

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

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

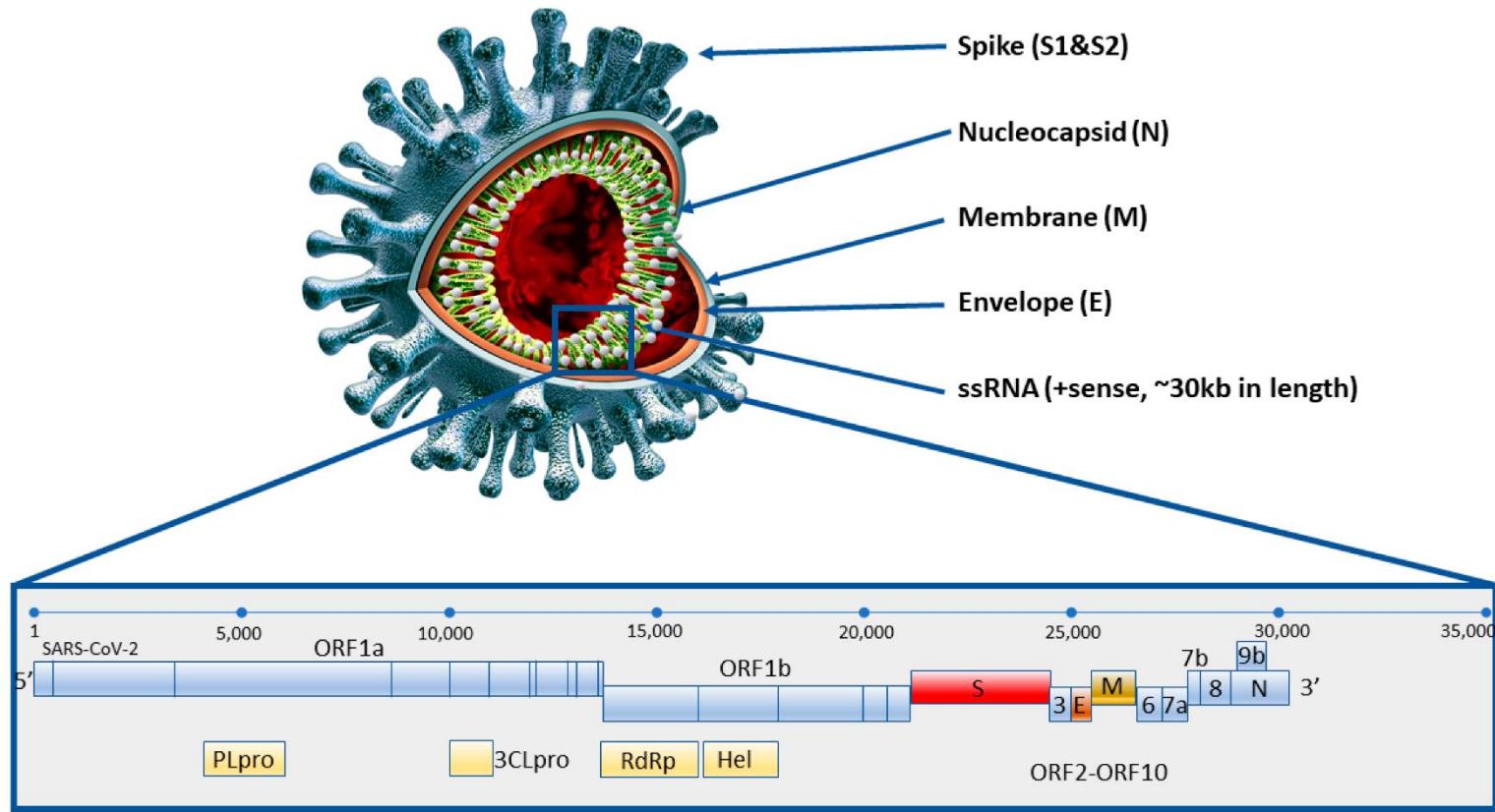
ORIGINAL ARTICLE BRIEF REPORT

A Novel Coronavirus from Patients with Pneumonia in China, 2019

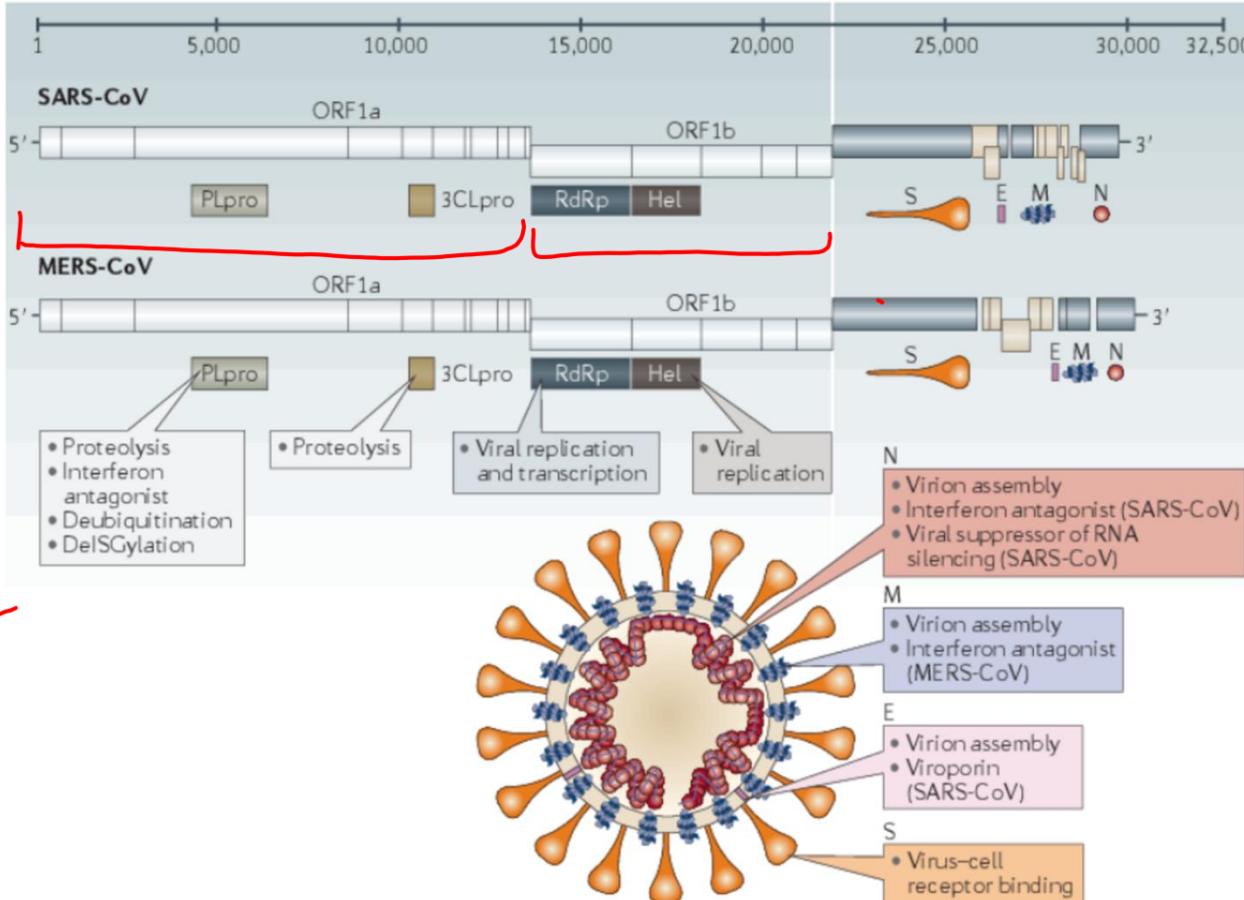
Na Zhu, Ph.D., Dingyu Zhang, M.D., Wenling Wang, Ph.D., Xingwang Li, M.D., Bo Yang, M.S., Jingdong Song, Ph.D., Xiang Zhao, Ph.D., Baoying Huang, Ph.D., Weifeng Shi, Ph.D., Roujian Lu, M.D., Peihua Niu, Ph.D., Faxian Zhan, Ph.D., et al., for the China Novel Coronavirus Investigating and

