wjpmr, 2021,7(4), 377 – 385.

WORLD JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.wjpmr.com

SJIF Impact Factor: 5.922

Review Article

ISSN 2455-3301

WJPMR

NATURE OF SECRET CODING IN CORONAVIRUS

Arunava Chandra Chandra*, Dr. Dhrubo Jyoti Sen and Dr. Beduin Mahanti

Department of Pharmaceutical Chemistry, School of Pharmacy, Techno India University, Salt Lake City, Sector-V, EM-4, Kolkata-700091, West Bengal, India.

***Corresponding Author: Arunava Chandra Chandra**
Department of Pharmaceutical Chemistry, School of Pharmacy, Techno India University, Salt Lake City, Sector-V, EM-4, Kolkata-700091, West Bengal, India.

Article Received on 19/02/2021 **Article Revised on 11/03/2021** **Article Accepted on 31/03/2021**

ABSTRACT

COVID-19 is a novel coronavirus with an outbreak of unusual viral pneumonia in Wuhan, China, and then pandemic. Based on its phylogenetic relationships and genomic structures the COVID-19 belongs to genera Betacoronavirus. Human Betacoronaviruses (SARS-CoV-2, SARS-CoV, and MERS-CoV) have many similarities, but also have differences in their genomic and phenotypic structure that can influence their pathogenesis. Coronavirus genome replication and transcription take place at cytoplasmic membranes and involve coordinated processes of both continuous and discontinuous RNA synthesis that are mediated by the viral replicase, a huge protein complex encoded by the 20-kb replicase gene. The replicase complex is believed to be comprised of up to 16 viral subunits and a number of cellular proteins. Besides RNA-dependent RNA polymerase, RNA helicase, and protease activities, which are common to RNA viruses, the coronavirus replicase was recently predicted to employ a variety of RNA processing enzymes that are not (or extremely rarely) found in other RNA viruses and include putative sequence-specific endoribonuclease, 3′-to-5′ exoribonuclease, 2′-O-ribose methyltransferase, ADP ribose 1′-phosphatase and, in a subset of group 2 coronaviruses, cyclic phosphodiesterase activities. Although many molecular details of the coronavirus life cycle remain to be investigated, the available information suggests that these viruses and their distant nidovirus relatives employ a unique collection of enzymatic activities and other protein functions to synthesize a set of 5′-leader-containing subgenomic mRNAs and to replicate the largest RNA virus genomes currently known.

KEYWORDS: Phylogenetic, Betacoronavirus, Human Betacoronaviruses, Genome replication, RNA synthesis, Pathogenesis, Nidovirus.

INTRODUCTION

Coronaviruses are involved in human and vertebrate's diseases. Coronaviruses are members of the subfamily Coronavirinae in the family Coronaviridae and the order Nidovirales. COVID-19 began as an epidemic in China, before making its way around the world in a matter of months and becoming a pandemic. But epidemics don't always become pandemics, and it's not always a fast or clear transition. For example, HIV was considered an epidemic in West Africa for decades before becoming a pandemic in the late 20th century. Now, thanks to advances in modern medicine.[1,2] HIV is considered endemic, which means the rate of the disease is stable and predictable among certain populations, according to the American Medical Association. Based on its phylogenetic relationships and genomic structures the COVID-19 belongs to genera Betacoronavirus which has a close similarity of the sequences of COVID-19 to that of severe acute respiratory syndrome-related coronaviruses (SARSr-CoV) and the virus uses ACE2 as the entry receptor-like SARS-CoV. These similarities of the SARS-CoV-2 to the one that caused the SARS outbreak (SARS-CoVs) the Coronavirus Study Group of the International Committee on Taxonomy of Viruses termed the virus as SARS-CoV-2. The sporadic emergence and outbreaks of new types of CoVs remind us that CoVs are a severe global health threat. It is highly likely that new CoV outbreaks are unavoidable in the future due to changes of the climate and ecology, and the increased interactions of human with animals. Thus, there is an urgent need to develop effective therapies and vaccines against CoVs.

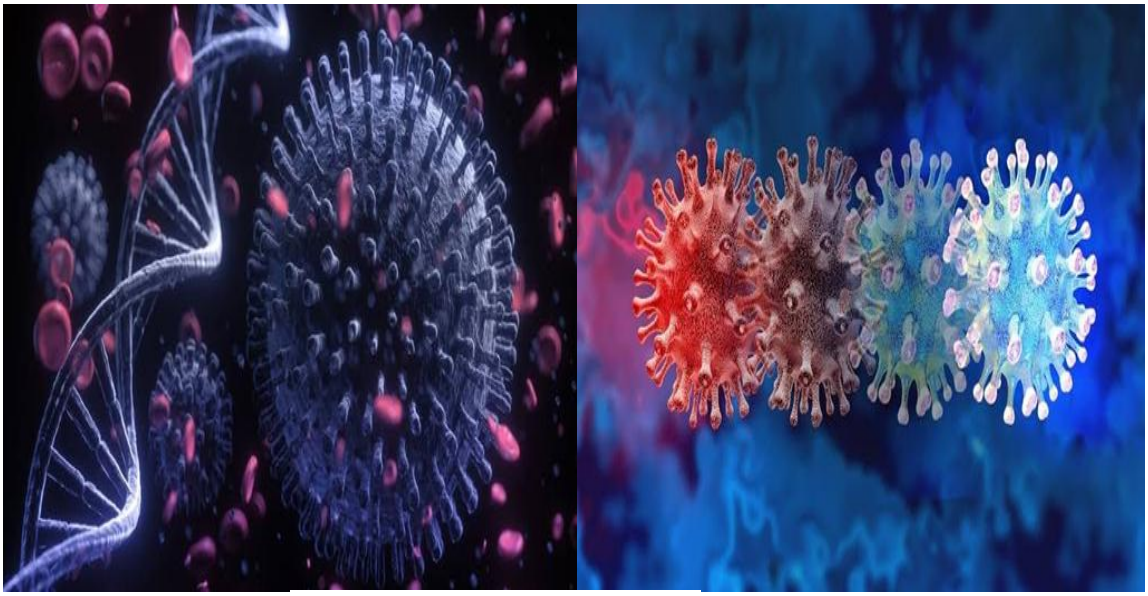

Figure-1: Genome & Replication of SARS-CoV-2.

