

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480007678.7

[51] Int. Cl.
C07H 21/04 (2006.01)
C12Q 1/16 (2006.01)

[43] 公开日 2006年5月3日

[11] 公开号 CN 1768074A

[22] 申请日 2004.3.24

[21] 申请号 200480007678.7

[30] 优先权

[32] 2003. 3. 24 [33] US [31] 60/457,031

[32] 2003. 3. 26 [33] US [31] 60/457,730

[32] 2003. 4. 2 [33] US [31] 60/459,931

[32] 2003. 4. 3 [33] US [31] 60/460,357

[32] 2003. 4. 8 [33] US [31] 60/461,265

[32] 2003. 4. 14 [33] US [31] 60/462,805

[32] 2003. 4. 23 [33] US [31] 60/464,886

[32] 2003. 5. 5 [33] US [31] 60/468,139

[32] 2003. 5. 16 [33] US [31] 60/471,200

[86] 国际申请 PCT/CN2004/000247 2004.3.24

[87] 国际公布 WO2004/085455 英 2004.10.7

[85] 进入国家阶段日期 2005.9.21

[71] 申请人 香港大学

地址 中国香港薄扶林道

[72] 发明人 陈国雄 管 轶 J·M·黎国思

J·S·M·裴伟士 潘烈文

袁国勇

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 王景朝

权利要求书 2 页 说明书 79 页 附图 94 页

[54] 发明名称

对引起严重急性呼吸道综合症(SARS)的人病毒的诊断性试验

[57] 摘要

本发明涉及对在人类中导致严重急性呼吸道综合症(SARS)的病毒(“hSARS 病毒”)的诊断性试验。具体地说,本发明涉及使用逆转录和聚合酶链反应的实时定量 PCR 试验,以检测 hSARS 病毒。明确地说,所述定量试验为 TaqMan[®] 试验。本发明进一步涉及包含用于检测 hSARS 病毒的核酸分子的诊断试剂盒。

1. 一种分离的核酸分子, 所述核酸分子基本由 SEQ ID NO: 2471、2472 核酸序列或其互补序列组成。
- 5 2. 一种分离的核酸分子, 所述核酸分子基本由 SEQ ID NO: 2474、2475 核酸序列或其互补序列组成。
3. 一种分离的核酸分子, 所述核酸分子基本由 SEQ ID NO: 2473、2476 核酸序列或其互补序列组成。
4. 一种分离的核酸分子, 所述核酸分子在严格条件下与具有 SEQ
10 ID NO: 2471、2472、2473、2474、2475 或 2476 核酸序列或其互补序
列的核酸分子杂交。
5. 一种分离的多肽, 所述多肽由权利要求 1-4 任一项的核酸分子
编码。
6. 一种抗体或其抗原结合片段, 所述抗体或其抗原结合片段免疫
15 特异性结合 SEQ ID NO: 2471、2472 或 2473 核酸序列编码的多肽。
7. 一种抗体或其抗原结合片段, 所述抗体或其抗原结合片段免疫
特异性结合 SEQ ID NO: 2474、2475 或 2476 核酸序列编码的多肽。
8. 一种在生物样品中检测 hSARS 病毒存在的方法, 所述方法包
括:
20 (a)使用具有 SEQ ID NO: 2471 和/或 2472 核酸序列的引物扩增
hSARS 病毒的核酸;
(b) 使用具有 SEQ ID NO: 2473 核酸序列的探针检测核酸。
9. 一种在生物样品中检测 hSARS 病毒存在的方法, 所述方法包
括:
25 (a)使用具有 SEQ ID NO: 2474 和/或 2475 核酸序列的引物扩增
hSARS 病毒的核酸;
(b) 使用具有 SEQ ID NO: 2476 核酸序列的探针检测核酸。
10. 一种鉴定 hSARS 病毒感染对象的方法, 所述方法包括:

- (a) 从来自对象的生物样品获得总 RNA;
 - (b) 逆转录总 RNA 以获得 cDNA;
 - (c) 用一套来源于 hSARS 核苷酸序列的引物对 cDNA 进行 PCR 试验。
- 5 11. 一种鉴定 hSARS 病毒感染对象的方法, 所述方法包括:
- (a) 从来自对象的生物样品获得总 RNA;
 - (b) 逆转录总 RNA 以获得 cDNA;
 - (c) 用一套具有 SEQ ID NO: 2471 和/或 2472 核酸序列的引物对 cDNA 进行 PCR 试验。
- 10 12. 权利要求 11 的方法, 所述方法进一步包括(d)用探针检测 PCR 试验的产物。
13. 权利要求 12 的方法, 其中所述探针为具有 SEQ ID NO: 2473 核苷酸序列的核酸分子。
14. 一种鉴定 hSARS 病毒感染对象的方法, 所述方法包括:
- 15 (a) 从来自对象的生物样品获得总 RNA;
- (b) 逆转录总 RNA 以获得 cDNA;
 - (c) 用一套具有 SEQ ID NO: 2474 和/或 2475 核酸序列的引物对 cDNA 进行 PCR 试验。
15. 权利要求 14 的方法, 所述方法进一步包括(d)用探针检测 PCR
- 20 试验的产物。
16. 权利要求 15 的方法, 其中所述探针为具有 SEQ ID NO: 2476 核苷酸序列的核酸分子。
17. 一种试剂盒, 所述试剂盒在一个或多个容器中包含一种或多种分离的核酸分子, 所述核酸分子包含选自 SEQ ID NO: 2471、SEQ ID
- 25 NO: 2472 和 SEQ ID NO: 2473 的核苷酸序列。
18. 一种试剂盒, 所述试剂盒在一个或多个容器中包含一种或多种分离的核酸分子, 所述核酸分子包含选自 SEQ ID NO: 2474、SEQ ID NO: 2475 和 SEQ ID NO: 2476 的核苷酸序列。

对引起严重急性呼吸道综合症(SARS)的人病毒的诊断性试验

5 本申请要求以下申请的优先权：2003年3月24日提交的美国临时申请 60/457,031、2003年3月26日提交的美国临时申请 60/457,730、
2003年4月2日提交的美国临时申请 60/459,931、2003年4月3日
提交的美国临时申请 60/460,357、2003年4月8日提交的美国临时
申请 60/461,265、2003年4月14日提交的美国临时申请 60/462,805、
10 2003年4月23日提交的美国临时申请 60/464,886、2003年5月5日
提交的美国临时申请 60/468,139 和 2003年5月16日提交的美国临时
申请 60/471,200，它们各自通过引用整体结合到本文中。

 本申请包含一份长序列表，该序列表通过一式三份的 CD-R 代
替印刷的纸质文本一并提交，其通过引用整体结合到本文中。所述
15 CD-R 于 2004年3月22日刻录，分别标记为“CRF”、“Copy 1”
和“Copy 2”，每个只含有一个相同的 1.58 MB 文件(V9661078.APP)。

1. 发明领域

 本发明涉及对在人类中导致严重急性呼吸道综合症(SARS)的病
20 毒(“hSARS 病毒”)的诊断性试验。具体地说，本发明涉及使用逆转
录和聚合酶链反应(RT-PCR)，测定 hSARS 病毒、其天然或人工变体、
类似物或衍生物的定量试验方法。明确地说，所述定量试验是
TaqMan[®]试验。本发明还涉及包含核酸分子的诊断试剂盒，用于检测
hSARS 病毒。

25

2. 背景

 最近，中国大陆的广东省爆发了非典型肺炎。在 2002 年 11 月
到 2003 年 3 月之间，报道了 792 例病例，其中 31 例死亡(WHO. Severe
Acute Respiratory Syndrome (SARS) *Weekly Epidemiol Rec.* 2003; 78:

86)。为应对此次危机，香港医院管理局加强了对严重非典型肺炎患者的监视。在此次调查过程中，确认了多个卫生保健工作者患有此病。另外，与此病感染者密切接触的人中出现多个肺炎病例。尽管采用了针对已知通常与非典型肺炎有关的细菌病原体的典型抗生素治疗方法，此病的严重程度和发展仍非同寻常。本发明的发明者是参与调查这些患者的小组之一。在这些患者中，对常见病毒和细菌的所有鉴定试验结果均为阴性。该疾病按严重急性呼吸道综合症(Severe Acute Respiratory Syndrome)的首字母命名(“SARS”)。在本发明发明者从 SARS 患者分离出 hSARS 病毒之后，才了解了此病的病源因子。本发明提供了本文公开的快速和特异性实时定量 PCR 方法。本发明可用于临床和科学研究用途。

3. 发明概述

本发明涉及分离的 hSARS 病毒的序列信息在诊断方法中的用途。在一个优选实施方案中，分离的 hSARS 病毒保藏在 Genbank, NCBI, 其检索号为 AY278491 (SEQ ID NO: 15), 其通过引用结合到本文中。分离的 hSARS 病毒于 2003 年 4 月 2 日保藏在中国典型培养物中心(CCTCC), 被给予检索号 CCTCC-V200303, 如下面第 7 节中所述, 其通过引用结合到本文中。

在一个具体实施方案中，本发明提供对 hSARS 病毒、其天然或人工变体、类似物或衍生物的诊断性试验方法。具体地说，本发明涉及使用逆转录和聚合酶链反应(RT-PCR)，测定 hSARS 病毒核酸分子的定量试验方法。明确地说，所述定量试验是 TaqMan[®]试验。本发明中还提供适合与 hSARS 核酸杂交的核酸分子，例如包括但不限于 PCR 引物、逆转录酶引物、用于 Southern 分析或其它核酸杂交分析以检测 hSARS 核酸的探针。所述 hSARS 核酸包含或由核酸序列组成，所述核酸序列为 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序

列或其互补物、类似物、衍生物、片段或部分。在一个优选实施方案中，所述引物包含 SEQ ID NO: 2471 和/或 2472 核酸序列。在一个优选实施方案中，所述引物包含 SEQ ID NO: 2474 和/或 2475 核酸序列。在一个最优选的实施方案中，所述核酸分子包含 SEQ ID NO: 2473 核酸序列或其部分，在 RT-PCR 试验中使用包含 SEQ ID NO: 2471 和/或 2472 核酸序列的核酸分子作为引物可用于检测 hSARS 病毒。在另一个最优选实施方案中，所述核酸分子包含 SEQ ID NO: 2476 核酸序列或其部分，在 RT-PCR 试验中使用包含 SEQ ID NO: 2474 和/或 2475 核酸序列的核酸分子作为引物可用于检测 hSARS 病毒。在又一个最优选实施方案中，所述试验为 TaqMan[®] 定量试验。

在一个实施方案中，本发明提供在生物材料中检测 hSARS 病毒或其天然或人工变体、类似物、衍生物的存在或表达的方法，所述生物材料例如为细胞、血液、血清、血浆、唾液、尿、粪、痰、鼻咽抽吸物等。样品中 hSARS 病毒活性或表达相对对照样品的升高或降低可通过使生物材料接触可直接或间接检测 hSARS 病毒存在或表达的试剂而测定。在一个具体实施方案中，检测试剂为本发明的核酸分子。在另一个具体实施方案中，检测核酸分子固定化在 DNA 微阵列芯片上。

在一个具体实施方案中，本发明提供诊断试剂盒，其中包含适合用于检测 hSARS 病毒或其天然或人工变体、类似物、衍生物的核酸分子。在一个具体实施方案中，所述核酸分子具有 SEQ ID NO: 2471 和/或 2472 核酸序列。在具体实施方案中，所述核酸分子具有 SEQ ID NO: 2473 核酸序列。在另一个具体实施方案中，所述核酸分子具有 SEQ ID NO: 2474 和/或 2475 核酸序列。在具体实施方案中，所述核酸分子具有 SEQ ID NO: 2476 核酸序列。