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

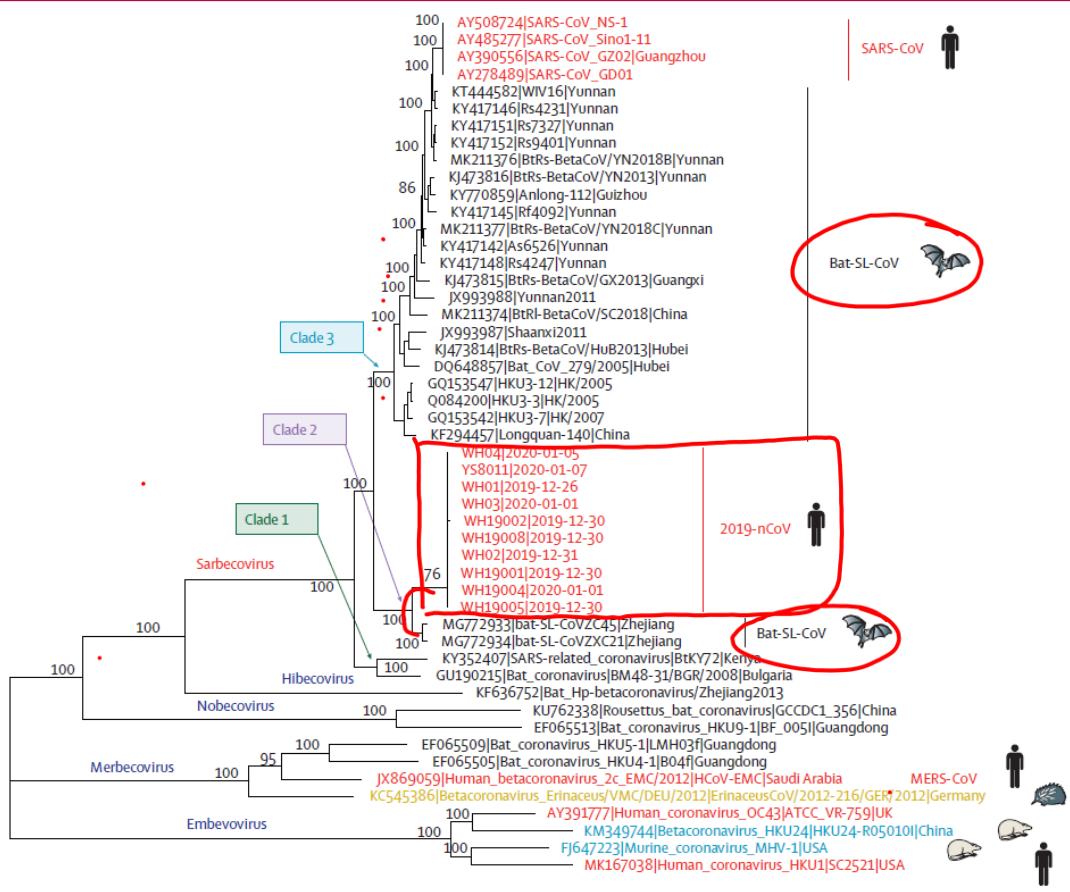
Genomic structure



SARS-CoV-2 proteins



