

新冠病例分析



提问



如何快速确认病原体与病毒序列分析。

如何进行大规模检测？

核酸检测和抗体检测的应用。

病毒的突变位点有哪些？这些突变对病毒的感染能力和检测结果有什么影响？

疫苗的种类。

如何进行基因治疗？

QUESTION1

如何快速确认病原体与病毒序列分析。



第一代测序原理

其基本原理是，聚丙烯酰胺凝胶电泳能够把长度只差一个核苷酸的单链DNA分子区分开来。一代测序实验的起始材料是均一的单链DNA分子。

缺点：测序成本高，通量低

第二代测序方法



- 基本原理：边合成边测序
- 宏基因组二代测序的检测流程可以大致分为5个步骤：核酸提取、文库构建、上机测序、生物信息学分析与报告解读

病毒序列分析

NCBI Resources How To Sign in to NCBI

Gene Gene COVID Search

Create RSS Save search Advanced Help

Gene sources Tabular 20 per page Sort by Relevance Send to: Hide sidebar >>

Genomic

Categories

- Alternatively spliced
- Annotated genes
- Non-coding
- Protein-coding

Sequence content

- CCDS
- Ensembl
- RefSeq
- RefSeqGene

Status

- Current

Clear all

Show additional filters

clear

REFERENCE GENOME SEQUENCE Was this helpful? [thumbs up/down]

[Severe acute respiratory syndrome coronavirus 2 \(SARS-CoV-2\) reference genome](#)

[Severe acute respiratory syndrome coronavirus 2 \(Host: human,vertebrates\)](#)

ssRNA(+)

RefSeq: NC_045512.2

[RefSeq genome \(1\)](#) [RefSeq Proteins \(38\)](#) [NCBI SARS-CoV-2 resources](#)

5' 3'

ORF1ab 3a M 7b 10

ORF1a S E 6 N

7a

8

Download

Assembly and annotation statistics +

Filters: [Manage Filters](#)

Results by taxon

Top Organisms [Tree](#)

- Homo sapiens (110)
- Severe acute respiratory syndrome coronavirus 2 (12)
- Chlorocebus sabaeus (2)
- Mus musculus (1)
- Rattus norvegicus (1)
- All other taxa (1)

More...

Find related data

Database: Select

Find items

Search details

COVID[All Fields] AND alive[prop]

Search See more...

Recent activity

基因分类

1. ORF1ab: 编码PP1ab和PP1a两种多蛋白，多蛋白可切割成多种非结构蛋白。例如，SARS-CoV-2非结构蛋白负责病毒转录、复制、蛋白酶处理、宿主免疫反应抑制和宿主基因表达抑制。RNA依赖RNA聚合酶是抗病毒疗法的目标。
2. S: 编码表面糖蛋白（结构蛋白，也叫尖峰蛋白）。这种糖蛋白调解病毒粒子的附着和进入宿主细胞。S蛋白是疫苗开发、抗体疗法和诊断抗原检测的重要靶点。
3. ORF3a: 编码一种具有离子通道活性（维罗戈林）的蛋白质，能够激活NLRP3炎症体（炎症小体的一种，能够识别病原相关分子模式或者宿主来源的危险信号分子，招募和激活促炎症蛋白酶Caspase-1）。ORF3a 也可能在病毒复制和发病机制中发挥作用。

4. E: 编码**包络蛋白**，存在于病毒颗粒中。这种蛋白质有助于病毒的组装和释放，并且具有发病机制所需的离子通道活性。

5. M: 编码**膜糖蛋白**，是一种跨膜糖蛋白，是病毒颗粒中最丰富的结构蛋白，**已被推荐为抗病毒药物靶点**。

6. ORF6a: 编码**ORF6a蛋白**，基于它与其他冠状病毒蛋白的相似性，ORF6蛋白被认为在病毒发病机制中发挥作用。

7. ORF7a: 编码**ORF7a蛋白**，基于它与其他冠状病毒蛋白的相似性，ORF7a蛋白被认为是一种第一型跨膜蛋白。

8. ORF7b: 编码**ORF7b蛋白**，基于与其他冠状病毒的相似性，被认为集中在高尔基体上

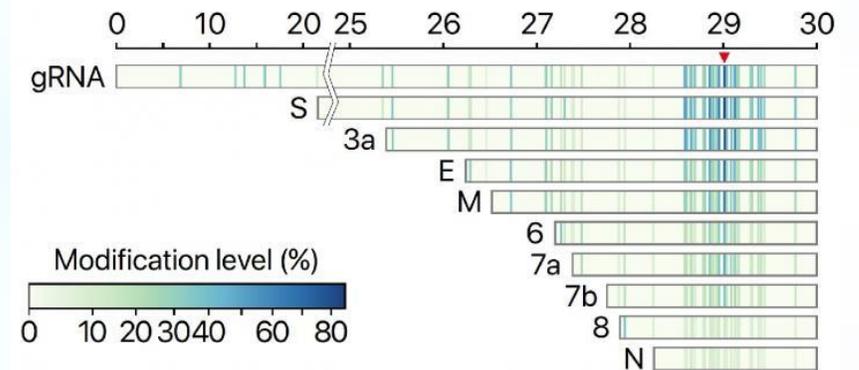
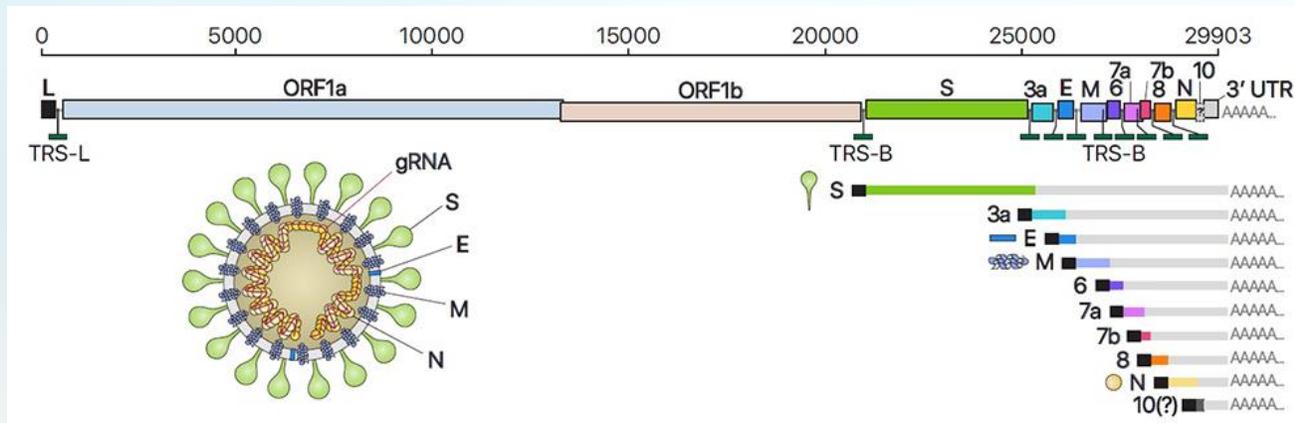
9. ORF8: 编码蛋白功能未知

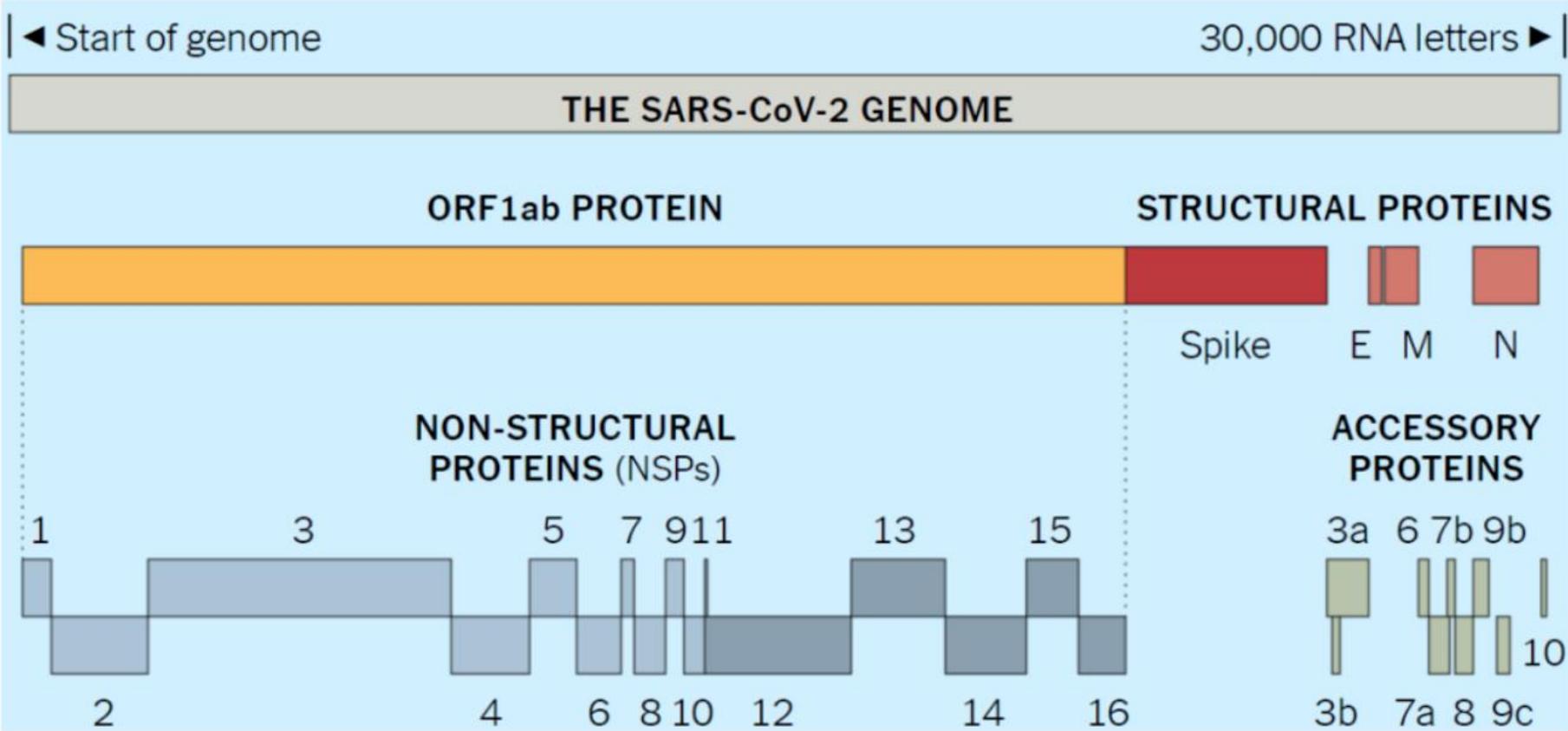
10. N: 编码**核衣壳磷蛋白**，是一种结构蛋白，与病毒RNA基因组结合，保护病毒RNA基因组，并参与将RNA包装成病毒颗粒。**N蛋白已被建议作为抗病毒药物靶点。**

