



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.



Short communication

Deletion in the C-terminal region of the envelope glycoprotein in some of the Indian SARS-CoV-2 genome

Ballamoole Krishna Kumar^a, Anusha Rohit^b, Kattapuni Suresh Prithvisagar^a, Praveen Rai^a, Indrani Karunasagar^a, Iddya Karunasagar^{c,*}^a Division of Infectious Diseases, Nitte University Centre for Science Education and Research, Nitte (Deemed to be University), Deralakatte, Mangaluru, 575018, India^b Department of Microbiology, The Madras Medical Mission, 4-A, Dr, Mogappair, Chennai, Tamil Nadu, 600037, India^c Nitte (Deemed to be University), University Enclave, Medical Sciences Complex, Deralakatte, Mangaluru, 575018, India

ARTICLE INFO

Keywords:

COVID-19
SARS-CoV-2
Viral assembly
Genome analysis
E protein

ABSTRACT

The envelope glycoprotein (E) is the smallest structural component of SARS-CoVs; plays an essential role in the viral replication starting from envelope formation to assembly. The *in silico* analysis of 2086 whole genome sequences from India performed in this study provides the first observation on the extensive deletion of amino acid residues in the C-terminal region of the envelope glycoprotein in 34 Indian SARS-CoV-2 genomes. These amino acid deletions map to the homopentameric interface and PDZ binding motif (PBM) present in the C-terminal region of E protein as well as immediately after the reverse primer binding region as per Charité protocol in 26 of these genomes, hence, their detection through RT-qPCR may not be hampered and therefore E gene-based RT-qPCR would still detect these isolates. Eight genomes from the State of Odisha had deletion even in the primer binding site. It is possible that the deletions in the C-terminal region of E protein of these genomes are a result of adapting to a newer geographical area and host. The information on the clinical status was available only for 9 out of 34 cases and these were asymptomatic. However, further studies are indispensable to understand the functional consequences of amino acid deletion in the C terminal region of SARS-CoV-2 envelope protein in the viral pathogenesis and host adaptation.

The coronavirus disease 2019 (COVID-19) caused by SARS-CoV-2 become a pandemic affecting over 200 countries, causing more than 33.8 million cases and associated with over 1.01 million deaths worldwide. The disease spreads mostly through droplets released when an infected person coughs or speaks or sneezes, and the disease spreads rapidly among contacts. SARS-CoV-2 is an enveloped, single-stranded positive-sense RNA virus belonging to the genus β -coronavirus (Lu et al., 2020; Malik et al., 2020). Similar to SARS-CoV and MERS-CoV, the 29.9 kb genome of SARS-CoV-2 encodes four major structural proteins - spike (S), envelope (E), membrane (M), as well as nucleocapsid (N) protein, 16 non-structural proteins (nsp1–16), and five to eight accessory proteins (Wu et al., 2020). Based on the currently available pieces of evidence, it's noteworthy to mention that the severity of COVID-19 differs significantly within populations and geographical locations. Episodes of asymptomatic infection were also reported in more than 80 % of the tested positive COVID-19 cases in India (Acharya and Porwal, 2020). Genomic data on Indian isolates of SARS-CoV-2 is coming from different laboratories, and interim analysis shows the introduction of

this virus to India from multiple sources such as China, Europe, USA, Canada and the Middle East (Potdar et al., 2020; Singh and Sharma, 2020; Somasundaram et al., 2020). Analysis of 361 genomes of SARS-CoV-2 from India revealed that there were 5 clusters, 4 of these being known clades identified by Nextstrain: A2a, A3, B, and B4. 62 % of the genomes belonged to A2a clade, but 29 % belonged to a distinct cluster designated clade I/A3i, not reported outside India (Banu et al., 2020). It has been suggested that the evolution of Clade I/A3i is primarily determined by changes in the structural proteins, N and E, while the evolution in globally predominant A2a clade is determined by changes in Spike (S) and Membrane (M) proteins. Since E gene is the target for amplification in Reverse Transcription Real-Time PCR (RT-qPCR) used for diagnostic purposes in India (Alagarasu et al., 2020), we analyzed the sequence of E genes to see if variations affected the primer binding sites.