在一个方面，本发明涉及分离的 hSARS 病毒在诊断方法中的用途。在一个具体实施方案中，本发明提供在生物材料中检测本发明 hSARS 病毒 mRNA 或基因组 RNA 的方法，所述生物材料例如为细

胞、血液、血清、血浆、唾液、尿、粪、痰、鼻咽抽吸物等。样品中 hSARS 病毒 mRNA 或基因组 RNA 相对对照样品的升高或降低可通过使生物材料接触可直接或间接检测 hSARS 病毒 mRNA 或基因组 RNA 的试剂而测定。在一个具体实施方案中，检测试剂为本发明的核酸分子。在另一个具体实施方案中，检测核酸分子固定化在 DNA 微阵列芯片上。

在另一方面，本发明涉及分离的 hSARS 病毒在诊断方法中的用途，例如在生物样品中检测免疫特异性结合 hSARS 病毒的抗体。在一个具体实施方案中，所述检测试剂为 hSARS 病毒，例如保藏号 CCTCC-V200303 的病毒，或具有 SEQ ID NO: 15 基因组核酸序列的病毒，或 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列编码的多肽。

在又一个方面，本发明提供抗体或其抗原结合片段，其免疫特异性结合 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核苷酸序列编码的本发明多肽、包含在严格条件下与 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核苷酸序列杂交的核苷酸序列的核酸编码的本发明多肽、和/或具有一种或多种本发明多肽生物活性的任何 hSARS 表位。这种抗体包括但不限于多克隆抗体、单克隆抗体、双特异性抗体、多特异性抗体、人抗体、人源化抗体、嵌合抗体、单链抗体、Fab 片段、F(ab')₂ 片段、二硫键连接的 Fv、胞内抗体以及含有特异性结合本发明多肽的 VL 或 VH 域乃至互补决定区(CDR)的片段。

本发明还涉及鉴别 hSARS 病毒或其天然或人工变体、类似物、衍生物感染对象的方法。在一个具体实施方案中，所述方法包括从来自对象的生物样品获得总 RNA；逆转录总 RNA 以获得 cDNA；和用一套来源于 hSARS 病毒核苷酸序列的引物使 cDNA 进行 PCR 测定。

本发明进一步涉及包含引物和核酸探针的诊断试剂盒，以检测 hSARS 病毒的 mRNA 或基因组 RNA。

3.1 定义

5 本文使用的术语“变体”指 hSARS 病毒的天然遗传变体或重组制备的 hSARS 病毒变体，与 CCTCC-V200303 的 hSARS 病毒相比，它们各自在其基因组中包含一种或多种突变。术语“变体”还可指特定肽的天然变体或重组制备的特定肽或蛋白的变体，其中一个或多个氨基酸残基通过氨基酸置换、插入或缺失被修饰。

10 本文使用的术语“类似物”对于非蛋白类似物指具有与第一种有机或无机分子相似或相同功能的第二种有机或无机分子，并且其结构与第一种有机或无机分子相似。

 本文使用的术语“衍生物”对于非蛋白类衍生物指基于第一种有机或无机分子的结构形成的第二种有机或无机分子。有机分子的衍生物包括但不限于例如通过加入或除去羟基、甲基、乙基、羧基或胺基而修饰的分子。有机分子也可酯化、烷基化和/或磷酸化。

15 本文使用的术语“突变体”指与野生型生物体相比核苷酸序列中存在突变的生物体。

 本文使用的术语“抗体”指单克隆抗体、双特异性抗体、多特异性抗体、人抗体、人源化抗体、嵌合抗体、骆驼源化(camelised)抗体、单域抗体、单链 Fv (scFv)、单链抗体、Fab 片段、F(ab')片段、二硫键连接的 Fv (sdFv)、抗独特型(抗 Id)抗体(包括例如抗本发明抗体的抗 Id 抗体)和上述任何抗体的表位结合片段。具体地说，抗体包括免疫球蛋白分子和免疫球蛋白分子的免疫活性片段(即包含抗原结合位点的分子)。免疫球蛋白分子可为任何型(例如 IgG、IgE、IgM、IgD、IgA 和 IgY)、类(例如 IgG1、IgG2、IgG3、IgG4、IgA1 和 IgA2)或亚类。

25 本文使用的术语“抗体片段”指免疫特异性结合 hSARS 病毒或

hSARS 病毒任何表位的抗体片段。抗体片段可通过任何本领域技术人员已知的技术产生。例如, Fab 和 F(ab')₂ 片段可通过使用酶如木瓜蛋白酶(产生 Fab 片段)或胃蛋白酶(产生 F(ab')₂ 片段)酶解免疫球蛋白分子产生。F(ab')₂ 片段含有完整的轻链和重链的可变区、CHI 区和铰链区。抗体片段还可通过重组 DNA 技术产生。抗体片段可为抗体的一个或多个互补决定区(CDR)。

本文使用的术语“免疫特异性结合本发明多肽的抗体或抗体片段”指抗体或其片段, 其免疫特异性结合 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其互补物、类似物、衍生物、片段或其部分编码的多肽, 或免疫特异性结合具有 SEQ ID NO: 2、12、14、17-239、241-736、738-1107、1109-1589、1591-1964 或 1966-2470 氨基酸序列或其变体、类似物、衍生物或片段的多肽, 并且不非特异性结合其它多肽。免疫特异性结合本发明多肽的抗体或其片段可与其它抗原交叉反应。优选免疫特异性结合本发明多肽的抗体或其片段不与其它抗原交叉反应。免疫特异性结合本发明多肽的抗体或其片段可通过例如免疫测定或其它本领域技术人员已知的技术鉴别。

本文使用的术语“表位”指在动物、优选哺乳动物、最优选人类中具有抗原性或免疫原性的 hSARS 病毒、多肽或蛋白的片段。具有免疫原性的表位是在动物中引起抗体应答的多肽的片段。具有抗原性的表位是抗体与其免疫特异性结合的多肽或蛋白的片段, 这可通过本领域已知的任何方法测定, 例如通过本文所述的免疫测定。抗原性表位不是必须有免疫原性。

本文使用的术语“抗原性”指物质(例如异物、微生物、药物、抗原、蛋白质、肽、多肽、核酸、DNA、RNA 等)在具体生物体、组织和/或细胞中引起免疫应答的能力。有时术语“抗原性的”与术语“免疫原性的”是同义词。

本文使用的术语“免疫原性”指物质(例如异物、微生物、药物、

抗原、蛋白质、肽、多肽、核酸、DNA、RNA 等)在生物体中引起免疫应答的特性。免疫原性部分取决于讨论的物质的大小,部分取决于该物质与宿主分子不相似的程度。高度保守的蛋白倾向具有较低免疫原性。

5 “分离的”核酸分子是与其天然来源中存在的其它核酸分子分离的核酸分子。此外,“分离的”核酸分子,例如 cDNA 分子,当通过重组技术生产时可基本不含其它细胞材料或培养基,当通过化学法合成时可基本不含化学前体或其它化学物质。在本发明的一个优选实施方案中,编码本发明多肽/蛋白质的核酸分子是分离的或纯化的。术语“分离的”核酸分子不包括属于文库的成员、还未与含有其它核酸分子的其它文库克隆分离的核酸。

10 本文使用的术语“在严格条件下杂交”描述相互之间具有至少 30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%或 95%一致性的核苷酸序列通常保持互相杂交的杂交和洗涤条件。这种杂交条件描述于例如但不限于 Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989),6.3.1-6.3.6.; Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., N.Y. (1986), 第 75-78 页和第 84-87 页;及 Molecular Cloning, Cold Spring Harbor Laboratory, N.Y. (1982), 第 387-389 页,对本领域技术人员来说是熟知的。严格杂交条件的优选非限制性实例是在约 68°C 下 6X 氯化钠/柠檬酸钠(SSC)、0.5% SDS 中杂交,然后在室温下 2X SSC、0.5% SDS 中洗涤一次或多次。严格杂交条件的另一个优选的非限制性实例是在约 45°C 下 6X SSC 中杂交,然后在约 50-65°C 下 0.2X SSC、0.1% SDS 中洗涤一次或多次。

25 “分离的”或“纯化的”肽或蛋白基本不含来自所述蛋白质的细胞来源或组织来源的细胞材料或其它污染蛋白质,或当它们通过化学法合成时,基本不含化学前体或其它化学物质。措词“基本不含细胞材料”包括多肽/蛋白质制品,其中多肽/蛋白质从其分离或重

组生产的细胞的细胞成分中分离出来。因此，基本不含细胞材料的多肽/蛋白质包括含少于约 30%、20%、10%、5%、2.5%或 1% (干重) 污染蛋白质的多肽/蛋白质制品。当多肽/蛋白质为重组生产时，还优选基本不含培养基，即培养基占蛋白质制品的体积少于约 20%、10% 或 5%。当多肽/蛋白质通过化学合成生产时，优选基本不含化学前体或其它化学物质，即与用于进行蛋白质合成的化学前体或其它化学物质分离。因此，这种多肽/蛋白质制品含少于约 30%、20%、10%、5% (干重) 不属于目的多肽/蛋白质片段的化学前体或化合物。在本发明优选的实施方案中，多肽/蛋白质是分离的或纯化的。

10 本文使用的术语“分离的”病毒是与其天然来源中存在的其它生物体分离的病毒，所述来源例如为生物材料如细胞、血液、血清、血浆、唾液、尿、粪、痰、鼻咽抽吸物等。分离的病毒可用于感染对象。

15 本文使用的术语“具有本发明多肽生物活性”指具有共同生物活性的多肽或蛋白质的特性，其与以下多肽相比具有类似或相同结构域和/或具有足够的氨基酸一致性：SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核苷酸序列或其互补物、类似物、衍生物、片段或部分编码的多肽；或具有 SEQ ID NO: 2、12、14、17-239、241-736、738-1107、1109-1589、1591-1964 或 1966-2470 氨基酸序列或其变体、类似物、20 衍生物或片段的多肽。本发明多肽的这种共同生物活性包括抗原性和免疫原性。

25 本文使用的术语“部分”或“片段”指含有相关核酸分子长度的至少约 25、30、35、40、45、50、60、70、80、90、100、150、200、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1,000、1,050、1,100、1,150、1,200、2,000、3,000、4,000、5,000、6,000、7,000、8,000、9,000、10,000、11,000、12,000、13,000、14,000、15,000、16,000、17,000、18,000、19,000、20,000、21,000、

22,000、23,000、24,000、25,000、26,000、27,000、28,000、29,000 或更多个连续核酸，且具有相关核酸分子的至少一种功能特征(或其编码的蛋白质具有相关核酸分子所编码的蛋白质的一种功能特征)的核酸分子片段；或指含有相关蛋白质或多肽长度的至少 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、90、100、120、140、160、180、200、220、240、260、280、300、320、340、360、400、500、600、800、1,000、2,000、3,000、4,000、5,000、6,000、7,000、8,000、9,000、9,500 或更多个氨基酸残基，且具有相关蛋白质或多肽的至少一种功能特征的蛋白质或多肽片段。