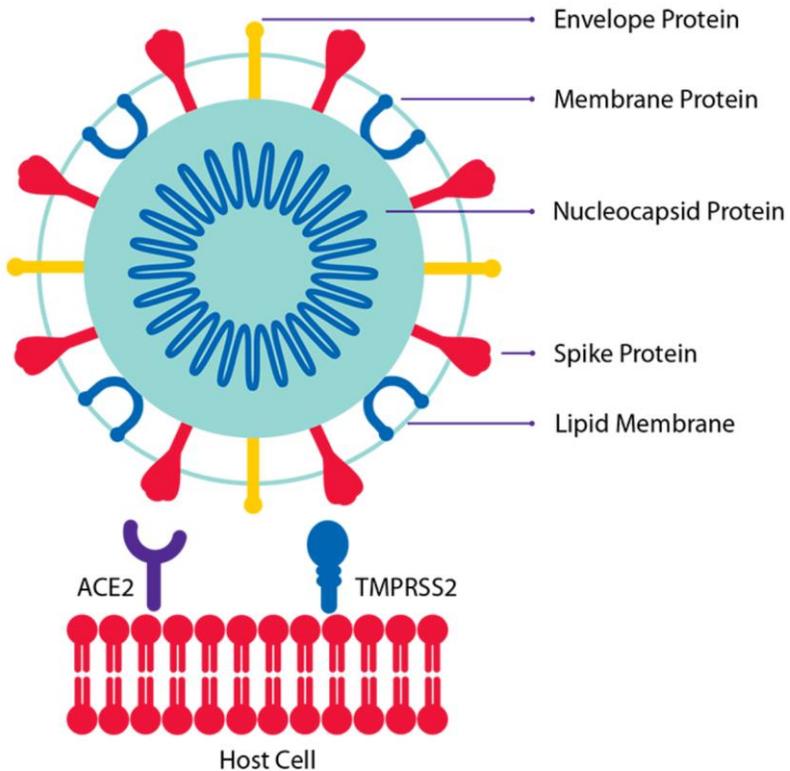
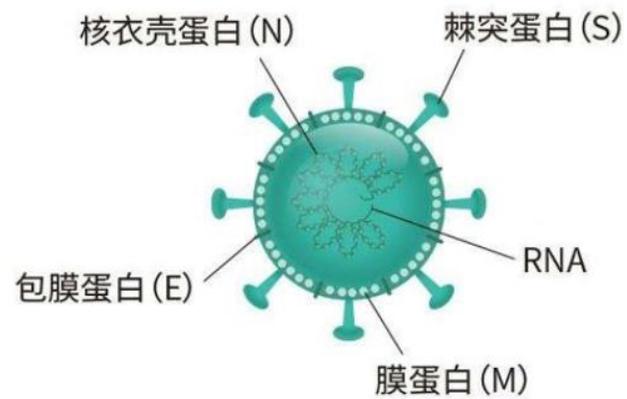
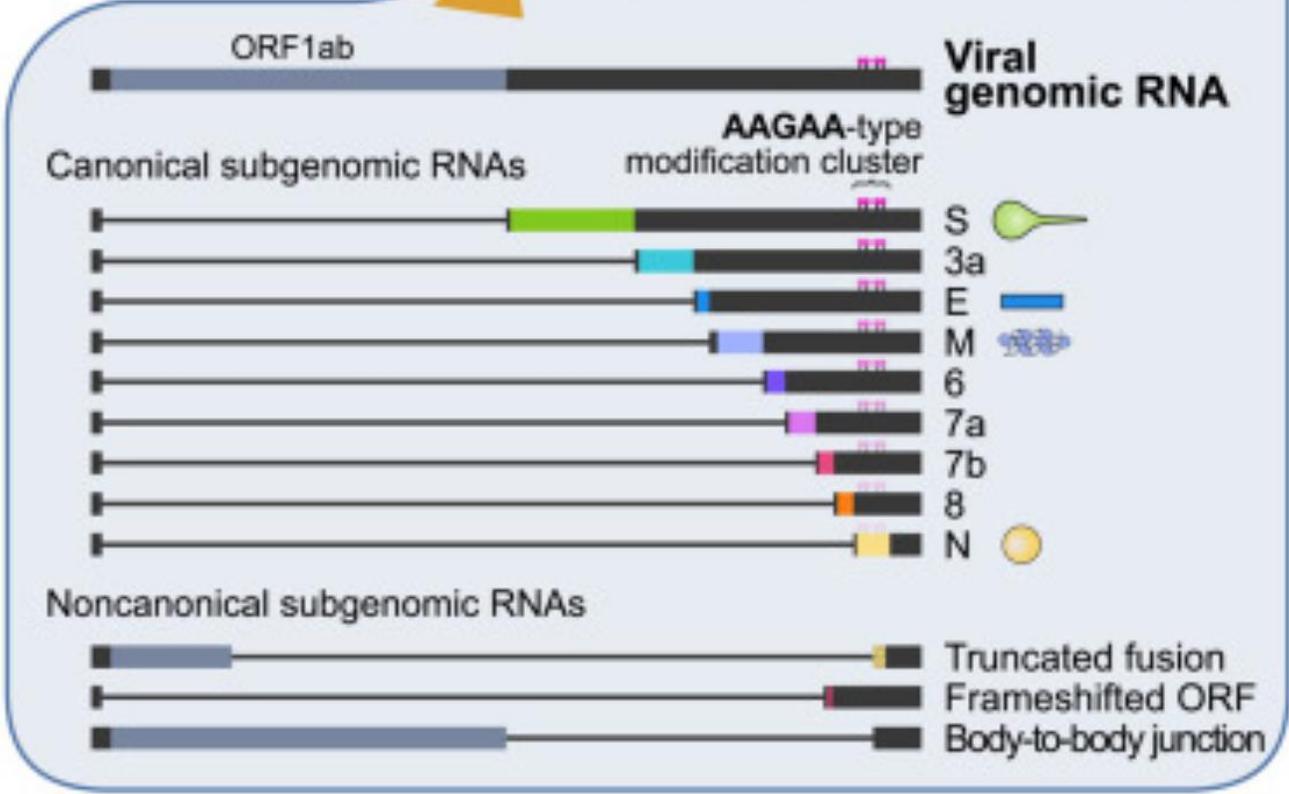
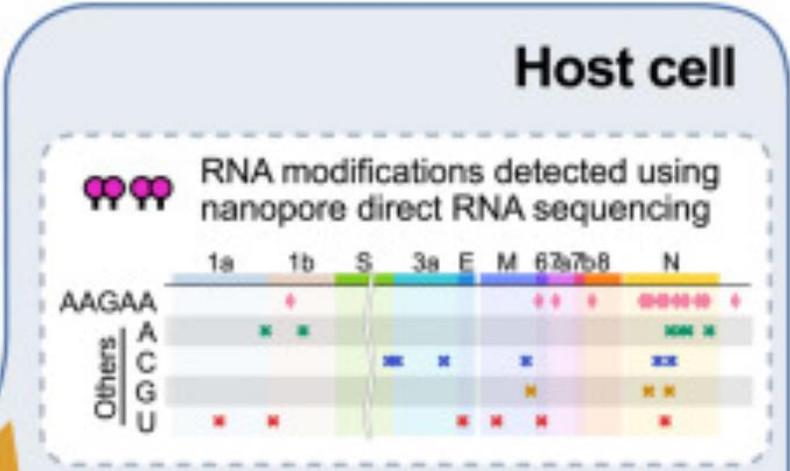
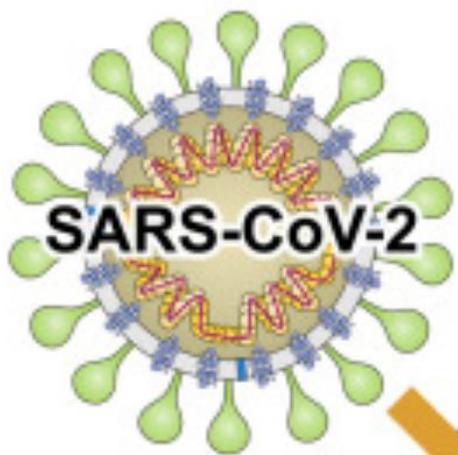
11. ORF10: 编码功能未知

序列分析

此前**已知10个亚基因组RNA构成病毒颗粒结构**。然而，研究小组证实了9个亚基因组RNA的存在，而剩下的一个亚基因组RNA是无效的。研究人员还发现，由于RNA融合和缺失事件，存在数十种未知的亚基因组RNA。“尽管需要进一步研究，但这些分子事件可能导致冠状病毒相对快速的进化。此外，**我们发现病毒RNA有多种未知的化学修饰**。目前尚不清楚这些修改的作用，但有一种可能性是，它们**可能有助于病毒避免来自宿主的攻击**。”







病毒序列分析

```
FEATURES             Location/Qualifiers
source                1..29903
                    /organism="Severe acute respiratory syndrome coronavirus 2"
                    /mol_type="genomic RNA"
                    /isolate="Wuhan-Hu-1"
                    /host="Homo sapiens"
                    /db_xref="taxon:2697049"
                    /country="China"
                    /collection_date="Dec-2019"
5' UTR               1..265
gene                 266..21555
                    /gene="ORF1ab"
                    /locus_tag="GU280_gp01"
                    /db_xref="GeneID:43740578"
CDS                  join(266..13468,13468..21555)
                    /gene="ORF1ab"
                    /locus_tag="GU280_gp01"
                    /ribosomal_slippage
                    /note="pp1ab; translated by -1 ribosomal frameshift"
                    /codon_start=1
                    /product="ORF1ab polyprotein"
                    /protein_id="YP_009724389.1"
                    /db_xref="GeneID:43740578"
                    /translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ
HLKDGTCGLVEVEKGVLPQLEQPYVFKRSRDARTAPHGHVMVELVALEGIQYGRSGE
TLGVLVPHVGEIPVAYRKVLLRKNKNGKAGGHSYGADLKSFDLGDELGTDYPYEDFQEN
WNTKHSVGTRELMRELNGGAYTRYVDNFCGPDGYPLECIKDLLARAGKASCTLSEQ
LDFIDTKRQVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDITFNGECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVSPNECNQMCLSTLMKCDHCGETSQWTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYPACHNSEVGEPEHSLAEYHNESG
LKTILRKGGRITIAFGGCVFYSYVGCNKAAYVWPRASANIGCNHTGVVGESEGLNDL
LEILQKEKVNINIVGDFKLNIEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN
FKVTKGKAKKGAWNIGEQKLSILSPLYAFASEAARVRSIFSRLETAQNSVRVLQKAA
ITILDGISQYSRLRIDAMMFTSDLATNNLVVMAYITGGVQLTSQWLTNIFGTVYEKLV
KPVLDWLEEKFKEGVEFLRDCWEIVKFIISTCACEIVGGQIVTCAKEIKESVQTFPKLV
NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKVKRSREETGLMLPKAPKEII
```

非翻译区域

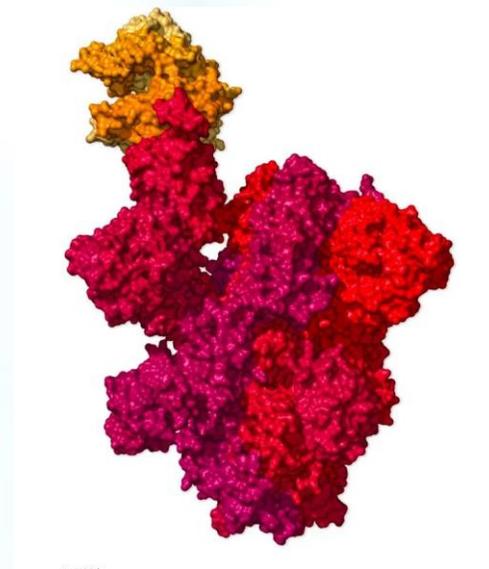
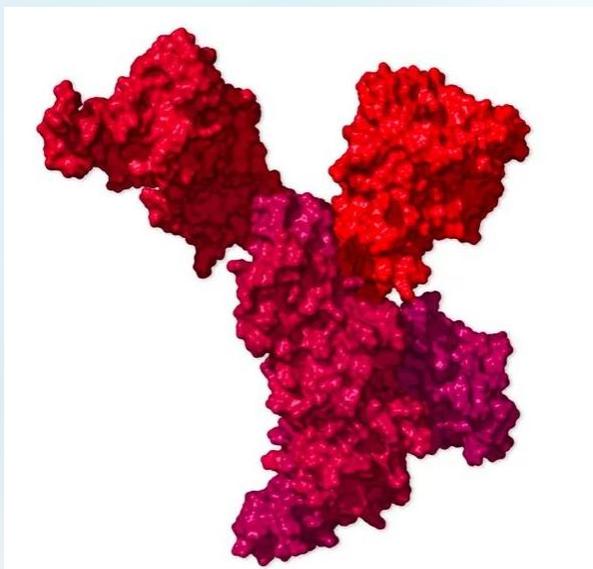
编码序列

```
ORIGIN
1 attaaagggt tataccttcc caggaataca accaaccaac ttctgatctc ttgatagatc
61 gttctotaaa cgaactttaa aatctgtgtg gctgtcactc ggctgcatgc ttatgtcact
121 caccgagtat aattaataac taattactgt cgttgacagg acacgagtaa ctcgtctatc
181 ttctgcagcg tgettacggt ttctgctcgt ttgcagccga tcatcagcac atctaggttt
241 cgtccgggtg tgaccgaaag gtaagatgga gacacctgtc cctggtttca acgagaaaac
301 acacgtccaa ctcagtttgc ctgttttaca ggttcgcgac gctgctgtac gtggcittgg
361 agactccgtg gaggaggtct taccagagcc acgtcaacat cttaaagatg gcacctgtgg
421 cttagtagaa gttgaaaaa gctgttttgc tcaacttgaa cagccctatg ttttcatcaa
481 acgttcggat gctcgaactg cacctcatgg tcatgttatg gttgagctgg tagcagaact
541 cgaaggcatt cagtaacggt gtagtggtga gacacttggg gctctgttcc ctcactgtgg
601 cgaataacca gtggcttacc gcaaggttct tcttctgaag aacggtaata aaggagctgg
661 tggccatagt tacggcgccg atctaaagtc atttgactta ggcgcagcgc ttggcactga
721 tcctatgaa gattttcaag aaaactggaa cactaaacat agcagttgtg ttaccctgta
781 actcatgctg gagcttaacg gaggggcata cactcgtatg gtcgataaca acttctgtgg
841 cctcatgctg tacctcttgg agtgcatcaa agaccttcta gcactgtgctg gtaaagcttc
901 atgcactttg tccgaacaac tggactttat tgacactaag aggggtgtat actgctgccg
961 tgaacatgag catgaaattg ctgtgtacac ggaactttct gaaaagagct atgaattgca
1021 gacacctttt gaaattaaat tggcaaaaga atttgacacc ttcaatgggg aatgtccaaa
1081 ttttgtattt cccttaaaat ccataatcaa gactattcaa ccaagggttg aaaaagaaaa
1141 gcttgatggc tttatgggta gaattcgtatc tgtctatcca gttgcgtcac caaatgaatg
1201 caacaaaatg tgcctttcaa ctctcatgaa gttgtgatcat tgtgtgtaaa cttcatggca
1261 gacggcgcat tttgttaaag ccacttgcca attttggcc actgagaatt tgactaaaga
1321 aggtgccact acttgtggtt acttaccoca aaatcgtgtt gttaaaattt atgtccagc
1381 atgtcaacat tcagaagtag gacctgagca tagtcttggc gaataccata atgaatctgg
1441 cttgaaaacc attcttcgta aggggtgctg cactattgcc tttggaggct gtgtgttctc
1501 ttatgttggg tgccataaca agtgtgccta ttgggttcca cgtgctagcg ctaacatagg
1561 ttgtaacatc acaggtgttg ttggagaagg ttccgaaggt cttaatgaca accttcttga
1621 aatactccaa aaagagaaag tcaacatcaa tattgttggg gactttaaac ttaatgaaga
1681 gatcgcocatt attttggcat cttttctgct ttccacaagt gcttttggg aaactgtgaa
1741 aggtttggat tataaagcat tcaacaaaat tgttgaatcc tgtgtgaatt ttaaagttac
1801 aaaaagaaaa gctaaaaaaag gtgcctggaa tattgtgtaa cagaaatcaa tcaatgagtc
1861 tcittatgca tttgcatcag aggcgtgctg tttgttagca tcaattttct cccgcaactc
1921 tgaactgctc caaaattctg tgcgtgtttt acagaaggcc gctataacaa tactagatgg
1981 aatttcacag tattcactga gactcattga tgctatgatg ttcacatctg atttggctac
2041 taacaactca gttgtaatgg cctacattac aggtgtgtgt gttcagttga cttcgcagtg
2101 gctaactaac atctttggca ctgtttatga aaaaactcaa cccgtccttg atttgcttga
```

