

核酸构建体及其应用

本公开要求如下专利申请的优先权：于 2022 年 04 月 27 日提交，申请号为 CN202210456261.9，发明名称为“核酸构建体及其应用”的中国专利申请；上述  
5 专利申请的全部内容通过引用结合在本公开中。

技术领域

本公开属于核酸药物领域，具体涉及包含工程化 UTR 的核酸构建体及其用于  
预防或治疗疾病(例如，病毒感染)的用途。

10

背景技术

疫苗是重要的医药研究领域，已有的疫苗包括灭活疫苗、腺病毒载体疫苗、  
减毒流感病毒载体疫苗、重组蛋白疫苗、核酸疫苗(包括 RNA 疫苗和 DNA 疫苗)。  
mRNA 是 RNA 疫苗的活性成分，其主要包括帽子结构(Cap)、5'非翻译区  
15 (Untranslatedregion, UTR)、编码抗原蛋白的开放阅读框(Open reading frame, ORF)、  
3'UTR 和 Poly(A)尾结构。其中，5'UTR 是将核糖体招募到 mRNA 并启动密码子  
选择的关键，在调控翻译效率和塑造细胞蛋白质组方面发挥着重要作用(Ivanov,  
et al. Science. 2016, 352(6292): 1413-1416.)。真核生物的 3'UTR 具有多种调控基  
元(regulatory motifs)，可被 microRNAs (miRNAs)和 rbp 识别，以控制 mRNA 的稳  
20 定性、定位和翻译(Mazumder et al., 2003; Mayr, 2017)。polyA 尾巴(位于 mRNA  
3'末端)是另一个决定 mRNA 稳定性和蛋白质水平的元素。可选的 3'UTR 不仅影  
响 mRNA 的稳定性和翻译,而且还控制 mRNA 的定位(Tushev 等人, 2018)。mRNA  
的设计与优化，特别是其 UTR 的选择和使用，是整个 mRNA 疫苗产品制备流程  
中的关键一环。因此，开发高效的 UTR，仍是 mRNA 药物领域的迫切需要。

25 本公开提供了具有新 UTR 结构 mRNA 疫苗，具有使得目的基因稳定、高效  
表达的优点，可普遍性的应用于 mRNA 疫苗(例如，流感或新冠疫苗)，用于调控  
目的基因表达，具有优异的临床药物应用前景。

发明内容

30 本公开提供可调节目的基因表达的核酸元件，以及核酸构建体。其中，所述  
核苷酸构建体包含至少一个可调节目的基因表达的核酸元件。

核酸元件

一些实施方案中，所述核酸元件为 5'非翻译区元件(5'UTR)。

一些实施方案中，5'UTR 选自源自或为 Rho GTPase 激活蛋白(ARHGAP)、热  
35 休克 27kD 蛋白 1(heat shock 27kDa protein 1, HSPB1)、β 珠蛋白(hemoglobin subunit  
beta, HBB)、趋化因子 c-c-基元配体 13 (C-C motif chemokine ligand 13, CCL13)

等任一基因的 5'UTR 或其衍生序列。

一些实施方案中，所述核酸构建体中的 5'UTR 是源自或为 ARHGAP 基因的 5'UTR 序列或其衍生序列。在一些实施方案中，所述 ARHGAP 包括 ARHGAP1、ARHGAP2、ARHGAP3、ARHGAP4、ARHGAP5、ARHGAP6、ARHGAP7 (DLC1)、  
5 ARHGAP8、ARHGAP9、ARHGAP10、ARHGAP12、ARHGAP13 (SRGAP1)、ARHGAP14 (SRGAP2)、ARHGAP15、ARHGAP17 (RICH1)、ARHGAP18、ARHGAP19、ARHGAP20、ARHGAP21、ARHGAP22、ARHGAP23、ARHGAP24、ARHGAP25、ARHGAP26。

一些实施方案中，所述 5'UTR 是源自或为 ARHGAP15 基因的 5'UTR 序列或其衍生序列。一些实施方案中，所述 ARHGAP15 基因是来源任意物种的，如人类  
10 ARHGAP15、狓狓 ARHGAP15、猴 ARHGAP15、小鼠 ARHGAP15 等。

一些实施方案中，所述 5'UTR 是源自或为人类 ARHGAP15 基因的 5'UTR 序列或其衍生序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或为如  
SEQ ID NO: 1 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、  
15 97%、98%、99%、100% 同一性的核苷酸序列。

一些实施方案中，所述 5'UTR 包含至少一个可用于抑制 UTR 内部 ATG 启动翻译的点突变。一些实施方案中，所述点突变选自 5'UTR 中 ATG 序列的 A、T 或 G 中的任意一个或多个位点上的突变。一些实施方案中，所述点突变为 5'UTR 中 ATG 序列的 A 位点发生突变。一些实施方案中，所述突变为 A 突变为 G、C 或 T。  
20 一些实施方案中，所述突变为 ATG 突变为 GTG、CTG 或 TTG。进一步地，所述突变可用于抑制 UTR 内部 ATG 启动翻译。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含至少一个可用于抑制 UTR 内部 ATG 启动翻译的点突变。一些实施方案中，所述点突变选自 5'UTR 中 ATG 序列的 A、T 或 G 中的任意一个或多个位点上的突变。一些实施方案中，  
25 所述点突变为 5'UTR 中 ATG 序列的 A 位点发生突变。一些实施方案中，所述突变为 A 突变为 G、C 或 T。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 SEQ ID NO: 44 所示的核苷酸序列，或，所述 5'UTR 包含与 SEQ ID NO: 44 具有至少 80%、85%、90%、95%、96%、97%、98%、99% 同一性的核苷酸序列，其中 N<sub>1</sub> 选自 A、G、C  
30 或 T。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为 GTG 的核苷酸序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 2 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100% 同一性的核苷酸序列。

35 一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为 CTG 的核苷酸序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或

为如 SEQ ID NO: 42 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为 TTG 的核苷酸序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或  
5 为如 SEQ ID NO: 43 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，5'UTR 为 5'UTR 截短体。一些实施方案中，所述 5'UTR 截短体仍然保持与天然 5'UTR 相似的活性功能，例如仍然保持调控 ORF 编码目的基因表达蛋白的功能。一些实施方案中，所述 5'UTR 截短体与天然 5'UTR 相比，具  
10 有增强的活性功能，例如增强的调控 ORF 编码目的基因表达蛋白的功能。

一些实施方案中，所述 5'UTR 截短体的截短方式包括：从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；并且，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。一些实施方案中，所述  
15 5'UTR 截短体还包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体不包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体保留了相比于天然 5'UTR 序列长度的至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、1%长度的序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含或为  
25 ARHGAP15 5'UTR 截短体。一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，缺失 5'末端连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸序列。

一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，  
30 5'UTR 截短体的截短方式包括从 ARHGAP15 5'UTR 的 3'至 5'末端的方向上，删除 3'末端连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，5'UTR 截短体的截短方式是从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列；并且，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸序列。

一些实施方案中，所述截短体保留了相比于天然 ARHGAP15 5'UTR(SEQ ID NO: 1)至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、  
35

1%长度的序列。

一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中至少 5 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中 5-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中 7-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 所示序列中 7-59 个连续的核苷酸序列；示例性地，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中 5、7、8、9、10、11、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 个连续的核苷酸序列。

一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸(也即，5'UTR 截短体的起始核苷酸)是 SEQ ID NO: 2 或 44 所示序列按自然计数的 1-57 位的任一位置处的核苷酸。一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸是 SEQ ID NO: 2 或 44 所示序列按自然计数的第 1-13 或第 17-29 位中的任一位置处的核苷酸。例如，第 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57 的任一位置处的核苷酸。

示例性地，所述 5'UTR 截短体的 5'末端的核苷酸，及所述 5'UTR 截短体包含的 SEQ ID NO: 2 所示序列中连续核苷酸序列的长度如下表所示：

| 起始核苷酸   | SEQ ID NO: 2 所示序列中连续核苷酸序列的长度                                                                                                                                               |
|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 第 1 位，G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 |
| 第 2 位，A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 |
| 第 3 位，G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60 或 61    |
| 第 4 位，T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、                                                                                   |

|           |                                                                                                                                                                 |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
|           | 36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60                                                                                      |
| 第 5 位, T  | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59 |
| 第 9 位, A  | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55             |
| 第 13 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51                         |
| 第 17 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47                                     |
| 第 21 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43                                                 |
| 第 23 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41                                                       |
| 第 25 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39                                                             |
| 第 29 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35                                                                         |
| 第 33 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31                                                                                     |
| 第 37 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27                                                                                                 |
| 第 41 位, C | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23                                                                                                             |
| 第 45 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19                                                                                                                         |
| 第 49 位, G | 5、6、7、8、9、10、11、12、13、14、15                                                                                                                                     |

|           |                 |
|-----------|-----------------|
| 第 53 位, A | 5、6、7、8、9、10、11 |
| 第 57 位, C | 5、6、7           |

一些实施方案中,所述 5'UTR 截短体的长度为 59bp、55bp、51bp、47bp、43bp、39bp、35bp、31bp、27bp、23bp、19bp、15bp、11bp、7bp。

一些实施方案中,所述 ARHGAP15 5'UTR 截短体还包含前述任意实施方案的  
5 点突变。一些实施方案中,所述 ARHGAP15 5'UTR 截短体还包含 ATG 突变为 GTG 的核苷酸序列。一些实施方案中,前述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 28-30 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,所述 ARHGAP15 5'UTR 截短体不包含前述任意实施方案的  
10 点突变。一些实施方案中,前述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 31-40 所示的核苷酸序列,包含 CTATAAT;或与前述任一序列具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,所述 5'UTR 截短体至少包含 CTATAAT 或 SEQ ID NO: 193 所示的核苷酸序列。

15 一些实施方案中,所述 ARHGAP15 5'UTR 截短体包含或为如 SEQ ID NO: 41、76、137-196 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 HSPB1 基因的 5'UTR 包含或为如 SEQ ID NO: 3-5 任一所示的核苷酸序列或与 SEQ ID NO: 3-5 任一具有至少 80%、85%、90%、95%、  
20 96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,5'UTR 包含或为如 SEQ ID NO: 1-2、3-5、28-43、76、137-196 任一所示的核苷酸序列,或如下序列:CTATAAT,或与前述任一序列具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

核酸构建体

25 本公开提供一种核酸构建体,其包括:

- (a)开放阅读框(ORF); 和
- (b)5'非翻译区元件(5'UTR)。

一些实施方案中,所述 ORF 为目的基因蛋白的编码多核苷酸序列。

一些实施方案中,所述目的基因为异源的。另一些实施方案中,所述目的基  
30 因为内源的。

一些实施方案中,所述核酸构建体中的 5'UTR 位于开放阅读框的上游。一些实施方案中,所述核酸构建体中的 5'UTR 位于开放阅读框的 5'末端。

一些实施方案中,所述核酸构建体中的 5'UTR 选自源自或为 Rho GTPase 激活蛋白(ARHGAP)、热休克 27kD 蛋白 1(heat shock 27kDa protein 1, HSPB1)、β 珠蛋

白(hemoglobin subunit beta, HBB)、趋化因子 c-c-基元配体 13 (C-C motif chemokine ligand 13, CCL13)等任一基因的 5'UTR 或其衍生序列。

一些实施方案中,所述核酸构建体中的 5'UTR 是源自或为 ARHGAP 基因的 5'UTR 序列或其衍生序列。在一些实施方案中,所述 ARHGAP 包括 ARHGAP1、  
5 ARHGAP2、ARHGAP3、ARHGAP4、ARHGAP5、ARHGAP6、ARHGAP7 (DLC1)、  
ARHGAP8、ARHGAP9、ARHGAP10、ARHGAP12、ARHGAP13 (SRGAP1)、  
ARHGAP14 (SRGAP2)、ARHGAP15、ARHGAP17 (RICH1)、ARHGAP18、  
ARHGAP19、ARHGAP20、ARHGAP21、ARHGAP22、ARHGAP23、ARHGAP24、  
ARHGAP25、ARHGAP26。

10 一些实施方案中,所述核酸构建体中的 5'UTR 是源自或为 ARHGAP15 基因的 5'UTR 序列或其衍生序列。一些实施方案中,所述 ARHGAP15 基因是来源任意物种的,如人类 ARHGAP15、狒狒 ARHGAP15、猴 ARHGAP15、小鼠 ARHGAP15 等。

一些实施方案中,所述核酸构建体中的 5'UTR 是源自或为人类 ARHGAP15 基  
15 因的 5'UTR 序列或其衍生序列。一些实施方案中,所述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 1 所示的核苷酸序列或与之具有至少 80%、85%、90%、  
95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,所述核酸构建体中的 5'UTR 包含至少一个可用于抑制 UTR  
内部 ATG 启动翻译的点突变。一些实施方案中,所述点突变选自 5'UTR 中 ATG  
20 序列的 A、T 或 G 中的任意一个或多个位点上的突变。一些实施方案中,所述点  
突变为 5'UTR 中 ATG 序列的 A 位点发生突变。一些实施方案中,所述突变为 A  
突变为 G、C 或 T。一些实施方案中,所述突变为 ATG 突变为 GTG、CTG 或 TTG。  
进一步地,所述突变可用于抑制 UTR 内部 ATG 启动翻译。

一些实施方案中,前述源自 ARHGAP15 基因的 5'UTR 序列包含至少一个可用  
25 于抑制 UTR 内部 ATG 启动翻译的点突变。一些实施方案中,所述点突变选自 5'UTR  
中 ATG 序列的 A、T 或 G 中的任意一个或多个位点上的突变。一些实施方案中,  
所述点突变为 5'UTR 中 ATG 序列的 A 位点发生突变。一些实施方案中,所述突变为 A  
突变为 G、C 或 T。

一些实施方案中,前述源自 ARHGAP15 基因的 5'UTR 序列包含 SEQ ID NO:  
30 44 所示的核苷酸序列,或,所述 5'UTR 包含与 SEQ ID NO:44 具有至少 80%、85%、  
90%、95%、96%、97%、98%、99%同一性的核苷酸序列,其中 N<sub>i</sub> 选自 A、G、C  
或 T。

一些实施方案中,前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为  
GTG 的核苷酸序列。一些实施方案中,所述的 ARHGAP15 基因的 5'UTR 包含或  
35 为如 SEQ ID NO: 2 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、  
96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为 CTG 的核苷酸序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 42 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

5 一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 ATG 突变为 TTG 的核苷酸序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 43 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，本公开核酸构建体中的 5'UTR 为 5'UTR 截短体。一些实施方案中，所述 5'UTR 截短体仍然保持与天然 5'UTR 相似的活性功能，例如仍然保持调控 ORF 编码目的基因表达蛋白的功能。一些实施方案中，所述 5'UTR 截短体与天然 5'UTR 相比，具有增强的活性功能，例如增强的调控 ORF 编码目的基因表达蛋白的功能。

一些实施方案中，所述 5'UTR 截短体的截短方式包括：从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；并且，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。

一些实施方案中，所述 5'UTR 截短体还包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体不包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体保留了相比于天然 5'UTR 序列长度的至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、1%长度的序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含或为 ARHGAP15 5'UTR 截短体。一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，缺失 5'末端连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸序列。

一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，5'UTR 截短体的截短方式包括从 ARHGAP15 5'UTR 的 3'至 5'末端的方向上，删除 3'末端连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，5'UTR 截短体的截短方式是从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列；并且，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸

序列。

一些实施方案中，所述截短体保留了相比于天然 ARHGAP15 5'UTR(SEQ ID NO: 1)至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、1%长度的序列。

5       一些实施方案中，所述 5'UTR 截短体在在 SEQ ID NO: 2 或 44 所示核苷酸序列中至少 5 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中 5-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 或 44 所示序列中 7-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 2 所示序列中 7-59 个连续的核苷酸  
10   序列；示例性地，所述 5'UTR 截短体包含 SEQ ID NO: 2 所示序列中 5、7、8、9、10、11、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 个连续的核苷酸序列。

15       一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸(也即，5'UTR 截短体的起始核苷酸)是 SEQ ID NO: 2 或 44 所示序列按自然计数的 1-57 位的任一位置处的核苷酸。一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸是 SEQ ID NO: 2 或 44 所示序列按自然计数的第 1-13 或第 17-29 位中的任一位置处的核苷酸。例如，第 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、  
20   20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57 的任一位置处的核苷酸。

示例性地，所述 5'UTR 截短体的 5'末端的核苷酸，及所述 5'UTR 截短体包含的 SEQ ID NO: 2 所示序列中连续核苷酸序列的长度如下表所示：

| 起始核苷酸   | SEQ ID NO: 2 所示序列中连续核苷酸序列的长度                                                                                                                                               |
|---------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 第 1 位，G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 |
| 第 2 位，A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 |
| 第 3 位，G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、                                      |

|           |                                                                                                                                                                    |
|-----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|           | 51、52、53、54、55、56、57、58、59、60 或 61                                                                                                                                 |
| 第 4 位, T  | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60 |
| 第 5 位, T  | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59    |
| 第 9 位, A  | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55                |
| 第 13 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51                            |
| 第 17 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47                                        |
| 第 21 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43                                                    |
| 第 23 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41                                                          |
| 第 25 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39                                                                |
| 第 29 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35                                                                            |
| 第 33 位, G | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31                                                                                        |
| 第 37 位, T | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27                                                                                                    |
| 第 41 位, C | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、                                                                                                                        |

|           |                                         |
|-----------|-----------------------------------------|
|           | 21、22、23                                |
| 第 45 位, A | 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19 |
| 第 49 位, G | 5、6、7、8、9、10、11、12、13、14、15             |
| 第 53 位, A | 5、6、7、8、9、10、11                         |
| 第 57 位, C | 5、6、7                                   |

一些实施方案中,所述 5'UTR 截短体的长度为 59bp、55bp、51bp、47bp、43bp、39bp、35bp、31bp、27bp、23bp、19bp、15bp、11bp、7bp。

一些实施方案中,所述 ARHGAP15 5'UTR 截短体还包含前述任意实施方案的点突变。一些实施方案中,所述 ARHGAP15 5'UTR 截短体还包含 ATG 突变为 GTG 的核苷酸序列。一些实施方案中,前述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 28-30 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,所述 ARHGAP15 5'UTR 截短体不包含前述任意实施方案的点突变。一些实施方案中,前述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 31-40 所示的核苷酸序列、CTATAAT 或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,所述 5'UTR 截短体至少包含 CTATAAT 或 SEQ ID NO: 193 所示的核苷酸序列。

一些实施方案中,所述 ARHGAP15 5'UTR 截短体包含或为如 SEQ ID NO: 41、76、137-196 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 HSPB1 基因的 5'UTR 包含或为如 SEQ ID NO: 3-5 任一所示的核苷酸序列或 SEQ ID NO: 3-5 任一具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,本公开所述核酸构建体中的 5'UTR 包含或为如 SEQ ID NO: 1-2、3-5、28-43、76、137-196 任一所示的核苷酸序列,或如下所示序列: CTATAAT; 或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些具体实施方案中,如上任一项核酸构建体中所述的 5'UTR 包含或为如 SEQ ID NO: 1-5、28-43、76、137-196 任一所示的核苷酸序列,或如下所示序列: CTATAAT。

一些实施方案中,本公开核酸构建体还进一步包含:(c) 3'非翻译区元件(3'UTR)。

一些实施方案中,本公开所述开放阅读框与所述 5'UTR 和/或所述 3'UTR 源自不同的基因。一些实施方案中,本公开核酸构建体包含至少一个开放阅读框、至少一个 5'UTR 或至少一个 3'UTR。一些实施方案中,本公开核酸构建体中的所述

5'UTR 和 3'UTR 是相同或不同来源的,例如源自相同或不同的基因。一些实施方案中,本公开核酸构建体中的所述 5'UTR 和 3'UTR 源自相同物种或不同物种。

一些实施方案中,本公开核酸构建体中的 3'UTR 位于开放阅读框的下游。一些实施方案中,所述核酸构建体中的 3'UTR 位于开放阅读框的 3'末端。

5 一些实施方案中,本公开核酸构建体中的所述 3'UTR 选自源自或为  $\beta$  珠蛋白(hemoglobin subunit beta, HBB)、ARHGAP15、冠蛋白 1A(coronin 1A, CORO1A)、血红素结合蛋白(hemopexin, HPX)等任一基因的 3'UTR 或其衍生序列。

一些实施方案中,其中所述的 3'UTR 是源自或为 HBB 或 ARHGAP15 基因的 3'UTR 序列或其衍生序列。

10 一些实施方案中,前述源自 HBB 基因的 3'UTR 包含或为如 SEQ ID NO: 7 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 ARHGAP15 基因的 3'UTR 包含或为如 SEQ ID NO: 8 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、  
15 99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 CORO1A 基因的 3'UTR 包含或为如 SEQ ID NO: 9 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 HPX 基因的 3'UTR 包含或为如 SEQ ID NO: 10  
20 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,本公开核酸构建体中的 3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。一些具体实施方案中,其中所述的 3'UTR 包含  
25 或为如 SEQ ID NO: 7 或 8 所示的核苷酸序列。

一些实施方案中,所述核酸构建体中包含 5'UTR 和 3'UTR,其中:

5'UTR 选自源自或为 ARHGAP(例如 ARHGAP15)、HSPB1、HBB、CCL13 等任一基因的 5'UTR 或其衍生序列,和 3'UTR 选自源自或为 HBB、ARHGAP15、CORO1A、HPX 等任一基因的 3'UTR 或其衍生序列。

30 一些实施方案中,所述核酸构建体中包含 5'UTR 和 3'UTR,所述 5'UTR 和 3'UTR 选自如下任意一项:

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列,和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列;

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列,和 3'UTR 源自或为  
35 ARHGAP15 基因的 3'UTR 或其衍生序列;

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列,和 3'UTR 源自或为

CORO1A 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；或

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；或

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列。

一些实施方案中，本公开核酸构建体中包含 5'UTR 和 3'UTR，所述 5'UTR 和 3'UTR 选自如下任意一项：

5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、

100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 8 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

5 5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 9 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

10 5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 10 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

15 5'UTR 包含或为如 SEQ ID NO: 3 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

20 5'UTR 包含或为如 SEQ ID NO: 3 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 8 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

25 5'UTR 包含或为如 SEQ ID NO: 3 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 9 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

30 5'UTR 包含或为如 SEQ ID NO: 3 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 10 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 4 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

35 5'UTR 包含或为如 SEQ ID NO: 4 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 8 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、

96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 4 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 9 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、

5 96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 4 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 10 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

10 5'UTR 包含或为如 SEQ ID NO: 5 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 5 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 8 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

15 5'UTR 包含或为如 SEQ ID NO: 5 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 9 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

5'UTR 包含或为如 SEQ ID NO: 5 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 10 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

25 一些实施方案中，所述核酸构建体中包含 5'UTR 和 3'UTR，其中：

5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列，和 3'UTR 源自或为源自 HBB、ARHGAP15、CORO1A、HPX 基因的 3'UTR 或其衍生序列。

30 一些具体实施方案中，所述核酸构建体中包含 5'UTR 和 3'UTR，其中：

5'UTR 包含或为如 SEQ ID NO: 1-2、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列。

一些实施方案中，本公开核酸构建体还进一步包含：(d) 多聚腺苷酸(poly-A)尾巴。

35 一些实施方案中，所述核酸构建体中的 poly-A 尾巴位于 3'UTR 的下游。一些实施方案中，所述核酸构建体中的 poly-A 尾巴位于 3'UTR 的 3'末端。一些实施方

案中，所述 poly-A 尾巴在所述核酸构建体的 3'末端。一些实施方案中，poly-A 尾巴的长度为至少约 50、100、150、200、300、400、500 个核苷酸。

一些实施方案中，所述 poly-A 尾巴包括但不限于 HGH polyA、SV40polyA、BGH polyA、rbGlob polyA 或 SV40late polyA。

5 一些具体实施方案中，所述 poly-A 尾巴包含或为如 SEQ ID NO: 16 或 135 所示的核苷酸序列。

一些实施方案中，从 5'至 3'方向上，如上任一项所述的核酸构建体含有 i)-iv) 中任一：

- i) 5'UTR，和开放阅读框(ORF)；
- 10 ii) 开放阅读框(ORF)，和 3'UTR；
- iii) 5'UTR，开放阅读框(ORF)，和 3'UTR；
- iv) 5'UTR，开放阅读框(ORF)，3'UTR，和 poly-A 尾巴；

其中所述 ORF 与所述 5'UTR 和/或所述 3'UTR 源自不同的基因。

一些实施方案中，i)、iii)至 iv)中的 5'UTR 选自源自 ARHGAP、HSPB1、HBB、  
15 CCL13 等任一基因的 5'UTR 或其衍生序列(例如 ARHGAP15 的 5'UTR 或其衍生序列)。

一些实施方案中，ii)至 iv)中的 3'UTR 选自源自 HBB、ARHGAP15、CORO1A、HPX 等任一基因的 3'UTR 或其衍生序列(例如 HBB 或 ARHGAP15 的 3'UTR 或其衍生序列)。

20 一些实施方案中，iv)至 iv)中的 poly-A 尾巴包含例如 SEQ ID NO: 16 或 135 所示的核苷酸序列。

一些实施方案中，i)中的 5'UTR 包含或为如 SEQ ID NO: 1-5、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列。

一些实施方案中，ii)中的 3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列。  
25

一些实施方案中，iii)中的 5'UTR 包含或为如 SEQ ID NO: 1-5、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列。

一些实施方案中，iv)中的 5'UTR 包含或为如 SEQ ID NO: 1-5、28-43、76、  
30 137-196、CTATAAT 任一所示的核苷酸序列，3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列，和 poly-A 尾巴包含或为如 SEQ ID NO: 16 或 135 所示的核苷酸序列。

本公开中，上述任一项核酸构建体中，所述 ORF 包含编码至少一种多肽或蛋白的核苷酸序列。一些实施方案中，所述核苷酸序列可以是密码子优化的核苷酸  
35 序列。

一些实施方案中，所述 ORF 编码的多肽或蛋白是荧光蛋白或荧光素酶

(luciferase)。

一些实施方案中，所述 ORF 编码的多肽或蛋白是病毒抗原。示例性地，病毒抗原包括但不限于流感病毒、冠状病毒、呼吸道合胞病毒、人类免疫缺陷病毒、单纯疱疹病毒、狂犬病病毒或 EB 病毒等的抗原。

5 一些实施方案中，所述病毒抗原是冠状病毒抗原。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中，所述冠状病毒是 SARS-COV-2。一些实施方案中，所述冠状病毒抗原是结构蛋白质。一些实施方案中，所述结构蛋白质选自刺突蛋白(Spike protein, S 蛋白或 Spike 蛋白)、包膜蛋白(envelope protein, E 蛋白)、膜蛋白(membrane protein, M 蛋白)和核衣壳蛋白(nucleocapsid protein, N 蛋白)。一些实施方案中，所述结构蛋白质为刺突蛋白。一些实施方案中，所述刺突蛋白为 SARS-COV-2 刺突蛋白。一些实施方案中，所述 SARS-COV-2 刺突蛋白选自 SARS-COV-2(例如野生型 SARS-COV-2)、SARS-COV-2 Alpha(B.1.1.7)、SARS-COV-2 Beta(B.1.351)、  
10 SARS-COV-2 Gamma(P.1)、SARS-COV-2 Kappa(B.1.617.1)、SARS-COV-2 Delta(B.1.617.2)、SARS-COV-2 Omicron(B.1.1.529)、SARS-COV-2 Omicron(B.A.4) 等任一病毒株的刺突蛋白。

一些实施方案中，其中所述 SARS-COV-2 刺突蛋白包含或为如 SEQ ID NO: 12、14、21、23、25、80、136 任一所示的氨基酸序列或与之具有至少 80%、85%、  
20 90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中，所述 ORF 包含密码子优化的核苷酸序列，其包含与编码 SARS-COV-2 抗原的野生型核苷酸序列(例如野生型 SARS-COV-2, SEQ ID NO: 81)具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

25 一些实施方案中，其中编码所述多肽或蛋白的 ORF 包含或为如 SEQ ID NO: 13、15、20、22、24、81、97 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，所述 ORF 编码流感病毒抗原。一些实施方案中，所述流感病毒选自 A 型流感病毒或 B 型流感病毒，示例性地，流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、A 型流感病毒 H3N8、A 型流感病毒 H2N2、A 型流感病毒 H5N1、A 型流感病毒 H9N2、A 型流感病毒 H7N7，B 型流感病毒/Victoria (例如，B 型流感病毒/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。一些实施方案中，流感病毒抗原为流感病毒的结构蛋白，例如，血凝素(hemagglutinin, HA)、神经氨酸酶(neuraminidase, NA)、离子通道蛋白 M2(M2 ion channel)、基质蛋白 M1(matrix protein)、核蛋白 NP(nucleoprotein)等。  
30 一些具体的实施方案中，所述流感病毒抗原为流感病毒的 HA 蛋白，例如 A 型流

感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria (例如, Influenza B/Washington/02/2019), B 型流感病毒 Yamagata (例如, Influenza B/Phuket/3073/2013)的 HA 蛋白。一些具体的实施方案中,所述流感病毒抗原为流感病毒的 NA 蛋白,例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria (例如, Influenza B/Washington/02/2019), B 型流感病毒 Yamagata (例如, Influenza B/Phuket/3073/2013)的 NA 蛋白。

一些实施方案中,所述 HA 蛋白包含或为如 SEQ ID NO: 83-86、98-101 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中,所述 ORF 包含密码子优化的核苷酸序列,其包含与编码 HA 抗原的野生型核苷酸序列具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

一些实施方案中,所述 ORF 包含或为 SEQ ID NO: 87-90、102-105 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

一些实施方案中,所述 ORF 包含或为 SEQ ID NO: 93-96 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

一些实施方案中,本公开提供一种核酸构建体,从 5'至 3'方向上,所述核酸构建体依次包含 5'UTR、ORF、3'UTR 和 poly-A 尾巴。一些实施方案中,所述 5'UTR 包含或为如 SEQ ID NO: 1-5、28-43、76、137-196、CTATAAT 任一所示的核苷酸序列,所述 ORF 包含或为如 SEQ ID NO: 13、15、20、22、24、81、97、87-90、93-96、102-105 任一所示的核苷酸序列,所述 3'UTR 包含或为如 SEQ ID NO: 7-10 任一所示的核苷酸序列,和所述 poly-A 尾巴包含或为如 SEQ ID NO: 16 或 135 所示的核苷酸序列。

示例性地,所述核酸构建体包含如 SEQ ID NO: 18-19 所示的核苷酸序列。

本公开中,包括其任何实施方案,所述核酸构建体是一种核酸分子,例如 cDNA。

RNA 分子

本公开提供一种 RNA 分子,其包括:

- (a)开放阅读框(ORF); 和
- (b)5'非翻译区元件(5'UTR)。

一些实施方案中,所述 ORF 为目的基因蛋白的编码多核苷酸序列。

一些实施方案中,所述目的基因为异源的。另一些实施方案中,所述目的基因为内源的。

一些实施方案中,所述 RNA 分子中的 5'UTR 位于开放阅读框的上游。一些实施方案中,所述 RNA 分子中的 5'UTR 位于开放阅读框的 5'末端。

一些实施方案中，所述 RNA 分子中的 5'UTR 选自源自或为 Rho GTPase 激活蛋白 (ARHGAP)、热休克 27kD 蛋白 1 (heat shock 27kDa protein 1, HSPB1)、β 珠蛋白 (hemoglobin subunit beta, HBB)、趋化因子 c-c-基元配体 13 (C-C motif chemokine ligand 13, CCL13) 等任一基因的 5'UTR 或其衍生序列。

5 一些实施方案中，所述 RNA 分子中的 5'UTR 是源自或为 ARHGAP 基因的 5'UTR 序列或其衍生序列。在一些实施方案中，所述 ARHGAP 包括 ARHGAP1、ARHGAP2、ARHGAP3、ARHGAP4、ARHGAP5、ARHGAP6、ARHGAP7 (DLC1)、ARHGAP8、ARHGAP9、ARHGAP10、ARHGAP12、ARHGAP13 (SRGAP1)、ARHGAP14 (SRGAP2)、ARHGAP15、ARHGAP17 (RICH1)、ARHGAP18、  
10 ARHGAP19、ARHGAP20、ARHGAP21、ARHGAP22、ARHGAP23、ARHGAP24、ARHGAP25、ARHGAP26。

一些实施方案中，其中所述的 5'UTR 是源自或为 ARHGAP15 基因的 5'UTR 序列或其衍生序列。一些实施方案中，所述 ARHGAP15 基因是来源任意物种的，如人类 ARHGAP15、狒狒 ARHGAP15、小鼠 ARHGAP15 等。

15 一些实施方案中，其中所述的 5'UTR 是源自或为人类 ARHGAP15 基因的 5'UTR 序列或其衍生序列。一些实施方案中，所述的 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 45 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100% 同一性的核苷酸序列。

一些实施方案中，本公开 RNA 分子中的 5'UTR 包含至少一个点突变，所述点  
20 突变包括前述核酸构建体任意实施方案中的点突变。一些实施方案中，所述突变为 A 突变为 G、C 或 U。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 SEQ ID NO: 79 所示的核苷酸序列，或，所述 5'UTR 包含与 SEQ ID NO: 79 具有至少 80%、85%、90%、95%、96%、97%、98%、99% 同一性的核苷酸序列，其中 N<sub>2</sub> 选自 A、G、C  
25 或 U。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 AUG 突变为 GUG 的核苷酸序列。一些实施方案中，所述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 46 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100% 同一性的核苷酸序列。

30 一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 AUG 突变为 CUG 的核苷酸序列。一些实施方案中，所述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 77 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100% 同一性的核苷酸序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含 AUG 突变为  
35 UUG 的核苷酸序列。一些实施方案中，所述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 78 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、

96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，本公开 RNA 分子中的 5'UTR 为 5'UTR 截短体。一些实施方案中，所述 5'UTR 截短体仍然保持与天然 5'UTR 相似的活性功能，例如仍然保持调控 ORF 编码目的基因表达蛋白的功能。

5 一些实施方案中，所述 5'UTR 截短体的截短方式包括：从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，所述 5'UTR 截短体的截短方式包括从 5'至 3'末端的序列方向上，删除 5'末端的连续的核苷酸序列；并且，所述 5'UTR 截短体的截短方式包括从 3'至 5'末端的序列方向上，删除 3'末端的连续的核苷酸序列。一些实施方案中，所述 5'UTR 截短体还包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体不包含前述任意实施方案的点突变。一些实施方案中，所述 5'UTR 截短体保留了相比于天然 5'UTR 序列长度的至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、1%长度的序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 序列包含或为 ARHGAP15 5'UTR 截短体。一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，缺失 5'末端连续的核苷酸序列；和/或，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸序列。

一些实施方案中，其截短方式包括从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列，保留 3'末端的核苷酸序列。一些实施方案中，5'UTR 截短体的截短方式包括从 ARHGAP15 5'UTR 的 3'至 5'末端的方向上，删除 3'末端连续的核苷酸序列，保留 5'末端的核苷酸序列。一些实施方案中，5'UTR 截短体的截短方式是从 ARHGAP15 5'UTR 的 5'至 3'末端的方向上，删除 5'末端连续的核苷酸序列；并且，从 3'至 5'末端的序列方向上，缺失 3'末端的连续的核苷酸序列。

一些实施方案中，所述截短体保留了相比于天然 ARHGAP15 5'UTR(SEQ ID NO: 45)至少 90%、80%、70%、60%、50%、40%、30%、20%、15%、10%、5%、1%长度的序列。

一些实施方案中，所述 5'UTR 截短体在在 SEQ ID NO: 46 或 79 所示核苷酸序列至少 5 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 46 或 79 所示序列中 5-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 46 或 79 所示序列中 7-62 个连续的核苷酸序列；一些实施方案中，所述 5'UTR 截短体包含 SEQ ID NO: 46 或 79 所示序列中 7-59 个连续

的核苷酸序列；示例性地，所述 5'UTR 截短体包含 SEQ ID NO: 2 所示序列中 5、7、8、9、10、11、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61 或 62 个连续的核苷酸序列。

一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸(也即，5'UTR 截短体的起始核苷酸)是 SEQ ID NO: 46 或 79 所示序列按自然计数的 1-57 位的任一位置处的核苷酸。一些实施方案中，所述 5'UTR 截短体的 5'末端的核苷酸是 SEQ ID NO: 46 或 79 所示序列按自然计数的第 1-13 或第 17-29 位中的任一位置处的核苷酸。例如，第 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57 的任一位置处的核苷酸。

一些实施方案中，所述 5'UTR 截短体的长度为 59bp、55bp、51bp、47bp、43bp、39bp、35bp、31bp、27bp、23bp、19bp、15bp、11bp、7bp。

一些实施方案中，所述 ARHGAP15 5'UTR 截短体还包含前述任意实施方案的点突变。一些实施方案中，所述 ARHGAP15 5'UTR 截短体还包含 AUG 突变为 GUG 的核苷酸序列。一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 63-65 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，所述 ARHGAP15 5'UTR 截短体不包含前述任意实施方案的点突变。一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 包含或为如 SEQ ID NO: 66-75 所示的核苷酸序列、CUAUAAU 或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 ARHGAP15 基因的 5'UTR 包含或为如包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T)或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 HSPB1 基因的 5'UTR 包含或为如 SEQ ID NO: 47 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 HBB 基因的 5'UTR 包含或为如 SEQ ID NO: 48 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，前述源自 CCL13 基因的 5'UTR 包含或为如 SEQ ID NO: 49 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,本公开所述 RNA 分子中的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。或者, RNA 分子中的 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T)或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些具体实施方案中,如上任一项 RNA 分子中所述的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列。或者, RNA 分子中的 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T)。

一些实施方案中,本公开 RNA 分子还进一步包含:(c) 3'非翻译区元件(3'UTR)。

一些实施方案中,本公开所述开放阅读框与所述 5'UTR 和/或所述 3'UTR 源自不同的基因。一些实施方案中,本公开 RNA 分子包含至少一个开放阅读框、至少一个 5'UTR 或至少一个 3'UTR。一些实施方案中,本公开 RNA 分子中的所述 5'UTR 和 3'UTR 是相同或不同来源的,例如源自相同或不同的基因。一些实施方案中,本公开 RNA 分子中的所述 5'UTR 和 3'UTR 源自相同或不同的物种。

一些实施方案中,本公开 RNA 分子中的 3'UTR 位于开放阅读框的下游。一些实施方案中,所述 RNA 分子中的 3'UTR 位于开放阅读框的 3'末端。一些实施方案中,本公开 RNA 分子中的所述 3'UTR 选自源自或为  $\beta$  珠蛋白(hemoglobin subunit beta, HBB)、ARHGAP15、冠蛋白 1A(coronin 1A, CORO1A)、血红素结合蛋白(hemopexin, HPX)等任一基因的 3'UTR 或其衍生序列。

一些实施方案中,其中所述的 3'UTR 是源自或为 HBB 或 ARHGAP15 基因的 3'UTR 序列或其衍生序列。

一些实施方案中,前述源自 HBB 基因的 3'UTR 包含或为如 SEQ ID NO: 50 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 ARHGAP15 基因的 3'UTR 包含或为如 SEQ ID NO: 51 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 CORO1A 基因的 3'UTR 包含或为如 SEQ ID NO: 52 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,前述源自 HPX 基因的 3'UTR 包含或为如 SEQ ID NO: 53 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中,本公开 RNA 分子中的 3'UTR 包含或为如 SEQ ID NO: 50-53

任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些具体实施方案中，其中所述的 3'UTR 包含或为如 SEQ ID NO: 50 或 51 所示的核苷酸序列。

5 一些实施方案中，所述 RNA 分子中包含 5'UTR 和 3'UTR，其中：

5'UTR 选自源自或为 ARHGAP(例如 ARHGAP15)、HSPB1、HBB、CCL13 等任一基因的 5'UTR 或其衍生序列，和 3'UTR 选自源自或为 HBB、ARHGAP15、CORO1A、HPX 等任一基因的 3'UTR 或其衍生序列。

10 一些实施方案中，所述 RNA 分子中包含 5'UTR 和 3'UTR，所述 5'UTR 和 3'UTR 选自如下任意一项：

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

15 5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 ARHGAP15 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；

20 5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；

25 5'UTR 源自或为 HBB 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

30 5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 HSPB1 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列；或

35 5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列，和 3'UTR 源自或为 HBB 基因的 3'UTR 或其衍生序列；

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列, 和 3'UTR 源自或为 ARHGAP15 基因的 3'UTR 或其衍生序列;

5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列, 和 3'UTR 源自或为 CORO1A 基因的 3'UTR 或其衍生序列; 或

5 5'UTR 源自或为 CCL13 基因的 5'UTR 或其衍生序列, 和 3'UTR 源自或为 HPX 基因的 3'UTR 或其衍生序列。

一些实施方案中, 本公开 RNA 分子中包含 5'UTR 和 3'UTR, 所述 5'UTR 和 3'UTR 选自如下任意一项:

5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、  
10 100%同一性的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T) 或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列; 和 3'UTR 包含或为如 SEQ ID NO: 50 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、  
15 97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、  
100%同一性的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T) 或与之具有至少 80%、85%、90%、  
20 95%、96%、97%、98%、99%、100%同一性的核苷酸序列; 和 3'UTR 包含或为如 SEQ ID NO: 51 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、  
100%同一性的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T) 或与之具有至少 80%、85%、90%、  
25 95%、96%、97%、98%、99%、100%同一性的核苷酸序列; 和 3'UTR 包含或为如 SEQ ID NO: 52 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、  
100%同一性的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T) 或与之具有至少 80%、85%、90%、  
30 95%、96%、97%、98%、99%、100%同一性的核苷酸序列; 和 3'UTR 包含或为如 SEQ ID NO: 53 所所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、  
35 97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 47 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 50 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 47 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 51 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 47 所示的核苷酸序列或与之具有至少 80%、  
10 85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR  
包含或为如 SEQ ID NO: 52 所示的核苷酸序列或与之具有至少 80%、85%、90%、  
95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 47 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 15 包含或为如 SEQ ID NO: 53 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 48 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 50 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 48 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 51 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

25 5'UTR 包含或为如 SEQ ID NO: 48 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 52 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 48 所示的核苷酸序列或与之具有至少 80%、  
30 85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR  
包含或为如 SEQ ID NO: 53 所所示的核苷酸序列或与之具有至少 80%、85%、90%、  
95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5'UTR 包含或为如 SEQ ID NO: 49 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列，和 3'UTR 包含或为如 SEQ ID NO: 50 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列；

5'UTR 包含或为如 SEQ ID NO: 49 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 51 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

5 5'UTR 包含或为如 SEQ ID NO: 49 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 52 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

10 5'UTR 包含或为如 SEQ ID NO: 49 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列, 和 3'UTR 包含或为如 SEQ ID NO: 53 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列;

一些实施方案中, 所述 RNA 分子中包含 5'UTR 和 3'UTR, 其中:

15 5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T); 和 3'UTR 源自或为 HBB、ARHGAP15、CORO1A、HPX 任一基因的 3'UTR 或其衍生序列。

一些具体实施方案中, 所述 RNA 分子中包含 5'UTR 和 3'UTR, 其中:

20 5'UTR 包含或为如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T); 和 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列。

一些实施方案中, 本公开 RNA 分子还进一步包含: (d) 多聚腺苷酸(poly-A) 尾巴。

25 一些实施方案中, 所述 RNA 分子中的 poly-A 尾巴位于 3'UTR 的下游。一些实施方案中, 所述 RNA 分子中的 poly-A 尾巴位于 3'UTR 的 3'末端。一些实施方案中, 所述 poly-A 尾巴在所述 RNA 分子的 3'末端。一些实施方案中, poly-A 尾巴的长度为至少约 50、100、150、200、300、400、500 个核苷酸。

30 一些实施方案中, 所述 poly-A 尾巴包括但不限于 HGH polyA、SV40polyA、BGH polyA、rbGlob polyA 或 SV40late polyA。

一些具体实施方案中, 所述 poly-A 尾巴包含或为如 SEQ ID NO: 56 或 135 所示的核苷酸序列。

一些实施方案中, 本公开 RNA 分子还进一步包含: (e) 5'帽结构(5'Cap)。

35 一些实施方案中, 所述 RNA 分子中的 5'帽结构位于 5'UTR 的上游。一些实施方案中, 所述 RNA 分子中的 5'帽结构位于 5'UTR 的 5'末端。在一些实施方案中, 5'帽结构是本领域技术人员已知的帽结构, 如 Cap0(第一个核碱基的甲基化, 例如

m7GpppN)、Cap1(m7GpppN 的相邻核苷酸的核糖的额外甲基化, 例如 m7G(5')ppp(5')(2'OMeA)pG)、Cap2(m7GpppN 下游第 3 个核苷酸的核糖的额外甲基化)、Cap3(m7GpppN 下游第 3 个核苷酸的核糖的额外甲基化)、Cap4(m7GpppN 下游第 4 个核苷酸的核糖的额外甲基化)、ARCA(抗反向帽类似物)、修饰的 ARCA(例如, 硫代磷酸酯修饰的 ARCA)、肌苷、N1-甲基-鸟苷、2'-氟代-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷和 2-叠氮基-鸟苷。

一些实施方案中, 使用化学 RNA 合成或 RNA 体外转录(共转录加帽)形成 5'-帽结构(如 Cap0 或 Cap1)。

一些实施方案中, 使用加帽酶(例如牛痘病毒加帽酶和/或帽依赖性 2'-O 甲基转移酶)经由酶促加帽来形成 5'-帽结构(如 Cap0 或 Cap1)。一些实施方案中, 使用固定化加帽酶, 添加 5'帽结构(Cap0 或 Cap1)。此处全文引入 WO2016/193226 中的加帽方法和手段。

一些实施方案中, 所述 5'帽结构包括但不限于 ARCA、3'OMe-m7G(5')ppp(5')G、m7G(5')ppp(5')(2'OMeA)pU、m7Gppp(A2'O-MOE)pG、m7G(5')ppp(5')(2'OMeA)pG、m7G(5')ppp(5')(2'OMeG)pG、m7(3'OMeG)(5')ppp(5')(2'OMeG)pG 或 m7(3'OMeG)(5')ppp(5')(2'OMeA)pG。

一些具体实施方案中, 所述 5'帽结构是 m7G(5')ppp(5')(2'OMeA)pG。还可以使用其他 5'帽结构或帽结构类似物。

一些实施方案中, 从 5'至 3'方向上, 如上任一项所述的 RNA 分子含有 i)-v)中任一:

- i) 5'UTR, 和开放阅读框(ORF);
  - ii) 开放阅读框(ORF), 和 3'UTR;
  - iii) 5'UTR, 开放阅读框(ORF), 和 3'UTR;
  - iv) 5'UTR, 开放阅读框(ORF), 3'UTR, 和 poly-A 尾巴;
  - v) 5'帽结构(5'Cap), 5'UTR, 开放阅读框(ORF), 3'UTR, 和 poly-A 尾巴;
- 其中所述 ORF 与所述 5'UTR 和/或所述 3'UTR 源自不同的基因。

一些实施方案中, i)、iii)至 v)中的 5'UTR 选自源自 ARHGAP、HSPB1、HBB、CCL13 等任一基因的 5'UTR 或其衍生序列(例如 ARHGAP15 的 5'UTR 或其衍生序列)。

一些实施方案中, ii)至 v)中的 3'UTR 选自源自 HBB、ARHGAP15、CORO1A、HPX 等任一基因的 3'UTR 或其衍生序列(例如 HBB 或 ARHGAP15 的 3'UTR 或其衍生序列)。

一些实施方案中, iv)至 v)中的 poly-A 尾巴包含例如 SEQ ID NO: 56 所示的核苷酸序列。

一些实施方案中, v)中的 5'帽结构包括但不限于 Cap0、Cap1(例如 m7G(5')ppp(5')(2'OMeA)pG)、Cap2、Cap3、Cap4、ARCA。

一些实施方案中, i)中的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列; 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T)。

5 一些实施方案中, ii)中的 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列。

一些实施方案中, iii)中的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T); 和 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列。

10 一些实施方案中, iv)中的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T); 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列, 和 poly-A 尾巴包含或为如 SEQ ID NO: 56 所示的核苷酸序列。

15 一些实施方案中, v)中的 5'UTR 包含或为如 SEQ ID NO: 45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列, 或者, 5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T); 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列, poly-A 尾巴包含或为如 SEQ ID NO: 56 所示的核苷酸序列, 和 5'帽结构为 m7G(5')ppp(5')(2'OMeA)pG。

20 本公开中, 上述任一项 RNA 分子中, 所述 ORF 包含编码至少一种多肽或蛋白的核苷酸序列。一些实施方案中, 所述核苷酸序列可以是密码子优化的核苷酸序列。

一些实施方案中, 所述 ORF 编码的多肽或蛋白是荧光蛋白或荧光素酶(luciferase)。

25 一些实施方案中, 所述 ORF 编码的多肽或蛋白是病毒抗原。示例性地, 病毒抗原包括但不限于流感病毒、冠状病毒、呼吸道合胞病毒、人类免疫缺陷病毒、单纯疱疹病毒、狂犬病病毒或 EB 病毒等的抗原。

一些实施方案中, 所述病毒抗原是冠状病毒抗原。一些实施方案中, 所述冠状病毒为感染人的冠状病毒, 例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中, 所述冠状病毒是 SARS-COV-2。一些实施方案中, 所述冠状病毒抗原是结构蛋白质。一些实施方案中, 所述结构蛋白质选自刺突蛋白(Spike protein, S 蛋白或 Spike 蛋白)、包膜蛋白(envelope protein, E 蛋白)、膜蛋白(membrane protein, M 蛋白)和核衣壳蛋白(nucleocapsid protein, N 蛋白)。一些实施方案中, 所述结构蛋白质为刺突蛋白。一些实施方案中, 所述刺突蛋白为 SARS-COV-2 刺突蛋白。一些实施方案中, 所述 SARS-COV-2 刺突蛋白选自 SARS-COV-2(例如野生型

30

35

SARS-COV-2)、SARS-COV-2 Alpha(B.1.1.7)、SARS-COV-2 Beta(B.1.351)、SARS-COV-2 Gamma(P.1)、SARS-COV-2 Kappa(B.1.617.1)、SARS-COV-2 Delta(B.1.617.2)、SARS-COV-2 Omicron(B.1.1.529)、SARS-COV-2 Omicron(BA.4)等任一病毒株的刺突蛋白。一些实施方案中,其中所述 SARS-COV-2 刺突蛋白包  
5 含或为如 SEQ ID NO: 12、14、21、23、25、80、136 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中,所述 ORF 包含密码子优化的核苷酸序列,其包含与编码 SARS-COV-2 抗原的野生型核苷酸序列(例如野生型 SARS-COV-2, SEQ ID NO:  
10 82)具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

一些实施方案中,其中编码所述多肽或蛋白的 ORF 包含或为如 SEQ ID NO: 54-55、59-61、82、130 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

15 一些实施方案中,所述 ORF 编码流感病毒抗原。一些实施方案中,所述流感病毒选自 A 型流感病毒或 B 型流感病毒,示例性地,流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、A 型流感病毒 H3N8、A 型流感病毒 H2N2、A 型流感病毒 H5N1、A 型流感病毒 H9N2、A 型流感病毒 H7N7, B 型流感病毒/Victoria (例如, Influenza B/Washington/02/2019), B 型流感病毒/Yamagata (例如, Influenza  
20 B/Phuket/3073/2013)等。一些实施方案中,流感病毒抗原为流感病毒的结构蛋白,例如,血凝素(hemagglutinin, HA)、神经氨酸酶(neuraminidase, NA)、离子通道蛋白 M2(M2 ion channel)、基质蛋白 M1(matrix protein)、核蛋白 NP(nucleoprotein)等。一些具体的实施方案中,所述流感病毒抗原为流感病毒的 HA 蛋白,例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria (例如, Influenza  
25 B/Washington/02/2019), B 型流感病毒/Yamagata (例如, Influenza B/Phuket/3073/2013)的 HA 蛋白。一些具体的实施方案中,所述流感病毒抗原为流感病毒的 NA 蛋白,例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria (例如, Influenza B/Washington/02/2019), B 型流感病毒 Yamagata (例如, Influenza B/Phuket/3073/2013)的 NA 蛋白。一些实施方案中,所述 HA 蛋白包含或  
30 为如 SEQ ID NO: 83-86、98-101 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中,所述 ORF 包含密码子优化的核苷酸序列,其包含与编码 HA 抗原的野生型核苷酸序列具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

35 一些实施方案中,其中编码所述多肽或蛋白的 ORF 包含或为如 SEQ ID NO: 126-129、131-134 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、

96%、97%、98%、99%、100%同一性的序列。

一些实施方案中，本公开提供一种 RNA 分子，从 5'至 3'方向上，所述 RNA 分子依次包含 5'UTR、ORF、3'UTR 和 poly-A 尾巴。一些实施方案中，所述 5'UTR 包含或为如 SEQ ID NO: 45-49、45-49、63-75、77-78、CUAUAAU 任一所示的核苷酸序列，或者，5'UTR 包含 SEQ ID NO: 41、76、137-196 所示序列所对应的序列(以 U 替换序列中的 T)；所述 ORF 包含或为如 SEQ ID NO: 54-55、59-61、82、126-134 所示的核苷酸序列，所述 3'UTR 包含或为如 SEQ ID NO: 50-53 任一所示的核苷酸序列，和所述 poly-A 尾巴包含或为如 SEQ ID NO: 56 或 135 所示的核苷酸序列。一些具体实施方案中，所述 RNA 分子包含如 SEQ ID NO: 57-58、106-125 所示的核苷酸序列。

本公开中，包括其任何实施方案，所述 RNA 分子可以是 mRNA。

本公开首次使用源自人类 Rho GTPase 激活蛋白(ARHGAP)的 5'UTR 序列用于目的基因(如 SARS-CoV-2 病毒蛋白)的重组表达，筛选、设计并成功验证了源自 ARHGAP15 5'UTR 及其改造修饰的 5'UTR 可以启动不同目的基因的高效表达。Rho GTP 酶激活蛋白 15(Rho GTPase activating protein 15, ARHGAP15)属于 ARHGAP 家族，是一种 Rac1 特异性 GTP 酶激活蛋白(GAPs)，其是 Rho 家族 GTP 酶活性的主要负向调节因子。本公开提供了更多的、具有优异表达效率的 5'UTR 和 3'UTR 选择和组合，同时，提供了具有高效预防新冠病毒感染、流感病毒感染潜力，特别是针对其各种变异株的 mRNA 疫苗。

