

# Case Study: SARS-CoV-2

---



INDRAPRASTHA INSTITUTE *of*  
INFORMATION TECHNOLOGY **DELHI**

Dr. Jaspreet Kaur Dhanjal  
Assistant Professor, Center for Computational Biology  
Email ID: [jaspreet@iiitd.ac.in](mailto:jaspreet@iiitd.ac.in)

*September 16, 2025*

# Identification of SARS-CoV-2

ORIGINAL ARTICLE BRIEF REPORT

## A Novel Coronavirus from Patients with Pneumonia in China, 2019

Na Zhu, Ph.D., Dingyu Zhang, M.D., Wenling Wang, Ph.D., Xingwang Li, M.D., Bo Yang, M.S., Jingdong Song, Ph.D., Xiang Zhao, Ph.D., Baoying Huang, Ph.D., Weifeng Shi, Ph.D., Roujian Lu, M.D., Peihua Niu, Ph.D., Faxian Zhan, Ph.D., et al., for the China Novel Coronavirus Investigating and

### VIRAL GENOME SEQUENCING

RNA extracted from bronchoalveolar-lavage fluid and culture supernatants was used as a template to clone and sequence the genome. We used a combination of Illumina sequencing and nanopore sequencing to characterize the virus genome. Sequence reads were assembled into contig maps (a set of overlapping DNA segments) with the use of CLC Genomics software, version 4.6.1 (CLC Bio). Specific primers were subsequently designed for PCR, and 5'- or 3'-RACE (rapid amplification of cDNA ends) was used to fill genome gaps from conventional Sanger sequencing. These PCR products were purified from gels and sequenced with a BigDye Terminator v3.1 Cycle Sequencing Kit and a 3130XL Genetic Analyzer, in accordance with the manufacturers' instructions.

