

猴痘病毒核酸疫苗

技术领域

本发明涉及生物医药和病毒学领域，具体涉及预防和/或治疗猴痘病毒感染的核酸疫苗(优选 mRNA 疫苗)。

背景技术

猴痘是一种由猴痘病毒(Mpox Virus, MPXV)引起的人畜共患病。猴痘病毒是一类有囊膜的 DNA 病毒，直径约 200-250 nm，是痘病毒科，正痘病毒属的成员之一(参见，例如 Alakunle E et al., Monkeypox Virus in Nigeria: Infection Biology, Epidemiology, and Evolution. Viruses. 2020 Nov 5;12(11):1257.和 Cho CT, Wenner HA. Monkeypox virus. Bacteriol Rev. 1973 Mar;37(1):1-18.)。虽然在 2023 年 5 月 12 日，世界卫生组织(World Health Organization, WHO)宣布猴痘疫情不再构成国际关注的突发公共卫生事件，但全球范围内的猴痘病例数仍在保持增长。

猴痘病毒的传播途径包括直接接触传播和飞沫传播等，其潜伏期为 7-14 天，最长可达 21 天(参见，例如 Moore MJ et al., Mpox (Monkeypox). 2023 May 3. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan.)。猴痘的主要症状包括高热、头痛、皮疹、肌肉疼痛和局部淋巴结肿大等，多为一过性，病程持续 2-4 周即可自行恢复，但出现皮疹的患者在痂皮脱落后可能遗留红斑、色素沉着甚至瘢痕，瘢痕持续时间可长达数年。对于免疫力低下的人群(如艾滋病患者)，严重的猴痘感染也可引起死亡。

猴痘病毒、高致病性天花病毒和牛痘病毒同属正痘病毒属，三种病毒的抗原蛋白之间具有较高的同源性。考虑到猴痘病毒需在生物安全三级试验室中操作，而牛痘病毒和猴痘病毒之间具有良好的交叉反应性(参见，例如 Zeng J et al., Mpox multi-antigen mRNA vaccine candidates by a simplified manufacturing strategy afford efficient protection against lethal orthopoxvirus challenge. Emerg Microbes Infect. 2023 Dec;12(1):2204151.)，故在目前已发表的科研文章中多用牛痘病毒来评价猴痘疫苗分子的免疫原性。此外也有证据表明，接种一种正痘病毒疫苗可提供针对其他正痘病毒的交叉保护性免疫(参见，例如 Lum FM et al., Monkeypox: disease epidemiology, host immunity and clinical interventions. Nat Rev Immunol. 2022 Oct;22(10):597-613.)。

目前仅有两款基于牛痘病毒抗原开发的天花疫苗 ACAM2000 和 JYNNEOS 被美国食品药品监督管理局(Food and Drug Administration, FDA)批准用于预防猴痘的免疫接种，然而，由于这两款疫苗在安全性方面表现不佳，与心肌炎、心包炎等严重的不良事件有关，故其仅被批准用于高危人群的暴露前免疫。自 1980 年天花疫苗接种计划停止后，普通大众对正痘病毒的群体免疫力下降，有研究显示猴痘在未接种天花疫苗的群体中具有更高的发病率(参见，例如 Shah HH et al., Human monkeypox - After 40 years, an unintended consequence of smallpox eradication. Front Public Health. 2023 Jan 5;10:1082586.和 Pattnaik H et al., Making Sense of Monkeypox: A Comparison of Other Poxviruses to the Monkeypox. Cureus. 2023 Apr 24;15(4):e38083.)。因此，一款针对猴痘病毒开发的安全有效的疫苗仍是迫切所需的。

发明内容

在第一方面，本发明提供一种多肽或多肽组合，其包含一个或多个肽段，肽段各自独立地选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段；任选地，多肽或多肽组合还包含一个或多个接头、IgE 信号肽和/或 HA 多肽标签。

5 在一些实施方案中，多肽或多肽组合包含 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列或者与 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列具有至少 80% 相同性的氨基酸序列。

10 在一些实施方案中，多肽或多肽组合包含两个或更多个选自 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 的氨基酸序列或者两个或更多个与选自 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 的氨基酸序列具有至少 80% 相同性的氨基酸序列。

15 在第二方面，本发明提供一种核酸或核酸组合，其包含编码本发明的多肽或多肽组合的多核苷酸。

20 在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列或者与 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，所述多核苷酸包含 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列或者与 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列。

25 在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列具有至少 80% 相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 的核苷酸序列具有至少 80% 相同性的核苷酸序列。

35 在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 之一的核苷酸序列或者与 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，所述多核苷酸包含 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 之一的核苷酸序列或者与 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、

82、83、84、85、86、87 和 88 之一的核苷酸序列具有至少 80%相同性的核苷酸序列。

在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列具有至少 80%相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 的核苷酸序列具有至少 80%相同性的核苷酸序列。

在第三方面，本发明提供一种组合物，其包含本发明的核酸或核酸组合。

在一些实施方案中，组合物包含第一脂质、第二脂质、磷脂和类固醇，其各自如本文所定义。在一些实施方案中，第一脂质为阳离子脂质。在一些实施方案中，组合物被配制为将本发明的核酸或核酸组合包封在脂质中的脂质纳米颗粒。

在第四方面，本发明提供一种药物组合物或试剂盒，其包含本发明的多肽或多肽组合、核酸或核酸组合或者组合物。

在第五方面，本发明提供一种表达载体，其包含本发明的核酸或核酸组合。

在第六方面，本发明提供一种宿主细胞，其包含本发明的核酸或核酸组合或者表达载体。

在第七方面，本发明提供本发明的多肽或多肽组合、核酸或核酸组合、组合物、药物组合物或试剂盒在制备用于预防和/或治疗受试者猴痘病毒感染的疫苗中的用途。

附图说明

图 1 示出检测候选猴痘疫苗在小鼠体内诱导的牛痘病毒交叉中和抗体效价的结果。

图 2A 和图 2B 示出检测小鼠脾脏 T 细胞对猴痘病毒混合肽库的免疫应答的结果。

具体实施方案

一般定义和术语

本文引用的所有专利、专利申请、科学出版物、制造商的说明书和指南等，无论上文或下文，均整体援引加入本文。本文中的任何内容均不应理解为承认本公开无权先于这样的公开。

除非另有说明，否则本文中使用的科学和技术名词具有本领域技术人员所通常理解的含义。并且，本文中所用的蛋白和核酸化学、分子生物学、细胞和组织培养、微生物学相关术语均为相应领域内广泛使用的术语(参见，例如，*Molecular Cloning: A Laboratory Manual*, 2nd Edition, J. Sambrook et al. eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989)。同时，为了更好地理解本发明，下面提供相关术语的定义和解释。

如本文所用，表述“包括”、“包含”、“含有”和“具有”是开放式的，表示包括所列举的元素、步骤或组分但不排除其他未列举的元素、步骤或组分。表述“由……组成”不包括未指定的任何元素、步骤或组分。表述“基本上由……组成”是指范围限于指定的元素、

步骤或组分，加上不显著影响要求保护的主题的基本和新特性的任选存在的元素、步骤或组分。应当理解，表述“基本上由.....组成”和“由.....组成”涵盖在表述“包含”的含义之内。

如本文所用，除非上下文另外指明，单数形式的表述“一个”、“一种”或“这个”包括复数指代。术语“一个(种)或多个(种)”或者“至少一个(种)”涵盖 1、2、3、4、5、6、7、8、9 个(种)或更多个(种)。

如本文所用，多个述及的元素之间的连接术语“和/或”应理解为包括单独的和组合的选项。换言之，“和/或”包括“和”以及“或”。例如，A 和/或 B 包括 A、B 以及 A+B。A、B 和/或 C 包括 A、B、C 及其任意组合，例如 A+B、A+C、B+C 以及 A+B+C。以“和/或”限定的更多元素以类似方式理解，并且包括其中的任一种及其任意组合。

本文中值的范围的列举仅为了用作单独提到落在所述范围内的每个不同值的速记方法。除非本文另有说明，否则每个单独的值如其在本文中单独列举地加入本说明书。除非明确指出相反，在本文示出的数值或范围均由“约”修饰，表示所列举或声称的数值或范围 $\pm 20\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 或 $\pm 3\%$ 。

如本文所用，“核酸组合”可以表示包含一条以上的核酸的组合或集合。这样的核酸可以分别存在，例如存在于不同的组合物、药物组合物或试剂盒的不同部分或者不同的载体中，也可以同时存在于同一组合物、药物组合物或试剂盒的同一部分或者同一载体中。核酸组合中的核酸可以相同也可以各自不同，可以各自编码相同的多肽(部分)或者可以编码不同的多肽(部分)。

如本文所用，“多肽组合”可以表示包含一条以上的多肽的组合或集合。这样的多肽可以分别存在，例如存在于不同的药物组合物或试剂盒的不同部分中，也可以同时存在于同一药物组合物或试剂盒的同一部分中。多肽组合中的多肽可以相同也可以各自不同，可以各自包含相同的氨基酸序列(部分)或者可以包含不同的氨基酸序列(部分)。

如本文所用，术语“野生型”表示该序列是天然存在的并且未经人为修饰的，包括天然存在的突变体。

如本文所用，关于序列的术语“%相同性”是指在待比较的序列之间的最佳比对中相同的核苷酸或氨基酸的百分比。两个序列之间的差异可以分布在待比较序列的局部区域(区段)或整个长度上。通常在对区段或“比较窗口”最佳比对之后，确定两个序列之间的相同性。最佳比对可以手动进行，或者借助于本领域已知算法，包括但不限于 Smith and Waterman, 1981, *Adv. App. Math.* 2, 482 和 Needleman and Wunsch, 1970, *J. Mol. Biol.* 48, 443 描述的局部同源性算法，Pearson and Lipman, 1988, *Proc. Natl Acad. Sci. USA* 88, 2444 描述的相似性搜索方法，或使用计算机程序，例如 Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis. 中的 GAP、BESTFIT、FASTA、BLASTP、BLASTN 和 TFASTA 进行。例如，可以利用美国国家生物技术信息中心(NCBI)网站公共可用的 BLASTN 或 BLASTP 算法确定两个序列的百分比相同性。

通过确定待比较的序列对应的相同位置的数目，用这个数目除以比较的位置数目(例如，参考序列中的位置数目)，并将这个结果乘以 100，获得%相同性。在一些实施方案中，至少约 50%、至少约 55%、至少约 60%、至少约 65%、至少约 70%、至少约 75%、至少约 80%、至少约 85%、至少约 90%、至少约 95%或约 100%的区域给出相同性程度。

在一些实施方案中，对参考序列的整个长度给出相同性程度。可以用本领域已知的工具进行确定序列相同性的比对，优选利用最佳序列比对，例如，利用 Align，利用标准设置，优选 EMBOSS:needle、Matrix: Blosum62、Gap Open 10.0、Gap Extend 0.5。

在本文中，“核苷酸”包括脱氧核糖核苷酸和核糖核苷酸及其衍生物。如本文所用，
5 “核糖核苷酸”是核糖核酸(RNA)的构成物质，由一分子碱基，一分子五碳糖，一分子磷酸组成，其是指在 β -D-呋喃核糖(β -D-ribofuranosyl)基团的 2' 位置具有羟基的核苷酸。而“脱氧核糖核苷酸”是脱氧核糖核酸(DNA)的构成物质，也由一分子碱基，一分子五碳糖，一分子磷酸构成，其是指在 β -D-呋喃核糖(β -D-ribofuranosyl)基团的 2' 位置的羟基被氢取代的核苷酸，是染色体的主要化学成分。“核苷酸”通常由代表其中碱基的单字母来指代：
10 “A (a)”指含有腺嘌呤的脱氧腺苷酸或腺苷酸，“C (c)”指含有胞嘧啶的脱氧胞苷酸或胞苷酸，“G (g)”指含有鸟嘌呤的脱氧鸟苷酸或鸟苷酸，“U (u)”指含有尿嘧啶的尿苷酸，“T (t)”指含有胸腺嘧啶的脱氧胸苷酸。

如本文所用，术语“多核苷酸”和“核酸”可以互换使用，用于指脱氧核糖核苷酸的聚合物(脱氧核糖核酸，DNA)或核糖核苷酸的聚合物(核糖核酸，RNA)。“多核苷酸序列”、
15 “核酸序列”和“核苷酸序列”可以互换使用，用来表示多核苷酸中核苷酸的排序。本领域人员应当理解，DNA 编码链(有义链)与其编码的 RNA 可以看作具有相同的核苷酸序列，DNA 编码链序列中的脱氧胸苷酸对应其编码的 RNA 序列中的尿苷酸。

如本文所用，术语“表达”包括核苷酸序列的转录和/或翻译。因此，表达可以涉及转录物和/或多肽的产生。术语“转录”涉及将 DNA 序列中的遗传密码转录为 RNA(转录物)
20 的过程。术语“体外转录”指在不含细胞的系统中(例如在适当的细胞提取物中)体外合成 RNA，特别是 mRNA(参见，例如 Pardi N., Muramatsu H., Weissman D., Karikó K. (2013). In: Rabinovich P. (eds) Synthetic Messenger RNA and Cell Metabolism Modulation. Methods in Molecular Biology (Methods and Protocols), vol 969. Humana Press, Totowa, NJ.)。可以用于产生转录物的载体又称为“转录载体”，其中包含转录所需的调控序列。术语“转录”涵盖“体外转录”。
25

如本文所用，“编码”是指核酸中特定核苷酸序列的固有特性，比如基因，cDNA 或者 mRNA 都能作为模板去合成其他生物过程中的多聚物和大分子，只要已经有明确的核苷酸序列或者有明确的氨基酸序列。因此一个基因编码蛋白是指基因的 mRNA 通过转录和翻译在细胞中或其它生物系统中产生蛋白。

30 如本文所用，术语“宿主细胞”指用于接受、保持、复制、表达核酸或载体的细胞。在一些实施方案中，宿主细胞可以是在其中表达本发明的多肽或多肽组合的细胞。

术语“氨基酸”旨在涵盖所有分子，无论是天然的还是合成的，其既包括氨基官能团又包括酸官能团并且能够被包括在天然存在的氨基酸的聚合物中。示例性的氨基酸包括天然存在的氨基酸；类似物、衍生物和同类物；具有变异侧链的氨基酸类似物；以及前述的任一种的所有立体异构体。如本文所用，术语“氨基酸”包括 D-或 L-旋光异构体和拟肽。
35

本文可互换地使用的术语“多肽链”、“多肽”、“肽”和“蛋白”(如果是单链的话)是指任何长度的氨基酸的聚合物。该聚合物可以是直链或支链的，它可以包含修饰的氨基酸，并且可以被非氨基酸中断。该术语还涵盖已被修饰的氨基酸聚合物，修饰可以是例如，

二硫键形成、糖基化、脂化、乙酰化、磷酸化或任何其他操作，例如与标记组分的缀合。多肽可以从天然来源中分离，可以通过重组技术从真核或原核宿主中产生，或者可以是合成方法的产物。

术语“信号肽”和“信号序列”在本文中可互换使用，其通常位于蛋白的氨基端，是引导新合成的蛋白向分泌通路转移的短肽链，长度约 5-30 个氨基酸。

如本文所用，“多联体”是指由 2、3、4、5、6、7、8、9、10 或更多多个多肽相连构成的多肽，其中每个多肽与另一个多肽之间可包含或不包含接头。

如本文所用，“抗原”是指进入机体后可以引起机体获得性免疫应答的分子，此免疫应答可能涉及抗体的产生，或特定的免疫原活性细胞，或者两者同时发生。本领域技术人员应当理解任何大分子，几乎包括所有蛋白或肽段可以作为抗原。更进一步，抗原可以来自重组或者基因组 DNA 或 RNA。本领域技术人员应当理解任何在本文的 DNA 或 RNA，它们的核苷酸序列或者部分核苷酸序列可以编码能引起机体获得性免疫的蛋白。更进一步，本领域技术人员应当理解一个抗原不需要单独地只编码一个基因的全长核苷酸序列。显而易见地，本发明包括但不限于使用超过一个基因的部分核苷酸序列，并且这些核苷酸序列形成不同的混合物去诱导应答的发生。更进一步，本领域技术人员应当理解抗原不需要完全被一个基因所编码。显而易见地，抗原可以被合成生成，也可以来源于生物样本。生物样本包括但不限于组织样本，肿瘤样本，细胞或生物流体。

如本文所用，“抗体”是指机体由于抗原的刺激而产生的具有保护作用的蛋白。它是一种免疫球蛋白，由 B 淋巴细胞产生。抗体的单体是一个 Y 形的分子，有 4 条多肽链组成。其中包括两条相同的重链，以及两条相同的轻链，之间由双硫键连接在一起。每条重链约 50kDa，每条轻链约 25kDa，轻重链间存在二硫键链接。其独特之处在于对结合伙伴的高亲和力和特异性。

如本文所用，“疫苗”是指包含活性成分(例如本发明的多肽或多肽组合或者核酸或核酸组合)的组合物，它能通过接种引起接种对象免疫应答。在具体实施方案中，它诱导的免疫应答能提供免疫保护，并足以预防和/或减轻与病原体或疾病感染相关的至少一种症状。根据本发明，本文所述的核酸或核酸组合或者组合物可以作为疫苗，用于在有需要的受试者中提供针对猴痘病毒的预防性和/或治疗性免疫。如本文所用，“猴痘病毒 RNA 疫苗(例如 mRNA 疫苗、circRNA 疫苗或 saRNA 疫苗)”提供基于肽的疫苗或 DNA 疫苗的独特治疗替代方案。本文所述的猴痘病毒 RNA 疫苗(例如 mRNA 疫苗、circRNA 疫苗或 saRNA 疫苗)包括至少一种具有编码至少一种猴痘病毒抗原性多肽或其免疫原性片段(例如能够诱导对猴痘病毒的免疫应答的免疫原性片段)的开放阅读框的核糖核酸(RNA)多核苷酸。

如本文所用，术语“中和抗体”是指能够中和，即防止、抑制、降低或干扰病原体在宿主(例如宿主生物体或宿主细胞)中引发和/或保持感染的能力的抗体或其片段。根据本发明，用本发明的疫苗进行接种的受试者中可以产生针对猴痘病毒抗原蛋白(或抗原多肽)的中和抗体，例如在受试者的免疫血清中。可以使用本领域已知的方法测量免疫血清中的中和抗体滴度水平。

术语例如“处理”或“治疗”或“以治疗”或“缓解”或“以缓解”指的是治愈、减缓、减少现存的已经诊断的病理病症或紊乱的症状，和/或截停或减缓现存的已经诊断的病理病症

或紊乱的进展的治疗性措施。术语例如“预防”、“防御”、“避免”、“遏制”等指的是预防未经诊断的目标病理病症或紊乱的进展的预防性的或预防用的措施。因此，“有需要的受试者”可包括已经患有疾病的受试者；易于患有疾病的受试者；和需要预防疾病的受试者。

5 术语“预防或治疗有效量”是指足以预防或抑制疾病或症状的发生和/或减缓、减轻、延迟疾病或症状的发展或严重程度。预防或治疗有效量受到包括但不限于以下因素的影响：疾病或症状的发展速度和严重程度，受试者的年龄、性别、体重和生理状况，治疗的持续时间以及具体施用途径。预防或治疗有效量可以在一个或多个剂量中施用。预防或治疗有效量可以通过持续或间断施用实现。

10 如本文所用，术语“受试者”描述了可以对其提供使用本发明多肽或多肽组合、核酸或核酸组合、组合物或药物组合物治疗的生物体，例如哺乳动物。优选地，受试者是啮齿类动物和人。

如本文所用，术语“第一多肽(链)”以及类似使用的“第二多肽(链)”、“第三多肽(链)”、“第四多肽(链)”、“第五多肽(链)”、“第六多肽(链)”等仅作为区分目的使用，并不意图限
15 定这些多肽(链)的顺序和连接方式。各个多肽(链)自身和之间可以相同也可以不同。若需要，也可以包含两个或更多个第一多肽(链)，其可以相同也可以不同。在不同的方案中使用的上述术语可以相同，也可以不同。

如本文所用，术语“第一多核苷酸”以及类似使用的“第二多核苷酸”、“第三多核苷酸”、“第四多核苷酸”、“第五多核苷酸”、“第六多核苷酸”等仅作为区分目的使用，并不意图
20 限定这些多核苷酸的顺序和连接方式。各个多核苷酸自身和之间可以相同也可以不同。若需要，也可以包含两个或更多个第一多核苷酸，其可以相同也可以不同。在不同的方案中使用的上述术语可以相同，也可以不同。

如本文所用并且除非另有说明，否则术语“脂质”是指一组有机化合物，其包括但不限于脂肪酸酯，并且一般以难溶于水但可溶于许多非极性有机溶剂中为特征。尽管脂质
25 一般具有弱水溶性，但是某些类别的脂质(例如经极性基团改性的脂质，例如 DMG-PEG2000)具有有限的水溶性并且在某些条件下可溶于水。已知的脂质类型包括生物分子，诸如脂肪酸、蜡、固醇、脂溶性维生素、单酸甘油酯、二酸甘油酯、三酸甘油酯和磷脂。脂质至少可分为三类：(1)“简单脂质”，包括脂肪和油，以及蜡；(2)“复合脂质”，包括磷脂和糖脂(例如 DMPE-PEG2000)；以及(3)“衍生脂质”，诸如类固醇。此外，如本文所
30 用，脂质还包括类脂质化合物。术语“类脂质化合物”又简称为“类脂质”，是指脂质样化合物(例如具有脂质样物理性质的两亲性化合物)。

术语“脂质纳米颗粒”或“LNP”是指具有至少一个纳米(nm)级尺寸(例如 1 至 1,000 nm)的颗粒，其含有一种或多种类型的脂质分子。本文所提供的 LNP 可进一步含有至少一种非脂质有效负载分子(例如一种或多种核酸分子)。在一些实施方案中，LNP 包含部分
35 或完全包封于脂质壳内的非脂质有效负载分子。特定来说，在一些实施方案中，其中有效负载是带负电荷的分子(例如 mRNA、circRNA 或 saRNA)，并且 LNP 的脂质组分包含至少一种阳离子脂质。不受理论束缚，预期阳离子脂质可与带负电荷的有效负载分子相互作用，并且在 LNP 形成期间促进有效负载并入和/或包封至 LNP 中。可形成如本文所提供的 LNP 的一部分的其它脂质包括但不限于中性脂质和带电脂质，诸如类固醇、聚

合物结合的脂质和各种两性离子性脂质。在某些实施方案中,根据本公开的 LNP 包含系列 01、02、03 和 04 的脂质中的一种或多种,例如根据本文所述的式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-II) (及其子式)的脂质中的一种或多种。

术语“阳离子脂质”是指在其环境的任何 pH 值或氢离子活性下带正电荷,或能够响应于其环境(例如其预定使用的环境)的 pH 值或氢离子活性而带正电荷的脂质。因此,术语“阳离子”涵盖“永久性阳离子”和“可阳离子化”。在某些实施方案中,阳离子脂质中的正电荷由季氮原子的存在引起。在某些实施方案中,阳离子脂质包含两性离子性脂质,所述两性离子性脂质在其预定使用的环境中(例如在生理 pH 值下)带正电荷。在某些实施方案中,阳离子脂质包含系列 01、02、03 和 04 的脂质中的一种或多种,例如根据本文所述的式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-II) (及其子式)的脂质中的一种或多种。术语“阴离子脂质”是指在其环境的任何 pH 值或氢离子活性下带负电荷,或能够响应于其环境(例如其预定使用的环境)的 pH 值或氢离子活性而带负电荷的脂质。示例性阴离子脂质包括一个或多个例如在生理 pH 值下带负电荷的磷酸基团。

术语“聚合物结合的脂质”是指既包含脂质部分又包含聚合物部分的分子。聚合物结合的脂质的实例是聚乙二醇化脂质(PEG-脂质),其中聚合物部分包含聚乙二醇。

术语“中性脂质”涵盖在所选 pH 值下或在所选 pH 值范围内以不带电形式或以中性两性离子形式存在的任何脂质分子。在一些实施方案中,所选的有用 pH 值或范围对应于预定脂质用途的环境中的 pH 条件,诸如生理 pH 值。作为非限制性实例,可与本公开结合使用的中性脂质包括但不限于磷脂酰胆碱,诸如 1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)、1,2-二棕榈酰基-sn-甘油-3-磷酸胆碱(DPPC)、1,2-二肉豆蔻酰基-sn-甘油-3-磷酸胆碱(DMPC)、1-棕榈酰基-2-油酰基-sn-甘油-3-磷酸胆碱(POPC)、1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC); 磷脂酰乙醇胺,诸如 1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)、2-((2,3-双(油酰氧基)丙基)二甲基铵基)磷酸氢乙酯(DOCP); 鞘磷脂(SM); 神经酰胺。本文所提供的中性脂质可为合成的或者衍生自天然来源或化合物(从其分离或改性)。

术语“类固醇”是指以环戊烷多氢菲为母核的一种脂质,包括固醇和固醇衍生物。类固醇化合物不含结合的脂肪酸,是非皂化性脂质。在本文中,类固醇可以作为纳米颗粒组合物脂质组分中的结构脂质。作为非限制性实例,可与本公开结合使用的类固醇包括但不限于胆固醇、粪固醇、谷固醇、麦角固醇、菜油固醇、豆固醇和其混合物。

术语“带电脂质”涵盖在所选 pH 值下或在所选 pH 值范围内以带正电荷或带负电荷形式存在的任何脂质分子。在一些实施方案中,所选 pH 值或范围对应于预定脂质用途的环境中的 pH 条件,诸如生理 pH 值。作为非限制性实例,可与本公开结合使用的带电脂质包括但不限于磷脂酰丝氨酸、磷脂酸、磷脂酰甘油、磷脂酰肌醇、固醇半琥珀酸酯、二烷基三甲基铵-丙烷(例如 DOTAP、DOTMA)、二烷基二甲基氨基丙烷、乙基磷酸胆碱、二甲基氨基乙烷氨基甲酰基固醇(例如 DC-Chol)、1,2-二油酰基-sn-甘油-3-磷酸-L-丝氨酸钠盐(DOPS-Na)、1,2-二油酰基-sn-甘油-3-磷酸-(1'-外消旋-甘油)钠盐(DOPG-Na)和 1,2-二油酰基-sn-甘油-3-磷酸钠盐(DOPA-Na)。本文所提供的带电脂质可为合成的或者衍生自天然来源或化合物(从其分离或改性)。

如本文所用并且除非另有说明,否则术语“烷基”是指仅由碳和氢原子组成的饱和直

链或支链烃链基团。在一实施方案中，烷基具有例如 1 至 24 个碳原子(C₁-C₂₄ 烷基)、4 至 20 个碳原子(C₄-C₂₀ 烷基)、6 至 16 个碳原子(C₆-C₁₆ 烷基)、6 至 9 个碳原子(C₆-C₉ 烷基)、1 至 15 个碳原子(C₁-C₁₅ 烷基)、1 至 12 个碳原子(C₁-C₁₂ 烷基)、1 至 8 个碳原子(C₁-C₈ 烷基)或 1 至 6 个碳原子(C₁-C₆ 烷基)，并且通过单键连接至分子其余部分。烷基的实例包括但不限于甲基、乙基、正丙基、1-甲基乙基(异丙基)、正丁基、正戊基、1,1-二甲

5 基乙基(叔丁基)、3-甲基己基、2-甲基己基等。除非另有说明，否则烷基任选地经取代。

如本文所用并且除非另有说明，否则术语“烯基”是指仅由碳和氢原子组成的直链或支链烃链基团，其含有一个或多个碳-碳双键。如本领域普通技术人员所理解，术语“烯基”还涵盖具有“顺式”和“反式”构型或替代地，“E”和“Z”构型的基团。在一实施方案中，

10 烯基具有例如 2 至 24 个碳原子(C₂-C₂₄ 烯基)、4 至 20 个碳原子(C₄-C₂₀ 烯基)、6 至 16 个碳原子(C₆-C₁₆ 烯基)、6 至 9 个碳原子(C₆-C₉ 烯基)、2 至 15 个碳原子(C₂-C₁₅ 烯基)、2 至 12 个碳原子(C₂-C₁₂ 烯基)、2 至 8 个碳原子(C₂-C₈ 烯基)或 2 至 6 个碳原子(C₂-C₆ 烯基)，并且通过单键连接至分子其余部分。烯基的实例包括但不限于乙烯基、丙-1-烯基、丁-1-烯基、戊-1-烯基、戊-1,4-二烯基等。除非另有说明，否则烯基任选地经取代。

如本文所用并且除非另有说明，否则术语“炔基”是指仅由碳和氢原子组成的直链或支链烃链基团，其含有一个或多个碳-碳三键。在一实施方案中，炔基具有例如 2 至 24 个碳原子(C₂-C₂₄ 炔基)、4 至 20 个碳原子(C₄-C₂₀ 炔基)、6 至 16 个碳原子(C₆-C₁₆ 炔基)、6 至 9 个碳原子(C₆-C₉ 炔基)、2 至 15 个碳原子(C₂-C₁₅ 炔基)、2 至 12 个碳原子(C₂-C₁₂ 炔基)、2 至 8 个碳原子(C₂-C₈ 炔基)或 2 至 6 个碳原子(C₂-C₆ 炔基)，并且通过单键连接至

20 分子其余部分。炔基的实例包括但不限于乙炔基、丙炔基、丁炔基、戊炔基等。除非另有说明，否则炔基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚烷基”或“亚烷基链”是指将分子其余部分连接至基团的直链或支链二价烃链，其仅由碳和氢组成并且为饱和的。在一实施方案中，亚烷基具有例如 1 至 24 个碳原子(C₁-C₂₄ 亚烷基)、1 至 15 个碳原子(C₁-C₁₅ 亚烷基)、

25 1 至 12 个碳原子(C₁-C₁₂ 亚烷基)、1 至 8 个碳原子(C₁-C₈ 亚烷基)、1 至 6 个碳原子(C₁-C₆ 亚烷基)、2 至 4 个碳原子(C₂-C₄ 亚烷基)、1 至 2 个碳原子(C₁-C₂ 亚烷基)。亚烷基的实例包括但不限于亚甲基、亚乙基、亚丙基、亚正丁基等。亚烷基链通过单键连接至分子其余部分，并且通过单键连接至基团。亚烷基链与分子其余部分和与基团的连接点可通过链内的一个碳或任何两个碳。除非另有说明，否则亚烷基链任选地经取代。

如本文所用并且除非另有说明，否则术语“亚烯基”是指将分子其余部分连接至基团的直链或支链二价烃链，其仅由碳和氢组成并且含有一个或多个碳-碳双键。在一实施方案中，亚烯基具有例如 2 至 24 个碳原子(C₂-C₂₄ 亚烯基)、2 至 15 个碳原子(C₂-C₁₅ 亚烯基)、2 至 12 个碳原子(C₂-C₁₂ 亚烯基)、2 至 8 个碳原子(C₂-C₈ 亚烯基)、2 至 6 个碳原子(C₂-C₆ 亚烯基)或 2 至 4 个碳原子(C₂-C₄ 亚烯基)。亚烯基的实例包括但不限于亚乙烯基、

30 亚丙烯基、亚正丁烯基等。亚烯基通过单键或双键连接至分子其余部分，并且通过单键或双键连接至基团。亚烯基与分子其余部分和与基团的连接点可通过链内的一个碳或任何两个碳。除非另有说明，否则亚烯基任选地经取代。

如本文所用并且除非另有说明，否则术语“环烷基”是指仅由碳和氢原子组成的非芳族饱和单环或多环烃基。环烷基可包括稠合或桥连环系统。在一实施方案中，环烷基具

有例如 3 至 15 个环碳原子(C₃-C₁₅ 环烷基)、3 至 10 个环碳原子(C₃-C₁₀ 环烷基)或 3 至 8 个环碳原子(C₃-C₈ 环烷基)。环烷基通过单键连接至分子其余部分。单环环烷基的实例包括但不限于环丙基、环丁基、环戊基、环己基、环庚基和环辛基。多环环烷基的实例包括但不限于金刚烷基、降冰片基、十氢萘基、7,7-二甲基-双环[2.2.1]庚基等。除非另有说明，否则环烷基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚环烷基”是二价环烷基。除非另有说明，否则亚环烷基任选地经取代。

如本文所用并且除非另有说明，否则术语“环烯基”是指仅由碳和氢原子组成并且包括一个或多个碳-碳双键的非芳族单环或多环烃基。环烯基可包括稠合或桥连环系统。在一实施方案中，环烯基具有例如 3 至 15 个环碳原子(C₃-C₁₅ 环烯基)、3 至 10 个环碳原子(C₃-C₁₀ 环烯基)或 3 至 8 个环碳原子(C₃-C₈ 环烯基)。环烯基通过单键连接至分子其余部分。单环环烯基的实例包括但不限于环丙烯基、环丁烯基、环戊烯基、环己烯基、环庚烯基、环辛烯基等。除非另有说明，否则环烯基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚环烯基”是二价环烯基。除非另有说明，否则亚环烯基任选地经取代。

如本文所用并且除非另有说明，否则术语“杂环基”是指含有一个或多个(例如一个、一个或两个、一个至三个、或一个至四个)独立地选自氮、氧、磷和硫的杂原子的非芳族基团单环或多环部分。杂环基可在任何杂原子或碳原子处连接至主结构。杂环基可为单环、双环、三环、四环或其它多环系统，其中多环系统可为稠合、桥连或螺环系统。杂环基多环系统可在一个或多个环中包含一个或多个杂原子。杂环基可为饱和或部分不饱和的。饱和杂环烷基可称为“杂环烷基”。部分不饱和的杂环烷基在杂环基含有至少一个双键时可称为“杂环烯基”，或在杂环基含有至少一个三键时可称为“杂环炔基”。在一实施方案中，杂环基具有例如 3 至 18 个环原子(3 元至 18 元杂环基)、4 至 18 个环原子(4 元至 18 元杂环基)、5 至 18 个环原子(5 元至 18 元杂环基)、4 至 8 个环原子(4 元至 8 元杂环基)或 5 至 8 个环原子(5 元至 8 元杂环基)。当在本文中出现时，数字范围，诸如“3 至 18”是指指定范围中的每个整数；例如，“3 至 18 个环原子”意指杂环基可由 3 个环原子、4 个环原子、5 个环原子、6 个环原子、7 个环原子、8 个环原子、9 个环原子、10 个环原子等(至多并且包括 18 个环原子)组成。杂环基的实例包括但不限于咪唑基、咪唑烷基、噁唑基、噁唑烷基、噻唑基、噻唑烷基、吡唑基、吡唑基、异噁唑烷基、异噁唑基、异噻唑烷基、异噻唑基、吗啉基、吡咯基、吡咯烷基、呋喃基、四氢呋喃基、噻吩基、吡啶基、哌啶基、喹啉基和异喹啉基。除非另有说明，否则杂环基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚杂环基”是二价杂环基。除非另有说明，否则亚杂环基任选地经取代。

如本文所用并且除非另有说明，否则术语“芳基”是指含有至少一个芳族烃环的单环芳族基团和/或多环单价芳族基团。在某些实施方案中，芳基具有 6 至 18 个环碳原子(C₆-C₁₈ 芳基)、6 至 14 个环碳原子(C₆-C₁₄ 芳基)或 6 至 10 个环碳原子(C₆-C₁₀ 芳基)。芳基的实例包括但不限于苯基、萘基、茚基、蒽基、蒽基、菲基、芘基、联苯基和联三苯基。术语“芳基”还指双环、三环或其它多环烃环，其中至少一个环是芳族环，并且其它环可为饱和、部分不饱和或芳族环，例如二氢萘基、茚基、二氢茚基或四氢萘基。

(tetrahydronaphthyl/tetralinyl)。除非另有说明，否则芳基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚芳基”是二价芳基。除非另有说明，否则亚芳基任选地经取代。

如本文所用并且除非另有说明，否则术语“杂芳基”是指含有至少一个芳族环的单环芳族基团和/或多环芳族基团，其中至少一个芳族环含有一个或多个(例如一个、一个或两个、一个至三个、或一个至四个)独立地选自 O、S 和 N 的杂原子。杂芳基可在任何杂原子或碳原子处连接至主结构。在某些实施方案中，杂芳基具有 5 至 20 个、5 至 15 个或 5 至 10 个环原子。术语“杂芳基”还指双环、三环或其它多环，其中至少一个环是芳族环，并且其它环可为饱和、部分不饱和或芳族环，其中至少一个芳族环含有一个或多个独立地选自 O、S 和 N 的杂原子。单环杂芳基的实例包括但不限于吡咯基、吡唑基、吡唑啉基、咪唑基、噁唑基、异噁唑基、噻唑基、噻二唑基、异噻唑基、呋喃基、噻吩基、噁二唑基、吡啶基、吡嗪基、嘧啶基、哒嗪基和三嗪基。双环杂芳基的实例包括但不限于吲哚基、苯并噻唑基、苯并噁唑基、苯并噻吩基、喹啉基、四氢异喹啉基、异喹啉基、苯并咪唑基、苯并吡喃基、吲哚嗪基、苯并呋喃基、异苯并呋喃基、色酮基、香豆素基、噌啉基、喹噁啉基、吡啶基、嘌呤基、吡咯并吡啶基、呋喃并吡啶基、噻吩并吡啶基、二氢异吲哚基和四氢喹啉基。三环杂芳基的实例包括但不限于咔唑基、苯并吡啶基、菲咯啉基、吡啶基、菲啶基和咕吨基。除非另有说明，否则杂芳基任选地经取代。

如本文所用并且除非另有说明，否则术语“亚杂芳基”是二价杂芳基。除非另有说明，否则亚杂芳基任选地经取代。

当本文所述的基团被称为“经取代”时，其可经一个或多个任何适当的取代基取代。取代基的说明性实例包括但不限于本文所提供的示例性化合物和实施方案中可见的那些，以及：卤素原子，诸如 F、Cl、Br 或 I；氰基；氧代基(=O)；羟基(-OH)；烷基；烯基；炔基；环烷基；芳基；-(C=O)OR'；-O(C=O)R'；-C(=O)R'；-OR'；-S(O)_xR'；-S-SR'；-C(=O)SR'；-SC(=O)R'；-NR'R'；-NR'C(=O)R'；-C(=O)NR'R'；-NR'C(=O)NR'R'；-OC(=O)NR'R'；-NR'C(=O)OR'；-NR'S(O)_xNR'R'；-NR'S(O)_xR'；以及-S(O)_xNR'R'，其中：R'在每次出现时独立地是 H、C₁-C₁₆ 烷基或环烷基，并且 x 是 0、1 或 2。在一些实施方案中，取代基是 C₁-C₁₂ 烷基。在其它实施方案中，取代基是环烷基。在其它实施方案中，取代基是卤基，诸如氟基。在其它实施方案中，取代基是氧代基。在其它实施方案中，取代基是羟基。在其它实施方案中，取代基是烷氧基(-OR')。在其它实施方案中，取代基是羧基。在其它实施方案中，取代基是氨基(-NR'R')。

如本文所用并且除非另有说明，否则术语“任选存在的”或“任选地”(例如任选地经取代)意指随后描述的事件或情况可能会发生或可能不会发生，并且该描述包括所述事件或情况发生的情况和其不发生的情况。举例来说，“任选地经取代的烷基”意指烷基可经取代或可不经取代，并且该描述包括经取代的烷基和不具有取代的烷基二者。

如本文所用并且除非另有说明，否则术语生物活性化合物的“前药”是指可在生理条件下或通过溶剂分解而转化为生物活性化合物的化合物。在一实施方案中，术语“前药”是指生物活性化合物的药学上可接受的代谢前体。前药在施用至有需要的受试者时可为无活性的，但在体内转化为生物活性化合物。前药通常在体内迅速转型以产生母体生物活性化合物，例如通过在血液中水解而转型。前药化合物在哺乳动物生物体中通常提供

溶解性、组织相容性或延迟释放的优点(参见 Bundgard, H., Design of Prodrugs (1985), 第 7-9 页, 第 21-24 页(Elsevier, Amsterdam)). 关于前药的论述提供于 Higuchi, T. 等人, A.C.S. Symposium Series, 第 14 卷; 以及 Bioreversible Carriers in Drug Design, Edward B. Roche 编辑, American Pharmaceutical Association and Pergamon Press, 1987 中。

5 在一实施方案中, 术语“前药”还意图包括任何共价键结的载体, 当将此类前药施用哺乳动物受试者时, 其在体内释放活性化合物。化合物的前药可通过修饰化合物中存在的官能团来制备, 其方式是使得修饰可在常规操作中或在体内裂解而得到母体化合物。前药包括羟基、氨基或巯基键结至任何基团的化合物, 当将化合物的前药施用哺乳动物受试者时, 所述基团裂解而分别形成游离羟基、游离氨基或游离巯基。

10 前药的实例包括但不限于本文所提供化合物中的醇官能团的乙酸酯、甲酸酯和苯甲酸酯衍生物或胺官能团的酰胺衍生物。

如本文所用并且除非另有说明, 否则术语“药学上可接受的盐”包括酸加成盐和碱加成盐二者。

药学上可接受的酸加成盐的实例包括但不限于盐酸、氢溴酸、硫酸、硝酸、磷酸等; 15 以及有机酸, 诸如但不限于乙酸、2,2-二氯乙酸、己二酸、褐藻酸、抗坏血酸、天冬氨酸、苯磺酸、苯甲酸、4-乙酰胺基苯甲酸、樟脑酸、樟脑-10-磺酸、癸酸、己酸、辛酸、碳酸、肉桂酸、柠檬酸、环拉酸(cyclamic acid)、十二烷基硫酸、乙烷-1,2-二磺酸、乙烷磺酸、2-羟基乙烷磺酸、甲酸、延胡索酸、半乳糖二酸、龙胆酸、葡庚糖酸、葡萄糖酸、葡萄糖醛酸、谷氨酸、戊二酸、2-氧代戊二酸、甘油磷酸、乙醇酸、马尿酸、异丁酸、乳 20 酸、乳糖酸、月桂酸、马来酸、苹果酸、丙二酸、扁桃酸、甲烷磺酸、粘液酸、萘-1,5-二磺酸、萘-2-磺酸、1-羟基-2-萘甲酸、烟酸、油酸、乳清酸、草酸、棕榈酸、帕莫酸(pamoic acid)、丙酸、焦谷氨酸、丙酮酸、水杨酸、4-氨基水杨酸、癸二酸、硬脂酸、琥珀酸、酒石酸、硫氰酸、对甲苯磺酸、三氟乙酸、十一碳烯酸等。

药学上可接受的碱加成盐的实例包括但不限于通过将无机碱或有机碱添加至游离 25 酸化合物而制备的盐。衍生自无机碱的盐包括但不限于钠盐、钾盐、锂盐、铵盐、钙盐、镁盐、铁盐、锌盐、铜盐、锰盐、铝盐等。在一实施方案中, 无机盐是铵盐、钠盐、钾盐、钙盐和镁盐。衍生自有机碱的盐包括但不限于以下的盐: 伯胺、仲胺和叔胺; 经取代胺, 包括天然存在的经取代胺; 环胺和碱性离子交换树脂, 诸如氨、异丙胺、三甲胺、二乙胺、三乙胺、三丙胺、二乙醇胺、乙醇胺、丹醇(deanol)、2-二甲基氨基乙醇、2-二 30 乙基氨基乙醇、二环己胺、赖氨酸、精氨酸、组氨酸、咖啡因、普鲁卡因(procaine)、哈胺(hydrabamine)、胆碱、甜菜碱、苄乙苄胺(benethamine)、苄星(benzathine)、乙二胺、葡萄糖胺、甲基葡萄糖胺、可可碱(theobromine)、三乙醇胺、氨基丁三醇、嘌呤、腺嘌呤、N-乙基腺嘌呤、聚胺树脂等。在一实施方案中, 有机碱是异丙胺、二乙胺、乙醇胺、三甲胺、二环己胺、胆碱和咖啡因。

35 本文所提供的化合物可含有一个或多个不对称中心, 因此可产生对映异构体、非对映异构体和其它立体异构形式, 所述形式可根据绝对立体化学定义为(R)-或(S)-或对于氨基酸可定义为(D)-或(L)-。除非另有说明, 否则本文所提供的化合物意图包括所有此类可能的异构体, 以及它们的外消旋和光学纯形式。光学活性(+)与(-)、(R)-与(S)-或(D)-与(L)-异构体可使用手性合成子或手性试剂制备, 或使用常规技术, 例如色谱法和分步结晶法

拆分。用于制备/分离个别对映异构体的常规技术包括从适合的光学纯前体手性合成或使用例如手性高压液相色谱法(HPLC)拆分外消旋物(或盐或衍生物的外消旋物)。当本文所述化合物含有烯属双键或其它几何不对称中心时,除非另有说明,否则所述化合物意图包括 E 和 Z 几何异构体。同样,还意图包括所有互变异构形式。

5 如本文所用并且除非另有说明,否则术语“异构体”是指具有相同分子式的不同化合物。“立体异构体”是仅原子在空间中的排列方式不同的异构体。“阻转异构体”是由围绕单键的受阻旋转得到的立体异构体。“对映异构体”是一对互为不可重叠的镜像的立体异构体。一对对映异构体的任何比例的混合物可称为“外消旋”混合物。“非对映异构体”是具有至少两个不对称原子但不互为镜像的立体异构体。

10 “立体异构体”还可包括 E 和 Z 异构体或其混合物,以及顺式和反式异构体或其混合物。在某些实施方案中,本文所述的化合物经分离为 E 或 Z 异构体。在其它实施方案中,本文所述的化合物是 E 和 Z 异构体的混合物。

“互变异构体”是指化合物的彼此平衡的异构形式。异构形式的浓度将取决于化合物所处的环境,并且可根据例如化合物是固体还是在有机溶液或水溶液中而有所不同。

15 还应注意,本文所述的化合物可在一个或多个原子处含有非天然比例的原子同位素。举例来说,化合物可用放射性同位素,诸如氚(^3H)、碘-125 (^{125}I)、硫-35 (^{35}S)或碳-14 (^{14}C)进行放射性标记,或者可为同位素富集的,诸如氘(^2H)、碳-13 (^{13}C)或氮-15 (^{15}N)。如本文所用,“同位素体”是同位素富集的化合物。术语“同位素富集”是指原子的同位素组成不同于该原子的天然同位素组成。“同位素富集”还可指化合物含有的至少一个原子的同位素组成不同于该原子的天然同位素组成。术语“同位素组成”是指给定原子存在的每种同位素的量。经放射性标记和同位素富集的化合物可用作治疗剂,例如癌症治疗剂;研究试剂,例如结合分析试剂;以及诊断剂,例如体内成像剂。本文所述化合物的所有同位素变化形式,无论是否具有放射性,都意图涵盖在本文所提供的实施方案的范围内。在一些实施方案中,提供本文所述化合物的同位素体,例如同位素体是氘、碳-13 和/或氮-15 富集的。如本文所用,“氘代”是指化合物中的至少一个氢(H)已经氘(以 D 或 ^2H 表示)置换,即,化合物在至少一个位置处富含氘。

25 应注意,如果所描绘的结构与该结构的名称之间存在差异,则应以所描绘的结构为准。

如本文所用并且除非另有说明,否则术语“药学上可接受的载剂、稀释剂或赋形剂”包括但不限于被美国食品与药物管理局(United States Food and Drug Administration)批准可接受用于人类或家畜的任何佐剂、载剂、赋形剂、助流剂、甜味剂、稀释剂、防腐剂、染料/着色剂、风味增强剂、表面活性剂、润湿剂、分散剂、助悬剂、稳定剂、等渗剂、溶剂或乳化剂。

术语“组合物”意图涵盖含有任选地选用的指定量的指定成分(例如本文所提供的 mRNA 分子)的产品。

35 本文所用的核酸构建体名称(如 0201、0302、0402、0505、0603、0702、0803、0201-J0、0301-J0、0501-J0、0701-J0、0901-J0、1101、1201、1301、1401、1501、1601、1701、1801、1901、2001 等)仅为区分或标识核酸、构建体、序列或产品所用,并不意图表示这样的标识是本发明的核酸或核酸组合、构建体、序列或产品的特征。本领域技术人员

应当理解，例如出于区分或标识的目的，其他核酸或核酸组合、构建体、序列或产品同样也有可能使用这样的标识，但并非是指相同或等价的核酸或核酸组合、构建体、序列或产品。类似地，实施例中所用的类似编号或标识也仅仅是用于示例方便，本发明的核酸或核酸组合、构建体、序列或产品由所附权利要求中所描述的特征限定。

5

猴痘病毒

如本文所用，术语“猴痘病毒(Mpox Virus, MPXV)”属于痘病毒科正痘病毒属，同属的还有天花病毒、牛痘病毒，MPXV 与其他已知的正痘病毒具有相似的形态。MPXV 呈椭圆形或砖形，被脂蛋白外膜包裹，其基因组为线性双链 DNA，基因组大小约为 197kb。

10 MPXV 可通过口咽、鼻咽或皮内途径进入宿主。病毒在感染部位复制，然后扩散到区域淋巴结。虽然 MPXV 是一种 DNA 病毒，但其生命周期发生在细胞质中。MPXV 复制过程产生两种感染性病毒颗粒：成熟病毒粒子(MV)和包膜病毒粒子(EV)；其中，MV 通过胞吞或膜融合进入细胞，而 EV 通过膜融合进入。在细胞质中完成病毒 DNA 的复制和病毒蛋白的组装后，MV 经细胞裂解被释放，而 EV 经胞吐被释放(参见，例如 Kumar N et al., The 2022 outbreak and the pathobiology of the monkeypox virus. J Autoimmun. 2022 Jul;131:102855.)。

如本文所用，术语“A29L 蛋白”与“A29L 肽”、“A35R 蛋白”与“A35R 肽”、“B2R 蛋白”与“B2R 肽”、“B6R 蛋白”与“B6R 肽”、“C15L 蛋白”与“C15L 肽”、“M1R 蛋白”与“M1R 肽”、“H3L 蛋白”与“H3L 肽”、“E8L 蛋白”与“E8L 肽”可各自互换使用，其都为猴痘病毒的主要抗原蛋白(或抗原多肽)。猴痘病毒的“A29L 蛋白”、“A35R 蛋白”、“B2R 蛋白”、“B6R 蛋白”、“M1R 蛋白”、“E8L 蛋白”、“H3L 蛋白”、“C15L 蛋白”分别与牛痘病毒的“A27L 蛋白”、“A33R 蛋白”、“A56R 蛋白”、“B5R 蛋白”、“L1R 蛋白”、“D8L 蛋白”、“H3L 蛋白”、“F9L 蛋白”同源。

据报道，“A29L 蛋白”为位于成熟病毒粒子的表面膜融合蛋白，可以与细胞表面乙酰肝素结合；“A35R 蛋白”为位于包膜病毒粒子的包膜蛋白，参与含肌动蛋白微绒毛的形成和病毒粒子的细胞间扩散；“B2R 蛋白”为位于包膜病毒粒子的膜糖蛋白血凝素，可以防止细胞融合；“B6R 蛋白”为位于包膜病毒粒子的棕榈酰化糖蛋白，是有效细胞扩散所必需的，参与补体调节；“M1R 蛋白”为位于成熟病毒粒子的肉豆蔻酰化表面膜蛋白，在病毒进入宿主过程中发挥重要作用；“E8L 蛋白”为位于成熟病毒粒子的表面膜吸附蛋白，可以与硫酸软骨素结合；“H3L 蛋白”为位于成熟病毒粒子的表面膜蛋白，可以与乙酰肝素和细胞表面蛋白聚糖结合；“C15L 蛋白”位于成熟病毒粒子(参见，例如 Papukashvili D et al., Strategy of developing nucleic acid-based universal monkeypox vaccine candidates. Front Immunol. 2022 Oct 27;13:1050309.)。

如本文所用，A29L 蛋白如 NCBI 登录号 YP_010377135.1 所示；A35R 蛋白如 NCBI 登录号 YP_010377142.1 所示；B2R 蛋白如 NCBI 登录号 YP_010377158.1 所示；B6R 蛋白如 NCBI 登录号 NP_536594.1 所示；C15L 蛋白如 NCBI 登录号 URF91574.1 所示；M1R 蛋白如 NCBI 登录号 URF91574.1 或 NP_536507.1 所示；H3L 蛋白如 NCBI 登录号 UVB87333.1 所示；E8L 蛋白如 NCBI 登录号 YP_010377102.1 所示。

多肽

在一方面，本发明提供一种多肽或多肽组合，其包含一个或多个肽段，肽段各自独立地选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。

在一些实施方案中，多肽或多肽组合包含 1、2、3、4、5、6、7、8、9、10、11、
5 12、13、14、15、16、17、18、19、20、21、22 个肽段，所述肽段各自独立地选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。

在一些实施方案中，多肽或多肽组合包含 1 种肽段，所述肽段选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段；或者多肽或多肽组合包含 2、3、4、5、6 或 7 种各自源自不同蛋白的肽段，所述肽段选自猴痘病毒 A29L、A35R、
10 B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。

在一实施方案中，多肽或多肽组合包含 3 种各自源自不同蛋白的肽段，所述肽段选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。

在一优选实施方案中，多肽或多肽组合包含 6 种各自源自不同蛋白的肽段，所述肽段选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。

15 在一些实施方案中，多肽或多肽组合还包含一个或多个接头、IgE 信号肽和/或 HA 多肽标签。

在一实施方案中，多肽或多肽组合包含 A29L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A35R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B2R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B6R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 H3L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A35R、B2R 和 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A29L、B6R 和 M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B6R、M1R 和 H3L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B6R、M1R 和 A35R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A29L、B2R 和 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B2R、H3L 和 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 B2R、A35R 和 B6R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A29L、M1R 和 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 H3L、M1R 和 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段。
20 25 30

在一实施方案中，A29L 蛋白中的肽段包含 A29L 肽全长。在一实施方案中，A29L 蛋白中的肽段包含 A29L 肽全长，A29L 肽全长包含 SEQ ID NO: 111 或 113 的氨基酸序列。
35

在一实施方案中，A35R 蛋白中的肽段包含 A35R 肽全长。在一实施方案中，A35R 肽全长包含 SEQ ID NO: 114 的氨基酸序列。在一实施方案中，A35R 蛋白中的肽段包含 A35R 胞外段。在一实施方案中，A35R 胞外段包含 SEQ ID NO: 115 的氨基酸序列。

在一实施方案中，B2R 蛋白中的肽段包含 B2R 信号肽。在一实施方案中，B2R 信

- 号肽包含 SEQ ID NO: 116 的氨基酸序列。在一实施方案中, B2R 蛋白中的肽段包含除 B2R 信号肽部分以外的 B2R 肽全长。在一实施方案中, B2R 蛋白中的肽段包含除 B2R 信号肽部分以外的 B2R 肽全长, B2R 蛋白中的肽段包含 SEQ ID NO: 117 的氨基酸序列。在一实施方案中, B2R 蛋白中的肽段包含 B2R 肽全长。在一实施方案中, B2R 肽全长
- 5 包含 SEQ ID NO: 118 的氨基酸序列。在一实施方案中, B2R 蛋白中的肽段包含 B2R 胞外段。在一实施方案中, B2R 胞外段包含 SEQ ID NO: 119 的氨基酸序列。在一实施方案中, B2R 蛋白中的肽段包含 B2R 信号肽和 B2R 胞外段。在一实施方案中, B2R 蛋白中的肽段包含 B2R 信号肽和 B2R 胞外段, B2R 蛋白中的肽段包含 SEQ ID NO: 120 的氨基酸序列。
- 10 在一实施方案中, B6R 蛋白中的肽段包含 B6R 信号肽。在一实施方案中, B6R 信号肽包含 SEQ ID NO: 121 的氨基酸序列。在一实施方案中, B6R 蛋白中的肽段包含除 B6R 信号肽部分以外的 B6R 肽全长。在一实施方案中, B6R 蛋白中的肽段包含除 B6R 信号肽部分以外的 B6R 肽全长, B6R 蛋白中的肽段包含 SEQ ID NO: 122 的氨基酸序列。在一实施方案中, B6R 蛋白中的肽段包含 B6R 肽全长。在一实施方案中, B6R 肽全长
- 15 包含 SEQ ID NO: 123 的氨基酸序列。在一实施方案中, B6R 蛋白中的肽段包含 B6R 胞外段。在一实施方案中, B6R 胞外段包含 SEQ ID NO: 124 的氨基酸序列。在一实施方案中, B6R 蛋白中的肽段包含 B6R 信号肽和 B6R 胞外段。在一实施方案中, B6R 蛋白中的肽段包含 B6R 信号肽和 B6R 胞外段, B6R 蛋白中的肽段包含 SEQ ID NO: 125 的氨基酸序列。
- 20 在一实施方案中, C15L 蛋白中的肽段包含 C15L 肽全长。在一实施方案中, C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中, C15L 蛋白中的肽段包含 C15L 胞外段。在一实施方案中, C15L 胞外段包含 SEQ ID NO: 127 的氨基酸序列。
- 在一实施方案中, M1R 蛋白中的肽段包含 M1R 肽全长。在一实施方案中, M1R 肽全长包含 SEQ ID NO: 130 的氨基酸序列。在一实施方案中, M1R 蛋白中的肽段包含
- 25 M1R 胞外段。在一实施方案中, M1R 胞外段包含 SEQ ID NO: 128 的氨基酸序列。
- 在一实施方案中, H3L 蛋白中的肽段包含 H3L 肽全长。在一实施方案中, H3L 肽全长包含 SEQ ID NO: 131 的氨基酸序列。在一实施方案中, H3L 蛋白中的肽段包含 H3L 胞外段。在一实施方案中, H3L 胞外段包含 SEQ ID NO: 132 的氨基酸序列。
- 在一实施方案中, 多肽或多肽组合包含 IgE 信号肽、A29L 蛋白中的肽段以及 HA 多肽标签。在一实施方案中, 多肽或多肽组合包含 A35R 蛋白中的肽段以及 HA 多肽标签。在一实施方案中, 多肽或多肽组合包含 B2R 蛋白中的肽段以及 HA 多肽标签。在一实施方案中, 多肽或多肽组合包含 B6R 蛋白中的肽段、接头以及 HA 多肽标签。在一实施方案中, 多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 C15L 蛋白中的肽段。在一实施方案中, 多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 M1R
- 30 蛋白中的肽段。在一实施方案中, 多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 H3L 蛋白中的肽段。在一实施方案中, 多肽或多肽组合包含 A35R、B2R 和 C15L 蛋白中的肽段以及接头。在一实施方案中, 多肽或多肽组合包含 A29L、B6R 和 M1R 蛋白中的肽段、IgE 信号肽以及接头。在一实施方案中, 多肽或多肽组合包含 B6R、M1R 和 H3L 蛋白中的肽段以及接头。在一实施方案中, 多肽或多肽组合包含 B6R、M1R 和

A35R 蛋白中的肽段以及接头。在一实施方案中，多肽或多肽组合包含 A29L、B2R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头。在一实施方案中，多肽或多肽组合包含 B2R、H3L 和 C15L 蛋白中的肽段以及接头。在一实施方案中，多肽或多肽组合包含 B2R、A35R 和 B6R 蛋白中的肽段以及接头。在一实施方案中，多肽或多肽组合包含 A29L、M1R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头。在一实施方案中，多肽或多肽组合包含 H3L、M1R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头。在一实施方案中，多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段、IgE 信号肽以及接头。在一实施方案中，多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段、IgE 信号肽、HA 多肽标签以及接头。在一实施方案中，多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段、IgE 信号肽、HA 多肽标签以及接头。在一实施方案中，多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段、IgE 信号肽以及接头。

接头可以包含任何长度的氨基酸序列，特别是 1-50，优选 1-30，例如 1-10 个氨基酸残基的氨基酸序列。示例性的接头可以包括但不限于聚甘氨酸(G)、聚丙氨酸(A)、聚丝氨酸(S)或其组合，例如 GGAS、GGGS、GGGSG 或者(G4S)_n，其中 n 为 1-30、优选 1-10 的整数。接头还可以是铰链区或其功能性等同物。其它合适的接头可以是一般适用于药用蛋白中的有机化合物或聚合物，包括但不限于聚乙二醇。

如本文所用，术语“HA 多肽标签”和“HA 标签蛋白”可互换使用，其源于流感病毒的红细胞凝集素表面抗原决定簇，其对外源靶蛋白的空间结构影响小，容易构建成标签蛋白融合到 N 端或者 C 端。“HA 多肽标签”和“HA 标签蛋白”被广泛用于蛋白的检测、分离和纯化。

在一实施方案中，多肽或多肽组合包含 IgE 信号肽。在一实施方案中，IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 HA 多肽标签。在一实施方案中，HA 多肽标签包含 SEQ ID NO: 135 或 136 的氨基酸序列。在一实施方案中，多肽或多肽组合包含接头。在一实施方案中，接头为 GS 接头。在一优选实施方案中，GS 接头包含 SEQ ID NO: 137 或 138 的氨基酸序列。

在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列或者与 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列具有至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同性的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段以及 HA 多肽标签。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段以及 HA 多肽标签；IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，A29L 蛋白中的肽段包含 SEQ ID NO: 111 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 135 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 89 的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 A35R 蛋白中的肽段以及 HA

多肽标签。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 A35R 肽全长以及 HA 多肽标签，A35R 肽全长包含 SEQ ID NO: 114 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 90 的氨基酸序列。

5 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 蛋白中的肽段以及 HA 多肽标签。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、除信号肽部分以外的 B2R 肽全长以及 HA 多肽标签。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、除信号肽部分以外的 B2R 肽全长以及 HA 多肽标签，B2R 信号肽包含 SEQ ID NO: 116 的氨基酸序列，除信号肽部分以外的 B2R 肽全长包含 SEQ
10 ID NO: 117 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 肽全长以及 HA 多肽标签。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 肽全长以及 HA 多肽标签，B2R 肽全长包含 SEQ ID NO: 118 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 91 的氨基酸序列。

15 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 蛋白中的肽段、接头、HA 多肽标签以及 B6R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、GS 接头、HA 多肽标签以及除信号肽部分以外的 B6R 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、GS 接头、HA 多肽标签以及除信号肽部分以外的 B6R 肽全长，B6R 信号肽包含 SEQ ID NO: 121 的氨基酸序列，GS 接头包含 SEQ ID NO: 137 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列，除信号肽部分以外的 B6R 肽全长包含 SEQ ID NO: 122 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 92 的氨基酸序列。

20 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、接头、HA 多肽标签以及 C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 C15L 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 C15L 肽全长，IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，GS 接头包含 SEQ ID NO: 137 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列，C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 93 的氨基酸序列。

25 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、接头、HA 多肽标签以及 M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 M1R 胞外段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 M1R 胞外段，IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，GS 接头包含 SEQ ID NO: 137 的氨基酸序列，
30 HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列，M1R 胞外段包含 SEQ ID NO: 128 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 94 的氨基酸序列。

35 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、接头、HA 多肽标签以及 H3L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 H3L 肽全长。在一实施方案中，多肽或多肽组合

从 N 端至 C 端包含 IgE 信号肽、GS 接头、HA 多肽标签以及 H3L 肽全长，IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，GS 接头包含 SEQ ID NO: 137 的氨基酸序列，HA 多肽标签包含 SEQ ID NO: 136 的氨基酸序列，H3L 肽全长包含 SEQ ID NO: 131 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 95 的氨基酸序列。

- 5 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段、接头、A35R 蛋白中的肽段、接头、B6R 蛋白中的肽段、接头、B2R 蛋白中的肽段、接头、C15L 蛋白中的肽段、接头、M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、A35R 胞外段、GS 接头、B6R 胞外段、GS 接头、B2R 胞外段、GS 接头、C15L 胞外段、GS 接头、M1R 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、A35R 胞外段、GS 接头、B6R 胞外段、GS 接头、B2R 胞外段、GS 接头、C15L 胞外段、GS 接头、M1R 肽全长；IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，A29L 肽全长包含 SEQ ID NO: 113 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，A35R 胞外段包含 SEQ ID NO: 115 的氨基酸序列，B6R 胞外段包含 SEQ ID NO: 124 的氨基酸序列，B2R 胞外段包含 SEQ ID NO: 119 的氨基酸序列，C15L 胞外段包含 SEQ ID NO: 127 的氨基酸序列，M1R 肽全长包含 SEQ ID NO: 130 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 101 的氨基酸序列。
- 10
- 15

- 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 蛋白中的肽段、接头、A35R 蛋白中的肽段、接头、C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、A35R 胞外段、GS 接头、C15L 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、A35R 胞外段、GS 接头、C15L 肽全长；B2R 信号肽包含 SEQ ID NO: 116 的氨基酸序列，B2R 胞外段包含 SEQ ID NO: 119 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，A35R 胞外段包含 SEQ ID NO: 115 的氨基酸序列，C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 102 的氨基酸序列。
- 20
- 25

- 在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段、接头、M1R 蛋白中的肽段、接头、B6R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、M1R 胞外段、GS 接头、除信号肽部分以外的 B6R 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、M1R 胞外段、GS 接头、除信号肽部分以外的 B6R 肽全长；IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，A29L 肽全长包含 SEQ ID NO: 113 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，M1R 胞外段包含 SEQ ID NO: 128 的氨基酸序列，除信号肽部分以外的 B6R 肽全长包含 SEQ ID NO: 122 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 103 的氨基酸序列。
- 30
- 35

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 蛋白中的肽段、接头、M1R 蛋白中的肽段、接头、H3L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、B6R 胞外段、GS 接头、M1R 胞外段、GS 接头、H3L 肽全

长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、B6R 胞外段、GS 接头、M1R 胞外段、GS 接头、H3L 肽全长；B6R 信号肽包含 SEQ ID NO: 121 的氨基酸序列，B6R 胞外段包含 SEQ ID NO: 124 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，M1R 胞外段包含 SEQ ID NO: 128 的氨基酸序列，H3L 肽全长包含
5 SEQ ID NO: 131 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 104 的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 蛋白中的肽段、接头、A35R 蛋白中的肽段、接头、M1R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、B6R 胞外段、GS 接头、A35R 胞外段、GS 接头、M1R
10 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B6R 信号肽、B6R 胞外段、GS 接头、A35R 胞外段、GS 接头、M1R 肽全长；B6R 信号肽包含 SEQ ID NO: 121 的氨基酸序列，B6R 胞外段包含 SEQ ID NO: 124 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，A35R 胞外段包含 SEQ ID NO: 115 的氨基酸序列，M1R 肽全长包含 SEQ ID NO: 130 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO:
15 105 的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段、接头、B2R 蛋白中的肽段、接头、C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、B2R 胞外段、GS
20 接头、C15L 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、B2R 胞外段、GS 接头、C15L 肽全长；IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列，A29L 肽全长包含 SEQ ID NO: 113 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，B2R 胞外段包含 SEQ ID NO: 119 的氨基酸序列，C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 106 的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 蛋白中的肽段、接头、H3L 蛋白中的肽段、接头、C15L 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、H3L 胞外段、GS 接头、C15L 肽
25 全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、H3L 胞外段、GS 接头、C15L 肽全长；B2R 信号肽包含 SEQ ID NO: 116 的氨基酸序列，B2R 胞外段包含 SEQ ID NO: 119 的氨基酸序列，GS 接头包含 SEQ ID NO: 138 的氨基酸序列，H3L 胞外段包含 SEQ ID NO: 132 的氨基酸序列，C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中，多肽或多肽组合包含 SEQ ID NO: 107 的氨基酸序列。

在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 蛋白中的肽段、接头、A35R 蛋白中的肽段、接头、B6R 蛋白中的肽段。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、A35R 胞外段、GS 接头、除信号
35 肽部分以外的 B6R 肽全长。在一实施方案中，多肽或多肽组合从 N 端至 C 端包含 B2R 信号肽、B2R 胞外段、GS 接头、A35R 胞外段、GS 接头、除信号肽部分以外的 B6R 肽全长；B2R 信号肽包含 SEQ ID NO: 116 的氨基酸序列，B2R 胞外段包含 SEQ ID NO:

119 的氨基酸序列, GS 接头包含 SEQ ID NO: 138 的氨基酸序列, A35R 胞外段包含 SEQ ID NO: 115 的氨基酸序列, 除信号肽部分以外的 B6R 肽全长包含 SEQ ID NO: 122 的氨基酸序列。在一实施方案中, 多肽或多肽组合包含 SEQ ID NO: 108 的氨基酸序列。

5 在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 蛋白中的肽段、接头、M1R 蛋白中的肽段、接头、C15L 蛋白中的肽段。在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、M1R 胞外段、GS 接头、C15L 肽全长。在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、A29L 肽全长、GS 接头、M1R 胞外段、GS 接头、C15L 肽全长; IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列, A29L 肽全长包含 SEQ ID NO: 113 的氨基酸序列, GS 接头包含 SEQ ID NO: 138 的氨基酸序列, M1R 胞外段包含 SEQ ID NO: 128 的氨基酸序列, C15L 肽全长包含 SEQ ID NO: 126 的氨基酸序列。在一实施方案中, 多肽或多肽组合包含 SEQ ID NO: 109 的氨基酸序列。

15 在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、H3L 蛋白中的肽段、接头、C15L 蛋白中的肽段、接头、M1R 蛋白中的肽段。在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、H3L 胞外段、GS 接头、C15L 胞外段、GS 接头、M1R 肽全长。在一实施方案中, 多肽或多肽组合从 N 端至 C 端包含 IgE 信号肽、H3L 胞外段、GS 接头、C15L 胞外段、GS 接头、M1R 肽全长; IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列, H3L 胞外段包含 SEQ ID NO: 132 的氨基酸序列, GS 接头包含 SEQ ID NO: 138 的氨基酸序列, C15L 胞外段包含 SEQ ID NO: 127 的氨基酸序列, M1R 肽全长包含 SEQ ID NO: 130 的氨基酸序列。在一实施方案中, 多肽或多肽组合包含 SEQ ID NO: 110 的氨基酸序列。

25 在一实施方案中, 多肽或多肽组合包含两个或更多个选自 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 的氨基酸序列或者两个或更多个与选自 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 的氨基酸序列具有至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 相同性的氨基酸序列。

30 在一实施方案中, 多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽, 其中第一多肽包含 SEQ ID NO: 89 的氨基酸序列, 第二多肽包含 SEQ ID NO: 90 的氨基酸序列, 第三多肽包含 SEQ ID NO: 91 的氨基酸序列, 第四多肽包含 SEQ ID NO: 92 的氨基酸序列, 第五多肽包含 SEQ ID NO: 93 的氨基酸序列, 第六多肽包含 SEQ ID NO: 94 的氨基酸序列。

35 在一实施方案中, 多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽, 其中第一多肽包含 SEQ ID NO: 90 的氨基酸序列, 第二多肽包含 SEQ ID NO: 91 的氨基酸序列, 第三多肽包含 SEQ ID NO: 92 的氨基酸序列, 第四多肽包含 SEQ ID NO: 93 的氨基酸序列, 第五多肽包含 SEQ ID NO: 94 的氨基酸序列, 第六多肽分包含 SEQ ID NO: 95 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含 SEQ ID NO: 102 的氨基酸序列, 第二多肽包含 SEQ ID NO: 103 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含

SEQ ID NO: 105 的氨基酸序列, 第二多肽包含 SEQ ID NO: 106 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含 SEQ ID NO: 108 的氨基酸序列, 第二多肽包含 SEQ ID NO: 109 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含
5 SEQ ID NO: 102 的氨基酸序列, 第二多肽包含 SEQ ID NO: 104 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含 SEQ ID NO: 105 的氨基酸序列, 第二多肽包含 SEQ ID NO: 107 的氨基酸序列。

在一实施方案中, 多肽或多肽组合包含第一多肽和第二多肽, 其中第一多肽包含 SEQ ID NO: 110 的氨基酸序列, 第二多肽包含 SEQ ID NO: 108 的氨基酸序列。

10 本发明的多肽或多肽组合可以作为多肽抗原用于在受试者中诱导针对猴痘病毒感染的保护性免疫应答。

核酸

在一方面, 本发明提供编码本文所述多肽或多肽组合的核酸或核酸组合。在一些实
15 施方案中, 核酸或核酸组合包含编码本文所述的多肽或多肽组合的多核苷酸。核酸可以是单链的或双链的。核酸包括但不限于 DNA、cDNA、RNA (例如 mRNA、circRNA 或 saRNA)、重组产生和化学合成的核酸。本发明的核酸或核酸组合可以包含在载体中。本发明的核酸可以包括天然存在的、合成的和修饰的核苷酸。在一些实施方案中, 本发明的核酸或核酸组合用于在细胞中表达本文所述的多肽或多肽组合以提供多肽抗原。在
20 一些实施方案中, 多肽抗原可以在合适的受试者中诱导针对猴痘病毒的免疫应答。

核酸可以包含一个或多个区段(核苷酸片段)(例如 1、2、3、4、5、6、7、8 或更多个区段)。核酸可以包含编码感兴趣多肽 (例如本文所述多肽和多肽抗原)的区段。在特定实施方案中, 核酸可以包含感兴趣多肽的编码序列以及调控序列(包括但不限于转录和翻译调控序列)。在一实施方案中, 调控序列包含以下的一个或多个: 启动子序列、5'非
25 翻译区(5' UTR)序列、3'非翻译区(3' UTR)序列和 poly(A)序列。

编码序列

如本文所用, “编码序列”是指核酸中可以作为模板用于在生物过程中合成具有确定的核苷酸序列(例如 tRNA 和 mRNA)或确定的氨基酸序列的核苷酸序列。编码序列可以
30 是 DNA 序列或 RNA 序列。如果对应于 DNA 序列(包括与 mRNA 序列相同的编码链和与之互补链的模板链)的 mRNA 在生物过程中翻译成多肽, 可以认为所述 DNA 序列或 mRNA 序列编码所述多肽。

如本文所用, “密码子”指核酸中三个连续的核苷酸序列(又称三联体密码), 其编码特定的氨基酸。同义密码子(编码相同氨基酸的密码子)在不同物种中使用的频率不同,
35 称为“密码子偏好性”。通常认为, 对于给定物种, 使用其偏好的密码子的编码序列可以在该物种表达系统中具有较高的翻译效率和准确率。因此, 可以对核酸进行“密码子优化”, 即改变核酸中的密码子以反映宿主细胞偏好的密码子, 而优选不改变其编码的氨基酸序列。本领域技术人员会理解, 由于密码子的简并性, 本发明的核酸或核酸组合可以包含这样的编码序列, 其与本文所述编码序列不同(例如与本文所述编码序列具有约