The envelope (E) protein of SARS-CoV is the smallest structural component plays an essential role in the viral replication starting from envelope formation, to assembly (Schoeman and Fielding, 2019).

* Corresponding author.

E-mail address: iddya.karunasagar@nitte.edu.in (I. Karunasagar).<https://doi.org/10.1016/j.virusres.2020.198222>

Received 24 June 2020; Received in revised form 21 October 2020; Accepted 29 October 2020

Available online 6 November 2020

0168-1702/© 2020 Elsevier B.V. All rights reserved.

Table 1
Patient characteristics of the COVID-19 cases used for the genome analysis of SARS-CoV-2.

Sl. No	Accession No. (NCBI/GISAID)	Isolate name	Sequencing Technology	Coverage	Geographical Location	Age	Gender	Patient status
1.	NC_045512 - WU	Wuhan-Hu-1	Illumina	NA	China/Wuhan	NA	NA	Ref. seq.
2.	MT012098 - KL	SARS-CoV-2/human/IND/29/2020	Illumina	NA	India/Kerala	NA	NA	NA
3.	EPI_ISL_428482	hCoV-19/India/nimh-0182/2020	Oxford nanopore technology MinION	1,900x	India/Karnataka	26	Male	NA
4.	EPI_ISL_436447	hCoV-19/India/KA-NCDC-4055/2020	Oxford nanopore technology MinION	1,294.84x	India/Karnataka	18	Male	NA
5.	EPI_ISL_428485	hCoV-19/India/KA-nimh-0834/2020	Oxford nanopore technology MinION	700x	India/Karnataka	32	Male	Asymptomatic
6.	EPI_ISL_486405	hCoV-19/India/KA-nimh-14709/2020	Oxford nanopore technology MinION	NA	India/Karnataka	34	Male	Asymptomatic
7.	EPI_ISL_516078	hCoV-19/India/KA-nimh-1600/2020	Oxford nanopore technology MinION	1814.56x	India/Karnataka	21	Female	Released
8.	EPI_ISL_428484	hCoV-19/India/KA-nimh-0351/2020	Oxford nanopore technology MinION	1200x	India/Karnataka	55	Male	NA
9.	EPI_ISL_486402	hCoV-19/India/KA-nimh-11076/2020	Oxford nanopore technology MinION	NA	India/Karnataka	30	Female	Asymptomatic
10.	EPI_ISL_428480	hCoV-19/India/KA-nimh-0116/2020	Oxford nanopore technology MinION	900x	India/Karnataka	28	Female	NA
11.	EPI_ISL_436139	hCoV-19/India/KA-nimh-3970/2020	Oxford nanopore technology MinION	NA	India/Karnataka	20	Male	Symptomatic
12.	EPI_ISL_436140	hCoV-19/India/KA-nimh-4376/2020	Oxford nanopore technology MinION	NA	India/Karnataka	52	Female	Asymptomatic
13.	EPI_ISL_436138	hCoV-19/India/KA-nimh-3952/2020	Oxford nanopore technology MinION	NA	India/Karnataka	38	Male	Asymptomatic
14.	EPI_ISL_428481	hCoV-19/India/KA-nimh-0130/2020	Oxford nanopore technology MinION	3000x	India/Karnataka	55	Male	NA
15.	EPI_ISL_428487	hCoV-19/India/KA-nimh-1071/2020	Oxford nanopore technology MinION	1600x	India/Karnataka	25	Male	NA
16.	EPI_ISL_428486	hCoV-19/India/KA-nimh-0996/2020	Oxford nanopore technology MinION	2000X	India/Karnataka	43	Male	NA
17.	EPI_ISL_428483	hCoV-19/India/KA-nimh-0318/2020	Oxford nanopore technology MinION	4600x	India/Karnataka	19	Male	NA