#### 20 多核苷酸

本公开还提供一种分离的多核苷酸，所述多核苷酸编码一种或多种多肽或蛋白，其中所述多肽或蛋白是冠状病毒抗原。一些实施方案中，所述冠状病毒是 SARS-COV-2。一些实施方案中，所述冠状病毒抗原是结构蛋白质。一些实施方案中，所述结构蛋白质选自刺突蛋白(Spike protein, S 蛋白或 Spike 蛋白)、包膜蛋白(envelope protein, E 蛋白)、膜蛋白(membrane protein, M 蛋白)和核衣壳蛋白(nucleocapsid protein, N 蛋白)。一些实施方案中，所述结构蛋白质为刺突蛋白。一些实施方案中，所述刺突蛋白为 SARS-COV-2 刺突蛋白。一些实施方案中，所述 SARS-COV-2 刺突蛋白选自 SARS-COV-2(例如野生型 SARS-COV-2)、SARS-COV-2 Alpha(B.1.1.7)、SARS-COV-2 Beta(B.1.351)、SARS-COV-2 Gamma(P.1)、SARS-COV-2 Kappa(B.1.617.1)、SARS-COV-2 Delta(B.1.617.2)、SARS-COV-2 Omicron(B.1.1.529)、SARS-COV-2 Omicron(BA.4)等任一病毒株的刺突蛋白。一些实施方案中，其中所述 SARS-COV-2 刺突蛋白包含或为如 SEQ ID NO: 12、14、21、23、25、80、136 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中，所述 ORF 编码流感病毒抗原。一些实施方案中，所述流感病毒选自 A 型流感病毒或 B 型流感病毒，示例性地，流感病毒为 A 型流感病毒

H1N1、A 型流感病毒 H3N2、A 型流感病毒 H3N8、A 型流感病毒 H2N2、A 型流感病毒 H5N1、A 型流感病毒 H9N2、A 型流感病毒 H7N7，B 型流感病毒/Victoria (例如，Influenza B/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。一些实施方案中，流感病毒抗原为流感病毒的结构蛋白，  
 5 例如，血凝素(hemagglutinin, HA)、神经氨酸酶(neuraminidase, NA)、离子通道蛋白 M2(M2 ion channel)、基质蛋白 M1(matrix protein)、核蛋白 NP(nucleoprotein)等。一些具体的实施方案中，所述流感病毒抗原为流感病毒的 HA 蛋白，例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria (例如，Influenza B/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza  
 10 B/Phuket/3073/2013)的 HA 蛋白。一些具体的实施方案中，所述流感病毒抗原为流感病毒的 NA 蛋白，例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria (例如，Influenza B/Washington/02/2019)，B 型流感病毒 Yamagata (例如，Influenza B/Phuket/3073/2013)的 NA 蛋白。一些实施方案中，所述流感病毒 HA 蛋白包含或为如 SEQ ID NO: 83-86、98-101 任一所示的氨基酸序列或与之具有至少  
 15 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。

一些实施方案中，所述 ORF 包含密码子优化的核苷酸序列，其包含与编码 SARS-COV-2 抗原的野生型核苷酸序列(例如野生型 SARS-COV-2 mRNA，SEQ ID NO: 82)具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的序列。

20 一些实施方案中，所述分离的多核苷酸包含或为如 SEQ ID NO: 54-55、59-61、82、126-134 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，所述多核苷酸是 RNA，例如 mRNA。一些实施方案中，所述多核苷酸为 mRNA 的 ORF 区。

25 一些实施方案中，本公开提供前述任意一种核酸构建体或 RNA 分子在如下任一项中的用途：(1)制备疫苗；(2)在受试者体内或体外编码病毒抗原；(3)制备在受试者体内或体外编码病毒抗原的药物；(4)用于制备药物。

一些实施方案中，本公开中的核酸构建体或 RNA 分子编码目标蛋白，其中开放阅读框(ORF)具有增强的目标蛋白表达。

30

### 修饰

为了进一步改善本公开的 RNA 或多核苷酸分子的蛋白质表达的稳定性，所述 RNA 或多核苷酸分子可进一步包含一种或多种修饰(包括化学修饰)，例如骨架修饰、糖修饰、碱基修饰和/或脂质修饰等。在一些实施方案中，所述 RNA 或多核苷酸分子被均匀地修饰成某个特定修饰(例如，在整个序列中完全修饰)。例如，可以用假尿苷均匀地修饰 RNA，使得序列中的每个 U 是假尿苷。

35

与本公开有关的骨架修饰是指本公开 RNA 或多核苷酸分子中包含的核苷酸的骨架的磷酸酯的化学修饰。一些实施方案中,所述骨架修饰包括但不限于用修饰的磷酸酯完全取代骨架中未修饰的磷酸酯部分,例如可以通过用不同的取代基取代一个或多个氧原子来修饰主链的磷酸基团。一些实施方案中,所述修饰的磷酸酯包括但不限于硫代磷酸酯、亚磷酸硒酸酯、硼烷磷酸酯、硼烷磷酸酯、膦酸氢酯、氨基磷酸酯、烷基或芳基膦酸酯和磷酸三酯。

与本公开有关的糖修饰是指本公开 RNA 或多核苷酸分子中包含的核苷酸的糖的化学修饰。一些实施方案中,所述糖修饰包括但不限于将 RNA 分子的 2'羟基(OH)修饰或替换为许多不同的“氧基”或“脱氧”取代基。一些实施方案中,所述“氧基”修饰包括但不限于烷氧基、芳氧基、聚乙二醇(PEG)等的取代修饰。一些实施方案中,“脱氧”修饰包括但不限于氢、氨基(例如 NH<sub>2</sub>、烷基氨基、二烷基氨基、杂环基、芳基氨基、二芳基氨基、杂芳基氨基、二杂芳基氨基或氨基酸)修饰。

与本公开有关的碱基修饰是指本公开 RNA 或多核苷酸分子中包含的核苷酸的碱基部分的化学修饰。一些实施方案中,所述碱基修饰包括对核苷酸中腺嘌呤、鸟嘌呤、胞嘧啶和尿嘧啶的修饰。例如,本文所述的核苷和核苷酸可在主凹槽表面上被化学修饰。一些实施方案中,主要的凹槽化学修饰可包括氨基、硫醇基、烷基或卤素基团。一些实施方案中,所述碱基修饰包括但不限于用假尿苷、1-甲基-假尿苷、5-氮杂胞苷、5-甲基胞嘧啶-5'-三磷酸或 2-甲氧基腺嘌呤修饰。一些实施方案中,所述碱基修饰为假尿苷修饰。一些具体的实施方案中,可以用假尿苷均匀地修饰 RNA,使得序列中的每个 U 是假尿苷。

与本公开有关的脂质修饰是指本公开 RNA 或多核苷酸分子中包含脂质修饰。一些实施方案中,所述脂质修饰包括但不限于本公开的 RNA 或多核苷酸分子共价连接至少一个接头,以及相应的接头与至少一个脂质共价连接。一些实施方案中,所述脂质修饰包括但不限于本公开的 RNA 或多核苷酸分子与至少一个脂质共价连接(无接头)。

### UTRs

UTRs(5'UTR 和/或 3'UTR)可以作为侧翼区域提供给本公开的核酸构建体、RNA 或多核苷酸分子。UTRs 可以与在本公开的核酸构建体、RNA 或多核苷酸分子中的编码区同源或异源。侧翼区域可以包含一个或多个 5'UTR 和/或 3'UTR,所述 UTRs 可以是相同的或不同的序列。侧翼区域的任何部分可以进行密码子优化。在密码子优化之前和/或之后,侧翼区域的任何部分可以独立地包含一个或多个不同的结构或化学修饰。

为了改变本公开的核酸构建体、RNA 或多核苷酸的一种或多种特性,将与本公开的 ORF 异源的 UTRs 引入或工程化合成到本公开的核酸构建体、RNA 或多核苷酸中。然后将所述重组核酸构建体、RNA 或多核苷酸施用于细胞、组织或生物体,并测量结果,如蛋白质水平、定位和/或半衰期,以评估异源 UTR 对本公开的、

RNA 或多核苷酸产生的有益影响。一些实施方案中，所述 UTR 包括野生 UTR 或其变体，所述 UTR 变体包括在末端添加或删除一个或多个核苷酸，包括 A、T、C 或 G。一些实施方案中，所述 UTR 变体也包括任何方式进行的密码子优化或修改。一些实施方案中，所述 UTR 变体也包括本公开任何实施方式的衍生序列，例如在天然 UTR 序列的基础上，将 ATG 点突变为 GTG，用于抑制 UTR 内部 ATG 启动翻译。

# 载体

本公开还提供一种载体，其包含前述任一项所述的核酸构建体、RNA 或多核苷酸。其中核酸构建体、RNA 或多核苷酸可存在于载体中和/或可为载体的一部分，该载体例如质粒、粘端质粒、YAC 或病毒载体。载体可为表达载体，即可提供核酸构建体、RNA 或多核苷酸编码多肽表达的载体。该表达载体通常包含至少一种本公开的核酸，其可操作地连接至一个或多个适合的表达调控元件(例如启动子、终止子等)。针对在特定宿主中的表达对所述元件及其序列进行选择为本领域技术人员的常识。对本公开的编码多肽的表达有用或必需的调控元件及其他元件例如为启动子、终止子、选择标记物、前导序列、报告基因等。

本公开的核酸构建体可基于本公开的核苷酸序列的信息通过已知的方式(例如通过自动 DNA 合成和/或重组 DNA 技术)制备或获得，和/或可从适合的天然来源加以分离。

一些实施方案中，本公开的载体还包含启动子，例如所述启动子在所述核酸构建体 5'UTR 的 5'末端，例如所述启动子为 T7 启动子、T7 lac 启动子、Tac 启动子、Lac 启动子、Trp 启动子。

一些具体实施方案中，其中所述的 T7 启动子包含或为如 SEQ ID NO: 17 所示的核苷酸序列。

# 宿主细胞

本公开还提供一种宿主细胞，其包含前述任一项所述的核酸构建体、RNA 或多核苷酸。一些实施方案中，所述细胞能够表达一种或多种本公开核酸构建体、RNA 或多核苷酸编码的多肽。一些实施方案中，所述宿主细胞为细菌细胞、真菌细胞或哺乳动物细胞。

细菌细胞例如包括革兰氏阴性细菌菌株(例如大肠杆菌(*Escherichia coli*)菌株、变形杆菌属(*Proteus*)菌株及假单胞菌属(*Pseudomonas*)菌株)及革兰氏阳性细菌菌株(例如芽孢杆菌属(*Bacillus*)菌株、链霉菌属(*Streptomyces*)菌株、葡萄球菌属(*Staphylococcus*)菌株及乳球菌属(*Lactococcus*)菌株)的细胞。

真菌细胞例如包括木霉属(*Trichoderma*)、脉孢菌属(*Neurospora*)及曲菌属(*Aspergillus*)的物种的细胞；或者包括酵母属(*Saccharomyces*)(例如酿酒酵母(*Saccharomyces cerevisiae*))、裂殖酵母属(*Schizosaccharomyces*)(例如粟酒裂殖酵母(*Schizosaccharomyces pombe*))、毕赤酵母属(*Pichia*)(例如巴斯德毕赤酵母(*Pichia*

*pastoris*)及嗜甲醇毕赤酵母(*Pichia methanolica*))及汉森酵母属(*Hansenula*)的物种的细胞。

哺乳动物细胞例如包括例如 HEK293 细胞、CHO 细胞、BHK 细胞、HeLa 细胞、COS 细胞等。

5 然而,本公开也可使用两栖类细胞、昆虫细胞、植物细胞及本领域中用于表达异源蛋白的任何其他细胞。

生产或制备方法

本公开提供一种制备本公开核酸构建体、RNA 或多核苷酸的方法,以及制备其编码多肽的方法。

10 用于制备产生核酸构建体、RNA 或多核苷酸,及其编码多肽的方法及试剂,例如特定适合载体、转化或转染方法、选择标记物、诱导蛋白表达的方法、培养条件等在本领域中是已知的。类似地,适用于制造本公开的编码多肽的方法中的蛋白分离及纯化技术为本领域技术人员所公知。

一些实施方案中,所述制备核酸构建体、RNA 或多核苷酸的方法包括培养前述的宿主细胞,并从培养物中回收产生的核酸构建体、RNA 或多核苷酸。本公开的核酸构建体、RNA 或多核苷酸,及其编码多肽也可以通过本领域已知的其它产生方法获得,例如化学合成,包括固相或液相合成。

一些实施方案中,制备 RNA 分子的方法包括:制备核酸构建体或载体,然后利用所述核酸构建体或载体进行逆转录,得到 RNA 分子。一些具体的实施方案中,20 所述方法还包括对所述 RNA 分子的 5'端添加 5'Cap。

疫苗

本公开还提供一种疫苗,其包含前述任一项所述的核酸构建体、前述任一项所述的 RNA 和/或前述任一项所述的多核苷酸。一些实施方案中,所述核酸构建体、RNA 或多核苷酸编码一种或多种病毒株的一个或多个抗原。本公开提供的疫苗25 可以是单价疫苗、多价疫苗或联合疫苗。

单价疫苗

本公开提供了一种单价疫苗,其包含编码一种生物体的一个抗原。一些实施方案中,单价疫苗包含编码一种病毒株的一个抗原。

多价/联合疫苗

30 一些实施方案中,所述疫苗可包括编码相同或不同物种的两种或更多种抗原的核酸构建、RNA 或多核苷酸分子,或多个核酸构建、RNA 或多核苷酸分子。一些实施方案中,所述疫苗包括编码相同或不同病毒株的两种或更多种抗原的 RNA 或多个 RNA。在一些实施方案中,所述 RNA 可以编码 1、2、3、4、5、6、7、8、9、10、11、12 或更多个病毒抗原。

35 一些实施方案中,在上述单价疫苗和多价/联合疫苗中,所述抗原是冠状病毒抗原,所述冠状病毒例如为 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中,所述

冠状病毒抗原是结构蛋白质，例如选自刺突蛋白(Spike protein, S 蛋白或 Spike 蛋白)、包膜蛋白(envelope protein, E 蛋白)、膜蛋白(membrane protein, M 蛋白)和核衣壳蛋白(nucleocapsid protein, N 蛋白)。一些实施方案中，所述结构蛋白质为刺突蛋白，例如为 SARS-COV-2 刺突蛋白。一些实施方案中，所述 SARS-COV-2 刺突蛋白选自 SARS-COV-2(例如野生型 SARS-COV-2 mRNA)、SARS-COV-2 Alpha(B.1.1.7)、SARS-COV-2 Beta(B.1.351)、SARS-COV-2 Gamma(P.1)、SARS-COV-2 Kappa(B.1.617.1)、SARS-COV-2 Delta(B.1.617.2)、SARS-COV-2 Omicron(B.1.1.529)、SARS-COV-2 Omicron(BA.4)等任一病毒株的刺突蛋白。一些实施方案中，其中所述 SARS-COV-2 刺突蛋白包含或为如 SEQ ID NO: 12、14、21、23、25、80、136 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。一些实施方案中，编码所述 SARS-COV-2 刺突蛋白的 DNA 序列包含或为如 SEQ ID NO: 13、15、20、22、24、81、97 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。一些实施方案中，编码所述 SARS-COV-2 刺突蛋白的 RNA 序列包含或为如 SEQ ID NO: 54-55、59-61、82、130 所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。

一些实施方案中，所述 ORF 编码流感病毒抗原。一些实施方案中，所述流感病毒选自 A 型流感病毒或 B 型流感病毒，示例性地，流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、A 型流感病毒 H3N8、A 型流感病毒 H2N2、A 型流感病毒 H5N1、A 型流感病毒 H9N2、A 型流感病毒 H7N7、B 型流感病毒/Victoria (例如，Influenza B/Washington/02/2019)、B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。一些实施方案中，流感病毒抗原为流感病毒的结构蛋白，例如，血凝素(hemagglutinin, HA)、神经氨酸酶(neuraminidase, NA)、离子通道蛋白 M2(M2 ion channel)、基质蛋白 M1(matrix protein)、核蛋白 NP(nucleoprotein)等。一些具体的实施方案中，所述流感病毒抗原为流感病毒的 HA 蛋白，例如 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria (例如，Influenza B/Washington/02/2019)、B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)的 HA 蛋白。一些实施方案中，所述 HA 蛋白包含或为如 SEQ ID NO: 83-86、98-101 任一所示的氨基酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%的氨基酸同一性的序列。一些实施方案中，编码所述 HA 蛋白的 DNA 序列包含或为如 SEQ ID NO: 87-90、93-97、102-105 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序列。一些实施方案中，编码所述 HA 蛋白的 RNA 序列包含或为如 SEQ ID NO: 126-129、131-134 任一所示的核苷酸序列或与之具有至少 80%、85%、90%、95%、96%、97%、98%、99%、100%同一性的核苷酸序

列。

一些实施方案中，可以在同一脂质纳米颗粒中配制两种或更多种不同的 RNA(例如，mRNA)。其他实施方案中，两个或更多个不同的 RNA 可以分别配制在单独的脂质纳米颗粒中，然后将脂质纳米颗粒组合作为单个疫苗组合物  
5 (例如,包括编码多个抗原的多个 RNA)，或者可以单独给药。

本公开还提供了多价/联合疫苗，其包含编码一种或多种冠状病毒的 RNA 或一种或多种不同生物体的抗原。即本公开的疫苗可以是多价/联合疫苗，其靶向相同毒株/物种的一种或多种抗原，或不同毒株/物种的一种或多种抗原。

#### 递送系统

10 本公开还提供一种递送系统，其包含前述任一项所述的核酸构建体，或前述任一项所述的 RNA 分子；其中所述递送媒介物是阳离子脂质递送颗粒。一些实施方案中，其中所述颗粒是纳米颗粒。一些实施方案中，所述递送媒介物是纳米脂质颗粒。本公开中的 RNA 分子可以使用本领域任意类型的纳米脂质颗粒实现向细胞内和/或体内的递送，示例性地，纳米脂质颗粒包括但不限于 WO2017075531、  
15 WO2018081480A1、WO2017049245A2、WO2017099823A1、WO2022245888A1、WO2022150717A1、CN101291653A、CN102119217A、WO2011000107A1、CN107028886A 中所公开的脂质颗粒，上述专利通过引用全文并入本公开。

#### 药物组合物

本公开还提供一种药物组合物，其包含前述任一项所述的核酸构建体、前述  
20 任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体，和/或前述任一项所述的递送媒介物，以及药学上可接受的载体、稀释剂或赋形剂，具体地，所述药物组合物为固体制剂、注射剂、外用制剂、喷剂、液体制剂、或复方制剂。

#### 产品或试剂盒

25 本公开提供一种产品或试剂盒，其包含前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体、前述任一项所述的递送媒介物，和/或前述任一项所述的药物组合物。所述试剂盒可用于提供相关检测或诊断用途。

#### 治疗和/或预防疾病的方法和制药用途

30 本公开还提供一种向有需要的受试者使用治疗和/或预防有效量的前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体、前述任一项所述的递送媒介物，前述任一项所述的药物组合物，和/或前述任一项所述的产品或试剂盒，在制备用于治疗或/或预防疾病的药物的用途。一些实施方案中，所述疾病包括病毒感染性疾病或病毒感染相关的呼吸  
35 系统疾病。一些实施方案中，所述病毒为冠状病毒。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、

HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中，所述冠状病毒为 SARS-CoV-2。一些实施方案中，所述病毒为流感病毒。一些实施方案中，所述冠状病毒为感染人的流感病毒，例如 A 型流感病毒或 B 型流感病毒。示例性地，流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria  
5 (例如，Influenza B/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。一些实施方案中，所述病毒感染相关的呼吸系统疾病(例如 SARS-CoV-2 感染相关的呼吸系统疾病)包括单纯性感染如发热、咳嗽和咽痛、头痛、鼻炎等、肺炎、急性呼吸道感染、严重急性呼吸道感染(SARI)、低氧性呼吸衰竭及急性呼吸窘迫综合征、脓毒症和脓毒性休克、重症急性呼吸综合征(SARS)  
10 等。

本公开还提供一种治疗和/或预防疾病的方法，包括向有需要的受试者施用治疗和/或预防有效量的前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体、前述任一项所述的递送媒介物，前述任一项所述的药物组合物，和/或前述任一项所述的产品或试剂盒。一些实施方案中，所述疾病包括病毒感染性疾病或病毒感染相关的呼吸系统疾病。一些实施方案中，所述病毒为冠状病毒。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中，所述冠状病毒为 SARS-CoV-2。一些实施方案中，所述病毒为流感病毒。一些实施方案中，所述冠状病毒为感染人的流感病毒，例如 A 型流感病毒或 B 型流感病毒。示例性地，流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria (例如，Influenza B/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。一些实施方案中，所述病毒感染相关的呼吸系统疾病(例如 SARS-CoV-2 感染相关的呼吸系统疾病)包括单纯性感染如发热、咳嗽和咽痛、头痛、鼻炎等、肺炎、急性呼吸道感染、严重急性呼吸道感染(SARI)、低氧性呼吸衰竭及急性呼吸窘迫综合征、脓毒症和脓毒性休克、重症急性呼吸综合征(SARS)  
20 等。

本公开还提供了一种方法，包括向有需要的受试者施用有效量的前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体、前述任一项所述的递送媒介物，前述任一项所述的药物组合物，和/或前述任一项所述的产品或试剂盒，其能够诱导受试者中和抗体反应和/或 T 细胞免疫应答，例如针对病毒抗原的中和抗体反应，例如针对病毒抗原的 CD4+ 和/或 CD8+T 细胞免疫反应。其中相对于用针对该抗原的传统疫苗的预防性有效剂量接种的受试者的抗原抗体滴度而言，受试者的抗原抗体滴度在疫苗接种后有所增加。一些实施方案中，所述病毒为冠状病毒。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、  
30 增加。一些实施方案中，所述病毒为冠状病毒。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、  
35 增加。一些实施方案中，所述病毒为冠状病毒。一些实施方案中，所述冠状病毒为感染人的冠状病毒，例如 SARS-CoV-2(COVID-19)、SARS-CoV、HCoV-229E、

HCoV-OC43、HCoV-NL63、HCoV-HKU1 或 MERS-CoV。一些实施方案中，所述冠状病毒为 SARS-CoV-2。一些实施方案中，所述病毒为流感病毒。一些实施方案中，所述冠状病毒为感染人的流感病毒，例如 A 型流感病毒或 B 型流感病毒。示例性地，流感病毒为 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒/Victoria  
5 (例如，Influenza B/Washington/02/2019)，B 型流感病毒/Yamagata (例如，Influenza B/Phuket/3073/2013)等。

一些实施方案中，受试者是免疫的。一些实施方案中，受试者具有肺部疾病。一些实施方案中，受试者为 5 岁或更年轻，或 65 岁或以上。

一些实施方案中，该方法包括向受试者施用至少一剂、两剂、三剂、四剂及其以上的前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任  
10 项所述的疫苗、前述任一项所述的载体、前述任一项所述的递送媒介物，前述任一项所述的药物组合物，或前述任一项所述的产品或试剂盒。

一些实施方案中，在给予前述任一项所述的核酸构建体、前述任一项所述的多核苷酸、前述任一项所述的疫苗、前述任一项所述的载体、前述任一项所述的  
15 递送媒介物，前述任一项所述的药物组合物，或前述任一项所述的产品或试剂盒的 1-72 小时后，在受试者的血清中产生可检测水平的病毒抗原(如冠状病毒抗原)。一些实施方案中，在给药 1-72 小时后的所述受试者的血清中产生了至少 100 NU/mL、200 NU/mL、300 NU/mL、400 NU/mL、500 NU/mL、600 NU/mL、700 NU/mL、800 NU/mL、900 NU/mL 或 1000 NU/mL 的中和抗体滴度。

一些实施方案中，在受试者中产生的抗体滴度相对于对照增加至少 1 log。例如，相对于对照，在受试者中产生的抗体滴度可以增加至少 2、3、4、5、6、7、8、  
20 9 或 10 log。

一些实施方案中，在受试者中产生的抗体滴度相对于对照增加至少 2 倍。例如，相对于对照，受试者产生的抗体滴度增加至少 3、4、5、6、7、8、9 或 10 倍。  
25 一些实施例中，几何平均值是  $n$  个数的乘积的  $n$  次方，通常用于描述比例增长。一些实施方案中，几何平均值用于表征在受试者中产生的抗体滴度。一些实施方案中，对照可以是未接种疫苗的受试者，或施用减毒活病毒疫苗、灭活病毒疫苗或蛋白质亚单位疫苗的受试者。

### 疫苗疗效

一些实施方案中，抗原特异性免疫应答的特征在于测量接受前述实施方案的给药后受试者中产生的抗(冠状)病毒抗原的抗体滴度。抗体滴度是对受试者内的抗体的量的测量，例如，特定于抗原的特定抗原或表位的抗体。抗体滴度通常表示为提供阳性结果的最大稀释的倒数。酶联免疫吸附测定(ELISA)是用于确定抗体滴度的常见测定。  
30

一些实施方案中，抗体滴度用于评估受试者是否已经感染或确定是否需要免疫。在一些实施方案中，抗体效价用于确定自身免疫反应的强度、确定是否需要  
35

加强免疫、确定以前的疫苗是否有效以及鉴定任何近期或之前的感染。根据本公开，抗体效价可用于确定由免疫组合物(例如，RNA 疫苗)在受试者中诱导的免疫反应的强度。

一些实施方案中，在受试者中产生的抗(冠状)病毒抗原的抗体滴度相对于对照增加至少 1 log。例如，相对于对照，在受试者中产生的抗体滴度可以增加至少 2、3、4、5、6、7、8、9 或 10 log。

一些实施方案中，在受试者中产生的抗体滴度相对于对照增加至少 2 倍。例如，相对于对照，受试者产生的抗体滴度增加至少 3、4、5、6、7、8、9 或 10 倍。

一些实施方案中，抗原特异性免疫反应是以血清中和抗体滴度与冠状病毒的几何平均滴度(GMT)的比率来衡量的，称为几何平均比率(GMR)。几何平均滴度(GMT)是一组受试者的平均抗体滴度，计算方法是将所有数值相乘并取其 n 次方根，其中 n 是有可用数据的受试者数量。

一些实施方案中，对照可以是未接种疫苗的受试者，或施用减毒活病毒疫苗、灭活病毒疫苗或蛋白质亚单位疫苗的受试者。

一些实施方案中，可使用标准分析法评估疫苗效力(efficacy)(例如，参见 Weinberg et al., J Infect Dis. 2010 Jun 1 ;201 ( 11 ) : 1607 -10)。例如，疫苗效力可以通过双盲、随机、临床对照试验来衡量。疫苗效力可表示为未接种疫苗(ARU)和接种疫苗(ARV)的研究队列之间疾病发作率(AR)的比例减少，并可使用以下公式从接种疫苗组中疾病的相对风险(RR)计算出来。

效力(efficacy)=(ARU-ARV)/ARUx100; 和  
效力(efficacy)=(1-RR)x100

另一些实施方案中，可使用标准分析来评估疫苗的有效性(effectiveness)(例如，参见 Weinberg et al., J Infect Dis. 2010 Jun 1 ;201 ( 11 ) : 1607 -10)。疫苗有效性是对一种疫苗(可能已被证明具有较高的疫苗效力)如何减少人群中的疾病进行评估。这项措施可以评估在自然条件下而不是在受控的临床试验中，疫苗接种计划的有利和不利影响的净平衡，而不仅仅是疫苗本身。疫苗有效性与疫苗效力成正比，但也受人群中目标群体的免疫程度，以及其他如住院、门诊就诊或成本等非疫苗相关因素的影响。例如，可以采用回顾性的病例对照分析，即比较一组感染病例和适当的对照组中的疫苗接种率。疫苗的有效性可表示为比率差异，使用接种疫苗后仍发生感染的几率(OR)。

有效性(effectiveness)=(1-OR)x100。

#### 附图说明

图 1: 含有 T7 启动子、5'UTR、3'UTR、polyA 尾巴和外源基因 CDS 序列的质粒模板示意图。

图 2: 以线性化质粒为模板体外转录合成的 mRNA 芯片电泳图。其中，L 为

RNA 分子量标记 (RNA Ladder)，1 为经纯化的 mRNA 样品。

图 3: ELISA 检测不同 5'UTR 和 3'UTR 组合导致的 SAS-Cov-2 B.1.351 刺突蛋白表达水平差异图。其中 5'UTR 和 3'UTR 简写示意图表 3。

图 4: ELISA 检测不同 5'UTR 和 3'UTR 组合导致的 SAS-Cov-2 B.1.1.7 刺突蛋白表达水平差异图。其中 5'UTR 和 3'UTR 简写示意图表 3。

图 5: 5'UTR I 点突变为 5'UTR I'示意图。

图 6: 蛋白印记 (Western Blot) 检测 5'UTR I 点突变前后导致的 SAS-Cov-2 B.1.617.2 刺突蛋白表达水平差异图。其中，样品 1 和 3 使用的 5'UTR 为 I，样品 2 和 4 使用的 5'UTR 为 I'；样品 1 和 2 mRNA 合成模板来源于质粒小提，样品 3 和 4 mRNA 合成模板来源于质粒大提；样品 5 为未转染细胞上清对照；M 为分子量标记。图 6A 为蛋白印记检测图，图 6B 为蛋白印记检测图量化后的样品比值图。

图 7: 蛋白印记检测 5'UTR I'对不同异源基因表达水平影响图。其中，图 7A 为蛋白印记检测图，图 7B 为蛋白印记检测图量化后的样品相对值图。

图 8: 相比于 BNT162b2 UTR 组合，5'UTR I'-3'UTR B 和 5'UTR I-3'UTR B 介导的荧光素酶表达。

图 9: 5'UTR I'截短体和 5'UTR I'突变体调控荧光素酶表达的结果图。

图 10: 5'UTR I'与对照 5'UTR 调控下 HA mRNA 表达水平差异结果。

具体实施方式

术语

为了更容易理解本公开，以下具体定义了某些技术和科学术语。除非在本文中另有明确定义，本文使用的所有其它技术和科学术语都具有本公开所属领域的一般技术人员通常理解的含义。

“2019 新型冠状病毒(2019-nCoV)”的正式分类名为严重急性呼吸综合征冠状病毒 2(severe acute respiratory syndrome coronavirus 2, SARS-CoV-2)。

“2019 新型冠状病毒(2019-nCoV)引起的疾病”的正式名称为 COVID-19。

“核酸”或“核苷酸”包括 RNA、DNA 和 cDNA 分子。应当理解，由于遗传密码的简并性，可以产生编码给定蛋白质的大量核苷酸序列。术语核酸可以与术语“多核苷酸”互换使用。“寡核苷酸”是短链核酸分子。“引物”是寡核苷酸，无论是天然存在于纯化的限制酶切消化中还是合成产生，当置于诱导与核酸链互补的引物延伸产物的合成的条件(即在核苷酸和诱导剂例如 DNA 聚合酶的存在下，并在合适的温度和 pH 下)下能够充当合成起始点。为了最大化扩增效率，引物优选是单链的，但另外可选地可以是双链的。如果是双链的，则在用于制备延伸产物之前，首先对引物进行处理以分离其链。优选地，引物是脱氧核糖核苷酸。引物必须足够长以在诱导剂的存在下引发延伸产物的合成。引物的确切长度将取决于许多因素，包括温度、引物来源和使用的方法。

“载体”或“表达载体”是指复制子，例如质粒、杆粒、噬菌体、病毒、病毒体或粘粒，可连接另一个 DNA 区段，即“插入物”以实现连接区段在细胞中的复制。载体可以是设计用于递送至宿主细胞或用于在不同宿主细胞之间转移的核酸构建体。如本文所用，载体在起源和/或最终形式上可以是病毒或非病毒的，如本文所用的

5 PUC57 DNA 载体。术语“载体”涵盖与适当的控制元件结合时能够复制并且可以将基因序列转移至细胞的任何遗传元件。在一些实施方案中，载体可以是表达载体或重组载体。

“启动子”是指通过驱动核酸序列的转录来调节另一核酸序列表达的任何核酸序列，其可以是编码蛋白质或 RNA 的异源靶基因。启动子可以是组成型的、诱导型的、阻遏型的、组织特异性的或其任何组合。启动子是核酸序列的控制区域，

10 在此核酸序列的其余部分的启动和转录速率是受控的。

“基因”是指涉及产生多肽链的 DNA 片段，可以包括或不包括在前和在后的编码区域，例如，5'非翻译(5'UTR)或“先导”序列和 3'UTR 或“非转录尾区”序列，以及在各自编码片段(外显子)之间的插入序列(内含子)。

15 “重组”是指多核苷酸是克隆、限制或连接步骤的各种组合的产物，以及导致与天然存在的多核苷酸差异和/或不同的构建体的其他步骤的产物。

“引入”，在将核酸序列插入细胞的内容中，意思为“转染”、“转化”或“转导”和包括参考核酸序列并入真核或原核细胞中，其中核酸序列可以并入细胞的基因组(例如，染色体、质粒、质体，或线粒体 DNA)，转入自主复制，或暂时表达(例如，

20 转染 mRNA)。

“核酸构建体”是指单链或双链的核酸分子，例如 DNA 片段，其被修饰或合成为以天然本来不存在的方式包含核酸区段，所述核酸分子包含一个或多个控制序列或调控元件。在本公开的上下文中，核酸构建体含有重组核苷酸序列，该重组核苷酸序列基本上由任选地一个、两个、三个或多个分离的核苷酸序列组成：包

25 括 5'UTR、开放阅读框(ORF)、3'UTR。在涉及包括两个或更多个序列的构建体的实施方式中，序列在构建体中彼此可操作地连接。

“衍生序列”是指与本公开 UTR 序列高度同源(例如与本公开 UTR 序列至少 80%、85%、88%、90%、93%、95%、96%、97%、98%、99%、100%同一性)且仍保留与本公开 UTR 序列具有相同或相似功能活性的核苷酸序列。一些实施方案

30 中，所述衍生序列为在天然 UTR 序列的基础上经过一个或多个核苷酸的取代、缺失或添加得到的核苷酸序列。一些实施方案中，所述衍生序列为在天然 UTR 序列的基础上经过截短的得到的核苷酸序列。一些实施方案中，所述衍生序列为在天然 UTR 序列的基础上，将 ATG 点突变为 GTG、CTG 或 TTG，用于抑制 UTR 内部 ATG 启动翻译。

35 “可操作地连接”在本文中定义为如下结构：其中控制序列即启动子序列和/或 5'UTR 序列，被适当地置于相对于编码 DNA 序列的位置处，使得控制序列指导编

码序列的转录和 mRNA 翻译成由编码 DNA 编码的多肽序列。

“开放阅读框架”缩写为“ORF”，是指编码多肽的 mRNA 分子的片段或区域。ORF 包括连续的非重叠的框架内密码子，从起始密码子开始并用终止密码子结束，由核糖体翻译。

5 “内源的”是指来自或在生物体、细胞、组织或系统内部产生的任何物质。

“外源的”是指从生物体、细胞、组织或系统外部引入或产生的任何物质。

“同一性的序列”或“序列同一性”是指基因或蛋白质之间分别在核苷酸或氨基酸水平上的序列同一性。“同一性的序列”或“序列同一性”是蛋白质之间在氨基酸水平上的同一性量度以及核酸之间在核苷酸水平上的同一性量度。蛋白质序列同一性可以通过在比对序列时比较每个序列中给定位置的氨基酸序列来确定。类似地，核酸序列同一性可以通过在比对序列时比较每个序列中给定位置的核苷酸序列来确定。用于比对序列以供比较的方法是本领域所熟知的，此类方法包括 GAP、BESTFIT、BLAST、FASTA 和 TFASTA。BLAST 算法计算序列同一性百分比并对两个序列之间的相似性进行统计学分析。用于进行 BLAST 分析的软件可通过国家生物技术信息中心(National Center for Biotechnology Information, NCBI)网站公开获得。

“同源性”或“同源”定义为，在对准序列且必要时引入空位以实现最大序列一致性百分比之后，与靶染色体上的相应序列中的核苷酸残基一致的核苷酸残基的百分比。为了确定核苷酸序列同源性百分比的比对可以用本领域技术范围内的各种方式来实现，例如使用公开可用的计算机软件，例如 BLAST、BLAST-2、ALIGN、ClustalW2 或 Megalign(DNASTAR)软件。一些具体的实施方案中，本公开基于 BLAST 计算序列同源性百分比。本领域技术人员可以确定用于比对序列的合适参数，包括在所比较的序列的全长上实现最大比对所需的任何算法。在一些实施方案中，当同源臂的例如核酸序列(例如 DNA 序列)与宿主细胞的相应原生或未经编辑的核酸序列(例如基因组序列)至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 或更多一致时，所述序列被视为“同源”。

“替换”定义为氨基酸或核苷酸序列的如下变化：与参比多肽的氨基酸序列或核苷酸序列相比，分别由不同氨基酸或核苷酸替换一个或多个氨基酸或核苷酸产生。如果替换是保守的，则替换成多肽的氨基酸具有与其替换的氨基酸相似的结构或化学性质(例如，电荷、极性、疏水性等)。在一些实施方案中，多肽变体可具有“非保守”变化，其中替换的氨基酸在结构和/或化学性质上不同。

“缺失”定义为氨基酸或核苷酸序列的如下变化：与参比多肽的氨基酸序列或核苷酸序列相比，分别缺少一个或多个氨基酸或核苷酸残基。在多肽或多核苷酸序列的情况下，考虑到被修饰的多肽或多核苷酸序列的长度，缺失可涉及缺失 2 个、5 个、10 个、高至 20 个、高至 30 个或高至 50 个或更多个氨基酸或核苷酸残基。

“插入”或“添加”是指氨基酸或核苷酸序列的如下变化：与参比多肽的氨基酸序列或核苷酸序列相比，该变化导致分别加入了一个或多个氨基酸或核苷酸残基。 “插入”通常是指在多肽的氨基酸序列内添加一个或多个氨基酸残基(或多核苷酸内的核苷酸残基)，而“添加”可以是插入或指在多肽的 N-或 C-末端添加的氨基酸残基  
5 (或在多核苷酸的 5'或 3'末端添加的核苷酸残基)。在多肽或多核苷酸序列的情况下，插入或添加可以是高至 10 个、高至 20 个、高至 30 个、高至 50 个或更多个氨基酸(或核苷酸残基)。

“密码子优化”是指将目标序列中存在的在给定物种的高度表达的基因中一般罕见的密码子替换为在这类物种的高度表达的基因中一般常见的密码子，而替换  
10 前后的密码子编码相同的氨基酸。各种物种对特定氨基酸的某些密码子表现出特定的偏好。密码子偏好(生物体之间密码子使用的差异)通常与信使 RNA(mRNA)的翻译效率相关，继而认为所述信使 RNA 尤其取决于翻译的密码子的特性和具体的转运 RNA(tRNA)分子的利用度。选择的 tRNA 在细胞中的优势通常是肽合成中最常使用的密码子的反映。因此，基于密码子优化，基因可以针对给定生物体中的  
15 最佳基因表达进行修改。因此，最佳密码子的选择取决于宿主基因组的密码子使用偏好。

“抗原抗体”是指一种与抗原特异性结合的血清抗体。

“细胞”或“宿主细胞”包括易于被本公开内容的核酸构建体或载体转化、转染、转导等的任何细胞类型。作为非限制性实例，宿主细胞可以是分离的原代细胞、  
20 多能干细胞、CD34+细胞、诱导的多能干细胞或许多永生化细胞系(例如 HepG2 细胞)中的任何一种。或者，宿主细胞可以是组织、器官或生物体中的原位或体内细胞。

“治疗”意指给予患者内用或外用治疗剂，例如包含本公开的任一种核酸构建体的组合物，所述患者具有一种或多种疾病症状，而已知所述治疗剂对这些症状具有治疗作用。通常，在受治疗患者或群体中以有效缓解一种或多种疾病症状的量  
25 给予治疗剂，以诱导这类症状退化或抑制这类症状发展到任何临床可测量的程度。有效缓解任何具体疾病症状的治疗剂的量(也称作“治疗有效量”)可根据多种因素变化，例如患者的疾病状态、年龄和体重，以及药物在患者产生需要疗效的能力。通过医生或其它专业卫生保健人士通常用于评价该症状的严重性或进展状况的任何临床检测方法，可评价疾病症状是否已被减轻。  
30

“有效量”或“药理学上的有效量”包含足以改善或预防医学疾病的症状或病症的量。有效量还意指足以允许或促进诊断的量。用于特定患者或兽医学受试者的有效量可依据以下因素而变化：例如，待治疗的病症、患者的总体健康情况、给药的方法途径和剂量以及副作用严重性。有效量可以是避免显著副作用或毒性作用  
35 的最大剂量或给药方案。一些实施方案中，“有效量”是一种有效的 RNA 的剂量，可以产生抗原特异性免疫应答。

“预防性有效剂量”是指在临床上可接受的水平的预防病毒感染的有效剂量。在一些实施方案中，有效剂量是疫苗包装说明书中列出的剂量。本文所用的传统疫苗是指除本公开的 mRNA 疫苗以外的疫苗。例如，传统疫苗包括但不限于活的微生物疫苗、杀死的微生物疫苗、亚单元疫苗、蛋白质抗原疫苗、DNA 疫苗、类病毒颗粒(VLP)疫苗等。在示例性的实施方案中，传统疫苗是一种已经取得监管部门批准和/或由国家药品监管机构注册的疫苗，例如中国食品和药品监督管理局。

“药学上可接受的”是指这些治疗剂、材料、组合物和/或剂型，在合理的医学判断范围内，适用于与患者组织接触而没有过度毒性、刺激性、过敏反应或其他问题或并发症，具有合理的获益/风险比，并且对预期的用途是有效。

“单价疫苗”是指由一种病原生物的一个血清型抗原所制成的用于免疫接种的疫苗制品，例如，只含有本公开一种 SARS-CoV-2 毒株的一个抗原的疫苗是单价疫苗。

“多价疫苗”是指由相同病原生物的多个血清型抗原所制成的用于免疫接种的疫苗制品。例如，含有一种 SARS-CoV-2 变异株的抗原与其他不同亚型 SARS-CoV-2 变异株的一种或多种抗原混合制成的疫苗是多价疫苗。

“联合疫苗”是指由不同病原生物的多个血清型抗原所制成的用于免疫接种的疫苗制品。例如，含有本公开 SARS-CoV-2 毒株抗原与可以预防和/或治疗其他病毒的几种病原微生物混合制成的疫苗是联合疫苗。联合疫苗一般是由同一制作方式制备出单价疫苗，再混合于同一针剂中供免疫使用；或者，将几种病原微生物的抗原混合后，再制备成疫苗制品。一些具体的实施方案中，本公开以首先混合病原微生物抗原，再制备为疫苗制品。联合、多价疫苗是未来发展趋势，其优点在于：(1)节省疫苗接种的成本；(2)节省多个单个疫苗分别包装、物流和组装的成本；(3)减少对接种者，尤其是婴儿的伤害；(4)增加疫苗接种覆盖率；(5)节约疫苗接种时间；(6)减少疫苗储存空间等。联合或多价疫苗可以节省包括多次就诊、遗漏接种等在内的直接和间接成本等，如目前常用的多价联合疫苗包括肺炎链球菌、脑膜炎球菌、脊灰病毒、轮状病毒、流感、HPV 等疫苗。

本公开的疫苗的“免疫反应”是指受试者对疫苗中存在的(一种或多种)病毒蛋白(如冠状病毒)产生体液和/或细胞免疫反应。就本公开内容而言，“体液”免疫反应是指由抗体分子介导的免疫反应，包括例如分泌性 IgA 或 IgG 分子，而“细胞”免疫反应是由 T 淋巴细胞(例如 CD4+辅助和/或 CD8+T 细胞(例如 CTL)和/或其他白细胞介导的免疫反应。细胞免疫的一个重要方面涉及细胞溶解性 T 细胞(CTLs)的抗原特异反应。CTLs 对肽类抗原具有特异性，这些抗原与主要组织相容性复合体(MHC)编码的蛋白质联合呈现，并在细胞表面表达，CTLs 有助于诱导和促进细胞内微生物的破坏或溶解感染了这些微生物的细胞。细胞免疫的另一个方面涉及辅助 T 细胞的抗原特异性反应，辅助 T 细胞的作用是帮助刺激功能，并使非特异性效应细胞的活动集中在与 MHC 分子相关的肽类抗原的细胞上。细胞免疫反应还导

致细胞因子、趋化因子和其他由激活的 T 细胞和/或其他白细胞产生的此类分子。

实施例

以下结合实施例进一步描述本公开，但这些实施例并非限制着本公开的范围。

5 本公开实施例中未注明具体条件的实验方法，通常按照常规条件，如细胞培养手册，分子克隆手册；或按照原料或商品制造厂商所建议的条件。未注明具体来源的试剂，为市场购买的常规试剂。

示例性给出如下常规实验操作 1)至 4):

1) mRNA 产生

10 为了产生体外转录的 mRNA，使用 BspQ 1(Vazyme，DD4302-PC，中国)将质粒在聚腺苷酸尾下游线性化，并通过 PCR 纯化试剂盒(QIAGEN，28106，德国)进行纯化。以纯化后的线性化质粒作为模板，使用 T7 体外转录试剂盒(Invitrogen，AM1333，美国)进行体外转录。体外转录合成的 mRNA 经过 MEGAclean™ Kit(Invitrogen，AM1908，美国)纯化得到相对高纯度的 mRNA，用于后续体外细胞  
15 转染等实验。

2) 芯片电泳

为了检验经体外转录所得到的 mRNA 的完整性与纯度，使用芯片电泳仪器 (Agilent 2100，美国)和 RNA 6000 Nano 试剂盒(Agilent，5067-1511，美国)对 mRNA 进行分析。实验方法具体如下：将待测 mRNA 进行变性处理(70℃条件下孵育 2min，  
20 迅速置于冰上)，依次将处理好的样品点在芯片的对应孔位，然后运行机器，等待分析结果产生。

3) 细胞培养

材料：FBS(Gibco)、DMEM 培养基(Gibco)、Opti-MEM(Gibco)、PBS(Gibco)、胰蛋白酶-EDTA(Gibco)、双抗(Pen/Strep,Gibco)和 Lipofectamine™2000(Invitrogen，  
25 52887)。

使人胚肾细胞(293T，ATCC)在补充有 10%FBS 和 1%双抗的 DMEM 培养基中，并在 5% CO<sub>2</sub> 的潮湿气体环境下生长。

4) 体外转染

转染前 8 小时，按  $8 \times 10^5$  个细胞/孔的标准，将细胞接种在 12 孔细胞培养板中。  
30 使用市售转染试剂 Lipofectamine™2000 按照每 1μg mRNA 与 2.5μL Lipofectamine™2000 混合的比例进行细胞转染，每一个孔的 mRNA 转染剂量为 2μg。实验方法具体如下：将 Lipofectamine™2000 和 mRNA 分别稀释在 Opti-MEM 培养基中，最终总体积分别为 100μL，然后将 Lipofectamine™2000 溶液和 mRNA 溶液混合，室温条件下孵育 10-15 分钟，形成含有 mRNA 的脂质体复合物。在去除细胞  
35 培养孔板的 DMEM 培养基后，将含有不同 mRNA 的脂复合物加到对应孔板中，然后在每个孔中补充 400μL 含有 10% FBS 的 Opti-MEM 培养基，细胞在 37℃(5%

CO<sub>2</sub> 水平)培养 4-6 个小时。此后，去除转染培养基，在每个细胞培养孔板中加入 1mL 补充有 10%FBS 和 1%双抗的 DMEM 培养基，将细胞在 37°C(5%CO<sub>2</sub> 水平)培养 48 小时。与此同时，按照上述方式对部分细胞转染不含 mRNA 的 Lipofectamine™2000 溶液，作为阴性对照。

5            这里对本公开所使用的目的基因序列进行总结，见表 1。

表 1. 本公开的目的基因序列信息

| 目的基因编码的蛋白               | 多肽序列         | DNA 序列       | mRNA 序列      | 经优化的 DNA 序列  | 经优化的 mRNA 序列 | 蛋白较野生型 SARS-COV-2 突变位点 (KV-PP) |
|-------------------------|--------------|--------------|--------------|--------------|--------------|--------------------------------|
| B.1.1.7(Alpha) 刺突蛋白     | SEQ ID NO:23 | /            | /            | SEQ ID NO:22 | SEQ ID NO:60 | 983-984                        |
| B.1.351(Beta) 刺突蛋白      | SEQ ID NO:21 | /            | /            | SEQ ID NO:20 | SEQ ID NO:59 | 983-984                        |
| B.1.617.2(Delta) 刺突蛋白   | SEQ ID NO:12 | /            | /            | SEQ ID NO:13 | SEQ ID NO:54 | 984-985                        |
| B.1.1.529(Omicron) 刺突蛋白 | SEQ ID NO:14 | /            | /            | SEQ ID NO:15 | SEQ ID NO:55 | 983-985                        |
| SARS-COV-2 刺突蛋白         | SEQ ID NO:25 | /            | /            | SEQ ID NO:24 | SEQ ID NO:61 | 986-987                        |
| 野生型 SARS-COV-2 刺突蛋白     | SEQ ID NO:80 | SEQ ID NO:81 | SEQ ID NO:82 | /            | /            | /                              |

SARS-COV-2 刺突蛋白的氨基酸 KV 突变为 PP 能使刺突蛋白由不稳定的融合前构象变成稳定的融合后构象 (Structure-based design of prefusion-stabilized SARS-CoV-2 spikes, Science, VOL. 369, NO. 6510)，使刺突蛋白在构象上处于稳定状态，有利于刺突蛋白作为免疫原进行疫苗设计和生产。

10

实施例 1、5'UTR 序列的筛选和制备

1、UTR 的筛选

为获得能在树突细胞 (DC, dendritic cell)中高表达目的蛋白的调控元件，本实施例经过四级筛选获得新结构 UTR。

15

方法包括：一：建立蛋白高表达基因库，包括单核细胞中蛋白表达量高基因、在 DC 中比单核细胞高蛋白表达基因、蛋白水平较高的基因以及经测试会增加蛋白表达量的 UTR 的基因，得到 2140 个基因。二：建立 UTR 库，从 Genbank 数据库(<https://www.ncbi.nlm.nih.gov/genbank/>)下载蛋白水平高表达的基因库中所有

序列后，先分析完整性，之后通过重复合并、CDS(coding sequence，编码序列)截取等方法分析，获得 UTR 数据，得到 940 个 UTR。三：建立 UTR 精选库，对 UTR 库进行分析，筛选出序列长度以及自由能合适的 UTR 序列，其中 5'UTR 序列长度在 40 到 70 之间，3'UTR 长度在 100 到 150 之间或者 300 到 400 之间，将  
5 序列逐一进行自由能预测，挑选 5'UTR 自由能较高序列，挑选 3'UTR 自由能较低序列，从该精选库中得到 64 个 UTR。四：确定最终候选序列，筛选策略为 5'UTR 自由能在-10 以上，3'UTR 自由能很低，最终确定 20 个 5'UTR 和 3'UTR 进行筛选和细胞内验证。

基于以上，我们认为源自人类 Rho GTPase 激活蛋白 15(Rho GTPase activating  
10 protein 15，简称 ARHGAP15)的 5'UTR 序列(命名为 5'UTR I，DNA 序列如 SEQ ID NO： 1 所示，RNA 序列如 SEQ ID NO： 45 所示)具有在人体细胞(特别是 DC)内高效表达目的蛋白的潜力。

2、UTR 的制备

本实施例示例性给出含 5'UTR I 的 mRNA 的制备过程。首先，人工合成含有  
15 5'UTR、3'UTR、polyA 尾巴和外源目的基因 CDS 的序列，并将其插入到 pUC57 载体上，制备出模板载体，其示意图见图 1。经测序，载体中插入的目的基因序列正确。

如图 1 的 pUC57 Delta13 载体所示，其中 5'UTR 为 SEQ ID NO： 1；3'UTR 为人类血红蛋白 β 亚基(HBB) 3'UTR 序列(SEQ ID NO： 7，命名为 3'UTR B)；CDS  
20 序列为编码 B.1.617.2(Delta) 刺突蛋白的经密码子优化的碱基序列(SEQ ID NO： 13)；polyA 尾巴为多聚腺苷酸序列(DNA 序列如 SEQ ID NO： 16 所示，RNA 序列如 SEQ ID NO： 56 所示)；T7 启动子(SEQ ID NO： 17)。

为了产生体外转录的 mRNA，使用 BspQ I(Vazyme，DD4302-PC，中国)将质粒在聚腺苷酸尾下游线性化，并通过 PCR 纯化试剂盒(QIAGEN，28106，德国)进  
25 行纯化。以纯化后的线性化质粒作为模板，按照下列过程进行体外转录合成 mRNA。合成 100μL 的 mRNA 反应体系，反应体系见表 2。反应体系中，线性化模板为 2μg，用无核苷酶水补足 100μL，37°C 反应 4 小时，DNase I 消化 30min。经检测，合成的 mRNA 中副产物 dsRNA 少，产量优。加帽过程采用化学加帽方式在体外转录合成过程中实现，帽子结构为 m7G(5')ppp(5')(2'OMeA)pG·NH<sub>4</sub> (Hongene，ON-134)。  
30 经过 MEGAclear™ Kit(Invitrogen，AM1908,美国)纯化得到相对高纯度的 mRNA，用于后续体外细胞转染等实验。

表 2. mRNA 反应体系

| 成分            | 货号    | 厂家      | 体系(μL) |
|---------------|-------|---------|--------|
| ATP 溶液(100mM) | R1331 | Hongene | 4.41   |
| CTP 溶液(100mM) | R3331 | Hongene | 5      |
| GTP 溶液(100mM) | R2331 | Hongene | 2.79   |

|                |                          |         |      |
|----------------|--------------------------|---------|------|
| 10×转录缓冲液       | DD4101R-02               | Vazyme  | 10   |
| T7 RNA 聚合酶     | DD4101-PC-02             | Vazyme  | 6.6  |
| 小鼠 RNase 抑制剂   | DD4102-PA                | Vazyme  | 2.75 |
| 焦磷酸酶           | DD4103-PC-01             | Vazyme  | 0.66 |
| Pseudo-UTP     | R5-022                   | Hongene | 2.79 |
| Cap-GAG Trimer | m7G(5')ppp(5')(2'OMeA)pG | Hongene | 1    |

经过层层筛选和功能验证，本公开获得了具有高效表达目的蛋白潜力的 5'UTR I，其序列为 SEQ ID NO: 1；且含有 5'UTR I 的构建可以高效转录出对应的 mRNA，经芯片电泳鉴定与理论大小一致(见图 2)。

5           **实施例 2、筛选高效表达异源蛋白的不同 5'UTR 和 3'UTR 的组合**

本实施例中，对不同的 5'UTR 和不同的 3'UTR 的组合进行筛选，以获得能高效表达目的基因的 UTR 组合。

方法为：人工合成构建体含有相同目的基因的 CDS 的不同 UTR 组合的质粒。以实施例 1 的方法构建基础质粒，通过人工合成和 HindIII/XhoI 酶切的方式重组构建替换目的基因 CDS 序列，将重新构建的质粒按照实施例 1 的方式线性化并合成纯化 mRNA。通过 Lipofectamine™ 2000 将 mRNA 转染 293T 细胞，48h 后，收取 293T 上清，分别进行蛋白印记和 ELISA 检测，检测分泌出的细胞上清中刺突蛋白的水平。