DNA转录成mRNA，mRNA经剪接等加工后翻译出蛋白质，所谓CDS就是与蛋白质序列一一对应的DNA序列，且该序列中间不含其它非该蛋白质对应的序列，不考虑mRNA加工等过程中的序列变化

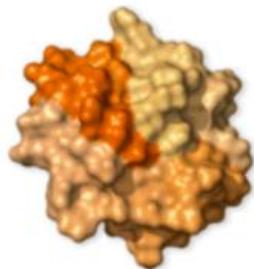
刺突蛋白·S

刺突蛋白是形成冠状病毒外层，并保护内部RNA的四种结构蛋白S，E，M和N之一的蛋白。通过形成三聚体，S蛋白在病毒表面上形成明显的刺突。刺突的一部分可以延伸并附着在ACE2的蛋白质上（图中为黄色），然后病毒可以入侵细胞。**SARS-CoV-2**的刺突蛋白基因中插入了12个碱基：**ccucggcgggca**。这种突变或有助于刺突与人类细胞更加紧密结合。目前许多科研团队正在设计药物用来阻止刺突蛋白附着在人体细胞上。



包膜蛋白•E

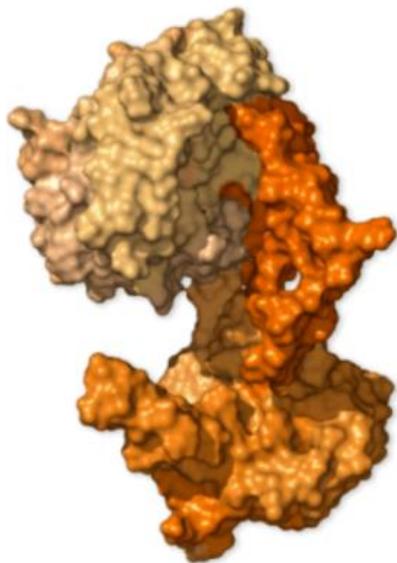
包膜蛋白是一种结构蛋白。一旦病毒进入细胞内部，它可以锁定蛋白质，从而帮助打开和关闭被感染细胞的基因组。



包膜蛋白

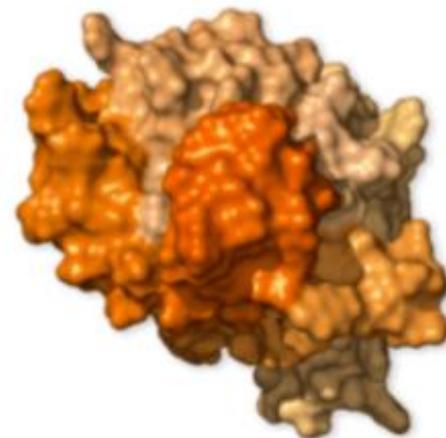
核衣壳蛋白•N

N蛋白可以保护病毒RNA，使其在病毒内部保持稳定。许多N蛋白以长螺旋形连接在一起，包裹RNA。



膜蛋白•M

形成病毒外壳的另一种结构蛋白。

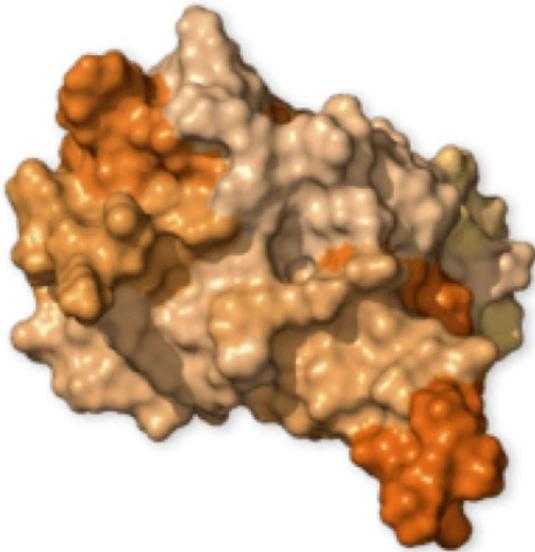


ORF1编码的 16个蛋白

被SARS-CoV-2感染的细胞内部产生的第一个病毒蛋白实际上是16种蛋白质结合在一起的蛋白链(见上图)。其中的2两种蛋白像剪刀一样把16种蛋白从蛋白链上释放出来。

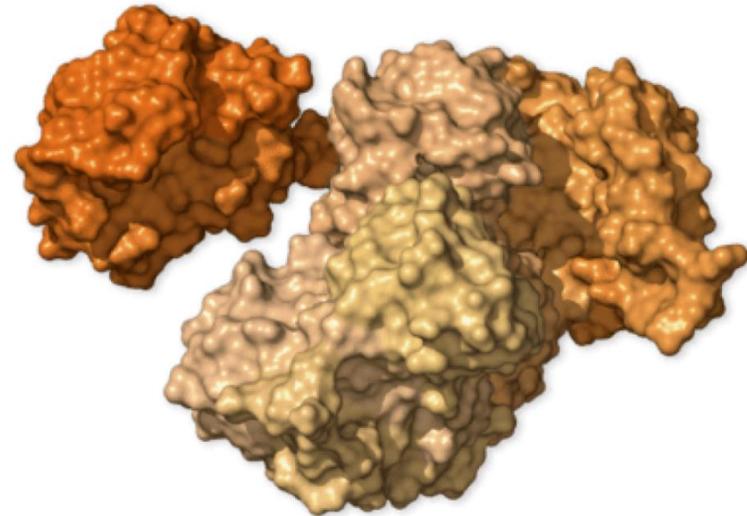
细胞破坏分子•NSP1

这种蛋白质会减慢受感染细胞自身蛋白质的产生。这种破坏行为会迫使细胞产生更多的病毒蛋白。



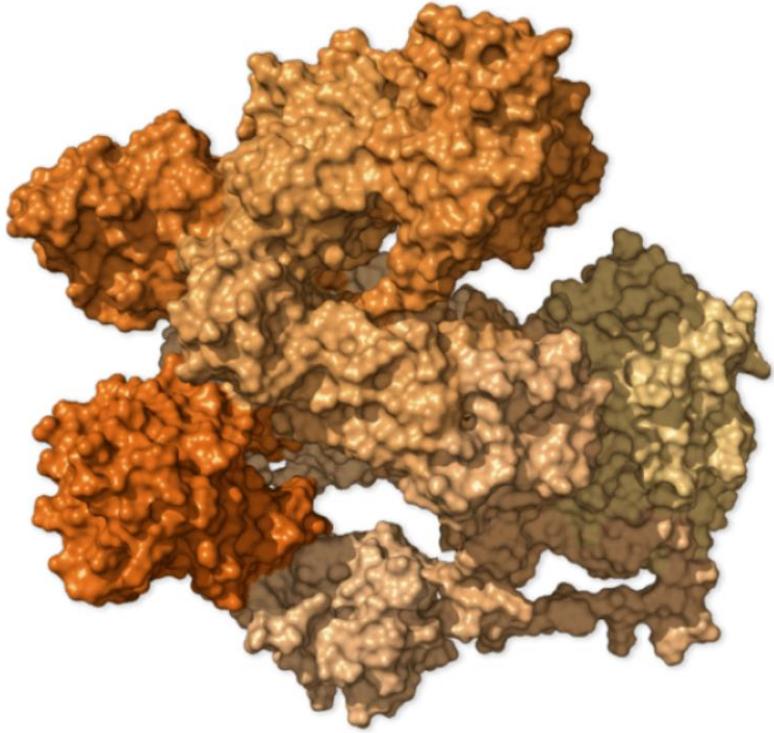
神秘蛋白•NSP2

尚不能确定NSP2的功能。它附着的蛋白质可能会提供一些线索，其中两个有助于内体的移动。



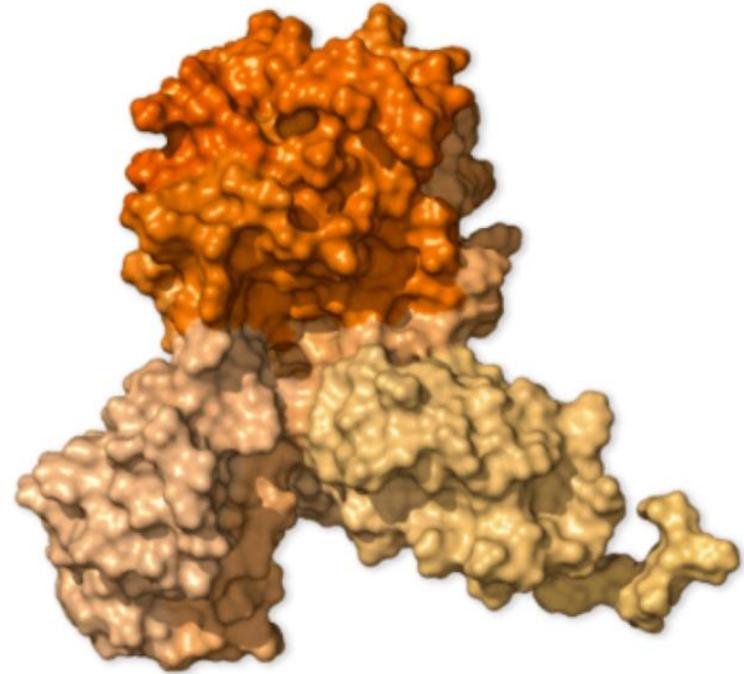
取消标记和切割蛋白·NSP3

NSP3是一种具有两个重要功能的蛋白质，一是切割其他病毒蛋白，以便行使各自的功能。它还会改变许多被感染细胞的蛋白质。通常健康的细胞会标记旧蛋白质以进行降解。但是SARS-CoV-2病毒可以去除这些标签，改变蛋白质的平衡，降低细胞抵抗病毒的能力。



