninsula [22]. The MERS-CoV is known to have genetic diversity. The Spike gene plays a significant role in host cell attachment and the entry of the virus in the host cells [23]. The RBD of Spike protein gene mediates the virus interaction to the host cell and binds with dipeptidyl peptidase 4 (DPP4, CD26) known as a cellular receptor which favors the viral entry into the cell and is immunodominant and induces neutralizing antibodies [24–26].

Based on the literature, it is well known that MERS-CoV Spike protein gene has significant genetic variability isolated from both human and camel. In South Korea, a recent outbreak occurred with a high fatality rate. The Spike gene diversity was identified in many samples and showed interspecific variation with MERS-CoV isolates from South Korea [27]. A novel recombinant MERS-CoV has already been identified from Saudi Arabia [28]. Recently, in another study, total 530 nucleotides deletion was observed in Spike gene from serum samples collected from Taif, Saudi Arabia and a novel genetic variant of MERS-CoV was designated as a quasispecies [29]. Multiple substitutions of amino acids were observed in RBD, part of Spike gene from a bat sample collected from Uganda and the recombination in the S1 subunit of the Spike gene was observed and it was expected that this variation likely to play an important role in the emergence of MERS-CoV causing disease in human [30]. Recently, the MERS-CoV has been genetically and phenotypically characterized from Africa [31] and South Korea [32].

The nucleotide substitutions/ amino acid variation of Spike gene has significantly affected the virus transmission, disease spread to

extended hosts and their evolution in different geographical locations. Based on the published information, there is a lack of detail information about Spike gene sequence variability of MERS-CoV from Saudi Arabia infecting human and camel. So, there is an urgent need to identify the genetic variability of Spike protein gene so that it can determine and established the link that how the virus is moving from infected camel to human. Based on the above information, this study was conducted. The detection of MERS-CoV in human and camel determining the genetic diversity among Spike gene will further help researchers as well as health authority to design and develop an effective disease management and control strategies in the Kingdom of Saudi Arabia.

Materials and methods

Sample collection and cell culture

The MERS-CoV samples were collected from both human and camel and stored at Special Infectious Agents Unit (SIAU), King Fahd Medical Research Centre (KFMRC), King Abdulaziz University (KAU), Jeddah, Saudi Arabia. The samples (blood and nasal swabs) were isolated from the six patients just after one day after admission into the hospital. All nasal swabs were properly collected and maintained by immersing into viral transport medium in a cold container. All the collected samples were stored for further analysis after proper processing at BSL-3 lab in SIAU, KFMRC, KAU, Jeddah, Saudi Arabia. The freshly prepared Vero cells (ATCCCL-81) were used to inoculate by using 100 µl nasal swab and maintained in complete DMEM following the described protocol [7]. The inoculated cells were further for virus infection and development of cytopathic effect by incubating at 37 °C with 5% CO₂. After complete cytopathic effect, the supernatant from cell-culture was collected.

Primer designing, virus detection and amplification of the spike protein gene

The complete sequences of MERS-CoV were retrieved from National Center for Biotechnology Information (NCBI/PubMed) database. The retrieved sequences were aligned by using BioEdit 7.2 software (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>). The MERS-CoV Spike protein gene-specific primers were designed by using the selected sequences to amplify the complete Spike protein gene. The viral RNA was purified from culture supernatants and nasal swab by QIAamp Viral RNA Mini Kit (Qiagen). The MERS-CoV was detected by real-time RT-PCR using upE gene and further confirmed by ORF1a primers, as described previously. Briefly, A 25 µl reaction was set up containing 5 µl of RNA, 12.5 µl of 2 X reaction buffer from the Superscript III one step RT-PCR system with Platinum Taq Polymerase 1 µl of reverse transcriptase/Taq mixture from the kit, 0.4 µl of a 50 mM MgCl₂ solution (Invitrogen), 1 µg of non-acetylated bovine serum albumin (Sigma), 400 nM of primers EMC-Orf1a-Fwd and EMC-Orf1a-Rev, as well as 200 nM of probe EMCOrf1a-Prb (6-carboxyfluorescein (FAM)-TTGCAAATTGGCTTGCCCCACT - 6-carboxy-N,N,N,N-tetramethylrhodamine (TAMRA). Thermal cycling was performed at 55 °C for 20 min for the RT, followed by 95 °C for 3 min and then 45 cycles of 95 °C for 15 s, 58 °C for 30 s. [7,46]. The purified Viral RNA was used to amplify the MERS-CoV Spike protein gene. The purified viral RNA was reverse transcribed, and the Spike protein gene was amplified by RT-PCR. The amplified product was visualized on 1% agarose gel. The PCR product was eluted from the gel and purified with a gel purification kit (Qiagen).

Sequencing, and diversity analysis

The PCR amplicon was gel eluted and purified and further sequenced by BIVeriti thermal cycler (Applied Biosystems) using Spike protein gene-specific primers in SIAU. The sequence alignment was performed using BioEdit, version 7.2.5. and the genetic diversity was determined by analyzing the sequence identity matrix with selected MERS-CoV sequences from Saudi Arabia. To explore the phylogenetic relationship of generated MERS-CoV sequences with sequences were analyzed by MEGAX software programme and a phylogenetic tree was constructed [33].

Recombination analysis

To analyze the pattern of recombination among the Spike protein gene sequences, we have used the Recombination Detection Program (RDP4, v.4.70) [34]. The multiple sequences alignment file was imported to RDP software for recombination detection. The major and minor parent and possible recombinant with recombination breakpoints and hot spots with their start and end sequences were also identified by default settings which include GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq, to detect putative recombination events in the Spike gene sequence of MERS-CoV. The putative recombination events were identified and considered significant with the cut off *P*-value (≤ 0.05) with standard Bonferroni corrections.