所使用的目的基因为：编码新冠 B.1.351 刺突蛋白的基因(CDS 如 SEQ ID NO: 20 所示)，或编码新冠 B.1.1.7 刺突蛋白的基因(CDS 如 SEQ ID NO: 22 所示)。所使用的不同 5'UTR 与 3'UTR 的信息及其组合方式如表 3 和表 4 所示。

蛋白印记方法：为了检测各构建所得到的 mRNA 的体外表达水平，转染完成后收集孔板中的细胞上清，通过聚丙烯酰胺凝胶电泳、转膜、一抗和二抗孵育和显色，最终得到各构建的 mRNA 的体外表达水平。其中，一抗使用 SARS-CoV-2 (2019-nCoV) 刺突 RBD 抗体 (Sino Biology, 40592-T62)，二抗使用 HRP-linked 的抗兔 IgG(Transgene, HS101)。

ELISA 方法：转染完成后，收集相应的细胞上清，按照 1:50 至 1: 100 的比例对细胞上清进行稀释，然后使用 ELISA 试剂盒(SARS-CoV-2 (2019-nCoV) 刺突检测 ELISA 试剂盒, KIT40591, Sino Biology)对各构建的 mRNA 在体外的表达水平进行检测，读取 450nm 的吸光值并计算相关抗原的浓度。

表 3. 5'UTR 与 3'UTR 来源及序列

| UTR   | 简称 | 来源         | GenBank ID     | DNA 序列<br>(SEQ ID NO: ) | RNA 序列<br>(SEQ ID NO: ) |
|-------|----|------------|----------------|-------------------------|-------------------------|
| 5'UTR | I  | 人 ARHGAP15 | XM_003822828.3 | 1                       | 45                      |

|       |   |                                     |             |    |    |
|-------|---|-------------------------------------|-------------|----|----|
|       | G | 人热休克 27kDa 蛋白<br>1(HSPB1)           | BC012292.1  | 3  | 47 |
|       | A | 人 DNA, 含有 HBB 基因                    | LC507571.1  | 4  | 48 |
|       | J | 人 C-C 基序趋化因子配体<br>13 (CCL13)        | NM_005408.3 | 5  | 49 |
| 3'UTR | B | 人 HBB                               | MK476504.1  | 7  | 50 |
|       | F | 人 ARHGAP15                          | NM_018460.4 | 8  | 51 |
|       | E | 人肌动蛋白结合蛋白<br>1A(coronin 1A, CORO1A) | NM_007074.4 | 9  | 52 |
|       | D | 人血液结合素<br>(hemopexin, HPX)          | NM_000613.3 | 10 | 53 |

表 4. 不同 5'UTR 与 3'UTR 组合

| 名称  | 5'UTR | 3'UTR |
|-----|-------|-------|
| A-B | A     | B     |
| A-E | A     | E     |
| A-D | A     | D     |
| A-F | A     | F     |
| G-E | G     | E     |
| G-B | G     | B     |
| G-D | G     | D     |
| I-E | I     | E     |
| I-B | I     | B     |
| I-F | I     | F     |
| I-D | I     | D     |
| J-E | J     | E     |

表 4 中 5'UTR 的 A、G、I 和 J 分别如 SEQ ID NO: 4、3、1 和 5 所示; 3'UTR 的 B、E、D 和 F 分别如 SEQ ID NO: 7、9、10 和 8 所示。

5 不同 5'UTR 与 3'UTR 组合对新冠 B.1.351 刺突蛋白、新冠 B.1.1.7 刺突蛋白表达的影响分别如图 3 和表 5、图 4 和表 6 所示。

表 5. 不同 UTR 组合对 B.1.351 刺突蛋白表达量的调控结果(Elisa 检测)

| 样品编号                      | #1: A-B   | #3: A-E | #5: A-D  | #8: G-B   | #10: G-D | #11: I-E  | #13: I-B   |
|---------------------------|-----------|---------|----------|-----------|----------|-----------|------------|
| 刺突蛋白浓度<br>(mean±SD ng/mL) | 48.2±13.4 | 6.3±5.4 | 2.9±0.28 | 36.7±10.5 | 26±0.14  | 43.4±2.97 | 69.4±10.88 |

(样品编号规则为, 例如, #1: A-B 代表的是#1 号样品, 其是 A-B 的 UTR 组合)

表 6. 不同 UTR 组合对 B.1.1.7 刺突蛋白表达量的调控结果(Elisa 检测)

| 样品编号 | UK-#11: I-E | UK-#13: I-B | UK-#14: I-F | UK-#15: I-D | NC<br>(Lipo 对照) |
|------|-------------|-------------|-------------|-------------|-----------------|
|------|-------------|-------------|-------------|-------------|-----------------|

|                           |             |             |            |             |            |
|---------------------------|-------------|-------------|------------|-------------|------------|
| 刺突蛋白浓度<br>(mean±SD ng/mL) | 118.35±3.61 | 206.1±19.66 | 227.6±4.38 | 153.55±7.99 | -9.65±0.64 |
|---------------------------|-------------|-------------|------------|-------------|------------|

(样品编号规则为，例如，UK-#11: I-E 代表的是 UK-#11 号样品，其是 I-E 的 UTR 组合)

根据图 3 和表 5 可知，不同 UTR 组合介导的 B.1.351 刺突蛋白表达水平比较如下：I-B>A-B>G-B，I-B>I-E。对于 5'UTR，I 优于 A 和 G；对于 I 的相关组合，I-B 优于 I-E。根据图 4 和表 6 可知，不同 UTR 组合介导的 B.1.1.7 刺突蛋白表达水平比较如下：I-F>I-B>I-D>I-E。综合 B.1.351 刺突和 B.1.1.7 刺突蛋白表达水平的结果分析，I-B 的 UTR 组合具有更显著的优势，可介导目的基因高效表达，此外，I-F 组合也显示了高水平表达目的基因的优势。

进一步地，我们制备含有 I-B 或 I-F 调控下的 B.1.617.2 刺突 mRNA，经细胞转染，ELISA 检测，对刺突蛋白表达量的检测结果如表 7 所示。

表 7. I-B 和 I-F UTR 调控的 B.1.617.2 刺突蛋白表达量(ELISA 检测)

| 样品            | I-B  | I-F  |
|---------------|------|------|
| 刺突蛋白浓度(ng/mL) | 57.1 | 58.8 |

不同于 I-B 组合中的 5'UTR 和 3'UTR 来自不同基因，I-F 组合中的 5'UTR 和 3'UTR 来自同一基因 ARHGAP15。表 7 结果显示，I-B (SEQ ID NO: 1 和 SEQ ID NO: 7)和 I-F (SEQ ID NO: 1 和 SEQ ID NO: 8)组合对目的基因表达的调控水平相当，进一步说明了 5'UTR I 的通用性，其可以和不同的 3'UTR 组合，介导目的基因(例如，刺突蛋白的编码基因)的高效表达。

实施例 3、5'UTR I 的序列改造和初步功能验证

1) 生物信息学分析和突变设计

5'UTR I(SEQ ID NO: 1)为天然存在人的 Rho GTPase 激活蛋白家族启动子。实施例 1-2 实验数据支持 5'UTR I 协助高效表达新冠刺突蛋白，但 5'UTR I 内部含有一个 ATG，可能会表达一个 27aa 的短肽，影响正常的目的基因蛋白表达。为了避免核糖体从 UTR 内部 ATG 启动翻译的现象，增加从 CDS 区 ATG 翻译的概率，增加表达量，本公开进行了点突变(A→G)，获得了突变体 5'UTR I'(DNA 序列如 SEQ ID NO: 2 所示，RNA 序列如 SEQ ID NO: 46 所示)。突变位置见图 5。

2) 5'UTR I 通过点突变为 5'UTR I'能有效增加目的基因的表达

采用实施例 1 方法制备分别含有 5'UTR I-B 和 5'UTR I'-B 调控目的基因表达的质粒，所述目的基因为编码 B.1.617.2 刺突蛋白的基因(SEQ ID NO:13)，所述质粒分别命名为 Delta 13、Delta 24，经 PCR 转录合成、纯化获得 mRNA(其中不含帽的 I'-Delta24-B mRNA 序列如 SEQ ID NO: 57 所示)，然后将制备出的 mRNA 转染 293T 细胞，再采用实施例 2 中的蛋白印记的方法检测 B.1.617.2 刺突蛋白的

表达。并通过 Image 软件进行灰度扫描，计算灰度值，以样品 1(Delta 13)的表达量计为 1，归一化处理后，得到表格 8 中的比值。

结果如图 6A 和 6B，表 8 所示。结果可知，无论质粒小提(未去除内毒素)作为模板，还是质粒大提(去除内毒素)作为模板制备的 mRNA，点突变后的 5'UTR I'的构建中刺突蛋白的表达量均大于含有 5'UTR I 的构建体。因此，5'UTR I 经过实

5 施例 3 中点突变改造有效提高了异源蛋白的表达。

表 8. I-B 和 I'-B UTR 调控下目的基因的表达水平(蛋白印记检测)

| 样品               | 质粒小提               |                    | 质粒大提               |                    |
|------------------|--------------------|--------------------|--------------------|--------------------|
|                  | 样品 1<br>(Delta 13) | 样品 2<br>(Delta 24) | 样品 3<br>(Delta 13) | 样品 4<br>(Delta 24) |
| 与样品 1 表达量的<br>比值 | 1.00               | 4.88               | 3.98               | 5.44               |

实施例 4、5'UTR I'调控下的不同目的基因 mRNA 的表达检测

10 通过酶切方式替换目的基因，保留 5'UTR I'和 3'UTR B，构建分别编码 B.1.617.2(Delta) 刺突蛋白的基因(SEQ ID NO:13)和 B.1.1.529(Omicron) 刺突蛋白的基因(SEQ ID NO:15)的质粒，所述质粒含有的目的基因序列分别是 SEQ ID NO: 13 和 SEQ ID NO: 15，均是密码子优化后的序列，其编码如 SEQ ID NO: 12 和 SEQ ID NO: 14 所示的靶蛋白序列。采用实施例 1 的方式制备和纯化 mRNA，通

15 过 Lipofectamine™ 2000 转染 mRNA 2μg 至 293T 细胞，2d 后收取细胞上清。采用实施例 2 中的蛋白印记的方法检测 I-B 调控下的 B.1.617.2 刺突蛋白 mRNA(命名为 Delta 13)、I'-B 调控下的 B.1.617.2 刺突蛋白 mRNA(命名为 Delta 24)和 I'-B 调控下的 B.1.1.529 刺突蛋白 mRNA(命名为 Omicron 24)的表达情况，其中，不含帽结构的 Delta 24 mRNA、Omicron 24 mRNA 序列分别如 SEQ ID NO: 57 所示

20 和 SEQ ID NO: 58 所示。通过 Image 软件进行灰度扫描，计算灰度值，以 Delta 13 的表达量计为 1，归一化处理后，得到表 9 中的比值。

图 7A 和 7B 中显示在 I'-B 的 UTR 组合调控下，新冠基因 B.1.617.2 和 B.1.1.529 均可高效表达。因此提示 5'UTR I'具有通用性，可以启动不同目的基因的高效表达。对比 5'UTR I 和 I'，Delta 24 刺突蛋白表达量明显高于 Delta 13 刺突蛋白表

25 达量，提示了点突变改造 5'UTR I'有效提高了蛋白表达效率，此结果进一步证实了实施例 3 的结论。

表 9. I-B 和 I'-B UTR 组合调控下目的基因的表达水平(蛋白印记检测)

| 样品                  | Delta 13 | Delta 24 | Omicron 24 |
|---------------------|----------|----------|------------|
| 与样品 Delta 13 表达量的比值 | 1.00     | 1.69     | 1.48       |

实施例 5、I'-B 的 UTR 组合调控下的荧光素酶(luciferase)的表达量检测

构建含有荧光素酶 CDS 序列(SEQ ID NO: 26)的载体, 包含与前述含有新冠基因载体中相同的 T7 启动子(SEQ ID NO: 17), polyA 尾。分别构建 UTR 不同的三个载体 Luc, Luc13 和 Luc24, 载体包含的 UTR 见表 10。其中, Luc 中的 BNT162b2 的 5'UTR 和 3'UTR 组合作为对照, 其来自 CN113521269A。通过与实施例 1 相同的方法进行体外转录合成含有帽子结构的 mRNA, 体外转染, 收取上清后, 加荧光素酶底物检测, 读板, 结果见表 11 和图 8。

表 10. 载体结构

| 样品编号<br>元件 | Luc                     | Luc13 | Luc24 |
|------------|-------------------------|-------|-------|
| 5'UTR      | BNT162b2(SEQ ID NO: 6)  | I     | I'    |
| 3'UTR      | BNT162b2(SEQ ID NO: 11) | B     | B     |

表 11. I'-B 和 I-B 调控下的荧光素酶表达量

| 样品  | Luc    | Luc13  | Luc24   |
|-----|--------|--------|---------|
| 实验一 | 774.81 | 992.36 | 1619.76 |
| 实验二 | 841.14 | 923.83 | 1519.50 |

注释: 参数以相对光单位(relative light unit, RLU)表示。

结果显示, 经点突变后的 5'UTR I'调控下的荧光素酶蛋白表达量高于 5'UTR I。与 Luc 对照中的 BNT162b2 的 5'UTR 和 3'UTR 组合相比, 本公开的 I-B 组合使得目的基因的表达量更高, I'-B 组合使得目的基因的表达量显著升高。而所述使用荧光素酶作为目的基因的检测结果也再次说明, 5'UTR I 或 5'UTR I'均可以通用性的调控任意目的蛋白的高效表达。

实施例 6、5'UTR I'截短体和突变体调控荧光素酶(luciferase)的表达量检测

6.1 对 5'UTR I'的 5'末端进行连续截短, 通过生物信息学方法(Nat Biotechnol. 2019 Jul;37(7):803-809.), 分析不同长度的 5'UTR I'截短体调控蛋白表达的效率。计算 5'UTR I'截短体与 5'UTR I'的相对 MRL 值, 结果如下表 12 所示。由表 12 结果可知, 对 5'UTR I'中 5'末端的核苷酸序列进行连续截短, 不会引起 MRL 值的显著降低, 说明 5'UTR I'截短体能够保留调控目标蛋白表达的功能活性。SEQ ID NO: 153 所示的 5'UTR I'截短体的 3'末端的核苷酸序列进行连续截短, 计算 3'末端截短后截短体与 SEQ ID NO: 153 所示序列的相对 MRL 值, 结果如表 13 所示。由表 13 可知, 3'末端截短未对 5'UTR I'截短体的调控效率产生显著影响。

表 12 5'UTR I'截短体与 5'UTR I'的相对 MRL 值

| 5'UTR I'截短体                                                        | MRL 相对值 | SEQ ID NO:     |
|--------------------------------------------------------------------|---------|----------------|
| AGTTAGGAGTTGGTGGCAGTTTCAATAACAGG<br>TCATTGCCGAGAAAAGGATAGCACTATAAT | 1       | SEQ ID NO: 137 |
| GTTAGGAGTTGGTGGCAGTTTCAATAACAGGT<br>CATTGCCGAGAAAAGGATAGCACTATAAT  | 1       | SEQ ID NO: 138 |
| TTAGGAGTTGGTGGCAGTTTCAATAACAGGTC<br>ATTGCCGAGAAAAGGATAGCACTATAAT   | 1       | SEQ ID NO: 139 |
| TAGGAGTTGGTGGCAGTTTCAATAACAGGTCA<br>TTGCCGAGAAAAGGATAGCACTATAAT    | 1       | SEQ ID NO: 28  |
| AGGAGTTGGTGGCAGTTTCAATAACAGGTCAT<br>TGCCGAGAAAAGGATAGCACTATAAT     | 1       | SEQ ID NO: 140 |
| GGAGTTGGTGGCAGTTTCAATAACAGGTCATT<br>GCCGAGAAAAGGATAGCACTATAAT      | 1       | SEQ ID NO: 141 |
| GAGTTGGTGGCAGTTTCAATAACAGGTCATTG<br>CCGAGAAAAGGATAGCACTATAAT       | 1       | SEQ ID NO: 142 |
| AGTTGGTGGCAGTTTCAATAACAGGTCATTGC<br>CGAGAAAAGGATAGCACTATAAT        | 1       | SEQ ID NO: 29  |
| GTTGGTGGCAGTTTCAATAACAGGTCATTGCC<br>GAGAAAAGGATAGCACTATAAT         | 1       | SEQ ID NO: 143 |
| TTGGTGGCAGTTTCAATAACAGGTCATTGCCG<br>AGAAAAGGATAGCACTATAAT          | 1       | SEQ ID NO: 144 |
| TGGTGGCAGTTTCAATAACAGGTCATTGCCGA<br>GAAAAGGATAGCACTATAAT           | 1       | SEQ ID NO: 145 |
| GGTGGCAGTTTCAATAACAGGTCATTGCCGAG<br>AAAAGGATAGCACTATAAT            | 1       | SEQ ID NO: 30  |
| GTGGCAGTTTCAATAACAGGTCATTGCCGAGA<br>AAAGGATAGCACTATAAT             | 1       | SEQ ID NO: 146 |
| TGGCAGTTTCAATAACAGGTCATTGCCGAGAA<br>AAGGATAGCACTATAAT              | 0.8     | SEQ ID NO: 147 |
| GGCAGTTTCAATAACAGGTCATTGCCGAGAAA<br>AGGATAGCACTATAAT               | 1       | SEQ ID NO: 148 |
| GCAGTTTCAATAACAGGTCATTGCCGAGAAAA<br>GGATAGCACTATAAT                | 1.02    | SEQ ID NO: 31  |
| CAGTTTCAATAACAGGTCATTGCCGAGAAAAG                                   | 1.02    | SEQ ID NO: 149 |

|                                                   |      |                |
|---------------------------------------------------|------|----------------|
| GATAGCACTATAAT                                    |      |                |
| AGTTTCAATAACAGGTCATTGCCGAGAAAAGG<br>ATAGCACTATAAT | 1.02 | SEQ ID NO: 150 |
| GTTTCAATAACAGGTCATTGCCGAGAAAAGGA<br>TAGCACTATAAT  | 1    | SEQ ID NO: 151 |
| TTTCAATAACAGGTCATTGCCGAGAAAAGGAT<br>AGCACTATAAT   | 1.01 | SEQ ID NO: 32  |
| TTCAATAACAGGTCATTGCCGAGAAAAGGATA<br>GCACTATAAT    | 1.01 | SEQ ID NO: 152 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAG<br>CACTATAAT     | 1.01 | SEQ ID NO: 153 |
| CAATAACAGGTCATTGCCGAGAAAAGGATAGC<br>ACTATAAT      | 1.02 | SEQ ID NO: 154 |
| AATAACAGGTCATTGCCGAGAAAAGGATAGCA<br>CTATAAT       | 1.02 | SEQ ID NO: 33  |
| ATAACAGGTCATTGCCGAGAAAAGGATAGCAC<br>TATAAT        | 1.02 | SEQ ID NO: 155 |
| TAACAGGTCATTGCCGAGAAAAGGATAGCACT<br>ATAAT         | 1.01 | SEQ ID NO: 156 |
| AACAGGTCATTGCCGAGAAAAGGATAGCACTA<br>TAAT          | 1.01 | SEQ ID NO: 157 |
| ACAGGTCATTGCCGAGAAAAGGATAGCACTAT<br>AAT           | 1    | SEQ ID NO: 34  |
| CAGGTCATTGCCGAGAAAAGGATAGCACTATA<br>AT            | 1    | SEQ ID NO: 158 |
| AGGTCATTGCCGAGAAAAGGATAGCACTATAA<br>T             | 1    | SEQ ID NO: 159 |
| GGTCATTGCCGAGAAAAGGATAGCACTATAAT                  | 0.97 | SEQ ID NO: 160 |
| GTCATTGCCGAGAAAAGGATAGCACTATAAT                   | 0.98 | SEQ ID NO: 35  |
| TCATTGCCGAGAAAAGGATAGCACTATAAT                    | 0.99 | SEQ ID NO: 161 |
| CATTGCCGAGAAAAGGATAGCACTATAAT                     | 1    | SEQ ID NO: 162 |
| ATTGCCGAGAAAAGGATAGCACTATAAT                      | 1    | SEQ ID NO: 163 |
| TTGCCGAGAAAAGGATAGCACTATAAT                       | 0.98 | SEQ ID NO: 36  |
| TGCCGAGAAAAGGATAGCACTATAAT                        | 0.93 | SEQ ID NO: 164 |
| GCCGAGAAAAGGATAGCACTATAAT                         | 0.99 | SEQ ID NO: 165 |

|                          |      |                |
|--------------------------|------|----------------|
| CCGAGAAAAGGATAGCACTATAAT | 1    | SEQ ID NO: 166 |
| CGAGAAAAGGATAGCACTATAAT  | 1    | SEQ ID NO: 37  |
| GAGAAAAGGATAGCACTATAAT   | 1    | SEQ ID NO: 167 |
| AGAAAAGGATAGCACTATAAT    | 1.01 | SEQ ID NO: 168 |
| GAAAAGGATAGCACTATAAT     | 1    | SEQ ID NO: 169 |
| AAAAGGATAGCACTATAAT      | 1    | SEQ ID NO: 38  |
| AAAGGATAGCACTATAAT       | 1    | SEQ ID NO: 170 |
| AAGGATAGCACTATAAT        | 0.99 | SEQ ID NO: 171 |
| AGGATAGCACTATAAT         | 0.99 | SEQ ID NO: 172 |
| GGATAGCACTATAAT          | 0.97 | SEQ ID NO: 39  |

表 13 5'UTR I'截短体在 3'末端截短后与未截短的相对 MRL 值

| 5'UTR I'截短体                                  | MRL 相对值 | SEQ NO ID:     |
|----------------------------------------------|---------|----------------|
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCA           | 0.99    | SEQ ID NO: 173 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTA        | 0.98    | SEQ ID NO: 174 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCACT         | 1       | SEQ ID NO: 175 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT<br>A  | 0.99    | SEQ ID NO: 176 |
| TCAATAACAGGTCATTGCCGAGAAAAG                  | 0.97    | SEQ ID NO: 177 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATA              | 0.97    | SEQ ID NO: 178 |
| TCAATAACAGGTCATTGCCGAGAAAAGGAT               | 0.98    | SEQ ID NO: 179 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGC            | 0.99    | SEQ ID NO: 180 |
| TCAATAACAGGTCATTGCCGAGAAAA                   | 0.95    | SEQ ID NO: 181 |
| TCAATAACAGGTCATTGCCGAGAAAAGG                 | 0.95    | SEQ ID NO: 182 |
| TCAATAACAGGTCATTGCCGAGAAA                    | 0.96    | SEQ ID NO: 183 |
| TCAATAACAGGTCATT                             | 0.93    | SEQ ID NO: 184 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT<br>AA | 1.01    | SEQ ID NO: 185 |
| TCAATAACAGGTCATTGCCGAGAAAAGGA                | 0.97    | SEQ ID NO: 186 |
| TCAATAACAGGTCATTGCCGAG                       | 0.94    | SEQ ID NO: 187 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT       | 0.93    | SEQ ID NO: 188 |
| TCAATAACAGGTCATTGCCGAGAA                     | 0.96    | SEQ ID NO: 189 |
| TCAATAACAGGTCATTG                            | 0.9     | SEQ ID NO: 190 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAG             | 0.98    | SEQ ID NO: 191 |
| TCAATAACAGGTCATTGCCGAGA                      | 0.94    | SEQ ID NO: 192 |
| TCAATAACAGGTCAT                              | 0.91    | SEQ ID NO: 193 |

|                                     |      |                |
|-------------------------------------|------|----------------|
| TCAATAACAGGTCATTGCC                 | 0.92 | SEQ ID NO: 194 |
| TCAATAACAGGTCATTGCCGAGAAAAGGATAGCAC | 0.94 | SEQ ID NO: 195 |
| TCAATAACAGGTCATTGCCG                | 0.93 | SEQ ID NO: 196 |
| TCAATAACAGGTCATTGCCGA               | 0.92 | SEQ ID NO: 41  |
| TCAATAACAGGTCATTGC                  | 0.85 | SEQ ID NO: 76  |

6.2 构建 16 个不同长度的 5'UTR I'截短体(SEQ ID NO: 28-40, CTATAAT)和 2 个不同于 A→G 点突变的 A→C、A→T 点突变(SEQ ID NO: 42-43), 并由苏州金唯智公司合成, 然后将 16 种合成序列通过 HindIII和 ScaI 酶切、连接插入到实施例 5 的 Luc 24 中, 进而获得相应的荧光素酶 5'UTR I'截短体和点突变的构建体, 命名为 LUC1-16, 其对应的 5'UTR I'截短体(1-I'~14-I')和突变体(15-I'~16-I')序列如 SEQ ID NO: 63-78 所示。通过与实施例 1 相同的方法进行体外转录合成含有帽子结构的 mRNA, 体外转染 293T 细胞 24h 后, 加荧光素酶底物检测, 读板, 结果见图 9 所示, Luc24(I'-B 的组合)及其 5'UTR I'的截短体和突变体介导的荧光素酶的表达量均显著高于 Luc(BNT162b2 的 5'和 3' UTR)。5'UTR I'截短体可以保持 5'UTR I'调控目标蛋白表达的活性, 其中, 部分 5'UTR I'的截短体(1-I'、2-I'、3-I'、4-I'、5-I'、6-I'、7-I'、9-I'、10-I'、11-I'、13-I')具有提高的调控目标蛋白表达的活性。同时, 5'UTR I'的突变体 15-I'和 16-I'使得目的基因的表达水平优于 5'UTR I', 说明对 5'UTR I'的 5'端进行截断或对内部进行点突变可以进一步增强 5'UTR 调控目的基因表达的功能活性。

### 实施例 7、对照 5'UTR 与 5'UTR I'调控下的流感 HA mRNA 表达水平差异

通过酶切方式替换目的基因, 保留 5'UTR I'和 3'UTR B, 分别构建编码 A 型流感病毒 H1N1 (A/Wisconsin/588/2019) HA 蛋白的基因(SEQ ID NO:83), A 型流感病毒 H3N2 (A/Cambodia/e0826360/2020) HA 蛋白的基因(SEQ ID NO:84), B 型流感病毒 (Washington/02/2019) HA 蛋白的基因(SEQ ID NO:85)、和 B 型流感病毒 (Phuket/3073/2013) HA 蛋白的基因(SEQ ID NO:86)的质粒, 所述质粒含有的目的基因序列分别是 SEQ ID NO: 87-90, 均是密码子优化后的序列。参考专利 WO 2022/245888 A1 和 WO2022/150717A1 的 5'和 3'UTR 序列(SEQ ID NO:91-92)合成编码 SEQ ID NO: 83-86 的四种 HA CDS 序列, 分别是 SEQ ID NO: 93-96。采用实施例 1 的方式制备和纯化 mRNA, 通过 Lipofectamine™ 2000 转染 mRNA 1μg 至 293T 细胞, 1d 后收取细胞裂解液。采用实施例 2 中的蛋白印记的方法检测 I'-B 调控下的四种流感 HA 蛋白 mRNA 的表达情况, 其中, 一抗使用 A 型流感病毒 Hemagglutinin/HA 抗体 (Sino Biology, 86001-RM01)和 B 型流感病毒 Hemagglutinin/HA 抗体(Sino Biology, 11053-R004), 二抗使用 HRP-linked 的抗兔 IgG(Transgene, HS101)。结果如图 10 和 11 所示, 相比对照 5'UTR, 5'UTR-I'调

控下的四种流感 HA 的表达量明显更高。

本公开部分序列如下所示：

- >5'UTR I(63bp)  
5 GAGTTAGGAGTTGATGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT  
AAT  
SEQ ID NO: 1
- >5'UTR I'(63bp)  
10 GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT  
AAT  
SEQ ID NO: 2
- >5'UTR G  
GCACTTTTCTGAGCAGACGTCCAGAGCAGAGTCAGCCGCCACC  
SEQ ID NO: 3
- 15 >5'UTR A  
ACATTGCTTCTGACACAACGTGTGTTCACTAGCAACCTCAAACGCCGCCACC  
SEQ ID NO: 4
- >5'UTR J  
AGAGAGGCAAAGAAACATTGTGAAATCTCCAACCTCTTAACCTTCAAGCCGCCACC  
20  
SEQ ID NO: 5
- >5' UTR BNT162b2  
GAATAAACTAGTATTCTTCTGGTCCCCACAGACTCAGAGAGAACCCGCCACC  
SEQ ID NO: 6
- >3'UTR B  
25 GCTCGCTTTCTTGCTGTCCAATTTCTATTAAAGGTTCCCTTGTTCCTAAGTCCAACCTACTA  
AACTGGGGGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTAT  
TTTCATTGCAA  
SEQ ID NO: 7
- >3'UTR F  
30 CAGACAAGACAAGCTACTGAATACGTTACATCTGTCTTGATGCCTAATATTTTTACATTT  
CTGTAAACATATTTCTGAAATATTTTTGCCTTTCAAGCGACAGATGCCTCATTTTGTGAA  
AACTTAATGATGATTTTGTGTTTAAAGTTCCAAACATTTGAATAAAATAATTGACAATA  
SEQ ID NO: 8
- >3'UTR E  
35 CCTGAGGCCTCAGAACACAGGTTTCAGATTGATAGGCCTGCAGGTCTCCAGGCAGCAAC  
CAGCTGAGCGACTAAAGGGCCCAAGGCCAGGGCTCTAGGGATGGGGCTCAGCAGAGGC  
TGGGGTAAGGGGAGCCAGGGAGGAGCTGGGCCTAATGCAGCACCGGGTCCCCAGGATG  
CCGTGTCTCGGCTGGAGGAGGAGATGCGGAAGCTCCAGGCCACGGTGCAGGAGCTCCA  
GAAGCGCTTGGACAGGCTGGAGGAGACAGTCCAGGCCAAGTAGAGCCCCGAGGGCCT  
40 CCAGCAGGGTCAGCCATTCACACCCATCCACTCACCTCCCATTCCCAGCCACATGGCAGA  
GAAAAAAATCATAATAAAATGGCTTTATTTTCTGGTA  
SEQ ID NO: 9
- >3'UTR D  
45 GGGGCCCTTCTGACATGAGTCTGGCCTGGCCCCACCTCCTAGTTCCTCATAATAAAGACAG  
ATTGCTTCTTCGCTTCTCACTGAGGGGCCTTCTGACATGAGTCTGGCCTGGCCCCACCTC  
CCCAGTTTCTCATAATAAAGACAGATTGCTTCTTCACTTGAAGCTCGCTTCTTGCTGTCC  
AATTTCTATTAAAGGTTCCCTTGTTCCTAAGTCCAACCTACTAAACTGGGGGATATTATGA  
AGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTGCAA  
SEQ ID NO: 10
- 50 >3'UTR BNT162b2  
CTCGAGCTGGTACTGCATGCACGCAATGCTAGCTGCCCCCTTTCCCGTCCTGGGTACCCC  
GAGTCTCCCCCGACCTCGGGTCCCAGGTATGCTCCACCTCCACCTGCCCCACTCACCA  
CCTCTGCTAGTTCCAGACACCTCCCAAGCACGCAGCAATGCAGCTCAAACGCTTAGCC  
TAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAA

CTAAGCTATACTAACCCCAGGGTTGGTCAATTTCTGTGCCAGCCACACCCTGGA

SEQ ID NO: 11

&gt;B.1.617.2(Delta)刺突蛋白氨基酸序列

MFVFLVLLPLVSSQCVNLRTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT  
5 WFHAIHVSNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSTQSLNINATNV  
VIKVFCEQFCNDPFLDVYYHKNNKSWMESGVYSSANNCTFEYVSQPFMDLEGKQGNFKN  
LREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTTPGD  
SSSGWTAGAAAYVGYLQPRIFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQT  
10 SNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKC  
YGVSPTKLNDLCFTNVYADSFVIRGDEVQIAPGQTGKIADYNYKLPPDFTGCVIWNSSN  
LDSKVGGNYNYRYRLFRKSNLKPFERDISTEIIYQAGSKPCNGVEGFNCYFPLQSYGFQPTNG  
VGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQ  
QFGRDIADTTDAVRDPQTLEILDITPCSFSGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHAD  
15 QLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSECDIPIGAGICASYQTQNSRRRARSVAS  
QSIIAYTMSLGAENSVYSSNIAIPTNFTISVTTEILPVSMKTSDVCTMYICGDSTECNLLL  
QYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFI  
EDLLFNKVTADAGFIKQYGDCLGDIAARDLCAQKFNGLTVLPPLTDEMIAQYTSALLAG  
TITSGWTFGAGAAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTAS  
20 ALGKLQNVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDLPPEAEVQIDRLITGRLQSLQT  
YVATQQLIRAAEIRASANLAATKMSECVLQSKRVDFCGKGYHLSMFPQSAPHGVVFLHVTY  
VPAQEKNFHTTAPAICHGDKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVV  
IGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKN  
LNEIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCCLKGCCSCGSCC  
KFDEDDSEPVKGVKLHYT

SEQ ID NO: 12

&gt;B.1.617.2(Delta)刺突蛋白 CDS

ATGTTCTGTTCCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGAGGACA  
AGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGGCGTGACTACCCCGA  
CAAGGTGTTTCAAGAGCAGCGTGCTGCACTCCACCCAAGACCTGTTTCTGCCCTTCTTCA  
30 GCAACGTGACCTGGTTCACGCCATCCAGCTCAGCGGCACCAACGGCACCAAGAGATT  
GACAACCCCGTGCTGCCCTTCAACGACGCGCTGTACTTCGCTAGCACCGAAAAAGAGCAA  
CATCATCAGAGGCTGGATCTTCGGCACCACTTAGACTCCAAGACACAAAGCCTGCTGAT  
CGTTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACC  
CCTTCTGGACGTGTACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGGGGTCTAC  
35 AGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGGACCTGGA  
GGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTCGTGTTCAAGAACATCGACGGCT  
ACTTCAAGATCTACAGCAAGCACACCCCATCAACCTGGTGAGAGACCTGCCCCAAGGC  
TTCAGCGCCCTGGAGCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAAGATTTACG  
ACCTGTCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCTCTGTGCTGGACG  
40 GCCGGCGCTGCCGCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTAC  
AACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTGAGCGAGAC  
CAAGTGACCCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTATCAGACAAGCAACTTCA  
GAGTGACGCCCCACGAGAGCATCGTGAGATTCCCCAACATCACCAACCTGTGCCCTTC  
GGCGAGGTGTTCAACGCCACAAGATTGCTAGCGGTGTACGCGTGGAACAGAAAGAGAAT  
45 CAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCACCTTCA  
AGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCC  
GACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCGGGCAAAACCGGCA  
AGATCGCCGACTACAACCTACAAGCTGCCCGACGACTTCACCGGCTGCGTGATCGCGTGG  
AACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACAACCTACCGGTACAGACTGTT  
50 CAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTACCAAGCC  
GGCAGCAAGCCCTGCAACGGCGTGGAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTA  
CGGCTTTCAGCCACCAACGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCT  
TCGAGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCAAAAAGAGCACCAACCTGGT  
GAAGAACAAGTGCGTGAACCTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTGACCG  
55 AGAGCAACAAGAAGTTCTGCCATTTAGCAATTCGGCAGAGACATCGCCGACACCACC  
GACGCGGTGAGAGACCCTCAGACCCTGGAGATCCTGGACATCACCCCTGTAGCTTCGG  
CGCGTGAGCGTGATCACCCCGGCACTAACACAAGCAACCAAGTGCCGTGCTGTACC

AAGGCGTGAAGTGCACCGAGGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCCCACC  
TGGAGAGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCGG  
CGCCGAGCACGTGAACAACAGCTACGAGTGCAGATCCCCATCGGCGCCGGCATCTGCG  
CTAGCTATCAGACACAGACCAACAGCCGCCGGAGAGCTAGAAGCGTGCTTCTCAGAGC  
5 ATCATCGCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGCCCTACAGCAACAACAG  
CATCGCCATCCCCACCAACTTCACCATCAGCGTGACCACCGAGATCCTGCCCCGTAAGTAT  
GACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCGAGTGCAGC  
AACCTGCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGCAT  
CGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTGCCCCAAGTGAAGCAGATCTAC  
10 AAGACCCCCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCAGCCCT  
AGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGC  
CGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAGACC  
TGATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGAGA  
TGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTTG  
15 GCGCCGGGGGCCGACTGCAAATCCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCA  
TCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAAATCAGTTCAAC  
AGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAGCT  
GCAAAACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTGAGC  
AGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGTCTAGACTGGACCCCCC  
20 GAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCAGACCT  
ACGTGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGGCCGCC  
ACCAAGATGAGCGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGG  
GCTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTG  
ACCTACGTGCCCGCCCAAGAGAAGAATTCACACCGCTCCCGCCATCTGCCACGACGG  
25 CAAGGCCCACTTCCCTAGAGAGGGCGTGTTCTGAGCAACGGCACCCACTGGTTCGTGA  
CACAGAGAACTTCTACGAGCCTCAGATCATCACACCGACAACACCTTCGTGAGCGGC  
AACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTATACGACCCCCCTGCAGCCGA  
GCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCCCGACG  
TGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTGAGAAGGAGATC  
30 GACCGGCTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGC  
TGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCG  
GCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCAGCT  
GCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGACAGCGA  
GCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 13

>B.1.1.529(Omicron) 刺突蛋白氨基酸序列  
MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT  
WFHVISGTNGTKRFDNPVLPFNDGVYFASIEKSNIIRGWIFGTTLDSKTQSLIVNNATNVVIK  
VCEFQFCNDPFLDHKNKSWMESEFRVYSSANNCTFEYVSQPFMDLEGKQGNFKNLREF  
40 VFKNIDGYFKIYSKHTPIIVREPEDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLTGPDSSS  
GWTAGAAAYYVGYLQPRTFLLKYENGTITDAVDCALDPLSETKCTLSFTVEKGIYQTSNF  
RVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNKRISNCVADYSVLYNLAPFFTFKCYG  
VSPTKLNLCFTNVYADSFVIRGDEVRIAPGQTGNIADYNYKLDDFTGCVIWNNSNKLDS  
KVSNGYNYLYRLFRKSNLKPFERDISTEYQAGNKPCNGVAGFNCYFPLRSYSFRPTYGVGH  
45 QPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLKGTGVLTESNKKFLPFQQFGR  
DIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTP  
TWRVYSTGSNVFQTRAGCLIGAEYVNNSECDIPIGAGICASYQTQTKSHRRARSVASQSIIA  
YTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQYGS  
FCTQLKRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKYFGGFNFSQILPDPSKPSKRSFIEDLL  
50 FNKVTLADAGFIKQYGDCLGDIAARDLICAQKFKGLTVLPLLTDEMIAQYTSALLAGTITSG  
WTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDLSSTASALGK  
LQDVVNHNAQALNTLVKQLSSKFGAISSVLNDFSRDPPEAEVQIDRLITGRLQSLQTYVTQ  
QLIRAAEIRASANLAATKMSECVLGQSKRVDGCGKGYHLMSPQSAHPGVVFLHVTYVPAQ  
EKNFTTAPAICHDGKAHFPREGVVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVN  
55 NTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESL  
IDLQELGKYEYQIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDED  
DSEPVLLKGVKLHYT

SEQ ID NO: 14

>B.1.1.529(Omicron) 刺突蛋白 CDS

ATGTTCTGTTCCCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGACCACA  
AGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGGCGTGTACTACCCCGA  
CAAGGTGTTTCTGAGAGCAGCGTGTGCTGCACTCCACCCAAGACCTGTTTCTGCCCTTCTTCA  
5 GCAACGTGACCTGGTTCACGTGATCAGCGGCACCAACGGCACCAGAGATTTCGACAAC  
CCCGTGCTGCCCTTCAACGACGGCGTGTACTTCGCTAGCATCGAAAAGAGCAACATCATC  
AGAGGCTGGATCTTCGGCACCACCCTAGACTCCAAGACACAAAAGCCTGCTGATCGTTAA  
CAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACCCCTTCC  
TGGACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTTCAGAGTGTACAGCAGCGC  
10 CAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGGACCTGGAGGGCAAGC  
AAGGCAATTTCAAGAACCTAAGAGAGTTTCGTGTTCAAGAACATCGACGGCTACTTCAAG  
ATCTACAGCAAGCACACCCCATCATCGTGAGAGAGCCCGAGGACCTGCCCCAAGGCTT  
CAGCGCCCTGGAGCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAGATTTTCA  
CCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCTCTGGCTGGACGG  
15 CCGGCGCTGCCGCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTACA  
ACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTGAGCGAGACC  
AAGTGCACCCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTATCAGACAAGCAACTTCAG  
AGTGCAGCCCACCGAGAGCATCGTGAGATTCCCCAACATCACCACCTGTGCCCTTCCG  
ACGAGGTGTTCAACGCCACCAGATTTGCTAGCGTGTACGCTGGAACCGGAAGAGAATC  
20 AGCAACTGCGTGCCGACTACAGCGTGTGTACAACCTGGCCCCCTTCTTACCTTCAA  
GTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCG  
ACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATCGCCCCCGGGCAGACCGGCAA  
CATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACCGGCTGCGTGATCGCCTGGA  
ACAGCAACAAGCTGGATAGCAAGGTGAGCGGCAACTACAATTACCTGTACAGACTGTTT  
25 AGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCAGAGATCTACCAAGCCG  
GCAACAAGCCCTGCAACGGCGTGGCGGGCTTCAACTGCTACTTCCCCCTCAGGACTAC  
AGCTTCAGACCCACCTACGGCGTGGGCCATCAGCCCTACAGAGTGGTGTGCTGAGCTT  
CGAGCTGCTGCACGCCCCTGCCACAGTGTGCGGCCCAAAAAGAGCACCAACCTGGTG  
AAGAACAAGTGCGTGAACCTTCAACTTCAACGGCCTGAAGGGCACCGGCGTGTGACCG  
30 AGAGCAACAAGAAGTTTCTGCCATTTTCAAGCAATTCGGCAGAGACATCGCCGACACCACC  
GACGCGGTGAGAGACCCTCAGACCCTGGAGATCCTGGACATCACCCTCTGAGCTTCGG  
CGGCGTGAGCGTGATCACCCTCGGCACTAACACAAGCAACCAAGTGCCGCTGTGTACC  
AAGGCGTGAACCTGCACCGAGGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCCAAC  
TGGAGAGTGTATAGCACCAGGCAACGTGTTTCAGACAAGAGCCGGCTGCTGATCGG  
35 CGCCGAGTACGTGAACAACAGCTACGAGTGCAGATCGGACATCCCCATCGGCGCCGGCATCTGCG  
CTAGCTATCAGACACAGACCAAGAGCCACCGGAGAGCTAGAAGCGTGGCTAGCCAAAG  
CATCATCGCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACAGCAACAACA  
GCATCGCCATCCCCACCAACTTCAACATCAGCGTGACCACCGAGATCCTGCCCCGTAAGTA  
TGACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCAGGTGCAGC  
40 AACCTGCTCCTGCAGTACGGCAGCTTCTGCACACAGCTGAAGAGAGCCCTGACCGGCAT  
CGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTAC  
AAGACCCCCCATCAAGTACTTCGGCGGCTTCAACTTACGCCAAATCCTGCCCCGACCT  
AGCAAGCCTAGCAAGCGGAGCTTATCAGGACCTGCTGTTCAACAAGGTGACCCTGGC  
CGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAGACC  
45 TGATCTGCGCTCAGAAGTTCAAGGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGAG  
ATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTT  
GGCGCCGGGGCCGCACTGCAAATCCCCCTCGCCATGCAGATGGCCTACAGATTCAACGG  
CATCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCA  
ACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCCTAGCGCCCTGGGCAA  
50 GCTGCAAGACGTGGTGAACCACAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTG  
AGCAGCAAGTTTCGGCGCCATCAGCAGCGTGTGATGACATCTTCTTAGACTGGACCC  
CCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCAG  
ACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGGC  
CGCCACCAAGATGAGCGAGTGCCTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGC  
55 AAGGGCTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTCTCTGCA  
CGTGACCTACGTGCCCCGCCAAGAGAAGAACTTCAACACCGCTCCCGCCATCTGCCACG  
ACGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTC  
GTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCACCGACAACACCTTCGTGAG

CGGCAACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTATACGACCCCCCTGCAGC  
CCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCCC  
GACGTGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTGAGAAGGA  
GATCGACCGGCTGAATGAGGTGGCCAAGAACCCTGAACGAGAGCCTGATCGACCTGCAA  
5 GAGCTGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATC  
GCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGC  
AGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGACA  
CGGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 15

10 >poly A  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
SEQ ID NO: 16

>T7 启动子  
15 TAATACGACTCACTATAAGG

SEQ ID NO: 17

>I'-Delta24-B  
GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT  
AATATGTTTCGTGTTCTTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGAGG  
20 ACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGGCGTGTACTACCC  
CGACAAGGTGTTTCAAGAGCAGCGTGCTGCACTCCACCCAAGACCTGTTTCTGCCCTTCT  
TCAGCAACGTGACCTGGTTCCACGCCATCCACGTCAGCGGCACCAACGGCACCAAGAGA  
TTCGACAACCCCGTGCTGCCCTTCAACGACGCGCTGTACTTCGCTAGCACCGAAAAGAG  
CAACATCATCAGAGGCTGGATCTTCGGCACCACCCTAGACTCCAAGACACAAAGCCTGC  
25 TGATCGTTAAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAAC  
GACCCCTTCTGGACGTGTACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGGGG  
TCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGGACC  
TGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTTCGTGTTCAAGAACATCGAC  
GGCTACTTCAAGATCTACAGCAAGCACACCCCATCAACCTGGTGAGAGACCTGCCCA  
30 AGGCTTCAGCGCCCTGGAGCCCTGGTGACCTGCCATCGGCATCAACATCACAAGAT  
TTCAGACCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCCTCTGGCT  
GGACGGCCGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCTGCTG  
AAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTGAG  
CGAGACCAAGTGACCCCTCAAGAGCTTCACCGTGAGAAAGGGCATCTATCAGACAAGCA  
35 ACTTCAGAGTGCAGCCACCGAGAGCATCGTGAGATTCCCAACATCACCAACCTGTGC  
CCCTTCGGCGAGGTGTTCAACGCCACAAGATTTCGCTAGCGTGTACGCGTGGAACAGAAA  
GAGAATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCA  
CCTTCAAGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGT  
ACGCCGACAGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCGGGCAAAAC  
40 CGGCAAGATCGCCGACTACAATAAGCTGCCGACGACTTCACCGCTGCGTGATCG  
CGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACAATAACCGGTACAGA  
CTGTTTCAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTACCA  
AGCCGGCAGCAAGCCCTGCAACGGCGTGAGGGCTTCAACTGCTACTTCCCCCTGCAGA  
GCTACGGCTTTCAGCCACCAACGGCGTGCGGCTATCAGCCCTACAGAGTGGTGGTGCTG  
45 AGCTTCGAGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCAAAAAGAGCACCAACC  
TGGTGAAGAACAAGTGCGTGAACCTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTG  
ACCGAGAGCAACAAGAAGTTCTGCCATTTACGCAATTTCGGCAGAGACATCGCCGACAC  
CACCGACGCGGTGAGAGACCCTCAGACCTGGAGATCCTGGACATCACCCCTGTAGCT  
TCGGCGGCGTGAGCGTGATCACCCCGGCACTAACACAAGCAACCAAGTGGCCGTGCTG  
50 TACCAAGGCGTGAACCTGCACCGAGGTGCCCGTGCCATCCACGCCGATCAGCTGACCCC  
CACCTGGAGAGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGA  
TCGGCGCCGAGCACGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATC  
TGCGCTAGCTATCAGACACAGACCAACAGCCGCCGAGAGCTAGAAGCGTGGCTTCTCA  
GAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACAGCAACA  
55 ACAGCATCGCCATCCCCACCAACTTACCATCAGCGTGACCAACGAGATCCTGCCCGTAA  
GTATGACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGACCGAGTG  
AGCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGG

CATCGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTGCCCCAAGTGAAGCAGATC  
TACAAGACCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGA  
CCCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCC  
TGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGA  
5 GACCTGATCTGCGCTCAGAAAGTTCAATGGACTCACCGTGCTCCCCCCCCCTGCTGACCGAC  
GAGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGAC  
ATTTGGCGCGGGGGCCGCACTGCAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAA  
CGGCATCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAAATCAGT  
TCAACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGC  
10 AAGCTGCAAAACGTGGTGAATCAGAACGCCAAGCCCTGAACACCCTGGTGAAGCAGC  
TGAGCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGTCTAGACTGGACC  
CCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCA  
GACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGG  
CCGCCACCAAGATGAGCGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGAATTCTGCGG  
15 CAAGGGCTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGC  
ACGTGACCTACGTGCCCCGCCAAGAGAAGAACTTACCACCGCTCCCGCCACTGTCCAC  
GACGGAAGGCCCACTTCCCTAGAGAGGGCGTGTTCGTGAGCAACGGCACCCCACTGGTT  
CGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCAACGACAACACCTTCGTGA  
GCGGCAACTGCGACGTGGTGTATCGGCATAGTGAACAACACCGTATACGACCCCTGCAG  
20 CCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCC  
CGACGTGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTGAGAAGG  
AGATCGACCGGTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCA  
AGAGCTGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCA  
TCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTT  
25 GCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGAC  
AGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGACTCGAGGCTCGCTT  
TCTTGCTGTCCAATTTCTATTAAAGGTTCCCTTTGTTCCCTAAGTCCAATACTAACTGGG  
GGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTG  
CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAA  
30 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AA

SEQ ID NO: 18

>I-OC24-B  
GAGTTAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTAT  
35 AATATGTTTCGTGTTCTTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGACC  
ACAAGAACACAGCTGCCCCCGCCTACACCAACAGCTTCAAGAGGCGGTGACTACCC  
CGACAAGGTGTTTCAAGAGCAGCGTGCTGCACCTCCACCCAAGACCTGTTTCTGCCCTTCT  
TCAGCAACGTGACCTGGTTCCACGTGATCAGCGGCACCAACGGCACCAAGAGATTCGAC  
AACCCCGTGCTGCCCTTCAACGACGGCGTGTACTTCGCTAGCATCGAAAAGAGCAACAT  
40 CATCAGAGGCTGGATCTTCGGCACCACCCTAGACTCCAAGACACAAAGCCTGCTGATCG  
TTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACCCC  
TTCTTGACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGAGTGTACAGCA  
GCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCTGATGGACCTGGAGGGC  
AAGCAAGGCAATTTCAAGAACCTAAGAGAGTTTCGTGTTCAAGAACATCGACGGCTACTT  
45 CAAGATCTACAGCAAGCACACCCCATCATCGTGAGAGAGCCCGAGGACCTGCCCAAG  
GCTTCAGCGCCCTGGAGCCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAGATTC  
AGACCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCTCTGGCTGGA  
CGGCCGGCGCTGCCGCCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCTGCTGAAGT  
ACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTGAGCGAG  
50 ACCAAGTGACCCCTCAAGAGCTTACCGTGAGAGAGGGCATCTATCAGACAAGCAACTT  
CAGAGTGCAGCCACCGAGAGCATCGTGAGATTCCCCAACATCACCAACCTGTGCCCT  
TCGACGAGGTGTTCAACGCCACCAGATTTGCTAGCGTGTACGCCTGGAACCGGAAGAGA  
ATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAACCTGGCCCCCTTCTTACCTTC  
AAGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTACCAACAGCTGTACCG  
55 CGACAGCTTCTGTGATCAGAGGCGACGAGGTGAGACAGATCGCCCCCGGCGAGACCGGC  
AACATCGCCGACTACAACCTACAAGCTGCCCGACGACTTACCGGCTGCGTGATCGCCTG  
GAACAGCAACAAGCTGGATAGCAAGGTGAGCGGCAACTACAATTACCTGTACAGACTGT  
TCAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTACCAAGC

CGGCAACAAGCCCTGCAACGGCGTGGCGGGCTTCAACTGCTACTTCCCCCTCAGGAGCT  
ACAGCTTCAGACCCACCTACGGCGTGGGCCATCAGCCCTACAGAGTGGTCGTGCTGAGC  
TTCGAGCTGCTGCACGCCCCCTGCCACAGTGTGCGGCCCAAAAAGAGCACCAACCTGGT  
GAAGAACAAGTGCGTGAACCTCAACTTCAACGGCCTGAAGGGCACCGGCGTGCTGACC  
5 GAGAGCAACAAGAAGTTCCTGCCATTTAGCAATTCGGCAGAGACATCGCCGACACCAC  
CGACGCGGTGAGAGACCCCTCAGACCCTGGAGATCCTGGACATCACCCCTGTAGCTTCG  
GCGGCGTGAGCGTGATCACCCCGGCACTAACACAAGCAACCAAGTGGCCGTGCTGTAC  
CAAGGCGTGAACCTGCACCGAGGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCAC  
CTGGAGAGTGTATAGCACCGGCAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGATCG  
10 GCGCCGAGTACGTGAACAACAGCTACGAGTGGGACATCCCCATCGGCGCCGGCATCTGC  
GCTAGCTATCAGACACAGACCAAGAGCCACCGGAGAGCTAGAAGCGTGGCTAGCCAAA  
GCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACAGCAACAAC  
AGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCACCGAGATCCTGCCCCGTAAGT  
ATGACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCGAGTGCAG  
15 CAACCTGCTCCTGCAGTACGGCAGCTTCTGCACACAGCTGAAGAGAGCCCTGACCGGCA  
TCGCCGTGGAGCAAGACAAGAACACCCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTA  
CAAGACCCCCCATCAAGTACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGACCC  
TAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCCTGG  
CCGACGCGCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAGAC  
20 CTGATCTGCGCTCAGAAGTTCAGGGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGA  
GATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATT  
TGGCGCCGGGGCCGCACTGCAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAACG  
GCATCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTC  
AACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCA  
25 AGCTGCAAGACGTGGTGAACCACAACGCCCAAGCCCTGAACACCCCTGGTGAAGCAGCT  
GAGCAGCAAGTTCGGCGCCATCAGCAGCGTGCTGAATGACATCTTCTCTAGACTGGACC  
CCCCCGAGGCGGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCA  
GACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACCTGG  
CCGCCACCAAGATGAGCGAGTGCCTGCTGGGCGAGAGCAAGAGAGTGGACTTCTGCGG  
30 CAAGGGCTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCAGCGCGTGGTGTCTCTGC  
ACGTGACCTACGTGCCCCGCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCAC  
GACGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTCTGTGAGCAACGGCACCCACTGGTT  
CGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACACCAGACAACACCTTCGTGA  
GCGGCAACTGCGACGTGGTGTATCGGCATAGTGAACAACACCGTATACGACCCCTGCAG  
35 CCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCC  
CGACGTGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTGAGAAGG  
AGATCGACCGGTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCA  
AGAGCTGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCA  
TCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTT  
40 GCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGAC  
AGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGACTCGAGGCTCGCTT  
TCTTGCTGTCCAATTTCTATTAAAGGTTCTTCTTCCCTAAGTCCAACCTACTAAACTGGG  
GGATATTATGAAGGGCCTTGAGCATCTGGATTCTGCCTAATAAAAAACATTTATTTTCATTG  
CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCATATGACTAAAAAAAAAAAAAAAA  
45 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AA