18.	EPI_ISL_516077	hCoV-19/India/KA-nimh-1453/2020	Oxford nanopore technology MinION	1000.31x	India/Karnataka	23	Male	Released
19.	EPI_ISL_515948	hCoV-19/India/KA-nimh-11074/2020	Oxford nanopore technology MinION	1,510x (average)	India/Karnataka	10	Male	Released
20.	EPI_ISL_515953	hCoV-19/India/KA-nimh-12171/2020	Oxford nanopore technology MinION	726x (average)	India/Karnataka	35	Female	Released
21.	EPI_ISL_436156	hCoV-19/India/KA-nimh-1596/2020	Oxford nanopore technology MinION	NA	India/Karnataka	55	Female	Asymptomatic
22.	EPI_ISL_436141	hCoV-19/India/KA-nimh-4378/2020	Oxford nanopore technology MinION	NA	India/Karnataka	27	Female	Asymptomatic
23.	EPI_ISL_436137	hCoV-19/India/KA-nimh-2873/2020	Oxford nanopore technology MinION	NA	India/Karnataka	28	Female	Asymptomatic
24.	EPI_ISL_436157	hCoV-19/India/KA-nimh-1598/2020	Oxford nanopore technology MinION	NA	India/Karnataka	50	Female	Asymptomatic
25.	EPI_ISL_428479	hCoV-19/India/KA-nimh-0113/2020	Oxford nanopore technology MinION	3000x	India/Karnataka	68	Male	NA
26.	EPI_ISL_528426	hCoV-19/India/MH-IGIB-D2/2020	Oxford nanopore technology MinION	NA	India/Maharashtra	24	Female	NA
27.	EPI_ISL_436438	hCoV-19/India/MH-NCDC-3835/2020	Oxford nanopore technology MinION	1,467.2x	India/Maharashtra	NA	Male	NA
28.	EPI_ISL_436446	hCoV-19/India/MH-NCDC-3985/2020	Oxford nanopore technology MinION	673.53x	India/Maharashtra	25	Male	NA
29.	EPI_ISL_436442	hCoV-19/India/MH-NCDC-3950/2020	Oxford nanopore technology MinION	1,345.3x	India/Maharashtra	31	Male	NA
30.	EPI_ISL_436437	hCoV-19/India/DL-NCDC-3831/2020	Oxford nanopore technology MinION	1,531.56x	India/Delhi	56	Male	NA
31.	EPI_ISL_436443	hCoV-19/India/MP-NCDC-3961/2020	Oxford nanopore technology MinION	1,346.03x	India/Madhya Pradesh	33	Male	NA
32.	EPI_ISL_455773	hCoV-19/India/OR-RMRC22/2020	Illumina	179x	India/Odisha	29	Male	NA
33.	EPI_ISL_463046	hCoV-19/India/OR-ILSCV16495/2020	Illumina	355x	India/Odisha	37	Male	NA
34.	EPI_ISL_463080	hCoV-19/India/OR-ILSCV20598/2020	Illumina	403x	India/Odisha	58	Male	NA
35.	EPI_ISL_463091	hCoV-19/India/OR-ILSCV27412/2020	Illumina	47x	India/Odisha	20	Female	NA
36.	EPI_ISL_463048	hCoV-19/India/OR-ILSCV16508/2020	Illumina	275x	India/Odisha	35	Male	NA
37.	EPI_ISL_463021	hCoV-19/India/OR-ILSCV13695/2020	Illumina	764x	India/Odisha	29	Male	NA

(continued on next page)

Table 1 (continued)

Sl. No	Accession No. (NCBI/GISAID)	Isolate name	Sequencing Technology	Coverage	Geographical Location	Age	Gender	Patient status
38.	EPI_ISL_463041	hCoV-19/India/OR-ILSCV15937/2020	Illumina	638x	India/Odisha	50	Male	NA
39.	EPI_ISL_463093	hCoV-19/India/OR-ILSCV28955/2020	Illumina	84x	India/Odisha	55	Male	NA

NA = Not Available.

