Genetic drift of MERS-CoV suggests that camel may not be the sources of human infection

Yuan-Qing Pan[1], Fan Guo[1], and Li Xing[1]

[1]Affiliation not available

August 14, 2020

Abstract

Middle East respiratory syndrome (MERS) is caused by MERS-CoV. To explore the conservation of non-coding 5'-UTR of MERS-CoV and its implication in epidermiology, 5'-UTRs from 252 human MERS-CoV and 207 camel and dromedary camel MERS-CoV were analyzed. We identified two conserved pyrimidine nucleotides that flank identical UAAU element in the loop of stem loop 2 of MERS-CoV 5'-UTR. These conserved pyrimidine nucleotides can be used as a novel genetic signature to re-genotype MERS-CoV into 3 types, i.e. U—-U, C—-U, and C—-C type viruses. Human MERS-CoV displays a genetic drift from U—-U, C—-U, to C—-C during period of 2012-2019. Camel virus only displayed a genetic drift from U—-U to C—-U, particularly in a delayed way when compared with human virus. The discrepancy in genetic drift suggests that camel may not necessarily be the natural reservior for human infection.

Genetic drift of MERS-CoV suggests that camel may not be the sources of human infection

Running Title: Genetic drift of MERS-CoV

Yuan-Qing Pan[1 a] , Fan Guo[1a] , and Li Xing [a *]

[a]*Institute of Biomedical Sciences, Shanxi University, 92 Wucheng Road, Taiyuan 030006, Shanxi province, People's Republic of China*

[1]Contributed equally to this study

[*]**Corresponding author**

Li Xing, PhD.

Email: *xingli107@gmail.com*

Format: Short communications

Word account: Abstract: 186 words; Text: 1200 words;

Figure No. 2

References: 14

Summary

Middle East respiratory syndrome (MERS) is caused by MERS-CoV. To explore the conservation of non-coding 5'-UTR of MERS-CoV and its implication in epidermiology, 5'-UTRs from 252 human MERS-CoV and 207 camel and dromedary camel MERS-CoV were analyzed. We identified two conserved pyrimidine nucleotides that flank identical UAAU element in the loop of stem loop 2 of MERS-CoV 5'-UTR. These conserved pyrimidine nucleotides can be used as a novel genetic signature to re-genotype MERS-CoV into 3

types, i.e. U—-U, C—-U, and C—-C type viruses. Human MERS-CoV displays a genetic drift from U—-U, C—-U, to C—-C during period of 2012-2019. Camel virus only displayed a genetic drift from U—-U to C—-U, particularly in a delayed way when compared with human virus. The discrepancy in genetic drift suggests that camel may not necessarily be the natural reservior for human infection.

Key words: MERS-CoV; Genetic drift; Pyrimidine; Saudi Arabia; 5'-UTR

Introduction

Middle East respiratory syndrome (MERS) was first reported in Saudi Arabia in 2012 (1). By the end of November 2019, a total of 2494 laboratory-confirmed cases of MERS were reported globally, which include 858 associated deaths (case–fatality rate is 34.4%) (2). 84.2% of these cases were from Saudi Arabia (2102 cases, including 780 related deaths). 27 countries have been affected by MERS (2). The causative pathogen of MERS is MERS-related coronavirus (MERS-CoV) that infect humans and camels (3). Like Severe acute respiratory syndrome coronavirus-1 and -2 (SARS-CoV-1 and SARS-CoV-2) that caused outbreak of SARS in April 2003 and ongoing pandemic of coronavirus disease 2019 (COVID-19) in 2020, respectively, MERS-CoV is also a member of the genus β-coronavirus (4). MRES-CoV enters host cell by binding to cellular receptor dipeptidyl peptidase 4 (DPP4, also known as CD26) (5). The genome of MERS-CoV is a positive-sense, single-stranded RNA with the length of over 30,000 nucleotides (nt), which contains 10 predicted open reading frames (ORFs) flanked by untranslated regions at 5'- and 3'-ends of viral genome (named 5'-UTR and 3'-UTR, respectively). MERS-CoV encodes 2 replicase polyproteins (pp1a and pp1b), 4 structural proteins including envelope (E), membrane (M), nucleocapsid (N), and spike glycoprotein (S), and 4 non-structural proteins (6). Coding sequences of MERS-CoV, in particular those for RNA polymerase, N and S proteins have been well analyzed in phylogenetic analysis (7, 8) and used as targets in molecular diagnostic assay (9, 10). Non-coding sequence in the 5'- and 3'-UTR was less studied. Herein, the secondary structure and sequence conservation of 5'-UTR were focused in this report. We found two conserved pyrimidine nucleotides in the loop region of 5'-UTR stem loop 2 and explored its potential implication as a novel genetic signature in epidemiology.

