ACTA SCIENTIFIC MICROBIOLOGY (ISSN: 2581-3226)

Volume 3 Issue 10 October 2020

Mini Review

An Insight into the Coronavirus

Priyanka Kulkarni and Jayant Londhe*

Department of Biotechnology, Sinhgad College of Science, Ambegaon, Pune, India

***Corresponding Author:** Jayant Londhe, Assistant Professor, Department of Biotechnology, Sinhgad College of Science, Ambegaon, Pune, India.

Received: August 08, 2020

Published: September 24, 2020

Abstract

The Coronavirus outbreak has brought the Science fraternity to put some light on the history and evolution of the virus and also the design of the effective drug and its potential target. Coronavirus consists of a positive strand RNA genome and four characteristic structural proteins, active and viable below 56^0C. Surprisingly, all the strains of Coronaviruses are found to be residing in bats for long period of time along with the Malayan pangolins, a group of mammals. The SARS-CoV-2 has significant similarities with SARS-CoV and SARsSr- RaTGL3 in the sequences of conserved non-structural proteins and exhibit 96.2% genome similarity with SARsSr- RaTGL3 virus. However, a unique feature of SARS-CoV2 is the presence of the four amino acid residue sequence at S1/S2 junction which is required to be cleaved by Furin-like enzymes for efficient membrane fusion of host and the virus. This distinguishing feature gives a reason to become a promising factor for drug target studies. However, the infectious diseases could be effectively controlled by vaccine. Hence, there are some promising candidate vaccines entered into Phase III of Clinical trials designed by Sinovac, Moderna Inc and AstraZeneca/Oxford University in collaboration with Serum Institute of India.

Keywords: Coronavirus; SARS-CoV2; SARsSr- RaTGL3; Spike Proteins; Drug Target; Vaccine

Introduction

The study of Coronavirus has become essential to focus on potential treatment and vaccine design to make deterioration in the pandemic. The Coronavirus outbreak has created a havoc which brings us to the study of origin, the evolution and the structure of the virus. The first detected Coronavirus in 1930's was the avian Infectious bronchitis virus (IBV) [3]. Coronavirus belongs to Family *Coronaviridae*, subfamily Coronvirinae and the order *Nidovirales* [5,8] which also includes *Arteriviridae* family. The Coronaviridae family consists of two subfamilies Letovirinae and Orthocoronavirinae. The subfamily Orthocoronavirinae consists of four genera- Alphacoronaviruses, Betacoronaviruses, Gammacoronaviruses and Deltaviruses [1,5]. Alphacoronaviruses and Betacoronaviruses are known to be infecting the mammals, whereas, gammacoronaviruses and deltacoronaviruses infect the avians [5]. The coronavirus despite of being detected in 1930's, they were considered as a serious pathogen in 2002 outbreak in Guangdong province in China. There were no reports on presence of Human coronavirus causing serious illnesses until the SARS-CoV causing serious respiratory syndrome. Until then, there were findings of two human coronaviruses OC43 and 229E causing common cold [15]. The coronavirus since then have been found to be thriving in bats as their natural hosts, whereas transmission to humans could be through the immediate hosts like mammals [16]. Bats being the widely known host of the coronaviruses are predicted to be holding an integral position in the evolution of coronaviruses. It has been studied that a considerable percentage of bats from diverse geographical locations possess different species of coronaviruses [14]. It has also been observed that the same species of bats possess same species of coronavirus enhancing the role on bats in the importance of survival of coronaviruses [14].

Molecular structure of coronavirus

The coronaviruses are enveloped viruses ranging from 60 to 220 nm and comprising of positive sense RNA strand [13]. The virus is known to be having four characteristic proteins building up the structure of the coronavirus. These proteins abbreviated as (N) for nucleoproteins, (S) for spike proteins, (M) for membrane and (E) for envelope proteins are found to be the diverse structural proteins categorising the coronavirus species in four major groups which are Group 1 coronaviruses, Severe acute respiratory syndrome (SARS) and SARS-like coronaviruses and independent group of bat coronaviruses [14]. The Spike glycoproteins characteristically club-shaped are arranged systematically for the host cell attachment by membrane fusion or endocytosis. The N- terminal of these glycoproteins contains a signal sequence for endoplasmic reticulum. In some species of coronaviruses, these S proteins are cleaved by proteases in two domains S1 and S2. S1 is the cell attachment domain and S2 is a stalk like structure attached to S1. The M proteins are transmembrane proteins that do not necessarily contain a signal sequence but are targeted to ER similar to S proteins and contain glycosylated N-terminal and a large C-terminal. The Membrane proteins define the shape of the viruses. E proteins are present in small quantities of which C- terminal possess an ion channel activity. The ion channel activity is believed to be involved in pathogenicity of the coronaviruses. N proteins are made up of two domains which are C-terminal domain (CTD) and N-terminal domain (NTD). Both the domains are required for efficient binding to RNA. N proteins are highly phosphorylated which is essential for particular recognition of viral RNA for attachment [7]. Along with these structural proteins defining the morphology of the coronaviruses, there are certain accessory proteins and non- structural proteins. The characteristic feature of coronaviruses is that having a nested set of subgenomic mRNA's which are extended towards the 5′ direction [13].