SEQ ID NO: 19

&gt;B.1.351 刺突蛋白 CDS

ATGTTCTGTTCCTGGTGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTTCACGACA  
50 AGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGGCGTGTAACCCCGA  
CAAGGTGTTTCAAGAGCAGCGTGCTGCACTCCACCCAAGACCTGTTTCTGCCCTTCTTCA  
GCAACGTGACCTGGTTCCACGCCATCCACGTGAGCGGCACCAACGGCACCAAGAGATTTC  
GCCAACCCCGTGCTGCCCTTCAACGACGGCGTGACTTCGCTAGCACCCGAAAAGAGCAA  
CATCATCAGAGGCTGGATCTTCGGCACCAACCTAGACTCCAAGACACAAAGCCTGCTGAT  
55 CGTTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACC  
CCTTCTGGGCGTGACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCCGT  
GTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTAGCCCTTCTGATGGAC  
CTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTCTGTTCAGAACATCGA

CGGCTACTTCAAGATCTACAGCAAGCACACCCCCATCAACCTGGTGAGAGGCCTGCCCC  
AAGGCTTCAGCGCCCTGGAGCCCCTGGTGACCTGCCCATCGGCATCAACATCACAAGA  
TTTCAGACCCTGCACCGTAGCTACCTAACCCCCGGCGATAGTTCTCTGGCTGGACGGCC  
GGCGCTGCCGCTACTACGTGGGCTACCTGCAGCCTAGAACCTTCCTGCTGAAGTACAAC  
5 GAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCCTGAGCGAGACCAA  
GTGCACCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTATCAGACAAGCAACTTCAGAG  
TGCAGCCCACCGAGAGCATCGTGAGATTCCCCAACATCACCAACCTGTGCCCTTCGGC  
GAGGTGTTCAACGCCACAAGATTGCTAGCGTGTACGCGTGGAACAGAAAGAGAATCAG  
CAACTGCGTGGCCGACTACAGCGTGTGTACAACAGCGCTAGCTTCAGCACCTTCAAGT  
10 GCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGTACGCCGAC  
AGCTTCGTGATCAGAGGCGACGAGGTGAGACAGATTGCCCCGGGCAAACCGGCAACA  
TCGCCGACTACAACCTACAAGCTGCCCGACGACTTCACCGGTGCGTGATCGCGTGGAAC  
AGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACAACCTACCTGTACAGACTGTTTCA  
AAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTACCAAGCCGGC  
15 AGCACCCCCTGCAACGGCGTGAAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTACGG  
CTTTCAGCCCACCTACGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGAGCTTCGA  
GCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCCAAAAGAGCACCAACCTGGTGAAG  
AACAAGTGCCTGAACCTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTGACCGAGAG  
CAACAAGAAGTTCTGCCATTTAGCAATTTCGGCAGAGACATCGCCGACACCACCGACG  
20 CGGTCAGAGACCCCTCAGACCCTGGAGATCCTGGACATCACCCCCTGTAGCTTCGGCGGC  
GTGAGCGTGATCACCCCCGGCACTAACACAAGCAACCAAGTGGCCGTGCTGTACCAAGG  
CGTGAACCTGCACCGAGGTGCCCCGTGGCCATCCACGCCGATCAGCTGACCCCCACCTGGA  
GAGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCTGATCGGCGCC  
GAGCACGTGAACAACAGCTACGAGTGCAGATCCCCATCGGCGCCGGCATCTGCGCTAG  
25 CTATCAGACACAGACCAACAGCCCCGGAGAGCTAGAAGCGTGGCTTCTCAGAGCATCA  
TCGCCTACACCATGAGCCTGGGCGTCGAGAACAGCGTGGCCTACAGCAACAACAGCATC  
GCCATCCCCACCAACTTCACCATCAGCGTGACCACCGAGATCCTGCCCGTAAGTATGACC  
AAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCGAGTGCAGCAACCT  
GCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGCATCGCCG  
30 TGGAGCAAGACAAGAACACCCAAGAGGTGTTTCGCCCAAGTGAAGCAGATCTACAAGAC  
CCCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCGACCCTAGCAA  
GCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCCGACG  
CCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAGACCTGATCT  
GCGCTCAGAAGTTCAATGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGAGATGATC  
35 GCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTTGGCGC  
CGGGGCGCCTGCAAATCCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCATCGG  
CGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACAGCG  
CCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCAAGCTGCAA  
GACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTGAGCAGCA  
40 ACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGTCTAGACTGGACCCCCCGAGG  
CCGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCAGACCTACGTG  
ACACAGCAGCTGATCAGAGCCCGGAGATCAGAGCTAGCGCCAACCTGGCCGCCACCA  
AGATGAGCGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTA  
CCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGACCTA  
45 CGTCCCCGCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGACGGCAAGG  
CCCACTTCCCTAGAGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTTCGTGACACAG  
AGAAACTTCTACGAGCCTCAGATCATCACCACCGACAACACCTTCGTGAGCGGCAACTG  
CGACGTGGTGATCGGCATAGTGAACAACACCGTATACGACCCCCCTGCAGCCCAGCTGG  
ACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGAC  
50 CTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTCAGAAGGAGATCGACCG  
GCTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGGC  
AAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTG  
ATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCAGCTGCCTG  
AAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGACAGCGAGCCCCG  
55 TGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 20

&gt;B.1.351 刺突蛋白氨基酸序列

MFVFLVLLPLVSSQCVNFTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT

WFHAIHVS G T N G T K R F A N P V L P F N D G V Y F A S T E K S N I I R G W I F G T T L D S K T Q S L L I V N N A T N V  
V I K V C E F Q C N D P F L G V Y Y H K N N K S W M E S E F R V Y S S A N N C T F E Y V S Q P F L M D L E G K Q G N F  
K N L R E F V F K N I D G Y F K I Y S K H T P I N L V R G L P Q G F S A L E P L V D L P I G I N I T R F Q T L H R S Y L T P G D S  
S S G W T A G A A Y Y V G Y L Q P R T F L L K Y N E N G T I T D A V D C A L D P L S E T K C T L K S F T V E K G I Y Q T S  
5 N F R V Q P T E S I V R F P N I T N L C P F G E V F N A T R F A S V Y A W N R K R I S N C V A D Y S V L Y N S A S F S T F K C Y  
G V S P T K L N D L C F T N V Y A D S F V I R G D E V R Q I A P G Q T G N I A D Y N Y K L P D D F T G C V I A W N S N N L  
D S K V G G N Y N Y L R L F R K S N L K P F E R D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T Y G V  
G Y Q P Y R V V L S F E L L H A P A T V C G P K K S T N L V K N K C V N F N F N G L T G T G V L T E S N K K F L P F Q Q F  
G R D I A D T T D A V R D P Q T L E I L D I T P C S F G G V S V I T P G T N T S N Q V A V L Y Q G V N C T E V P V A I H A D Q L  
10 T P T W R V Y S T G S N V F Q T R A G C L I G A E H V N N S Y E C D I P I G A G I C A S Y Q T Q T N S P R R A R S V A S Q S I I  
A Y T M S L G V E N S V A Y S N N S I A I P T N F T I S V T T E I L P V S M T K T S V D C T M Y I C G D S T E C S N L L L Q Y G  
S F C T Q L N R A L T G I A V E Q D K N T Q E V F A Q V K Q I Y K T P I K D F G G F N F S Q I L P D P S K P S K R S F I E D L  
L F N K V T L A D A G F I K Q Y G D C L G D I A A R D L I C A Q K F N G L T V L P P L T D E M I A Q Y T S A L L A G T I T S  
G W T F G A G A A L Q I P F A M Q M A Y R F N G I G V T Q N V L Y E N Q K L I A N Q F N S A I G K I Q D S L S S T A S A L G  
15 K L Q D V V N Q N A Q A L N T L V K Q L S S N F G A I S S V L N D I L S R L D P P E A E V Q I D R L I T G R L Q S L Q T Y V T  
Q Q L I R A A E I R A S A N L A A T K M S E C V L G Q S K R V D F C G K G Y H L M S F P Q S A P H G V V F L H V T Y V P A  
Q E K N F T T A P A I C H D G K A H F P R E G V F V S N G T H W F V T Q R N F Y E P Q I I T D N T F V S G N C D V V I G I V  
N N T V Y D P L Q P E L D S F K E E L D K Y F K N H T S P D V D L G D I S G I N A S V V N I Q K E I D R L N E V A K N L N E  
S L I D L Q E L G K Y E Q Y I K W P W Y I W L G F I A G L I A I V M V T I M L C C M T S C C S C L K G C C S C G S C C K F D  
20 E D D S E P V L K G V K L H Y T

SEQ ID NO: 21

## &gt;B.1.1.7 刺突蛋白 CDS

A T G T T C G T G T T C C T G G T G C T G C T G C C C T G G T G A G C T C T C A G T G C G T G A A C C T G A C C A C A  
A G A A C A C A G C T G C C C C C G C C T A C A C C A A C A G C T T C A C A A G A G G C G T G T A C T A C C C C G A  
25 C A A G G T G T T C A G A A G C A G C G T G C T G C A C T C C A C C C A A G A C C T G T T T C T G C C T T C T T C A  
G C A A C G T G A C T G G T T C C A C G C C A T C A G C G G C A C C A A C G G C A C C A A G A G A T T C G A C A A C  
C C C G T G C T G C C C T T C A A C G A C G G C G T G T A C T T C G C T A G C A C C G A A A A G A G A C A A C A T C A T  
C A G A G G C T G G A T C T T C G G C A C C A C C C T A G A C T C C A A G A C A C A A A G C C T G C T G A T C G T T A  
A C A A C G C G A C C A A C G T G G T G A T C A A G G T G T G C G A G T T T C A G T T C T G C A A C G A C C C C T T C  
30 C T G G G C G T G T A C C A C A A G A A C A A C A A G A G C T G G A T G G A G A G C G A G T T C C G T G T C T A C A  
G C A G C G C C A A C A A C T G C A C C T T C G A G T A C G T G T C T C A G C C C T T C C T G A T G G A C C T G G A G  
G G C A A G C A A G G C A A T T T C A A G A A C C T A A G A G A G T T C G T G T T C A A G A A C A T C G A C G G C T A  
C T T C A A G A T C T A C A G C A A G C A C A C C C C C A T C A A C C T G G T G A G A G A C C T G C C C A A G G C T  
T C A G C G C C C T G G A G C C C C T G G T G G A C C T G C C C A T C G G C A T C A A C A T C A C A A G A T T T C A G  
35 A C C T G C T G G C C C T G C A C C G T A G C T A C C T A A C C C C G G C G A T A G T T C C T T G C T G G A G C G  
A C C G G C G T G C C G C C T A C T A C G T G G G C T A C C T G C A G C C T A G A A C C T T C C T G C T G A A G T A C  
A A C G A G A A C G G C A C T A T C A C T G A T G C C G T G G A C T G T G C C C T G G A C C C C C T G A G C G A G A C  
C A A G T G C A C C C T C A A G A G C T T C A C C G T G G A G A A G G G C A T C T A T C A G A C A A G C A A C T T C A  
G A G T G C A G C C C A C C G A G A G C A T C G T G A G A T T C C C C A A C A T C A C C A A C C T G T G C C C C T T C  
40 G G C G A G G T G T T C A A C G C C A C A A G A T T C G T A G C G T G T A C G C G T G G A A C A G A A A G A G A A T  
C A G C A A C T G C G T G G C C G A C T A C A G C G T G C T G T A C A A C A G C G C T A G C T T C A G C A C C T T C A  
A G T G C T A T G G C G T C A G C C C C A C C A A G C T G A A C G A C C T G T G C T T C A C C A A C G T G T A C G C C  
G A C A G C T T C G T G A T C A G A G G C G A C G A G T G A G A C A G A T T G C C C C G G G C A A A C C G G C A  
A G A T C G C C G A C T A C A A C T A C A A G C T G C C C G A C G A C T T C A C C G G C T G C G T G A T C G C G T G G  
45 A A C A G C A A T A A C C T G G A T T C C A A G G T G G G C G G C A A C T A C A A C T A C C T G T A C A G A C T G T T  
C A G A A A G A G C A A C C T G A A G C C C T T C G A G A G A G A C A T C A G C A C C G A G A T C T A C C A A G C C  
G G C A G C A C C C C T G C A A C G G C G T G G A G G G C T T C A A C T G C T A C T T C C C C C T G C A G A G C T A  
C G G C T T T C A G C C C A C C T A C G G C G T G G G C T A T C A G C C C T A C A G A G T G G T G G T G C T G A G C T T  
C G A G C T G C T G C A C G C C C C G C A A C C G T G T G C G G C C C C A A A A G A G C A C C A A C C T G G T G  
50 A A G A A C A A G T G C G T G A A C T T C A A C T T T A A T G G C C T T A C C G G C A C G G G C G T G C T G A C C G A  
G A G C A A C A A G A A G T T C C T G C C A T T T C A G C A A T T C G G C A G A G A C A T C G A C G A C A C C A C C G  
A C G C G T C A G A G A C C C T C A G A C C C T G G A C A T C C T G G A C A T C A C C C C T G T A G C T T C G G C  
G G C G T G A G C G T A T C A C C C C G G C A C T A A C A C A A G C A A C C A A G T G G C C G T G C T G T A C C A  
A G G C G T G A A C T G C A C C G A G G T G C C C T G G C C A T C C A C G C C G A T C A G C T G A C C C C A C C T  
55 G G A G A G T A T A C A G C A C C G G A A G C A A C G T G T T T C A G A C A A G A G C C G G C T G C C T G A T C G G C  
G C C G A G C A C G T G A A C A A C A G C T A C G A G T G C G A C A T C C C C A T C G G C G C C G G C A T C T G C G C  
T A G C T A T C A G A C A C A G A C C A A C A G C C A C C G G A G A G C T A G A A G C G T G G C T T C T C A G A G C A  
T C A T C G C C T A C A C C A T G A G C C T G G G C G C C G A G A A C A G C G T G G C C T A C A G C A A C A A C A G C

ATCGCCATCCCCATCAACTTCACCATCAGCGTGACCACCGAGATCCTGCCCCGTAAGTATG  
 ACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCAGGTGCAGCA  
 ACCTGCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGCATC  
 GCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTCGCCCAAGTGAAGCAGATCTACA  
 5 AGACCCCCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCCGACCCTA  
 GCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGCC  
 GACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAGACCT  
 GATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCTCCCCCCCCCTGCTGACCGACGAGAT  
 GATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACATTTGG  
 10 CGCCGGGGCCGCACTGCAAATCCCCTTCGCCATGCAGATGGCCTACAGATTCAACGGCAT  
 CGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTCAACA  
 GCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCCTAGCGCCCTGGGCAAGCTG  
 CAAGACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCTGAGCA  
 GCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGGCTAGACTGGACCCCCCG  
 15 AGGCCGAGGTGCAGATCGACAGACTGATACCGGCAGACTGCAGAGCCTGCAGACCTA  
 CGTGACACAGAGCTGATCAGAGCCGCGGATCAGAGCTAGCGCCAACCTGGCCGCCA  
 CCAAGATGAGCGAGTGCCTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGGCAAGGG  
 CTACCACCTGATGAGCTTCCCTCAGAGCGCCCCCACGGCGTGGTGTTCCTGCACGTGAC  
 CTACGTGCCCGCCCAAGAGAAGAACTTCACCACCGCTCCCGCCATCTGCCACGACGGCA  
 20 AGGCCCACTTCCCTAGAGAGGGCGTGTTCGTGAGCAACGGCACCCACTGGTTCGTGACA  
 CAGAGAACTTCTACGAGCCTCAGATCATCACCAACCAACACCTTCGTGAGCGGCAA  
 CTGCGACGTGGTGAATCAGGACATAGTGAACAACACCGTATACGACCCCTGCAGCCGAGC  
 TGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCCCGACGTG  
 GACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTGAGAAGGAGATCGA  
 25 CCGCTGAATGAGGTGGCCAAGAACCCTGAACGAGAGCCTGATCGACCTGCAAGAGCTG  
 GGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGC  
 CTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTTGCAGCTGC  
 CTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGACAGCGAGC  
 CCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 22

>B.1.1.7 刺突蛋白氨基酸序列  
 MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT  
 WFHAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDLSDKTSLLIVNNATNVVI  
 KVCEFFQCNDFLGVYHKNNKSWMESEFRVYSSANNCTFEYVSQPFMDLEGKQGNFKNL  
 35 REFVFNIDGYFKIYKHTPINLVRDLPPQGSALPLVDLPIGINITRFQTLALHRSYLTYPQDS  
 SSGWTAGAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGITYQS  
 NFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNKRISNCVADYSVLYNSASFSTFKCY  
 GVSPTKLNLDLCTNRYADSFVIRGDEVRIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNL  
 DSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTYGV  
 40 GYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQF  
 GRDIDDTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQL  
 TPTWRVYSTGSNVFQTRAGCLIGAEHVNNSEYCDIPIGAGICASYQTQTNSHRRARSVASQSI  
 IAYTMSLGAENSVAYSNNIAIPINFITISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYG  
 SFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDL  
 45 LFNKVTLADAGFIKQYGDCLGDIAARDLCAQKFNGLTVLPPLTDEMIAQYTSALLAGTITS  
 GWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDLSSTASALG  
 KLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRQLSLQTYV  
 TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSPQSAPHGVVFLHVTYVP  
 AQEKNFTTAPAICHGDKAHFHPREGVVFVSNGTHWFTVQRNFYEPQIITHTNFVSGNCDVVIGI  
 50 VNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDIGINASVUNIKEIDRLNEVAKNLN  
 ESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMTIMLCCMTSCCSCCLKGCCSCGSCCKF  
 DEDDSEPVLGKVKLHYT

SEQ ID NO: 23

>SARS-COV-2 刺突蛋白 CDS  
 55 ATGTTCTGTGTTCCCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGACCACA  
 AGAACACAGCTGCCCCCGCCTACACCAACAGCTTCACAAGAGGCGTGTAACCTACCCCGA  
 CAAGGTGTTTCAAGAGCAGCGTGTGCACTCCACCCAAGACCTGTTTCTGCCCTTCTTCA  
 GCAACGTGACCTGGTTCACGCCATCCACGTGAGCGGCACCAACGGCACCAAGAGATTG

GACAACCCCGTGCTGCCCTTCAACGACGGCGTGTA CTTCGCTAGCACCGAAAAGAGCAA  
CATCATCAGAGGCTGGATCTTCGGCACCACCCTAGACTCCAAGACACAAAGCCTGCTGAT  
CGTTAACAACGCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACC  
CCTTCCTGGGCGTGTA CTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCCGT  
5 GTCTACAGCAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGGAC  
CTGGAGGGCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTTCGTGTTCAAGAACATCGA  
CGGCTACTTCAAGATCTACAGCAAGCACACCCCCATCAACCTGGTGAGAGACCTGCCCC  
AAGGCTTCAGCGCCCTGGAGCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAAG  
TTTCAGACCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCCGGCGATAGTTCTCTGGC  
10 TGGACGGCCGGCGCTGCCGCTACTACGTGGGTACCTGCAGCCTAGAACCTTCCTGCTG  
AAGTACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCCTGAG  
CGAGACCAAGTGCACCCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTATCAGACAAGCA  
ACTTCAGAGTGCAGCCACCGAGAGCATCGTGAGATTCCCCAACATCACCAACCTGTGC  
CCCTTCGGCGAGGTGTTCAACGCCACAAGATTTCGTAGCGTGTA CGCGTGGAACAGAAA  
15 GAGAATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAACAGCGCTAGCTTCAGCA  
CCTTCAAGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTCACCAACGTGT  
ACGCCGACAGCTTCGTGATCAGAGGGCGACGAGGTGAGACAGATTGCCCCCGGGCAAAAC  
CGGCAAGATCGCCGACTACA ACTACAAGCTGCCCGACGACTTCACCGGCTGCGTGATCG  
CGTGGAACAGCAATAACCTGGATTCCAAGGTGGGCGGCAACTACA ACTACCTGTACAGA  
20 CTGTTTCAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTACCA  
AGCCGGCAGCACCCCCTGCAACGGCGTGGAGGGCTTCAACTGCTACTTCCCCCTGCAGA  
GCTACGGCTTTCAGCCCACCAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGA  
GCTTCGAGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCAAAAAGAGCACCAACCT  
GGTGAAGAACAAGTGCGTGA ACTTCAACTTTAATGGCCTTACCGGCACGGGCGTGCTGA  
25 CCGAGAGCAACAAGAAGTTCTGECATTCAGCAATTCGGCAGAGACATCGCCGACACC  
ACCGACGCGGTGAGAGACCCTCAGACCCTGGAGATCCTGGACATCACCCCTGTAGCTT  
CGGCGGCGTGAGCGTGATCACCCCGGCACCTAACACAAGCAACCAAGTGGCCGTGCTGT  
ACCAAGACGTGA ACTGCACCGAGGTGCCCGTGGCCATCCACGCCGATCAGCTGACCCCC  
ACCTGGAGAGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCCGGCTGCCTGAT  
30 CGGCGCCGAGCACGTGAACAACAGCTACGAGTGCGACATCCCCATCGGCGCCGGCATCT  
GCGCTAGCTATCAGACACAGACCAACAGCCCTCGGAGAGCTAGAAGCGTGGCTTCTCAG  
AGCATCATCGCTACACCATGAGCCTGGGCGCCGAGAACAGCGTGGCCTACAGCAACAA  
CAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACACCGAGATCCTGCCCGTAAG  
TATGACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCGACAGCACCGAGTGCA  
35 GCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAACAGAGCCCTGACCGGC  
ATCGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTCCGCCAAGTGAAGCAGATCT  
ACAAGACCCCCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAAATCCTGCCCGAC  
CCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCT  
GGCCGACGCGGCTTCATCAAGCAGTACGGCGACTGCCTAGGCGACATCGCAGCTAGAG  
40 ACCTGATCTGCGCTCAGAAAGTTCAATGGACTCACCGTGCTCCCCCCCCTGCTGACCGACG  
AGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAACCATAACAAGCGGTTGGACAT  
TTGGCGCCGGGGCCGCACTGCAAATCCCCTTGCCCATGCAGATGGCCTACAGATTCAACG  
GCATCGGCGTGACACAGAACGTGCTGTACGAGAATCAGAAGCTGATCGCCAATCAGTTC  
AACAGCGCCATCGGCAAGATCCAAGACAGCCTGAGCAGCACCGCTAGCGCCCTGGGCA  
45 AGCTGCAAGACGTGGTGAATCAGAACGCCCAAGCCCTGAACACCCTGGTGAAGCAGCT  
GAGCAGCAACTTCGGCGCCATCAGCAGCGTATTGAATGACATCCTGAGCAGACTGGACC  
CCCCCGAGGCCGAGGTGCAGATCGACAGACTGATCACCGGCAGACTGCAGAGCCTGCA  
GACCTACGTGACACAGCAGCTGATCAGAGCCGCCGAGATCAGAGCTAGCGCCAACTGG  
CCGCCACCAAGATGAGCGAGTGCGTGCTGGGGCAGAGCAAGAGAGTGGACTTCTGCGG  
50 CAAGGGCTACACCTGATGAGCTTCCCTCAGAGCGCCCCCACCAGCGGTGGTCTCTGCG  
ACGTGACCTACGTGCCCGCCCAAGAGAAGAACTTACCACCCTGCCCATCTGCCAC  
GACGGCAAGGCCCACTTCCCTAGAGAGGGCGTGTTCTGTGAGCAACGGCACCCACTGGTT  
CGTGACACAGAGAACTTCTACGAGCCTCAGATCATCACCAACGACAACACCTTCGTGA  
GCGGCAACTGCGACGTGGTGATCGGCATAGTGAACAACACCGTATACGACCCCCTGCAG  
55 CCGAGCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAAAACCATACTAGCCC  
CGACGTGGACCTGGGAGACATCTCTGGCATCAACGCTAGCGTGGTGAACATTAGAAGG  
AGATCGACCGGTGAATGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCA  
AGAGCTGGGCAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCA

TCGCCGGCCTGATCGCCATCGTGATGGTGACCATCATGCTGTGCTGCATGACAAGCTGTT  
GCAGCTGCCTGAAAGGTTGCTGCAGCTGTGGCTCGTGCTGCAAGTTCGACGAGGACGAC  
AGCGAGCCCGTGCTGAAGGGCGTGAAGCTGCACTACACCTGATGA

SEQ ID NO: 24

5 >SARS-COV-2 刺突蛋白氨基酸序列  
MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT  
WFHAIHVSNGTKRFDNPNVLPFNDGVYFASTEKSNIIRGWIFGTTLDSTQSLIVNNATNV  
VIKVFCEQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFMDLEGKQGNF  
10 KNLREFVFNIDGYFKIYSKHTPINLVRDLPGFSALEPLVDLPIGINITRFQTLALHRSYLT  
GDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIY  
QTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTF  
KCYGVSPTKLNLDLCTNVYADSFVIRGDEVRIAPGQTGKIADYNYKLDDFTGCVIAWNS  
NNLDSKVGNNYLYRLFRKSNLKPFRDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPT  
15 NGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLP  
FQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIH  
ADQLTPTWRVYSTGSNVFQTRAGCLIGAHEVNNSYECDIPIGAGICASYQTQTNPRRARSV  
ASQSIIAYTMSLGAENSVAYSNNISIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNL  
LLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRS  
FIEDLLFNKVTADAGFIKQYGDCLGDIAARDLICAQKFNGLTVPPLLTDEMIAQYTSALLA  
20 GTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTVQNVLYENQKLIANQFNSAIGKIQDSLST  
ASALGKLQDVVNQNAQALNTLVKQLSSNFGAIVSLNDILSRLDPPEAEVQIDRLITGRLQSL  
QTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHV  
TYVPAQEKNFHTTAPAICHGDKAHFPREGVFVSNGTHWFVTQRNFYEPQIITDNTFVSGNCD  
VVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVUNIQKEIDRLNEVA  
25 KNLNESLIDLQELGKYEYIKWPWYIWLGLFIAGLIAIVMTIMLCCMTSCCCLKGCCSCGS  
CCKFDEDDSEPVKGVKLHYT

SEQ ID NO: 25

&gt;荧光素酶 CDS

ATGGAGGACGCCAAGAACATCAAGAAAGGCCCCGCCCCCTTCTACCCCCTGGAGGACG  
30 GCACCGCCGGCGAGCAGCTGCACAAAGCCATGAAAAGATACGCTCTGGTGCCCGGCAC  
CATCGCCTTCACCGACGCCACATCGAGGTGAACATCACCTACGCCGAGTACTTCGAGA  
TGAGCGTGAGACTGGCCGAGGCTATGAAGAGATACGGCTTGAACACCAACCACAGAAT  
CGTGGTGTGCAGCGAGAACAGCCTGCAGTTCTTCATGCCCGTGCTGGGCGCCCTGTTCA  
TCGGCGTGCCGTTGCCCCGCCAACGACATCTACAACGAGAGAGAGCTGCTGAACAG  
35 CATGAACATCTCTCAGCCCACCGTGGTTTTCGTGTCCAAAAAGGGCCTGCAGAAGATCC  
TGAACGTGCAGAAGAAGCTGCCCATCATTAGAAGATCATCATGACAGCAAGACC  
GACTACCAAGGCTTTCAGAGCATGTACACCTTCGTGACAAGCCACCTGCCCCCGGCTT  
CAACGAGTACGACTTCGTGCCCGAGAGCTTCGACAGAGACAAGACCATCGCCCTGATC  
ATGAACAGCAGCGGCAGCACCGGCCTGCCCAAGGGCGTGCCCTGCCCCACAGAACC  
40 GCCTGCGTGAGATTAGCCACGCTAGAGACCCCATCTTCGGCAATCAGATCATCCCCGA  
CACCGCCATCCTGAGCGTAGTGCTTTCCATCATGGCTTCGGCATGTTACCAACCTGGG  
CTACCTGATCTGCGGCTTCAGAGTGGTGCTGATGTACAGATTGAGAGGAGGCTGTTCC  
TGAGAAGCCTGCAAGACTACAAGATTAGAGCGCCCTGCTGGTGCCCAACCTGTTTCAG  
CTTCTTCGCCAAGAGCACCTGATCGACAAGTATGACCTGAGCAACCTGCACGAAATTG  
45 CGAGCGGCGGAGCGCCGCTCAGCAAAGAGGTGGGCGAGGCCGTGGCCAAGAGATTCC  
ACCTGCCCCGCATCAGACAAGGCTACGGCCTGACCGAGACCACAAGCGCCATCCTGAT  
CACCCCCGAGGGCGACGATAAGCCCCGGCGCCGTGGGCAAGGTGGTACCATTCTTCGAG  
GCCAAGGTGGTGACCTGGACACCGCAAGACCCTGGGCGTGAATCAGAGAGGCGAG  
CTGTGCGTGAGAGGCCCATGATCATGAGCGGCTACGTGAACAACCCCCGAGGCCACCA  
50 ACGCCCTCATCGATAAGGATGGGTGGCTGCACAGCGGCGACATCGCCTACTGGGACGA  
GGACGAGCACTTCTTCATCGTGACAGACTCAAGAGCCTGATCAAGTACAAGGGCTAC

CAAGTGGCCCCCGCCGAGTTGGAGAGCATCTGCTGCAGCACCCCAATATCTTCGACGC  
TGGCGTGGCCGGACTACCCGACGACGACGCCGGCGAGTTACCCGCCGCCGTGGTGGTG  
CTGGAGCACGGCAAGACCATGACCGAGAAGGAGATCGTGGACTACGTGGCTAGCCAAG  
TGACCACCGCCAAGAAGCTGAGAGGCGGCGTGGTGTTCGTTGACGAAGTGCCCAAGG  
5 GCCTGACCGGCAAGCTGGACGCTAGAAAGATCAGAGAGATCCTGATCAAGGCCAAGAA  
GGGCGGCAAGAGCAAGCTGTGA

SEQ ID NO: 26

&gt;荧光素酶氨基酸序列

MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAHIEVNITYAEYFEMSVR  
10 LAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVAPANDIYNERELNSMNISQPT  
VVFVSKKGLQKILNVQKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPESFD  
RDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQIIPDTAILSVPFHHGFGMFT  
TLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASG  
GAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPFPEAKVVDLDT  
15 GKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDKDGWLHSGDIAYWDEDEHFFIVDRL  
KSLIKYKGYQVAPAELESILLQHPNIFDAGVAGLPDDDAGELPAVVVLEHGKTMTEKEIVD  
YVASQVTTAKLRGGVVVFDEVPKGLTGKLDARKIREILIKAKKGGKSKL

SEQ ID NO: 27

&gt;LUC1-1'截短体

20 TAGGAGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 28

&gt;LUC2-1'截短体

AGTTGGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 29

25 &gt;LUC3-1'截短体

GGTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 30

&gt;LUC4-1'截短体

GCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

30 SEQ ID NO: 31

&gt;LUC5-1'截短体

TTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 32

&gt;LUC6-1'截短体

35 AATAACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 33

&gt;LUC7-1'截短体

ACAGGTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 34

40 &gt;LUC8-1'截短体

GTCATTGCCGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 35

&gt;LUC9-1'截短体

TTGCCGAGAAAAGGATAGCACTATAAT

45 SEQ ID NO: 36

&gt;LUC10-1'截短体

CGAGAAAAGGATAGCACTATAAT

SEQ ID NO: 37

&gt;LUC11-1'截短体

50 AAAAGGATAGCACTATAAT

SEQ ID NO: 38

- >LUC12- I'截短体  
GGATAGCACTATAAT  
SEQ ID NO: 39
- >LUC13- I'截短体  
5 AGCACTATAAT  
SEQ ID NO: 40
- >LUC14- I'截短体  
CTATAAT
- 10 >LUC15- I'突变体  
GAGTTAGGAGTTGCTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTA  
TAAT  
SEQ ID NO: 42
- >LUC16- I'突变体  
15 GAGTTAGGAGTTGTTGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTA  
TAAT  
SEQ ID NO: 43
- >5'UTR I 通式  
20 GAGTTAGGAGTTGN<sub>1</sub>TGGCAGTTTCAATAACAGGTCATTGCCGAGAAAAGGATAGCACTA  
TAAT  
SEQ ID NO: 44
- N<sub>1</sub> 选自 A、G、C 或 T  
以下 SEQ ID NO: 45-79 为 RNA 序列。
- >5'UTR I(63bp)  
25 GAGUUAGGAGUUGAUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA  
CUAUAUU  
SEQ ID NO: 45
- >5'UTR I'(63bp)  
30 GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA  
CUAUAUU  
SEQ ID NO: 46
- >5'UTR G  
GCACUUUUCUGAGCAGACGUCCAGAGCAGAGUCAGCCGCCACC  
SEQ ID NO: 47
- 35 >5'UTR A  
ACAUUUGCUUCUGACACAACUGUGUUCACUAGCAACCUCAAACGCCGCCACC  
SEQ ID NO: 48
- >5'UTR J  
40 AGAGAGGCAAAGAAACAUUGUGAAAUCUCCAACUCUUAACCUUCAAGCCGCCACC  
SEQ ID NO: 49
- >3'UTR B  
GCUCGCUUUCUUGCUGUCCAAUUCUAUUAAAAGGUUCCUUGUUCUUUAGGUCCAAC  
UACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAA  
AAACAUUUAUUUCAUUGCAA  
45  
SEQ ID NO: 50
- >3'UTR F  
CAGACAAGACAAGCUACUGAAUACGUUCACAUCUGUCUUGAUGCCUAAUAUUUUUAC  
AUUUCUGUAAACAUAUUUCUGAAAUAUUUUUUGCCUUUCAAGCGACAGAUGCCUCA  
UUUUGUGAAAACUUAUGAUGAUUUUGUGUUUAAGUCCAAACAUAUUGAAUAAAAU  
50 AAUUGACAAUA  
SEQ ID NO: 51
- >3'UTR E  
CCUGAGGCCUCAGAACACAGGUUUCAGAUUGAUAGGCCUGCAGGUCUCCAGGCAGCA

ACCAGCUGAGCGACUAAAGGGCCCAAGGCCAGGGCUCUAGGGAUGGGGCUCAGCAGA  
GGCUGGGGUAAGGGGAGCCAGGGAGGAGCUGGGCCUAAUGCAGCACCGGGUCCCCAG  
GAUGCCGUGUCUGGGCUGGAGGAGGAGAUGCGGAAGCUCAGGCCACGGUGCAGGAG  
CUCCAGAAGCGCUUGGACAGGCUGGAGGAGACAGUCCAGGCCAAGUAGAGCCCCGCA  
5 GGGCCUCCAGCAGGGUCAGCCAUUCACACCCAUCCACUCACCUCUCCAUUCCAGCCAC  
AUGGCAGAGAAAAAAUCAUAAUAAAAUGGCUUUAUUUUCUGGUA

SEQ ID NO: 52

>3'UTR D  
GGGGCCUUCUGACAUGAGUCUGGGCCUGGCCCCACCUCUAGUUCUCAAUAAUAAAGA  
10 CAGAUUGCUUCUUCGCUUCUCACUGAGGGGCCUUCUGACAUGAGUCUGGGCCUGGGCC  
CACCUCUCCAGUUUCUCAAUAAUAAAGACAGAUUGCUUCUUCACUUGAAGCUCGCUUU  
CUUGCUGUCCAAUUUCUAAUAAAGGUUCCUUUGUUCUCCUAAAGUCCAACUACUAAACU  
GGGGGAUUAUUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACAUUA  
UUUUCAUUGCAA

15 SEQ ID NO: 53

>B.1.617.2(Delta) 刺突蛋白 mRNA  
AUGUUCGUGUCCUGGUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACCCUGAGG  
ACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGUGUACUAC  
20 CCGACAAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUUUUUCUGCCC  
UUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUCAGCGGCACCAACGGCACC  
AAGAGAUUCGACAACCCCGUGCUGCCCUUCAAACGACGGCGUGUACUUCGCUAGCACC  
GAAAAGAGCAACAUCAUCAGAGGCUGGAUCUUCGGCACCACCCUAGACUCCAAGACA  
CAAAGCCUGCUGAUCGUUAAACAACGCGACCAACGUGGUGAUCAAGGUGUGCGAGUUU  
25 CAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAAGAGCUGG  
AUGGAGAGCGGGGUCUACAGCAGCGCCAACAACUGCACCUCUGAGUACGUGUCUCAG  
CCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUCAAGAACCUAAGAGAGUUU  
GUGUUCAAGAACAUCGACGGCUACUUAAGAUUCUACAGCAAGCACACCCCAUCAAC  
CUGGUGAGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCC  
30 AUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGGCCUGCACCUGUAGCUACCUA  
ACCCCGGGCGAUAGUCCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUGGGC  
UACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCACUGAU  
GCCGUGGACUGUGCCCUGGACCCCUAGAGCGAGACCAAGUGCACCCUCAAGAGCUUC  
ACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACCGAGAGC  
35 AUCGUGAGAUUCCCCAACAUACCAACCUUGUGCCCCUUCGGCGAGGUGUUAACGCC  
ACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUCGCGUGGCC  
GACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUC  
AGCCCCACCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGACAGCUUCGUG  
AUCAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAAGAUCCGGAC  
40 UACAACUACAAGCUGCCCGACGACUUCACCGGCGUGGUGAUCGCGUGGAACAGCAAU  
AACCUGGAUUCCAAGGUGGGCGGCAACUACAACUACCGGUACAGACUGUUCAGAAAG  
AGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUUCUACCAAGCCGGCAGC  
AAGCCUUGCAACGGCGUGGAGGGCUUCAAUCGCUACUUCUCCUUGCAGAGCUACGGC  
UUUCAGCCCAACACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUC  
45 GAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCAAAAAGAGCACCAACCUGGUG  
AAGAACAAGUGCGUGAACUUAACUUAUAAUGGCCUUAACGGCACGGGCGUGCUGACC  
GAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCGACACC  
ACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCUUGUAGC  
UUCGGCGGGGUGAGCGUGAUCACCCCGGCACUAAACAAAGCAACCAAGUGGCCGUG  
50 CUGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCAUCCACGCCGAUCAGCUG  
ACCCUACCCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGC  
UGCCUGAUUCGGCGCCGAGCAGGUGAACAACAGCUACGAGUGCGACAUCUCCUACGGC  
GCCGGCAUCGCGCUAGCUAUCAGACACAGACCAACAGCCGCCGAGAGCUAGAAGC  
GUGGCUUUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUG  
55 GCCUACAGCAACAACAGCAUCGCCAUCCCAACCAACUUCACCAUCAGCGUGACCACCG  
AGAUCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCG  
GCGACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGC  
UGAACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCAAGAGGUGU

UCGCCCCAAGUGAAGCAGAUUCUACAAGACCCCCCAUCAAGGACUUCGGCGGCUUCA  
ACUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGG  
ACCUGCUGUUCAACAAGGUGACCCUGGCCGACGCCGGCUUCAUCAAGCAGUACGGCG  
ACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUUCAUUGGAC  
5 UCACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCU  
GCUGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAU  
CCCCUUCGCCAUGCAGAUUGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGU  
GCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUCAACAGCGCCAUCGGCAAGAU  
CCAAGACAGCCUGAGCAGCACCUCUAGCGCCCCUGGGCAAGCUGCAAAACGUGGUGAA  
10 UCAGAACGCCAAGCCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGC  
CAUCAGCAGCGUAUUGAAUGACAUCUUGUCUAGACUGGACCCCCCGAGGCCGAGGU  
GCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACUGAGACACA  
GCAGCUGAUCAGAGCCGCCGAGAUCAGAGCUAGCGCCAACCUUGGCCGCCACCAAGAU  
GAGCGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACC  
15 ACCUGAUGACGUUCCCUAGAGCGCCCCCACGGCGUGGUGUUCUUGCAGCUGACCU  
ACGUGCCCCGCCAAGAGAAGAACUUCACCACCGCUCUCCGCCAUCUGCCAGCGGCA  
AGGCCACUUCUCCUAGAGAGGGCGUGUUCUGUGAGCAACGGCACCCACUGGUUCGUGA  
CACAGAGAAACUUCUACGAGCCUCAGAUCAUACCAACCGACAACACCUUCGUGAGCG  
GCAACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGC  
20 CCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCC  
CCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAG  
AGGAGAUUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACC  
UGCAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCUGGUACAUCUGGCUGG  
GCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGA  
CAAGCUGUUGCAGCUGCCUGAAAGGUUCUGCAGCUGUGGCUCGUGCAAGUUCG  
25 ACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 54

&gt;B.1.1.529(Omicron) 刺突蛋白 mRNA

AUGUUCGUGUUCUUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACCCUGACC  
30 ACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGUGUACUAC  
CCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUUUUCUGCCC  
UUCUUCAGCAACGUGACCUGGUUCCACGUGAUCAGCGGCACCAACGGCACCAAGAGA  
UUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCAUCGAAAAG  
AGCAACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGC  
35 CUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUCAGUUC  
UGCAACGACCCCUUCCUGGACCAACAAGAACAAGAGCUGGAUGGACGCGAGUUC  
AGAGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGUCUCAGCCCUUCCUG  
AUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAAGAGAGUUCGUGUUA  
GAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCCAUCAUCGUGAGAGA  
40 GCCCCGAGGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACCUGCCAU  
CGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCGUGCACCGUAGCUACCUAAC  
CCCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCGUGCCGCCUACUACGUGGGCUA  
CCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCACUGAUGC  
CGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCCUCAAGAGCUUCAC  
45 CGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAU  
CGUGAGAUUCCCCAACAUACCAACCUUGGCCCUUCGACGAGGUGUUAACGCCAC  
CAGAUUUGCUAGCGUGUACGCCUGGAACCGGAAGAGAAUCAGCAACUGCGUGGCCGA  
CUACAGCGUGCUGUACAACCUUGGCCCCCUUCUUCACCUUCAAGUGCUAUGGCGUCAG  
CCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAU  
50 CAGAGGCGACGAGGUGAGACAGAUUCGCCCCGGGCAGACCGGCAACUUCGCCGACUA  
CAACUACAAGCUGCCCCGACGACUUCACCGGCUGCGUGAUCGCCUGGAACAGCAACAA  
GCUGGAUAGCAAGGUGAGCGGCAACUACAUAUACCUGUACAGACUGUUCAGAAAGA  
GCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCAGAGAUUCAACAGCCGGCAACA  
AGCCCUGCAACGGCGUGGGCGGCUUACAACUGCUACUUCUCCCCUCAGGAGCUACAGCU  
55 UCAGACCCACCUACGGCGUGGGCCAUCAGCCCUACAGAGUGGUCGUGCUGAGCUUCG  
AGCUGCUGCACGCCCCUGCCACAGUGUGCGGCCCCAAAAAGAGCACCAACCUGGUGA  
AGAACAAGUGCGUGAACUUAACUUAACGGCCUGAAGGGCACCGGCGUGCUGACCG  
AGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCGACACCA

CCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCCCUGUAGCU  
 UCGGCGGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCGUGC  
 UGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUCAGCUGA  
 CCCCCACCUGGAGAGUGUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAGCCGGCU  
 5 GCCUGAUCGGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCG  
 CCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUAGAAGCG  
 UGGCUAGCCAAAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUGG  
 CCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCACCGA  
 GAUCCUGCCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGG  
 10 CGACAGCACCCGAGUGCAGCAACCUGCUCCUGCAGUACGGCAGCUUCUGCACACAGCU  
 GAAGAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUU  
 CGCCCAAGUGAAGCAGAUUCAAGACCCCCCAUCAAGUACUUCGGCGGCUUCA  
 CUUCAGCCAAAUCUGCCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGGA  
 CCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUAAGCAGUACGGCGA  
 15 CUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUUCAAGGGACU  
 CACCGUGCUCCCCCCCCUGCUGACCCGACGAGAUAGUAGCUCAGUACACCUUGCCCCUG  
 CUGGCGGGAACCAUAACAAGCGGUUGGACAUUUUGCGCCGGGGCCGCAUCGAAAU  
 CCCUUCGCCAUGCAGAUUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUG  
 CUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCCGCAAGAUC  
 20 CAAGACAGCCUGAGCAGCACCGCUAGCGCCUGGGCAAGCUGCAAGACGUGGUGAAC  
 CACAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAAGUUCGGCGCC  
 AUCAGCAGCGUGCUGAAUGACAUCUUCUCUAGACUGGACCCCCCGAGGCCGAGGUG  
 CAGAUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAG  
 CAGCUGAUCAGAGCCGCCGAGAUACAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUG  
 25 AGCGAGUGCGUGCGGGGACAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCAC  
 CUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUAC  
 GUGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGACGGCAAG  
 GCCCACUUCUUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACA  
 CAGAGAAACUUCUACGAGCCUCAGAUCAUACCAACCGACAACACCUUCGUGAGCGGC  
 30 AACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCC  
 GAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCC  
 GACGUGGACCUUGGGAGACAUUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAAG  
 GAGAUACCGGCGUGAAUGAGGUGGCCAAGAACGCUAGCGUGGUGAACAUUCAGAAG  
 CAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCCUUGGUACAUCUGGCGGGC  
 35 UUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACA  
 AGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGAC  
 GAGGACGACAGCGAGCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 55

>poly A  
 40 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 56

>I-Delta24-B mRNA  
 GAGUUAAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA  
 45 CUUAAUAUGUUCGUGUCCUGGUGCUGCGUCCCCUGGUGAGCUCUCAGUGCGUGAA  
 CCUGAGGACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGU  
 GUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUU  
 UCGCCCCUUCUUCAGCAACGUGACCUUGGUUCCACGCCAUCCACGUCAGCGGCACCAA  
 CGGCACCAAGAGAUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGC  
 50 UAGCACCGAAAAGAGCAACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUC  
 CAAGACACAAAGCCUGCUGAUCGUUACAACGCGACCAACGUGGUGAUCAAGGUGUG  
 CGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACA  
 GAGCUGGAUGGAGAGCGGGUCUACAGCAGCGCCAACAACUGCACCUCGAGUACGU  
 GUCUCAGCCCUUCCUGAUGGACCUAGGGGCAAGCAAGGCAUUUCAAGAACCUAAG  
 55 AGAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCC  
 CAUCAACCUGGUGAGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGA  
 CCUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCUGCACCUGAG  
 CUACCUAACCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUA

CGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAU  
CACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCCUCAA  
GAGCUUACCCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCAC  
CGAGAGCAUCGUGAGAUUCCCCAACAUACCAACCUUGUGCCCCUUCGGCGAGGUGUU  
5 CAACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUG  
CGUGGGCCGACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUA  
UGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG  
CUUCGUGAUCAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAAGAU  
CGCCGACUACAACUACAAGCUGCCCCGACGACUUCACCGGCUGCGUGAUCGCGUGGAA  
10 CAGCAAUAACCUUGAUUCCAAGGUGGGCGGCAACUACAACUACCGGUACAGACUGUU  
CAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUUCUACCAAGC  
CGGCAGCAAGCCUUGCAACGGCGUGGAGGGCUUCAACUGCUACUCCCCCUGCAGAG  
CUACGGCUUUCAGCCCACCAACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCU  
GAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCCAAAAAGAGCACCAA  
15 CCUGGUGAAGAACAAGUGCGUGAACUUAACUUUAUGGCCUUAACGGCACGGGCGCU  
GCUGACCGAGAGCAACAAGAAGUUCUGGCCAUUUCAGCAAUUCGGCAGAGACAUCGC  
CGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUACCCCC  
CUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGU  
GGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGA  
20 UCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAG  
AGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAACAAACAGCUACGAGUGCGACAUCCC  
CAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCGCCGGAGAGC  
UAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAA  
CAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUACGCGU  
25 GACCACCGAGAUCCUGCCCCGUAAGUAUGACCAAGACAAGCGUGGACUAGCACCAGUA  
CAUCUGCGGCGACAGCACCGAGUGCAGCAACCUUGCUGCUGCAGUACGGCAGCUUCUG  
CACACAGCUGAACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCA  
AGAGGUGUUCGCCCCAAGUGAAGCAGAUUCAAGACCCCCCCCCAUCAAGGACUUCGG  
CGGCUUCAACUUCAGCCAAAUCUGCCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUU  
30 CAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGGCCGACGCCGGCUUCAUAAGCA  
GUACGGCGACUGCCUAGGGGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUU  
CAAUGGACUACACCGUGCUCCCCCCCCUGCUGACCGACGAGAUUAUCGCUACAGUACAC  
CUCUGCCUGCUGGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCCCGGGGGCCGC  
ACUGCAAAUCCCCUUCGCCAUGCAGAUUGGCCUACAGAUUCAACGGCAUCGGCGUGAC  
35 ACAGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAU  
CGGCAAGAUAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAAA  
CGUGGUGAAUCAGAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAA  
CUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCUGUCUAGACUGGACCCCCCGGA  
GGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUA  
40 CGUGACACAGCAGCUGAUCAGAGCCGCCGAGAUACAGAGCUAGCGCCAACCUGGCCGC  
CACCAAGAUGAGCGAGUGCGUGCUGGGGCGAGAGCAAGAGAGUGGACUUCUGCGGCA  
AGGGCUACCACCUGAUGAGCUUCCCUACAGAGCGCCCCCACGGCGUGGUGUCCUGC  
ACGUGACCUACGUGCCCCGCCCAAGAGAAGAAUCUACCAACCGCUCCCCGCCAUCUGCC  
ACGACGGCAAGGCCACUUCUUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACU  
45 GGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACCAACCGACAACACCU  
UCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAACAAACCCGUUAACGACC  
CCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUAACAAACC  
AUACUAGCCCCGACGUGGACCUUGGGAGACAUCUCUGGCAUACCGCUAGCGUGGUGA  
ACAUUCAGAAGGAGAUUCGACCGGCUAGAUAGAGGUGGCCAAGAACCUGAACGAGAGCC  
50 UGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCCUUGGUACA  
UCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU  
GCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCU  
GCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACA  
CCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAUUUCUAUUAAAGGUUCCUUUGU  
55 UCCCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGA  
UUCUGCCUAAUAAAAACAUUUAUUUUAUUGCAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 57

&gt;I-OC24-B mRNA

GAGUUAAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAAGGAUAGCA  
CUAUAUAUUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA  
5 CCUGACCACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGU  
GUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUU  
UCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGUGAUCAGCGGCACCAACGGCAC  
CAAGAGAUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCAU  
CGAAAAGAGCAACAUCAUACAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGAC  
10 ACAAAGCCUGCUGAUUCGUUAAACAACGCGACCAACGUGGUGAUCAAGGUGUGCGAGUU  
UCAGUUCUGCAACGACCCCUUCCUGGACCACAAGAACAACAAGAGCUGGAUGGAGAG  
CGAGUUCAGAGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGUCUCAGCC  
CUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAUUUCAAAGAACCUAAGAGAGUUCG  
UGUUCAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCCAUCAUCG  
15 UGAGAGAGCCCAGGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUGGACC  
UGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCUGCACCCGUAGCU  
ACCUAACCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACG  
UGGGCUACCUUGCAGCCUAGAACCUUCCUGUGAAGUACAACGAGAACGGCACUAUCA  
CUGAAGCCGUGGACUGGCCUGGACCCCGUAGCGAGACCAAGUGCACCCUCAAGA  
20 GCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCCACCG  
AGAGCAUCGUGAGAUUCCCCAACAUACCAACCUUGGCCCUUCGACGAGGUGUUA  
ACGCCACCAGAUUUGCUAGCGUGUACGCCUGGAACCGGAAGAGAAUCAGCAACUGCG  
UGGCCGACUACAGCGUGCUGUACAACCUUGGCCCUUUCUUCACCUUCAAGUGCUAUG  
GCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGACAGCU  
25 UCGUGAUCAGAGGCGACGAGGUGAGACAGAUUCGCCCCGGGCAGACCGGCAACAUCG  
CCGACUACAACUACAAGCUGCCCCGACGACUUCACCGGCUGCGUGAUCGCCUGGAACA  
GCAACAAGCUGGAUAGCAAGGUGAGCGGCAACUACAUAUACCUGUACAGACUGUUA  
GAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUUCACCAAGCCG  
GCAACAAGCCUGCAACGGCGUGGCGGGCUUACAACUGCUACUUCCCCCUCAGGAGCU  
30 ACAGCUUCAGACCCACCUACGGCGUGGGCCAUCAGCCCUACAGAGUGGUCGUGCUGA  
GCUUCGAGCUGCUGCACGCCCUUGCCACAGUGUGCGGCCCAAAAAGAGCACCAACC  
UGGUGAAGAACAAGUGCGUGAACUUAACUUAACGGCCUGAAGGGCACCGGCGUGC  
UGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCG  
ACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUACCCCCU  
35 GUAGCUUCGGCGGCGUGAGCGUGAUCACCCCGGCACUAAACACAAGCAACCAAGUGG  
CCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGAUC  
AGCUGACCCCCACCUGGAGAGUGUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAG  
CCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACAUCCCCA  
UCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUA  
40 GAAGCGUGGCUAGCCAAAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACA  
GCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGA  
CCACCGAGAUCCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACA  
UCUGCGGCGACAGCACCGAGUGCAGCAACCUUGCUCUCCUGCAGUACGGCAGCUUCUGCA  
CACAGCUGAAGAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAG  
45 AGGUGUUCGCCAAGUGAAGCAGAUUCAAGACCCCCCAUCAAGUACUUCGGCG  
GCUUCAACUUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUA  
UCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUAAGCAGU  
ACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUUA  
AGGGACUACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUACAGUACACCU  
50 CUGCCCUUGCUGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCAC  
UGCAAAUCCCCUUCGCCAUGCAGAUUGCCUACAGAUUCAACGGCAUCGGCGUGACAC  
AGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCG  
GCAAGAACCAAGACAGCCUGAGCAGCAGCCGUAAGCGCCUGGGCAAGCUGCAAGACG  
UGGUGAACCACAACGCCCAAGCCUGAACACCCUUGGUGAAGCAGCUGAGCAGCAAGU  
55 UCGGCGCCAUCAGCAGCGUGCUGAUAUGACAUCUUCUCUAGACUGGACCCCCCGAGG  
CCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACG  
UGACACAGCAGCUGAUCAGAGCCGCCGAGAUCAGAGCUAGCGCCAACCUUGGCCGCCA  
CCAAGAUGAGCGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAG

GGCUACCACCUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCAC  
GUGACCUACGUGCCCCGCCAAGAGAAGAACUUCACCACCGCUGCCGCCAUCUGCCAC  
GACGGCAAGGCCACUUCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGG  
UUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACCACCGACAACACCUUC  
5 GUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCC  
CUGCAGCCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAU  
ACUAGCCCCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAAC  
AUUCAGAAGGAGAUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCU  
GAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGGCCUGGUACAU  
10 CUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUG  
CUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUG  
CAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACAC  
CUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUU  
CCCUAAGUCCAACUACUAAACUGGGGGAUUAUUGAAGGGCCUUGAGCAUCUGGAU  
15 UCUGCCUAAUAAAAACAUAUUUUUUAUUGCAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 58