10 本文使用的术语“类似物”对于蛋白类物质(例如蛋白、多肽、肽和抗体)指蛋白类物质其具有与第二种蛋白类物质相似或相同的功能，但不是必须包含与第二种蛋白类物质相似或相同的氨基酸序列，或具有与第二种蛋白类物质相似或相同的结构。在一个具体实施方案中，抗体类似物与该类似物的来源抗体免疫特异性结合相同的表位。在另一个实施方案中，抗体类似物与该类似物的来源抗体免疫特异性结合不同的表位。具有相似氨基酸序列的蛋白类物质指满足至少一项以下条件的第二种蛋白类物质：(a)蛋白类物质其氨基酸序列与第二种蛋白类物质的氨基酸序列至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99% 相同；(b)蛋白类物质其编码核苷酸序列在严格条件下与核苷酸序列杂交，所述核苷酸序列编码第二种蛋白类物质的至少 5 个连续氨基酸残基、至少 10 个连续氨基酸残基、至少 15 个连续氨基酸残基、至少 20 个连续氨基酸残基、至少 25 个连续氨基酸残基、至少 40 个连续氨基酸残基、至少 50 个连续氨基酸残基、至少 60 个连续氨基酸残基、至少 70 个连续氨基酸残基、至少 80 个连续氨基酸残基、至少 90 个连续氨基酸残基、至少 100 个连续氨基酸残基、至少 125 个连续氨基酸残基或至少 150 个连续氨基酸残基；(c)蛋白类物质其

编码核苷酸序列与编码第二种蛋白类物质的核苷酸序列至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95%或至少 99%相同。与第二种蛋白类物质具有相似结构的蛋白类物质指与第二种蛋白类物质具有类似的二级、三级或四级结构的蛋白类物质。蛋白类物质的结构可通过本领域技术人员已知的方法测定，包括但不限于肽测序、X 射线晶体学、核磁共振、圆二色性和晶体电子显微镜。

为测定两种氨基酸序列或两种核酸序列的一致性百分比，为达到最佳比较目的而对齐序列(例如在第一种氨基酸序列或核酸序列中引入空位以与第二种氨基酸或核酸序列达到最佳对齐)。然后在相应的氨基酸位置或核苷酸位置比较氨基酸残基或核苷酸。当第一种序列的位置被与第二种序列相应位置相同的氨基酸或核苷酸占据时，所述分子在该位置是相同的。两种序列之间的一致性百分比是两种序列共有相同位置的数量的函数(即%一致性=相同重叠位置的数量/位置总数量×100%)。在一个实施方案中，两种序列具有相同长度。

两种序列之间一致性百分比的测定可使用数学算法完成。用于比较两种序列的数学算法的优选非限制性实例是 Karlin 和 Altschul, 1990, Proc. Natl. Acad. Sci. U.S.A. 87: 2264 2268 的算法，在 Karlin 和 Altschul, 1993, Proc. Natl. Acad. Sci. U.S.A. 90: 5873 5877 中被修改。该算法被结合到 Altschul 等, 1990, J. Mol. Biol. 215: 403 的 NBLAST 和 BLAST 程序中。BLAST 核苷酸检索可用 NBLAST 核苷酸程序进行，参数设定在例如分值=100，字长=12，以获得与本发明核酸分子同源的核苷酸序列。BLAST 蛋白检索可用 XBLAST 程序进行，参数设定在例如分值=50，字长=3，以获得与本发明蛋白质分子同源的氨基酸序列。为获得用于比较的有空位对齐，可使用 Altschul 等, 1997, Nucleic Acids Res. 25: 3389 3402 中所述的 Gapped BLAST。或者，可使用 PSI BLAST 进行检测分子间远缘关系(Id.)的叠代搜索。

当使用 BLAST、Gapped BLAST 和 PSI Blast 程序时,可使用各程序(例如 XBLAST 和 NBLAST)默认的参数(参见例如 NCBI 网站)。用于比较序列的数学算法的另一优选非限制性实例是 Myers 和 Miller, 1988, CABIOS 4: 11 17 的算法。该算法被结合到 ALIGN 程序(2.0 版)中, 5 其为 GCG 序列对齐软件包的部分。当使用 ALIGN 程序比较氨基酸序列时,可使用 PAM120 加权残基表、空位长度罚分 12 和空位罚分 4。

两种序列间一致性百分比可使用与上述技术相似的技术测定, 允许或不允许空位。在计算一致性百分比时,一般只计算完全匹配。

10 本文使用的术语“衍生物”对于蛋白类物质(例如蛋白质、多肽、肽和抗体)指包含通过引入氨基酸残基置换、缺失和/或插入而改变的氨基酸序列的蛋白类物质。本文使用的术语“衍生物”也指被修饰的蛋白类物质,即通过将任何类型的分子共价连接至蛋白类物质。例如但不以任何方式限制,抗体可被修饰,例如通过糖基化、乙酰化、PEG 化、磷酸化、酰胺化、通过已知保护/封闭基团的衍化、蛋 15 白水解、连接细胞配体或其它蛋白等。蛋白类物质的衍生物可通过使用本领域技术人员已知技术化学修饰而产生,所述技术包括但不限于特异性化学裂解、乙酰化、甲酰化、衣霉素的代谢合成等。此外,蛋白类物质的衍生物可包含一种或多种非典型氨基酸。蛋白类 20 物质的衍生物具有与其来源的蛋白类物质相似或相同的功能。

本文使用的术语“对象”和“患者”可互换使用。本文使用的术语“对象”指动物,优选包括非灵长目动物(例如牛、猪、马、山羊、绵羊、猫、狗、鸟类和啮齿类)和灵长目动物(例如猴如猕猴 (cynomolgous monkey)和人)的哺乳动物,更优选人。

25

4. 附图简述

图 1 显示从 SARS 病毒获得的部分 DNA 序列(SEQ ID NO: 1)及其推断的氨基酸序列(SEQ ID NO: 2),所述 SARS 病毒与已知冠状

病毒的 RNA 依赖性 RNA 聚合酶蛋白具有 57% 的同源性。

图 2 显示具有类似冠状病毒的形态学特征的新型 hSARS 病毒的电子显微照片。

图 3 显示与感染新型冠状病毒科(*Coronaviridae*)人呼吸道病毒的 FrHK-4 细胞结合的 IgG 抗体的免疫荧光染色。

图 4 显示生长于细胞培养物中、在 pH7.0 用 3% 磷钨酸钾负染色的 hSARS 病毒的超离心沉积物的电子显微照片。

图 5A 显示 SARS 患者的肺活组织切片检查的薄层切片电子显微照片；图 5B 显示 hSARS 病毒感染细胞的薄层切片电子显微照片。

图 6 显示 hSARS 病毒(GenBank 检索号 AY268070)的部分蛋白序列(215 氨基酸; SEQ ID NO: 2)的系统发生学分析结果。系统树是通过邻接法构建的。水平线距离表示两个相比较的序列不同的位点的数目。自引导值是从 500 个重复中推断得出的。

图 7A 显示在能定量检测样品中 hSARS 病毒的实时定量 PCR 试验中荧光强度对 PCR 循环的扩增图。指出了反应中输入质粒 DNA 的拷贝数。X 轴表示定量 PCR 试验的循环数, Y 轴表示超过背景的荧光强度(FI)。图 7B 显示临床样品的 PCR 产物的解链曲线分析结果。指出了得自阳性(+ve)样品、阴性(-ve)样品和水对照(水)的信号。X 轴表示温度(°C), Y 轴表示超过背景的荧光强度(FI)。

图 8 显示另一种从 hSARS 病毒获得的部分 DNA 序列(SEQ ID NO: 11)及其推断的氨基酸序列(SEQ ID NO: 12)。

图 9 显示又一种从 hSARS 病毒获得的部分 DNA 序列(SEQ ID NO: 13)及其推断的氨基酸序列(SEQ ID NO: 14)。

图 10 显示 hSARS 病毒的完整基因组 DNA 序列(SEQ ID NO: 15)。

图 11 显示以三种读框(参见 SEQ ID NO: 16、240 和 737)从 SEQ ID NO: 15 获得的推断氨基酸序列。星号(*)表示标志肽的末端的终止密码子。第一读框氨基酸序列: SEQ ID NO: 17-239; 第二读框氨

氨基酸序列: SEQ ID NO: 241-736; 第三读框氨基酸序列: SEQ ID NO: 738-1107。

图 12 显示以三种读框(参见 SEQ ID NO: 1108、1590 和 1965)从 SEQ ID NO: 15 的互补物获得的推断氨基酸序列。星号(*)表示标志肽的末端的终止密码子。第一读框氨基酸序列: SEQ ID NO: 1109-1589; 第二读框氨基酸序列: SEQ ID NO: 1591-1964; 第三读框氨基酸序列: SEQ ID NO: 1966-2470。

图 13 显示正向引物(SEQ ID NO: 2471 和 2474)、反向引物(SEQ ID NO: 2472 和 2475)和杂交探针(SEQ ID NO: 2473 和 2476)的核酸序列, 用于定量 TaqMan[®] 试验以测定 hSARS 病毒。

图 14 显示用于实时定量 RT-PCR 试验 SAS-CoV 的标准曲线。阈循环数(Ct)是反应荧光强度达到预定阈值所需的 PCR 循环数。Ct 与质粒 DNA 起始浓度的对数成反比。给出了相关系数。在标准扩增图中通过最大曲率法对不同起始拷贝数基于计算阈值计算 Ct。X-轴表示标准品拷贝数的对数, Y-轴表示 Ct。

图 15 显示对于从 SARS 患者分离的 NPA 样品, 使用本发明的修改 RT-PCR 检测方法, 将荧光强度对 PCR 循环数作的典型扩增图。通过修改的 RNA 提取方法, 50 NPA 样品中的 40 个在实时试验中是阳性的。对于在第一代 RT-PCR 试验中是阴性的样品, 通过本发明的检测方法发现其中都含有很少量的病毒 RNA。X-轴表示 PCR 循环数, Y-轴表示超过背景的荧光强度(ΔR_n)。

图 16 显示临床样品中 SARS-CoV 的病毒负载量与开始日期的关系图。结果显示, 病毒负载量随疾病进展而增加。某些在第一代试验中为阳性的样品被发现含有很大的病毒 RNA。X-轴表示开始日期, Y-轴表示样品中每个反应的拷贝数。

5. 发明详述

本发明涉及分离的 hSARS 病毒的序列信息在诊断方法中的用

途。具体地说,本发明提供在生物样品中检测 hSARS 病毒或其天然或人工变体、类似物或衍生物的核酸分子存在与否的方法。所述方法包括从各种来源获得生物样品,使样品与可检测 hSARS 病毒或其天然或人工变体、类似物、衍生物的核酸(例如 mRNA、基因组 RNA)的化合物或试剂接触,从而检测样品中 hSARS 病毒或其天然或人工变体、类似物、衍生物的存在。优选检测 hSARS mRNA 或基因组 RNA 的试剂是能与 mRNA 或基因组 RNA 杂交的标记核酸探针。在一个优选实施方案中,所述核酸探针为包含或由 SEQ ID NO: 2473 或 2476 核酸序列或其部分组成的核酸分子,其足以在严格条件下特异性与 hSARS mRNA 或基因组 RNA 杂交。在一个优选的具体实施方案中,在样品中通过逆转录聚合酶链反应(RT-PCR),使用基于部分 hSARS 病毒核苷酸序列构建的引物,检测 hSARS 病毒或其天然或人工变体、类似物或衍生物的存在。在一个非限制性具体实施方案中,优选 RT-PCR 方法中使用的引物为:

5'-CAGAACGCTGTAGCTTCAAAAATCT-3' (SEQ ID NO: 2471)和 5'-TCAGAACCCTGTGATGAATCAACAG-3' (SEQ ID NO: 2472), 在 $MgCl_2$ 存在下,热循环例如为但不限于 $50^{\circ}C$ 2 分钟、 $95^{\circ}C$ 10 分钟,然后 $95^{\circ}C$ 15 秒、 $60^{\circ}C$ 1 分钟的循环 45 个(也参见以下 6.7、6.8、6.9 节)。在优选实施方案中,所述引物包含 SEQ ID NO: 2471 和 2472 核酸序列。在另一个非限制性具体实施方案中,优选 RT-PCR 方法中使用的引物为: 5'-ACCAGAATGGAGGACGCAATG-3' (SEQ ID NO: 2474)和 5'-GCTGTGAACCAAGACGCAGTATTAT-3' (SEQ ID NO: 2475), 在 $MgCl_2$ 存在下,热循环例如为但不限于 $50^{\circ}C$ 2 分钟、 $95^{\circ}C$ 10 分钟,然后 $95^{\circ}C$ 15 秒、 $60^{\circ}C$ 1 分钟的循环 45 个(也参见以下 6.7、6.8、6.9 节)。在优选实施方案中,所述引物包含 SEQ ID NO: 2474 和 2475 核酸序列。

本发明方法可包括实时定量 PCR 试验。在一个优选实施方案中,本发明中使用的定量 PCR 为 TaqMan[®] 试验(Holland 等, *Proc Natl Acad*

5 *Sci USA* 88(16): 7276 (1991))。该测定可在用于进行这种试验的仪器上进行，例如可获自 Applied Biosystems (Foster City, CA)的仪器。在更优选的具体实施方案中，本发明提供实时定量 PCR 试验，通过使来自样品的提取总 RNA 逆转录，将获得的 cDNA 用特异性引物进行 PCR 反应，使用探针检测扩增产物的存在，从而在生物样品中检测 hSARS 病毒或其天然或人工变体、类似物或衍生物的存在。在优选实施方案中，所述探针为 TaqMan[®]探针，其由带 5'-报道染料和 3'-猝灭染料的寡核苷酸组成。在一个优选实施方案中，所述探针具有 5'- TCTGCGTAGGCAATCC-3' (SEQ ID NO: 2473)的核苷酸序列。10 在另一个优选实施方案中，所述探针具有 5'-ACCCCAAGGTTTACC-3' (SEQ D NO: 2476)的核苷酸序列。荧光报道染料，例如 FAM[®]染料，与寡核苷酸探针的 5'端共价连接。其它染料，例如 TET[®]染料或 VIC 可用作报道染料。各报道染料被 3'端 TAMRA[®]染料或非荧光猝灭剂猝灭。在一个优选实施方案中，3'端用 NFQ-MGB 标记。由于经过一定数量的热循环产生了 PCR 产物，在延长步骤结束时捕获这些反应的荧光信号，因此可基于扩增图定量测定样品中病毒负载量。

也可使用其它检测 RNA 的技术。例如，检测 mRNA 的体外技术包括 Northern 杂交、原位杂交、RT-PCR 和 RNA 酶保护。检测基因组 RNA 的体外技术包括 Northern 杂交、RT-PCT 和 RNA 酶保护。

20 如上所述，在一个优选实施方案中，hSARS 病毒的多核苷酸可在其被检测前扩增。术语“扩增”指从单个多核苷酸分子制备该核酸多个拷贝的过程。多核苷酸的扩增可在体外通过本领域技术人员已知的生物化学方法进行。扩增试剂可为任何可完成引物延长产物合成功能的化合物或体系，包括酶。用于该目的合适酶包括例如 25 大肠杆菌 DNA 聚合酶 I、Taq 聚合酶、大肠杆菌 DNA 聚合酶 I 的 Klenow 片段、T4 DNA 聚合酶、其它可获得的 DNA 聚合酶、聚合酶突变蛋白、逆转录酶、连接酶和包括热-稳定酶(即在升高到足以导致变性的温度后进行引物延长的酶)的其它酶。合适的酶促进核苷酸以合适的

方式结合，以形成与各突变体核苷酸链互补的引物延长产物。在一个优选实施方案中，所述酶为获自 Applied Biosystems 的 AmpliTaq Gold® DNA 聚合酶。通常，该合成从各引物的 3'-端开始，沿模板链向 5'-方向进行，直到合成结束，产生不同长度的分子。但是，某些
5 扩增试剂使用与上述相同的过程，在 5'-端开始合成，向另一方向进行。在任何情况下，本发明的方法不限于本文所述扩增的具体实施方案。

本发明可使用的一种体外扩增方法是美国专利 4,683,202 和 4,683,195 中所述的聚合酶链反应(PCR)。术语“聚合酶链反应”指使用
10 热稳定 DNA 聚合酶和两种寡核苷酸引物扩增 DNA 碱基序列的方法，所述引物一种在待扩增序列的一端与(+)链互补，另一种在另一端与(-)链互补。因为新合成的 DNA 链可随后作为相同引物序列的额外模板，引物退火、链延长和解离的连续循环使所需序列产生快速和高特异性扩增。聚合酶链反应用于在样品中检测编码细胞因子的
15 多核苷酸的存在。许多聚合酶链方法是本领域技术人员已知的，可用于本发明的方法。例如，DNA 可在热循环仪中经历 30-35 个下述扩增循环：95°C 30 秒、52-60°C 1 分钟和 72°C 1 分钟，最终延长步骤为 72°C 5 分钟。又例如，DNA 可在热循环仪中经历 35 个聚合酶链反应循环，变性温度为 95°C 30 秒，然后可变退火温度 54-58°C 1
20 分钟，延长步骤为 70°C 1 分钟，最终延长步骤为 70°C 5 分钟。

用于扩增 hSARS 病毒 mRNA 或基因组 RNA 的引物可使用任何合适的方法制备，例如传统的磷酸三酯和磷酸二酯方法或其自动化
25 实施方案，只要引物能与目标多核苷酸杂交。一种在修饰的固体载体上合成寡核苷酸的方法如美国专利 4,458,066 中所述。引物的准确长度取决于许多因素，包括温度、缓冲液和核苷酸组成。引物必须在扩增诱导试剂存在下启动延长产物的合成。

本发明方法使用的引物与待扩增核苷酸序列的各链互补。术语“互补”指引物必须在允许试剂进行聚合的条件下与它们各自的链

杂交。换句话说，与侧面序列互补的引物与侧面序列杂交，允许扩增核苷酸序列。优选延续的引物 3'端具有与侧面链互补的完全碱基配对互补性。可使用已知方法结合本公开来开发用于 hSARS 病毒多核苷酸的引物和探针。在优选实施方案中，引物根据 TaqMan[®]引物方案(Applied Biosystems)设计。所述引物使用 Primer Express User Bulletin (Applied Biosystems)中所述的 Primer Express 软件设计。简单地说，当设计引物时，其应在探针后选择。引物优选与探针尽可能接近而不与探针重叠。引物的 G-C 含量应在 20%-80%的范围内。优选避免多个相同的核苷酸。尤其是对于鸟嘌呤，优选避免 4 个或更多连续的 G。各引物的解链温度优选为 58°C-60°C。各引物 3'端的 5 个核苷酸中优选 G 和/或 C 碱基不超过 2 个。

可使用 Primer Express User Bulletin (P/N 4317594) (Applied Biosystems)中所述的 Primer Express 软件设计探针。简单地说，优选保持 G-C 含量在 20%-80%的范围内。优选避免多个相同的核苷酸。尤其是对于鸟嘌呤，优选避免 4 个或更多连续的 G。优选不将 G 碱基放在 5'端。优选选择使探针中 C 超过 G 的链。优选两种探针在相同链上。对于单探针试验，解链温度优选为 68°C-70°C。

本领域普通技术人员了解各种可用于增加目标核酸拷贝数的扩增方法。本发明方法中检测的多核苷酸可在溶液或结合到固体载体后进一步被评价、检测、克隆、测序等，这可通过任何通常用于检测具体核苷酸序列的方法，如另一聚合酶链反应、寡聚物限制性酶切(Saiki 等, *Bio/Technology* 3: 1008-1012 (1985))、等位基因特异性寡核苷酸(ASO)探针分析(Conner 等, *Proc. Natl. Acad. Sci. USA* 80: 278 (1983)、寡核苷酸连接试验(OLA) (Landegren 等, *Science* 241: 1077 (1988))、RNA 酶保护试验等。已有 DNA 分析的分子技术的综述(Landegren 等, *Science* 242: 229-237 (1988))。DNA 扩增后，反应产物通过 DNA 杂交分析检测，而不使用放射性探针。在这种方法中，例如，含从组织或对象获得的多核苷酸的少量 DNA 样品被扩增，通

过 DNA 印迹技术分析。高度放大的信号促进了非放射性探针或标记的使用。本发明一个实施方案中，一种三磷酸核苷被放射性标记，从而允许通过放射自显影法直接显示扩增产物。在另一个实施方案中，扩增引物被荧光标记，电泳通过电泳体系。通过激光检测后计算机辅助图形显示而不是放射性信号，显示扩增产物。

用于扩增 hSARS 病毒 mRNA 或基因组 RNA 的部分的引物，其大小为至少 10、15、20、25 或 30 个核苷酸长度。优选 GC 比例应高于 30%、35%、40%、45%、50%、55% 或 60%，以防止引物上的发夹结构。此外，扩增子应有足够的长度以通过标准分子生物学方法检测。优选扩增子长度为至少 20、30、40、50、60、70、80、90、100、110、120、130、140、150、175、200、250、300、350、400、450、500、550、600、700、800 或 1000 碱基对。

在一个具体实施方案中，本发明方法进一步涉及从对照对象获得对照样品，将对照样品与能检测样品中 mRNA 或基因组 RNA 存在的化合物或试剂接触，并将对照样品中 mRNA 或基因组 RNA 存在与待测样品中 mRNA 或基因组 DNA 存在相比较。

本发明还包括在待测样品中检测 hSARS 病毒核酸存在的试剂盒。该试剂盒例如可包含能在待测样品中检测核酸分子的标记化合物或试剂，在某些实施方案中包含检测样品中 mRNA 量的工具(结合 DNA 或 mRNA 的寡核苷酸探针)。

对于基于寡核苷酸的试剂盒，该试剂盒可例如包含：(1)寡核苷酸，例如可检测的标记寡核苷酸，其与 hSARS 病毒的核苷酸序列杂交；和/或(2)用于扩增含 hSARS 病毒序列的核酸分子的引物对。该试剂盒还可包含，例如，缓冲试剂、防腐剂或蛋白稳定试剂。该试剂盒还可包含检测可检测物必须的成分(例如酶或底物)。该试剂盒还可包含对照样品或一系列对照样品，其可被检测并与待测样品比较。该试剂盒的各种成分通常封装入单独的容器，所有各种容器和使用说明一起封在单个包装物内。

