Proteomic basis of antigenic shift of Coronavirus from animal to human hosts

Khaneja Poorna, Angel Annette*, Angel Bennet, Joshi Vinod and Buvvaji Mohammed Shareef
Centre of Excellence in Virology and Immunology, Sharda University, Greater Noida, Uttar Pradesh, 201310, INDIA
*annetteangel156@gmail.com

Abstract

Coronaviruses are RNA viruses which were first identified as chicken infections way back in the year 1920 from North America. Of the seven Human Coronavirus species, three viz. SARS-CoV-1, MERS-CoV and SARS-CoV-2 have caused severe diseases. Low titre of neutralizing antibodies was seen in case of SARS-COV-2 as found in the previous infection of 2002-2003 i.e. SARS-CoV-1. When the genome profile of the current SARS-CoV-2 was compared with the older strains, it showed forty percent similarity with the MERS CoV. This indicates that origin of SARS-CoV-2 is basically from the non-human reservoirs viz. bat, camel, bovine, mice, snake etc. We aimed to study the number of nucleotide substitutions throughout the course of evolution in the coronaviruses. Clustal Omega multiple alignment tool was used for studying mutations in nucleotide and amino acid sequences. The structure of spike protein was designed using Swiss modelling and then interactions were studied by PyMOL.

On comparing the results of genome sequences in animal reservoirs, it was seen that in the Receptor Binding Domain (RBD) of the Spike Protein of SARS-CoV-2, number of mutations have taken place enabling the zoonotic strain to find its way to internalize into human cells by using human Angiotensin Converting Enzyme 2 (ACE-2) as host receptor. Open Reading frame (ORF) region of SARS-CoV-2 had shown more number of mutation in coronavirus as compared to viral genome in bats. These mutations in Coronavirus of bats have enabled them as a serious diseases causing pathogens to humans.

Keywords: Coronavirus, Evolution, Zoonotic, Protein, Bioinformatics.

Introduction

Number of zoonotic diseases reported so far have proven to be the deadliest diseases known. The recently emerged SARS-CoV-2 is latest example. This is a RNA virus possessing 96% sequence similarity with Bat Coronavirus[3]. Coronaviruses are known to be disease causing pathogens to both mammals as well as birds. They are composed of a large group of viruses that can also occasionally spread among humans. They belong to Coronaviridae family which are majorly involved in respiratory infections[8] as well as in gastrointestinal infections in domestic and wild animals including pigs, rodents, birds etc[2]. These viruses were first identified as chicken infections way back in the year 1920 from North America while in human population, it first appeared in the year 1960, from England[4].

Historically, in the year 1920, a new group of animal RNA virus was discovered which was morphologically similar to MHV. The virus was similar to corona of sun and was thus named as coronavirus by group of virologists and the name was subsequently accepted for the nomenclature of the virus[8]. The clinical symptoms of the coronavirus were similar to the Infectious Bronchitis Virus (IBV) which was first reported from the chicken. A study done in 1930s concluded that the humans working in the poultry were infected with this disease[8].

Subsequently, Estola, 1920 characterized the disease among poultry workers being caused by coronavirus[7,9,10]. The first human strain of the corona group came from a common cold case using human embryonic tracheal and nasal organ culture and was named B814 in 1960s. Another strain 229E was identified by using human embryo kidney cells. Both B814 and 229E were morphologically similar to IBV[1].

As of today, thirty-nine species of coronavirus including pig, dogs, cats, rodents, cows, horses, camels, beluga whales, birds and bats (Fig.1) have been reported out of which only seven species are affecting humans namely, HCoV 229E, HCoV OC43, SARS-CoV, HCoV NL63, HCoV HKUI, MERS-CoV and SARS-CoV-2[5,6]. Out of these seven, only three viz. SARS-CoV, MERS-CoV and SARS-CoV2 have shown severe clinical severities where SARS-COV-2 resulted in severe global health problems due to exposure to many mutants[11]. Various epidemiological studies have also been undertaken since then for better understanding of transmission, predisposing factors and spectrum of illness caused by this virus[12].

However, the genomic and proteomic basis of transformation of zoonotic to human interactions by coronaviruses has still not been studied. Present study reviews the genomic and proteomic basis of evolution of SARS-CoV-2 from zoonotic to human host manifesting the severity and mortality in later. A study of trends in the evolution of the coronaviruses from causing acute to severe infections was undertaken among animal, bird and human population. The objective of this study is to understand the basis transmission of the virus from zoonotic reservoirs to humans for the long term public health measures.