Coronavirus possesses positive strand RNA as their genome sized ~30 Kb. The genome consists of a 5' cap and a 3' polyadenylated tail. The 5' end of the genome contains the leader sequence and untranslated region (UTR) required for replication and transcription. The 3' UTR end of the genome contains the structure required for replication and production of the virus. Additionally, every structural and accessory gene contains a regulatory sequence at the initiation to regulate the transcription process. Hence, the structure of the genome of Coronaviruses makes it no less than mRNA. It consists of the replicase polyprotein gene initiating the RNA replication process within the host cell. The four structural genes are interfered with the genes of accessory proteins and non-structural genes in regular intervals [7].

Temperature tolerance of coronavirus

The Coronavirus outbreak has been efficiently lethal due to its incredible ability of survival in large range of temperatures. It has been declared as Pandemic as it has conquered majority of the countries and has shown its effect exponentially. The structure of coronavirus itself has been very unique due to the presence of helically arranged nucleocapsid and its symmetry as this structural arrangement is usually observed in negative sense RNA viruses [7]. Therefore, it is significant that the temperature tolerance of Coronaviruses is studied and made some revealing conclusions. According to a study, the suitable temperature for the survival of these viruses has been room temperature and lower temperatures with lower humidity. It has also been studied that, the virus can be effectively inactivated at higher temperatures i.e. at 56^0C with higher humidity i.e. upto 95%. It is well known that viruses are obligatory parasites, they need a host for their replication and amplification, however, in case of Coronaviruses, it has been seen that they survive in dried form for upto five days in room temperatures on the environmental surfaces. This could be a plausible reason of effective transmission of Coronaviruses at certain places like hospitals. Coronaviruses are transmitted primarily by respiratory droplets; however, they can be settled on the surfaces maintaining their viability which is as lethal as the respiratory droplets [4].

Replication of coronavirus

The spike glycoproteins (One of the structural proteins) specifically bind to angiotensin converting enzyme 2 (ACE2) [5] receptor with the penetration of the viral genome into a mammalian cell. The genome then binds with the ribosome translating polymerase to initiate the replication of the genome. The RNA replication leads to a daughter strand where the nested subgenomic transcription initiates. Subgenomic transcription is nothing but the transcription commenced from the reverse direction, i.e. from 3′ to 5′ end. The subgenomic mRNA's are found to be much smaller than a normal transcribed mRNA. These subgenomic mRNA's consists the information of the major structural proteins. The translation of M, E and S proteins occurs at endoplasmic reticulum whereas the translation of N proteins occurs at cytoplasm. The proteins are assembled together with the genome forming a new viral particle which is eventually released from the cell.

Structural and genomic similarities of SARS-Cov-2 to other members of the family

SARS-CoV is reported to be the cause of the pandemic in the year 2002 - 2003. Focusing on the evolution and emergence of SARS-Cov-2 (Novel Coronavirus), it has been reported that it has similarities with SARS-Cov and SARsSr- RaTGL3 viruses. It has been widely known that the SARS-Cov-2 may have been resided in the bats for long period of time, although the immediate host is unknown. The comparative analysis of the whole genome of the bats coronaviruses and SARS-Cov-2 shows 96% of identity. Moreover, the sequences of the seven conserved non-structural proteins puts SARS-Cov-2 in the *SARSSr-Cov* species.

Focusing on the structural and genomic similarities between SARS-Cov-2, SARS-CoV and RaTGL3 viruses, it has been seen that SARS-Cov-2 and SARS-Cov has less than 80% of nucleotide sequence similarity. However, the complete phylogenetic analysis exhibited that the gene sequence for Spike protein (S gene) and RNA dependent RNA polymerase gene shows highest similarities with Bat RaTGL3 coronavirus. The S gene does not show a considerable match (< 75% similarity) with any of the SARS-CoV viruses but with Bat RaTGL3 virus. Overall, the whole genome sequence analysis of SARS-Cov-2 shows about 96.2% similarities with RaTGL3 virus [17].