# Identification of SARS-CoV-2

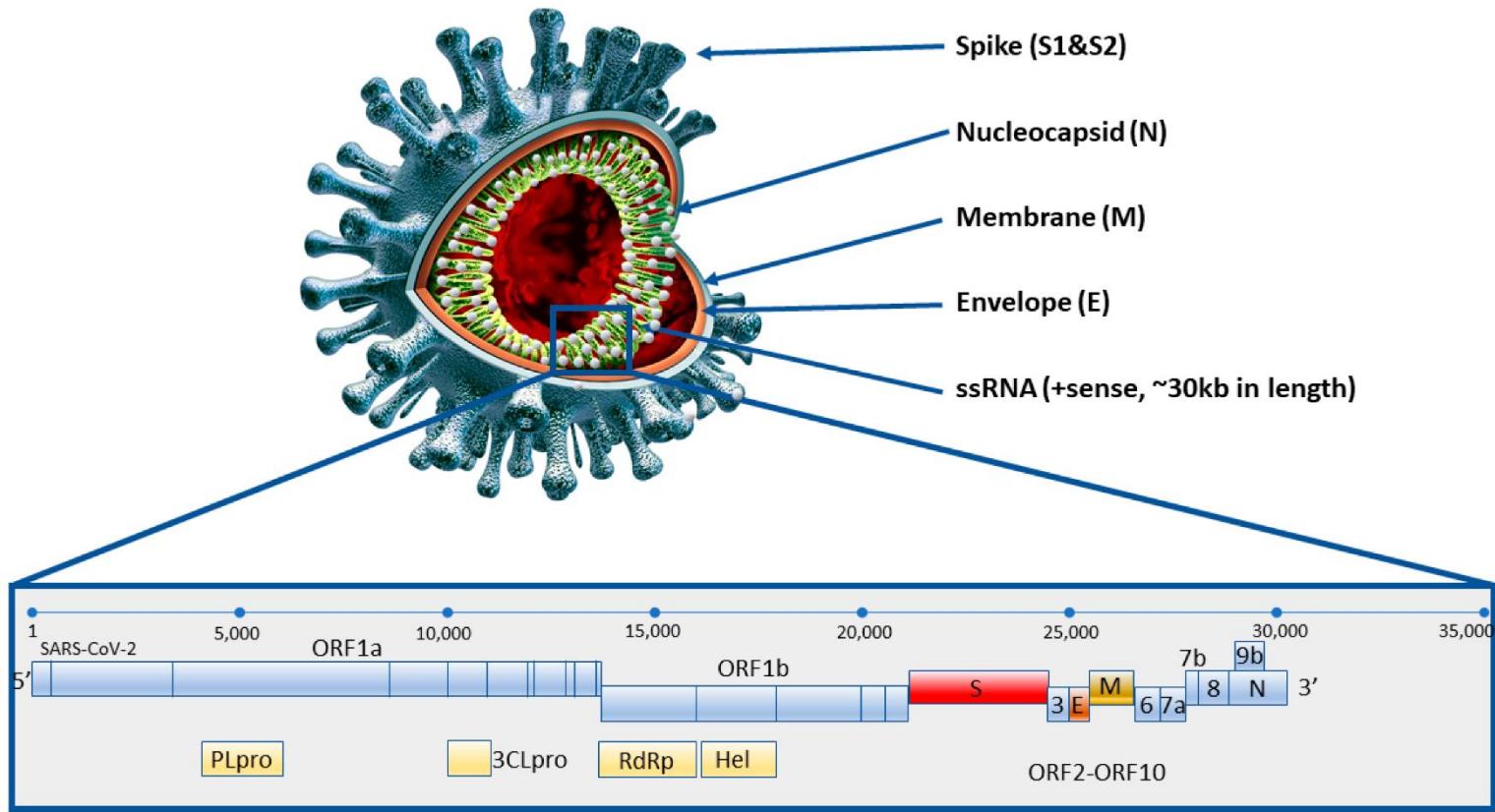
ORIGINAL ARTICLE BRIEF REPORT

## A Novel Coronavirus from Patients with Pneumonia in China, 2019

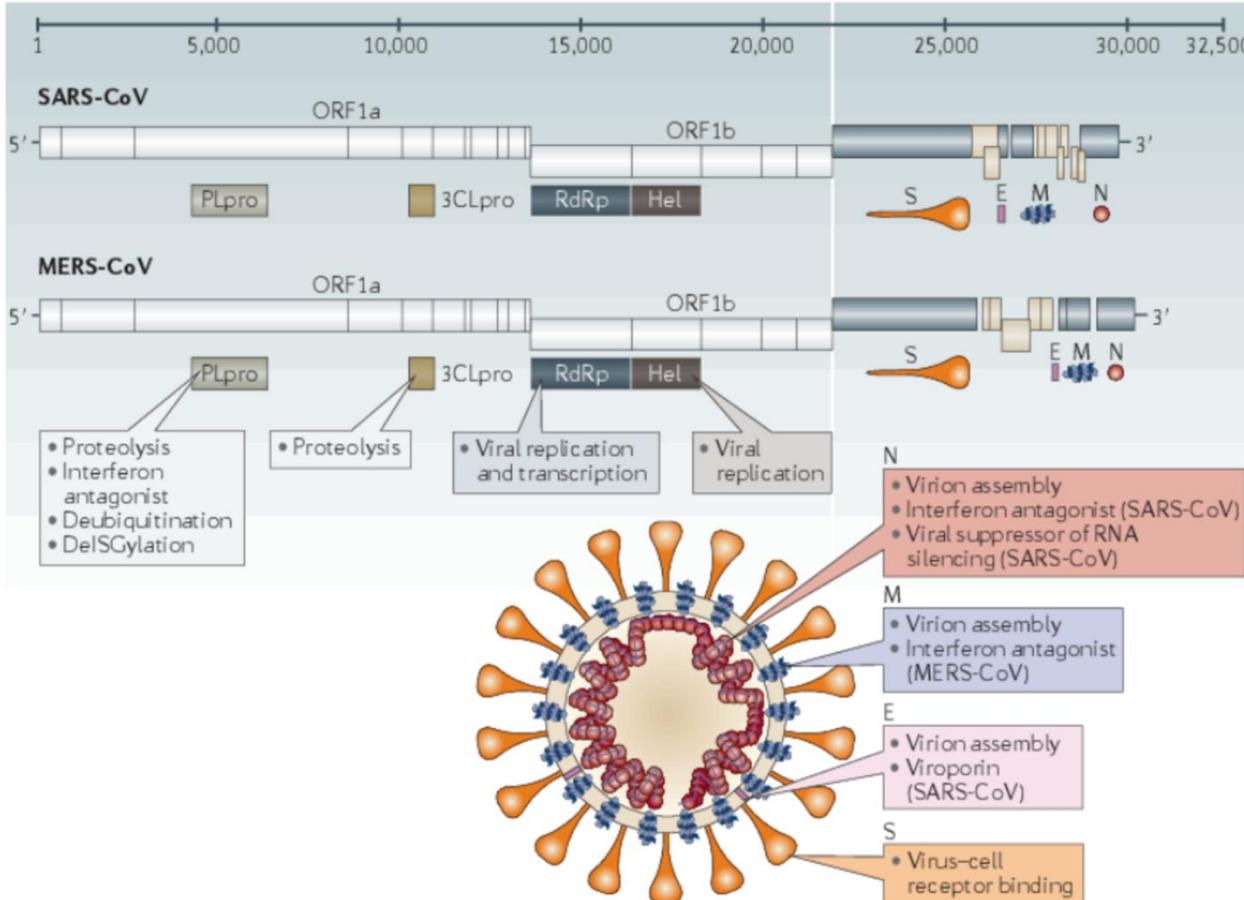
Na Zhu, Ph.D., Dingyu Zhang, M.D., Wenling Wang, Ph.D., Xingwang Li, M.D., Bo Yang, M.S., Jingdong Song, Ph.D., Xiang Zhao, Ph.D., Baoying Huang, Ph.D., Weifeng Shi, Ph.D., Roujian Lu, M.D., Peihua Niu, Ph.D., Faxian Zhan, Ph.D., et al., for the China Novel Coronavirus Investigating and