Materials and Methods

Analysis of sequence and secondary structure conservation of MERS-CoV 5'-UTR.

GenBank at NCBI is the NIH genetic sequence database that has compiled complete MERS-CoV genomes. 252 complete viral genomes were isolated from humans in 11 countries including Saudi Arabia, United Arab Emirates, Jodan, Quatar, Oman, Egypt, United Kingdom, USA, South Korea, Thailand, and China. 74 complete genomes were isolated from camels in Saudi Arabia and 132 from dromedary camels in United Arab Emirates. The initial alignment of sequence was performed using EMBL-EBI Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). The most complete sequence for 5'-UTR were further analyzed using LocARNA program (11, 12)(*http://rna.informatik.uni-freiburg.de/*) to determine the conservation of both sequence and secondary structure of viral RNA. LocARNA performs simultaneous alignment and folding of multiple RNAs. It implements the state-of-the-art light-weight alignment algorithm with ensemble-based sparsification to improve the accuracy of analysis.

Real-time tracking of prevalence of MERS-CoV from 2012-2019

Sequence analysis reveals 2 conserved pyrimidine nucleotides in the loop of stem loop 2 of MERS-CoV 5'-UTR. According to this newly identified genetic signature, MERS-CoV can be re-genotyped into 3 types, U—-U, C—-U, and C—-C types. The prevalence of each virus type during 2012 to 2019 was determined by plotting the numbers of case reported in NCBI GenBank. The genetic drift between human and camel viruses was compared.

Results

Conserved pyrimidine nucleotides indicate a genetic drift of human MERS-CoV from 2012 to 2019

The conservation of both sequence and secondary structure of 5'-UTR from 252 human MERS-CoV reported from 2012-2019 was evaluated by using LocARNA program (11, 12). 5'-UTR contains 287 nucleotides in length (according to NCBI Reference Sequence: NC_019843.3) and forms 5 conserved stem loop regions (SL1-5). We found 2 conserved pyrimidine nucleotides in the longest stem loop region-2 (SL2) of MERS-CoV 5'-UTR (Figure 1). These two conserved pyrimidine nucleotides flank 4 consecutive nucleotides UAAU in the loop of SL2 (Figure 1). According to the conserved pyrimidine nucleotides, human MERS-CoV can be re-genotyped into 3 types, i.e. UUAAUU, CUAAUU, and CUAAUC (referred to as U—-U, C—-U, and C—-C types, respectively). In human population, MERS-CoV displays a U to C genetic drift over 8 years period from 2012 to 2019 in Saudi Arabia (Figure 2A) and in a global scale (Figure 2B). The first detected MERS-CoV is U—-U type virus that persisted from April 2012 until June 2015. C—-U type virus then emerged in 2013 and persisted until 2016. C—-C type virus emerged in April 2015 and then soon became dominant in Saudi Arabia until 2019. Only this type virus was detected after October 2016.