>B.1.351 刺突蛋白 mRNA  
20 AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACUUCACG  
ACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGUGUACUAC  
CCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUUUUCUGCCC  
UUCUUCAGCAACGUGACCUUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACC  
AAGAGAUUCGCCAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACC  
25 GAAAAGAGCAACAUCAUCAGAGGCUGGAUCUUCGGCACCACCCUAGACUCCAAGACA  
CAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUU  
CAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGACGUGG  
AUGGAGAGCGAGUUCGUGUCUACAGCAGCGCCAACAACUGCACCUCGAGUACGUG  
UCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUCAAGAACCUAAGA  
30 GAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCC  
AUCACCCUGGUGAGAGGGCCUGCCCCAAGGCUUCAGCGCCUGGAGCCCCUGGUGGAC  
CUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCACCUGAGCUACCUAACC  
CCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCGUGCCGCCUACUACGUGGGCUAC  
CUGCAGCCUAGAACCUUCCUGCUAAGUACAACGAGAACGGCACUAUCACUGAUGCC  
35 GUGGACUGUGCCCUGGACCCCUAGCGAGACCAAGUGCACCCUACAAGAGCUUCACC  
GUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCACCGAGAGCAUC  
GUGAGAUUCCCCAACAUACCAACCUUGUGCCCCUUCGGCGAGGUGUUAACGCCACA  
AGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGGCCGAC  
UACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGGGUCAGC  
40 CCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCUUCGUGAUC  
AGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCAAACCGGCAACAUCCGGACUAC  
AACUACAAGCUGCCCCGACGACUUCACCGGCUGCGUGAUCGCGUGGAACAGCAAUAAC  
CUGGAUCCAAGGUGGGCGGCAACUACAACUACCUGUACAGACUGUUCAGAAAGAGC  
AACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUUCACCAAGCCGGCAGCACC  
45 CCCUGCAACGGCGUGAAGGGCUUAACUGCUACUUCCCCCUGCAGAGCUACGGCUUU  
CAGCCCACCUACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUCGAG  
CUGCUGCACGCCCCGCAACCGUGUGCGGCCCAAAAAGAGCACCAACCUUGGUGAAG  
AACAAGUGCGUGAACUUAACUUUAAUGGCCUUAACGGCACGGGCGUGCUGACCGAG  
AGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGCCGACACCACC  
50 GACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCCCUGUAGCUUC  
GGCGGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCGGUGCUG  
UACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCAUCCACGCCGAUCAGCUGACC  
CCCACCGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCGUC  
CUGAUCGGCGCCGAGCACGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCGCC  
55 GGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCCCGGAGAGCUAGAAGCGUG  
GCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGUCGAGAACAGCGUGGCC  
UACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGACCACCGAG  
AUCCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCGGC

GACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUGCACACAGCUG  
AACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUC  
GCCCAAGUGAAGCAGAUUCAAGACCCCCCAUCAAGGACUUCGGCGGCUUACAC  
UUCAGCCAAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGGAC  
5 CUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUCAAGCAGUACGGCGAC  
UGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUUCAAUGGACUC  
ACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUCAGUACACCUCUGCCCUGC  
UGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAUUC  
CCUUCGCCAUGCAGAUUGGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGC  
10 UGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUC  
AAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGUGGUGAAUC  
AGAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGCCA  
UCAGCAGCGUAUUGAAUGACAUCUCCUGUCUAGACUGGACCCCCCGAGGCCGAGGUGC  
AGAUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGC  
15 AGCUGAUCAGAGCCGCCGAGAUACAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGA  
GCGAGUGCGUGCUGGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUGACACC  
UGAUGAGCUUCCCUACAGAGCGCCCCACGGCGUGGUGUUCUUGCAGCUGGACCUACG  
UGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGACGGCAAGG  
CCCACUCCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACAC  
20 AGAGAAACUUCUACGAGCCUCAGAUCAUACACCACCGACAACACCUUCGUGAGCGGCA  
ACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCG  
AGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCG  
ACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAAGG  
AGAUCGACCGGCUGAUUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGC  
25 AAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCUUGGUACAUCUGGCGGGCU  
UCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAA  
GCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACG  
AGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA

SEQ ID NO: 59

30 &gt;B.1.1.7 刺突蛋白 mRNA

AUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACCUGACC  
ACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGUGUACUAC  
CCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCGUUCUUCUGCCC  
UUCUUCAGCAACGUGACCUGGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGAGA  
35 UUCGACAACCCCGUGCUGCCCUCAACGACGGCGUGUACUUCGCUAGCACCAAGAG  
AGCAACAUCAGAGGCUAGAUUCUUGGCACCAACCUAGACUCCAAGACCAAGAGC  
CUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUCAGUUC  
UGCAACGACCCCUUCCUGGGCGUGUACCACAAGAACAACAAGAGCUGGAUGGAGAGC  
GAGUUCGUGUCUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGUCUACAGCCC  
40 UUCUGAUGGACCUGGAGGGCAAGCAAGGCAUUUCAAGAACCUAAGAGAGUUCGU  
GUUCAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCCAUCAACCU  
GGUGAGAGACCUGCCCCAAGGCUUCAGCGCCCCUGGAGCCCCUGGUGGACCUGCCCAU  
CGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCUUGCACCUGAGCUACCUAAC  
CCCCGGCGAUAGUCCUCUGGCUGGACGGCCGGCGUGCCGCCUACUACGUGGGCUA  
45 CCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUCACUGAUGC  
CGUGGACUGUGCCCUGGACCCCUAGCGGAGACCAAGUGCACCCUCAAGAGCUUCAC  
CGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACCGAGAGCAU  
CGUGAGAUUCCCCAACAUACCAACCUUGGCCCUUCGGCGAGGUGUUAACGCCAC  
AAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGCGUGGCCG  
50 ACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAUGGCGUCA  
GCCCCACCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGACAGCUUCGUGA  
UCAGAGGCGACGAGGUGAGACAGAUUUGCCCCGGGCAACCGGCAAGAUCGCCGACU  
ACAAUCUACAAGCUGCCCCGACGACUUCACCGGCGUGGUGAUCGCGUGGAACAGCAAUA  
ACCGGAUUCCAAGGUGGGCGGCAACUACAACUACCGUACAGACUGUUCAGAAAGA  
55 GCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUUCAACAGCCGGCAGCA  
CCCCUGCAACGGCGUGGAGGGCUUACAACUGCUACUUCCCCCUGCAGAGCUACGGCU  
UUCAGCCCACCUACGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCUGAGCUUCG  
AGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCCAAAAGAGCACCAACCUGGUGA

AGAACAAGUGCGUGAACUUAACUUUAAUGGCCUUACCGGCACGGGCGUGCUGACCG  
AGAGCAACAAGAAGUCCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGACGACACCA  
CCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCUGGACAUCACCCCCUGUAGCU  
UCGGCGGGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGUGGCCGUGC  
5 UGUACCAAGGCGUGAACUGCACCAGGUGCCCCGUGGCCAUCCACGCCGAUCAGCUGA  
CCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAGAGCCGGCU  
GCCUGAUCGGCGCCGAGCAGUGAACAACAGCUACGAGUGCGACAUCCCCAUCGGCG  
CCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCACCGGAGAGCUAGAAGCG  
UGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACAGCGUGG  
10 CCUACAGCAACAACAGCAUCGCCAUCCCCAUCAACUUCACCAUCAGCGUGACCACCG  
AGAUCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACAUCUGCG  
GCGACAGCACCGAGUGCAGCAACUUGCUGCUGCAGUACGGCAGCUUCUGCACACAGC  
UGAACAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAGAGGUGU  
UCGCCCAAGUGAAGCAGAUCAACAGACCCCCCAUCAAGGACUUCGGCGGCUUCA  
15 ACUUCAGCCAAAUCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGG  
ACCUGUGUUCAACAAGGUGACCCUGGCCGACGCCGGCUUCAUCAAGCAGUACGGCG  
ACUGCCUAGGGCAGCAUCGCGAGCUAGAGACCUAGUUCUGCGCUCAGAAAGUCAAUGGAC  
UCACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUACAGUACACCUUGCCCU  
GCUGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAU  
20 CCCCUCGCCAUUGCAGAUUGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGU  
GCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAU  
CCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGUGGUGAA  
UCAGAACGCCAAGCCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAACUUCGGCGC  
CAUCAGCAGCGUAUUGAAUGACAUCUUGGCUAGACUGGACCCCCCGAGGCCGAGGU  
25 GCAGAUCCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACUGACACA  
GCAGCUGAUCAGAGCCGCCGAGAUAGAGCUAGCGCCAACCUUGGCCGCCACCAAGAU  
GAGCGAGUGCGUGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACC  
ACCUGAUGAGCUUCCCUAGAGCGCCCCCACGGCGUGGUGUCCUGCAGCUGACCU  
ACGUGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGACGGCA  
30 AGGCCACAUUCCCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGA  
CACAGAGAAACUUCUACGAGCCUCAGAUCAUCACCACCCACAACACCUUCGUGAGCG  
GCAACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGC  
CCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUAACCAACUACUAGCC  
CCGACGUGGACCUUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAG  
35 AGGAGAUUCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACC  
UGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUUGGUACAUCUGGCUGG  
GCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGA  
CAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCG  
ACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGA  
40 SEQ ID NO: 60

>SARS-COV-2 刺突蛋白 mRNA  
AUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACCUGACC  
ACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGUGUACUAC  
CCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCC  
45 UUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUGAGCGGCACCAACGGCACC  
AAGAGAUUCGACAACCCCGUGCUGCCCUUACAACGACGGCGUGUACUUCGCUAGCACC  
GAAAAGAGCAACAUCAUCAGAGGCUUGGAUCUUCGGCACCAACCUAGACUCCAAGACA  
CAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUCAGGUGUGCGAGUUU  
CAGUUCUGCAACGACCCCUUCCUGGGCGUGUACUACCACAAGAACAACAAGAGCUGG  
50 AUGGAGAGCGAGUUCGUGUCUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUG  
UCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAAGA  
GAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCC  
AUCAACCUGGUGAGAGACCUGCCCCAAGGCUUCAGCGCCUGGAGCCCCUGGUGGAC  
CUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGC  
55 UACCUAACCCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUAC  
GUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAUC  
ACUGAUGCCGUGGACUGUGCCUUGGACCCCUAGCGAGACCAAGUGCACCCUCAAG  
AGCUUACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACC

GAGAGCAUCGUGAGAUUCCCCAACAUACCAACCUGUGCCCCUUCGGCGAGGUGUUC  
AACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUGC  
GUGGCCGACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUAU  
GGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGACAGC  
5 UUCGUGAUACAGAGGCGACGAGGUGAGACAGAUUGCCCCCGGGCA5AACC GGCAAGAU  
CGCCGACUACAACUACAAGCUGCCCCGACGACUUCACCGGCUUGCUGAUUCGCGUGGAA  
CAGCAAUAACCUGGAUUCGAAGGUGGGCGGCAACUACAACUACCUGUACAGACUGUU  
CAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUCUACCAAGC  
CGGCAGCACCCCCUGCAACGGCGUGGAGGGCUUCAACUGCUACUUCCCCCUGCAGAG  
10 CUACGGCUUUCAGCCCACCAUUGGCGUGGGCUAUCAGCCCUACAGAGUGGUGGUGCU  
GAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGGCCCCAAAAAGAGCACCAA  
CCUGGUGAAGAACAAGUGCGUGAACUUAACUUUAAUGGCCUUACCGGCACGGGCGU  
GCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACAUCGC  
CGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCC  
15 CUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCCGGCACUAACACAAGCAACCAAGU  
GGCCGUGCUUACCAAGACGUGAACUGCACCGAGGUGCCCCGUGGCCAUCCACGCCGA  
UCAGCUGACCCCCACCCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAG  
AGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAACAAACAGCUACGAGUGCGACAUCCC  
CAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCUCGGAGAGC  
20 UAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAA  
CAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGU  
GACCACCGAGAUCCUGCCCCUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUA  
CAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUG  
CACACAGCUGAACAGAGCCUGACCGGCAUCGCGGUGGAGCAAGACAAGAACCCCCA  
25 AGAGGUGUUCGCCCAAGUGAAGCAGAUCAACAAGACCCCCCAUCAAGGACUUCGG  
CGGCUUCAACUUCAGCCAAAUCCUGCCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUU  
CAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGGCGACGCCGGCUUCAUAAGCA  
GUACGGCGACUGCCUAGGGGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUU  
CAAUGGACUACCGUGCUCCCCCCCCUGCUGACCGACGAGAUAGUAGCUACAGUACAC  
30 CUCUGCCCUGCUGGCGGGAACCAUAACAAGCGGUUGGACAUAUUGGCGCCGGGGCCGC  
ACUGCAAAUCCCCUUCGCCAUGCAGAUUGCCUACAGAUUCAACGGCAUCGGCGUGAC  
ACAGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAU  
CGGCAAGAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCUGGGCAAGCUGCAAGA  
CGUGGUGAAUACAGAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAA  
35 CUUCGGCGCCAUCAGCAGCGUAUUGAAUGACAUCUGAGCAGACUGGACCCCCCGGA  
GGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUA  
CGUGACACAGCAGCUGAUCAGAGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGC  
CACCAAGAUGAGCGAGUGCGUGCUGGGGCGAGAGCAAGAGAGUGGACUUCUGCGGCA  
AGGGCUACCACUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGUUCUGC  
40 ACGUGACCUACGUGCCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCCGCCAUCUGCC  
ACGACGGCAAGGCCACUUCUUUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACU  
GGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACACCACCGACAACACCU  
UCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAACAAACCCGUAUACGACC  
CCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUAACAAACC  
45 AUACUAGCCCCGACGUGGACCUGGGAGACAUCUCUGGCAUACACGCUAGCGUGGUGA  
ACAUUCAGAAGGAGAUUCGACCGGCUAUGAGGUGGCCAAGAACCUGAACGAGAGCC  
UGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCUGGUACA  
UCUGGCUUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU  
GCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUUGCUGCU  
50 GCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACA  
CCUGAUGA

SEQ ID NO: 61

&gt;荧光素酶 mRNA

AUGGAGGACGCCAAGAACAUCAAGAAAGGCCCGCCCCCUUACCCCCUGGAGGA  
55 CGGCACCGCCGGCGAGCAGCUGCACAAAGCCAUGAAAAGAUACGCUCUGGUGCCCG  
GCACCAUCGCCUUCACCGACGCCCAUCGAGGUGAACAUACCUACGCCGAGUACU  
UCGAGAUGAGCGUGAGACUGGCCGAGGCUAUGAAGAGAUACGGCUUGAACACCAAC

CACAGAAUCGUGGUGUGCAGCGAGAACAGCCUGCAGUUCUUAUGCCCGUGCUGGG  
CGCCUGUUAUCGGCGUGGCCGUUGCCCCGCCAACGACAUCUACAACGAGAGAG  
AGCUGCUGAACAGCAUGAACAUUCUCAGCCCACCGUGGUUUUCGUGUCCAAAAAG  
GGCCUGCAGAAGAUCCUGAACGUGCAGAAAGAAGCUGCCCAUCAUUCAGAAGAUCAU  
5 CAUCAUGGACAGCAAGACCGACUACCAAGGCUUUCAGAGCAUGUACACCUUCGUGA  
CAAGCCACCUGCCCCCGGCUUACAGGAGUACGACUUCGUGCCCGAGAGCUUCGACA  
GAGACAAGACCAUCGCCCUGAUCAUGAACAGCAGCGGCAGCACCGGCCUGCCCAAG  
GGCGUGGCCUGCCCCACAGAACCGCCUGCGUGAGAUUCAGCCACGCUAGAGACCCC  
AUCUUCGGCAAUCAGAUCAUCCCCGACACCGCCAUCUGAGCGUAGUGCCUUUCCA  
10 UCAUGGCUUCGGCAUGUUCACCACCCUGGGCUACCUGAUCUGCGGCUUCAGAGUGG  
UGCUGAUGUACAGAUUCGAGGAGGAGCUGUUCUGAGAAGCCUGCAAGACUACAAG  
AUUCAGAGCGCCCUGCUGGUGCCCACCCUGUUCAGCUUCUUCGCCAAGAGCACCCUG  
AUCGACAAGUAUGACCUGAGCAACCUGCACGAAAUUGCGAGCGGCGGAGCGCCGCU  
CAGCAAAGAGGUGGGCGAGGCCGUGGCCAAGAGAUUCCACCUGCCCGGCAUCAGAC  
15 AAGGCUACGGCCUGACCGAGACCACAAGCGCCAUCUGAUCACCCCGAGGGCGACG  
AUAAGCCCCGGCGCCGUGGGCAAGGUGGUACCAUUCUUCGAGGCCAAGGUGGUGGAC  
CUGGACACCGGCAAGACCCUGGGCGUGAAUCAGAGAGGCGAGCUGUGCGUGAGAGG  
CCCCAUGAUCAUGAGCGGCUACGUGAACAACCCCGAGGCCACCAACGCCCUCAUCGA  
UAAGGAUGGGUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACU  
20 UCUUCAUCGUGGACAGACUCAAGAGCCUGAUCAAGUACAAGGGCUACCAAGUGGCC  
CCCGCCGAGUUGGAGAGCAUCCUGCUGCAGCACCCCAAUAUCUUCGACGCUGGCGU  
GGCCGACUACCCGACGACGACGCCGGCGAGUUACCCGCCGCCGUGGUGGUGCUGG  
AGCACGGCAAGACCAUGACCGAGAAGGAGAUUCGUGGACUACGUGGCUAGCCAAGUG  
ACCACCGCCAAGAAGCUGAGAGGGCGGCGUGGUUCGUUGACGAAGUGCCCAAGGG  
25 CCUGACCGGCAAGCUGGACGCUAGAAAGAUCAGAGAGAUCCUGAUCAAGGCCAAGA  
AGGGCGGCAAGAGCAAGCUGUGA

SEQ ID NO: 62

>LUC1-1'截短体  
UAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAU  
30 AAU

SEQ ID NO: 63

>LUC2-1'截短体  
AGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU  
35

SEQ ID NO: 64

>LUC3-1'截短体  
GGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU  
40

SEQ ID NO: 65

>LUC4-1'截短体  
GCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU  
40

SEQ ID NO: 66

>LUC5-1'截短体  
UUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU  
45

SEQ ID NO: 67

>LUC6-1'截短体  
AAUAACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU  
45

SEQ ID NO: 68

>LUC7-1'截短体  
ACAGGUCAUUGCCGAGAAAAGGAUAGCACUAUAAU

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |               |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 69 |
|    | >LUC8-1'截短体<br>GUCAUUGCCGAGAAAAGGAUAGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |               |
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 70 |
| 5  | >LUC9-1'截短体<br>UUGCCGAGAAAAGGAUAGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |               |
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 71 |
|    | >LUC10-1'截短体<br>CGAGAAAAGGAUAGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |               |
| 10 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 72 |
|    | >LUC11-1'截短体<br>AAAAGGAUAGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |               |
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 73 |
|    | >LUC12-1'截短体<br>GGAUAGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |               |
| 15 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 74 |
|    | >LUC13-1'截短体<br>AGCACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |               |
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 75 |
| 20 | >LUC14-1'截短体<br>CUAAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |               |
|    | >LUC15-1'突变体<br>GAGUUAGGAGUUGCUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC<br>ACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |               |
| 25 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 77 |
|    | >LUC16-1'突变体<br>GAGUUAGGAGUUGUUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC<br>ACUAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |               |
| 30 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 78 |
|    | >5'UTR I 通式<br>GAGUUAGGAGUUGN <sub>2</sub> UGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA<br>CUAAUAAU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |               |
|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | SEQ ID NO: 79 |
| 35 | N <sub>2</sub> 选自 A、G、C 或 U<br>>野生型 SARS-COV-2 刺突氨基酸序列<br>MFVFLVLLPLVSSQCVNLTTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVT<br>WFHAIHVSNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNATN<br>VVIKVCEFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQG<br>40 NFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLALHRSY<br>LTPGDSSSGWTAGAAAYYVGYLQPRTELLKYNENGTITDAVDCALDPLSETKCTLKSTVE<br>KGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSA<br>SFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVI<br>AWNSNNLDSKVGNNYLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSY<br>45 GFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNGLTGTGVLTESN<br>KKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTE<br>VPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNNSYECDIPIGAGICASYQTQTNSP<br>RRARSVASQSIIAYTMSLGAENSVAYSNNISAIPTNFTISVTTEILPVSMTKTSVDCTMYICGD |               |

STECNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPD  
PSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPLLTDemia  
QYTSALLAGTITSGWTFGAGAAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNsaig  
KIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQID  
5 RLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQS  
APHGVVFLHVTYVPAQEKNFHTTAPAICHGKAHFPREGVSVSNGTHWFVTQRNFYEPQIIT  
TDNTFVSGNCDVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLDGDISGINASVV  
NIQKEIDRLNEVAKNLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTS  
CCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

10 SEQ ID NO: 80

>野生型 SARS-COV-2 刺突蛋白 CDS (NC\_045512.2)  
ATGTTTGTCTTTCTTGTCTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCA  
GAAGTCAATTACCCCTGCATACACTAATTCTTTCACACGTGGTGTCTTATTACCCTGACA  
AAGTTTTAGATCCTCAGTTTTACATTCAACTCAGGACTTGTCTTACCTTTCTTTTCCAA  
15 TGTTACTTGGTTCCTATGCTATACATGTCTCTGGGACCAATGGTACTAAGAGGTTTGATAA  
CCCTGTCCTACCATTAAATGATGGTGTCTTATTTGCTTCCACTGAGAAGTCTAACATAATA  
AGAGGCTGGATTTTTGGTACTACTTTAGATTCGAAGACCCAGTCCCTACTTATTGTTAATA  
ACGCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTGGG  
TGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTATTCTA  
20 GTGCGAATAATTGCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAA  
AACAGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAA  
AATATATTCTAAGCACACGCCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCT  
TTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTTCAAACCTTACTTG  
CTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTGGACAGCTGGTGCTG  
25 CAGCTTATTATGTGGGTATCTTCAACCTAGGACTTTTCTATTAAAAATAAATGAAAATGG  
AACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGACGT  
TGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTTCTAACTTTAGAGTCCAACCA  
ACAGAATCTATTGTTAGATTTCCTAATATTACAACTTGTGCCCTTTTGGTGAAGTTTTTA  
ACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTT  
30 GCTGATTATTCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTGTATGGAGTGTC  
TCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAG  
GTGATGAAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTAT  
AAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGAATTTCAACAATCTTGATTCTA  
AGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTT  
35 TGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTG  
AAGGTTTTAATTGTTACTTTCCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTG  
TTACCAACCATACAGAGTAGTAGTACTTTCTTTTGAACCTTCTACATGCACCAGCAACTGT  
TTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAA  
TGGTTTAAACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCCTTTCCAAC  
40 AATTTGGCAGAGACATTGCTGACACTACTGATGCTGTCCGTGATCCACAGACACTTGAG  
ATTCTTGACATTACACCATGTTCTTTTGGTGGTGTGTCAGTGTATAACACCAGGAACAAAT  
ACTTCTAACCAGGTTGCTGTTCTTTATCAGGATGTAACTGCACAGAAGTCCCTGTTGCT  
ATTCATGCAGATCAACTTACTCCTACTTGGCGTGTATTATTCTACAGGTTCTAATGTTTTTC  
AAACACGTGCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGAC  
45 ATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGG

GCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAAT  
TCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTACTATTAGTGTTACCA  
CAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACATTTGTG  
GTGATTCAACTGAATGCAGCAATCTTTTGTGCAATATGGCAGTTTTTGTACACAATTAA  
5 ACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACAAAAACACCCAAGAAGTTTTTGC  
ACAAGTCAAACAAATTTACAAAACACCACCAATTAAGATTTTGGTGGTTTTAATTTTTC  
ACAAATATTACCAGATCCATCAAAACCAAGCAAGAGGTCATTTATTGAAGATCTACTTTT  
CAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTG  
ATATTGCTGCTAGAGACCTCATTGTGCACAAAAGTTAACGGCCTTACTGTTTTGCCAC  
10 CTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCA  
CTTCTGGTTGGACCTTTGGTGCAGGTGCTGCATTACAAATACCATTGCTATGCAAATGG  
CTTATAGGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGA  
TTGCCAACCAATTTAATAGTGCTATTGGCAAATTTCAAGACTCACTTTCTCCACAGCAA  
GTGCACTTGGAAAACCTCAAGATGTGGTCAACCAAAATGCACAAGCTTTAAACACGCT  
15 TGTTAAACAACCTTAGCTCCAATTTTGGTGAATTTCAAGTGTTTTAAATGATATCCTTTCA  
CGTCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCA  
AAGTTTGCAGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCTG  
CTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATT  
TTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAGTCAGCACCTCATGGTGTAGTCT  
20 TCTTGCATGTGACTTATGTCCCTGCACAAGAAAAGAACTTCACAACGTCTCCTGCCATTT  
GTCATGATGGAAAAGCACACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACAC  
TGGTTTGTAAACACAAAGGAATTTTATGAACCACAAATCATTACTACAGACAACACATTT  
GTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTG  
CAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTAAGAATCATAATCA  
25 CCAGATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACATTCAAAAA  
GAAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAA  
GAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATA  
GCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCAGTTGCTGTA  
GTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCAAATTTGATGAAGACGACTCT  
30 GAGCCAGTGCTCAAAGGAGTCAAATTACATTACACATAA

SEQ ID NO: 81

&gt;野生型 SARS-COV-2 刺突蛋白 mRNA

AUGUUUUGUUUUUCUUGUUUUUUAUUGCCACUAGUCUCUAGUCAGUGUGUUAUCUUAC  
AACCAGAACUCAUUACCCUGCAUACACUAAUUCUUUCACACGUGGUGUUUAUU  
35 ACCUGACAAAGUUUUCAGAUCCUCAGUUUUACAUUCAACUCAGGACUUGUUCUUA  
CCUUUCUUUCCAAUGUUACUUGGUUCCAUGCUAUACAUGUCUCUGGGACCAUUGG  
UACUAAGAGGUUUGAUAAACCCUGUCCUACCAUUUAAUGAUGGUGUUUAUUUUGCUU  
CCACUGAGAAGUCUAACAUAUAAGAGGCUGGAUUUUUGGUACUACUUUAGAUUCG  
AAGACCCAGUCCCUACUUUUUGUUAAUAACGCUACUAAUGUUGUUUUAAAGUCUG  
40 UGAAUUUCAUUUUUGUAAUGAUCCAUUUUUGGGUGUUUAUUACCACAAAACAACA  
AAAGUUGGAUGGAAAGUGAGUUCAGAGUUUAUUCUAGUGCGAAUAAUUGCACUUU  
UGAAUAUGUCUCUCAGCCUUUUUCUUAUGGACCUUGAAGGAAAACAGGGUAAUUUCA  
AAAAUCUUAGGGAAUUUGUGUUUAAGAAUAUUGAUGGUUAAUUUAAAAUAUAUUC  
UAAGCACACGCCUAUUAAUUUAGUGCGUGAUCUCCUCAGGGUUUUUCGGCUUUAG  
45 AACCAUUGGUAGAUUUUGCCAAUAGGUUAUAACAUCACUAGGUUUCAAACUUUACUU

GCUUUACAUAAGAAGUUAUUUGACUCCUGGUGAUUCUUCUUCAGGUUGGACAGCUGG  
UGCUGCAGCUUAUUAUGUGGGUUAUCUUCAACCUAGGACUUUUCUAUUAAAAUAUA  
AUGAAAAUGGAACCAUACAGAUUGCUGUAGACUGUGCACUUGACCCUCUCUCAGAA  
ACAAAGUGUACGUUGAAAUCCUUCACUGUAGAAAAAGGAAUCUAUCAAACUUCUAA  
5 CUUUAGAGUCCAACCAACAGAAUCUAUUGUUAGAUUCCUAAUAAUACAAACUUGU  
GCCCCUUUGGUGAAGUUUUAAACGCCACCAGAUUUGCAUCUGUUUAUGCUUGGAAC  
AGGAAGAGAAUCAGCAACUGUGUUGCUGAUUAUUCUGUCCUAUAUAAUUCGCAUC  
AUUUUCCACUUUUAAAGUGUUAUGGAGUGUCUCCUACUAAAUUAAAUGAUCUCUGCU  
UUACUAAUGUCUAUGCAGAUUCAUUUGUAAUUAAGAGGUGAUGAAGUCAGACAAAU  
10 CGCUCCAGGGCAAACUGGAAAGAUUGCUGAUUAUAAUUAUAAAUUACCAGAUGAUU  
UUACAGGCUGCGUUUAAGCUUGGAAUUCUACAAUCUUGAUUCUAAAGGUUGGUGGU  
AAUUAUAAUUAACUGUAUAGAUUGUUUAGGAAGUCUAAUCUCAAAACUUUUGAGA  
GAGAUAAUUAACUGAAAUCUAUCAGGCCGGUAGCACACCUUGUAAUGGUGUUGAA  
GGUUUAAUUGUUACUUUCCUUUACAAUCAUAUGGUUCCAACCCACUAAUGGUGU  
15 UGGUUAACCAACCAUACAGAGUAGUAGUACUUUCUUUUGAACUUCUACAUGCACCAG  
CAACUGUUUGUGGACCUAAAAAGUCUACUAAUUUGGUUAAAAACAAUUGUGUCAAU  
UUCAACUCAAUGGUUUAACAGGCACAGGUGUUCUACUGAGUCUAAACAAAAGUU  
UCUGCCUUUCCAACAAUUUGGCAGAGACAUUGCUGACACUACUGAUGCUGUCCGUG  
AUCCACAGACACUUGAGAUUCUUGACAUAUACCAUGUUCUUUUGGUGGUGUCAGU  
20 GUUAUAACACCAGGAACAAUACUUCUAACCAGGUUGCUGUUCUUUAUCAGGAUGU  
UACUGCACAGAAGUCCUGUUGCUAUUCAUGCAGAUCAACUACUCCUACUUGGC  
GUGUUUAUUCUACAGGUUCUAAUGUUUUUCAAACACGUGCAGGCUGUUUAAUAGGG  
GCUGAACAUUGUCAACAACUCAUAUGAGUGUGACAUAACCAUUGGUGCAGGUUAUUG  
CGCUAGUUAUCAGACUCAGACUAAUUCUCCUCGGCGGGCACGUAGUGUAGCUAGUC  
25 AAUCCAUAUUGCCUACACUAUGUCACUUGGUGCAGAAAAUUCAGUUGCUUACUCU  
AAUAAUCUCUAUUGCCAUAACCCACAAAUUUUACUAAUAGUGUUACCACAGAAAUUCU  
ACCAGUGUCUAUGACCAAGACAUCAGUAGAUUGUACAAUGUACAUAUUGUGGUGAUU  
CAACUGAAUGCAGCAAUCUUUUGUUGCAAUAUGGCAGUUUUUGUACACAAUUAAC  
CGUGCUUUAACUGGAAUAGCGUUGAACAAGACAAAAACACCCAAGAAGUUUUUGC  
30 ACAAGUCAAAACAAAUUUACAAAACACCACCAAUUAAAGAUUUUGGUGGUUUAAU  
UUUCACAAAUUUUACCAGAUCCAUAACAAACCAAGCAAGAGGUCAUUUAUUGAAGAU  
CUACUUUUCAACAAAGUGACACUUGCAGAUUGCUGGCUUCAUCAAACAAUAUGGUGA  
UUGCCUUGGUGAUUAUUGCUGCUAGAGACCUCUAAUUGUGCACAAAAGUUUAACGGCC  
UUACUGUUUUGCCACCUUUGCUCACAGAUGAAAUGAUUGCUCAAUACACUUCUGCA  
35 CUGUUAGCGGGUACAAUCACUUCUGGUUGGACCUUUGGUGCAGGUGCUGCAUUA  
AAUACCAUUUGCUAUGCAAAUGGCUUAUAGGUUUAUUGGUAUUGGAGUUACACAG  
AAUGUUCUCUAUGAGAACCACAAAAUUGAUUGCCAACCAAUUAAUAGUGCUAUUGG  
CAAAAUUCAAGACUCACUUUCUCCACAGCAAGUGCACUUGGAAAACUUCAAGAUG  
UGGUCAACCAAAUUGCACAAGCUUUAACACGCUUGUUAACAACUAGCUCCA  
40 UUUGGUGCAAUUUCAAGUGUUUUAAAUGAUAUCCUUUCACGUCUUGACAAAGUUGA  
GGCUGAAGUGCAAUUGAUAGGUUGAUACAGGCAGACUUCAAAGUUUGCAGACAU  
AUGUGACUCAACAAUUAUUAAGAGCUGCAGAAAUCAGAGCUUCUGCUAAUUCUUGCU  
GCUACUAAAAUGUCAGAGUGUGUACUUGGACAAUCAAAGAGUUGAUUUUUGUG  
GAAAGGGCUAUCUUCUUAUGUCCUUCUCCUCAGUCAGCACCUCAUGGUGUAGUCUUC  
45 UUGCAUGUGACUUAUGUCCCGCACAAAGAAAAGAACUUCACAACUGCUCCUGCCAU

UUGUCAUGAUGGAAAAGCACACUUUCCUCGUGAAGGUGUCUUUGUUUCAAUUGGCA  
CACACUGGUUUUGUAACACAAAGGAAUUUUUAUGAACCACAAUUAUACUACAGAC  
AACACAUUUUGUGUCUGGUAACUGUGAUGUUGUAAUAGGAAUUGUCAACAACACAGU  
UUAUGAUCCUUUGCAACCUGAAUUAGACUCAUUAAGGAGGAGUUAGAUAAAUAUU  
5 UUAAGAAUCAUACAUCACCAGAUGUUGAUUUAGGUGACAUCUCUGGCAUUAUUGCU  
UCAGUUGUAAACAUUCAAAAAGAAAUUGACCGCCUCAUUGAGGUUGCCAAGAAUUU  
AAAUGAAUCUCUCAUCGAUCUCCAAGAACUUGGAAAGUAUGAGCAGUAUUAUAAA  
GGCCAUGGUACAUUUUGGCUAGGUUUUAUAGCUGGCUUGAUUGCCAUGUAAUUGGUG  
ACAAUUAUGCUUUGCUGUAUGACCAGUUGCUGUAGUUGUCUCAAGGGCUGUUGUUC  
10 UUGUGGAUCCUGCUGCAAAUUUGAUGAAGACGACUCUGAGCCAGUGCUCAAAGGAG  
UCAAAUUAUACAUAACAUAA

SEQ ID NO: 82

>A 型流感病毒 H1N1 (A/Wisconsin/588/2019) HA 氨基酸序列  
MKAILVVMlyTFTTANADTLcIGYHANNSTDTVDTVLEKNVTVTHSVNlLEDKHNGKLCK  
15 LRGVAPLHLGKCNIAGWILGNPECESLSTARSWSYIVETSNSDNGTCYPGDFINYEELREQL  
SSVSSFERFEIFPKTSSWPNHSDNGVTAACPHAGAKSFYKNLIWLVKKGKSYPKINQTYIN  
DKGKEVLVLWGIHHPPTIADQQSLYQNADAYVFGTSRYSKKFKPEIATRPKVRDQEGRM  
NYYWTLVEPGDKITFEATGNLVAPRYAFTMERDAGSGIIISDTPVHDCNTTCQTPEGAINTS  
PFQNVHPITIGKCPKYVKSTKLRLATGLRNVPSIQSRGLFGAIAAGFIEGGWTGMVDGWYGY  
20 HHQNEQSGSYAADLKSTQNAIDKITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKV  
DDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRNQLKNNAKEIGNGCFEFYHKC  
DNTCMESVKNGTYDYPKYSEEAKLNREKIDGVKLDSTRIYQILAIYSTVASSLVLVVSLGAI  
SFWMCSNGSLQCRICI

SEQ ID NO: 83

>A 型流感病毒 H3N2 (A/Cambodia/e0826360/2020) HA 氨基酸序列  
MKTIIALSylCLVFAQKIPGNDNSTATLCLGHHAVPngTIVKtITNDRIEVTNATELVQNSSI  
GEICDSPHQILDGGNCTLIDALLGDPQCDGFQNKewDLfVERSrANSNCYPYDVPDYASLR  
SLVASSGTLEfKNESFNWTGVKQNGTSSACIRGSSSSFFSRLNWLTHLNyTYPALNVTMPN  
NEQFDKLYIWGVHHPSTDKDQISLFAQPSGRITVSTKRSQqAVIPNIGSRPRIRDIPSRIyWT  
30 IVKPGDILLINSTGNLIAPRGYfKIRSGKSSIMRSDaPIGKCKSECITPngSIPNDKPFQNVNRI  
TYGACPRYVKQSTLKLATGMRNVPEKQTRGIFGAIAAGFIENGWEGMVDGWYgFRHQNSE  
GRGQAADLKSTQAAIDQINGKLNRlIGKTNEKFHqIEKEfSEVEGRVQDLEKYVEDTKIDL  
WSYNAELLVALENQHTIDLTdSEMnKLFekTKKQLRENAEDMGNGCFKIYHKCDNACIGS  
IRNETYdHNVYRDEALNNRFQIKGVELKSGYKdWILWISFAMSCfLLCiallGFIMWACQK  
35 GNIRCNICI

SEQ ID NO: 84

>B 型流感病毒 Victoria 谱系 (B/Washington/02/2019) HA 氨基酸序列  
MKAIIVLLMVVtSNADRICTGITSSNSPHVVKtATQGEVNVTGVIPLTTTPtKSHFANLKGt  
ETRgKLCpKCLNCTDLdVALGRPKCTGKIPSARVSILHEVRPVTSGCFpIMHDRTKIRQLPN  
40 LLRGYEHVRLSTHNVINAEDAPGRPYEIGTSgSCPnITNGNGFFATMAWAVPKNKtATNPLT  
IEVPYICtEGEDQITVWGFHSDNETQMAKLYGDSKpQKFTSSANGVTTHYVSQIGGFpNQ  
EDGGLPQSGRIVVDYmVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSPLIGEADCL  
HEKYGGLNKSKPYTGEHAKAIGNCPIWVKtPLKLANGTKYRPPAKLLKERGFfGAIAGF  
LEGGWEGMIAGWHGYTSHGAHGVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGA  
45 MDELHNEILELDEKVDdLRADTISSQIELAVLLSNEGIINSEdHLLALERKLKKMLGPSAVE

IGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSL NITAASLNDDGLDNHTILLYSTA  
ASSLAVTLMIAIFVVMVSRDNVSCSICL

SEQ ID NO: 85

>B 型流感病毒 Yamagata 谱系 (B/Phuket/3073/2013) HA 氨基酸序列

5 MKAIIVLLMVVTSNADRICTGITSSNSPHVVKTATQGEVNVTVIPLTTPTKSYFANLKGT  
RTRGKLCPDCLNCTDLVALGRPMC VGTTPSAKASILHEVRPVTSGCFPI MHDRTKIRQLPN  
LLRGYEKIRLSTQNVIDAEKAPGGPYRLGTSGSCP NATSKIGFFATMAWAVPKDNYKNATNP  
LTVEVPYICTEGEDQITVWGFHSDNKTQMKSLYGDSNPQKFTSSANGVTTHYVSQIGDFPD  
QTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSPLIGEA  
10 DCLHEEYGG LNKSKPYT GKHAKAIGNCPIWVK TPLKLANGTKYRPPAKLLKERGFFGAI  
AGFLEGGWEGMIAGWHGYTSHGAHGVAADLKSTQEAINKITKNLSLSELEVKNLQRL  
SGAMDELHNEILELDEKVDDL RADTISSQIELAVLLSNEGIINSEDEHLLALERKLLKMLGP  
SAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSL NITAASLNDDGLDNHTILLY  
YSTAASSLAVTLMIAIFIVM VSRDNVSCSICL

15 SEQ ID NO: 86

>A 型流感病毒 H1N1 (A/Wisconsin/588/2019) HA CDS

ATGAAGGCCATCCTGGTGGTGATGCTGTACACCTCACCACCGCCAACGCCGACACCCT  
GTGCATCGGCTACCACGCCAACACAGCACCGACACCGTGACACCGTGCTGGAGAAG  
AACGTGACCGTGACCCACAGCGTGAACCTGCTGGAGGACAAGCACAAACGGCAAGCTG  
20 TGCAAGCTGCGGGGCGTGGCCCCCTGCACCTGGGCAAGTGCAACATCGCCGGCTGGA  
TCCTGGGCAACCCCGAGTGCAGTCCCTGAGCACCGCCCGCTCCTGGAGCTACATCGT  
GGAGACCTCCAACAGCGACAACGGCACCTGCTACCCCGGGGACTTCATCAACTACGAG  
GAGCTGCGGGAGCAGCTGTCCAGCGTGTCCAGCTTCGAGCGCTTCGAGATCTTCCCCA  
AGACCTCCAGCTGGCCCAACCACGACTCCGACAACGGCGTGACCGCCGCCTGCCCCCA  
25 CGCCGGGGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAAGTCC  
TACCCCAAGATCAACCAGACCTACATCAACGACAAGGGGAAGGAGGTCTCTGCTGCTGT  
GGGGCATCCACCACCCCCCACCATCGCCGACCAGCAGAGCCTCTACCAGAACGCCGA  
CGCCTACGTCTTCGTGGGGACCTCCCGGTACAGCAAGAAGTTCAAGCCCGAGATCGCC  
ACCCGCCCAAGGTCCGGGACCAGGAGGGCCGCATGAAC TACTACTGGACCCTGGTGG  
30 AGCCGGGGGACAAGATCACCTTCGAGGCCACCGGCAACCTCGTCGCCCCGCGGTACGC  
CTTCACGATGGAGCGCGACGCCGGGTCCGGCATCATCATCAGCGACACCCCGTGCAC  
GACTGCAACACGACCTGCCAGACGCCCCGAGGGGGCGATCAACACCTCCCTGCCGTTCC  
AGAACGTCCACCCGATCACGATCGGCAAGTGCCCCAAGTACGTGAAGTCCACCAAGCT  
CCGGCTGGCCACGGGGCTCCGCAACGTCCCCTCCATCCAGAGCCGGGGCTGTTCTGGG  
35 GCGATCGCCGGCTTCATCGAGGGGGGCTGGACCGGGATGGTGGACGGCTGGTACGGGT  
ACCACCACCAGAACGAGCAGGGCTCCGGGTACGCGGCCGACCTCAAGTCGACGCAGA  
ACGCGATCGACAAGATCACCACAAGGTCAACAGCGTGATCGAGAAGATGAACACGCA  
GTTACCCGCCGTCCGCAAGGAGTTCAACCACCTGGAGAAGCGCATCGAGAACCTCAAC  
AAGAAGGTGGACGACGGGTTCTTGACATCTGGACGTACAACGCGGAGCTCCTGGTCC  
40 TCCTGGAGAACGAGCGGACCCTCGACTACCACGACTCCAACGTGAAGAACCTGTACGA  
GAAGGTCCGCAACCAGCTCAAGAACAACGCCAAGGAGATCGGCAACGGGTGCTTCGA  
GTTCTACCACAAGTGCGACAACACGTGCATGGAGTCGGTGAAGAACGGCACCTACGAC  
TACCCGAAGTACAGCGAGGAGGCGAAGCTGAACCGGGAGAAGATCGACGGGGTCAAG  
CTCGACTCCACGCGCATCTACCAGATCCTGGCCATCTACTCGACCGTGGCGAGCTCCCT  
45 CGTCCTGGTGGTCTCGCTCGGCGCCATCAGCTTCTGGATGTGCTCCAACGGGTGCTGCTG

AGTGCCGGATCTGCATCTGA

SEQ ID NO: 87

&gt;A 型流感病毒 H3N2 (A/Cambodia/e0826360/2020) HA CDS

ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGAAGATCCC  
5 CGGCAACGACAACAGCACCGCCACCCCTGTGCCTGGGCGCCACACGCCGTGCCCAACGGC  
ACCATCGTGAAGACCATCACCAACGACCGGATCGAGGTGACCAACGCCACCGAGCTGG  
TGCAGAACTCCAGCATCGGCGAGATCTGCGACTCCCCCACCAGATCCTGGACGGCGG  
CAACTGCACCCTGATCGACGCCCTGCTGGGCGACCCCCAGTGCGACGGGTTCAGAAC  
AAGGAGTGGGACCTGTTCGTGGAGCGGAGCCGCGCCAACTCCAAGTGTACCCCTACG  
10 ACGTGCCCCGACTACGCCAGCCTGCGGTCCCTGGTGGCCAGCTCCGGCACCCCTGGAGTT  
CAAGAACGAGAGCTTCAACTGGACCGGGGTCAAGCAGAACGGCACCTCCAGCGCCTG  
CATCCGCGGGTCCAGCTCCAGCTTCTTCTCCCGGCTGAACTGGCTCACCCACCTGAACT  
ACACCTACCCCGCCCTCAACGTGACCATGCCCAACAACGAGCAGTTCGACAAGCTGTA  
CATCTGGGGCGTCCACCACCCAGCACCGACAAGGACCAGATCTCCCTCTTCGCCAG  
15 CCCAGCGGGCGCATCACCGTGTCCACGAAGCGGAGCCAGCAGGCCGTATCCCCAACA  
TCGGTCCCCGCGCGGATCCGCGACATCCCCAGCCGGATCTCCATCTACTGGACGATC  
GTGAAGCCGGGGACATCCTGCTCATCAACAGCACCGGCAACCTGATCGCCCCCGCG  
GGTACTTCAAGATCCGGTCCGGCAAGAGCTCGATCATGCGCTCCGACGCCCCGATCGGG  
AAGTGCAAGAGCGAGTGCATACCCCCAACGGCTCGATCCCGAACGACAAGCCCTTCC  
20 AGAACGTCAACCGGATCACGTACGGGGCTGCCGCGCTACGTGAAGCAGTCCACGCT  
CAAGCTGGCCACCGCATGCGGAACGTCCCCGAGAAGCAGACCCGCGGGATCTTCGGC  
GCCATCGCGGGGTTCATCGAGAACGGCTGGGAGGGGATGGTGGACGGCTGGTACGGGT  
TCCGGCACCAAGAACAGCGAGGGCCGCGGGCAGGCCGCGGACCTCAAGTCGACGCAGG  
CCGCGATCGACAGATCAACGGCAAGCTGAACCGGCTCATCGGAAGACGAACGAGA  
25 AGTTCCACCAGATCGAGAAGGAGTTCTCCGAGGTTCGAGGGCCGCGTGCAGGACCTGGA  
GAAGTACGTTCGAGGACACCAAGATCGACCTCTGGAGCTACAACGCCGAGCTGCTCGTG  
GCGCTGGAGAACCAGCACACCATCGACCTCACGGACTCGGAGATGAACAAGCTGTTTCG  
AGAAGACGAAGAAGCAGCTCCGGGAGAACGCCGAGGACATGGGGAACGGCTGCTTCA  
AGATCTACCACAAGTGCGACAACGCGTGCATCGGGTCCATCCGCAACGAGACCTACGA  
30 CCACAACGTCTACCGGGACGAGGCCCTGAACAACCGCTTCCAGATCAAGGGCGTGGAG  
CTCAAGAGCGGGTACAAGGACTGGATCCTGTGGATCTCGTTCGCGATGTCTTGCTTCT  
CCTGTGCATCGCCCTCCTGGGCTTCATCATGTGGGCGTGCCAGAAGGGGAACATCCGGT  
GCAACATCTGCATCTGA

SEQ ID NO: 88

35 &gt;B 型流感病毒 Victoria 谱系 (B/Washington/02/2019) HA CDS

ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
CGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCCACCCAGGGCGAGGTG  
AACGTGACCGGCGTGATCCCCCTGACCACACCCCCACCAAGTCCCACTTCGCCAAC  
TGAAGGGCACCGAGACCCGGGGCAAGCTGTGCCCAAGTGCCGTAAGTGCACCGACCT  
40 GGACGTGGCCCTGGGCGGCCCAAGTGACCGGCAAGATCCCCAGCGCCCGGGTGTCC  
ATCCTGCACGAGGTGCGCCCCGTGACCAGCGGCTGCTTCCCATCATGCACGACCGCAC  
CAAGATCCGGCAGCTGCCAACCTGCTGCGGGGGTACGAGCACGTCCGCTGTCCACC  
CACAACGTGATCAACGCCGAGGACGCCCCGGCGGCCCTACGAGATCGGGACCGAGCG  
GCTCCTGCCCCAACATCACCAACGGGAACGGCTTCTTCGCCACCATGGCCTGGGCGGTC  
45 CCCAAGAACAAGACCGCCACCAACCCCTGACCATCGAGGTGCCGTACATCTGCACCG

AGGGGGAGGACCAGATCACCGTCTGGGGCTTCCACAGCGACAACGAGACGCAGATGG  
CCAAGCTGTACGGGGACTCCAAGCCCCAGAAGTTCACGAGCTCCGCCAACGGCGTGAC  
CACCCTACTAGTCAGCCAGATCGGGGGCTTCCCGAACCAGACGGAGGACGGGGGCTG  
CCCCAGTCCGGGCGCATCGTGGTCGACTACATGGTGCAGAAGTCCGGCAAGACGGGGA  
5 CCATCACGTACCAGCGGGGCATCCTGCTGCCGCAGAAGGTCTGGTGCGCCTCCGGGCG  
CAGCAAGGTGATCAAGGGCTCCCTGCCCCATCGGGGAGGCCGACTGCCTGCACGAG  
AAGTACGGCGGGCTCAACAAGAGCAAGCCGTACTACACGGCGAGCACGCCAAGGCC  
ATCGGGAAGTGGCCATCTGGGTCAAGACCCGCTGAAGCTCGCCAACGGCACGAAGT  
ACCGGCCCCCGGCCAAGCTGCTCAAGGAGCGCGGGTTCTTCGGCGCCATCGCCGGGTT  
10 CCTGGAGGGCGGGTGGGAGGGCATGATCGCCGGGTGGCACGGCTACACGTCCCACGGG  
GCCACGGCGTGCGGTGCGCCGGACCTCAAGAGCACCCAGGAGGCCATCAACAAG  
ATCAGGAAGAACCTGAACTCCCTCAGCGAGCTGGAGGTGAAGAACCTCCAGCGCCTGT  
CCGGGGCGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGACGAGAAGGTGACG  
ACCTGCGGGCCGACACCATCAGCTCGCAGATCGAGCTCGCGGTGCTGCTCTCCAACGA  
15 GGGCATCATCAACAGCGAGGACGAGCACCTGCTCGCCCTGGAGCGGAAGCTCAAGAA  
GATGCTGGGGCCCTCGGCGGTGAGATCGGCAACGGGTGCTTCGAGACCAAGCACAAAG  
TGCAACCAGACGTGCCTCGACCGCATCGCCGCGGGCACGTTGACGCCGGGGAGTTCT  
CCCTGCCGACCTTCGACAGCCTCAACATCAGGCGGCCTCGCTGAACGACGACGGCCT  
CGACAACCACACCATCCTGCTCTACTACTCCACCGCGGCCAGCTCGCTGGCGGTGACGC  
20 TCATGATCGCCATCTTCGTCGTGTACATGGTCTCCCGCGACAACGTGAGCTGCTCGATCT  
GCCTGTGA

SEQ ID NO: 89

>B 型流感病毒/Yamagata 谱系 (B/Phuket/3073/2013) HA CDS  
ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
25 CGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCCACCCAGGGCGAGGTG  
AACGTGACCGGCGTGATCCCCCTGACCACCACCCCCACCAAGTCTACTTCGCCAACCT  
GAAGGGCACCCGGACCCGCGGCAAGCTGTGCCCCGACTGCCTGAACTGCACCGACCTG  
GACGTGGCCCTGGGCCGGCCCATGTGCGTGGGCACCACCCCCAGCGCCAAGGCCTCCA  
TCCTGCACGAGGTGCGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACC  
30 AAGATCCGGCAGCTGCCAACCTGCTGCGGGGTACGAGAAGATCCGCCTGTCCACCC  
AGAACGTATCGACGCCGAGAAGGCCCCCGCGGGGCCCTACCGGCTGGGCACCAGCG  
GGTCTTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCACCATGGCCTGGGCCGTG  
CCCAAGGACAACCTACAAGAACGCCACCAACCCCTGACCGTCGAGGTGCCCTACATCT  
GCACCGAGGGGGAGGACCAGATCACGGTCTGGGGCTTCCACTCCGACAACAAGACCC  
35 AGATGAAGTCCCTGTACGGGGACTCCAACCCGAGAAGTTCACGAGCTCCGCCAACGG  
CGTGACCACGCACTACGTCAGCCAGATCGGGGACTTCCCGGACCAGACCGAGGACGGC  
GGGCTGCCCCAGTCCGGCCGCATCGTGGTCGACTACATGATGCAGAAGCCCGGGAAGA  
CGGGCACCATCGTGTAACAGCGGGGGGTCTGCTGCCGAGAAGGTGTGGTGCGCCAG  
CGGCCGCTCCAAGGTCATCAAGGGGAGCCTGCCCCATCGGCGAGGCCGACTGCCTG  
40 CACGAGGAGTACGGGGGCTCAACAAGTCCAAGCCGTACTACACGGGGAAGCACGCC  
AAGGCCATCGGCAACTGCCCAGTCTGGGTGAAGACCCCCCTGAAGCTCGCCAACGGGA  
CGAAGTACCGGCCGCCGCCAAGCTGCTCAAGGAGCGCGGCTTCTTCGGGGCCATCGC  
CGGCTTCTGGAGGGGGGCTGGGAGGGGATGATCGCCGGCTGGCACGGGTACACCAGC  
CACGGCGCCACGGGGTCGCGGTGGCCGCGGACCTCAAGTCCACGCAGGAGGCCATCA  
45 ACAAGATACCAAGAACCTGAACAGCCTCTCCGAGCTGGAGGTCAAGAACCTCCAGCG

CCTGAGCGGCGGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGACGAGAAGGTG  
GACGACCTGCGGGCCGACACGATCTCGTCCCAGATCGAGCTCGCGGTCTGCTCAGCA  
ACGAGGGGATCATCAACTCGGAGGACGAGCACCTGCTCGCCCTGGAGCGGAAGCTCAA  
GAAGATGCTGGGCCCCTCCGCGGTGGACATCGGGAACGGCTGCTTCGAGACCAAGCAC  
5 AAGTGCAACCAGACGTGCCTCGACCCGATCGCCGCGGGGACCTTCAACGCCGGCGAGT  
TCAGCCTGCCCACGTTTCGACTCGCTCAACATCACCGCGGCCTCCCTGAACGACGACGG  
GCTCGACAACCACACGATCCTGCTCTACTACAGCACCGCGGCCTCGTCCCTGGCGGTCA  
CGTCTATGCTGGCCATCTTCATCGTGTACATGGTCAGCCGCGACAACGTGTCGTGCTCCA  
TCTGCCTCTGA  
10  
SEQ ID NO: 90  
>5'UTR-Moderna  
GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGACCCCGGCGCCGCCACC  
SEQ ID NO: 91  
>3'UTR-Moderna  
15 GCTGGAGCCTCGGTGGCCTAGCTTCTTGCCCCCTGGGCCTCCCCCAGCCCCCTCCTCCC  
CTTCCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC  
SEQ ID NO: 92  
>A 型流感病毒 H1N1 (A/Wisconsin/588/2019) HA Moderna  
ATGAAGGCCATCCTGGTTCGTGATGCTGTACACCTTCACCACCGCCAACGCCGACACCCT  
20 GTGCATCGGCTACCACGCCAACAACAGCACCGACACCGTGGACACCGTGCTGGAGAAG  
AACGTGACCGTGACCCACAGCGTGAACCTGCTGGAGGACAAGCACAACGGCAAGCTG  
TGCAAGCTGAGGGGAGTGGCACCCCTGCACCTGGGCAAGTGCAACATCGCCGGCTGGA  
TCCTGGGCAACCCCGAGTGCGAGAGCCTGAGCACAGCCCGGAGCTGGAGCTACATCGT  
GGAGACCAGCAACAGCGACAACGGCACCTGTTACCCCGGCGACTTCATCAACTACGAG  
25 GAGCTGCGGGAGCAGCTGAGCAGCGTGAGCAGCTTCGAGCGGTTTCGAGATCTTCCCCA  
AGACCAGCAGCTGGCCCAACCACGACAGCGACAACGGCGTGACAGCAGCCTGTCCAC  
ACGCCGGAGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAAGA  
GCTACCCCAAGATCAACCAGACCTACATCAACGACAAGGGCAAGGAGGTGCTGGTGT  
30 GTGGGGCATCCACCACCCACCTACCATCGCCGACCAGCAGAGCCTGTACCAGAACGCC  
GACGCCTACGTGTTTCGTGGGCACCAGCCGGTACAGCAAGAAGTTCAAGCCAGAGATCG  
CCACCCGGCCCAAGGTGAGAGACCAGGAGGGCCGGATGAACTACTACTGGACCCTGGT  
GGAGCCCGGAGACAAGATTACCTTCGAGGGCCACCGGCAACCTGGTGGCCCCCTCGGTAC  
GCCTTACCATGGAACGGGACGCTGGCAGCGGCATCATCATCAGCGACACTCCCGTGCA  
CGACTGCAACACCACCTGCCAGACTCCCGAGGGCGCTATCAACACCAGCCTGCCCTTC  
35 CAGAACGTGCACCCCATCACCATCGGCAAGTGCCCCAAGTACGTAAAGAGCACCAAAT  
TGCGGCTGGCCACCGGACTCAGGAACGTGCCCAGCATCCAAAGCCGGGGCCTGTTTGG  
CGCAATCGCCGGCTTCATCGAGGGCGGCTGGACTGGCATGGTGGACGGCTGGTACGGC  
TACCACCACCAGAACGAACAGGGGAGCGGTACGCAGCTGACCTGAAGAGCACCCAG  
AACGCCATCGACAAGATCACCAACAAGGTGAACAGCGTGATCGAGAAGATGAACACCC  
40 AGTTCACCGCCGTGGGCAAGGAGTTCAACCACCTGGAGAAGCGGATCGAGAACCTGA  
ACAAGAAGGTGGACGACGGCTTCCTGGACATCTGGACCTACAACGCCGAGCTGCTGGT  
TCTGCTGGAGAACGAGCGGACCCTGGACTATCACGACAGCAACGTGAAGAACCTGTAC  
GAGAAGGTGCGGAACCAGCTGAAGAACAACGCCAAGGAGATCGGCAACGGCTGCTTC  
GAGTTCTACCACAAGTGCGACAACACCTGCATGGAGAGCGTGAAGAACGGCACCTACG  
45 ACTACCCCAAGTACAGCGAGGAGGCCAAGCTGAACCGGGAGAAGATCGACGGCGTGA

AGCTGGACAGCACCCGGATCTACCAGATCCTGGCCATCTACAGCACCGTGGCCAGCAG  
CCTGGTGCTGGTGGTGAGCCTGGGCGCCATCAGCTTCTGGATGTGCAGCAACGGCAGC  
CTGCAGTGCCGGATCTGCATC

SEQ ID NO: 93

5 >A 型流感病毒 H3N2 (A/Cambodia/e0826360/2020) HA Moderna  
ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGAAGATCCC  
CGGCAACGATAACAGCACCGCCACCCTGTGTCTGGGACACCACGCCGTGCCAACGGC  
ACCATCGTGAAGACTATCACCAACGACCGGATCGAGGTGACCAACGCCACCGAGCTGG  
TGCAGAACAGCAGCATCGGCGAGATCTGCGACAGCCCTCACCAGATCCTGGACGGCGG  
10 CAACTGCACCCTGATCGACGCACTGCTGGGCGACCCCTCAGTGCGACGGCTTTCAGAAC  
AAGGAGTGGGACCTGTTCTGTGGAGAGATCGCGGGCCAACAGCAACTGCTACCCCTACG  
ACGTCCCCGACTACGCAAGCCTGAGAAGCCTCGTGGCCTCAAGCGGCACCCTGGAGTT  
CAAGAACGAGAGCTTCAACTGGACCGGCGTGAAGCAGAACGGCACCTCAAGCGCCTG  
CATCCGGGGCTCCAGCAGCAGCTTCTTCTCACGGCTGAACTGGCTGACCCACCTGAACT  
15 ACACCTACCCCGCCCTGAACGTGACCATGCCCCAACACGAGCAGTTCGACAAGCTGTA  
CATCTGGGGAGTGCACCATCCCAGCACCGACAAGGACCAGATTAGCCTGTTGCCCCAG  
CCCAGCGGCCGGATCACCGTGAGCACCAAGCGGAGCCAGCAGGCCGTGATCCCCAACA  
TCGGCTCTCGGCCCAGAATCCGGGACATCCCCAGCCGGATCAGCATCTACTGGACCATT  
GTGAAGCCCGGCGACATCCTGCTGATCAACTCCACCGGCAACCTGATCGCCCCCTCGGGG  
20 CTATTTCAAGATCCGGAGCGGCAAGAGCAGCATCATGCGGAGCGACGCCCTATCGGCA  
AGTGCAAGAGCGAGTGCATCACACCCAACGGAAGCATCCCCAACGACAAGCCCTTCCA  
GAACGTGAACCGGATAACCTACGGCGCCTGCCCTAGATACGTGAAGCAGAGCACCCCTG  
AAGCTGGCCACCGGCATGCGGAACGTGCCCGAGAAGCAGACTCGGGGCATCTTCGGCG  
CCATCGCCGGCTTCATCGAGAACGGCTGGGAGGGCATGGTGGACGGCTGGTACGGCTT  
25 CCGGCACCAGAACTCTGAGGGCAGAGGACAGGCCGACAGCCTGAAGAGCACCCAGGC  
CGCCATCGACCAGATCAACGGCAAGCTGAACCGGCTGATCGGCAAGACCAACGAGAAG  
TTCCACCAGATCGAGAAGGAGTTCAGCGAGGTGGAGGGCAGGGTACAGGACCTGGAG  
AAGTACGTGGAGGACACCAAGATCGACCTGTGGAGCTACAACGCCGAGCTGCTGGTAG  
CCCTGGAGAACCAGCACACCATCGACCTGACCGACAGCGAGATGAACAAGCTGTTCTGA  
30 GAAGACCAAGAAGCAGCTGCGGGAGAACGCCGAGGACATGGGCAACGGCTGCTTCAA  
GATCTACCACAAGTGCGACAACGCCTGCATCGGCAGCATCCGGAACGAGACCTACGAC  
CACAACGTGTACCGGGACGAGGCCCTGAACAACCGGTTCCAGATCAAGGGCGTGGAGC  
TGAAGAGCGGCTACAAGGACTGGATCCTGTGGATCAGCTTCGCCATGTCTGCTTCCTG  
CTGTGCATCGCCCTGCTGGGTTTCATCATGTGGGCCTGCCAGAAGGGCAACATCCGGTG  
35 CAACATCTGCATC

SEQ ID NO: 94

>B 型流感病毒 Victoria 谱系 (B/Washington/02/2019) HA CDS Moderna  
ATGAAGGCCATCATCGTGCTGTTAATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
CGGCATCACCTCTAGCAACAGCCCTCACGTGGTGAAGACCGCCACACAGGGCGAGGTG  
40 AACGTGACCGGCGTGATTCCCCTGACCACCAACCCTACCAAGAGCCACTTCGCCAACCT  
GAAGGGAACCGAGACCCGGGGCAAGCTGTGTCCCAAGTGCCTGAACTGCACCGACCT  
GGACGTGGCCCTGGGCAGACCCAAGTGACCGGCAAGATCCCAGCGCCCGGGTGTCT  
ATCCTGCACGAAGTGCGGCCCGTGACTAGCGGCTGCTTCCCCATCATGCACGACCGGAC  
CAAGATCCGGCAGCTGCCAACCTGCTGCGGGGCTACGAGCACGTGCGGCTGAGCACC  
45 CACAACGTGATCAACGCCGAAGACGCACCCGGGAGACCATAACGAGATCGGCACCGAGC

GCTCTTGCCCCAACATCACCAACGGCAACGGCTTCTTCGCTACCATGGCCTGGGCGGTG  
CCAAAGAACAAGACTGCCACCAACCTCTGACCATCGAGGTGCCCTACATCTGCACCG  
AGGGCGAGGACCAGATCACCGTGTGGGGCTTCCACAGCGACAGCGAGACCCAGATGG  
CCAAGCTGTACGGCGACAGCAAGCCCCAGAAGTTCACCAGCAGCGCCAACGGCGTGA  
5 CCACCCACTACGTGAGCCAGATCGGCGGCTTCCCCAACAGACCGAGGACGGCGGCTT  
ACCCAGAGCGGCCGGATCGTGGTGGACTACATGGTGCAGAAGAGCGGCAAGACCGG  
CACCATCACCTACCAGCGGGGCATCCTGCTGCCACAGAAGGTGTGGTGCGCCTCAGGG  
CGGTCAAAGGTGATCAAGGGCAGCCTGCCACTGATTGGCGAGGCCGACTGCCTGCACG  
AGAAGTACGGCGGCCTGAACAAGAGCAAGCCCTACTACACCGGCGAGCACGCCAAGG  
10 CAATCGGCAACTGCCCCATCTGGGTGAAGACACCCCTGAAGCTGGCCAACGGCACCAA  
GTACCGGCCACCCGCCAACTGCTGAAGGAGCGGGGCTTCTTCGGCGCCATTGCCGGC  
TTCTCGAAGGCGGTTGGGAGGGCATGATCGCCGGCTGGCACGGCTACACTAGCCACG  
GCGCACACGGAGTAGCAGTGGCCGCCGACCTGAAGAGCACCCAGGAGGCCATCAACA  
AGATCACCAAGAACCTGAACAGCCTGAGCGAGCTGGAGGTGAAGAATCTGCAGCGGCT  
15 GTCTGGCGCTATGGACGAGCTGCACAACGAGATCCTGGAGCTGGACGAGAAGGTGGAC  
GACTTACGGGCGGACACCATCAGCAGCCAGATCGAGCTGGCCGTGCTGCTCTCCAACG  
AGGGCATCATCAACAGCGAGGACGAGCACCTGCTCGCCCTGGAGCGGAAGCTGAAGA  
AGATGCTGGGCGCTTCTGCCGTGGAGATCGGTAACGGCTGCTTCGAGACCAAGCACAA  
GTGCAACCAGACCTGCCTGGATCGGATCGCAGCCGGCACCTTTGACGCGGGGAGTTC  
20 AGCCTGCCCACCTTCGACAGCCTGAACATCACCGCCGCCAGCCTGAACGACGACGGCC  
TGGACAACCACACCATCCTGCTGTACTACTCTACAGCCGCTAGCAGCCTGGCCGTGACC  
CTGATGATCGCCATCTTCGTGGTGTACATGGTGAAGCCGGGACAACGTGAGCTGCAGCAT  
CTGCCTG

SEQ ID NO: 95

25 >B 型流感病毒 Yamagata 谱系 (B/Phuket/3073/2013) HA CDS Moderna  
ATGAAGGCCATCATCGTGCTACTGATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
CGGCATCACCAGCAGCAACAGCCCGCACGTGGTGAAGACCGCCACCCAAGGCGAGGT  
GAACGTGACCGGCGTGATCCCCTGACCACCACTCCCACCAAGAGCTACTTCGCCAAC  
CTGAAGGGCACACGGACTCGGGGCAAGCTGTGCCCCGACTGCCTGAACTGCACCGACC  
30 TGGACGTGGCCCTGGGCAGACCCATGTGCGTGGGCACCACCCCTTCTGCCAAGGCCAG  
CATCTGCACGAGGTGAGACCCGTGACCAGCGGGTGCTTCCCCATCATGCACGACCGG  
ACCAAGATCCGGCAGCTGCCAACCTGCTGCGGGGCTACGAGAAGATCCGGCTGAGCA  
CCCAGAACGTGATCGACGCCGAGAAGGCCCTGGAGGTCCCTACCGGCTGGGCACCA  
CGGAAGCTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCACCATGGCCTGGGCT  
35 GTGCCCCAAGGACAACCTACAAGAAGCCACCAATCCCCTGACCGTGGAGGTGCCCTACA  
TCTGCACCGAGGGCGAGGACCAGATCACCGTGTGGGGCTTCCACAGCGACAACAAGAC  
CCAGATGAAGAGCCTGTACGGCGACAGCAATCCCCAGAAGTTCACAAGCAGCGCCAAC  
GGCGTGACCACCCACTACGTGAGCCAGATCGGCGACTTCCCCGACCAGACCGAGGACG  
GAGGGCTGCCTCAGAGTGGCCGGATCGTGGTGGACTACATGATGCAGAAGCCCGGCAA  
40 GACCGGCACCATCGTGTAACAGCGGGGCGTGCTGTTGCCCTCAGAAAGTTTGGTGTGCC  
AGCGGCAGGAGCAAGGTGATCAAGGGCAGCCTGCCCCTGATCGGCGAGGCAGACTGC  
CTCCACGAGGAGTACGGCGGCCTGAACAAGAGCAAGCCCTACTACACCGGCAAGCACG  
CCAAGGCCATCGGCAACTGCCCCATCTGGGTGAAGACCCCTCTGAAGCTGGCCAACGG  
CACCAAGTACCGGCCACCAGCCAAGCTGCTGAAGGAGCGGGGCTTCTTTGGCGCCATT  
45 GCCGGCTTCTCGAGGGAGGCTGGGAGGGCATGATCGCCGGCTGGCACGGCTACACAA

GCCACGGCGCACACGGAGTGGCTGTGGCTGCCGACCTGAAGAGCACCCAGGAGGCCA  
TCAACAAGATCACCAAGAACCTGAACAGCCTGAGCGAGCTGGAGGTGAAGAACCTGC  
AGCGGCTGTCAGGCGCCATGGACGAGCTGCACAACGAGATCCTGGAGCTGGACGAGA  
AGGTGGACGACCTGCGTGCCGACACCATCAGCAGCCAGATCGAGCTGGCCGTGCTGCT  
5 GAGCAACGAGGGCATCATCAACAGCGAGGACGAGCACCTGCTGGCCCTGGAGCGGAA  
ACTGAAGAAGATGCTGGGACCCTCTGCCGTGGACATCGGCAACGGCTGCTTCGAGACC  
AAGCACAAGTGCAACCAGACCTGCCTGGATCGCCGCCGGAACCTTCAACGCCG  
GCGAGTTCAGCCTGCCACCTTCGACAGCCTGAACATCACCGCCGCCAGCCTGAACGA  
CGACGGCCTGGACAACCACACCATCCTGCTGTACTACAGCACTGCCGCCTCAAGCCTG  
10 GCCGTGACCCTGATGCTGGCCATCTTCATCGTGTACATGGTGAGCCGGGACAACGTGAG  
CTGCAGCATCTGCCTG

SEQ ID NO: 96

&gt; SARS-COV-2 BA.4 刺突蛋白 CDS

ATGTTCTGTTCCTGGTGCTGCTGCCCCCTGGTGAGCTCTCAGTGCGTGAACCTGATCAC  
15 AAGAACACAGTCGTACACCAACAGCTTCACAAGAGGCGTGTAACCCCGACAAGGTG  
TTCAGAAGCAGCGTGCTGCACTCCACCCAAGACCTGTTTCTGCCCTTCTTCAGCAACGT  
GACCTGGTTCCACGCCATCAGCGGCACCAACGGCACCAAGAGATTGACAACCCCGTG  
CTGCCCTTCAACGACGGCGTGTAACCTCGCTAGCACCGAAAAGAGCAACATCATCAGAG  
GCTGGATCTTCGGCACCAACCTAGACTCCAAGACACAAAGCCTGCTGATCGTTAACAAC  
20 GCGACCAACGTGGTGATCAAGGTGTGCGAGTTTCAGTTCTGCAACGACCCCTTCCTGG  
ACGTGTACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGAGTCTACAG  
CAGCGCCAACAACCTGCACCTTCGAGTACGTGTCTCAGCCCTTCCTGATGGACCTGGAGG  
GCAAGCAAGGCAATTTCAAGAACCTAAGAGAGTTCGTGTTCAAGAACATCGACGGCTA  
CTTCAAGATCTACAGCAAGCACACCCCATCAACCTGGGGAGAGACCTGCCCCAAGGC  
25 TTCAGCGCCCTGGAGCCCCCTGGTGGACCTGCCCATCGGCATCAACATCACAAGATTTCA  
GACCCTGCTGGCCCTGCACCGTAGCTACCTAACCCCGGCGATAGTTCTCTGGCTGGA  
CGGCCGGCGCTGCCGCCTACTACGTGGGTACCTGCAGCCTAGAACCTTCCTGCTGAAG  
TACAACGAGAACGGCACTATCACTGATGCCGTGGACTGTGCCCTGGACCCCTGAGCG  
AGACCAAGTGACCCCTCAAGAGCTTCACCGTGGAGAAGGGCATCTATCAGACAAGCAA  
30 CTTCAAGAGTGACGCCCACCGAGAGCATCGTGAGATTCCCCAACATCACCAACCTGTGCC  
CCTTCGACGAGGTGTTCAACGCCACAAGATTCGCTAGCGGTGACGCGTGGAACAGAAA  
GAGAATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAACCTCGCTCCCTTCTTCG  
CCTTCAAGTGCTATGGCGTCAGCCCCACCAAGCTGAACGACCTGTGCTTACCAACGTG  
TACGCCGACAGCTTCGTGATCAGAGGCAACGAGGTGAGCCAGATTGCCCCCGGGCAAA  
35 CCGGCAACATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACCGGTGCGTGATC  
GCGTGGAACAGCAATAAGCTGGATTCCAAGGTGGGCGGCAACTACAACCTACCGGTACA  
GACTGTTCAAGAAAGAGCAACCTGAAGCCCTTCGAGAGAGACATCAGCACCGAGATCTA  
CCAAGCCGGCAACAAGCCCTGCAACGGCGTGGCGGGCGTCAACTGCTACTTCCCCCTG  
CAGAGCTACGGCTTTAGGCCCCACCTACGGCGTGGGCCATCAGCCCTACAGAGTGGTGGT  
40 GCTGAGCTTCGAGCTGCTGCACGCCCCCGCAACCGTGTGCGGCCCCAAAAAGAGCACC  
AACCTGGTGAAGAACAAGTGCGTGAACCTTCAACTTTAATGGCCTTAAGGGCACGGGCG  
TGCTGACCGAGAGCAACAAGAAGTTCCTGCCATTTCAAGCAATTCGGCAGAGACATCGC  
CGACACCAACCGACGCGGTACAGAGACCTCAGACCCTGGAGATCCTGGACATCACCCCC  
TGTAGCTTCGGCGGGCGTGAGCGTGATCACCCCGGCACTAACACAAGCAACCAAGTGG  
45 CCGTGCTGTACCAAGGCGTGAACCTGCACCGAGGTGCCCGTGGCCATCCACGCCGATCA

GCTGACCCCCACCTGGAGAGTATACAGCACCGGAAGCAACGTGTTTCAGACAAGAGCC  
GGCTGCCTGATCGGCGCCGAGTACGTGAACAACAGCTACGAGTGCGACATCCCCATCG  
GCGCCGGCATCTGCGCTAGCTATCAGACACAGACCAAGAGCCACCGGAGAGCTAGAAG  
CGTGGCTTCTCAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAACAGCGTG  
5 GCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCACCGA  
GATCCTGCCCCGTAAGTATGACCAAGACAAGCGTGGACTGCACCATGTACATCTGCGGCG  
ACAGCACCGAGTGCAGCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACACAGCTGAA  
GAGAGCCCTGACCGGCATCGCCGTGGAGCAAGACAAGAACACCCAAGAGGTGTTTCGC  
CCAAGTGAAGCAGATCTACAAGACCCCCCCCCATCAAGTACTTCGGCGGGCTTCAACTTCA  
10 GCCAAATCCTGCCCCGACCCTAGCAAGCCTAGCAAGCGGAGCTTCATCGAGGACCTGCT  
GTTCAACAAGGTGACCCTGGCCGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTA  
GGCGACATCGCAGCTAGAGACCTGATCTGCGCTCAGAAGTTCAATGGACTCACCGTGCT  
CCCCCCCCCTGCTGACCGACGAGATGATCGCTCAGTACACCTCTGCCCTGCTGGCGGGAA  
CCATAACAAGCGGTTGGACATTTGGCGCCGGGGCCGCACTGCAAATCCCCCTTCGCCATG  
15 CAGATGGCCTACAGATTCAACGGCATCGGCGTGACACAGAACGTGCTGTACGAGAATC  
AGAAGCTGATCGCCAATCAGTTCAACAGCGCCATCGGCAAGATCCAAGACAGCCTGAG  
CAGCACCGCTAGCGCCCTGGGCAAGCTGCAAGACGTGGTGAATCACAACGCCCAAGCC  
CTGAACACCCTGGTGAAGCAGCTGAGCAGCAAGTTTCGGCGCCATCAGCAGCGTATTGA  
ATGACATCCTGTCTAGACTGGACCCCCCGAGGCCGAGGTGCAGATCGACAGACTGATC  
20 ACCGGCAGACTGCAGAGCCTGCAGACCTACGTGACACAGCAGCTGATCAGAGCCGCCG  
AGATCAGAGCTAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTGGGGCA  
GAGCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGCTTCCCTCAGAGC  
GCCCCCACGGCGTGGTGTTCCTGCACGTGACCTACGTGCCCGCCCAAGAGAAGAACT  
TCACCACCGCTCCCGCCATCTGCCACGACGGCAAGGCCCACTTCCCTAGAGAGGGCGT  
25 GTTCGTGAGCAACGGCACCCACTGGTTTCGTGACACAGAGAACTTCTACGAGCCTCAG  
ATCATCACACCACGACAACACCTTCGTGAGCGGCAACTGCGACGTGGTGATCGGCATAGT  
GAACAACACCGTATACGACCCCCTGCAGCCCAGCTGGACAGCTTCAAGGAGGAGCTG  
GACAAGTACTTCAAAAACCATACTAGCCCCGACGTGGACCTGGGAGACATCTCTGGCAT  
CAACGCTAGCGTGGTGAACATTGAGAAGGAGATCGACCGGCTGAATGAGGTGGCCAAG  
30 AACCTGAACGAGAGCCTGATCGACCTGCAAGAGCTGGGCAAGTACGAGCAGTACATCA  
AGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTGATCGCCATCGTGATGGTG  
ACCATCATGCTGTGCTGCATGACAAGCTGTTGACAGCTGCCTGAAAGGTTGCTGCAGCTG  
TGGCTCGTGCTGCAAGTTCGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAA  
GCTGCACTACACCTGATGA

35

SEQ ID NO: 97

&gt;A 型流感病毒 H1N1 HA0 突变体氨基酸序列

MKAILVVMlyTFTTANADTLcIGYHANNSTDTVDTVLEKNVTVTHSVNlLEDKHNGKLCK  
LRGVAPLHLGKcNIAGWILGNPECESLSTARSWSYIVETSNSDNGTCYPGDFINYEELREQL  
SSVSSFERFEIFPKTSSWPNHDSNGVTAACPHAGAKSFYKNLIWLVKKGKSYPKINQTYIN  
40 DKGKEVLVLWGIHHPPTIADQQSLYQNADAYVFGVTSRYSKKFKPEIATRPKVRDQEGRM  
NYYWTLVEPGDKITFEATGNLVAPRYAFTMERDAGSGIIISDTPVHDCNTTCQTPEGAINSL  
PFQNVHPITIGKCPKYVKSTKLRLATGLRNVPSIQSQGLFGAIAGFIEGGWTGMVDGWYGY  
HHQNEQGSGYAADLKSTQNAIDKITNKVNSVIEKMNTQFTAVGKEFNHLEKRIENLNKKV  
DDGFLDIWTYNAELLVLENERLTLDYHDSNVKNLYEKVRNQLKNNAKEIGNGCFEFYHKC  
45 DNTCMESVKNGTYDYPKYSEEAKLNREKIDGVKLDSTRIYQILAIYSTVASSLVLVVSLGAI

SFWMCSNGSLQCRICI

SEQ ID NO: 98

&gt;A 型流感病毒 H3N2 HA0 突变体氨基酸序列

MKTIIALSYILCLVFAQKIPGNDNSTATLCLGHHAVPNGTIVKTITNDRIEVTNATELVQNSS  
5 IGEICDSPHQILDGGNCTLDALLGDPQCDGFQNKEDLFFVERSRANSNCYPYDVPDYASL  
RSLVASSGTLEFKNESFNWTGVKQNGTSSACIRGSSSSFFSRLNWLTHLNYTYPALNVTMP  
NNEQFDKLYIWGVHHPSTDQISLFAQPSGRITVSTKRSQQA VIPNIGSRPRIRDIPSRSIY  
WTIVKPGDILLINSTGNLIAPRGYFKIRSGKSSIMRSDAPIGKCKSECITPNGSIPNDKPFQNV  
NRITYGACPRYVKQSTLKLATGMRNVPEKQTQGIFGAIAGFIENGWEGMVDGWYGFRRHQ  
10 NSEGRGQAADLKSTQAAIDQINGKLNRLIGKTNEKFHQIEKEFSEVEGRVQDLEKYVEDTK  
IDLWSYNAELLVALENQHTIDLTDSEMKNLFKTKKQLRENAEDMGNGCFKIYHKCDNA  
CIGSIRNETYDHNVYRDEALNNRFQIKGVELKSGYKDWILWISFAMSCFLLCIALLGFMW  
ACQKGNIRCNICI

SEQ ID NO: 99

15 &gt;B 型流感病毒/Victoria HA0 突变体氨基酸序列

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATATQGEVNVTVGIPLTTTPTKSHFANLKGT  
ETRGLCPKCLNCTDLVALGRPKCTGKIPSARVSILHEVRPVTSGCFPIIMHDRTKIRQLPN  
LLRGYEHVRLSTHNINAEDAPGRPYEIGTSGSCPNTNGNGFFATMAWAVPKNKATATNPLT  
IEVPYICTEGEDQITVWGFHSDNETQMAKLYGDSKPQKFTSSANGVTTHYVSQIGGFNPQT  
20 EDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKVWCASGRSKVIKGSPLIGEADCL  
HEKYGGLNKSPPYTGEGHAKAIGNCPIWVKTPKLKLANGTKYRPPAKLLKEQGFFGAIAGF  
LEGGWEGMIAGWGHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGA  
MDELHNEILELDEKVDDLADTISQIELAVLLSNEGIINSEDEHLLALERKLKMLGPSAVE  
IGNGCFETKHKCNQTCLDRIAAGTFDAGEFSLPTFDSLNTAASLNDDGLDNHTILLYSTA  
25 ASSLAVTLMIAIFVYVMVSRDNVSCSICL

SEQ ID NO: 100

&gt;B 型流感病毒/Yamagata HA0 突变体氨基酸序列

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATATQGEVNVTVGIPLTTTPTKSYFANLKGT  
RTRGKLCPCDCLNCTDLVALGRPMC VGTTPSAKASILHEVRPVTSGCFPIIMHDRTKIRQLPN  
30 LLRGYEKIRLSTQNVIDAEKAPGGPYRLGTSGSCP NATSKIGFFATMAWAVPKDNYKNATNP  
LTVEVPYICTEGEDQITVWGFHSDNKTQMSLYGDSNPQKFTSSANGVTTHYVSQIGDFPD  
QTEDGGLPQSGRIVVDYMMQKPGKTGTIVYQRGVLLPQKVWCASGRSKVIKGSPLIGEA  
DCLHEEYGGNKSPPYTGKHAKAIGNCPIWVKTPKLKLANGTKYRPPAKLLKEQGFFGAI  
AGFLEGGWEGMIAGWGHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRL  
35 SGAMDELHNEILELDEKVDDLADTISQIELAVLLSNEGIINSEDEHLLALERKLKMLGP  
SAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDSLNTAASLNDDGLDNHTILLY  
YSTAASSLAVTLMIAIFIVYVMVSRDNVSCSICL

SEQ ID NO: 101

&gt;A 型流感病毒 H1N1 HA0 突变体 CDS

40 ATGAAGGCCATCCTGGTGGTGATGCTGTACACCTTCACCACCGCCAACGCCGACACCTT  
GTGCATCGGCTACCACGCCAACACAGCACCGACACCGTGGACACCGTGTCTGGAGAAG  
AACGTGACCGTGACCCACAGCGTGAACCTGCTGGAGGACAAGCACACGGAAGCTG  
TGCAAGCTGCGGGGCGTGGCCCCCTGCACCTGGGCAAGTGCAACATCGCCGGCTGGA  
TCCTGGGCAACCCCGAGTGCGAGTCCCTGAGCACCGCCCGCTCCTGGAGCTACATCGT  
45 GGAGACCTCCAACAGCGACAACGGCACCTGCTACCCCGGGGACTTCATCAACTACGAG

GAGCTGCGGGAGCAGCTGTCCAGCGTGTCAGCTTCGAGCGCTTCGAGATCTTCCCCA  
AGACCTCCAGCTGGCCCAACCACGACTCCGACAACGGCGTGACCGCCGCCTGCCCCCA  
CGCCGGGGCCAAGAGCTTCTACAAGAACCTGATCTGGCTGGTGAAGAAGGGCAAGTCC  
TACCCCAAGATCAACCAGACCTACATCAACGACAAGGGGAAGGAGGTCCCTCGTGCTGT  
5 GGGGCATCCACCACCCCCCACCATCGCCGACCAGCAGAGCCTCTACCAGAACGCCGA  
CGCCTACGTCTTCGTGGGGACCTCCCGGTACAGCAAGAAGTTCAAGCCCGAGATCGCC  
ACCCGCCCAAGGTCCGGGACCAGGAGGGCCGCATGAACTACTACTGGACCCTGGTGG  
AGCCGGGGGACAAGATCACCTTCGAGGCCACCGGCAACCTCGTCGCCCCGCGGTACGC  
CTTCACGATGGAGCGCGACGCCGGGTCCGGCATCATCATCAGCGACACCCCCGTGCAC  
10 GACTGCAACACGACCTGCCAGACGCCCAGGGGGGCGATCAACACCTCCCTGCCGTTCC  
AGAACGTCCACCCGATCACGATCGGCAAGTGCCCCAAGTACGTGAAGTCCACCAAGCT  
CCGGCTGGCCACGGGGCTCCGCAACGTCCCCCTCCATCCAGAGCCAGGGCCTGTTCTGGG  
GCGATCGCCGGCTTCATCGAGGGGGGCTGGACCGGGATGGTGGACGGCTGGTACGGGT  
ACCACCACCAGAACGAGCAGGGCTCCGGGTACGCGGCCGACCTCAAGTCGACGCAGA  
15 ACGCGATCGACAAGATCACCAACAAGGTCAACAGCGTGATCGAGAAGATGAACACGCA  
GTTACCCGCCGTCGGCAAGGAGTTCAACCACCTGGAGAAGCGCATCGAGAACCTCAAC  
AAGAAGGTGGACGACGGGTTCTTGACATCTGGACGTACAACGCGGAGCTCCTGGTCC  
TCCTGGAGAACGAGCGGACCCTCGACTACCACGACTCCAACGTGAAGAACCTGTACGA  
GAAGGTCCGCAACCAGCTCAAGAACAACGCCAAGGAGATCGGCAACGGGTGCTTCGA  
20 GTTCTACCACAAGTGCGACAACACGTGCATGGAGTCGGTGAAGAACGGCACCTACGAC  
TACCCGAAGTACAGCGAGGAGGCGAAGCTGAACCGGGAGAAGATCGACGGGGTCAAG  
CTCGACTCCACGCGCATCTACCAGATCCTGGCCATCTACTCGACCGTGGCGAGCTCCCT  
CGTCCTGGTGGTCTCGCTCGGCGCCATCAGCTTCTGGATGTGCTCCAACGGGTGCTGCG  
AGTGCCGGATCTGCATCTGA

25

SEQ ID NO: 102

&gt;A 型流感病毒 H3N2 HA0 突变体 CDS

ATGAAGACCATCATCGCCCTGAGCTACATCCTGTGCCTGGTGTTCGCCCAGAAGATCCC  
CGGCAACGACAACAGCACCGCCACCCTGTGCCTGGGCCACCACGCCGTGCCCAACGGC  
ACCATCGTGAAGACCATCACCAACGACCGGATCGAGGTGACCAACGCCACCGAGCTGG  
30 TGCAGAACTCCAGCATCGGCGAGATCTGCGACTCCCCCACCAGATCCTGGACGGCGG  
CAACTGCACCCTGATCGACGCCCTGTGGGCGACCCCCAGTGGACGGGTTCAGAAC  
AAGGAGTGGGACCTGTTCTGTGGAGCGGAGCCGCGCCAACCTCCAACCTGCTACCCCTACG  
ACGTGCCCCGACTACGCCAGCCTGCGGTCCCTGGTGGCCAGCTCCGGCACCCCTGGAGTT  
CAAGAACGAGAGCTTCAACTGGACCGGGGTCAAGCAGAACGGCACCTCCAGCGCCTG  
35 CATCCGCGGGTCCAGCTCCAGTTCTTCTCCCGGTGAACTGGCTCACCCACCTGAACT  
ACACCTACCCCGCCCTCAACGTGACCATGCCCAACAACGAGCAGTTCGACAAGCTGTA  
CATCTGGGGCGTCCACCACCCCAGCACCGACAAGGACCAGATCTCCCTCTTCGCCCAG  
CCCAGCGGGGCGCATCACCGTGTCCACGAAGCGGAGCCAGCAGGCCGTATCCCCAACA  
TCGGTCCCCGCGCGGATCCGCGACATCCCCAGCCGGATCTCCATCTACTGGACGATC  
40 GTGAAGCCGGGGGACATCTGCTCATCAACAGCACCGGCAACCTGATCGCCCCCGCG  
GGTACTTCAAGATCCGGTCCGGCAAGAGCTCGATCATGCGCTCCGACGCCCCGATCGGG  
AAGTGCAAGAGCGAGTGCATACCCCCAACGGCTCGATCCCGAACGACAAGCCCTTCC  
AGAACGTCAACCGGATCACGTACGGGGCTGCCGCGCTACGTGAAGCAGTCCACGCT  
CAAGCTGGCCACCGGCATGCGGAACGTCCCCGAGAAGCAGACCCAGGGGATCTTCGGC  
45 GCCATCGCGGGGTTTCATCGAGAACGGCTGGGAGGGGATGGTGGACGGCTGGTACGGGT

TCCGGCACCAGAACAGCGAGGGCCGCGGGCAGGCCGCGGACCTCAAGTCGACGCAGG  
CCGCGATCGACCAGATCAACGGCAAGCTGAACCGGCTCATCGGAAGACGAACGAGA  
AGTTCACCAGATCGAGAAGGAGTTCTCCGAGGTCGAGGGCCGCGTGCAGGACCTGGA  
GAAGTACGTCGAGGACACCAAGATCGACCTCTGGAGCTACAACGCCGAGCTGCTCGTG  
5 GCGCTGGAGAACCAGCACACCATCGACCTCACGGACTCGGAGATGAACAAGCTGTTTCG  
AGAAGACGAAGAAGCAGCTCCGGGAGAACGCCGAGGACATGGGGAACGGCTGCTTCA  
AGATCTACCACAAGTGGGACAACGCGTGCATCGGGTCCATCCGCAACGAGACCTACGA  
CCACAACGTCTACCGGGACGAGGCCCTGAACAACCGCTTCCAGATCAAGGGCGTGAG  
CTCAAGAGCGGGTACAAGGACTGGATCCTGTGGATCTCGTTCGCGATGTCCTGCTTCCT  
10 CCTGTGCATCGCCCTCCTGGGCTTCATCATGTGGGCGTGCCAGAAGGGGAACATCCGGT  
GCAACATCTGCATCTGA

SEQ ID NO: 103

&gt;B 型流感病毒/Victoria HA0 突变体 CDS

ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
15 CGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCCACCCAGGGCGAGGTG  
AACGTGACCGGCGTGATCCCCCTGACCACCACCCCCACCAAGTCCCACTTCGCCAACCC  
TGAAGGGCACCGAGACCCGGGGCAAGCTGTGCCCCAAGTGCCCTGAAGTGCACCGACCT  
GGACGTGGCCCTGGGCGGCCCAAGTGACCCGGCAAGATCCCCAGCGCCCGGGTGTCC  
ATCCTGCACGAGGTGCGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCAC  
20 CAAGATCCGGCAGCTGCCAACCTGCTGCGGGGGTACGAGCACGTCCGCCTGTCCACC  
CACAACGTGATCAACGCCGAGGACGCCCCGGCGGGCCCTACGAGATCGGGACCAGCG  
GCTCCTGCCCAACATCACCAACGGGAACGGCTTCTTCGCCACCATGGCCTGGGCGCTC  
CCCAAGAACAAGACCGCCACCAACCCCCCTGACCATCGAGGTGCCGTACATCTGCACCG  
AGGGGGAGGACCAGATCACCGTCTGGGGCTTCCACAGCGACAACGAGACGCAGATGG  
25 CCAAGCTGTACGGGGACTCCAAGCCCCAGAAAGTTCACGAGCTCCGCCAACGGCGTGAC  
CACCCACTACGTACGCCAGATCGGGGGCTTCCCGAACCAGACGGAGGACGGGGGCCTG  
CCCCAGTCCGGGCGCATCGTGGTCGACTACATGGTGCAGAAAGTCCGGCAAGACGGGGA  
CCATCACGTACCAGCGGGGCATCCTGCTGCCGCGAAGGTCTGGTGCGCCTCCGGGCG  
CAGCAAGGTGATCAAGGGCTCCCTGCCCCCTCATCGGGGAGGCGGACTGCCTGCACGAG  
30 AAGTACGGCGGGCTCAACAAGAGCAAGCCGTACTACACCGGCGAGCACGCCAAGGCC  
ATCGGGAAGTGCCCCATCTGGGTCAAGACCCCGCTGAAGCTCGCCAACGGCACGAAGT  
ACCGGCCCCCGGCCAAGCTGCTCAAGGAGCAGGGGTTCTTCGGCGCCATCGCCGGGTT  
CCTGGAGGGCGGGTGGGAGGGCATGATCGCCGGGTGGCACGGCTACACGTCCCACGGG  
GCCACGGCGTGCGGTGCGCCGCGGACCTCAAGAGCACCCAGGAGGCCATCAACAAG  
35 ATCACGAAGAACCTGAACTCCCTCAGCGAGCTGGAGGTGAAGAACCTCCAGCGCCTGT  
CCGGGGCGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGACGAGAAGGTGACG  
ACCTGCGGGCCGACACCATCAGCTCGCAGATCGAGCTCGCGGTGCTGCTCTCCAACGA  
GGGCATCATCAACAGCGAGGACGAGCACCTGCTCGCCCTGGAGCGGAAGCTCAAGAA  
GATGCTGGGGCCCTCGGCGGTGAGATCGGCAACGGGTGCTTCGAGACCAAGCACAAAG  
40 TGCAACCAGACGTGCCCTGACCGCATCGCCGCGGGCACGTTTCGACGCCGGGGAGTTCT  
CCCTGCCGACCTTCGACAGCCTCAACATCACGGCGGCCTCGCTGAACGACGACGGCCT  
CGACAACCACACCATCCTGCTCTACTACTCCACCGCGGCCAGCTCGCTGGCGGTGACGC  
TCATGATCGCCATCTTCGTCGTGTACATGGTCTCCCGCGACAACGTGAGCTGCTCGATCT  
GCCTGTGA

SEQ ID NO: 104

>B 型流感病毒/Yamagata HA0 突变体 CDS

ATGAAGGCCATCATCGTGCTGCTGATGGTGGTGACCAGCAACGCCGACCGGATCTGCAC  
CGGCATCACCAGCTCCAACAGCCCCACGTGGTGAAGACCGCCACCCAGGGCGAGGTG  
AACGTGACCGGCGTGATCCCCCTGACCACCACCCCCACCAAGTCCTACTTCGCCAACCT  
5 GAAGGGCACCCGGACCCGCGGCAAGCTGTGCCCCGACTGCCTGAACTGCACCGACCTG  
GACGTGGCCCTGGGCCGCCCCATGTGCGTGGGCACCACCCCCAGCGCCAAGGCCTCCA  
TCCTGCACGAGGTGCGCCCCGTGACCAGCGGCTGCTTCCCCATCATGCACGACCGCACC  
AAGATCCGGCAGCTGCCCCAACCTGCTGCGGGGGTACGAGAAGATCCGCCTGTCCACCC  
AGAACGTTCATCGACGCCGAGAAGGCCCCCGCGGGGCCCTACCGGCTGGGCACCAGCG  
10 GGTCTTGCCCCAACGCCACCAGCAAGATCGGCTTCTTCGCCACCATGGCCTGGGCCGTG  
CCCAAGGACAACCTACAAGAAGCCACCAACCCCCCTGACCGTCGAGGTGCCCTACATCT  
GCACCGAGGGGGAGGACCAGATCACGGTCTGGGGCTTCCACTCCGACAACAAGACCC  
AGATGAAGTCCCTGTACGGGGACTCCAACCCGAGAAAGTTCACGAGCTCCGCCAACGG  
CGTGACCACGCACTACGTCAGCCAGATCGGGGACTTCCCGGACCAGACCGAGGACGGC  
15 GGGCTGCCCCAGTCCGGCCGCATCGTGGTCTGACTACATGATGCAGAAGCCCGGAAGA  
CGGGCACCATCGTGTAACAGCGGGGGTCTGCTGCCGAGAAGGTGTGGTGCGCCAG  
CGGCCGCTCCAAGGTCATCAAGGGGAGCCTGCCCCATCGGCGAGGCCGACTGCCTG  
CACGAGGAGTACGGGGGCCTCAACAAGTCCAAGCCGTAACACGGGGAAGCACGCC  
AAGGCCATCGGCAACTGCCCCGATCTGGGTGAAGACCCCCCTGAAGCTCGCCAACGGGA  
20 CGAAGTACCGGCCGCCGCAAGCTGCTCAAGGAGCAGGGCTTCTTCGGGGCCATCGC  
CGGCTTCTGGAGGGGGGCTGGGAGGGGATGATCGCCGGCTGGCACGGGTACACCAGC  
CACGGCGCCACGGGGTCGCGGTGGCCGCGGACCTCAAGTCCACGCAGGAGGCCATCA  
ACAAGATCACCAAGAACCTGAACAGCCTCTCCGAGCTGGAGGTCAAGAACCTCCAGCG  
CCTGAGCGGCGCGATGGACGAGCTCCACAACGAGATCCTGGAGCTCGACGAGAAGGTG  
25 GACGACCTGCGGGCCGACACGATCTCGTCCCAGATCGAGCTCGCGGTCTGCTCAGCA  
ACGAGGGGATCATCAACTCGGAGGACGAGCACCTGCTCGCCCTGGAGCGGAAGCTCAA  
GAAGATGCTGGGCCCGTCCGCGGTGGACATCGGGAACGGCTGCTTCGAGACCAAGCAC  
AAGTGCAACCAGACGTGCCTCGACCGCATCGCCGCGGGGACCTTCAACGCCGGCGAGT  
TCAGCCTGCCCACGTTCTGACTCGCTCAACATCACCGCGGCCTCCCTGAACGACGACGG  
30 GCTCGACAACCACACGATCCTGCTCTACTACAGCACCGCGGCCTCGTCCCTGGCGGTCA  
CGTCATGCTGGCCATCTTCATCGTGATCATGGTCAGCCGCGACAACGTGTCTGTCTCA  
TCTGCCCTCTGA

SEQ ID NO: 105

>I-Delta24-B-120A mRNA

35 GAGUUAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA  
CUAUAUAUUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA  
CCUGAGGACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGU  
GUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUUU  
UCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCCACGUCAGCGGCACCAA  
40 CGGCACCAAGAGAUUCGACAACCCCGUGCUGCCCUUAACGACGGCGUGUACUUCGC  
UAGCACCGAAAAGAGCAACAUCUACAGAGGCUUGAUUCUUCGGCACCAACCUAGACUC  
CAAGACACAAAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUCAAGGUGUG  
CGAGUUUCAGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAA  
GAGCUGGAUGGAGAGCGGGGUCUACAGCAGCGCCAACAACUGCACCUCUGAGUACGU  
45 GUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAAG  
AGAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCC  
CAUCAACCUGGUGAGAGACCUGCCCCAAGGCUUCAGCGCCCUUGGAGCCCCUGGUGGA  
CCUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAG

CUACCUAACCCCGGCGAUAGUUCUCUGGCUUGACGGCCGGCGUGCCGCCUACUA  
CGUGGGCUACCUAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCACUAU  
CACUGAUGCCGUGGACUGUGCCCUGGACCCCUAGCGAGACCAAGUGCACCCUCAA  
GAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCAC  
5 CGAGAGCAUCGUGAGAUUCCCCAACAUACCAACCUUGUGCCCUUCGCGGAGGUGU  
CAACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACUG  
CGUGGCCGACUACAGCGUGCUGUACAACAGCGCUAGCUUCAGCACCUUCAAGUGCUA  
UGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAG  
CUUCGUGAUACAGAGGCGACGAGGUGAGACAGAUUGCCCCGGGCAAACCGGCAAGAU  
10 CGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCGUGGAA  
CAGCAAUAACCUUGAUUCCAAGGUGGGCGGCAACUACAACUACCGGUACAGACUGU  
CAGAAAGAGCAACCUGAAGCCUUCGAGAGAGACAUCAGCACCGAGAUCUACCAAGC  
CGGCAGCAAGCCUGCAACGGCGUGGAGGGCUUCAACUGCUACUCCCCCUGCAGAG  
CUACGGCUUCAGCCCACCAACGGCGUGGGCUAUCAGCCUACAGAGUGGUGGUGCU  
15 GAGCUUCGAGCUGCUGCACGCCCCCGAACCGUGUGCGGCCCAAAAAGAGCACCAA  
CCUGGUGAAGAACAAGUGCGUGAACUUAACUUAUUGGCCUUAACCGGACGGGCGU  
GCUGACCGAGAGCAACAAGAAGUCCUGCCAUAUUCAGCAAUUCGGCAGAGACAUCGC  
CGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUACCCC  
CUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCGGCACUAAACAAGCAACCAAGU  
20 GGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCCCGUGGCCAUCCACGCCGA  
UCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACGUGUUUCAGACAAG  
AGCCGGCUGCCUGAUCGGCGCCGAGCACGUGAACAAACAGCUACGAGUGCGACAUCCC  
CAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAACAGCCGCCGGAGAGC  
UAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAA  
25 CAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGU  
GACCACCGAGAUCCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUA  
CAUCUGCGGCGACAGCACCGAGUGCAGCAACCUGCUGCUGCAGUACGGCAGCUUCUG  
CACACAGCUGAACAGAGCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCA  
AGAGGUGUUCGCCAAGUGAAGCAGAUUCUACAAGACCCCCCAUCAAGGACUUCGG  
30 CGGCUUCAACUUCAGCCAAAUCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUU  
CAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUAAGCA  
GUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAGACCUGAUCUGCGCUCAGAAGUU  
CAAUGGACUACCGUGCUCUCCCCCUGCUGACCGGACGAGAUAGCUGCUGAUCAC  
CUCUGCCCUGCUGGCGGGAACCAUAACAAGCGGUUGGACAUAUUGGCCCGGGCGC  
35 ACUGCAAAUCCCCUUCGCCAUGCAGAUUGCCUACAGAUUCAACGGCAUCGGCGUGAC  
ACAGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCAU  
CGGCAAGAUCCAAGACAGCCUGAGCAGCACCGCUAGCGCCCUGGGCAAGCUGCAAAA  
CGUGGUGAAUCAGAACGCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAA  
CUUCGGCGCAUCAGCAGCGUAUUGAAUGACAUCUGUCUAGACUGGACCCCCCGA  
40 GGCCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUA  
CGUGACACAGCAGCUGAUCAGAGCCGCCGAGAUACAGAGCUAGCGCCAACCUGGCCG  
CACCAGAUGAGCGAGUGCGUGCGUGGCGGACAGCAAGAGAGUGGACUUCUGCGGA  
AGGGCUACCACUGAUGAGCUUCCCUAGAGCGCCCCCACGGCGUGGUGUUCUGC  
ACGUGACCUACGUGCCCGCCCAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCC  
45 ACGACGGCAAGGCCACUUCUUCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACU  
GGUUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACACCACGACAACACCU  
UCGUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAACAAACCCGUAUACGACC  
CCCUGCAGCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUAACAAACC  
AUACUAGCCCCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGA  
50 ACAUUCAGAAGGAGAUAGACCGGCUAGAUAGGUGGCCAAGAACCUGAACGAGAGCC  
UGAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCUGGUACA  
UCUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGU  
GCUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCU  
GCAAGUUCGACGAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACA  
55 CCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAUAUUAUUAAGGUUCCUUUGU  
UCCCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGA  
UUCUGCCUAAUAAAAACAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 106

&gt;I'-OC24-B-120A mRNA

5 GAGUUAAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGCA  
CUAUAUAUUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAA  
CCUGACCACAAGAACACAGCUGCCCCCGCCUACACCAACAGCUUCACAAGAGGCGU  
GUACUACCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUU  
UCUGCCCUUCUUCAGCAACGUGACCUGGUUCCACGUGAUACAGCGGCACCAACGGCAC  
CAAGAGAUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCAU  
10 CGAAAAGAGCAACAUCAUACAGAGGCUGGAUCUUCGGCACCACCCUAGACUCCAAGAC  
ACAAAGCCUGCUGAUCGUAACAACGCGACCAACGUGGUGAUCAAGGUGUGCGAGUU  
UCAGUUCUGCAACGACCCCUUCCUGGACCACAAGAACAACAAGAGCUGGAUUGGAGAG  
CGAGUUCAGAGUGUACAGCAGCGCCAACAACUGCACCUUCGAGUACGUGUCUCAGCC  
CUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAAGAGAGUUCG  
15 UGUUCAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCAUCAUCG  
UGAGAGAGCCCGAGGACCUGCCCCAAGGCUUCAGCGCCUGGAGCCCUUGGUGGACC  
UGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGCU  
ACCUAACCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACG  
UGGGCUACCUUGCAGCCUAGAACCUCUGCUGAAGUACAACGAGAACGGCACUAUCA  
20 CUGAUGCCGUGGACUGGCCUGGACCCCUAGCGGAGACCAAGUGCACCCUCAAGA  
GCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCACCG  
AGAGCAUCGUGAGAUUCCCCAACUACCAACCUGUGCCCUUCGACGAGGUGUUA  
ACGCCACCAGAUUUGCUAGCGUGUACGCCUGGAACCGGAAGAGAAUCAGCAACUGCG  
UGGCCGACUACAGCGUGCUGUACAACCUGGCCCCUUCUUCACCUUCAAGUGCUAUG  
25 GCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUACGCCGACAGCU  
UCGUGAUCAGAGGCGACGAGGUGAGACAGAUCCGCCCGGGCAGACCGGCAACAUCG  
CCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCGUGCGUGAUCGCCUGGAACA  
GCAACAAGCUGGAUAGCAAGGUGAGCGGCAACUACAUAUACCGUACAGACUGUUA  
GAAAGAGCAACCUGAAGCCUUCGAGAGAGACAUACGACCCGAGAUUAACCAAGCCG  
30 GCAACAAGCCUGCAACGGCGUGGCGGGCUUACUACUGCUACUUCCCCCUCAGGAGCU  
ACAGCUUCAGACCCACCUACGGCGUGGGCAUCAGCCCUACAGAGUGGUCGUGCUGA  
GCUUCGAGCUGCUGCACGCCCUGCCACAGUGUGCGGCCCAAAAAGAGCACCAACC  
UGGUGAAGAACAAGUGCGUGAACUUAACUUAACGGCCUGAAGGGCACCGGCGUGC  
UGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAUUCGGCAGAGACUUCGCCG  
35 ACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCUGGACAUCACCCCU  
GUAGCUUCGGCGGCGUGAGCGUGAUCACCCCGGCACUAAACAAGCAACCAAGUGG  
CCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGGCCCGUGGCCAUCCACGCCGAC  
AGCUGACCCCCACCUGGAGAGUGUAUAGCACCGGCAGCAACGUGUUUCAGACAAGAG  
CCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAACAGCUACGAGUGCGACAUCCCCA  
40 UCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAAGAGCCACCGGAGAGCUA  
GAAGCGUGGCUAGCCAAAGCAUCAUCGCCUACACCAUGAGCCUGGGCGCCGAGAACA  
GCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCAACUUCACCAUCAGCGUGA  
CCACCGAGAUCCUGCCCGUAAGUAUGACCAAGACAAGCGUGGACUGCACCAUGUACA  
UCUGCGGCGACAGCACCCGAGUGCAGCAACCUGCUCCUGCAGUACGGCAGCUUCUGCA  
45 CACAGCUGAAGAGAGCCCUGACCGGCAUCGCCGUGGAGCAAGACAAGAACACCCAAG  
AGGUGUUCGCCAAGUGAAGCAGAUUCAAGACCCCCCAUCAAGUACUUCGGCG  
GCUUCAACUUCAGCCAAUCCUGCCCGACCCUAGCAAGCCUAGCAAGCGGAGCUUCA  
UCGAGGACCUGCUGUUAACAAGGUGACCCUGGCCGACGCCGGCUUCAUAAGCAGU  
ACGGCGACUGCCUAGGCGACAUCGCGAGCUAGAGACCUGAUCUGCGCUCAGAAGUUA  
50 AGGGACUACCGUGCUCCCCCCCCUGCUGACCGACGAGAUGAUCGCUACUACACCU  
CUGCCCUUGCUGGCGGGAACCAUAACAAGCGGUUGGACAUUUGGCGCCGGGGCCGCAC  
UGCAAUCCCCUUCGCCAUGCAGAUUGGCCUACAGAUUCAACGGCAUCGGCGUGACAC  
AGAACGUGCUGUACGAGAAUCAGAAGCUGAUCGCCAAUCAGUUAACAGCGCCCAUCG  
GCAAGAUCCAAGACGCCUGAGCAGCACCGCUAGCGCCCUUGGGCAAGCUGCAAGCAG  
55 UGGUGAACCACAACGCCCAAGCCUGAACACCCUGGUGAAGCAGCUGAGCAGCAAGU  
UCGGCGCCAUCAGCAGCGUGCUGAUAUGACAUCUUCUCUAGACUGGACCCCCCGAGG  
CCGAGGUGCAGAUUCGACAGACUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACG  
UGACACAGCAGCUGAUCAGAGCCGCCGAGAUACAGAGCUAGCGCCAACCUGGCCGCCA