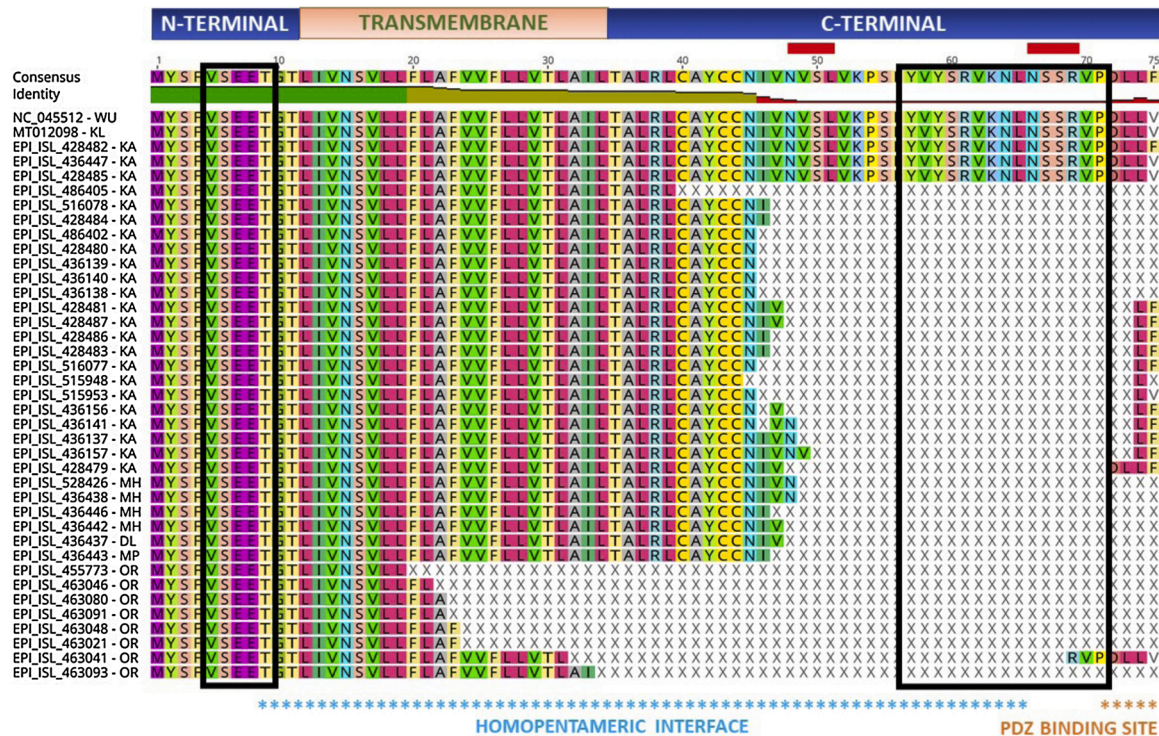


Fig. 1. Representative image showing the amino acid residues and domains in the envelope (E) protein of SARS-CoV-2. Homopentameric interface region and PDZ binding domains are indicated with star marks coloured in blue and orange. The location of N-glycosylation sites are marked with a straight line (—) at the top of the residue. The predicted antigenic determinant regions are highlighted in the box.

Studies have reported that mutation or deletions in the E protein significantly affects the viral maturation, incompetent viral progeny leading to reduced viral titer upon infection (Cohen et al., 2011; Schoeman and Fielding, 2019). In this study nucleotide sequence of E genes from Indian isolates were examined to see if there are any changes in primer binding sites. Further, *in silico* analysis of the Indian SARS-CoV-2 E proteins were performed to investigate the amino acid composition, domains involved in the pathogenicity to obtain better insight into the epidemiology and pathogenicity of Indian SARS-CoV-2.