Genetic drift in camel and dromedary camel MERS-CoV was delayed

By far, most of MERS-CoV outbreak occurred in Saudi Arabia and almost all human infections were linked directly or indirectly to Arabian Peninsula. Camels and dromedary camels are supposed to be sources of human zoonotic infection of MERS-CoV since the serum of camels contain specific antibodies against MERS-CoV spike protein (13, 14). We also analyzed 74 complete genome sequences of camel MERS-CoV and 132 dromedary camel MERS-CoV in Saudi Arabia and United Arab Emirates, respectively. Both countries are located in Arabian Peninsula. Only U—-U type virus was detected in 2013 in Saudi Arabia (Figure 2C) or in 2014 in United Arab Emirates (Figure 2D). C—-U type virus then emerged and became dominant in Saudi Arabia in 2014. One year later, this type virus emerged in United Arab Emirates. While the camel MERS-CoV displayed a genetic drift from U—-U to C—-U, the timeframe between human and camel viruses is not in consistence. Noticeably, in 2015, almost 43% of human MERS-CoVs were C—-C type virus in Saudi Arabia, but this type virus was never ever detected in camels in Saudi Arabia (Figure 2C) or United Arab Emirates (Figure 2D) at the same period. In contrast to SARS-CoV-1 and -2, direct human-to-human transmission of MERS-CoV seems to be limited (10). While the main known exposures to MERS-CoV are health care-associated infections or contact with camels (Drosten et al. 2015), the source of infection for most primary human cases are still not known (Fanoy et al. 2014). Camels are believed to be the natural reservoir for human infection (13, 14). However, this notion may not be supported by the results shown here. While we can not exclude the species-cross transmission between human and camels, the results in this report suggest that human infection may have independent unknown source of infection.

Discussion

The phylogenetic analysis shows that camel and human MERS-CoVs are clustered together. The human-camel MERS-CoV cluster is further divided into two clades, i.e. clade A and clade B. The earliest cases of MERS including EMC/2012 and Jordan-N3/2012 belong to clade A clusters, and all other cases from 2012-2015 are in clade B (7, 8). Being distinct from phylogenetic analysis, which is biasedly based on the coding sequence, characterization of conserved pyrimidine nucleotide in the non-coding 5'-UTR of MERS-CoV may represent another possible mechanism in virus adaptive evolution since the conserved nucleotide sequence and structure of 5'-UTR directly interact with the host cellular factors during virus replication to regulate transcription, translation, and genomic RNA synthesis. A total of 10 complete genomes of MERS-CoV were reported in 2012 and classified into 2 clades (clade A and clade B) in phylogenetic analysis (7). However, all of them are the U—-U type virus according to the conserved pyrimidine nucleotide in 5'-UTR. Thus, the conserved pyrimidine nucleotides in 5'-UTR can serve as a novel genetic signature in evolution of MERS-CoV.

There are no specifically effective therapeutics and vaccines for MERS. In general, surveillance for MERS-CoV in humans and camels in affected countries is still a priority for infection control. In addition to the coding sequence, the variations in non-coding region, particularly in the conserved stem-loop region of 5'-UTR can offer novel implication in epidemiology. Its importance needs an attention without contempt in

both tracing the chain of infection transmission and algorithm of phylogenetic analysis.

Ethics Statement

Not applicable.

Data Availability Statement

The data that support the findings of this study are available in The NCBI GenBank database(https://www.ncbi.nlm.nih.gov/).

Conflict of interest

No conflict of interest declared.

Funding

Not applicable.

Acknowledgements

The authors are grateful to the Freiburg Bioinformatics Group, and all the persons who contributed to the production of MRES-CoV sequence data in National Center for Biotechnology Information (NCBI), USA.

Authors' contributions

LX conceived the study. YQP and FG collected data. LX, PTS, and YQP analyzed the data. LX wrote the paper.

References

1. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814-20.

2. WHO. 2019. MERS situation update.

3. Wong ACP, Li X, Lau SKP, Woo PCY. 2019. Global Epidemiology of Bat Coronaviruses. Viruses 11.

4. Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J, Quan L, Xia Z, Tan W, Cheng G, Jiang T. 2020. Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. Cell Host Microbe 27:325-328.

5. Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, Muth D, Demmers JA, Zaki A, Fouchier RA, Thiel V, Drosten C, Rottier PJ, Osterhaus AD, Bosch BJ, Haagmans BL. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 495:251-4.

6. de Wit E, van Doremalen N, Falzarano D, Munster VJ. 2016. SARS and MERS: recent insights into emerging coronaviruses. Nat Rev Microbiol 14:523-34.

7. Zhang Z, Shen L, Gu X. 2016. Evolutionary Dynamics of MERS-CoV: Potential Recombination, Positive Selection and Transmission. Sci Rep 6:25049.