Structural prediction of spike protein gene

The structure was predicted by using SWISS modeling software (<https://swissmodel.expasy.org/>) utilizing the Spike protein gene of both human and camel samples with selected MERS-CoV. As it has already been published about the recombinant MERS-CoV, we selected a few sequences of recombinant virus to compare the protein structure with samples collected from human and camel samples.

Results

Sample collection, virus detection and amplification of the spike protein gene

The nasal swabs collected from both human and camels were found positive by RT-PCR. The Vero cells were inoculated with samples obtained from human and camels showed a cytopathic effect after 3 days of inoculation. The culture supernatants were used to detect the virus by RT-PCR for *npE*, *ORF1a*, and *ORF1b* regions. The Spike protein gene was amplified by using Spike gene-specific primers and visualized on 1% Agarose gel.

Genetic diversity analysis

The full-genome of Spike protein gene was sequenced bidirectionally from both human and camel samples at SIAU, KFMRK, KAU, Jeddah and tentatively designated as sample1390-Hu-Jed and sample-31-Cam-Jed. The generated sequences have been submitted in GenBank with their accession numbers MN403101, MN403102 and used for diversity analysis with previously published sequences. The sample1390-Hu-Jed was used to analyze the similarity with other sequences and the highest nucleotides (98.4%) and amino acid similarity (99.8%) was observed with many isolates of MERS-CoV and the lowest nucleotides (64.7%) and amino acid (65.7%) similarity with 3 selected isolates collected from human samples with their accession numbers (KT624243-Riyadh-2014, KT624242-Madinah-2014, KT624241-Dammam-2014). Interest-

ingly, the samples collected from camel also showed higher nucleotides similarity ranged from 81.7%–98.4% and amino acid similarity 83.0%–99.7% with selected virus isolates (KY706247-KFU-HKU-2014, KT368853-Jeddah-2014, KT368860-Jeddah-2015, KP405227-AI-Ahsa-2015) Table 1.

The phylogenetic analysis with Spike protein gene sequences (Nucleotide and Amino acid) of selected MERS-CoV sequences deposited in GenBank was performed. The phylogenetic tree based on nucleotide sequences showed various clusters. The sample1390-Hu-Jed formed closed cluster with an isolate from Buraydah (KT806006-2014) and Taif (KR912196) while the sample-31-Cam-Jed formed the closest cluster with camel isolate from Hufuf (KFU-HKU-KY706247-2014) there mixed clustering of human and camel isolates were observed in a phylogenetic relationship, Fig. 1. The phylogenetic relationship was also analyzed by using amino acid sequences. The phylogenetic tree was constructed, and an almost similar pattern of clustering was observed with selected virus isolates (Fig. 2).

The amino acid sequence variations were observed at multiple locations. Total of 28 samples sequences showed sequence transversion of amino acids. Interestingly, two common variations were also observed in all 35 samples analyzed as compared to human and camel sample collected and used in this study (Table 2). An attempt was made to analyze the nucleotide and amino acid sequence variations with selected recombinant MERS-CoV as compared to sample 31-camel-Jeddah and total nineteen nucleotide sequence variations were observed scattered throughout the Spike protein gene, and only three amino acid in camel while two amino acid variations were observed in human sample (Fig. 3A, B). Our findings were supported by earlier reports and a unique amino acid substitution was observed in RBD which affecting the binding efficiency (Kossyvakis et al., 2015). Additionally, in another study, only five mutations were detected in consensus sequences 473 intrahost and single nucleotide variants were identified (Borucki et al., 2016).

Analysis of recombination

The Spike protein gene sequences were used to analyze the pattern of recombination in selected MERS-CoV from Saudi Arabia. Putative recombination events were identified using Recombination Detection Programme (RDP4, v.4.70) with the default settings [34]. Based on RDP parameters, no significant recombination was observed. But when the recombination pattern was analyzed by using BootScan, 3Seq, and parameters, the recombination was observed in eight, seven and one sequences at multiple positions with the average *P*-value of 6.870×10^{-02} , 8.031×10^{-03} and 4.589×10^{-02} respectively with all the sequences analyzed (Fig. 4). Recently, a genetic diversity analysis study was conducted from UAE and 10 recombination events were observed in the camel samples. The most frequent recombination breakpoint was the junctions between *ORF1b* and Spike protein gene [20]. Our data also supported by above reports that most of the recombination and sequences diversity have been observed in the Spike protein gene.

Structural prediction of Spike protein gene

The structure of Spike protein using amino acid sequences generated in this study from both human and camel samples and was compared to protein structure with recombinant MERS-CoV reported recently. The predicted structures of both recombinant as well as our generated sequences from both human and camel sample sequences, are presented in Fig. 5. During modeling, no significant variations were observed among the sequences and 99.4% similarity was observed with the available template in the protein database, but 3 amino acid variations were observed in human and 2 variations in camel samples in this study. The mutation in the

Table 1
Sequence identity matrix of spike protein gene (sample1390-Hu-Jed) with selected MERS-CoV isolates from Saudi Arabia.