Coronavirus genome structure and life cycle: COVID-19 is a spherical or pleomorphic enveloped particle containing single-stranded (positive-sense) RNA associated with a nucleoprotein within a capsid comprised of matrix protein. The envelope bears club-shaped glycoprotein projections. Some coronaviruses also contain a hem agglutinin-esterase protein (HE).[3,4]

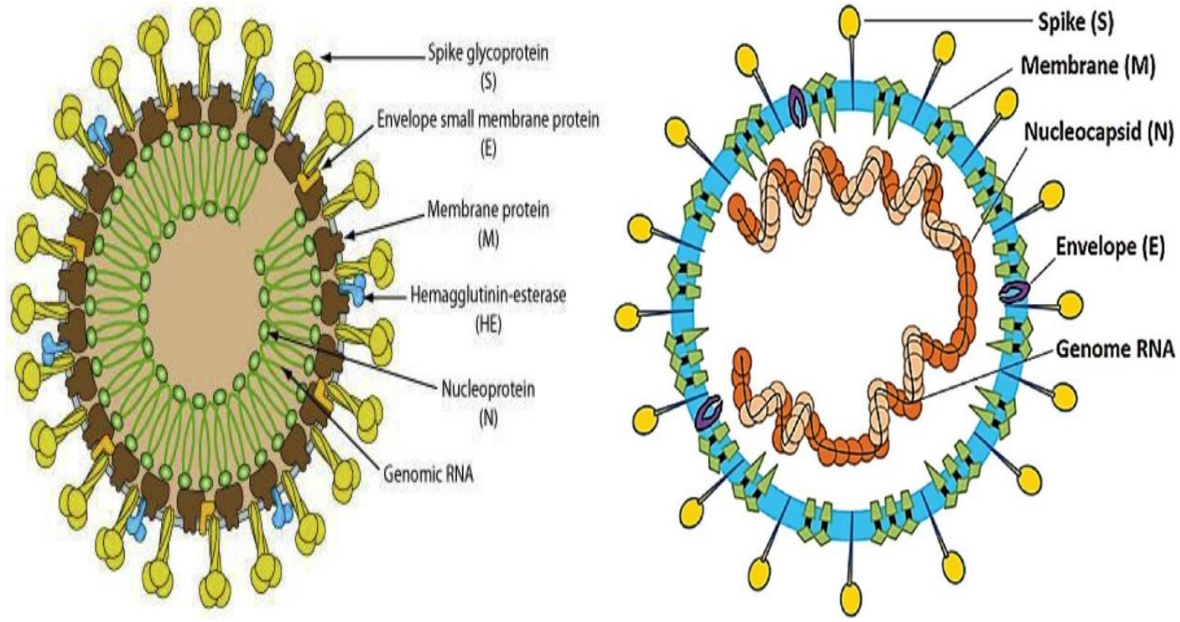

Figure-2: Schematic of a coronavirus – this new virus probably looks a lot like Coronaviruses possess the largest genomes [26.4 kb (ThCoV HKU12) to 31.7 kb (SW1)] among all known RNA viruses].

The large genome has given this family of virus extra plasticity in accommodating and modifying genes. The G + C contents of coronavirus genomes vary from 32% (HCoV-HKU1) to 43% (Pi-BatCoV HKU5 and MunCoV HKU13). Both the 5′ and 3′ ends of coronavirus genomes contain short untranslated regions. For the coding regions, the genome organizations of all coronaviruses are similar, with the characteristic gene order 5′-replicase ORF1ab, spike (S), envelope (E), membrane (M), nucleocapsid (N)-3′, although variable numbers of additional ORFs are present in each subgroup of coronavirus. A transcription regulatory sequence (TRS) motif is present at the 3′ end of the leader sequence preceding most ORFs. The TRS motifs are thought to be important for a "copy-choice" mechanism that mediates the unique random template switching during RNA replication, resulting in a high frequency of homologous RNA recombination in coronaviruses.[5,20]

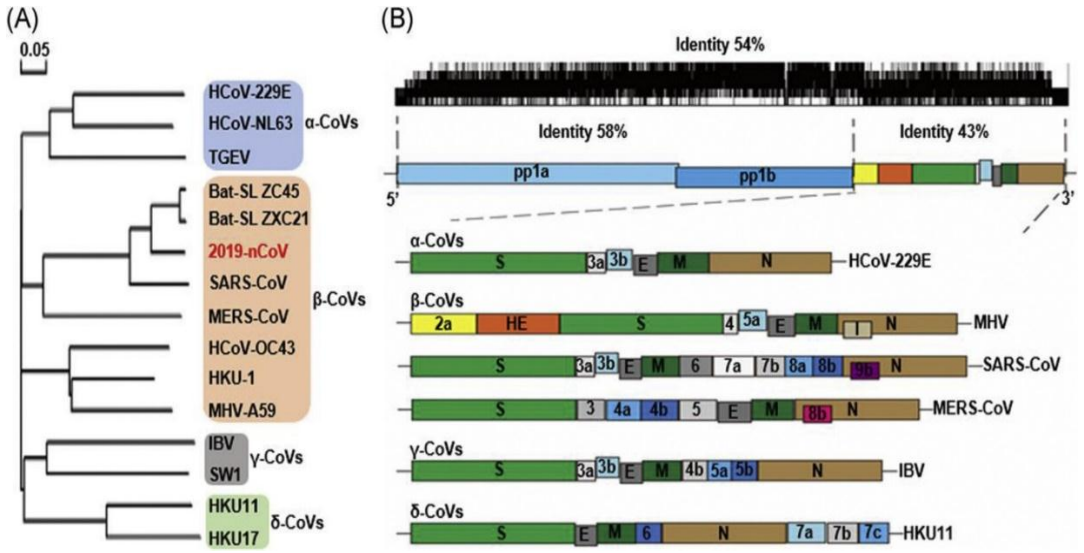

Figure-3: The genomic structure and phylogenetic tree of coronaviruses: A, the phylogenetic tree of representative CoVs, with the new coronavirus COVID-19 shown in red. B.