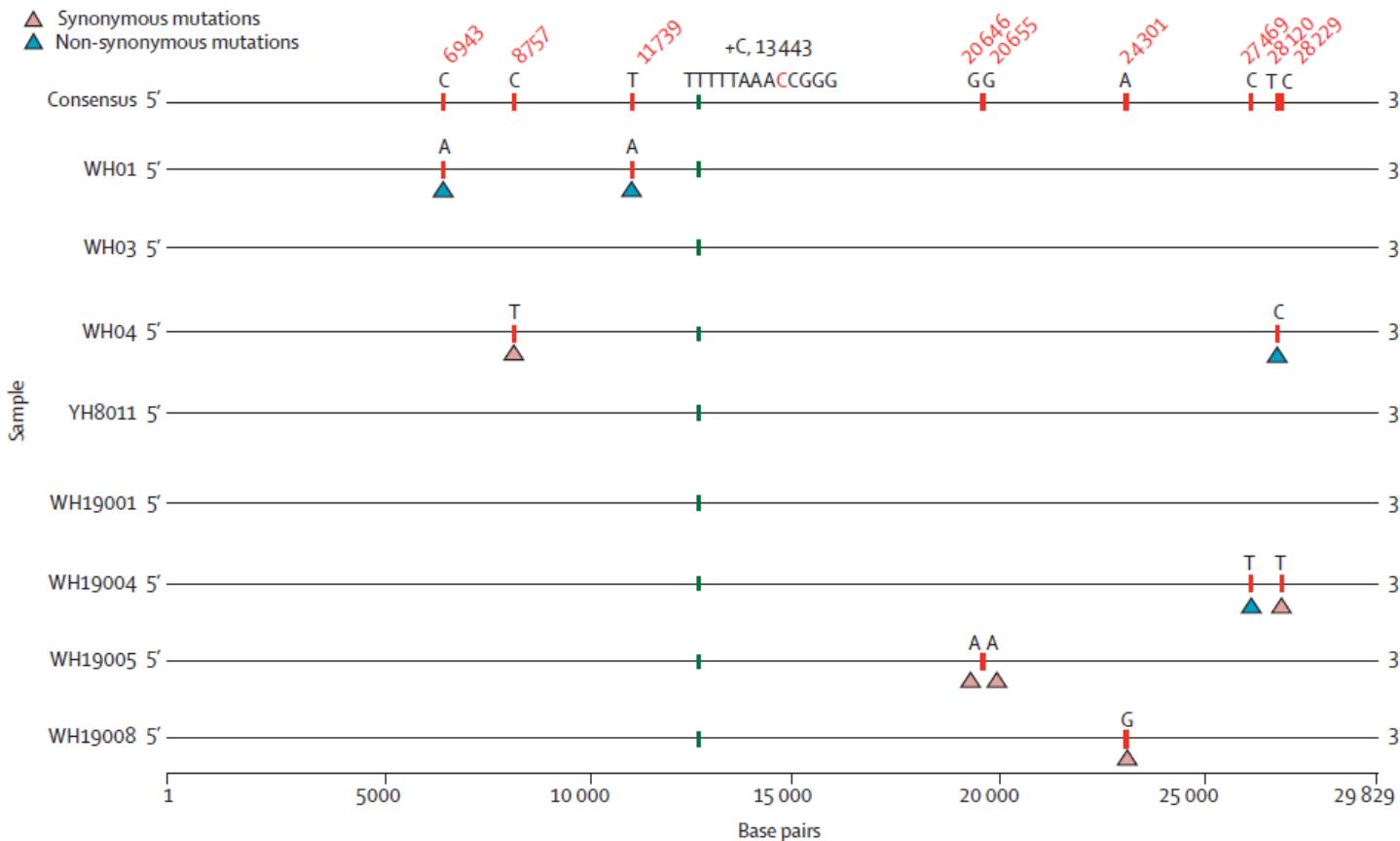
Phylogenetic analysis of SARS-CoV-2



Phylogenetic analysis

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