E gene nucleotide and protein sequences available for the 2086 Indian SARS-CoV-2 genome as of Sep 11th 2020 and Reference sequence were downloaded from the GISAID as well as NCBI virus database and aligned using Geneious Prime. These genome sequences were derived from both illumina sequencing technology with coverage of more than 47 X (47X-764 X) and ARTIC Network Primal Scheme on MinION with coverage of more than 673 X (673 X- 4600X) (Table 1). The detailed information on the sequences used and their metadata is given in Supplementary Table S1. As the success of PCR based molecular diagnostics mainly depends on the efficient primers and/or probes to specifically amplify the target gene, any genetic variations, especially in the 3' primer or probe binding regions, can result in potential mismatches and false-negative results. The primer/probes described in Charité-E protocols (WHO, 2020) were mapped to the multiple sequence alignment of E gene to see if there are any changes in primer binding sites (data not shown).

The analysis revealed that none of the E gene sequences analysed had variations in the primer/probe binding regions except in 8 out of 168 sequences from the State of Odisha where reverse primer binding found to be abolished. Hence such strains may be negative for E gene in RT PCR. This may explain observations from some diagnostic laboratories in India using GeneXpert and Trunat systems that occasional samples can be negative for E gene but positive for N gene or ORF1 gene (Rohit et al., unpublished observations). SARS-CoV-2 E protein has 75 amino acid residues with a predicted molecular weight of 8.3 kDa and has an N-terminal cytoplasmic segment and C-terminal non-cytoplasmic region separated by a hydrophobic transmembrane domain (aa 12-aa 35) as predicted using XtalPred Server (<http://xtalpred.godziklab.org/>) InterPro Scan search (<http://www.ebi.ac.uk/interpro/>). Multiple sequence alignment revealed that amino acid residues of E proteins are highly conserved in all the sequences except that there was a characteristic in-frame deletion of 25–55 amino acid residues in the E protein C-terminal region of the 34 SARS-CoV-2 genomes available from Karnataka (20/128), Maharashtra (4/325), Delhi (1/121), Madhya Pradesh (1/41) and Odisha (8/168) (Fig. 1). These amino acid deletions map to the homopentameric interface and PDZ binding motif (PBM) present in the C-terminal region as well as immediately after the reverse primer binding region as per Charité protocol (WHO, 2020) in 26 genomes (except the ones from the State of Odisha discussed above), hence, their detection through RT-qPCR may not be hampered and therefore E gene based RT-qPCR would still detect these isolates. Among them, 9 individuals

were asymptomatic (Table 1) and the clinical condition of other individuals was not available in the GISAID database. When we analyzed the patient data for these cases, none of them had a travel history to COVID-19 infected countries; hence their exact source of the contract cannot be adequately traced. We also examined the lineage of these sequences using the methodology of Pattabiraman et al. (2020) and as indicated in Table S1, of the 34 sequences with E gene deletions, 15 belonged to 19A, one belonged to 19B, 4 belonged to 20A and 20B each and remaining sequences could not be categorised into any of the clades defined by Nextstrain. Our data suggests that C-terminal deletion in the E-gene of SARS-CoV-2 was spread across different lineages and this deletion event would have occurred independently in different lineages and geographical locations.