The genome structure of four genera of coronaviruses: two long polypeptides 16 nonstructural proteins have proceeded from Pp1a and pp1b represent. S, E, M, and N are represented of the four structural proteins spike, envelope, membrane, and nucleocapsid. COVID-19; CoVs, coronavirus; HE, hemagglutinin-esterase. Viral names: HKU, coronaviruses identified by Hong Kong University; HCoV, human coronavirus; IBV, infectious bronchitis virus; MHV, murine hepatitis virus; TGEV, transmissible gastroenteritis virus.

A typical CoV contains at least six ORFs in its genome. Except for Gammacoronavirus that lakes nsp1, the first ORFs (ORF1a/b), about two-thirds of the whole genome length, encode 16 nsps (nsp1-16). ORF1a and ORF1b contain a frameshift in between which produces two polypeptides: pp1a and pp1ab. These polypeptides are processed by virally encoded chymotrypsin-like protease (3CLpro) or main protease (Mpro) and one or two papain-like proteases into 16 nsps. All the structural and accessory proteins are translated from the sgRNAs of CoVs. Four main structural proteins contain spike (S), membrane (M), envelope (E), and nucleocapsid (N) proteins are encoded by ORFs 10, 11 on the one-third of the genome near the 3′-terminus. Besides these four main structural proteins, different CoVs encode special structural and accessory proteins, such as HE protein, 3a/b protein, and 4a/b protein (lower panel). These mature proteins are responsible for several important functions in genome maintenance and virus replication.[10]

There are three or four viral proteins in the coronavirus membrane. The most abundant structural protein is the membrane (M) glycoprotein; it spans the membrane bilayer three times, leaving a short NH2-terminal domain outside the virus and a long COOH terminus (cytoplasmic domain) inside the virion. The spike protein (S) as a type I membrane glycoprotein constitutes the peplomers. In fact, the main inducer of neutralizing antibodies is S protein. Between the envelope proteins with exist a molecular interaction that probably determines the formation and composition of the coronaviral membrane. M plays a predominant role in the intracellular formation of virus particles without requiring S. In the presence of tunicamycin coronavirus grows and produces spikeless, noninfectious virions that contain M but devoid of S.

ORF1ab: ORF1ab of coronaviruses occupy about two thirds of their genomes. It encodes the replicase polyprotein and is translated from ORF1a (11826 to 13425 nt) and ORF1b (7983 to 8157 nt). In all coronaviruses, a slippery sequence (UUUAAAC), followed by sequences that form a putative pseudoknot structure, are present at the junction between ORF1a and ORF1b. Translation occurs by a -1 RNA-mediated ribosomal frameshift at the end of the slippery sequence. Instead of reading the transcript as UUUAAACGGG, it will be read as UUUAAACCGGG. The replicase polyprotein is cleaved by papain-like protease(s) (PL^{pro}) and 3C-like protease ($3CL^{pro}$), proteins encoded by ORF1ab of the coronavirus genome, at consensus cleavage sites, into 15 to 16 non-structural proteins (nsps) named nsp1, nsp2, nsp3, etc. As the number of coronavirus genomes is expanding, novel cleavage sites have been discovered. Some of these non-structural proteins encode proteins of essential functions, such as PL^{pro} (nsp3), $3CL^{pro}$ (nsp5), RNA-dependent RNA polymerase (Pol) (nsp12) and helicase (nsp13). The genomes of all known members of *Alphacoronavirus* and *Betacoronavirus* subgroup A possess two PL^{pro} ($PL1^{pro}$ and $PL2^{pro}$), while those of all known members of *Betacoronavirus* subgroup B, C and D and *Gammacoronavirus* possess only one PL^{pro}. The gene sequences that encode these conserved proteins are frequently used for phylogenetic analysis.[12,14]

Table-1 Studies on estimation of dates of divergence of SARSr-CoV.

References	Gene	No. of SARSr-CoV strains				Estimated mean substitution rate (no. of substitutions per site per year)	Methods for estimating TMRCA	TMRCA of human/civet SARSr-CoV (95% HPD)	TMRCA of (human/civet)/ Bat Rp3 SARSr-CoV (95% HPD)	TMRCA of (human/civet/Bat Rp3 SARSr-CoV)/ SARSr-Rh-BatCoV (95% HPD)
		Human	Civet	Bat Rp3	SARSr-Rh-BatCoV					
Zeng *et al.* 2003 [38]	Spike	139				-	Linear regression	Dec 2002 (Sep 2002, Jan 2003)	-	-
Salemi *et al.* 2004 [39]	ORF1ab	10				$4/35\times10^{-4}$[b]	Molecular clock model	-	-	-
Zhao *et al.* 2004 [40]	Genome	16				$8\text{-}23.8\times10^{-4}$	Three strategies described by the author	Spring 2002	-	-
Song *et al.* 2005 [32]	CDSs[a]	3	5			2.92×10^{-3}	Linear regression	Nov 2002	-	-
Vijaykrishna *et al.* 2007 [35]	Helicase	3	3	1	5	2.0×10^{-2}, 1.7×10^{-2}[c]	Relaxed clock model	1999 (1990-2003)	1986 (1964-2002)	1961 (1918-1995)
Hon *et al.* 2008 [33]	ORF1ab	13	6	1	4	2.79×10^{-3}	Various clock models	2002.63 (2002.14-2002.96)	1998.51 (1993.55-2001.32)	~1985[d]
Lau *et al.* 2010 [37]	ORF1ab	8	8	1	15	2.82×10^{-3}	Relaxed clock model	2001 (1999.16-2002.14)	1995.10 (1986.53-2000.13)	1972.39 (1935.28-1990.63)

Table-1: Studies on estimation of dates of divergence of SARSr-CoV.