70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性)但编码相同的氨基酸序列。在特定实施方案中,本发明的核酸包含针对宿主(例如受试者,特别是人)细胞优化的密码子,使得本发明的多肽在宿主(例如受试者,特别是人)中最佳表达。在特定实施方案中,本发明的 RNA 包含针对宿主(例如受试者,特别是人)细胞优化的密码子,使得本发明的多肽在宿主(例如受试者,特别是人)中最佳表达。

在一些实施方案中,本发明的核酸或核酸组合包含如本文所述多肽的编码序列。在一些实施方案中,本发明的核酸或核酸组合包含如本文所述多肽抗原的编码序列。在一些实施方案中,本发明的核酸或核酸组合包含与本文所述编码序列互补的核苷酸序列。在一些实施方案中,编码序列在其 5'端包含起始密码子,并且在其 3'端包含终止密码子。在一些实施方案中,编码序列包含本文所述的开放阅读框(ORF)。

在一些实施方案中,本发明的编码序列编码本发明的任意多肽抗原。

在一实施方案中,本发明的编码序列编码多肽抗原,多肽抗原包含:

(1) SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列; (2) 与 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列具有至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同性的氨基酸序列; (3) SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列的免疫原性片段; (4) 与 SEQ ID NO: 89、90、91、92、93、94、95、101、102、103、104、105、106、107、108、109 和 110 之一的氨基酸序列具有至少 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%相同性的氨基酸序列的免疫原性片段。

在一实施方案中,本发明的编码序列编码多肽,多肽包含如上所述多肽抗原的氨基酸序列。

在一实施方案中,本文所述的多肽或多肽抗原的编码序列包含核苷酸序列,核苷酸序列包含: (1) SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列; (2) 与 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列; (3) SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列; 或者 (4) 与 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中,本文所述的多肽或多肽组合的编码序列包含核苷酸序列,核苷酸序列包含: (1) 两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列; (2) 两个或更多个与选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列; (3) 两个或更多个选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、

41、42、43 和 44 的核苷酸序列；或者(4) 两个或更多个与选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸、第二多核苷酸、
5 第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 1 的核苷酸序列或者与 SEQ ID NO: 1 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 2 的核苷酸序列或者与 SEQ ID NO: 2 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、
10 98%、99%相同性的核苷酸序列；第三多核苷酸的核苷酸序列包含：SEQ ID NO: 3 的核苷酸序列或者与 SEQ ID NO: 3 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第四多核苷酸的核苷酸序列包含：SEQ ID NO: 4 的核苷酸序列或者与 SEQ ID NO: 4 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第五多核苷酸的
15 核苷酸序列包含：SEQ ID NO: 5 的核苷酸序列或者与 SEQ ID NO: 5 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第六多核苷酸的核苷酸序列包含：SEQ ID NO: 6 的核苷酸序列或者与 SEQ ID NO: 6 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 23 的核苷酸序列或者与 SEQ ID NO: 23 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 24 的核苷酸序列或者
25 与 SEQ ID NO: 24 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第三多核苷酸的核苷酸序列包含：SEQ ID NO: 25 的核苷酸序列或者与 SEQ ID NO: 25 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第四多核苷酸的核苷酸序列包含：SEQ ID NO: 26 的核苷酸序列或者与 SEQ ID NO: 26 的核苷酸序列具有至少 70%、
30 75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第五多核苷酸的核苷酸序列包含：SEQ ID NO: 27 的核苷酸序列或者与 SEQ ID NO: 27 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；第六多核苷酸的核苷酸序列包含：SEQ ID NO: 28 的核苷酸序列或者与 SEQ ID NO: 28 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、
35 98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 2 的核苷酸序列或者与 SEQ ID NO: 2 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性

的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 3 的核苷酸序列或者与 SEQ ID NO: 3 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第三多核苷酸的核苷酸序列包含：SEQ ID NO: 4 的核苷酸序列或者与 SEQ ID NO: 4 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第四多核苷酸的核苷酸序列包含：SEQ ID NO: 5 的核苷酸序列或者与 SEQ ID NO: 5 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第五多核苷酸的核苷酸序列包含：SEQ ID NO: 6 的核苷酸序列或者与 SEQ ID NO: 6 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第六多核苷酸的核苷酸序列包含：SEQ ID NO: 7 的核苷酸序列或者与 SEQ ID NO: 7 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 24 的核苷酸序列或者与 SEQ ID NO: 24 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 25 的核苷酸序列或者与 SEQ ID NO: 25 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第三多核苷酸的核苷酸序列包含：SEQ ID NO: 26 的核苷酸序列或者与 SEQ ID NO: 26 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第四多核苷酸的核苷酸序列包含：SEQ ID NO: 27 的核苷酸序列或者与 SEQ ID NO: 27 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第五多核苷酸的核苷酸序列包含：SEQ ID NO: 28 的核苷酸序列或者与 SEQ ID NO: 28 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第六多核苷酸的核苷酸序列包含：SEQ ID NO: 29 的核苷酸序列或者与 SEQ ID NO: 29 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 14 的核苷酸序列或者与 SEQ ID NO: 14 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 15 的核苷酸序列或者与 SEQ ID NO: 15 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 36 的核苷酸序列或者与 SEQ ID NO: 36 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 37 的核苷酸序列或者与 SEQ ID NO: 37 的核苷酸序列具有至少 70%、75%、80%、85%、

90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 17 的核苷酸序列或者与 SEQ ID NO: 17 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
5 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 18 的核苷酸序列或者与 SEQ ID NO: 18 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 39 的核苷酸序列或者与 SEQ ID NO: 39 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
10 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 40 的核苷酸序列或者与 SEQ ID NO: 40 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 20 的核苷酸序列或者与 SEQ ID NO: 20 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
15 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 21 的核苷酸序列或者与 SEQ ID NO: 21 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 42 的核苷酸序列或者与 SEQ ID NO: 42 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
20 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 43 的核苷酸序列或者与 SEQ ID NO: 43 的核苷酸序列具有至少 70%、75%、80%、85%、
25 90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 14 的核苷酸序列或者与 SEQ ID NO: 14 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
30 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 16 的核苷酸序列或者与 SEQ ID NO: 16 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 36 的核苷酸序列或者与 SEQ ID NO: 36 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、
35 97%、98%、99%相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 38 的核苷酸序列或者与 SEQ ID NO: 38 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 17 的核苷酸序列或

者与 SEQ ID NO: 17 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 19 的核苷酸序列或者与 SEQ ID NO: 19 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

- 5 在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 39 的核苷酸序列或者与 SEQ ID NO: 39 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 41 的核苷酸序列或者与 SEQ ID NO: 41 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

- 10 在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 22 的核苷酸序列或者与 SEQ ID NO: 22 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 20 的核苷酸序列或者与 SEQ ID NO: 20 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

- 15 在一实施方案中，多肽或多肽组合的编码序列包含第一多核苷酸和第二多核苷酸的核苷酸序列；其中，第一多核苷酸的核苷酸序列包含：SEQ ID NO: 44 的核苷酸序列或者与 SEQ ID NO: 44 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；第二多核苷酸的核苷酸序列包含：SEQ ID NO: 42 的核苷酸序列或者与 SEQ ID NO: 42 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

- 20 在一实施方案中，多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽，其中第一多肽包含 SEQ ID NO: 89 的氨基酸序列，第二多肽包含 SEQ ID NO: 90 的氨基酸序列，第三多肽包含 SEQ ID NO: 91 的氨基酸序列，第四多肽包含 SEQ ID NO: 92 的氨基酸序列，第五多肽包含 SEQ ID NO: 93 的氨基酸序列，第六多肽包含 SEQ ID NO: 94 的氨基酸序列；本文所述的编码序列包含编码第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 1 的核苷酸序列或者与 SEQ ID NO: 1 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 2 的核苷酸序列或者与 SEQ ID NO: 2 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第三多肽的核苷酸序列包含：SEQ ID NO: 3 的核苷酸序列或者与 SEQ ID NO: 3 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第四多肽的核苷酸序列包含：SEQ ID NO: 4 的核苷酸序列或者与 SEQ ID NO: 4 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第五多肽的核苷酸序列包含：SEQ ID NO: 5 的核苷酸序列或者与 SEQ ID NO: 5 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第六

多肽的核苷酸序列包含：SEQ ID NO: 6 的核苷酸序列或者与 SEQ ID NO: 6 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、
5 第五多肽和第六多肽，其中第一多肽包含 SEQ ID NO: 89 的氨基酸序列，第二多肽包含 SEQ ID NO: 90 的氨基酸序列，第三多肽包含 SEQ ID NO: 91 的氨基酸序列，第四多肽包含 SEQ ID NO: 92 的氨基酸序列，第五多肽包含 SEQ ID NO: 93 的氨基酸序列，第六多肽包含 SEQ ID NO: 94 的氨基酸序列；本文所述的编码序列包含编码第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽的核苷酸序列；其中，编码第一多肽
10 的核苷酸序列包含：SEQ ID NO: 23 的核苷酸序列或者与 SEQ ID NO: 23 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 24 的核苷酸序列或者与 SEQ ID NO: 24 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第三多肽的核苷酸序列包含：SEQ ID NO: 25 的核苷酸
15 序列或者与 SEQ ID NO: 25 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第四多肽的核苷酸序列包含：SEQ ID NO: 26 的核苷酸序列或者与 SEQ ID NO: 26 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第五多肽的核苷酸序列包含：SEQ ID NO: 27 的核苷酸序列或者与 SEQ ID NO: 27 的核苷酸序列具有至少
20 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第六多肽的核苷酸序列包含：SEQ ID NO: 28 的核苷酸序列或者与 SEQ ID NO: 28 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、
25 第五多肽和第六多肽，其中第一多肽包含 SEQ ID NO: 90 的氨基酸序列，第二多肽包含 SEQ ID NO: 91 的氨基酸序列，第三多肽包含 SEQ ID NO: 92 的氨基酸序列，第四多肽包含 SEQ ID NO: 93 的氨基酸序列，第五多肽包含 SEQ ID NO: 94 的氨基酸序列，第六多肽分包含 SEQ ID NO: 95 的氨基酸序列；本文所述的编码序列包含编码第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽的核苷酸序列；其中，编码第一多肽
30 的核苷酸序列包含：SEQ ID NO: 2 的核苷酸序列或者与 SEQ ID NO: 2 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 3 的核苷酸序列或者与 SEQ ID NO: 3 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第三多肽的核苷酸序列包含：SEQ ID NO: 4 的核苷酸序列或者
35 与 SEQ ID NO: 4 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第四多肽的核苷酸序列包含：SEQ ID NO: 5 的核苷酸序列或者与 SEQ ID NO: 5 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第五多肽的核苷酸序列包含：SEQ ID NO: 6 的核苷酸序列或者与 SEQ ID NO: 6 的核苷酸序列具有至少 70%、

75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第六多肽的核苷酸序列包含：SEQ ID NO: 7 的核苷酸序列或者与 SEQ ID NO: 7 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

- 5 在一实施方案中，多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽，其中第一多肽包含 SEQ ID NO: 90 的氨基酸序列，第二多肽包含 SEQ ID NO: 91 的氨基酸序列，第三多肽包含 SEQ ID NO: 92 的氨基酸序列，第四多肽包含 SEQ ID NO: 93 的氨基酸序列，第五多肽包含 SEQ ID NO: 94 的氨基酸序列，第六多肽分包含 SEQ ID NO: 95 的氨基酸序列；本文所述的编码序列包含编码第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 24 的核苷酸序列或者与 SEQ ID NO: 24 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 25 的核苷酸序列或者与 SEQ ID NO: 25 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第三多肽的核苷酸序列包含：SEQ ID NO: 26 的核苷酸序列或者与 SEQ ID NO: 26 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第四多肽的核苷酸序列包含：SEQ ID NO: 27 的核苷酸序列或者与 SEQ ID NO: 27 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第五多肽的核苷酸序列包含：SEQ ID NO: 28 的核苷酸序列或者与 SEQ ID NO: 28 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；编码第六多肽的核苷酸序列包含：SEQ ID NO: 29 的核苷酸序列或者与 SEQ ID NO: 29 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

- 25 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 102 的氨基酸序列，第二多肽包含 SEQ ID NO: 103 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 14 的核苷酸序列或者与 SEQ ID NO: 14 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；
30 编码第二多肽的核苷酸序列包含：SEQ ID NO: 15 的核苷酸序列或者与 SEQ ID NO: 15 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

- 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 102 的氨基酸序列，第二多肽包含 SEQ ID NO: 103 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 36 的核苷酸序列或者与 SEQ ID NO: 36 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列；
35 编码第二多肽的核苷酸序列包含：SEQ ID NO: 37 的核苷酸序列或者与 SEQ ID NO: 37 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相

同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 105 的氨基酸序列，第二多肽包含 SEQ ID NO: 106 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 17 的核苷酸序列或者与 SEQ ID NO: 17 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 18 的核苷酸序列或者与 SEQ ID NO: 18 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

10 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 105 的氨基酸序列，第二多肽包含 SEQ ID NO: 106 的氨基酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 39 的核苷酸序列或者与 SEQ ID NO: 39 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 40 的核苷酸序列或者与 SEQ ID NO: 40 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

20 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 108 的氨基酸序列，第二多肽包含 SEQ ID NO: 109 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 20 的核苷酸序列或者与 SEQ ID NO: 20 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 21 的核苷酸序列或者与 SEQ ID NO: 21 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

25 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 108 的氨基酸序列，第二多肽包含 SEQ ID NO: 109 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 42 的核苷酸序列或者与 SEQ ID NO: 42 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 43 的核苷酸序列或者与 SEQ ID NO: 43 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

35 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 102 的氨基酸序列，第二多肽包含 SEQ ID NO: 104 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 14 的核苷酸序列或者与 SEQ ID NO: 14 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 16 的核苷酸序列或者与 SEQ ID NO: 16 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相

同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 102 的氨基酸序列，第二多肽包含 SEQ ID NO: 104 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 36 的核苷酸序列或者与 SEQ ID NO: 36 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 38 的核苷酸序列或者与 SEQ ID NO: 38 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

10 在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 105 的氨基酸序列，第二多肽包含 SEQ ID NO: 107 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 17 的核苷酸序列或者与 SEQ ID NO: 17 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；
15 编码第二多肽的核苷酸序列包含：SEQ ID NO: 19 的核苷酸序列或者与 SEQ ID NO: 19 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 105 的氨基酸序列，第二多肽包含 SEQ ID NO: 107 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 39 的核苷酸序列或者与 SEQ ID NO: 39 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 41 的核苷酸序列或者与 SEQ ID NO: 41 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。
25 同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 110 的氨基酸序列，第二多肽包含 SEQ ID NO: 108 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 22 的核苷酸序列或者与 SEQ ID NO: 22 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 20 的核苷酸序列或者与 SEQ ID NO: 20 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列。

在一实施方案中，多肽或多肽组合包含第一多肽和第二多肽，其中第一多肽包含 SEQ ID NO: 110 的氨基酸序列，第二多肽包含 SEQ ID NO: 108 的氨基酸序列；本文所述的编码序列包含编码第一多肽和第二多肽的核苷酸序列；其中，编码第一多肽的核苷酸序列包含：SEQ ID NO: 44 的核苷酸序列或者与 SEQ ID NO: 44 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% 相同性的核苷酸序列；编码第二多肽的核苷酸序列包含：SEQ ID NO: 42 的核苷酸序列或者与 SEQ ID NO: 42

的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%、99%相同性的核苷酸序列。

RNA

5 在一些实施方案中，本发明的核酸或核酸组合是 RNA。在一些实施方案中，本发明的多核苷酸是 RNA。如本文所用，“RNA”的定义涵盖单链、双链、线性和环状 RNA。本发明的 RNA 可以是非扩增 RNA 或自我扩增 RNA。本发明的 RNA 可以通过化学合成的、重组产生的和体外转录的 RNA。在一些实施方案中，本发明的 RNA 用于在宿主细胞中表达本发明的多肽。在一些实施方案中，本发明的 RNA 用于在宿主细胞中表
10 达本发明的多肽组合。

在一些实施方案中，本发明的 RNA 是单链 RNA。在一些实施方案中，本发明的 RNA 是体外转录的 RNA (IVT-RNA)。IVT-RNA 可以通过 RNA 聚合酶利用 DNA 模板进行体外转录获得(例如如本文所述)。

在一些实施方案中，本发明的 RNA 是信使 RNA (mRNA)。一般而言，mRNA 可以
15 包含 5'-UTR 序列、多肽(例如本发明的多肽或多肽组合)的编码序列、3'-UTR 序列和任选存在的 poly(A)序列。mRNA 可以例如通过体外转录或化学合成产生。在一实施方案中，本发明的 mRNA 通过 RNA 聚合酶(例如 T7 RNA 聚合酶)利用 DNA 模板进行体外转录获得。在一实施方案中，本发明的 mRNA 包含(1) 5'-UTR、(2) 编码序列、(3) 3'-UTR 和(4) 任选存在的 poly(A)序列。5'-UTR、编码序列、3'-UTR 和 poly(A)序列如本文
20 所述。在一实施方案中，本发明的 mRNA 是核苷修饰的 mRNA。在一实施方案中，本发明的 mRNA 包含任选存在的 5'帽。在一实施方案中，本发明的 mRNA 包含帽近端序列。

在一些实施方案中，本发明的 RNA 是环状 RNA (circRNA)。如本文所用，术语“circRNA”或“环状 RNA”是指通过共价键形成环状结构的 RNA。一般而言，circRNA 是
25 一种单链 RNA，其可以包含 UTR (5'-UTR 和/或 3'-UTR)和开放阅读框，但是不包含 5'帽和 poly(A)尾。circRNA 可以在 UTR 中包含促进帽-非依赖性翻译的元件，例如内部核糖体进入位点(IRES)以及任选存在的聚腺苷酸(polyA)束或其中散置有胞苷酸的聚腺苷酸(polyAC)。circRNA 可以利用酶促或化学方法产生。一般而言，酶促方法可以包括例如利用 T4 连接酶催化的线性前体 RNA 的末端连接和利用核酶在 IVT 过程中形成
30 circRNA。利用核酶制备 circRNA 的方法可以包括例如基于 I 型内含子的方法、基于 II 型内含子的方法和基于发夹核酶的方法(参见例如 Wesselhoeft R. A. et al., Nature Communications, (2018) 9: 2629; Petkovic, S. et al., Nucleic Acids Research, 2015, Vol. 43, No. 4, 2454-2465; Chen, XJ et al., Front. Bioeng. Biotechnol., 30 November 2021; 和 Lambowitz, A.M. et al., Cold Spring Harb Perspect Biol. 2011 Aug; 3(8): a003616)。

35 在一些实施方案中，本发明的 RNA 是自我扩增 RNA (saRNA)。如本文所用，术语“自我复制 RNA”、“自我扩增 RNA (saRNA)”和“复制子”可以互换地使用来表示可以自我扩增和翻译的 RNA。一般而言，saRNA 可以源自或包含病毒(例如正义单链 RNA 病毒，如甲病毒，如委内瑞拉马脑炎病毒(VEEV))基因组的自我复制亚基因组，其包括编码形成 RNA 依赖性 RNA 聚合酶(RdRP)的非结构蛋白 nsP1、nsP2、nsP3 和 nsP4 的病毒基

因，而编码结构蛋白的病毒基因被编码感兴趣多肽的基因代替。saRNA 可以包含由亚基因组启动子隔开的两个开放阅读框(ORF)，即在编码 nsP1-nsP4 的第一 ORF 的 3'端跟随亚基因组启动子驱动下的编码感兴趣多肽(例如本文所述的多肽或多肽组合)的第二 ORF。saRNA 还可以在 5'端包含 5'帽和 5'-UTR，并且在 3'端包含 3'-UTR 和 poly(A)序列。此外，saRNA 的非翻译区还可以包含供 RdRP 识别的自我扩增所需的 5'和 3'保守序列元件(CSE)。与 mRNA 类似，saRNA 也可以通过体外转录或化学合成产生。关于 saRNA 的详细描述可以参见例如 WO2016135675A1 和 WO2021183564A1。

在一些实施方案中，本发明的 RNA (例如 mRNA、circRNA 和 saRNA)包含如本文所述多肽的编码序列。在一些实施方案中，本发明的 RNA (例如 mRNA、circRNA 和 saRNA)包含如本文所述多肽抗原的编码序列。在一些实施方案中，本发明的 RNA (例如 mRNA、circRNA 和 saRNA)包含如本文所述多肽组合的编码序列。

在一些实施方案中，本发明的 RNA 还包含有助于提高 RNA 的稳定性和/或翻译效率的结构元件，包括但不限于 5'帽、帽近端序列、5'-UTR、3'-UTR 和 poly(A)序列。

如本文所用，术语“非翻译区(UTR)”一般指 RNA 中(如 mRNA)中不翻译为氨基酸序列的区域(非编码区)，或者 DNA 中的相应区域。通常，位于开放阅读框(起始密码子)的 5'端(上游)的 UTR 可以称为 5'非翻译区 5'-UTR；位于开放阅读框(终止密码子)的 3'端(下游)的 UTR 可以称为 3'-UTR。在 5'帽存在的情况下，5'-UTR 位于 5'帽的下游，例如，与 5'帽直接相邻。在特定实施方案中，可以在 5'-UTR 中，例如在临近起始密码子的位置，包含优化的“Kozak 序列”以提高翻译效率。在 poly(A)序列存在的情况下，3'-UTR 位于 poly(A)序列的上游，例如与 poly(A)序列直接相邻。

在一些实施方案中，本发明的 RNA 包含 5'-UTR。在一优选实施方案中，5'-UTR 包含 SEQ ID NO: 139 或 148 的核苷酸序列。在一些实施方案中，本发明的 RNA 包含 3'-UTR。在一优选实施方案中，3'-UTR 包含 SEQ ID NO:140、141 或 149 的核苷酸序列。在一些实施方案中，本发明的 RNA 包含 5'-UTR 和 3'-UTR。在一实施方案中，5'-UTR 包含 SEQ ID NO: 139 或 148 的核苷酸序列，3'-UTR 包含 SEQ ID NO: 140、141 或 149 的核苷酸序列。在一具体实施方案中，5'-UTR 包含 SEQ ID NO: 139 的核苷酸序列，3'-UTR 包含 SEQ ID NO: 140 或 141 的核苷酸序列。在一具体实施方案中，5'-UTR 包含 SEQ ID NO: 148 的核苷酸序列，3'-UTR 包含 SEQ ID NO: 149 的核苷酸序列。

如本文所用，术语“poly(A)序列”或“poly(A)尾”是指包含连续或不连续腺苷酸的核苷酸序列。poly(A)序列通常位于 RNA 的 3'端，例如 3'-UTR 的 3'端(下游)。在一些实施方案中，poly(A)序列在其 3'端不包含腺苷酸以外的核苷酸。Poly(A)序列可以在制备 IVT-RNA 期间，由 DNA 依赖性 RNA 聚合酶根据 DNA 模板的编码序列转录产生，或者通过不依赖于 DNA 的 RNA 聚合酶(poly(A)聚合酶)连接至 IVT-RNA 的游离 3'端，例如 3'-UTR 的 3'端。

在一些实施方案中，本发明的 RNA 包含 poly(A)序列。在一实施方案中，poly(A)序列包含连续的腺苷酸。在一实施方案中，poly(A)序列可以包含至少 20、30、40、50、60、70、75、80、85、95 或 100 以及多达 120、150、180、200、300 个腺苷酸。在一实施方案中，poly(A)序列中的连续腺苷酸序列被包含 U、C 或 G 核苷酸的序列中断。在一实施方案中，poly(A)序列包含至少 50 个核苷酸。在一实施方案中，poly(A)序列包含至少 80

个核苷酸。在一实施方案中，poly(A)序列包含至少 100 个核苷酸。在一些实施方案中，poly(A)序列包含约 70、80、90、100、120 或 150 个核苷酸。在一优选实施方案中，poly(A)序列包含 121 个核苷酸。在一优选实施方案中，poly(A)序列包含 SEQ ID NO: 145 的核苷酸序列。

5 如本文所用，术语“5'帽”一般涉及通过 5'至 5'三磷酸键连接至 mRNA 的 5'端的 N7-甲基鸟苷结构(又称为“m⁷G 帽”、“m⁷Gppp-”)。5'帽可以在体外转录中共转录加至 RNA 中(例如使用抗反向帽类似物“ARCA”)，或者可以利用加帽酶在转录后连接至 RNA。在一些实施方案中，使用加帽酶(例如牛痘病毒加帽酶)将 RNA (例如 IVT-RNA)修饰为“Cap0 RNA”。在一些实施方案中，Cap0 RNA 中紧邻 m⁷G 帽的核苷酸的核糖 2'-O 位置
10 上发生额外的甲基化(例如通过 2'-O-甲基转移酶)，产生“Cap1 RNA”。在一些实施方案中，使用帽类似物来产生 5'帽修饰的 RNA。关于“帽类似物”的描述可以参见例如 Contreas, R. et al. (1982). Nucl. Acids Res. 10, 6353-6363 和 US7074596B2。帽类似物的实例包括但不限于 N7-甲基鸟苷-5'-三磷酸-5'鸟苷(m⁷G(5')ppp(5')G)、N7-甲基鸟苷-5'-三磷酸-5'-腺苷(m⁷G(5')ppp(5')A)和 3'-O-Me-m⁷G(5')ppp(5')G (ARCA)。

15 本发明的 RNA (例如 mRNA 和 saRNA)可以是包含 Cap0 (m⁷G 的相邻核苷酸的核糖的不发生甲基化)、Cap1 (m⁷G 的相邻核苷酸的核糖的甲基化)或 Cap2 (m⁷G 下游的第二核苷酸的核糖的甲基化)结构的 RNA。在一优选实施方案中，本发明的 mRNA 是 Cap1 mRNA。在一优选实施方案中，本发明的 mRNA 包含 ⁷MeGpppG_{2'}OMe。

 在一些实施方案中，本发明的 RNA 还包含帽近端序列。在一些实施方案中，5' UTR
20 包含帽近端序列。在一些实施方案中，5' UTR 不包含帽近端序列。在一些实施方案中，帽近端序列包含与 5'帽相邻的序列。在一实施方案中，帽近端序列包含 SEQ ID NO: 147 的核苷酸序列。在一实施方案中，帽近端序列中的一个或多个残基可以因为已包括在帽实体(例如，Cap1 结构等)中而包括在 RNA 中；或者，在一实施方案中，帽近端序列中的至少一些残基可以酶促添加(例如，通过聚合酶如 T7 聚合酶)。

25 在一些实施方案中，本发明的核酸或核酸组合包含编码本文所述的多肽或多肽组合的多核苷酸。

 在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列或者与 SEQ ID NO: 1、2、3、4、5、6、7、
30 13、14、15、16、17、18、19、20、21 和 22 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 之一的核苷酸序列或者与 SEQ ID NO: 45、46、47、48、49、50、51、
35 57、58、59、60、61、62、63、64、65 和 66 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。

 在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列或者两个或更多个与选自 SEQ

ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。在一些实施方案中，多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，所述多核苷酸包含两个或更多个选自 SEQ ID NO: 45、46、47、48、
5 49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第
10 四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 1 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 2 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 3 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 4 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 5 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 6 的核苷酸序列的核苷酸序列。在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三
15 多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 45 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 46 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 47 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 48 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 49 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 50 的核苷酸序列的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第
20 四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 2 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 3 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 4 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 5 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 6 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 7 的核苷酸序列。
25 在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 46 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 47 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 48 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 49 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 50 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 51 的核苷酸序列的核苷酸序列。
30

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 14 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 15 的核苷酸序列。
在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 58 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 59 的核苷酸序列。

35 在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 17 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 18 的核苷酸序列。
在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 61 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 62 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷

酸包含 SEQ ID NO: 20 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:21 的核苷酸序列。
在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 64 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 65 的核苷酸序列。

在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 14 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:16 的核苷酸序列。
在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 58 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 60 的核苷酸序列。

在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 17 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:19 的核苷酸序列。
在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 61 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 63 的核苷酸序列。

在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 22 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:20 的核苷酸序列。
在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 66 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 64 的核苷酸序列。

修饰的核苷酸

在一些实施方案中,本发明的 RNA (例如 mRNA)中的核苷酸可以是天然存在的核苷酸(例如天然存在的核糖核苷酸)和修饰的核苷酸。修饰的核苷酸可以是例如不存在于天然存在的 RNA 中的核苷酸,例如非标准核苷酸或脱氧核苷酸。核苷酸的修饰可以发生在核苷上,例如在核糖部分和/或核碱基部分上。修饰的核苷酸可以在转录(例如体外转录)过程中掺入,也可以在 RNA 化学合成的过程中添加。

在一些实施方案中, RNA 通过包含一个或多个修饰的核苷进行修饰。在一些实施方案中,本发明的 RNA (例如 mRNA)通过包含一个或多个修饰的核碱基进行修饰。在一些实施方案中,修饰的核碱基包括修饰的胞嘧啶、修饰的尿嘧啶、修饰的腺嘌呤、修饰的鸟嘌呤或其组合。

具有经修饰的尿嘧啶的示例性核碱基和核苷可以包括但不限于:假尿苷(ψ)、吡啶-4-酮核糖核苷、5-氮杂-尿嘧啶、6-氮杂-尿嘧啶、2-硫代-5-氮杂-尿嘧啶、2-硫代-尿嘧啶(s2U)、4-硫代尿嘧啶(s4U)、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基-尿嘧啶(ho5U)、5-氨基烯丙基-尿嘧啶、5-卤基-尿嘧啶(例如,5-碘-尿嘧啶或5-溴-尿嘧啶)、3-甲基尿嘧啶(m3U)、5-甲氧基-尿嘧啶(mo5U)、尿嘧啶 5-氧基乙酸(cmo5U)、尿嘧啶 5-氧基乙酸甲酯(mcmo5U)、5-羧甲基-尿嘧啶(cm5U)、1-羧甲基-假尿苷、5-羧基羟甲基-尿嘧啶(chm5U)、5-羧基羟甲基-尿嘧啶甲酯(mchm5U)、5-甲氧基羰基甲基-尿嘧啶(mcm5U)、5-甲氧基羰基甲基-2-硫代-尿嘧啶(mcm5s2U)、5-氨基甲基-2-硫代-尿嘧啶(nm5s2U)、5-甲基氨基甲基-尿嘧啶(mnm5U)、5-甲基氨基甲基-2-硫代-尿嘧啶(mnm5s2U)、5-甲基氨基甲基-2-硒代-尿嘧啶(mnm5se2U)、5-氨甲酰基甲基-尿嘧啶(ncm5U)、5-羧甲基氨基甲基-尿嘧啶(cmm5U)、5-羧甲基氨基甲基-2-硫代-尿嘧啶(cmm5s2U)、5-丙炔基-尿嘧啶、1-丙炔基-假尿嘧啶、5-牛磺酸甲基-尿嘧啶(τ m5U)、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代-尿嘧啶(τ m5s2U)、1-牛磺酸甲基-4-硫代-假尿苷、5-甲基-尿嘧啶(m5U,即具有核碱

基脱氧胸腺嘧啶)、1-甲基-假尿苷(m1ψ)、1-乙基-假尿苷(Et1ψ)、5-甲基-2-硫代-尿嘧啶(m5s2U)、1-甲基-4-硫代-假尿苷(m1s4ψ)、4-硫代-1-甲基-假尿苷、3-甲基-假尿苷(m3ψ)、2-硫代-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-1-脱氮-假尿苷、二氢尿嘧啶(D)、二氢假尿苷、5,6-二氢尿嘧啶、5-甲基-二氢尿嘧啶(m5D)、2-硫代-二氢尿嘧啶、2-硫代-二氢假尿苷、2-甲氧基-尿嘧啶、2-甲氧基-4-硫代-尿嘧啶、4-甲氧基-假尿苷、4-甲氧基-2-硫代-假尿苷、N1-甲基-假尿苷、3-(3-氨基-3-羧基丙基)尿嘧啶(acp3U)、1-甲基-3-(3-氨基-3-羧基丙基)假尿苷(acp3ψ)、5-(异戊烯基氨基甲基)尿嘧啶(m5U)、5-(异戊烯基氨基甲基)-2-硫代-尿嘧啶(m5s2U)、5,2'-O-二甲基-尿苷(m5Um)、2-硫代-2'-O-甲基-尿苷(s2Um)、5-甲氧基羰基甲基-2'-O-甲基-尿苷(mcm5Um)、5-氨甲酰基甲基-2'-O-甲基-尿苷(ncm5Um)、5-羧甲基氨基甲基-2'-O-甲基-尿苷(cmm5Um)、3,2'-O-二甲基-尿苷(m3Um)和5-(异戊烯基氨基甲基)-2'-O-甲基-尿苷(inm5Um)、1-硫代-尿嘧啶、脱氧胸苷、5-(2-甲氧羰基乙烯基)-尿嘧啶、5-(氨甲酰基羧基甲基)-尿嘧啶、5-氨甲酰基甲基-2-硫代-尿嘧啶、5-羧甲基-2-硫代-尿嘧啶、5-氰基甲基-尿嘧啶、5-甲氧基-2-硫代-尿嘧啶和5-[3-(1-E-丙烯基氨基)]尿嘧啶。

具有经修饰的胞嘧啶的示例性核碱基和核苷可以包括但不限于:5-氮杂胞嘧啶、6-氮杂胞嘧啶、假异胞苷、3-甲基胞嘧啶(m3C)、N4-乙酰基胞嘧啶(ac4C)、5-甲酰基胞嘧啶(f5C)、N4-甲基-胞嘧啶(m4C)、5-甲基-胞嘧啶(m5C)、5-卤基-胞嘧啶(例如5-碘-胞嘧啶)、5-羟甲基-胞嘧啶(hm5C)、1-甲基-假异胞苷、吡咯并胞嘧啶、吡咯并假异胞苷、2-硫代胞嘧啶(s2C)、2-硫代-5-甲基胞嘧啶、4-硫代假异胞苷、4-硫代-1-甲基-假异胞苷、4-硫代-1-甲基-1-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、泽布拉林(zebrarine)、5-氮杂-泽布拉林、5-甲基-泽布拉林、5-氮杂-2-硫代-泽布拉林、2-硫代-泽布拉林、2-甲氧基-胞嘧啶、2-甲氧基-5-甲基-胞嘧啶、4-甲氧基-假异胞苷、4-甲氧基-1-甲基-假异胞苷、立西啉(lysidine)(k2C)、5,2'-O-二甲基-胞苷(m5Cm)、N4-乙酰基-2'-O-甲基-胞苷(ac4Cm)、N4,2'-O-二甲基-胞苷(m4Cm)、5-甲酰基-2'-O-甲基-胞苷(f5Cm)、N4,N4,2'-O-三甲基-胞苷(m42Cm)、1-硫代-胞嘧啶、5-羟基-胞嘧啶、5-(3-叠氮基丙基)-胞嘧啶和5-(2-叠氮基乙基)-胞嘧啶。

具有经修饰的腺嘌呤的示例性核碱基和核苷可以包括但不限于:2-氨基-嘌呤、2,6-二氨基嘌呤、2-氨基-6-卤基-嘌呤(例如2-氨基-6-氯-嘌呤)、6-卤基-嘌呤(例如6-氯-嘌呤)、2-氨基-6-甲基-嘌呤、8-叠氮基-腺嘌呤、7-脱氮-腺嘌呤、7-脱氮-8-氮杂-腺嘌呤、7-脱氮-2-氨基-嘌呤、7-脱氮-8-氮杂-2-氨基-嘌呤、7-脱氮-2,6-二氨基嘌呤、7-脱氮-8-氮杂-2,6-二氨基嘌呤、1-甲基-腺嘌呤(m1A)、2-甲基-腺嘌呤(m2A)、N6-甲基-腺嘌呤(m6A)、2-甲硫基-N6-甲基-腺嘌呤(ms2m6A)、N6-异戊烯基-腺嘌呤(i6A)、2-甲硫基-N6-异戊烯基-腺嘌呤(ms2i6A)、N6-(顺-羟基异戊烯基)腺嘌呤(io6A)、2-甲硫基-N6-(顺-羟基异戊烯基)腺嘌呤(ms2io6A)、N6-甘氨酸基氨基甲酰基-腺嘌呤(g6A)、N6-苏氨酸基氨基甲酰基-腺嘌呤(t6A)、N6-甲基-N6-苏氨酸基氨基甲酰基-腺嘌呤(m6t6A)、2-甲硫基-N6-苏氨酸基氨基甲酰基-腺嘌呤(ms2g6A)、N6,N6-二甲基-腺嘌呤(m62A)、N6-羟基正缬氨酸基氨基甲酰基-腺嘌呤(hn6A)、2-甲硫基-N6-羟基正缬氨酸基氨基甲酰基-腺嘌呤(ms2hn6A)、N6-乙酰基-腺嘌呤(ac6A)、7-甲基-腺嘌呤、2-甲硫基-腺嘌呤、2-甲氧基-腺嘌呤、N6,2'-O-二甲基-腺苷(m6Am)、N6,N6,2'-O-三甲基-腺苷(m62Am)、1,2'-O-二甲基-腺苷(m1Am)、2-氨基-N6-甲

基-嘌呤、1-硫代-腺嘌呤、8-叠氮基-腺嘌呤、N6-(19-氨基-五氧杂十九烷基)-腺嘌呤、2,8-二甲基-腺嘌呤、N6-甲酰基-腺嘌呤和 N6-羟甲基-腺嘌呤。

具有经修饰的鸟嘌呤的示例性核碱基和核苷可以包括但不限于：肌苷(I)、1-甲基-肌苷(m1I)、怀俄苷(wyosine)(imG)、甲基怀俄苷(mimG)、4-脱甲基-怀俄苷(imG-14)、异怀俄苷(imG2)、怀丁苷(wybutosine)(yW)、过氧怀丁苷(o2yW)、羟基怀丁苷(OHyW)、修饰不足(undermodified)的羟基怀丁苷(OHyW*)、7-脱氮-鸟嘌呤、癸苷(queuosine)(Q)、环氧癸苷(oQ)、半乳糖基-癸苷(galQ)、甘露糖基-癸苷(manQ)、7-氰基-7-脱氮-鸟嘌呤(preQO)、7-氨基甲基-7-脱氮-鸟嘌呤(preQ1)、古嘌呤(archaeosine)(G+)、7-脱氮-8-氮杂-鸟嘌呤、6-硫代-鸟嘌呤、6-硫代-7-脱氮-鸟嘌呤、6-硫代-7-脱氮-8-氮杂-鸟嘌呤、7-甲基-鸟嘌呤(m7G)、6-硫代-7-甲基-鸟嘌呤、7-甲基-肌苷、6-甲氧基-鸟嘌呤、1-甲基-鸟嘌呤(m1G)、N2-甲基-鸟嘌呤(m2G)、N2,N2-二甲基-鸟嘌呤(m22G)、N2,7-二甲基-鸟嘌呤(m2,7G)、N2,N2,7-二甲基-鸟嘌呤(m2,2,7G)、8-氧代-鸟嘌呤、7-甲基-8-氧代-鸟嘌呤、1-甲基-6-硫代-鸟嘌呤、N2-甲基-6-硫代-鸟嘌呤、N2,N2-二甲基-6-硫代-鸟嘌呤、N2-甲基-2'-O-甲基-鸟苷(m2Gm)、N2,N2-二甲基-2'-O-甲基-鸟苷(m22Gm)、1-甲基-2'-O-甲基-鸟苷(m1Gm)、N2,7-二甲基-2'-O-甲基-鸟苷(m2,7Gm)、2'-O-甲基-肌苷(Im)、1,2'-O-二甲基-肌苷(m1Im)、1-硫代-鸟嘌呤和 O-6-甲基-鸟嘌呤。

在一些实施方案中，本发明的 RNA 中修饰的核碱基的比例为 10%-100%，即本发明的 RNA 可以通过用修饰的核碱基代替其中 10%-100%的核碱基来修饰。

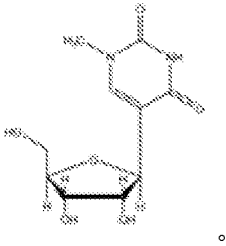
在一些实施方案中，本发明的 RNA (例如 mRNA)通过用修饰的尿嘧啶代替一个或多个尿嘧啶进行修饰。在一些实施方案中，修饰的尿嘧啶包括 1-甲基假尿嘧啶、假尿嘧啶、5-甲基-尿嘧啶或其组合。在一实施方案中，修饰的尿嘧啶包括假尿嘧啶。在一实施方案中，修饰的尿嘧啶包括 5-甲基-尿嘧啶。在一实施方案中，修饰的尿嘧啶包括 1-甲基-假尿嘧啶。

在一实施方案中，RNA 通过用修饰的尿嘧啶代替至少一个尿嘧啶进行修饰。在一实施方案中，RNA 通过用修饰的尿嘧啶代替所有尿嘧啶进行修饰。在一实施方案中，RNA 中修饰的尿嘧啶的比例为 10%-100%，例如 10%、20%、30%、40%、50%、60%、70%、80%、90%、95%或 100%。在一实施方案中，RNA 中修饰的尿嘧啶的比例为 20%-100%。在一实施方案中，RNA 中 20%-100%的尿嘧啶被 1-甲基假尿嘧啶代替。在优选的实施方案中，RNA 中 100%的尿嘧啶被 1-甲基假尿嘧啶代替。

在一些实施方案中，本发明的 RNA (例如 mRNA)通过用修饰的尿苷代替一个或多个尿苷进行修饰。在一些实施方案中，修饰的尿苷包括 1-甲基假尿苷、假尿苷、5-甲基-尿苷或其组合。在一实施方案中，修饰的尿苷包括假尿苷。在一实施方案中，修饰的尿苷包括 5-甲基-尿苷。在一实施方案中，修饰的尿苷包括 1-甲基-假尿苷。

在一实施方案中，RNA 通过用修饰的尿苷代替至少一个尿苷进行修饰。在一实施方案中，RNA 通过用修饰的尿苷代替所有尿苷进行修饰。在一实施方案中，RNA 中修饰的尿苷的比例为 10%-100%，例如 10%、20%、30%、40%、50%、60%、70%、80%、90%、95%或 100%。在一实施方案中，RNA 中修饰的尿苷的比例为 20%-100%。在一实施方案中，RNA 中 20%-100%的尿苷被 1-甲基假尿苷代替。在优选的实施方案中，RNA 中 100%的尿苷被 1-甲基假尿苷代替。

1-甲基-假尿苷具有如下结构：



在一具体实施方案中，本发明的 RNA 包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列，并且其中 100% 的尿嘧啶被 1-甲基假尿嘧啶代替。

在一具体实施方案中，本发明的 RNA 包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列，并且其中 100% 的尿苷被 1-甲基假尿苷代替。

在一具体实施方案中，本发明的 mRNA 包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列，并且其中 100% 的尿嘧啶被 1-甲基假尿嘧啶代替。

在一具体实施方案中，本发明的 mRNA 包含 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列，并且其中 100% 的尿苷被 1-甲基假尿苷代替。

DNA

在一些实施方案中，本发明的核酸或核酸组合是 DNA。在一些实施方案中，本发明的多核苷酸是 DNA。这样的 DNA 可以是例如用于在体外转录本发明的 RNA 的 DNA 模板或者用于在宿主细胞中表达多肽或者多肽抗原的 DNA。DNA 可以是双链、单链、线性和环状 DNA。

DNA 模板可以在合适的转录载体中提供。一般而言，DNA 模板可以是双链复合物，其包含与本文所述编码序列相同的核苷酸序列(编码链)和与本文所述编码序列互补的核苷酸序列(模板链)。如本领域技术人员已知的，DNA 模板可以包含启动子，5'-UTR、编码序列、3'-UTR 和任选存在的 poly(A)序列相应的 DNA 序列。启动子可以是本领域技术人员已知的合适 RNA 聚合酶(特别是 DNA 依赖性 RNA 聚合酶)可用的启动子，包括但不限于 SP6、T3 和 T7 RNA 聚合酶的启动子。在一些实施方案中，DNA 模板中的 5'-UTR、编码序列、3'-UTR 和 poly(A)序列为本文所述 RNA 中包含的相应序列或者与之互补。DNA 可以在质粒载体(例如环状质粒载体)中提供。

在一些实施方案中，本发明的 DNA 包含如本文所述多肽的编码序列。在一些实施方案中，本发明的 DNA 包含如本文所述多肽抗原的编码序列。在一些实施方案中，本发明的 DNA 包含如本文所述多肽组合的编码序列。

在一些实施方案中，本发明的 DNA 包含 5'-UTR 序列对应的 DNA 序列。在一优选实施方案中，5'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 142 或 150 的核苷酸序列。

在一些实施方案中，本发明的 DNA 包含 3'-UTR 序列对应的 DNA 序列。在一优选实施方案中，3'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 143、144 或 151 的核苷酸序列。在一些实施方案中，本发明的 DNA 包含 5'-UTR 和 3'-UTR 序列对应的 DNA 序列。在一实施方案中，5'-UTR 对应的 DNA 序列包含 SEQ ID NO: 142 或 150 的核苷酸序列，
5 3'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 143、144 或 151 的核苷酸序列。在一具体实施方案中，5'-UTR 对应的 DNA 序列包含 SEQ ID NO: 142 的核苷酸序列，3'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 143 或 144 的核苷酸序列。在一具体实施方案中，5'-UTR 对应的 DNA 序列包含 SEQ ID NO: 150 的核苷酸序列，3'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 151 的核苷酸序列。

10 在一些实施方案中，本发明的 DNA 包含 poly(A)序列对应的 DNA 序列。在一实施方案中，poly(A)序列对应的 DNA 序列包含连续的脱氧腺苷酸。在一实施方案中，poly(A)序列对应的 DNA 序列可以包含至少 20、30、40、50、60、70、75、80、85、95 或 100 以及多达 120、150、180、200、300 个脱氧腺苷酸。在一实施方案中，poly(A)序列对应的 DNA 序列包含至少 50 个核苷酸。在一实施方案中，poly(A)序列对应的 DNA 序列包
15 含至少 80 个核苷酸。在一实施方案中，poly(A)序列对应的 DNA 序列包含至少 100 个核苷酸。在一些实施方案中，poly(A)序列对应的 DNA 序列包含约 70、80、90、100、120 或 150 个核苷酸。在一优选实施方案中，poly(A)序列对应的 DNA 序列包含 121 个核苷酸。在一优选实施方案中，poly(A)序列对应的 DNA 序列包含 SEQ ID NO: 146 的核苷酸序列。

20 在一些实施方案中，多核苷酸为 DNA，多核苷酸包含 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列或者与 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99% 相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，多核苷酸包含 SEQ ID NO:
25 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 之一的核苷酸序列或者与 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 之一的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。

30 在一些实施方案中，多核苷酸为 DNA，多核苷酸包含两个或更多个选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 23、24、25、26、27、28、29、35、36、37、38、39、40、41、42、43 和 44 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。在一些实施方案中，多核苷酸为 DNA，多核苷酸包含两个或更多个选自 SEQ ID NO: 67、68、69、70、71、72、73、79、
35 80、81、82、83、84、85、86、87 和 88 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 67、68、69、70、71、72、73、79、80、81、82、83、84、85、86、87 和 88 的核苷酸序列具有至少 70%、75%、80%、85%、90%、95%、96%、97%、98%或 99%相同性的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第

四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 23 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 24 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 25 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 26 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 27 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 28 的核苷酸序列的核苷酸序列。在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 67 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 68 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 69 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 70 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 71 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 72 的核苷酸序列的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 24 的核苷酸序列，第二多核苷酸包含 SEQ ID NO:25 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 26 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 27 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 28 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 29 的核苷酸序列。在一实施方案中，多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 68 的核苷酸序列，第二多核苷酸包含 SEQ ID NO:69 的核苷酸序列，第三多核苷酸包含 SEQ ID NO: 70 的核苷酸序列，第四多核苷酸包含 SEQ ID NO: 71 的核苷酸序列，第五多核苷酸包含 SEQ ID NO: 72 的核苷酸序列，并且第六多核苷酸包含 SEQ ID NO: 73 的核苷酸序列的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 36 的核苷酸序列，第二多核苷酸包含 SEQ ID NO:37 的核苷酸序列。
在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包
25 含 SEQ ID NO: 80 的核苷酸序列，第二多核苷酸包含 SEQ ID NO:81 的核苷酸序列。

在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 39 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:40 的核苷酸序列。在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 83 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 84 的核苷酸序列。

30 在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 42 的核苷酸序列,第二多核苷酸包含 SEQ ID NO:43 的核苷酸序列。在一实施方案中,多核苷酸包含第一多核苷酸和第二多核苷酸;其中,第一多核苷酸包含 SEQ ID NO: 86 的核苷酸序列,第二多核苷酸包含 SEQ ID NO: 87 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 36 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 38 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 80 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 82 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 39 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 41 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 83 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 85 的核苷酸序列。

在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 44 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 42 的核苷酸序列。

5 在一实施方案中，多核苷酸包含第一多核苷酸和第二多核苷酸；其中，第一多核苷酸包含 SEQ ID NO: 88 的核苷酸序列，第二多核苷酸包含 SEQ ID NO: 86 的核苷酸序列。

载体和宿主细胞

在又一方面，本发明还提供了包含本发明的核酸或核酸组合的表达载体。表达载体
10 可以进一步包含额外的多核苷酸序列，例如调控序列和抗生素抗性基因。本发明的核酸或核酸组合可以存在于一种或多种表达载体中。

本发明还提供了一种宿主细胞，其包含本发明的核酸或核酸组合或表达载体。可以采用本领域已知的各种方法将本发明的核酸或核酸组合或者表达载体导入合适的宿主细胞中。这类方法包括但不限于脂质体转染、电穿孔、病毒转导和磷酸钙转染等。在优
15 选的实施方案中，宿主细胞用于表达本发明的多肽、多肽抗原或多肽组合。宿主细胞的实例包括但不限于原核细胞(例如细菌，例如大肠杆菌)和真核细胞(例如酵母、昆虫细胞、哺乳动物细胞)。适合于表达的哺乳动物宿主细胞包括但不限于人宫颈癌细胞(HeLa 细胞)、人类胚胎肾细胞(HEK 细胞，例如 HEK293 细胞)、中国仓鼠卵巢(CHO)细胞和其他适于表达的哺乳动物细胞。

20

组合物

在一方面，本文所述的核酸或核酸组合被配制用于体外和体内递送。在一方面，本发明提供一种组合物，其包含本发明的核酸或核酸组合。特别地，在一些实施方案中，核酸或核酸组合被配制成含脂质组合物。在一些实施方案中，组合物包含包封核酸或核
25 酸组合的脂质。在一些实施方案中，含脂质组合物形成将核酸或核酸组合封闭在脂质壳内的脂质纳米颗粒。在一些实施方案中，脂质壳保护核酸或核酸组合免于降解。在一些实施方案中，脂质纳米颗粒还促进将封入的核酸或核酸组合运输至细胞内隔室和/或机构中以发挥预期治疗功能。在某些实施方案中，当存在于脂质纳米颗粒中时，核酸在水溶液中抵抗核酸酶的降解。包含核酸的脂质纳米颗粒和其制备方法为此项技术中已知的，
30 诸如在例如美国专利公布第 2004/0142025 号、美国专利公布第 2007/0042031 号、PCT 公布第 WO 2017/004143 号、PCT 公布第 WO 2015/199952 号、PCT 公布第 WO 2013/016058 号、PCT 公布第 WO 2013/086373 号和 PCT 公布第 WO 2023/098842 号中公开的那些，这些公布中的每一者的完整公开内容出于所有目的以引用的方式整体并入本文中。

35 可与本公开结合使用的纳米颗粒组合物包括例如脂质纳米颗粒(LNP)、纳米脂蛋白颗粒、脂质体、脂质囊泡和脂质复合物(lipoplex)。在一些实施方案中，纳米颗粒组合物是包括一个或多个脂质双层的囊泡。在一些实施方案中，纳米颗粒组合物包括经水性隔室隔开的两个或更多个同心双层。脂质双层可经官能化和/或彼此交联。脂质双层可包括一种或多种配体、蛋白质或通道。

在一些实施方案中，组合物包含至少第一脂质。在一些实施方案中，组合物包含至少第一脂质，第一脂质为阳离子脂质。在一些实施方案中，组合物包含至少第一脂质(例如阳离子脂质)，第一脂质包含如本文所述的脂质系列 01、02、03 和 04 的化合物中的一种或多种，例如根据式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-III)(及其子式)的化合物中的一种或多种。在一些方案中，第一脂质包含表 01-1、02-1、03-1 和 04-1 所列的化合物中的一种或多种。在一些实施方案中，组合物还包含第二脂质，第二脂质为聚合物结合的脂质。在一些方案中，第二脂质包含根据式 05-I 的化合物中的一种或多种。在一些实施方案中，组合物被配制为将核酸或核酸组合包封在脂质中的脂质纳米颗粒。

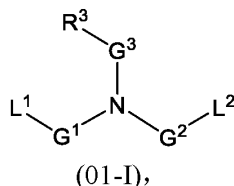
在一些实施方案中，纳米颗粒组合物包含脂质组分，所述脂质组分包含至少一种脂质，诸如如本文所述的脂质系列 01、02、03、04 和 05 的化合物中的一种或多种，例如根据式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)、(04-III)和(05-I)(及其子式)的化合物中的一种或多种。例如，在一些实施方案中，纳米颗粒组合物可包含脂质组分，所述脂质组分包括本文提供的化合物之一。纳米颗粒组合物还可包含一种或多种其它脂质或非脂质组分，示例性的其它脂质或非脂质组分可以参见例如 PCT 公布第 WO 2023/098842 号。

阳离子脂质

阳离子脂质包括以下脂质系列 01-04 (及其子式)。

1) 脂质系列 01

在一实施方案中，阳离子脂质包含式(01-I)化合物：



或其药学上可接受的盐、前药或立体异构体中的一种或多种，其中：

G^1 和 G^2 各自独立地是键、 $\text{C}_2\text{-C}_{12}$ 亚烷基或 $\text{C}_2\text{-C}_{12}$ 亚烯基，其中所述亚烷基或亚烯基中的一个或多个 $-\text{CH}_2-$ 任选地经 $-\text{O}-$ 替代；

L^1 是 $-\text{OC}(=\text{O})\text{R}^1$ 、 $-\text{C}(=\text{O})\text{OR}^1$ 、 $-\text{OC}(=\text{O})\text{OR}^1$ 、 $-\text{C}(=\text{O})\text{R}^1$ 、 $-\text{OR}^1$ 、 $-\text{S}(\text{O})_x\text{R}^1$ 、 $-\text{S}-\text{SR}^1$ 、 $-\text{C}(=\text{O})\text{SR}^1$ 、 $-\text{SC}(=\text{O})\text{R}^1$ 、 $-\text{NR}^a\text{C}(=\text{O})\text{R}^1$ 、 $-\text{C}(=\text{O})\text{NR}^b\text{R}^c$ 、 $-\text{NR}^a\text{C}(=\text{O})\text{NR}^b\text{R}^c$ 、 $-\text{OC}(=\text{O})\text{NR}^b\text{R}^c$ 、 $-\text{NR}^a\text{C}(=\text{O})\text{OR}^1$ 、 $-\text{SC}(=\text{S})\text{R}^1$ 、 $-\text{C}(=\text{S})\text{SR}^1$ 、 $-\text{C}(=\text{S})\text{R}^1$ 、 $-\text{CH}(\text{OH})\text{R}^1$ 、 $-\text{P}(=\text{O})(\text{OR}^b)(\text{OR}^c)$ 、 $-(\text{C}_6\text{-C}_{10}\text{ 亚芳基})-\text{R}^1$ 、 $-(6\text{ 至 }10\text{ 元亚杂芳基})-\text{R}^1$ 或 R^1 ；

L^2 是 $-\text{OC}(=\text{O})\text{R}^2$ 、 $-\text{C}(=\text{O})\text{OR}^2$ 、 $-\text{OC}(=\text{O})\text{OR}^2$ 、 $-\text{C}(=\text{O})\text{R}^2$ 、 $-\text{OR}^2$ 、 $-\text{S}(\text{O})_x\text{R}^2$ 、 $-\text{S}-\text{SR}^2$ 、 $-\text{C}(=\text{O})\text{SR}^2$ 、 $-\text{SC}(=\text{O})\text{R}^2$ 、 $-\text{NR}^d\text{C}(=\text{O})\text{R}^2$ 、 $-\text{C}(=\text{O})\text{NR}^e\text{R}^f$ 、 $-\text{NR}^d\text{C}(=\text{O})\text{NR}^e\text{R}^f$ 、 $-\text{OC}(=\text{O})\text{NR}^e\text{R}^f$ 、 $-\text{NR}^d\text{C}(=\text{O})\text{OR}^2$ 、 $-\text{SC}(=\text{S})\text{R}^2$ 、 $-\text{C}(=\text{S})\text{SR}^2$ 、 $-\text{C}(=\text{S})\text{R}^2$ 、 $-\text{CH}(\text{OH})\text{R}^2$ 、 $-\text{P}(=\text{O})(\text{OR}^e)(\text{OR}^f)$ 、 $-(\text{C}_6\text{-C}_{10}\text{ 亚芳基})-\text{R}^2$ 、 $-(6\text{ 至 }10\text{ 元亚杂芳基})-\text{R}^2$ 或 R^2 ；

R^1 和 R^2 各自独立地是 $\text{C}_6\text{-C}_{32}$ 烷基或 $\text{C}_6\text{-C}_{32}$ 烯基；

R^a 、 R^b 、 R^d 和 R^e 各自独立地是 H 、 $\text{C}_1\text{-C}_{24}$ 烷基或 $\text{C}_2\text{-C}_{24}$ 烯基；

R^c 和 R^f 各自独立地是 $\text{C}_1\text{-C}_{32}$ 烷基或 $\text{C}_2\text{-C}_{32}$ 烯基；

G^3 是 $\text{C}_2\text{-C}_{24}$ 亚烷基、 $\text{C}_2\text{-C}_{24}$ 亚烯基、 $\text{C}_3\text{-C}_8$ 亚环烷基或 $\text{C}_3\text{-C}_8$ 亚环烯基；

R^3 是 $-N(R^4)R^5$;

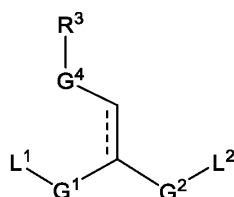
R^4 是 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、4 至 8 元杂环基或 C_6 - C_{10} 芳基; 或 R^4 、 G^3 或 G^3 的一部分与它们所连接的氮一起形成环状部分;

R^5 是 C_1 - C_{12} 烷基或 C_3 - C_8 环烷基; 或 R^4 、 R^5 与它们所连接的氮一起形成环状部分;

5 x 是 0、1 或 2; 并且

其中每个烷基、烯基、环烷基、环烯基、杂环基、芳基、亚烷基、亚烯基、亚环烷基、亚环烯基、亚芳基、亚杂芳基和环状部分独立地任选地经取代。

在一实施方案中, 阳离子脂质包含式(01-II)化合物:



10 (01-II),

或其药学上可接受的盐、前药或立体异构体中的一种或多种, 其中:

-----是单键或双键;

G^1 和 G^2 各自独立地是键、 C_2 - C_{12} 亚烷基或 C_2 - C_{12} 亚烯基, 其中所述亚烷基或亚烯基中的一个或多个 $-CH_2-$ 任选地经 $-O-$ 替代;

15 L^1 是 $-OC(=O)R^1$ 、 $-C(=O)OR^1$ 、 $-OC(=O)OR^1$ 、 $-C(=O)R^1$ 、 $-OR^1$ 、 $-S(O)_xR^1$ 、 $-S-SR^1$ 、 $-C(=O)SR^1$ 、 $-SC(=O)R^1$ 、 $-NR^aC(=O)R^1$ 、 $-C(=O)NR^bR^c$ 、 $-NR^aC(=O)NR^bR^c$ 、 $-OC(=O)NR^bR^c$ 、 $-NR^aC(=O)OR^1$ 、 $-SC(=S)R^1$ 、 $-C(=S)SR^1$ 、 $-C(=S)R^1$ 、 $-CH(OH)R^1$ 、 $-P(=O)(OR^b)(OR^c)$ 、 $-(C_6-C_{10}$ 亚芳基)- R^1 、 $-(6$ 至 10 元亚杂芳基)- R^1 或 R^1 ;

20 L^2 是 $-OC(=O)R^2$ 、 $-C(=O)OR^2$ 、 $-OC(=O)OR^2$ 、 $-C(=O)R^2$ 、 $-OR^2$ 、 $-S(O)_xR^2$ 、 $-S-SR^2$ 、 $-C(=O)SR^2$ 、 $-SC(=O)R^2$ 、 $-NR^dC(=O)R^2$ 、 $-C(=O)NR^eR^f$ 、 $-NR^dC(=O)NR^eR^f$ 、 $-OC(=O)NR^eR^f$ 、 $-NR^dC(=O)OR^2$ 、 $-SC(=S)R^2$ 、 $-C(=S)SR^2$ 、 $-C(=S)R^2$ 、 $-CH(OH)R^2$ 、 $-P(=O)(OR^e)(OR^f)$ 、 $-(C_6-C_{10}$ 亚芳基)- R^2 、 $-(6$ 至 10 元亚杂芳基)- R^2 或 R^2 ;

R^1 和 R^2 各自独立地是 C_6 - C_{32} 烷基或 C_6 - C_{32} 烯基;

R^a 、 R^b 、 R^d 和 R^e 各自独立地是 H 、 C_1 - C_{24} 烷基或 C_2 - C_{24} 烯基;

25 R^c 和 R^f 各自独立地是 C_1 - C_{32} 烷基或 C_2 - C_{32} 烯基;

G^4 是键、 C_1 - C_{23} 亚烷基、 C_2 - C_{23} 亚烯基、 C_3 - C_8 亚环烷基或 C_3 - C_8 亚环烯基;

R^3 是 $-N(R^4)R^5$;

R^4 是 C_1 - C_{12} 烷基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、4 至 8 元杂环基或 C_6 - C_{10} 芳基; 或 R^4 、 G^3 或 G^3 的一部分与它们所连接的氮一起形成环状部分;

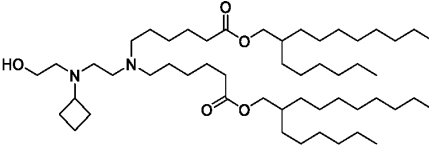
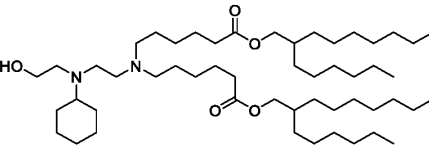
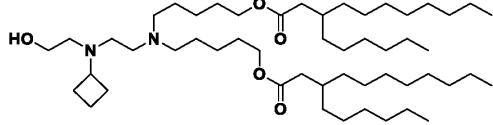
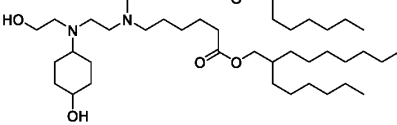
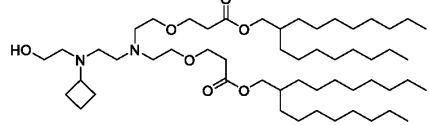
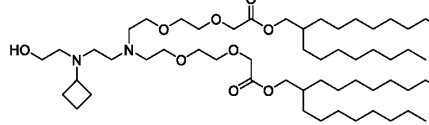
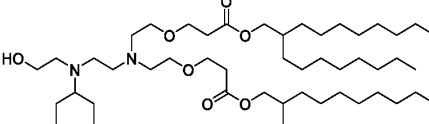
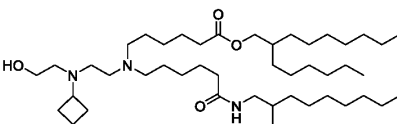
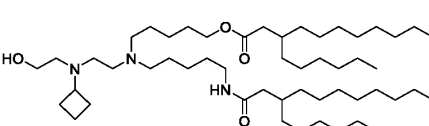
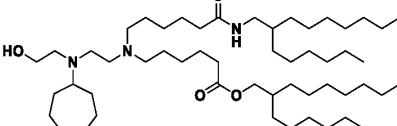
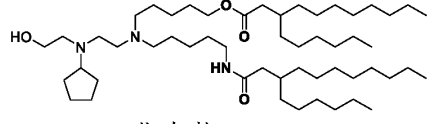
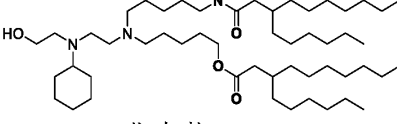
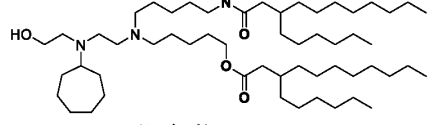
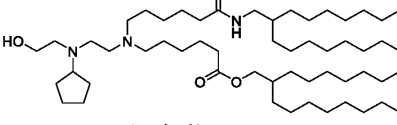
30 R^5 是 C_1 - C_{12} 烷基或 C_3 - C_8 环烷基; 或 R^4 、 R^5 与它们所连接的氮一起形成环状部分;

x 是 0、1 或 2; 并且

其中每个烷基、烯基、环烷基、环烯基、杂环基、芳基、亚烷基、亚烯基、亚环烷基、亚环烯基、亚芳基、亚杂芳基和环状部分独立地任选地经取代。

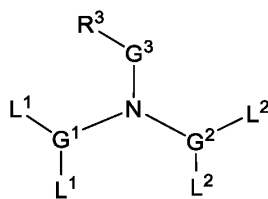
35 在一实施方案中, 阳离子脂质包含表 01-1 中的化合物, 或其药学上可接受的盐、前药或立体异构体中的一种或多种。

表 01-1.

 化合物 01-1	 化合物 01-2
 化合物 01-20	 化合物 01-21
 化合物 01-102	 化合物 01-103
 化合物 01-104	 化合物 01-106
 化合物 01-108	 化合物 01-128
 化合物 01-130	 化合物 01-131
 化合物 01-132	 化合物 01-136

2) 脂质系列 02

在一实施方案中，阳离子脂质包含式(02-I)化合物：



(02-I)

或其药学上可接受的盐、前药或立体异构体中的一种或多种，其中：

G^1 和 G^2 各自独立地是 C_2 - C_{12} 亚烷基或 C_2 - C_{12} 亚烯基，其中 G^1 和 G^2 中的一个或多个 - CH_2 - 任选地经 -O-、-C(=O)O- 或 -OC(=O)- 替代；

每个 L^1 独立地是 -OC(=O) R^1 、-C(=O)OR¹、-OC(=O)OR¹、-C(=O) R^1 、-OR¹、-S(O)_x R^1 、-S-SR¹、-C(=O)SR¹、-SC(=O) R^1 、-NR^aC(=O) R^1 、-C(=O)NR^b R^c 、-NR^aC(=O)NR^b R^c 、-OC(=O)NR^b R^c 、-NR^aC(=O)OR¹、-SC(=S) R^1 、-C(=S)SR¹、-C(=S) R^1 、-CH(OH) R^1 、-P(=O)(OR^b)(OR^c)、-NR^aP(=O)(OR^b)(OR^c)；

每个 L^2 独立地是 -OC(=O) R^2 、-C(=O)OR²、-OC(=O)OR²、-C(=O) R^2 、-OR²、-S(O)_x R^2 、-S-SR²、-C(=O)SR²、-SC(=O) R^2 、-NR^dC(=O) R^2 、-C(=O)NR^e R^f 、-NR^dC(=O)NR^e R^f 、-OC(=O)NR^e R^f 、-NR^dC(=O)OR²、-SC(=S) R^2 、-C(=S)SR²、-C(=S) R^2 、-CH(OH) R^2 、-P(=O)(OR^e)(OR^f)、-NR^dP(=O)(OR^e)(OR^f)；

R^1 和 R^2 各自独立地是 C_6 - C_{24} 烷基或 C_6 - C_{24} 烯基；

R^a 、 R^b 、 R^d 和 R^e 各自独立地是 H、 C_1 - C_{24} 烷基或 C_2 - C_{24} 烯基；

R^c 和 R^f 各自独立地是 C_1 - C_{24} 烷基或 C_2 - C_{24} 烯基；

G^3 是 C_2 - C_{12} 亚烷基或 C_2 - C_{12} 亚烯基，其中亚烷基或亚烯基的一部分或全部任选地经 C_3 - C_8 亚环烷基或 C_3 - C_8 亚环烯基替代；

R^3 是 -N(R^4) R^5 、-OR⁶ 或 -SR⁶；

R^4 是 C_1 - C_{12} 烷基、 C_2 - C_{12} 烯基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、 C_6 - C_{10} 芳基或 4 至 8 元杂环烷基；

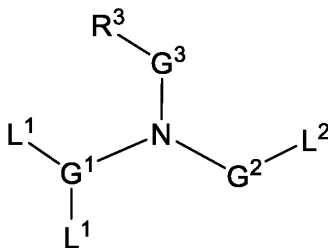
R^5 是 H、 C_1 - C_{12} 烷基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、 C_6 - C_{10} 芳基或 4 至 8 元杂环烷基；

R^6 是氢、 C_1 - C_{12} 烷基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基或 C_6 - C_{10} 芳基；

x 是 0、1 或 2；并且

其中每个烷基、烯基、环烷基、环烯基、杂环烷基、芳基、亚烷基、亚烯基、亚环烷基和亚环烯基独立地任选地经取代。

在一实施方案中，阳离子脂质为式(02-II)化合物：



(02-II),

或其药学上可接受的盐、前药或立体异构体，其中：

G^1 和 G^2 各自独立地为 C_2 - C_{12} 亚烷基或 C_2 - C_{12} 亚烯基，其中 G^1 和 G^2 中的一个或多个 $-CH_2-$ 任选地经 $-O-$ 、 $-C(=O)O-$ 或 $-OC(=O)-$ 置换；

各 L^1 独立地为 $-OC(=O)R^1$ 、 $-C(=O)OR^1$ 、 $-OC(=O)OR^1$ 、 $-C(=O)R^1$ 、 $-OR^1$ 、 $-S(O)_xR^1$ 、
5 $-S-SR^1$ 、 $-C(=O)SR^1$ 、 $-SC(=O)R^1$ 、 $-NR^aC(=O)R^1$ 、 $-C(=O)NR^bR^c$ 、 $-NR^aC(=O)NR^bR^c$ 、
 $-OC(=O)NR^bR^c$ 、 $-NR^aC(=O)OR^1$ 、 $-SC(=S)R^1$ 、 $-C(=S)SR^1$ 、 $-C(=S)R^1$ 、 $-CH(OH)R^1$ 、
 $P(=O)(OR^b)(OR^c)$ 、 $-NR^aP(=O)(OR^b)(OR^c)$ ；

各 L^2 独立地为 $-OC(=O)R^2$ 、 $-C(=O)OR^2$ 、 $-OC(=O)OR^2$ 、 $-C(=O)R^2$ 、 $-OR^2$ 、 $-S(O)_xR^2$ 、
10 $-S-SR^2$ 、 $-C(=O)SR^2$ 、 $-SC(=O)R^2$ 、 $-NR^dC(=O)R^2$ 、 $-C(=O)NR^eR^f$ 、 $-NR^dC(=O)NR^eR^f$ 、
 $-OC(=O)NR^eR^f$ 、 $-NR^dC(=O)OR^2$ 、 $-SC(=S)R^2$ 、 $-C(=S)SR^2$ 、 $-C(=S)R^2$ 、 $-CH(OH)R^2$ 、
 $P(=O)(OR^e)(OR^f)$ 、 $-NR^dP(=O)(OR^e)(OR^f)$ ；

R^1 和 R^2 各自独立地为 C_6 - C_{24} 烷基或 C_6 - C_{24} 烯基；

R^a 、 R^b 、 R^d 和 R^e 各自独立地为 H、 C_1 - C_{24} 烷基或 C_2 - C_{24} 烯基；

R^c 和 R^f 各自独立地为 C_1 - C_{24} 烷基或 C_2 - C_{24} 烯基；

15 G^3 为 C_2 - C_{12} 亚烷基或 C_2 - C_{12} 亚烯基，其中亚烷基或亚烯基的一部分或全部任选地
经 C_3 - C_8 亚环烷基或 C_3 - C_8 亚环烯基置换；

R^3 为 $-N(R^4)R^5$ 、 $-OR^6$ 或 $-SR^6$ ；

R^4 为 C_1 - C_{12} 烷基、 C_2 - C_{12} 烯基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、 C_6 - C_{10} 芳基或 4 元至
8 元杂环烷基；

20 R^5 为 H、 C_1 - C_{12} 烷基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基、 C_6 - C_{10} 芳基或 4 元至 8 元杂环
烷基；

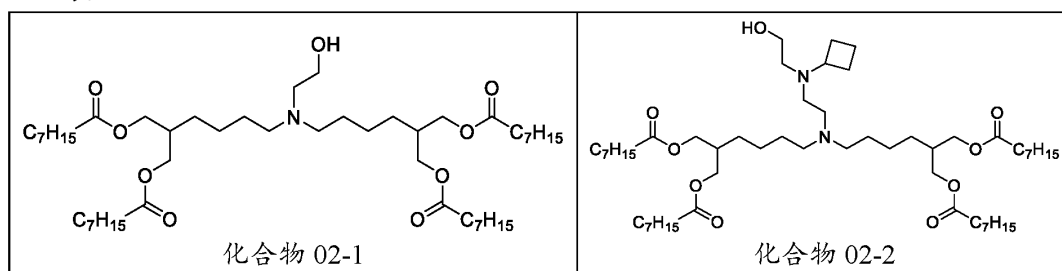
R^6 为氢、 C_1 - C_{12} 烷基、 C_3 - C_8 环烷基、 C_3 - C_8 环烯基或 C_6 - C_{10} 芳基；

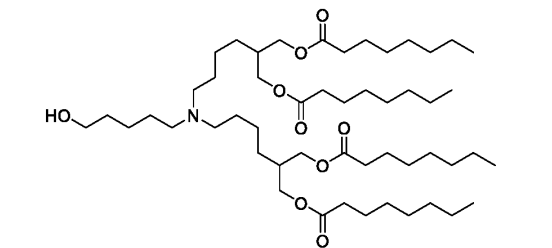
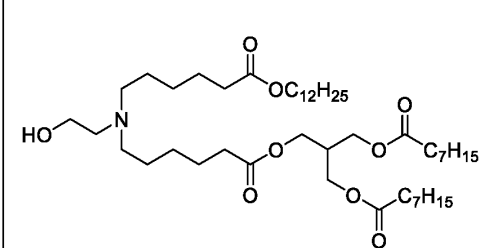
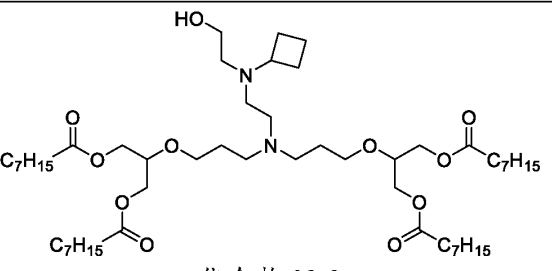
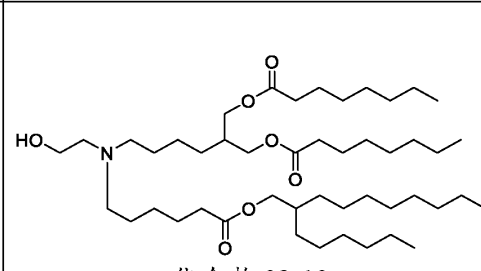
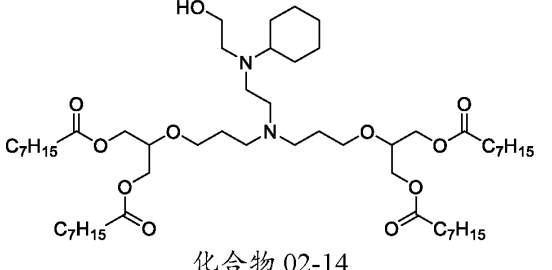
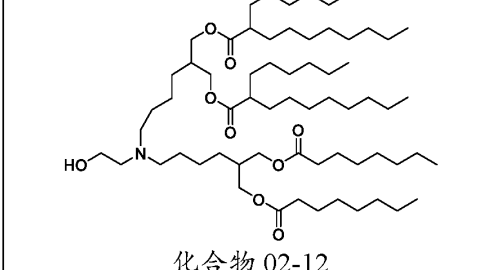
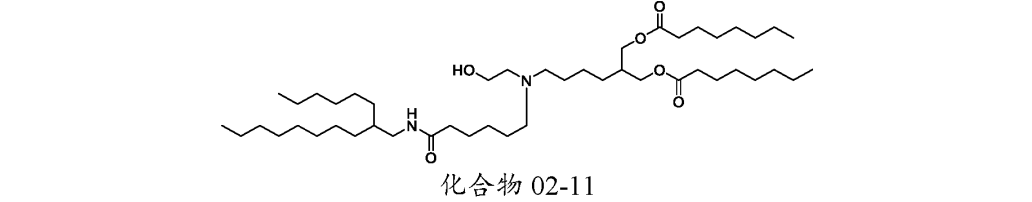
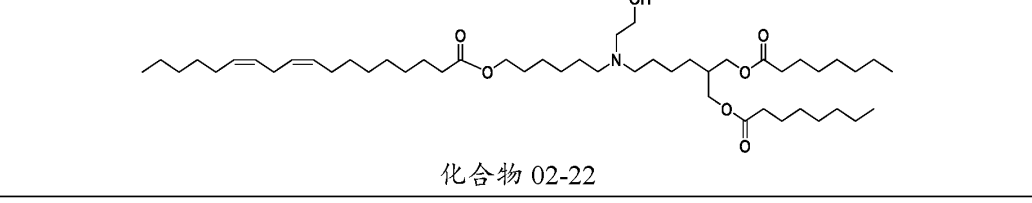
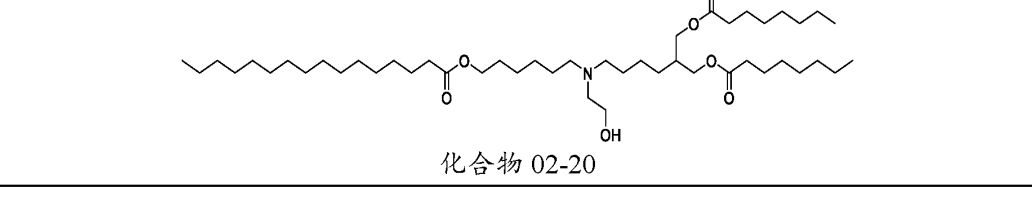
x 为 0、1 或 2；并且

其中每个烷基、烯基、环烷基、环烯基、杂环烷基、芳基、亚烷基、亚烯基、亚环
25 烷基和亚环烯基独立地任选地经取代。

在一实施方案中，阳离子脂质包含表 02-1 中的化合物，或其药学上可接受的盐、前
药或立体异构体中的一种或多种。

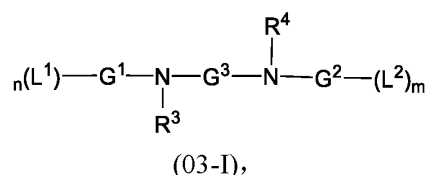
表 02-1.