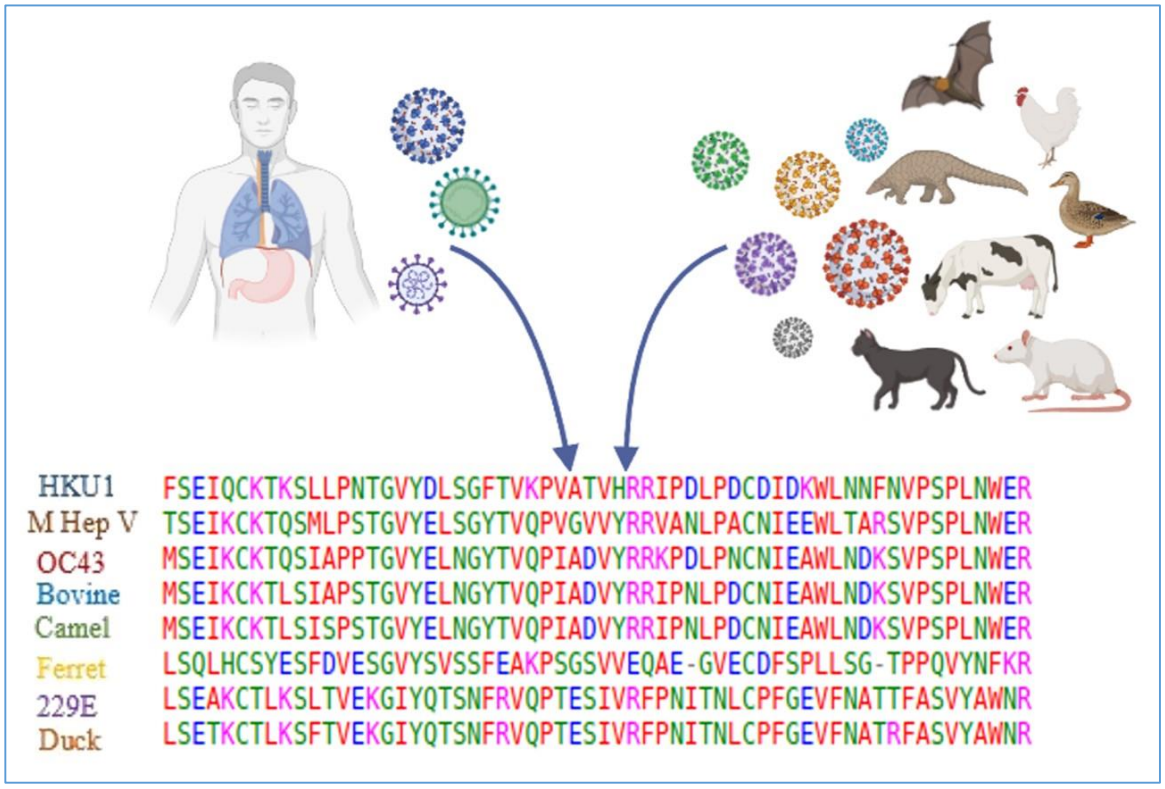

Fig. 1: Graphical representation of transmission of Coronavirus

Material and Methods

The genomic and proteomic sequences of coronavirus: SARS-CoV-1 (NC_002645.1), MERS-CoV (OL622035.1) and SARS-CoV-2 (NC_045512.2) were obtained using National Centre for Biotechnology Information (NCBI) Virus database and their genomic and proteomic comparison was done by using Clustal Omega multiple alignment tool. Also, the genomic and proteomic comparison of bat (NC_048212.1), pangolin (MT121216.1), bovine (UVJ47473.1), civet cat (AY572034.1), camel (ALA50080.1), raccoon dog (QPI18714.1), chicken (AXQ05191.1), duck (YP_009825008.1), murine hepatitis virus (YP_009824982.1), pig (QGV12781.1), feline (NC_002306), ferret (BAS25710.1), HCoV HKU1 (AXT92557.1), HCoV NL63 (AAS58177.1), HCoV OC43 (AMK59677.1), HCoV 229E (QOP39313.1) was done by using Clustal Omega multiple alignment tool to look for the mutations in nucleotide and amino acid sequences.

The molecular weight of these hosts was compared to compare the pathogenicity of the virus. Their furin cleavage site was identified using the ProP – 1.0, DTU Health Tech tool to identify the basic amino acid sequence sites cleaved by furin. The 3D model of the spike protein of these above-mentioned hosts of coronavirus was developed by using Swiss modelling and their interaction sites were studied by PyMOL visualizing tool.