S.No	Accession Nos	Hosts	Locations	Year	% Identity	
					NT	AA
1	KT806044	Human	Jeddah	2015	99.8	99.8
2	KT806045	Human	Jeddah	2015	99.8	99.8
3	KT805993	Human	Jeddah	2015	99.9	99.8
4	KF958702	Human	Jeddah	2014	98.1	99.7
5	KT805999	Human	Jeddah	2015	98.1	99.8
6	KT624243	Human	Riyadh	2014	64.7	65.7
7	KT806021	Human	Riyadh	2015	98.4	99.8
8	KT806015	Human	Makkah	2015	98.3	99.7
9	KM027261	Human	Makkah	2015	98.2	99.8
10	KT624242	Human	Madinah	2014	64.7	65.7
11	KT624241	Human	Dammam	2014	64.7	65.7
12	KT806041	Human	Hufuf	2015	98.3	99.8
13	KT806023	Human	Aflaj	2015	98.4	99.8
14	KT806022	Human	Qunfidhah	2015	98.3	99.7
15	KT806017	Human	Taif	2015	98.4	99.8
16	KT806016	Human	Duba	2015	98.3	99.7
17	KT806008	Human	Shaqra	2015	98.3	99.8
18	KT806006	Human	Buraydah	2015	98.0	99.6
19	KT806003	Human	Rass	2015	98.4	99.7
20	KT806000	Human	Unayzah	2015	98.1	99.7
21	KT805996	Human	Najran	2015	98.4	99.8
22	KT805992	Human	Aseer	2015	98.1	99.6
23	KT805988	Human	Buqayq	2015	98.0	99.7
24	KT805984	Human	Dhahran	2015	98.3	99.8
25	KT805966	Human	Quwaiyah	2015	98.0	99.7
26	KR912196	Human	Taif	2014	98.1	99.7
27	KT357805	Human	HafrAlBatin	2014	98.1	99.7
28	MK462258	Human	Qaseem	2018	98.2	99.6
29	KY688118	Human	Khobar	2015	98.3	99.8
30	KT806053	Human	Kharj	2015	98.4	99.8
31	KT368853	Camel	Jeddah	2014	98.0	99.6
32	KT368860	Camel	Jeddah	2015	98.4	99.7
33	KP405227	Camel	Al-Ahsa	2015	98.1	99.7
34	KY706247	Camel	KFU-HKU	2014	81.7	83.0
35	Sample- 31	Camel	Jeddah	2014	96.5	99.6

Table 2
Variations of amino acids (AA) in spike protein gene of MERS-CoV as compared to sample 31-Cam-Jed.

S.No	Accession Nos	Hosts	Locations	Year	AA positions in genome	Changes of AA
1	MK462258	Human	Qaseem	2018	26,133,191	V/L, S/R, S/P
2	KT806017	Human	Taif	2015	84	M/I
3	KT806008	Human	Shaqra	2015	84	M/I
4	KT368853	Camel	Jeddah	2014	143,992,1224	I/F, E/D, G/S
5	KT805984	Human	Dhahran	2015	200	S/F
6	KT806006	Human	Buraydah	2015	222,556,1224	N/Y, A/V, G/S
7	KT357805	Human	HafrAlBatin	2014	411	N/F
8	KT805988	Human	Buqayq	2015	424	T/I
9	KT805966	Human	Quwaiyah	2015	424	T/I
10	KP405227	Camel	Al-Ahsa	2015	424	T/I
11	KT806003	Human	Rass	2015	465	I/F
12	KY706247	Camel	KFU-HKU	2014	473	F/S
13	KT806016	Human	Duba	2015	507	L/P
14	KT806021	Human	Riyadh	2015	534	V/I
15	KR912196	Human	Taif	2014	536	E/K
16	KT806015	Human	Makkah	2015	556,833,1120	A/V, Q/R, T/I
17	KT805992	Human	Aseer	2015	626,746,1083	R/L, L/S, A/S
18	KT806022	Human	Qunfidhah	2015	920,1333	A/V, C/F
19	KF958702	Human	Jeddah	2014	1020	R/H
20	KT806016	Human	Duba	2015	1183	E/D
21	Sample 1390	Human	Jeddah	2019	1224	G/S
22	KT805999	Human	Jeddah	2015	1224	G/S
23	KT806000	Human	Unayzah	2015	1224	G/S
24	KR912196	Human	Taif	2014	1224	G/S
25	KT806041	Human	Hufuf	2015	1246	I/X
26	KT368860	Camel	Jeddah	2015	1333	C/F
27	All 35 samples				705	S/G
28	All 35 samples				1301	V/L