RespiFinderSmart22kit. RNA extracted from bronchoalveolar-lavage fluid from the patients was used as a template to clone and sequence a genome using a combination of Illumina sequencing and nanopore sequencing. More than 20,000 viral reads from individual specimens were obtained, and most contigs matched to the genome from lineage B of the genus betacoronavirus — showing more than 85% identity with a bat SARS-like CoV (bat-SL-CoVZC45, MG772933.1) genome published previously. Positive

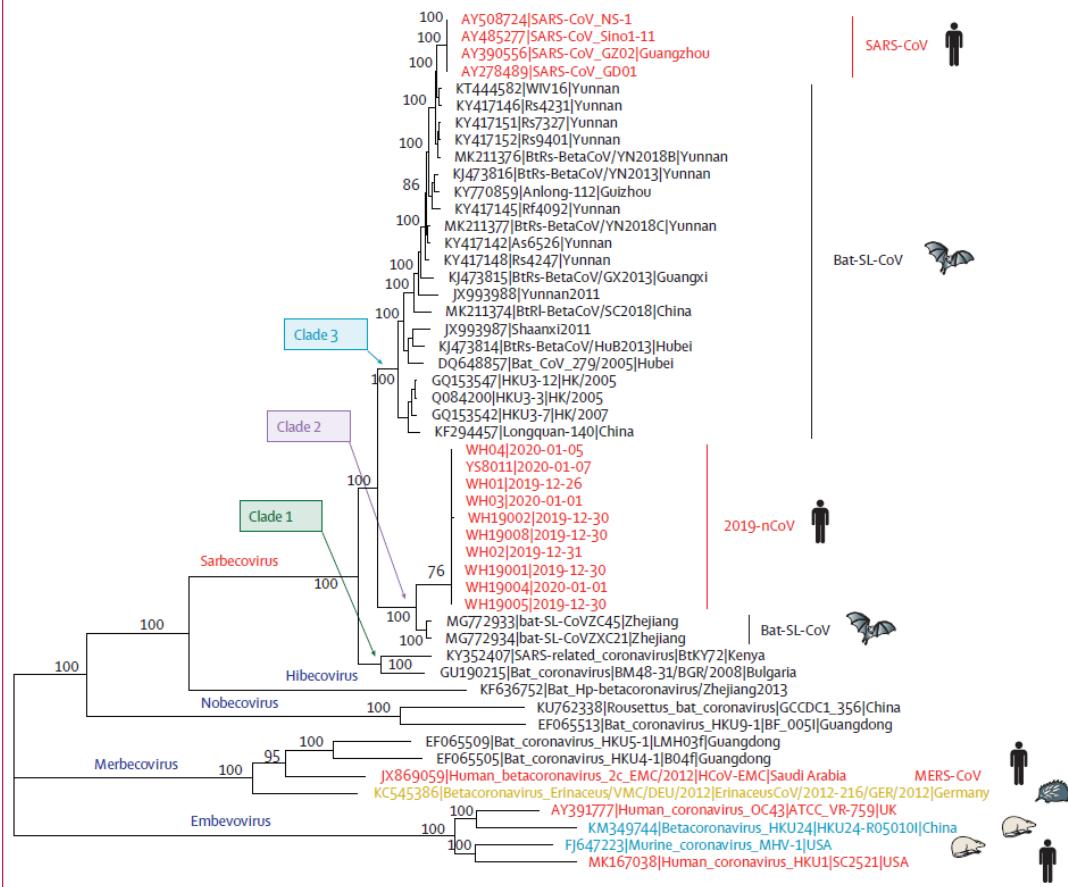
# Genomic structure



# SARS-CoV-2 proteins



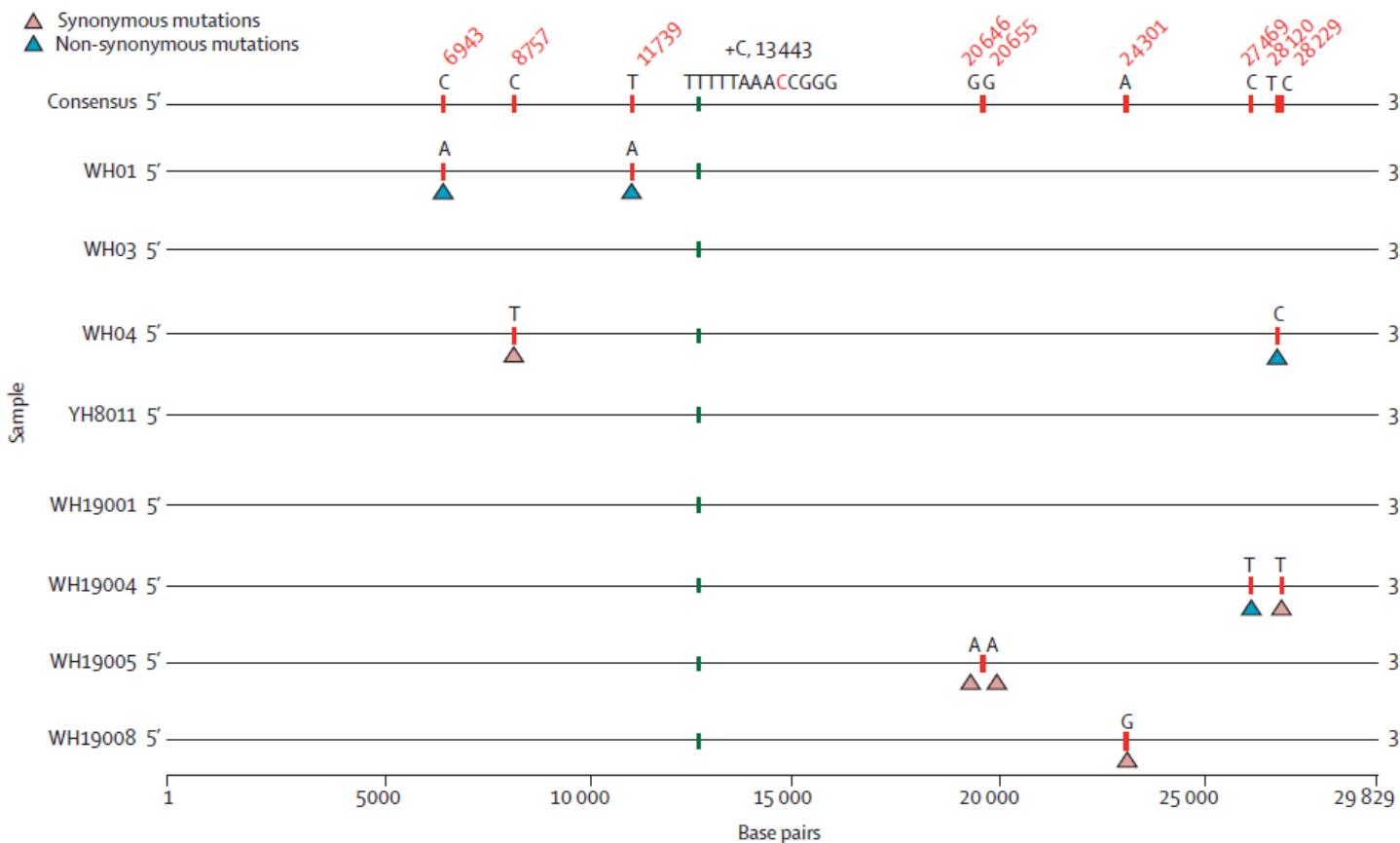
# Phylogenetic analysis of SARS-CoV-2



## Phylogenetic analysis

Sequence alignment of 2019-nCoV with reference sequences was done with Mafft software (version 7.450).<sup>21</sup> Phylogenetic analyses of the complete genome and major coding regions were done with RAxML software (version 8.2.9)<sup>22</sup> with 1000 bootstrap replicates, employing the general time reversible nucleotide substitution model.