5.1 hSARS 病毒的核酸序列

本发明涉及分离的病毒的序列信息在诊断和治疗方法中的用途。hSARS 病毒 CCTCC-V200303 的完整基因组序列公开于 2004 年 3 月 24 日与本文同时提交的美国专利申请，律师档案号 V9661.0069，其通过引用整体结合到本文中。在一个具体实施方案中，本发明提供 hSARS 病毒 CCTCC-V200303 的完整核苷酸序列 SEQ ID NO: 15 或其互补物、类似物、衍生物、片段或部分。此外，本发明涉及在严格条件下与 hSARS 病毒 CCTCC-V200303 基因组 SEQ ID NO: 15 的任何部分杂交的核酸分子。在一个具体实施方案中，本发明提供适合用作引物的核酸分子，该核酸分子包含或由 SEQ ID NO: 1、3、4、11 或 13 核苷酸序列或其互补物、类似物、衍生物、片段或部分组成。在优选的具体实施方案中，所述引物包含 SEQ ID NO: 2471、2472、2474 或 2475 核酸序列。在另一个具体实施方案中，本发明提供适合用作杂交探针的核酸分子，用于检测包含或由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其互补物、类似物、衍生物、片段或部分组成的、编码本发明多肽的核酸。在另一个实施方案中，本发明涉及包含其核酸序列为 SEQ ID NO: 2471 和/或 2472 的引物的试剂盒，以检测 hSARS 病毒或其天然或人工变体、类似物或衍生物。在一个优选实施方案中，所述试剂盒还包含具有 SEQ ID NO: 2473 核酸序列的探针。在另一个实施方案中，本发明涉及包含其核酸序列为 SEQ ID NO: 2474 和/或 2475 的引物的试剂盒，以检测 hSARS 病毒或其天然或人工变体、类似物或衍生物。在一个优选实施方案中，所述试剂盒还包含具有 SEQ ID NO: 2476 核酸序列的探针。在另一个优选实施方案中，所述试剂盒还包含试剂，用于检测作为阴性对照的 hSARS 病毒中不存在的基因。本发明进一步包括嵌合病毒或重组病毒或由所述核苷酸序列编码的病毒蛋白质。

本发明也涉及 hSARS 病毒的分离的核酸分子，所述核酸分子包含或由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其互补物、类似物、衍生物、片段或部分组成。在另一个具体实施方案中，本发明提供在本文定义的严格条件下与具有 SEQ ID NO: 1、11、15、13、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列的核酸分子、或冠状病毒科已知成员的特定基因、或其互补物、类似物、衍生物、片段或部分杂交的分离核酸分子。在另一个具体实施方案中，本发明提供核酸分子编码的分离多肽或蛋白质，所述核酸分子包含 SEQ ID NO: 1 核酸序列或其互补物、类似物、衍生物或片段的至少约 5、10、15、20、25、30、35、40、45、100、150、200、300、350、400、450、500、550、600 或更多个连续核苷酸的核苷酸序列。在另一个具体实施方案中，本发明提供核酸分子编码的分离多肽或蛋白质，所述核酸分子包含 SEQ ID NO: 11 核酸序列或其互补物、类似物、衍生物或片段的至少约 5、10、15、20、25、30、35、40、45、100、150、200、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200 或更多个连续核苷酸的核苷酸序列。在又一个具体实施方案中，本发明提供核酸分子编码的分离多肽或蛋白质，所述核酸分子包含 SEQ ID NO: 13 核酸序列或其互补物、类似物、衍生物或片段的至少约 5、10、15、20、25、30、35、40、45、100、150、200、300、350、400、450、500、550、600、650、700 或更多个连续核苷酸的核苷酸序列。在又一个具体实施方案中，本发明提供核酸分子编码的分离多肽或蛋白质，所述核酸分子包含 SEQ ID NO: 15 核酸序列或其互补物、类似物、衍生物或片段的至少约 5、10、15、20、25、30、35、40、45、100、150、200、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、2000、3000、4000、5000、6000、7000、8000、

9000、10000、11000、12000、13000、14000、15000、16000、17000、18000、19000、20000、21000、22000、23000、24000、25000、26000、27000、28000、29000 或更多个连续核苷酸的核苷酸序列。多肽包括图 11 (SEQ ID NO: 17-239、241-736 和 738-1107)和 12 (SEQ ID NO: 1109-1589、1591-1964 和 1966-2470)中所示的多肽。本发明多肽或蛋白优选具有以下蛋白质的一种或多种生物活性：由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列编码的蛋白质；或含有由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列编码的氨基酸序列的天然病毒蛋白质。

本发明进一步提供能特异性结合由 SEQ ID NO: 1、11、13、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其片段编码的本发明多肽或任何 hSARS 表位的抗体。本发明还提供特异性结合 SEQ ID NO: 15 核酸序列或其片段编码的本发明多肽或任何 hSARS 表位的抗体。这种抗体包括但不限于多克隆抗体、单克隆抗体、双特异性抗体、多特异性抗体、人抗体、人源化抗体、嵌合抗体、单链抗体、Fab 片段、F(ab')₂ 片段、二硫键连接的 Fv、胞内抗体和含有能与本发明多肽特异性结合的 VL 或 VH 结构域乃至互补决定区(CDR)的片段。

在另一个实施方案中，本发明提供包含 hSARS 病毒或其天然或人工变体、类似物或衍生物的疫苗制品。在又一个实施方案中，本发明提供包含 hSARS 病毒的重组形式和嵌合形式或病毒亚单位的疫苗制品。在一个具体实施方案中，本发明疫苗制品包含活的但减毒的 hSARS 病毒，有或没有药物可接受赋形剂，包括佐剂。在另一个具体实施方案中，本发明疫苗制品包含灭活或死的 hSARS 病毒，有或没有药物可接受载体，包括佐剂。本发明疫苗制品还可包含佐剂。因此，本发明进一步提供制备 hSARS 病毒的重组形式或嵌合形式的

方法。在另一个具体实施方案中，本发明疫苗制品包含一种或多种包含或由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其片段组成的核酸分子。在另一个实施方案中，本发明提供包含一种或多种
5 本发明多肽的疫苗制品，所述多肽由包含或由 SEQ ID NO: 1、11、13、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列或其片段组成的核苷酸序列编码。在另一个实施方案中，本发明提供包含一种或多种本发明多肽的疫苗制品，所述多肽由包含或由 SEQ ID NO: 15 核酸序列或其片段组成的核苷酸
10 序列编码。此外，本发明提供通过单独或与以下药物联合给予本发明疫苗制品或抗体来治疗、改善、控制或预防 SARS 的方法：抗病毒剂[例如金刚烷胺、金刚乙胺、更昔洛韦、阿昔洛韦、利巴韦林、喷昔洛韦、奥塞米韦、膦甲酸、齐多夫定(AZT)、去羟肌苷(ddI)、拉米夫定(3TC)、扎西他滨(ddC)、司他夫定(d4T)、奈韦拉平、地位韦啉、茚地那韦、利托那韦、阿糖腺苷、奈非那韦、沙奎那韦、扎那米韦、磷酸奥塞米韦、普来可那立、干扰素等]、类固醇和皮质类固醇(如泼尼松、可的松、氟替卡松)和糖皮质激素、抗生素、止痛剂、
15 支气管扩张药或其它用于呼吸道和/或病毒感染的治疗药物。

此外，本发明提供包含本发明抗病毒剂和药物可接受载体的药物组合物。本发明还提供包含本发明药物组合物的药盒。
20

在另一方面，本发明提供筛选能抑制 hSARS 病毒或其天然或人工变体、类似物或衍生物传染性或复制的抗病毒剂的方法。

在一个实施方案中，本发明提供在生物材料如细胞、血液、血清、血浆、唾液、尿、粪、痰、鼻咽抽吸物等中检测本发明 hSARS
25 病毒或其天然或人工变体、类似物或衍生物存在、活性或表达的方法。可通过使生物材料接触可直接或间接检测 hSARS 病毒或其天然或人工变体、类似物或衍生物的存在试剂，测定样品中 hSARS 病毒或其天然或人工变体、类似物或衍生物的存在。在一个具体实施

方案中，检测试剂为本发明的抗体。在另一个实施方案中，检测试剂为本发明的核酸。

5.2. hSARS 病毒

5 5.2.1. hSARS 病毒的天然变体

本发明基于本发明发明人对来自患 SARS 对象的新型病毒的分离和鉴定。分离的 hSARS 病毒于 2003 年 4 月 2 日保藏在中国典型培养物中心(CCTCC)，被给予检索号 CCTCC-V200303。本发明也涉及保藏检索号 CCTCC-V200303 的 hSARS 病毒的天然变体。

10 hSARS 病毒的天然变体具有与 hSARS 病毒基因组序列不同的序列，这是由于基因组序列的一种或多种自发突变，包括但不限于点突变、重排、插入、缺失等，所述突变会或不会导致发生表型改变。优选所述变体相对于 hSARS 病毒包含少于 25、20、15、10、5、4、3、或 2 个氨基酸的置换、重排、插入和/或缺失。

15 保守或非保守氨基酸置换可在一个或多个氨基酸残基进行。在优选实施方案中，变体在一个或多个预测的非必需氨基酸残基(即对于病毒生物活性如感染性、复制能力、蛋白合成能力、装配能力和致细胞病变效应的表达不是决定性的氨基酸残基)上具有保守氨基酸置换。在其它实施方案中，变体在一个或多个预测的非必需氨基酸

20 残基(即对于病毒生物活性如感染性、复制能力、蛋白合成能力、装配能力和致细胞病变效应的表达不是决定性的氨基酸残基)上具有非保守氨基酸置换。

“保守氨基酸置换”是其中氨基酸残基被侧链电荷相似的氨基酸残基置换。“非保守氨基酸置换”是其中氨基酸残基被侧链电荷

25 相反的氨基酸残基置换。侧链电荷相似的氨基酸残基的家族在本领域是已定义的。遗传编码的氨基酸可被分为四个家族：(1)酸性=天冬氨酸、谷氨酸；(2)碱性=赖氨酸、精氨酸、组氨酸；(3)非极性=丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、

色氨酸; (4)无电荷极性=甘氨酸、天冬酰胺、谷氨酰胺、半胱氨酸、丝氨酸、苏氨酸、酪氨酸。以相似的方式,全部氨基酸可被分类为(1)酸性=天冬氨酸、谷氨酸; (2)碱性=赖氨酸、精氨酸、组氨酸; (3)脂族=甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、丝氨酸、苏氨酸, 5 丝氨酸和苏氨酸任选被单独分类为脂族-羟基; (4)芳族=苯丙氨酸、酪氨酸、色氨酸; (5)酰胺=天冬酰胺、谷氨酰胺; (6)含硫=半胱氨酸和甲硫氨酸。(参见例如 Biochemistry, 第4版, L. Stryer, WH Freeman 和 Co 编: 1995)。

本发明进一步涉及 hSARS 病毒突变体。在一个实施方案中,突 10 变可沿 hSARS 病毒或其变体的所有或部分编码序列随机引入,例如通过饱和诱变,可筛选所得突变体的生物活性以鉴定保留活性的突变体。也可使用本领域已知的诱变技术,包括但不限于定点突变、化学诱变、体外定点诱变,使用例如 QuikChange 定点诱变试剂盒 (Stratagene)等。这种修饰的非限制性实例包括氨基酸置换为半胱氨酸 15 而形成二硫键; 氨基酸置换为酪氨酸并随后化学处理多肽以形成二酪氨酸键,如本文中详细公开的; 一个或多个氨基酸置换和/或生物或化学修饰以产生小分子(底物或抑制剂)结合袋; 和/或引入侧链特异性标记(例如以鉴别分子相互作用或捕获蛋白质-蛋白质相互作用伴侣 (protein-protein interaction partner))。在一个具体实施方案中,生物修 20 饰包括烷基化、磷酸化、硫酸化、氧化或还原、ADP-核糖基化、羟基化、糖基化、葡糖磷脂酰肌醇加成、泛素化。在另一个具体实施方案中,化学修饰包括改变重组病毒的电荷。在又一个实施方案中,正电荷或负电荷被化学加入到氨基酸残基中,其中带电荷的氨基酸残基被修饰成不带电荷的残基。