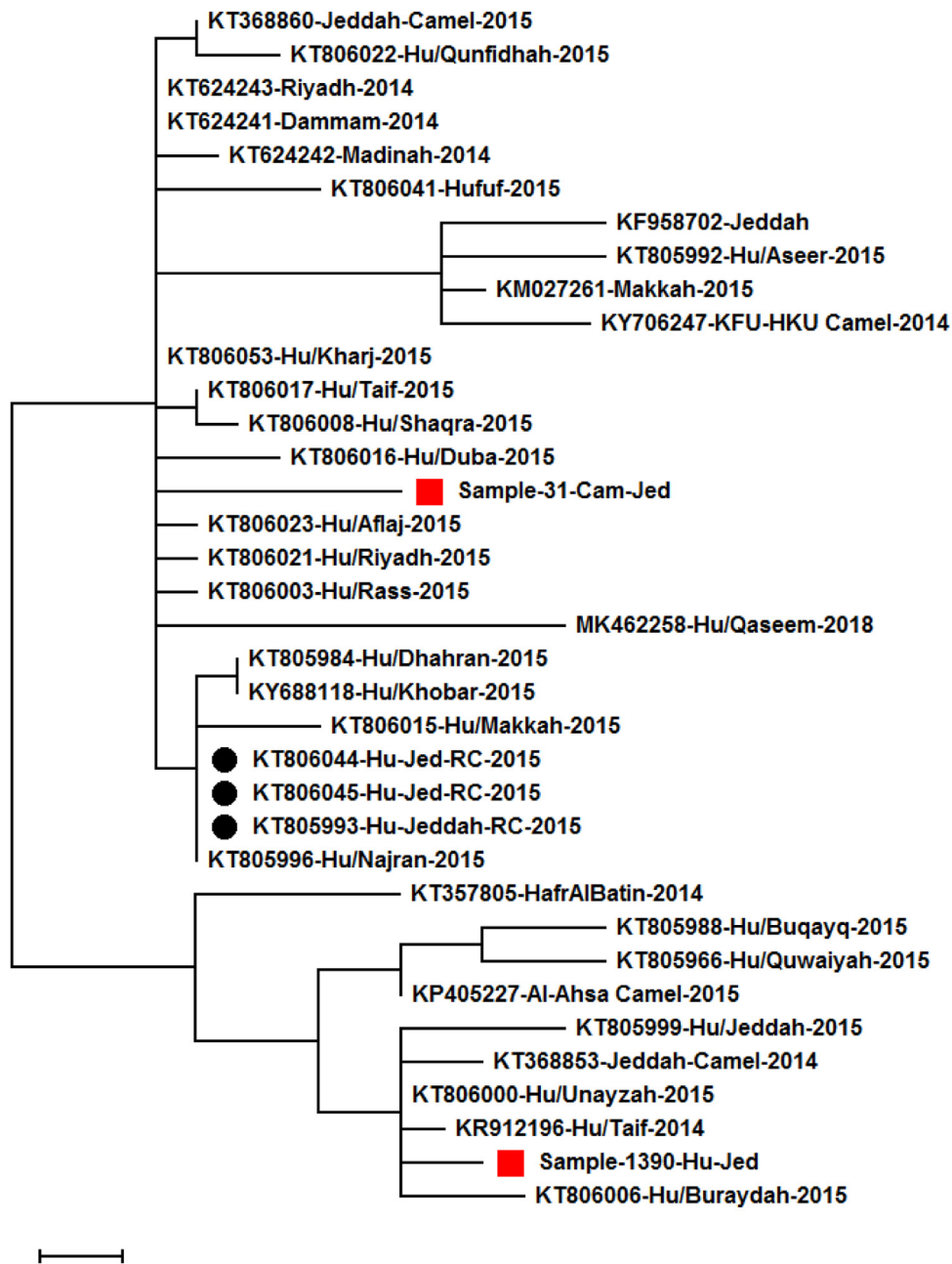


Fig. 1. Phylogenetic tree based on nucleotide sequences of spike protein gene with selected MERS-CoV.

RBD of Spike protein gene affects the interactions with the human receptor, CD26. The similar kind of mutations and changes in amino acid have been reported earlier and the structural changes has been presented [16,17].

Discussion

The MERS-CoV was identified from Jeddah, Saudi Arabia since 2012 and causes respiratory disease to a human. This virus has spread to 27 countries. The genetic diversity has been reported among CoVs. MERS-CoV also has been reported to have genetic diversity across the whole genome and especially in Spike protein gene. Our study has provided the genetic diversity of the Spike protein gene isolated from both camel and human samples from Jeddah, Kingdom of Saudi Arabia. The highest sequence identity was observed with the previously reported sequences submitted in

the GenBank. The phylogenetic tree relationships were also formed a closed cluster with earlier known viruses from various locations in the Kingdom. The nucleotide and amino acid variations were scattered throughout the full Spike protein genome. Interestingly, two common amino acid changes were observed in all the 35 samples analyzed. As it has already been reported about the emergence of the recombinant virus from Saudi Arabia [19,28] and mutant virus from South Korea [16,17] and other intrahost variants reported [35]. The transversion of three amino acids was observed in our sample collected from human and two amino acid variation was observed in the camel sample as compared to selected recombinant virus sequences from Saudi Arabia. The comparative structure of the Spike protein gene has been predicted and presented in this study to show how the structural changes as compared to recombinant virus appear. The effect of a mutation in the Spike protein gene with the viral entry to the host cell by using CD26 binding