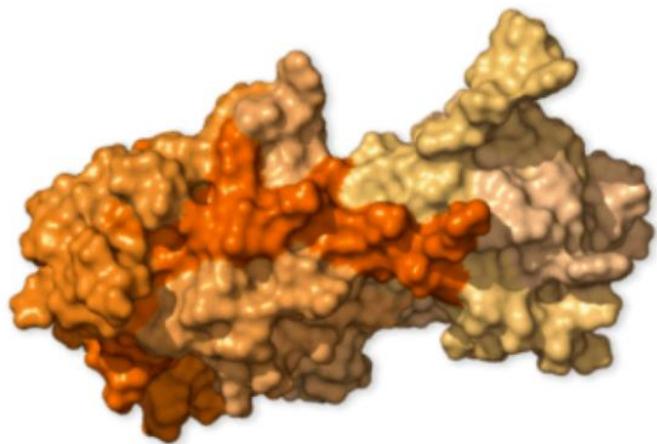
液泡制造蛋白·NSP4

NSP4与其他蛋白质结合，有助于在受感染的细胞内形成充满液体的泡泡结构，这里也是构建部分新病毒的场所。



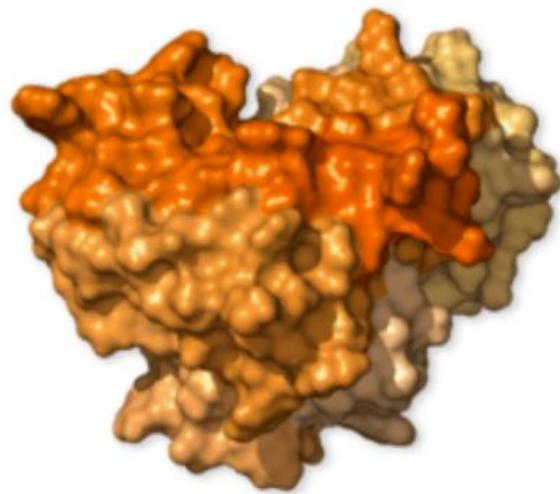
蛋白质剪刀·NSP5

这种蛋白质可以使大多数其他NSP蛋白质从蛋白链上解放出来。



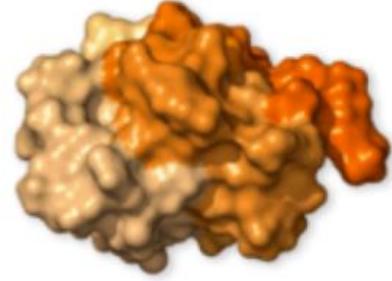
液泡工厂·NSP6

与NSP3和NSP4配合构建生产病毒的泡泡结构。

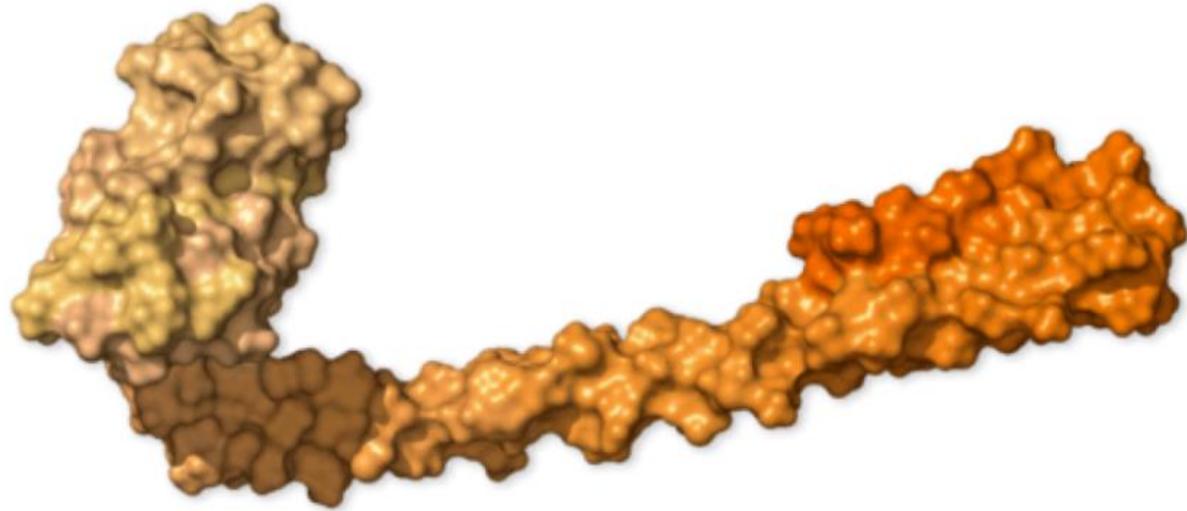


复制助手·NSP7和NSP8

这两种蛋白质可以帮助NSP12生成RNA基因组的新拷贝，用于组装新病毒。

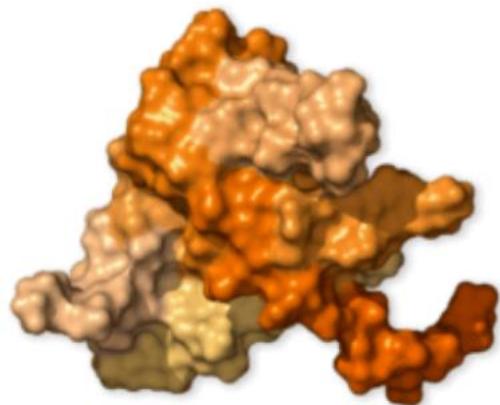


NSP7



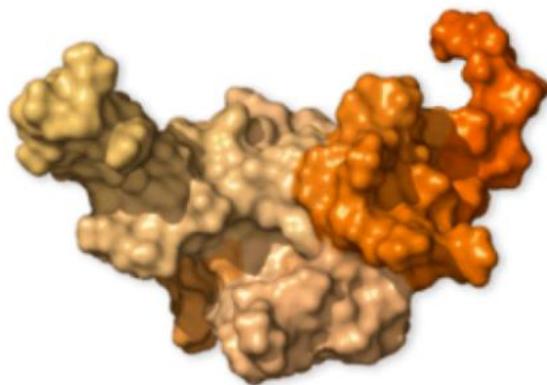
细胞中心·NSP9

这种蛋白质会通过微小通道渗入被感染细胞的核中，它可能能够影响分子进出核的运动-但是出于什么目的，还未知。



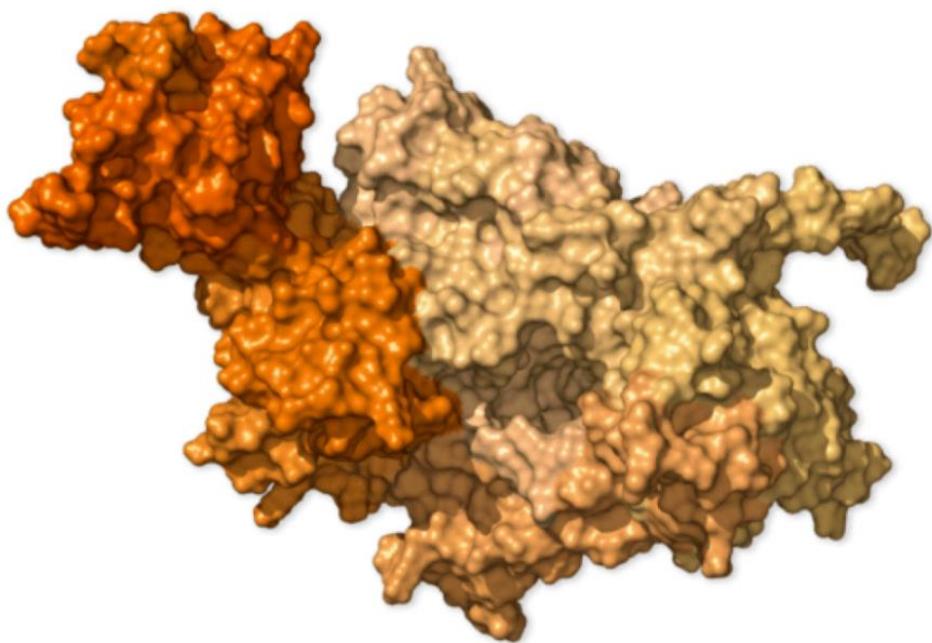
伪装蛋白·NSP10

人类的细胞具有抗病毒蛋白，可以识别病毒RNA并将其降解。NSP10可与NSP16结合，以伪装病毒的基因组，从而不会受到攻击。



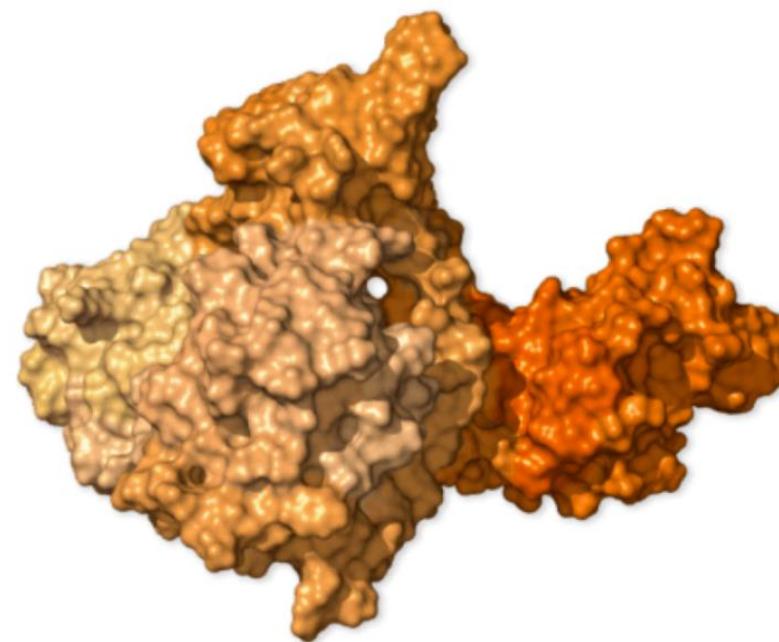
复制机器•NSP12

这种蛋白质协助病毒基因组复制。研究人员已经发现在其他冠状病毒中，抗病毒药remdesivir会干扰NSP12。另一个序列NSP11与编码NSP12的RNA的一部分重叠。但是尚不清楚该基因编码的微小蛋白质是否具有功能。



RNA解开蛋白•NSP13

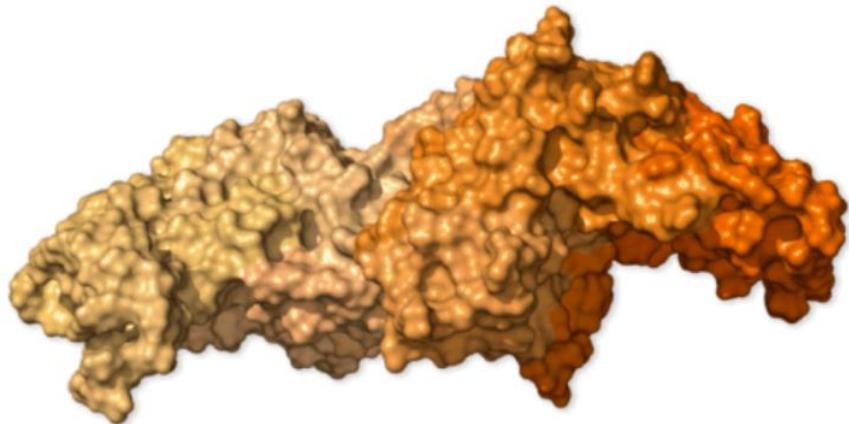
病毒RNA会缠绕起来。研究学者推测NSP13可以解开缠绕的RNA链，以便复制或者蛋白翻译。



NCD12

病毒校对器·NSP14

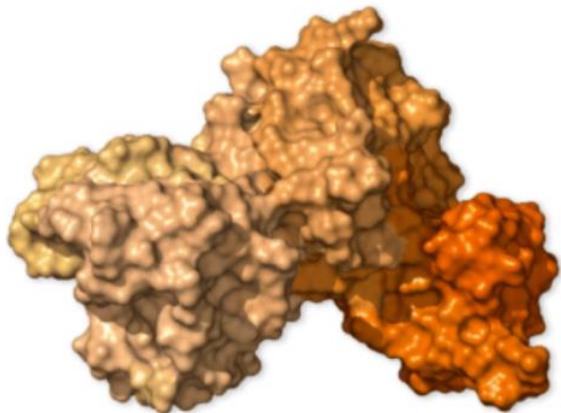
NSP12协助病毒基因组复制时，有时会出错。NSP14可以纠正这些错误。



NSP14

清理蛋白·NSP15

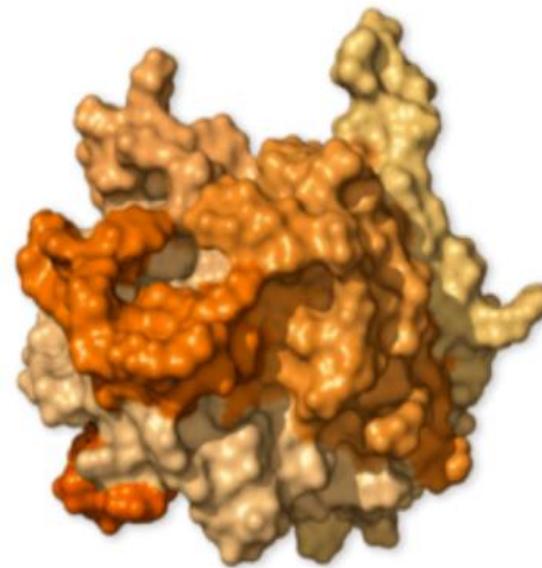
研究人员推测这种蛋白质会切碎剩余的病毒RNA，以躲避机体的抗病毒防御反应。



NSP15

隐藏蛋白·NSP16

NSP16与NSP10一起可将病毒的基因组隐藏起来，以躲避切割病毒RNA的蛋白质。



如何进行大规模检测？



- 大规模新冠病毒的检测可以通过基于RT-qPCR的试剂盒进行。qPCR诊断耗时短，检测样本多，结果判定简单快速，检测成本低，特别适合本次大面积流行的新冠肺炎疫情。CT影像诊断能作为辅助手段检测新冠病毒，在一定程度上规避假阴性。

qPCR推荐选用针对新型冠状病毒的开放读码框1ab（open reading frame,ORF1ab）、核壳蛋白（nucleoprotein, N）基因区域的引物和探针。



中国疾病预防控制中心
病毒病预防控制所
National Institute For Viral Disease Control and Prevention

首页 机构信息 党群建设 疾病控制 科技工作 教育培训 实验室管理 国际合作

新型冠状病毒核酸检测引物和探针序列 (Specific primers and probes for detection 2019 novel coronavirus)
来源：病毒病所 发布时间：2020-01-21

检测方法

新型冠状病毒核酸检测（实时荧光RT-PCR方法）

推荐选用针对新型冠状病毒的开放读码框1ab（open reading frame,ORF1ab）、核壳蛋白（nucleoprotein, N）基因区域的引物和探针。

Target 1（ORF1ab）：

正向引物（F）：CCCTGTGGGTTTTACACTTAA

反向引物（R）：ACGATTGTGCATCAGCTGA

荧光探针（P）：5'-FAM-CCGTCTGCGGTATGTGGAAAGGTTATGG-BHQ1-3'

Target 2（N）：

正向引物（F）：GGGGAAGTTCTCCTGCTAGAAT

反向引物（R）：CAGACATTTTGCTCTCAAGCTG

荧光探针（P）：5'-FAM-TTGCTGCTGCTTGACAGATT-TAMRA-3'

[Download](#) ▾

[GenBank](#) [Graphics](#)

Sort by: E value



[◀ Descriptions](#)

Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/UT-UPHL-2101998754/2020, complete genome

Sequence ID: [MW476657.1](#) Length: 29901 Number of Matches: 2

Range 1: 23591 to 23610 [GenBank](#) [Graphics](#)