25 5.2.2. 重组和嵌合 hSARS 病毒

本发明还包括由来源于 hSARS 病毒或其天然变体的基因组的病毒载体编码的重组或嵌合病毒。在一个具体实施方案中,重组病毒来源于保藏号为 CCTCC-V200303 的 hSARS 病毒。在一个具体实施

方案中，病毒具有 SEQ ID NO: 15 核酸序列。在另一个具体实施方案中，重组病毒来源于 hSARS 病毒的天然变体。hSARS 病毒的天然变体具有与 hSARS 病毒 CCTCC-V200303 的基因组序列(SEQ ID NO: 15)不同的序列，这是由于基因组序列的一种或多种自发突变，包括但不限于点突变、重排、插入、缺失、置换等，所述突变会或不会导致发生表型改变。根据本发明，来源于 hSARS 病毒 CCTCC-V200303 基因组的病毒载体含有编码至少 hSARS 病毒一个 ORF 的一部分的核酸序列。在一个具体实施方案中，所述 OFR 包含或由 SEQ ID NO: 1、11 或 13 核酸序列或其片段组成。在一个具体实施方案中，在 SEQ ID NO: 15 核苷酸序列或其片段中存在多于一个的 OFR，如图 11 (SEQ ID NO: 16、240 和 737)及图 12 (SEQ ID NO: 1108、1590 和 1965)中所示。在另一实施方案中，ORF 编码的多肽包含或由 SEQ ID NO: 2、12 或 14 氨基酸序列或其片段、或如图 11 (参见 SEQ ID NO: 17-239、241-736 或 738-1107)及图 12 (参见 SEQ ID NO: 1109 -1589、1591-1964 或 1966-2470)中所示多肽或其片段组成。根据本发明，这些病毒载体可包含或不包含不是天然病毒基因组的核酸。

在另一个具体实施方案中，本发明嵌合病毒是进一步包含异源核苷酸序列的重组 hSARS 病毒。根据本发明，嵌合病毒可由核苷酸序列编码，其中已向基因组添加了异源核苷酸序列或其中内源或天然核苷酸序列已被异源核苷酸序列置换。

根据本发明，嵌合病毒由进一步包含异源核苷酸序列的本发明病毒载体编码。根据本发明，嵌合病毒由可包含或不包含不是天然病毒基因组的核酸的病毒载体编码。根据本发明，嵌合病毒由病毒载体编码，其中已添加、插入异源核苷酸序列或已置换天然或非天然序列。根据本发明，嵌合病毒可由来源于 hSARS 病毒的不同毒株或变体的核苷酸序列编码。具体地说，嵌合病毒由核苷酸序列编码，所述核苷酸序列编码来源于 hSARS 病毒的不同毒株或变体的抗原多肽。

嵌合病毒对于生产抗两种或多种病毒的重组疫苗是特别有用的 (Tao 等, *J. Virol.* 72, 2955-2961; Durbin 等, 2000, *J. Virol.* 74, 6821-6831; Skiadopoulos 等, 1998, *J. Virol.* 72, 1762-1768; Teng 等, 2000, *J. Virol.* 74, 9317-9321)。例如, 可以设想, 来源于 hSARS 病毒、表达 hSARS 病毒变体的一种或多种蛋白质(反之亦然)的病毒载体将保护接种了这种载体的对象免受天然 hSARS 病毒及其变体的感染。对于用活疫苗接种的目的, 可以象其它病毒一样使用减毒和复制缺陷型病毒。(参见 PCT WO 02/057302 第 6 页和第 23 页, 其通过引用结合到本文中)。

10 根据本发明, 待掺入编码本发明重组或嵌合病毒的病毒载体的异源序列包括从 hSARS 的不同毒株或变体获得或衍生的序列。

在某些实施方案中, 本发明嵌合或重组病毒由来源于病毒基因组的病毒载体编码, 其中一个或多个序列、基因间区、末端序列或 ORF 的整体或部分已被异源或非天然序列取代。在本发明的某些实施方案中, 本发明嵌合病毒由来源于病毒基因组的病毒载体编码, 其中一种或多种异源序列已被插入或添加到载体中。

病毒载体的选择可取决于待治疗病毒感染或保护免受病毒感染的对象的种类。如果对象是人类, 则可用减毒 hSARS 病毒来提供抗原序列。

20 根据本发明, 可对病毒载体进行人工改造, 以提供对 hSARS 病毒、其天然或人工变体、类似物或衍生物的感染带来保护作用的抗原序列。可对病毒载体进行人工改造, 以提供一种、两种、三种或更多种抗原序列。根据本发明, 抗原序列可来源于同一病毒、同一种类病毒的不同毒株或变体、或不同病毒。

25 根据本发明获得的表达产物和/或重组或嵌合病毒颗粒可有利地应用在疫苗制品中。可对本发明表达产物和嵌合病毒颗粒进行人工改造, 以产生抗多种病原体的疫苗, 所述病原体包括病毒和细菌抗原、肿瘤抗原、变应原抗原和与自身免疫病有关的自身抗原。具体

地说, 可对本发明嵌合病毒颗粒进行人工改造, 以产生能保护对象免受 hSARS 病毒或其天然或人工变体、类似物或衍生物感染的疫苗。

在某些实施方案中, 可对本发明表达产物和重组或嵌合病毒颗粒进行人工改造, 以产生抗多种病原体的疫苗, 所述病原体包括病毒抗原、肿瘤抗原和与自身免疫病有关的自身抗原。实现此目标的一个方法包括对现有的 hSARS 基因进行修饰, 使得在所述基因各自的外结构域中包含外源序列。在异源序列为病原体的表位或抗原的情况下, 这些嵌合病毒可用来诱导针对衍生这些决定簇的疾病因子的保护性免疫应答。

因此, 本发明涉及使用病毒载体和重组或嵌合病毒来制备抗多种病毒和/或抗原的疫苗。本发明还包括包含病毒载体的重组病毒, 所述病毒载体来源于 hSARS 病毒或其天然或人工变体、类似物或衍生物, 含有能使病毒具有更适合用于疫苗制品的表型(例如减毒表型或增强的抗原性)的序列。变异和修饰可出现在病毒的编码区、基因间区和前导序列和尾随序列。

本发明提供包含本发明核酸或载体的宿主细胞。含有 hSARS 病毒聚合酶成分的质粒载体或病毒载体在原核细胞中产生, 以在相关的细胞类型(细菌、昆虫细胞、真核细胞)中表达所述成分。含有 hSARS 基因组全长拷贝或部分拷贝的质粒或病毒载体在原核细胞中产生, 以在体外或体内表达病毒核酸。后者载体可含有其它病毒序列以产生嵌合病毒或嵌合病毒蛋白, 可缺少病毒基因组的某些部分以产生复制缺陷型病毒, 并可含有突变、缺失、置换或插入以产生减毒病毒。

本发明也提供包含本发明核酸分子的宿主细胞。此外, 本发明提供感染 hSARS 病毒的宿主细胞, 例如保藏号为 CCTCC-V200303 的 hSARS 病毒或其天然或人工变体、类似物或衍生物。在一个具体实施方案中, 本发明包括 hSARS 病毒感染的传代细胞系。优选该细胞系为灵长目动物细胞系。这些细胞系可用已知细胞培养技术培养

和维持，例如在 Celis, Julio 编，1994，Cell Biology Laboratory Handbook, Academic Press, N. Y. 中所述。这些细胞的各种培养条件，包括关于具体营养的培养基配方、氧气、张力、二氧化碳和减少血清水平，可由本领域技术人员选择和优化。

- 5 优选本发明的细胞系为真核细胞系，优选灵长目动物细胞系，更优选猴细胞系，最优选胎恒河猴肾细胞系(例如 FRhK-4)，其短暂或稳定表达一种或多种全长或部分 hSARS 蛋白。这种细胞可通过转染(蛋白质载体或核酸载体)、感染(病毒载体)或转导(病毒载体)来制成，可用于与所述野生型、减毒型、复制缺陷型或嵌合型病毒互补。
- 10 本发明中使用的细胞系可用本领域技术人员熟悉的已知细胞培养技术克隆。可从单个细胞用市售商品化培养基在适合细胞增殖的已知条件下培养和扩展细胞。

例如，本发明细胞系保持冷冻直到使用，可在约 37°C 温度下加热，然后加入合适生长培养基，例如含有 3%胎牛血清(FBS)的

15 DMEM/F-12 (Life Technologies, Inc.)。所述细胞可在湿润恒温箱中约 5% CO₂、37°C 温度下孵育，直到融合。为使所述细胞传代，除去生长培养基，向细胞加入 0.05%胰蛋白酶和 0.53mM EDTA。分离所述细胞，将细胞悬液收集到离心管中，离心为细胞沉淀。可除去胰蛋白酶溶液，细胞沉淀重悬在新的生长培养基中。然后细胞可进一步

20 在另外的生长瓶内增殖至所需的密度。

根据本发明，传代细胞系包括无限增殖化细胞，其在体外可维持至少 5、10、15、20、25 或 50 代。

hSARS(野生型、减毒型、复制缺陷型或嵌合型)的传染性拷贝可根据上述现有技术共表达聚合酶成分时产生。

- 25 此外，可使用短暂或稳定表达一种或多种全长或部分 hSARS 蛋白质的真核细胞。这种细胞可通过转染(蛋白质载体或核酸载体)、感染(病毒载体)或转导(病毒载体)来制成，可用于与所述野生型、减毒型、复制缺陷型或嵌合型病毒互补。

本发明病毒载体和嵌合病毒可用于通过刺激体液免疫应答、细胞免疫应答或通过刺激对抗原的耐受性而调节对象的免疫系统。本文所用的对象指：人、灵长目动物、马、牛、绵羊、猪、山羊、狗、猫、鸟类和啮齿动物。

5

5.3. 疫苗和抗病毒剂

在一个优选实施方案中，本发明提供由本发明核酸编码的蛋白质分子或 hSARS 病毒特异性病毒蛋白及其功能片段。有用的蛋白质分子例如来源于可从本发明病毒衍生的任何基因或基因组片段，包括包膜蛋白(E 蛋白)、膜内在蛋白质(M 蛋白)、刺突蛋白(S 蛋白)、核壳蛋白(N 蛋白)、血凝素酯酶(HE 蛋白)和 RNA 依赖性 RNA 聚合酶。本文所提供的这种分子或其抗原片段例如可用于诊断方法或试剂盒中，以及用于药物组合物如亚单位疫苗中。特别有用的是由 SEQ ID NO: 1、11、13、15、2471、2472、2473、2474、2475 或 2476 核酸序列编码的多肽、图 11 (SEQ ID NO: 17-239、241-736 或 738-1107) 及图 12 (SEQ ID NO: 1109-1589、1591-1964 或 1966-2470) 中所示的多肽或其抗原片段，供混入作为抗原或亚单位免疫原，但也可使用灭活的全病毒。还尤其有用的是由 hSARS 基因组的重组核酸片段编码的蛋白质物质，更优选的是在 ORF 的优选界限内、尤其是在体内(例如出于保护目的或治疗目的，或用于提供诊断性抗体)或体外(例如通过噬菌体展示技术或其它用于产生合成抗体的技术)都引发 hSARS 特异性抗体或 T 细胞应答的蛋白质物质。