Further, sequential B cell epitopes on the E protein was predicted using BepiPred-2.0 (<http://www.cbs.dtu.dk/services/BepiPred/index.php>). This showed the presence of highly conserved antigenic determinant regions in the N-terminal region (SEET) and C-terminal region (YVYSRVKLNLSRVP) of E protein (Fig. 1). It is also important to highlight that; we were unable to map the C terminal antigenic determinant region in those 34 isolates of SARS-CoV-2 E proteins as there was a deletion of 25–55 amino acid residues in its C-terminal region. Based on PROSITE analyses (<https://prosite.expasy.org/>), it was predicted that SARS-CoV-2 E protein has two N-glycosylation sites at N48 and N66 but lacking in all 34 isolates of SARS-CoV-2 which had an extensive deletion/gaps in C-terminal region amino acid residues. Recent study of Sun et al. (2020) reported 12 bp deletions in the E gene of SARS-CoV-2 genome at the position of 26320–26331 and mutant strain had higher spike protein compared to wildtype without any difference in viral titer. In SARS-CoV, PDZ binding motif of E protein is considered to be the major determinant of virulence and known to participate in interactions with several host proteins like syntenin and PALS1 during viral infection (Jimenez-Guardeño et al., 2014; Teoh et al., 2010). SARS-CoVs lacking PBM domain of E protein have abrogated expression of inflammatory cytokines and were attenuated, causing the least damage to lungs without the mortality of the infected animal (Jimenez-Guardeño et al., 2014). DeDiego et al. (2008) showed that elimination of SARS-CoV E protein has the deleterious effect of viral replication where it grew about 100-fold lower compared to the wild type virus in lung cells of the infected mice. In many coronaviruses, E protein is targeted towards the Golgi complex, where it participates in viral assembly and budding. Studies have shown that Golgi complex targeting information is located in the C-terminal region of the E protein and therefore, either truncation of C-terminal region or mutation of conserved amino acid residues in the C-terminal would disrupt the Golgi complex targeting the region of the SARS-CoV E protein leading to crippled viral maturation and incompetent viral progeny (Cohen et al., 2011). It has also been shown that the deletion of N-glycosylation sites in the envelope protein of many viruses had a significant consequence on infectivity potential and antibody-mediated neutralization by the host immune system. It was also observed that mutation of the glycosylation site at N66 resulted in the formation of oligomers of E protein with higher molecular weight, thereby affecting the functionality of E protein in the viral replication (Liao et al., 2006; Schoeman and Fielding, 2019).

The *in silico* analysis performed in this study provides the first observation on the extensive deletion of amino acid residues in the C-terminal region of the envelope glycoprotein in some of the Indian SARS-CoV-2 genomes. It is possible that the deletions in the C-terminal region of E protein of these genomes are a result of adapting to a newer geographical area and host. Sometimes sequencing of samples with low viral titers could lead to assemblies with low coverage and/or spurious gaps in the genome, but this is unlikely in this case since gaps in the same region in 34 isolates from are different geographical location and sequences by different laboratories. It is also possible that these have reduced virulence since nine of the thirty-four individuals from whom isolates were obtained were asymptomatic and clinical information for others were not available. Lack of travel history in the infected

individuals suggests that the virus isolates might be circulating in this region for some time. However, further studies are indispensable to understand the functional consequences of amino acid deletion in the C terminal region of SARS-CoV-2 envelope protein in the viral pathogenesis and host adaptation.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgement

The authors are grateful to Nitte (Deemed to be University) for providing computing infrastructure for the execution of this research. We are also very grateful to all the data contributors of the GISAID Initiative and NCBI Virus database.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.virusres.2020.198222>.