# SARS-CoV-2 variants



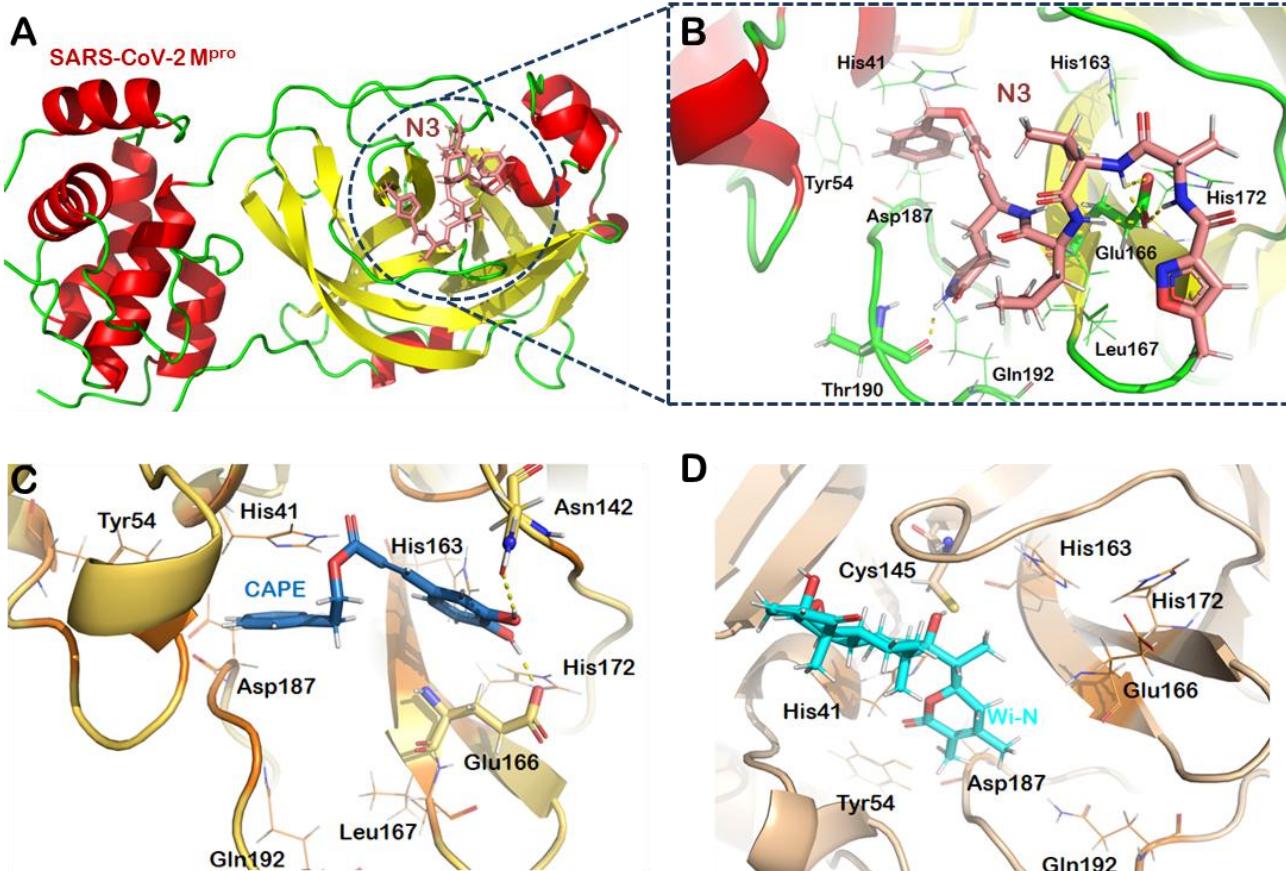
# Drug repurposing for SARS-CoV-2

CLUSTAL O(1.2.4) multiple sequence alignment

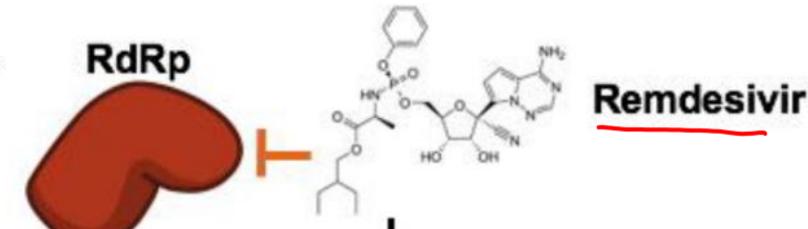
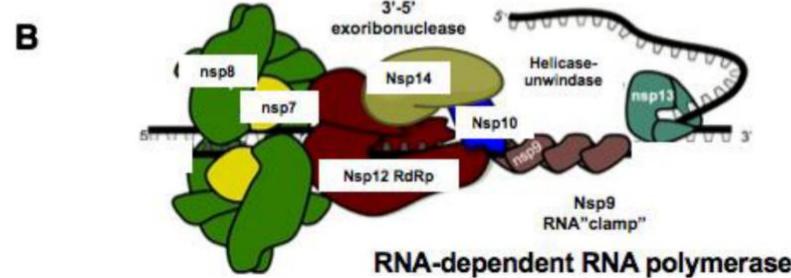
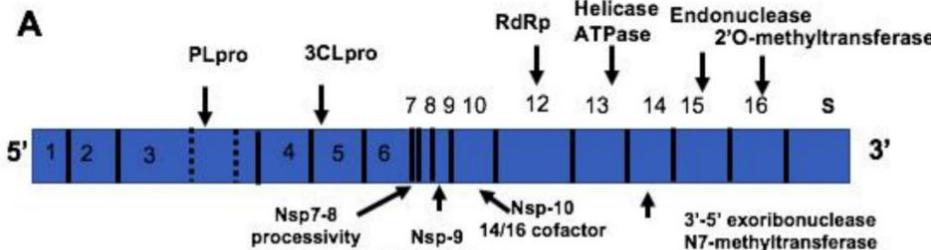
3TLO:A PDBID CHAIN SEQUENCE	-----SGLKKMAOPSGCVERCVRVCYGSTVNLNGWILGDTVTCPPFVIAPISTT-VLID	52
1P9S:A PDBID CHAIN SEQUENCE	-----AGLRKMAQPSGFVEKCVRVCYGNVLNLGLWLDIVYCPRHIAASNTT-SAID	52
6LU7:A PDBID CHAIN SEQUENCE	-----SGFRKMAFPSSGVKEGOMQVTCGTTNLGLWLDDVVYCPRHICITSEDMLNPN	53
3IWM:A PDBID CHAIN SEQUENCE	-----SGFRKMAFPSSGVKEGOMQVTCGTTNLGLWLDDVVYCPRHICITAEDMLNPN	53
3D23:B PDBID CHAIN SEQUENCE	-----ASSGIKMNSPTSKIEPC1CVS1TYGS1MTLNGWLDDKVKYF1C1SSSNNEPD	55
5WKK:A PDBID CHAIN SEQUENCE	MHHHHHHSGLVKMSHPSGGDVACM1QVTCGS1MTLNGWLDDNTWCPRHICPAQDSLDPN	60
	: *: ** *: : *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: :	
3TLO:A PDBID CHAIN SEQUENCE	VDHAYSTMRHLNHSVS--HNGVFLGVGVTVHGSVSLRIKVQSNSVHTPKHVFKTLKPGD	109
1P9S:A PDBID CHAIN SEQUENCE	DHEVYIMRLNHSVS---SGTAFLGVVGTMHVTKLIIKVSGTMNHTPRHSFRTLSKG	109
6LU7:A PDBID CHAIN SEQUENCE	YEDILLRKSHNHLFVLQ---AGNVLRLVIGHSMQNCVLLKVDTANPKTPKVKFVR1QPGQ	110
3IWM:A PDBID CHAIN SEQUENCE	YEDILLRKSHNHSVLQ---AGNVLRLVIGHSMQNCVLLKVDTANPKTPKVKFVR1QPGQ	110