SARS-CoV-2 variants



Drug repurposing for SARS-CoV-2

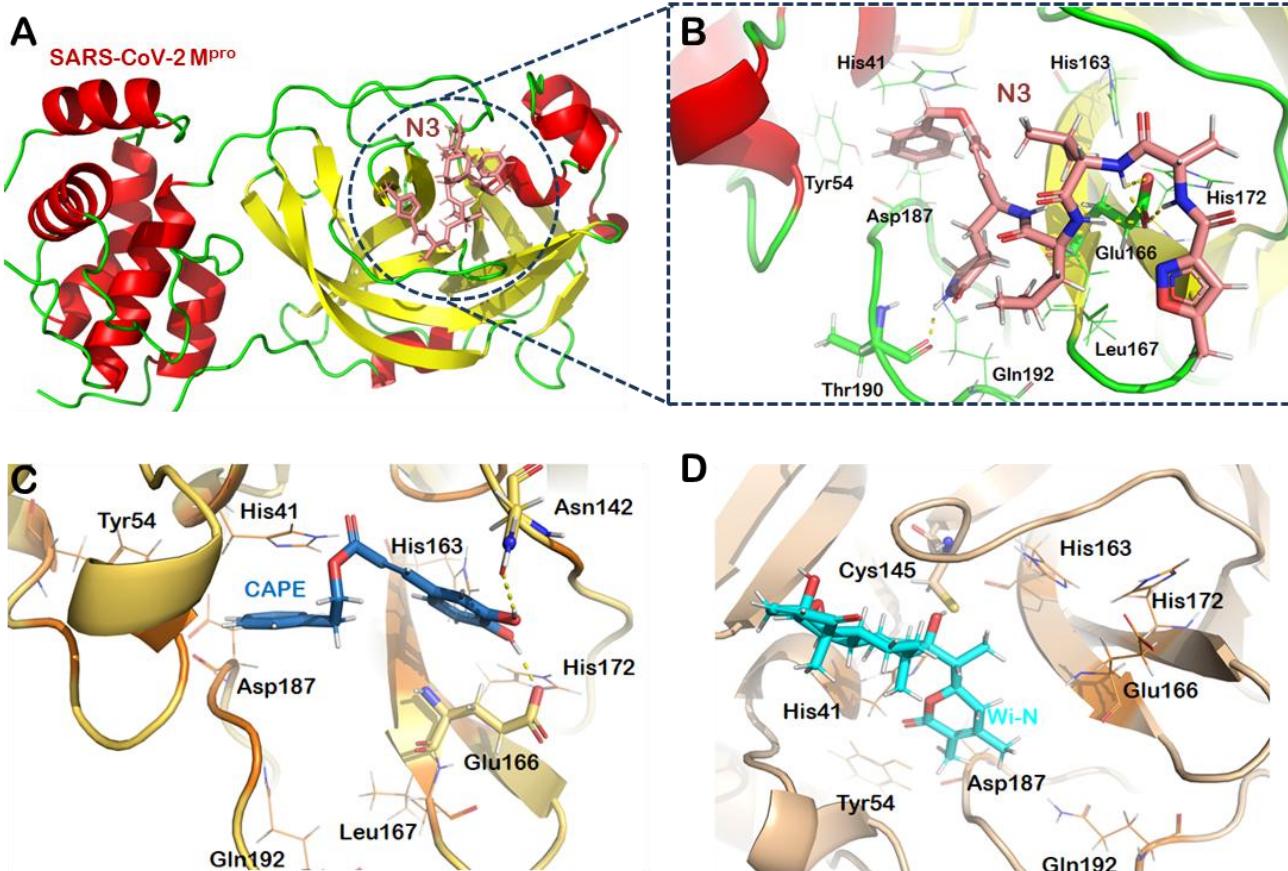
CLUSTAL O(1.2.4) multiple sequence alignment