References

- Acharya, R., Porwal, A., 2020. A vulnerability index for the management of and response to the COVID-19 epidemic in India: an ecological study. *Lancet Glob. Health* 8 (9), e1142–51, 1.
- Alagarasu, K., Choudhary, M.L., Lole, K.S., Abraham, P., Potdar, V., Team, N.I.C., 2020. Evaluation of RdRp & ORF-1b-nsp14-based real-time RT-PCR assays for confirmation of SARS-CoV-2 infection: an observational study. *Indian J. Med. Res.* 151, 483–485.
- Banu, S., Jolly, B., Mukherjee, P., Singh, P., Khan, S., Zaveri, L., Shambhavi, S., Gaur, N., Mishra, R.K., Scaria, V., Sowpati, D.T., 2020. A distinct phylogenetic cluster of Indian SARS-CoV-2 isolates. *bioRxiv*. <https://doi.org/10.1101/2020.05.31.126136>.
- Cohen, J.R., Lin, L.D., Machamer, C.E., 2011. Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein. *J. Virol.* 85, 5794–5803.
- DeDiego, M.L., Pewe, L., Alvarez, E., Rejas, M.T., Perlman, S., Enjuanes, L., 2008. Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice. *Virology* 376, 379–389.
- Jimenez-Guardeño, J.M., Nieto-Torres, J.L., DeDiego, M.L., Regla-Nava, J.A., Fernandez-Delgado, R., Castaño-Rodríguez, C., Enjuanes, L., 2014. The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS Pathog.* 10, e1004320.
- Liao, Y., Yuan, Q., Torres, J., Tam, J.P., Liu, D.X., 2006. Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. *Virology* 349, 264–275.
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., 2020. Genomic characterization and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 395, 565–574.
- Malik, Y.S., Sircar, S., Bhat, S., Sharun, K., Dhama, K., Dadar, M., Tiwari, R., Chaicumpa, W., 2020. Emerging novel coronavirus (2019-nCoV)—current scenario, evolutionary perspective based on genome analysis and recent developments. *Vet. Q.* 40, 68–76.
- Pattabiraman, C., Habib, S., Harsha, P.K., Rasheed, R., Reddy, V., Dinesh, P., Damodar, T., Hosallimath, K., George, A.K., Reddy, N.V.K., John, B., Pattanaik, A., Kumar, N., Mani, R.S., Venkataswamy, M.M., Hameed, S.K.S., Kumar, B.G.P., Desai, A., Vasanthapuram, R., 2020. Genomic Epidemiology Reveals Multiple Introductions and Spread of SARS-CoV-2 in the Indian State of Karnataka. <https://doi.org/10.1101/2020.07.10.20150045>.
- Potdar, V., Cherian, S.S., Deshpande, G.R., Ullas, P.T., Yadav, P.D., Choudhary, M.L., Gughe, R., Vipat, V., Jadhav, S., Patil, S., Nyayanit, D., 2020. Genomic analysis of SARS-CoV-2 strains among Indians returning from Italy, Iran & China, & Italian tourists in India. *Indian J. Med. Res.* 151, 255–260.
- Schoeman, D., Fielding, B.C., 2019. Coronavirus envelope protein: current knowledge. *Virology* 16, 1–22.
- Singh, S., Sharma, B.B., 2020. Severe acute respiratory syndrome-coronavirus 2 and novel coronavirus disease 2019: an extraordinary pandemic. *Lung India* 37, 268–271.
- Somasundaram, K., Mondal, M., Lawarde, A., 2020. Genomics of Indian SARS-CoV-2: implications in genetic diversity, possible origin and spread of virus. *medRxiv*. <https://doi.org/10.1101/2020.04.25.20079475>. Jan 1st.
- Sun, Y., Xu, F., An, Q., Chen, C., Yang, Z., Lu, H., Chen, J., Yao, P., Jiang, J., Zhu, H., 2020. A SARS-CoV-2 variant with the 12-bp deletion at E gene. *Emerg. Microbes Infect.* <https://doi.org/10.1080/22221751.2020.1837017>.
- Teoh, K.T., Siu, Y.L., Chan, W.L., Schlüter, M.A., Liu, C.J., Peiris, J.M., Bruzzone, R., Margolis, B., Nal, B., 2010. The SARS coronavirus E protein interacts with PALS1 and

- alters tight junction formation and epithelial morphogenesis. *Mol. Biol. Cell* 201 (21), 3838–3852.
- WHO, 2020. Diagnostic Detection of 2019-nCoV by Real-Time RT-PCR. accessed on Jun 6th 2020. Available from: <https://www.who.int/docs/default-source/coronavirus/e/protocol-v2-1.pdf>.
- Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., Sheng, J., 2020. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host Microbe* 27, 325–328.