3D23:B PDBID CHAIN SEQUENCE	YSALLCRVTLGDF1IM---SGRMSLTVVSYQMGQLVLTVSLQNPYPTKVTGNNVPGE	112
5WKK:A PDBID CHAIN SEQUENCE	YDALLISMTHHSFSVQKHIGAPANLRVVGHAMQGTLKLKTVDAVNPSTPAYTF1TTVKPGA	120
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	SFN1LACYEGIASGVFGVNLRTNFTIKGSFINGACGSPGYVNVRNDGTVEFCYLH1E1GS	169
1P9S:A PDBID CHAIN SEQUENCE	GFN1LACYDGCACQGVFVNHRNTWIRGSFINGACGSPGYVN1LNK-GEVEFVYH1E1GS	168
6LU7:A PDBID CHAIN SEQUENCE	TFSVLACYNGSPGVYQCAMRPNFTIKGSFLNGSCSGVFNIDY-DCVSCFCYH1ME1PT	169
3IWM:A PDBID CHAIN SEQUENCE	TFSVLACYNGSPGVYQCAMRPNHTIKGSFLNGSCSGVFNIDY-DCVSCFCYH1ME1PT	169
3D23:B PDBID CHAIN SEQUENCE	TF1FLAAYINGRPQAFHVTMRSSYT1KGSFLCGSCGSVGVYL1G-DSV1KFVYH1LE1ST	171
5WKK:A PDBID CHAIN SEQUENCE	AFSVLACYNGRPTGFTTVMMRPN1TY1KGSFLCGSCGSVGVY1KEG-SV1NF1CYH1ME1AN	179
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	GAT1GSDFTGSVYGNFDODPSLQESEN1LMSDN1VAFLYA1LNGCRW1L1CSTRVN1VD	229
1P9S:A PDBID CHAIN SEQUENCE	GS1GSSFDGV1HYGGFED1DPL1QVESEN1QMLTVN1VAFLYA1LNGCT1W1LKGE1L1FVEH	228
6LU7:A PDBID CHAIN SEQUENCE	GV1GTDLEG1FYGP1F1D1QTA1QAG1D1T1TV1W1L1W1LY1A1V1NG1DR1W1L1N1F1T1L1D	229
3IWM:A PDBID CHAIN SEQUENCE	GV1GTDLEG1FYGP1F1D1QTA1QAG1D1T1TV1W1L1W1LY1A1V1NG1DR1W1L1N1F1T1L1D	229
3D23:B PDBID CHAIN SEQUENCE	GCH1GTDFT1G1F1G1P1D1Q1A1Q1D1P1V1Q1D1Q1V1T1W1L1W1LY1A1L1N1CA1W1F1Q1D1V1C1S1T1D	231