3TLO:A PDBID CHAIN SEQUENCE	-----SGLKKMAOPSGCVERCIVRVCYGSTVNLNGWILGDTVTCPPFVIAPISTT-VLID	52
1P9S:A PDBID CHAIN SEQUENCE	-----AGLRKMAQPSGFVEKCVRVCYGNTRVLNLGLWDIVYCPRHIASTTT-SAID	52
6LU7:A PDBID CHAIN SEQUENCE	-----SGFRKMAFPSSGVKEGOMVQTCGTTTNGLWLDDVVYCPRHICITSEDMNPN	53
3IWM:A PDBID CHAIN SEQUENCE	-----SGFRKMAFPSSGVKEGOMVQTCGTTTNGLWLDDVVYCPRHICITSEDMNPN	53
3D23:B PDBID CHAIN SEQUENCE	-----ASSGIKMNPSPTSKIEPCIVSVTYGSMTNLGLWLDDTVYCPRHICSSSNNEPD	55
5WKK:A PDBID CHAIN SEQUENCE	MHHHHHHSGLVKMSHPSGGDVACMIVQTCGSMTNLGLWLDDTVYCPRHICPAQDSLDPN	60
	: *: ** *: : *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: *: :	
3TLO:A PDBID CHAIN SEQUENCE	VDHAYSTMRHLNHSVS--HNGVFLGVGVVTMHGSVSLRIKVQSNSVHTPKHVKFTLKGPD	109
1P9S:A PDBID CHAIN SEQUENCE	DHEYTSIMRLNHSVS---SGTAFLGVVGTMHVTLKIKVSQTMHPTRHSFRTLSKGE	109
6LU7:A PDBID CHAIN SEQUENCE	YEDLLIRKSNIHFLVQ---AGNVLRLVIGHSMQNCVLLKLVDTANPKTPKVKFVRIPQGQ	110
3IWM:A PDBID CHAIN SEQUENCE	YEDLLIRKSNIHFLVQ---AGNVLRLVIGHSMQNCVLLKLVDTANPKTPKVKFVRIPQGQ	110
3D23:B PDBID CHAIN SEQUENCE	YSALLCRVTLGDFTIM---SGRMSLTVVSYQMGQLVLTVSLQNPYPTKVTGNGVPKG	112
5WKK:A PDBID CHAIN SEQUENCE	NDALLISMTHHSFSVQKHIGAPANLRVVGHAMQGTLLKTRDVANPSTPAYTFTTVKPGA	120
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	SFNILACYEGIASGVFGVNLRTNTIKGSFINGACGSPGYVNVRNDGTVEFCYLHIEIGS	169
1P9S:A PDBID CHAIN SEQUENCE	GFNILACYDGCACQGVFGVNRNTRNTIAGCSPGYVNLLN-GEVEFVYHIEIGS	168
6LU7:A PDBID CHAIN SEQUENCE	TFSVLACYNGSPGVYQCAMPRNTIKGSFLNGSCSGVFNIDY-DCVSCFYHMEIPT	169
3IWM:A PDBID CHAIN SEQUENCE	TFSVLACYNGSPGVYQCAMPRNTIKGSFLNGSCSGVFNIDY-DCVSCFYHMEIPT	169
3D23:B PDBID CHAIN SEQUENCE	TFTRLAAYNGRPQAFHVTRRSSSYTIKGSFLCGSCGSVGVYLTG-DSVKFVYHMEIPT	171
5WKK:A PDBID CHAIN SEQUENCE	AFSVLACYNGRPTGFTTVMMRPNYTIKGSFLCGSCGSVGYTKLEG-SVINFCYHMEIAN	179
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	GATHGSDFTGSVYGNFDODPSLQESENANLMSDNVAFLYAALNLGCRWMLCSTRNVND	229
1P9S:A PDBID CHAIN SEQUENCE	GSFNGSSFDGVHYGGFEDDPNQLESENANLMSDNVAFLYAALNLGCTWNLKGEKLFEH	228
6LU7:A PDBID CHAIN SEQUENCE	GVHGTDLEGIFYGPFDIQTQAAGTDTTITVNLWALWLYAAVINGDRWFLNRFFTTLND	229
3IWM:A PDBID CHAIN SEQUENCE	GVHGTDLEGIFYGPFDIQTQAAGTDTTITVNLWALWLYAAVINGDRWFLNRFFTTLND	229
3D23:B PDBID CHAIN SEQUENCE	GCHTGTDFTGNYFPGYFDQAVQVLPVQDVKYDQVTVNVIAWLYAAILNWCASFVQNDVCSTED	231
5WKK:A PDBID CHAIN SEQUENCE	GHTGSAFDGTHYGAFFDQVHQQLTDKYCSMVNVVAVLYAAILNGCAWFVKPNRTSVVS	239
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	FNIEWAMANGTYSVSSVE--CYSIILAAKTGVSVEQQLASIQH-LHEGFGGKNILGYSSLCD	286
1P9S:A PDBID CHAIN SEQUENCE	YNEWAQANGFTAMNGED--AFSIIALAATGVVERLHLAIQH-LWNGFGKQILGYSSLND	285
6LU7:A PDBID CHAIN SEQUENCE	FNLVAMCYNEYPLTQDHVDI1GPLSAQTGIAVLDCIASLKELLQNQNGRILGSALEED	289
3IWM:A PDBID CHAIN SEQUENCE	FNLVAMCYNEYPLTQDHVDI1GPLSAQTGIAVLDCIAALKELLQNQNGRILGSALEED	289
3D23:B PDBID CHAIN SEQUENCE	FNIVWAMANGFSQVKADL--VLDALASMTGVSIETLAAIHKR-LYMGFQGRQILGSCTFED	288
5WKK:A PDBID CHAIN SEQUENCE	SWALANQFTEFVGTOQ-SVDMALAVKTGVATEQQLYAIQQ-LYTFQGKQILGSTHLED	296
	: *: :	
3TLO:A PDBID CHAIN SEQUENCE	EFTLAEVVKQHYGVNLQ	303
1P9S:A PDBID CHAIN SEQUENCE	EFSINEVVKQMFGVN--	300
6LU7:A PDBID CHAIN SEQUENCE	EFTPFDVVRQCSGVTFQ	306
3IWM:A PDBID CHAIN SEQUENCE	EFTPFDVVRQCSGVTFQ	306
3D23:B PDBID CHAIN SEQUENCE	ELAPSDVYQQLAGV---	302
5WKK:A PDBID CHAIN SEQUENCE	EFTPEDVNNQIMGVMQ	313
	: *: *: *: *	