 化合物 02-3	 化合物 02-4
 化合物 02-9	 化合物 02-10
 化合物 02-14	 化合物 02-12
 化合物 02-11	
 化合物 02-22	
 化合物 02-20	

3) 脂质系列 03

在一实施方案中，阳离子脂质包含式(03-I)化合物：



5 或其药学上可接受的盐、前药或立体异构体，其中：

G^1 和 G^2 各自独立地是键、 $\text{C}_2\text{-C}_{12}$ 亚烷基或 $\text{C}_2\text{-C}_{12}$ 亚烯基，其中 G^1 和 G^2 中的一个或多个 $\text{-CH}_2\text{-}$ 任选地经 -O- 替代；

每个 L^1 独立地是 -OC(=O)R^1 、 -C(=O)OR^1 、 -OC(=O)OR^1 、 -C(=O)R^1 、 -OR^1 、 $\text{-S(O)}_x\text{R}^1$ 、 -S-SR^1 、 -C(=O)SR^1 、 -SC(=O)R^1 、 $\text{-NR}^a\text{C(=O)R}^1$ 、 $\text{-C(=O)NR}^b\text{R}^c$ 、 $\text{-NR}^a\text{C(=O)NR}^b\text{R}^c$ 、 $\text{-OC(=O)NR}^b\text{R}^c$ 、 $\text{-NR}^a\text{C(=O)OR}^1$ 、 -SC(=S)R^1 、 -C(=S)SR^1 、 -C(=S)R^1 、 -CH(OH)R^1 、 $\text{-P(=O)(OR}^b\text{)(OR}^c\text{)}$ 、 $\text{-NR}^a\text{P(=O)(OR}^b\text{)(OR}^c\text{)}$ 、 $\text{-(C}_6\text{-C}_{10}\text{ 亚芳基)-R}^1$ 、 $\text{-(6 至 10 元亚杂芳基)-R}^1$ 、 $\text{-(4 至 8 元亚杂环基)-R}^1$ 或 R^1 ；

每个 L^2 独立地是 -OC(=O)R^2 、 -C(=O)OR^2 、 -OC(=O)OR^2 、 -C(=O)R^2 、 -OR^2 、 $\text{-S(O)}_x\text{R}^2$ 、 -S-SR^2 、 -C(=O)SR^2 、 -SC(=O)R^2 、 $\text{-NR}^d\text{C(=O)R}^2$ 、 $\text{-C(=O)NR}^e\text{R}^f$ 、 $\text{-NR}^d\text{C(=O)NR}^e\text{R}^f$ 、 $\text{-OC(=O)NR}^e\text{R}^f$ 、 $\text{-NR}^d\text{C(=O)OR}^2$ 、 -SC(=S)R^2 、 -C(=S)SR^2 、 -C(=S)R^2 、 -CH(OH)R^2 、 $\text{-P(=O)(OR}^e\text{)(OR}^f\text{)}$ 、 $\text{-NR}^d\text{P(=O)(OR}^e\text{)(OR}^f\text{)}$ 、 $\text{-(C}_6\text{-C}_{10}\text{ 亚芳基)-R}^2$ 、 $\text{-(6 至 10 元亚杂芳基)-R}^2$ 、 $\text{-(4 至 8 元亚杂环基)-R}^2$ 或 R^2 ；

R^1 和 R^2 各自独立地是 $\text{C}_6\text{-C}_{24}$ 烷基或 $\text{C}_6\text{-C}_{24}$ 烯基；

R^a 、 R^b 、 R^d 和 R^e 各自独立地是 H 、 $\text{C}_1\text{-C}_{24}$ 烷基或 $\text{C}_2\text{-C}_{24}$ 烯基；

20 R^c 和 R^f 各自独立地是 $\text{C}_1\text{-C}_{24}$ 烷基或 $\text{C}_2\text{-C}_{24}$ 烯基；

G^3 是 $\text{C}_2\text{-C}_{12}$ 亚烷基或 $\text{C}_2\text{-C}_{12}$ 亚烯基，其中亚烷基或亚烯基的一部分或全部任选地经 $\text{C}_3\text{-C}_8$ 亚环烷基、 $\text{C}_3\text{-C}_8$ 亚环烯基、 $\text{C}_3\text{-C}_8$ 亚环炔基、4 至 8 元亚杂环基、 $\text{C}_6\text{-C}_{10}$ 亚芳基或 5 至 10 元亚杂芳基替代；

R^3 是氢、 $\text{C}_1\text{-C}_{12}$ 烷基、 $\text{C}_2\text{-C}_{12}$ 烯基、 $\text{C}_2\text{-C}_{12}$ 炔基、 $\text{C}_3\text{-C}_8$ 环烷基、 $\text{C}_3\text{-C}_8$ 环烯基、 $\text{C}_3\text{-C}_8$ 环炔基、4 至 8 元杂环基、 $\text{C}_6\text{-C}_{10}$ 芳基或 5 至 10 元杂芳基；或 R^3 、 G^1 或 G^2 的一部分与它们所连接的氮一起形成环状部分；或 R^3 、 G^3 或 G^2 的一部分与它们所连接的氮一起形成环状部分；

R^4 是 $\text{C}_1\text{-C}_{12}$ 烷基或 $\text{C}_3\text{-C}_8$ 环烷基；

x 是 0、1 或 2；

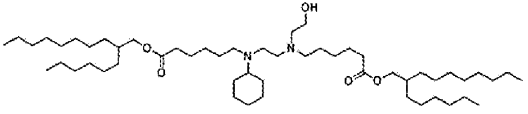
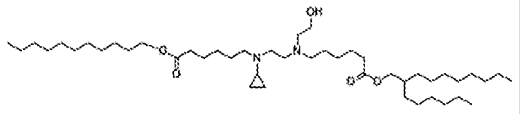
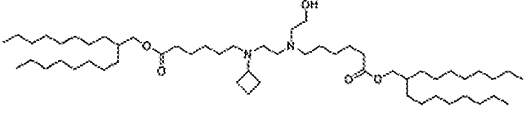
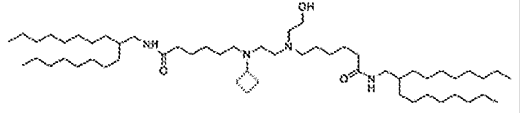
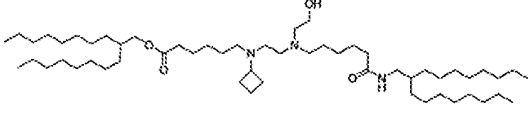
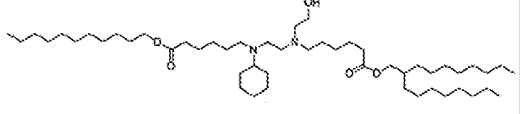
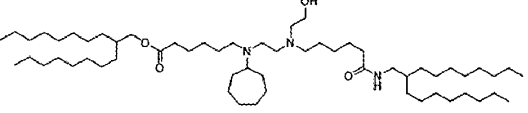
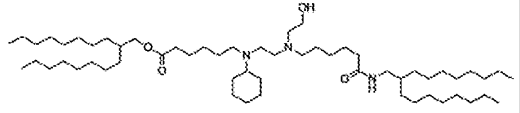
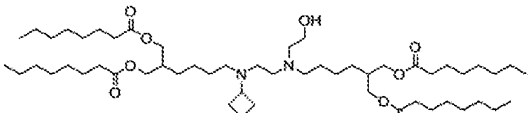
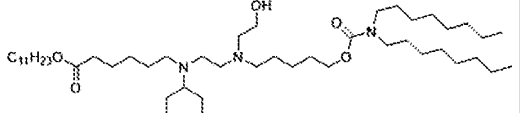
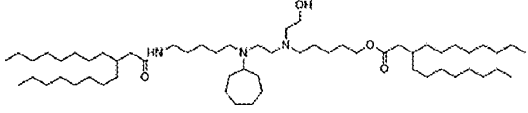
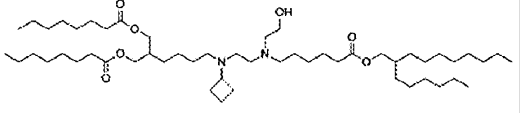
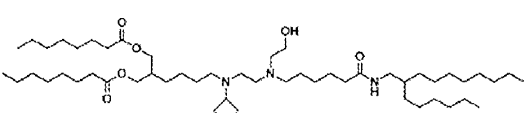
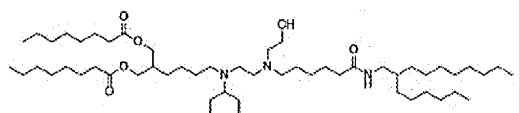
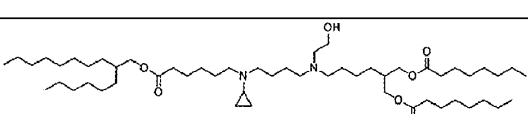
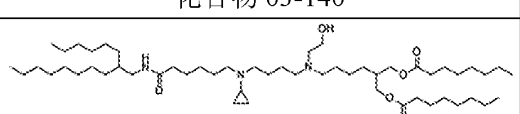
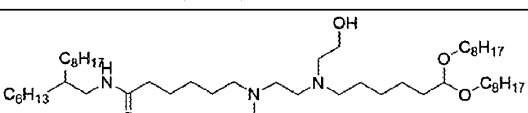
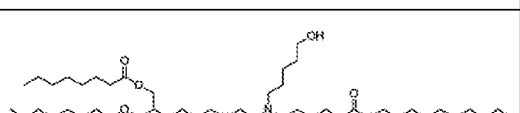
30 n 是 1 或 2；

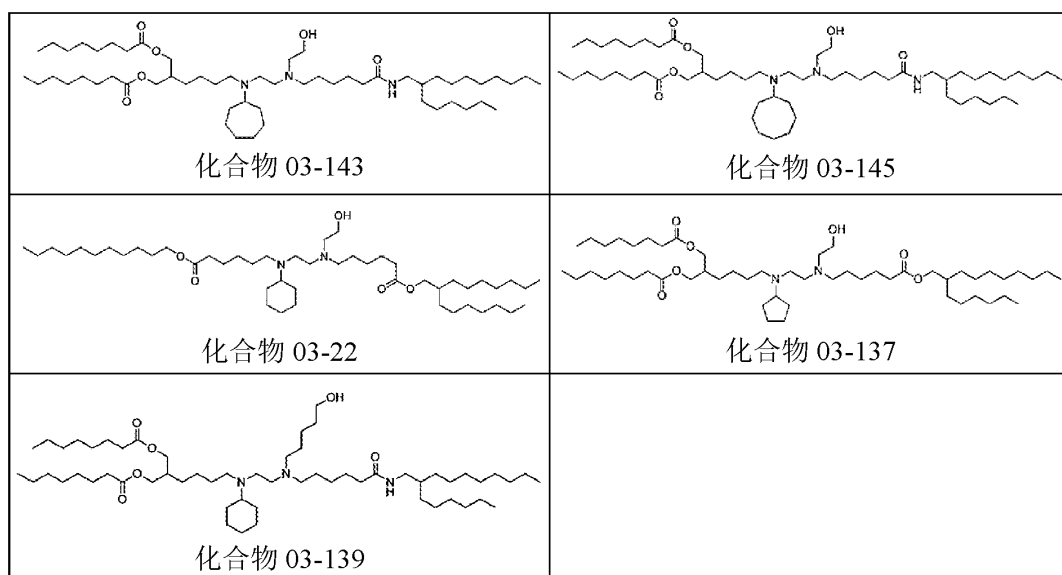
m 是 1 或 2；并且

其中每个烷基、烯基、炔基、环烷基、环烯基、环炔基、杂环基、芳基、杂芳基、亚烷基、亚烯基、亚环烷基、亚环烯基、亚环炔基、亚杂环基、亚芳基、亚杂芳基和环状部分独立地任选地经取代。

35 在一实施方案中，阳离子脂质包含表 03-1 中的化合物，或其药学上可接受的盐、前药或立体异构体中的一种或多种。

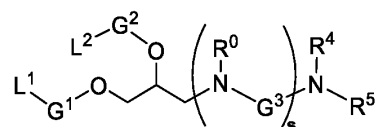
表 03-1.

 化合物 03-1	 化合物 03-3
 化合物 03-10	 化合物 03-11
 化合物 03-15	 化合物 03-21
 化合物 17	 化合物 03-23
 化合物 03-71	 化合物 99
 化合物 03-74	 化合物 03-135
 化合物 03-136	 化合物 03-140
 化合物 03-190	 化合物 03-195
 化合物 202	 化合物 03-209



4) 脂质系列 04

在一实施方案中，阳离子脂质包含式(04-I)化合物：



5

(04-I),

或其药学上可接受的盐、前药或立体异构体中的一种或多种，其中：

G^1 和 G^2 各自独立地是键、 $\text{C}_2\text{-C}_{12}$ 亚烷基或 $\text{C}_2\text{-C}_{12}$ 亚烯基；

L^1 是 $-\text{OC}(=\text{O})\text{R}^1$ 、 $-\text{C}(=\text{O})\text{OR}^1$ 、 $-\text{OC}(=\text{O})\text{OR}^1$ 、 $-\text{C}(=\text{O})\text{R}^1$ 、 $-\text{OR}^1$ 、 $-\text{S}(\text{O})_x\text{R}^1$ 、 $-\text{S-SR}^1$ 、 $-\text{C}(=\text{O})\text{SR}^1$ 、 $-\text{SC}(=\text{O})\text{R}^1$ 、 $-\text{NR}^a\text{C}(=\text{O})\text{R}^1$ 、 $-\text{C}(=\text{O})\text{NR}^b\text{R}^c$ 、 $-\text{NR}^a\text{C}(=\text{O})\text{NR}^b\text{R}^c$ 、 $-\text{OC}(=\text{O})\text{NR}^b\text{R}^c$ 、
 10 $-\text{NR}^a\text{C}(=\text{O})\text{OR}^1$ 、 $-\text{SC}(=\text{S})\text{R}^1$ 、 $-\text{C}(=\text{S})\text{SR}^1$ 、 $-\text{C}(=\text{S})\text{R}^1$ 、 $-\text{CH}(\text{OH})\text{R}^1$ 、 $-\text{P}(=\text{O})(\text{OR}^b)(\text{OR}^c)$ 、 $-(\text{C}_6\text{-C}_{10}$ 亚芳基)- R^1 、 $-(6$ 至 10 元亚杂芳基)- R^1 或 R^1 ；

L^2 是 $-\text{OC}(=\text{O})\text{R}^2$ 、 $-\text{C}(=\text{O})\text{OR}^2$ 、 $-\text{OC}(=\text{O})\text{OR}^2$ 、 $-\text{C}(=\text{O})\text{R}^2$ 、 $-\text{OR}^2$ 、 $-\text{S}(\text{O})_x\text{R}^2$ 、 $-\text{S-SR}^2$ 、 $-\text{C}(=\text{O})\text{SR}^2$ 、 $-\text{SC}(=\text{O})\text{R}^2$ 、 $-\text{NR}^d\text{C}(=\text{O})\text{R}^2$ 、 $-\text{C}(=\text{O})\text{NR}^e\text{R}^f$ 、 $-\text{NR}^d\text{C}(=\text{O})\text{NR}^e\text{R}^f$ 、 $-\text{OC}(=\text{O})\text{NR}^e\text{R}^f$ 、
 15 $-\text{NR}^d\text{C}(=\text{O})\text{OR}^2$ 、 $-\text{SC}(=\text{S})\text{R}^2$ 、 $-\text{C}(=\text{S})\text{SR}^2$ 、 $-\text{C}(=\text{S})\text{R}^2$ 、 $-\text{CH}(\text{OH})\text{R}^2$ 、 $-\text{P}(=\text{O})(\text{OR}^e)(\text{OR}^f)$ 、 $-(\text{C}_6\text{-C}_{10}$ 亚芳基)- R^2 、 $-(6$ 至 10 元亚杂芳基)- R^2 或 R^2 ；

R^1 和 R^2 各自独立地是 $\text{C}_5\text{-C}_{32}$ 烷基或 $\text{C}_5\text{-C}_{32}$ 烯基；

R^a 、 R^b 、 R^d 和 R^e 各自独立地是 H 、 $\text{C}_1\text{-C}_{24}$ 烷基或 $\text{C}_2\text{-C}_{24}$ 烯基；

R^c 和 R^f 各自独立地是 $\text{C}_1\text{-C}_{32}$ 烷基或 $\text{C}_2\text{-C}_{32}$ 烯基；

R^0 是 $\text{C}_1\text{-C}_{12}$ 烷基、 $\text{C}_2\text{-C}_{12}$ 烯基、 $\text{C}_3\text{-C}_8$ 环烷基、 $\text{C}_3\text{-C}_8$ 环烯基、 $\text{C}_6\text{-C}_{10}$ 芳基或 4 至 8
 20 元杂环烷基；

G^3 是 $\text{C}_2\text{-C}_{12}$ 亚烷基或 $\text{C}_2\text{-C}_{12}$ 亚烯基；

R^4 是 $\text{C}_1\text{-C}_{12}$ 烷基、 $\text{C}_2\text{-C}_{12}$ 烯基、 $\text{C}_3\text{-C}_8$ 环烷基、 $\text{C}_3\text{-C}_8$ 环烯基、 $\text{C}_6\text{-C}_{10}$ 芳基或 4 至 8 元杂环烷基；

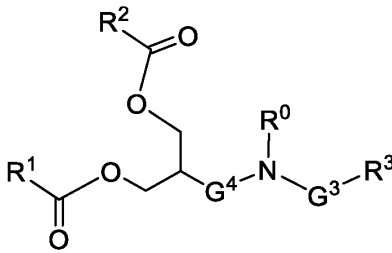
R⁵ 是 C₁-C₁₂ 烷基、C₃-C₈ 环烷基、C₃-C₈ 环烯基、C₆-C₁₀ 芳基或 4 至 8 元杂环烷基；

x 是 0、1 或 2；

s 是 0 或 1；并且

其中每个烷基、烯基、环烷基、环烯基、杂环烷基、芳基、亚烷基、亚烯基、亚芳基和亚杂芳基独立地任选地经取代。

在一实施方案中，阳离子脂质为式(04-III)化合物：



(04-III),

或其药学上可接受的盐、前药或立体异构体，其中：

R¹ 和 R² 各自独立地为 C₅-C₃₂ 烷基或 C₅-C₃₂ 烯基；

R⁰ 为 C₁-C₁₂ 烷基、C₂-C₁₂ 烯基、C₃-C₈ 环烷基、C₃-C₈ 环烯基、C₆-C₁₀ 芳基或 4 元至 8 元杂环烷基；

G³ 为 C₂-C₁₂ 亚烷基或 C₂-C₁₂ 亚烯基；

G⁴ 为 C₂-C₁₂ 亚烷基或 C₂-C₁₂ 亚烯基；

R³ 为 -N(R⁴)R⁵ 或 -OR⁶；

R⁴ 为 C₁-C₁₂ 烷基、C₂-C₁₂ 烯基、C₃-C₈ 环烷基、C₃-C₈ 环烯基、C₆-C₁₀ 芳基或 4 元至 8 元杂环烷基；

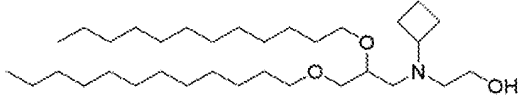
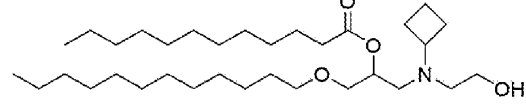
R⁵ 为 C₁-C₁₂ 烷基、C₃-C₈ 环烷基、C₃-C₈ 环烯基、C₆-C₁₀ 芳基或 4 元至 8 元杂环烷基；或 R⁴、R⁵ 与其所连接的氮一起形成环状部分；

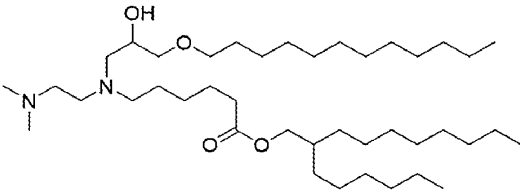
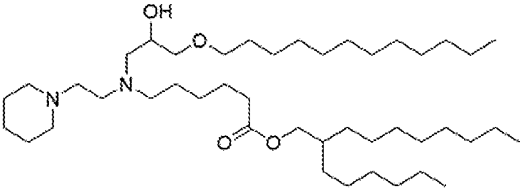
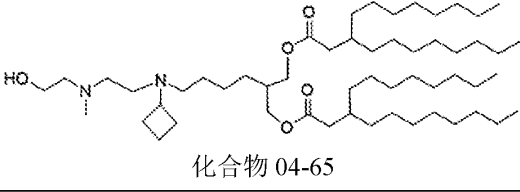
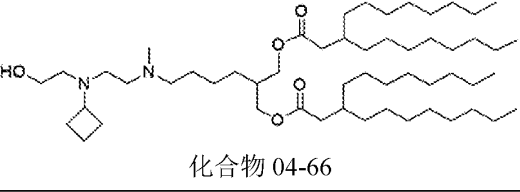
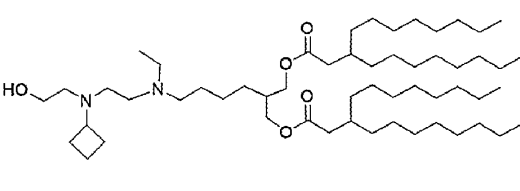
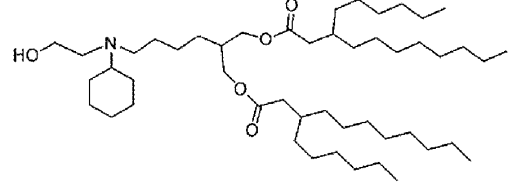
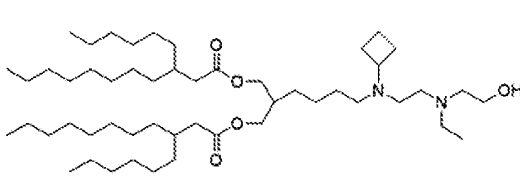
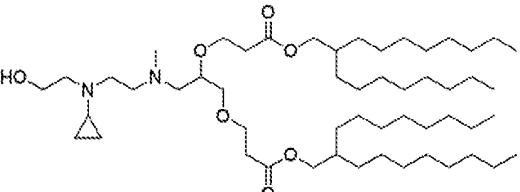
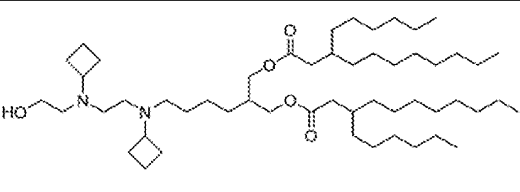
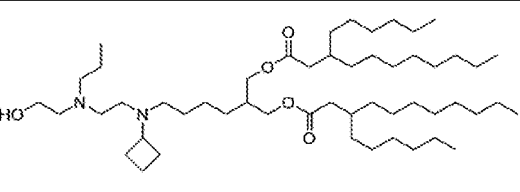
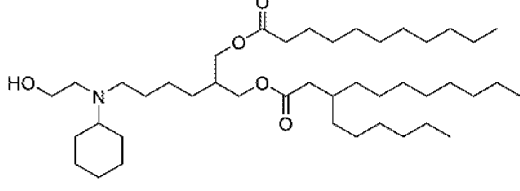
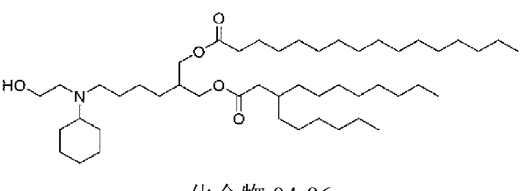
R⁶ 为氢、C₁-C₁₂ 烷基、C₃-C₈ 环烷基、C₃-C₈ 环烯基或 C₆-C₁₀ 芳基；并且

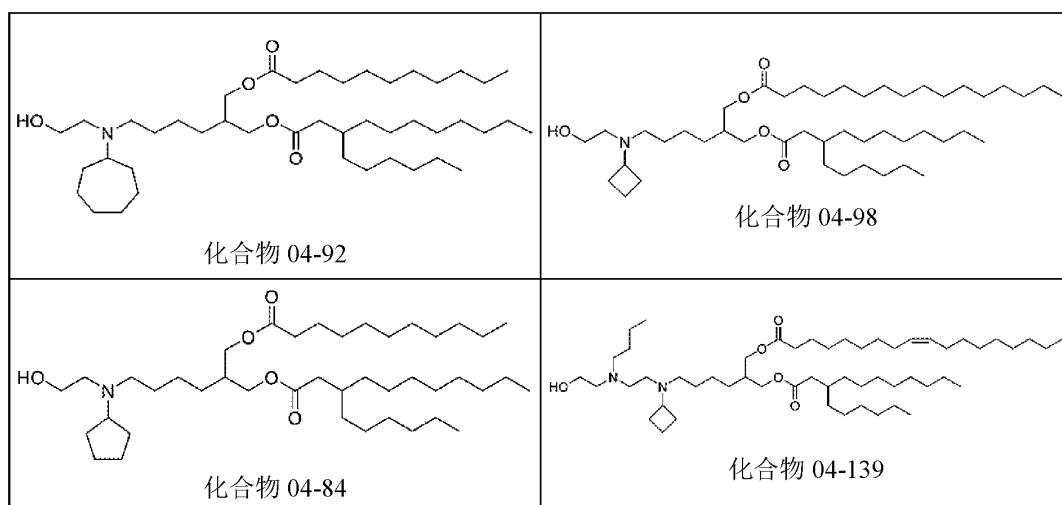
其中每个烷基、烯基、环烷基、环烯基、杂环烷基、芳基、亚烷基、亚烯基和环状部分独立地任选地经取代。

在一实施方案中，阳离子脂质包含表 04-1 中的化合物，或其药学上可接受的盐、前药或立体异构体。

表 04-1.

 化合物 04-1	 化合物 04-2
---	--

 化合物 04-7	 化合物 04-8
 化合物 04-65	 化合物 04-66
 化合物 04-67	 化合物 04-68
 化合物 04-69	 化合物 04-70
 化合物 04-79	 化合物 04-80
 化合物 04-85	 化合物 04-86



在一优选实施方案中，阳离子脂质包含化合物 01-1。

应理解，如上所陈述的本文所提供的化合物的任何实施方案，以及如上所陈述的本文所提供的化合物的任何特定取代基和/或变量可独立地与化合物的其它实施方案和/或取代基和/或变量组合以形成以上未具体陈述的实施方案。此外，在列出任何特定基团或变量的取代基和/或变量的列表的情况下，应理解，可从特定实施方案和/或技术方案中删除每一个别取代基和/或变量，并且其余的取代基和/或变量列表将被认为在本文所提供的实施方案的范围内。

应理解，在本说明书中，所描绘各式的取代基和/或变量的组合仅在此类贡献产生稳定化合物时才为容许的。

其它可离子化脂质

如本文所述，在一些实施方案中，本文提供的纳米颗粒组合物除了包含系列 01、02、03 和 04 的脂质中的一种或多种，例如根据式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-III) (及其子式)的脂质中的一种或多种以外，还包含一种或多种带电或可离子化脂质。不受理论束缚，预期纳米颗粒组合物的某些带电或两性离子脂质组分类似于细胞膜中的脂质组分，由此可改善纳米颗粒的细胞摄取。可形成本发明纳米颗粒组合物的一部分的示例性其它带电脂质、可离子化脂质或阳离子脂质可以参见例如 PCT 公布第 WO 2023/098842 号，其内容整体援引加入本文。

此外，在一些实施方案中，可形成本发明纳米颗粒组合物的一部分的带电或可离子化脂质是包含环胺基团的脂质。适于本文所公开的制剂和方法的额外阳离子脂质包括 WO2015199952A1、WO2016176330A1 和 WO2015011633A1 中所述的那些阳离子脂质，这些文献的全部内容以引用的方式整体并入本文中。此外，在一些实施方案中，可形成本发明纳米颗粒组合物的一部分的带电或可离子化脂质是包含环胺基团的脂质。适于本文所公开的制剂和方法的额外阳离子脂质包括 WO2015199952A1、WO2016176330A1 和 WO2015011633A1 中所述的那些阳离子脂质，这些文献的全部内容以引用的方式整体并入本文中。

聚合物结合的脂质

在一些实施方案中，纳米颗粒组合物的脂质组分可包含一种或多种聚合物结合的脂质，诸如聚乙二醇化脂质(PEG 脂质)。不受理论束缚，预期纳米颗粒组合物中的聚合物结合的脂质组分可改善胶体稳定性和/或减少纳米颗粒的蛋白质吸收。可与本公开结合使用的示例性聚合物结合的脂质包括但不限于 PEG 改性的磷脂酰乙醇胺、PEG 改性的磷脂酸、PEG 改性的神经酰胺、PEG 改性的二烷基胺、PEG 改性的二酰基甘油、PEG 改性的二烷基甘油和其混合物。举例来说，PEG 脂质可为 PEG-c-DOMG、PEG-DMG、PEG-DLPE、PEG-DMPE、PEG-DPPC、PEG-DSPE、神经酰胺-PEG2000 或 Chol-PEG2000。

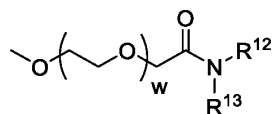
在一实施方案中，聚合物结合的脂质是聚乙二醇化脂质。举例来说，一些实施方案包括聚乙二醇化二酰基甘油(PEG-DAG)，诸如 1-(单甲氧基-聚乙二醇)-2,3-二肉豆蔻酰基甘油(PEG-DMG)；聚乙二醇化磷脂酰乙醇胺(PEG-PE)；PEG 琥珀酸酯二酰基甘油(PEG-S-DAG)，诸如 4-O-(2',3'-二(十四烷酰氧基)丙基-1-O-(ω -甲氧基(聚乙氧基)乙基)丁二酸酯(PEG-S-DMG)；聚乙二醇化神经酰胺(PEG-cer)；或 PEG 二烷氧基丙基氨基甲酸酯，诸如 ω -甲氧基(聚乙氧基)乙基-N-(2,3-二(十四烷氧基)丙基)氨基甲酸酯或 2,3-二(十四烷氧基)丙基-N-(ω -甲氧基)(聚乙氧基)乙基)氨基甲酸酯。

在一些实施方案中，聚合物结合的脂质(优选地，聚乙二醇化脂质)以在 0.5 mol% 至 5 mol% 范围内的浓度存在。在一实施方案中，聚合物结合的脂质以在 1.0 mol% 至 2.5 mol% 范围内的浓度存在。在一实施方案中，聚合物结合的脂质以在约 1.2 mol% 至 2 mol% 范围内的浓度存在。在一实施方案中，聚合物结合的脂质以约 1.7 mol% 的浓度存在。在一实施方案中，聚合物结合的脂质以约 1.5 mol% 的浓度存在。

在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 35:1 至约 25:1 范围内。在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 100:1 至约 20:1 范围内。

在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 35:1 至约 25:1 范围内。在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 100:1 至约 20:1 范围内。

在一实施方案中，聚乙二醇化脂质具有下式：



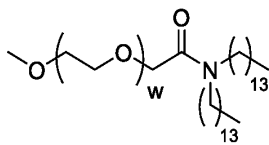
或其药学上可接受的盐、互变异构体或立体异构体，其中：

R^{12} 和 R^{13} 各自独立地是含有 10 至 30 个碳原子的直链或支链饱和或不饱和烷基链，其中所述烷基链任选地间杂有一个或多个酯键；并且

w 具有介于 30 至 60 范围内的平均值。

在一实施方案中， R^{12} 和 R^{13} 各自独立地是含有 12 至 16 个碳原子的直链饱和烷基链。在其它实施方案中，w 平均值在 42 至 55 的范围内，例如，w 平均值是 42、43、44、45、46、47、48、49、50、51、52、53、54 或 55。在一些具体实施方案中，w 平均值是约 49。

在一实施方案中，聚乙二醇化脂质具有下式：

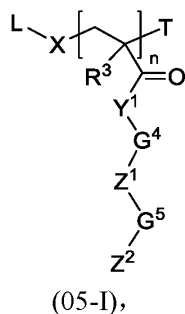


其中 w 平均值是约 49。

聚合物结合的脂质还包括以下根据脂质系列 05 (及其子式)的化合物中的一种或多种。

5) 脂质系列 05

在一实施方案中, 本文提供了式(05-I)化合物:



或其药学上可接受的盐或立体异构体中的一种或多种, 其中:

10 L 是脂质;

X 是连接子;

每个 R^3 独立地是 H 或 C_1-C_6 烷基;

每个 Y^1 独立地是键、O、S 或 NR^a ;

15 每个 G^4 独立地是键或 C_1-C_{12} 亚烷基, 其中一个或多个 $-CH_2-$ 独立地任选地经-O-、-S-或- NR^a -替代;

每个 G^5 独立地是键或 C_1-C_{12} 亚烷基, 其中一个或多个 $-CH_2-$ 独立地任选地经-O-、-S-或- NR^a -替代;

每个 R^a 独立地是 H、 C_1-C_{12} 烷基或 C_2-C_{12} 烯基;

Z^1 和 Z^2 中的一个带正电荷的部分, 并且 Z^1 和 Z^2 中的另一个是带负电荷的部分;

20 n 是 2 至 100 的整数;

T 是氢、卤素、烷基、烯基、-OR^{''}、-SR^{''}、-COOR^{''}、-OCOR^{''}、-NR^{''}R^{''}、-N⁺(R^{''})₃、-P⁺(R^{''})₃、-S-C(=S)-S-R^{''}、-S-C(=S)-O-R^{''}、-S-C(=S)-NR^{''}R^{''}、-S-C(=S)-芳基、氰基、叠氮基、芳基、杂芳基或靶向基团, 其中 R^{''} 在每次出现时独立地是氢或烷基; 并且

25 其中每个烷基、烯基、亚烷基、芳基和杂芳基独立地任选地经取代。应理解, 如上所陈述的本文所提供的化合物的任何实施方案, 以及如上所陈述的本文所提供的化合物的任何特定取代基和/或变量可独立地与化合物的其它实施方案和/或取代基和/或变量组合以形成以上未具体陈述的实施方案。此外, 在列出任何特定基团或变量的取代基和/或变量的列表的情况下, 应理解, 可从特定实施方案和/或权利要求中删除每一个别取代基和/或变量, 并且其余的取代基和/或变量列表将被认为在本文所提供的实施方案的范围

30 范围内。

应理解，在本说明书中，所描绘各式的取代基和/或变量的组合仅在此类贡献产生稳定化合物时才为容许的。

结构脂质

在一些实施方案中，纳米颗粒组合物的脂质组分可包含一种或多种结构脂质。不受理论束缚，预期结构脂质可使纳米颗粒的两亲结构，诸如但不限于纳米颗粒的脂质双层结构稳定。可与本公开结合使用的示例性结构脂质包括但不限于胆固醇、粪固醇、谷固醇、麦角固醇、菜油固醇、豆固醇、菜子固醇、番茄碱、番茄苷、熊果酸、 α -生育酚和其混合物。在某些实施方案中，结构脂质是胆固醇。在一些实施方案中，结构脂质包括胆固醇和皮质类固醇(诸如泼尼松龙(prednisolone)、地塞米松(dexamethasone)、泼尼松(prednisone)和氢化可的松(hydrocortisone))或其组合。

在一些实施方案中，结构脂质以在 20 mol%至 50 mol%范围内的浓度存在。在一些实施方案中，结构脂质以在 30 mol%至 50 mol%范围内的浓度存在。优选地，结构脂质以在 32 mol%至 46 mol%范围内的浓度存在。更优选地，结构脂质以在 34 mol%至 44 mol%范围内的浓度存在。还更优选地，结构脂质以在 36 mol%至 42 mol%范围内的浓度存在。

在一些实施方案中，结构脂质以约 30 mol%、31 mol%、32 mol%、33 mol%、34 mol%、35 mol%、36 mol%、37 mol%、38 mol%、38.5 mol%、39 mol%、40 mol%、41 mol%、42 mol%、43 mol%、44 mol%、45 mol%、46 mol%、47 mol%、48 mol%、49 mol%或 50 mol%的浓度存在。

在一实施方案中，本文所提供的脂质纳米颗粒包含类固醇或类固醇类似物。在一实施方案中，类固醇或类固醇类似物是胆固醇。

在一实施方案中，类固醇以在 39 mol%至 49 mol%、40 mol%至 46 mol%、40 mol%至 44 mol%、40 mol%至 42 mol%、42 mol%至 44 mol%或 44 mol%至 46 mol%范围内的浓度存在。在一实施方案中，类固醇以 40 mol%、41 mol%、42 mol%、43 mol%、44 mol%、45 mol%或 46 mol%的浓度存在。

在一实施方案中，阳离子脂质与类固醇的摩尔比在 1.0:0.9 至 1.0:1.2、或 1.0:1.0 至 1.0:1.2 范围内。在一实施方案中，阳离子脂质与胆固醇的摩尔比在约 5:1 至 1:1 范围内。在一实施方案中，类固醇以在 32 mol%至 40 mol%类固醇范围内的浓度存在。

在一实施方案中，阳离子脂质与类固醇的摩尔比在 1.0:0.9 至 1.0:1.2、或 1.0:1.0 至 1.0:1.2 范围内。在一实施方案中，阳离子脂质与胆固醇的摩尔比在约 5:1 至 1:1 范围内。在一实施方案中，类固醇以在 32 mol%至 40 mol%类固醇范围内的浓度存在。

磷脂

在一些实施方案中，纳米颗粒组合物的脂质组分可包含一种或多种磷脂，诸如一种或多种(多)不饱和脂质。不受理论束缚，预期磷脂可组装成一个或多个脂质双层结构。可形成本发明纳米颗粒组合物的一部分的示例性磷脂包括但不限于 1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)、1,2-二油酰基-sn-甘油-3-磷酸乙醇胺(DOPE)、1,2-二亚油酰基-sn-甘油-3-磷酸胆碱(DLPC)、1,2-二肉豆蔻酰基-sn-甘油-磷酸胆碱(DMPC)等。示例性磷脂可以参见例如 PCT 公布第 WO 2023/098842 号，其内容整体援引加入本文。在某些实施方案中，纳米颗粒组合物包含 DSPC。在某些实施方案中，纳米颗粒组合物包含 DOPE。

在一些实施方案中，纳米颗粒组合物包含 DSPC 和 DOPE 二者。

额外示例性磷脂包括例如二棕榈酰基磷脂酰甘油(DPPG)、棕榈酰基油酰基磷脂酰乙醇胺(POPE)和二油酰基磷脂酰乙醇胺 4-(N-马来酰亚胺基甲基)-环己烷-1-甲酸酯(DOPE-mal)、二棕榈酰基磷脂酰乙醇胺(DPPE)、二肉豆蔻酰基磷酸乙醇胺(DMPE)、二硬脂酰基-5 磷脂酰乙醇胺(DSPE)、16-O-单甲基 PE、16-O-二甲基 PE、18-1-反式 PE、1-硬脂酰基-2-油酰基磷脂酰乙醇胺(SOPE)和 1,2-二反油酰基-sn-甘油-3-磷酸乙醇胺(反式 DOPE)。在一实施方案中，中性脂质为 1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。在一实施方案中，中性脂质选自 DSPC、DPPC、DMPC、DOPC、POPC、DOPE 和 SM。在一实施方案中，磷脂为 1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱(DSPC)。在一实施方案中，磷脂选自 10 DSPC、DPPC、DMPC、DOPC、POPC、DOPE 和 SM。

在一实施方案中，磷脂是磷脂酰胆碱(PC)、磷脂酰乙醇胺(PE)、磷脂酰丝氨酸(PS)、磷脂酸(PA)或磷脂酰甘油(PG)。

此外，可形成本发明纳米颗粒组合物的一部分的磷脂还包括 WO2017/112865 中所述的那些磷脂，该案的全部内容以引用的方式整体并入本文中。

15 在一些实施方案中，磷脂以在 5 mol%至 40 mol%范围内的浓度存在。在一些实施方案中，磷脂以在 5 mol%至 15 mol%范围内的浓度存在。优选地，磷脂以在 7 mol%至 13 mol%范围内的浓度存在。更优选地，磷脂以在 9 mol%至 11 mol%范围内的浓度存在。

在一些实施方案中，磷脂以约 5 mol%、6 mol%、7 mol%、8 mol%、9 mol%、9.5 mol%、10 mol%、10.5 mol%、11 mol%、12 mol%、13 mol%、14 mol%或 15 mol%的浓度存在。 20

制剂

根据本公开，本文所述的纳米颗粒组合物可包含至少一种脂质组分和一种或多种额外组分，诸如核酸或核酸组合(例如，本文所述的治疗性核酸或核酸组合)。可针对一种或多种特定应用或靶来设计纳米颗粒组合物。纳米颗粒组合物的成分可基于特定应用 25 或靶，和/或基于一种或多种成分的功效、毒性、费用、易用性、可用性或其它特征来选择。类似地，纳米颗粒组合物的特定制剂可根据例如每种成分的特定组合的功效和毒性，针对特定应用或靶来选择。

纳米颗粒组合物的脂质组分可包含例如本文所述的系列 01、02、03 和 04 的脂质中的一种或多种，例如根据式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-III)(及其子式)的脂质中的一种或多种、磷脂(诸如不饱和脂质，例如 DOPE 或 DSPC)、PEG 脂质 30 和结构脂质。脂质组分的成分可以特定的分率提供。

在一实施方案中，本文提供了纳米颗粒组合物，所述纳米颗粒组合物包含本文提供的阳离子或可离子化脂质化合物、核酸或核酸组合和一种或多种其他脂质。在一实施方案中，阳离子或可离子化脂质化合物包含如本文所述的系列 01、02、03 和 04 的化合物 35 中的一种或多种，例如根据式(01-I)、(01-II)、(02-I)、(02-II)、(03-I)、(04-I)和(04-III)(及其子式)的化合物中的一种或多种，以及任选存在的一种或多种其它可离子化脂质化合物。在一实施方案中，一种或多种其他脂质选自中性脂质、类固醇和聚合物结合的脂质。在一实施方案中，核酸或核酸组合被包封在脂质纳米颗粒内或与脂质纳米颗粒缔合。

在一实施方案中，本文提供了纳米颗粒组合物(脂质纳米颗粒)，所述纳米颗粒组合

物包含：

- i) 介于 40 mol% 与 50 mol% 之间的阳离子脂质；
- ii) 中性脂质；
- iii) 类固醇；
- 5 iv) 聚合物结合的脂质；和
- v) 核酸或核酸组合。

如本文所用，“mol%”是指组分相对于 LNP 中所有脂质组分的总摩尔数(即阳离子脂质、中性脂质、类固醇和聚合物结合的脂质的总摩尔数)的摩尔百分率。除非特别指明，组合物中各脂质的量(mol%)的总和为 100 mol%，即阳离子脂质、磷脂、类固醇和聚乙二醇结合的脂质的量(mol%)的总和为 100 mol%。

在一实施方案中，脂质纳米颗粒包含 20 mol% 至 65 mol% 的阳离子脂质。更优选地，脂质纳米颗粒包含 25 mol% 至 60 mol% 的阳离子脂质。还更优选地，脂质纳米颗粒包含 30 mol% 至 55 mol% 的阳离子脂质。还更优选地，脂质纳米颗粒包含 35 mol% 至 55 mol% 的阳离子脂质。还更优选地，脂质纳米颗粒包含 40 mol% 至 52 mol% 的阳离子脂质。

15 在一实施方案中，脂质纳米颗粒包含 41 mol% 至 49 mol%、41 mol% 至 48 mol%、42 mol% 至 48 mol%、43 mol% 至 48 mol%、44 mol% 至 48 mol%、45 mol% 至 48 mol%、46 mol% 至 48 mol% 或 47.2 mol% 至 47.8 mol% 的阳离子脂质。在一实施方案中，脂质纳米颗粒包含约 47.0 mol%、47.1 mol%、47.2 mol%、47.3 mol%、47.4 mol%、47.5 mol%、47.6 mol%、47.7 mol%、47.8 mol%、47.9 mol% 或 48.0 mol% 的阳离子脂质。

20 在一实施方案中，中性脂质以 5 mol% 至 15 mol%、7 mol% 至 13 mol% 或 9 mol% 至 11 mol% 范围内的浓度存在。在一实施方案中，中性脂质以约 9.5 mol%、10 mol% 或 10.5 mol% 的浓度存在。在一实施方案中，阳离子脂质与中性脂质的摩尔比在约 4.1:1.0 至约 4.9:1.0、约 4.5:1.0 至约 4.8:1.0 或约 4.7:1.0 至 4.8:1.0 的范围内。

在一实施方案中，类固醇以在 39 mol% 至 49 mol%、40 mol% 至 46 mol%、40 mol% 25 至 44 mol%、40 mol% 至 42 mol%、42 mol% 至 44 mol% 或 44 mol% 至 46 mol% 范围内的浓度存在。在一实施方案中，类固醇以 40 mol%、41 mol%、42 mol%、43 mol%、44 mol%、45 mol% 或 46 mol% 的浓度存在。在一实施方案中，阳离子脂质与类固醇的摩尔比在 1.0:0.9 至 1.0:1.2 或 1.0:1.0 至 1.0:1.2 的范围内。在一实施方案中，类固醇是胆固醇。

在一实施方案中，LNP 中的核酸或核酸组合:脂质比(即，N/P，其中 N 代表阳离子脂质的摩尔数，并且 P 代表作为核酸主链的一部分存在的磷酸酯的摩尔数)在 2:1 至 30:1，30 例如 3:1 至 22:1 的范围内。在一实施方案中，N/P 在 6:1 至 20:1 或 2:1 至 12:1 的范围内。示例性 N/P 范围包括约 3:1、约 6:1、约 12:1 和约 22:1。

在一实施方案中，本文提供了脂质纳米颗粒，所述脂质纳米颗粒包含：

- i) 有效 pKa 大于 6.0 的阳离子脂质；
- 35 ii) 5 mol% 至 15 mol% 的中性脂质；
- iii) 1 mol% 至 15 mol% 的阴离子脂质；
- iv) 30 mol% 至 45 mol% 的类固醇；
- v) 聚合物结合的脂质；和
- vi) 核酸或核酸组合，

其中 mol% 是基于脂质纳米颗粒中存在的脂质的总摩尔数确定的。

在一实施方案中，本文提供的组合物包括：

(a) (i) 20 mol% 至 65 mol% 的阳离子脂质；(ii) 5 mol% 至 40 mol% 的磷脂；(iii) 20 mol% 至 50 mol% 的类固醇；和 (iv) 聚合物结合的脂质；

5 (b) (i) 40 mol% 至 55 mol% 的阳离子脂质；(ii) 5 mol% 至 15 mol% 的磷脂；(iii) 35 mol% 至 50 mol% 的类固醇；和 (iv) 2 mol% 至 10 mol% 的聚合物结合的脂质；

或

(c) (i) 45 mol% 至 55 mol% 的阳离子脂质；(ii) 6 mol% 至 10 mol% 的磷脂；(iii) 40 mol% 至 48 mol% 的类固醇；和 (iv) 1 mol% 至 2.5 mol% 的聚合物结合的脂质。

10 在一实施方案中，阳离子脂质可以是在选定 pH (诸如生理 pH) 下携带净正电荷的多种脂质种类中的任一种。示例性阳离子脂质在下文描述。在一实施方案中，阳离子脂质具有大于 6.25 的 pKa。在一实施方案中，阳离子脂质具有大于 6.5 的 pKa。在一实施方案中，阳离子脂质具有大于 6.1、大于 6.2、大于 6.3、大于 6.35、大于 6.4、大于 6.45、大于 6.55、大于 6.6、大于 6.65 或大于 6.7 的 pKa。

15 在一实施方案中，脂质纳米颗粒包含 40 mol% 至 45 mol% 的阳离子脂质。在一实施方案中，脂质纳米颗粒包含 45 mol% 至 50 mol% 的阳离子脂质。

在一实施方案中，阳离子脂质与中性脂质的摩尔比在约 2:1 至约 8:1 的范围内。在一实施方案中，脂质纳米颗粒包含 5 mol% 至 10 mol% 的中性脂质。

20 示例性阴离子脂质包括但不限于磷脂酰甘油、二油酰磷脂酰甘油(DOPG)、二棕榈酰磷脂酰甘油(DPPG)或 1,2-二硬脂酰基-sn-甘油-3-磷酸-(1'-外消旋-甘油) (DSPG)。

在一实施方案中，脂质纳米颗粒包含 1 mol% 至 10 mol% 的阴离子脂质。在一实施方案中，脂质纳米颗粒包含 1 mol% 至 5 mol% 的阴离子脂质。在一实施方案中，脂质纳米颗粒包含 1 mol% 至 9 mol%、1 mol% 至 8 mol%、1 mol% 至 7 mol% 或 1 mol% 至 6 mol% 的阴离子脂质。在一实施方案中，阴离子脂质与中性脂质的摩尔比在 1:1 至 1:10 的范围内。

25 在一实施方案中，类固醇是胆固醇。在一实施方案中，阳离子脂质与胆固醇的摩尔比在约 5:1 至 1:1 的范围内。在一实施方案中，脂质纳米颗粒包含 32 mol% 至 40 mol% 的类固醇。

30 在一实施方案中，中性脂质的 mol% 与阴离子脂质的 mol% 之和在 5 mol% 至 15 mol% 的范围内。在一实施方案中，其中性脂质的 mol% 与阴离子脂质的 mol% 之和在 7 mol% 至 12 mol% 的范围内。

在一实施方案中，阴离子脂质与中性脂质的摩尔比在 1:1 至 1:10 的范围内。在一实施方案中，中性脂质的 mol% 和类固醇的 mol% 之和在 35 mol% 至 45 mol% 的范围内。

在一实施方案中，脂质纳米颗粒包含：

35 i) 45 mol% 至 55 mol% 的阳离子脂质；
ii) 5 mol% 至 10 mol% 的中性脂质；
iii) 1 mol% 至 5 mol% 的阴离子脂质；和
iv) 32 mol% 至 40 mol% 的类固醇。

在一实施方案中，脂质纳米颗粒包含 1.0 mol% 至 2.5 mol% 的结合脂质。在一实施方

案中，聚合物结合的脂质以约 1.5 mol% 的浓度存在。

在一实施方案中，中性脂质以 5 mol% 至 15 mol%、7 mol% 至 13 mol% 或 9 mol% 至 11 mol% 范围内的浓度存在。在一实施方案中，中性脂质以约 9.5 mol%、10 mol% 或 10.5 mol% 的浓度存在。在一实施方案中，阳离子脂质与中性脂质的摩尔比在约 4.1:1.0 至约 4.9:1.0、约 4.5:1.0 至约 4.8:1.0 或约 4.7:1.0 至 4.8:1.0 的范围内。

在一实施方案中，类固醇是胆固醇。在一些实施方案中，类固醇以在 39 mol% 至 49 mol%、40 mol% 至 46 mol%、40 mol% 至 44 mol%、40 mol% 至 42 mol%、42 mol% 至 44 mol% 或 44 mol% 至 46 mol% 范围内的浓度存在。在一实施方案中，类固醇以 40 mol%、41 mol%、42 mol%、43 mol%、44 mol%、45 mol% 或 46 mol% 的浓度存在。在某些实施方案中，阳离子脂质与类固醇的摩尔比在 1.0:0.9 至 1.0:1.2 或 1.0:1.0 至 1.0:1.2 的范围内。

在一实施方案中，阳离子脂质与类固醇的摩尔比在 5:1 至 1:1 的范围内。

在一实施方案中，脂质纳米颗粒包含 1.0 mol% 至 2.5 mol% 的结合脂质。在一实施方案中，聚合物结合的脂质以约 1.5 mol% 的浓度存在。

在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 100:1 至约 20:1 的范围内。在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 35:1 至约 25:1 的范围内。

在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 100:1 至约 20:1 的范围内。在一实施方案中，阳离子脂质与聚合物结合的脂质的摩尔比在约 35:1 至约 25:1 的范围内。

在一些实施方案中，脂质纳米颗粒具有在 50 nm 至 180 nm 范围内的平均直径。在一些实施方案中，脂质纳米颗粒具有在 50 nm 至 150 nm 范围内的平均直径。在一些实施方案中，脂质纳米颗粒具有在 50 nm 至 120 nm 范围内的平均直径。在一实施方案中，脂质纳米颗粒具有在 50 nm 至 100 nm 或 60 nm 至 85 nm 范围内的平均直径。

在一实施方案中，组合物包含本文提供的阳离子脂质、DSPC、胆固醇和 PEG-脂质以及核酸或核酸组合。在一实施方案中，本文提供的阳离子脂质、DSPC、胆固醇和 PEG-脂质处于 50:10:38.5:1.5 的摩尔比。

在一优选实施方案中，阳离子脂质包含化合物 01-1。

在一优选实施方案中，PEG-脂质包含 DMG-PEG。

在一优选实施方案中，脂质纳米颗粒包含 30-55 mol% 化合物 01-1、0.5-5 mol% DMG-PEG (例如 DMG-PEG2000)、5-40 mol% DSPC 和 20-50 mol% 胆固醇。

在一最优先实施方案中，组合物包含 50 mol% 化合物 01-1、1.5 mol% DMG-PEG、10 mol% DSPC 和 38.5 mol% 胆固醇。

可针对一种或多种特定应用或靶来设计纳米颗粒组合物。举例来说，纳米颗粒组合物可设计成用于将核酸或核酸组合，诸如 RNA 或 RNA 组合递送至哺乳动物体内的特定细胞、组织、器官或系统或其群组。纳米颗粒组合物的物理化学特性可经改变以增加对特定身体靶的选择性。举例来说，可基于不同器官的开窗大小来调整粒径。纳米颗粒组合物中所包含的核酸或核酸组合还可基于一个或多个所需的递送靶进行选择。举例来说，核酸或核酸组合可针对特定适应症、疾患、疾病或病症和/或针对递送至特定

细胞、组织、器官或系统或其群组(例如局部或特异性递送)进行选择。在某些实施方案中,纳米颗粒组合物可包含编码目标多肽的核酸或核酸组合(特别是 mRNA),其能够在细胞内翻译以产生目标多肽。此类组合物可设计成特异性递送至特定器官。在某些实施方案中,组合物可设计成特异性递送至哺乳动物肝脏。

5 纳米颗粒组合物中核酸或核酸组合的量可取决于纳米颗粒组合物的大小、组成、期望靶和/或应用,或其它特性,以及核酸或核酸组合的特性。举例来说,可用于纳米颗粒组合物中的 RNA 的量可取决于 RNA 的大小、序列和其它特征。纳米颗粒组合物中核酸或核酸组合和其它成分(例如脂质)的相对量还可变化。在一些实施方案中,纳米颗粒组合物中脂质组分与核酸或核酸组合的 wt/wt 比率可为约 5:1 至约 60:1,诸如为约 5:1、
10 6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、22:1、25:1、30:1、35:1、40:1、45:1、50:1 和 60:1。举例来说,脂质组分与核酸或核酸组合的 wt/wt 比率可为约 10:1 至约 40:1。在某些实施方案中,wt/wt 比率为约 20:1。纳米颗粒组合物中核酸或核酸组合的量可例如使用吸收光谱法(例如紫外-可见光谱法)测量。

15 在一些实施方案中,纳米颗粒组合物包含一种或多种 RNA,并且可选择一种或多种 RNA、脂质和其量来提供特定 N:P 比。组合物的 N:P 比是指一种或多种脂质中的氮原子与 RNA 中磷酸酯基的数量的摩尔比。在一些实施方案中,选择较低的 N:P 比。可选择一种或多种 RNA、脂质和其量以提供约 2:1 至约 30:1,诸如 2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、12:1、14:1、16:1、18:1、20:1、22:1、24:1、26:1、28:1 或 30:1 的
20 N:P 比。在某些实施方案中,N:P 比可为约 2:1 至约 8:1。在其它实施方案中,N:P 比为约 5:1 至约 8:1。举例来说,N:P 比可为约 5.0:1、约 5.5:1、约 5.67:1、约 6.0:1、约 6.5:1 或约 7.0:1。举例来说,N:P 比可为约 5.67:1。

纳米颗粒组合物的其他表征方法和多个特征,诸如粒径、多分散性指数和 ζ 电位可以参见例如 PCT 公布第 WO 2023/098842 号,其内容整体援引加入本文。

25 核酸或核酸组合的包封效率描述相对于所提供的初始量,在制备后经纳米颗粒组合物包封或以其它方式与纳米颗粒组合物缔合的核酸或核酸组合的量。包封效率期望较高(例如接近 100%)。包封效率可例如通过比较在用一种或多种有机溶剂或清洁剂破坏纳米颗粒组合物之前与之后含有纳米颗粒组合物的溶液中核酸或核酸组合的量来测量。荧光可用于测量溶液中游离核酸或核酸组合(例如 RNA)的量。对于本文所述的纳米颗粒组
30 合物,核酸或核酸组合的包封效率可为至少 50%,例如 50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 100%。在一些实施方案中,包封效率可为至少 80%。在某些实施方案中,包封效率可为至少 90%。

纳米颗粒组合物可任选地包含一种或多种包衣。举例来说,可将纳米颗粒组合物配制成具有包衣的胶囊、膜片或片剂。包含本文所述组合物的胶囊、膜片或片剂可具有任
35 何有用的大小、抗拉强度、硬度或密度。

在一方面,本发明还提供一种疫苗制剂(又称为“疫苗试剂”)包含本文所述的核酸或核酸组合。

在一些实施方案中,疫苗制剂包含本文所述的组合物。

在一些实施方案中,疫苗制剂包含化合物 01-1、DMG-PEG、DSPC 和胆固醇;其

中，核酸或核酸组合编码本文所述的多肽或多肽组合。在一实施方案中，核酸或核酸组合包含 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列。

5 在一些实施方案中，疫苗制剂包含化合物 01-1、DMG-PEG、DSPC 和胆固醇；其中，核酸或核酸组合编码本文所述的多肽或多肽组合。在一实施方案中，核酸或核酸组合包含两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或者 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列。

10 在一实施方案中，疫苗制剂包含编码本文所述的多肽或多肽组合的核酸或核酸组合以及包封核酸或核酸组合的脂质，脂质包含化合物 01-1、DMG-PEG、DSPC 和胆固醇，核酸或核酸组合包含 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列。

15 在一实施方案中，疫苗制剂包含编码本文所述的多肽或多肽组合的核酸或核酸组合以及包封核酸或核酸组合的脂质，脂质包含化合物 01-1、DMG-PEG、DSPC 和胆固醇，核酸或核酸组合包含两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或者 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列。

20 在一实施方案中，疫苗制剂包含编码本文所述的多肽或多肽组合的核酸或核酸组合以及包封核酸或核酸组合的脂质，脂质包含 50 mol% 化合物 01-1、1.5 mol% DMG-PEG、10 mol% DSPC 和 38.5 mol% 胆固醇，核酸或核酸组合包含 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 中任一个的核苷酸序列。

25 在一实施方案中，疫苗制剂包含编码本文所述的多肽或多肽组合的核酸或核酸组合以及包封核酸或核酸组合的脂质，脂质包含 50 mol% 化合物 01-1、1.5 mol% DMG-PEG、10 mol% DSPC 和 38.5 mol% 胆固醇，核酸或核酸组合包含两个或更多个选自 SEQ ID NO: 1、2、3、4、5、6、7、13、14、15、16、17、18、19、20、21 和 22 或者 SEQ ID NO: 45、46、47、48、49、50、51、57、58、59、60、61、62、63、64、65 和 66 的核苷酸序列。
30

药物组合物/试剂盒

根据本公开，纳米颗粒组合物可整体或部分地配制成药物组合物。药物组合物可包含一种或多种纳米颗粒组合物。举例来说，药物组合物可包含一种或多种纳米颗粒组合物，所述一种或多种纳米颗粒组合物包含一种或多种不同的核酸或核酸组合。药物组合物还可包含一种或多种药学上可接受的载剂、稀释剂、赋形剂或辅助成分，诸如本文所述的那些。关于药物组合物和剂的配制和制造的一般指南可例如在 Remington, The Science and Practice of Pharmacy, 第 21 版, A. R. Gennaro; Lippincott, Williams & Wilkins, Baltimore, Md., 2006 中获得。常规载剂、稀释剂、赋形剂和辅助成分可用于任何药物组合物中，除非任何常规载剂、稀释剂、赋形剂或辅助成分与纳米颗粒组合物的一种或多

35

种组分不相容。如果载剂、稀释剂、赋形剂或辅助成分与纳米颗粒组合物的组分的组合会导致任何不希望的生物效应或其它有害效应，则载剂、稀释剂、赋形剂或辅助成分与纳米颗粒组合物的组分不相容。

5 在一些实施方案中，一种或多种载剂、稀释剂、赋形剂或辅助成分可构成包含纳米颗粒组合物的药物组合物的总质量或体积的大于 50%。举例来说，一种或多种载剂、稀释剂、赋形剂或辅助成分可构成药物组合物的 50%、60%、70%、80%、90%或更高百分比。在一些实施方案中，药学上可接受的载剂、稀释剂、赋形剂为至少 95%、至少 96%、至少 97%、至少 98%、至少 99%或 100%纯。在一些实施方案中，载剂、稀释剂、赋形剂经批准用于人类和兽医用途。在一些实施方案中，载剂、稀释剂、赋形剂得到美国食品与药物管理局批准。在 10 一些实施方案中，载剂、稀释剂、赋形剂是医药级的。在一些实施方案中，载剂、稀释剂、赋形剂符合美国药典(USP)、欧洲药典(EP)、英国药典和/或国际药典的标准。

根据本公开的药物组合物中的一种或多种纳米颗粒组合物、一种或多种药学上可接受的载剂、稀释剂、赋形剂和/或任何额外成分的相对量将取决于所治疗受试者的身份、15 体格和/或状况并且进一步取决于组合物的施用途径而变化。举例来说，药物组合物可包含在 0.1%与 100% (wt/wt)之间的一种或多种纳米颗粒组合物。

在某些实施方案中，本公开的纳米颗粒组合物和/或药物组合物经冷藏或冷冻储存和/或运输，冷藏或冷冻储存和/或运输条件可以参见例如 PCT 公布第 WO 2023/098842 号，其内容整体援引加入本文。

20 在某些实施方案中，本公开的药物组合物包含本文所公开的纳米颗粒组合物和药学上可接受的载剂、稀释剂、赋形剂，所述载剂、稀释剂、赋形剂选自以下中的一种或多种：Tris、乙酸盐(例如乙酸钠)、柠檬酸盐(例如柠檬酸钠)、生理食盐水、PBS 和蔗糖。在某些实施方案中，本公开的药物组合物的 pH 值在约 7 与 8 之间(例如 6.8、6.9、7.0、7.1、7.2、7.3、7.4、7.5、7.6、7.7、7.8、7.9 或 8.0，或在 7.5 与 8 之间或在 7 与 7.8 之 25 间)。举例来说，本公开的药物组合物包含本文所公开的纳米颗粒组合物、Tris、生理食盐水和蔗糖，并且具有约 7.5-8 的 pH 值，其适于在例如约-20°C下储存和/或运输。举例来说，本公开的药物组合物包含本文所公开的纳米颗粒组合物和 PBS，并且具有约 7-7.8 的 pH 值，其适于在例如约 4°C或更低温度下储存和/或运输。在本公开的上下文中，“稳定性”、“稳定化”和“稳定的”是指本文所公开的纳米颗粒组合物和/或药物组合物在给定 30 制造、制备、转运、储存和/或使用条件下，例如当施加应力，诸如剪切力、冷冻/解冻应力等时，对化学或物理变化(例如降解、粒径变化、聚集、包封的变化等)具有抗性。

可将纳米颗粒组合物和/或包含一种或多种纳米颗粒组合物的药物组合物施用于任何患者或受试者，包括可受益于通过将核酸或核酸组合递送至一种或多种特定细胞、组织、器官或系统或其群组的患者或受试者。尽管本文所提供的关于纳米颗粒组合物和包含 35 纳米颗粒组合物的药物组合物的描述主要针对适于施用于人的组合物，但本领域技术人员应理解，此类组合物一般适于施用于任何其它哺乳动物。为了使组合物适于施用于各种动物而对适于施用于人的组合物的改进是众所周知的，并且有普通技术的兽医药理学家长通过普通实验(如果有的话)即可设计和/或进行此类改进。经考虑，施用组合物的受试者包括但不限于人、其它灵长类动物和其它哺乳动物，包括商业上相关的哺乳动物，

诸如牛、猪、马、绵羊、猫、狗、小鼠和/或大鼠。

包含一种或多种纳米颗粒组合物的药物组合物可通过药理学领域中已知或以后将开发的任何方法制备。一般来说，此类制备方法包括使活性成分与载剂、稀释剂、赋形剂和/或一种或多种其它辅助成分结合，并且接着，如果需要或必要，则将产物分成、成型成和/或包装成所需的单剂量或多剂量单元。

根据本公开的药物组合物可以散装、作为单次单位剂量和/或作为多个单次单位剂量制备、包装和/或出售。如本文所用，“单位剂量”是包含预定量的活性成分(例如纳米颗粒组合物)的药物组合物的离散量。活性成分的量一般等于将被施用受试者的活性成分的剂量和/或这个剂量的便利部分，诸如这个剂量的一半或三分之一。

药物组合物可制备成适合多种施用途径和方法的多种形式。举例来说，药物组合物可制备成液体剂型(例如乳液、微乳液、纳米乳液、溶液、悬浮液、糖浆和酏剂)、可注射形式、固体剂型(例如胶囊、片剂、丸剂、粉剂和颗粒剂)、用于局部和/或经皮施用的剂型(例如软膏、糊剂、乳膏、洗剂、凝胶剂、粉剂、溶液、喷雾剂、吸入剂和贴片)、悬浮液、粉剂和其它形式。

经口和胃肠外施用的液体剂型包括但不限于药学上可接受的乳液、微乳液、纳米乳液、溶液、悬浮液、糖浆和/或酏剂。除活性成分外，液体剂型还可包含此项技术中常用的惰性稀释剂，例如水或其它溶剂、增溶剂和乳化剂，诸如乙醇、异丙醇、碳酸乙酯、乙酸乙酯、苯甲醇、苯甲酸苯甲酯、丙二醇、1,3-丁二醇、二甲基甲酰胺、油类(尤其为棉籽油、花生油、玉米油、胚芽油、橄榄油、蓖麻油和芝麻油)、甘油、四氢糠醇、聚乙二醇和脱水山梨醇脂肪酸酯，以及其混合物。除惰性稀释剂外，经口组合物还可包含额外治疗剂和/或预防剂、额外剂，诸如润湿剂、乳化剂和助悬剂、甜味剂、调味剂和/或芳香剂。在供胃肠外施用的某些实施方案中，将组合物与增溶剂混合，所述增溶剂诸如为Cremophor™、醇、油、改性油、二醇、聚山梨醇酯、环糊精、聚合物和/或其组合。

可注射制剂，例如无菌可注射水性或油性悬浮液，可根据已知技术，使用适合分散剂、润湿剂和/或助悬剂来配制。无菌可注射制剂可为在无毒胃肠外可接受的稀释剂和/或溶剂中的无菌可注射溶液、悬浮液和/或乳液，例如在1,3-丁二醇中的溶液。可使用的可接受的媒介物和溶剂包括水、林格氏溶液(U.S.P.)和等渗氯化钠溶液。无菌不挥发性油通常用作溶剂或悬浮介质。出于这个目的，可使用任何温和的不挥发性油，包括合成单酸甘油酯或二酸甘油酯。诸如油酸的脂肪酸可用于制备可注射剂。

可注射制剂可经灭菌，例如通过经由细菌截留过滤器过滤，和/或通过掺入呈无菌固体组合物形式的灭菌剂来灭菌，所述灭菌剂可在使用前溶解或分散于无菌水或其它无菌可注射介质中。

本发明提供向哺乳动物细胞或器官递送核酸或核酸组合，在哺乳动物细胞中产生抗原多肽，以及预防和/或治疗有需要的哺乳动物的疾病或病症的方法，所述方法包括向哺乳动物施用包含核酸或核酸组合的纳米颗粒组合物和/或使哺乳动物细胞与所述纳米颗粒组合物接触。

本公开的多肽/多肽组合、核酸/核酸组合或者组合物还可以以试剂盒的形式提供。因此，本公开还涉及一种试剂盒，其包含本文所述的多肽或多肽组合、核酸或核酸组合或者组合物。试剂盒还可以按需要包含使用说明或者包含合适的容器。在某些实施方案

中，试剂盒还包含给药的装置。试剂盒一般包括表明试剂盒内容物的预期用途和/或使用方法的标签。术语“标签”包括在试剂盒上或与试剂盒一起提供的或以其他方式随试剂盒提供的任何书面的或记录的材料。

治疗应用

5 本发明提供本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物、疫苗制剂、药物组合物或试剂盒，用于预防和/或治疗受试者猴痘病毒感染。

 本发明提供本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物、疫苗制剂、药物组合物或试剂盒在制备用于预防和/或治疗受试者猴痘病毒感染的药物中的用途。

10 本发明提供本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物、疫苗制剂、药物组合物或试剂盒在制备用于预防和/或治疗受试者猴痘病毒感染的疫苗中的用途。