5WKK:A PDBID CHAIN SEQUENCE	GT1HTGSA1FD1G1D1H1G1A1F1D1Q1V1Q1D1Q1V1L1K1Y1C1S1V1N1V1W1Y1A1L1N1CA1W1F1V1K1P1N1R1T1S1V1S	239
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	FN1WAMANG1T1V1S1V1E--CYS1LA1KT1G1V1S1E1Q1H1-L1H1E1F1G1G1K1N1L1G1Y1S1L1D	286
1P9S:A PDBID CHAIN SEQUENCE	Y1N1E1A1Q1A1G1F1T1M1G1E1D1--A1F1S1L1A1K1T1G1V1C1E1Q1H1-L1H1E1F1G1G1K1N1L1G1Y1S1L1D	285
6LU7:A PDBID CHAIN SEQUENCE	FN1L1V1A1M1C1Y1N1E1P1L1T1Q1D1H1G1S1A1S1L1K1E1Q1H1-Q1N1G1M1R1I1G1S1L1E1D	289
3IWM:A PDBID CHAIN SEQUENCE	FN1L1V1A1M1C1Y1N1E1P1L1T1Q1D1H1G1S1A1S1L1K1E1Q1H1-Q1N1G1M1R1I1G1S1L1E1D	289
3D23:B PDBID CHAIN SEQUENCE	FN1WAMANG1S1Q1V1K1A1L1--V1L1D1A1S1M1T1G1V1S1E1T1L1A1K1R1-L1Y1M1G1Q1R1Q1L1G1S1C1T1D	288
5WKK:A PDBID CHAIN SEQUENCE	FN1WALANQ1T1E1F1V1G1Q1--S1V1D1M1L1A1V1K1T1G1V1A1E1Q1L1Y1A1Q1-L1Y1G1F1Q1R1Q1L1G1S1M1H1D	296
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	EFT1LAEV1K1Q1H1G1V1N1Q1-	303
1P9S:A PDBID CHAIN SEQUENCE	EFS1N1E1V1K1Q1H1G1V1N1Q1-	300
6LU7:A PDBID CHAIN SEQUENCE	EFT1P1D1V1R1Q1C1S1G1V1T1F1Q1-	306
3IWM:A PDBID CHAIN SEQUENCE	EFT1P1D1V1R1Q1C1S1G1V1T1F1Q1-	306
3D23:B PDBID CHAIN SEQUENCE	ELAPSDV1Y1Q1L1A1G1V1--	302
5WKK:A PDBID CHAIN SEQUENCE	EFT1P1D1V1N1Q1M1Q1G1V1M1Q1-	313
	: *: :	