8. Chu DK, Poon LL, Gomaa MM, Shehata MM, Perera RA, Abu Zeid D, El Rifay AS, Siu LY, Guan Y, Webby RJ, Ali MA, Peiris M, Kayali G. 2014. MERS coronaviruses in dromedary camels, Egypt. Emerg Infect Dis 20:1049-53.

9. Smits SL, Raj VS, Pas SD, Reusken CB, Mohran K, Farag EA, Al-Romaihi HE, AlHajri MM, Haagmans BL, Koopmans MP. 2015. Reliable typing of MERS-CoV variants with a small genome fragment. J Clin Virol 64:83-7.

10. Dawson P, Malik MR, Parvez F, Morse SS. 2019. What Have We Learned About Middle East Respiratory Syndrome Coronavirus Emergence in Humans? A Systematic Literature Review. Vector Borne Zoonotic Dis 19:174-192.

11. Raden M, Ali SM, Alkhnbashi OS, Busch A, Costa F, Davis JA, Eggenhofer F, Gelhausen R, Georg J, Heyne S, Hiller M, Kundu K, Kleinkauf R, Lott SC, Mohamed MM, Mattheis A, Miladi M, Richter AS, Will S, Wolff J, Wright PR, Backofen R. 2018. Freiburg RNA tools: a central online resource for RNA-focused research and teaching. Nucleic Acids Res 46:W25-W29.

12. Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. 2012. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. RNA 18:900-14.

13. Memish ZA, Cotten M, Meyer B, Watson SJ, Alsahafi AJ, Al Rabeeah AA, Corman VM, Sieberg A, Makhdoom HQ, Assiri A, Al Masri M, Aldabbagh S, Bosch BJ, Beer M, Muller MA, Kellam P, Drosten C. 2014. Human infection with MERS coronavirus after exposure to infected camels, Saudi Arabia, 2013. Emerg Infect Dis 20:1012-5.

14. Reusken CB, Haagmans BL, Muller MA, Gutierrez C, Godeke GJ, Meyer B, Muth D, Raj VS, Smits-De Vries L, Corman VM, Drexler JF, Smits SL, El Tahir YE, De Sousa R, van Beek J, Nowotny N, van Maanen K, Hidalgo-Hermoso E, Bosch BJ, Rottier P, Osterhaus A, Gortazar-Schmidt C, Drosten C, Koopmans MP. 2013. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. Lancet Infect Dis 13:859-66.

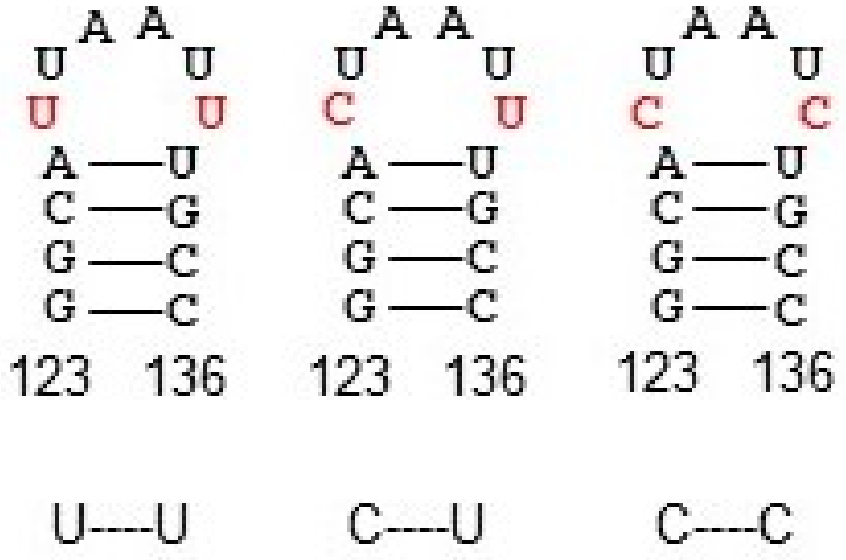

Figure 1. Sequence and secondary structure of stem-loop 2 in 5'-UTR of MERS-CoV. The conserved pyrimidine nucleotides are in red. Numbers indicating nucleotide position are according to a MERS-CoV strain HCoV-EMC reference sequence (NCBI Reference Sequence: NC_019843.3).