 本发明提供一种预防和/或治疗有需要的受试者猴痘病毒感染的方法，所述方法包括给有需要的受试者施用治疗有效量的本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物、疫苗制剂、药物组合物或试剂盒。在一实施方案中，所述方法包括施用治疗有效量的包含本发明的核酸或核酸组合的组合物。在一实施方案中，所述方法包括施用治疗有效量的包含本发明的 RNA 或 RNA 组合的组合物，特别是包含如本文所述的 LNP 的组合物。在一实施方案中，所述方法包括施用治疗有效量的包含本发明的 mRNA 或 mRNA 组合的组合物，特别是包含如本文所述的 LNP 的组合物。

20 在一些实施方案中，治疗有效量在一次或多次施用中提供。在一些实施方案中，治疗有效量在两次施用中提供。在一些实施方案中，治疗有效量在三次施用中提供。在一些实施方案中，治疗有效量在四次施用中提供。在一些实施方案中，治疗有效量在五次施用中提供。在一些实施方案中，治疗有效量在六次施用中提供。优选地，在一些实施方案中，治疗有效量在两次施用中提供。

25 在一些实施例中，本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物或药物组合物可以通过本领域技术人员已知的任何方法施用受试者，例如肠胃外、经口、经粘膜、经皮、肌肉内、静脉内、皮内、皮下或腹腔内。优选地，本发明的组合物通过肌肉注射施用。

30 **有益效果**

 本发明的多肽或多肽组合、核酸或核酸组合(特别是 RNA)、组合物、药物组合物或者试剂盒可以呈现出优异的效果，例如但不限于：(1) 体外表达效率高；(2) 在受试者中诱导较高水平的针对猴痘病毒抗原蛋白的结合抗体；(3) 在受试者中诱导较高水平的牛痘病毒交叉中和抗体；(4) 在受试者中诱导较高水平的 T 细胞免疫应答。

35

实施例

 通过参考以下实施例进一步描述本发明。应当理解，这些实施例仅作为示例，而不对本发明构成限制。以下材料和仪器均是可商购的或根据本领域公知的方法制备。以下

实验均按照制造商的说明书或根据本领域公知的方法和步骤进行。

实施例 1 mRNA 的制备

1.1. DNA 模板质粒的设计及线性化

- 5设计分别编码不同猴痘病毒抗原多肽单一体(包括目的多肽为 A29L、A35R、B2R、B6R、C15L、M1R 或者 H3L 的抗原多肽单一体)或者抗原多肽多联体(包括目的多肽为 A29L+A35R+B6R+B2R+C15L+M1R 、 B2R+A35R+C15L 、 A29L+M1R+B6R 、 B6R+M1R+H3L 、 B6R+A35R+M1R 、 A29L+B2R+C15L 、 B2R+H3L+C15L 、 B2R+A35R+B6R、 A29L+M1R+C15L 或者 H3L+C15L+ M1R 的抗原多肽多联体)的构建
- 10体的 DNA 开放阅读框(ORF)序列，将各构建体分别命名为 0201、0302、0402、0505、0603、0702、0803、1101、1201、1301、1401、1501、1601、1701、1801、1901、2001；其中各构建体的 DNA 开放阅读框(ORF)序列以及其所编码的氨基酸序列的信息分别如表 2 和表 1 所示。

表 1. 各构建体所编码的氨基酸序列

构建体名称	目的多肽	编码一种抗原多肽 (抗原多肽单一体)	编码多种抗原多肽 (抗原多肽多联体)	编码的全长多肽	编码的全长多肽的氨基酸序列(SEQ ID NO:)
0201	A29L	A29L 肽全长 (SEQ ID NO: 111)	/	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全长(SEQ ID NO: 111)+HA 多肽标签 (SEQ ID NO: 135)	89
0302	A35R	A35R 肽全长 (SEQ ID NO: 114)	/	A35R 肽全长(SEQ ID NO: 114)+HA 多肽标签(SEQ ID NO: 136)	90
0402	B2R	B2R 信号肽(SEQ ID NO: 116)、B2R 肽全长(除信号肽) (SEQ ID NO: 117)	/	B2R 信号肽(SEQ ID NO: 116)+B2R 肽全长(除信号肽) (SEQ ID NO: 117)+ HA 多肽标签(SEQ ID NO: 136)	91
0505	B6R	B6R 信号肽(SEQ ID NO: 121)、B6R 肽全长(除信号肽) (SEQ ID NO: 122)	/	B6R 信号肽(SEQ ID NO: 121)+GS 接头 (SEQ ID NO: 137)+HA 多肽标签(SEQ ID NO: 136)+B6R 肽全长(除信号肽) (SEQ ID NO: 122)	92

0603	C15L	C15L 肽全长 (SEQ ID NO: 126)	/	IgE 信号肽(SEQ ID NO: 134)+GS 接头 (SEQ ID NO: 137)+HA 多肽标签(SEQ ID NO: 136)+C15L 肽全长 (SEQ ID NO: 126)	93
0702	M1R	M1R 胞外段 (SEQ ID NO: 128)	/	IgE 信号肽(SEQ ID NO: 134)+GS 接头 (SEQ ID NO: 137)+HA 多肽标签(SEQ ID NO: 136)+ M1R 胞外段 (SEQ ID NO: 128)	94
0803	H3L	H3L 肽全长 (SEQ ID NO: 131)	/	IgE 信号肽(SEQ ID NO: 134)+GS 接头 (SEQ ID NO: 137)+HA 多肽标签(SEQ ID NO: 136)+H3L 肽全长(SEQ ID NO: 131)	95
0201-J0	A29L	A29L 肽全长 (SEQ ID NO: 112)	/	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全 长(SEQ ID NO: 112)	96
0301-J0	A35R	A35R 肽全长 (SEQ ID NO: 114)	/	IgE 信号肽(SEQ ID NO: 134)+A35R 肽全 长(SEQ ID NO: 114)	97
0501-J0	B6R	B6R 信号肽(SEQ ID NO: 121)、B6R 肽全长(除信号肽) (SEQ ID NO: 122)	/	IgE 信号肽(SEQ ID NO: 134)+B6R 信号肽 (SEQ ID NO: 121)+B6R 肽全长(除信 号肽) (SEQ ID NO: 122)	98
0701-J0	M1R	M1R 肽全长 (SEQ ID NO: 129)	/	IgE 信号肽(SEQ ID NO: 134)+M1R 肽全长 (SEQ ID NO: 129)	99
0901-J0	E8L	E8L 肽全长 (SEQ ID NO: 133)	/	IgE 信号肽(SEQ ID NO: 134)+E8L 肽全长 (SEQ ID NO: 133)	100

1101	A29L+A35R+B6R+B2R+C15L+M1R	/	A29L 肽全长(SEQ ID NO: 113)、A35R 胞外段(SEQ ID NO: 115)、B6R 胞外段(SEQ ID NO: 124)、B2R 胞外段(SEQ ID NO: 119)、C15L 胞外段(SEQ ID NO: 127)、M1R 肽全长(SEQ ID NO: 130)	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全长(SEQ ID NO: 113)+GS 接头(SEQ ID NO: 138)+ A35R 胞外段(SEQ ID NO: 115)+GS 接头(SEQ ID NO: 138)+B6R 胞外段(SEQ ID NO: 124)+GS 接头(SEQ ID NO: 138)+B2R 胞外段(SEQ ID NO: 119)+ GS 接头(SEQ ID NO: 138)+ C15L 胞外段(SEQ ID NO: 127)+GS 接头(SEQ ID NO: 138)+M1R 肽全长(SEQ ID NO: 130)	101
1201	B2R+A35R+C15L	/	B2R 信号肽(SEQ ID NO: 116)、B2R 胞外段(SEQ ID NO: 119)、A35R 胞外段(SEQ ID NO: 115)、C15L 肽全长(SEQ ID NO: 126)	B2R 信号肽(SEQ ID NO: 116)+B2R 胞外段(SEQ ID NO: 119)+GS 接头(SEQ ID NO: 138)+ A35R 胞外段(SEQ ID NO: 115)+GS 接头(SEQ ID NO: 138)+ C15L 肽全长(SEQ ID NO: 126)	102
1301	A29L+M1R+B6R	/	A29L 肽全长(SEQ ID NO: 113)、M1R 胞外段(SEQ ID NO: 128)、B6R 肽全长(除信号肽)(SEQ ID NO: 122)	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全长(SEQ ID NO: 113)+GS 接头(SEQ ID NO: 138)+ M1R 胞外段(SEQ ID NO: 128)+GS 接头(SEQ ID NO: 138)+ B6R 肽全长(除信号肽)(SEQ ID NO: 122)	103

1401	B6R+M1R+ H3L	/	B6R 信号肽(SEQ ID NO: 121)、B6R 胞外段(SEQ ID NO: 124)、M1R 胞外段(SEQ ID NO: 128)、H3L 肽全长(SEQ ID NO: 131)	B6R 信号肽(SEQ ID NO: 121)+B6R 胞外段(SEQ ID NO: 124)+GS 接头(SEQ ID NO: 138)+ M1R 胞外段(SEQ ID NO: 128)+GS 接头(SEQ ID NO: 138)+ H3L 肽全长(SEQ ID NO: 131)	104
1501	B6R+A35R +M1R	/	B6R 信号肽(SEQ ID NO: 121)、B6R 胞外段(SEQ ID NO: 124)、A35R 胞外段(SEQ ID NO: 115)、M1R 肽全长(SEQ ID NO: 130)	B6R 信号肽(SEQ ID NO: 121)+B6R 胞外段(SEQ ID NO: 124)+GS 接头(SEQ ID NO: 138)+ A35R 胞外段(SEQ ID NO: 115)+GS 接头(SEQ ID NO: 138)+ M1R 肽全长(SEQ ID NO: 130)	105
1601	A29L+B2R +C15L	/	A29L 肽全长(SEQ ID NO: 113)、B2R 胞外段(SEQ ID NO: 119)、C15L 肽全长(SEQ ID NO: 126)	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全长(SEQ ID NO: 113)+GS 接头(SEQ ID NO: 138)+B2R 胞外段(SEQ ID NO: 119)+ GS 接头(SEQ ID NO: 138)+ C15L 肽全长(SEQ ID NO: 126)	106
1701	B2R+H3L+ C15L	/	B2R 信号肽(SEQ ID NO: 116)、B2R 胞外段(SEQ ID NO: 119)、H3L 胞外段(SEQ ID NO: 132)、C15L 肽全长(SEQ ID NO: 126)	B2R 信号肽(SEQ ID NO: 116)+B2R 胞外段(SEQ ID NO: 119)+GS 接头(SEQ ID NO: 138)+H3L 胞外段(SEQ ID NO: 132)+ GS 接头(SEQ ID NO: 138)+ C15L 肽全长(SEQ ID NO: 126)	107

1801	B2R+A35R +B6R	/	B2R 信号肽(SEQ ID NO: 116)、B2R 胞外段(SEQ ID NO: 119)、A35R 胞外段(SEQ ID NO: 115)、B6R 肽全长(除信号肽)(SEQ ID NO: 122)	B2R 信号肽(SEQ ID NO: 116)+B2R 胞外段(SEQ ID NO: 119)+GS 接头(SEQ ID NO: 138)+ A35R 胞外段(SEQ ID NO: 115)+GS 接头(SEQ ID NO: 138)+ B6R 肽全长(除信号肽)(SEQ ID NO: 122)	108
1901	A29L+M1R +C15L	/	A29L 肽全长(SEQ ID NO: 113)、M1R 胞外段(SEQ ID NO: 128)、C15L 肽全长(SEQ ID NO: 126)	IgE 信号肽(SEQ ID NO: 134)+A29L 肽全长(SEQ ID NO: 113)+GS 接头(SEQ ID NO: 138)+ M1R 胞外段(SEQ ID NO: 128)+GS 接头(SEQ ID NO: 138)+ C15L 肽全长(SEQ ID NO: 126)	109
2001	H3L+C15L + M1R	/	H3L 胞外段(SEQ ID NO: 132)、C15L 胞外段(SEQ ID NO: 127)、M1R 肽全长(SEQ ID NO: 130)	IgE 信号肽(SEQ ID NO: 134)+ H3L 胞外段(SEQ ID NO: 132)+GS 接头(SEQ ID NO: 138)+ C15L 胞外段(SEQ ID NO: 127)+GS 接头(SEQ ID NO: 138)+ M1R 肽全长(SEQ ID NO: 130)	110

表 2. 各构建体的核酸序列(SEQ ID NO:)

构建体名称	RNA(ORF)	mRNA	DNA(ORF)	DNA
0201	1	45	23	67
0302	2	46	24	68
0402	3	47	25	69
0505	4	48	26	70
0603	5	49	27	71
0702	6	50	28	72
0803	7	51	29	73
0201-J0	8	52	30	74
0301-J0	9	53	31	75
0501-J0	10	54	32	76
0701-J0	11	55	33	77

0901-J0	12	56	34	78
1101	13	57	35	79
1201	14	58	36	80
1301	15	59	37	81
1401	16	60	38	82
1501	17	61	39	83
1601	18	62	40	84
1701	19	63	41	85
1801	20	64	42	86
1901	21	65	43	87
2001	22	66	44	88

将 5'-UTR 序列(对应的 DNA 序列如 SEQ ID NO: 142 或 150 所示)、DNA ORF、3'-UTR 序列(对应的 DNA 序列如 SEQ ID NO: 143、144 或 151 所示)插入用于 IVT 的质粒 pJ241 (内部开发, 含有卡那霉素(kanamycin)抗性基因、T7 启动子序列、poly(A)标记(对应的 DNA 序列如 SEQ ID NO: 146 所示)和在 poly(A)序列下游的独特 IIS 型限制位点), 进行 IVT 载体构建。

使用 IIS 型限制酶 Esp3I/BsmBI (Thermo, FD0294)或 Bsa I(恺佶, GMP-BSA-EE101-11)消化进行线性化。具体地, 将每 10 µg 质粒与 10 U 的限制酶混合, 在 37°C 孵育 4 小时以确保完全线性化。通过添加 1/10 体积的 3 M 乙酸钠(pH 5.5)和 2.5 倍体积的乙醇终止反应, 充分混合, 并在-20°C 冷却 1 小时。4°C, 13800 g 离心 15 分钟, 以沉淀线性化的 DNA, 用 70%乙醇洗涤两次, 重悬于无核酸酶 H₂O 中。

1.2. 从 DNA 模板体外转录 mRNA

表 3 示出 20 µL 体外转录反应体系, 根据此反应体系制备反应混合物进行体外转录。其中 UTP 采用 1-甲基假尿苷修饰。

表 3. 20 µL 体外转录反应体系

试剂		供应商	货号	用量
无核酸酶 H ₂ O		Invitrogen	10977-015	补足至 20 µL
RNA 酶抑制剂 (40U/µl)		恺佶	GMP-RNI-ME101-12	0.5 µL
rNTP Mix	A	武汉糖智	MR-1001	8 µL (各自最终 10 mM)
	U	武汉糖智	MR-1004	
	N1-Me-pUTP	武汉糖智	MR-3002	
	C	武汉糖智	MR-1003	
	G	武汉糖智	MR-1002	
10xIVT 反应缓冲液		恺佶	GMP-T7P-EE101-A	2 µL
1M MgCl ₂		Invitrogen	AM9530G	0.8 µL
0.1M DTT		Sigma	43816-10ML	2 µL
100U/mL 无机焦磷酸酶		恺佶	GMP-PYR-YE101-12	0.8 µL
100mM NaCl		Invitrogen	AM9760G	1 µL

T7RNA 聚合酶 (50U/μL)		恺佻	GMP-T7P-EE101-12	2 μL
线性化 DNA				1 μg

体外转录：将反应混合物在 37℃ 孵育 6 小时，然后添加 1 μL DNA 酶 I (无 RNA 酶，1 U/μL，恺佻，GMP-DNI-EE001-12)以除去 DNA 模板，在 37℃ 孵育 30 分钟。通过添加 0.5 体积的 7.5 M LiCl、50 mM EDTA 并在-20℃ 孵育 45 分钟，然后在 4℃，13800 g 离心 15 分钟沉淀 mRNA，以纯化合成的 RNA。除去上清液，并将沉淀物用 500 μL 的冰冷 70%乙醇冲洗两次。冲洗后，将 mRNA 重悬于无核酸酶 H₂O 中，将浓度调节至 1 mg/mL，并在-20℃ 储存。

mRNA 加帽：将每 10 μg 未加帽的 mRNA 在 65℃ 加热 10 分钟，在冰上放置 5 分钟，然后与 10 U 牛痘病毒加帽酶(恺佻，GMP-VCS-VE101-13)、50 U mRNA 帽 2'-O-甲基转移酶(恺佻，GMP-MEH-VE101-13)、0.2 mM SAM、0.5 mM GTP 和 1 U RNA 酶抑制剂混合，并在 37℃下孵育 60 分钟以生成 Cap 1 修饰结构。如前所述将经修饰的 mRNA 通过 LiCl 沉淀，并将 RNA 重悬于无核酸酶 H₂O 中，并在-20℃下储存。

HPLC 纯化：使用 C4 柱(5 μm)(10 mm×250 mm 柱)，通过高效液相色谱(HPLC)纯化 RNA(数据未示出)。缓冲液 A 含有 0.1 M 乙酸三乙铵(TEAA) (pH = 7.0)，并且缓冲液 B 含有 0.1 M TEAA (pH = 7.0)和 25%乙腈。

通过上述体外转录、加帽和纯化过程获得如表 2 所示的 mRNA 序列。

并且，通过上述方法制备分别编码 A29L、A35R、B6R、M1R、E8L 抗原多肽的构建体，将其分别命名为 0201-J0、0301-J0、0501-J0、0701-J0、0901-J0。各构建体的核酸序列以及其所编码的氨基酸序列的信息分别如表 2 和表 1 所示。

20

实施例 2 体外转染

本实施例使用 Lipofectamine 2000 转染试剂将实施例 1 中制备的编码不同猴痘病毒抗原多肽单一体或者抗原多肽多联体的 mRNA 转染至 HEK293T 细胞，并通过流式细胞术评估不同 mRNA 构建体的体外表达情况。具体体外转染方法和流式细胞检测方法如下所述。

体外转染：将 HEK293T 细胞(中科院上海细胞库，SCSP-502)以 1.5×10⁵ 个细胞/孔接种于 24 孔细胞培养板板中。于 37℃，5% CO₂ 过夜培养细胞，待细胞汇合度达到 70%-90%后进行转染。转染的示例性方法为：向 25 μL 不含血清的 Opti-MEM 培养基中加入 1 μg mRNA，吹打混匀后作为混合物 A。再另取 25 μL 不含血清的 Opti-MEM 培养基，向其中加入 1 μL Lipofectamine 2000 转染试剂，充分混匀后作为混合物 B。将混合物 A 加入混合物 B 中，轻柔混匀后室温静置 5 分钟。最后将混合液逐滴加入 HEK293T 细胞中，于 37℃，5% CO₂ 培养 18-24 小时。其中，每种 mRNA 转染 2 个复孔。另外准备只含 Lipofectamine 2000 不含 mRNA 的混合液以相同的条件转染 HEK293T 细胞并在相同条件下培养，以此作为空白对照。

收样：吸弃细胞培养上清，用 200 μL 的染色缓冲液(含有 2% FBS 的 PBS)将 HEK239T 细胞轻轻吹打下来并重悬混匀，转移至 96 孔板中。对于转染了编码抗原多肽多联体的

- mRNA 分子的细胞，每个 24 孔板中一个孔的细胞量应均分入 96 孔板中相应个数的孔内，例如，转染了 1101 mRNA(编码 6 个抗原多肽)的细胞应均分入 96 孔板中的 6 个孔内，转染了 1201、1301、1401、1501、1601、1701、1801、1901 或者 2001 mRNA(都分别编码 3 个抗原多肽)的细胞应均分入 96 孔板中的 3 个孔内。转移细胞后，300g 离心，
- 5 弃去上清，再用 300 μL 染色缓冲液洗细胞两遍。

染色：根据每种 mRNA 编码的抗原多肽选择相应的一抗二抗进行染色，以检测 mRNA 是否能在体外成功表达抗原多肽。表 4-1 示出所用的染色抗体。表 4-2 示出所用的染色抗体的名称、供应商和货号。

表 4-1. 流式细胞术所用的染色抗体

构建体名称	编码多肽	染色一抗	一抗浓度	染色二抗	二抗稀释度
0201	A29L	鼠抗 A29L	5 μg/mL	抗鼠-FITC	1:100
0302	A35R	人抗 A35R	2 μg/mL	抗人-PE	1:50
0402	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100
0505	B6R	人抗 B6R	2 μg/mL	抗人-PE	1:50
0603	C15L	兔抗 C15L	5 μg/mL	抗兔-FITC	1:100
0702	M1R	鼠抗 M1R	5 μg/mL	抗鼠-FITC	1:100
0803	H3L	鼠抗 H3L	5 μg/mL	抗鼠-FITC	1:100
1101	A29L	鼠抗 A29L	5 μg/mL	抗鼠-FITC	1:100
	A35R	人抗 A35R	2 μg/mL	抗人-PE	1:50
	B6R	人抗 B6R	2 μg/mL	抗人-PE	1:50
	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100
	C15L	兔抗 C15L	5 μg/mL	抗兔-FITC	1:100
	M1R	鼠抗 M1R	5 μg/mL	抗鼠-FITC	1:100
1201	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100
	A35R	人抗 A35R	2 μg/mL	抗人-PE	1:50
	C15L	兔抗 C15L	5 μg/mL	抗兔-FITC	1:100
1301	A29L	鼠抗 A29L	5 μg/mL	抗鼠-FITC	1:100
	M1R	鼠抗 M1R	5 μg/mL	抗鼠-FITC	1:100
	B6R	人抗 B6R	2 μg/mL	抗人-PE	1:50
1401	B6R	人抗 B6R	2 μg/mL	抗人-PE	1:50
	M1R	鼠抗 M1R	5 μg/mL	抗鼠-FITC	1:100
	H3L	鼠抗 H3L	5 μg/mL	抗鼠-FITC	1:100
1501	B6R	人抗 B6R	2 μg/mL	抗人-PE	1:50
	A35R	人抗 A35R	2 μg/mL	抗人-PE	1:50
	M1R	鼠抗 M1R	5 μg/mL	抗鼠-FITC	1:100
1601	A29L	鼠抗 A29L	5 μg/mL	抗鼠-FITC	1:100
	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100
	C15L	兔抗 C15L	5 μg/mL	抗兔-FITC	1:100
1701	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100
	H3L	鼠抗 H3L	5 μg/mL	抗鼠-FITC	1:100
	C15L	兔抗 C15L	5 μg/mL	抗兔-FITC	1:100
1801	B2R	兔抗 B2R	5 μg/mL	抗兔-FITC	1:100

	A35R	人抗 A35R	2 µg/mL	抗人-PE	1:50
	B6R	人抗 B6R	2 µg/mL	抗人-PE	1:50
1901	A29L	鼠抗 A29L	5 µg/mL	抗鼠-FITC	1:100
	M1R	鼠抗 M1R	5 µg/mL	抗鼠-FITC	1:100
	C15L	兔抗 C15L	5 µg/mL	抗兔-FITC	1:100
2001	H3L	鼠抗 H3L	5 µg/mL	抗鼠-FITC	1:100
	C15L	兔抗 C15L	5 µg/mL	抗兔-FITC	1:100
	M1R	鼠抗 M1R	5 µg/mL	抗鼠-FITC	1:100
空白对照	—	鼠抗 A29L	5 µg/mL	抗鼠-FITC	1:100
		人抗 A35R	2 µg/mL	抗人-PE	1:50
		兔抗 B2R	5 µg/mL	抗兔-FITC	1:100
		人抗 B6R	2 µg/mL	抗人-PE	1:50
		兔抗 C15L	5 µg/mL	抗兔-FITC	1:100
		鼠抗 M1R	5 µg/mL	抗鼠-FITC	1:100
		鼠抗 H3L	5 µg/mL	抗鼠-FITC	1:100

表 4-2. 流式细胞术所用染色抗体的抗体名称、供应商和货号

染色一抗	抗体名称	厂商	货号
鼠抗 A29L	抗猴痘病毒/MPXV A29L 抗体(SAA0286)	Antibody System	RVV13401
人抗 A35R	抗 MPXV A35R 单克隆抗体	近岸蛋白	NC083
兔抗 B2R	抗猴痘病毒/MPXV B2R/HA/血凝素多克隆抗体	Antibody System	PVV12501
人抗 B6R	抗 MPXV B6R 单克隆抗体	近岸蛋白	NC082
兔抗 C15L	兔抗 MPXV C15L 多抗	擎科生物	NA (定制产品)
鼠抗 M1R	抗猴痘病毒/MPXV M1R 抗体 (SAA0283)	Antibody System	RVV13301
鼠抗 H3L	抗猴痘病毒/MPXV H3L 抗体(SAA0282)	Antibody System	RVV12701
染色二抗	抗体名称	厂商	货号
抗人-PE	PE 抗人 IgG Fc	Biolegend	410708
抗鼠-FITC	荧光素 (FITC)-缀合的亲纯化山羊抗小鼠 IgG (亚类 1+2a+2b+3), 与 Fc+片段特异性反应(经人, 牛, 兔血清蛋白吸附)	Jacson	155-095-164
抗兔-FITC	FITC 驴抗兔 IgG	Biolegend	406403

5 将一抗用染色缓冲液稀释至工作浓度后重悬细胞，50 µL/孔，4℃ 避光染色 30 分钟。孵育结束后，300g 离心，弃去一抗稀释液，用染色缓冲液洗细胞三遍，弃上清。再用染色缓冲液将二抗稀释至工作浓度后重悬细胞，50 µL/孔，4℃ 避光染色 30 分钟。孵育结束后，300 g 离心，弃去二抗稀释液，用染色缓冲液洗细胞三遍，弃上清。最后用 200 µL/孔的染色缓冲液重悬细胞，使用流式细胞仪(BD biosciences, LSRFortessa)检测样

品。

流式细胞术检测体外表达情况：以空白对照组细胞作为基准画门，检测成功表达目的多肽的 HEK293T 细胞数占总细胞数的比值。

流式细胞术结果如表 5 所示(数据示为两个复孔表达阳性率的平均值)。除 0201、1101 和 1301 外，其余 mRNA 均可在体外成功表达目的多肽。其中，0302、1201 和 1401 的体外表达水平较高；0402、0505、0603、0702、0803、1601、1701 和 1901 体外表达水平中等，1501、1801 和 2001 体外表达水平较低。虽未在此实施例中检测到 0201 和 1101 的目的多肽的成功表达，但在后续实验将 mRNA 包裹入脂质体制备成 mRNA 疫苗后免疫小鼠，可在受免小鼠血清中检测到相应抗原多肽的特异性结合抗体，提示在此实施例中未检测到目的多肽表达并不一定代表 mRNA 分子不表达，而是可能由于检测方法的灵敏度不足或其他原因导致的假阴性。而 1301 既未在体外水平检测到目的蛋白的表达，也未在受免小鼠血清中检测到抗原蛋白的特异性结合抗体，提示 1301 无法成功表达目的蛋白。

表 5. 体外表达结果

构建体名称	编码多肽	A29L	A35R	B2R	B6R	C15L	M1R	H3L
0201	A29L	0.1	—	—	—	—	—	—
0302	A35R	—	95.3	—	—	—	—	—
0402	B2R	—	—	35.9	—	—	—	—
0505	B6R	—	—	—	49.3	—	—	—
0603	C15L	—	—	—	—	28.6	—	—
0702	M1R	—	—	—	—	—	27.7	—
0803	H3L	—	—	—	—	—	—	17.2
1101	A29L-A35R-B6R-B2R-C15L-M1R	0.1	0.6	1.3	0.3	0.3	0.2	—
1201	B2R-A35R-C15L	—	60.2	13.3	—	17.9	—	—
1301	A29L-M1R-B6R	0.1	—	—	0.3	—	0.2	—
1401	B6R-M1R-H3L	—	—	—	90.3	—	33.1	6.5
1501	B6R-A35R-M1R	—	7.1	—	4.2	—	1.5	—
1601	A29L-B2R-C15L	4.4	—	23.9	—	36.3	—	—
1701	B2R-H3L-C15L	—	—	18.3	—	28.3	—	5.0
1801	B2R-A35R-B6R	—	2.5	2.1	2.3	—	—	—
1901	A29L-M1R-C15L	3.2	—	—	—	19.3	11.7	—
2001	H3L-C15L-M1R	—	—	—	—	1.6	1.3	0.6
空白对照组	—	0.2	0.8	1.7	0.3	0.3	0.3	0.3

“—” 表示不适用或未检测

实施例 3 mRNA-LNP 制剂的制备

3.1. 实验材料

系列 01 的脂质参见专利 WO2021204175 中描述的方法进行合成；系列 02 的脂质参见专利 WO 2023/138611 中描述的方法进行合成；系列 03 的脂质参见专利 WO2022152109 中描述的方法进行合成；系列 04 的脂质参见专利 WO2022247755 中描

述的方法进行合成。

DSPC、胆固醇、mPEG2000-DMG (即 DMG-PEG 2000)通过常规商用渠道采购。

3.2. 脂质纳米颗粒的制备的示例性方法

简言之，将阳离子脂质(如 Compound 01-1)、DSPC、胆固醇和 PEG-脂质以 50:10:38.5:1.5 的摩尔比溶解在乙醇中，并将实施例 1 中制备的 mRNA 根据表 6 示出的单独或组合方式稀释在 10 至 50 mM 柠檬酸盐缓冲液(pH = 4)中。通过使用总流速在 9-30 mL/min 范围内的微流体装置(自研设备)将乙醇脂质溶液与 mRNA 水溶液以 1:3 的体积比混合，以大约 10:1 至 30:1 的总脂质与 mRNA 重量比制备 LNP。使用透析移除乙醇且以 DPBS 替代。最后，将脂质纳米颗粒通过 0.2 μm 无菌过滤器过滤。将最终获得的 mRNA-LNP 制剂(包含 mRNA 的脂质纳米颗粒)分别命名为 mRNA-LNP 制剂 G1、G2、G3、G4、G5、G6、G7、G8、G9 和 G10。其中，G10 所编码的抗原多肽与 Zhang RR et al., Rational development of multicomponent mRNA vaccine candidates against mpox. Emerg Microbes Infect. 2023 Dec;12(1):2192815.中的 mRNA 疫苗所编码的抗原多肽一致，因此，将 G10 作为后续实验的 mRNA 疫苗参照品。

表 6. 各 mRNA-LNP 制剂的组成

制剂名称	mRNA 或 mRNA 组合	备注
G1	0201 + 0302 + 0402 + 0505 + 0603 +0702	各组分与 G4-G6 组编码的抗原多肽保持等摩尔比
G2	0302 + 0402 +0505 + 0603 + 0702 +0803	各组分与 G7-G9 组编码的抗原多肽保持等摩尔比
G3	1101	—
G4	1201 +1301	等质量比混合
G5	1501 +1601	等质量比混合
G6	1801 + 1901	等质量比混合
G7	1201 + 1401	等质量比混合
G8	1501 +1701	等质量比混合
G9	2001 + 1801	等质量比混合
G10	0201-J0 + 0301-J0 + 0501-J0 + 0701-J0 + 0901-J0	等质量比混合

实施例 4 候选疫苗的免疫原性评估

本实施例对实施例 3 制备的 mRNA-LNP 制剂(疫苗) G1、G2、G3、G4、G5、G6、G7、G8、G9、G10 进行免疫原性评估。

本实施例将 55 只 6-8 周龄的雌性 BALB/c 小鼠(浙江维通利华)按照体重随机分为 11 组(以所给药的制剂命名的 G1、G2、G3、G4、G5、G6、G7、G8、G9 和 G10 组，以及

给药 PBS 作为安慰剂的 G11 组), 每组 5 只, 按照表 7 中的给药设计, 通过肌肉注射的方式向每只小鼠体内注射总剂量为 2 μg 的猴痘病毒 mRNA 疫苗或安慰剂, 50 μL/只。在初次免疫后第 3 周, 即初免后第 21 天进行一次等剂量的加强免疫。于初免后第 28 天对受试小鼠进行眼眶采血。分离血清, 并将血清置于 56°C 灭活 30 min, 冻存于-80°C 备用。

表 7. 给药设计表

组别	包含的 mRNA 或 mRNA 组合	总剂量 (μg/只)	各组分剂量 (μg/只)
G1	0201 + 0302 + 0402 + 0505 + 0603 + 0702	2	0.2+0.3+0.4+0.4+0.4+0.3
G2	0302 + 0402 + 0505 + 0603 + 0702 + 0803	2	0.26+0.37+0.38+0.30+0.28+0.4
G3	1101	2	2
G4	1201 + 1301	2	1
G5	1501 + 1601	2	1
G6	1801 + 1901	2	1
G7	1201 + 1401	2	1
G8	1501 + 1701	2	1
G9	2001 + 1801	2	1
G10	0201-J0 + 0301-J0 + 0501-J0 + 0701-J0 + 0901-J0	2	0.4
G11	PBS	—	—

4.1 疫苗诱导的血清结合抗体滴度的检测

用上述收集的小鼠血清检测疫苗诱导的血清结合抗体滴度, 其具体检测方法如下。

10 包被: 使用 1×PBS 缓冲液将各猴痘抗原蛋白(A29L 抗原蛋白: 近岸蛋白, DRA207; A35R 抗原蛋白: 近岸蛋白, DRA209; B2R 抗原蛋白: Antibody System, EVV12501; B6R 抗原蛋白: 近岸蛋白, DRA211; C15L 抗原蛋白: 擎科生物定制蛋白; M1R 抗原蛋白: 近岸蛋白, DRA210; H3L 抗原蛋白: Antibody System, EVV12701; E8L 抗原蛋白: Antibody System, YVV13201)分别稀释至 1 μg/mL 并加入 96 孔酶标板中, 100 μL/孔,

15 贴上封板膜, 4°C 过夜包被。包被完成后, 弃去孔内液体, 用 PBST(300 μL/孔)洗板 3 次。

封闭: 向孔内加入 5% BSA 用于封闭非特异性结合位点, 100 μL/孔, 贴上封板膜, 37°C 孵育 1 h。封闭结束后, 弃去孔内液体, 用 PBST(300 μL/孔)洗板 3 次。

加入待测样品: 用 5% BSA 将热灭活的受试小鼠血清从适当的稀释度开始, 三倍倍比稀释 4 个浓度梯度, 将稀释好的血清加入封闭完成的酶标板中, 100 μL/孔, 贴上封板

20 膜, 37°C 孵育 1 h。孵育结束后, 弃去孔内液体, 用 PBST(300 μL/孔)洗板 5 次。

孵育二抗: 用 5% BSA 将 Peroxidase-conjugated Affini Pure Donkey Anti-Mouse IgG

(H+L)(Jackson, 715-035-151)稀释 5 万倍并加入酶标板中, 100 μL/孔, 贴上封板膜, 37°C 孵育 1 h。孵育结束后, 弃去孔内液体, 用 PBST(300 μL/孔)洗板 5 次。

显色与终止: 向孔内加入 TMB 显色液, 100 μL/孔, 室温避光孵育 5 min。孵育结束后, 向孔内加入 ELISA 终止液(Solarbio, C1058), 100 μL/孔, 混匀后用多功能酶标仪

5 (Molecular Devices, ID5)读取 450 nm 波长下的吸光值。

结果判定: 以 PBS 组(G11)小鼠血清吸光值平均数的 2.1 倍作为临界值, 选取 OD₄₅₀ 读值大于临界值的最高稀释倍数, 根据以下公式计算各组小鼠血清中的结合抗体滴度: 血清结合抗体滴度(IgG titer)= 450 nm 处 OD 读值/临界值 × 血清稀释倍数。对于在 1:100 倍稀释条件下 OD₄₅₀ 读值仍低于临界值的组别, 人工定义其血清结合抗体滴度为起始稀

10 释度的一半, 即 50, 意味着在小鼠血清中未检测到针对包被抗原蛋白的特异性结合抗体滴度。

试验结果如表 8 所示, 除 G4 和 G8 组小鼠未产生针对部分抗原的特异性结合抗体之外, 其他组小鼠均可针对接种的疫苗所编码的抗原蛋白产生特异性结合抗体。

在体外表达试验(实施例 2)中, 未检测到 1301 mRNA 分子编码的 A29L、M1R 和

15 B6R 抗原多肽的成功表达, 同样在免疫有 1301 组分的 G4 组小鼠血清中, 也未检测到 A29L、M1R 和 B6R 蛋白的特异性结合抗体产生, 合理推测是由于 1301 mRNA 不表达所致。然而, 虽然在体外表达试验(实施例 2)中, 未检测到 0201、1101 mRNA 分子的成功表达, 但在免疫有 0201 或 1101 组分的小鼠血清中均可检测到针对 0201 或 1101 编码的各抗原多肽的结合抗体滴度, 提示 0201、1101 mRNA 分子成功表达但表达水平较低,

20 处于试验方法的检测限水平以下。

此外, 在所有组别中, G6和G7组小鼠血清中的M1R抗原特异性结合抗体滴度最高。并且, G7组小鼠血清具有最高水平的B6R结合抗体滴度。

表 8. 受试小鼠的血清结合抗体滴度

组别	受试小鼠血清中针对抗原多肽的特异性结合抗体滴度							
	A29L	A35R	B2R	B6R	C15L	M1R	H3L	E8L
G1	1897	400830	401541	51322	220048	385276	—	—
G2	—	416877	519825	26757	208430	458519	76099	—
G3	3979	13011	19583	6763	117417	13453	—	—
G4	50	42885	528079	50	313817	50	—	—
G5	29001	324	763742	69573	995947	41525	—	—
G6	12910	45044	349956	248	628080	1111940	—	—
G7	—	17427	302889	681074	133933	856499	50905	—
G8	—	50	887373	4936	758921	35196	113083	—

G9	—	10022	300280	549	359878	47304	13636	—
G10	28187	114777	—	121940	—	606369	—	114123
G11	50	50	50	50	50	50	50	50

“—” 表示未检测(因为疫苗中包含的 mRNA 不编码该抗原多肽)

4.2 疫苗诱导的血清中和抗体效价的检测

用收集的小鼠血清检测疫苗诱导的血清中和抗体效价，其具体检测方法如下。

5 准备细胞：将 Vero 细胞(中科院上海细胞库, SCSP-520)接种至 96 孔细胞培养板中，
2×10⁴ 个细胞/孔，使用维持培养基(含有 2% FBS 的 DMEM)，于 37°C、5% CO₂ 过夜培
养细胞 8-16 小时，待第二日细胞汇合度达到 90%-95%时，将其用于中和抗体的检测。

中和病毒：另取无菌的 96 孔板若干，将热灭活过的血清样本用维持培养基从适当
的稀释度开始，三倍倍比稀释 6 个浓度梯度；再将牛痘(Vaccinia)病毒(Elstree 株)(ATCC，
10 VR-1549)用维持培养基稀释至 2000 PFU/mL，随后与稀释好的血清样本等体积混合。每
个血清稀释度设置 2 个复孔，同时设置不含血清样本的病毒感染孔作为病毒对照组(12
个复孔)，设置维持培养基孔作为细胞对照组(12 个复孔)，将以上所有混合液或稀释液置
于 37°C 孵育 1 小时。

感染细胞：吸弃细胞培养上清，取 100 μl 孵育完毕的病毒/血清混合液、病毒稀释液
15 或单纯维持培养基加入到含有 Vero 细胞的 96 孔板中，在 37°C 感染 1 小时。在感染期
间，每隔 15 分钟摇晃一次细胞培养板，以使病毒均匀感染细胞。感染结束后，吸弃上
清，向孔内加入已预热至 37°C 的半固体培养基(含有 1.8%羧甲基纤维素钠和 2%FBS 的
DMEM 培养基)，200 μL/孔，使细胞于 37°C、5% CO₂ 培养 15-17 小时。

固定通透：培养结束后，吸弃细胞上清，使用 PBS 缓冲液洗细胞一遍(200μL/孔)，
20 再向细胞中加入 4%组织固定液，50 μL/孔，室温作用 30 分钟。随后将细胞换液至 0.25%
Triton X 中，50 μL/孔，室温通透 30 分钟，弃去孔内液体，用 PBST(300 μL/孔)洗细胞 3
遍。

封闭：向孔内加入 50 μL 10% BSA 用于封闭非特异性结合位点，37°C 孵育 1 h。封
闭完成后，弃去孔内液体，用 PBST(300 μL/孔)洗细胞 5 遍。

25 孵育抗体：用 5% BSA 将 HRP 偶联的抗牛痘病毒抗体(Antibodies-Online，
ABIN288495)稀释 2000 倍并加入 96 孔板中，100 μL/孔，37°C 孵育 1 h。孵育结束后，
弃去孔内液体，用 PBST(300 μL/孔)洗细胞 5 遍。

显色与读值：根据使用说明，向 96 孔板中加入 50μL/孔的 TrueBlue™过氧化物酶底
物(KPL, 5510-0030)，常温避光孵育 15 分钟，使病毒斑点显色。显色结束后，吸弃显色
30 液，将 96 孔板置于室温过夜晾干，在此过程中需注意避光。次日，用荧光酶联免疫斑
点分析仪(CTL)进行斑点计数。

结果计算：以检测孔的噬斑面积作为 A_x，病毒对照孔噬斑面积的均值为 A_v，细胞
对照孔噬斑面积的均值为 A_c，根据以下公式计算血清中和抑制率(%)：血清中和抑制率

(%)=1-(A_x-A_c)/(A_v-A_c)。将计算得到的血清中和抑制率与其对应的浓度对数(Log)值导入 Graphpad Prism 9.0 软件, 计算血清样本内的中和抗体滴度 NT₅₀, 即减少 50%噬斑产生所需的血清稀释度的倒数。通过单因素方差分析(One-way ANOVA)计算各组之间是否具有显著性差异。

5 试验结果如图 1 所示, 对于融合表达的 mRNA 疫苗, 不同的疫苗设计方式会对候选疫苗的免疫原性产生不同的影响。例如, G1、G3、G4、G5 和 G6 组小鼠免疫的 mRNA 疫苗(G1、G3、G4、G5、G6 mRNA 疫苗)都针对 6 种抗原多肽(A29L、A35R、B2R、B6R、C15L 和 M1R), 其中 G1 mRNA 疫苗包含 6 种单独表达的 mRNA 分子, 与之相比, G5 mRNA 疫苗所采用的融合表达设计方式(1501 + 1601)并未对候选疫苗的免疫原性造成明
10 显影响, 而 G6 mRNA 疫苗所采用的融合表达设计方式(1801 + 1901)可显著提升候选疫苗诱导产生中和抗体的能力。

类似地, G2、G7、G8 和 G9 组小鼠免疫的 mRNA 疫苗(G2、G7、G8、G9 mRNA 疫苗)也都针对 6 种抗原多肽(A35R、B2R、B6R、C15L、M1R 和 H3L), 其中 G7 mRNA 疫苗所采用的融合表达设计方式(1201 + 1401)可显著提升候选疫苗诱导产生中和抗体的能力。
15

此外, 在 mRNA 疫苗 G1-G10 中, G6 和 G7 候选疫苗诱导最高的血清中和抗体滴度, 其中和水平显著高于将单独表达的 mRNA 分子通过共同包封的方式制成的多价苗 G1 和/或 G2, 以及参照品疫苗 G10。并且通过融合表达设计减少制备疫苗所需的 mRNA 条数, 大大降低了疫苗生产及放行过程中所需的工作量及工作难度。

20

4.3 疫苗诱导的细胞免疫应答的检测

于初免后第 42 天安乐死受试小鼠, 采集小鼠脾脏组织, 研磨制成单细胞悬液后检测疫苗诱导的细胞免疫应答, 其具体检测方法如下。

制备脾细胞悬液和计数活细胞: 准备适量 6 孔板, 向每孔中加入 3 mL 已预冷的含有 2% FBS 的 1640 培养基, 将无菌细胞滤网置于孔中。将新鲜采集的脾脏组织置于细胞滤网上, 使用 5 mL 注射器的推柄轻轻研磨脾脏组织并通过细胞滤网过滤去除脂肪等组织, 形成脾细胞悬液。将脾细胞悬液转移到 15 mL 离心管中, 4°C, 300 g 离心 5 min, 弃去上清。用 2 mL 1×红细胞裂解液重悬脾细胞, 室温条件下作用 3-5 min, 然后立即加入 4 mL 含有 10% FBS 的 PBS 缓冲液并上下颠倒混匀, 终止裂解反应。通过细胞筛网
25 将裂红后的脾细胞过滤入另一新的 15 mL 离心管中, 去除细胞团, 随后在 4°C, 300 g 离心 5 min, 彻底去除上清。使用 1 mL 已预冷的含有 10% FBS 的 1640 培养基重悬细胞, 轻柔吹打混匀后, 取 10 μL 细胞与等体积的 AO/PI 染色液(Count Star, RE010212) 1:1 混合, 计算活细胞浓度。将细胞浓度调整至 1.5×10^7 个活细胞/mL, 然后取 0.1 mL 细胞加入 U 型底 96 孔细胞培养板中, 即每孔 1.5×10^6 个活细胞, 每个样品做两个复孔。
30

35 配制抗原刺激物工作液、阴性对照培养基工作液和细胞活化刺激物工作液(阳性对照):

1) 抗原刺激物工作液: 将 CD28 抗体(Biolegend, 122002)、CD49d(BD Biosciences, 553313)抗体和猴痘抗原混合肽库(金斯瑞)加入到已预冷的含有 10% FBS 的 1640 培养基中, 使得 CD28 抗体、CD49d 抗体和每个单个猴痘抗原肽库的终浓度均为 4 μg/mL。

2) 阴性对照培养基工作液: 将 CD28 抗体和 CD49d 抗体加入到已预冷的含有 10% FBS 的 1640 培养基中, 使得二者的终浓度均为 4 $\mu\text{g/mL}$ 。

3) 细胞活化刺激物工作液: 将 500 $\times$ 细胞激活混合物(Cell activation cocktail, Biolegend, 423302)用已预冷的含有 10% FBS 的 1640 培养基稀释 250 倍, 使其浓度变为 2 $\times$ 。

刺激: 对于每个样品的两个复孔, 向其中一个复孔中加入 100 μL 抗原刺激物工作液, 使得 CD28 抗体、CD49d 抗体和每个单个猴痘抗原肽库的终浓度均为 2 $\mu\text{g/mL}$; 向另一个复孔中加入 100 μL 阴性对照培养基工作液, 使得 CD28 抗体和 CD49d 的终浓度均为 2 $\mu\text{g/mL}$; 对于阳性对照孔, 向每孔中加入 100 μL 细胞活化刺激物工作液, 使得每孔中的 Cell activation cocktail 的工作浓度为 1 $\times$ 。将细胞样品置于 37 $^{\circ}\text{C}$ 、5% CO_2 刺激 2 h。

阻断: 用已预冷的含有 10% FBS 的 1640 培养基将 1000 $\times$ Brefeldin A 阻断剂(Biolegend, 420601)稀释 50 倍, 即稀释至 20 $\times$, 随后向每孔的 200 μL 细胞悬液内加入 10.5 μL 20 $\times$ Brefeldin A, 使其工作浓度为 1 $\times$, 轻柔吹打混匀。将加有阻断剂的细胞样品置于 37 $^{\circ}\text{C}$ 、5% CO_2 阻断 4 h。阻断完成后, 将细胞置于离心机中, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。使用 PBS 缓冲液洗细胞两遍, 200 μL /孔, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。

活死染料染色: 用 PBS 缓冲液将细胞活死染料 Zombie green(Biolegend, 423112)稀释 1000 倍后加入细胞中并重悬细胞, 50 μL /孔, 4 $^{\circ}\text{C}$ 避光孵育 30 min。孵育完成后, 将细胞置于离心机中, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。使用流式染色缓冲液(含有 2% FBS 的 PBS 缓冲液)洗细胞三遍, 200 μL /孔, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃去上清。

表面染色: 向流式染色缓冲液中依次加入表面染色抗体抗小鼠 CD45(Biolegend, 103138)、抗小鼠 CD3(Biolegend, 100216)、抗小鼠 CD4(Biolegend, 100434)和抗小鼠 CD8a(Biolegend, 100714), 使得每种抗体的终浓度均为 2 $\mu\text{g/mL}$ 。取 50 μL 表面染色抗体混合液加入孔中并重悬细胞, 4 $^{\circ}\text{C}$ 避光孵育 30 min。孵育完成后, 将细胞置于离心机中, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。使用流式染色缓冲液洗细胞三遍, 200 μL /孔, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃去上清。

固定、破膜和胞内染色: 向细胞内加入 50 μL 固定通透试剂盒(BD Biosciences, 554714)中的固定液并重悬细胞, 4 $^{\circ}\text{C}$ 避光孵育 20 min。固定完成后, 将细胞置于离心机中, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。用双蒸水将固定通透试剂盒中的 10 $\times$ 破膜液稀释至 1 $\times$, 随后使用 1 $\times$ 破膜液洗细胞三遍, 200 μL /孔, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃去上清。向 1 $\times$ 破膜液中依次加入胞内染色抗体抗小鼠 IFN- γ (Biolegend, 505808)、抗小鼠 TNF- α (Biolegend, 506328)、抗小鼠 IL-2 (Biolegend, 503810)和抗小鼠 IL-4(Biolegend, 504118), 使得每种抗体的终浓度均为 2 $\mu\text{g/mL}$, 取 50 μL 胞内染色抗体混合液加入孔中并重悬细胞, 4 $^{\circ}\text{C}$ 避光孵育 30 min。孵育完成后, 将细胞置于离心机中, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃上清。使用 1 $\times$ 破膜液洗细胞三遍, 200 μL /孔, 4 $^{\circ}\text{C}$, 300 g 离心 5 min, 弃去上清。用 200 μL 流式染色缓冲液重悬细胞, 使用流式细胞仪(BD biosciences, LSRFortessa)检测样品。

试验结果如图 2 所示, 相较于将单独表达的 mRNA 分子通过共同包封的方式制成的多价苗 G1, 融合表达设计的 mRNA 疫苗 G6 诱导更高的细胞免疫应答水平。对于表达另一抗原多肽组合的 mRNA 疫苗, 融合表达的 mRNA 疫苗 G7 与将单独表达的 mRNA

分子通过共同包封的方式制成的多价苗 G2 诱导的细胞免疫应答水平相当(数据未显示)。此外，实施例 3 中制备的所有 mRNA 疫苗均诱导 Th1 型免疫应答(部分数据未显示)。

序列表

SEQ ID NO:	描述	序列
1	0201 RNA ORF	AUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCCCACUAGGGUGCACAGC GACGGGACCCUGUUCCAGGGGACGACGACCUGGCCAUCCCCGCUACCGAGUUCU UCAGCACCAAGGCUGCCAAGAACCCCGAGACUAAGCGGGAGGCCAUCGUGAAGGC CUACGGGGACGAUAAUGAGGAGACGCGUAAGCAGCGUCUACCAAUCUGGAGAA GAAGAUUACUAAUAUACCAACCAAGUUCGAGCAGAUCCGAGAAGUGCUGUAAGAG GAAUGACGAGGUGCUGUUCGGCUGGAGAACCACGCCGAGACACUGCGGGCUGC CAUGAUCUCUCUUGCCAAGAAGAUCCAUGUGCAGACCGGCAGGCACCCUUAUUA CCUACGACGUGCCGAUUAUGCCGAGUGA
2	0302 RNA ORF	AUGAUGACUCCCGAGAAUGACGAGGAGCAGACCAGCGUGUUCUCCGCUACCGUG UACGGGGACAAGAUCCAGGGGAAGAAUAAGAGGAAGCGGGUGAUCGGCCUGUGC AUCAGGAUCAGCAUGGUGAUCAGCCUGCUGAGCAUGAUCACCAUGUCUGCAUUC CUGAUCGUGCGGCUGAAUCAGUGCAUGAGCGCCAACAAGGCCGCCAUCACUGAUU CAGCCGUGGGCCUGGGCCGCCGCUUCCUCUACUCACAGGAAGGUGGUGUCCUCCAC CACUCAGUACGAUCACAAGGAGAGCUGCAAUGGCCUGUAUUAACAGGGCUCUUGC UACAUCCUGCACAGCGAUUACAAGAGCUUUGAGGAUGCCAAGGCCAAUUGCGCCG CCGAGAGCAGCACUCUGCCUAACAAGAGCGACGUGCUGACUACUUGGCUGAUCGA CUAUGUGGAGGACACCUUGGGGGAGCGAUGGGAACCCCAUACCAAGACUACCAGC GACUAUCAGGACAGCGACGUGAGCCAGGAGGUGCGGAAGUACUUCUGCACCUACC CUUACGACGUGCCUGACUACGCCUGA
3	0402 RNA ORF	AUGACCCAGCUGCCUAUCCUGCUGCUGAUCAGCCUGGUGUACGCUACCCCUA GUCCCCAGACUCCAAGAAGAUCCGGGACGACGCCACUAUCAGCUGCAGCAGGAA UAACACCAACUACUACGUGGUGAUGAGCGCCUGGUUAUAGGAGCCUAACAGCAUC AUCCUGCUGGCCGCCAAGAGCGAUGUGCUGUACUUCGACAACUAUACCAAGGACA AGAUUCUUAACGAUAGCCCCUACGAUGACCUUGGUGACUACCAUACCAUAAGAG UCUGACCGCCGGGAUGCCGGGACCUAUUAUCUGUGCGUUUUUAUGACCUUAC AACAAACGACACAGAUAAAGUGGACUACGAGGAGUAUAGCAUCGAGCUGAUCGU GAACACCGACAGCGAGAGUACCAUCGACAUCAUCCUGAGCGGGAGUACUCCCGAG ACUAUCAGCGAGAAGCCAGAGGACAUAGCAUAGCAACUGCAGCAGCGUGUUCG AGAUCACUACUCCUGAGCCUAUUAACAGAUAAACGUGGAUGAUCACACCGACACCGU GACCUACACAGCGACAGCAUAAACCGUGAAUGCUAGCAGCGGUGAGAGUACC ACGGAUGAGAUCCCCGAGCCCAUCACGGAUAAGGAGGAGGAUCACACCGUGACCG ACACGGUGAGUUAACACACCGUGUCUACCUCCUAGGGAUUGUGACCACAAAGUC UACCACCGAUGAUGCCGAUCUGUAUGACACAUACAAGACAACGAUACCGUGCCG CCCACACCGUGGGCGGCAGCACAACAUCCAUAAGUAACUACAAGACUAAGGAUU UCGUGGAGAUUUUGGGAUACAACCCUGAUUAUCCUGAGCGCCGUGGCCAUCU UCUGUAUACCUACUUAUCUGUAUAAGCAUCCAGGAAGUAUAAGACCGAGAA CAAGGUGUAUCCGUACGACGUGCCGGACUACGCCUGA
4	0505 RNA ORF	AUGAAGACCAUUCUGGUGACCCUGCUGUGCGUGCUGCCAGCAGUGGUGUAC

	RNA		AGUGGCGGCGGGGGAGCUACCCCUACGACGUGCCCGACUACGCCACCUGUACGG UGCCAACCAUGAACAAACGCUAAGCUGACCUCCACAGAGACGAGUUUUAAUGACAA GCAGAAGGUGACCUUACCUUGCGACAGUGGCUACCACAGCUUGGACCCCAACGCU GUGUGCGAGACAGACAAGUGGAAGUACGAGAACCCGUGCAAGAAGAUUGGCACC GUGAGUGACUACGUCAGCGAGCUGUACGACAAGCCGCUUACGAGGUCAACUCCA CGAUGACACUCUCUUGCAACGGGGAGACAAAGUACUUCAGGUGCGAGGAGAAGA ACGGGAUJACCAGCUGGAAUGAUACCGUGACAUGCCCCAAUGCCGAGUGCCAGCC CCUGCAGCUCGAGCAUGGGAGCUGUCAGCCCGUGAAGGAGAAGUACUCCUUCGG GGAGUACAUGACCAUCAAUUGCGACGUGGGGUACGAGGUGAUCGGGGUGAGCUA CAUCUCUUGCACCGCCAAUAGCUGGAACGUGAUUCCCAGCUGUCAGCAGAAGUGC GACAUCCCUUCUCUGUCCAACGGGCUGAUCAGCGGGAGCACGUUCAGCAUUGGC GGGGUGAUCCACCUGAGCUGCAAGAGUGGGUUCACCCUGACCGGGUCCCCGAGC AGCACCUGCAUCGAUGGGAAGUGGAACCCGAUCCUGCCAACUUGCGUCCGGUCCA ACGAGGAGUUCGACCCCGUGGACGACGGACCAGACGACGAAACCGAUUCAGCAA GUUGAGCAAGGACGUGGUGCAGUACGAGCAGGAGAUCGAGAGCCUGGAGGCCAC CUACCACAUCAUCAUAGGCACUGACCAUCAUGGGCGUGAUUUUCUUGAUCAGC AUCAUUGUGCUGGUGUGUUCUUGCGACAAGAACAACGACCAGUACAAGUUUCAC AAGCUGCUGCCCUGA
5	0603 RNA	ORF	AUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCGCUACUCGGGUGCAUAGC GGCGGCGGGGGGUCCUACCCCUACGACGUGCCCGACUAGCCGCCGAGACAAAGG AGUUUAAGACCCUGUACAACCUGUUUAUCGACAGCUAUCUGCAGAAGCUGGCUC AGCACUCUAUUCCACCAUUGUGACAUGCGCAUUCACAUUGGUGAGGUGAUCG GCCAGUUAAGAACUGCGCCUGAGGAUCACGAAUAAGUGCAUGAGCAACAGCAG GCUGAGCUUACCCUGAUGGUGGAGUCCUUAUCGAGGUGAUCAGCCUGCUGAG CGAGAAGGACAGGAGGGCAUCGCCGAGGAGAUUCGGCAUCGACCUGGACGAUGU GCCCAGCGUGGUGAGCAAGCUGGAGAAGAACUGCAACGCAUUAUCUGAGGUGAA CAACAUCAUCGAUUAUCAGAAGCUGAAUAUCGGUGAGUGUCCCGUCCCCUGG GCAGCAUUAUCUGUUGCAGAUUCGUGAACACGGGCAGUGCCGAGGCCAACUGCGG GCUGCAGACCAUCGUGAAGUCACUGAAUAAGAUUACGUGCCACCCAUCAUCGAG AACCGGCUGCCCUACUACGACCCGUGGUUCCUGGUGGGGUGGCCAUUAUUCUG GUGAUUUACCGUGGCCAUUCGAGCAUCCGCAGGAACCUUGGCCUGAAGUACA GGUACGGGACCUUCCUGUACGUCUGA
6	0702 RNA	ORF	AUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCACCAGGGUGCACAGU GGCGGAGGAGGCAGCUACCCCUACGACGUGCCCGACUACGUGGGGUGCUGCCU CCAUCCAGACCACUGUGAACACCCUGAGUGAGCGGAUCAGCAGCAAGCUGGAGCA GGAGGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAUAGAGAUUGGCAUUUUCUA UAUCAGGCAGAACACGGGUGCAACAUCACCGUGAAGAAUAUGUCUCCGUGAC GCCGACGUCACGUGGACGCCUGCUGAGCGCGGCCACGGAGACAUUUCGGGUC UGACCCUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCACUGAACAU CCAGACCAGCGUGAACACCGUGGUGAGGGACUUCGAGAACUACGUGAAGCAGACA UGCAACAGCAGCGCGGUGGUGGAUAAUAAGCUCAAGAUCAGAAUGGAUCAUU GACGAGUGUUAUGGGGCCCUUGGAGCCCCACUAACCUCGAGUUUAUUAUACG GGCAGCAGCAAGGGCAACUGCGCAUCAAGGCCUGAUGCAGCUGACAACCAAGG CCACCACCAGAUUGCCCCAGGCAGGUGGCCGGUUGA

7	0803 RNA	ORF	AUGGACUGGACCUGGAUUCUGUUUCUGGUGGCCGCCGCCACACGGGUGCAUUC GGCGGCGGCGGCAGCUACCAUACGAUGUGCCUGACUACGCCGCCGCCGUGAAGA CACCCGUGAUCGUGGUGCCCGUGAUUGACCGGCCCCUAGCGAGACAUUCCCCAA UGUGCACGAGCACAUCAACGACCAGAAGUUCGACGACGUGAAGGACAACGAGGUG AUGCAGGAGAAGAGGGACGUGGUGAUUGUGAACGACGAUCCUGACCACUACAAG GACUACGUGUUAUCCAGUGGACAGGCGGGAACAUCAGGGACGACGACAAGUAU ACCCACUUCUUCUCCGGCUUCUGUAACACCAUGUGUACAGAGGAGACAAAGCGGA ACAUUGCUCGGCACCUGGCCUGUGGGACAGCAAGUUCUUCACCGAGCUGGAGA ACAAGAACGUGGAGUAUGUGGUGAUCAUCGAGAACGACAACGUCAUCGAGGACA UCACCUUUCUGCGCCUGUGCUGAAGGCCAUCCACGAUAAGAAGAUCGACAUCU GCAGAUCCGGGAGAUCAUACCCGGGAUAAGGUGAAGACCGAGCUGGUGAUCGA CAAGGAUCACGCCAUUCUACCUACACCGGCGGCUAUGAUGUGAGCCUGAGCGCC UACAUCAUCCGGGUGACCACCGCCUGAAUAUCGUGGAUGAGAUCAUCAAGUCCG GCGGGCUGAGCAGCGGCUUCUACUUCGAGAUCCGAGAAUCGAGAACGAGAUGA AGAUCAACAGGCAGAUCAUGGACAACAGCGCAAGUACGUGGAGCAGACCCUCG GCUGGUGGCCGAGCACCAGUUCGAGACUAUGAAGCCCAACUUCUGGUCCAGAAU CGGGACUGUGGCCGCCAAGCGGUACCCCGGCGUGAUGUACACAUUACCAACCCCA CUGAUCAGCUUCUUGGCCUGUUCGACAUCAACGUGAUCGGCCUGAUCGUGAUC CUGUUAUCAUGUUAUGCUGAUCUUAACGUGAAGUCCAAGCUGCUGUGGUUC CUGACCGGCACCUUCGUGACUGCCUUAUCUGA
8	0201-J0 RNA	ORF	AUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCCACGCGAGUCCACUCC AUGGACGGAACUCUUUUUCCCCGAGAUACGAUCUUGCAAUUCCAGCAACUGAA UUUUUCUCUACAAAGGCUGCUAAAAUCCAGAGACUAAACGCGAAGCAAUUGUU AAAGCCUAUGGAGACGAUAAUGAGGAAACUCUAAACAACGGCUAACUAAUUGG AAAAAAGAUUACUAAUAUAACAACAAGUUUGAACAAUAGAAAAGUGUUGUA AACGCAACGAUGAAGUUCUUAUUUAGGUUGGAAAAUCACGCUGAAACUCUAGAG CGGCUAUGAUUUCUGGCUAAAAAGAUUGAUGUACAGACUGGACGGCAUCCAU AUGAGUAA
9	0301-J0 RNA	ORF	AUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCCACGCGAGUCCACUCC AUGAUGACACCAGAAAACGACGAAGAGCAGACAUUCUGUGUUCUCCGCUACUGUU UACGGAGACAAAUUCAGGGAAAGAAUAAACGCAAACGCGUGAUUGGUCUAUGU AUUAGAAUUCUAUGGUUAUUUCACUACUACUUAUGAUUACCAUGUCCGCGUUU CUCAUAGUGCGCCUAAAUCAUUGCAUGUCUGCUAACAAGGCUAGUUAUACUGAC UCCGCUGUUGCCGUUGCUGCGGCAUCAUCUACUACUAGAAAGGUUGUGUCUAGC ACUACACAAUUGAUACAAAGAAAGCUGUAAUGGUUUUAUUAUACAGGGUUCU UGUUUAUUAUUAUUAUUCAGACUUAAGUCAUUCGAGGAUGCUAAAGCAAACUGC GCUGCGGAAUCAUACACUACCCAAUAAAUCCGAUGUCUUGACUACCGGCUCA UUGAUUAUGUUGAGGAUACUAGGGGAUCUGAUGGUAAUCCAAUUAACAAAACUA CAUCCGAUUAUCAAGAUUCUGAUGUAUCACAAGAAGUAGAAAGUAUUUUUGUA CAUAA
10	0501-J0 RNA	ORF	AUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCCACGCGAGUCCACUCC AUGAAAACGAUUUCCGUUGUUAUGUUAUGCGUACUACCGUCUGUUGUUUA UUCAACAUGUACUGUACCCACUAUGAAUACGCUAAAUAACGUCUACCGAAACA UCGUUUAAUGAUAAACAGAAAGUACGUUUACAUGUGAUUCAGGAUAUCAUUCU

		UUGGAUCCAAUUGCUGUCUGCGAAACAGAUAAAUGGAAUACGAAAAUCCAUGC AAGAAAAUGUGCACAGUUUCUGAUUAUGUCUCUGAACUUAUGAUAAAGCCAUUA UACGAAGUGAAUUCCACCAUGACACUAAGUUGCAACGGUGAAACAAAUAUUUUC GUUGUGAAGAAAAAAUUGGAAUACUUCUUGGAUGAUACUGUCACGUGUCCUA AUGCGGAAUGUCAACCUCUCAAUUAGAACACGGAUCGUGUCAACCAGUUAAG AAAAUACUCAUUUGGGGAAUAUGACUAUCAACUGUGAUGUUGGAUAUGAG GUUAUUGGUGUUCGUAUAUAAGUUGUACGGCUAAUUCUUGGAAUGUUAUUC AUCAUGUCAACAAAAAUGGAUAUACCGUCCCUAUCUAAUGGAUUAUUUCCGG AUCUACAUUUUCUAUCGGUGGCGUUAUACAUCUJAGUUGUAAAAGUGGUUUUA CACUAACGGGGUCCCAUCAUCCACAUGUAUCGACGGUAAAUGGAAUCCCAUACU CCCAACAUUGUGUACGAUCUAACGAAGAAUUUGAUCCAGUGGAUGAUGGUCCGA CGAUGAGACAGAUCUGAGCAAACUCUCGAAAGACGUUGUACAAUUGAACAAAGAA AUAGAAUCGUUAGAAGCAACUUAUCAUAUAUACAUAUUGGCGUUGACAAUUAUG GGUGUCAUAUUUCUAAUCUCCAUAUAGUAUJAGUUUGUUCUGUGACAAAAA UAAUGACCAUAUAAGUCCAUAUUUGCUACCGUGA
11	0701-J0 ORF RNA	AUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCCACGCGAGUCCACUCC AUGGGUGCCGCAGCAAGCAUACAGACGACUGUGAAUACACUCAGUGAACGUUAUCU CGUCUAAAUAJAGAACAAGAAGCGAACGCUAGUGCUAAACAAAUGUGAUUAJAGA AAUCGGAUUUUUAUAUCCGACAAAACCAUGGAUGUAACAUCACUGUUAUAAAA UAUGUGCUCUGCGGACGCGGAUGCUCAGUUGGAUGCUGUGUUAUCAGCCGCUAC AGAAACAUAUAGUGGAUUAACACCGGAACAAAAGCAUACGUACCAGCUAUGUUU ACUGCUGCGUUAACAUAUCAGACGAGUGUAAACACUGUUGUJAGAGAUUUUGAA AAUUAUGUGAAACAGACUUGUAAUUCUAGCGCUGUUGUCGAUAACAAAUUAAG AUACAAACGUAAUUAJAGAUAGAAUGUACGGAGCCCCAGGAUCUCCAACAAAUU UGGAUUUUAUAUAACAGGAUCUJAGCAAAGGAAAUUGUGCCAUAAGGCGUUGA UGCAAUUGACUACUAAGGCCACUACUCAAUAGCACCUAGACAAGUUGCUGGUAC AGGAGUUCAGUUUUUAUJAGAUUGUUAUCGGUGUUAUAUAUUGGACGCGUUGU UUAUGUACUAUGCCAAGCGUAUGCUGUUCACAUCCACCAUJAGUAAAAUCAAACU UAUUUUJAGCAAUAAGGAAACGUCCAUUGGACUACUUAACUJAGGACAUUCU UAGAACUUCUCCGAUGAUUAUUGCUACCACGGAUUAUACAAACUGA
12	0901-J0 ORF RNA	AUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCCACGCGAGUCCACUCC AUGCCUCAACAACUUCUCCUAUUAUAUJAGAAACUAAAAAGCAAUUCUGACG CGCGAUUGAAGACGUJAGACAUAUAUAUJAGAGUCGAAACCAACCACUAUCCA GAAUACUGGAAAACUAGUAAGGAUUAUUUUAAAGGAGGAUAUAUJAGUGGAG GGUUUCUCCCAUJAGAAUJAGUGUUAUCAACAAUACAUAUAUJAGGGGAAAGG AAGACGAUUAUGGAUCCAUCACUJAGAUJAGAUJAGUAUACAUAUACUCUGGAGAGA UUAUUCUGGUUCAUUGGAUAAGAAAAAUJAGUUCUUAUGAAGAGGCAAAAA AACACGAUGAUGGAUUAUCAUAUUGCUUAUUCUUAACAAGUAUCAGAUCAUA AAAAUGUAUAUUUCAAAGAUJAGUCAUAUUGGAUJAGAUUJAGAUCCGCUA AUJAGUCUGCGCAUUUGAUUCAGUAUUUUUAUCUJAGACAAUUGCUACCUAGUA CGUUGGAUUAUUUAUAUAUCUJAGGAACAACUAUACACACUCUGCAGACGCAGC AUGGAUAAUUUUCCAACCAUAUAAACAUUAUUCUJAGUAACUAUCUAAAUUC AGAACACUAUUGUCGUCGUCUAAUCAUGAGGGUAAACCUAUUAUAUAACAGAG AAUUAJAGAAAUCCGUUAUAAUUGAACGACGACGCAAGUAUAUAUUCUGGG

			GAGAUUAUACGAGCUGCAACCACCUCACCAGUUCGAGAGAACUAUUUUAUGAAAU GGUUGUCUGAUUUUGAGAGAGGCAUGUUUUUCAUUAUUAUAAAAUUAUCGAA GGGAAUAAAACAUUCGCAUUAUUGCCAUGAUUUCGUUUUAUACUCACCGCU AUUCUCUUUCUUAUGAGUCAACGAUUAUUCGCGAGAAAAACAAACUAG
13	1101 RNA	ORF	AUGGACUGGACGUGGAUUCUCUCCUGGUGGCAGCCGCCACUCGUGUGCACUCC GAUGGGACCUUGUUCUCCUGGGGACGACGACCUGGCCAUCCCGCUACGGAGUUC UUCAGCACCAAGGCUGCCAAGAACC CGGAGACUAAGCGGGAGGCCAUCGUCAAGG CCUACGGAGAUGACAACGAGGAGACGCUGAAGCAGCGGCUGACCAACCUGGAGAA GAAGAUACGAACAUACGACCAAGUUCGAGCAGAUUGAGAAGUGCUGCAAGCGG AACGACGAGGUGUUGUUCGCUUGGAGAACCACGCAGAGACUCUGCGGGCAGCC AUGAUCAGCUUGGCGAAGAAGAUUCGACGUGCAGACUGGACGCCACCCCUAUGAGG GAGGAGGCGGCAGCGGCGGCGGCGGCAGCCGGCUGAAUCAGUGCAUGAGCGCCA ACAAGGCGGCCAUACUGAUUCAGCCGUCGCCGUCGCUGCCGCCUCCUCCACUCAU AGGAAGGUGGUGUCCAGUACGACACAGUACGACCAUAAGGAGAGCUGUAAUGGG CUGUACUACCAGGGGUCGUGCUACAUCCUGCACAGCGACUACAAGUCGUUUGAG GAUGCGAAGGCGAACUGCGCAGCCGAGAGCAGCACGUUGCCUAACAAGAGCGACG UGCUGACUACUUGGCUGAUUGACUAUGUCGAGGAUACCUGGGGCUCCGAUGGG AACCCCAUCACGAAGACCACGUCUGACUACCAGGACUCCGAUGUGUCUCAGGAGG UGC GGAAGUACUUCUGUACUGGCGGCGGCGGUUCUGGUGGGGCGGCAGCACC UGCACCGUCCCCACCAUGAACAACGCCAAGCUGACCAGUACAGAGACUCCUUA CGAUAAAGCAGAAGGUGACAUUUAUGCGAUUCCGGCUACCACUCGUUGGACCCC AACGUCUGUCUGAGACAGACAAGUGGAAGUACGAGAAUCCAUGUAAGAAGAU UGCACGGUGUCCGACUAUGUUUCUGAGCUGUACGACAAGCCCCUUAACGAGGUG AACAGCACCAUGACUCUCAGCUGCAACGCGGAGACUAAGUACUUUCGGUGUGAG GAGAAGAACGGUAAUACGUCCUGGAACGAUACCGUUAUUGCCCCAACGCCGAGU GCCAGCCGUUGCAGCUGGAGCAUGGUAGCUGUCAGCCCGUGAAGGAGAAGUAUA GUUUCGGCGAGUACAUGACCAUUAUUGUGAUGUGGGUUAACGAGGUCAUCGGG GUGAGUUACAUCUCUUGCACUGCCAUAAGCUGGAACGUGAUUCCAGCUGUCAG CAGAAGUGUGACAUACCUUCUCUGUCCAACGGGCGUGAUCAGCGGGAGCACGUUC AGCAUUGGCGGGGUGAUCCACCUCAGCUGCAAGAGCGGGUUCACGUUGACCGGG UCGCCUCGUCGACCUGCAUCGACGGGAAGUGGAUCCAUUCUGCCACGUGCG UGC GGUCGAACGAGGAGUUCGACCCGUCGACGAUGGACCCGAUGACGAGACUG AUCUGUCUAAGUUGUCGAAGGACGUCGUUCAGUACGAGCAGGAGAUCGAGUCCC UGGAGGCCACGUACCACGGCGGUGGCGGGAGCGGGGAGGUGGCAGCAGUCCCC AGACUCCAAGAAGAUCCGGGACGACGCCACCAUCCUGCUCGCCCAACAACACU AACUACUACGUGGUUAUGUCCGCCUGGUACAAGGAGCCCAUCGAUCAUCCUGC UCGUCGUAAGAGCGACGUCCUGUACUUCGACAACUAUACCAAGGACAAGAUACG CUACGACUCACCCUACGAUGACCUCGUGACCACCAUCACAAUUAAGAGUCUGACCG CCGGGGACGCCGGCAUACAUUUGCGCCUUCUUAUGACGAGUACCACCAUUGA UACCGACAAGGUGGAUUAUGAGGAGUACUCGAUCGAGCUUAUCGUGAACACGGA CAGUGAGUCGACCAUCGACAUCAUCCUGUCCGGUAGCACCCUGAGACAAUCC GAGAAGCCGGAGGACAUCGACAAUAGCAAUUGCAGUUCGUCUUCGAGAUACGGA CCCCUGAGCCCAUACCGACAACGUGGAUGAUCACACCGACAGGUGACGUACACC AGCGACUCUAUUAACACCGUCAUUGCUUCCAGUGGGGAGUCCACCACUGACGAGA