On the other hand, Malayan Pangolins are also proved to be the reservoirs of Coronaviruses. The studies say that several species of coronaviruses found in the samples of Lungs, intestines and blood of Malayan Pangolins show identity to SARS-Cov-2. Along with these samples, the scale samples pangolins showed the presence of coronavirus designated as GS/P2S showed 72% similarities with the whole genome of SARS-Cov-2. Likewise, another isolate from ling samples of Pangolins showed 86.3% of identity with the SARS-Cov-2 genome. The receptor binding domains (RBD) of the pangolin coronaviruses and SARS-Cov-2 have higher similarities in their sequences than Bat RaTGL3 coronaviruses where pangolin coronaviruses share five identical amino acids with SARS-Cov-2 and RaTGL3 shares one identical amino acid [10].

Possible target for drug discovery and vaccine development

In order to put an end to the current pandemic of SARS-Cov-2 caused by human-to-human contact, the search for possible target for drug discovery is imperative. Since, the replication cycle of the SARS-Cov-2 and the general steps of replication of any virus is known, the target identification for the drug delivery can be quite clear. The spike proteins present on the viral membranes involved in receptor binding, critical for viral genome entry into the host can be a promising target to design an efficient drug system. The spike glycoproteins as studied in SARS-Cov are trimeric and divided into an ectodomain, a transmembrane anchor and an intracellular tail. The ectodomain is subdivided into S1 domain (N-terminus) and S2 domain (C-terminus). The S1 domain is known to be receptor recognition and binding domain and S2 is involved in viral and host membrane fusion for smooth entry of the viral genome into the host cell. The S1 domain is further subdivided into a core structure composed of five strands of antiparallel B sheets and a receptor binding motif, a slight concave structure composed of two strands of antiparallel B sheets forming an elevation by loops [11]. SARS-Cov-2 is differentiated from all its closely related members by the presence of three insertions of short stretch of sequence in S1 (N-terminal domain) and variations in four key amino acid residues in the Receptor binding motif [10,17]. This distinguishing sequence of four amino acid residues is a target for furin like enzymes for cleavage. Furin, is one of the nine proteases belonging to proprotein convertase (PC) Family. Furin enzyme is located on the membranes of Golgi apparatus and expressed at very low levels [9]. After extensive studies on SARS-Cov-2, it has been proved that the cleaving event at this site of distinct four amino acid sequence is vital for activation and membrane fusion of the virus with the membranes of host cell. Along with Furin proteases, PC1, another member of PC family, Trypsin proteases and Cathepsins are observed to be involved in the cleaving event. Another distinguishing feature of SARS-Cov-2 is the presence of Proline residue at the start of the distinct four amino acid residue. This amino acid residue as a part of S1/S2 cleavage junction holds a significance in such a way that it creates a turn in the cleavage loop which has an effect on the accessibility of the cleavage loop to the proteases for effective cleavage of the junction and activation [9]. Additionally, this Proline residue is thought to be functioning in addition of O linked glycans to the adjacent residues. These O linked glycans shield the possible epitopes preventing an immune response, thereby increasing the virulence [9]. Hence, the presence of these unique features at the very site of activation and fusion of the membranes for effective penetration of the viral genome makes them a possible a drug target. The drug targeted to the stretch of four amino acid residues at S1/S2 junction, the initiating proline residue and the Furin proteases can be considered as promising candidates for vaccine development or drug target.

The hope of deterioration of the SARS-Cov-2 pandemic is the vaccine development. To date, there are several promising vaccine candidates that have entered the human trials. Likewise, there were several vaccines designed and assessed in response to SARS-CoV pandemic in the year 2002-2003. That included live attenuated virus vaccine, inactivated virus vaccine, virus-like particles, DNA, recombinant viral vectors. Out of these, Inactivated, live attenuated and DNA based vaccine were held to be promising and effective. Similarly, the number of vaccine which are under trials against SARS-CoV-2 are based on live attenuated, inactivated whole virus, DNA based vaccine, viral-like particles, Protein subunits, Replicating viral vectors, RNA-based vaccine and non-viral replicating vectors [12]. Table 1 gives general information of the vaccines for SARS-Cov-2 under study as of 28 July, 2020 [2,6,12,16].