In addition to the nsps with essential functions, bioinformatics analysis of some other nsps revealed their putative functions. Downstream to PL^{pro} or $PL1^{pro}$ in nsp3 is the X domain which contains putative ADP-ribose 1″-phosphatase (ADRP) activity. In other microorganisms, such as *Saccharomyces cerevisiae* and other eukaryotes, ADRP and its functionally related enzyme cyclic nucleotide phosphodiesterase (CPDase), were important for tRNA processing. ADP-ribose 1″,2″-cyclic phosphate (Appr>p) is produced as a result of tRNA splicing. Appr>p is in turn converted to ADP-ribose 1″-phosphate (Appr-1″p) by CPDase and Appr-″p is then further processed by ADRP. As for nsp13, nsp14 and nsp15, they possess a putative 3′-to-5′ exonuclease (ExoN) domain of the DEDD superfamily, a putative poly(U)-specific endoribonuclease (XendoU) domain, and a putative S-adenosylmethionine-dependent ribose 2′-O-methyltransferase (2′-O-MT) domain of the RrmJ family respectively. ADRP, CPDase, ExoN, XendoU and 2′-O-MT are enzymes in RNA processing pathways. Contrary to the pre-tRNA splicing pathway that ADRP and CPDase belong to, ExoN, XendoU and 2′-O-MT are enzymes in a small nucleolar RNA processing and utilization pathway.

(A) Genome structure of SARS-CoV-2

(B) ORF1ab gene-coding proteins

ORF1ab-coding proteins	ORF1ab polyprotein				ORF1a polyprotein			
	Start	Stop	Length	Protein ID	Start	Stop	Length	Protein ID
Leader protein	1	540	540	YP_009725297.1	266	805	540	YP_009742608.1
Nsp2	541	2454	1914	YP_009725298.1	806	2719	1914	YP_009742609.1
Papain-like protease	2455.	8289	5835	YP_009725299.1	2720	8554	5835	YP_009742610.1
Nsp4	8290	9789	1500	YP_009725300.1	8555	10054	1500	YP_009742611.1
3C-like proteinase	9790	10707	918	YP_009725301.1	10055	10972	918	YP_009742612.1
Nsp6	10708	11577	870	YP_009725302.1	10973	11842	870	YP_009742613.1
Nsp7	11578	11826	249	YP_009725303.1	11843	12091	249	YP_009742614.1
Nsp8	11827	12420	594	YP_009725304.1	12092	12685	594	YP_009742615.1
Nsp9	12421	12759	339	YP_009725305.1	12686	13024	339	YP_009742616.1
Nsp10	12760	13176	417	YP_009725306.1	13025	13441	417	YP_009742617.1
Nsp11	-	-	-	-	13442	13480	39	YP_009725312.1
RNA-dependent RNA polymerase	13177	15971	2795	YP_009725307.1	-	-	-	-
Helicase	15972	17774	1803	YP_009725308.1	-	-	-	-
3'-to-5' Exonuclease	17775	19355	1581	YP_009725309.1	-	-	-	-
EndRNAse	19356	20393	1038	YP_009725310.1	-	-	-	-
2'-O-Ribose methyltransferase	20659	21552	894	YP_009725311.1	-	-	-	-

Figure-4: Studies on estimation of dates of divergence of SARSr-CoV.

Phylogeny: The first impression of the phylogenetic position of a strain or species of coronavirus is usually acquired by constructing a phylogenetic tree using a short fragment of a conserved gene, such as Pol or N.

However, this can sometimes be misleading because the results of phylogenetic analysis using different genes or characters can be different. When SARSr-CoV was first discovered, it was proposed that it constituted a fourth group of coronaviruses. However, analyses of the amino-terminal domain of S of SARSr-CoV revealed that 19 out of the 20 cysteine residues were spatially conserved with those of the consensus sequence for Betacoronavirus. On the other hand, only five of the cysteine residues were spatially conserved with those of the consensus sequences in Alphacoronavirus and Gammacoronavirus. Furthermore, subsequent phylogenetic analysis using both complete genome sequence and proteomic approaches, it was concluded that SARSr-CoV is probably an early split-off from the Betacoronavirus lineage, and SARSr-CoV was subsequently classified as Betacoronavirus subgroup B and the historical Betacoronavirus as Betacoronavirus subgroup A. Therefore, the phylogenetic position of a coronavirus is best appreciated and confirmed by constructing phylogenetic trees using different genes in the coronavirus genome. The most commonly used genes along the coronavirus genome for phylogenetic studies include chymotrypsin-like protease, helicase, S and N, because these genes are present in all coronavirus genomes and are of significant length. The envelope and membrane genes, although present in all coronavirus genomes, are too short for phylogenetic studies. It is noteworthy that the cluster formed by the three novel avian coronaviruses BuCoV HKU11, ThCoV HKU12 and MunCoV HKU13, which was originated proposed as group 3c, might represent a new coronavirus genus provisionally designated Deltacoronavirus.[6,7]