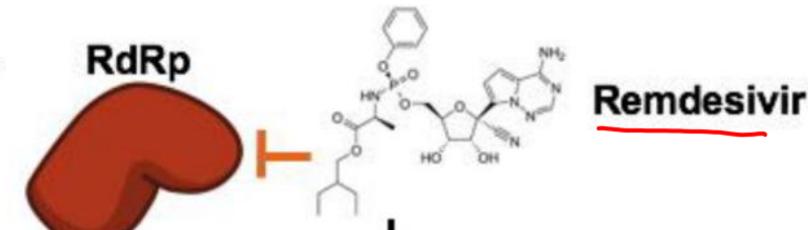
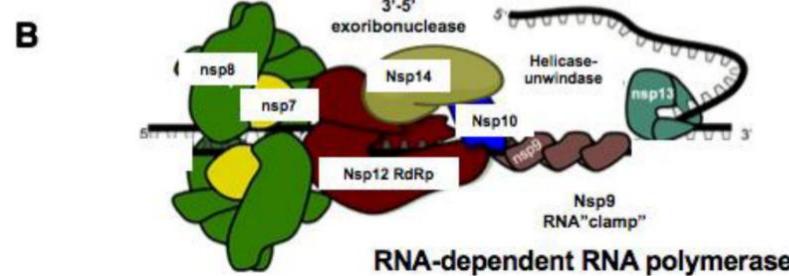
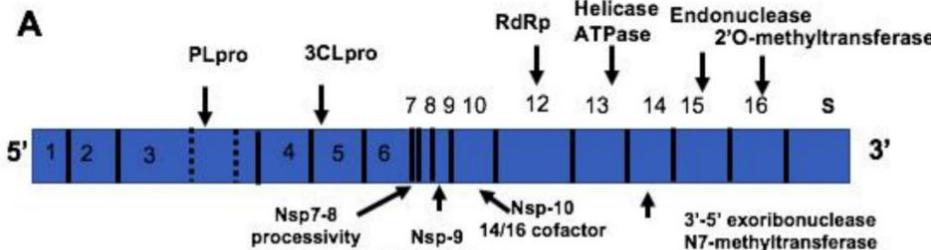
Multiple sequence alignment of Main protease enzyme of various human coronaviruses, including SARS-CoV-2

M *Res*

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

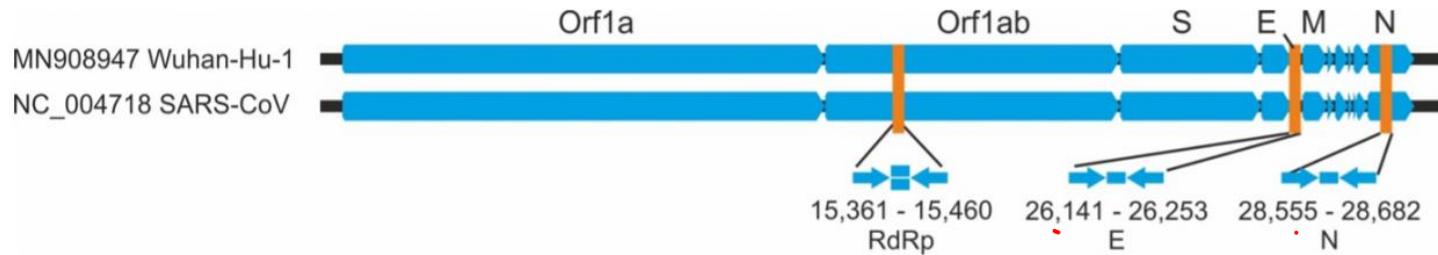
This puts a stop to Coronavirus genome replication