			UUCCGGAGCCCAUACCCGACAAGGAGGAGGAUACACCGUGACCGACACGGUGAG UUACACCACCGUGUCUACCUCUCUGGUAUUGUGACUACCAAGAGCACCACGGAC GACGCCGAUCUGUACGACACGUACAAUGACAACGACACCGUGCCGCCACCACCGU GGGCGGGUCCACGACGUCCAUAUAGUAACUACAAGACUAAGGACUUCGUGGGGGG CGGCGGUUCCGGUGGCGGGGGCAGCGCUGAAACCAAGGAGUUAAGACGCUGUA CAACUUGUUAUCGACUCUUAACUGCAGAAGUUGGCCCAGCACUCCAUCCCUACC AAUGUCACUUGUGCCAUCCACAUUGGAGAGGUGAUCGGCCAGUUAAGAACUGC GCCCUGAGGAUACCAACAAGUGCAUAGCAAUAGCCGCCUGUCCUUCACGCUCA UGGUGGAGUCCUUAUCGAGGUGAUCUCCUGCUGAGCGAGAAGGACAGGCGGG CUAUUGCUGAGGAGAUCCGUAAUUGACCUCGACGACGUGCCGUCGGUCGUCUCCA AGCUGGAGAAGAACUGCAACGCAUAUUGCUGAGGUGAACAACAUCAUCGAUUAUC AGAAGCUGAAUAUCGGUGAGUGUUCCGCUCCCCUGGGCAGCAUAUCGUGUUGC AGAUUGUCAAUACCGGGAGUGCUGAGGCCAACUGCGGGUUGCAGACCAUCGUGA AGAGUCUGAACAAGAUUUAACGUGCCCCGAUCAUCGAGAACCGCCUGCCCUACUA CGACCCCGUGGGGGCGGGUCUGGAGGAGGCGGCAGCGCGCUGCCGCCUCCA CCAGACCACGGUGAAUACGUUGUCCGAGCGGAUUUCGAGCAAGCUGGAGCAGGA GGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAAGAGAUUGGCAUUUCUAUAU CAGGCAGAACCACGGGUGCAACAUACCCGUGAAGAAUAUGUGCUCCGCGUACGCC GACGCUACGUGGACGCCGUGCUGAGCGCGGCCACGGAGACAUUUCGGUCUGA CCCCUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCUCUGAACAUCCA GACGUCCGUGAAUACCGUGGUGAGGGACUUCGAGAACUACGUGAAGCAGACGUG UAAUAGCAGCGCUGUCGUGCAUAAUAAGCUGAAGAUCCAGAACGUUAUUAUCGA CGAGUGCUACGGCGCCCCGGCUCUCCCGACCAACCUGGAGUUAUUAUACCGGU UCCUCCAAGGGGAACUGCGUAUUAAGGCGCUGAUGCAGCUGACUACGAAGGCG ACGACCCAGAUUCGCGCCUCGCCAGGUGGCGGGGACCGGGGUGCAGUUCUACAUGA UCGUGAUCGGUGUGAUAUCCUGGCAGCUCUCUUAUUGUACUACGCCAAGCGCA UGCUGUUCACCAGCACCAACGACAAGAUUAAGCUGAUCCUGCCAAUAAGGAGAA CGUCCACUGGACCACCUACAUGGACACAUUCUCCGCACCAGCCCUAUGAUUAUC GCCACCACUGACAUAUCAGAACUGA
14	1201 RNA	ORF	AUGACCCAGCUGCCUAUCCUGCUGCUGAUCAGUCUGGUGUACGCUACACCGU CUCCCCAGACUCCAAGAAGAUCCGGGACGACGCGACGAUCAGCUGCAGCAGGAA UAACACCAACUACUACGUGGUGAUGAGCGCCUGGUACAAGGAGCCGAUJAGCAUC AUUCUGCUGGCCGCCAAGUCUGACGUGCUGUACUUCGACAAUUAACACAAAGGACA AGAUCUCCUACGACAGUCCUUAUGAUGAUCUGGUGACCACGAUCACCAUUAAGUC UCUGACUGCUGGGGAUGCCGGUACGUUAUUCUGCGCCUUCUUAUGACGUGCCAC GACCAUUGACACUGAUAAAGGUGGAUUAACGAGGAGUACUCCAUCGAGCUGAUCGU GAACACCGACAGCGAGUCUACCAUCGACAUAUCCUGUCCGGGAGACCCCCGAGA CUAUCAGUGAGAAGCCUGAGGAUAUUGACAACAGCAAUUGCAGUAGUGUCUUCG AGAUACCAACCCCGAGCCUAUCACUGACAACGUGGACGAUCACACCGACACGGUG ACGUACACCAGUGAUAGCAUAACACCGUGAAUUCUAGCUCUCCGGGGAGUCCACGA CCGAUGAGAUCCCGAGCCUAUACCGAUAAAGGAGGAGGAUCACACCGUGACCGA CACGGUGAGUUAACACACCGUGUCUACCUCCUCCGGGAUCGUGACUACGAAGAGC ACUACUGACGAUGCUGAUCUGUACGAUACCUACAACGAUAACGAUACCGUGCCGC CCACCACCGUGGGCGGCAGCACCAACAGCAUCAGCAACUAUAAGACUAAGGACUUC

			GUGGGGGGCGGGGGAGCGGCGGCGGCGGCGAGCCGGCUGAAUCAGUGCAUGAGC GCCAACAGGCCGCCAUCACUGAUUCAGCCGUGGCUGUCGCCGCCAGCUCCAC CCACCGGAAGGUGGUGUCCAGCACCACCCAGUACGAUCAAGGAGUCCUGCAAU GGCCUGUAUUAACAGGGCUCUUGCUACAUCUUGCACAGCGAUUACAAGAGCUUU GAGGAUGCCAAGGCCAAUUGUGCAGCCGAGAGCAGCACGCUGCCUAACAAGAGCG ACGUGCUGACUACUUGGCUGAUCGACUAUGUGGAGGACACCUGGGGGAGUGAU GGGAACCCCAUCACUAAGACCACGUCCGACUAUCAGGACAGUGACGUGAGUCAGG AGGUGCGGAAGUACUUCUGCACCGGCGGCGGGGGCAGCGGCGGGGGCGGCAGCG CUGAGACAAAGGAGUUCAAGACCCUUACAACUUGUUAUCGAUAGUUAUCUGC AGAAGCUGGCCCAGCACAGCAUCCCCACCAUGUGACAUGCGCCAUUCACAUUGG UGAGGUGAUUGGCCAGUUAAGAACUGCGCCUGCGGAUAACUAACAAGUGUAU GUCUAAUUCGCGCCUGUCCUUCACGCUCAUGGUGGAGUCCUUAUCGAGGUGAU CUCCUGCUGAGCGAGAAGGACAGGCGGGCUAUCGCCGAGGAGAUCCGCAUCGAC CUGGAUGACGUCCCCAGCGUGGUGUCCAAGCUGGAGAAGAUCGCAACGCAUAU GCUGAGGUGAACAACAUCAUCAUGAUUUCAGAAGCUGAAUUAUCGGUGAGUGUCC GCUCCCCUGGGCAGCAUAUGCUGUUGCAGAUCAUUAACGGGUCUGCCGAG GCCAACUGCGGGCUGCAGACCAUCGUGAAGUCACUGAAUAAGAUUACGUGCCAC CCAUCAUCGAGAACCGGCGUGCCCUACUACGACCCGUGGUUCCUGGUGGGGGUGG CCAUUAUUCUGGUGAUUCUACGGUGGCCAUUCGAGCAUCCGAGGAACCUUGG CCUGAAGUACCGGUACGGUACCUUCCUCUACGUGUGA
15	1301 RNA	ORF	AUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCGCCACAAGGGUGCACUCC GAUGGGACCUUGUUCUCCUGGGGACGACGACCUGGCAUCCCGCUACGGAGUUC UUCAGCACCAAGGCGUCCAAGAACCCGGAGACUAAGCGGGAGGCCAUCGUCAAGG CCUACGGAGAUGACAAUGAGGAGACGCUGAAGCAGCGUCUACCAACCUUGAGAA GAAGAUACCAACAUCACUACGAAGUUUGAGCAGAUUGAGAAGUGCUGCAAGCGC AAUGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGAGACUCUGCGGGCUGCC AUGAUCAGCCUCGCCAAGAAGAUCAUGCAGGACGACCGGAAGGCACCCCUAUGAGG GGGGAGGAGGGAGCGGCGGGCGGGAGCGGGGCCGCGCCAGUAUCCAGACCA CCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAAGCUGGAGCAGGAGGCCAACGC CUCUGCUCAGACCAAGUGUGAUUAAGAGAUUGGCAAUUUCUAUAUCAGGCAGAA CCACGGGUGCAACAUACCGUGAAGAAUUGUGCUCGCGUGACGCCGACGCUAG CUGGACGCGGUGCUGAGCGCGGCCACGGAGACAUUUCGGUUCUGACCCUGAGC AGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCGCUGAAUUAUCCAGACGAGUG UUAACACGGUGGUGCGGGACUUUGAGAACUACGUGAAGCAGACGUGCAACAGCU CAGCGGUGGUGGAUAAUAAGCUCAAGAUCCAGAAUGUGAUCAUUGACGAGUGUU AUGGGGCCCCUGGGAGCCCCACUAACCUCGAGUUUAUUAUACGGGCAGCAGCAA GGGCAACUGCGCUAUCAGGCGCUGAUGCAGCUGACAACGAAGGCCACCACCCAG AUUGCUCCGCGGCAGGUGGCAGGAGGGGGAGGGGGUCCGGCGGGGGAGGCUC GACUUGCACGGUGCCGACGAUGAAUAACGCCAAGCUGACCUCCACCGAGACGUCG UUAACGACAAGCAGAAGGUGACGUUACCUUGCGACAGCGGCUAUCACAGCUUGG ACCCCAACCCGUCUGCGAGACGGACAAGUGGAAGUACGAGAACCCGUGCAAGAA GAUGUGCACGGUGUCCGACUAUGUGUCCGAGCUGUACGAUAAGCCGCUUACGA GGUGAACAGCACCAUGACGCUUGCUGCAACGGCGAGACAAAGUACUUCGGUGC GAGGAGAAGAACGGGAACACCAGCUGGAACGAUACAGUACCUGUCCCAACGCAG

			AGUGCCAGCCUCUGCAGUUGGAGCACGGGUCCUGUCAGCCGUUAAGGAGAAGU ACAGCUUUGGCGAGUACAUGACGAUUAACUGCGACGUCGGUUAACGAGGUGAUCG GCGUGAGUUAUCAUCUAUGUACCGCCAACAGCUGGAUUGUCAUACCUAGCUGCCA GCAGAAGUGUGACAUUCCGUCUCUCUCCAACGGGCGUGAUCAGCGGGUCCACGUU CUCUAUCGGCGGGCGUUAUUAUCUGUCCUGCAAGUCGGGCUUACCCUGACCGG GUCCCCUCUCCACUUGCAUCGACGGCAAGUGGAACCCAUCCUGCCCACCUUGCG UGCGGAGCAACGAGGAGUUCGACCCCGUGGAUGACGGGCCUGAUGAUGAGACUG AUCUCUCUAAGCUGUCUAAGGACGUGGUCCAGUACGAGCAGGAGAUCCAGUCGC UGGAGGCCACCUACCACAUCAUCAUUGGCCCUACCAUCAUGGGGGUGAUCUU CUUGAUCAGCAUCAUCGUGCUGGUGUGUUCUUGCGACAAGAACAACGACCAGUA CAAGUUUCACAAGCUGCUGCCUUGA
16	1401 RNA	ORF	AUGAAGACCAUCUCCGUGGUGACCCUGCUGUGCGUGCUGCCAGCAGUGGUGUAC AGCACCUGUACGGUGCCGACGAUGAAUAACGCCAAGCUGACCUCCACCGAGACGU CGUUAACGACAAGCAGAAGGUGACGUUACCUUGCGACAGCGGCUAUCACAGCUU GGACCCCAACGCCGUCUGCGAGACGGACAAGUGGAAGUACGAGAACCCGUGCAAG AAGAUUGCAGCGUGUCCGACUAUGUGUCCGAGCUGUACGAUAAGCCGUGUAC GAGGUGAACAGCACCAUGACGCGUGCGUACAACGGCGAGACAAAGUACUCCGG UGCGAGGAGAAGAACGGGAACACCAGCUGGAACGAUACAGUACCUUGUCCCAACG CAGAGUGCCAGCCUCUGCAGUUGGAGCACGGGUCCUGUCAGCCCGUUAAGGAGA AGUACAGCUUUGGCGAGUACAUGACGAUUAACUGCGACGUCGGUUAACGAGGUGA UCGGCGUGAGUUAUCUUAUGUACCGCCAACAGCUGGAUUGCAUACCUAGCUG CCAGCAGAAGUGUGACAUUCCGUCUCUCUCCAACGGGCUAUCAGCGGGUCCACG UUCUCUAUCGGCGGGCGUUAUUAUCUGUCUUGCAAGUCCGGCUUACGUUGACC GGGUCGCCUCUGUGACCUUGCAUCGACGGGAAGUGGAUCCAUUCUGCCCACGU GCGUGCGGUCGAACGAGGAGUUCGACCCGGUGCAGAUUGGGCCGGACGACGAGA CAGAUUCUGCAAAGUUGAGCAAGGAUGUGGUUCAGUAUGAGCAGGAGAUUCGAGA GCCUGGAGGCCACGUACCACGGGGGAGGGGGGAGCGGCGGCGGGGAGCGGGG CCGCGGCCAGUAUCCAGACCACCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAA GCUGGAGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAAGAGAUUGG CAAUUAUUAUACAGGCAGAACACGGGUGCAACAUCACCGUGAAGAAUUGUGC UCCGUGACGCCGACGUCAGCUGGACGCCGUGCUGAGCGCGGCCACGGAGACAU AUUCCGGUCUGACCCUUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUG CGCUGAAUUAUCAGACGAGUGUUAACACGGUGGUGCGGGACUUAUGAGAACUACG UGAAGCAGACGUGCAACUCCAGCGCCGUGGUGGAUAAUAAGCUCAAGAUCCAGAA UGUGAUCAUUGACGAGUGUUAUGGGGGCCUUGGAGCCCCACUAACCUAGAUU UAUUAUUAACGGUUCUCCAAGGGGAACUGCGCAUCAAGGCGCUGAUGCAGUU GACAACGAAGGCGACGACCCAGAUUCGCGCCUCGGCAGGUGGCCGGCGGGUGGC GGGUCAGGCGGGGGGGGAGCGCCCGUGAAGACGCCGUGAUCGUGGUGCCC GUGAUCGACCGGCCUCCGAGCGAGACGUUCCCCAACGUGCAUGAGCAUUAACG AUCAGAAGUUCGAUGACGUUAAGGACAACGAGGUGAUGCAGGAGAAGCGGGACG UGGUGAUCGUGAAUGACGACCCGGAUCACUACAAGGACUACGUGUUAUCCAGU GGACCGGAGGCAACAUCGGGAUGACGACAAGUACACCAUUUUUUUCCGGCUU CUGUAACACCAUGUGCAGAGGAGACGAAGCGGAACUUCGUCGCCACCUUGGCA CUGUGGGACUCCAAGUUCUUCACAGAGCUGGAGAACAAGAAUGUGGAGUACGUC

			GUCAUCAUCGAGAAUGAUAAUGUGAUUUGAGGACAUUACAUUUCUCAGGCCGGUG CUCAAGGCCAUCCAUGACAAGAAGAUUGGACAUCCUGCAGAUUGCGGGAGAUCAUCA CCGGUAAUAAGGUGAAGACUGAGCUGGUGAUUUGACAAGGACCACGCCAUUUAC CUAUACCGGUGGCUACGACGUCUCGUGAGCGGUAUAUCAUACGCGUGACCACC GCUCUUAUAUCGUCGACGAGAUCAUUAAGAGCGGUGGUCUGUCCUGGGCUUC UACUUCGAGAUAGCCCGAUUGGAGAACGAGAUAGAUAACCGCCAGAUCAUGG ACAACUCAGCGAAGUACGUCGAGCAUGAUCCCCGCCUGGUGGCCGAGACCGGUU UGAGACAAUGAAGCCCAACUUCUGGUCCCGAUAGGCACCGUGGCCGCCAAGCGG UAUCCGGGCGUCAUGUACACGUUACACCCCCGCUCAUCUAUUCUUGGCCUGU UCGACAUUAACGUAUCGGCCUGAUCGUGAUACUGUUAUCAUGUUAUGCUGA UCUUAACGUGAAGAGCAAGUUGCUGUGGUUCCUGACAGGAACCUUCGUGACUG CCUUAUUUGA
17	1501 RNA	ORF	AUGAAGACCAUCUCGGUGGUGACACUGCUGUGCGUGUUGCCGGCGGUUGUUUA UAGUACGUGCACCUGGCCGACUAUGAACACGCCAAGUUGACGUCUACUGAGACU UCUUUCAUGACAAGCAGAAGGUGACGUUACCUUGCGACAGCGGCUAUCACAGCU UGGACCCCAACCCGUCUGCGAGACGGACAAGUGGAAGUACGAGAACCCGUGCAA GAAGAUUGCAGCGUGUCCGACUAUGUGUCCGAGCUGUACGAUAAGCCGUGUA CGAGGUGAACAGCACCAUGACGCUUGCCUGCAACGGCGAGACAAAGUACUCCGU UGCGAGGAGAAGAACGGCAACACCAGCUGGAAUGACACCGUGACCUGUCCGAACG CCGAGUGCCAGCCCCUGCAGUUGGAGCAUGGCUCUGCCAGCCCGUGAAGGAGAA GUACUCCUUCGGGGAGUACAUGACUAUCAACUGCGACGUGGGCUACGAGGUGAU CGGGGUGUCUUAUAUCAGCUGCACCGCCAAUAGCUGGAACGUGAUUCCAGCUG UCAGCAGAAGUGUGACAUAACUUCUCUGUCCAACGGGCUGAUCAGCGGGAGCAC GUUCAGCAUUGGCGGGGUGAUCCACCUGAGUUGCAAGUCUGGCUUACGCGUAC CGGAUCUCCUAGCUCGACCUGCAUCGACGGGAAGUGGAACCCCAUCCUGCCAACU UGCGUCCGGUCCAACGAGGAGUUCGACCCGUGGACGACGACGACGACGACGAAA CCGAUCUGAGCAAGUUGAGCAAGGAUGUCGUGCAGUACGAGCAGGAGAUUGAGU CCUUGGAGGCGACCUAUCACGUGGGGGAGGGAGCGGCGGCGGCGGCUACGGC UGAAUCAGUGCAUGAGCGCCAACAAGGCCGCCAUACUGAUUACGCCGUGGCCGU CGCGCCGCUUCCUCCACCACCGGAAGGUCGUCUCCAGCACCAACCCAGUAUGAUC ACAAGGAGUCCUGCAACGGGCUGUACUAUCAGGGCAGUUGCUACAUCUGCAGAC CGACUACAAGUCGUUUGAGGAUGCUAAGGCCAACUGCGCCGCCGAAAGCUCCACG CUGCCGAAUAAGUCGGAUGUCCUGACCACGUGGCUAUCGACUACGUGGAGGAC ACCUGGGGCUCGGAUGGGAACCCCAUCACGAAGACCACGACGACUACAGGACA GCGAUGUGAGCCAGGAGGUGCGUAAGUACUUCUGCACCGGGGGCGGAGGGUCU GGAGGAGGCGGACGCGGCGUGCCGCCUCCAUCAGACCACUGUGAACACCCUGU CCGAGCGCAUUCUGUCCAAGCUGGAGCAGGAGGCCAACGCCUCCGCCAGACGAA GUGCGACAUCGAGAUAGGGAAUUCUACAUCGCCAGAACACGGGUGCAACAUC ACCGUGAAGAAUUGUGCUCCGUGACGCCGACGCUACGUGGACGCCGUGCUGA GCGCGCCACGGAGACAUAUCCGGUCUGACCCCGGAGCAGAAGGCUUACGUGCC UGCCAUGUUACGGCGGCCUGAAUAUACAGACUUCUGUGAAUACGGUGGUGCG GGACUUCGAGAAUACGUGAAGCAGACUUGCAACUCCAGCGCCGUGGUGGAUAA UAAGCUCAAGAUCCAGAAUGUGAUCAUUGACGAGUGUUAUGGGGCCCCUGGGAG CCCCACUAACCUCGAGUUUAUUAUACCGGUUCCUCCAAGGGGAACUGCGCCAUC

			AAGGCGCUGAUGCAGCUGACUACUAAGGCCACCACCAGAU CGCCCCACGUCAGG UGGCUGGCACCGGCGUUCAGUUCUACAUGAUCGUGAUCGGGGUGAUUUCUG GCUGCGCUGUUAUGUACUACGCCAAGCGCAUGCUCUUUACUUCGACCAACGAUA AGAUCAAACUUAUCUUGGCGAAUAAGGAGAAUGUGCAUUGGACGACGUACAUGG ACACGUUCUUCGGACGUCUCCUAUGAUCAUUGCGACCACCGAUUAUUCAGAACUG A
18	1601 RNA	ORF	AUGGACUGGACAUGGAUUCUGUUCUGGUGGCCGCCACCAGGGUGCACAGC GACGGCACCCUGUUC CAGGCGACGACGACCUGGCCAUCCCCGUACCGAGUUCU UCAGCACCAAGGCUGCCAAGAACCCCGAGACUAAGCGGGAGGCCAUCGUGAAGGC CUACGGGGGAUGAUAAACGAGGAGACUCUGAAGCAGCGGCUGACCAACCUUGGAGAA GAAGAUUACCAACAUUACCAACAGUUCGAGCAGAUUCGAGAAGUGCUGCAAGCGG AACGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGAGACACUGCGCGCCGCCA UGAUCAGCCUGGCCAAGAAGAU CGACGUGCAGACCGGCCGCCACCCUACGAGGG GGGCGGCGGCUCUGGGGGCGGGGAGCAGCCCCAGACUUC AAGAAGAU CGG GGAUGAUGCCACCAUCAGCUGCUCCCGAACAACCAACUACUACGUGGUGAUG UCCGCCUGGUACAAGGAGCCUAACAGCAUAUCCUGCUGGCCGCCAAGAGCGAUG UGCUGUACUUGACA AUUACCAAGGAUAAGAU CAGCUACGACAGCCCCUACGA CGAUCUGGUGACUACCAUCACCAUCAAGAGCCUGACAGCCGGCGACGCAGGCACC UACAUCUGCGCCUUCUUAUGACCAGCACCACCAUGACACCGACAAGGUGGAUU ACGAGGAGUACUCCAUCGAGCUGAUCGUGAAUACUGACUCAGAGUCUACCAUCGA CAUUAUCCUGUCCGGGAGCACCCCGAGACUACAGUGAGAAGCCUGAGGACAUC GACAACAGCAACUGCAGCAGUGUGUUCGAGAU CACCACCCAGAGCCUAUCACUG ACAACGUGGAUGAUCACACCGACACUGUGACAUAACCCAGUGAUAGCAUAACAC CGUGAAUGCUAGCUCUGGGGAGAGCACCACUGAUGAGAUCCCCGAGCCUAUCACC GACAAGGAGGAGGAUCACACCGUGACCGACACUGUGAGUUAACCCACAGUGUCUA CCUCCUCUGGGAUCGUGACCACCAAGAGCACCACCGACGAUGCCGACCUGUACGA UACAUAACAUGAUAAAGAUACAGUGCCUCCACCACCGUGGGAGGCAGCACCACCA GCAUUAUUAUUAAGACCAAGGACUUCGUGGGGGGCGGGGGGAGCGGGGGCG GGGGGAGCGCCGAAACCAAGGAGUUAAGACCCUGUACAACCUGUUAUCGAUAG CUACCUGCAGAAGCUGGCCCAGCACAGCAU UCCACCAAUUGACAUGCGCCAUU CACAUGGGGAGGUGAUCGGCCAGUUAAGAACUGCGCCUGAGGAUCACCAU AAGUGCAUGAGCAUAGCCGCCUGUCCUUCACUCUGAUGGUGGAGUCCUUAUC GAGGUGAUCUCCUGCUGUCAGAGAAGGACAGGCGGGCUAUUGCCGAGGAGAU GGGAUCGACCUGGACGACGUGCCAGCGUGGUGUCCAAGCUGGAGAAGAACUGC AACGCCUACGCCGAGGUGAACAACAUCAU CGAUUAUCAGAAGCUGAAUAUCGGUG AGUGUUCGCCCCUCCCGCCAGCACAUGCUGCUGCAGAU CGUAUACAGGGAG CGCUGAGGCCAACUGCGGGCUCCAGACCAUCGUGAAGAGCCUGAAUAAGAUUAC GUGCCUCCUAUCAUCGAGAAUCGGCUGCCCUACUACGACCCUGGUUCCUGGUGG GCGUGGCCAUUAUCCUGGUGAUCUUCACCGUGGCCAUUGUAGCAUCAGGCGGA ACCUGGCCCUGAAGUACAGGUACGGGACCUUCCUGUACGUGUGA
19	1701 RNA	ORF	AUGACCCAGCUGCCCAUCCUGCUGCUGCUGAUCAGCCUGGUGUACGCCACUCCA GUCCUCAGACCUCCAAGAAGAUUGGCGACGAUGCCACCAUCAGCUGCUCCCGGAA UAAUACCAACUACU AUGUGGUGAUGAGCGCCUGGUACAAGGAGCCUAACAGCAUC AUCCUGCUGGCCGCCAAGAGCGAUGUGCUGUACUUCGAUAACUACCAAGGACA

			AGAUCAGCUACGAUAGCCCCUACGACGAUCUGGUGACUACCAUACCAUCAAGAG UCUGACUGCCGGCGACGCCGGGACUUAUACUUGCGCCUUCUUAUGACCAGCACC ACCAAUGACACCGACAAGGUGGACUACGAGGAGUACAGCAUCGAGCUGAUUCGUGA ACACCGACAGCGAGAGCACCACUACGAUCAUCCUGUCUGGCAGCACCCCCGAGACU AUCAGCGAGAAGCCUGAGGACAUCGAUAACAGCAACUGUUCAGCGUUUUCGAG AUCACCACCCCGAGCCUACACCGAUAAACGUGGACGAUCACACCGACACCGUGAC CUACACCUCCGAUAGCAUAAACCGUGAAUGCCAGCAGCGGGGAGAGUACUACA GAUGAGAUCCCCGAGCCUACACCGAUAAAGGAGGAGGAUCACACCGUGACCGACA CUGUGAGUUACACCACAGUGUCUACCUCUCCGGCAUCGUGACUACCAAGAGUAC CACCGACGAUGCCGAUCUGUACGAUACCUACAACGAUAACGACACAGUGCCUCCCA CCACCGUGGGAGGCAGCACCACAGCAUCAGCAACUACAAGACCAAGGAUUUUUGU GGGCGGCGGGGCUCCGGCGGCGGAGGCUCCGCCGCGUGAAGACCCCCGUGAU UGUGGUGCCUGUGAUCGAUCGGCCCCUAGCGAGACAUUUCCAACGUGCACGA GCACAUCAACGAUCAGAAGUUCGAUGACGUGAAGGACAACGAGGUGAUGCAGGA GAAGCGGGACGUGGUGAUCGUGAACGACGAUCCUGACCACUACAAGGACUACGU GUUCAUCCAGUGGACCGGGGGCAACAUCCGGGAUGACGACAAGUACCCACUUC UUCUCCGGCUUCUGUAACACCAUGUGUACAGAGGAGACAAAGCGGAACAUUGCU CGGCACCUGGCCUGUGGGACAGCAAGUUCUUCACUGAGCUGGAGAACAAGAAC GUGGAGUAUGUGGUGAUCUACGAGAACGACAAGUGAUCGAGGACAUUACCUUC CUCAGGCCUGUGCUGAAGGCCAUCCACGAUAAGAAGAUCGACAUCCUGCAGAUUC GGGAGAUCAUCACCGGGAAUAAGGUGAAGACCGAGCUGGUGAUCGACAAGGAUC ACGCCAUCUUCACCUACACCGCGGCUAUGAUGUGAGCCUGAGCGCUUAUCAU CCGGGUGACCACCGCCUGAAUAUCGUGGAUGAGAUCAUUAAGUCCGGGGGGCU GUCCAGCGGGUUCUACUUCGAGAUUGCUAGGAUCGAGAACGAGAUAGAAGUCAA CCGGCAGAUCAUGGAUAAUUCUGCCAAGUACGUGGAGCACGACCCUCGGCUGGU GGCUGAGCAUCGCUUCGAGACUAGAAGCCCAACUUCUGGAGCAGGAUUGGGAC CGUGGCCGCCAAGCGGUACCCAGGGGUGAUGUACACCUUUACCACACCCUGAUC AGCGGCGGCGGGGCGAGCGGCGGGGCGGCAGCGCUGAGACAAAGGAGUUAAG ACCUGUACAACCUGUUUAUCGACAGCUAUCUGCAGAAGCUGGCUCAGCACUCUA UUCCACCAUUGUGACAUGCAGCAUUCACAUUGGGGAGGUGAUCGGCCAGUUA AGAACUGCGCCUGAGGAUACCAUAAGUGCAUGAGCAACAGCAGGCUAGAGCUU CACCCUGAUGGUGGAGUCCUUAUCGAGGUGAUCAGCCUGCUGAGCGAGAAGGA CAGGAGGGCCAUCGCCGAGGAGAUCCGCAUCGACCGGAGCAGUUGGCCAGCGUG GUGAGCAAGCUGGAGAAGAAUUGUAACGCCUACGCCGAGGUGAACAAUAUCAUC GAUAUUCAGAAGCUGAAUAUCGGGGAGUGCAGUGCUCCCCUGGGCAGCACAUG CUGCUGCAGAUUGUGAACACCGGCAGCGCCGAGGCCAAUUGCGGGCUGCAGACCA UUGUGAAGUCACUGAAUAAGAUUCACGUGCCACCAUCAUCGAGAACCGGCGGCC AUACUACGAUCCUGGUUCCUGGUGGGGGUGGCCAUUAUUCUGGUGAUCUUCAC AGUGGCCAUUCGAGCAUCAGGCGCAACCUGGCUCUGAAGUACAGGUACGGGACC UUCUGUACGUGUGA
20	1801 RNA	ORF	AUGACCCAGCUGCCUUAUCCUGCUGCUGCUGAUCAGCCUGGUGUACGCUACCCCUA GUCCCCAGACUUCGAAGAAGAUCCGGGACGACGCCACGAUCAGCUGCAGCAGGAA UAACACCAACUACUACGUGGUGAUGAGCGCCUGGUACAAGGAGCCAAUAGCAUC AUUCUGCUGGCCGCCAAGUCUGACGUGCUGUACUUCGACAUAUACACAAAGGACA

			AGAUCUCCUACGACAGUCCUUAUGAUGAUCUGGUGACCACGAUCACCAUUAAGUC UCUGACUGCUGGGGAUGCCGGUACCUAUAUCUGUGCGUUCUUAUGACCUCUAC AACGAACGACACAGAUAAAGGUGGACUACGAGGAGUAUAGCAUCGAGCUGAUCGU GAACACCGACAGCGAGAGCACAUAUGACAUUAUCCUCAGCGGGAGCACCCCGAG ACAAUCUCUGAGAAGCCCGAGGACAUUGAUAAUAGCAAUUGUAGCUCGUGUUC GAGAUUACUACUCCCGAGCCCAUUAACGAUAAUGUGGAUGAUCAUACAGAUACGG UGACCUACACCAGCGACAGCAUCAAACCGUGAAUGCUAGCAGCGGUGAGAGCAC CACCGAUGAGAUCCTCGAGCCUUAUACCGAUAAAGGAGGAGGAUCACACCGUGACC GACACGGUGAGUUACACCACCGUGUCUACCUCUCCGGGAUCGUGACUACCAAGA GCACCACGGAUGACGCCGAUCUGUAUGACACCUACAACGAUAAUGAUACCGUGCC GCCCACCACCGUGGGCGGCAGCACCAACAGCAUCAGCAACUAUAAGACUAAGGACU UCGUGGGGGGCGGGGGGAGCGGCGGCGGCGGCAGCCGCGUGAAUCAGUGCAUG AGCGCCAACAAGGCCGCAUCACUGAUUCAGCCGUGGCUGUCGCCGCCGCGAGCU CCACCCACCGAAGGUGGUGUCCAGCACCAACAGUACGAUCACAAGGAGUCCUG CAAUGGCCUGUAUUACAGGGCUCUUGCUACAUCUCCUGCACAGCGAUUACAAGAGC UUGAGGAUGCCAAGGCCAAUUGUGCAGCCGAGAGCAGCACGCUGCCUAACAAGA GCGACGUGCUGACUACUUGGCUGAUCGACUAUGUGGAGGACACCUUGGGGCGAGUG AUGGGAACCCCAUCACUAAGACCACCUUGACUACCAGGACUCCGACGUGAGUCA GGAGGUGAGGAAGUACUUCUGUACGGGCGGCGGCGGCAGCGGCGGCGGCGGUGC CACUGCACCGUGCCCAUGAACAACGCCAAGCUGACUUCUACCGAGACUUCUU UCAACGACAAGCAGAAGGUGACCUUACCUUGCGACAGCGGCUAUCAUUCUCCUGCA CCCUAACGCCGUGUGCGAGACAGACAAGUGGAAGUAUGAGAACCCGUGCAAGAAG AUGUGUACGGUGUCCGACUACGUCUCUGAGCUGUACGACAAGCCUCUGUACGAG GUCAACUCCACUAUGACUCUGUCUUGCAACGGCGAGACAAAGUACUUUAGGUGC GAGGAGAAGAACGGCAAUACCAGCUGGAACGACACCGUGACGUGCCCCAACGCCG AGUGUCAGCCUCUCCAGCUGGAGCACGGCAGCUGUCAGCCUGUAAGGAGAAGU ACAGCUUCGGCGAGUACAUGACCAUCAACUGUGAUGUGGGGUACGAGGUGAUCG GGGUGAGCUAUUACAGCUGCACCGCCAAUAGCUGGAACGUGAUUCCAGCUGUC AGCAGAAGUGCGACAUCUUUCUGUCCAACGGGCUGAUCAGCGGGAGCACGU UCAGCAUUGGCGGGGUGAUCCACCUAGCUGCAAGUCAGGGUUCACCCUGACUG GCUCACCCAGCAGCACCUCAUCGAUGGCAAGUGGAACCCCAUUCUGCCCACCUG UGUGAGAAGCAACGAGGAGUUCGAUCCUGUGGACGACGGACCAGACGACGAAACC GAUCUGUCCAAGCUGUCCAAGGAUGUGGUGCAGUACGAGCAGGAGAUUCGAGAGC CUGGAGGCCACCUACCACAUCAUCAUAGGCACUGACCAUCAUGGGCGUGAUUAU UCUUGAUCAGCAUCAUUGUGCUGGUGUGUUCUUGCGACAAGAACAACGACCAGU ACAAGUUUCAAGCUGCUGCCUGA
21	1901 RNA	ORF	AUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCGCCACAAGGGUGCACUCC GAUGGGACCUUGUUCUCCUGGGGACGACGACCUGGCCAUCCCCGCUACGGAGUUC UUCAGCACCAAGGUGCCAAGAACCCGGAGACUAAGCGGGAGGCAUCGUAAGG CCUACGGAGAUGACAAUGAGGAGACGCUGAAGCAGCGUCUACCAACCUUGAGAA GAAGAUACCAACAUCACUACGAAGUUUGAGCAGAUUGAGAAGUGCUGCAAGCGC AAUGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGAGACUCUGCGGGCUGCC AUGAUCAGCCUCGCCAAGAAGAUAGCAGUGCAGACCGGAAGGCACCCCUAUGAGG GGGGAGGAGGAGCGGCGGCGGCGGAGCGGGGCCGCGGCCAGUAUCCAGACCA

			CCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAAGCUGGAGCAGGAGGCCAACGC CUCUGCUCAGACCAAGUGUGAUUAGAGAUUGGCAAUUCUAUAUCAGGCAGAA CCACGGGUGCAACAUACCCGUGAAGAAUUGUGCUCGCGUGACGCCGACGCUCAG CUGGACGCCGUGCUGAGCGCGGCCACGGAGACAUUUCGGGUCUGACCCUGAGC AGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCGCUGAAUUAUCCAGACGAGUG UUAACACGGUGGUGCGGGACUUUGAGAACUACGUGAAGCAGACGUGCAACUCCA GCGCCGUGGUGGAUAAUAAGCUCAAGAUCCAGAAUGUGAUCAUUGACGAGUGUU AUGGGGCCCCUGGGAGCCCCACUAACCUCGAGUUUAUUAUACCGGUUCCUCCAA GGGGAACUGCGCCAUCAAGGCGCUGAUGCAGUUGACAACGAAGGCCACCACCCAG AUCGCCCCCAGGCAGGUUGCCGGGGGCGGGGGGGAGCGGUGGCGGGCGGCAGC GCUGAGACAAAGGAGUUAAGACCCUGUACAACUUGUUAUAGAUUCUUAUCUU CAGAAGUUGGCCCAGCACUCCAUCCCUACCAUUGUCACUUGUGCCAUCCACAUUG GAGAGGUGAUCGGCCAGUUAAGAACUGCGCCUGAGGAUACCAACAAGUGCAU GAGCAUAGCCGCCUGUCCUUCACGCUCAUGGUGGAGUCCUUAUCGAGGUGAU CUCCUGCUGAGCGAGAAGGACAGGCGGGCUAUUGCUGAGGAGAUCGGUUAUGA CCUCGACGACGUGCCGUCGUCUCCAAGCUGGAGAAGAACUGCAACGCAUUAU GCUGAGGUGAACAACAUCAUCGAUUAUCAGAAGCUGAAUUCGGUGAGUGUUC GCUCCCCUGGGCAGCAUUGCUGUUGCAGAUUGUCAAUACGGGGAGUGCUGAG GCCAACUGCGGGCUGCAGACCAUCGUGAAGUCACUGAAUAAGAUUCACGUGCCAC CCAUCAUCGAGAACAGGCUGCCGUACUACGACCCGUGGUUCUUGGUGGGGGUGG CCAUAUUCUGGUGAUCUUCACGGUGGCCAUCUGCAGCAUCCGCCGAACCUGGC UCUGAAGUACCGGUACGGGACCUUCCUGUACGUGUGA
22	2001 RNA	ORF	AUGGAUUGGACCUGGAUUCUGUCCUCGUUGCCGCCGCCACCAGGGUGCAUAGU GCGGCAGUGAAGACUCCCGUUAUCGUCGUCGCGGUGAUUGACCGCCGCCUCGG AGACGUUCCCCAACGUGCAUGAGCACAUAACGAUCAGAAGUUCGACGAUGUCA GGACAAUGAGGUGAUCGAGGAGAAGCGGGACGUGGUGAUCGUCAACGACGAUCC CGAUCACUACAAGGAUACGUCUUAUCCAGUGGACGGGCGGCAACAUCGGGAC GACGAUAAGUACACGCAUUCUUCUCCGGCUUCUGUAACACCAUGUGUACAGAG GAGACGAAGCGGAACAUCCGAGGCAUCUGGCGCUGUGGGACUCGAAGUUCUUC ACCGAGCUGGAGAAUAAGAACGUGGAGUACGUGGUGAUCAUAGAGAACGACAAC GUCAUCGAGGACAUACGUUUCUGCGGCCGGUGCUGAAGGCCAUCCACGACAAGA AGAUUCGACAUCCUGCAGAUUGCGGGAGAUCAUCACGGGUAUAAGGUGAAGACCG AGCUGGUGAUUGACAAGGACCACGCGAUUAUACCUACACUGGGGGGAUACGACGU GUCCUUGUCCGCCUAUAUUAUCCGGGUGACCACCGCUCUGAACAUUGGGAUGA GAUCAUCAAGUCGGGAGGCCUCAGCUCGGGCUUCUACUUCGAGAUCCGCCGAUC GAGAACGAGAUAGAUAACCGCCAGAUCAUGGACAACAGCGCCAAGUAUGUCG AGCACGACCCUCGGCUGGUGGCAGAGCACAGAUUCGAGACUAUGAAGCCCAACUU CUGGAGUAGGAUCGGCACUGUCGCCGCCAAGCGCUACCCAGGUGUGAUGUACACC UUCACCACCCCGUGAUCAGCGGGGGGGGGGUGAGCGGCGGGCGGGCAGUGCC GAGACAAAGGAGUUAAGACCCUGUACAACUUGUUAUCGACUCCUACCUCCAGA AGUUGGCUACGACUCGAUCCCCACCAUGUGACAUGCGCAUUCACAUUGGUGA GGUGAUCGGGCGAGUUAAGAACUGUGCUCUGCGGAUUAACCAACAAGUGUAUGAG UAAUUCGCCCGUGUCCUUCACGCUCAUGGUGGAGUCCUUAUCGAGGUGAUCUC CCUGCUGAGCGAGAAGGACAGGCGGGCCAUCGCCGAGGAGAUCCGCAUCGACCUG

			GACGAUGUUCGUCUGUCGUCUCCAAGCUGGAGAAGAACUGCAACGCAUAUGCU GAGGUGAACAAUCAUCAUGAUUUCAGAAAGCUGAAUAUCGGUGAGUGUUCGCU CCCCUGGGCAGCAUAUGCUGUUGCAGAUUCGUGAAUACCGGGAGUGCGGAGGCG AACUGCGGGCUCCAGACCAUCGUGAAGAGCCUGAACAAAGAUUACGUCCCGCCUA UCAUCGAGAAUCGGCUGCCGUACUAUGACCCUGGUGGGCGGCGAGCGGGGGG GGGGGUCGGGGCCGUGCUUCCAUCCAGACCACCGUGAACACUCUGUCUGAGC GUAUCAGCAGCAAGCUGGAGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUGUGA UAUAGAGAUUGGCAAUUUCUAUAUCAGGCAGAACACGGGUGCAACAUCACCGU GAAGAAUAUGUCUCCGCUACGCCGACGCUACGUGGACGCCGUGCUGAGCGCG GCCACGGAGACAUAUUCCGGUCUGACCCUGAGCAGAAGGGCUAUGUGCCUGCCA UGUUCACAGCUGCGCUGAAUAUCCAGACGAGUGUUAACACGGUGGUGCGGGAAU UCGAGAAUUAUGUGAAGCAGACCUGUAACAGCUCGGCUGUUGUGGACAAACAGC UCAAGAUCCAGAAUGUGAUCAUUGACGAGUGUUAACGGGGCCCCCGCUCGCCAC GAAUCUGGAGUUAUCAACACGGGCAGCAGCAAGGGCAACUGCGCCAUAAGGCG CUGAUGCAGCUGACCACUAAGGCCACAACCCAGAUCCGCCUAGGCAGGUGGCCG GGACUGGCGUGCAGUUCUAUAUGAUAGUGAUUGGCGUGAUCAUCCUGGCCGCC UGUUAUGUACUACGCCAAGCGCAUGCUCUUAUUCGACCAACGAUAAGAUCAA ACUUAUCUUGGCGAAUAAGGAGAAUGUGCAUUGGACGACGUACAUGGACACCU CUUUCGCACAUCACCUAUGAUCAUUGCCACGACUGACAUAUCAGAACUGA
23	0201 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTTGGTGGCGGCCGCTAGGGTGACAGCGACG GGACCCTGTTCCAGGGGACGACGACCTGGCCATCCCCGCTACCGAGTTCTTCAGCAC CAAGGCTGCCAAGAACCCCGAGACTAAGCGGGAGGCCATCGTGAAGGCCTACGGGGA CGATAATGAGGAGACGCTGAAGCAGCGTCTACCAATCTGGAGAAGAAGATTACTAAT ATCACCACCAAGTTCGAGCAGATCGAGAAGTGCTGTAAGAGGAATGACGAGGTGCTGT TTCGGCTGGAGAACCACGCCGAGACACTGCGGGCTGCCATGATCTCTTGTCCAAGAA GATCGATGTGCAGACCGGACGGCACCTTATTACCCTTACGACGTGCCCCGATTATGCCG AGTGA
24	0302 DNA	ORF	ATGATGACTCCCGAGAATGACGAGGAGCAGACCAGCGTGTCTCCGCTACCGTGACG GGGACAAGATCCAGGGGAAGAATAAGAGGAAGCGGGTGATCGGCCTGTGCATCAGG ATCAGCATGGTGATCAGCCTGCTGAGCATGATCACCATGTCTGCATTCTGATCGTGCGG CTGAATCAGTGATGAGCGCCAACAAGGCCGCCATCACTGATTACGCCGTGGCCGTGG CCGCCGCTTCTCTACTCACAGGAAGGTGGTGTCTCCACCACTCAGTACGATCACAAG GAGAGCTGCAATGGCCTGTATTACCAGGGCTCTTGCTACATCCTGCACAGCGATTACAA GAGCTTTGAGGATGCCAAGGCCAATTGCGCCGCCGAGAGCAGCACTCTGCCTAACAAAG AGCGACGTGCTGACTACTTGCTGATCGACTATGTGGAGGACACCTGGGGGAGCGATG GGAACCCCATCACAAGACTACCAGCGACTATCAGGACAGCGACGTGAGCCAGGAGG TGCGGAAGTACTTCTGCACCTACCCTTACGACGTGCCTGACTACGCCTGA
25	0402 DNA	ORF	ATGACCCAGCTGCCTATCCTGCTGCTGATCAGCCTGGTGACGCTACCCCTAGTCCC CAGACTTCCAAGAAGATCGGGGACGACGCCACTATCAGCTGCAGCAGGAATAACACCA ACTACTACGTGGTGATGAGCGCCTGGTATAAGGAGCCTAACAGCATCATCCTGCTGGCC GCCAAGAGCGATGTGCTGTACTTCGACAACTATACCAAGGACAAGATCTTTACGATAG CCCCTACGATGACCTGGTGACTACCATCACCATCAAGAGTCTGACCGCCGGGGATGCCG GGACCTATATCTGTGCGTTTTTATGACCTCTACAACAAACGACACAGATAAGGTGGACT ACGAGGAGTATAGCATCGAGCTGATCGTGAACACCGACAGCGAGAGTACCATCGACAT

			CATCCTGAGCGGGAGTACTCCCGAGACTATCAGCGAGAAGCCAGAGGACATCGACAAT AGCAACTGCAGCAGCGTGTTCGAGATCACTACTCCTGAGCCTATTACAGATAACGTGGA TGATCACACCGACACCGTGACCTACACCAGCGACAGCATCAACACCGTGAATGCTAGCA GCGGTGAGAGTACCACGGATGAGATCCCCGAGCCCATCACGGATAAGGAGGAGGATC ACACCGTGACCGACACGGTGAGTTACACCACCGTGTCTACCTCCTCAGGGATTGTGACC ACAAAGTCTACCACCGATGATGCCGATCTGTATGACACATAACAATGACAACGATACCGTG CCGCCCACCACCGTGGGCGGCAGCACAACATCCATTAGTAATAACAAGACTAAGGATTT CGTGAGATCTTTGGGATCACAACCTGATTATCCTGAGCGCCGTGGCGATCTTCTGTAT CACCTACTATATCTGTAATAAGCATCCCAGGAAGTATAAGACCGAGAACAAGGTGTATCC GTACGACGTGCCGACTACGCCTGA
26	0505 DNA	ORF	ATGAAGACCATCTCCGTGGTGACCTGCTGTGCGTGCTGCCAGCAGTGGTGACAGTG GCGGCGGGGGGAGCTACCCCTACGACGTGCCCGACTACGCCACCTGTACGGTGCCAAAC CATGAACAACGCTAAGCTGACCTCCACAGAGACGAGTTTAAATGACAAGCAGAAGGTG ACCTTCACCTGCGACAGTGGCTACCACAGCTTGGACCCCAACGCTGTGTGCGAGACAG ACAAGTGGAAGTACGAGAACCCGTGCAAGAAGATGTGCACCGTGAGTGACTACGTCA GCGAGCTGTACGACAAGCCGCTGTACGAGGTCAACTCCACGATGACACTCTTTGCAAC GGGGAGACAAAGTACTTCAGGTGCGAGGAGAAGAACGGGAATACCAGCTGGAATGAT ACCGTGACATGCCCAATGCCGAGTGCCAGCCCTGCAGCTCGAGCATGGGAGCTGTC AGCCCGTGAAGGAGAAGTACTCCTTCGGGGAGTACATGACCATCAATTGCGACGTGGG GTACGAGGTGATCGGGGTGAGCTACATCTTTGACCCGCAATAGCTGGAACGTGATTC CCAGCTGTACAGAGAAGTGCAGATCCCTTCTGTCCAACGGGCTGATCAGCGGGAG CACGTTGAGCATTGGCGGGGTGATCCACCTGAGCTGCAAGAGTGGGTTACCCCTGACC GGGTCCCCGAGCAGCACCTGCATCGATGGGAAGTGGAACCCGATCCTGCCAATTGCG TCCGGTCCAACGAGGAGTTGACCCCGTGACGACGACGACGACGACGAAACCGATC TGAGCAAGTTGAGCAAGGACGTGGTGAGTACGAGCAGGAGATCGAGAGCCTGGAG GCCACCTACCACATCATCATGCGACTGACCATCATGGGCGTGATATTCTTGATCAGC ATCATTGTGCTGGTGTGTTCTTGCACAGAACAACGACCAAGTACAAGTTTCACAAGCT GCTGCCCTGA
27	0603 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCTACTCGGGTGCATAGCGGCG GCGGGGGGTCTACCCCTACGACGTGCCCGACTATGCCGCCGAGACAAAGGAGTTTAA GACCCTGTACAACCTGTTTATCGACAGCTATCTGCAGAAGCTGGCTCAGCACTTATTC CACCAATGTGACATGCGCCATTACATTGGTGAGGTGATCGGCCAGTTCAAGAAGCTGCG CCCTGAGGATCACGAATAAGTGCATGAGCAACAGCAGGCTGAGCTTACCCTGATGGT GGAGTCCTTCATCGAGGTGATCAGCCTGCTGAGCGAGAAGGACAGGAGGGCCATCGC CGAGGAGATCGGCATCGACCTGGACGATGTGCCAGCGTGGTGAGCAAGCTGGAGAA GAACTGCAACGCATATGCTGAGGTGAACAACATCATCGATATTAGAAAGCTGAATATCG GTGAGTGTTCCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATCGTGAACACGGGCAG TGCCGAGGCCAACTGCGGGCTGCAGACCATCGTGAAGTCACTGAATAAGATCTACGTG CCACCCATCATCGAGAACCGGCTGCCCTACTACGACCCGTGGTTCTGGTGGGGGTGG CCATTATTCTGGTGATCTTACGGTGGCCATCTGCAGCATCCGAGGAACCTGGCCCTG AAGTACAGGTACGGGACCTTCTGTACGTCTGA
28	0702 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTGGTGGCTGCCGCCACCAGGGTGACAGTGGCG GAGGAGGCAGCTACCCCTACGACGTCCCCGACTACGCTGGGGCTGCTGCCTCCATCCA GACCACTGTGAACACCCGTGAGTGAGCGGATCAGCAGCAAGCTGGAGCAGGAGGCCAA

			CGCCTCTGCTCAGACCAAGTGTGATATAGAGATTGGCAATTCTATATCAGGCAGAACCA CGGGTGCAACATCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGAC GCCGTGCTGAGCGCGGCCACGGAGACATATCCGGTCTGACCCCTGAGCAGAAGGCG TATGTGCCTGCCATGTTACAGCTGCACTGAACATCCAGACCAGCGTGAACACCGTGTT GAGGGACTTCGAGAACTACGTGAAGCAGACATGCAACAGCAGCGCGGTGGTGATAA TAAGCTCAAGATCCAGAATGTGATCATTGACGAGTGTTATGGGGCCCTGGGAGCCCCA CTAACCTCGAGTTTATTAATACGGGCAGCAGCAAGGGCAACTGCGCCATCAAGGCCCTG ATGCAGCTGACAACCAAGGCCACCACCCAGATTGCCCCAGGCAGGTGGCCGGTTGA
29	0803 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTGGTGGCCGCCGCCACACGGGTGCATTCCGGCG GCGGCGGCAGCTACCCATACGATGTGCTGACTACGCCGCCCGTGAAGACACCCGT GATCGTGGTGCCCGTGATTGACCGGCCCTTAGCGAGACATTCCCAATGTGCACGAGC ACATCAACGACCAGAAGTTCGACGACGTGAAGGACAACGAGGTGATGCAGGAGAAGA GGGACGTGGTGATTGTGAACGACGATCCTGACCACTACAAGGACTACGTGTTTCATCCA GTGGACAGGCGGGAACATCAGGGACGACGACAAGTATACCACTTCTTCTCCGGCTTC TGTAACACCATGTGTACAGAGGAGACAAAGCGGAACATTGCTCGGCACCTGGCCCTGT GGGACAGCAAGTTCTTACCAGCTGGAGAACAAGAACGTGGAGTATGTGGTGATCAT CGAGAACGACAACGTCATCGAGGACATCACCTTCTGCGCCCTGTGCTGAAGGCCATCC ACGATAAGAAGATCGACATCCTGCAGATGCGGGAGATCATCACCGGAATAAGGTGAA GACCGAGCTGGTGATCGACAAGGATCACGCCATCTTACCTACACCGGCGGTATGATG TGAGCCTGAGCGCTACATCATCCGGGTGACCACCGCCCTGAATATCGTGGATGAGATC ATCAAGTCCGGCGGGCTGAGCAGCGGCTTCTACTTCGAGATCGCCAGAATCGAGAACG AGATGAAGATCAACAGGCAGATCATGGACAACAGCGCCAAGTACGTGGAGACGACC CTCGGCTGGTGGCCGAGCACCGGTTGAGACTATGAAGCCCACTTCTGGTCCAGAAT CGGGACTGTGGCCGCAAGCGGTACCCCGGCGTGATGTACACATTACCAACCCCACTG ATCAGCTTCTTTGGCCTGTTGACATCAACGTGATCGGCCTGATCGTGATCCTGTTTCATC ATGTTTCATGCTGATCTTCAACGTGAAGTCCAAGCTGCTGTGGTTCCTGACCGGCACCTTC GTGACTGCCTTCATCTGA
30	0201-J0 DNA	ORF	ATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGTCCACTCCATGGA CGGAACTCTTTTCCCGGAGATGACGATCTTGCAATTCAGCAACTGAATTTTCTCTAC AAAGGCTGCTAAAAATCCAGAGACTAAACGCGAAGCAATTGTAAAGCCTATGGAGAC GATAATGAGGAACTCTCAAACAACGGCTAACTAATTGGAAAAAAGATTACTAATATA ACAACAAAGTTTGAACAAATAGAAAAGTGTGTAAACGCAACGATGAAGTTCTATTAG GTTGAAAATCACGCTGAACTCTAAGAGCGGCTATGATATCTTGCTAAAAAGATTG ATGTACAGACTGGACGGCATCCATATGAGTAA
31	0301-J0 DNA	ORF	ATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGTCCACTCCATGAT GACACCAGAAAACGACGAAGAGCAGACATCTGTGTTCTCCGCTACTGTTTACGGAGAC AAAATTCAGGGAAAGAATAAACGCAACGCGTGATTGGTCTATGTATTAGAATATCTATG GTTATTTCACTACTATCTATGATTACCATGTCCGCGTTTCTCATAGTGCGCCTAAATCAATG CATGTCTGCTAACAAGGCTGCTATTACTGACTCCGCTGTTGCCGTTGCTGCGGCATCATC TACTCATAGAAAGGTTGTGTCTAGCACTACACAATATGATCACAAGAAAGCTGTAATGG TTTATATTACAGGGTTCTTGTATATATTACATTCAGACTATAAGTCATTCGAGGATGCTA AAGCAAACTGCGCTGCGGAATCATCAACTACCCAATAAATCCGATGTCTTGACTACCT GGCTCATTGATTATGTTGAGGATACATGGGGATCTGATGGTAATCCAATTACAAAACTA CATCCGATTATCAAGATTCTGATGTATCACAAGAAGTTAGAAAAGTATTTTGTACATAA

32	0501-J0 DNA	ORF	ATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGTCCACTCCATGAA AACGATTTCCGTTGTACGTTGTATGCGTACTACCTGCTGTTGTTTATCAACATGTACT GTACCCACTATGAATAACGCTAAATTAACGTCTACCGAAACATCGTTAATGATAAACAG AAAGTTACGTTTACATGTGATTGAGGATATCATTCTTTGGATCCAAATGCTGTCTGCGAA ACAGATAAATGGAAATACGAAAATCCATGCAAGAAAATGTGCACAGTTTCTGATTATGTC TCTGAACATATGATAAGCCATTATACGAAGTGAATTCACCATGACACTAAGTTGCAAC GGTGAAACAAAATATTTTCGTTGTGAAGAAAAAATGGAAATACTTCTTGAATGATAC TGTCACGTGTCCTAATGCGGAATGTCAACCTCTTCAATTAGAACACGGATCGTGTCAACC AGTTAAAGAAAAATACTCATTTGGGGAATATATGACTATCAACTGTGATGTTGGATATGA GGTTATTGGTGTTCGTATATAAGTTGTACGGCTAATTCTTGAATGTTATTCCATCATGTC AACAAAAATGTGATATACCGTCCCTATCTAATGGATTAATTTCCGGATCTACATTTCTATC GGTGGCGTTATACATCTTAGTTGTAAAAGTGGTTTTACTACTAACGGGGTCCCCATCATCC ACATGTATCGACGGTAAATGGAATCCCATCTCCCAACATGTGTACGATCTAACGAAGAA TTTGATCCAGTGGATGATGGTCCCGACGATGAGACAGATCTGAGCAAACCTCGAAAG ACGTTGTACAATATGAACAAGAAATAGAATCGTTAGAAGCAACTTATCATATAATCATAAT GGCGTTGACAATTATGGGTGTCATATTTCTAATCTCCATTATAGTATTAGTTTGTCTCTGTG ACAAAAATAATGACCAATATAAGTTCCATAAATTGCTACCGTGA
33	0701-J0 DNA	ORF	ATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGTCCACTCCATGG GTGCCGAGCAAGCATACAGACGACTGTGAATACACTCAGTGAACGTATCTCGTCTAAA TTAGAACAAGAAGCGAACGCTAGTGCTCAAACAAAATGTGATATAGAAATCGGAAATTT TTATATCCGACAAAACCATGGATGTAACATCACTGTTAAAAATATGTGCTCTGCGGACGC GGATGCTCAGTTGGATGCTGTGTTATCAGCCGCTACAGAAACATATAGTGGATTAACACC GGAACAAAAAGCATACGTACCAGCTATGTTTACTGCTGCGTTAAACATTCAGACGAGTG TAAACACTGTTGTTAGAGATTTTGAATTTATGTGAAACAGACTTGTAATCTAGCGCTG TTGTCGATAACAAATTAAAGATACAAAACGTAATTATAGATGAATGTTACGGAGCCCCAG GATCTCCAACAAATTTGGAATTTATTAATACAGGATCTAGCAAAGGAAATTTGTGCCATTA AGGCGTTGATGCAATTGACTACTAAGGCCACTACTCAATAGCACCTAGACAAGTTGCT GGTACAGGAGTTCAGTTTATATGATTGTTATCGGTGTTATAATATTGGCAGCGTTGTTTA TGTAATATGCCAAGCGTATGCTGTTACATCCACCAATGATAAAATCAAACCTATTTTAGC CAATAAGGAAAACGTCCATTGGACTACTTACATGGACACATTCTTTAGAACTTCTCCGAT GATTATTGCTACCACGGATATACAAAACCTGA
34	0901-J0 DNA	ORF	ATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGTCCACTCCATGCC TCAACAACATATCTCCTATTAATATAGAACTAAAAAAGCAATTTCTGACGCGCGATTGAA GACGTTAGACATACATTATAATGAGTCGAAACCAACCACTATCCAGAATACTGGAAACT AGTAAGGATTAATTTAAAGGAGGATATATAAGTGGAGGGTTTCTCCCCAATGAATATGT GTTATCAACAATACATATATATTGGGGAAAGGAAGACGATTATGGATCCAATCACTTGATA GATGTATACAAATACTCTGGAGAGATTAATCTGGTTCATTGGAATAAGAAAAAATATAGT TCTTATGAAGAGGCCAAAAAACACGATGATGGAATTATCATTATTGCTATATTCTTACAAG TATCAGATCATAAAAATGTATATTTTCAAAAGATAGTCAATCAATTGGATTCCATTAGATCC GCTAATATGTCTGCGCCATTTGATTGAGTATTTTATCTAGACAATTTGTACCTAGTACGTT GGATTATTTTACATATCTAGGAACAACATCAACCACTCTGCAGACGCAGCATGGATAATT TTTCCAACCAATAAACATTCATTCTGATCAACTATCTAAATTCAGAACACTATTGTGCT CGTCTAATCATGAGGGTAAACCTCATTATATAACAGAGAATTATAGAAATCCGTATAAATT GAACGACGACACGCAAGTATATTATTCTGGGGAGATTATACGAGCTGCAACCACCTCAC

			CAGTTCGAGAGAACTATTTTATGAAATGGTTGTCTGATTGAGAGAGGCATGTTTTTCAT ATTATCAAAAATATATCGAAGGGAATAAAACATTGCAATTATTGCCATAGTATTCGTATTT ATACTCACCGCTATTCTCTTTCTTATGAGTCAACGATATTCGCGAGAAAAACAAAACCTAG
35	1101 DNA	ORF	ATGGACTGGACGTGGATTCTCTTCTGGTGGCAGCCGCCACTCGTGTGCACTCCGATGG GACCTTGTTCCTGGGGACGACGACCTGGCCATCCCCGCTACGGAGTTCTTCAGCACC AAGGCTGCCAAGAACCCGGAGACTAAGCGGGAGGCCATCGTCAAGGCCTACGGAGAT GACAACGAGGAGACGCTGAAGCAGCGGCTGACCAACCTGGAGAAGAAGTACAGAA CATCAGACCAAGTTCGAGCAGATTGAGAAGTGCTGCAAGCGGAACGACGAGGTGTT GTTCCGCTTGGAGAACCACGCAGAGACTCTGCGGGCAGCCATGATCAGCTTGGCGAA GAAGATCGACGTGCAGACTGGACGCCACCCCTATGAGGGAGGAGGCGGCAGCGGCG GCGGCGGCAGCCGGCTGAATCAGTGCATGAGCGCCAACAAGGCCGCCATCACTGATTC AGCCGTCGCCGTCGCTGCCGCCTCTCCACTCATAGGAAGGTGGTGTCAGTACGACAC AGTACGACCATAAGGAGAGCTGTAATGGGCTGTACTACCAGGGGTCGTGTACATCCTG CACAGCGACTACAAGTCGTTTGAGGATGCGAAGGCGAACTGCGCAGCCGAGAGCAGC ACGTTGCCTAACAAGAGCGACGTGCTGACTACTTGGCTGATTGACTATGTCGAGGATAC CTGGGGCTCGGATGGGAACCCCATCACGAAGACCACGTCTGACTACCAGGACTCCGAT GTGTCTCAGGAGGTGCGGAAGTACTTCTGTACTGGCGGCGGCGGTTCTGGTGGGGGC GGCAGCACCTGCACCGTCCCCACCATGAACAACGCCAAGCTGACCAGTACAGAGACTT CCTTCAACGATAAGCAGAAGGTGACATTTACATGCGATTCCGGCTACCACTCGTTGGAC CCCAACGCTGTCTGTGAGACAGACAAGTGGAAGTACGAGAATCCATGTAAGAAGATGT GCACGGTGTCCGACTATGTTTCTGAGCTGTACGACAAGCCCCTTTACGAGGTGAACAGC ACCATGACTCTCAGCTGCAACGGCGAGACTAAGTACTTTCGGTGTGAGGAGAAGAACG GTAATACGTCCTGGAACGATACCGTTACTTGCCCCAACGCCGAGTGCCAGCCGTTGCAG CTGGAGCATGGTAGCTGTGAGCCCGTGAAGGAGAAGTATAGTTTCGGCGAGTACATGA CCATTAATTGTGATGTGGGTTACGAGGTCATCGGGGTGAGTTACATCTCTTGCACTGCCA ATAGCTGGAACGTGATTCCCAGCTGTGAGCAGAAGTGTGACATACCTTCTCTGTCCAAC GGGCTGATCAGCGGGAGCACGTTGAGCATTGGCGGGGTGATCCACCTCAGCTGCAAG AGCGGGTTCAGTTGACCGGGTCGCCCTCGTCGACCTGCATCGACGGGAAGTGAATC CCATTCTGCCACGTGCGTGCGGTGCAACGAGGAGTTGACCCGGTGCACGATGGACC CGATGACGAGACTGATCTGTCTAAGTTGTGGAAGGACGTCGTTAGTACGAGCAGGAG ATCGAGTCCCTGGAGGCCACGTACCACGGCGGTGGCGGGAGCGGGGGAGGTGGCAG CAGTCCCCAGACTTCCAAGAAGATCGGGGACGACGCCACCATCTCTGCTCCCGCAAC AACACTAATACTACGTGGTTATGTCCGCCTGGTACAAGGAGCCCAACTCGATCATCCTG CTCGCTGCTAAGAGCGACGTCTGTACTTCGACAATACTACCAAGGACAAGATCAGCTA CGACTACCCCTACGATGACCTCGTGACCACCATCACAATTAAGAGTCTGACCGCGGGG ACGCCGGCACATACATTTGCGCCTTCTTCATGACGAGTACCACCAATGATACCGACAAG GTGGATTATGAGGAGTACTCGATCGAGCTTATCGTGAACACGGACAGTGAGTCGACCAT CGACATCATCCTGTCCGGTAGCACCCCTGAGACAATCTCCGAGAAGCCGGAGGACATC GACAATAGCAATTGCAGTTCCGCTCTTCGAGATCACGACCCCTGAGCCCATTACCGACAA CGTGGATGATCACACCGACACGGTGACGTACACCAGCGACTCTATTAACACCGTCAATG CTTCCAGTGGGGAGTCCACCACTGACGAGATTCCGGAGCCCATACCGACAAGGAGG AGGATCACACCGTGACCGACACGGTGAGTTACACCACCGTGTCTACCTCCTGTTGATT GTGACTACCAAGAGACACCGACGACGCGGATCTGTACGACACGTACAATGACAACG ACACCGTGCCGCCACACCGTGGGCGGGTCCACGACGTCCATTAGTAACTACAAGAC

			TAAGGACTTCGTGGGGGGCGGCGGTTCCGGTGCGGGGGCAGCGCTGAAACCAAGG AGTTCAAGACGCTGTACAACCTGTTTCATCGACTCTTACCTGCAGAAGTTGGCCACGAC TCCATCCCTACCAATGTCACTTGTGCCATCCACATTGGAGAGGTGATCGGCCAGTTCAG AACTGCGCCCTGAGGATCACCAACAAGTGCATGAGCAATAGCCGCCTGTCTTCACGCT CATGGTGGAGTCCTTCATCGAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCGGGCT ATTGCTGAGGAGATCGGTATTGACCTCGACGACGTGCCGTCCGTCTCTCAAGCTGGA GAAGAACTGCAACGCATATGCTGAGGTGAACAACATCATCGATATTCAGAAGCTGAATA TCGGTGAGTGTTCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATTGTCAATACCGGG AGTGCTGAGGCCAACTGCGGGTTGCAGACCATCGTGAAGAGTCTGAACAAGATTTACG TGCCCCGATCATCGAGAACCGCCTGCCCTACTACGACCCCGGTGGGGGCGGGTCTGG AGGAGGCGGCAGCGCGCTGCCGCCTCCATCCAGACCACGGTGAATACGTTGTCCGA GCGGATTTGAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTA TATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCAACATCACCGTGAAGA ATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGA GACATATTCGGGTCTGACCCCTGAGCAGAAGGCGTATGTGCTGCCATGTTACAGCTG CTCTGAACATCCAGACGTCCGTGAATACCGTGGTGAGGGACTTCGAGAACTACGTGAA GCAGACGTGTAATAGCAGCGCTGTCGTCGATAATAAGCTGAAGATCCAGAACGTTATTAT CGACGAGTGCTACGGCGCCCCCGGCTCCCCGACCAACCTGGAGTTCATTAATACCGGTT CCTCCAAGGGGAACTGCGCTATTAAGGCGCTGATGCAGCTGACTACGAAGGCGACGAC CCAGATCGCGCCTCGCCAGGTGGCGGGGACCGGGGTGCAGTTCTACATGATCGTGATC GGTGTGATCATCTGGCAGCTCTCTTATGTACTACGCCAAGCGCATGCTGTTACCAGC ACCAACGACAAGATTAAGCTGATCTCGCCAATAAGGAGAACGTCCACTGGACCACCTA CATGGACACATTCTCCGCACCAGCCCTATGATTATCGCCACCACTGACATTCAGAACTG A
36	1201 DNA	ORF	ATGACCCAGCTGCCTATCCTGCTGCTGCTGATCAGTCTGGTGACGCTACACCGTCTCCC CAGACTTCCAAGAAGATCGGGGACGACGCGACGATCAGCTGCAGCAGGAATAACACC AACTACTACGTGGTGATGAGCGCCTGGTACAAGGAGCCGAATAGCATCATTCTGCTGGC CGCCAAGTCTGACGTGCTGTACTTCGACAATTACACAAAGGACAAGATCTCCTACGACA GTCCTTATGATGATCTGGTGACCACGATCACCATTAAGTCTCTGACTGCTGGGGATGCCG GTACGTATATCTGCGCCTTCTTCATGACGTCCACGACCAATGACACTGATAAGGTGGATT ACGAGGAGTACTCCATCGAGCTGATCGTGAACACCGACAGCGAGTCTACCATCGACATC ATCCTGTCCGGGAGCACCCCCGAGACTATCAGTGAGAAGCCTGAGGATATTGACAACA GCAATTGCAGTAGTGTCTTCGAGATCACCAACCCGGAGCCTATCACTGACAACGTGGAC GATCACACCGACACGGTGACGTACACCAAGTGATAGCATCAACACCGTGAATGCTAGCTC CGGGGAGTCCACGACCGATGAGATCCCGGAGCCTATCACCGATAAGGAGGAGGATCAC ACCGTGACCGACACGGTGAGTTACACCACCGTGTCTACCTCTCCGGGATCGTGACTAC GAAGAGCACTACTGACGATGCTGATCTGTACGATACCTACAACGATAACGATACCGTGCC GCCCACCACCGTGGGCGGCAGCACCACCAGCATCAGCAACTATAAGACTAAGGACTTC GTGGGGGGCGGGGGGAGCGGCGGCGGCGGCAGCCGGCTGAATCAGTGTCATGAGCG CCAACAAGGCCGCCATCACTGATTACGCCGTGGCTGTCGCCGCCGCCAGCTCCACCCAC CGGAAGGTGGTGTCCAGCACCACCCAGTACGATCACAAGGAGTCTGCAATGGCCTGT ATTACCAGGGCTCTTGCTACATCCTGCACAGCGATTACAAGAGCTTTGAGGATGCCAAG GCCAATTGTGCAGCCGAGAGCAGCAGCTGCCTAACAAGAGCGACGTGCTGACTACTT GGCTGATCGACTATGTGGAGGACACCTGGGGGAGTGATGGGAACCCCATCACTAAGAC