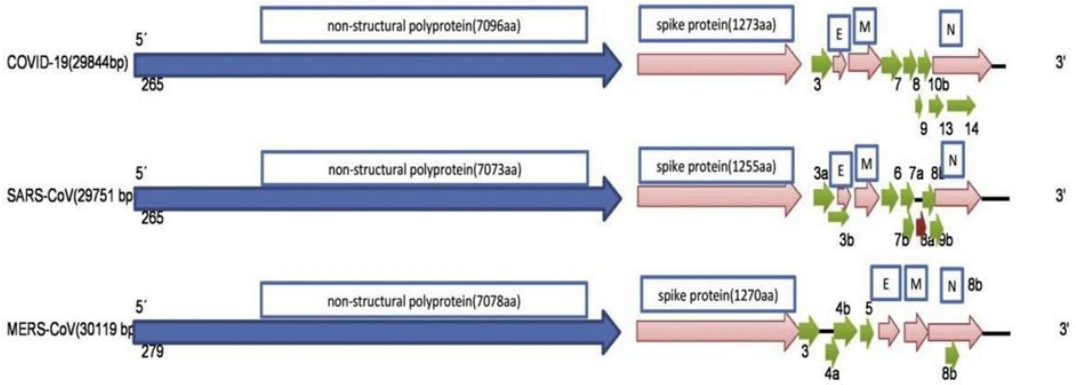

Figure-5: 5′ UTR and 3′ UTR and coding region of COVID-19, SARS-CoV, and MERS-CoV. The numbers of base pairs among betacoronaviruses are shown. This figure is modified from the sequence comparison and genomic organization of 2019-nCoV, 2020.The differences in the arrangement of the envelope (E), membrane (M), and nucleoprotein (N) among COVID-19, SARS-CoV, and MERS-CoV are shown at 3′ end.

Recombination analysis: As a result of their unique random template switching during RNA replication, thought to be mediated by a "copy-choice" mechanism, cronaviruses have a high frequency of homologous RNA recombination. Recombination in coronaviruses was first recognized between different strains of MHV and subsequently in other coronaviruses such as IBV, between MHV and BCoV, and between feline coronavirus (FCoV) type I and canine coronavirus (CCoV).[11,13] As shown below, such recombination can result in the generation of coronavirus species or different genotypes within a coronavirus species. In our experience, the possibility of homologous RNA recombination and the possible part of the genome that recombination has taken place are usually first appreciated using bootscan analysis or phylogenetic analysis using different parts of the coronavirus genome. Other methods for recombination analysis, such as those in the RDP3 package, are also available. Then, the exact site of homologous RNA recombination would be best revealed by multiple sequence alignment. The best documented example of generation of coronavirus species through homologous RNA recombination is the generation of FCoV type II by double recombination between FCoV type I and CCoV. It was first observed that the sequence of the S protein in FCoV type II was closely related to that of CCoV.[17,18] but the sequence downstream of the E gene in FCoV type II was more closely related to that of FCoV type I strain than to CCoV. This observation suggested that there might have been a homologous RNA recombination event between the genomes of CCoV and FCoV type I, resulting in the generation of FCoV type II. Further analysis by multiple sequence alignments pinpointed the site of recombination to a region in the E gene. A few years later, an additional recombination region in the Pol gene was also discovered, and it was concluded that FCoV type II originated from two recombination events between the genomes of CCoV and FCoV type I. As for the generation of different genotypes in a coronavirus species through homologous RNA recombination, the best documented example is HCoV-HKU1. The possibility of homologous RNA recombination was first suspected when a few strains of HCoV-HKU1 showed differential clustering when the Pol, S and N genes were used for phylogenetic tree construction. This observation has led to our subsequent study on complete genome sequencing of 22 strains of HCoV-HKU1.

Recombination analysis by bootscan analysis and phylogenetic analysis using different parts of the 22 complete genomes revealed extensive recombination in different parts of the genomes, resulting in the generation of three genotypes, A, B and C, of HCoV-HKU1. Using multiple sequence alignment, two sites of recombination were pinpointed. The first one was observed in a stretch of 143 nucleotides near the 3' end of nsp6, where recombination between HCoV-HKU1 genotype B and genotype C has generated genotype A; and the second one in another stretch of 29 nucleotides near the 3' end of nsp16, where recombination between HCoV-HKU1 genotype A and genotype B has generated genotype C.[11,15]

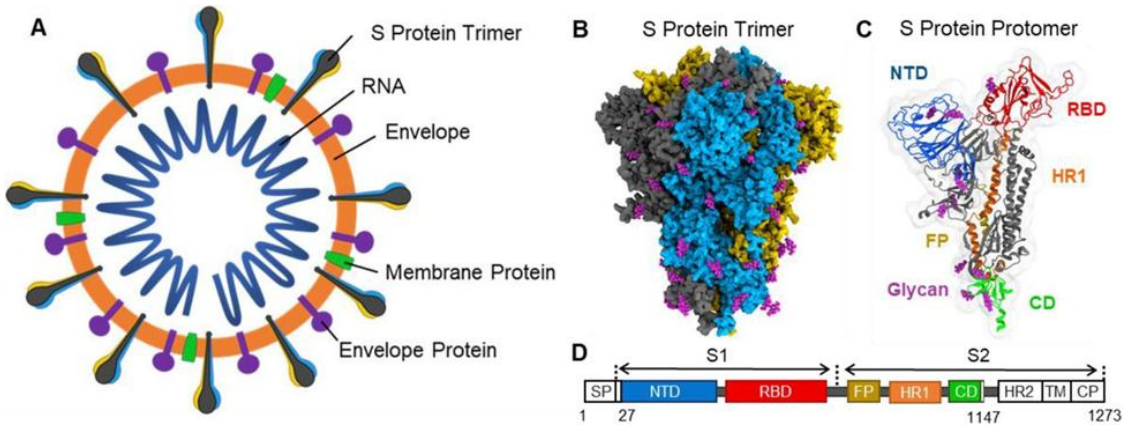

Figure-6: Genotype & Phenotype Comparison of SARS-CoV-2.