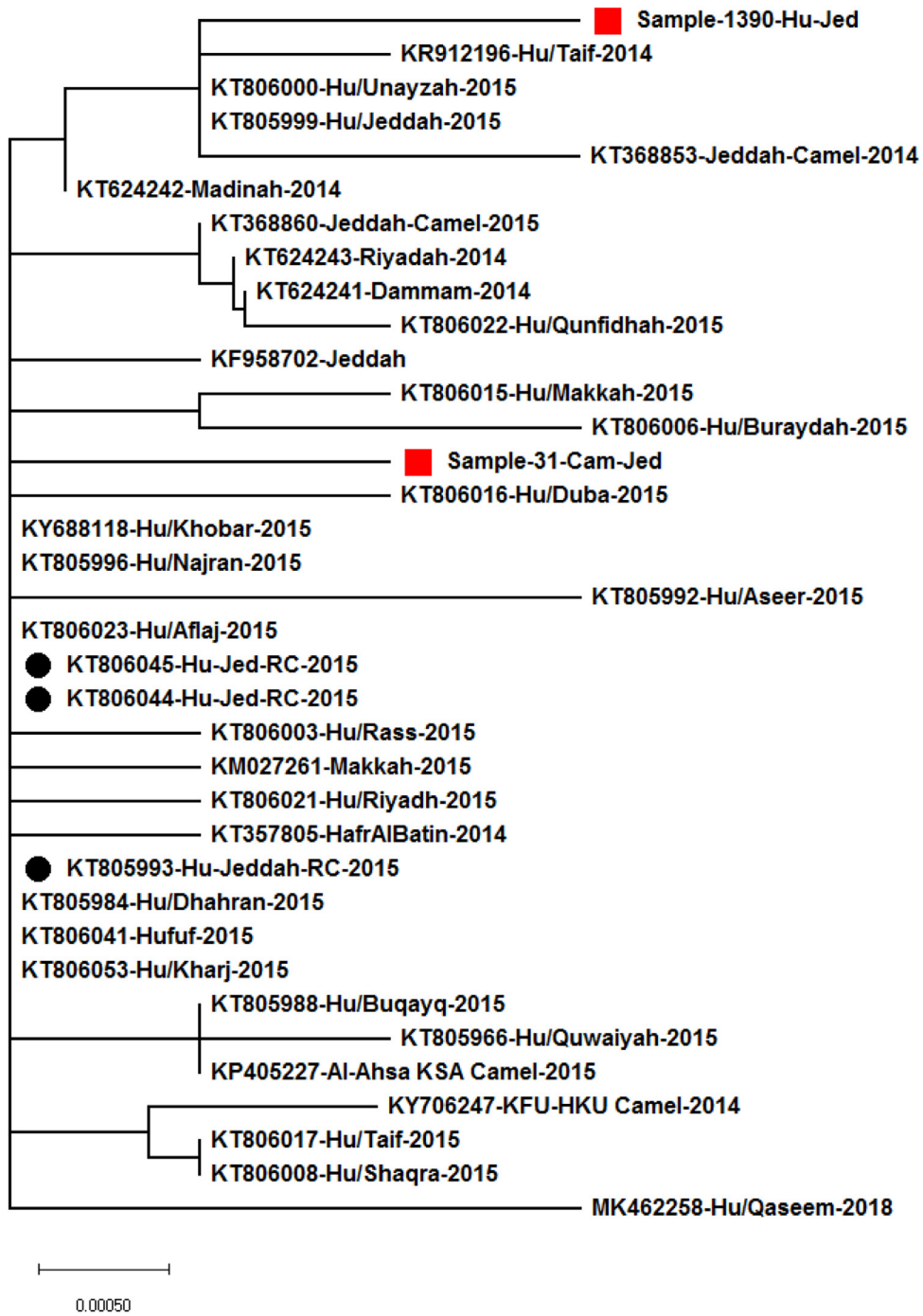


Fig. 2. Phylogenetic tree based on amino acid sequences of spike protein gene with selected MERS-CoV.

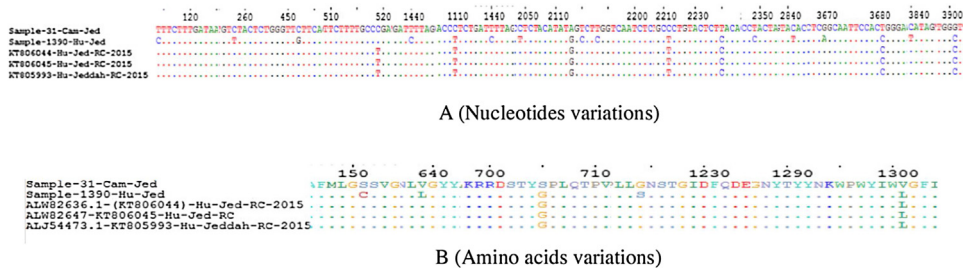


Fig. 3. Spike protein gene sequence variation of selected MERS-CoV. A: Position of Nucleotide sequences variations in Spike genome; B: Position of Amino acid sequences variations in Spike genome.