			CACGTCCGACTATCAGGACAGTGACGTGAGTCAGGAGGTGCGGAAGTACTTCTGCACC GGCGGCGGGGCGAGCGGGGGGGCGGCAGCGCTGAGACAAAGGAGTTCAAGACCC TTTACAACTTGTTTCATCGATAGTTATCTGCAGAAGCTGGCCCAGCACAGCATCCCCACCA ATGTGACATGCGCCATTACATTGGTGAGGTGATTGGCCAGTTCAAGAACTGCGCCCTG CGGATAACTAACAAGTGATGTCTAATTCCCGCCTGTCCTTCACGCTCATGGTGAGTCC TTCATCGAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCGGGCTATCGCCGAGGAG ATCGGCATCGACCTGGATGACGTCCCCAGCGTGGTGCCAAGCTGGAGAAGAAGTGA ACGCATATGCTGAGGTGAACAACATCATCGATATTAGAAGCTGAATATCGGTGAGTGT CCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATCGTCAATACGGGTCTGCCGAGGCC AACTGCGGGCTGCAGACCATCGTGAAGTCACTGAATAAGATCTACGTGCCACCCATCAT CGAGAACCGGCTGCCCTACTACGACCCGTGGTTCTGGTGCGGGGTGGCCATTATTCTG GTGATCTTCACGGTGGCCATCTGCAGCATCCGAGGAACCTGGCCCTGAAGTACCGGTA CGGTACCTTCTCTACGTGTGA
37	1301 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCCACAAGGGTGCACTCCGATG GGACCTTGTTCCCTGGGGACGACGACCTGGCCATCCCCGCTACGGAGTTCTTCAGCAC CAAGGCTGCCAAGAACC CGGAGACTAAGCGGGAGGCCATCGTCAAGGCCCTACGGAGA TGACAATGAGGAGACGCTGAAGCAGCGTCTACCAACCTTGAGAAGAAGATACCAAC ATCACTACGAAGTTTGAGCAGATTGAGAAGTGCTGCAAGCGCAATGATGAGGTGCTGT TTCGGCTGGAGAACCACGCCGAGACTCTGCGGGCTGCCATGATCAGCCTCGCCAAAGAA GATCGACGTGCAGACCGGAAGGCACCCCTATGAGGGGGGAGGAGGGAGCGGCGGCG GCGGGAGCGGGGCCGCGGCCAGTATCCAGACCACCGTGAACACTCTGTCTGAGCGTAT CAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTGATATAGA GATTGGCAATTTCTATATCAGGCAGAACCGGGTGCAACATCACCGTGAAGAATATGT GCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGAGACAT ATTCCGGTCTGACCCCTGAGCAGAAGGCGTATGTGCCTGCCATGTTACAGCTGCGCTG AATATCCAGACGAGTGTTAACACGGTGGTGCGGGACTTTGAGAAGTACGTGAAGCAGA CGTGCAACAGCTCAGCGGTGGTGATAATAAGCTCAAGATCCAGAATGTGATCATTGAC GAGTGTTATGGGGCCCCCTGGGAGCCCCACTAACCTCGAGTTTATTAATACGGGCAGCAG CAAGGGCAACTGCGCTATCAAGGCGCTGATGCAGCTGACAACGAAGGCCACCACCCA GATTGCTCCGCGGCAGGTGGCAGGAGGGGGAGGGGGGTCCGGCGGGGGAGGCTCG ACTTGACGGTGCCGACGATGAATAACGCCAAGCTGACCTCCACCGAGACGTGCTTCA ACGACAAGCAGAAGGTGACGTTACCTGCGACAGCGGCTATCACAGCTTGGACCCCAA CGCCGTCTGCGAGACGGACAAGTGGAAGTACGAGAACCCGTGCAAGAAGATGTGCAC GGTGTCGACTATGTGTCCGAGCTGTACGATAAGCCGCTGTACGAGGTGAACAGCACC ATGACGCTGTGTCGAACGGCGAGACAAAGTACTTCGGTGCGAGGAGAAGAACGGG AACACCAGCTGGAACGATACAGTGACCTGTCCCAACGCAGAGTGCCAGCCTCTGCAGT TGGAGCACGGGTCTGTGAGCCCGTTAAGGAGAAGTACAGCTTTGGCGAGTACATGAC GATTAAGTGCAGCTCGGTTACGAGGTGATCGGCGTGAGTTACATCTCATGTACCGCCA ACAGCTGGAATGTCATACCTAGCTGCCAGCAGAAGTGTGACATTCCGTCTCTCTCAAC GGGCTGATCAGCGGTCCACGTTCTCTATCGGCGGCGTTATTCATCTGTCTGCAAGTC GGGCTTACCCTGACCGGTCCCCCTTCCACTGTCATCGACGGCAAGTGGAACCCC ATCCTGCCACCTGCGTGCGGAGCAACGAGGAGTTCGACCCGTGGATGACGGGCTG ATGATGAGACTGATCTCTAAGCTGTCTAAGGACGTGGTCCAGTACGAGCAGGAGATC GAGTCGCTGGAGGCCACCTACCACATCATCATCATGGCCCTACCATCATGGGGGTGAT

			CTTCTTGATCAGCATCATCGTGCTGGTGTGTTCTTGCGACAAGAACAACGACCAGTACA AGTTTCACAAGCTGCTGCCTTGA
38	1401 DNA	ORF	ATGAAGACCATCTCCGTGGTGACCCTGCTGTGCGTGCTGCCAGCAGTGGTGTACAGCAC CTGTACGGTGCCGACGATGAATAACGCCAAGCTGACCTCCACCGAGACGTCGTTCAAC GACAAGCAGAAGGTGACGTTACCTGCGACAGCGGCTATCACAGCTTGGACCCCAACG CCGTCTGCGAGACGGACAAGTGGAAGTACGAGAACCCGTGCAAGAAGATGTGCACGG TGTCGACTATGTGTCCGAGCTGTACGATAAGCCGCTGTACGAGGTGAACAGCACCATG ACGCTGTCGTGCAACGGCGAGACAAAGTACTTCCGGTGCGAGGAGAAGAACGGGAAC ACCAGCTGGAACGATACAGTGACCTGTCCCAACGCAGAGTGCCAGCCTCTGCAGTTGG AGCACGGGTCCTGTGACCCCGTTAAGGAGAAGTACAGCTTTGGCGAGTACATGACGAT TAACTGCGACGTCGGTTACGAGGTGATCGGCGTGAGTTACATCTCATGTACCGCCAACA GCTGGAATGTCATACCTAGTGCCAGCAGAAGGTGTGACATTCGCTCTCTCTCAACGGG CTGATCAGCGGGTCCACGTTCTCTATCGGCGGCGTTATTCATCTGTCTTGCAAGTCCGGC TTCACGTTGACCGGGTCGCCCTCGTCGACCTGCATCGACGGGAAGTGGAATCCATTCT GCCCACGTGCGTGCGGTGCAACGAGGAGTTCGACCCGGTCGACGATGGGCCGGACGA CGAGACAGATCTGTCAAAGTTGAGCAAGGATGTGGTTCAGTATGAGCAGGAGATCGAG AGCCTGGAGGCCACGTACCACGGGGGAGGGGGGAGCGGCGGCGGGGAGCGGG GCCGCGGCCAGTATCCAGACCACCGTGAACACTCTGTCTGAGCGTATCAGCAGCAAGC TGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTGATATAGAGATTGGCAATTT CTATATCAGGCAGAACACGGGTGCAACATCACCGTGAAGAATATGTGCTCCGCTGACG CCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGAGACATATTCGGTCTGAC CCCTGAGCAGAAGGCGTATGTGCCTGCCATGTTACAGCTGCGCTGAATATCCAGACGA GTGTTAACACGGTGGTGCGGGACTTTGAGAACTACGTGAAGCAGACGTGCAACTCCAG CGCCGTGGTGGATAATAAGCTCAAGATCCAGAATGTGATCATTGACGAGTGTATGGGG CCCCTGGGAGCCCCACTAACCTCGAGTTTATTAATACCGGTTCTCTCAAGGGGAAGTGC GCCATCAAGGCGCTGATGCAGTTGACAACGAAGGCGACGACCCAGATCGCGCCTCGG CAGGTGGCCGGCGGCGGTGGCGGGTCAGGCGGCGGGGGGAGCGCCGCCGTGAAGA CGCCGGTGATCGTGGTGGCCGTGATCGACCGGCTCCGAGCGAGACGTTCCTCAACGT GCATGAGCATATCAACGATCAGAAGTTCGATGACGTTAAGGACAACGAGGTGATGCAG GAGAAGCGGGACGTGGTGATCGTGAATGACGACCCGGATCACTACAAGGACTACGTGT TTATCCAGTGGACCGGAGGCAACATCCGGGATGACGACAAGTACACCCATTTTCTCC GGCTTCTGTAACACCATGTGCACAGAGGAGACGAAGCGGAACATCGCTCGCCACCTGG CACTGTGGGACTCCAAGTTCTTCACAGAGCTGGAGAACAAGAATGTGGAGTACGTCGT CATCATCGAGAATGATAATGTGATTGAGGACATTACATTTCTAGGCCGGTGCTCAAGGC CATCCATGACAAGAAGATCGACATCCTGCAGATGCGGGAGATCATCACCGGTAATAAGG TGAAGACTGAGCTGGTCATTGACAAGGACCACGCCATCTTACCTATACCGGTGGCTAC GACGTCTCGCTGAGCGGTATATCATACGCGTGACCACCGCTCTTAATATCGTCGACGAG ATCATTAAAGAGCGGTGGTCTGTCTCGGGCTTCTACTTCGAGATAGCCCGGATCGAGAA CGAGATGAAGATCAACCGCCAGATCATGGACAACCTCAGCGAAGTACGTCGAGCATGAT CCCCGCCTGGTGGCCGAGACCGGTTTGAGACAATGAAGCCCAACTTCTGGTCCCGGA TAGGCACCGTGGCCGCCAAGCGGTATCCGGGCGTCATGTACACGTTACCAACCCCGCTC ATCTCATTCTTCGGCCTGTTTCGACATTAACGTCATCGGCCTGATCGTGATACTGTTTCATCA TGTTTCATGCTGATCTTCAACGTGAAGAGCAAGTTGCTGTGGTTCCTGACAGGAACCTTC GTGACTGCCTTCATTGA

39	1501 DNA	ORF	ATGAAGACCATCTCGGTGGTGACACTGCTGTGCGTGTGCGGCGGTTGTTTATAGTAC GTGCACCGTGCCGACTATGAACAACGCCAAGTTGACGTCTACTGAGACTTCTTTCAATG ACAAGCAGAAGGTGACGTTACCTGCGACAGCGGCTATCACAGCTTGGACCCCAACGC CGTCTGCGAGACGGACAAGTGAAGTACGAGAACCCGTGCAAGAAGATGTGCACGGT GTCCGACTATGTGTCGAGCTGTACGATAAGCCGCTGTACGAGGTGAACAGCACCATGA CGCTGTCTGCAACGGCGAGACAAAGTACTTCCGTTGCGAGGAGAAGAACGGCAACA CCAGCTGGAATGACACCGTGACCTGTCCGAACGCCGAGTGCCAGCCCTGCAGTTGGA GCATGGCTCCTGCCAGCCCGTGAAGGAGAAGTACTCCTTCGGGGAGTACATGACTATC AACTGCGACGTGGGCTACGAGGTGATCGGGGTGTCTTATATCAGCTGCACCGCCAATAG CTGGAACGTGATTCCCAGCTGTGACGAGAAGTGTGACATACCTTCTCTGTCCAACGGGC TGATCAGCGGGAGCACGTTGAGCATTGGCGGGGTGATCCACCTGAGTTGCAAGTCTGG CTTCACGCTGACCGGATCTCCTAGCTCGACCTGCATCGACGGGAAGTGGAAACCCATCC TGCCAACTTGCCTCCGGTCCAACGAGGAGTTCGACCCCGTGACGACGGACCAGACG ACGAAACCGATCTGAGCAAGTTGAGCAAGGATGTCGTGAGTACGAGCAGGAGATCG AGTCCCTGGAGGCGACCTATCACGTTGGGGGAGGGAGCGGCGGCGGCTCACGG CTGAATCAGTGCATGAGCGCCAACAAGGCCGCATCACTGATTCAGCCGTGGCCGTCG CCGCCGCTTCTCCACCCACCGGAAGGTGCTCTCCAGCACCCAGTATGATCACAAG GAGTCTGCAACGGGCTGTACTATCAGGGCAGTTGCTACATCCTGCACAGCGACTACAA GTCTTTGAGGATGCTAAGGCCAACTGCGCCGCCGAAAGCTCCACGCTGCCGAATAAG TCGGATGTCTGACCAGTGGCTCATCGACTACGTGGAGGACACCTGGGGCTCGGATG GGAACCCCATCAGGAAGACCACCGACTATCAGGACAGCGATGTGAGCCAGGAGG TGCGTAAGTACTTCTGCACCGGGGGCGGAGGGTCTGGAGGAGGCGGACGCGCGCT GCCGCTCCATCCAGACCACTGTGAACACCCTGTCCGAGCGCATTCGTCCTCAAGCTGGA GCAGGAGGCCAACGCCTCCGCCAGACGAAGTGCGACATCGAGATAGGGAATTTCTAC ATCCGCCAGAACCACGGGTGCAACATCACCGTGAAGAATATGTGCTCCGCTGACGCCG ACGCTCAGCTGGACGCCGTGCTGAGCGCGCCACGGAGACATATCCGGTCTGACCCC GGAGCAGAAGGCTTACGTGCCTGCCATGTTTACGGCGGCCCTGAATATACAGACTTCTG TGAATACGGTGGTGCGGGACTTCGAGAATTACGTGAAGCAGACTTGCAACTCCAGCGC CGTGGTGATAATAAGCTCAAGATCCAGAATGTGATCATTGACGAGTGTATGGGGCCC CTGGGAGCCCCACTAACCTCGAGTTTATTAATACCGGTTCTCCAAGGGGAAGTGCGCC ATCAAGGCGCTGATGCAGCTGACTACTAAGGCCACCCAGATCGCCCCAGTCAGGT GGCTGGCACCGGCGTTGAGTTCTACATGATCGTGATCGGGGTGATCATTCTGGCTGCGC TGTTTCATGTACTACGCCAAGCGCATGCTCTTACTTCGACCAACGATAAGATCAAATTAT CTTGCGCAATAAGGAGAATGTGATTGGACGACGTACATGGACACGTTCTCCGGACG TCTCCTATGATCATTGCGACCACCGATATTCAGAACTGA
40	1601 DNA	ORF	ATGGACTGGACATGGATTCTGTTCTGTTGGCCGCCGCCACCAGGGTGACAGCGACG GCACCCTGTTCCAGGCGACGACGACCTGGCCATCCCCGCTACCGAGTTCTTCAGCACC AAGGCTGCCAAGAACCCCGAGACTAAGCGGGAGGCCATCGTGAAGGCCTACGGGGAT GATAACGAGGAGACTCTGAAGCAGCGGCTGACCAACCTGGAGAAGAAGATTACCAAC ATTACCACCAAGTTCGAGCAGATCGAGAAGTCTGCAAGCGGAACGATGAGGTGCTGT TTCGGCTGGAGAACCACGCCGAGACACTGCGCGCCGCCATGATCAGCCTGGCCAAGA AGATCGACGTGCAGACCGGCCGCCACCCCTACGAGGGGGGCGGCGGCTCTGGGGGC GGCGGGAGCAGCCCCAGACTTCCAAGAAGATCGGGGATGATGCCACCATCAGCTGCT CCCGAACAACCAACTACTACGTGGTGATGTCCGCTGGTACAAGGAGCCTAACAG

			CATCATCCTGCTGGCCGCAAGAGCGATGTGCTGTACTTTGACAATTACACCAAGGATA AGATCAGCTACGACAGCCCTACGACGATCTGGTGA CTACCATCACCATCAAGAGCCTG ACAGCCGGCGACGCAGGCACCTACATCTGCGCTTCTTCATGACCAGCACCACCAATGA CACCGACAAGGTGGATTACGAGGAGTACTCCATCGAGCTGATCGTGAATACTGACTCAG AGTCTACCATCGACATTATCCTGTCCGGGAGCACCCCGAGACTATCAGTGAGAAGCCT GAGGACATCGACAACAGCAACTGCAGCAGTGTGTTGAGATCACCACCCAGAGCCTA TCACTGACAACGTGGATGATCACACCGACACTGTGACATACACCAGTGATAGCATCAAC ACCGTGAATGCTAGCTCTGGGGAGAGCACCAGTATGAGATCCCCGAGCCTATCACCG ACAAGGAGGAGGATCACACCGTGACCGACACTGTGAGTTACACCACAGTGTCTACCTC CTCTGGGATCGTGACCACCAAGAGCACCACCGACGATGCCGACCTGTACGATACATACA ATGATAACGATACAGTGCCTCCCACCACCGTGGGAGGCAGCACCACCAGCATTTCTAAT TACAAGACCAAGGACTTCGTGGGGGGCGGGGGAGCGGGGGCGGGGGGAGCGCCG AAACCAAGGAGTTCAAGACCCTGTACAACCTGTTTCATCGATAGCTACCTGCAGAAGCTG GCCCAGCACAGCATTCCACCAATGTGACATGCGCCATTACATTGGGGAGGTGATCGG CCAGTTCAAGAACTGCGCCTGAGGATCACCAATAAGTGCATGAGCAATAGCCGCTGT CCTTCACTCTGATGGTGGAGTCCTTCATCGAGGTGATCTCCCTGCTGTGAGAGAAGGAC AGGCGGGCTATTGCCGAGGAGATCGGGATCGACCTGGACGACGTGCCAGCGTGGTG TCCAAGCTGGAGAAGAACTGCAACGCCTACGCCGAGGTGAACAACATCATCGATATCA GAAGCTGAATATCGGTGAGTGTCCGCCCTCCCGGCCAGCACATGCTGCTGCAGATCG TGAATACAGGGAGCGCTGAGGCCAACTGCGGGCTCCAGACCATCGTGAAGAGCCTGA ATAAGATTACGTGCCTCTATCATCGAGAATCGGCTGCCCTACTACGCCCTGGTTCTCT GGTGGGCGTGCCATTATCCTGGTGATCTTACCCTGGCCATCTGTAGCATCAGGCGGA ACCTGGCCCTGAAGTACAGGTACGGGACCTTCTGTACGTGTGA
41	1701 DNA	ORF	ATGACCCAGCTGCCATCCTGCTGCTGCTGATCAGCCTGGTGACGCCACTCCCAGTCTCT CAGACCTCCAAGAAGATTGGCGACGATGCCACCATCAGCTGCTCCCGGAATAATACCAA CTACTATGTGGTGATGAGCGCCTGGTACAAGGAGCCTAACAGCATCATCTGCTGGCCG CCAAGAGCGATGTGCTGTACTTCGATAACTACCAAGGACAAGATCAGCTACGATAGC CCCTACGACGATCTGGTGACTACCATCACCATCAAGAGTCTGACTGCCGGCGACGCCG GACTTACATCTGCGCTTCTTCATGACCAGCACCACCAATGACACCGACAAGGTGGACT ACGAGGAGTACAGCATCGAGCTGATCGTGAACACCGACAGCGAGAGCACCATCGACAT CATCCTGTCTGGCAGCACCCCGAGACTATCAGCGAGAAGCCTGAGGACATCGATAACA GCAACTGTTCCAGCGTTTTCGAGATCACCACCCCGAGCCTATCACCATAACGTGGAC GATCACACCGACACCGTGACCTACACCTCCGATAGCATCAACACCGTGAATGCCAGCAG CGGGGAGAGTACTACAGATGAGATCCCCGAGCCTATCACCATAAGGAGGAGGATCAC ACCGTGACCGACACTGTGAGTTACACCACAGTGTCTACCTCCTCCGGCATCGTGACTAC CAAGAGTACCACCGACGATGCCGATCTGTACGATACCTACAACGATAACGACACAGTGC CTCCCACCACCGTGGGAGGCAGCACCACCAGCATCAGCAACTACAAGACCAAGGATT TGTGGGCGGGGGGGCTCCGGCGGGGAGGCTCCGCCGCGGTGAAGACCCCGTGA TTGTGGTGCTGTGATCGATCGGCCCTAGCGAGACATTTCCAACGTGCACGAGCAC ATCAACGATCAGAAGTTCGATGACGTGAAGGACAACGAGGTGATGAGGAGAAGCGG GACGTGGTGATCGTGAACGACGATCCTGACCACTACAAGGACTACGTGTTTCATCCAGTG GACCGGGGGCAACATCCGGGATGACGACAAGTACACCCACTTCTTCTCCGGCTCTGTA ACACCATGTGTACAGAGGAGACAAAGCGGAACATTGCTCGGCACCTGGCCCTGTGGG ACAGCAAGTTCTTCACTGAGCTGGAGAACAAGACGTGGAGTATGTGGTGATCATCGA

			GAACGACAATGTGATCGAGGACATTACCTTCCTCAGGCCTGTGCTGAAGGCCATCCACG ATAAGAAGATCGACATCCTGCAGATGCGGGAGATCATCACCGGGAATAAGGTGAAGAC CGAGCTGGTGATCGACAAGGATCACGCCATCTTCACCTACACCGGCGGCTATGATGTGA GCCTGAGCGCTTACATCATCCGGGTGACCACCGCCCTGAATATCGTGGATGAGATCATT AGTCCGGGGGGCTGTCCAGCGGTTCTACTTCGAGATTGCTAGGATCGAGAACGAGAT GAAGATCAACCGGCAGATCATGGATAATTCTGCCAAGTACGTGGAGCACGACCCTCGG CTGGTGGCTGAGCATCGCTTCGAGACTATGAAGCCCAACTTCTGGAGCAGGATTGGGA CCGTGGCCGCCAAGCGGTACCCAGGGGTGATGTACACCTTTACCACACCCCTGATCAGC GGCGGCGGGGCGAGCGGCGGGGGCGGCAGCGCTGAGACAAAGGAGTTCAAGACCC TGTAACCTGTTTATCGACAGCTATCTGCAGAAGCTGGCTCAGCACTCTATCCACCA ATGTGACATGCGCCATTACATTGGGGAGGTGATCGGCCAGTTCAAGAAGTGCGCCCT GAGGATCACCAATAAGTGCATGAGCAACAGCAGGCTGAGCTTCACCCTGATGGTGGAG TCCTTCATCGAGGTGATCAGCCTGCTGAGCGAGAAGGACAGGAGGGCCATCGCCGAG GAGATCGGCATCGACCTGGACGATGTGCCAGCGTGGTGAGCAAGCTGGAGAAGAAT TGTAACGCCTACGCCGAGGTGAACAATATCATCGATATTCAGAAGCTGAATATCGGGGA GTGAGTGCTCCCCCTGGGCAGCACATGCTGCTGCAGATTGTGAACACCGGCAGCGCC GAGGCCAATTGCGGGCTGCAGACCATGTGAAGTCACTGAATAAGATCTACGTGCCACC CATCATCGAGAACCGGCTGCCATACTACGATCCCTGGTTCTGGTGGGGGTGGCCATTA TTCTGGTGATCTTCACAGTGGCCATCTGCAGCATCAGGCGCAACCTGGCTCTGAAGTAC AGGTACGGGACCTTCCTGTACGTGTGA
42	1801 DNA	ORF	ATGACCCAGCTGCCTATCCTGCTGCTGCTGATCAGCCTGGTGTACGCTACCCCTAGTCCC CAGACTTCCAAGAAGATCGGGGACGACGCCACGATCAGCTGCAGCAGGAATAACACC AACTACTACGTGGTGATGAGCGCCTGGTACAAGGAGCCAAATAGCATCATTCTGCTGGC CGCCAAGTCTGACGTGCTGTACTTCGACAATTACACAAAGGACAAGATCTCCTACGACA GTCCTTATGATGATCTGGTGACCACGATCACCATTAAGTCTCTGACTGCTGGGGATGCCG GTACCTATATCTGTGCGTTCTTTATGACCTCTACAACGAACGACACAGATAAGGTGGACT ACGAGGAGTATAGCATCGAGCTGATCGTGAACACCGACAGCGAGAGCACAATTGACAT TATCCTCAGCGGGAGCACCCCGAGACAATCTCTGAGAAGCCCGAGGACATTGATAATA GCAATTGTAGCTCCGTGTTTCGAGATTACTACTCCCGAGCCATTACCGATAATGTGGATG ATCATACAGATACGGTGACCTACACCAGCGACAGCATCAACACCGTGAATGCTAGCAGC GGTGAGAGCACCACCGATGAGATCCCGGAGCCTATCACCGATAAGGAGGAGGATCACA CCGTGACCGACACGGTGAGTTACACCACCGTGTCTACCTCCTCCGGGATCGTGACTACC AAGAGCACACGGATGACGCCGATCTGTATGACACCTACAACGATAATGATACCGTGCC GCCCACCACCGTGGGCGGCAGCACCACGACATCAGCAACTATAAGACTAAGGACTTC GTGGGGGGCGGGGGGAGCGGCGGCGGCGGCAGCCGGCTGAATCAGTGCATGAGCG CCAACAAGGCCGCCATCACTGATTACGCCGTGGCTGTGCGCGCCGCCAGCTCCACCCAC CGGAAGGTGGTGCCAGCACCCAGTACGATCACAAGGAGTCTGCAATGGCCTGT ATTACCAGGGCTCTTGCTACATCCTGCACAGCGATTACAAGAGCTTTGAGGATGCCAAG GCCAATTGTGCAGCCGAGAGCAGCAGCTGCCTAACAAGAGCGACGTGCTGACTACTT GGCTGATCGACTATGTGGAGGACACCTGGGGCAGTGATGGGAACCCCATCTAAGAC CACCTCTGACTACCAGGACTCCGACGTGAGTCAGGAGGTGAGGAAGTACTTCTGTACG GGCGGCGGCGGCAGCGGCGGCGGCGGTCCACCTGCACCGTGCCCACTGAAACAAC GCCAAGCTGACTTCTACCGAGACTTCTTTCAACGACAAGCAGAAGGTGACCTTCACCT GCGACAGCGGCTATCATTCCCTCGACCCTAACGCCGTGTGCGAGACAGACAAGTGAA

			GTATGAGAACCCGTGCAAGAAGATGTGTACGGTGTCCGACTACGTCTCTGAGCTGTACG ACAAGCCTCTGTACGAGGTCAACTCCACTATGACTCTGTCTTGCAACGGCGAGACAAAG TACTTTAGGTGCGAGGAGAAGAACGGCAATACCAGCTGGAACGACACCGTGACGTGC CCCAACGCCGAGTGTACAGCTCTCCAGCTGGAGCACGGCAGCTGTACAGCTGTGAAGG AGAAGTACAGCTTCGGCGAGTACATGACCATCAACTGTGATGTGGGGTACGAGGTGAT CGGGGTGAGCTATATCAGCTGCACCGCCAATAGCTGGAACGTGATTCCAGCTGTCAGC AGAAGTGCGACATCCCTTCTGTGTTCAACGGGCTGATCAGCGGGAGCACGTTACAGCAT TGCGGGGTGATCCACCTGAGCTGCAAGTCAGGGTTCACCTGACTGGCTCACCAGC AGCACCTGCATCGATGGCAAGTGAACCCATTCTGCCACCTGTGTGAGAAGCAACG AGGAGTTCGATCCTGTGGACGACGGACCAGACGACGAAACCGATCTGTCCAAGCTGTC CAAGGATGTGGTGCAGTACGAGCAGGAGATCGAGAGCCTGGAGGCCACCTACCACAT CATCATCATGGCACTGACCATCATGGGCGTGATATTCTTGATCAGCATCATTGTGCTGGT GTGTTCTTGCACAAGAACAACGACCAGTACAAGTTTACAAGCTGCTGCCCTGA
43	1901 DNA	ORF	ATGGACTGGACCTGGATTCTGTTCTGTGGTGGCGGCCGCCACAAGGGTGCACTCCGATG GGACCTTGTTCCCTGGGACGACGACCTGGCCATCCCCGCTACGGAGTTCTTCAGCAC CAAGGCTGCCAAGAACCCGGAGACTAAGCGGGAGGCCATCGTCAAGGCCCTACGGAGA TGACAATGAGGAGACGCTGAAGCAGCGTCTACCAACCTTGAGAAGAAGATACCAAC ATCACTACGAAGTTTGAGCAGATTGAGAAGTGCTGCAAGCGCAATGATGAGGTGCTGT TTCGGCTGGAGAACCACGCCGAGACTCTGCGGGTGCCATGATCAGCCTCGCCAAAGAA GATCGACGTGCAGACCGGAAGGCACCCCTATGAGGGGGGAGGAGGGAGCGGCGGCG GCGGGAGCGGGGCCGCGGCCAGTATCCAGACCACCGTGAACACTCTGTCTGAGCGTAT CAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTGATATAGA GATTGGCAATTTCTATATCAGGCAGAACACGGGTGCAACATCACCGTGAAGAATATGT GCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGAGACAT ATTCCGGTCTGACCCTGAGCAGAAGGCGTATGTGCCTGCCATGTTACAGCTGCGCTG AATATCCAGACGAGTGTTAACACGGTGGTGCGGGACTTTGAGAACTACGTGAAGCAGA CGTGCAACTCCAGCGCGGTGGTGGATAATAAGCTCAAGATCCAGAATGTGATCATTGAC GAGTGTTATGGGGCCCTGGGAGCCCCACTAACCTCGAGTTTATTAATACCGGTTCTCTC AAGGGGAAGTGCGCCATCAAGGCGCTGATGCAGTTGACAACGAAGGCCACCCAG ATCGCCCCAGGCAGGTTGCCGGGGGCGGCGGGGGAGCGGTGGCGGCGGCAGCG CTGAGACAAAGGAGTTCAAGACCTGTACAACTTGTTATAGATTCTTATCTCAGAAGT TGCCCCAGCACTCCATCCCTACCAATGTCACTTGTCATCCACATTGGAGAGGTGATCG GCCAGTTCAAGAACTGCGCCTGAGGATACCAACAAGTGCATGAGCAATAGCCGCT GTCCTTCACGCTCATGGTGGAGTCCTTCATCGAGGTGATCCTGCTGAGCGAGAAGG ACAGGCGGGCTATTGCTGAGGAGATCGGTATTGACCTCGACGACGTGCCGTCCGTCGT CTCCAAGCTGGAGAAGAAGTGAACGCATATGCTGAGGTGAACAACATCATGATATTC AGAAGCTGAATATCGGTGAGTGTTCCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATT GTCAATACCGGGAGTGCTGAGGCCAACTGCGGGCTGCAGACCATCGTGAAGTCACTGA ATAAGATCTACGTGCCACCCATCATCGAGAACAGGCTGCCGTACTACGACCCGTGGTTCT TGGTGGGGGTGGCCATTATTCTGGTGATCTTACGGTGGCCATCTGCAGCATCCGCCGG AACCTGGCTCTGAAGTACCGGTACGGGACCTTCTGTACGTGTGA
44	2001 DNA	ORF	ATGGATTGGACCTGGATTCTGTTCTCGTTGCCGCCGCCACCAGGGTGATAGTGGCGC AGTGAAGACTCCCGTTATCGTCGTCGCCGTGATTGACCGCCCGCCCTCGGAGACGTTCC CCAACGTGCATGAGCACATCAACGATCAGAAGTTCGACGATGTCAAGGACAATGAGGT

		CATGCAGGAGAAGCGGGACGTGGTGATCGTCAACGACGATCCCGATCACTACAAGGAT TACGTCTTCATCCAGTGGACGGGCGGCAACATCCGGGACGACGATAAGTACACGCATTT CTTCTCCGGCTTCTGTAACACCATGTGTACAGAGGAGACGAAGCGGAACATCGCCAGG CATCTGGCGCTGTGGGACTCGAAGTTCTTCACCGAGCTGGAGAATAAGAACGTGGAGT ACGTGGTGATCATCGAGAACGACAACGTCATCGAGGACATCACGTTTCTGCGGCCGGT GCTGAAGGCCATCCACGACAAGAAGATCGACATCCTGCAGATGCGGGAGATCATCACG GGTAATAAGGTGAAGACCGAGCTGGTCATTGACAAGGACCACGCGATATTACCTACAC TGGGGGATACGACGTGTCTTGTCCGCCTATATTATCCGGGTGACCACCGCTCTGAACAT CGTGATGAGATCATCAAGTCGGGAGGCCTCAGCTCCGGCTTCTACTTCGAGATCGCCC GGATCGAGAACGAGATGAAGATCAACCGCCAGATCATGGACAACAGCGCCAAGTATGT CGAGCACGACCCTCGGCTGGTGGCAGAGCACAGATTCGAGACTATGAAGCCCAATTC TGGAGTAGGATCGGCACTGTGCGCCGAAGCGCTACCCAGGTGTGATGTACACCTTCAC CACCCCGCTGATCAGCGGGGGGGGGGTAGCGGCGGCGGCGGCAGTGCCGAGACAA AGGAGTTCAAGACCCTGTACAATTGTTTCATCGACTCCTACCTCCAGAAGTTGGCTCAG CACTCGATCCCCACCAATGTGACATGCGCCATTACATTGGTGAGGTGATCGGGCAGTT CAAGAACTGTGCTCTGCGGATTACCAACAAGTGATGAGTAATTCCCGCTGTCTTCAC GCTCATGGTGGAGTCTTCATCGAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCGG GCCATCGCCGAGGAGATCGGCATCGACCTGGACGATGTTCCGTCTGTCTCTCAAGCT GGAGAAGAACTGCAACGCATATGCTGAGGTGAACAACATCATCGATATTAGAAGCTG AATATCGGTGAGTGTCCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATCGTAATACC GGGAGTGCGGAGGCGAACTGCGGGCTCCAGACCATCGTGAAGAGCCTGAACAAGATC TACGTCCCGCTATCATCGAGAATCGGCTGCCGTACTATGACCCTGGTGCGGCGGCGAG CGGGGGGGGGGGTCTGGGGGCCGCTGCTTCCATCCAGACCACCGTGAACACTCTGTC TGAGCGTATCAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTG TGATATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCAACATCACCGTGA AGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCAC GGAGACATATTCCGGTCTGACCCTGAGCAGAAGGCGTATGTGCTGCCATGTTACAG CTGCGCTGAATATCCAGACGAGTGTTAACACGGTGGTGCGGGATTCGAGAATTATGTG AAGCAGACCTGTAACAGCTCGGCTGTTGTGGACAACAAGCTCAAGATCCAGAATGTGA TCATTGACGAGTGTTACGGGGCCCCCGGCTCCCCACGAATCTGGAGTTCATCAACACG GGCAGCAGCAAGGGCAACTGCGCCATTAAGGCGCTGATGCAGCTGACCACTAAGGCC ACAACCCAGATCGCCCTAGGCAGGTGGCCGGGACTGGCGTGCAAGTCTATATGATAGT GATTGGCGTGATCATCTGGCCGCCCTGTTTCATGTACTACGCCAAGCGCATGCTCTTAC TTCGACCAACGATAAGATCAAATTATCTTGGCGAATAAGGAGAATGTGCATTGGACGA CGTACATGGACACCTTCTTTCGCACATCACCTATGATCATTGCCACGACTGACATTCAGA ACTGA
45	0201 mRNA	AUCCAAGCUUAGCUUGCUUUGUUCUUUUUGCAGAAGCUAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGGACUGGACCUUGGAUUCUGUUCUGUGGCGGCCGCC ACUAGGGUGCACAGCGACGGGACCCUGUJCCCAGGGGACGACGACCUGGCCAUCC CCGUACCGAGUUCUUCAGCACCAAGGUGCCAAGAACCCCGAGACUAAGCGGGA GGCCAUCUGUAAGGCCUACGGGGACGAUAAUGAGGAGACGUGAAGCAGCGUCU CACCAAUCUGGAGAAGAAGAUUAUAUAUACCAACCAAGUUCGAGCAGAUUCGAG AAGUGCUGUAAGAGGAUAGCAGGUGCUGUJUCGGCUGGAGAACCACGCCGAG ACACUGCGGGCUGCCAUGAUCUCUCUUGCCAAGAAGAUCAUGUGCAGACCGGC

		AGGCACCCUUAUUACCCUUAACGACGUGCCGAUUUAUGCCGAGUGAUUAAGAAUUC UAGGCUUGGAGCCUCGGUGGCCAUGCUUCUUGCCCUUGGGCCUCCCCCAGCCCC UCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGG CGGCACUAGUAAA AA AAAAAAAAAAAAAAAAAAAAA
46	0302 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGAUGACUCCGAGAAUGACGAGGAGCAGACCAGCGUG UUCUCCGCUACCGUGUACGGGGACAAGAUCAGGGGAAGAAUAGAGGAAGCGG GUGAUCGGCCUGUGCAUCAGGAUCAGCAUGGUGAUCAGCCUGCUGAGCAUGAUC ACCAUGUCUGCAUUCUGAUCGUGCGGCUGAAUCAGUGCAUGAGCGCCAACAAG GCCGCCAUCACUGAUUCAGCCGUGGCCGUGGCCGCCGCUUCCUCUACUCACAGGA AGGUGGUGUCCUCCACCACUCAGUACGAUCACAAGGAGAGCUGCAAUGGCCUGU AUUACCAGGGCUCUUGCUACAUCUGCACAGCGAUUACAAGAGCUUUGAGGAUG CCAAGGCCAAUUGCGCCGCCGAGAGCAGCACUCUGCCUACAAGAGCGACGUGCU GACUACUUGGCUGAUCGACUAUGUGGAGGACACCUGGGGGAGCGAUGGGAACCC CAUCACCAAGACUACCAGCGACUAUCAGGACAGCGACGUGAGCCAGGAGGUGCGG AAGUACUUCUGCACCUACCCUUAACGACGUGCCUGACUACGCCUGAUAAAGAAUUCU AGGCUUGGAGCCUCGGUGGCCAUGCUUCUUGCCCUUGGGCCUCCCCCAGCCCCU CCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGC GGCACUAGUAAA AA AAAAAAAAAAAAAAAAAAAAA
47	0402 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGACCCAGCUGCCUAUCCUGCUGCUGCUGAUCAGCCUGG UGUACGCUACCCCUAGUCCCCAGACUCCAAGAAGAUCGGGGACGACGCCACUAU CAGCUGCAGCAGGAUAAACACCAACUACUACGUGGUGAUGAGCGCCUGGUUAUAA GGAGCCUAAACAGCAUCAUCCUGCUGGCCGCCAAGAGCGAUGUGCUGUACUUCGA CAACUAUACCAAGGACAAGAUCUCUUAACGAUAGCCCCUACGAUGACCUGGUGACU ACCAUCACCAUCAAGAGUCUGACCGCCGGGAUGCCGGGACCUAUUAUCUGUGCG UUUUUUUUGACCUCUACAACAAACGACACAGAUAAAGUGGACUACGAGGAGUAU AGCAUCGAGCUGAUCGUGAACACCGACAGCGAGAGUACCAUCGACAUCAUCCUGA GCGGGAGUACUCCCGAGACUAUCAGCGAGAAGCCAGAGGACAUCGACAAUAGCAA CUGCAGCAGCGUGUUCGAGAUACUACUCCUGAGCCUAUUACAGAUAAACGUGGA UGAUCACACCGACACCGUGACCUACACCAGCGACAGCAUCAACACCGUGAAUGCU AGCAGCGGUGAGAGUACCACGGAUGAGAUCCCGAGCCCAUCACGGAUAAGGAG GAGGAUCACACCGUGACCGACACGGUGAGUUACACCACCGUGUCUACCUCUCAG GGAUUGUGACCACAAAGUCUACCACCGAUGAUGCCGAUCUGUAUGACACAUACA AUGACAACGAUACCGUGCCGCCACCACCGUGGGCGGCAGCACAAUCCAUIAG UAACUACAAGACUAAGGAUUUCGUGGAGAUUUUGGGAUCACAACCCUGAUUAU CCUGAGCGCCGUGGCGAUUCUUGUAUCACCUACUUAUCUGUAUAAGCAUCC CAGGAAGUAUAAGACCGAGAACAAGGUGUAUCCGUACGACGUGCCGGACUACGC CUGAUAAAGAAUUCUAGGCUUGGAGCCUCGGUGGCCAUGCUUCUUGCCCUUGGGC CUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAA

		AGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
48	0505 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGAAGACCAUCUCCGUGGUGACCCUGCUGUGCGUGCUG CCAGCAGUGGUGUACAGUGGCGGCGGGGGAGCUACCCUACGACGUGCCCGAC UACGCCACCUGUACGGUGCCAACCAUGAACACGCUAAGCUGACCUCCACAGAGA CGAGUUUUAAUGACAAGCAGAAGGUGACCUUACCCUGCGACAGUGGCUACCACA GCUUGGACCCCAACGCUGUGUGCGAGACAGACAAGUGGAAGUACGAGAACCCGU GCAAGAAGAUUGCACCUGAGUGACUACGUCAGCGAGCUGUACGACAAGCCGC UGUACGAGGUCAACUCCACGAUGACACUCUCUUGCAACGGGGAGACAAAGUACU UCAGGUGCGAGGAGAAGAACGGAAUACCAGCUGGAAUGAUACCGUGACAUGCC CCAAUGCCGAGUGCCAGCCCCUGCAGCUCGAGCAUGGGAGCUGUCAGCCCGUGAA GGAGAAGUACUCCUUCGGGGAGUACAUGACCAUCAAUUGCGACGUGGGGUACGA GGUGAUCGGGUGAGCUACAUCUCUUGCACCGCCAAUAGCUGGAACGUGAUUCC CAGCUGUCAGCAGAAGUGCGACAUCUCCUUCUCUGUCCACGGGCUGAUCAGCGG GAGCACGUUCAGCAUUGGCGGGGUGAUCCACCUAGCUGCAAGAGUGGGUUCAC CCUGACCGGGUCCCCGAGCAGCACCUGCAUCGAUGGGAAGUGGAACCCGAUCCUG CCAACUUGCGUCCGGUCCAACGAGGAGUUCGACCCCGUGGACGACGGACCAGACG ACGAAACCGAUCUGAGCAAGUUGAGCAAGGACGUGGUGCAGUACGAGCAGGAGA UCGAGAGCCUGGAGGCCACCUACCACAUCAUCAUUGGCACUGACCAUCAUGGG CGUGAUUUCUUGAUCAGCAUCAUUGUGCUGGUGUGUUCUUGCGACAAGAACA ACGACCAGUACAAGUUUACAAGCUGCUGCCCUGAUAGAAUUCUAGGCUGGAG CCUCGGUGGGCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUUC CUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUA AAA AAA AAAAAAA
49	0603 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGGACUGGACCUUGGAUUCUGUUCUGGUGGCGGCCGU ACUCGGGUGCAUAGCGGCGGGGGGUCCUACCCUACGACGUGCCCGACUUAU GCCGCCGAGACAAAGGAGUUUAAGACCCUGUACAACCUGUUUAUCGACAGCUAU CUGCAGAAGCUGGCUCAGCACUCUAUCCACCAAUUGACAUUGCGCCAUUCACA UUGGUGAGGUGAUCGGCCAGUUCAAGAACUGCGCCUGAGGAUCACGAUAAGU GCAUGAGCAACAGCAGGCUGAGCUUCACCCUGAUGGUGGAGUCCUUAUCGAGG UGAUCAGCCUGCUGAGCGAGAAGGACAGGAGGGCCAUCGCCGAGGAGUUCGGCA UCGACCUGGACGAUGUGCCAGCGUGGUGAGCAAGCUGGAGAAGAACUGCAACG CAUAUGCUGAGGUGAACACAUCAUCGAUAUUCAGAAGCUGAAUAUCGGUGAGU GUUCCGCUCCCCUGGGCAGCAUAUGCUGUUGCAGAUUCUGAACACGGGCAGUG CCGAGGCCAACUGCGGGCUGCAGACCAUCGUGAAGUCACUGAAUAAGAUCUACG UGCCACCAUCAUCGAGAACC GGCGUCCCUACUACGACCCGUGGUUCCUGGUGGG GGUGGCCAUUAUUCUGGUGAUCUUCACGGUGGCCAUUCGAGCAUCCGAGGAA CCUGGCCUGAAGUACAGGUACGGGACCUUCCUGUACGUCUGAUAGAAUUCUA GGCUGGAGCCUCGGUGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUC

		CUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCG GCACUAGUAAA AAA AAAAAAAAAAAAAAAAAAAA
50	0702 mRNA	AUCCAAGCUUAGCUUGCUUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUGUUCUGGUGGCGCCGCC ACCAGGGUGCACAGUGGCGGAGGAGGCAGCUACCCUACGACGUCCCCGACUACG CUGGGGCGUGUGCCUCCAUCCAGACCACUGUGAACACCCUGAGUGAGCGGAUCAG CAGCAAGCUGGAGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAGA GAUUGGCAAUUUCUAUAUCAGGCAGAACACGGGUGCAACUACCCGUGAAGAA UAUGUGCUCCGUGACGCCGACGCUCAGCUGGACGCCGUGCUGAGCGCGGCCACG GAGACAUUUCGGUCUGACCCUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUC ACAGCUGCACUGAACAUCCAGACCAGCGUGAACACCGUGGUGAGGGACUUCGAG AACUACGUGAAGCAGACAUGCAACAGCAGCGCGGUGGUGGAUAAUAGCUCUAG AUCCAGAAUGUGAUCAUUGACGAGUGUUAUGGGGCCCCUGGGAGCCCCACUAA CUCGAGUUUAUUAAUACGGGCAGCAGCAAGGGCAACUGCGCCAUAAGGCCUG AUGCAGCUGACAACCAAGGCCACCACCCAGAUUGCCCCAGGCAGGUGGCCGCUU GAUAAGAAUUCUAGGCUGGAGCCUCGUGGCCAUGCUUCUUGCCCCUUGGGCCU CCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAG UCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
51	0803 mRNA	AUCCAAGCUUAGCUUGCUUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUGUUCUGGUGGCGCCGCC ACACGGGUGCAUUCGGCGGCGGCGGCAGCUACCAUACGAUGUGCCUGACUAC GCCGCCGCGUGAAGACACCCGUGAUCGUGGUGCCCGUGAUUGACCGGCCCCUA GCGAGACAUUCCCCAAUGUGCAGCAGCACAUAACGACCAGAAGUUCGACGACGU GAAGGACAACGAGGUGAUGCAGGAGAAGAGGGACGUGGUGAUUGUGAACGACG AUCCUGACCACUACAAGGACUACGUGUUAUCCAGUGGACAGGCGGGAACAUA GGGACGACGACAAGUAUACCCACUUCUUCUCCGGCUUCUGUAAACCAUGUGUA CAGAGGAGACAAAGCGGAACAUUGCUCGGCACCUGGCCUGUGGGACAGCAAGU UCUUCACCGAGCUGGAGAACAAGAACGUGGAGUAUGUGGUGAUCAUCGAGAACG ACAACGUAUCGAGGACAUCACCUUUCUGCGCCUGUGCUGAAGGCCAUCCACGA UAAGAAGAUUCGACAUCUGCAGAUCCGGGAGAUCAUACCGGGAAUAGGUGAA GACCGAGCUGGUGAUCGACAAGGAUACGCCAUUCUACCUACACCGGCGGCUAU GAUGUGAGCCUGAGCGCCUACAUAUCCGGGUGACCACCGCCUGAAUAUCGUG GAUGAGAUCAUAAGUCCGGCGGGCUGAGCAGCGGCUUCUACUUCGAGAUCCG AGAAUCGAGAACGAGAUAGAAGUAACAGGCAGAUCAUGGACAACAGCGCCAAG UACGUGGAGCACGACCCUCGGCUGGUGGCCGAGCACCGGUUCGAGACUUGAAG CCCAACUUCUGGUCCAGAAUCGGGACUGUGGCCGCAAGCGGUACCCGCGGUGA UGUACACAUUACCAACCCACUGAUCAGCUUCUUUGGCCUGUUCGACAUAACGU GAUCGGCCUGAUCGUGAUCUGUUAUUAUGUUAUGCUGAUUCUUAACGUGAA GUCCAAGCUGCUGUGGUUCCUGACCGGCACCUUCGUGACUGCCUUAUCUGAUA AGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCC

		CCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUGAAUAAAGUCU GAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
52	0201-J0 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCC ACGCGAGUCCACUCCAUGGACGGAACUCUUUCCCGGAGAUACGAUCUUGCA AUUCCAGCAACUGAAUUUUUCUCUACAAAGGCUGCUAAAAUCCAGAGACUAAA CGCGAAGCAAUUGUUAAGCCUAUGGAGACGAUAAUGAGGAAACUCUCAAACAA CGGCUAACUAAUUUGGAAAAAAGAUUACUAAUAUAACAACAAAGUUUGAACAA AUAGAAAAGUGUUGUAAACGCAACGAUGAAGUUCUAAUUAGGUUGGAAAAUCA CGCUGAAACUCUAAGAGCGGCUAUGAUUUCUCUGGCUAAAAAGAUUGAUGUACA GACUGGACGGCAUCCAUAUGAGUAAGAAUUCUAGGCUGGAGCCUCGGUGGCCAU GCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACC CCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAA AA AA
53	0301-J0 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCC ACGCGAGUCCACUCCAUGAUGACACCAGAAAACGACGAAGAGCAGACAUUGUGU UCUCCGCUACUGUUACGGAGACAAAUUCAGGGAAAGAAUAAACGCAACGCG UGAUUGGUCUAUGUAUUAGAAUAUCUUGGUUAUUUCACUACUACUUAUGAUU ACCAUGUCCGCGUUUCUUAUGUGCGCCUAAAUCAAUGCAUGUCGCUAACAAG GCUGCUAUUACUGACUCCGCGUUGCCGUUGCUGCGGCAUCAUCUACUUAAGA AAGGUUGUGUCUAGCACUACACAAUAUGAUCACAAAGAAAGCUGUAAUGGUUUA UAUUACCAGGGUUCUUGUUAUAUAUUACAUCAGACUAUAAGUCAUUCGAGGA UGC UAAAGCAAACUGCGCUGCGGAUUAUACAACUACCCAUAUAAUCCGAUGUC UUGACUACCUGGCUCUUAUGUUAUGUUGAGGAUACAUGGGGAUCUGAUGGUAA UCCAAUUAACAAAACUACAUCGGAUUAUCAAGAUUCUGAUGUAUCACAAGAAGU UAGAAAGUAUUUUUGUACAUAAGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGC UUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCC CGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAA AA AA
54	0501-J0 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUCUUCUUGGUGGCAGCAGCC ACGCGAGUCCACUCCAUGAAAACGAUUUCCGUUGUACGUUGUUAUGCGUACUA CCUGCUGUUGUUAUUAACAUGUACUGUACCCACUAUGAAUAAACGCUAAAUUA ACGUCUACCGAAACUCCGUUUAUGAUAAACAGAAAGUACGUUUACAUGUGAU UCAGGAUAUCAUUCUUGGAUCCAAUUGCUGUCUGCGAAACAGAUAAUUGGAAA UACGAAAAUCCAUGCAAGAAAAUGUGCACAGUUUCUGAUUAUGUCUCUGAACUA UAUGAUAAAGCCAUUAUACGAAGUGAAUUCACCAUGACACUAAGUUGCAACGGU GAAACAAAUAUUUUCGUUGUGAAGAAAAAUUGGAAUACUUCUUGGAUUGA UACUGUCACGUGUCCUAAUGCGGAUUGUCAACCUCUCAAUUAAGAACACGGAUC

		GUGUCAACCAGUUAAGAAAAAUACUCAUUUGGGGAAUAUAUGACUAUCAACUG UGAUGUUGGAUAUGAGGUUAUUGGUGUUUCGUUAUAAGUUGUACGGCUAAU UCUUGGAAUGUUAUCCAUAUGUCAACAAAAUGUGAUUAACGGUCCCUAUCU AAUGGAUUAUUUCCGGAUCUACAUUUUCUAUCGGUGGCGUUAUACAUCUUAAG UUGUAAAAGUGGUUUUACACUAACGGGUCCCCAUCAUCCACAUGUAUCGACGG UAAAUGGAAUCCCAUACUCCCAACAUGUGUACGAUCUAACGAAGAAUUUGAUCC AGUGGAUGAUGGUCCCGACGAUGAGACAGAUCUGAGCAAACUCUCGAAAGACGU UGUACAAUAUGAACAAGAAAUAGAAUCGUUAGAAGCAACUUAUCAUAUAUCAU AAUGGCGUUGACAAUUAUGGGUGUCAUAUUUCUAAUCUCCAUUAUAGUAUUAG UUUGUUCUGUGACAAAAUAAUGACCAUAUAAGUCCAUAUUUGCUACCGU GAGAAUUCUAGGCUGGAGCCUCGGUGGGCAUGCUUCUUGCCCCUUGGGCCUCC CCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
55	0701-J0 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACUGGAUUCUUCUUGGUGGCAGCAGCC ACGCGAGUCCACUCCAUGGGUGCCGAGCAAGCAUACAGACGACUGUGAAUACAC UCAGUGAACGUAUCUCGUCUAAAUAAGAACAAGAAGCGAACGCUAGUGCUCAAA CAAAUGUGAUUAAGAAUUCGGAUUUUUUAUAUCCGACAAAACCAUGGAUGUA ACAUACUGUUAUUAAUAUGUGCUCUGCGGACGCGGAUGCUCAGUUGGAUGCUG UGUUAUCAGCCGCUACAGAAACAUUAAGUGGAUUAACACCGGAACAAAAGCAU ACGUACCAGCUAUGUUUACUGCUGCGUUAACAUCAGACGAGUGUAAACACUG UUGUUAGAGAUUUUGAAAAUUAUGUGAAACAGACUUGUAAUUCUAGCGCUGUU GUCGAUACAAAUUAAAGAUACAAAACGUAAUUAUAGAUAGUUAACGGAGCC CCAGGAUCUCCAACAAAUUUGGAUUUAUUAUACAGGAUCUAGCAAAGGAAAU UGUGCCAUUAAGGCGUUGAUGCAAUUGACUACUAAGGCCACUACUCAAUAGCA CCUAGACAAGUUGCUGGUACAGGAGUUCAGUUUUUAUAGAUUGUUAUCGGUGU UAUAAUAUUGGCAGCGUUGUUUAUGUACUAUGCCAAGCGUAUGCUGUUCACAU CCACCAUAGUAAAAUCAACUUAUUUUAGCCAAUAGGAAAACGUCCAUUGGAC UACUUAUAGGACACAUUCUUAAGAUCUUCUCCGAUGAUUAUUGCUACCACGGA UAUACAAAACUGAGAAUUCUAGGCUGGAGCCUCGGUGGCAUGCUUCUUGCCCC UUGGGCCUCCCCCAGCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUU GAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
56	0901-J0 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGCGCCGCCACCAUGGACUGGACUGGAUUCUUCUUGGUGGCAGCAGCC ACGCGAGUCCACUCCAUGCCUACAACUAUCUCCUAUUAUAUAGAAACUAAAA AAGCAAUUCUGACGCGCAUUGAAGACGUUAGACAUACAUAUAUAGAGUCGA AACCAACCACUAUCCAGAAUACUGGAAAACUAGUAAGGAUUAUUUUAAAGGAG GAUAUAUAGUGGAGGGUUUCUCCCAAUGAAUUGUGUUAUCAACAUACAUA UAUAUUGGGGAAAGGAAGACGAUUAUGGAUCCAUCACUUGAUAGAUGUAUAC AAUACUCUGGAGAGAUUAUCUGGUUCAUUGGAUAAGAAAAAUUAAGUUC

		UUAUGAAGAGGCAAAAAACACGAUGAUGGAAUUAUCAUUAUUGCUAUAUUCU UACAAGUAUCAGAUCAUAAAAUGUAUUAUUUCAAAGAUAGUCAAUCAAUUGG AUUCCAUAJAGAUCCGCUAAUAUGUCUGCGCCAUUUGAUUCAGUAUUUUUAUCUA GACAAUUUGCUACCUAGUACGUUGGAUUAUUUUACAUAUCUAGGAACAACUAUC AACCACUCUGCAGACGCAGCAUGGAUAAUUUUUCCAACACCAAUAAACAUUCAUU CUGAUCAACUAUCUAAAUUCAGAACACUAUUGUCGUCGUCUAAUUCAGAGGGUA AACCUCAUUAUAUACAGAGAAUUAUAGAAAUCCGUAAAAUUGAACGACGACA CGCAAGUAUAUUAUUCUGGGGAGAUUAUACGAGCUGCAACCACCUCACCAGUUC GAGAGAACUAUUUUUAUGAAUUGGUUGUCUGAUUUGAGAGAGGCAUGUUUUUCA UAUUAUCAAAAAUAUAUCGAAGGGAAUAAAACAUUCGCAAUUAUUGCCAUAUGA UUCGUUUUAUACUCACCGCUAUUCUCUUUCUUAUGAGUCAACGAUAUUCGCGA GAAAAACAAACUAGGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCC CCUUGGGCCUCCCCCAGCCCCUCCUCCCUUCCUGCACCCGUACCCCCGUGGUCU UUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
57	1101 mRNA	AUCCAAGCUUAGCUUGCUUUGUUCUUUUUGCAGAAGCUAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGGACUGGACGUGGAUUCUCUUCUGGUGGCAGCCGCC ACUCGUGUGCACUCCGAUGGGACCUUGUUCUCCUGGGGACGACGACCUUGGCCAUC CCCGCUACGGAGUUCUUCAGCACCAAGGCUGCCAAGAACCCGGAGACUAAGCGGG AGGCCAUCGUCAAGGCCUACGGAGAUACAACGAGGAGACGCUAAGCAGCGGC UGACCAACCUGGAGAAGAAGAUACGAACAUCACGACCAAGUUCGAGCAGAUUGA GAAGUGCUGCAAGCGGAACGACGAGGUGUUGUUCGCUUGGAGAACCACGCAGA GACUCUGCGGGCAGCCAUGAUCAGCUUGGCGAAGAAGAUAGCAGUGCAGACUGG ACGCCACCCCUAUGAGGGGAGGAGGCGGCAGCGCGGCGGCGGCGAGCCGGCUGAA UCAGUGCAUGAGCGCCAACAAGGCCGCCAUCACUGAUUCAGCCGUCGCCGUCGCU GCCGCCUCCUCCACUCAUAGGAAGGUGGUGUCCAGUACGACACAGUACGACCAUA AGGAGAGCUGUAAUGGGCUGUACUACCAGGGGUCGUGCUACAUCUGCAGACGCG ACUACAAGUCGUUUGAGGAUGCGAAGGCGAACUGCGCAGCCGAGAGCAGCACGU UGCCUAACAAGAGCGACGUGCUGACUACUUGGCUGAUUGACUAUGUCGAGGAUA CCUGGGGCUCCGAUGGGAACCCCAUCACGAAGACCACGUCUGACUACCAGGACUC CGAUGUGUCUCAGGAGGUGCGGAAGUACUUCUGUACUGGCGGCGGCGGUUCUG GUGGGGGCGGCAGCACCUGCACCGUCCCCACCAUGAACAAACGCCAAGCUGACCAG UACAGAGACUCCUUAACGAUAAGCAGAAGGUGACAUUUAUUGCGAUUCCGG CUACCACUCGUUGGACCCCAACGCUGUCUGUGAGACAGACAAGUGGAAGUACGA GAUCCAUGUAAGAAGAUUGCACGGUGUCCGACUAUGUUUCUGAGCUGUACG ACAAGCCCCUUACGAGGUGAACAGCACCAUGACUCUCAGCUGCAACGGCGAGAC UAAGUACUUCGGUGUGAGGAGAAGAACGGUAAUACGUCCUGGAACGAUACCG UUACUUGCCCAACGCCGAGUGCCAGCCGUUGCAGCUGGAGCAUGGUAGCUGUC AGCCCGUGAAGGAGAAGUAUAGUUUCGGCGAGUACAUGACCAUUAUUGUGAU GUGGGUUACGAGGUCAUCGGGUGAGUUACAUCUCUUGCACUGCCAUAAGCUG GAACGUGAUUCCAGCUGUCAGCAGAAGUGUGACAUACCUUCUCUGUCCAACGG GCUGAUCAGCGGGAGCACGUUCAGCAUUGGCGGGGUGAUCCACCUCAGCUGCAA GAGCGGGUUCACGUUGACCGGGUCGCCUCGUCGACCUGCAUCGACGGGAAGUG

		GAAUCCCAUUCUGCCCACGUGCGUGCGGUCGAACGAGGAGUUCGACCCGGUUCGA CGAUGGACCCGAUGACGAGACUGAUCUGUCUAAGUUGUCGAAGGACGUCGUUCA GUACGAGCAGGAGAUUCGAGUCCUGGAGGCCACGUACCACGGCGGUGGCGGGAG CGGGGGAGGUGGCAGCAGUCCCCAGACUCCAAGAAGAUUCGGGGACGACGCCACC AUCUCCUGCUCCGCAACAACACUAACUACUACGUGGUUAUGUCCGCCUGGUACA AGGAGCCCAACUCGAUCAUCCUGCUCGUCGUAAGAGCGACGUCCUGUACUUCGA CAACUAUACCAAGGACAAGAUACGCUACGACUCACCCUACGAUGACCUCGUGACC ACCAUCACAUAUAAAGAGUCUGACCGCCGGGGACGCCGGCACAUAUAUUGCGCCU UCUUCAUGACGAGUACCACCAUUAUACCGACAAGGUGGAUUAUGAGGAGUACU CGAUCGAGCUUAUCGUGAACACGGACAGUGAGUCGACCAUCGACAUAUCCUGU CCGGUAGCACCCUGAGACAUAUCCGAGAAGCCGGAGGACAUCGACAUAAGCAA UUGCAGUUCGUCUUCGAGAUACGACCCUGAGCCCAUUAACGACAACGUGGAU GAUCACACCGACACGGUGACGUACACCAGCGACUCUAUAACACCGUCAUUGCUU CCAGUGGGGAGUCCACCACUGACGAGAUUCCGGAGCCCAUACCGACAAGGAGGA GGAUCACACCGUGACCGACACGGUGAGUUAACACCACCGUGUCUACCUCUCUGGU AUUGUGACUACCAAGAGCACACGGACGACGCCGAUCUGUACGACACGUACAUG ACAACGACACCGUGCCGCCACACCGUGGGCGGGUCCACGACGUCCAUAUAGUAA CUACAAGACUAAGGACUUCGUGGGGGGGCGGCGGUUCCGGUGGGCGGGGGCAGCGC UGAAACCAAGGAGUUAAGACGCGUGUACAACUUGUUAUCGACUCUUAACUGCA GAAGUUGGGCCAGCACUCCAUCCUACCAUUGUCACUUGUGCCAUCCACAUGGA GAGGUGAUCGGCCAGUUAAGAACUGCGCCUGAGGAUACCAACAAGUGCAUG AGCAAUAGCCGCCUGUCCUUCACGCUCAUGGUGGAGUCCUUAUCGAGGUGAUC UCCUGCUGAGCGAGAAGGACAGGCGGGCUAUUGCUGAGGAGAUCCGUUAUUGA CCUCGACGACGUGCCGUCGUCUCCAAGCUGGAGAAGAACUGCAACGCAUAU GCUAGGUGAACAACAUCAUCGAUAUUCAGAAGCUGAAUAUCGGUGAGUGUUC GCUCCCCUGGGCAGCAUAUGCUGUUGCAGAUUGUCAUAACGGGAGUGCUGAG GCCAACUGCGGGUUGCAGACCAUCGUGAAGAGUCUGAACAAGAUUUAACGUGCCC CCGAUCAUCGAGAACCGCCUGCCUACUACGACCCCGUGGGGGCGGGUCUGGAG GAGGCGGCAGCGGCGUGCCGCCUCCAUCAGACCACGGUGAAUACGUUGUCCGA GCGGAUUUCGAGCAAGCUGGAGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUG UGAUUAAGAGAUUGGCAAUUUCUAUAUCAGGCAGAACCACGGUGCAACAUCAC CGUGAAGAAUAUGUGCUCCGUGACGCCGACGCUAGCUGGACGCCGUGCUGAG CGCGGCCACGGAGACAUAUCCGGUCUGACCCUGAGCAGAAGGCGUAUGUGCC UGCCAUGUUCACAGCUGCUCUGAACAUCAGACGUCGUGAAUACCGUGGUGAG GGACUUCGAGAACUACGUGAAGCAGACGUGUAAUAGCAGCGCUGUCGUGCAUAA UAAGCUGAAGAUCCAGAACGUUAUAUUCGACGAGUGCUACGGCGCCCCCGGUCC CCGACCAACCGGAGUUAUAUAACCGGUUCCUCCAAGGGGAACUGCGCUAUU AAGGCGCUGAUGCAGCUGACUACGAAGGCGACGACCCAGAUUCGCGCCUGCCAGG UGGCGGGGACCGGGGUGCAGUUCUAUGAUCGUGAUCGGUGUGAUAUCCUG GCAGCUCUCUUAUGUACUACGCCAAGCGCAUGCUGUUCACCAGCACCAACGACA AGAUUAAGCUGAUCCUCGCCAAUAAGGAGAACGUCCACUGGACCACCUACAUGGA CACAUCUUCGCGACACGCCUUAUGAUUAUCGCCACCACUGACAUAUGAAGCUGA UAAGAAUUCUAGGCUUGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCUUCCUGACCCGUACCCCGUGGUCUUAUGAAUAAAGUC
--	--	--

		UGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
58	1201 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGACCCAGCUGCCUAUCCUGCUGCUGCUGAUCAGUCUG GUGUACGCUACACCGUCUCCCCAGACUCCAAGAAGAU CGGGGACGACGCGACGA UCAGCUGCAGCAGGAUAACACCAACUACUACGUGGUGAUGAGCGCCUGGUACA AGGAGCCGAAUAGCAUCAUUCUGCUGGCCGCCAAGUCUGACGUGCUGUACUUCG ACAAUACACAAAGGACAAGAUCUCCUACGACAGUCCUUAUGAUGAUCUGGUGA CCACGAUACCAUUAAGUCUCUGACUGCUGGGGAUGCCGGUACGUUAUUCUGCG CCUUCUUAUGACGUCCACGACCAUAGACACUGAUAAAGUGGAUUAACGAGGAGU ACUCCAUCGAGCUGAUCGUGAACACCGACAGCGAGUCUACCAUCGACAUCAUCCU GUCCGGGAGCACCCCGAGACUAUCAGUGAGAAGCCUGAGGAUUAUUGACAACAG CAAUUGCAGUAGUGUCUUCGAGAUACACCCCGAGCCUAUCACUGACAACGUG GACGAUCACACCGACACGGUGACGUACACAGUGAUAGCAUCAACACCGUGAAUG CUAGCUCCGGGGAGUCCACGACCGAUAGAUUCCCGGAGCCUAUCACCGAUAAAGGA GGAGGAUCACACCGUGACCGACACGGUGAGUUAACACACCGUGUCUACCUCCUC GGGAUCGUGACUACGAAGAGCACUACUGACGAUGCUGAUCUGUACGAUACCUAC AACGAUAACGAUACCGUGCCGCCACCACCGUGGGCGGCAGCACACCAGCAUCAG CAACUAUAAGACUAAGGACUUCGUGGGGGGCGGGGGGAGCGGCGGCGGCGGCA GCCGGCUGAAUACAGUGCAUGAGCGCCAACAAGGCCGAUACUGAUUACGCCGU GGCUGUCGCCGCCAGCUCCACCCACCGAAGGUGGUGUCCAGCACCCACGAG UACGAUCACAAGGAGUCCUGCAAUGGCCUGUAUUAACAGGGCUCUUGCUACAUC CUGCACAGCGAUUACAAGAGCUUUGAGGAUGCCAAGGCCAAUUGUGCAGCCGAG AGCAGCACGUGCCUAACAAGAGCGACGUGCUGACUACUUGGCUGAUCGACUUAU GUGGAGGACACCUGGGGGAGUGAUGGGAACCCCAUCACUAAGACCACGUCCGAC UAUCAGGACAGUGACGUGAGUCAGGAGGUGCGGAAGUACUUCUGACCGGCGGC GGGGGAGCGGCGGGGGGCGGCAGCGCUGAGACAAAGGAGUUAAGACCCUUUAC AACUUGUUAUCGAUAGUUAUCUGCAGAAGCUGGCCAGCACAGCAUCCCCACCA AUGUGACAUGCGCAUUCACAUUGGUGAGGUGAUUGGCCAGUUAAGAACUGCG CCUGCGGAUAACUAACAAGUGUAUGUCUAAUUCCGCCUGUCCUUCACGCUCA UGGUGGAGUCCUUAUCGAGGUGAUCUCCUGCUGAGCGAGAAGGACAGGCGG GCUAUCGCCGAGGAGAU CGGCAUCGACCUGGAUGACGUCCCCAGCGUGGUGUCC AAGCUGGAGAAGAACUGCAACGCAUUGCUGAGGUGAACAACAUCAUGAUUU CAGAAGCUGAAUUCGGUGAGUGUUCGCUCCCCUGGGCAGCAUUGCUGUUG CAGAU CGUCAAUACCGGGUCUGCCGAGGCCAACUGCGGGCUGCAGACCAUCGUG AAGUCACUGAAUAAGAUCUACGUGCCACCAUCAUCGAGAACC GGCGUCCCUACU ACGACCCGUGGUUCCUGGUGGGGGUGGCCAUUAUUCUGGUGAUCUUCACGGUG GCCAUCUGCAGCAUCCGAGGAACCUGGCCCUGAAGUACCGGUACGGUACCUUCC UCUACGUGUGAUAGAUAUUCUAGGCUGGAGCCUCGUGGCGCAUGCUUCUUGCCC CUUGGGCCUCCCCCAGCCCCUCCUCCCUUCCUGACCCGUACCCCGUGGUCUU UGAAUAAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

59	1301 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCGCC ACAAGGGUGCACUCCGAUGGGACCUUGUUCCUGGGGACGACGACCUGGCGCAUC CCCGCUACGGAGUUCUUCAGCACCAAGGCUGCCAAGAACCCGGAGACUAAGCGGG AGGCCAUCGUCAAGGCCUACGGAGAUGACAAUGAGGAGACGCUGAAGCAGCGUC UCACCAACCUUGAGAAGAAGAUACCAACAUCACUACGAAGUUUGAGCAGAUUGA GAAGUGCUGCAAGCGCAAUGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGA GACUCUGCGGGCUGCCAUGAUCAGCCUCGCCAAGAAGAUACGUGCAGACCCGGA AGGCACCCCUAUGAGGGGGGAGGAGGGAGCGGCGGCGGGAGCGGGGCCGCG GCCAGUAUCCAGACCACCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAAGCUGG AGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAUAGAGAUUGGCAAUU UCUAUUACAGGCAGAACCACGGGUGCAACAUCACCGUGAAGAAUUGUGCUCCG CUGACGCCGACGUCAGCUGGACGCCGUGCUGAGCGCGGCCACGGAGACAUUUC CGGUCUGACCCUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCGCU GAUUAUCCAGACGAGUGUUAACACGGUGGUGCGGGACUUUGAGAACUACGUGA AGCAGACGUGCAACAGCUCAGCGGUGGUGGAUAAUAAAGCUCAAGAUCCAGAAUG UGAUCAUUGACGAGUGUUAUGGGGCCCCUGGGAGCCCCACUAACCUCGAGUUUA UUAUACGGGCGAGCAGCAAGGGCAACUGCGCUAUAAGGCGCUGAUGCAGCUGA CAACGAAGGCCACCACCCAGAUUGCUCGCGGCAGGUGGCAGGAGGGGGAGGGG GGUCCGGCGGGGGAGGCUCGACUUGCACGGUGCCGACGAUGAAUAAACGCCAAGC UGACCUCACCGAGACGUCGUUACAACGACAAGCAGAAGGUGACGUUACCCUGCGA CAGCGGCUAUCACAGCUUGGACCCCCAACGCCGUCUGCGAGACGGACAAGUGGAAG UACGAGAACCCGUGCAAGAAGAUUGUCACGGUGUCCGACUAUGUGUCCGAGCUG UACGAUAAAGCCGUGUACGAGGUGAACAGCACCAUGACGCUUGCUGGCAACGGC GAGACAAAGUACUUCGGUGCGAGGAGAAGAACGGGAACACCAGCUGGAACGAU ACAGUGACCUGUCCAACGCAGAGUGCCAGCCUCUGCAGUUGGAGCACGGGUCCU GUCAGCCGUUAAGGAGAAGUACAGCUUUGGCGAGUACAUGACGAUUAACUGCG ACGUCGGUUAACGAGGUGAUCGGCGUGAGUUAUCUUAUGUACCGCCAACAGCU GGAUUGUCAUACCUAGCUGCCAGCAGAAGUGUGACAUUCCGUCUCUCUCCAACG GGCUGAUCAGCGGUCCACGUUCUCUAUCGGCGGCGUUAUUAUCUGUCCUGCA AGUCGGGCUUACCCUGACCGGUCCCCCUUCCACUUGCAUCGACGGCAAGUG GAACCCAUCCUGCCCACCGUGCGUGCGGAGCAACGAGGAGUUCGACCCGUGGAU GACGGGCCUGAUGAUGAGACUGAUCUCUCUAAGCUGUCUAAGGACGUGGUCCAG UACGAGCAGGAGAUCGAGUCGUGGAGGCCACCUACCACAUCAUCAUAGGCC UCACCAUCAUGGGGGUGAUCUUCUUGAUCAGCAUCAUCGUGCUGGUGUGUUCU UGCGACAAGAACAACGACCAGUACAAGUUACAAGCUGCUGCCUUGAUAAAGAA UUCUAGGCUUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAG CCCUCCUCCCUUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGU GGGCGGCACUAGUAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAA
60	1401 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGAAGACCAUCUCCGUGGUGACCCUGCUGUGCGUGCUG CCAGCAGUGGUGUACAGCACCGUACGGUGCCGACGAUGAAUAAACGCCAAGCUG