Results

On comparing amino acid compositions of Spike Protein of SARS-CoV-2 and host ACE-2 receptor protein with the alignment sequences of all the hosts, it was observed that there are drastic mutational changes from infectious bronchitis virus to the emergence of latest coronavirus SARS-CoV-2 (Fig. 2). The analysis of amino acid constitution of Spike Protein (SP) of zoonotic and human coronavirus showed 1456 amino acids in SP of Racoon dogs, 1269 amino acids in Bats, 1265 in Pangolin, 1255 in Civet Cat, 1162 amino acids in Chicken and 1273 amino acids in humans. The molecular weight of spike protein when compared between all the hosts showed similar molecular weight falling in range from 129 kDa to 160 kDa (Table 1). The host receptor Angiotensin Converting Enzyme (ACE-2) was composed of 805 amino acids in all the animal hosts (Table 1).

The furin cleavage site score was found to be maximizing in chicken 0.889 followed by HCoV HKU1 0.882, bovine 0.849 and murine hepatitis virus 0.802 which indicates that the virus got transmitted to animal host to human host as already explained by Estola and El Sayed et al[4] (Table 2 and fig. 3). The docking results showed that there are 8 amino acids in spike protein of infectious bronchitis virus interacting with 7 amino acids of human ACE2 which may facilitate the viral entry, of which some residues are located in receptor binding domain (Fig. 4). The polar contacts between the viral spike and host hACE2 are shown in table 3.

Discussion

Bats are largely found worldwide and so they are tolerant to many viral infections and also become reservoir of micro-organisms. The previous studies have reported the zoonotic travel of Coronavirus from bats to humans through

intermediate hosts such as civet cat, camels etc[5]. Receptor binding domain of spike protein is the main site interacting with host cell receptor to enter the cell along with its co-receptors. ACE2 regulates the body homeostasis involving in RAAS pathway and facilitates the entry of virus consisting of 805 amino acids.

Fig. 2: Alignment sequences of different host of Coronavirus

Table 1
Comparison of amino acids and molecular weight of Spike Protein (SP) among different hosts

Hosts	Number of Amino Acids in SP of Coronavirus	Molecular Weight (KDa)
Chicken (Avian)	1169	129.19
Pangolin	1265	140.26
Bovine	1363	140.26
Civet Cat	1255	139.10
Ferret	1435	157.45
Feline	1452	160.47
HCoV HKU1	1356	151.76
HCoV NL63	1356	149.85
HCoV OC43	1359	150.59
HCoV 229E	1173	128.46
SARS-CoV-2	1273	141.17
MERS-CoV	1353	149.39
Duck	1191	130.90
Murine Hepatitis Virus	1324	145.93
Bat	1269	139.70
Sus Scrofa	1386	151.71
Camel	1366	150.76
Raccoon Dog	1456	161.05

Table 2
Furin Cleavage sites in Spike protein among different hosts

Host	Position	Sequence	Score
Chicken (Avian)	544	GSRSRR | SV	0.889
Bovine	768	TKRRSRR | SI	0.849
Camel	771	IDRRSRR | AI	0.718
Duck	1189	EQIRPKK | FV	0.535
Feline	-	-	-
Ferret	-	-	-
HCoV 229E	-	-	-
HCoV NL63	863	LPQRNIR | SS	0.519
HCoV OC43	764	KTRRSRR | AI	0.720
HCoV HKU1	759	SSSRRKR | RS	0.672
	760	SSRRKRR | SI	0.882
MERS CoV	751	LTPRSVR | SV	0.563
	887	TGSRSAR | SA	0.707
	1113	VKAQSKR | SG	0.512
Murine Hepatitis Virus	717	KSRRAHR | SV	0.802
Pangolin	-	-	-
Sus Scrofa	-	-	-
Raccoon Dog	965	NSRRKYR | SA	0.693
Wuhan SARS-CoV-2	685	NSPRRAR | SV	0.620

Table 3
Interaction sites between Human ACE2 and Infectious Bronchitis Virus Spike glycoprotein

hACE2	Viral Spike Glycoprotein
600K	118D
258P	71H
257S	73G
615D	77N
735N	252K
735N	431C
734P	512R
742W	511S