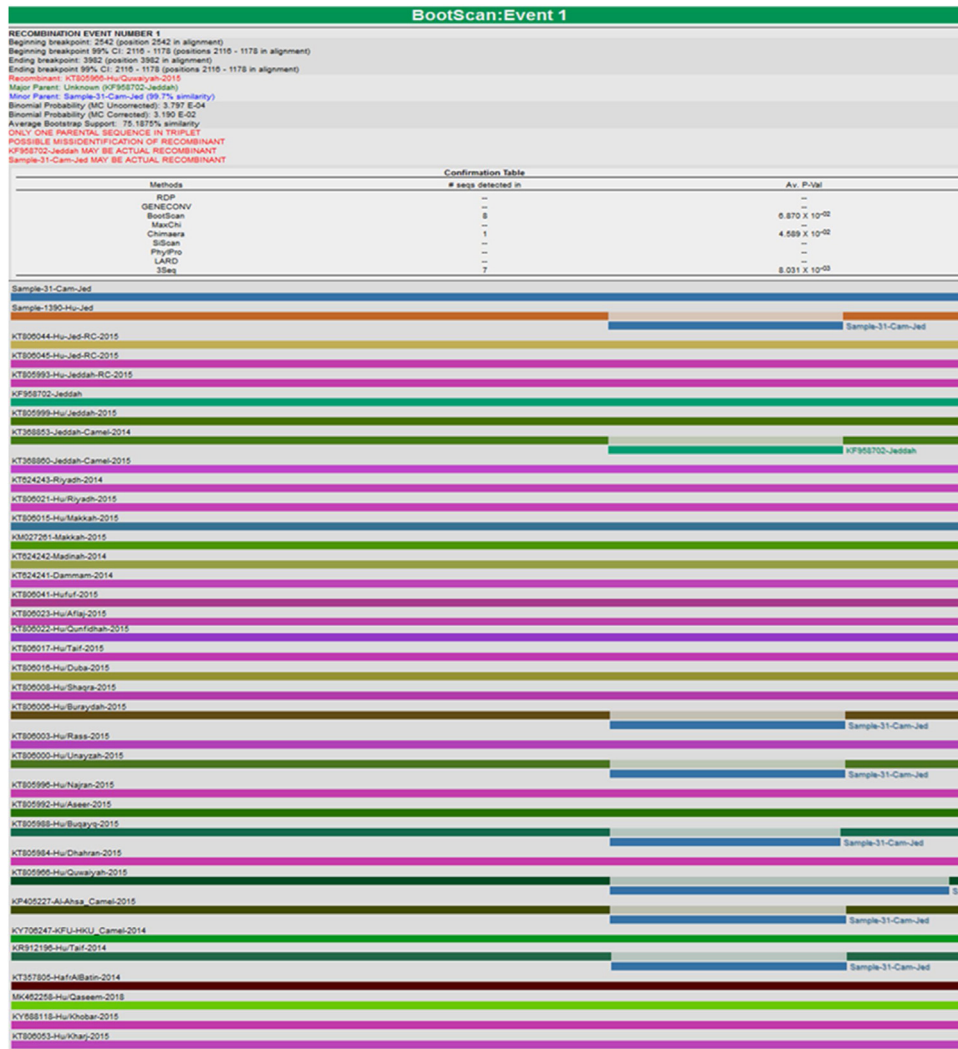


Fig. 4. Analysis of Recombination pattern of Spike protein gene with selected MERS-CoV from Saudi Arabia.

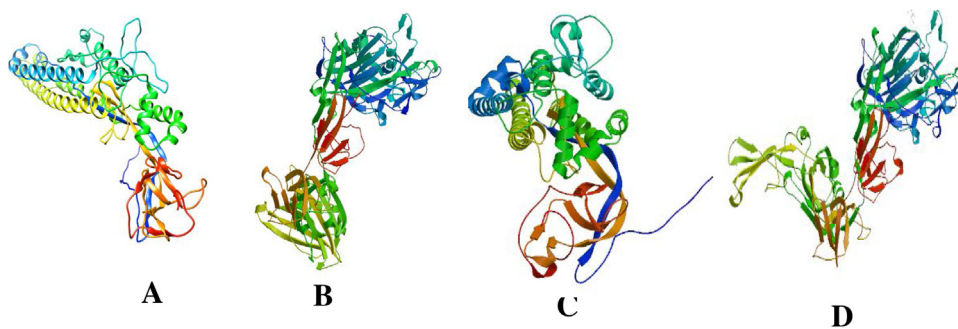


Fig. 5. *In silico* predicted structures of Spike protein gene of MERS-CoV. A: Camel sample-31; B: Human sample; C: Recombinant MERS-CoV (KT806044); D: Recombinant MERS-CoV (KT806045).

requires further detail study. These changes may affect the structural changes of RBD and interactions with the cognate human receptor, CD26 as it has been reported earlier [17].

The most important factor for the emergence of a novel virus is during the recombination step in the life cycle of the virus. The recombination happens with the co-circulating CoVs in multiple hosts which further increase the rate of recombinant virus emergence. The estimated rate of mutations in CoVs are known as moderate to high as compared to other known ssRNA viruses. The

rate of average substitution for CoVs has been reported as $\sim 10^{-4}$ substitutions per year per site while the average substitution of S gene in 229E was observed to be $\sim 3 \times 10^{-4}$ per site per year [36] and in SAARS-CoV it was estimated to be $0.80\text{--}2.38 \times 10^{-3}$ per site per year. In the case of MERS-CoV, the average rate of substitution was found to be 1.12×10^{-3} per site per year [37]. Recently, the co-circulation of multiple HCoV species in camel along with MERS-CoV has been reported from Saudi Arabia and resulted into an emergence of a recombinant virus which was responsible for an

outbreak in 2015 [28,38]. These results showed that many CoV are circulating in the wild animals originated from human and animal and by this, an opportunity is favored for the genetic recombination, evolution and emergence of potential and recombinant virus lethal to human [3].

By considering the variations in the MERS-CoV Spike protein gene reported so far, it is very important to consider the frequent appearance and conservation of RBD alterations in human infections. It is well reported that the interspecies transmission of CoVs is mainly mediated by mutations in Spike protein gene with a high affinity toward human receptors [28,39,40]. The RBD mutations increased the unexpected emergence with reduced affinity to human CD26 in South Korean MERS-CoV outbreak. The Identification of recombinant virus from Saudi Arabia [28] as well as unusual presentation and emergence of MERS-CoV resulted in an outbreak in Riyadh [19] raises several critical questions. Based on epidemiologic features, the zoonotic transmission of the MERS-CoV from the animal reservoir has been suggested by an intermediate animal host. The evidence of dromedary camel as a reservoir for MERS-CoV has been already reported based on the identical sequences obtained from camel and a patient with close contact with camel nasal secretion which directs cross-species transmission without any intermediate host [7,13]. Indeed, it has been reported that the interfacial residual difference in receptors of various mammalian host species is very important for interspecies transmission of CoVs [40]. The sequence diversity and homology play an important role in various functions of virus and host cells like cellular fusion and attachment [28,41]. The viruses have ability to cause various types of diseases including neurological disorders. [42,43]. The human genome sequence diversity due to gene flow from South east Asia to other locations also plays an important role in the disease transmission and spread of pathogens from one location to other as well as intrahost transmission. As it has been shown that the ancient east-Asian mtDNA HG-M and exhibit the highest nucleotide diversity [44,45]. However, our data showed no significant amino acid sequences variations in the Spike protein gene. Finally, there is not much information available about the role of RBD mutation linked with reduced affinity to host receptor CD26 and MERS-CoV transmissibility to infect human. The genetic diversity was significantly observed in the Spike protein gene sequences of both nucleotides and amino acids collected from Saudi Arabia. The information generated from this study provided an insight to the pattern of sequence diversity which plays an important role in the viral disease spread as well as movement of virus from one host to others. This diversity information will play an important role in designing and development of antiviral drugs, vaccines as well as antiviral compounds against MERS-CoV.