		ACCUCCACCGAGACGUCGUUCAACGACAAGCAGAAGGUGACGUUACCCUGCGACA GCGGCUAUCACAGCUUGGACCCCAACGCCGUCUGCGAGACGGACAAGUGGAAGU ACGAGAACCCGUGCAAGAAGAUUGGCACGGUGUCCGACUUAUGUUGCCGAGCUGU ACGAUAAGCCGCUGUACGAGGUGAACAGCACCAUGACGCGUGUCGUGCAACGGCG AGACAAAGUACUUCCGGUGCGAGGAGAAGAACGGGAACACCAGCUGGAACGAUA CAGUGACCUGUCCCAACGCAGAGUGCCAGCCUCUGCAGUUGGAGCACGGGUCCU GUCAGCCCGUUAAGGAGAAGUACAGCUUUGGCGAGUACAUGACGAUUAACUGCG ACGUCGGUACGAGGUGAUCGGCGUGAGUUAACUUCUACUGUACCGCCACAGCU GGAAUGUCAUACCUAGCUGCCAGCAGAAGUGUGACAUUCGGUCUCUCUCCAACG GGCUGAUCAGCGGGUCCAGUUCUCUAUCGGCGGCGUUAUUAUCUGUCUUGCA AGUCCGGCUUCACGUUGACCGGGUCGCCUCGUCGACCUUGCAUCGACGGGAAGU GGAAUCCAUUUCUGCCACGUGCGUGCGGUCGAACGAGGAGUUCGACCCGGUCCG ACGAUGGGCCGACGACGAGACAGAUCUGUCAAGUUGAGCAAGGAUGUGGUUC AGUAUGAGCAGGAGAUUGAGAGCCUGGAGGCCACGUACCACGGGGGAGGGGGGA GCGGGCGCGGGAGCGGGGCCGCGGCCAGUAUCCAGACCACCGUGAACACUCU GUCUGAGCGUAUCAGCAGCAAGCUGGAGCAGGAGGCCAACGCCUCUGCUCAGACC AAGUGUGAUUAAGAGAUUGGCAUUAUUAUACAGGCAGAACACGGGUGCAAC AUCACCGUGAAGAAUUAUGUCUCCGUGACGCCGACGCUAGCUGGACGCCGUGC UGAGCGCGGCCACGGAGACAUUAUCCGGUCUGACCCUGAGCAGAAGGCGUAUG UGCCUGCCAUGUUCACAGCUGCGCUGAAUAUCCAGACGAGUGUUAACACGGUGG UGCGGGACUUAUGAGAACUACGUGAAGCAGACGUGCAACUCCAGCGCCGUGGUGG AUAUAAGCUCAAGAUCCAGAAUGUGAUCAUUGACGAGUGUUAUGGGGCCCCUG GGAGCCCACUAACCUCGAGUUUAUUAUACCGGUUCCUCCAAGGGGAACUGCG CCAUAAGGCGCUGAUGCAGUUGACAACGAAGGCGACGACCCAGAUUCGCGCCUCG GCAGGUGGCCGCGCGGUGGCGGGUCAGGCGCGGGGGGAGCGCCGCCGUGA AGACGCCGGUGAUCGUGGUGCCCGUGAUCGACCGGCCUCCGAGCGAGACGUUCC CCAACGUGCAUGAGCAUAUACGAUCAGAAGUUCGAUGACGUUAAGGACAACG AGGUGAUGCAGGAGAAGCGGGACGUGGUGAUCGUGAUGACGACCCGGAUCACU ACAAGGACUACGUGUUUAUCCAGUGGACCGGAGGCAACAUCCGGGAUGACGACA AGUACACCCAUUUUUUCUCCGGCUUCUGUAACACCAUGUGCACAGAGGAGACGA AGCGGAACAUUCGUCGCCACCUGGCACUGUGGGACUCCAAGUUCUUCACAGAGCU GGAGAACAAAGAAUGUGGAGUACGUCGUCAUUCGAGAAUGAUAAUGUGAUUG AGGACAUUACAUUUCUAGGCCGGUGCUCAAGGCCAUCCAUGACAAGAAGAUUCG ACAUCCUGCAGAUUCGGGAGAUCAUACCGGUAAUAAGGUGAAGACUGAGCUGG UCAUUGACAAGGACCACGCCAUUCACCUAUACCGGUGGCUACGACGUCUCGCU GAGCGGUAUAUACUACGCGUGACCACCGCUCUUAUAUUCGUCGACGAGAUCAU UAAGAGCGGUGGUCUGUCCUGGGCUUCUACUUCGAGAUAGCCCGGAUCGAGAA CGAGAUGAAGAUCAACCGCCAGAUCAUGGACAACUCAGCGAAGUACGUCGAGCAU GAUCCCCGCCUGGUGGCCGAGCACCGGUUUGAGACAAUGAAGCCCAACUUCUGG UCCCGGAUAGGCACCGUGGCCGCAAGCGGUUACCGGGCGUACUGUACACGUUC ACCACCCCGCUCAUCUACUUCUUCGGCCUGUUCGACAUUAACGUCAUCGGCCUGA UCGUGAUACUGUUAUUAUGUUAUGCUGAUCUUAACGUGAAGAGCAAGUUG CUGUGGUUCCUGACAGGAACCUUCGUGACUGCCUUAUUUGAUAAAGAAUUCUAG GCUUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUC
--	--	--

		UCCCCUUCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG CACUAGUAAA AAA AAAAAAAAAAAAAAAAAAAA
61	1501 mRNA	AUCCAAGCUUAGCUUGCUUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGAAGACCAUCUCGGUGGUGACACUGCUGUGCGUGUUG CCGGCGGUUGUUUAUAGUACGUGCACCGUGCCGACUAUGAACAACGCCAAGUUG ACGUCUACUGAGACUUCUUUCAUGACAAGCAGAAGGUGACGUUACCUGCGAC AGCGGCUAUCACAGCUUGGACCCCAACGCCGUCUGCGAGACGGACAAGUGGAAG UACGAGAACCCGUGCAAGAAGAUUGGCACGGUGUCCGACUAUGUGUCCGAGCUG UACGAUAAGCCGUGUACGAGGUGAACAGCACCAUGACGCUGUCCUGCAACGGC GAGACAAAGUACUUCGGUUGCGAGGAGAAGAACGGCAACACCAGCUGGAAUGAC ACCGUGACCUGUCCGAACGCCGAGUGCCAGCCCCUGCAGUUGGAGCAUGGCUCCU GCCAGCCCGUGAAGGAGAAGUACUCCUUCGGGGAGUACAUGACUAUACAACUGCG ACGUGGGCUACGAGGUGAUCGGGGUGUCUUAUAUCAGCUGCACCGCCAAUAGCU GGAACGUGAUUCCCAGCUGUCAGCAGAAGUGUGACAUACCUUCUCUGUCCAACG GGCUGAUCAGCGGGAGCACGUUCAGCAUUGGCGGGGUGAUCCACCUGAGUUGCA AGUCUGGCUUCACGCUGACCGGAUCUCCUAGCUCGACCUGCAUCGACGGGAAGU GGAACCCCAUCCUGCCAACUUGCGUCCGGUCCAACGAGGAGUUCGACCCCGUGGA CGACGGACCAGACGACGAAACCGAUCUGAGCAAGUUGAGCAAGGAUGUCGUGCA GUACGAGCAGGAGAUUGAGUCCUGGAGGCGACCUAUCACGGUGGGGGAGGGAG CGGCGGCGGGCUCACGGCUGAAUACAGUGCAUGAGCGCCAACAAGGCCGCCAUC ACUGAUUCAGCCGUGGCCGUCGCCGCCCUUCCUCCACCCACCGGAAGGUCGUCU CCAGCACCACCCAGUAUGAUACAAGGAGUCCUGCAACGGGCUGUACUAUCAGGG CAGUUGCUACAUCUGCACAGCGACUACAAGUCGUUUGAGGAUGCUAAGGCCAA CUGCGCCGCCGAAAGCUCCACGCUGCCGAUAAGUCGGAUGUCCUGACCACGUGG CUCAUCGACUACGUGGAGGACACCUGGGGCUCGGAUGGGAACCCCAUCACGAAG ACCACCAGCGACUAUCAGGACAGCGAUGUGAGCCAGGAGGUGCGUAAGUACUUC UGCACCGGGGGCGGAGGGUCUGGAGGAGGCGGCAGCGGCGUGCCGCCUCCAUC CAGACCACUGUGAACACCCUGUCCGAGCGCAUUUCGUCCAAGCUGGAGCAGGAGG CCAACGCCUCCGCCCAGACGAAGUGCGACAUUGAGAUAGGGAUUUUCUAUCCG CCAGAACACGGGUGCAACAUACCGUGAAGAAUUGUGCUCCGCUGACGCCGAC GCUACGUGGACGCCGUGCUGAGCGCGGCCACGGAGACAUAUCCGGUCUGACCC CGGAGCAGAAGGCUUACGUGCCUGCCAUGUUUACGGCGGCCUGAAUUAACAGA CUUCUGUGAAUACGGUGGUGCGGGACUUCGAGAAUACGUGAAGCAGACUUGC AACUCCAGCGCCGUGGUGGAUAAUAAGCUCAAGAUAUCCAGAAUGUGAUCAUUGAC GAGUGUUAUGGGGCCCCUGGGAGCCCCACUAACCUCGAGUUUAUUAUACCGGU UCCUCCAAGGGGAACUGCGCAUCAAGGCGCUGAUGCAGCUGACUACUAAGGCCA CCACCCAGAUCGCCCACGUCAGGUGGCUGGCACCGGCGUUCAGUUCUACAUGAU CGUGAUCGGGGUGAUCAUUCUGGCUGCGCUGUUCUAGUACUACGCCAAGCGCAU GCUCUUUACUUCGACCAACGAUAAGAUCAAACUUAUCUUGGCGAAUAAGGAGAA UGUGCAUUGGACGACGUACAUGGACACGUUCUCCGGACGUCUCCUAUGAUCAU UGCGACCACCGAUUAUCAGAACUGAUAAAGAAUUCUAGGCUGGAGCCUCGGUGGC CAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCCUCCCUUCCUGCACCCGU

		ACCCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAA AA AA
62	1601 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC UUUGGGCGCCGCCACCAUGGACUGGACAUGGAUUCUGUUCUGGUGGCCGCCGCC ACCAGGGUGCACAGCGACGGCACCCUGUUCCCAGGCGACGACGACCUGGCCAUCC CCGCUACCGAGUUCUUCAGCACCAAGGCUGCCAAGAACCCCGAGACUAAGCGGGA GGCCAUCGUGAAGGCCUACGGGGAUGAUAAACGAGGAGACUCUGAAGCAGCGGCU GACCAACUGGAGAAGAAGAUUACCAACAUUACCACCAAGUUCGAGCAGAUUCGAG AAGUGCUGCAAGCGGAACGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGAG ACACUGCGCGCCCAUGAUCAGCCUGGCCAAGAAGAUUCGACGUGCAGACCGGCC GCCACCCUACGAGGGGGGGCGGCGGCUCUGGGGGCGGCGGGAGCAGCCCCAGAC UUCCAAGAAGAUCGGGGAUGAUGCCACCAUCAGCUGCUCGGAACAACACCAAC UACUACGUGGUGAUGUCCGCCUGGUACAAGGAGCCUAAACAGCAUCAUCCUGCUG GCCGCCAAGAGCGAUGUGCUGUACUUUGACAAUACACCAAGGAUAAGAUCAGC UACGACAGCCCCUACGACGAUCUGGUGACUACCAUCACCAUCAAGACCCUGACAG CCGGCGACGCAGGCACCUACAUCUGCGCCUUCUUAUGACCAGCACCAUAUGA CACCGACAAGGUGGAUUACGAGGAGUACUCCAUCGAGCUGAUCGUGAAUACUGA CUCAGAGUCUACCAUCGACAUAUUCUGUCCGGGAGCACCCCGAGACUAUCAGU GAGAAGCCUGAGGACAUCGACAACAGCAACUGCAGCAGUGUGUUCGAGAUACCA CCCCAGAGCCUAUCACUGACAACGUGGAUGAUCACACCGACACUGUGACAUACAC CAGUGAUAGCAUCAACACCGUGAAUGCUAGCUCUGGGGAGAGCACCACUGAUGA GAUCCCCGAGCCUAUCACCGACAAGGAGGAGGAUCACACCGUGACCGACACUGUG AGUUACACCACAGUGUCUACCUCUCUGGGAUUCGUGACCACCAAGAGCACCACCG ACGAUGCCGACCUGUACGAUACAUAACGAUAACGAUACAGUGCCUCCACCAC CGUGGGAGGCAGCACCACCAGCAUUUCUAAUACAAGACCAAGGACUUCGUGGG GGGCGGGGGAGCGGGGGCGGGGGAGCGCCGAAACCAAGGAGUUAAGACCCU GUACAACCUGUUAUCGAUAGCUACCUGCAGAAGCUGGCCAGCACAGCAUCCC ACCAAUGUGACAUGCGCAUUCACAUAUGGGGAGGUGAUCGGCCAGUUAAGAAC UGCGCCUGAGGAUACCAUAAGUGCAUGAGCAAUAGCCGCCUGUCCUUCACUC UGAUGGUGGAGUCCUUAUCGAGGUGAUCUCCUGCUGUCAGAGAAGGACAGGC GGGCUAUUGCCGAGGAGAUCGGGAUCGACCUGGACGACGUGCCCAGCGUGGUGU CCAAGCUGGAGAAGAACUGCAACGCCUACGCCGAGGUGAACACAUAUCGAUAU UCAGAAGCUGAAUAUCGGUGAGUGUUCGCCCCUCCGGCCAGCACAUUCGUCU GCAGAUUCGUGAAUACAGGGAGCGCUGAGGCCAACUGCGGGCUCAGACCAUCGU GAAGAGCCUGAAUAAGAUUUACGUGCCUCCUAUCAUCGAGAAUCGGCUGCCCUA CUACGACCCUGGUUCCUGGUGGGCGUGGCCAUUAUCCUGGUGAUCUUCACCGU GGCCAUCUGUAGCAUCAGGCGGAACCGGCCUGAAGUACAGGUACGGGACCUU CCUGUACGUGUGAUAAGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUG CCCCUUGGGCCUCCCCCAGCCCCUCCUCCCUUCCUGCACCCGUACCCCGUGGU CUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAA AA AA
63	1701 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUAAC

		UUUGGCGCCGCCACCAUGACCCAGCUGCCCAUCCUGCUGCUGCUGAUCAGCCUGG UGUACGCCACUCCCAGUCCUCAGACCUCCAAGAAGAUUGGCGACGAUGCCACCAU CAGCUGCUCCCGGAAUAAUACCAACUACUAUGUGGUGAUGAGCGCCUGGUACAA GGAGCCUAAACAGCAUCAUCCUGCUGGCCGCCAAGAGCGAUGUGCUGUACUUCGA UAACUACACCAAGGACAAGAUCAGCUACGAUAGCCCCUACGACGAUCUGGUGACU ACCAUCACCAUCAAGAGUCUGACUGCCGGCGACGCCGGGACUUACAUCUGCGCCU UCUUAUGACCAGCACCACCAUUGACACCGACAAGGUGGACUACGAGGAGUACAG CAUCGAGCUGAUCGUGAACACCGACAGCGAGAGCACCAUCGACAUCAUCCUGUCU GGCAGCACCCCCGAGACUAUCAGCGAGAAGCCUGAGGACAUCGAUAACAGCAACU GUUCCAGCGUUUUCGAGAUCAACCACCCCGAGCCUAUCACCGAUAAACGUGGACGA UCACACCGACACCGUGACCUACACCUCCGAUAGCAUCAACACCGUGAAUGCCAGCA GCGGGGAGAGUACUACAGAUGAGAUCCCCGAGCCUAUCACCGAUAAAGGAGGAGG AUCACACCGUGACCGACACUGUGAGUUACACCACAGUGUCUACCUCCUGCGCAU CGUGACUACCAAGAGUACACCGACGAUGCCGAUCUGUACGAUACCUACAACGAU AACGACACAGUGCCUCCACACCGUGGGAGGCAGCACACCAGCAUCAGCAACUA CAAGACCAAGGAUUUUGUGGGCGGCGGGGCUCCGGCGGCGGAGGCUCCGCCGC CGUGAAGACCCCGUGAUUGUGGUGCCUGUGAUCGAUCGGCCCCUAGCGAGAC AUUUCCCAACGUGCAGGAGCAUCAACGAUCAGAAGUUCGAUGACGUGAAGGA CAACGAGGUGAUGCAGGAGAAGCGGGACGUGGUGAUCGUGAACGACGAUCCUGA CCACUACAAGGACUACGUGUUAUCCAGUGGACCGGGGGCAACUCCGGGAUGA CGACAAGUACACCACUUCUUCUCCGGCUUCUGUAACACCAUGUGUACAGAGGAG ACAAAGCGGAACAUUGCUCGGCACCUGGCCUUGUGGACAGCAAGUUCUUCACU GAGCUGGAGAACAAGACGUGGAGUAUGUGGUGAUCAUAGAGAACGACAAUGU GAUCGAGGACAUUACCUCCUCAGGCCUGUGCUGAAGGCCAUCCACGAUAAGAA GAUCGACAUCUGCAGAUUGCGGGAGAUCAUACCGGGAAUAAGGUGAAGACCGA GCUGGUGAUCGACAAGGAUACGCAUCUUCACCUACACCGGCGGCUAUGAUGU GAGCCUGAGCGCUUACAUAUCCGGGUGACCACCGCCUGAAUAUCGUGGAUGA GAUCAUUAAGUCCGGGGGGGCUUGCCAGCGGGUUCUACUUCGAGAUUGCUAGGA UCGAGAACGAGAUAGAUAACCGGCAGAUCAUGGAUAAUUCGCCAAGUACG UGGAGCACGACCCUCGGCUGGUGGCUGAGCAUCGCUUCGAGACUAUGAAGCCCA ACUUCUGGAGCAGGAUUGGGACCGUGGCCGCCAAGCGGUACCCAGGGGUGAUGU ACACCUUUACCACACCCUGAUACGCGGCGGCGGGGGCAGCGGCGGGGGCGGCGAG CGCUGAGACAAAGGAGUUAAGACCCUGUACAACCUUUUAUCGACAGCUAUCU GCAGAAGCUGGCUCAGCACUCUAUCCCAACCAUGUGACAUGCGCCAUUCACAUU GGGGAGGUGAUCGGCCAGUUAAGAACUGCGCCUGAGGAUACCAAUAAGUGC AUGAGCAACAGCAGGCUGAGCUUACCCUGAUGGUGGAGUCCUUAUCGAGGUG AUCAGCCUGCUGAGCGAGAAGGACAGGAGGGCCAUCGCCGAGGAGAUCCGGAUC GACCUGGACGAUGUGCCAGCGUGGUGAGCAAGCUGGAGAAGAAUUGUAACGCC UACGCCGAGGUGAACAAUAUCAUGAUUAUCAGAAGCUGAAUAUCGGGGAGUGC AGUGCUCCCCUGGGCAGCACAUGCUGCUGCAGAUUGUGAACACCGGCAGCGCCG AGGCCAAUUGCGGGCUGCAGACCAUUGUGAAGUACUGAAUAAGAUACGUGC CACCCAUAUCGAGAACCGGCUGCCAUACUACGAUCCUGGUUCCUGGUGGGGG UGGCCAUUAUUCUGGUGAUCUUCACAGUGGCCAUUCGAGCAUCAGGCGCAACC UGGCUCUGAAGUACAGGUACGGGACCUUCCUGUACGUGUGAUAAGAAUUCUAG
--	--	--

		GCUGGAGCCUCGGUGGCCAUGCUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCC UCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGG CACUAGUAAA AA AAAAAAAAAAAAAA
64	1801 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGGCGCCGCCACCAUGACCCAGCUGCCUAUCCUGCUGCUGCUGAUCAGCCUGG UGUACGCUACCCUAGUCCCCAGACUCCAAGAAGAUCGGGGACGACGCCACGAU CAGCUGCAGCAGGAAUACACCAACUACUACGUGGUGAUGAGCGCCUGGUACAA GGAGCCAAUAGCAUCAUUCUGCUGGCCGCCAAGUCUGACGUGCUGUACUUCGA CAAUJACACAAAGGACAAGAUCUCCUACGACAGUCCUUAUGAUGAUCUGGUGAC CACGAUCACCAUUAAGUCUCUGACUGCUGGGGAUGCCGGUACCUAUUUCUGUGC GUUCUUUAUGACCUCUACAACGAACGACACAGAUAGGUGGACUACGAGGAGUA UAGCAUCGAGCUGAUCGUGAACACCGACAGCGAGAGCACAUAUGACAUUAUCCUC AGCGGGAGCACCCCGAGACAUCUCUGAGAAGCCCGAGGACAUUGAUAAUAGCA AUUGUAGCUCCGUGUUCGAGAUUACUACUCCCGAGCCAUUACCGAUAAUGUGG AUGAUCAUACAGAUACGGUGACCUACACCAGCGACAGCAUACACCGUGAAUGC UAGCAGCGGUGAGAGACACCACCGAUGAGAUCCCGAGCCUUAUACCGAUAAAGGAG GAGGAUCACACCGUGACCGACACGGUGAGUUAACACCACCGUGUCUACCUCUCCG GGAUCGUGACUACCAAGAGCACACCGAUGACGCCGAUCUGUAUGACACCUACAA CGAUAAUGAUACCGUGCCGCCACCACCGUGGGCGGCAGCACACCAGCAUCAGC AACUAUAAGACUAAGGACUUCGUGGGGGGCGGGGGAGCGGCGGCGGGCAG CCGGCUGAAUCAGUGCAUGAGCGCCAACAAGGCCCAUACUGAUUACGCCGUG GCUGUCGCCGCCGCCAGCUCCACCCACCGAAGGUGGUGUCCAGCACCCACAGU ACGAUCACAAGGAGUCCUGCAAUGGCCUGUAUUAACAGGGCUCUUGCUACAUC UGCACAGCGAUUACAAGAGCUUUGAGGAUGCCAAGGCCAAUUGUGCAGCCGAGA GCAGCACGUGCCUAACAAGAGCGACGUGCUGACUACUUGGUGAUCGACUAUG UGGAGGACACCUGGGGCAGUGAUGGGAACCCCAUCACUAAGACCACCUUGACUA CCAGGACUCCGACGUGAGUCAGGAGGUGAGGAAGUACUUCUGUACGGGCGGCGG CGGCAGCGGCGGCGGCGGUGCCACCUGCACCGUGCCCACCAUGAACACGCCAAG CUGACUUCUACCGAGACUUCUUUAACGACAAGCAGAAGGUGACCUUACCCUGCG ACAGCGGCUAUCAUUCUCCUGACCCUAACGCCGUGUGCGAGACAGACAAGUGGAA GUAUGAGAACCCGUGCAAGAAGAUGUGUACGGUGUCCGACUACGUCUCUGAGCU GUACGACAAGCCUCUGUACGAGGUCAACUCCACUAUGACUCUGUCUUGCAACGG CGAGACAAAGUACUUAAGGUGCGAGGAGAAGAACGGCAAUACAGCUGGAACGA CACCGUGACGUGCCCCAACGCCGAGUGUCAGCCUCUCCAGCUGGAGCACGGCAGC UGUCAGCCUGUGAAGGAGAAGUACAGCUUCGGCGAGUACAUGACCAUCAAUCUGU GAUGUGGGGUACGAGGUGAUCGGGGUGAGCUAUACAGCUGCACCGCCAUAUG CUGGAACGUGAUUCCAGCUGUCAGCAGAAGUGCGACAUCCUUCUCUGUCCAAC GGGCUGAUCAGCGGGAGCACGUUCAGCAUUGGCGGGGUGAUCCACCUAGAGCUGC AAGUCAGGGUUCACCCUGACUGGCUCACCCAGCAGCACCUGCAUCGAUGGCAAGU GGAACCCCAUUCUGCCACCUUGUGUGAGAAGCAACGAGGAGUUCGAUCCUGUGG ACGACGGACCAGACGACGAAACCGAUCUGUCCAAGCUGUCCAAGGAUGUGGUGC AGUACGAGCAGGAGAUCGAGAGCCUGGAGGCCACCUACCACAUCAUCAUAGGC

		ACUGACCAUCAUGGGCGUGAUUUCUUGAUCAGCAUCAUUGUGCUGGUGUGUU CUUUGCGACAAGAACAACGACCAGUACAAGUUUCACAAGCUGCUGCCCUGAUAAGA AUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCA GCCCCUCCUCCCCUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAG UGGGCGGCACUAGUAAA AAA AAA
65	1901 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC UUUGGCGCCGCCACCAUGGACUGGACCUGGAUUCUGUUCUGGUGGCGGCCGCC ACAAGGGUGCACUCCGAUGGGACCUUGUUCUCCUGGGGACGACGACCUGGCAUC CCCGCUACGGAGUUCUUCAGCACCAAGGCUGCCAAGAACCCGGAGACUAAGCGGG AGGCCAUCGUAAGGCCUACGGAGAUGACAAUGAGGAGACGCUGAAGCAGCGUC UCACCAACCUUGAGAAGAAGAUACCAACAUCACUACGAAGUUUGAGCAGAUUGA GAAGUGCUGCAAGCGCAAUGAUGAGGUGCUGUUUCGGCUGGAGAACCACGCCGA GACUCUGCGGGCUGCCAUGAUCAGCCUCGCAAGAAGAUACGACGUGCAGACCGGA AGGCACCCUAUGAGGGGGGAGGAGGGAGCGGCGGCGGGAGCGGGGCCGCG GCCAGUAUCCAGACCACCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAAGCUGG AGCAGGAGGCCAACGCCUCUGCUCAGACCAAGUGUGAUUAGAGAUUGGCAUU UCUAUUAUCAGGCAGAACACGGGUGCAACAUCACCGUGAAGAAUUGUGCUCCG CUGACGCCGACGUCAGCUGGACGCCGUGCUGAGCGCGGCCACGGAGACUAUUC CGGUCUGACCCUAGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCGCU GAUAUCCAGACGAGUGUUAACACGGUGGUGCGGGACUUUGAGAACUACGUGA AGCAGACGUGCAACUCCAGCGCCGUGGUGGAUAAUAGCUCAGAUCCAGAAUG UGAUCAUUGACGAGUGUUAUGGGGCCCCUGGGAGCCCCACUACCUCGAGUUUA UUAUACCGGUUCCUCCAAGGGGAACUGCGCCAUAAGGCGCUGAUGCAGUUGA CAACGAAGGCCACCACCCAGAUCGCCCCAGGCAGGUUGCCGGGGCGGCGGGGG GAGCGGUGGCGGCGGCAGCGCUGAGACAAAGGAGUUAAGACCCUGUACAACUU GUUCAUAGAUUCUUAUCUUCAGAAGUUGGCCCAGCACUCCAUCCUACCAAUGU CACUUGUGCCAUCCACAUUGGAGAGGUGAUCGGCCAGUUAAGAACUGCGCCCU GAGGAUCACCAACAAGUGCAUGAGCAAUAGCCGCCUGUCCUUCACGCUCAUGGU GGAGUCCUUAUCGAGGUGAUCUCCUGCUGAGCGAGAAGGACAGGCGGGCUAU UGCUGAGGAGAU CGGUUUGACCUCGACGACGUGCCGUCGUCUCCAAGCU GGAGAAGAACUGCAACGCAUAUGCUGAGGUGAACACAUAUCGAUAUUCAGAA GCUGAAUAUCGGUGAGUGUUCGCUCCCCUGGGCAGCAUAUGCUGUUGCAGAU UGUCAAUACCGGGAGUGCUGAGGCCAACUGCGGGCUGCAGACCAUCGUGAAGUC ACUGAAUAGAUACUGGCCACCCAUAUCGAGAACAGGCUGCCGUACUACGAC CCGUGGUUCUUGGUGGGGGUGGCCAUUAUUCUGGUGAUCUUCACGGUGGCCAU CUGCAGCAUCCGCCGGAACCUGGCUCUGAAGUACCGGUACGGGACCUUCCUGUAC GUGUGAUAAAGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCUUG GGCCUCCCCCAGCCCCUCCUCCCUUCCUGCACCCGUACCCCGUGGUCUUUGAA UAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
66	2001 mRNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC

		UUUGGCGCCGCCACCAUGGAUUGGACCUGGAUUCUGUUCUUCGUUGCCGCCGCC ACCAGGGUGCAUAGUGCGGCAGUGAAGACUCCGUUAUCGUCGUCCCGUGAUU GACCGCCCGCCUCGAGACGUUCCCCAACGUGCAUGAGCACAUC AACGAUCAGA AGUUCGACGAUGUCAAGGACAAUGAGGUCAUGCAGGAGAAGCGGGACGUGGUG AUCGUCAACGACGAUCCCGAUCACUACAAGGAUUAACGUCUUAUCCAGUGGACG GGCGGCAACAUCGGGACGACGAUAAGUACACGCAUUUCUUCUCCGGCUUCUGU AACACCAUGUGUACAGAGGAGACGAAGCGGAACAUCGCCAGGCAUCUGGCGCUG UGGGACUCGAAGUUCUUCACCGAGCUGGAGAAUAAGAACGUGGAGUACGUGGU GAUCAUCGAGAACGACAACGUAUCGAGGACAUCACGUUUCUGCGGCCGGUGCU GAAGGCCAUCCACGACAAGAAGAU CGACAUCUGCAGAU GCGGGAGAUCAUCACG GGUAAUAAGGUGAAGACCGAGCUGGUCAUUGACAAGGACCACGCAUAUUAUACC UACACUGGGGGAUACGACGUGUCCUUGUCCGCCUAUAUUAUCCGGGUGACCACC GCUCUGAACAU CGUGGAUGAGAUCAUCAAGUCGGGAGGCCUCAGCUCCGGCUUC UACUUCGAGAU CGCCCGAUCGAGAACGAGAU GAAGAUCAACCGCCAGAUCAUG GACAACAGCGCCAAGUAUGUCGAGCACGACCCUCGGCUGGUGGCAGAGCACAGAU UCGAGACUAUGAAGCCCAACUUCUGGAGUAGGAUCGGCACUGUCGCCCAAGC GCUACCCAGGUGUGAUGUACACCUUACCAACCCCGCUGAUCAGCGGGGGGGGG GUAGCGGCGGCGGCGGCGAGUGCCGAGACAAAGGAGUUAAGACCCUGUAACAUCU UGUUAUCGACUCCUACCUCAGAAAGUUGGCUAGCACUCGAUCCCCACCAUUGU GACAUGCGCAUUCACAUUGGUGAGGUGAUCGGGCAGUUAAGAACUGUGCUCU GCGGAUUAACCAACAAGUGUAUGAGUAAUUCGCCUGUCCUUCACGCUAUGGU GGAGUCCUUAUCGAGGUGAUCUCCUGCUGAGCGAGAAGGACAGGCGGGCCAUC CGCCGAGGAGAU CGGCAUCGACCUGGACGAUGUCCGUCUGUCGUCUCCAAGCU GGAGAAGAACUGCAACGCAUAUGCUGAGGUGAACACAUCAU CGAUUAUCAGAA GCUGAUAUCGGUGAGUGUUCGCUCCCCUGGGCAGCAUAUGCUGUUGCAGAU CGUGAAUACCGGGAGUGCGGAGGCGAACUGCGGGCUCCAGACCAUCGUGAAGAG CCUGAACAAAGAUCAUGUCCCGCCUAUCAUCGAGAAUCGGCUGCCGUACUAUGAC CCUGGUGGCGGCGGCGAGCGGGGGGGGGGGGUCGGGGGCCGUCGUUCCAUCCA GACCACCGUGAACACUCUGUCUGAGCGUAUCAGCAGCAAGCUGGAGCAGGAGGC CAACGCCUCUGCUCAGACCAAGUGUGAUUAAGAGAUUGGCAUUUCUAUAUCAG GCAGAACACGGGUGCAACAUACCGUGAAGAAUAUGUGCUCCGUGACGCCGAC GCUAGCUGGACGCCGUGCUGAGCGCGGCCACGGAGACAUUAUCCGGUCUGACC CUGAGCAGAAGGCGUAUGUGCCUGCCAUGUUCACAGCUGCGCUGAAUAUCCAGA CGAGUGUUAACACGGUGGUGCGGGAUUCGAGAAUAUGUGAAGCAGACCUGU AACAGCUCGGCUGUUGUGGACAACAAGCUCAAGAUCCAGAAUGUGAUCAUUGAC GAGUGUUAACGGGGCCCCCGGCUCCCCACGAUUCGGAGUUAUCAACACGGGCA GCAGCAAGGGCAACUGCGCAUUAAGGCGCUGAUGCAGCUGACCACUAAGGCCAC AACCCAGAU CGCCCUAGGCAGGUGGCCGGGACUGGCGUGCAGUUCUAUAUGAU AGUGAUUGGCGUGAUCAUCCUGGCCCGCCUGUUAUGUACUACGCCAAGCGCAU GCUCUUUAUCUACCAACGAUAAGAUCAAACUUAUCUUGGCGAAUAAGGAGAA UGUGCAUUGGACGACGUACAUGGACACCUUCUUCGCACAUCACCUAUGAUCAU UGCCACGACUGACAUUCAGAACUGAUAAAGAAUUCUAGGCUGGAGCCUCGGUGGC CAUGCUCUUGCCCCUUGGGCCUCCCCCAGCCCCUCUCCCCUUCUGACCCCGU ACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGCGGCACUAGUAAAAAAAAA
--	--	--

		AA AA
67	0201 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCGCCACTAGGGT GCACAGCGACGGGACCCTGTTCCAGGGGACGACGACCTGGCCATCCCCGCTACCGA GTTCTTCAGCACCAAGGCTGCCAAGAACCCCGAGACTAAGCGGGAGGCCATCGTGAA GGCCTACGGGGACGATAATGAGGAGACGCTGAAGCAGCGTCTACCAATCTGGAGA AGAAGATTACTAATATCACCACCAAGTTCGAGCAGATCGAGAAGTGCTGTAAGAGGA ATGACGAGGTGCTGTTTCGGCTGGAGAACCACGCCGAGACACTGCGGGCTGCCATGA TCTCTTTGCCAAGAAGATCGATGTGCAGACCGGCAGGCACCTTATTACCCTTACGA CGTGCCCGATTATGCCGAGTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCT TCTTGCCCTTGCGCCTCCCCCAGCCCTCCTCCCTTCTGCAACCGTACCCCGTG GTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
68	0302 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGATGACTCCCAGAAATGACGAGGAGCAGACCAGCGTGTCTCCGC TACCGTGTACGGGGACAAGATCCAGGGGAAGAATAAGAGGAAGCGGCTGATCGGCC TGTGCATCAGGATCAGCATGGTGATCAGCCTGCTGAGCATGATCACCATGCTGCATT CCTGATCGTGCGGCTGAATCAGTGCATGAGCGCCAACAAGGCCGCCATCACTGATTC AGCCGTGGCCGTGGCCGCCGCTTCTCTACTCACAGGAAGGTGGTGTCTCCACCACT CAGTACGATACAAGGAGAGCTGCAATGGCCTGTATTACCAGGGCTCTTGCTACATCC TGACAGCGATTACAAGAGCTTTGAGGATGCCAAGGCCAATTGCGCCGCCGAGAGCA GCACTCTGCCTAACAAGAGCGACGTGCTGACTACTTGGCTGATCGACTATGTGGAGG ACACCTGGGGGAGCGATGGGAACCCCATCACAAGACTACCAGCGACTATCAGGACA GCGACGTGAGCCAGGAGGTGCGGAAGTACTTCTGCACCTACCCTTACGACGTGCCTG ACTACGCTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTG GGCCTCCCCCAGCCCTCCTCCCTTCTGCAACCGTACCCCGTGCTTCTTGATAA AGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAA AAA
69	0402 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGACCCAGCTGCCTATCCTGCTGCTGCTGATCAGCCTGGTGACGCT ACCCCTAGTCCCAGACTTCCAAGAAGATCGGGGACGACGCCACTATCAGCTGCAGC AGGAATAACACCAACTACTACGTGGTGATGAGCGCCTGGTATAAGGAGCCTAACAGC ATCATCCTGCTGGCCGCCAAGAGCGATGTGCTGTACTTCGACAACCTATACCAAGGACA AGATCTCTTACGATAGCCCCTACGATGACCTGGTGACTACCATCACCATCAAGAGTCT GACCGCCGGGGATGCCGGGACCTATATCTGTGCGTTTTTATGACCTTACAACAAAC GACACAGATAAGGTGGACTACGAGGAGTATAGCATCGAGCTGATCGTGAACACCGA CAGCGAGAGTACCATCGACATCATCCTGAGCGGGAGTACTCCCGAGACTATCAGCGA GAAGCCAGAGGACATCGACAATAGCAACTGCAGCAGCGTGTTCGAGATCACTACTCC TGAGCCTATTACAGATAACGTGGATGATCACACCGACACCGTGACCTACACAGCGAC AGCATCAACACCGTGAATGCTAGCAGCGGTGAGAGTACCACGGATGAGATCCCCGAG CCCATCACGGATAAGGAGGAGGATCACACCGTGACCGACCGGTGAGTTACACCACC

		GTGTCTACCTCCTCAGGGATTGTGACCACAAAGTCTACCACCGATGATGCCGATCTGT ATGACACATACAATGACAACGATACCGTGCCGCCACCACCGTGGGCGGCAGCACAA CATCCATTAGTAACTACAAGACTAAGGATTTCGTGGAGATCTTTGGGATCACAAACCT GATTATCCTGAGCGCCGTGGCGATCTTCTGTATCACCTACTATATCTGTAATAAGCATC CCAGGAAGTATAAGACCGAGAACAAGGTGTATCCGTACGACGTGCCGGACTACGCCT GATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCC CCAGCCCCTCCTCCCTTCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGT GGGCGGCACTAGTAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
70	0505 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGAAGACCATCTCCGTGGTGACCCTGCTGTGCGTGCTGCCAGCAGT GGTGTACAGTGGCGGCGGGGGAGCTACCCCTACGACGTGCCGACTACGCCACCT GTACGGTGCCAACCATGAACAACGCTAAGCTGACCTCCACAGAGACGAGTTTAAATG ACAAGCAGAAGGTGACCTTCACCTGCGACAGTGCGTACCACAGCTTGACCCCAACG CTGTGTGCGAGACAGACAAGTGGAAGTACGAGAACCCGTGCAAGAAGATGTGCACC GTGAGTGACTACGTCAGCGAGCTGTACGACAAGCCGCTGTACGAGGTCAACTCCACG ATGACACTCTCTTGCAACGGGGAGACAAAGTACTTCAGGTGCGAGGAGAAGAACGG GAATACCAGCTGGAATGATACCGTGACATGCCCAATGCCGAGTGCCAGCCCTGCA GCTCGAGCATGGGAGCTGTGAGCCGTGAAGGAGAAGTACTCCTTCGGGGAGTACAT GACCATCAATTGCGACGTGGGGTACGAGGTGATCGGGGTGAGCTACATCTCTTGAC CGCCAATAGCTGGAACGTGATTCCCAGCTGTGACGAGAAGTGCGACATCCCTTCTCTG TCCAACGGGCTGATCAGCGGGAGCACGTTTACGATTGGCGGGGTGATCCACCTGAGC TGCAAGAGTGGGTTACCCCTGACCGGGTCCCCGAGCAGCACCTGCATCGATGGGAAG TGGAACCCGATCTGCCAATTGCGTCCGGTCCAACGAGGAGTTGACCCCGTGGAC GACGGACCAGACGACGAAACCGATCTGAGCAAGTTGAGCAAGGACGTGGTGCAGTA CGAGCAGGAGATCGAGAGCCTGGAGGCCACCTACCACATCATCATATGGCACTGAC CATCATGGGCGTGATATTCTTGATCAGCATCATTGTGCTGGTGTGTTCTTGCGACAAG AACAACGACCAGTACAAGTTTACAAGCTGCTGCCCTGATAAGAATTCTAGGCTGGA GCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCCTCTCCCTTCTCT GCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAA AA AA AAA
71	0603 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCTACTCGGGT GCATAGCGGCGGCGGGGGTCTACCCCTACGACGTGCCGACTATGCCGCCGAGAC AAAGGAGTTTAAGACCCTGTACAACCTGTTTATCGACAGCTATCTGCAGAAGCTGGCT CAGCACTCTATTCCACCAATGTGACATGCGCCATTACATTGGTGAGGTGATCGGCC AGTTCAAGAACTGCGCCCTGAGGATCACGAATAAGTGCATGAGCAACAGCAGGCTGA GCTTACCCTGATGGTGGAGTCCTTCATCGAGGTGATCAGCCTGCTGAGCGAGAAGG ACAGGAGGGCCATCGCCGAGGAGATCGGCATCGACCTGGACGATGTGCCAGCGTG GTGAGCAAGCTGGAGAAGAACTGCAACGCATATGCTGAGGTGAACAACATCATCGAT ATTCAGAAGCTGAATATCGGTGAGTGTTCCGCTCCCCCTGGGCAGCATATGCTGTTGC

		AGATCGTGAACACGGGCAGTGCCGAGGCCAACTGCGGGCTGCAGACCATCGTGAAG TCACTGAATAAGATCTACGTGCCACCCATCATCGAGAACCGGCTGCCCTACTACGACC CGTGGTTCCTGGTGGGGGTGGCCATTATTCTGGTGATCTTCACGGTGGCCATCTGCA GCATCCGCAGGAACCTGGCCCTGAAGTACAGGTACGGGACCTTCTGTACGTCTGAT AAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCA GCCCCTCTCCCTTCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGG CGGCACTAGTAA AA AAAAAAAAAAAAAAAAAAAA
72	0702 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTCTGGTGGCTGCCGCCACCAAGGT GCACAGTGGCGGAGGAGGCAGCTACCCCTACGACGTCCCCGACTACGCTGGGGCTG CTGCCTCCATCCAGACCACTGTGAACACCCTGAGTGAGCGGATCAGCAGCAAGCTGG AGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTGATATAGAGATTGGCAATTTCT ATATCAGGCAGAACCACGGGTGCAACATCACCGTGAAGAATATGTGCTCCGCTGACG CCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGAGACATATTCCGGTCTGA CCCCTGAGCAGAAGGCGTATGTGCCTGCCATGTTACAGCTGCACTGAACATCCAGAC CAGCGTGAACACCGTGGTGAGGGACTTCGAGAACTACGTGAAGCAGACATGCAACA GCAGCGCGGTGGTGGATAATAAGCTCAAGATCCAGAATGTGATCATTGACGAGTGTT ATGGGGCCCCTGGGAGCCCCACTAACCTCGAGTTTATTAATACGGGCAGCAGCAAGG GCAACTGCGCCATCAAGGCCCTGATGCAGCTGACAACCAAGGCCACCACCCAGATTG CCCCAGGCAGGTGGCCGTTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGC TTCTTGCCCTTGGGCCTCCCCCAGCCCTCCTCCCTTCTGCACCCGTACCCCGTG GTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAA AA AA
73	0803 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTTCTGGTGGCCGCCGCCACAGGT GCATTCCGGCGGCGGCGGCAGCTACCCATACGATGTGCCTGACTACGCCGCCGCCGT GAAGACACCCGTGATCGTGGTGCCCGTGATTGACCGGCCCCCTAGCGAGACATTCCC CAATGTGCACGAGCACATCAACGACCAGAAGTTCGACGACGTGAAGGACAACGAGG TGATGCAGGAGAAGAGGGACGTGGTGATTGTGAACGACGATCCTGACCACTACAAG GACTACGTGTTTCATCCAGTGGACAGGCGGGAACATCAGGGACGACGACAAGTATACC CACTTCTTCTCCGGCTTCTGTAACACCATGTGTACAGAGGAGACAAAGCGGAACATTG CTCGGCACCTGGCCCTGTGGGACAGCAAGTTCTTACCGAGCTGGAGAACAAGAACG TGGAGTATGTGGTGATCATCGAGAACGACAACGTCATCGAGGACATCACCTTTCTGC GCCCTGTGCTGAAGGCCATCCACGATAAGAAGATCGACATCCTGCAGATGCGGGAGA TCATCACCGGGAATAAGGTGAAGACCGAGCTGGTGATCGACAAGGATCACGCCATCT TCACCTACACCGGCGGCTATGATGTGAGCCTGAGCGCTACATCATCCGGGTGACCAC CGCCCTGAATATCGTGGATGAGATCATCAAGTCCGGCGGGCTGAGCAGCGGCTTCTA CTTCGAGATCGCCAGAATCGAGAACGAGATGAAGATCAACAGGCAGATCATGGACA ACAGCGCCAAGTACGTGGAGCACGACCCTCGGCTGGTGGCCGAGCACCGGTTTCGAG ACTATGAAGCCCAACTTCTGGTCCAGAATCGGGACTGTGGCCGCCAAGCGGTACCCC GGCGTGATGTACACATTCACCACCCCACTGATCAGCTTCTTTGGCCTGTTGACATCAA

		CGTGATCGGCCTGATCGTGATCCTGTTTCATCATGTTTCATGCTGATCTTCAACGTGAAGT CCAAGCTGCTGTGGTTCTGACCGGCACCTTCGTGACTGCCTTCATCTGATAAGAATT CTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCTCCCCCAGCCCCTC CTCCCCTTCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCAC TAGTAA AA AAAAAAAAAAAAA
74	0201-J0 mRNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGT CCACTCCATGGACGGAACCTTTTCCCGGAGATGACGATCTTGCAATTCCAGCAACT GAATTTTTCTCTACAAAGGCTGCTAAAAATCCAGAGACTAAACGCGAAGCAATTGTTA AAGCCTATGGAGACGATAATGAGGAACTCTCAAACAACGGCTAACTAATTTGGAAA AAAAGATTACTAATATAACAACAAAGTTTGAACAAATAGAAAAGTGTGTAAACGCA ACGATGAAGTTCTATTTAGGTTGGAAAATCACGCTGAACTCTAAGAGCGGCTATGAT ATCTCTGGCTAAAAAGATTGATGTACAGACTGGACGGCATCCATATGAGTAAGAATTC TAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCTCCCCCAGCCCCTCC TCCCCTTCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACT AGTAA AA AAAAAAAAAAAAA
75	0301-J0 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGT CCACTCCATGATGACACCAGAAAACGACGAAGAGCAGACATCTGTGTTCTCCGCTACT GTTTACGGAGACAAAATTCAGGGAAAGAATAAACGCAAACGCGTGATTGGTCTATGT ATTAGAATATCTATGGTTATTTCACTACTATCTATGATTACCATGTCCGCGTTTCTCATA GTGCGCTAAATCAATGCATGTCTGCTAACAAGGCTGCTATTACTGACTCCGCTGTTG CCGTTGCTGCGGCATCATCTACTCATAGAAAGTTGTGTCTAGCACTACACAATATGA TCACAAAGAAAGCTGTAATGGTTTATATTACCAGGGTCTTGTTATATATTACATTGAG ACTATAAGTCATTGAGGATGCTAAAGCAAACGCGCTGCGGAATCATCAACACTACC CAATAAATCCGATGTCTTGACTACCTGGCTCATTGATTATGTTGAGGATACATGGGGA TCTGATGGTAATCCAATTACAAAACACTACATCCGATTATCAAGATTCTGATGTATCACA AGAAGTTAGAAAGTATTTTGTACATAAGAATTCTAGGCTGGAGCCTCGGTGGCCAT GCTTCTTGCCCTTGGGCTCCCCCAGCCCCTCTCCCCTTCTGCACCCGTACCCCG TGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAA AA AAA
76	0501-J0 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGT CCACTCCATGAAAACGATTTCCGTTGTTACGTTGTTATGCGTACTACCTGCTGTTGTTT ATTCAACATGTACTGTACCCACTATGAATAACGCTAAATTAACGTCTACCGAAACATC GTTAATGATAAACAGAAAGTTACGTTTACATGTGATTGAGGATATCATTCTTTGGAT CCAAATGCTGTCTGCGAAACAGATAAATGAAATACGAAAATCCATGCAAGAAAATG TGACAGTTTCTGATTATGTCTCTGAACTATATGATAAGCCATTATACGAAGTGAATTC CACCATGACACTAAGTTGCAACGGTGAAACAAAATATTTTCGTTGTGAAGAAAAAAT

		GGAAATACTTCTTGAATGATACTGTCACGTGCTTAATGCGGAATGTCAACCTCTTC AATTAGAACACGGATCGTGTCAACCAAGTTAAAGAAAAATACTCATTTGGGGAATATAT GACTATCAACTGTGATGTTGGATATGAGGTTATTGGTGTTCGTATATAAGTTGTACG GCTAATTCTTGAATGTTATTCCATCATGTCAACAAAAATGTGATATACCGTCCCTATC TAATGGATTAATTTCCGGATCTACATTTTCTATCGGTGGCGTTATACATCTTAGTTGTA AAAGTGGTTTTACTAACGGGGTCCCCATCATCCACATGTATCGACGGTAAATGGAA TCCCATACTCCCAACATGTGTACGATCTAACGAAGAATTTGATCCAGTGGATGATGGT CCCGACGATGAGACAGATCTGAGCAAACTCTCGAAAGACGTTGTACAATATGAACAA GAAATAGAATCGTTAGAAGCACTTATCATATAATCATAATGGCGTTGACAATTATGG GTGTCATATTTCTAATCTCCATTATAGTATTAGTTTGTTCCTGTGACAAAAATAATGAC CAATATAAGTTCATAAATTGCTACCGTGAGAATTCTAGGCTGGAGCCTCGGTGGCCA TGCTTCTTGCCCTTGGGCCTCCCCCAGCCCCTCTCCCCTTCTGCACCCGTACCCCC GTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAA AA AA
77	0701-J0 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGT CCACTCCATGGGTGCCGACGAAGCATACAGACGACTGTGAATACACTCAGTGAACG TATCTCGTCTAAATTAGAACAAGAAGCGAACGCTAGTGCTCAAACAAATGTGATATA GAAATCGGAAATTTTATATCCGACAAAACCATGGATGTAACATCACTGTTAAAAATA TGTGCTCTGCGGACGCGGATGCTCAGTTGGATGCTGTGTTATCAGCCGTACAGAAA CATATAGTGGATTAACACCGGAACAAAAAGCATACGTACCAGCTATGTTTACTGCTGC GTTAAACATTCAGACGAGTGTAACACTGTTGTTAGAGATTTTGAATATATGTGAAA CAGACTTGTAATTCTAGCGCTGTTGTCGATAACAAATTAAGATACAAACGTAATTA TAGATGAATGTTACGGAGCCCCAGGATCTCCAACAAATTTGGAATTTATTAATACAGG ATCTAGCAAAGGAAATTGTGCCATTAAGGCGTTGATGCAATTGACTACTAAGGCCACT ACTCAAATAGCACCTAGACAAGTTGCTGGTACAGGAGTTCAGTTTTATATGATTGTTA TCGGTGTTATAATATTGGCAGCGTTGTTTATGTAATGATGCAAGCGTATGCTGTTTAC ATCCACCAATGATAAAATCAAATTTATTTAGCCAATAAGGAAAACGTCCATTGGACT ACTTACATGGACACATTTTGAACCTTCTCCGATGATTATTGCTACCACGGATATACA AACTGAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCC CCCCAGCCCCTCTCCCCTTCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGA GTGGGCGGCACTAGTAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
78	0901-J0 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTCTTCTTGGTGGCAGCAGCCACGCGAGT CCACTCCATGCCTCAACAATATCTCTATTAATATAGAACTAAAAAGCAATTTCTG ACGCGCGATTGAAGACGTTAGACATACATTATAATGAGTCGAAACCAACCACTATCCA GAATACTGGAAAAGTAGTAAGGATTAATTTAAAGGAGGATATATAAGTGGAGGGTT TCTCCCAATGAATATGTGTTATCAACAATACATATATATTGGGGAAAGGAAGACGAT TATGGATCCAATCACTTGATAGATGTATACAAATACTCTGGAGAGATTAATCTGGTTC ATTGGAATAAGAAAAATATAGTTCTTATGAAGAGGCAAAAAACACGATGATGGAA TTATCATTATTGCTATATTCTTACAAGTATCAGATCATAAAATGTATATTTCAAAG

		ATAGTCAATCAATTGGATTCCATTAGATCCGCTAATATGTCTGCGCCATTGATTCACT ATTTTATCTAGACAATTTGCTACCTAGTACGTTGGATTATTTTACATATCTAGGAACAA CTATCAACCACTCTGCAGACGCAGCATGGATAATTTTCCAACACCAATAAACATTCAT TCTGATCAACTATCTAAATTCAGAACACTATTGTCGTCGTCTAATCATGAGGGTAAACC TCATTATATAACAGAGAATTATAGAAATCCGTATAAATTGAACGACGACACGCAAGTA TATTATTCTGGGGAGATTATACGAGCTGCAACCACCTCACCAGTTCGAGAGAACTATT TTATGAAATGGTTGTCTGATTGAGAGAGGCATGTTTTTCATATTATCAAAAATATATC GAAGGGAATAAAACATTCGCAATTATTGCCATAGTATTCTGATTATTACTACCGCTAT TCTCTTCTTATGAGTCAACGATATTGCGGAGAAAAACAACTAGGAATTCTAGGCT GGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCTCCTCCCTT CCTGCACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAA AA AA AAAAA
79	1101 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACGTGGATTCTTCTCTGGTGGCAGCCGCCACTCGTGTG CACTCCGATGGGACCTTGTTCCTGGGGACGACGACCTGGCCATCCCCGCTACGGAG TTCTTCAGCACCAAGGCTGCCAAGAACCCGGAGACTAAGCGGGAGGCCATCGTCAAG GCCTACGGAGATGACAACGAGGAGACGCTGAAGCAGCGGCTGACCAACCTGGAGAA GAAGATCACGAACATCACGACCAAGTTCGAGCAGATTGAGAAGTGTGCAAGCGGA ACGACGAGGTGTTGTTCCGCTTGGAGAACCACGCAGAGACTCTGCGGGCAGCCATGA TCAGCTTGGCGAAGAAGATCGACGTGCAGACTGGACGCCACCCCTATGAGGGAGGA GGCGGCAGCGGCGGCGGCGGAGCCGGCTGAATCAGTGCATGAGCGCAACAAGG CCGCCATCACTGATTAGCCGTCGCCGTCGCTGCCGCTCCTCCACTCATAGGAAGGT GGTGTCCAGTACGACACAGTACGACCATAAGGAGAGCTGTAATGGGCTGTACTACCA GGGGTCTGTCTACATCCTGCACAGCGACTACAAGTCGTTTGAGGATGCGAAGGCGAA CTGCGCAGCCGAGAGCAGCACGTTGCCTAACAGAGCGACGTGCTGACTACTTGGCT GATTGACTATGTCGAGGATACCTGGGGCTCGGATGGGAACCCCATCACGAAGACCAC GTCTGACTACCAGGACTCCGATGTGTCTCAGGAGGTGCGGAAGTACTTCTGTACTGG CGGCGGCGGTTCTGGTGGGGGCGGCGAGCACCTGCACCGTCCCCACCATGAACAACGC CAAGCTGACCAGTACAGAGACTTCCTTCAACGATAAGCAGAAGGTGACATTTACATG CGATTCCGGCTACCACTCGTTGGACCCCAACGCTGTCTGTGAGACAGACAAGTGAA GTACGAGAATCCATGTAAGAAGATGTGCACGGTGTCCGACTATGTTTCTGAGCTGTAC GACAAGCCCCTTTACGAGGTGAACAGCACCATGACTCTCAGCTGCAACGGCGAGACT AAGTACTTTCGGTGTGAGGAGAAGAACGTAATACGTCCTGGAACGATACCGTTACT TGCCCCAACGCCGAGTGCCAGCCGTTGCAGCTGGAGCATGGTAGCTGTCAGCCCGTG AAGGAGAAGTATAGTTTCGGCGAGTACATGACCATTAAATTGTGATGTGGGTTACGAG GTCATCGGGGTGAGTTACATCTCTTGCACTGCCAATAGCTGGAACGTGATTCCCAGCT GTCAGCAGAAGTGTGACATACCTTCTGTGTCAACGGGCTGATCAGCGGGAGCACGT TCAGCATTGGCGGGGTGATCCACCTCAGCTGCAAGAGCGGGTTACGTTGACCGGGT CGCCCTCGTCGACCTGCATCGACGGGAAGTGAATCCATTCTGCCACGTGCGTGC GGTCGAACGAGGAGTTCGACCCGGTCGACGATGGACCCGATGACGAGACTGATCTG TCTAAGTTGTGCAAGGACGTCGTTCACTACGAGCAGGAGATCGAGTCCCTGGAGGCC ACGTACCACGGCGGTGGCGGGAGCGGGGAGGTGGCAGCAGTCCCCAGACTTCCAA

		GAAGATCGGGGACGACGCCACCATCTCCTGCTCCCGCAACAACACTAACTACTACGTG GTTATGTCCGCCTGGTACAAGGAGCCCAACTCGATCATCCTGCTCGCTGCTAAGAGCG ACGTCCTGTACTTCGACAACATATACCAAGGACAAGATCAGCTACGACTCACCTACGA TGACCTCGTGACCACCATCACAATTAAGAGTCTGACCGCCGGGGACGCCGGGCACATA CATTTGCGCCTTCTTCATGACGAGTACCACCAATGATACCGACAAGGTGGATTATGAG GAGTACTCGATCGAGCTTATCGTGAACACGGACAGTGAGTCGACCATCGACATCATC CTGTCCGGTAGCACCCCTGAGACAATCTCCGAGAAGCCGGAGGACATCGACAATAGC AATTGCAGTTCCTGCTTCGAGATCACGACCCCTGAGCCCATTACCGACAACGTGGATG ATCACACCGACACGGTGACGTACACCAGCGACTCTATTAACACCGTCAATGCTTCCAG TGGGGAGTCCACCACTGACGAGATTCCGGAGCCCATACCGACAAGGAGGAGGATC ACACCGTGACCGACACGGTGAGTTACACCACCGTGTCTACCTCCTCTGGTATTGTGAC TACCAAGAGCACCACGGACGACGCCGATCTGTACGACACGTACAATGACAACGACAC CGTGCCGCCACCAACCGTGGGCGGGTCCACGACGTCCATTAGTAACTACAAGACTAA GGACTTCGTGGGGGGCGGCGGTTCCGGTGGCGGGGGCAGCGCTGAAACCAAGGAG TTCAAGACGCTGTACAACCTGTTTCATCGACTCTTACCTGCAGAAGTTGGCCAGCACT CCATCCCTACCAATGTCACTTGTCATCCACATTGGAGAGGTGATCGGCCAGTTCAA GAACTGCGCCCTGAGGATCACCAACAAGTGCATGAGCAATAGCCGCCTGTCTTCAC GCTCATGGTGGAGTCTTCATCGAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCG GGCTATTGCTGAGGAGATCGGTATTGACCTCGACGACGTGCCGTCCGTCGTCTCCAA GCTGGAGAAGAAGTCAACGCATATGCTGAGGTGAACAACATCATCGATATTAGAA GCTGAATATCGGTGAGTGTTCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATTGTC AATACCGGGAGTGTGAGGCCAACTGCGGGTTGCAGACCATCGTGAAGAGTCTGAA CAAGATTTACGTGCCCCGATCATCGAGAACCGCCTGCCCTACTACGACCCCGGTGGG GGCGGGTCTGGAGGAGCGGCAGCGCGCTGCCGCCTCCATCCAGACCACGGTGAA TACGTTGTCCGAGCGGATTCGAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCA GACCAAGTGTGATATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCAA CATCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCT GAGCGCGGCCACGGAGACATATTCGGTCTGACCCCTGAGCAGAAGCGGTATGTGCC TGCCATGTTACAGCTGCTCTGAACATCCAGACGTCCGTGAATACCGTGGTGAGGGA CTTCGAGAAGTACGTGAAGCAGACGTGTAATAGCAGCGCTGTCGTCGATAATAAGCT GAAGATCCAGAACGTTATTATCGACGAGTGCTACGGCGCCCCCGGCTCCCCGACCAA CCTGGAGTTCATTAATACCGTTCTCCAAGGGGAACTGCGCTATTAAGGCGCTGATG CAGCTGACTACGAAGGCGACGACCCAGATCGCGCCTGCCAGGTGGCGGGGACCGG GGTGCAGTTCTACATGATCGTGATCGGTGTGATCATCCTGGCAGCTCTCTTTATGACT ACGCCAAGCGCATGCTGTTACCCAGCACCAACGACAAGATTAAGCTGATCCTCGCCAA TAAGGAGAACGTCCACTGGACCACCTACATGGACACATTCTCCGACCAGCCCTATG ATTATCGCCACCACTGACATTGAGAACTGATAAGAATTCTAGGCTGGAGCCTCGGTGG CCATGCTTCTTGCCCTTGGGCTCCCCCAGCCCTCCTCCCTTCTGCAACCCGTACC CCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAA AA AA
80	1201 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTGG CGCCGCCACCATGACCCAGCTGCCTATCCTGCTGCTGCTGATCAGTCTGGTGTACGCT ACACCGTCTCCCAGACTTCCAAGAAGATCGGGGACGACGCGACGATCAGCTGCAGC

		AGGAATAACACCAACTACTACGTGGTGATGAGCGCCTGGTACAAGGAGCCGAATAGC ATCATTCTGCTGGCCGCAAGTCTGACGTGCTGTACTTCGACAATTACACAAAGGACA AGATCTCCTACGACAGTCCTTATGATGATCTGGTGACCACGATCACCATTAAGTCTCTG ACTGCTGGGGATGCCGGTACGTATATCTGCGCCTTCTTCATGACGTCCACGACCAATG ACACTGATAAGGTGGATTACGAGGAGTACTCCATCGAGCTGATCGTGAACACCGACA GCGAGTCTACCATCGACATCATCTGTCCGGGAGCACCCCCGAGACTATCAGTGAGA AGCCTGAGGATATTGACAACAGCAATTGCAGTAGTGTCTTCGAGATCACCACCCCGG AGCCTATCACTGACAACGTGGACGATCACACCGACACGGTGACGTACACCACTGATA GCATCAACACCGTGAATGCTAGCTCCGGGGAGTCCACGACCGATGAGATCCCGGAGC CTATCACCGATAAGGAGGAGGATCACACCGTGACCGACACGGTGAGTTACACCACCG TGCTACCTCCTCCGGGATCGTGACTACGAAGAGCACTACTGACGATGCTGATCTGTA CGATACCTACAACGATAACGATAACCGTGCCGCCCACCACCGTGGGCGGCAGCACCAC CAGCATCAGCAACTATAAGACTAAGGACTTCGTGGGGGGCGGGGGGAGCGGCGGC GGCGGCAGCCGGCTGAATCAGTGATGAGCGCCAACAAGGCCCATCACTGATTCA GCCGTGGCTGTGCGCCGCCGACGTCCACCCACCGGAAGGTGGTGTCAGCACCACC CAGTACGATCACAAGGAGTCTGCAATGGCCTGTATTACCAGGGCTCTTGCTACATCC TGACAGCGATTACAAGAGCTTTGAGGATGCCAAGGCCAATTGTGCAGCCGAGAGCA GCACGCTGCCTAACAAGAGCGACGTGCTGACTACTTGGCTGATCGACTATGTGGAGG ACACCTGGGGGAGTGATGGGAACCCCATCTAAGACCACGTCCGACTATCAGGACA GTGACGTGAGTCAGGAGGTGCGGAAGTACTTCTGCACCGCGCGGGGGGAGCGGC GGGGGCGGCAGCGCTGAGACAAAGGAGTTCAAGACCCTTTACAACCTGTTTCATCGAT AGTTATCTGCAGAAGCTGGCCAGCACAGCATCCCCACCAATGTGACATGCGCCATTTC ACATTGGTGAGGTGATTGGCCAGTTCAAGAACTGCGCCTGCGGATAACTAACAAGT GTATGTCTAATTCGCGCTGTCTTCACGCTCATGGTGGAGTCTTCATCGAGGTGATC TCCCTGCTGAGCGAGAAGGACAGGCGGGCTATCGCCGAGGAGATCGGCATCGACCT GGATGACGTCCCCAGCGTGGTGCCAAGCTGGAGAAGAACTGCAACGCATATGCTGA GGTGAACAACATCATCGATATTAGAACTGAATATCGGTGAGTGTTCCGCTCCCCCT GGGCAGCATATGCTGTTGCAGATCGTCAATACCGGGTCTGCCGAGGCCAACTGCGGG CTGCAGACCATCGTGAAGTCACTGAATAAGATCTACGTGCCACCCATCATCGAGAACC GGCTGCCCTACTACGACCGTGTTCTGGTGGGGGTGGCCATTATTCTGGTGATCTT CACGGTGGCCATCTGCAGCATCCGACGAACTGGCCCTGAAGTACCGGTACGGTAC CTTCCTTACGTGTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCC CCTTGGGCCTCCCCCAGCCCCTCCTCCCTTCTGCACCCGTACCCCCGTGGTCTTTG AATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA AA
81	1301 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCCACAAGGT GCACTCCGATGGGACCTTGTTCCTGGGGACGACGACCTGGCCATCCCCGCTACGGA GTTCTTCAGCACCAAGGTGCCAAGAACCCGGAGACTAAGCGGGAGGCCATCGTCAA GGCCTACGGAGATGACAATGAGGAGACGCTGAAGCAGCGTCTACCAACCTTGAGA AGAAGATACCAACATCACTACGAAGTTTGAGCAGATTGAGAAGTGCTGCAAGCGCA ATGATGAGGTGCTGTTTCGGCTGGAGAACCACGCCGAGACTCTGCGGGCTGCCATGA TCAGCCTCGCCAAGAAGATCGACGTGCAGACCGGAAGGCACCCCTATGAGGGGGGA

		GGAGGGAGCGGCGGCGGGAGCGGGGCCGCGGCCAGTATCCAGACCACCGTGA ACACTCTGTCTGAGCGTATCAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTC AGACCAAGTGTGATATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCA ACATCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGC TGAGCGCGGCCACGGAGACATATTCCGGTCTGACCCCTGAGCAGAAGGCGTATGTGC CTGCCATGTTACAGCTGCGCTGAATATCCAGACGAGTGTTAACACGGTGGTGCGGG ACTTTGAGAACTACGTGAAGCAGACGTGCAACAGCTCAGCGGTGGTGATAATAAGC TCAAGATCCAGAATGTGATCATTGACGAGTGTTATGGGGCCCCCTGGGAGCCCCACTA ACCTCGAGTTTATTAATACGGGCAGCAGCAAGGGCAACTGCGCTATCAAGGCGCTGA TGCAGCTGACAACGAAGGCCACCACCCAGATTGCTCCGCGGCAGGTGGCAGGAGGG GGAGGGGGGTCCGGCGGGGGAGGCTCGACTTGACGGTGCCGACGATGAATAACG CCAAGCTGACCTCCACCGAGACGTCGTTCAACGACAAGCAGAAGGTGACGTTACCT GCGACAGCGGCTATCACAGCTTGACCCCAACGCCGTCTGCGAGACGGACAAGTGG AAGTACGAGAACCCGTGCAAGAAGATGTGCACGGTGTCGACTATGTGTCCGAGCTG TACGATAAGCCGCTGTACGAGGTGAACAGCACCATGACGCTGTCTGTGCAACGGCGAG ACAAAGTACTTCCGGTTCGAGGAGAAGAACGGGAACACCAGCTGGAACGATACAGT GACCTGTCCAACGCAGAGTGCCAGCCTCTGCAGTTGGAGCACGGGTCTGTGACGCC CGTTAAGGAGAAGTACAGCTTTGGCGAGTACATGACGATTAAGTGCACGTCGGTTA CGAGGTGATCGGCGTGAGTTACATCTCATGTACCGCCAACAGCTGGAATGTCATACCT AGCTGCCAGCAGAAGGTGTGACATTCCGTCTCTCTCAACGGGCTGATCAGCGGGTCC ACGTTCTCTATCGGCGGCGTTATTCATCTGTCTCTGCAAGTCGGGCTTACCCTGACCG GGTCCCCCTCTTCACTTGTCATCGACGGCAAGTGGAACCCCATCTGCCACCTGCGT GCGGAGCAACGAGGAGTTCGACCCCGTGATGACGGGCTGATGATGAGACTGATC TCTCTAAGCTGTCTAAGGACGTGGTCCAGTACGAGCAGGAGATCGAGTCGCTGGAGG CCACCTACCACATCATCATATGGCCCTCACCATCATGGGGGTGATCTTCTTGATCAGC ATCATCGTGTGGTGTGTCTTTCGACAAGAACAACGACCAAGTACAAGTTTCACAAGC TGCTGCCTTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTG GGCCTCCCCCAGCCCTCTCCCTTCTGCAACCGTACCCCGTGGTCTTTGAATAA AGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AA AA
82	1401 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTTCGAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGAAGACCATCTCCGTGGTGACCCTGCTGTGCGTGCTGCCAGCAGT GGTGTACAGCACCTGTACGGTGCCGACGATGAATAACGCCAAGCTGACCTCCACCGA GACGTCGTTCAACGACAAGCAGAAGGTGACGTTACCTGCGACAGCGGCTATCACAG CTTGACCCCAACGCCGTCTGCGAGACGGACAAGTGGAAGTACGAGAACCCGTGCAA GAAGATGTGCACGGTGTCCGACTATGTGTCCGAGCTGTACGATAAGCCGCTGTACGA GGTGAACAGCACCATGACGCTGTCTGCAACGGCGAGACAAAGTACTTCCGGTGCGA GGAGAAGAACGGGAACACCAGCTGGAACGATACAGTGACCTGTCCCAACGCAGAGT GCCAGCCTCTGCAGTTGGAGCACGGGTCTGTGACCCGTTAAGGAGAAGTACAGCT TTGGCGAGTACATGACGATTAAGTGCACGTCGGTTACGAGGTGATCGGCGTGAGTT ACATCTCATGTACCGCCAACAGCTGGAATGTCATACCTAGCTGCCAGCAGAAGTGTGA CATTCCGTCTCTCTCAACGGGCTGATCAGCGGGTCCACGTTCTCTATCGGCGGCGTT ATTCATCTGTCTTGCAAGTCCGGCTTACGTTGACCGGGTCCGCTCGTGCACCTGCAT

		CGACGGGAAGTGGAATCCATTCTGCCACGTGCGTGCGGTGCGAAGGAGGAGTTGCA CCCGGTGACGATGGGCGGACGACGAGACAGATCTGTCAAAGTTGAGCAAGGATG TGGTTCAGTATGAGCAGGAGATCGAGAGCCTGGAGGCCACGTACCACGGGGGAGGG GGGAGCGGCGGCGGGAGCGGGCCGCGGCCAGTATCCAGACCACCGTGAACA CTCTGTCTGAGCGTATCAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGA CCAAGTGTGATATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCAACA TCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGA GCGCGGCCACGGAGACATATTCGGGTCTGACCCCTGAGCAGAAGGCGTATGTGCCTG CCATGTTACAGCTGCGCTGAATATCCAGACGAGTGTTAACACGGTGGTGCGGGACT TTGAGAACTACGTGAAGCAGACGTGCAACTCCAGCGCCGTGGTGGATAATAAGCTCA AGATCCAGAATGTGATCATTGACGAGTGTTATGGGGCCCTGGGAGCCCCACTAACC TCGAGTTTATTAATACCGGTTCTCCAAGGGAACTGCGCCATCAAGGCGCTGATGCA GTTGACAACGAAGGCGACGACCCAGATCGCGCCTCGGCAGGTGGCGGCGCGGGTG GCGGGTCAGGCGGCGGGGGGAGCGCGCCGTGAAGACGCCGGTGATCGTGGTGCC CGTGATCGACCGCCTCCGAGCGAGACGTTCCCCAACGTGCATGAGCATATCAACGA TCAGAAGTTCGATGACGTTAAGGACAACGAGGTGATGCAGGAGAAGCGGGACGTGG TGATCGTGAATGACGACCCGGATCACTACAAGGACTACGTGTTATCCAGTGACCG GAGGCAACATCCGGGATGACGACAAGTACACCCATTTTTCTCCGGCTTCTGTAACAC CATGTGCACAGAGGAGACGAAGCGGAACATCGCTCGCCACCTGGCACTGTGGGACTC CAAGTTCTTACAGAGCTGGAGAACAAGAATGTGGAGTACGTGTCATCATCGAGAA TGATAATGTGATTGAGGACATTACATTTCTCAGGCCGGTGCTCAAGGCCATCCATGAC AAGAAGATCGACATCCTGCAGATGCGGGAGATCATCACCGGTAATAAGGTGAAGACT GAGCTGGTCATTGACAAGGACCACGCCATCTTACCTATACCGGTGGCTACGACGTCT CGCTGAGCGCGTATATCATACGCGTGACCAACCGCTCTTAATATCGTCGACGAGATCAT TAAGAGCGGTGGTCTGTCTCGGGCTTCTACTTCGAGATAGCCCGGATCGAGAACGA GATGAAGATCAACCGCCAGATCATGGACAACCTAGCGAAGTACGTGAGCATGATCC CCGCTGGTGGCCGAGCACCAGTTTGAACAATGAAGCCCACTTCTGGTCCCGGAT AGGCACCGTGCCGCAAGCGGTATCCGGGCGTCATGTACACGTTACCACCCCGCT CATCTATTCTTCCGGCTGTTGACATTAACGTGATCGGCCTGATCGTGATACTGTTCA TCATGTTGATGCTGATCTTCAACGTGAAGAGCAAGTTGCTGTGGTTCCTGACAGGAAC CTTCGTGACTGCCTTCATTTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTT CTTGCCCCTTGGGCCTCCCCCAGCCCCCTCTCCCTTCTGACCCGTACCCCGTG TCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAAAAAA AAA AAA
83	1501 DNA	ATCCAAGCTTAGCTTGCTGTTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTGG CGCCGCCACCATGAAGACCATCTCGGTGGTGACACTGCTGTGCGTGTTGCCGGCGGT TGTTTATAGTACGTGCACCGTGCCGACTATGAACAACGCCAAGTTGACGTCTACTGAG ACTTCTTTCAATGACAAGCAGAAGGTGACGTTACCTGCGACAGCGGCTATCACAGCT TGGACCCCAACCGGTCTGCGAGACGGACAAGTGGAAGTACGAGAACCCGTGCAAG AAGATGTGCACGGGTGCGACTATGTGTCGAGCTGTACGATAAGCCGCTGTACGAG GTGAACAGCACCATGACGCTGTCCTGCAACGGCGAGACAAAGTACTTCCGTTGCGAG GAGAAGAACGGCAACACCAGCTGGAATGACACCGTGACCTGTCCGAACGCCGAGTG CCAGCCCCTGCAGTTGGAGCATGGCTCCTGCCAGCCCGTGAAGGAGAAGTACTCCTT