▾ [Next Match](#) ▲ [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	0.36	20/20(100%)	0/20(0%)	Plus/Plus
Query 1	GACTAATTCTCCTCGGCGGG	20		
Sbjct 23591	GACTAATTCTCCTCGGCGGG	23610		

Range 2: 24184 to 24203 [GenBank](#) [Graphics](#)

▾ [Next Match](#) ▲ [Previous Match](#) ▲ [First Match](#)

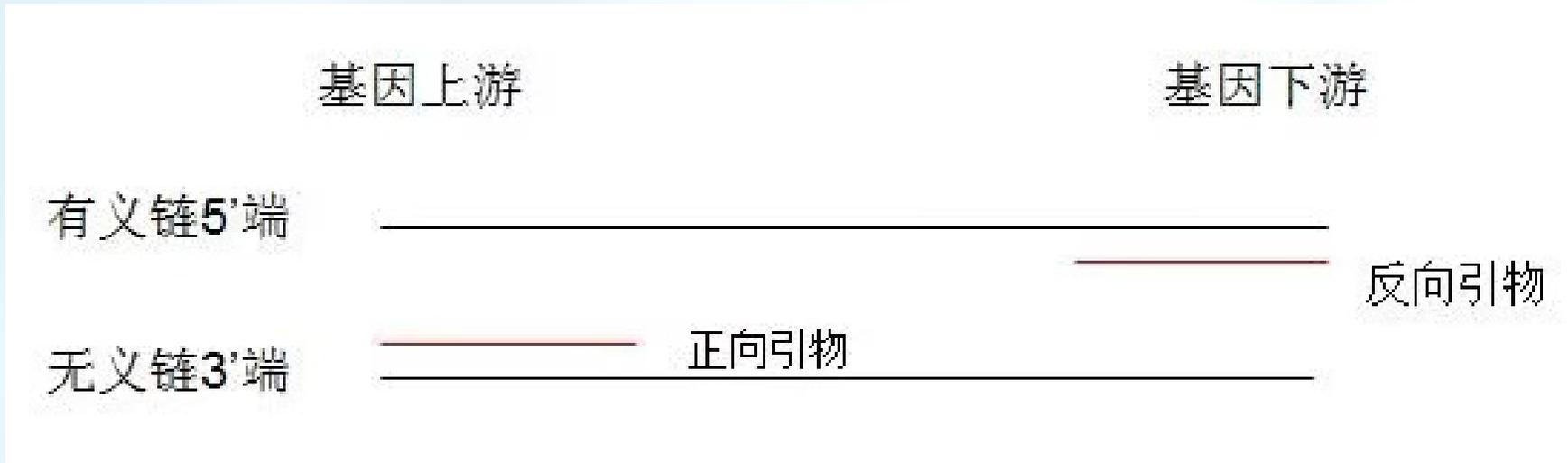
Score	Expect	Identities	Gaps	Strand
37.4 bits(40)	0.36	20/20(100%)	0/20(0%)	Plus/Minus
Query 51	TGTACCCGCTAACAGTGCAG	70		
Sbjct 24203	TGTACCCGCTAACAGTGCAG	24184		

Target 1 (ORF1ab) :

正向引物 (F) : CCCTGTGGGTTTTACACTTAA (20个碱基)

反向引物 (R) : ACGATTGTGCATCAGCTGA (18个碱基)

荧光探针 (P) : 5'-FAM-CCGTCTGCGGTATGTGGAAAGGTTATGG-BHQ1-3



由稀释公式(稀释到10 μ M) : $V_{\text{水}}/\mu\text{L} = \text{OD数} \times 33 \times 100000 / \text{分子量}$
引物分子量为6051, 以1 OD计算, 稀释到10 μ M需加**545 μ L**双蒸水。

温度: 引物 $T_m = 4^\circ\text{C}(\text{G} + \text{C}) + 2^\circ\text{C}(\text{A} + \text{T})$, 经计算, 本次使用的正向引物的 T_m 值为58 $^\circ\text{C}$, 根据 $T_m - 5$ 得53 $^\circ\text{C}$, 反向引物的 T_m 值为56 $^\circ\text{C}$, 根据 $T_m - 5 = 51^\circ\text{C}$ 。

检测方法

结果判断:

阴性:无Ct值或Ct为40。

阳性: Ct值<37, 可报告为阳性。

可疑: Ct值在37-40之间, 建议重复实验, 若重做结果Ct值<40, 扩增曲线有明显起峰, 该样本判断为阳性, 否则为阴性。

直接检测

目前实验室广泛应用的针对病毒的核酸检测，检测新型冠状病毒基因中某些特定核酸序列的存在，是目前公认的新型冠状病毒肺炎实验室诊断金标准方法。



核酸检测

优点：核酸检测的**阳性预测值可达100%**，也就是说核酸检测的**假阳性率极低**。

缺点：

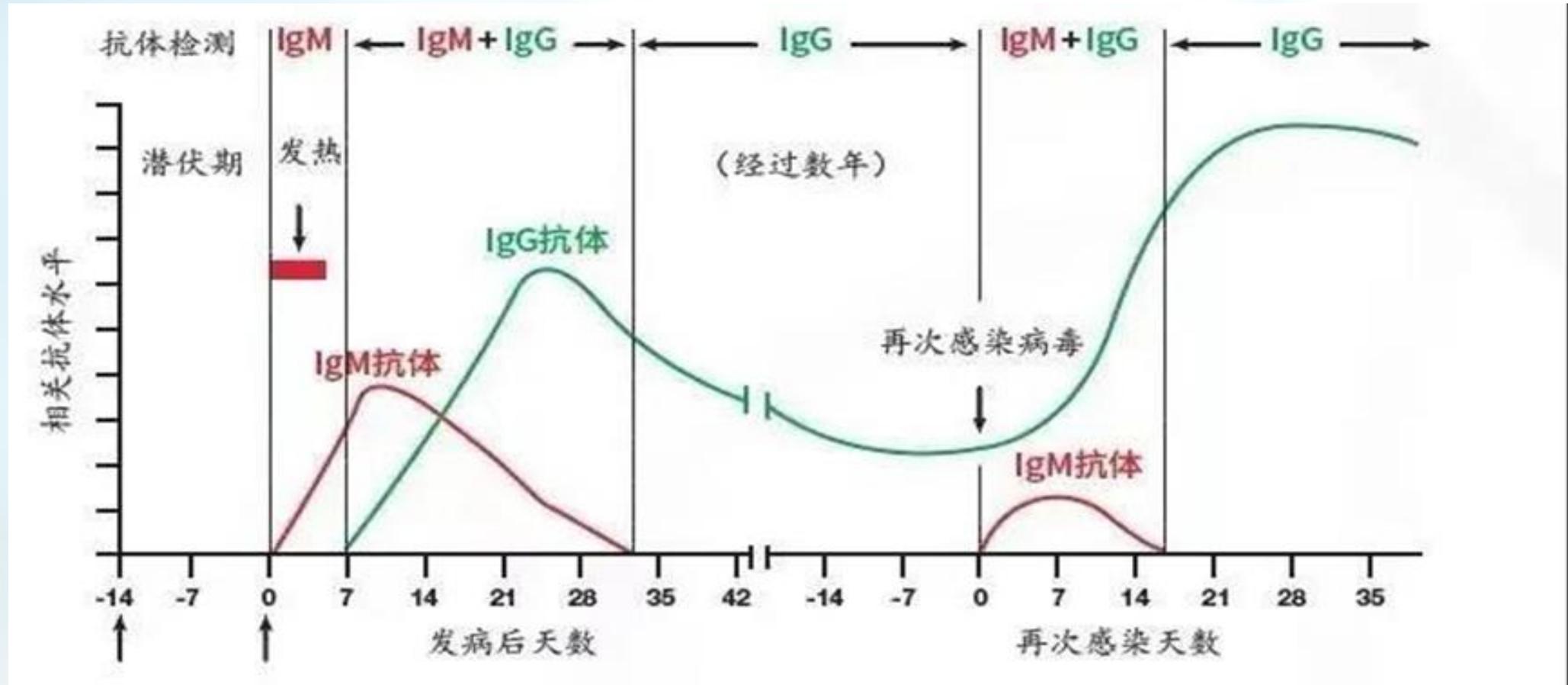
- 1、**采样部位影响检测结果**，新冠病毒在下呼吸道及肺部病毒载量最高，进行咽拭子采样可能导致假阴性结果的出现。
- 2、**采样时机影响检测结果**，患者出现症状后第一周内唾液病毒载量最高，其后随着时间，病毒载量逐步下降；
- 3、采样人员的技术直接影响检测结果；
- 4、试剂的质量影响检测结果。

间接检验

间接检测试剂分为针对病毒抗体检测和针对病毒抗原的检测。抗体检测试剂是检测血清中由病毒进入人体后刺激人体产生的IgM或IgG抗体，IgM抗体出现较早，IgG抗体出现较晚。病毒抗原检测主要是检测病毒表面的一些蛋白质。



抗体检测



抗体检测

优点：具有**采样方便**（静脉血）、**检测效率高**（检测周期短）的特点；

缺点：

- 1、抗体检测的**窗口期问题**（病毒进入机体到机体产生特异性的抗体需要一段时间）；
- 2、**血液标本中的某些干扰物质**可能导致抗体检测的假阳性、假阴性的问题。
- 3、**抗原交叉反应**的影响。



抗体检测



抗体检测是核酸检测的补充，不能单独作为诊断的筛查方法。核酸的检测是早期检测，因为是直接查病毒的检测。抗体的产生一般在抗原（病毒）进入机体后1到2周后才能产生，所以不适合早期检查，因为即便查不到也不代表没感染。但是一阵大流行后，通过抗体检测可以判断此人是否已经感染，属于回顾式检查，当核酸检测阴性时，将IgM和IgG抗体检测增加进去，可以弥补核酸检测容易造成漏诊的缺点。

目前来说核酸检测最靠谱，抗体检测的优势是快速，但假阳性率要比核酸检测高。

QUESTION4



病毒的突变位点有哪些？这些突变对病毒的感染能力和检测结果有什么影响？

病毒的突变位点有哪些？

Region (ORF)	Nucleotide (nt)			Amino acid (aa)		
	Start and end	Length	Homology (%) ^a	Start and end	Length	Homology (%) ^a
Full length	1-29870	29 870	99.99 (99.91-100)	1-9744	9744	99.99 (99.79-100)
lab	266-21555	21 306	100 (99.91-100)	1-7096	7096	100 (99.80-100)
1a	266-13483	13 218	99.99 (99.88-100)	1-4401	4401	100 (99.73-100)
1b	13468-21555	8088	100 (99.93-100)	4402-7096	2695	100 (99.85-100)
S	21563-25384	3822	100 (99.82-100)	7097-8369	1273	100 (99.53-100)
3a	25393-26220	828	100 (99.76-100)	8370-8644	275	100 (99.27-100)
E	26245-26472	228	100 (100-100)	8645-8719	75	100 (100-100)
M	26523-27191	669	100 (99.70-100)	8720-8941	222	100 (99.95-100)
6	27202-27387	186	100 (100-100)	8942-9002	61	100 (100-100)
7a	27394-27759	366	100 (99.73-100)	9003-9123	121	100 (99.17-100)
7b	27756-27887	132	100 (100-100)	9124-9166	43	100 (100-100)
8	27894-28259	366	100 (99.45-100)	9167-9287	121	100 (98.35-100)
N	28274-29533	1260	100 (99.84-100)	9288-9706	419	100 (99.76-100)
10	29558-29674	117	100 (99.15-100)	9707-9744	38	100 (97.37-100)

Abbreviations: E, Envelope; M, Membrane; N, Nucleoprotein; ORF, open-reading frame; S, Spike; SARS-COV-2, severe acute respiratory syndrome coronavirus 2; 1ab, open-reading frames 1ab; 1a, open-reading frames 1a; 1b, open-reading frames 1b; 3a, open-reading frames 3a; 6, open-reading frames 6; 7a, open-reading frames 7a; 7b, open-reading frames 7b; 8, open-reading frames 8; 10, open-reading frames 10.

^aMedian (min-max).