Vaccine Platform	Company/Research Institute/University	Stage
Inactivated Virus Vaccine	Sinovac	Clinical Trial (Phase III)
Live attenuated virus vaccine	Codagenix Inc. In collaboration with Serum Institute of India	Preclinical Phase
Subunit Vaccine using S protein	University of Queensland	Clinical trial (Phase I)
Subunit Vaccine using S protein	Clover Biopharmaceuticals	Clinical Trial (Phase I)
mRNA Vaccine	CureVac AG	Clinical trial (Phase I)
mRNA Vaccine	Stermirna Therapeutics, Tongji University, CanSino Biologics Inc, Fudan University collaboration with Shanghai Jiaotong University	Preclinical phase
mRNA Vaccine	Moderna Inc.	Clinical trial (Phase III)
DNA Vaccine	LineaRx, in collaboration with Takis Biotech	Preclinical phase
DNA Vaccine	Inovio Pharmaceuticals	Clinical trial (Phase I/II)
Replicating Live Vector Vaccine	Tonix Pharmaceuticals	Preclinical
DNA Plasmid Vaccine	Zydus Cadila Healthcare limited	Clinical trials (Phase I/II)
Non Replicating Viral Vector	Oxford University/AstraZeneca collaborate with Serum Institute of India	Clinical trial (Phase III)
Non replicating Viral vector	Bharat Biotech in Collaboration with Indian Council of Medical Research (ICMR)	Clinical phase (Phase II)

Table 1: Ongoing studies for vaccine development against SARS-CoV-2.

Conclusion

According to a study, the suitable temperature for the survival of these viruses has been room temperature and lower temperatures with lower humidity. It has also been studied that, the virus can be effectively inactivated at higher temperatures i.e. at 56^0C with higher humidity i.e. upto 95%. It is well known that viruses are obligatory parasites, they need a host for their replication and amplification, however, in case of Coronaviruses, it has been seen that they survive in dried form for upto five days in room temperatures on the environmental surfaces. This could be a plausible reason of effective transmission of Coronaviruses at certain places like hospitals. Coronaviruses are transmitted primarily by respiratory droplets; however, they can be settled on the surfaces maintaining their viability which is as lethal as the respiratory droplets

Bibliography

1. Banerjee., *et al*. "Bats and Coronaviruses". *Viruses* 11.1 (2019): 1-15.

2. Businesstoday.in "Coronavirus update: 7 Indian firms in race to find COVID-19 vaccine". Business Today (2020).

3. Cavanagh D. "Coronaviridae: a review of coronaviruses and toroviruses". Coronaviruses with Special Emphasis on First Insights Concerning SARS, edited by A. Schmidt, M.H. Wolff and O. Weber (2005): 1-55.

4. Chan K., *et al*. "The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus". *Advances in Virology* (2011): 7.

5. Cui J., *et al*. "Origin and evolution of pathogenic coronaviruses". *Nature Reviews Microbiology* 17 (2019): 181-192.

6. DRAFT Landscape of COVID-19 Candidate Vaccines. World Health Organization (2020).

7. Fehr A and Perlman S. "Coronaviruses: An Overview of Their Replication and Pathogenesis". Coronaviruses: Methods and Protocols Methods in Molecular Biology, edited by Helena Maier, Erica Bickerton Paul Britton (2015): 1-23.

8. Hiscox J., *et al*. "The Coronavirus Infectious Bronchitis Virus Nucleoprotein Localizes to the Nucleolus". *Journal of Virology* 75.1 (2000): 506-512.

9. Jaimes J., *et al*. "Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of Novel S1/S2 Site". *Science* 23.6 (2020): 101212.

10. Lam T., *et al*. "Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins". *Nature* 583 (2020): 282-285.

11. Li F. "Structure, Function, and Evolution of Coronavirus Spike Proteins". *Annual Reviews of Virology* 3 (2016): 237-261.

12. Padron- Regalado E. "Vaccines for SARS-CoV-2: Lessons from Other Coronavirus Strains". *Infectious Diseases and Therapy* 9 (2020): 255-274.

13. Sidell S., *et al*. "Coronaviridae". *Intervirology* 20 (1983): 181-189.

14. Tang X., *et al*. "Prevalence and Genetic Diversity of Coronaviruses in Bats from China". *Journal of Virology* 80.15 (2006): 7481-7490.

15. Weiss S and Navas-Martin S. "Coronavirus Pathogenesis and the Emerging Pathogen Severe Acute Respiratory Syndrome Coronavirus". *Microbiology and molecular Biology Reviews* 69.4 (2005): 635-644.

16. Zhang J., *et al*. "Progress and Prospects on Vaccine Development against SARS-CoV-2". *Vaccines* 8. 2 (2020): 1-12.

17. Zhou P., *et al*. "A pneumonia outbreak associated with a new coronavirus of probable bat origin". *Nature* 579 (2020): 270-273.