10

15

20

5.3.1 hSARS 病毒及其变体的减毒

可对本发明 hSARS 病毒或其变体进行基因工程改造，以显示出减毒表型。具体地说，本发明病毒在所述病毒作为疫苗给予的对象中显示出减毒表型。减毒可通过普通技术人员熟知的任何方法来实现。非为理论所囿，可例如通过使用本质上在预定宿主物种中不能很好地复制的病毒来产生本发明病毒的减毒表型，例如相对于病毒

25

的野生型毒株，通过减少病毒基因组的复制，通过降低病毒感染宿主细胞的能力，或通过降低病毒蛋白质装配成传染性病毒颗粒的能力。

5 在一个实施方案中，病毒的感染性被降低 10000 倍、9000 倍、8000 倍、7000 倍、6000 倍、5000 倍、4000 倍、3000 倍、2500 倍、2000 倍、1500 倍、1250 倍、1000 倍、900 倍、800 倍、700 倍、600 倍、500 倍、400 倍、300 倍、200 倍、100 倍、50 倍、25 倍、10 倍、5 倍、1 倍或 90%、80%、70%、60%、50%、40%、30%、20%或 10%。本文使用的术语“感染性”指病毒进入、存活和在敏感宿主中复制的能力。在一个具体实施方案中，当在人类宿主中生长时，如果 hSARS 病毒或其变体在人类宿主中的生长与非减毒 hSARS 病毒或其变体相比是降低的，hSARS 病毒的感染性被称为是降低的或减弱的。病毒的感染性可使用各种方法测定，例如但不限于蛋白质印迹(蛋白质)、RNA 印迹(RNA)、DNA 印迹(DNA)、噬斑形成试验、比色法、显微镜和化学发光技术。病毒的感染性可在动物细胞中测定，优选灵长目动物细胞，更优选猴细胞，最优选人细胞。

15 在另一个实施方案中，病毒的复制能力被降低 10000 倍、9000 倍、8000 倍、7000 倍、6000 倍、5000 倍、4000 倍、3000 倍、2500 倍、2000 倍、1500 倍、1250 倍、1000 倍、900 倍、800 倍、700 倍、600 倍、500 倍、400 倍、300 倍、200 倍、100 倍、50 倍、25 倍、10 倍、5 倍、1 倍或 90%、80%、70%、60%、50%、40%、30%、20%或 10%。本文使用的术语“复制能力”指病毒复制、繁殖和/或生殖的能力。复制能力可使用病毒的倍增时间、复制速率、生长速率和/或半寿期来测定。在一个具体实施方案中，当在人类宿主中生长时，如果 hSARS 病毒或其变体在人类宿主中的生长与非减毒 hSARS 病毒或其变体相比是降低的，hSARS 病毒的复制能力被称为是降低的或减弱的。病毒的复制能力可使用各种方法测定，例如但不限于蛋白质印迹(蛋白质)、RNA 印迹(RNA)、DNA 印迹(DNA)、噬斑形成

试验、比色法、显微镜和化学发光技术。在某些情况下，复制和转录可是同义。病毒的复制能力可在动物细胞中测定，优选灵长目动物细胞，更优选猴细胞，最优选人细胞。

5 在另一个实施方案中，病毒的蛋白质合成能力被降低 10000 倍、9000 倍、8000 倍、7000 倍、6000 倍、5000 倍、4000 倍、3000 倍、2500 倍、2000 倍、1500 倍、1250 倍、1000 倍、900 倍、800 倍、700 倍、600 倍、500 倍、400 倍、300 倍、200 倍、100 倍、50 倍、25 倍、10 倍、5 倍、1 倍或 90%、80%、70%、60%、50%、40%、30%、20%或 10%。本文使用的术语“蛋白质合成能力”指病毒合成蛋白质的能力，所述蛋白质例如但不限于包膜蛋白(E 蛋白)、膜内在蛋白质(M 10 蛋白)、刺突蛋白(S 蛋白)、核壳蛋白(N 蛋白)、血凝素酯酶(HE 蛋白)和 RNA 依赖性 RNA 聚合酶。蛋白质合成能力可通过蛋白质合成速率(例如转录水平、翻译水平)以及病毒合成的蛋白质的类型和量来测定。在一个具体实施方案中，当在人类宿主中生长时，如果 hSARS 15 病毒或其变体在人类宿主中的生长与非减毒 hSARS 病毒或其变体相比是降低的，hSARS 病毒的蛋白质合成能力被称为是降低的或减弱的。病毒的蛋白质合成能力可使用各种方法测定，例如但不限于蛋白质印迹(蛋白质)、RNA 印迹(RNA)、DNA 印迹(DNA)、噬斑形成 20 试验、比色法、显微镜和化学发光技术。病毒的蛋白质合成能力可在动物细胞中测定，优选灵长目动物细胞，更优选猴细胞，最优选人细胞。

在另一个实施方案中，病毒的装配能力被降低 10000 倍、9000 倍、8000 倍、7000 倍、6000 倍、5000 倍、4000 倍、3000 倍、2500 25 倍、2000 倍、1500 倍、1250 倍、1000 倍、900 倍、800 倍、700 倍、600 倍、500 倍、400 倍、300 倍、200 倍、100 倍、50 倍、25 倍、10 倍、5 倍、1 倍或 90%、80%、70%、60%、50%、40%、30%、20% 或 10%。本文使用的术语“装配能力”指病毒装配必需蛋白或蛋白组件成为病毒颗粒的能力。在一个具体实施方案中，当在人类宿主

中生长时，如果 hSARS 病毒或其变体在人类宿主中的生长与非减毒 hSARS 病毒或其变体相比是降低的，hSARS 病毒的装配能力被称为是降低的或减弱的。病毒的装配能力可使用各种方法测定，例如但不限于蛋白质印迹(蛋白质)、RNA 印迹(RNA)、DNA 印迹(DNA)、噬斑形成试验、比色法、显微镜和化学发光技术。病毒的装配能力可在动物细胞中测定，优选灵长目动物细胞，更优选猴细胞，最优选人细胞。

在另一个实施方案中，病毒的致细胞病变效应被降低 10000 倍、9000 倍、8000 倍、7000 倍、6000 倍、5000 倍、4000 倍、3000 倍、2500 倍、2000 倍、1500 倍、1250 倍、1000 倍、900 倍、800 倍、700 倍、600 倍、500 倍、400 倍、300 倍、200 倍、100 倍、50 倍、25 倍、10 倍、5 倍、1 倍或 90%、80%、70%、60%、50%、40%、30%、20%或 10%。本文使用的术语“致细胞病变效应”指感染病毒引起的对被感染宿主细胞的损害。病毒感染可导致细胞异常(生物化学的和形态学的)和/或细胞死亡(例如裂解)。在一个具体实施方案中，当在人类宿主中生长时，如果 hSARS 病毒或其变体在人类宿主中的生长与非减毒 hSARS 病毒或其变体相比是降低的，hSARS 病毒的致细胞病变效应被称为是降低的或减弱的。病毒的致细胞病变效应可使用各种方法测定，例如但不限于蛋白质印迹(蛋白质)、RNA 印迹(RNA)、DNA 印迹(DNA)、噬斑形成试验、比色法、显微镜和化学发光技术。病毒的致细胞病变效应可在动物细胞中测定，优选灵长目动物细胞，更优选猴细胞，最优选人细胞。

可对本发明病毒进行减毒，以使病毒的一种或多种功能特征受损。hSARS 病毒及其变体的减毒表型可通过普通技术人员熟知的任何方法来检测。候选病毒可例如检测其感染宿主的能力或其在细胞培养系统中的复制速率。在某些实施方案中，用不同温度下的生长曲线来检测病毒的减毒表型。例如，减毒病毒能在 35°C 下生长，但不能在 39°C 或 40°C 下生长。在某些实施方案中，可用不同的细胞

系评价病毒的减毒表型。例如，减毒病毒可只能在猴细胞系中生长，但却不能在人细胞系中生长，或者减毒病毒在不同细胞系中可达到的病毒滴度不同。在某些实施方案中，病毒在小型动物模型(包括但不限于仓鼠、棉鼠、小鼠、豚鼠)的呼吸道中的复制被用来评价病毒的减毒表型。在其它实施方案中，病毒诱导的免疫应答，包括但不限于抗体滴度(例如通过蚀斑减少中和试验或 ELISA 来试验)被用来评价病毒的减毒表型。在一个具体实施方案中，蚀斑减少中和试验或 ELISA 在低剂量下进行。在某些实施方案中，可检测出 hSARS 病毒在动物模型中引起病理症状的能力。病毒在动物模型中引起病理症状的能力降低是其减毒表型的指示。在一个具体实施方案中，检测候选病毒在猴模型中对鼻的感染，以黏液的产生为指标。

在某些其它实施方案中，减毒是通过与减毒病毒来源的病毒野生型毒株作比较来测定的。在其它实施方案中，减毒是通过比较减毒病毒在不同宿主系统中的生长来确定的。因此，作为非限制性实例，当在人宿主中生长时，如果 hSARS 病毒或其变体与非减毒 hSARS 或其变体相比在人宿主中的生长降低，则 hSARS 病毒或其变体被称为减毒的。

在某些实施方案中，本发明减毒病毒能够感染宿主，或者能够在宿主中复制，从而产生传染性病毒颗粒。但是，与野生型毒株相比，减毒株生长得到的滴度低，或生长得更缓慢。可用普通技术人员熟知的任何方法确定减毒病毒的生长曲线并将其与野生型病毒的生长曲线相比较。

在某些实施方案中，本发明减毒病毒在人细胞中复制得不如野生型病毒好。但是，减毒病毒在缺乏干扰素功能的细胞系如 Vero 细胞中可复制良好。

在其它实施方案中，本发明减毒病毒能够感染宿主，能够在宿主中复制，以及能够使本发明病毒的蛋白质嵌入到胞质膜中，但是减毒病毒不会引起宿主产生新的传染性病毒颗粒。在某些实施方案

中，减毒病毒感染宿主、在宿主中复制、并且导致病毒蛋白质嵌入到宿主的胞质膜中的效率与野生型 hSARS 一样。在其它实施方案中，减毒病毒相对于野生型病毒其导致病毒蛋白质嵌入到宿主细胞的胞质膜中的能力降低。在某些实施方案中，减毒 hSARS 病毒相对于野生型病毒其在宿主中复制的能力降低。可使用普通技术人员熟知的任何技术来确定病毒是否能够感染哺乳动物细胞，是否能够在宿主中复制以及是否能够导致病毒蛋白质嵌入到宿主的胞质膜中。

在某些实施方案中，本发明减毒病毒能够感染宿主。但与野生型 hSARS 病毒相反的是，减毒 hSARS 病毒不能在宿主中复制。在一个具体实施方案中，减毒 hSARS 病毒能感染宿主，能导致宿主将病毒蛋白质嵌入到其胞质膜中，但减毒病毒不能够在宿主中复制。可使用普通技术人员熟知的任何方法来检测减毒 hSARS 病毒是否已感染宿主以及是否已导致宿主将病毒蛋白质嵌入到其胞质膜中。