Conclusion

Based on the data generated in this study, it is concluded that the genetic diversity of the Spike protein gene plays an important role in interaction with CD26. The genetic diversity emerges with the high frequency of recombination events in the CoVs resulted in the emergence of novel recombinant viruses with unpredictable changes in the virulence during human infection. Finally, to elucidate the evolutionary pathway, more detail mutation analysis of Spike protein gene needs to be done using more sequences from Saudi Arabia and we should take lessons from the outbreak caused by MERS-CoV and SAARS-CoV and we should take necessary measures to combat any further outbreak.

Author contributions

SSS collected and processed samples, designed the study, analyzed the data, and conceived the idea. EIA supervised the research and reviewed the manuscript.

Funding

Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia.

Competing interests

None declared.

Ethical approval

Not required.

Acknowledgments

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant number (G-203-141-40). The authors, therefore, acknowledge with thanks DSR technical and financial support. The author is also thankful to Special Infectious Agents Unit, King Fahd Medical Research Center (KFMRC), King Abdulaziz University, Jeddah, Saudi Arabia for providing necessary research facilities.

References

- [1] Brian DA, Baric RS. Coronavirus genome structure and replication. *Curr Top Microbiol Immunol* 2005;287:1–30.
- [2] Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev* 2005;69:635–64.
- [3] Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol* 2016:490–502.
- [4] Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 2012;367:1814–20.
- [5] WHO: Global alert and response. <https://www.who.int/emergencies/mers-cov/en/2019>. [Accessed 31 August 2019].
- [6] Memish ZA, Mishra N, Olival KJ, Fagbo SF, Kapoor V, Epstein JH, et al. Middle east respiratory syndrome coronavirus in bats, Saudi Arabia. *Emerging Infect Dis* 2013;19:1819–23.
- [7] Azhar EI, El-Kafrawy SA, Farraj SA, Hassan AM, Al-Saeed MS, Hashem AM, et al. Evidence for camel-to-human transmission of MERS coronavirus. *N Engl J Med* 2014;26:2499–505.
- [8] Oboho IK, Tomczyk SM, Al-Asmari AM, Banjar AA, Al-Mugti H, Aloraini MS. MERS-CoV outbreak in Jeddah — a link to health care facilities. *N Eng J Med* 2015;372:846–54.
- [9] Assiri A, Al-Tawfiq JA, Al-Rabeeh AA, Al-Rabiah FA, Al-Hajjar S, Al-Barrak A, et al. Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. *Lancet Infect Dis* 2013;9:752–61.
- [10] Drosten C, Seilmaier M, Corman VM, Hartmann W, Scheible G, Sack S, et al. Clinical features and virological analysis of a case of Middle East respiratory syndrome coronavirus infection. *Lancet Infect Dis* 2013;9:745–51.
- [11] Drosten C, Meyer B, Müller MA, Corman VM, Al-Masri M, Hossain R, Madani H, et al. Transmission of MERS-coronavirus in household contacts. *N Engl J Med* 2014;371:828.
- [12] Lee SS, Wong NS. Probable transmission chains of MERS-CoV and the multiple generations of secondary infections in South Korea. *Int J Infect Dis* 2015;38:65–7.
- [13] Azhar EI, Hashem AM, El-Kafrawy SA, Sohrab SS, Aburizaiza AS, Farraj SA, et al. Detection of the Middle East respiratory syndrome coronavirus genome in an air sample originating from a camel barn owned by an infected patient. *MBio* 2014;4, e01450-14.
- [14] Chen Y, Chan KH, Kang Y, Chen H, Luk HK, Poon RW, et al. A sensitive and specific antigen detection assay for Middle East respiratory syndrome coronavirus. *Emerg Microbes Infect* 2015;4:e26.