SARS-CoV-2参考序列的建立和变异分析

病毒的突变位点

S区主要的突变位点:

spike	C5388A	A1708D
	T6954C	I2230T
	11288-11296 deletion	SGF 3675- 3677 deletion
	<u>21765-21770 deletion</u>	<u>HV 69-70</u> ← deletion
	21991-21993 deletion	Y144 deletion
	<u>A23063T</u>	<u>N501Y</u>
	C23271A	A570D
	<u>C23604A</u>	<u>P681H</u>
	C23709T	T716I
	T24506G	S982A

第一个是69-70del，这是古普塔在他的英国剑桥患者身上发现的一种基因缺失，该患者体内的新冠病毒似乎因为这一个点位的缺失逃避了免疫系统。它导致刺突蛋白中两个氨基酸的缺失。在实验室实验中，古普塔发现工程改造后携带这种缺失刺突蛋白的病毒具有两倍的感染力。

病毒的突变位点

S区主要的突变位点:

spike	C5388A	A1708D
	T6954C	I2230T
	11288-11296 deletion	SGF 3675- 3677 deletion
	<u>21765-21770 deletion</u>	<u>HV 69-70 deletion</u>
	21991-21993 deletion	Y144 deletion
	<u>A23063T</u>	<u>N501Y</u> ←
	C23271A	A570D
	<u>C23604A</u>	<u>P681H</u>
	C23709T	T716I
	T24506G	S982A

第二个是N501Y，这是弗雷德·哈钦森癌症研究中心的进化生物学家布鲁姆（Jesse Bloom）所发现的病毒突变，可以增加病毒蛋白与ACE2受体（这是新冠病毒进入人体细胞的“大门”）的结合程度。

病毒的突变位点

S区主要的突变位点:

spike	C5388A	A1708D
	T6954C	I2230T
	11288-11296 deletion	SGF 3675- 3677 deletion
	<u>21765-21770 deletion</u>	<u>HV 69-70 deletion</u>
	21991-21993 deletion	Y144 deletion
	<u>A23063T</u>	<u>N501Y</u> ←
	C23271A	A570D
	<u>C23604A</u>	<u>P681H</u>
	C23709T	T716I
	T24506G	S982A

该突变也存在于501Y.V2中，这是南非研究人员在对三个沿海省份迅速暴发的疫情进行调查后发现的新冠状病毒变体。夸祖鲁-纳塔尔大学（University of KwaZulu-Natal）的病毒学家Tulio de Oliveira说，他们发现这一谱系似乎传播得更快了，他的工作首先使英国科学家意识到了N501Y的重要性。

病毒的突变位点

S区主要的突变位点:

spike	C5388A	A1708D
	T6954C	I2230T
	11288-11296 deletion	SGF 3675- 3677 deletion
	<u>21765-21770 deletion</u>	<u>HV 69-70 deletion</u>
	21991-21993 deletion	Y144 deletion
	<u>A23063T</u>	<u>N501Y</u>
	C23271A	A570D
	<u>C23604A</u>	<u>P681H</u> ←
	C23709T	T716I
	T24506G	S982A

第三个令人担忧的变异位点是P681H，它改变了必须剪切刺突蛋白才能进入人体细胞的位点。这也是新冠病毒和SARS病毒刺突细胞的不同位点这一。柏林Charité大学医院的病毒学家Christian Drosten说，这一位点上的突变和N501Y一样重要。

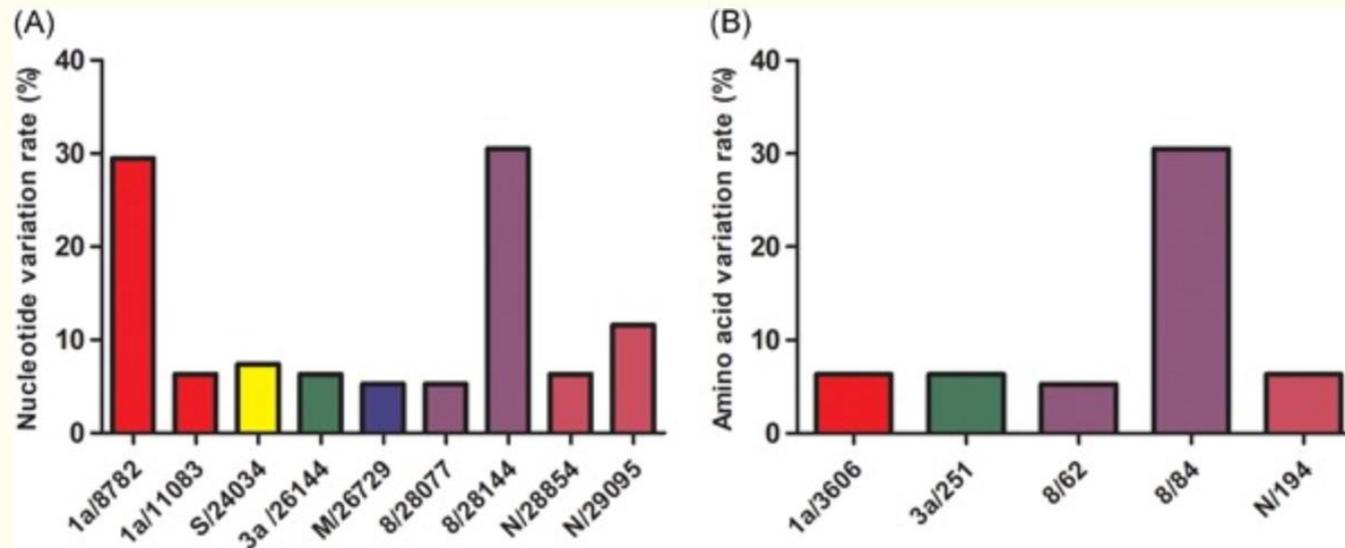


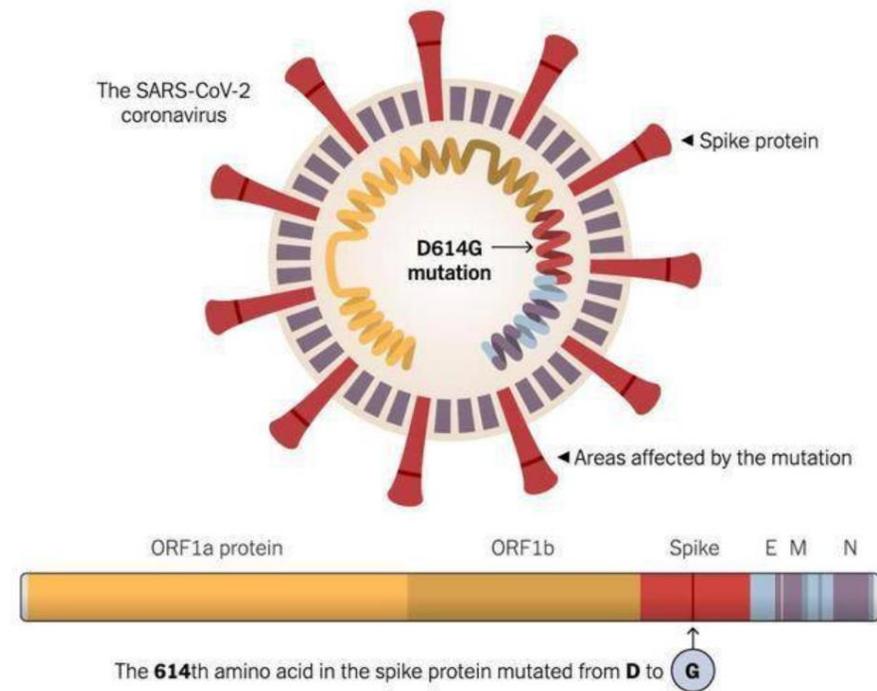
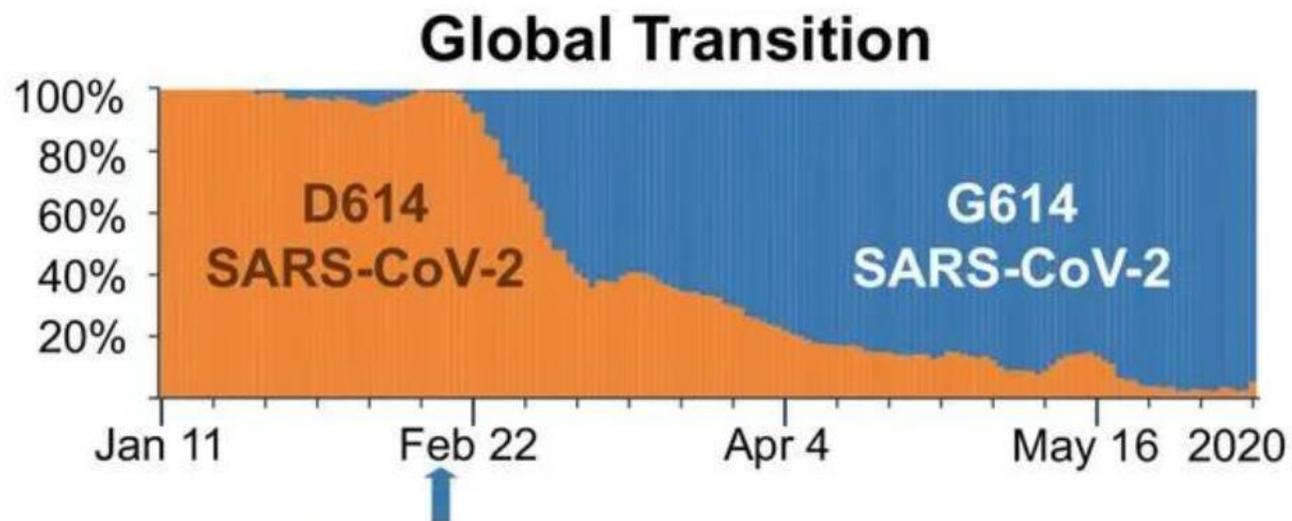
图3

SARS+COV+2 中常见突变的位点和频率 ($\geq 5/95$)。A, 核苷酸。B, 氨基酸。SARS+COV+2, 严重急性呼吸系统综合症冠状病毒2

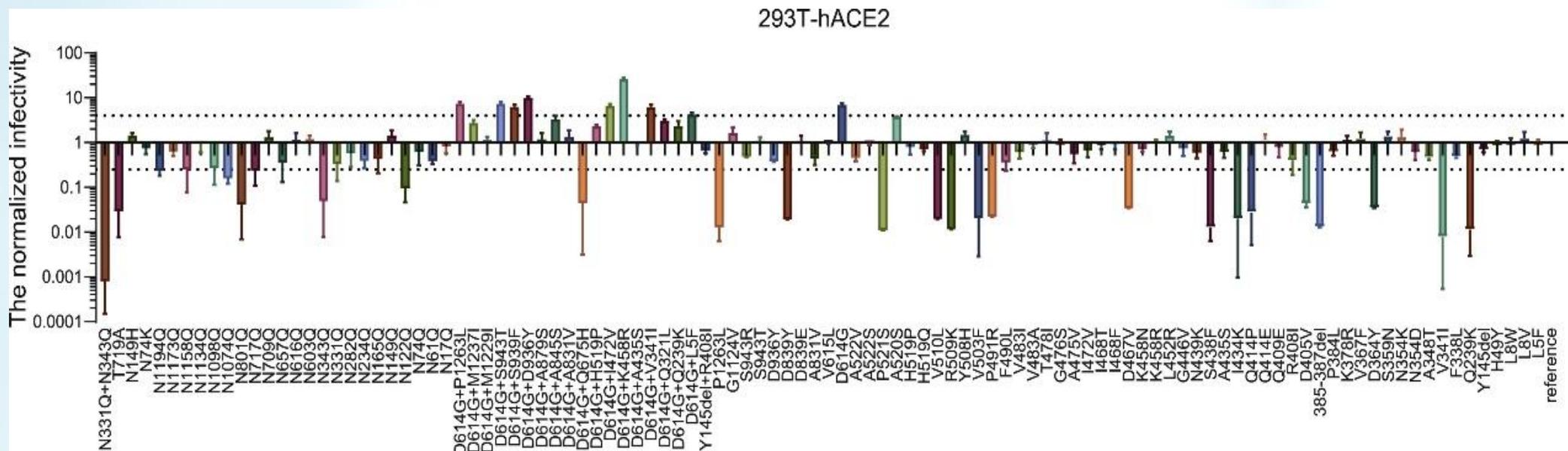
另外一个非常重要的突变（目前主要的突变），新冠病毒第614氨基酸位点D（天冬氨酸）到G（甘氨酸）发生突变，即D614G突变，他们把这种拥有新突变的病毒版本称为“G614”。

该突变是一个错译突变（改变氨基酸的变异），并且位于新冠病毒的刺突蛋白（spike protein, S蛋白）上，该蛋白是新冠病毒入侵人体细胞的核心武器，也是目前许多疫苗和疗法所重点针对的目标。相关突变很可能会改变刺突蛋白的结构、性质和活力，进而改变病毒的传染性与致病性。

G614主要突变产生的影响



- 经过4个多月的传播，G614成为目前新冠传播的主要基因型。WHO7月3日表示，29%的新冠病毒样本都出现了该变异，前段时间北京疫情反弹中也发现了这一突变株。
- 研究人员表示，新版病毒的载量更高，且在人肺上皮细胞、hACE2细胞中发现其感染能力更强。据杜克大学实验室的测试情况，变异后病毒的传染性要比早期病毒强3至9倍。



与所有参考病毒株相比，单个D614G及其与D614G的组合变体显示出增加的感染力。而单个D614G和D614G组合变体之间没有发现差异，这表明增强的传染性更可能归因于D614G本身

	A组	B组	C组
变体或突变体数量	29	51	26
传染性增加	D614G, D614G + L5F, D614G + D936Y, D614G + S939F, D614G + S943T	D614G + V341I, D614G + K458R, D614G + I472V ^{一种}	没有
传染性降低	Q239K, D839Y, P1263L, D614G + Q675H	V341I, D364Y, 385-387del, D405V, Q414P, I434K, S438F, D467V, P491R, V503F, R509K, V510L, P521S	N122Q, N343Q, N717Q, T719A, N801Q, N1074Q, N331Q + N343Q
对中和单克隆抗体的敏感性增加	没有	V367F, Q409E, Q414E, I468F, I468T, Y508H, A522V	N165Q, N709Q
对中和单克隆抗体的敏感性降低	A831V	N439K, L452R, A475V, V483A, F490L, Y508H, D614G + A435S, D614G + I472V ^{一种}	N234Q
对恢复期血清的敏感性增加	V615L	F338L, V367F, I468F, I468T	N149H, N149Q, N165Q, N331Q, N354D, N709Q, N1173Q
对恢复期血清的敏感性降低	Y145del, A831V, D614G + A831V, D614G + A879S, D614G + M1237I	Q414E, N439K, G446V, K458N, I472V, A475V, T478I, V483I, F490L, H519P, D614G + Q321L, D614G + I472V ^{一种}	没有