		CGGGGAGTACATGACTATCAACTGCGACGTGGGCTACGAGGTGATCGGGGTGTCTTA TATCAGCTGCACCGCCAATAGCTGGAACGTGATTCCCAGCTGTCAGCAGAAGTGTGA CATACCTTCTCTGTCCAACGGGCTGATCAGCGGGAGCACGTTTCAGCATTGGCGGGGT GATCCACCTGAGTTGCAAGTCTGGCTTCACGCTGACCGGATCTCCTAGCTCGACCTGC ATCGACGGGAAGTGGAACCCCATCCTGCCAATTGCGTCCGGTCCAACGAGGAGTTC GACCCCGTGGACGACGGACCAGACGACGAAACCGATCTGAGCAAGTTGAGCAAGGA TGTCGTGCAGTACGAGCAGGAGATCGAGTCCCTGGAGGCGACCTATCAGGTGGGG GAGGGAGCGGCGGCGGGCTCACGGCTGAATCAGTGCATGAGCGCCAACAAGGCC GCCATCACTGATTAGCCGTGGCCGTGCGCGCGCTTCTCCACCCACCGGAAGGTCG TCTCCAGCACCAACCCAGTATGATCACAAGGAGTCTGCAACGGGCTGTACTATCAGGG CAGTTGCTACATCCTGCACAGCGACTACAAGTCGTTTGAGGATGCTAAGGCCAAGTGC GCCGCCGAAAGCTCCACGCTGCCGAATAAGTCGGATGTCCTGACCACGTGGCTCATC GACTACGTGGAGACACCTGGGGCTCGGATGGGAACCCCATCAGCAAGACCACAG CGACTATCAGGACAGCGATGTGAGCCAGGAGGTGCGTAAGTACTTCTGCACCGGGG GCGGAGGGTCTGGAGGAGGCGGCAGCGGCGCTGCCGCTCCATCCAGACCACTGTG AACACCCTGTCCGAGCGCATTTCTGTCGAAGCTGGAGCAGGAGGCCAACGCCTCCGCC CAGACGAAGTGCGACATCGAGATAGGGAATTTCTACATCCGCCAGAACCACGGGTGC AACATCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACGCCGTG CTGAGCGCGGCCACGGAGACATATTCCGCTCTGACCCCGGAGCAGAAGGCTTACGTG CCTGCCATGTTTACGGCGGCCCTGAATATACAGACTTCTGTGAATACGGTGGTTCGGG GACTTCGAGAATTACGTGAAGCAGACTTGCAACTCCAGCGCGGTGGTGGATAATAAG CTCAAGATCCAGAATGTGATCATTGACGAGTGTATGGGGCCCTGGGAGCCCCACT AACCTCGAGTTTATTAATACCGTTCTCCAAGGGGAACTGCGCCATCAAGGCGCTGA TGCAGCTGACTACTAAGGCCACCAACCCAGATCGCCCCACGTCAGGTGGCTGGCACCG GCGTTCAGTTCTACATGATCGTGATCGGGGTGATCATTCTGGCTGCGCTGTTTCATGTA CTACGCCAAGCGCATGCTCTTTACTTCGACCAACGATAAGATCAAACCTTATCTTGGCG AATAAGGAGAATGTGATTGGACGACGTACATGGACACGTTCTCCGGACGTCTCCT ATGATCATTGCGACCACCGATATTCAGAACTGATAAGAATTCTAGGCTGGAGCCTCGG TGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCCCTCTCCCCCTTCTGCACCCGT ACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAA AA AA
84	1601 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTGG CGCCGCCACCATGGACTGGACATGGATTCTGTTCTGGTGGCCGCCGCCACCAAGGT GCACAGCGACGGCACCTGTTCCAGGCGACGACGACCTGGCCATCCCGCTACCGA GTTCTTCAGACCAAGGCTGCCAAGAACCCGAGACTAAGCGGGAGGCCATCGTGAA GGCCTACGGGGATGATAACGAGGAGACTCTGAAGCAGCGGCTGACCAACCTGGAGA AGAAGATTACCAACATTACCACCAAGTTCGAGCAGATCGAGAAGTGCTGCAAGCGGA ACGATGAGGTGCTGTTTCGGCTGGAGAACCACGCCGAGACACTGCGCGCCGCCATGA TCAGCCTGGCCAAGAAGATCGACGTGCAGACCGGCCGCCACCCCTACGAGGGGGGC GGCGGCTCTGGGGGCGGCGGGAGCAGCCCCAGACTTCAAGAAGATCGGGGATGA TGCCACCATCAGCTGCTCCCGGAACAACCAACTACTACGTGGTGATGTCCGCTGG TACAAGGAGCCTAACAGCATCATCCTGCTGGCCGCCAAGAGCGATGTGCTGTACTTG ACAATTACCAAGGATAAGATCAGCTACGACAGCCCTACGACGATCTGGTGACTAC

		CATCACCATCAAGAGCCTGACAGCCGGCGACGCAGGCACCTACATCTGCGCCTTCTTC ATGACCAGCACCACCAATGACACCGACAAGGTGGATTACGAGGAGTACTCCATCGAG CTGATCGTGAATACTGACTCAGAGTCTACCATCGACATTATCCTGTCCGGGAGCACCC CCGAGACTATCAGTGAGAAGCCTGAGGACATCGACAACAGCAACTGCAGCAGTGTGT TCGAGATCACCACCCAGAGCCTATCACTGACAACGTGGATGATCACACCGACACTGT GACATACACCAGTGATAGCATCAACACCGTGAATGCTAGCTCTGGGGAGAGCACCAC TGATGAGATCCCCGAGCCTATCACCAGACAAGGAGGAGGATCACACCGTGACCGACAC TGTGAGTTACACCACAGTGTCTACCTCCTCTGGGATCGTGACCACCAAGAGCACCACC GACGATGCCGACCTGTACGATACATACAATGATAACGATACAGTGCCTCCCACCACCG TGGGAGGCAGCACCACAGCATTTCTAATTACAAGACCAAGGACTTCGTGGGGGGCG GGGGGAGCGGGGGCGGGGGGAGCGCCGAAACCAAGGAGTTCAAGACCCTGTACAA CCTGTTCATCGATAGCTACCTGCAGAAGCTGGCCAGCACAGCATTCCCACCAATGTG ACATGCGCCATTACATTGGGGAGGTGATCGGCCAGTTCAAGAACTGCGCCTGAGG ATCACCATAAGTGCATGAGCAATAGCCGCTGTCTTCACTCTGATGGTGGAGTCTT TCATCGAGGTGATCTCCTGCTGTGAGAGAAGGACAGGCGGGCTATTGCCGAGGAGA TCGGGATCGACCTGGACGACGTGCCAGCGTGGTGTCCAAGCTGGAGAAGAACTGC AACGCCTACGCCGAGGTGAACAACATCATCGATATTCAGAAGCTGAATATCGGTGAG TGTTCCGCCCTCCCGGCCAGCACATGCTGCTGCAGATCGTGAATACAGGGAGCGCT GAGGCCAACTGCGGGCTCCAGACCATCGTGAAGAGCCTGAATAAGATTTACGTGCCT CCTATCATCGAGAATCGGCTGCCCTACTACGACCCCTGGTTCCTGGTGGGCGTGGCCA TTATCCTGGTGATCTTCACCGTGGCCATCTGTAGCATCAGGCGGAACCTGGCCCTGAA GTACAGGTACGGGACCTTCTGTACGTGTGATAAGAATTCTAGGCTGGAGCCTCGGT GGCCATGCTTCTTGCCCTTGGGCTCCCCCAGCCCTCCTCCCTTCTGCACCCGT ACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAA AA AA
85	1701 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTGG CGCCGCCACCATGACCCAGCTGCCATCCTGCTGCTGCTGATCAGCCTGGTGTACGCC ACTCCCAGTCTCAGACCTCCAAGAAGATTGGCGACGATGCCACCATCAGCTGCTCCC GGAATAATACCAACTACTATGTGGTGATGAGCGCCTGGTACAAGGAGCCTAACAGCA TCATCCTGCTGGCCGCAAGAGCGATGTGCTGTACTTCGATAACTACACCAAGGACAA GATCAGCTACGATAGCCCTACGACGATCTGGTGACTACCATCACCATCAAGAGTCTG ACTGCCGGCGACGCCGGGACTTACATCTGCGCCTTCTTCATGACCAGCACCACCAATG ACACCGACAAGGTGGACTACGAGGAGTACAGCATCGAGCTGATCGTGAACACCGAC AGCGAGAGCACCATCGACATCATCCTGTCTGGCAGCACCCCGAGACTATCAGCGAG AAGCCTGAGGACATCGATAACAGCAACTGTTCCAGCGTTTTTCGAGATCACCACCCCG AGCCTATCACCATAACGTGGACGATCACACCGACACCGTGACCTACACCTCCGATAG CATCAACACCGTGAATGCCAGCAGCGGGGAGAGTACTACAGATGAGATCCCCGAGCC TATCACCATAAGGAGGAGGATCACACCGTGACCGACACTGTGAGTTACACCACAGT GTCTACCTCCTCCGCATCGTGACTACCAAGAGTACCACCGACGATGCCGATCTGTAC GATACCTACAACGATAACGACACAGTGCCTCCCACCACCGTGGGAGGCAGCACCACC AGCATCAGCAACTACAAGACCAAGGATTTGTGGGCGGCGGGGGCTCCGGCGGCGG AGGCTCCGCCCGCGTGAAGACCCCGTGATTGTGGTGCCTGTGATCGATCGCCCCCT AGCGAGACATTTCCCAACGTGCACGAGCATCAACGATCAGAAGTTCGATGACGTG

		AAGGACAACGAGGTGATGCAGGAGAAGCGGGACGTGGTGATCGTGAACGACGATCC TGACCACTACAAGGACTACGTGTTTCATCCAGTGGAACCGGGGGCAACATCCGGGATGA CGACAAGTACACCCACTTCTTCTCCGGCTTCTGTAACACCATGTGTACAGAGGAGACA AAGCGGAACATTGCTCGGCACCTGGCCCTGTGGGACAGCAAGTTCTTCACTGAGCTG GAGAACAAGAACGTGGAGTATGTGGTGATCATCGAGAACGACAATGTGATCGAGGA CATTACCTTCTCAGGCCTGTGCTGAAGGCCATCCACGATAAGAAGATCGACATCCTG CAGATGCGGGAGATCATCACCGGGAATAAGGTGAAGACCGAGCTGGTGATCGACAA GGATCACGCCATCTTACCTACACCGGCGGCTATGATGTGAGCCTGAGCGCTTACATC ATCCGGGTGACCACCGCCCTGAATATCGTGGATGAGATCATTAAAGTCCGGGGGGCTG TCCAGCGGGTTCTACTTCGAGATTGCTAGGATCGAGAACGAGATGAAGATCAACCGG CAGATCATGGATAATTCTGCCAAGTACGTGGAGCACGACCCCTCGGCTGGTGGCTGAG CATCGCTTCGAGACTATGAAGCCCACTTCTGGAGCAGGATTGGGACCGTGGCCGCC AAGCGGTACCCAGGGGTGATGTACACCTTACCACACCCCTGATCAGCGGCGGCGGG GGCAGCGGCGGGGCGGCAGCGCTGAGACAAAGGAGTTCAAGACCTGTACAACCT GTTTATCGACAGCTATCTGCAGAAGCTGGCTCAGCACTCTATCCCAACCAATGTGACA TGCGCCATTACATTGGGGAGGTGATCGGCCAGTTCAAGAACTGCGCCCTGAGGATC ACCAATAAGTGCATGAGCAACAGCAGGCTGAGCTTCACCCTGATGGTGGAGTCTTC ATCGAGGTGATCAGCCTGCTGAGCGAGAAGGACAGGAGGGCCATCGCCGAGGAGAT CGGCATCGACCTGGACGATGTGCCAGCGTGGTGAGCAAGCTGGAGAAGAAATTGTA ACGCCTACGCCGAGGTGAACAATATCATCGATATTCAGAAGCTGAATATCGGGGAGT GCAGTGCTCCCCCTGGGCAGCACATGCTGCTGCAGATTGTGAACACCGGCAGCGCCG AGGCCAATTGCGGGCTGCAGACCATTGTGAAGTCACTGAATAAGATCTACGTGCCAC CCATCATCGAGAACCGGCTGCCATACTACGATCCCTGGTTCCTGGTGGGGGTGGCCAT TATTCTGGTGATCTTCACAGTGGCCATCTGCAGCATCAGGCGCAACCTGGCTCTGAAG TACAGGTACGGGACCTTCTGTACGTGTGATAAGAATTCTAGGCTGGAGCCTCGGTG GCCATGCTTCTTGCCCTTGGGCTCCCCCAGCCCTCCTCCCCTTCTGCACCCGTAC CCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAAAAAAA AA AA
86	1801 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGACCCAGCTGCCTATCCTGCTGCTGCTGATCAGCCTGGTGTACGCT ACCCCTAGTCCCCAGACTTCCAAGAAGATCGGGGACGACGCCACGATCAGCTGCAGC AGGAATAACACCAACTACTACGTGGTGATGAGCGCCTGGTACAAGGAGCCAAATAGC ATCATTCTGCTGGCCGCAAGTCTGACGTGCTGTACTTCGACAATTACACAAAGGACA AGATCTCCTACGACAGTCCTTATGATGATCTGGTGACCACGATCACCATTAAAGTCTCTG ACTGCTGGGGATGCCGTACCTATATCTGTGCGTTCTTATGACCTCTACAACGAACG ACACAGATAAGGTGGACTACGAGGAGTATAGCATCGAGCTGATCGTGAACACCGACA GCGAGAGCACAATTGACATTATCCTCAGCGGGAGCACCCCGAGACAATCTCTGAGA AGCCCGAGGACATTGATAATAGCAATTGTAGCTCCGTGTTGAGATTACTACTCCCGA GCCATTACCGATAATGTGGATGATCATAAGATACGGTGACCTACACCAGCGACAG CATCAACACCGTGAATGCTAGCAGCGGTGAGAGCACCAACCGATGAGATCCCGGAGCC TATCACCGATAAGGAGGAGGATCACACCGTGACCGACACGGTGAGTTACACCACCGT GTCTACCTCCTCCGGGATCGTGACTACCAAGAGCACCACGGATGACGCCGATCTGTAT GACACCTACAACGATAATGATACCGTGCCGCCACCACCGTGGGCGGCAGCACCAACC

		AGCATCAGCAACTATAAGACTAAGGACTTCGTGGGGGGCGGGGGGAGCGGCGGCG GCGGCAGCCGGCTGAATCAGTGCATGAGCGCCAACAAGGCCGCCATCACTGATTCAG CCGTGGCTGTCGCCGCCGCCAGTCCACCCACCGGAAGGTGGTGTCCAGCACCACCC AGTACGATCACAAGGAGTCCTGCAATGGCCTGTATTACCAGGGCTCTTGCTACATCCT GCACAGCGATTACAAGAGCTTTGAGGATGCCAAGGCCAATTGTGCAGCCGAGAGCA GCACGCTGCCTAACAAGAGCGACGTGCTGACTACTTGGCTGATCGACTATGTGGAGG ACACCTGGGGCAGTGATGGGAACCCCATCACTAAGACCACCTCTGACTACCAGGACT CCGACGTGAGTCAGGAGGTGAGGAAGTACTTCTGTACGGGCGGCGGCGGCAGCGGC GGCGGCGGGTCCACCTGCACCGTGCCACCATGAACAACGCCAAGCTGACTTCTACC GAGACTTCTTTCAACGACAAGCAGAAGGTGACCTTCACCTGCGACAGCGGCTATCATT CCCTCGACCCTAACGCCGTGTGCGAGACAGACAAGTGGAAGTATGAGAACCCGTGCA AGAAGATGTGTACGGTGTCCGACTACGTCTCTGAGCTGTACGACAAGCCTCTGTACG AGGTCAACTCCACTATGACTCTGTCTTGCAACGGCGAGACAAAGTACTTTAGGTGCGA GGAGAAGAACGGCAATACCAGCTGGAACGACACCGTGACGTGCCCCAACGCCGAGT GTCAGCCTCTCCAGCTGGAGCACGGCAGCTGTGAGCCTGTGAAGGAGAAGTACAGCT TCGGCGAGTACATGACCATCAACTGTGATGTGGGGTACGAGGTGATCGGGGTGAGC TATATCAGCTGCACCGCCAATAGCTGGAACGTGATTCCCAGCTGTGAGCAGAAGTGC GACATCCCTTCTCTGTCCAACGGGCTGATCAGCGGGAGCACGTTTACGATTGGCGGG GTGATCCACCTGAGCTGCAAGTCAGGGTTCACCTGACTGGCTCACCCAGCAGCACCT GCATCGATGGCAAGTGGAACCCCATTTCTGCCACCTGTGTGAGAAGCAACGAGGAGT TCGATCCTGTGGACGACGGACCAGACGACGAAACCGATCTGTCCAAGCTGTCCAAGG ATGTGGTGAGTACGAGCAGGAGATCGAGAGCCTGGAGGCCACCTACCACATCATCA TCATGGCACTGACCATCATGGGCGTGATATTCTTGATCAGCATCATTGTGCTGGTGTG TTCTTGCGACAAGAACAACGACCAGTACAAGTTTACAAGCTGCTGCCCTGATAAGAA TTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCTCCCCCAGCCCCT CCTCCCCTTCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCA CTAGTAAA AA AAAAAAAAAAAAAAAA
87	1901 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGACTGGACCTGGATTCTGTTCTGGTGGCGGCCGCCACAAGGT GCACTCCGATGGGACCTTGTTCCTGGGGACGACGACCTGGCCATCCCCGCTACGGA GTTCTTCAGCACCAAGGTGCCAAGAACCCGGAGACTAAGCGGGAGGCCATCGTCAA GGCCTACGGAGATGACAATGAGGAGACGCTGAAGCAGCGTCTACCAACCTTGAGA AGAAGATCACCAACATCACTACGAAGTTTGAGCAGATTGAGAAGTGCTGCAAGCGCA ATGATGAGGTGCTGTTTCGGCTGGAGAACCACGCCGAGACTCTGCGGGCTGCCATGA TCAGCCTGCCAAGAAGATCGACGTGCAGACCGGAAGGCACCCCTATGAGGGGGGA GGAGGGAGCGGCGGCGGGGAGCGGGGCCGCGGCCAGTATCCAGACCACCGTGGA ACACTCTGTCTGAGCGTATCAGCAGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTC AGACCAAGTGTGATATAGAGATTGGCAATTTCTATATCAGGCAGAACCACGGGTGCA ACATCACCGTGAAGAATATGTGCTCCGCTGACGCCGACGCTCAGCTGGACCGGTGC TGAGCGCGGCCACGGAGACATATTCCGGTCTGACCCCTGAGCAGAAGGCGTATGTGC CTGCCATGTTACAGCTGCGCTGAATATCCAGACGAGTGTTAACACGGTGGTGCGGG ACTTTGAGAACTACGTGAAGCAGACGTGCAACTCCAGCGCCGTGGTGATAATAAGC

		TCAAGATCCAGAATGTGATCATTGACGAGTGTTATGGGGCCCCCTGGGAGCCCCACTA ACCTCGAGTTTATTAATACCGGTTCTCCAAGGGGAAGTGCGCCATCAAGGCGCTGAT GCAGTTGACAACGAAGGCCACCACCCAGATCGCCCCAGGCAGGTTGCCGGGGGCG GCGGGGGGAGCGGTGGCGGGCGCAGCGCTGAGACAAAGGAGTTCAAGACCTGTA CAACTTGTTCATAGATTCTTATCTTCAGAAGTTGGCCAGCACTCCATCCCTACCAATG TCACTTGTGCCATCCACATTGGAGAGGTGATCGGCCAGTTCAGAAGTGCGCCCTGA GGATCACCAACAAGTGCATGAGCAATAGCCGCCTGTCTTCACGCTCATGGTGGAGT CCTTCATCGAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCGGGCTATTGCTGAGG AGATCGGTATTGACCTCGACGACGTGCCGTCCGTCTCTCAAGCTGGAGAAGAACT GCAACGCATATGCTGAGGTGAACAACATCATCGATTCAGAAGCTGAATATCGGTG AGTGTTCCGCTCCCCCTGGGCAGCATATGCTGTTGCAGATTGTCAATACCGGGAGTGC TGAGGCCAACTGCGGGCTGCAGACCATCGTGAAGTCACTGAATAAGATCTACGTGCC ACCCATCATCGAGAACAGGCTGCCGTACTACGACCCGTGGTCTTGGTGGGGGTGGC CATTATTCTGGTGATCTTCACGGTGGCCATCTGCAGCATCCGCCGAACCTGGCTCTG AAGTACCGGTACGGGACCTTCTGTACGTGTGATAAGAATTCTAGGCTGGAGCCTCG GTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCTCCTCCCTTCTGCACCC GTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCACTAGTAAAAAAAAAAAA AA AA
88	2001 DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACCATGGATTGGACCTGGATTCTGTTCTCTGTTGCCGCCGCCACAGGGTG CATAGTGCGGCAGTGAAGACTCCCGTTATCGTCGTCCCGGTGATTGACCGCCCGCCT CGGAGACGTTCCCAACGTGCATGAGCACATCAACGATCAGAAGTTCGACGATGTCA AGGACAATGAGGTCATGCAGGAGAAGCGGGACGTGGTGATCGTCAACGACGATCCC GATCACTACAAGGATTACGTCTTCATCCAGTGGACGGGCGGCAACATCCGGGACGAC GATAAGTACACGATTTCTTCTCCGGCTTCTGTAACACCATGTGTACAGAGGAGACGA AGCGGAACATCGCCAGGCATCTGGCGCTGTGGGACTCGAAGTTCTTACCGAGCTGG AGAATAAGAACGTGGAGTACGTGGTGATCATCGAGAACGACAACGTATCGAGGAC ATCACGTTTCTGCGGCCGGTGCTGAAGGCCATCCACGACAAGAAGATCGACATCCTG CAGATGCGGGAGATCATCACGGGTAATAAGGTGAAGACCGAGCTGGTCATTGACAA GGACCACGCGATTTACCTACACTGGGGGATACGACGTGCTTGTCCGCCTATATT ATCCGGGTGACCACCGCTCTGAACATCGTGGATGAGATCATCAAGTCGGGAGGCCTC AGCTCCGGCTTCTACTTCGAGATCGCCCGGATCGAGAACGAGATGAAGATCAACCGC CAGATCATGGACAACAGCGCCAAGTATGTCGAGCACGACCTCGGCTGGTGGCAGAG CACAGATTCGAGACTATGAAGCCCACTTCTGGAGTAGGATCGGCACTGTCGCCGCC AAGCGCTACCCAGGTGTGATGTACACCTTACCACCCCGCTGATCAGCGGGGGGGGG GGTAGCGGCGGCGGCGGCAGTGCCGAGACAAAGGAGTTCAAGACCCTGTACAACTT GTTTCATCGACTCCTACCTCCAGAAGTTGGCTCAGCACTCGATCCCCACCAATGTGACAT GCGCCATTACATTGGTGAGGTGATCGGGCAGTTCAAGAACTGTGCTCTGCGGATTA CCAACAAGTGTATGAGTAATTCCCGCTGTCTTCACGCTCATGGTGGAGTCTTCATC GAGGTGATCTCCCTGCTGAGCGAGAAGGACAGGCGGGCCATCGCCGAGGAGATCGG CATCGACCTGGACGATGTTCCGTCTGTCGTCTCAAGCTGGAGAAGAAGTGAACGC ATATGCTGAGGTGAACAACATCATCGATTCAGAAGCTGAATATCGGTGAGTGTTCC GCTCCCCCTGGGCAGCATATGCTGTTGCAGATCGTGAATACCGGGAGTGCAGGAGCG

		AACTGCGGGCTCCAGACCATCGTGAAGAGCCTGAACAAGATCTACGTCCCGCCTATCA TCGAGAATCGGCTGCCGTACTATGACCCTGGTGGCGGCGGCAGCGGGGGGGGGGG GTCGGGGGCCGCTGCTTCCATCCAGACCACCGTGAACACTCTGTCTGAGCGTATCAGC AGCAAGCTGGAGCAGGAGGCCAACGCCTCTGCTCAGACCAAGTGTGATATAGAGATT GGCAATTTCTATATCAGGCAGAACCACGGGTGCAACATCACCGTGAAGAATATGTGC TCCGCTGACGCCGACGCTCAGCTGGACGCCGTGCTGAGCGCGGCCACGGAGACATAT TCCGGTCTGACCCCTGAGCAGAAGGCGTATGTGCCTGCCATGTTACAGCTGCGCTGA ATATCCAGACGAGTGTTAACACGGTGGTGC GGGATTTCGAGAATTATGTGAAGCAGA CCTGTAACAGCTCGGCTGTTGTGGACAACAAGCTCAAGATCCAGAATGTGATCATTGA CGAGTGTTACGGGGCCCCGGCTCCCCACGAATCTGGAGTTCATCAACACGGGCAG CAGCAAGGGCAACTGCGCCATTAAGGCGCTGATGCAGCTGACCACTAAGGCCACAAC CCAGATCGCCCTAGGCAGGTGGCCGGGACTGGCGTGCACTTCTATATGATAGTGAT TGGCGTGATCATCCTGGCCGCCCTGTTCTGTACTACGCCAAGCGCATGCTCTTACTT CGACCAACGATAAGATCAAATTATCTTGGCGAATAAGGAGAATGTGCATTGGACGA CGTACATGGACACCTTCTTTCGCACATCACCTATGATCATTGCCACGACTGACATTAG AACTGATAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTCCCCCTTGGGCCT CCCCCAGCCCCTCTCCCTTCTGACCCGTACCCCGTGGTCTTTGAATAAAGTCT GAGTGGGCGGCACTAGTAAA AA AAAAAAAAAAAAAAAAAAAAAAAAAAAA
89	0201 编码的氨基酸序列	MDWTWILFLVAAATRVHSDGTLFPGDDDLAIPATEFFSTKAANKPETKREAIVKAYGDDN EETLKQRLTNLEKKITNITTKFEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGR HPYYPYDVPDYAE
90	0302 编码的氨基酸序列	MMTPENDEEQTSVFATVYGDKIQGKNKRKRVIGLCIRISMVISLLSMITMSAFLIVRLNQ CMSANKAAITDSAVAVAAASSTHRKVVSSTTQYDHKESCNGLYYQGSYILHSDYKSFED AKANCAAESSTLPNKSDVLTWLDYVEDTWGSDGNPITKTTSDYQSDSVSQEVRYFCT YPYDVPDYA
91	0402 编码的氨基酸序列	MTQLPILLLISLVYATPSPQTSKKIGDDATISCSRNTNYYVMSAWYKEPNSIILAAKSD VLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTYICAFMTSTTNDTKVDYEEYSIELI VNTDSESTIDIILSGSTPETISEKPEDIDNSNCSSVFEITPEPITDNVDDHTDVTYTSDSINT VNASSGESTTDEIPEPITDKEEDHTVTDVSYTTVSTSSGIVTKSTDDADLYDTYNDNDT VPPTTVGGSTTSISNYKTKDFEIFGITTILLSAVAIFCITYICNKHPRKYKTENKVYPYDVP DYA
92	0505 编码的氨基酸序列	MKTISVVTLLCVLPAVVYSGGGGSYPYDVPDYATCTVPTMNNAKLTSTETSFNDKQKVTF TCDSGYHSLDPNAV CETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGET KYFRCEEKNGNTSWNDT VTC PNAECQPLQLEHGSCQPVKEKYSFGEYMTINCDVGYEVI GVSYSCTANSWNVIPSCQKCDIPSLSNGLISGSTFSIGGVIHLSCSGFTLTGSPSSTCIDG KWNPIPTCVRSNEEFDPVDDGPDDDETDL SKLSKDVVQYEQEIESLEATYHIIIMALTIMGV IFLISIIVLVCSCKNNDQYKFHKLLP
93	0603 编码的氨基酸序列	MDWTWILFLVAAATRVHSGGGGSYPYDVPDYAAETKEFKTLYNLFIDSYLQKLAQHSIPT NVTCAIHIGEVIQFKNCALRITNKCMSNSRSLFTLMVESFIEVILLSEKDRRAIAEEIGIDL DDVPSVSKLEKNCNAYAEVNNIIDIQKLNIGEC SAPPGQHMLLQIVNTGSAEANCGLQTI VKSLNKIYVPIIENRLPYDPWFLVGVAIILVIFTVAICSI RRNLALKYRYGTFLYV
94	0702 编码的氨基酸序列	MDWTWILFLVAAATRVHSGGGGSYPYDVPDYAGAAASIQTTVNTLSERISSKLEQEANA

	氨基酸序列	SAQTKCDIEIGNFYIRQNHGCNITVKNMCSADADAQLDAVLSAATETYSGLTPEQKAYVP AMFTAALNIQTSVNTVVRDFENYVKQTCNSSAVVDNKLKIQNVIIIDECYGAPGSPTNLEFI NTGSSKGNCAIKALMQLTTKATTQIAPRQVAG
95	0803 编码的氨基酸序列	MDWTWILFLVAAATRVHSGGGGSYPYDVPDYAAAVKTPVIVVPVIDRPPSETFPNVHEHI NDQKFDDVKDNEVMQEKRDDVIVNDPDDHYKDYVFIQWTGGNIRDDDKYTHFFSGFC NTMCTEETKRNIARHLALWDSKFFTELENKNVEYVVIENDNVIEDITFLRPVLKAIHDKKID ILQMREIITGNKVKTTELVIDKDHAIFTYTGgyDVLSAYIIRVTALNIVDEIISKGGLSGGFYF EIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAAKRYPGVMY TFTTPLISFFGLFDINVIGLIVLIFMFMILFNVKSKLLWFLTGTFTVAFI
96	0201-J0 编码的氨基酸序列	MDWTWILFLVAAATRVHSM DGTLP GDDDLAIPATEFFSTKAAKNPETKREAI VKAYGD DNEETLKQRLTNLEKKITNITTKFEQIEKCKRNDEVLFRL ENHAETLRAAMISLAKKIDVQT GRHPYE
97	0301-J0 编码的氨基酸序列	MDWTWILFLVAAATRVHSMMPENDEEQTSVFSATVYGDKIQGKNKRKRIVGLCIRISM VISLLSMITMSAFLIVRLNQCMSANKAAITDSAVAVAAASSTHRKVVSSTTQYDHKESCN LYYQGSCYILHSDYKSFEDAKANCAAESSTLPNKSDVLTTLWIDYVEDTWGSDGNPITKTT SDYQSDVSQEVRYKFC
98	0501-J0 编码的氨基酸序列	MDWTWILFLVAAATRVHSMKTSISVVTLLCVLPAVVYSTCTVPTMNNAKLTSTETSFNDK QKVFTFTCDSGYHSLDPNAV CETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLS CNGETKYFRCEEKNGNTSWNDTVTCPNAECQPLQLEHGSCQPVKEKYSFGGYMTINCDV GYEVIGVSYISCTANSWNVIPSCQKCDIPSLNGLISGSTFSIGGVHLSCKSGFTLTGSPSS TCIDGKWNPIPTCVRSNEEFDPVDDGPDDDETDL SKLSKDVVQYEQEIESLEATYHIIIMAL TIMGVIFLISIIVLCSCDKNNDQYKFHKLLP
99	0701-J0 编码的氨基酸序列	MDWTWILFLVAAATRVHSMGAAASIQTTVNTLSERISSKLEQEANASQTKCDIEIGNFYI RQNHGCNITVKNMCSADADAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVN TVVRDFENYVKQTCNSSAVVDNKLKIQNVIIIDECYGAPGSPTNLEFI NTGSSKGNCAIKAL MQLTTKATTQIAPRQVAGTG VQFYMIVIGVILALFMYAKRMLFTSTNDKIKLILANKE NVHWTTYMDTFFRTSPMIATTDIQN
100	0901-J0 编码的氨基酸序列	MDWTWILFLVAAATRVHSM PQQLSPINIETKKAISDARLKTLDIHYNESKPTTIQNTGKLV RINFKGGYISGGFLPNEYVLSTIH YWGKEDDYGSNHLIDVYKYSGEINLVHWNKKKYSSYE EAKKHDDGIIIIAIFLQVSDHKNVYFQKIVNQLDSIRSANMSAPFDSVFYLDNLLPSTLDYFT YLGTTINHSADAAWIFPTPINIHS DQLSKFRLLSSSNHEGKPHYITENYRNPYKLNDDTQ VYYSGEIIRAATTSPVRENYFMKWLSDLREACFSYYQKYIEGKNTFAIIAIVFVFILTAIFLMS QRY SREKQN
101	1101 编码的氨基酸序列	MDWTWILFLVAAATRVHSDGTLP GDDDLAIPATEFFSTKAAKNPETKREAI VKAYGDDN EETLKQRLTNLEKKITNITTKFEQIEKCKRNDEVLFRL ENHAETLRAAMISLAKKIDVQTGR HPYEGGGGSGGGGSRLNQCMSANKAAITDSAVAVAAASSTHRKVVSSTTQYDHKESCN GLYYQGSCYILHSDYKSFEDAKANCAAESSTLPNKSDVLTTLWIDYVEDTWGSDGNPITKT TSDYQSDVSQEVRYKFC TGGGGSGGGGSTCTVPTMNNAKLTSTETSFNDKQKVFTFTCD SGYHSLDPNAV CETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFR CEEKNGNTSWNDTVTCPNAECQPLQLEHGSCQPVKEKYSFGGYMTINCDVGYEVIGVSYI SCTANSWNVIPSCQKCDIPSLNGLISGSTFSIGGVHLSCKSGFTLTGSPSSTCIDGKWN PIPTCVRSNEEFDPVDDGPDDDETDL SKLSKDVVQYEQEIESLEATYHGGGGSGGGGSSP QTSKKIGDDATISCSRNTNYYVVM SAWYKEPNSIILAAKSDVLYFDNYTKDKISYDSPYD DLVTTITIKSLTAGDAGTYICAFMTSTTNDTKVDYEEYSIELIVNTDSESTIDIILSGSTPETI

		SEKPEDIDNSNCSSVFEITTPPEITDNVDDHTDVTYTSDSINTVNASSGESTTDEIPEPITDK EEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDVPPTTVGGSTTSISNYKTKD FVGGGGSGGGGAETKEFKTLYNLFIDSYLQKLAQHSIPTNVTCIAIHIGEVIQFKNCALRI TNKCMSNSRSLFTLMVESFIEVISLLSEKDRRAIAEEIGIDLDDVPSVSVSKLEKNCNAYAEVN NIIDIQKLNIGECSPAPGQHMLLQIVNTGSAEANCGLQTIVKSLNKIYVPIIENRPLPYDPG GGGSGGGSGAAASIQTTVNTLSERISSKLEQEANASATKCDIEIGNFYIRQNHGCNITV KNMCSADADAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTVVRDFENYV KQTCNSSAVVDNKLKIQNVIIDECYAGPSPTNLEFINTGSSKGNCALMQLTTKATTQI APRQVAGTGVQFYMIVIGVILALFMYAKRMLFTSTNDKIKLILANKENVHWTYMDT FFRTSPMIIATTDIQN
102	1201 编码的氨基酸序列	MTQLPILLLLISLVYATPSPQTSKIGDDATISCSRNNNTNYYVMSAWYKEPNSIILAAKSD VLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTYICAFMTSTTNDTKVDYEEYSIELI VNTDSESTIDIILSGSTPETISEKPEDIDNSNCSSVFEITTPPEITDNVDDHTDVTYTSDSINT VNASSGESTTDEIPEPITDKEEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDT VPPTTVGGSTTSISNYKTKDFVGGGGSGGGGSRNQCMSANKAAITDSAVAVAAASSTH RKVVSSTTQYDHKESCNGLYYQGSYILHSDYKSFEDAKANCAEESTLPNKSVDLTTWLI DYVEDTWGSDGNPITKTSYQSDVSQEVRYFCTGGGGSGGGGAETKEFKTLYNLFI DSYLQKLAQHSIPTNVTCIAIHIGEVIQFKNCALRITNKCMSNSRSLFTLMVESFIEVISLLSE KDRRAIAEEIGIDLDDVPSVSVSKLEKNCNAYAEVNNIIDIQKLNIGECSPAPGQHMLLQIVN TGSAEANCGLQTIVKSLNKIYVPIIENRPLPYDPWFLVGVAIILVIFTVAICSIRRNALKYRY GTFLYV
103	1301 编码的氨基酸序列	MDWTWILFLVAAATRVHSDGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKAYGDDN EETLKQRLTNLEKKITNITTKFEQIEKCKRNDVLFRLNHAETLRAAMISLAKKIDVQGR HPYEGGGSGGGSGAAASIQTTVNTLSERISSKLEQEANASATKCDIEIGNFYIRQNHG CNITVKNMCSADADAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTVVRD FENYVKQTCNSSAVVDNKLKIQNVIIDECYAGPSPTNLEFINTGSSKGNCALMQLTT KATTQIAPRQVAGGGGGSGGGGSTCTVPTMNNAKLTSTETSFNDKQKVTFTCDSGYHSL DPNAVCECDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKN GNTSWNDVTCNAECQPLQLEHGSCQPVKEKYSFGEYMTINCDVGVEVIGVSYISCTAN SWNVIPSCQKCDIPSLNGLISGSTFSIGGVHLSCKSGFTLTGSPSSTCIDGKWNPIPTC VRSNEEFDPVDDGPDDETDLKLSKDVVQYEQEIESLEATYHIIIMALTIMGVIFLISIVLVC SCDKNNDQYKFHKLLP
104	1401 编码的氨基酸序列	MKTISVVTLLCVLPAVVYSTCTVPTMNNAKLTSTETSFNDKQKVTFTCDSGYHSLDPNAV ETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWN DVTCPNAECQPLQLEHGSCQPVKEKYSFGEYMTINCDVGVEVIGVSYISCTANSWNVIPS CQKCDIPSLNGLISGSTFSIGGVHLSCKSGFTLTGSPSSTCIDGKWNPIPTCVRSEEF DPVDDGPDDETDLKLSKDVVQYEQEIESLEATYHGGGGSGGGSGAAASIQTTVNTLSE RISSKLEQEANASATKCDIEIGNFYIRQNHGCNITVKNMCSADADAQLDAVLSAATETYS GLTPEQKAYVPAMFTAALNIQTSVNTVVRDFENYVKQTCNSSAVVDNKLKIQNVIIDECY GAPGSPTNLEFINTGSSKGNCALMQLTTKATTQIAPRQVAGGGGGSGGGGSAAVKT PVIVVPVIDRPPSETFPNVHEHINDQKFDVVDNEVMQEKRDVVVNDPDHYKDYVFI QWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARHLALWDSKFFTELENKNVEYVVIIE NDNVIEDITFLRPVLKAIHDKKIDILQMRIITGNKVTELVIDKDHAIFTYTGGYDVSLSAYII RVTTALNIVDEIISKGLSSGFYFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRFETM

		KPNFWSRIGTVAAKRYPGVMYFTFTPLISFFGLFDINVIGLIVLFIMFMLIFNVKSKLLWFLT GTFVTAFI
105	1501 编码的氨基酸序列	MKTISVVTLLCVLPVVYSTCTVPTMNNAKLTSTETSFNDKQKVFTFTCDSGYHSLDPNAVCE ETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWN DTVTCPNAECQPLQLEHGSCQPVKEKYSFGYMTINCDVGVEVIGVSYISCTANSWNVIPS CQQKCDIPSLSNGLISGSTFSIGGVIHLSCSGFTLTGSPSSTCIDGKWNPIPTCVRSNEEF DPVDDGPDDETDLKSLSKDVVQYEQIESLEATYHGGGGSGGGGSRLNQCMSANKAAIT DSAVAVAAAASSTHRKVVSSTTQYDHKESCNGLYYQGSCYILHSDYKSFEDAKANCAAESST LPNKSDVLTWLDYVEDTWGSDGNPITKTTSDYQSDVSQEVRYKFTGGGGSGGGGS GAAASIQTTVNTLSERISSKLEQEANASATKCDIEIGNFYIRQNHGCNITVKNMCSADAD AQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTTVRDFENYVKQTCNSSAVV DNKLIKQNVIIDECYAGPSPTNLEFINTGSSKGNCAIKALMQLTTKATTQIAPRQVAGTG VQFYMIVIGVILAAALFMYAKRMLFTSTNDKIKLILANKENVHWTYMDTFFRTSPMIAT TDIQN
106	1601 编码的氨基酸序列	MDWTWILFLVAAATRVHSDGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKVAYGDDN EETLKQRLTNLEKKITNITTKFEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGR HPYEGGGGGSGGGSSPQTSKKIGDDATISCSRNNNTNYYVMSAWYKEPNSIILAAKSDV LYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTYICAFFMTSTTNDTDKVDYEEYSIELIV NTDSESTIDIILSGSTPETISEKPEDIDNSNCSSVFEITTPITDNVDDHTDTVTYTSDSINTV NASSGESTTDEIPEPITDKEDHTVTDTVSYTTVSTSSGIVTTKSTTDDADLYDTYNDNDTV PPTTVGGSTTSISNYKTKDFVGGGGSGGGGSAETKEFKTLYNLFIDSYLQKLAQHSIPTNV CAIHIGEIVIGQFKNCALRITNKCMSNSRLSFTLMVESFIEVISLLSEKDRRAIAEEIGIDLDDV PSVVSKLEKNCNAYAENVNIIDIQKLNIGECsAPPGQHMLLQIVNTGSAEANCGLQTVKKS LNKIYVPIIENRLPYDPWFLVGVAIILVIFTVAICSIRRNALKYRYGTFLYV
107	1701 编码的氨基酸序列	MTQLPILLLLISLVYATPSPQTSKKIGDDATISCSRNNNTNYYVMSAWYK EPNSIILAAKSDVLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTY ICAFFMTSTTNDTDKVDYEEYSIELIVNTDSESTIDIILSGSTPETISEKPEDI DNSNCSSVFEITTPITDNVDDHTDTVTYTSDSINTVNASSGESTTDEIPE PITDKEDHTVTDTVSYTTVSTSSGIVTTKSTTDDADLYDTYNDNDTVPP TTVGGSTTSISNYKTKDFVGGGGSGGGGSAAVKTPVIVVPVIDRPPSETF PNVHEHINDQKFDDVKDNEVMQEKRDVVIVNDDPDHYKDYVFIQWTG GNIRDDDKYTHFFSGFCNTMCETEETKRNIARHLALWDSKFFTELENKNV EYVVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREIITGNKVKTTELVIDK DHAIFTYGGYDVSL SAYIIRVTTALNIVDEIISGGLSSGFYFEIARIENE MKINRQIMDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAAKRYPG VMYFTFTPLISGGGGSGGGGSAETKEFKTLYNLFIDSYLQKLAQHSIPTN VTCAIHIGEIVIGQFKNCALRITNKCMSNSRLSFTLMVESFIEVISLLSEKDR RAIAEEIGIDLDDVPSVVSKLEKNCNAYAENVNIIDIQKLNIGECsAPPGQ HMLLQIVNTGSAEANCGLQTVKKS LNKIYVPIIENRLPYDPWFLVGVA IILVIFTVAICSIRRNALKYRYGTFLYV
108	1801 编码的氨基酸序列	MTQLPILLLLISLVYATPSPQTSKKIGDDATISCSRNNNTNYYVMSAWYK EPNSIILAAKSDVLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTY ICAFFMTSTTNDTDKVDYEEYSIELIVNTDSESTIDIILSGSTPETISEKPEDI DNSNCSSVFEITTPITDNVDDHTDTVTYTSDSINTVNASSGESTTDEIPE

		PITDKEDHTVTDTVSYTTVSTSSGIVTTKSTTDDADLYDTYNDNDTVPP TTVGGSTTSISNYKTKDFVGGGGSGGGGSRLNQCMSANKAAITDSAVA VAAASSTHRKVVSSTTQYDHKESCNGLYYQGSCYILHSDYKSFEDAKA NCAAESSTLPNKSDVLTWTLIDYVEDTWGSDGNPITKTTSDYQSDVSQ EVRKYFCTGGGGSGGGGSTCTVPTMNNAKLTSTETSFNDKQKVTFTCD SGYHSLDPNAV CETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNST MTLSCNGETKYFRCEEKNGNTSWNDTVTCPNAECQPLQLEHGSCQPVK EKYSFGEYMTINCDVGVEVIGVSYISCTANSWNVIPSCQKQCDIPSLNG LISGSTFSIGGVIHLSCKSGFTLTGSPSSCIDGKWNPIPTCVRSEEFDPV DDGPDDDETDL SKLSKDVVQYEQEIESLEATYHIIIMALTIMGVIFLISIIVL VCSCDKNNDQYKFHKLPL
109	1901 编码的氨基酸序列	MDWTWILFLVAAATRVHSDGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKAYGDDN EETLKQRLTNLEKKITNITTKFEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGR HPYEGGGSGGGSGGAAASIQTTVNTLSERISSKLEQEANASATKCDIEIGNFYIRQNHG CNITVKNMCSADADAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTVVRD FENYVKQTCNSSAVVDNKLKIQNVIIDECYAGPSPTNLEFINTGSSKGNCAIKALMQLTT KATTQIAPRQVAGGGGGSGGGGSAETKEFKTLNLFIDSYLQKLAQHSIPTNVTCIHIGE VIGQFKNCALRITNKCMSNRLSFTLMVESFIEVISLLSEKDRRAIAEEIGIDLDDVPSVSKL EKNCNAYAEVNNIIDIKLNIGECAPPQGHMLLQIVNTGSAEANCGLQITVKSILNKIYVP PIENRPPYDPWFLVGVAAILVIFTVAICSIRRNALKYRYGTFLYV
110	2001 编码的氨基酸序列	MDWTWILFLVAAATRVHSAAVKTPVIVVPVIDRPPSETFPNVHEHINDQKFDDVKDNEV MQEKRDVVIVNDPDHYKDYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIAR HLALWDSKFFTELENKNVEYVVIENDNVIEDITFLRPVLKAIHDKKIDILQMREITGNKVKT ELVIDKDHAIFTYTGGYDVLSAYIIRVTTALNIVDEIISKGGSSGFYFEIARIENEMKINRQI MDNSAKYVEHDPRLVAEHRFETMKPNFWSRIGTVAAKRYPGVMYFTTTPISGGGGSG GGGSAETKEFKTLNLFIDSYLQKLAQHSIPTNVTCIHIGEIVIGQFKNCALRITNKCMSNS RLSFTLMVESFIEVISLLSEKDRRAIAEEIGIDLDDVPSVSKLEKNCNAYAEVNNIIDIKLNI GECAPPQGHMLLQIVNTGSAEANCGLQITVKSILNKIYVPIIENRPPYDPGGGGSGGG GSGAAASIQTTVNTLSERISSKLEQEANASATKCDIEIGNFYIRQNHGCNITVKNMCSAD ADAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTVVRDFENYVKQTCNSSA VVDNKLKIQNVIIDECYAGPSPTNLEFINTGSSKGNCAIKALMQLTTKATTQIAPRQVAG TGVQFYMIVIGVILAALFMYAKRMLFTSTNDKIKLILANKENVHWTYMDTFFRTSPMII ATTDIQN
111	A29L 肽-1	DGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKAYGDDNEETLKQRLTNLEKKITNITK FEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGRHPY
112	A29L 肽-2	MDGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKAYGDDNEETLKQRLTNLEKKITNIT TKFEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGRHPY
113	A29L 肽-3	DGTLFPGDDDLAIPATEFFSTKAAKNPETKREAIKAYGDDNEETLKQRLTNLEKKITNITK FEQIEKCKRNDEVLFRLNHAETLRAAMISLAKKIDVQTGRHPY
114	A35R 肽	MMTPENDEEQTSVFSATVYGDKIQGKNKRKRVIGLCIRISMVISLLSMITMSAFLIVRLNQ CMSANKAAITDSAVAVAAASSTHRKVVSSTTQYDHKESCNGLYYQGSCYILHSDYKSFED AKANCAAESSTLPNKSDVLTWTLIDYVEDTWGSDGNPITKTTSDYQSDVSQEVRYFCT
115	A35R 胞外段	RLNQCMSANKAAITDSAVAVAAASSTHRKVVSSTTQYDHKESCNGLYYQGSCYILHSDYK SFEDAKANCAAESSTLPNKSDVLTWTLIDYVEDTWGSDGNPITKTTSDYQSDVSQEVRYK

		YFCT
116	B2R 信号肽	MTQLPILLLLISLVYATP
117	B2R 肽全长(除信号肽)	SPQTSKKIGDDATISCSRNTNYYYVMSAWYKEPNSIILAAKSDVLYFDNYTKDKISYDSP YDDLVTITIKSLTAGDAGTYICAFMTSTTNDTDKVDYEEYSIELIVNTDSESTIDIILSGSTP ETISEKPEDIDNSNCSSVFEITTEPITDNVDDHTDVTYTSDSINTVNASSGESTTDEIPEPI TDKEEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDTVPPTTVGGSTTSISNYK TKDFVEIFGITLLIILSAVAIFCITYYICNKHPRKYKTENKV
118	B2R 肽全长	MTQLPILLLLISLVYATPSPQTSKKIGDDATISCSRNTNYYYVMSAWYKEPNSIILAAKSD VLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTYICAFMTSTTNDTDKVDYEEYSIELI VNTDSESTIDIILSGSTPETISEKPEDIDNSNCSSVFEITTEPITDNVDDHTDVTYTSDSINT VNASSGESTTDEIPEPITDKEEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDT VPPTTVGGSTTSISNYKTKDFVEIFGITLLIILSAVAIFCITYYICNKHPRKYKTENKV
119	B2R 胞外段	SPQTSKKIGDDATISCSRNTNYYYVMSAWYKEPNSIILAAKSDVLYFDNYTKDKISYDSP YDDLVTITIKSLTAGDAGTYICAFMTSTTNDTDKVDYEEYSIELIVNTDSESTIDIILSGSTP ETISEKPEDIDNSNCSSVFEITTEPITDNVDDHTDVTYTSDSINTVNASSGESTTDEIPEPI TDKEEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDTVPPTTVGGSTTSISNYK TKDFV
120	B2R 信号肽+胞外段	MTQLPILLLLISLVYATPSPQTSKKIGDDATISCSRNTNYYYVMSAWYKEPNSIILAAKSD VLYFDNYTKDKISYDSPYDDLVTITIKSLTAGDAGTYICAFMTSTTNDTDKVDYEEYSIELI VNTDSESTIDIILSGSTPETISEKPEDIDNSNCSSVFEITTEPITDNVDDHTDVTYTSDSINT VNASSGESTTDEIPEPITDKEEDHTVTDVSYTTVSTSSGIVTTKSTDDADLYDTYNDNDT VPPTTVGGSTTSISNYKTKDFV
121	B6R 信号肽	MKTISVVTLLCVLPAVVYS
122	B6R 肽全长(除信号肽)	TCTVPTMNNAKLTSTETSFNDKQKVFTFTCDSGYHSLDPNAVCECTDKWKYENPCKKMCTV SDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWNDTVTCPNAECQPLQLEH GSCQPVKEKYSFGEYMTINCDVGVEIGVSYISCTANSWNVIPSCQKCDIPSLSNGLISGS TFSIGGVIHLSCSGFTLTGSPSSTCIDGKWNPIPTCVRNNEEFDPVDDGPDDDETDLKLSK DVVQYEQEIESLEATYHIIIMALTIMGVIFLISIIVLCSCDKNNDQYKFHKLLP
123	B6R 肽全长	MKTISVVTLLCVLPAVVYSTCTVPTMNNAKLTSTETSFNDKQKVFTFTCDSGYHSLDPNAVCE ETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWN DTVTCPNAECQPLQLEHGSCQPVKEKYSFGEYMTINCDVGVEIGVSYISCTANSWNVIPSC CQKCDIPSLSNGLISGSTFSIGGVIHLSCSGFTLTGSPSSTCIDGKWNPIPTCVRNNEE DPVDDGPDDDETDLKLSKDVVQYEQEIESLEATYHIIIMALTIMGVIFLISIIVLCSCDKNND QYKFHKLLP
124	B6R 胞外段	TCTVPTMNNAKLTSTETSFNDKQKVFTFTCDSGYHSLDPNAVCECTDKWKYENPCKKMCTV SDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWNDTVTCPNAECQPLQLEH GSCQPVKEKYSFGEYMTINCDVGVEIGVSYISCTANSWNVIPSCQKCDIPSLSNGLISGS TFSIGGVIHLSCSGFTLTGSPSSTCIDGKWNPIPTCVRNNEEFDPVDDGPDDDETDLKLSK DVVQYEQEIESLEATYH
125	B6R 信号肽+胞外肽	MKTISVVTLLCVLPAVVYSTCTVPTMNNAKLTSTETSFNDKQKVFTFTCDSGYHSLDPNAVCE ETDKWKYENPCKKMCTVSDYVSELYDKPLYEVNSTMTLSCNGETKYFRCEEKNGNTSWN DTVTCPNAECQPLQLEHGSCQPVKEKYSFGEYMTINCDVGVEIGVSYISCTANSWNVIPSC CQKCDIPSLSNGLISGSTFSIGGVIHLSCSGFTLTGSPSSTCIDGKWNPIPTCVRNNEE DPVDDGPDDDETDLKLSKDVVQYEQEIESLEATYH

126	C15L 肽全长	AETKEFKTLYNLFIDSYLQKLAQHSIPTNVTCAIHIGEIVIGQFKNCALRITNKCMSNSRSLSTL MVESFIEVISLLSEKDRRAIAEEIGIDLDDVPSVVSKEKNCNAYAENVNIIQKLNIGECSA PPGQHMLLQIVNTGSAEANCGLQTIVKSLNKIYVPPHIIENRLPYDPWFLVGVAILVIFTVA ICSIRRNALKYRYGTFLYV
127	C15L 胞外段	AETKEFKTLYNLFIDSYLQKLAQHSIPTNVTCAIHIGEIVIGQFKNCALRITNKCMSNSRSLSTL MVESFIEVISLLSEKDRRAIAEEIGIDLDDVPSVVSKEKNCNAYAENVNIIQKLNIGECSA PPGQHMLLQIVNTGSAEANCGLQTIVKSLNKIYVPPHIIENRLPYDP
128	M1R 胞外段	GAAASIQTTVNTLSERISSKLEQEANASAQTKCDIEIGNFYIRQNHGCNITVKNMCSADAD AQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTTVRDFENYVKQTCNSSAVV DNKLKIQNVIIDECYGAPGSPTNLEFINTGSSKGNCAIKALMQLTTKATTQIAPRQVAG
129	M1R 肽全长-1	MGAAASIQTTVNTLSERISSKLEQEANASAQTKCDIEIGNFYIRQNHGCNITVKNMCSADA DAQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTTVRDFENYVKQTCNSSAV VDNKLKIQNVIIDECYGAPGSPTNLEFINTGSSKGNCAIKALMQLTTKATTQIAPRQVAGT GVQFYMIVIGVILALFMYAKRMLFTSTNDKIKLILANKENVHWTTTMDTFFRTSPMIIA TTDIQN
130	M1R 肽全长-2	GAAASIQTTVNTLSERISSKLEQEANASAQTKCDIEIGNFYIRQNHGCNITVKNMCSADAD AQLDAVLSAATETYSGLTPEQKAYVPAMFTAALNIQTSVNTTVRDFENYVKQTCNSSAVV DNKLKIQNVIIDECYGAPGSPTNLEFINTGSSKGNCAIKALMQLTTKATTQIAPRQVAGTG VQFYMIVIGVILALFMYAKRMLFTSTNDKIKLILANKENVHWTTTMDTFFRTSPMIIAT TDIQN
131	H3L 肽全长	AAVKTPVIVPVIDRPPSETFPNVHEHINDQKFDDVDKNEVMQEKRDRVIVNDPDPHYK DYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARHLALWDSKFFTELENKNVEY VVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREIITGNKVKTTELVIDKDHAIFTYTGgyDVSL SAYIIRVTTALNIVDEIHKSGGLSSGFYFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRF ETMKPNFWSRIGTVAAKRYPGVMYFTTTP LISFFGLFDINVLIVILFIMFMLIFNVKSKLL WFLTGTFTVAFI
132	H3L 胞外段	AAVKTPVIVPVIDRPPSETFPNVHEHINDQKFDDVDKNEVMQEKRDRVIVNDPDPHYK DYVFIQWTGGNIRDDDKYTHFFSGFCNTMCTEETKRNIARHLALWDSKFFTELENKNVEY VVIIENDNVIEDITFLRPVLKAIHDKKIDILQMREIITGNKVKTTELVIDKDHAIFTYTGgyDVSL SAYIIRVTTALNIVDEIHKSGGLSSGFYFEIARIENEMKINRQIMDNSAKYVEHDPRLVAEHRF ETMKPNFWSRIGTVAAKRYPGVMYFTTTP LIS
133	E8L 肽全长	MPQQLSPINIETKKAISDARLKTLDIHYNESKPTTIQNTGKLVIRINFKGGYISGGFLPNEYVL STIHIIYWGKEDDYGSNHLIDVYKYSGEINLVHWNKKYSSYEEAKKHDDGIIIIAIFLQVSDH KNVYFQKIVNQLDSIRSANMSAPFDSVFYLDNLLPSTLDYFTYLGTTINHSADAAWIIPTPI NIHSDQLSKFRTLSSSNHEGKPHYITENYRNPYKLNDDTQVYYSGEIIRAATTSPVRENYF MKWLSDLREACFSYYQKYIEGNKTFI AIVFVILTALFLMSQRYRSREKQN
134	IgE 信号肽	MDWTWILFLVAAATRVS
135	HA 多肽标签-1	YPYDVPDYAE
136	HA 多肽标签-2	YPYDVPDYA
137	GS 接头-1	GGGGS
138	GS 接头-2	GGGSGGGGS
139	5'UTR RNA	AUCCAAGCUUAGCUUGCUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAAC

		UUUGGCGCCGCCACC
140	3'UTR RNA-1	UAAGAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCC CCCCAGCCCCUCCUCCCUUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUC UGAGUGGGCGGCACUAGU
141	3'UTR RNA-2	GAAUUCUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCC CAGCCCCUCCUCCCUUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUG AGUGGGCGGCACUAGU
142	5'UTR DNA	ATCCAAGCTTAGCTTGCTTGTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGG CGCCGCCACC
143	3'UTR DNA-1	TAAGAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCC AGCCCTCCTCCCCTTCTGACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGG GCGGCACTAGT
144	3'UTR DNA-2	GAATTCTAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGC CCCTCCTCCCCTTCTGACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCG GCACTAGT
145	PolyA 尾 RNA	AA AA AAAAAAAAA
146	PolyA 尾 DNA	AA AA AAAAAAAAA
147	帽近端序列	GGGGGGG
148	5'UTR RNA-2	CUUGUUCUUUUUGCAGAAGCUCAGAAUAAACGCUCAACUUUGGC
149	3'UTR RNA-3	UAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCAGCCCC UCCUCCCUUCCUGCACCCGUACCCCGUGGUCUUUGAAUAAAGUCUGAGUGGGC GGC
150	5'UTR DNA-2	CTTGTTCTTTTGCAGAAGCTCAGAATAAACGCTCAACTTTGGC
151	3'UTR DNA-3	TAGGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCTTGGGCCTCCCCCAGCCCCTCT CCCCTTCTGACCCGTACCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGC

权 利 要 求 书

1. 一种多肽或多肽组合，其包含一个或多个肽段，所述肽段各自独立地选自猴痘病毒 A29L、A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段；
任选地，所述多肽或多肽组合还包含一个或多个接头、IgE 信号肽和/或 HA 多肽标签。
2. 权利要求 1 的多肽或多肽组合，其中
- (1) 所述多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段；
 - (2) 所述多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段；
 - (3) 所述多肽或多肽组合包含 A35R、B2R 和 C15L 蛋白中的肽段；
 - (4) 所述多肽或多肽组合包含 B6R、M1R 和 H3L 蛋白中的肽段；
 - (5) 所述多肽或多肽组合包含 B2R、A35R 和 B6R 蛋白中的肽段；
 - (6) 所述多肽或多肽组合包含 A29L、M1R 和 C15L 蛋白中的肽段；
 - (7) 所述多肽或多肽组合包含 A29L、B6R 和 M1R 蛋白中的肽段；
 - (8) 所述多肽或多肽组合包含 B6R、M1R 和 A35R 蛋白中的肽段；
 - (9) 所述多肽或多肽组合包含 A29L、B2R 和 C15L 蛋白中的肽段；
 - (10) 所述多肽或多肽组合包含 B2R、H3L 和 C15L 蛋白中的肽段；
 - (11) 所述多肽或多肽组合包含 H3L、M1R 和 C15L 蛋白中的肽段；
 - (12) 所述多肽或多肽组合包含 A29L 蛋白中的肽段；
 - (13) 所述多肽或多肽组合包含 A35R 蛋白中的肽段；
 - (14) 所述多肽或多肽组合包含 B2R 蛋白中的肽段；
 - (15) 所述多肽或多肽组合包含 B6R 蛋白中的肽段；
 - (16) 所述多肽或多肽组合包含 C15L 蛋白中的肽段；
 - (17) 所述多肽或多肽组合包含 M1R 蛋白中的肽段；或者
 - (18) 所述多肽或多肽组合包含 H3L 蛋白中的肽段。
3. 权利要求 1 或 2 的多肽，其中
- 所述 A29L 蛋白中的肽段包含 SEQ ID NO: 111 或 113 的氨基酸序列；
所述 A35R 蛋白中的肽段包含 SEQ ID NO: 114 或 115 的氨基酸序列；
所述 B2R 蛋白中的肽段包含 SEQ ID NO: 116、117、118、119 或 120 的氨基酸序列；
所述 B6R 蛋白中的肽段包含 SEQ ID NO: 121、122、123、124 或 125 的氨基酸序列；
所述 C15L 蛋白中的肽段包含 SEQ ID NO: 126 或 127 的氨基酸序列；
所述 M1R 蛋白中的肽段包含 SEQ ID NO: 128 或 130 的氨基酸序列；或者

所述 H3L 蛋白中的肽段包含 SEQ ID NO: 131 或 132 的氨基酸序列。

4. 权利要求 1-3 中任一项的多肽或多肽组合，其中

- (1) 所述多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段、以及接头；
- (2) 所述多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段、IgE 信号肽以及接头；
- (3) 所述多肽或多肽组合包含 A29L、A35R、B2R、B6R、C15L 和 M1R 蛋白中的肽段、IgE 信号肽、HA 多肽标签以及接头；
- (4) 所述多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段、IgE 信号肽、HA 多肽标签以及接头；
- (5) 所述多肽或多肽组合包含 A35R、B2R、B6R、C15L、M1R 和 H3L 蛋白中的肽段、IgE 信号肽以及接头；
- (6) 所述多肽或多肽组合包含 A35R、B2R 和 C15L 蛋白中的肽段以及接头；
- (7) 所述多肽或多肽组合包含 B6R、M1R 和 H3L 蛋白中的肽段以及接头；
- (8) 所述多肽或多肽组合包含 B2R、A35R 和 B6R 蛋白中的肽段以及接头；
- (9) 所述多肽或多肽组合包含 A29L、M1R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头；
- (10) 所述多肽或多肽组合包含 A29L、B6R 和 M1R 蛋白中的肽段、IgE 信号肽以及接头；
- (11) 所述多肽或多肽组合包含 B6R、M1R 和 A35R 蛋白中的肽段以及接头；
- (12) 所述多肽或多肽组合包含 A29L、B2R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头；
- (13) 所述多肽或多肽组合包含 B2R、H3L 和 C15L 蛋白中的肽段以及接头；
- (14) 所述多肽或多肽组合包含 H3L、M1R 和 C15L 蛋白中的肽段、IgE 信号肽以及接头；
- (15) 所述多肽或多肽组合包含 IgE 信号肽、A29L 蛋白中的肽段以及 HA 多肽标签；
- (16) 所述多肽或多肽组合包含 A35R 蛋白中的肽段以及 HA 多肽标签；
- (17) 所述多肽或多肽组合包含 B2R 蛋白中的肽段以及 HA 多肽标签；
- (18) 所述多肽或多肽组合包含 B6R 蛋白中的肽段、接头以及 HA 多肽标签；
- (19) 所述多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 C15L 蛋白中的肽段；
- (20) 所述多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 M1R 蛋白中的肽段；或者
- (21) 所述多肽或多肽组合包含 IgE 信号肽、接头、HA 多肽标签以及 H3L 蛋白中的肽段。

5. 权利要求 1-4 中任一项的多肽或多肽组合，其中

所述 IgE 信号肽包含 SEQ ID NO: 134 的氨基酸序列；

所述 HA 多肽标签包含 SEQ ID NO: 135 或 136 的氨基酸序列；和/或
所述接头为 GS 接头，优选地，所述 GS 接头包含 SEQ ID NO: 137 或 138 的氨基酸序列。

6. 权利要求 1-5 中任一项的多肽或多肽组合，其中

所述多肽或多肽组合包含 SEQ ID NO: 102、104、108、109、89、90、91、92、93、94、95、101、103、105、106、107 和 110 之一的氨基酸序列或者与 SEQ ID NO: 102、104、108、109、89、90、91、92、93、94、95、101、103、105、106、107 和 110 之一的氨基酸序列具有至少 80% 相同性的氨基酸序列。

7. 权利要求 1-5 中任一项的多肽或多肽组合，其中

所述多肽或多肽组合包含两个或更多个选自 SEQ ID NO: 102、104、108、109、89、90、91、92、93、94、95、101、103、105、106、107 和 110 的氨基酸序列或者两个或更多个与选自 SEQ ID NO: 102、104、108、109、89、90、91、92、93、94、95、101、103、105、106、107 和 110 的氨基酸序列具有至少 80% 相同性的氨基酸序列。

8. 权利要求 7 的多肽或多肽组合，其中

(1) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 102 和 104 的氨基酸序列；

(2) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 108 和 109 的氨基酸序列；

(3) 所述多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽，其中所述第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽分别包含 SEQ ID NO: 89、90、91、92、93 和 94 的氨基酸序列；

(4) 所述多肽或多肽组合包含第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽，其中所述第一多肽、第二多肽、第三多肽、第四多肽、第五多肽和第六多肽分别包含 SEQ ID NO: 90、91、92、93、94 和 95 的氨基酸序列；

(5) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 102 和 103 的氨基酸序列；

(6) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 105 和 106 的氨基酸序列；

(7) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 105 和 107 的氨基酸序列；或者

(8) 所述多肽或多肽组合包含第一多肽和第二多肽，其中所述第一多肽和第二多肽分别包含 SEQ ID NO: 110 和 108 的氨基酸序列。

9. 一种核酸或核酸组合，其包含编码权利要求 1-8 中任一项的多肽或多肽组合的多核苷酸。

10. 权利要求 9 的核酸或核酸组合，

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

所述多核苷酸包含 SEQ ID NO: 14、16、20、21、1、2、3、4、5、6、7、13、15、17、18、19 和 22 之一的核苷酸序列或者与 SEQ ID NO: 14、16、20、21、1、2、3、4、5、6、7、13、15、17、18、19 和 22 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列；或者

其中所述多核苷酸为 DNA，

所述多核苷酸包含 SEQ ID NO: 36、38、42、43、23、24、25、26、27、28、29、35、37、39、40、41 和 44 之一的核苷酸序列或者与 SEQ ID NO: 36、38、42、43、23、24、25、26、27、28、29、35、37、39、40、41 和 44 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列。

11. 权利要求 9 的核酸或核酸组合，

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

所述多核苷酸包含两个或更多个选自 SEQ ID NO: 14、16、20、21、1、2、3、4、5、6、7、13、15、17、18、19 和 22 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 14、16、20、21、1、2、3、4、5、6、7、13、15、17、18、19 和 22 的核苷酸序列具有至少 80% 相同性的核苷酸序列；或者

其中所述多核苷酸为 DNA，

所述多核苷酸包含两个或更多个选自 SEQ ID NO: 36、38、42、43、23、24、25、26、27、28、29、35、37、39、40、41 和 44 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 36、38、42、43、23、24、25、26、27、28、29、35、37、39、40、41 和 44 的核苷酸序列具有至少 80% 相同性的核苷酸序列。

12. 权利要求 11 的核酸或核酸组合，其中

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

(1) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 14 和 16 的核苷酸序列；

(2) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 20 和 21 的核苷酸序列；

(3) 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 1、2、3、4、5 和 6 的核苷酸序列；

(4) 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 2、3、4、5、6 和 7 的核苷酸序列；

(5) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第

二多核苷酸分别包含 SEQ ID NO: 14 和 15 的核苷酸序列；

(6) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 17 和 18 的核苷酸序列；

(7) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 17 和 19 的核苷酸序列；或者

(8) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 22 和 20 的核苷酸序列；

或者

其中所述多核苷酸为 DNA，

(1') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 36 和 38 的核苷酸序列；

(2') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 42 和 43 的核苷酸序列；

(3') 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 23、24、25、26、27 和 28 的核苷酸序列；

(4') 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 24、25、26、27、28 和 29 的核苷酸序列；

(5') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 36 和 37 的核苷酸序列；

(6') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 39 和 40 的核苷酸序列；

(7') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 39 和 41 的核苷酸序列；或者

(8') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 44 和 42 的核苷酸序列。

13. 权利要求 9-12 中任一项的核酸或核酸组合，

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

所述多核苷酸进一步包含 5'-UTR 序列；优选地，所述 5'-UTR 序列包含 SEQ ID NO: 139 或 148 的核苷酸序列；

所述多核苷酸进一步包含 3'-UTR 序列；优选地，所述 3'-UTR 序列包含 SEQ ID NO: 140、141 或 149 的核苷酸序列；和/或

所述多核苷酸进一步包含 poly(A)序列；优选地，所述 poly(A)序列包含 SEQ ID NO: 145 的核苷酸序列；或者

其中所述多核苷酸为 DNA，

所述多核苷酸进一步包含 5'-UTR 序列对应的 DNA 序列；优选地，所述 5'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 142 或 150 的核苷酸序列；

所述多核苷酸进一步包含 3'-UTR 序列对应的 DNA 序列；优选地，所述 3'-UTR 序列对应的 DNA 序列包含 SEQ ID NO: 143、144 或 151 的核苷酸序列；和/或

所述多核苷酸进一步包含 poly(A)序列对应的 DNA 序列；优选地，所述 poly(A)序列对应的 DNA 序列包含 SEQ ID NO: 146 的核苷酸序列。

14. 权利要求 9-13 中任一项的核酸或核酸组合，

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

所述多核苷酸包含 SEQ ID NO: 58、60、64、65、45、46、47、48、49、50、51、57、59、61、62、63 和 66 之一的核苷酸序列或者与 SEQ ID NO: 58、60、64、65、45、46、47、48、49、50、51、57、59、61、62、63 和 66 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列；或者

其中所述多核苷酸为 DNA，

所述多核苷酸包含 SEQ ID NO: 80、82、86、87、67、68、69、70、71、72、73、79、81、83、84、85 和 88 之一的核苷酸序列或者与 SEQ ID NO: 80、82、86、87、67、68、69、70、71、72、73、79、81、83、84、85 和 88 之一的核苷酸序列具有至少 80% 相同性的核苷酸序列。

15. 权利要求 9-13 中任一项的核酸或核酸组合，

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

所述多核苷酸包含两个或更多个选自 SEQ ID NO: 58、60、64、65、45、46、47、48、49、50、51、57、59、61、62、63 和 66 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 58、60、64、65、45、46、47、48、49、50、51、57、59、61、62、63 和 66 的核苷酸序列具有至少 80% 相同性的核苷酸序列；或者

其中所述多核苷酸为 DNA，

所述多核苷酸包含两个或更多个选自 SEQ ID NO: 80、82、86、87、67、68、69、70、71、72、73、79、81、83、84、85 和 88 的核苷酸序列或者两个或更多个与选自 SEQ ID NO: 80、82、86、87、67、68、69、70、71、72、73、79、81、83、84、85 和 88 的核苷酸序列具有至少 80% 相同性的核苷酸序列。

16. 权利要求 15 的核酸或核酸组合，其中

其中所述多核苷酸为 RNA，优选为 mRNA、circRNA 或 saRNA，更优选为 mRNA，

(1) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 58 和 60 的核苷酸序列；

(2) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 64 和 65 的核苷酸序列；

(3) 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核

苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 45、46、47、48、49 和 50 的核苷酸序列；

(4) 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 46、47、48、49、50 和 51 的核苷酸序列；

(5) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 58 和 59 的核苷酸序列；

(6) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 61 和 62 的核苷酸序列；

(7) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 61 和 63 的核苷酸序列；或者

(8) 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 66 和 64 的核苷酸序列；

或者

其中所述多核苷酸为 DNA，

(1') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 80 和 82 的核苷酸序列；

(2') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 86 和 87 的核苷酸序列；

(3') 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 67、68、69、70、71 和 72 的核苷酸序列；

(4') 所述多核苷酸包含第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸，其中所述第一多核苷酸、第二多核苷酸、第三多核苷酸、第四多核苷酸、第五多核苷酸和第六多核苷酸分别包含 SEQ ID NO: 68、69、70、71、72 和 73 的核苷酸序列；

(5') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 80 和 81 的核苷酸序列；

(6') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 83 和 84 的核苷酸序列；

(7') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 83 和 85 的核苷酸序列；或者

(8') 所述多核苷酸包含第一多核苷酸和第二多核苷酸，其中所述第一多核苷酸和第二多核苷酸分别包含 SEQ ID NO: 88 和 86 的核苷酸序列。

17. 一种组合物，其包含权利要求 9-16 中任一项的核酸或核酸组合。

18. 权利要求 17 的组合物，其包含包封所述核酸或核酸组合的脂质。
19. 权利要求 17 或 18 的组合物，其包含至少第一脂质，
其中所述第一脂质包含根据式 01-I 或 01-II 的化合物中的一种或多种，优选地所述第一脂质包含表 01-1 中所列的化合物中的一种或多种；
其中所述第一脂质包含根据式 02-I 或 02-II 的化合物中的一种或多种，优选地所述第一脂质包含表 02-1 中所列的化合物中的一种或多种；
其中所述第一脂质包含根据式 03-I 的化合物中的一种或多种，优选地所述第一脂质包含表 03-1 中所列的化合物中的一种或多种；或者
其中所述第一脂质包含根据式 04-I 或 04-III 的化合物中的一种或多种，优选地所述第一脂质包含表 04-1 中所列的化合物中的一种或多种；
更优选地，所述第一脂质包含化合物 01-1。
20. 权利要求 19 的组合物，其还包含第二脂质，
其中所述第二脂质为聚合物结合的脂质；优选地，其包含根据式 05-I 的化合物中的一种或多种。
21. 权利要求 17-20 中任一项的组合物，其包含
第一脂质、第二脂质、磷脂和类固醇。
22. 权利要求 17-21 中任一项的组合物，其中所述组合物被配制为将所述核酸或核酸组合包封在脂质中的脂质纳米颗粒。
23. 一种药物组合物或试剂盒，其包含权利要求 1-8 中任一项的多肽或多肽组合、权利要求 9-16 中任一项的核酸或核酸组合或者权利要求 17-22 中任一项的组合物。
24. 一种表达载体，其包含权利要求 9-16 中任一项的核酸或核酸组合。
25. 一种宿主细胞，其包含权利要求 9-16 中任一项的核酸或核酸组合或者权利要求 24 的表达载体。
26. 权利要求 1-8 中任一项的多肽或多肽组合、权利要求 9-16 中任一项的核酸或核酸组合、权利要求 17-22 中任一项的组合物、权利要求 23 的药物组合物或试剂盒在制备用于预防和/或治疗受试者猴痘病毒感染的疫苗中的用途。
27. 权利要求 26 的用途，其中所述受试者是哺乳动物；优选地，所述受试者是人。