Codon usage bias: Recently, using the complete genome sequences of the 19 coronavirus genomes, we analysed the codon usage bias in coronaviruses as well as selection of CpG suppressed clones by the immune system and cytosine deamination being the two major independent biochemical and biological selective forces that has shaped such codon usage bias. In the study, we showed that the mean CpG relative abundance in the coronavirus genomes is markedly suppressed. However, we observed that only CpG containing codons in the context of purine-CpG (ACG and GCG), pyrimidine-CpG (UCG and CCG) and CpG-purine (CGA and CGG); but not CpG-pyrimidine (CGU and CGC); are suppressed. However, when trinucleotide frequencies were analyzed in the 19 coronavirus genomes, all the eight trinucleotides with CpG were suppressed. These indicate that another force that has led to an increase use of CGU and CGC as codons for arginine, but does not act on trinucleotides over the whole genome in general, is probably present. Furthermore, this force is probably unrelated to the relative abundance of the corresponding tRNA molecules in the hosts of the coronaviruses, as the pattern of bias in the hosts is not the same as that in the coronaviruses. In addition to CpG suppression, marked cytosine deamination is also observed in all the 19 coronavirus genomes. Using the six amino acids that are only encoded by NNC or NNU (asparagine, histidine, aspartic acid, tyrosine, cysteine and phenylalanine), hence excluding most other pressures that may affect the relative abundance of cytosine and uracil, it was observed that all NNU are markedly over represented with usage fractions of more than 0.700, whereas the usage fractions of all NNC are less than 0.300. For all codons that encode the same amino acid and with either U or C in any position, the usage fraction of the codon that uses U is always higher than the one that uses C in all coronaviruses. Furthermore, the percentage of U showed strong inverse relationships with the percentage of C in the coronavirus genomes.[12] These suggest that cytosine deamination is another important biochemical force that shaped coronavirus evolution. Interestingly, among all the 19 coronaviruses, HCoV-HKU1 showed the most extreme codon usage bias. HCoV-HKU1 is the only coronavirus that had effective number of codons outside the mean ± 2 standard deviations range. In addition, HCoV-HKU1 also possessed the lowest G + C content, highest GC skew, lowest percentages of G and C and highest percentage of U among all coronavirus genomes. Furthermore, HCoV-HKU1 showed extremely high NNU/NNC ratio of 8.835. The underlying mechanism for the extreme codon usage bias, cytosine deamination and G + C content in HCoV-HKU1 is intriguing.

Database (PDB): Rapid and accurate batch sequence retrieval is always the cornerstone and bottleneck for all kinds of comparative genomics and bioinformatics analysis. During the process of batch sequence retrieval for comparative genomics and other bioinformatics analysis of the coronavirus genomes that we have sequenced, we encountered a number of major problems about the coronavirus sequences in GenBank and other coronavirus databases. First, in GenBank, the non-structural proteins encoded by ORF1ab are not annotated. Second, in all databases, the annotations for the non-structural proteins encoded by ORFs downstream to ORF1ab are often confusing because they are not annotated using a standardized system. Third, multiple accession numbers are often present for reference sequences. These problems will often lead to confusion during sequence retrieval. Fourth, coronaviruses, especially SARSr-CoV, amplified from

different specimens may contain the same gene or genome sequences, which will lead to redundant work when they are analysed. In view of these problems, we have developed a comprehensive database, CoVDB, of annotated coronavirus genes and genomes, which offers rapid, efficient and user-friendly batch sequence retrieval and analysis. In CoVDB, first, annotations on all non-structural proteins in the polyprotein encoded by ORF1ab of every single sequence were performed. Second, annotation was performed for the non-structural proteins encoded by ORFs downstream to ORF1ab using a standardized system. Third, all sequences with identical nucleotide sequences were labelled and one can choose to show or not to show strains with identical sequences. Fourth, this database contains not only complete coronavirus genome sequences, but also incomplete genomes and their genes. This is useful because some genes of coronaviruses, such as Pol, S and N, are sequenced much more frequently than others because they are either most conserved or least conserved, and therefore are particularly important for primers design for RT-PCR assays and evolutionary studies.

RNA-Dependent RNA Polymerase: As discussed above for other coronavirus pp1a/pp1ab proteins, the RdRp domain also differs substantially from its homologs in other +RNA viruses. Coronavirus RdRps and their nidovirus relatives have been classified as an outgroup of SF1 RdRps. The coronavirus RdRp domain comprising the finger, palm, and thumb subdomains occupies the C-terminal two-thirds of nsp12. Recent data suggest that replication complex association of the RdRp may occur through interactions of the nsp12 (located upstream of the RdRp core domain in nsp12) with ORF1a-encoded proteins, such as nsp5 (3CLpro), nsp8, and nsp9. Consistent with the presumed RdRp activity of nsp12, a mutation in nsp12 was found to cause an RNA-negative phenotype in an MHV ts mutant, Alb ts22. Thus, when infected cultures of Alb ts22 were shifted to the restrictive temperature at 40C, both plus- and minus-strand RNA synthesis ceased immediately.[8,9]