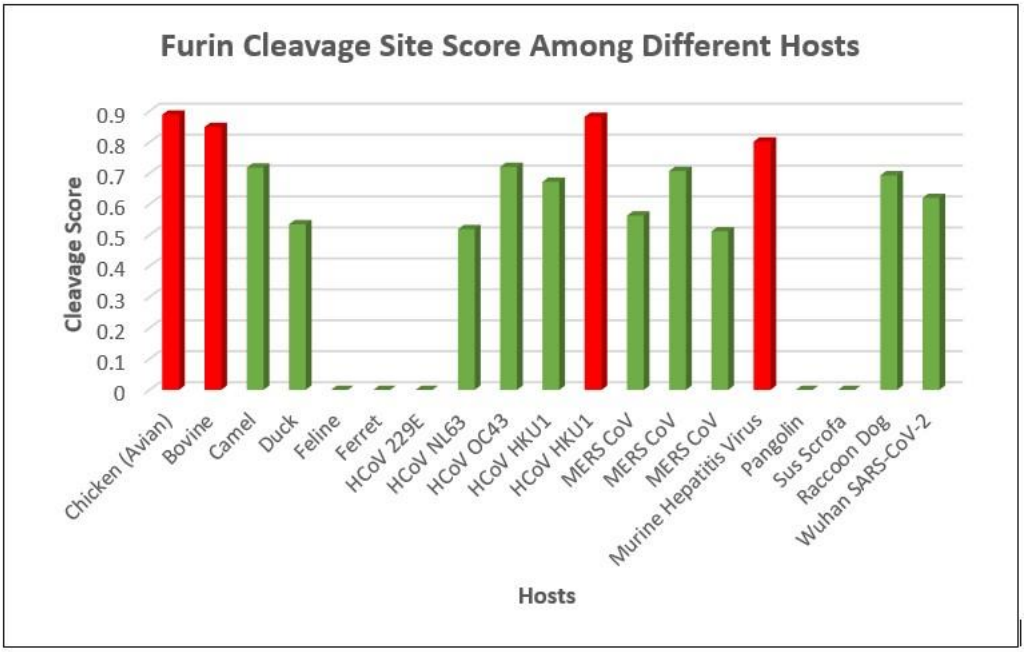

Fig. 3: Furin cleavage site among different host of SARS-CoV-2

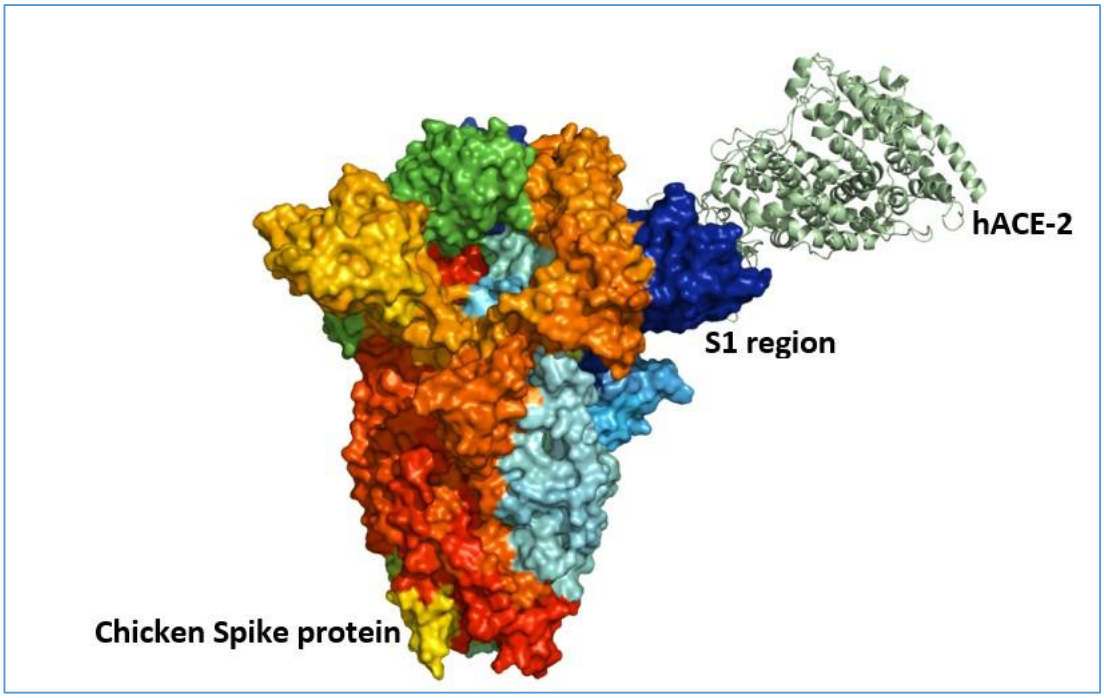

Fig. 4: Docking result of spike glycoprotein of Infectious Bronchitis Virus and human ACE2 receptor