CCAAGAUGAGCGAGUGCGUGCGUGGGGCGAGAGCAAGAGAGUGGACUUCUGCGGCAAG  
GGCUACCACCUGAUGAGCUUCCUCAGAGCGCCCCCACGGCGUGGUGUUCUGGCAC  
GUGACCUACGUGCCCCGCCAAGAGAAGAACUUCACCACCGCUCCCCGCCAUCUGCCAC  
GACGGCAAGGCCACUUCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGG  
5 UUCGUGACACAGAGAAACUUCUACGAGCCUCAGAUCAUACACCACCGACAACACCUUC  
GUGAGCGGCAACUGCGACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCC  
CUGCAGCCCCGAGCUGGACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAU  
ACUAGCCCCGACGUGGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAAC  
AUUCAGAAGGAGAUUCGACCGGCGUAAUGAGGUGGCAAGAACCUGAACGAGAGCCU  
10 GAUCGACCUGCAAGAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGGCCUGGUACA  
CUGGCUGGGCUUCAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUG  
CUGCAUGACAAGCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUG  
CAAGUUCGACGAGGACGACAGCGAGCCCCGUGCUGAAGGGCGUGAAGCUGCACUACAC  
CUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUGUU  
15 CCCUAAGUCCAACUACUAAACUGGGGGAUUAUUAUGAAGGGCCUUGAGCAUCUGGAU  
UCUGCCUAAUAAAAAAAAACAUUUUUUCAUUGCAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 107

20 &gt;I'-BA4-B-120A mRNA

GAGUUAAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUG  
AACCUGAUCACAAGAACACAGUCGUACACCAACAGCUUCACAAGAGGCGUGUACUA  
CCCCGACAAGGUGUUCAGAAAGCAGCGUGCUGCACUCCACCCAAGACCUGUUUCUGC  
25 CCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGA  
GAUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCGAA  
AAGAGCAACAUCAUACAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACA  
AAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUC  
AGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAAGAGCUGG  
30 AUGGAGAGCGAGUUCAGAGUCUACAGCAGCGCCAACAACUGCACCUCGAGUACGU  
GUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAA  
GAGAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACC  
CCCAUCAACCUGGGGAGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUG  
GACCUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGU  
35 AGCUACCUAACCCCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCGUGCCGCCUAC  
UACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCAC  
UAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCC  
UCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAG  
CCCACCGAGAGCAUCGUGAGAUUCCCCAACAUCACCAACCUGUGCCCCUUCGACGAG  
40 GUGUUAACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAG  
CAACUGCGUGGCCGACUACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGCCUUA  
AGUGCUAUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUAC  
GCCGACAGCUUCGUGAUCAGAGGCAACGAGGUGAGCCAGAUUGCCCCCGGGCAAAC  
CGGCAACAUCGCCGACUACAACUACAAGCUGCCCCGACGACUUCACCGGUGCGUGA  
45 UCGCGUGGAACAGCAAUAAGCUGGAUUCGAAGGUGGGCGGCAACUACAACUACCGG  
UACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGA  
GAUCUACCAAGCCGGCAACAAGCCCUGCAACGGCGUGGCGGGCGUCAACUGCUACU  
UCCCCUGCAGAGCUACGGCUUUAAGGCCACCUACGGCGUGGGCCAUCAGCCCUACA  
GAGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCC

AAAAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUUAAUGGCCU  
UAAGGGCACGGGCGUGCUGACCGAGAGCAACAAGAAGUCCUGCCAUUUCAGCAAU  
UCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAG  
AUCCUGGACAUCACCCCCUGUAGCUUCGGCGGGCUGAGCGUGAUCACCCCCGGCACU  
5 AACACAAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCC  
CGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAA  
GCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAAAC  
AGCUACGAGUGCGACAUCUCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACA  
GACCAAGAGCCACCGGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACA  
10 CCAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUC  
CCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCCUGCCCCGUAAGUAUGACCAAG  
ACAAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCU  
GCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCCUGACCGGCAUCG  
CCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAGUGAAGCAGAUUCAC  
15 AAGACCCCCCCCCAUCAAGUACUUCGGCGGGCUUCAACUUCAGCCAAAUCUGCCCCGAC  
CCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGAC  
CCUGGCCGACGCCGGCUUCAUAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAG  
CUAGAGACCUGAUCUGCGCUCAGAAGUUAAGGACUCACCGUGCUCUCCCCCCCCUGC  
UGACCGACGAGAUGAUCGCUCAGUACACCUUCGCCUGCUGGCGGGAACCAUAACA  
20 AGCGGUUGGACAUUUGGCGCCGGGGCCGACUGCAAAUCCCCUUCGCCAUGCAGAU  
GGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACGAGAAUCAGA  
AGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUAACAGACAGCCUGAGC  
AGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGUGGUGAAUACAACGCCCAAGC  
CCUGAACACCCUGGUGAAGCAGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUAU  
25 UGAAUGACAUCUUGUCUAGACUGGACCCCCCGAGGCGGAGGUGCAGAUUCGACAGA  
CUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAG  
AGCCGCCGAGAUCAGAGCUAGCGCCAACCUGGCCGCCACCAAGAUGAGCGAGUGCG  
UGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGC  
UCCCCUCAGAGCGCCCCCACGGCGUGGUGUCCUGCACGUGACCUACGUGCCCCGCC  
30 CAAGAGAAGAACUUCACCACCGCUCUCCGCCAUCUGCCACGACGGCAAGGCCACUUC  
CCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAA  
CUUCUACGAGCCUCAGAUCAUACACCACCGACAACACCUUCGUGAGCGGCAACUGCG  
ACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCCGAGCUG  
GACAGCUUCAAGGAGGAGCUGGACAAGUACUUAACCAUACUAGCCCCGACGU  
35 GGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAAGGAGA  
UCGACCGGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGCAA  
GAGCUGGGCAAGUACGAGCAGUACAUAAGUGGCCCUUGGUACAUCUGGCUGGGCUU  
CAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAA  
GCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGAC  
40 GAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACU  
CGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUUGUCCCCUAAGU  
CCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUA  
AUAAAAAACAUUUAUUUUCAUUGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
45 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 108

&gt;I-BA4-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGUUCGUGUUCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUG  
5 AACCUGAUCACAAGAACACAGUCGUACACCAACAGCUUCACAAGAGGCGUGUACUA  
CCCCGACAAGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUUUCUGC  
CCUUCUUCAGCAACGUGACCUGGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGA  
GAUUCGACAACCCCGUGCUGCCCUUCAACGACGGCGUGUACUUCGCUAGCACCGAA  
AAGAGCAACAUCAUACAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACA  
10 AAGCCUGCUGAUCGUUAACAACGCGACCAACGUGGUGAUC AAGGUGUGCGAGUUUC  
AGUUCUGCAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAAGAGCUGG  
AUGGAGAGCGAGUUCAGAGUCUACAGCAGCGCCAACAACUGCACCUUCGAGUACGU  
GUCUCAGCCCUUCCUGAUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAA  
GAGAGUUCGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACC  
15 CCCAUAACCUUGGGGAGAGACCUGCCCCAAGGCUUCAGCGCCCUGGAGCCCCUGGUG  
GACCUGCCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCGUCACCGU  
AGCUACCUAACCCCGGCGAUAGUUCUCUGGCUGGACGGCCGGCGCUGCCGCCUAC  
UACGUGGGCUACCUGCAGCCUAGAACCUUCCUGCUGAAGUACAACGAGAACGGCAC  
UAUCACUGAUGCCGUGGACUGUGCCCUGGACCCCCUGAGCGAGACCAAGUGCACCC  
20 UCAAGAGCUUCACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAG  
CCCACCGAGAGCAUCGUGAGAUUCCCCAACAUCACCAACCUGUGCCCUUCGACGAG  
GUGUUAACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAG  
CAACUGCGUGGCCGACUACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGCCUUA  
AGUGCUAUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUCACCAACGUGUAC  
25 GCCGACAGCUUCGUGAUCAGAGGCAACGAGGUGAGCCAGAUUGCCCCCGGGCAAAC  
CGGCAACAUCGCCGACUACAACUACAAGCUGCCCCGACGACUUCACCGGCGUGGUGA  
UCGCGUGGAACAGCAAUAAGCUGGAUUC AAGGUGGGCGGCAACUACAACUACCGG  
UACAGACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGA  
GAUCUACCAAGCCGGCAACAAGCCUGCAACGGCGUGGCGGGCGUCAACUGCUACU  
30 UCCCCUGCAGAGCUACGGCUUUAGGCCCAACUACGGCGUGGGCCAUCAGCCCUACA  
GAGUGGUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCGCAACCGUGUGCGGCCCC  
AAAAAGAGCACCAACCUGGUGAAGAACAAGUGCGUGAACUUAACUUAUUGGCCU  
UAAGGGCACGGGCGUGCUGACCGAGAGCAACAAGAAGUUCUGCCAUUUCAGCAAU  
UCGGCAGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAG  
35 AUCCUGGACAUCACCCCCUGUAGCUUCGGCGGCGUGAGCGUGAUCACCCCCGGCACU  
AACACAAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCC  
CGUGGCCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAA  
GCAACGUGUUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAAC  
AGCUACGAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACA  
40 GACCAAGAGCCACCGGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACA  
CCAUGAGCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUC  
CCCACCAACUUCACCAUCAGCGUGACCACCGAGAUCCUGCCCGUAAGUAUGACCAAG  
ACAAGCGUGGACUGCACCAUGUACAUCUGCGGCGACAGCACCGAGUGCAGCAACCU  
GCUGCUGCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCCUGACCGGCAUCG  
45 CCGUGGAGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAGUGAAGCAGAUAC

AAGACCCCCCAUCAAGUACUUCGGCGGCUUCAACUUCAGCCAAAUCCUGCCCCGAC  
CCUAGCAAGCCUAGCAAGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGAC  
CCUGGCCGACGCCGGCUUCAUCAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAG  
CUAGAGACCUGAUCUGCGCUCAGAAGUUCAAUUGGACUCACCGUGCUCCCCCCCCUGC  
5 UGACCGACGAGAUGAUCGCUCAGUACACCUUCUGCCCUGCUGGCGGGAACCAUAACA  
AGCGGUUGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCAGAU  
GGCCUACAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACGAGAAUCAGA  
AGCUGAUCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUAACAAGACAGCCUGAGC  
AGCACCGCUAGCGCCCUGGGCAAGCUGCAAGACGUGGUGAAUCACAACGCCCAAGC  
10 CCUGAACACCCUGGUGAAGCAGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUAU  
UGAAUGACAUCUUGUCUAGACUGGACCCCCCGAGGCGGAGGUGCAGAUUCGACAGA  
CUGAUCACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAG  
AGCCGCCGAGAUCAGAGCUAGCGCCAACCUUGGCCGCCACCAAGAUGAGCGAGUGCG  
UGCUGGGGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGC  
15 UUCCUCAGAGCGCCCCCACGGCGUGGUGUUCUUGCAGCUGACCUACGUGCCCCGCC  
CAAGAGAAGAACUUCACCACCGCUCCCGCCAUCUGCCACGACGGCAAGGCCACUUC  
CCUAGAGAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAA  
CUUCUACGAGCCUCAGAUCAUACACCACCGACAACACCUUCGUGAGCGGCAACUGCG  
ACGUGGUGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCUG  
20 GACAGCUUCAAGGAGGAGCUGGACAAGUACUUCAAAAACCAUACUAGCCCCGACGU  
GGACCUGGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUUCAGAAGGAGA  
UCGACCGGCGUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUCAA  
GAGCUGGGCAAGUACGAGCAGUACAUCAAGUGGCCCUUGGUACAUCUGGCUGGGCUU  
CAUCGCCGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAA  
25 GCUGUUGCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGAC  
GAGGACGACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACACCUGAUGACU  
CGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUCCUUUGUCCCUAAGU  
CCAACUACUAAACUGGGGGAUUAUAUGAAGGGCCUUGAGCAUCUGGAUUCUGCCUA  
AUAAAAACAUAUUAUUUUAUUGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
30 AGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 109

&gt;I'-H1N1-B mRNA

GAGUUAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
35 ACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUACACCAGCCAAAC  
GCCGACACCCUGUGCAUCGGCUACCACGCCAACAACAGCACCGACACCGUGGACACC  
GUGCUGGAGAAGAACGUGACCGUGACCCACAGCGUGAACCUGCUGGAGGACAAGCA  
CAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGCCCCCUGCACCUGGGCAAGUGCA  
ACAUCGCCGGCUGGAUCCUGGGCAACCCCGAGUGCGAGUCCUGAGCACCGCCCGCU  
40 CCUGGAGCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUUGCUACCCCGGG  
GACUUAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGA  
GCGCUUCGAGAUCUCCCCAAGACCUCAGCUGGCCCAACCACGACUCCGACAACGG  
CGUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUG  
GCUGGUGAAGAAGGGCAAGUCCUACCCCAAGAUCAACCAGACCUACAUCAACGACA  
45 AGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCCACCUCGCCGACC

AGCAGAGCCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCGCGGUAC  
AGCAAGAAGUUCAAGCCCGAGAUCCGACCCGCCCAAGGUCCGGGACCAGGAGGG  
CCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUACCCUUCGAGG  
CCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGU  
5 CCGGCAUCAUCAUCAGCGACACCCCCGUGCAGCAGUCGCAACACGACCUGCCAGACGC  
CCGAGGGGGCGAUCAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCG  
GCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGC  
AACGUCCCCUCCAUCAGAGCCGGGGCCUGUUCGGGGCGAUCCGCCGCUUCAUCGA  
GGGGGGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACGAGC  
10 AGGGCUCGCGGUACGCGGCCGACCUCUAGUCGACGCAGAACGCGAUCGACAAGAUC  
ACCAACAAGGUCAACAGCGUGAUUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGG  
CAAGGAGUUAACACCUGGAGAAGCGCAUCGAGAACCUCUACAAGAAGGUGGACG  
ACGGGUUCCUGGACAUCUGGACGUACAACGCGGAGCUCCUGGUCCUCCUGGAGAAC  
GAGCGGACCCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCG  
15 CAACCAGCUCAAGAACAACGCCAAGGAGAUCCGGCAACGGGUGCUUCGAGUUCUACC  
ACAAGUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUACCCG  
AAGUACAGCGAGGAGGCGAAGCUGAACCGGGAGAAGAUCCGACGGGGUCAAGCUCGA  
CUCCACGCGCAUCUACCAGAUCCUGGCCAUCUACUCGACCGUGGCGAGCUCCUCGU  
CCUGGUGGUCUCGCUCCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCUGC  
20 AGUGCCGGAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUA  
UUAAGGUUCCUUGUCCCUAAGUCCAACUACUAAACUGGGGGGAUAUUAUGAAGG  
GCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUAUUAUUUUAUUGCAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
25 SEQ ID NO: 110  
>I'-H3N2-B mRNA  
GAGUUAAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGACCAUCAUCGCCCUGAGCUACAUCUCCUGUGCCUGGUGUUCGCC  
CAGAAGAUCCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCC  
30 GUGCCCAACGGCACCAUCGUGAAGACCAUACCAACGACCGGAUCGAGGUGACCAA  
CGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGAGAUUCGCGACUCCCCCACCA  
GAUCCUGGACGGCGGCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCCAGUG  
CGACGGGUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACU  
CCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCUGGUGGCCA  
35 GCUCCGGCACCCUGGAGUUAAGAACGAGAGCUUACAACUGGACCGGGGUCAAGCAG  
AACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCUCCCGGCGUG  
AACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAAC  
AACGAGCAGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCGACAA  
GGACCAGAUUCUCCUUCGCCCAGCCAGCGGGCGCAUACCGUGUCCACGAAGCG  
40 GAGCCAGCAGGCCGUAUCCCCAACAUCCGCUCCCGCCCGCGGAUCCGCGACAUCCC  
CAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCUGCUCAUCA  
ACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUAAGAUCGGGUCCGGCAAGA  
GCUCGAUAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCC  
CCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGUCAACCGGAUCACGUAC  
45 GGGGCCUGCCCGCGCUACGUGAAGCAGUCCACGCUAAGCUGGCCACCGGCAUGCG

GAACGUCCCCGAGAAGCAGACCCGCGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCG  
AGAACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAACAGC  
GAGGGCCGCGGGCAGGCCGCGGACCUC AAGUCGACGCAGGCCGGAUCGACCAGAU  
CAACGGCAAGCUGAACCGGCUC AUCGGGAAGACGAACGAGAAGUCCACCAGAUUCG  
5 AGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAG  
GACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGCUCGUGGCGCUGGAGAA  
CCAGCACACCAUCGACCUCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGA  
AGAAGCAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUUCAC  
CACAAGUGCGACAACCGUGCAUCGGGUCCAUCCGCAACGAGACCUACGACCACAA  
10 CGUCUACCGGGACGAGGCCCUGAACAACCGCUCCAGAUCAAGGGCGUGGAGCUCA  
AGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCCUC  
CUGUGCAUCGCCCUCCUGGGCUUCAUCAUGUGGGCGUGCCAGAAGGGGAACAUCCG  
GUGCAACAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAU  
UAAAGGUUCCUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGG  
15 CCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUUCAUUGCAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAGCAUAGACUAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 111

>I'-B-Victoria-B mRNA  
20 GAGUUAAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACC GGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCCUGACCACCACCCCCACCAAGUCC  
CACUUCGCCAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCU  
25 GAACUGCACC GACCUGGACGUGGCCCUGGGCCGCCCAAGUGCACC GGCAAGAUCUCC  
CAGCGCCCCGGGUGUCCAUCUUGCAGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC  
CAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGGGGUACG  
AGCACGUCGCCUGUCCACCCACAACGUGAUAACGCCGAGGACGCCCCCGGCCGGC  
CCUACGAGAUCGGGACCAGCGGCUCUUGCCCCAACAUACCAACGGGAACGGCUUCU  
30 UCGCCACCAUGGCCUGGGCCGUCUCCCAAGAACAAGACCGCCACCAACCCCCUGACCA  
UCGAGGUGCCGUACAUCUGCACCGAGGGGGAGGACCAGAUACCGUCUGGGGCUUC  
CACAGCGACAACGAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAA  
GUUCACGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCU  
UCCCGAACCCAGACGGAGGACGGGGGGCCUGCCCCAGUCCGGGGCGAUCGUGGUCGAC  
35 UACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCCU  
GCUGCCGCAGAAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUAAGGGCUCCC  
UGCCCCUCAUCGGGGAGGCCGACUGCCUGCACGAGAAGUACGGCGGGCUCAACAAG  
AGCAAGCCGUACUACACCGGCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUG  
GGUCAAGACCCCGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCU  
40 GCUCAAGGAGCGCGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGUGGG  
AGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCACGGGGCCCACGGCGUGGCG  
GUCGCCCGGACCUC AAGAGCACCCAGGAGGCCAUACAAGAUCACGAAGAACCU  
GAACUCCUCAGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGG  
ACGAGCUCCACAACGAGAUCUUGGAGCUCGACGAGAAGGUCGACGACCUGCGGGCC  
45 GACACCAUCAGCUCGCAGAUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAU

CAACAGCGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGG  
GGCCCUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGCAAC  
CAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGAGUUCUCCUG  
CCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUCGAC  
5 AACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGCGGUGACGCUC  
AUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUCGUCGAU  
CUGCCUGUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUC  
CUUUGUCCCCUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCA  
UCUGGAUUCUGCCUAAUAAAAACAUUUAUUUCAUUGCAAAAAAAAAAAAAAAAAA  
10 AAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 112

>I'-B-Yamagata-B mRNA  
GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
15 ACUAAUAUGAAGGCCAUCAUCGUCGUCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACC GGCAUACACAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAACAGUCC  
UACUUCGCCAACCUGAAGGGCACCCGACCCGCGCAAGCUGUGCCCCGACUGCCUG  
AACUGCACC GACCUGGACGUGGCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCC  
20 AGCGCCAAGGCCUCCAUCUGCAGGAGUGCGCCCCGUGACCAGCGGCUCGUCCCC  
AUCAUGCACGACCGCACCAAGAUCGCGCAGCUGCCCAACCUGCUGCGGGGGUACGA  
GAAGAUCGCCUGUCCACCCAGAACGUAUCGACGCCGAGAAGGCCCCCGCGGGCC  
CUACCGGCGUGGGCACCGCGGUCCUGCCCCAACGCCACCAGCAAGAUCGGCUUCUU  
CGCCACCAUGGCCUGGGCCGUGCCCAAGGACAACUACAAGAACGCCACCAACCCCU  
25 GACCGUCGAGGUGCCUACAUCUGCACCAGGGGGAGGACCAGAUACGGUCUGGG  
GCUUCCACUCCGACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCG  
AGAAGUUCACGAGCUCGCGCAACGGCGUGACCACGCACUACGUCAGCCAGAUCCGG  
GACUCCCCGGACCAAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUC  
GACUACAUGAUGCAGAAGCCCGGAAGACGGGCACCAUCGUGUACCAGCGGGGGU  
30 CCUGCUGCCGAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGA  
GCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCAGGAGUACGGGGGCCUACAC  
AAGUCCAAGCCGUACUACACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCGAU  
CUGGGUGAAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCGCCAA  
GCUGCUAAGGAGCGCGGUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCU  
35 GGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUC  
GCGGUGGCCGCGGACCUAAGUCCACGCAGGAGGCCAUACAAGAUAACCAAGAA  
CCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGCGCGA  
UGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUGGACGACCUGCGG  
GCCGACACGAUCUCGUCCCAGAUAGCUGCUGCGGUCCUGCUCAGCAACGAGGGGAU  
40 CAUCAACUCGGAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUAAGAAGAU  
UGGGCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAGUGC  
AACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUCAACGCCGGCGAGUUCAGC  
CUGCCCACGUUCGACUCGCUACAACACCGCGGCCUCCUGAACGACGACGGGCU  
GACAACCACACGAUCCUGCUCUACUACAGCACCGCGGCCUCGUCCUGGCGGUCACG  
45 CUCAUGCUGGCCAUCUUCAUUCGUGUACAUGGUCAGCCGCGACAACGUGUCGUCUC

CAUCUGCCUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGG  
UUCUUUGUUCUCCUAAGUCCAACUACUAAACUGGGGGAUUAUGAAGGGCCUUGA  
GCAUCUGGAUUCUGCCUAAUAAAAACAUAUUUUUCAUUGCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAA  
5 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 113

&gt;I'-H1N1-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUACACCGCCAAC  
10 GCCGACACCCUGUGCAUCGGCUACCACGCCAACACAGCACCGACACCGUGGACACC  
GUGCUGGAGAAGAAGCUGACCGUGACCCACAGCGUGAACCUGCUGGAGGACAAGCA  
CAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCA  
ACAUCGCCGGCUGGAUCCUGGGCAACCCCGAGUGCGAGUCCUGAGCACCGCCCGCU  
CCUGGAGCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGGG  
15 GACUUAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGA  
GCGCUUCGAGAUCUUCCTCAAGACCUCCAGCUGGCCCAACCACGACUCCGACAACGG  
CGUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUG  
GCUGGUGAAGAAGGGCAAGUCCUACCTCAAGAUCAACCAGACCUACAUCAACGACA  
AGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCCACCUCGCCGACC  
20 AGCAGAGCCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCTCGGUAC  
AGCAAGAAGUUCAAGCCCGAGAUCGCCACCCGCCCAAGGUCCGGGACCAGGAGGG  
CCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUCAACCUUCGAGG  
CCACCGGCAACCUUCGUCGCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGU  
CCGGCAUCAUCAUCAGCGACACCCCGUGCACGACUGCAACACGACCUGCCAGACGC  
25 CCGAGGGGGCGAUCAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCG  
GCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCUCCTGGCUGGCCACGGGGCUCCTGC  
AACGUCCCCUCCAUCACAGAGCCAGGGCCUGUUCGGGGCGAUCCGCCGGCUUCAUCGA  
GGGGGGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACAGAACGAGC  
AGGGCUCCTGGGUACGCGGCCGACCUCUAGUCGACGCAGAACGCGAUCCGACAAGAUC  
30 ACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGG  
CAAGGAGUUAACACCUUGGAGAAGCGCAUCGAGAACCUCUACAAGAAGGUGGACG  
ACGGGUUCCUGGACAUUCUGGACGUACAACGCGGAGCUCCUGGUCCUCCUGGAGAAC  
GAGCGGACCCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCG  
CAACCAGCUCAAGAACAACGCCAAGGAGAUCCGCAACGGGUGCUUCGAGUUCUACC  
35 ACAAGUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUACCCG  
AAGUACAGCGAGGAGGCGAAGCUGAACCGGGAGAAGAUCGACGGGGUCAAGCUCGA  
CUCCACGCGCAUCUACCAGAUCCUGGCCAUUCUACUCGACCGUGGCGAGCUCCUCGU  
CCUGGUGGUCUCGCUCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCUGC  
AGUGCCGGAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUA  
40 UUAAGGUUCCUUUGUUCUCCUAAGUCCAACUACUAAACUGGGGGAUUAUGAAGG  
GCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUAUUUUUCAUUGCAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 114

45 &gt;I'-H3N2-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGACCAUCAUCGCCCUGAGCUACAUCCUGUGCCUGGUGUUCGCC  
CAGAAGAUCCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCC  
GUGCCCAACGGCACCAUCGUGAAGACCAUACCAACGACCGGAUCGAGGUGACCAA  
5 CGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGAGAUUCUGCGACUCCCCCACCA  
GAUCCUGGACGGCGGCAACUGCACCCUGAUCGACGCCCUGCUGGGCGACCCCCAGUG  
CGACGGGUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACU  
CCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCUGGUGGCCA  
GCUCCGGCACCCUGGAGUUCAAGAACGAGAGCUUCAAACUGGACCGGGGUCAAGCAG  
10 AACGGCACCUCCAGCGCCUGCAUCCGCGGUCCAGCUCCAGCUUCUUCUCCCGGCUG  
AACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAAC  
AACGAGCAGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCGACAA  
GGACCAGAUCUCCCUUCGCCCAGCCCAGCGGGCGCAUCACCGUGUCCACGAAGCG  
GAGCCAGCAGGCCGUAUCCCCAACUCCGCUCCCGCCCGCGGAUCCGCGACAUCCC  
15 CAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCCUGCUCAUCA  
ACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUAAGAUCCGGUCCGGCAAGA  
GCUCGAUCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCC  
CCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGUAACCGGAUCACGUAC  
GGGGCCUGCCCGCGCUACGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCG  
20 GAACGUCCCCGAGAAGCAGACCCAGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCG  
AGAACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAAGACAGC  
GAGGGCCGCGGGCAGGCCGCGGACCUGAAGUCGACGCAGGCCGCGAUCCAGAGAU  
CAACGGCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUCCACCAGAUCCG  
AGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAG  
25 GACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGCUCGUGGCGCUGGAGAA  
CCAGCACACCAUCGACCUCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGA  
AGAAGCAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUUCAC  
CACAAGUGCGACAACGCGUGCAUCGGGUCCAUCCGCAACGAGACCUACGACCACAA  
CGUCUACCGGGACGAGGCCUGAACAAACCGCUUCCAGAUCAAGGGCGUGGAGCUCA  
30 AGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCCUC  
CUGUGCAUCGCCCUCUGGGCUUCAUCAUGUGGGCGUGCCAGAAGGGGAACAUCCG  
GUGCAACAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAU  
UAAAGGUUCCUUUGUUCUUAAAGUCCAACUACUAAACUGGGGGAUUAUUAUGAAGGG  
CCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUUAUUGCAAAAAAAA  
35 AAAAAAAAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 115

&gt;I'-B-Victoria-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
40 ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACCGGCAUACACAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCC  
CACUUCGCCAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCU  
GAACUGCACCGACCGGACGUGGGCCUGGGCCGCCCAAGUGCACCGGCAAGAUC  
45 CAGCGCCCGGGUGUCCAUCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCC

CAUCAUGCACGACCGCACCAAGAUCGGGAGCUGCCCAACCUGCUGCGGGGGUACG  
AGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGGACGCCCCGGCCGGC  
CCUACGAGAUCGGGACCAGCGGCUCCUGCCCCAACAUACCAACGGGAACGGCUUCU  
UCGCCACCAUGGCCUGGGCCGUGCCCCAAGAACAAGACCGCCACCAACCCCCUGACCA  
5 UCGAGGUGCCGUACAUCUGCACCGAGGGGAGGACCAGAUACCGUCUGGGGCUUC  
CACAGCGACAACGAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAA  
GUUCACGAGCUCGGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCU  
UCCCCAACCAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGAC  
UACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCCU  
10 GCUGCCGCAGAAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUAAGGGCUCCC  
UGCCCCUCAUCGGGGAGGCGGACUGCCUGCACGAGAAGUACGGCGGGCUCAACAAG  
AGCAAGCCGUACUACACCGGCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUG  
GGUCAAGACCCCGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCGGCCAAGCU  
GCUCAAGGAGCAGGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGUGGG  
15 AGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCACGGGGCCCACGGCGUGGCG  
GUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUAACAAGAUCACGAAGAACCU  
GAACUCCUCACGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGG  
ACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUCGACGACCUGCGGGCC  
GACACCAUCAGCUCGAGAUCCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAU  
20 CAACAGCGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGG  
GGCCUCGGCGGUGCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGCAAC  
CAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCGGGGAGUUCUCCUG  
CCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUCGAC  
AACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGCGGUGACGCUC  
25 AUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUCGUCGAU  
CUGCCUGUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUC  
CUUUGUUCUUAAAGUCCAACUACUAAACUGGGGGAUUAUUAUGAAGGGCCUUGAGCA  
UCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUCAUUGCAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
30 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 116

&gt;I-B-Yamagata-HA0-mut-B mRNA

GAGUUAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
35 CGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCC  
UACUUCGCCAACCUGAAGGGCACCCGACCCGCGGCAAGCUGUGCCCCGACUGCCUG  
AACUGCACCGACCUUGGACGUGGCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCC  
AGCGCCAAGGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCGUCUCCCC  
40 AUCAUGCACGACCGCACCAAGAUCGGCAGCUGCCCAACCUGCUGCGGGGGUACGA  
GAAGAUCGCCUGUCCACCCAGAACGUAUCGACGCCGAGAAGGCCCCCGGGGGCC  
CUACCGGCUUGGGCACCGCGGUGCCUUGCCCCAACGCCACCAGCAAGAUCGGCUUCUU  
CGCCACCAUGGCCUGGGCCGUGCCCAAGGACAACUACAAGAAGCCACCAACCCCU  
GACCGUCGAGGUGCCCUACAUCUGCACCGAGGGGGAGGACCAGAUACGGUCUGGG  
45 GCUUCCACUCCGACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGC

AGAAGUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAUCCGG  
GACUCCCCGGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUC  
GACUACAUGAUGCAGAAGCCCGGAAGACGGGCACCAUCGUGUACCAGCGGGGGGU  
CCUGCUGCCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGA  
5 GCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCACGAGGAGUACGGGGGGCCUCAAC  
AAGUCCAAGCCGUACUACACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAU  
CUGGGUGAAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCGCCAA  
GCUGCUC AAGGAGCAGGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCU  
GGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUC  
10 GCGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAUCAACAAGAUACCAAGAA  
CCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGGCGGA  
UGGACGAGCUCCACAACGAGAUCUGGAGCUCGACGAGAAGGUGGACGACCUGCGG  
GCCGACACGAUCUCGUCCCAGAU CGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAU  
CAUCAACUCGGAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAU GC  
15 UGGGCCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAUGGC  
AACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUCAACGCCGGCGAGUUCAGC  
CUGCCCACGUUCGACUCGCUCAACAUCACCGCGGGCUCCUGAACGACGACGGGCUC  
GACAACCACACGAUCCUGCUCUACUACAGCACCGCGGCCUCGUCCUGGCGGUCACG  
CUCAUGCUGGCCAUUCUCAUCGUGUACAUGGUCAGCCGCGACAACGUGUCGUGCUC  
20 CAUCUGCCUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGG  
UUCUUUGUUCUUAAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGA  
GCAUCUGGAUUCUGCCUAAUAAAAACAUAUUAUUUCAUUGCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAGCAUAUGACUAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
25 SEQ ID NO: 117  
>I'-H1N1-B-120A mRNA  
GAGUUAAGGAGUUGGUGGCAGUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUACACCGCCAAC  
GCCGACACCCUGUGCAUCGGCUACCACGCCAACAACAGCACCGACACCGUGGACACC  
30 GUGCUGGAGAAGAACGUGACCGUGACCCACAGCGUGAACCUGCUGGAGGACAAGCA  
CAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGCCCCCUGCACCUGGGCAAGUGCA  
ACAUCGCCGGCUGGAUCCUGGGCAACCCGAGUGCGAGUCCUGAGCACCGCCCGCU  
CCUGGAGCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGGG  
GACUUCAUCAACUACGAGGAGCUCGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGA  
35 GCGCUUCGAGAUCUUCCTCAAGACCUCAGCUGGCCCAACCACGACUCCGACAACGG  
CGUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUG  
GCUGGUGAAGAAGGGCAAGUCCUACCTCAAGAUCAACCAGACCUACAUCAACGACA  
AGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCCACCUCGCCGACC  
AGCAGAGCCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCTGGUAC  
40 AGCAAGAAGUUC AAGCCCGAGAU CGCCACCCGCCCAAGGUCCGGGACCAGGAGGG  
CCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUCACCUUCGAGG  
CCACCGGCAACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGU  
CCGGCAUCAUCAUCAGCGACACCTCCGUGCAGCAGUGCAACACGACCUGCCAGACGC  
CCGAGGGGGCGAUCAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCG  
45 GCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGC

AACGUCCCCUCCAUCCAGAGCCGGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGA  
GGGGGGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACGAGC  
AGGGCUCGCGGUACGCGGCCGACCUC AAGUCGACGCAGAACGCGAUCGACAAGAUC  
ACCAACAAGGUCAACAGCGUGAUCGAGAAGAU GAACACGCAGUUCACCGCCGUCGG  
5 CAAGGAGUUCAACCACCUGGAGAAGCGCAUCGAGAACCUCAACAAGAAGGUGGACG  
ACGGGUUCCUGGACAUCUGGACGUACAACGCGGAGCUCUGGUCCUCCUGGAGAAC  
GAGCGGACCCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCG  
CAACCAGCUCAAGAACAACGCCAAGGAGAU CGGCAACGGGUGCUUCGAGUUCUACC  
ACAAGUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUACCCG  
10 AAGUACAGCGAGGAGGCGAAGCUGAACC GGGAGAAGAU CGACGGGGUCAAGCUCGA  
CUCCACGCGCAUCUACCAGAUCCUGGCCAU CUACUCGACCGUGGCGAGCUCUCCUCGU  
CCUGGUGGUCUCGCUCCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCUGC  
AGUGCCGGAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUA  
UUAAGGUUCCUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGG  
15 GCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACA UUUUAUUUUAUUGCAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAA

SEQ ID NO: 118

20 >I'-H3N2-B-120A mRNA  
GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAUAUGAAGACCAUCAUCGCCCUGAGCUACAUCUGGCCUGGUGUUCGCC  
CAGAAGAUCCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCC  
GUGCCCAACGGCACCAUCGUGAAGACCAUCACCAACGACCGGAUCGAGGUGACCAA  
25 CGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGAGAUUCUGCGACUCCCCCACC  
GAUCCUGGACGGCGGCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCCAGUG  
CGACGGGUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAAUC  
CCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCUGGUGGCCA  
GCUCCGGCACCCUGGAGUUAAGAACGAGAGCUUAACUGGACCGGGGUCAAGCAG  
30 AACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCUCCCGGCUG  
AACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAAC  
AACGAGCAGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCGACAA  
GGACCAGAUUCUCCUUCGCCCAGCCAGCGGGCGCAUCACCGUGUCCACGAAGCG  
GAGCCAGCAGGCCGUAUCCCCAACAU CGGCUCUCCGCCCCGCGGAUCCGCGACAUCCC  
35 CAGCCGGAUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCUGCUCAUCA  
ACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUAAGAUCGGUCCGGCAAGA  
GCUCGAUAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCC  
CCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGUAACCGGAUCACGUAC  
GGGGCCUGCCCGCGCUACGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCG  
40 GAACGUCCCCGAGAAGCAGACCCGCGGAUCUUCGGCGCCAUCGCGGGGUUCAUCG  
AGAACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAAGACG  
GAGGGCCGCGGGCAGGCCGCGGACCUC AAGUCGACGCAGGCCGCGAUCGACCAGAU  
CAACGGCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUCCACCAGAU CG  
AGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAG  
45 GACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUCGUCUGGCGCUGGAGAA

CCAGCACACCAUCGACCUCACGGACUCGGAGAUGAACAAAGCUGUUCGAGAAGACGA  
AGAAGCAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUAC  
CACAAGUGCGACAACGCGUGCAUCGGGUCCAUCCGCAACGAGACCUACGACCACAA  
CGUCUACCGGGACGAGGCCCUGAACAACCGCUCCAGAUCAAGGGCGUGGAGCUCA  
5 AGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCCUC  
CUGUGCAUCGCCCUCUGGGCUUCAUCAUGUGGGCGUGCCAGAAGGGGAACAUCCG  
GUGCAACAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUUCUAU  
UAAAGGUUCCUUUGUUCCCUAAGUCCAACUACUAAACUGGGGGAUUAUUGAAGGG  
CCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUAUUGCAAAAAAA  
10 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAA

SEQ ID NO: 119

>I-B-Victoria-B-120A mRNA  
15 GAGUUAAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCC  
CACUUCGCCAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCU  
20 GAACUGCACCGACCUGGACGUGGCCCCUGGGCCGCCCAAGUGCACCGGCAAGAUGCC  
CAGCGCCCCGGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC  
CAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGGGGUACG  
AGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGGACGCCCCGGCCGGC  
CCUACGAGAUCCGGACAGCGGCUCUCCUGCCCCAACAUACCAACGGGAACGGCUUCU  
25 UCGCCACCAUGGCCUGGGCCGUCCCCAAGAACAGACCGCCACCAACCCCUAGACCA  
UCGAGGUGCCGUACAUCUGCACCGAGGGGGAGGACCAGAUACCGUCUGGGGCUUC  
CACAGCGACAACGAGACGCAGAUUGCCAAGCUGUACGGGGACUCCAAGCCCCAGAA  
GUUCACGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCU  
UCCCGAACACGAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGAC  
30 UACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCCU  
GCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUAAGGGCUCCC  
UGCCCCUCAUCGGGGAGGCCGACUGCCUGCACGAGAAGUACGGCGGGCUACAAG  
AGCAAGCCGUACUACACCGGCAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUG  
GGUCAAGACCCCGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCU  
35 GCUCAAGGAGCGCGGGUUCUUCGGCGCAUCGCCGGGUUCCUGGAGGGCGGGUGGG  
AGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCCACGGGGCCCACGGCGUGGCG  
GUCGCCCGGGACCUCAAGAGCACCCAGGAGGCCAUACAAGAUCACGAAGAACCU  
GAACUCCUCAGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGG  
ACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUCGACGACCUCCGGGCC  
40 GACACCAUCAGCUCGCGAGAUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAU  
CAACAGCGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGG  
GGCCUCGCGGGUGGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGCAAC  
CAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCGGGGAGUUCUCCUG  
CCGACCUUCGACAGCCUCAACAUCAGGCGGCCUCGCUGAACGACGACGGCCUCGAC  
45 AACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCUUGGCGGUGACGCUC

AUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUCGUCGAU  
CUGCCUGUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGGUUC  
CUUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUUAUGAAGGGCCUUGAGCA  
UCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAAAAAAAAA  
5 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
SEQ ID NO: 120

>I'-B-Yamagata-B-120A mRNA

GAGUUAAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
10 ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCCUGACCACCACCCCCACCAAGUCC  
UACUUCGCCAACCUGAAGGGCACCCGACCCGCGGCAAGCUGUGCCCCGACUGCCUG  
AACUGCACCGACCUUGGACGUGGCCCUUGGGCCGGCCCAUGUGCGUGGGCACACCCCC  
15 AGCGCCAAGGCCUCCAUCUGCAGGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC  
AUCAUGCACGACCGCACCAAGAUCGCGCAGCUGCCCAACCUGCUGCGGGGGUACGA  
GAAGAUCCGCCUGUCCACCCAGAACGUAUCGACGCCGAGAAGGCCCCCGCGGGGCC  
CUACCGGUGGGCACCGCGGUGCCUGCCCCAACGCCACCAGCAAGAUCGGCUUCUU  
CGCCACCAUGGCCUGGGCCGUGCCCAAGGACAACUACAAGAAGCCACCAACCCCCU  
20 GACCGUCGAGGUGCCCUACAUCUGCACCGAGGGGGAGGACCAGAUACGGUCUGGG  
GCUUCCACUCCGACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGC  
AGAAGUUCACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAUCCGG  
GACUCCCCGACACGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGAUCGUGGUC  
GACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAUCGUGUACCAGCGGGGGGU  
25 CCUGCUGCCGAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGA  
GCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCACGAGGAGUACGGGGGGCCUAC  
AAGUCCAAGCCGUACUACACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAU  
CUGGGUGAAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCGCCAA  
GCUGCUAAGGAGCGCGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCU  
30 GGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUC  
GCGGUGGCCGCGGACCUCAAGUCCACGCAGGAGGCCAUACAAGAUCACCAAGAA  
CCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGCGCGA  
UGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUGGACGACCUGCGG  
GCCGACACGAUCUCGUCCCAGAUUGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAU  
35 CAUCAACUCGGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAU  
UGGGCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAGUGC  
AACCAGACGUGCCUGCACCGAUCGCCGCGGGGACCUCAACGCCGCGGAGUUCAGC  
CUGCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUCCUGAACGACGACGGGCUC  
GACAACCACACGAUCCUGCUCUACUACAGCACCGCGGCCUGUCCUGGCGGUCACG  
40 CUCAUGCUGGCCAUUCUACUUGGUACAUGGUCAGCCGCGACAACGUGUCGUGCUC  
CAUCUGCCUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAUUAAGG  
UUCUUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUUAUGAAGGGCCUUGA  
GCAUCUGGAUUCUGCCUAAUAAAAACAUUUAUUUUCAUUGCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
45 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 121

>I'-H1N1-HA0-mut-B-120A mRNA  
GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUCACCACCGCCAAC  
5 GCCGACACCCUGUGCAUCGGCUACCACGCCAACAAACAGCACCGACACCGUGGACACC  
GUGCUGGAGAAGAACGUGACCGUGACCCACAGCGUGAACCUGCUGGAGGACAAGCA  
CAACGGCAAGCUGUGCAAGCUGCGGGGCGUGGCCCCCUGCACCUGGGCAAGUGCA  
ACAUCGCCGGCUGGAUCCUGGGCAACCCCGAGUGCGAGUCCUGAGCACCGCCCGCU  
CCUGGAGCUACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGGG  
10 GACUUCAUCAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGA  
GCGCUUCGAGAUCUUCCTCAAGACCUCAGCUGGCCCAACCACGACUCCGACAACGG  
CGUGACCGCCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUG  
GCUGGUGAAGAAGGGCAAGUCCUACCTCAAGAUAACCAGACCUACAUCAACGACA  
AGGGGAAGGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCACCAUCGCCGACC  
15 AGCAGAGCCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCTCGUAC  
AGCAAGAAGUUCAAGCCCGAGAUCCGACCCGCCCCAAGGUCCGGGACCAGGAGGG  
CCGCAUGAACUACUACUGGACCCUGGUGGAGCCGGGGACAAGAUCACCUUCGAGG  
CCACCGGCAACCUCGUCGCCCCGCGUACGCCUUCACGAUGGAGCGCGACGCCGGGU  
CCGGCAUCAUCAUCAGCGACACCTCCGUGCAGCAGCUGCAACACGACCUGCCAGACGC  
20 CCGAGGGGGCGAUAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCG  
GCAAGUGCCCCAAGUACGUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGC  
AACGUCCCCUCCAUCAGAGCCAGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGA  
GGGGGGCUGGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAAGACGAGC  
AGGGCUCCGGGUACGCGGCCGACCUCAGUCGACGCAGAACGCGAUCGACAAGAUC  
25 ACCAACAAGGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGG  
CAAGGAGUUAACACCUUGGAGAAGCGCAUCGAGAACCUCACCAAGAAGGUGGACG  
ACGGGUUCCUGGACAUCUGGACGUACAACGCGGAGCUCCUGGUCCUCCUGGAGAAC  
GAGCGGACCCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCG  
CAACCAGCUCAAGAACAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACC  
30 ACAAGUGCGACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUACCCG  
AAGUACAGCGAGGAGGCGAAGCUGAACCGGGAGAAGAUCGACGGGGUCAAGCUCGA  
CUCCACGCGCAUCUACCAGAUCCUGGCCAUUCUACUGACCGUGGCGAGCUCCUCGU  
CCUGGUGGUCUCGUCGCGGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCGUC  
AGUGCCGGAUCUGCAUCUGAUGACUCGAGGCUUCGUUUCUUGCUGUCCAAUUCUA  
35 UUAAGGUUCCUUUGUUCCTUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGG  
GCCUUGAGCAUCUGGAUUCUGCCUAAUAAAAACAUAUUAUUUCAUUGCAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAA

40

SEQ ID NO: 122

>I'-H3N2-HA0-mut-B-120A mRNA  
GAGUUAGGAGUUGGUGGCAGUUUCAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGACCAUCAUCGCCCUGAGCUACAUCUGGUGCCUGGUGUUCGCC  
CAGAAGAUCCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCC  
45 GUGCCCAACGGCACCAUCGUGAAGACCAUCACCAACGACCGGAUCGAGGUGACCAA

CGCCACCGAGCUGGUGCAGAACUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCA  
GAUCCUGGACGGCGGCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCAAGUG  
CGACGGGUUCCAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACU  
CCAACUGCUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCUGGUGGCCA  
5 GCUCCGGCACCCUGGAGUUCAAGAACGAGAGCUUCAACUGGACCGGGGUCAAGCAG  
AACGGCACCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCUCCCGGCUG  
AACUGGCUCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAAC  
AACGAGCAGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCGACAA  
GGACCAGAUUCUCCUCUUCGCCCAGCCAGCGGGCGCAUACCGUGUCCACGAAGCG  
10 GAGCCAGCAGGCCGUAUCCCCAACUCCGCUCCCGCCCGCGGAUCCGCGACAUCCC  
CAGCCGGAUUCUUAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCCUGCUCAUCA  
ACAGCACCGGCAACCUGAUCGCCCCCGCGGGUACUUAAGAUCGGGUCCGGCAAGA  
GCUCGAUCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCC  
CCCAACGGCUCGAUCCCGAACGACAAGCCCUUCCAGAACGUCAACCGGAUCACGUAC  
15 GGGGCCUGCCCGCGCUACGUGAAGCAGUCCACGCUCAAGCUGGCCACCGGCAUGCG  
GAACGUCCCCGAGAAGCAGACCCAGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCG  
AGAACGGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAAGACG  
GAGGGCCGCGGGCAGGCCGCGGACCUCAGUUCGACGACGCGGCGGAUCCGACAGAU  
CAACGGCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUCCACCAGAUCCG  
20 AGAAGGAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAG  
GACACCAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGCUCGUGGCGCUGGAGAA  
CCAGCACACCAUCGACCUCACGGACUCGGAGAUGAACAGCUGUUCGAGAAGACGA  
AGAAGCAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUUCAC  
CACAAGUGCGACAACGCGUGCAUCGGGUCCAUCGCAACGAGACCUACGACCACAA  
25 CGUCUACCGGGACGAGGCCCUGAACAACCGCUUCCAGAUCAAGGGCGUGGAGCUCA  
AGAGCGGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCCUC  
CUGUGCAUCGCCCUCUGGGCUUCAUCAUGUGGGCGUGCCAGAAGGGGAACAUCCG  
GUGCAACAUCUGCAUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUCUAU  
UAAAGGUUCCUUGUUCUUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGG  
30 CCUUGAGCAUCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUUAUUGCAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAA

SEQ ID NO: 123

35 >I-B-Victoria-HA0-mut-B-120A mRNA  
GAGUUAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCC  
40 CACUUCGCCAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCAAGUGCCU  
GAACUGCACCGACCUGGACGUGGGCCUGGGCCGCCCAAGUGCACCGGCAAGAUCUCC  
CAGCGCCCGGGUGUCCAUCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCC  
CAUCAUGCACGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGGGUACG  
AGCACGUCCGCCUGUCCACCCACAACGUGAUAACGCCGAGGACGCCCCGGCCGGC  
45 CCUACGAGAUCCGGACAGCGGCUCCUGCCCCAACAUACCAACGGGAACGGCUUCU

UCGCCACCAUGGCCUGGGCCGUCCCCAAGAACAAGACCGCCACCAACCCCCUGACCA  
UCGAGGUGCCGUACAUCUGCACCGAGGGGGAGGACCAGAUCACCGUCUGGGGCUUC  
CACAGCGACAACGAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCCAGAA  
GUUCACGAGCUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCU  
5 UCCCCAACCAGACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGAC  
UACAUGGUGCAGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCCU  
GCUGCCGCAGAAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUAAGGGCUC  
UGCCCCUCAUCGGGGAGGCCGACUGCCUGCACGAGAAGUACGGCGGGCUCAACAAG  
AGCAAGCCGUACUACACCGGCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUG  
10 GGUCAAGACCCCGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCCGGCCAAGCU  
GCUCAAGGAGCAGGGGUUCUUCGGCGCCAUCGCCGGGUUCCUGGAGGGCGGGUGGG  
AGGGCAUGAUCGCCGGGUGGCACGGCUACACGUCCACGGGGCCCACGGCGUGGCG  
GUCGCCGCGGACCUCAAGAGCACCCAGGAGGCCAUAACAAGAUCACGAAGAACCU  
GAACUCCUCAGCGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGG  
15 ACGAGCUCCACAACGAGAUCUGGAGCUCGACGAGAAGGUCGACGACCUGCGGGCC  
GACACCAUCAGCUCGCAGAUCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAU  
CAACAGCGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAGAAGAUGCUGG  
GGCCUCGGCGGUCGAGAUCGGCAACGGGUGCUUCGAGACCAAGCACAAGUGCAAC  
CAGACGUGCCUCGACCGCAUCGCCGCGGGCACGUUCGACGCCGGGGAGUUCUCCCU  
20 CCGACCUUCGACAGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUCGAC  
AACCACACCAUCCUGCUCUACUACUCCACCGCGGCCAGCUCGCGUGGCGGUGACGCUC  
AUGAUCGCCAUCUUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUCGCUCAU  
CUGCCUGUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCA AUUUAUUAAGGUUC  
CUUUGUUCUUAAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGAGCA  
25 UCUGGAUUCUGCCUAAUAAAAAACAUUUAUUUCAUUGCAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 124

&gt;I-B-Yamagata-HA0-mut-B-120A mRNA

30 GAGUUAAGGAGUUGGUGGCAGUUUCAAUAACAGGUCAUUGCCGAGAAAAGGAUAGC  
ACUAUAAUAUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGAC  
CGGAUCUGCACCGGCAUCACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACC  
CAGGGCGAGGUGAACGUGACCGGCGUGAUCCCCCUGACCACCACCCCCACCAAGUCC  
UACUUCGCCAACCUGAAGGGCACCCGACCCGCGGCAAGCUGUGCCCCGACUGCCUG  
35 AACUGCACCGACCUUGGACGUGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCACCCC  
AGCGCCAAGGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCC  
AUGAUGCACGACCGCACCAAGAUCGCGCAGCUGCCCAACCUGCUGCGGGGGUACGA  
GAAGAUCCGCCUGUCCACCCAGAACGUAUCGACGCCGAGAAGGCCCCCGCGGGGCC  
CUACCGGCUUGGGACACGCGGGUCCUGCCCCAACGCCACCAGCAAGAUCGGCUUCU  
40 CGCCACCAUGGCCUGGGCCGUGCCCCAAGGACAACUACAAGAAGCCACCAACCCCCU  
GACCGUCGAGGUGCCCUACAUCUGCACCGAGGGGGAGGACCAGAUCACGGUCUGGG  
GCUUCCACUCCGACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCCG  
AGAAGUUCACGAGCUCGCGCAACGGCGUGACCACGCACUACGUCAGCCAGAUCGGG  
GACUCCCCGGACACGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUC  
45 GACUACAUGAUGCAGAAGCCCGGGAAGACGGGCACCAUCGUGUACCAGCGGGGGU

CCUGCUGCCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGA  
GCCUGCCCCUCAUCGGCGAGGCCGACUGCCUGCACGAGGAGUACGGGGGCCUCAAC  
AAGUCCAAGCCGUACUACACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAU  
CUGGGUGAAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCGCCAA  
5 GCUGCUCAAGGAGCAGGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCU  
GGGAGGGGAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUC  
GCGGUGGCCGCGGACCUCAGUCCACGCAGGAGGCCAUAACAAGAUACCAAGAA  
CCUGAACAGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGCGGA  
UGGACGAGCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUGGACGACCUGCGG  
10 GCCGACACGAUCUCGUCCCAGAUCCGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAU  
CAUCAACUCGGAGGACGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAU  
UGGGCCCGUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAGUGC  
AACCAGACGUGCCUCGACCGCAUCGCCGCGGGGACCUCAACGCCGCGGAGUUCAGC  
CUGCCCACGUUCGACUCGCUCAACAUCACCGCGGCCUCCUGAACGACGACGGGCU  
15 GACAACCACACGAUCCUGCUCUACUACAGCACCAGCGGCCUCGUCCUGGCGGUCACG  
CUCAUGCUGGCCAUUCUACUUCGUGUACAUGGUCAGCCGCGACAACGUGUCGUGCUC  
CAUCUGCCUCUGAUGACUCGAGGCUCGCUUUCUUGCUGUCCAAUUAUUAAGG  
UUCUUUGUUCUUUAAGUCCAACUACUAAACUGGGGGAUAUUAUGAAGGGCCUUGA  
GCAUCUGGAUUCUGCCUAAUAAAAACAUAUUAUUUUAUUGCAAAAAAAAAAAAAA  
20 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID NO: 125