突变对检测结果的影响

SARS-CoV-2是一种单股正链RNA病毒，属冠状病毒科β冠状病毒属中的 Sarbecovirus亚属。其基因组包含11个基因，序列大小约29 882 nt。其中功能较为重要的5个基因分别是orf1ab含有的rdrp基因、s基因编码刺突蛋白、e基因编码囊膜蛋白、m基因编码膜蛋白、n基因编码核蛋白。

根据SARS-CoV-2核酸序列特点，我国疾病预防控制中心推荐的检测基因为orf1ab基因和n基因，而美国疾控中心检测的基因为n基因，世界卫生组织推荐的检测基因为orf1ab基因、n基因和e基因。

国家生物信息中心对基于人体提取的65株病毒全基因组序列进行分析，发现SARS-CoV-2群体内的序列变异主要发生在5个基因上，即s基因、n基因、orf8基因、orf3a基因和最大的orf1ab基因。

荧光定量PCR法检测SARS-CoV-2有可能因为病毒株核酸序列变异产生假阴性结果。

病毒的检测

世卫组织表示，全球目前所使用的绝大多数荧光定量检测都使用多个不同基因目标，因此变异病毒不会对检测有效性造成太大的影响。

中国选择核酸检测的两个探针并不包含变异区域，所以仍然能够有效检测病毒。

“两个探针的位置相对保守，目前没有出现检测敏感性降低的情况。”

所谓探针是检测新冠病毒核酸的关键组分，如果探针区域发生了变异，对变异后毒株的检测可能会发生漏检。但是核酸检测试剂的探针设计一般会选择保守区段，即不太容易发生变异的核酸序列，以确保检出率。

新冠疫苗

为全民
免费提供



(资料图)

央视
新闻

新华社北京电

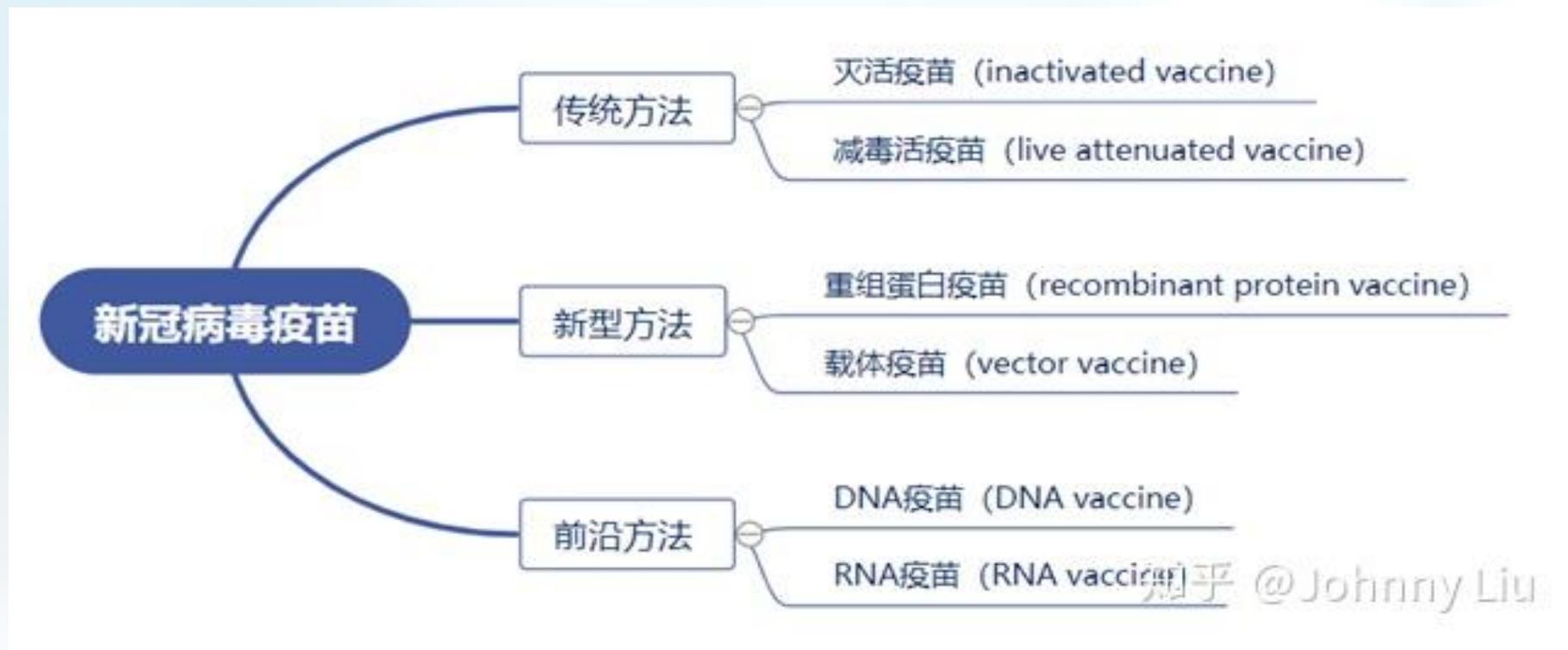
新冠病毒疫苗的种类可以概括为第一类是此前无同类疫苗获批过的新型疫苗，主要是指核酸疫苗，分为RNA（核糖核酸）疫苗和DNA（脱氧核糖核酸）疫苗，这类疫苗是将编码抗原蛋白的RNA或DNA片段直接导入人体细胞内。第二类是此前已得到广泛应用的传统类型疫苗，包括灭活病毒疫苗、基因工程亚单位疫苗、重组病毒载体疫苗等，多数在研新冠病毒疫苗都属于此类。比如中国团队研发的“重组新冠疫苗”就属于重组病毒载体疫苗，采用5型腺病毒作载体向人体内输送表达新冠病毒刺突蛋白的基因。



基因工程疫苗是通过基因工程对抗原基因与其他细胞、病毒、质粒重组或将其作为载体，甚至通过技术平台直接合成基因序列制备的疫苗，包括了**病毒载体疫苗**、**核酸疫苗**（DNA 疫苗和 mRNA 疫苗）等。

病毒载体疫苗是指以无致病性或改造过的病毒如腺病毒、痘病毒、甲病毒等为载体，搭载具有免疫原性的病原体基因的疫苗，通过病毒载体进入人体产生免疫反应。我国目前已有腺病毒载体新型冠状病毒疫苗在进行三期临床试验。

核酸疫苗是将编码病原体抗原蛋白的外源基因(DNA 或 RNA) 导入人体细胞内， 在细胞内表达合成抗原蛋白， 诱导人体产生对该抗原蛋白的免疫应答， 以达到预防疾病的目的。



传统方法：

	灭活疫苗	减毒活疫苗
定义	在体外培养新冠病毒，然后将其灭活，使之没有毒性，但这些病毒的“尸体”仍能刺激人体产生抗体，使免疫细胞记住病毒的模样。	繁殖能力大幅减弱的新冠病毒，复制过程不像灭活疫苗那样完全不发生，而是被限制在一定程度内，不产生严重疾病，但是可以模拟自然感染过程而诱导免疫反应。它的生产主要是通过不利环境培养或者人工修饰完成的。
优点	技术成熟，生产相对容易	鼻内给药。突变不敏感
缺点	产量受限，要求较高	要求较高，消耗时间长

新型方法：

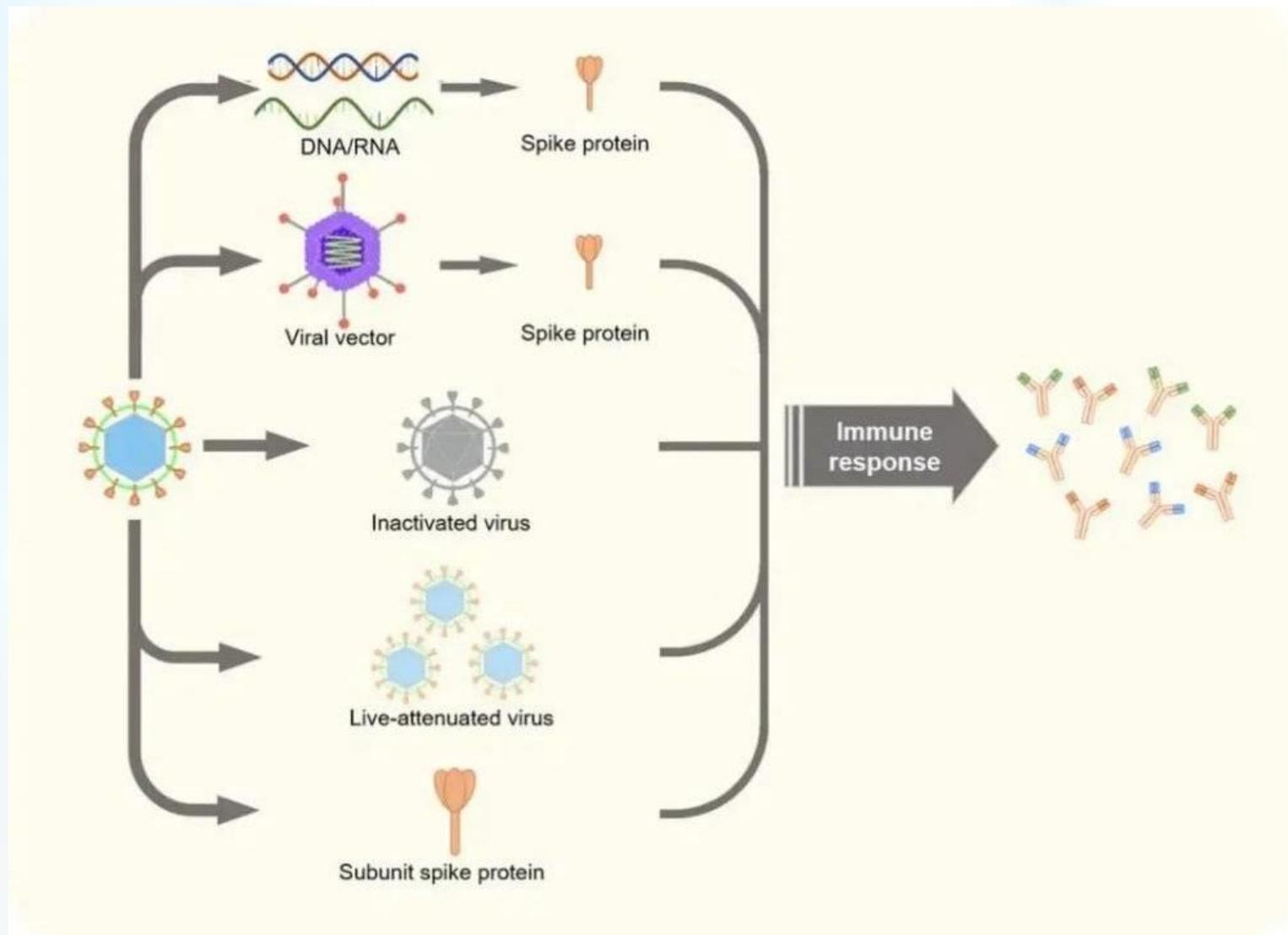
	重组蛋白疫苗	载体疫苗
定义	不同细胞表达体系中产生的重组蛋白，可以理解为借助别人的工作室（其它细胞）生产自己的产品（新冠病毒的蛋白质）	以其它病毒为载体来运送新冠病毒的蛋白质，这些作为载体的病毒已经通过生物工程技术加以改造，因此可以表达新冠病毒的刺突蛋白，从而在体内产生免疫反应。
优点	安全，投入使用	安全，投入使用，兼顾体液和细胞免疫
缺点	蛋白难以表达，抗原转移	抗体综合

前沿方法：

	DNA疫苗	RNA疫苗
定义	利用细菌的质粒DNA表达新冠病毒的刺突蛋白，这些质粒一般含有哺乳动物的表达启动子，因而可以通过运输而在体内表达。	将表达抗体的RNA运送到细胞中，RNA会在接种者体内的细胞中表达抗体，从而产生免疫反应，这个运送的过程需要脂质纳米颗粒（LNP）的帮助。
优点	大批量生产（工程菌）	大批量生产（活体外部）
缺点	免疫原性低，要求较高	技术不成熟

六种类型疫苗的区别：

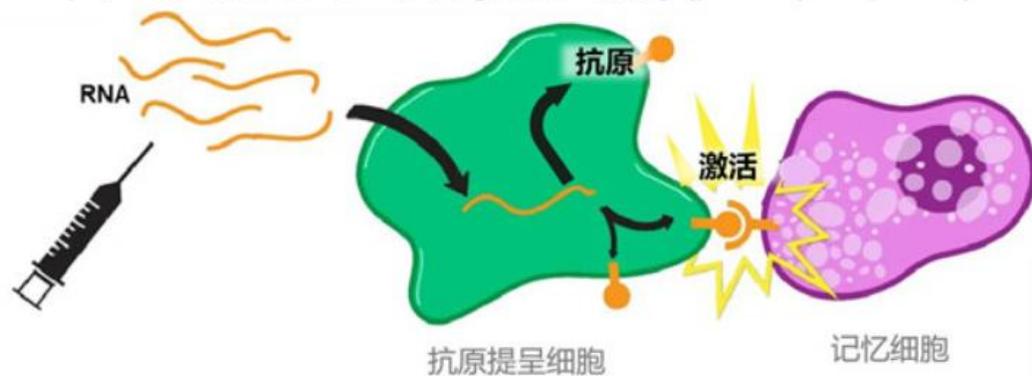
引发免疫应答的方式



mRNA疫苗工作机制

图2. mRNA预防疾病的原理

1) 创造免疫记忆 (首次免疫应答)