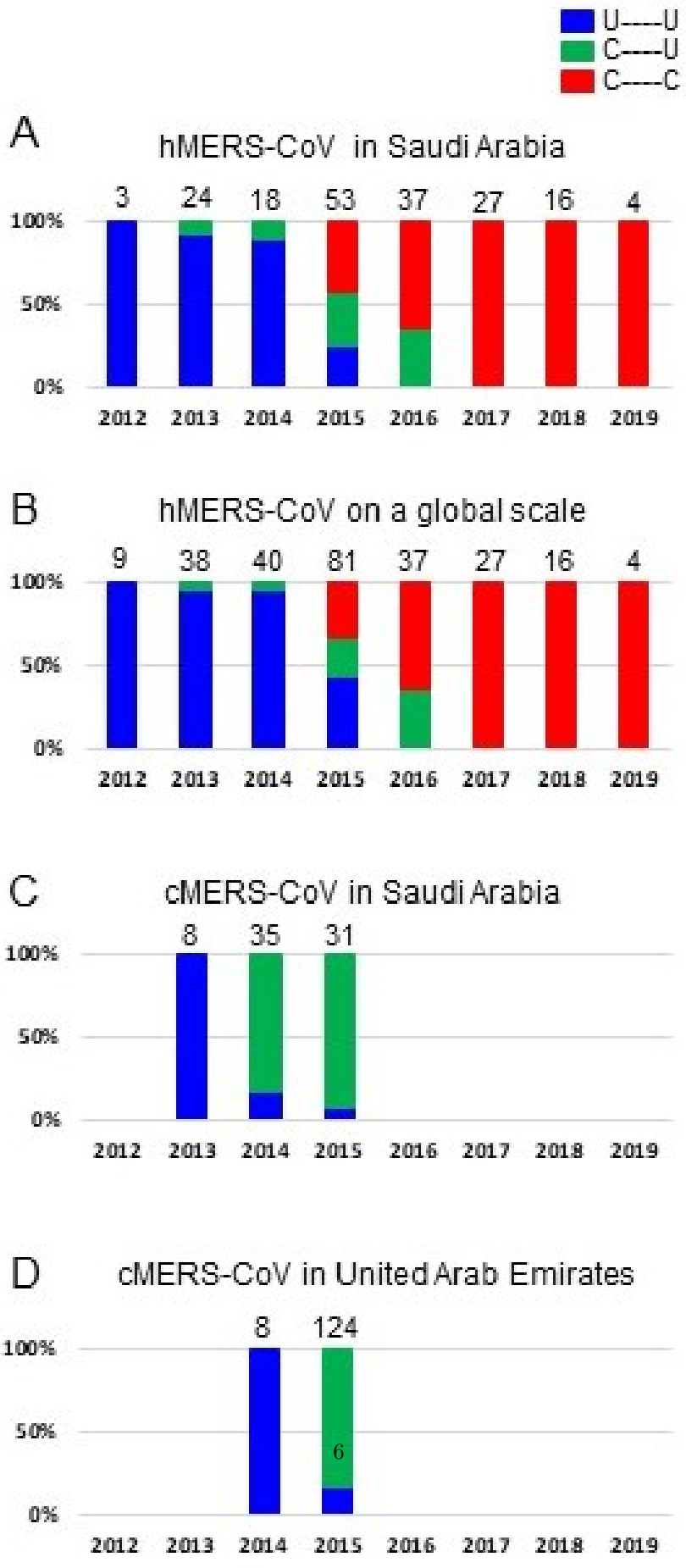

Figure 2. Prevalence of U—-U, C—-U, and C—-C types of MERS-CoV during 2012-2019 in percentage. Numbers on the top of each column indicate the total numbers of complete genome that are available in NCBI database and included in each column. hMERS-CoV, human MERS-CoV; cMERS-CoV, camel or dromedary camel MERS-CoV.