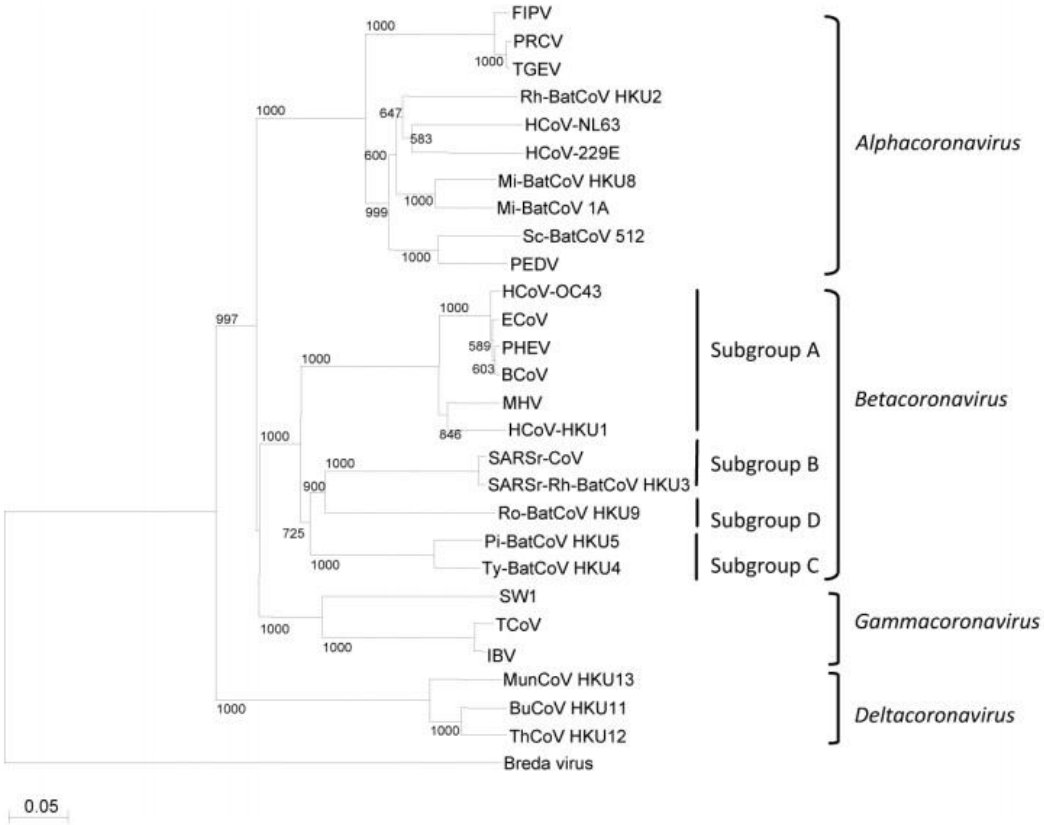

Figure-7: Phylogenetic analysis of RNA-dependent RNA polymerases (Pol) of coronaviruses with complete genome sequences available. The tree was constructed by the neighbor-joining method and rooted using Breda virus polyprotein. Bootstrap values were calculated from 1000 trees. 1118 amino acid positions in Pol were included. The scale bar indicates the estimated number of substitutions per 20 amino acids.

Even at the permissive temperature, the mutant synthesized 4–5 times less RNA compared with revertants. The defect of this mutant in RNA synthesis can easily be explained by the fact that His868 is part of the predicted thumb subdomain of the MHV RdRp that, in other RNA polymerases, has been implicated in polymerase activity. The Cys/His-rich nsp10 that immediately precedes RdRp in pp1ab has also been implicated in RNA synthesis. An MHV ts mutant, Alb ts6, encoding a mutant form of nsp10, was shown to have a defect in minus-strand RNA synthesis.[16,17] Thus, when the temperature was shifted to 40C, minus-strand

synthesis stopped immediately but plus-strand synthesis continued at the same level as was occurring at the time of temperature shift. Plus-strand RNA synthesis gradually declined over 3–4 h (starting at 30–60 min after the shift to 40C) because the minus strands produced at the permissive temperature were turned over and, because of the defect in their synthesis, were not replenished at the restrictive temperature. Nsp10 and nsp12 (RdRp) are adjacent domains in pp1ab. Peptide cleavage data have shown that, most likely because of a replacement of the conserved P2 Leu residue, the nsp10|nsp12 cleavage site is less efficiently cleaved than other SARS-CoV 3CLpro sites. Also, the nsp10|nsp12 sites of other coronaviruses have the P2 position occupied by noncanonical residues. It is thus tempting to speculate that the nsp10|nsp12 site has to be cleaved more slowly than other sites, probably to attain a specific activity mediated by a nsp10–nsp12- containing intermediate. The IBV nsp10 has been reported to form dimers. It localizes to membranes near the site of viral RNA synthesis.[19]

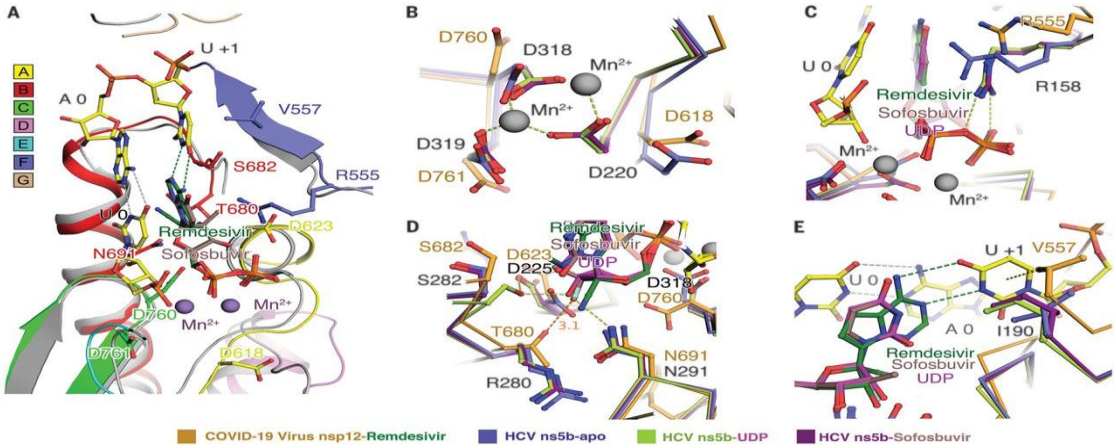

Figure-8: RNA dependent RNA Polymerase of SARS-CoV-2.

Concluding Remarks: Although much has been learned about coronavirus replicase organization, localization, proteolytic processing, and some of the viral replicaive enzymes (e.g., proteinases and helicases), there are still major gaps in our knowledge. Given the availability of full-length clones of coronaviruses, directed genetic analysis is now possible. In-vivo studies as well as biochemical and structural information should yield important new information on the molecular details of coronaviral RNA synthesis. In this context, it will be of particular interest to define the proteins that are responsible for the unique features of coronavirus RNA synthesis, for example, the production of an extensive set of 50 - and 30 -coterminal subgenomic RNAs and the synthesis and maintenance of RNA genomes of this unique size. Studies on coronavirus replicases and their homologs on closely related viruses may also help to determine the structural and functional constraints that have driven the evolution of nidoviruses and enable them to infect a broad range of vertebrate and invertebrate hosts. Furthermore, the relationship of the recently identified coronavirus RNA processing activities with cellular proteins may reveal interesting insights into similarities and differences (or even an interplay) between coronaviral and cellular RNA metabolism pathways. In the long term, the unique structural properties of coronavirus replicative enzymes may allow the development of very selective enzyme inhibitors and possibly even drugs suitable to combat coronavirus infections.