&gt;A 型流感病毒 H1N1 (A/Wisconsin/588/2019) HA mRNA

AUGAAGGCCAUCCUGGUGGUGAUGCUGUACACCUUCACCACCGCCAACGCCGACAC  
25 CCUGUGCAUCGGCUACCACGCCAACACAGCACCAGACACCGUGGACACCGUGCUGGA  
GAAGAACGUGACCGUGACCCACAGCGUGAACCUUGCUGGAGGACAAGCACAACGGCA  
AGCUGUGCAAGCUGCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCC  
GGCUGGAUCCUGGGCAACCCCGAGUGCGAGUCCUGAGCACCGCCCGCUCCUGGAGC  
UACAUCGUGGAGACCUCCAACAGCGACAACGGCACCUGCUACCCCGGGGACUUAU  
30 CAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGAGCGCUUCG  
AGAUCUUCCTCAAGACCUCAGCUGGCCCAACCACGACUCCGACAACGGCGUGACCG  
CCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUGGCUGGUG  
AAGAAGGGCAAGUCCUACCCCAAGAUCAACCAGACCUACAUAACGACAAGGGGAA  
GGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCACCAUCGCCGACCAGCAGAG  
35 CCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCGCCGUACAGCAAGA  
AGUUCAAGCCCGAGAUCCGACCCGCCCCAAGGUCCGGGACCAGGAGGGCCGCAUG  
AACUACUACUGGACCCUGGUGGAGCCGGGGACAAGAUAACCUUCGAGGCCACCGG  
CAACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGUCCGGCAU  
CAUCAUCAGCGACACCCCGUGCACGACUGCAACACGACCUGCCAGACGCCCGAGGG  
40 GGCGAUCAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUG  
CCCCAAGUACGUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCCC  
CUCCAUCACAGAGCCGGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGGGGCU  
GGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACAGAACGAGCAGGGCUCC  
GGGUACGCGGCCGACCUCAGUCCGACGACGAGAACGCGAUCGACAAGAUAACCAACA  
45 GGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGGCAAGGAGU

UCAACCACCUUGGAGAAGCGCAUCGAGAACCUCAACAAGAAGGUGGACGACGGGUUC  
CUGGACAUCUGGACGUACAACGCGGAGCUCCUGGUCCUCCUGGAGAACGAGCGGAC  
CCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGC  
UCAAGAACAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGC  
5 GACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUCACGACUACCCGAAGUACAG  
CGAGGAGGCGAAGCUGAACC GGGAGAAGAU CGACGGGUCAAGCUCGACUCCACGC  
GCAUCUACCAGA UCCUGGCCAU CUACUCGACCGUGGCGAGCUCCUCGUCCUGGUG  
GUCUCGCUCGGCGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCUGCAGUGCCG  
GAUCUGCAUCUGA

10

SEQ ID NO: 126

>A 型流感病毒 H3N2 (A/Cambodia/e0826360/2020) HA mRNA  
AUGAAGACCAUCAUCGCCCCUGAGCUACAUCUUGGCCUGGUGUUCGCCCAGAAGAU  
CCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCCGUGCCCAA  
CGGCACCAUCGUGAAGACCAUACCAACGACCGGAUCGAGGUGACCAACGCCACCG  
15 AGCUGGUGCAGAACUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGA UCCUG  
GACGGCGGCAACUGCACCCUGAUCGACGCCCUGCUUGGGCGACCCCCAGUGCGACGGG  
UUC CAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACUCCAACUG  
CUACCCCUACGACGUGCCCCGACUACGCCAGCCUGCGGUCCCUUGGUGGCCAGCUCCGG  
CACCCUGGAGUUAAGAACGAGAGCUUAACUGGACCGGGGUCAAGCAGAACGGCA  
20 CCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCUCCCGGCUGAACUGGC  
UCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAACAACGAGC  
AGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCGACAAGGACCAAG  
AUCUCCCUUCUUCGCCCAGCCCAGCGGGCGCAUCACCGUGUCCACGAAGCGGAGCCAG  
CAGGCCGUAUCCCCAACAU CGGCUCCCGCCCGCGGAUCCGCGACA UCCCCAGCCGG  
25 AUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACA UCCUGCUCAUAACAGCAC  
CGGCAACCUGAUCGCCCCCGCGGGUACUUAAGA UCCGGUCCGGCAAGAGCUCGA  
UCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCCCCAAC  
GGCUCGAUCCCGAACGACAAGCCCUCCAGAACGUAACCGGAUCACGUACGGGGC  
CUGCCCCGCGCUACGUGAAGCAGUCCACGCUAAGCUGGCCACCGGCAUGCGGAACG  
30 UCCCCGAGAAGCAGACCCGCGGGAUUCUUCGGCGCCAUCGCGGGGUUCAUCGAGAAC  
GGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAACAGCGAGGG  
CCGCGGGCAGGCCGCGGACCUC AAGUCGACGACGAGCCGCGAU CGACCAGAUCAACG  
GCAAGCUGAACC GGCUCAUCGGGAAGACGAACGAGAAGU UCCACCAGAU CGAGAAG  
GAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAGGACAC  
35 CAAGAUCGACCUCUGGAGCUACAACGCCGAGCUGCUCGUGGCGCUGGAGAACCAGC  
ACACCAUCGACCUCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAG  
CAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUAAGAUCUACCACAA  
GUGCGACAACGCGUGCAUCGGGUCCAUCGCAACGAGACCUACGACCACAACGUCU  
ACCGGGACGAGGCCUGAACAAACCGCUUCCAGAUCAAGGGCGUGGAGCUCAAGAGC  
40 GGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUGUCCUGCUUCCUCCUGUG  
CAUCGCCCUCUUGGGCUUCAUGUGGGCGUGCCAGAAAGGGGAACA UCCGGUGCA  
ACAUCUGCAUCUGA

SEQ ID NO: 127

>B 型流感病毒 Victoria 谱系 (B/Washington/02/2019) HA mRNA  
45 AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACCGGAUCUG

CACCGGCAUACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACCCAGGGCGA  
GGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCCACUUCGC  
CAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCA  
CCGACCUGGACGUGGGCCUGGGCCGCCCCAAGUGCACCGGCAAGAUCCCCAGCGCCC  
5 GGGUGUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGC  
ACGACCGCACCAAGAUCCGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGCACGUCC  
GCCUGUCCACCCACAACGUGAUCAACGCCGAGGACGCCCCGGCCGGCCUACGAGA  
UCGGGACCAGCGGCUCCUGCCCCAACAUACCAACGGGAACGGCUUCUUCGCCACCA  
UGGCCUGGGCCGUCCCCAAGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGC  
10 CGUACAUCUGCACCGAGGGGGAGGACCAGAUCACCGUCUGGGGCUUCCACAGCGAC  
AACGAGACGCAGAUUGGCAAGCUGUACGGGGACUCCAAGCCCCAGAAGUUCACGAG  
CUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCGGGGGCUUCCCGAACCA  
GACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGACUACAUGGUGC  
AGAAGUCCGGCAAGACGGGGACCAUCACGUACCAGCGGGGCAUCCUGCUGCCGAG  
15 AAGGUCUGGUGCGCCUCCGGGCGCAGCAAGGUGAUCAAGGGCUCCUGCCCCUCAU  
CGGGGAGGCCGACUGCCUGCACGAGAAGUACGGCGGGCUCAACAAGAGCAAGCCGU  
ACUACACCGGCGAGCACGCCAAGGCCAUCGGGAACUGCCCCAUCUGGGUCAAGACCC  
CGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCCGGCCAAGCUGCUCAAGGAGC  
GCGGGUUCUUCGGCGCAUCGCCGGGUUCCUGGAGGGCGGGUGGGAGGGCAUGAUC  
20 GCCGGGUGGCACGGCUACACGUCCCACGGGGCCCACGGCGUGGCGGUCGCCGCGGAC  
CUCAAGAGCACCCAGGAGGCCAUCAACAAGAUACGAAGAACCUGAACUCCUCAG  
CGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGGACGAGCUCCACA  
ACGAGAUCCUGGAGCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGC  
UCGCAGAUCCGAGCUCGCGGUGCUGCUCUCCAACGAGGGCAUCAUAACAGCGAGGA  
25 CGAGCACCUCCUGGCCUGGAGCGGAAGCUCUAGAAGAUGCUGGGGGCCUCGGCGG  
UCGAGAUCCGCAACGGGUGCUUCGAGACCAAGCACAAAGUGCAACCAGACGUGCCUC  
GACCGCAUCGCCCGGGGACGUUCGACGCCGGGAGUUCUCCUGCCGACCUUCGAC  
AGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUCGACAACCACACCAUC  
CUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUC  
30 UUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGA

SEQ ID NO: 128

&gt;B 型流感病毒/Yamagata 谱系 (B/Phuket/3073/2013) HA mRNA

AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACCGGAUCUG  
CACCGGCAUACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACCCAGGGCGA  
35 GGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCAAGUCCUACUUCGC  
CAACCUGAAGGGCACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCAC  
CGACCUGGACGUGGGCCUGGGCCGGCCCAUGUGCGUGGGCACCAACCCAGCGCCAA  
GGCCUCCAUCCUGCACGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCA  
CGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGAAGAUGC  
40 GCCUGUCCACCCAGAACGUCAUCGACGCCGAGAAGGCCCCCGGCGGGCCUACCGGC  
UGGGCACAGCGGGUCCUGCCCCAACGCCACCAGCAAGAUCGGCUUCUUCGCCACCA  
UGGCCUGGGCCGUGCCCCAAGGACAACUACAAGAAGCCACCAACCCUGACCGUCG  
AGGUGCCCUACAUCUGCACCGAGGGGGAGGACCAGAUCACGGUCUGGGGCUUCCAC  
UCCGACAACAAGACCCAGAUGAAGUCCUGUACGGGGACUCCAACCCGCAGAAGUU  
45 CACGAGCUCCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAUCGGGGACUCCCC

GGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUCGACUACA  
UGAUGCAGAAGCCCGGGAAGACGGGCACCAUCGUGUACCAGCGGGGGGUCCUGCUG  
CCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCC  
CCUCAUCGGCGAGGCCGACUGCCUGCACGAGGAGUACGGGGGCCUCAACAAGUCCA  
5 AGCCGUACUACACGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCC GAUCUGGGUG  
AAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCUGCUC  
AAGGAGCGCGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGUGGGAGGG  
GAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCCACGGGGUCGCGGUGG  
CCGCGGACCUCAAGUCCACGCAGGAGGCCAUACAAGAUCACCAAGAACCUGAAC  
10 AGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGCGCGAUGGACGA  
GCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACA  
CGAUCUCGUCCCAGAUUCGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUAAC  
UCGGAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUUCUGGGCCC  
GUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAUGCAACCAGA  
15 CGUGCCUCGACCGCAUCGCCGCGGGGACCUUAACGCCGCGGAGUUCAGCCUGCCCA  
CGUUCGACUCGCUCAACAUCACCGCGGCCUCCUGAACGACGACGGGCUCGACAACC  
ACACGAUCCUGCUCUACUACAGCACC GCGGCCUCGUCCUGGCGGUCACGCUCAUGC  
UGGCCAUUCUUAUCGUGUACAUGGUCAGCCGCGACAACGUGUCGUGCUCUACUGC  
CUCUGA

20

SEQ ID NO: 129

&gt; SARS-COV-2 BA.4 刺突 mRNA

AUGUUCGUGUCCUGGUGCUGCUGCCCCUGGUGAGCUCUCAGUGCGUGAACCUGAU  
CACAAGAACACAGUCGUACACCAACAGCUUCACAAGAGGCGUGUACUACCCCGACA  
25 AGGUGUUCAGAAGCAGCGUGCUGCACUCCACCCAAGACCUGUUUCUGCCCUUCUUC  
AGCAACGUGACCUGGUUCCACGCCAUCAGCGGCACCAACGGCACCAAGAGAUUCGA  
CAACCCCGUGCUGCCCUUAACGACGGCGUGUACUUCGCUAGCACCGAAAAGAGCA  
ACAUCAUCAGAGGCUGGAUCUUCGGCACCAACCUAGACUCCAAGACACAAAGCCUG  
CUGAUCGUUAACAACGCGACCAACGUGGUGAUAAGGUGUGCGAGUUUCAGUUCUG  
30 CAACGACCCCUUCCUGGACGUGUACUACCACAAGAACAACAAGAGCUGGAUGGAGA  
GCGAGUUCAGAGUCUACAGCAGCGCCAACAACUGCACCUCGAGUACGUGUCUCAG  
CCCUUCCUGAUUGGACCUGGAGGGCAAGCAAGGCAAUUUAAGAACCUAAGAGAGUU  
CGUGUUAAGAACAUCGACGGCUACUUAAGAUCUACAGCAAGCACACCCCCAUCA  
ACCUGGGGAGAGACCUGCCCCAAGGCUUCAGCGCCUGGAGCCCCUGGUGGACCUG  
35 CCCAUCGGCAUCAACAUCACAAGAUUUCAGACCCUGCUGGCCCUGCACCGUAGCUAC  
CUAACCCCGGCGAUAGUUCUUCUGGCUGGACGGCCGGCGCUGCCGCCUACUACGUG  
GGCUACCUGCAGCCUAGAACCUCUGCUGAAGUACAACGAGAACGGCACUAUCAC  
UGAUGCCGUGGACUGUGCCUGGACCCUGAGCGAGACCAAGUGCACCCUCAAGA  
GCUUACCGUGGAGAAGGGCAUCUAUCAGACAAGCAACUUCAGAGUGCAGCCCACC  
40 GAGAGCAUCGUGAGAUUCCCCAACAUACCAACCUGUGCCCCUUCGACGAGGUGUU  
CAACGCCACAAGAUUCGCUAGCGUGUACGCGUGGAACAGAAAGAGAAUCAGCAACU  
GCGUGGGCCGACUACAGCGUGCUGUACAACUUCGCUCCCUUCUUCGCCUUAAGUGC  
UAUGGCGUCAGCCCCACCAAGCUGAACGACCUGUGCUUACCAACGUGUACGCCGA  
CAGCUUCGUGAUCAGAGGCAACGAGGUGAGCCAGAUUGCCCCCGGGCAAACCGGCA  
45 ACAUCGCCGACUACAACUACAAGCUGCCCGACGACUUCACCGGCUGCGUGAUCGCG

UGGAACAGCAAUAAGCUGGAUUGCAAGGUGGGCGGCAACUACAACUACCGGUACAG  
ACUGUUCAGAAAGAGCAACCUGAAGCCCUUCGAGAGAGACAUCAGCACCGAGAUCU  
ACCAAGCCGGCAACAAGCCUGCAACGGCGUGGCGGGCGUCAACUGCUACUUCCCCC  
UGCAGAGCUACGGCUUUAGGCCACCUCACGGCGUGGGCCAUCAGCCCUACAGAGUG  
5 GUGGUGCUGAGCUUCGAGCUGCUGCACGCCCCCGCAACCGUGUGCGGCCCCAAAA  
GAGACCAACCUUGGUGAAGAACAAGUGCGUGAACUUAACUUUAUGGCCUUAAGG  
GCACGGGCGUGCUGACCGAGAGCAACAAGAAGUCCUGCCAUUUCAGCAAUUCGGC  
AGAGACAUCGCCGACACCACCGACGCGGUCAGAGACCCUCAGACCCUGGAGAUCCU  
GGACAUCACCCCCUGUAGCUUCGGCGGGCUGAGCGUGAUCACCCCCGGCACUAACAC  
10 AAGCAACCAAGUGGCCGUGCUGUACCAAGGCGUGAACUGCACCGAGGUGCCCCGUGG  
CCAUCCACGCCGAUCAGCUGACCCCCACCUGGAGAGUAUACAGCACCGGAAGCAACG  
UGUUUCAGACAAGAGCCGGCUGCCUGAUCGGCGCCGAGUACGUGAACAAACAGCUAC  
GAGUGCGACAUCCCCAUCGGCGCCGGCAUCUGCGCUAGCUAUCAGACACAGACCAA  
GAGCCACCGGAGAGCUAGAAGCGUGGCUUCUCAGAGCAUCAUCGCCUACACCAUGA  
15 GCCUGGGCGCCGAGAACAGCGUGGCCUACAGCAACAACAGCAUCGCCAUCCCCACCA  
ACUUCACCAUCAGCGUGACCACCGAGAUCUCCUGCCCGUAAGUAUGACCAAGACAAGC  
GUGGACUGCACCAUGUACAUCUGCGGGCGACAGCACCGAGUGCAGCAACCUGCUGCU  
GCAGUACGGCAGCUUCUGCACACAGCUGAAGAGAGCCUGACCGGCAUCGCCGUGG  
AGCAAGACAAGAACACCCAAGAGGUGUUCGCCCAAGUGAAGCAGAUCAACAAGACC  
20 CCCCCCAUCAAGUACUUCGGCGGCUUAACUUCAGCCAAAUCCUGCCCGACCCUAGC  
AAGCCUAGCAAGCGGAGCUUCAUCGAGGACCUGCUGUUAACAAGGUGACCCUGGC  
CGACGCCGGCUUCAUAAGCAGUACGGCGACUGCCUAGGCGACAUCGCAGCUAGAG  
ACCUGAUCUGCGCUCAGAAGUUAUUGGACUCACCGUGCUCCCCCCCCUGCUGACCG  
ACGAGAUGAUCGCUCAGUACACCUCUGCCCUGCUGGCGGGGAACCAUAACAAGCGGU  
25 UGGACAUUUGGCGCCGGGGCCGCACUGCAAAUCCCCUUCGCCAUGCAGAUGGCCUA  
CAGAUUCAACGGCAUCGGCGUGACACAGAACGUGCUGUACGAGAAUCAGAAGCUGA  
UCGCCAAUCAGUUAACAGCGCCAUCGGCAAGAUAAGACAGCCUGAGCAGCACC  
GCUAGCGCCCCUGGGCAAGCUGCAAGACGUGGUGAAUACAACGCCCAAGCCCUGAA  
CACCCUGGUGAAGCAGCUGAGCAGCAAGUUCGGCGCCAUCAGCAGCGUAUUGAAUG  
30 ACAUCCUGUCUAGACUGGACCCCCCGAGGCCGAGGUGCAGAUUCGACAGACUGAUC  
ACCGGCAGACUGCAGAGCCUGCAGACCUACGUGACACAGCAGCUGAUCAGAGCCGC  
CGAGAUCAGAGCUAGCGCCAACCUUGGCCGCCACCAAGAUGAGCGAGUGCGUGCUGG  
GGCAGAGCAAGAGAGUGGACUUCUGCGGCAAGGGCUACCACCUGAUGAGCUUCCCU  
CAGAGCGCCCCCACGGCGUGGUGUUCUGCAGCUGACCUACGUGCCCGCCCCAAGAG  
35 AAGAACUUCACCACCGCUCCCGCCAUCUGCCACGACGGCAAGGGCCACUUCCCUAGA  
GAGGGCGUGUUCGUGAGCAACGGCACCCACUGGUUCGUGACACAGAGAAACUUCUA  
CGAGCCUCAGAUCAUCACCACCGACAACACCUUCGUGAGCGGCAACUGCGACGUGG  
UGAUCGGCAUAGUGAACAACACCGUAUACGACCCCCUGCAGCCCGAGCUGGACAGC  
UUAAGGAGGAGCUGGACAAGUACUUAACCAUACUAGCCCCGACGUGGACCU  
40 GGGAGACAUCUCUGGCAUCAACGCUAGCGUGGUGAACAUCAGAAGGAGAUCGACC  
GGCUGAAUGAGGUGGCCAAGAACCUGAACGAGAGCCUGAUCGACCUGCAAGAGCUG  
GGCAAGUACGAGCAGUACAUAAGUGGCCUGGUACAUCUGGCUGGGCUUCAUCGC  
CGGCCUGAUCGCCAUCGUGAUGGUGACCAUCAUGCUGUGCUGCAUGACAAGCUGUU  
GCAGCUGCCUGAAAGGUUGCUGCAGCUGUGGCUCGUGCUGCAAGUUCGACGAGGAC  
45 GACAGCGAGCCCGUGCUGAAGGGCGUGAAGCUGCACUACCCUGAUGA

SEQ ID NO: 130

&gt;A 型流感病毒 H1N1 HA0 突变体 mRNA

AUGAAGGCCAUCCUGGUGUGAUGCUGUACACCUUCACCACCGCCAACGCCGACAC  
CCUGUGCAUCGGCUACCACGCCAACACAGCACCGACACCGUGGACACCGUGCUGGA  
5 GAAGAACGUGACCGUGACCCACAGCGUGAACCUUGCUGGAGGACAAGCACAAACGGCA  
AGCUGUGCAAGCUGCGGGGCGUGGGCCCCCUGCACCUGGGCAAGUGCAACAUCGCC  
GGCUGGAUCCUGGGCAACCCCGAGUGCGAGUCCUGAGCACCGCCCGCUCCUGGAGC  
UACAUCGUGGAGACCUCAACAGCGACAACGGCACCUGCUACCCCGGGGACUUAU  
CAACUACGAGGAGCUGCGGGAGCAGCUGUCCAGCGUGUCCAGCUUCGAGCGCUUCG  
10 AGAUCUUCCTCAAGACCUCCAGCUGGCCCCAACCACGACUCCGACAACGGCGUGACCG  
CCGCCUGCCCCACGCCGGGGCCAAGAGCUUCUACAAGAACCUGAUCUGGCUGGUG  
AAGAAGGGCAAGUCCUACCTCAAGAUCAACCAGACCUACAUCAACGACAAGGGGAA  
GGAGGUCCUCGUGCUGUGGGGCAUCCACCACCCCCCACCACUGCCGACCAGCAGAG  
CCUCUACCAGAACGCCGACGCCUACGUCUUCGUGGGGACCUCCTGGUACAGCAAGA  
15 AGUUCAAGCCCGAGAUCGCCACCCGCCCAAGGUCCGGGACCAGGAGGGCCGCAUG  
AACUACUACUGGACCCUGGUGGAGCCGGGGGACAAGAUCACCUUCGAGGCCACCGG  
CAACCUCGUCGCCCCGCGGUACGCCUUCACGAUGGAGCGCGACGCCGGGUCCGGCAU  
CAUCAUCAGCGACACCCCGUGCAGCAGUGCAACACGACCUGCCAGACGCCCGAGGG  
GGCGAUCAACACCUCCUGCCGUUCCAGAACGUCCACCCGAUCACGAUCGGCAAGUG  
20 CCCCCAGUACGUGAAGUCCACCAAGCUCCGGCUGGCCACGGGGCUCCGCAACGUCC  
CUCCAUCAGAGCCAGGGCCUGUUCGGGGCGAUCGCCGGCUUCAUCGAGGGGGGCU  
GGACCGGGAUGGUGGACGGCUGGUACGGGUACCACCACCAGAACGAGCAGGGCUCC  
GGGUACGCGGCCGACCUCAGUCGACGCAGAACGCGAUCGACAAGAUCACCAACAA  
GGUCAACAGCGUGAUCGAGAAGAUGAACACGCAGUUCACCGCCGUCGGCAAGGAGU  
25 UCAACCACCUUGGAGAAGCGCAUCGAGAACCUCACAAGAAGGUGGACGACGGGUUC  
CUGGACAUCUGGACGUACAACGCGGAGCUCCUGGUCCUCCUGGAGAACGAGCGGAC  
CCUCGACUACCACGACUCCAACGUGAAGAACCUGUACGAGAAGGUCCGCAACCAGC  
UCAAGAACAACGCCAAGGAGAUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGC  
GACAACACGUGCAUGGAGUCGGUGAAGAACGGCACCUACGACUACCCGAAGUACAG  
30 CGAGGAGGCGAAGCUGAACCGGGAGAAGAUACGACGGGGUCAAGCUCGACUCCACGC  
GCAUCUACCAGAUCUUGGCCAUUCUACUCGACCGUGGCGAGCUCCUCUGUCCUGGUG  
GUCUCGUCGCGGCCAUCAGCUUCUGGAUGUGCUCCAACGGGUCGCUAGUGCCG  
GAUCUGCAUCUGA

SEQ ID NO: 131

35 &gt;A 型流感病毒 H3N2 HA0 突变体 mRNA

AUGAAGACCAUCAUCGCCCUGAGCUACAUCUUGUGCCUGGUGUUCGCCCAGAAGAU  
CCCCGGCAACGACAACAGCACCGCCACCCUGUGCCUGGGCCACCACGCCGUGCCCAA  
CGGCACCAUCGUGAAGACCAUCACCAACGACCGGAUCGAGGUGACCAACGCCACCG  
AGCUGGUGCAGAACUCCAGCAUCGGCGAGAUCUGCGACUCCCCCACCAGAUCCUG  
40 GACGGCGGCAACUGCACCCUGAUCGACGCCUGCUGGGCGACCCCCAGUGCGACGGG  
UUCGAGAACAAGGAGUGGGACCUGUUCGUGGAGCGGAGCCGCGCCAACUCCAACUG  
CUACCCCUACGACGUGCCCGACUACGCCAGCCUGCGGUCCUGGUGGCCAGCUCCGG  
CACCCUGGAGUUAAGAACGAGAGCUUAACUGGACCGGGGUCAAGCAGAACGGCA  
CCUCCAGCGCCUGCAUCCGCGGGUCCAGCUCCAGCUUCUUCUCCCGGCUGAACUGGC  
45 UCACCCACCUGAACUACACCUACCCCGCCCUAACGUGACCAUGCCCAACAACGAGC

AGUUCGACAAGCUGUACAUCUGGGGCGUCCACCACCCAGCACCAGACAAGGACCAG  
AUCUCCUCUUCGCCCAGCCCAGCGGGCGCAUCACCGUGUCCACGAAGCGGAGCCAG  
CAGGCCGUAUCCCCAACAUCCGGCUCCCGCCCGCGGAUCCGCGACAUCCCCAGCCGG  
AUCUCCAUCUACUGGACGAUCGUGAAGCCGGGGGACAUCUUGCUCUACAACAGCAC  
5 CGGCAACCUGAUCGCCCCCGCGGGUACUUCAAGAUCGCGUCCGGCAAGAGCUCGA  
UCAUGCGCUCCGACGCCCCGAUCGGGAAGUGCAAGAGCGAGUGCAUACCCCCAAC  
GGCUCGAUCCCGAACGACAAGCCUUCAGAACGUAACCGGAUACGUACGGGGC  
CUGCCCGCGCUACGUGAAGCAGUCCACGCUAAGCUGGCCACCGGAUGCGGAACG  
UCCCCGAGAAGCAGACCCAGGGGAUCUUCGGCGCCAUCGCGGGGUUCAUCGAGAAC  
10 GGCUGGGAGGGGAUGGUGGACGGCUGGUACGGGUUCCGGCACCAGAACAGCGAGGG  
CCGCGGGCAGGCCGCGGACCUCAGUUCGACGACGAGGCCGCGAUCCGACCAGAUCAACG  
GCAAGCUGAACCGGCUCAUCGGGAAGACGAACGAGAAGUUCACCAGAUCCGAGAAG  
GAGUUCUCCGAGGUCGAGGGCCGCGUGCAGGACCUGGAGAAGUACGUCGAGGACAC  
CAAGAUCCGACCUCUGGAGCUACAACGCCGAGCUGCUCUGGGCGUGGAGAACCAGC  
15 ACACCAUCGACCUCACGGACUCGGAGAUGAACAAGCUGUUCGAGAAGACGAAGAAG  
CAGCUCCGGGAGAACGCCGAGGACAUGGGGAACGGCUGCUUCAAGAUCUACCACAA  
GUGCGACAACGCGUGCAUCGGGUCCAUCGCAACGAGACCUACGACCACAACGUCU  
ACCGGGACGAGGCCUGAACAACCGCUUCAGAUCAAGGGCGUGGAGCUCAAGAGC  
GGGUACAAGGACUGGAUCCUGUGGAUCUCGUUCGCGAUUGUCCUGCUUCCUCCUGUG  
20 CAUCGCCCUCUGGGCUUCAUGUGGGCGUGCCAGAAGGGGAACAUCGCGUGCA  
ACAUCUGCAUCUGA

SEQ ID NO: 132

>B 型流感病毒/Victoria HA0 突变体 mRNA  
AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACCGGAUCUG  
25 CACCGGCAUACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACCCAGGGCGA  
GGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCACCAAGUCCACUUCGC  
CAACCUGAAGGGCACCGAGACCCGGGGCAAGCUGUGCCCCAAGUGCCUGAACUGCA  
CCGACCUGGACGUGGCCUGGGCCGCCCAAGUGCACCAGGCAAGAUCCCCAGCGCCC  
GGGUGUCCAUCUGCAGGAGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGC  
30 ACGACCGCACCAAGAUCGCGCAGCUGCCCAACCUUGCUGCGGGGGUACGAGCACGUCC  
GCCUGUCCACCCACAACGUGAUAACGCCGAGGACGCCCCGGCCGGCCUACGAGA  
UCGGGACCAGCGGUCCUGCCCCAACAUACCAACGGGAACGGCUUCUUCGCCACCA  
UGGCCUGGGCCGUCCCCAAGAACAAGACCGCCACCAACCCCCUGACCAUCGAGGUGC  
CGUACAUCUGCACCGAGGGGGAGGACCAGAUACCGUCUGGGGCUUCCACAGCGAC  
35 AACGAGACGCAGAUGGCCAAGCUGUACGGGGACUCCAAGCCCAGAGUUCACGAG  
CUCCGCCAACGGCGUGACCACCCACUACGUCAGCCAGAUCCGGGGCUUCCCGAACCA  
GACGGAGGACGGGGGCCUGCCCCAGUCCGGGCGCAUCGUGGUCGACUACAUGGUGC  
AGAAGUCCGGCAAGACGGGGACCAUACGUACCAGCGGGGAUCCUGCUGCCGCAG  
AAGGUCUGGUGCGCUCCGGGCGCAGCAAGGUGAUCAAGGGCUCCUUGCCCCUCAU  
40 CGGGGAGGCCGACUGCCUGCACGAGAAGUACGGCGGGCUCAACAAGAGCAAGCCGU  
ACUACACCGGCGAGCACGCAAGGCCAUCCGGGAACUGCCCCAUUGGGUCAAGACCC  
CGCUGAAGCUCGCCAACGGCACGAAGUACCGGCCCGGCCAAGCUGCUCAAGGAGC  
AGGGGUUCUUCGGCGCAUCGCCGGGUUCCUGGAGGGCGGGUGGGAGGGCAUGAUC  
GCCGGGUGGCACGGCUACACGUCCACGGGGCCCACGGCGUGGCGGUCGCCGCGGAC  
45 CUCAAGAGCACCCAGGAGGCCAUCAACAAGAUACGAAGAACCUGAACUCCUCCAG

CGAGCUGGAGGUGAAGAACCUCAGCGCCUGUCCGGGGCGAUGGACGAGCUCACACA  
ACGAGAUCCUGGAGCUCGACGAGAAGGUCGACGACCUGCGGGCCGACACCAUCAGC  
UCGCAGAUCCGAGCUCGCGGUGCUCUCUCCAACGAGGGCAUCAUAACAGCGAGGA  
CGAGCACCUGCUCGCCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGGCCUCGGCGG  
5 UCGAGAUCCGCAACGGGUGCUUCGAGACCAAGCACAAAGUGCAACCAGACGUGCCUC  
GACCGCAUCGCCGCGGGCACGUUCGACGCCGGGAGUUCUCCCUGCCGACCUUCGAC  
AGCCUCAACAUCACGGCGGCCUCGCUGAACGACGACGGCCUCGACAACCACACCAUC  
CUGCUCUACUACUCCACCGCGGCCAGCUCGCUGGCGGUGACGCUCAUGAUCGCCAUC  
UUCGUCGUGUACAUGGUCUCCCGCGACAACGUGAGCUGCUCGAUCUGCCUGUGA  
10 SEQ ID NO: 133

>B 型流感病毒/Yamagata HA0 突变体 mRNA  
AUGAAGGCCAUCAUCGUGCUGCUGAUGGUGGUGACCAGCAACGCCGACCGGAUCUG  
CACCGGCAUACCAGCUCCAACAGCCCCACGUGGUGAAGACCGCCACCCAGGGCGA  
GGUGAACGUGACCGGCGUGAUCCCCUGACCACCACCCCCACCAAGUCCUACUUCGC  
15 CAACCUGAAGGGCACCCGGACCCGCGGCAAGCUGUGCCCCGACUGCCUGAACUGCAC  
CGACCUGGACGUGGCCCUGGGCCGGCCCAUGUGCGUGGGCACCAACCCCCAGCGCCAA  
GGCCUCCAUCUGCAGGAGGUGCGCCCCGUGACCAGCGGCUGCUUCCCCAUCAUGCA  
CGACCGCACCAAGAUCGGGCAGCUGCCCAACCUGCUGCGGGGGUACGAGAAGAUGC  
GCCUGUCCACCCAGAACGUAUCGACGCCGAGAAGGCCCCCGCGGGGCCUACCGGC  
20 UGGGCACCAGCGGGUCCUGCCCCAACGCCACCAGCAAGAUCCGCUUCUUCGCCACCA  
UGGCCUGGGCCGUGCCCAAGGACAACUACAAGAAGCCACCAACCCCCUGACCGUCG  
AGGUGCCCUACAUCUGCACCAGGGGGGAGGACCAGAUACGGUCUGGGGCUUCCAC  
UCCGACAACAAGACCCAGAUAGAAGUCCUGUACGGGGACUCCAACCCGCAGAAGUU  
CACGAGCUCGCCAACGGCGUGACCACGCACUACGUCAGCCAGAUCCGGGGACUCCCC  
25 GGACCAGACCGAGGACGGCGGGCUGCCCCAGUCCGGCCGCAUCGUGGUCGACUACA  
UGAUGCAGAAAGCCCGGAAGACGGGCACCAUCGUGUACCAGCGGGGGGUCCUGCUG  
CCGCAGAAGGUGUGGUGCGCCAGCGGCCGCUCCAAGGUCAUCAAGGGGAGCCUGCC  
CCUCAUCGGCGAGGCCGACUGCCUGCAGGAGGAGUACGGGGGCCUCAACAAGUCCA  
AGCCGUACUACAGGGGAAGCACGCCAAGGCCAUCGGCAACUGCCCGAUCUGGGUG  
30 AAGACCCCCUGAAGCUCGCCAACGGGACGAAGUACCGGCCGCCCGCCAAGCUGCUC  
AAGGAGCAGGGCUUCUUCGGGGCCAUCGCCGGCUUCCUGGAGGGGGGCUUGGAGGG  
GAUGAUCGCCGGCUGGCACGGGUACACCAGCCACGGCGCCACGGGGUCGCGGUGG  
CCGCGGACCUCAGUCCACGCAGGAGGCCAUACAAGAUAACCAAGAACCUGAAC  
AGCCUCUCCGAGCUGGAGGUCAAGAACCUCAGCGCCUGAGCGGCGCGAUGGACGA  
35 GCUCCACAACGAGAUCCUGGAGCUCGACGAGAAGGUGGACGACCUGCGGGCCGACA  
CGAUCUCGUCCAGAUCCGAGCUCGCGGUCCUGCUCAGCAACGAGGGGAUCAUAAC  
UCGGAGGACGAGCACCUGCUCGCCUGGAGCGGAAGCUCAAGAAGAUGCUGGGCCC  
GUCCGCGGUGGACAUCGGGAACGGCUGCUUCGAGACCAAGCACAAAGUGCAACCAGA  
CGUGCCUCGACCGCAUCGCCGCGGGGACCUCAACGCCGGCGAGUUCAGCCUGCCCA  
40 CGUUCGACUCGCUCAACAUCACCGCGGCCUCCUGAACGACGACGGGCUCGACAACC  
ACACGAUCCUGCUCUACUACAGCACCGCGGCCUCGUCCCUGGCGGUCACGCUCAUGC  
UGGCCAUUCUUAUCGUGUACAUGGUCAGCCGCGACAACGUGUCGUGCUCAUCUGC  
CUCUGA

SEQ ID NO: 134

45 &gt;Poly A

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

SEQ ID NO: 135

5 > SARS-COV-2 BA.4 刺突 氨基酸序列  
MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWF  
HAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLIVNNATNVVIKQ  
CEFQFCNDPFLDVYHKNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNL  
REFVFKNIDGYFKIYSKHTPINLGRDLPQGFSALEPLVDLPIGINITRFQTLALHRSYLT  
10 XSSGWTAGAAAYYVGYLQPRTFLLKYNENGTTITDAVDCALDPLSETKCTLKSFTVEKGIYQ  
TSNFRVQPTESIVRFPNITNLCPFDEVFNATRFASVYAWNRRKRISNCVADYSVLYNFAPFFAF  
KCYGVSPTKLNDLCFTNVYADSFVIRGNEVSQIAPGQTGNIADYNYKLDDFTGCVIAWNS  
NKLDSKVGNNYNYRYRLFRKSNLKPFERDISTEIQAGNKPCNGVAGVNCYFPLQSYGFRP  
TYGVGHQPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLKGTGVLTESNKKF  
15 LPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPV  
AIHADQLTPTWRVYSTGSNVFQTRAGCLIGA EYVNNSYECDIPIGAGICASYQTQTKSHRR  
ARSVASQSIIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVXMTKTSVDCTMYICGDST  
ECSNLLLQYGSFCTQLKRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKYFGGFNFSQILPDPS  
KPSKRSFIEDLLFNKVTADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQ  
20 YTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN  
SAIGKIQDSLSTASALGKLQDVVNHNALNTLVKQLSSKFGAISSVLNDILSRDPPEAEVQIDR  
LITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSA  
PHGVVFLHVTYVPAQEKNFTTAPAICHGKAHFPREGVFVSNGTHWFTVQRNFYEPQIITT  
DNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVLGDISGINASVVNI  
25 QKEIDRLNEVAKNLNESLIDLQELGKYEYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSC  
CSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

SEQ ID NO: 136

权利要求书:

1. 核酸构建体, 其包含:
  - (a) 开放阅读框(ORF), 和
  - (b) 5'非翻译区元件(5'UTR);
- 5 其中, 所述 5'UTR 源自 ARHGAP 基因;
 

优选地, 所述 ARHGAP 为 ARHGAP15;

更优选地, 所述 5'UTR 包含 SEQ ID NO: 44 所示核苷酸序列或其截短体。
2. 如权利要求 1 所述的核酸构建体, 其中, 所述 5'UTR 包含至少一个可用于
 10 抑制 UTR 内部 ATG 启动翻译的点突变;
 

优选地, 所述点突变选自 5'UTR 中 ATG 序列的 A、T 或 G 中的任意一个或多个位点上的突变;

更优选地, 所述突变为 A 突变为 G、C 或 T。
- 15 3. 如权利要求 1 或 2 所述的核酸构建体, 其中, 所述 5'UTR 截短体仍然保持调控 ORF 编码蛋白表达的功能; 或者, 所述 5'UTR 截短体相比于天然 5'UTR 具有增强的调控 ORF 编码蛋白表达的功能;
 

优选地, 所述 5'UTR 截短体的截短方式包括: 从 5'至 3'末端的序列方向上, 缺失 5'末端的连续的核苷酸序列; 和/或, 从 3'至 5'末端的序列方向上, 缺失 3'末端的连续的核苷酸序列。
- 20 4. 如权利要求 1-3 任一项所述的核酸构建体, 其中, 所述 5'UTR 截短体是从 5'至 3'末端的序列方向上, 缺失 5'末端的连续的核苷酸序列; 和/或, 从 3'至 5'末端的序列方向上, 缺失 3'末端的连续的核苷酸序列;
 

优选地, 所述 5'UTR 截短体包含 SEQ ID NO: 44 所示序列中至少 5 个连续的核苷酸序列, 更优选包含 SEQ ID NO: 2 所示序列中至少 7 个连续的核苷酸序列;

优选地, 所述 5'UTR 截短体的 5'末端的核苷酸是 SEQ ID NO: 44 所示序列按自然计数的第 1-57 位中的任一位置处的核苷酸, 更优选是 SEQ ID NO: 2 所示序列按自然计数的第 1-13 或第 17-29 位中的任一位置处的核苷酸;
- 30 5. 如权利要求 1-3 任一项所述的核酸构建体, 其中, 所述 5'UTR 截短体包含 CTATAAT 或 SEQ ID NO: 193 所示的核苷酸序列。
6. 如权利要求 1-3 任一项所述的核酸构建体, 其中, 所述 5'UTR 包含 SEQ ID NO: 2、28-43、76、137-196 任一项所示的核苷酸序列, 或包含 CTATAAT; 或与前述任一序列具有至少 80%同一性的核苷酸序列。

6. 如权利要求 1 至 5 任一项所述的核酸构建体, 其进一步包含:

(c) 3'非翻译区元件(3'UTR);

优选地, 所述 3'UTR 包含源自 HBB、ARHGAP、CORO1A 或 HPX 任一基因的 3'UTR;

5 更优选地, 所述 3'UTR 包含源自 ARHGAP15 的 3'UTR。

7. 如权利要求 6 所述的核酸构建体, 其中,

所述 3'UTR 包含 SEQ ID NO: 7-10 中任一所示的核苷酸序列, 或, 所述 3'UTR 包含与 SEQ ID NO: 7-10 任一具有至少 80%同一性的核苷酸序列。

10

8. 如权利要求 1 至 7 任一项所述的核酸构建体, 其进一步包含:

(d) 多聚腺苷酸(poly-A)尾巴。

9. 如权利要求 8 所述的核酸构建体, 其中,

15 所述 poly-A 尾巴选自 HGH polyA、SV40polyA、BGH polyA、rbGlob polyA 或 SV40late polyA;

优选地, 所述 poly-A 尾巴包含 SEQ ID NO: 16 或 SEQ ID NO: 135 所示或与之具有至少 80%同一性的核苷酸序列。

20 10. 如权利要求 1 至 9 任一项所述的核酸构建体, 其中, 所述 ORF 编码病毒抗原; 优选地, 所述 ORF 编码冠状病毒抗原或流感病毒抗原;

所述冠状病毒优选为 SARS-COV-2, 所述冠状病毒抗原优选为刺突蛋白;

更优选地, 所述刺突蛋白选自 SARS-COV-2、SARS-COV-2 Alpha、SARS-COV-2 Beta、SARS-COV-2 Gamma、SARS-COV-2 Kappa、SARS-COV-2  
25 Delta 或 SARS-COV-2 Omicron 任一病毒株的刺突蛋白;

所述流感病毒优选为 A 型流感病毒或 B 型流感病毒, 所述抗原为血凝素蛋白(HA)和/或神经酰胺酶(NA);

更优选地, 所述流感病毒抗原选自 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria、B 型流感病毒 Yamagata 任一病毒株的血凝素蛋白和/或神经  
30 氨酸酶。

11. 如权利要求 1 至 10 任一项所述的核酸构建体, 其中,

所述 ORF 所编码的多肽序列包含 SEQ ID NO: 12、14、21、23、25、80、83-86、98-101 和 136 中任一所示的氨基酸序列;

35 优选地, 所述 ORF 包含 SEQ ID NO: 13、15、20、22、24、81、87-90、97、102-105 中任一所示的核苷酸序列, 或与 SEQ ID NO: 13、15、20、22、24、81、

87-90、97、102-105 任一所示具有至少 80%、85%、90%、95%、96%、97%、98%、99%同一性的核苷酸序列。

12. 如前述任一项所述的核酸构建体，其包含如 SEQ ID NO: 18-19 所示的核苷酸序列或与之具有至少 80%同一性的核苷酸序列。

13. RNA 分子，其包括：

(a) 开放阅读框(ORF)；

(b) 5'非翻译区元件(5'UTR)；和

(c) 3'非翻译区元件(3'UTR)；

其中，所述 5'UTR 选自源自 ARHGAP、HSPB1、HBB 或 CCL13 任一基因的 5'UTR，所述 3'UTR 选自源自 HBB、ARHGAP、CORO1A 或 HPX 任一基因的 3'UTR；

所述 ARHGAP 优选为 ARHGAP15，所述 5'UTR 优选包含 SEQ ID NO: 79 所示的核苷酸序列或其截短体。

15

14. 如权利要求 13 所述的 RNA 分子，所述 5'UTR 和 3'UTR 选自如下任一组合：

1) 5'UTR 源自 ARHGAP 的 5'UTR，3'UTR 源自 HBB、ARHGAP、CORO1A 或 HPX 任一的 3'UTR；

2) 5'UTR 源自 HSPB1 的 5'UTR，3'UTR 源自 HBB、ARHGAP、CORO1A 或 HPX 任一的 3'UTR；

3) 5'UTR 源自 HBB 的 5'UTR，3'UTR 源自 HBB、ARHGAP、CORO1A 或 HPX 任一的 3'UTR；

4) 5'UTR 源自 CCL13 的 5'UTR，3'UTR 源自 HBB、ARHGAP、CORO1A 或 HPX 任一的 3'UTR；

25

优选地，

所述 ARHGAP 的 5'UTR 包含如 SEQ ID NO: 45-46、63-75、77-78、CUAUAAU 任一所示的核苷酸序列，或包含 CUAUAAU，或与前述任一序列具有至少 80%同一性的核苷酸序列；

所述源自 HSPB1 的 5'UTR 包含如 SEQ ID NO: 47 所示或与之具有至少 80%同一性的核苷酸序列；

所述源自 HBB 的 5'UTR 包含如 SEQ ID NO: 48 所示或与之具有至少 80%同一性的核苷酸序列；

所述源自 CCL13 的 5'UTR 包含如 SEQ ID NO: 49 所示或与之具有至少 80%同一性的核苷酸序列；

35

所述源自 HBB 的 3'UTR 包含如 SEQ ID NO: 50 所示或与之具有至少 80%同

一性的核苷酸序列；

所述源自 ARHGAP 的 3'UTR 包含如 SEQ ID NO: 51 所示或与之具有至少 80% 同一性的核苷酸序列；

所述源自 CORO1A 的 3'UTR 包含如 SEQ ID NO: 52 所示或与之具有至少 80% 同一性的核苷酸序列；

所述源自 HPX 的 3'UTR 包含如 SEQ ID NO: 53 所示或与之具有至少 80% 同一性的核苷酸序列。

15. 如权利要求 13 至 14 任一项所述的 RNA 分子，其中，所述 ORF 编码病毒抗原；优选地，所述 ORF 编码冠状病毒抗原或流感病毒抗原；

所述冠状病毒优选为 SARS-COV-2，所述冠状病毒抗原优选为刺突蛋白；

更优选地，所述刺突蛋白选自 SARS-COV-2、SARS-COV-2 Alpha、SARS-COV-2 Beta、SARS-COV-2 Gamma、SARS-COV-2 Kappa、SARS-COV-2 Delta 或 SARS-COV-2 Omicron 任一病毒株的刺突蛋白；或者，

所述流感病毒选自 A 型流感病毒或 B 型流感病毒，所述流感病毒抗原为血凝素蛋白(HA)和/或神经氨酸酶(NA)；

更优选地，所述流感病毒抗原选自 A 型流感病毒 H1N1、A 型流感病毒 H3N2、B 型流感病毒 Victoria、B 型流感病毒 Yamagata 任一病毒株的血凝素蛋白和/或神经氨酸酶。

16. 如权利要求 15 所述的 RNA 分子，其中，

所述 ORF 所编码的多肽序列包含 SEQ ID NO: 12、14、21、23、25、80、83-86、98-101 和 136 中任一所示的氨基酸序列；

优选地，所述 ORF 的核苷酸序列包含 SEQ ID NO: 54-55、59-61、82 和 126-134 任一所示，或与 SEQ ID NO: 54-55、59-61、82、126-134 任一所示具有至少 80%、85%、90%、95%、96%、97%、98%、99% 同一性的核苷酸序列。

17. 如权利要求 13 至 16 任一项所述的 RNA 分子，其进一步包含：

(d) 多聚腺苷酸(poly-A)尾巴。

18. 如权利要求 17 所述的 RNA 分子，其中，

所述 poly-A 尾巴选自 HGH polyA、SV40polyA、BGH polyA、rbGlob polyA 或 SV40late polyA；

优选地，所述 poly-A 尾巴包含 SEQ ID NO: 56 或 SEQ ID NO: 135 所示的核苷酸序列或与之具有至少 80% 同一性的核苷酸序列。

19. 如权利要求 13 至 18 任一项所述的 RNA 分子, 其包含(a) 开放阅读框(ORF), (b) 5'非翻译区元件(5'UTR), (c) 3'非翻译区元件(3'UTR), 和(d) 多聚腺苷酸(poly-A)尾巴;

5 优选地, 所述 RNA 分子包含如 SEQ ID NO: 57-58、106-125 任一所示的核苷酸序列。

20. 如权利要求 13 至 19 任一项所述的 RNA 分子, 其进一步包含:  
(e) 5'帽结构(5'Cap)。

10 21. 如权利要求 20 所述的 RNA 分子, 其中,

所述 5'Cap 选自 Cap0、Cap1、Cap2、Cap3、Cap4、ARCA、修饰的 ARCA、肌苷、N1-甲基-鸟苷、2'-氟代-鸟苷、7-脱氮-鸟苷、8-氧代-鸟苷、2-氨基-鸟苷、LNA-鸟苷和 2-叠氮基-鸟苷;

15 优选地, 所述 5'Cap 选自 ARCA、3'OMe-m7G(5')ppp(5')G、m7G(5')ppp(5')(2'OMeA)pU、m7Gppp(A2'O-MOE)pG、m7G(5')ppp(5')(2'OMeA)pG、m7G(5')ppp(5')(2'OMeG)pG、m7(3'OMeG)(5')ppp(5')(2'OMeG)pG 或 m7(3'OMeG)(5')ppp(5')(2'OMeA)pG;

优选地, 所述 5'Cap 为 m7G(5')ppp(5')(2'OMeA)pG。

20 22. 如权利要求 13 至 21 中任一项所述的 RNA 分子, 所述 RNA 分子是 mRNA。

23. 如权利要求 13 至 22 中任一项所述的 RNA 分子, 其进一步包含一种或多种修饰,

优选地, 所述修饰包括骨架修饰、糖修饰、碱基修饰和/或脂质修饰;

25 更优选地, 所述碱基修饰为假尿苷修饰。

24. RNA 分子, 其包含 SEQ ID NO: 54-55、59-61、82、126-134 任一所示的核苷酸序列或与 SEQ ID NO: 54-55、59-61、82、126-134 任一所示具有至少 80% 同一性的核苷酸序列;

30 优选地, 所述 RNA 分子进一步包含一种或多种修饰, 更优选地, 所述修饰为假尿苷修饰。

25. 权利要求 1-12 任一项所述的核酸构建体, 或权利要求 13-24 任一项所述的 RNA 分子在如下任一项中的用途: (1)制备疫苗; (2)在受试者体内或体外编码病毒抗原; (3)制备在受试者体内或体外编码病毒抗原的药物; (4)用于制备药物。  
35

26. 疫苗，其包含权利要求 13 至 24 任一项所述的 RNA 分子，优选地，所述 RNA 分子编码一种或多种病毒株的一个或多个抗原。

27. 载体，其包含权利要求 1 至 12 任一项所述的核酸构建体，或权利要求 13 至 24 任一项所述的 RNA 分子。

28. 宿主细胞，其包含权利要求 27 所述的载体。

29. 权利要求 13-24 任一项所述的 RNA 分子的制备方法，包括：将权利要求 1 至 12 任一项所述的核酸构建体，或权利要求 27 所述的载体进行逆转录，得到 RNA 分子；优选地，所述方法还包含对所述 RNA 分子的 5'端添加 5'Cap。

30. 递送系统，其包含权利要求 1 至 12 任一项所述的核酸构建体或权利要求 13 至 24 任一项所述的 RNA 分子，其中，所述递送系统的递送媒介物是阳离子脂质递送颗粒，优选地，所述递送媒介物是纳米脂质颗粒。

31. 药物组合物，其包含：

药学上可接受的载体、稀释剂或赋形剂，和

选自以下任意一项：

权利要求 1 至 12 任一项所述的核酸构建体、权利要求 13 至 24 任一项所述的 RNA 分子、权利要求 26 所述的疫苗，和/或权利要求 30 所述的递送系统。

32. 产品或试剂盒，其包含权利要求 1 至 12 任一项所述的核酸构建体、权利要求 13 至 24 任一项所述的 RNA 分子、权利要求 26 所述的疫苗、权利要求 30 所述的递送系统，和/或权利要求 31 所述的药物组合物。

33. 一种制备用于治疗 and/或预防疾病的药物的用途，包括向有需要的受试者施用有效量的权利要求 1 至 12 任一项所述的核酸构建体、权利要求 13 至 24 任一项所述的 RNA 分子、权利要求 26 所述的疫苗、权利要求 30 所述的递送系统、权利要求 31 所述的药物组合物，和/或权利要求 32 所述的产品或试剂盒；所述疾病为病毒感染性疾病或病毒感染相关的呼吸系统疾病；

优选地，所述病毒为冠状病毒或流感病毒，所述病毒更优选为 SARS-CoV-2、A 型流感病毒或 B 型流感病毒；

优选地，所述病毒感染相关呼吸系统疾病包括单纯性感染、发热、咳嗽、咽痛、鼻炎、头痛、肺炎、急性呼吸道感染、严重急性呼吸道感染(SARI)、低氧性呼吸衰竭、急性呼吸窘迫综合征、脓毒症、脓毒性休克、重症急性呼吸综合征

(SARS)。

34. 一种治疗和/或预防疾病的方法，包括向有需要的受试者施用有效量的权利要求 1 至 12 任一项所述的核酸构建体、权利要求 13 至 24 任一项所述的 RNA 分子、权利要求 26 所述的疫苗、权利要求 30 所述的递送系统、权利要求 31 所述的药物组合物，和/或权利要求 32 所述的产品或试剂盒；所述疾病为病毒感染性疾病或病毒感染相关的呼吸系统疾病；

优选地，所述病毒为冠状病毒或流感病毒，所述病毒更优选为 SARS-CoV-2、A 型流感病毒或 B 型流感病毒；

- 10 优选地，所述病毒感染相关呼吸系统疾病包括单纯性感染、发热、咳嗽、咽痛、鼻炎、头痛、肺炎、急性呼吸道感染、严重急性呼吸道感染(SARI)、低氧性呼吸衰竭、急性呼吸窘迫综合征、脓毒症、脓毒性休克、重症急性呼吸综合征(SARS)。

- 15 35. 一种诱导受试者产生中和抗体反应和/或 T 细胞免疫应答方法，包括向有需要的受试者施用有效量的权利要求 1 至 12 任一项所述的核酸构建体、权利要求 13 至 24 任一项所述的 RNA 分子、权利要求 26 所述的疫苗、权利要求 30 所述的递送系统、权利要求 31 所述的药物组合物，和/或权利要求 32 所述的产品或试剂盒；

- 20 优选地，所述中和抗体反应为针对病毒抗原的中和抗体反应，所述 T 细胞免疫应答包括 CD4+和/或 CD8+T 细胞免疫反应；

优选地，所述病毒为冠状病毒或流感病毒，所述病毒更优选为 SARS-CoV-2、A 型流感病毒或 B 型流感病毒。