Multiple sequence alignment of Main protease enzyme of various human coronaviruses, including SARS-CoV-2

# Drug repurposing for SARS-CoV-2



# Remdesivir against RdRp



RdRp can incorporate remdesivir, which resembles an RNA building block ATP, into new RNA strands

After binding of remdesivir, RdRp stops being able to incorporate RNA subunits

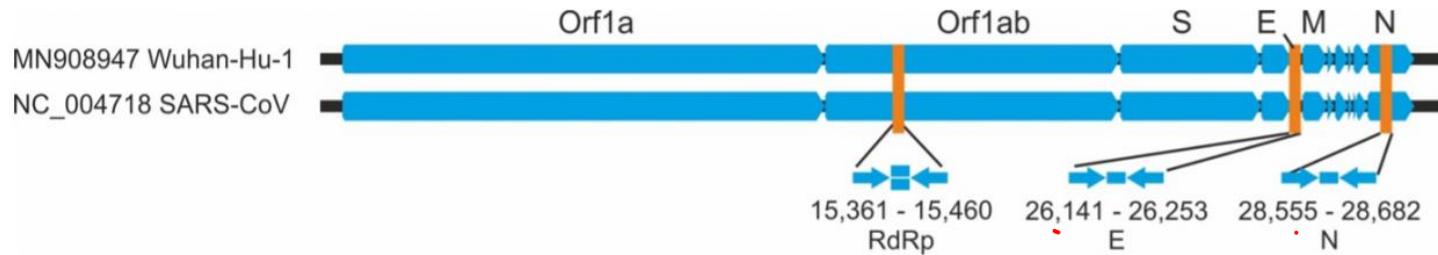
This puts a stop to Coronavirus genome replication