RTPCR test for SARS-CoV-2

First line screening assay: E gene assay

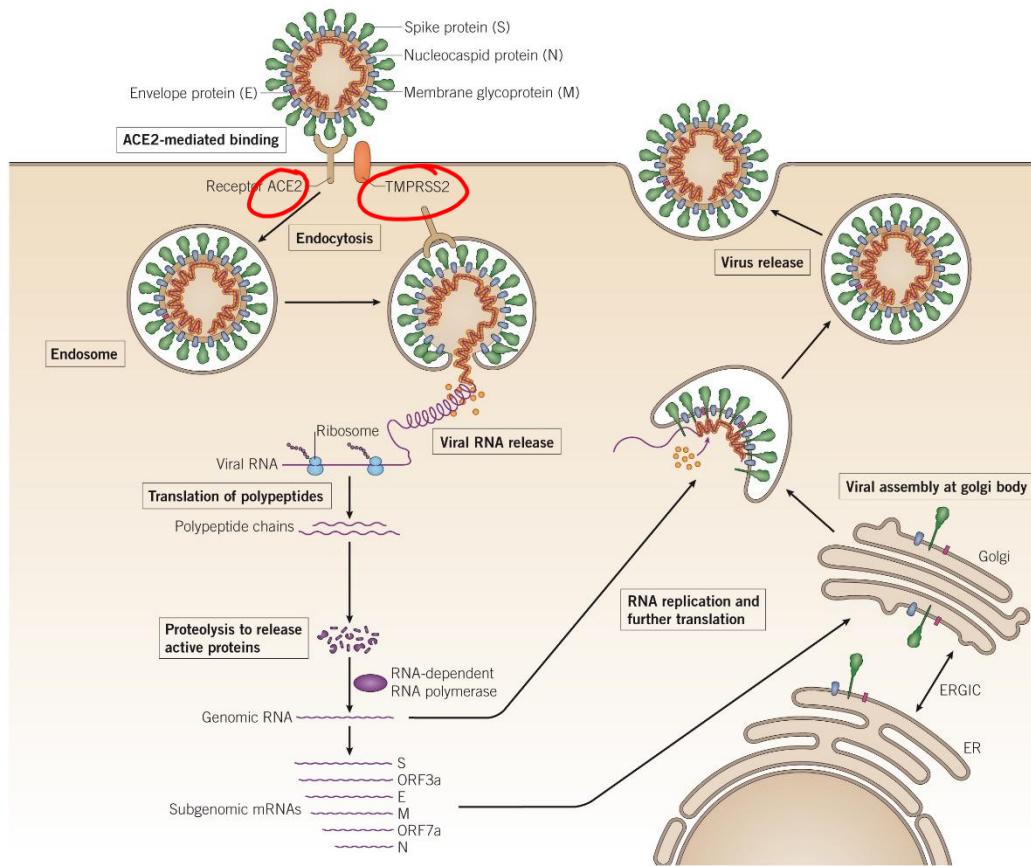
Confirmatory assay: RdRp gene assay

Additional confirmatory assay: N gene assay



Transmission of SARS-CoV-2

31nM
4.1 nM.

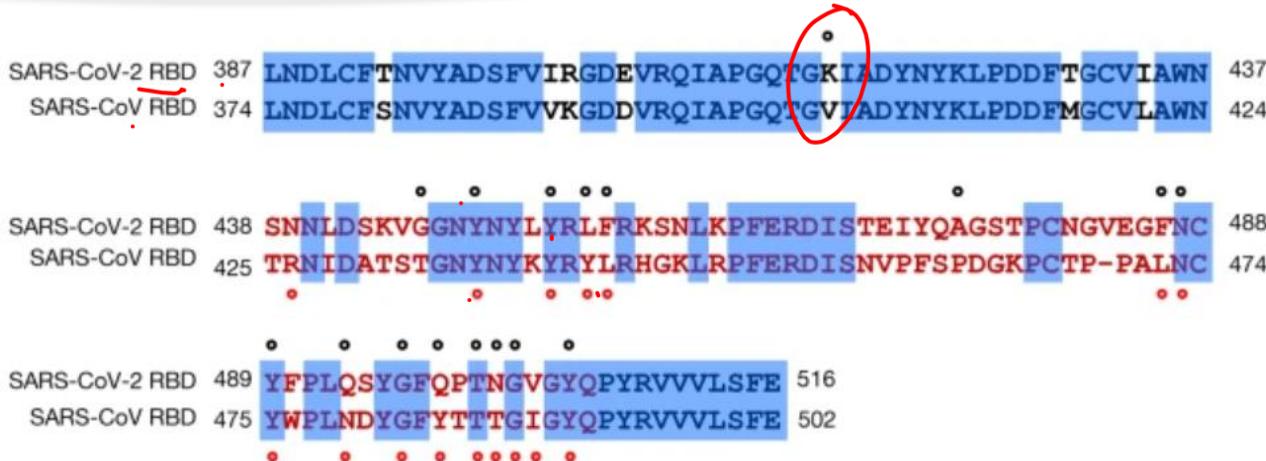


Transmission of SARS-CoV-2

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.