- [15] Ali M, El-Shesheny R, Kandeil A, Shehata M, Elsokary B, Gomaa M, et al. Cross-sectional surveillance of Middle East respiratory syndrome coronavirus (MERS-CoV) in dromedary camels and other mammals in Egypt, August 2015 to January 2016. *Euro Surveill* 2017;11:30487.
- [16] Kim DW, Kim YJ, Park SH, Yun MR, Yang JS, Kang HJ, et al. Variations in spike glycoprotein gene of MERS-CoV, South Korea, 2015. *Emerg Infect Dis* 2016;22(1):100–4.
- [17] Kim Y, Cheon S, Min CK, Sohn KM, Kang YJ, et al. Spread of mutant Middle East respiratory syndrome coronavirus with reduced affinity to human CD26 during the South Korean outbreak. *MBio* 2016;7(2):e00019.
- [18] Ali S, Omrani Jaffar A, Al-Tawfiq, Memish Ziad A. Middle East respiratory syndrome coronavirus (MERS-CoV): animal to human interaction. *Pathog Glob Health* 2016;109(8):354–62.
- [19] Amer H, Alqahtani AS, Alzoman H, Algerian N, Memish ZA. Unusual presentation of Middle East respiratory syndrome coronavirus leading to a large outbreak in Riyadh during 2017. *Am J Infect Control* 2018;46(9):1022–5.
- [20] Yusof MF, Queen K, Eltahir YM, Paden CR, Al Hammadi ZMAH, Tao Y, et al. Diversity of Middle East respiratory syndrome coronaviruses in 109 dromedary camels based on full-genome sequencing, Abu Dhabi, United Arab Emirates. *Emerg Microbes Infect* 2017;6(11):e101.
- [21] Paden CR, Yusof MFBM, Al Hammadi ZM, Queen K, Tao Y, Eltahir YM, et al. Zoonotic origin and transmission of Middle East respiratory syndrome coronavirus in the UAE. *Zoonoses Public Health* 2018;65(3):322–33.
- [22] Dudas G, Carvalho LM, Rambaut A, Bedford T. MERS-CoV spillover at the camel-human interface. *Elife* 2018;7(January):e31257.
- [23] Belouzard S, Millet JK, Licitra BN, Whittaker GR. Mechanisms of coronavirus cell entry mediated by the viral Spike protein. *Viruses* 2012;4:1011–33.
- [24] Raj VS, Mou H, Smits SL, Dekkers DH, Müller MA, Dijkman R, et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* 2013;495(7440):251–4.
- [25] Wang N, Shi X, Jiang L, Zhang S, Wang D, Tong P, et al. Structure of MERS-CoV Spike receptor-binding domain complexed with human receptor DPP4. *Cell Res* 2013;23:986–93.
- [26] Coleman CM, Venkataraman T, Liu YV, Glenn GM, Smith GE, Flyer DC, et al. MERS-CoV Spike nanoparticles protect mice from MERS-CoV infection. *Vaccine* 2017;35(12):1586–9.
- [27] Kim DW, Kim YJ, Park SH, Yun MR, Yang JS, Kang HJ, et al. Variations in spike glycoprotein gene of MERS-CoV, South Korea, 2015. *Emerg Infect Dis* 2016;22(1):100–4.
- [28] Assiri AM, Midgley CM, Abedi GR, Bin Saeed A, Almasri MM, Lu X, et al. Epidemiology of a novel recombinant MERS-CoV in humans in Saudi Arabia. *J Infect Dis* 2016;214:712–21.
- [29] Xiaoyan Lu, Lori A, Rowe MF, James S, Glen R, Abedi OE, et al. Spike gene deletion quasispecies in serum of patient with acute MERS-CoV infection. *J Med Virol* 2017;89:542–5.
- [30] Anthony SJ, Gilardi K, Menachery VD, Goldstein T, Ssebide B, Mbabazi, et al. Further evidence for bats as the evolutionary source of Middle East respiratory syndrome coronavirus. *MBio* 2017;8(2):e00373–417.
- [31] Chu DKW, Hui KPY, Perera RAPM, Miguel E, Niemeyer D, Zhao J, et al. MERS-CoV from camels in Africa exhibit region-dependent genetic diversity. *Proc Natl Acad Sci U S A* 2018;115(12):3144–9.
- [32] Chung YS, Kim JM, Man Kim H, Park KR, Lee A, Lee NJ, et al. Genetic characterization of Middle East respiratory syndrome coronavirus, South Korea, 2018. *Emerg Infect Dis* 2019;25(May (5)):958–62.
- [33] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–4.
- [34] Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 2015;1(1), vev003. eCollection 2015.
- [35] Borucki MK, Lao V, Hwang M, Gardner S, Adney D, Munster V, et al. Middle east respiratory syndrome coronavirus intra-host populations are characterized by numerous high frequency variants. *PLoS One* 2016;11(1):e0146251.
- [36] Pyrc K, Dijkman R, Deng L, Jebbink MF, Ross HA, Berkhout B, et al. Mosaic structure of human coronavirus NL63, one thousand years of evolution. *J Mol Biol* 2006;364(5):964–73.
- [37] Cotten M, Watson SJ, Zumla AI, Makhdoom HQ, Palsler AL, Ong SH, et al. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *MBio* 2014;5(1), e01062–13.
- [38] Sabir JS, Lam TT, Ahmed MM, Li L, Shen Y, Abo-Aba SE, et al. Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science* 2016;351:81–4, 2016.
- [39] Woo PC, Lau SK, Huang Y, Yuen KY. Coronavirus diversity, phylogeny and inter-species jumping. *Exp Biol Med (Maywood)* 2009;234:1117–27.
- [40] Lu G, Wang Q, Gao GF. Bat-to-human: spike features determining “host jump” of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol* 2015;23:468–78.
- [41] Gorry PR, Ahmad F, Mohl J, Alkhatib G. Low levels of HIV-1 envelope-mediated fusion are associated with long-term survival of an infected CCR5^{-/-} patient. *AIDS* 2018;32(16):2269–78.
- [42] Ahmad Fahim. ZIKV leads to microcephaly. *EC Microbiol* 2016;4:12–3.
- [43] Ahmad F, Siddiqui A, Kamal MA, Sohrab SS. Inhibition of Neurogenesis by Zika virus infection. *CNS Neurol Disord Drug Targets* 2018;2:78–86.
- [44] Ahmad F, Siddiqui A, Raut AA, Pandey PK. Austro Asiatic tribes are original native inhabitants of India. *Biotechnol Int* 2009;2:91–115.
- [45] Borker M, Ahmad F, Khan F, Agrawal S. Paleolithic spread of Y-chromosomal lineage of tribes in Eastern and Northeastern India. *Ann Hum Biol* 2011;6:736–46.
- [46] Corman VM, Müller MA, Costabel U, Timm J, Binger T, Meyer B, et al. Assays for laboratory confirmation of novel human coronavirus (hCoV-EMC) infections. *Euro Surveill* 2012;49:20334.