# RTPCR test for SARS-CoV-2

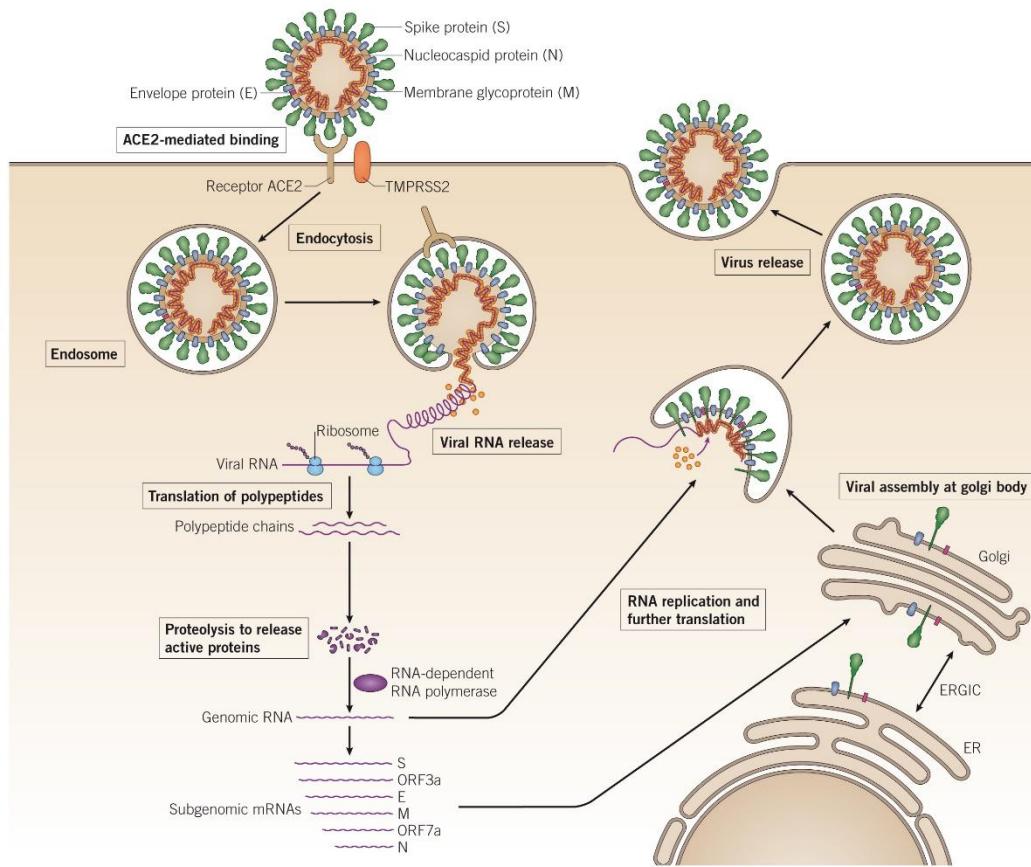
First line screening assay: E gene assay

Confirmatory assay: RdRp gene assay

Additional confirmatory assay: N gene assay



# Transmission of SARS-CoV-2

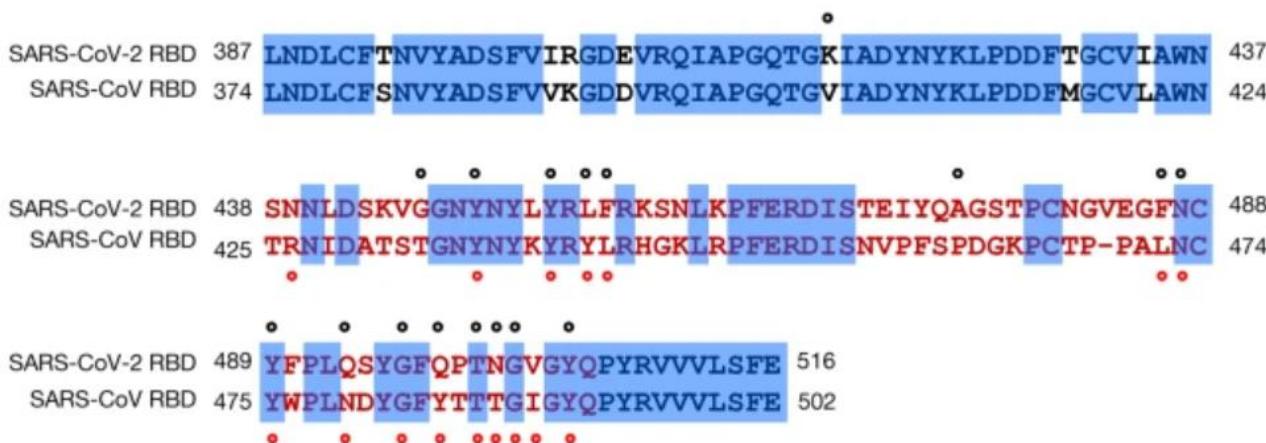


# Transmission of SARS-CoV-2

## Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor

Jun Lan, Jiwan Ge, Jinfang Yu, Sisi Shan, Huan Zhou, Shilong Fan, Qi Zhang, Xuanling Shi, Qisheng Wang,  
Linqi Zhang & Xinquan Wang

*Nature* 581, 215–220 (2020) | [Cite this article](#)



Sequence alignment of the SARS-CoV-2 and SARS-CoV RBDs.