The S protein in case of SARS-CoV-2 virus is much bigger than in each of the host because this is an enzymatic protein and will get digested and simplified by ACE2 which is an enzyme which will fasten the reaction of simplification. Furin is a transmembrane protein involved in cleavage of basic amino acids present in the S1 region of viral spike glycoprotein. The cleavage of S1 region enables the exposure of Fusion Protein (FP) which facilitates the release of viral genome into host cell cytoplasm. If we see the evolution of the viruses from the mammals to the current and when the amino acid in the host is fixed, it means that the virus has made its coating much more robust as compared in terms of number of amino acids. We therefore hypothesize that people living in close proximity with animals are more prone to get infected with the virus. The virus is not much pathogenic to animals but is life-threatening to humans.

Conclusion

Our present knowledge on all the existing viruses teach us that these pathogens erupt from the common belts where the fringes of forest and human habitations meet. The close and consistent sharing of man and forest animals provide sylvatic birth to such infections and here lies our primary priorities

of research. Human habitations in each country are surrounded by the fauna and flora specific to that particular country. Each of country therefore represents a unique disease ecological zone with merging belts of human and forest borders containing their own flora and fauna.

A pro-active and comprehensive inter-sectoral research is thus needed for each country to undertake pathogen and sero-surveillance of man and animals in joint bands of forest and human habitations endings. A robust virological surveillance is needed to generate a predictive data base for possibilities of future epidemics and pandemics.

Acknowledgement

We thank Indian Council of Medical Research, India for providing us funds to do the study (grant No.: 2021-9858).

References

1. Al-Imam A., Khalid U. and Al-Doori H., Clustering analysis of coronavirus disease 2019 pandemic, *Asian J. Med. Sc*., **12**, 108-113 **(2021)**

2. AL-Khikani F.H.O., Surveillance 2019 novel coronavirus (COVID-19) spreading: Is a terrifying pandemic outbreak is soon?, *Biomed. Biotech. Res. J.*, **4(1)**, 81-82 **(2020)**

3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, *Nature Microbio.*, **5**, 536–544 **(2020)**

4. El-Sayed A. and Kamel M., Coronaviruses in humans and animals: the role of bats in viral evolution, *Environ. Sc. Pollut. Res.*, **28**, 19589-19600 **(2021)**

5. Fenton M.B., Mubareka S., Tsang S.M., Simmons N.B. and Becker D.J., COVID-19 and threats to bats, *FACETS*, **5**, 349–352 **(2020)**

6. Hu B., Guo H. and Zhou P., Characteristics of SARS-CoV-2 and COVID-19, *Nature Rev. Microbio.*, **19**, 41-154 **(2021)**

7. Knapp A., The Secret History of the First Coronavirus, Forbes, https://www.forbes.com/sites/alexknapp/2020/04/11/the-secret-history-of-the-first-coronavirus-229e/ **(2020)**

8. Koppolu P., Genady E.M., Albdeirat L.M., Sebai F.A., Alrashdi D.M., Lingam A.S.R., Alsada F.A., Al-Khalifa F.I. and Abdelrahim R.K., Association between severity of COVID-19, Periodontal health and disease in Riyadh subpopulation, *Internat. J. Mycobacterio.*, **12(1)**, 33-37 **(2023)**

9. Lalchhandama K., The chronicles of coronaviruses: the bronchitis, the hepatitis and the common cold, *Sc. Vision*, **20(1)**, 43–53 **(2020)**

10. Mahase E., Covid-19: Coronavirus was first described in The BMJ in 1965, *BMJ* (Clinical Research Ed.), **369**, m1547 **(2020)**

11. Mortaz E.A. and Ian M.I., Immune System Diversity Against SARS-CoV-2 Infection and Vaccines, *J. Prevent. Diag. Treat. Strat. Med*., **1(3)**, 153-155 **(2022)**

12. Shrivastava S., Shrivastava R.B.L. and Saurabh P., Coronavirus Disease 2019 Pandemic: Envisaging Capacity Building in the Field of Research to Accomplish Effective Control, *Biomed. Biotech. Res. J.*, **4(3)**, 266-267 **(2020)**.

(Received 29th August 2024, accepted 14th October 2024)