2) 二次免疫应答--对真的病原体

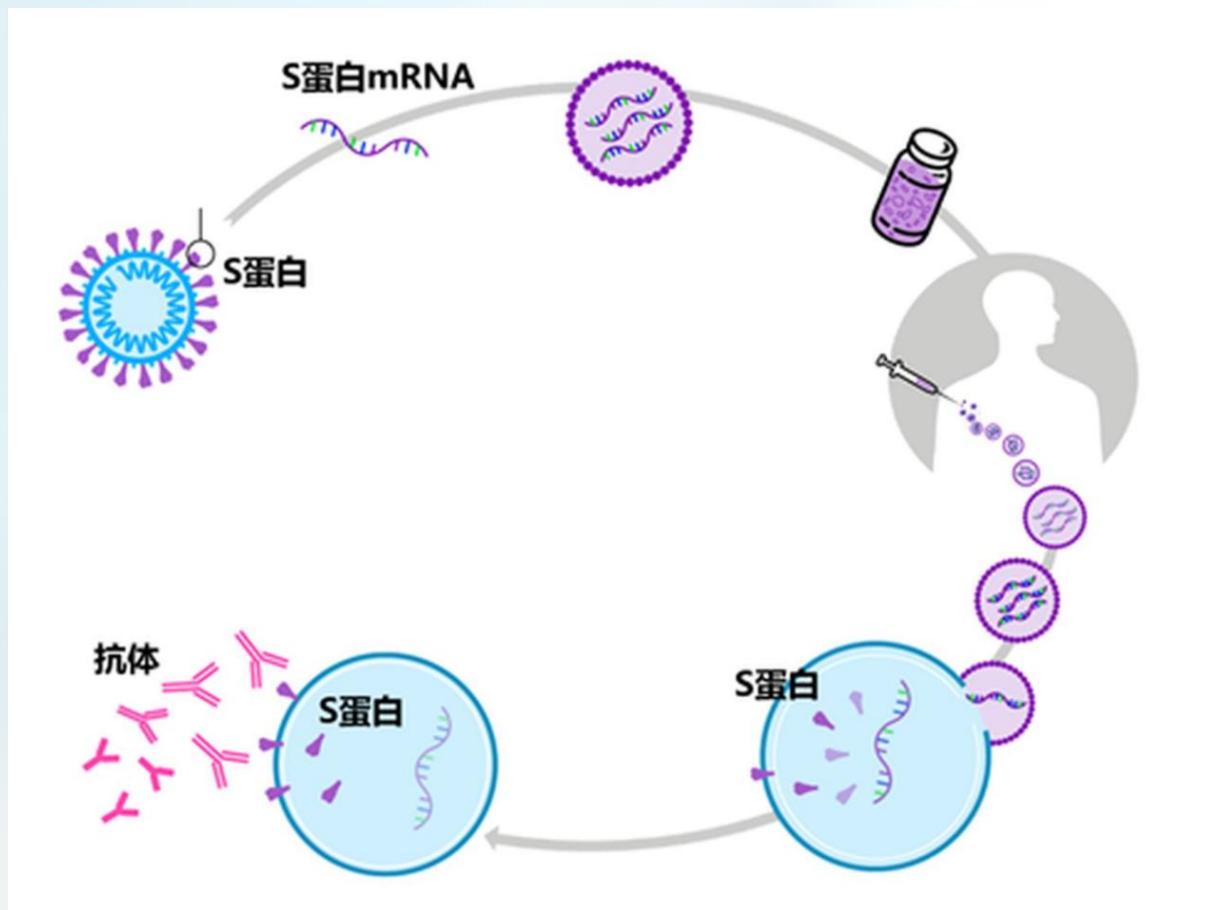


当一个人接受了编码病毒蛋白的mRNA注射后，这个人的细胞就会翻译它们，在体内表达病毒蛋白。预计在mRNA注射后24-48小时内，细胞内的蛋白质水平达到峰值，并可能持续几天。蛋白质抗原由抗原提呈细胞提呈给免疫系统，激活特定免疫细胞，触发免疫应答反应。

最早两家获批上市的新冠疫苗（辉瑞和Moderna）都属于mRNA疫苗，而mRNA疫苗由于之前从来没有过上市大规模使用的先例，属于比较新的一种疫苗研发技术

- mRNA疫苗是化学方法制作的，而不是生物合成的。同时，因为不需要收集和纯化蛋白（接种病毒到鸡蛋或者动物细胞株中）这些实验操作，制备mRNA疫苗变得更快且容易操作。
- 因为mRNA疫苗的温度敏感特点，高于冰点就容易被分解，所以开发mRNA疫苗最具挑战的仍然是其不稳定性。这就是为什么辉瑞公司的mRNA疫苗需要在-70度的温度储存，而Moderna公司声称，它的疫苗可以在大多数家庭或医疗冰箱温度下保存长达六个月。

mRNA疫苗制作



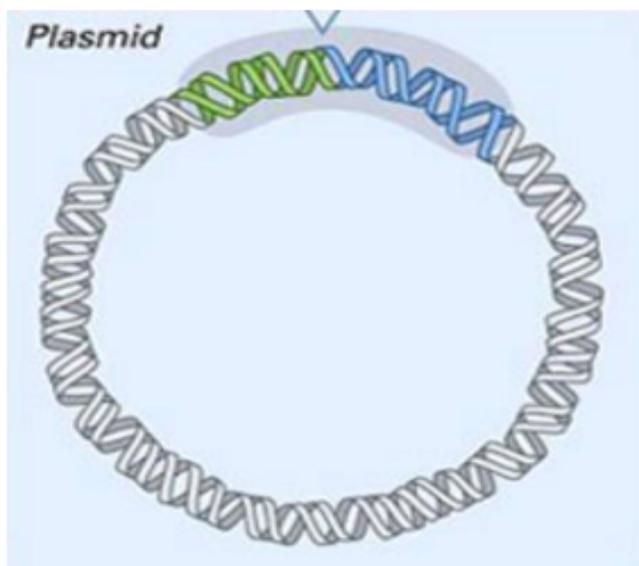
- 采用新冠病毒表面的突刺蛋白（S）作为抗原，反推得出其对应的mRNA，再加入盐溶液（与身体的盐度类似，主要起传递的作用，属于很天然物质）以及其他佐剂，用脂质纳米粒包裹制成疫苗
- 疫苗接种主要是以胳膊注射的形式，所以疫苗中的mRNA主要是进入我们的肌肉细胞，当然也可能会有极少数进入到一些特殊的免疫细胞中
- mRNA进入细胞后，首先会在短时间内被翻译很多次，从而产生足够多的S蛋白在细胞表面表达，激活我们的自体免疫反应。通常情况下，RNA在人体内只能够存活数分钟到数小时，有少部分可能可以存活几天，而疫苗中的mRNA在稳定性上因为是特意设计过的，所以平均会在体内存活数天。

使用的疫苗种类 —— “CanSino’ s replication defective Ad5 vectored vaccine expressing the spike glycoprotein of SARS-CoV-2” ↵

(CanSino 指的是天津的康希诺生物股份公司。这篇论文中所用的疫苗是由北京生物技术研究院 (Beijing Institute of Biotechnology) 和康希诺生物股份公司 (CanSino Biologics) 提供的。) ↵

Ad5 的全称是 Adenovirus type-5 vector。因为从 Adenovirus 中提取的病毒载体不仅可以进入正在复制的细胞，也可以进入停止复制的细胞，所以是制作疫苗时常用的一种载体。↵

spike glycoprotein of SARS-CoV-2 指的是新冠病毒的棘突蛋白，这个蛋白是镶嵌在新冠包膜上，辅助新冠入侵人体细胞的重要蛋白。生物学家将新冠病毒内编译这段蛋白的 RNA 提取出来，将其转化为 DNA，嵌入 Ad5 这个病毒载体中 (如下图)。在 Ad5 进入人体细胞的时候，这段属于新冠的 DNA 也会一同进入。↵



(忽视上方的 plasmid，这个灰色的环是 Ad5，蓝绿色的是上

QUESTION6

如何进行基因治疗？

在**Covid-19**肆虐的当下，针对它的基因治疗方案仍处于萌芽阶段。

于是，我们检索了针对**SARS-Cov**的较为成熟的方案，希望能为**Covid-19**的基因治疗提供方向。

基因治疗

思路一：RNAi

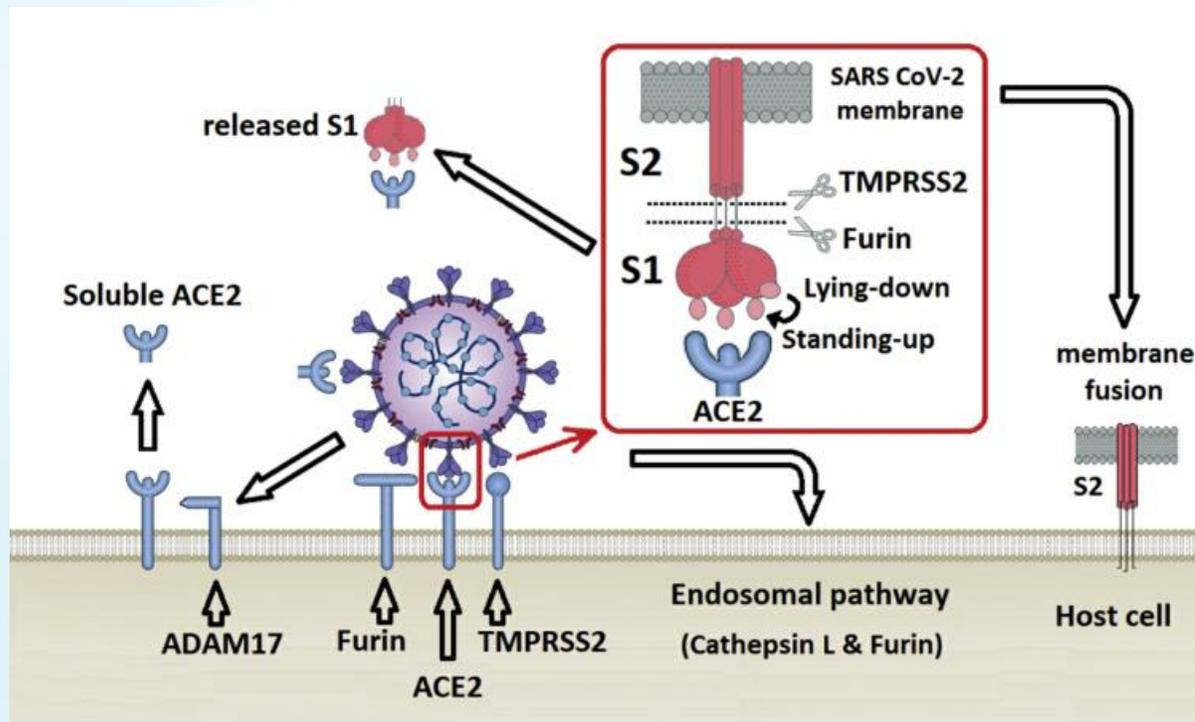
siRNA的设计合成及其对SARS病毒部分复制酶基因的抑制作用。由于复制酶基因的表达是SARS病毒启动转录的前提, 而刺突蛋白在病毒入侵宿主细胞的过程中发挥着重要作用, 因此选择SARS病毒复制酶基因(rep)和刺突蛋白基因(s)的保守序列作为靶基因。

因此, 我们也可以选择covid-19的保守序列, 譬如ORF1ab, N基因, S基因作为RNA干扰的靶标。

思路二：CRISPR-Cas13d系统

Cas13d使用CRISPR相关RNA (crRNA), crRNA含有可定制的长22个核苷酸 (nt) 间隔序列, 从而可以将Cas13d蛋白引导到特定的RNA分子上以便进行靶向RNA降解。

Cas13d在人细胞中的高催化活性为靶向Covid-19以进行特异性病毒RNA基因组降解和病毒基因表达抑制提供了一种潜在的机制。



参与SARS-CoV-2进入宿主细胞的蛋白酶示意图

多功能蛋白血管紧张素转换酶2（ACE2）被认为是其进入宿主细胞的主要靶点。SARS-CoV-2的细胞进入可以通过两种不同的途径进行。如果SARS-S激活蛋白酶TMPRSS2与ACE2在靶细胞表面共表达，可以寻求第一种激活途径。与ACE2结合并由TMPRSS2处理（见插图）被认为允许在细胞表面融合。弗林主要位于晚期内体和高尔基体。然而，它被分泌到细胞外液中，也可以被运输到细胞膜。

第二种途径涉及SARS-CoV-2病毒粒相关三聚体S蛋白与ACE2的结合。结果，病毒颗粒被吸收到内体中，在内体中，SARS-S蛋白被依赖于pH值的半胱氨酸蛋白酶组织蛋白酶L和弗林蛋白酶切割和激活。SARS-S蛋白可以激活ADAM17，ADAM17切割ACE2，导致ACE2（可溶性ACE2）脱落。S1顶端的粉红色卵圆形结构表示受体结合域。



THANK YOU