REFERENCES

1. Arunava Chandra Chandra, Dr. Dhrubo Jyoti Sen and Dr. Beduin Mahanti; *Protein data bank of corona virus→A global pandemic→Stay safe→Stay tuned*; World Journal of Pharmaceutical Research, 2020; 9(5): 1971-1994. (ISSN: 2277-7105, Impact Factor: 8.084, Index Copernicus Value: 90.8)
2. Dr. Dhrubo Jyoti Sen, Dr. Sudip Kumar Mandal, Arpita Biswas, Dipra Dastider and Dr. Beduin Mahanti; Corona is culprit of ridiculous offensive nonsense air: World Journal of Pharmaceutical Research, 2020; 9(4): 503-518.
3. Y. Chen, Q. Liu, D. Guo; Emerging coronaviruses: genome structure, replication, and pathogenesis; J Med Virol, 2020; 92: 418-423.
4. C.A.M. de Haan, L. Kuo, P.S. Masters, H. Vennema, P.J.M. Ro ttier; Coronavirus particle assembly: primary structure requirements of the membrane protein. J Virol, 1998; 72(8): 6838-6850.
5. Woo, P.C.; Lau, S.K.; Lam, C.S.; Lai, K.K.; Huang, Y.; Lee, P.; Luk, G.S.; Dyrting, K.C.; Chan, K.H.; Yuen, K.Y. Comparative analysis of complete genome sequences of three avian coronaviruses reveals a novel group 3c coronavirus. J. Virol, 2009; 83: 908-917.
6. Mihindukulasuriya, K.A.; Wu, G.; St Leger, J.; Nordhausen, R.W.; Wang, D. Identification of a novel coronavirus from a beluga whale by using a panviral microarray. J. Virol, 2008; 82: 5084-5088.

7. Eickmann, M.; Becker, S.; Klenk, H.D.; Doerr, H.W.; Stadler, K.; Censini, S.; Guidotti, S.; Masignani, V.; Scarselli, M.; Mora, M.; Donati, C.; Han, J.H.; Song, H.C.; Abrignani, S.; Covacci, A.; Rappuoli, R. Phylogeny of the SARS coronavirus. Science, 2003; 302: 1504-1505.
8. Chen Y, Guo D. Molecular mechanisms of coronavirus RNA capping and methylation. *Virol Sin*., 2016; 31(1): 3- 11.
9. Ogando NS, Ferron F, Decroly E, Canard B, Posthuma CC, Snijder EJ. The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. *Front Microbiol*, 2019; 10: 1813.
10. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya AE. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J Mol Biol.*, 2003; 331: 991–1004.
11. Herrewegh, A.A.; Smeenk, I.; Horzinek, M.C.; Rottier, P.J.; de Groot, R.J. Feline coronavirus type II strains 79-1683 and 79-1146 originate from a double recombination between feline coronavirus type I and canine coronavirus. J. Virol, 1998; 72: 4508-4514.
12. Herrewegh AA, Vennema H, Horzinek MC, Rottier PJ, de Groot RJ. The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes. *Virology*, 1995; 212: 622–631.
13. Lavi, E.; Haluskey, J.A.; Masters, P.S. The pathogenesis of MHV nucleocapsid gene chimeric viruses. Adv. Exp. Med. Biol., 1998; 440: 537-541.
14. Woo PC, Lau SK, Tsoi HW, Huang Y, Poon RW, Chu CM, Lee RA, Luk WK, Wong GK, Wong BH, Cheng VC, Tang BS, Wu AK, Yung RW, Chen H, Guan Y, Chan KH, Yuen KY. Clinical and molecular epidemiological features of coronavirus HKU1-associated community-acquired pneumonia. *J Infect Dis.*, 2005; 192: 1898–1907.
15. Woo PC, Lau SK, Yip CC, Huang Y, Tsoi HW, Chan KH, Yuen KY. Comparative analysis of 22 coronavirus HKU1 genomes reveals a novel genotype and evidence of natural recombination in coronavirus HKU1. *J Virol*, 2006; 80: 7136–7145.
16. Huang Y, Lau SK, Woo PC, Yuen KY. CoVDB: a comprehensive database for comparative analysis of coronavirus genes and genomes. *Nucleic Acids Res.*, 2008; 36: D504–511.
17. Motokawa, K.; Hohdatsu, T.; Aizawa, C.; Koyama, H.; Hashimoto, H. Molecular cloning and sequence determination of the peplomer protein gene of feline infectious peritonitis virus type I. Arch. Virol, 1995; 140: 469-480.
18. Wesseling, J.G.; Vennema, H.; Godeke, G.J.; Horzinek, M.C.; Rottier, P.J. Nucleotide sequence and expression of the spike (S) gene of canine coronavirus and comparison with the S proteins of feline and porcine coronaviruses. J. Gen. Virol, 1994; 75(Pt 7): 1789-1794.
19. Woo PC, Lau SK, Huang Y, Yuen KY. Coronavirus diversity, phylogeny and interspecies jumping. *Exp Biol Med (Maywood)*, 2009; 234: 1117–1127.
20. Salemi, M.; Fitch, W.M.; Ciccozzi, M.; Ruiz-Alvarez, M.J.; Rezza, G.; Lewis, M.J. Severe acute respiratory syndrome coronavirus sequence characteristics and evolutionary rate estimate from maximum likelihood analysis. J. Virol, 2004; 78: 1602-1603.