在某些实施方案中，减毒病毒感染宿主的能力与野生型病毒感染相同宿主的能力相比降低了。可使用普通技术人员熟知的任何技术来确定病毒是否能够感染宿主。

在某些实施方案中，突变(例如错义突变)被引入到病毒的基因组中，例如被引入到 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 核酸序列中，以产生具有减毒表型的病毒。突变(例如错义突变)可被引入到 hSARS 病毒的结构基因和/或调节基因中。突变可以是增添、置换、缺失或它们的组合。可筛选这种 hSARS 病毒变体的预期功能性，如在细胞培养物中的传染性、复制能力、蛋白质合成能力、装配能力以及致细胞病变效应。在一个具体实施方案中，错义突变是冷敏突变。在另一实施方案中，错义突变是热敏突变。在另一实施方案中，错义突变防止病毒蛋白质的正常加工或剪切。

在其它实施方案中，缺失被引入到 hSARS 病毒的基因组中，导致病毒的减毒。

在某些实施方案中，病毒的减毒是通过用不同种、不同亚群或不同变体的病毒基因置换野生型病毒的基因来实现的。在另一方面，病毒的减毒是通过用来源于不同种病毒的相应蛋白质的结构域置换野生型病毒的蛋白质的一个或多个特定结构域来实现的。在某些其它实施方案中，病毒的减毒是通过缺失野生型病毒蛋白质的一个或多个特定结构域来实现的。

当使用活减毒疫苗时，还必须考虑其安全性。所述疫苗必须不会导致疾病。本领域熟知的能使疫苗安全的任何技术均可在本发明中使用。除减毒技术外，也可使用其它技术。一个非限制性实例是使用不能被掺入到病毒颗粒膜中的可溶性异源基因。例如，可使用单拷贝的缺乏跨膜结构域和胞质结构域的病毒跨膜蛋白质可溶性形式。

可使用多种试验检测疫苗的安全性。例如，可使用蔗糖梯度试验和中和试验来检测安全性。蔗糖梯度试验可用来确定异源蛋白质是否被插入到病毒颗粒中。如果异源蛋白质被插入到病毒颗粒中，则应检测病毒颗粒在适当的动物模型中导致症状的能力，因为病毒可能已经获得新的、可能致病的性质。

5.3.2 疫苗的配制

本发明提供用于预防或治疗 hSARS 病毒感染的疫苗制品。在某些实施方案中，本发明疫苗包含 hSARS 病毒的重组和嵌合病毒。在某些实施方案中，病毒是减毒的、灭活的或死的。

在本发明此方面的另一实施方案中，灭活疫苗制品可通过使用常规技术“杀死”嵌合病毒来制备。在其传染性已被破坏的意义上，灭活疫苗是“死的”。理想的是，病毒的传染性被破坏，但不影响其免疫原性。为制备灭活疫苗，可使嵌合病毒在细胞培养物中或在鸡胚的尿囊中生长，通过区带超离心纯化，用甲醛或 β -丙醇酸内酯灭活，收集。所得疫苗通常通过肌肉内接种。

灭活病毒可用合适的佐剂配制，以增强免疫应答。这种佐剂可

包括但不限于无机凝胶，例如氢氧化铝；表面活性物质如溶血卵磷脂、聚醚多元醇(pluronic polyol)、聚阴离子；肽；油乳液；及具有潜在用途的人佐剂如 BCG 和小棒杆菌(*Corynebacterium parvum*)。

5 本发明疫苗可为多价或单价的。多价疫苗制备自表达超过一种抗原的重组病毒。

在另一方面，本发明还提供 DNA 疫苗制品，其中包含 hSARS 病毒(例如保藏号为 CCTCC -V200303 的病毒)的核酸或其片段，或具有 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、2473、2474、2475 或 2476 序列或其互补物、类似物、
10 衍生物、片段或部分的核酸分子。在另一个具体实施方案中，本发明 DNA 疫苗制品包含编码免疫特异性结合 hSARS 病毒的抗体的核酸或其片段。在 DNA 疫苗制品中，DNA 疫苗包含带有插入片段的病毒载体(例如来源于 hSARS 病毒、细菌质粒或其它表达载体)，所述插入片段包含可操作性地与一种或多种控制元件相连的本发明核酸分子，
15 从而使得由所述核酸分子编码的接种蛋白质可以在接种对象中表达。这种载体可用重组 DNA 技术制备成携带本发明核酸分子的重组或嵌合病毒载体(参见上文 5.1 节)。

已描述了各种供 DNA 接种以抗病毒感染的异源载体。例如，以下参考文献描述的载体可用来表达 hSARS 序列而不是所描述的病毒
20 或其它病原体的序列；尤其描述用于以下的载体：乙肝病毒(Michel, M.L.等,1995,DAN-mediated immunization to the hepatitis B surface antigen in mice: Aspects of the humoral response mimic hepatitis B viral infection in humans, Proc. Natl. Aca Sci. USA 92: 5307-5311; Davis, H.L.等,1993, DNA-based immunization induces continuous secretion of hepatitis B surface antigen and high levels of circulating antibody, Human Molec. Genetics 2: 1847-1851)、HIV 病毒(Wang, B.等, 1993, Gene inoculation generates immune responses against human immunodeficiency virus type 1, Proc. Natl. Acad. Sci. USA 90: 4156-4160; Lu, S.等, 1996,
25

Simian immunodeficiency virus DNA vaccine trial in macques, J. Virol. 70: 3978-3991; Letvin, N.L.等, 1997, Potent, protective anti-HIV immune responses generated by bimodal HIV envelope DNA plus protein vaccination, Proc Natl Acad Sci USA. 94(17): 9378-83)和流感病毒 (Robinson, HL 等, 1993, Protection against a lethal influenza virus challenge by immunization with a haemagglutinin-expressing plasmid DNA, Vaccine 11: 957-960; Ulmer, J.B.等, Heterologous protection against influenza by injection of DNA encoding a viral protein, Science 259: 1745-1749), 以及细菌感染如结核病(Tascon, R. E.等, 1996, Vaccination against tuberculosis by DNA injection, Nature Med. 2: 888-892; Huygen, K.等, 1996, Immunogenicity and protective efficacy of a tuberculosis DNA vaccine, Nature Med.,2: 893-898)和寄生虫感染如疟疾(Sedegah,M., 1994, Protection against malaria by immunization with plasmid DNA encoding circumsporozoite protein, Proc. Natl. Acad. Sci. USA 91: 9866-9870; Doolan, D.L.等, 1996, Circumventing genetic restriction of protection against malaria with multigene DNA immunization: CD8⁺ T cell-interferon δ , and nitric oxide-dependent immunity, J. Exper. Med., 1183: 1739-1746)。

可使用多种方法来输入上述疫苗制品。这些方法包括但不限于口、皮内、肌内、腹膜内、静脉内、皮下和鼻内途径, 以及通过划痕法(例如使用分叉针刮破皮肤表层)。

另外, 优选通过疫苗针对其设计的病原体的自然感染途径来输入嵌合病毒疫苗制品。本发明 DNA 疫苗可以盐水溶液形式, 通过用注射器和针头注射入肌肉或皮肤内来给予(Wolff J.A.等, 1990, Direct gene transfer into mouse muscle in vivo, Science 247: 1465-1468; Raz, E., 1994, Intradermal gene immunization: The possible role of DNA uptake in the induction of cellular immunity to viruses, Proc. Natl. Acad. Sci. USA 91: 9519-9523)。另一种给予 DNA 疫苗的方式称为“基因枪”方法, 通过该法, 涂有目的 DNA 分子的微小金珠被射入细胞中

(Tang, D.等, 1992, Genetic immunization is a simple method for eliciting an immune response, Nature 356: 152-154)。有关 DNA 疫苗的方法的全面综述参见 Robinson, H.L., 1999, DNA vaccines: basic mechanism and immune responses (综述), Int. J. Mol. Med. 4(5): 549-555; Barber, B., 1997, Introduction: Emerging vaccine strategies, Seminars in Immunology 9(5): 269-270; 和 Robinson, H.L.等, 1997, DNA vaccines, Seminars in Immunology 9(5): 271-283。

被给予疫苗的患者优选为哺乳动物, 最优选为人, 但也可为非人动物, 包括但不限于牛、马、绵羊、猪、禽(例如鸡)、山羊、猫、狗、仓鼠、小鼠和大鼠。

5.3.3. 佐剂和载体分子

在某些实施方案中, hSARS 相关抗原与一种或多种佐剂一起给药。在一个实施方案中, hSARS 相关抗原与无机盐佐剂或无机盐凝胶佐剂一起给药。这种无机盐佐剂和无机盐凝胶佐剂包括但不限于氢氧化铝(ALHYDROGEL、REHYDRAGEL)、磷酸铝凝胶、羟基磷酸铝(ADJU-PHOS)和磷酸钙。

在另一实施方案中, hSARS 相关抗原与免疫刺激佐剂一起给药。这类佐剂包括但不限于细胞因子(例如白介素-2、白介素-7、白介素-12、粒细胞-巨噬细胞集落刺激因子(GM-CSF)、干扰素- γ 、白介素-1 β (IL-1 β)和 IL-1 β 肽或 Sclavo 肽)、含细胞因子的脂质体、三萜类糖苷或皂苷(例如 QuilA 和 QS-21, 亦在商标 STIMULON、ISCOPREP 下出售)、胞壁酰二肽(MDP)衍生物如 N-乙酰-胞壁酰-L-苏氨酸-D-异谷氨酰胺(苏氨酸-MDP, 在商标 TERMURTIDE 下出售)、GMDP、N-乙酰-去甲胞壁酰-L-丙氨酸-D-异谷氨酰胺、N-乙酰胞壁酰-L-丙氨酸-D-异谷氨酰胺酰-L-丙氨酸-2-(1'-2'-二棕榈酰-sn-甘油基-3-羟基磷酰氧基)乙胺、胞壁酰三肽磷脂酰乙醇胺(MTP-PE)、未甲基化的 CpG 二核苷酸和寡核苷酸如细菌 DNA 及其片段、LPS、一磷酰脂质 A (3D-MLA, 在商标 MPL 下出售)和聚磷腈。

在另一实施方案中，所用的佐剂是特殊的佐剂，包括但不限于乳液，例如弗氏完全佐剂、弗氏不完全佐剂、例如与嵌段共聚物如 L-121 (聚氧丙烯/聚氧乙烯，在商标 PLURONIC L-121 下出售)制备的角鲨烯或角鲨烷水包油佐剂制剂如 SAF 和 MF59、脂质体、病毒体、
5 脂质体卷(cochleate)和在商标 ISCOM 下出售的免疫刺激复合物。

在另一实施方案中，使用微粒佐剂。微粒佐剂包括但不限于生物可降解和生物相容性聚酯、乳酸的均聚物(PLA)及乙醇酸的均聚物(PGA)和它们的共聚物、丙交酯-乙交酯共聚物(PLGA)微粒、能自缔合成微粒的聚合物(泊洛沙姆颗粒)、可溶性聚合物(聚磷腈)和病毒样
10 颗粒(VLP)如重组蛋白质微粒，例如乙肝表面抗原(HbsAg)。

可以使用的又一类佐剂包括黏膜佐剂，包括但不限于来自大肠杆菌(*Escherichia coli*)的不耐热肠毒素(LT)、来自霍乱弧菌(*Vibrio cholerae*)的霍乱全毒素(CT)和霍乱毒素 B 亚单位、突变株毒素(例如 LTK63 和 LTR72)、微粒和聚合脂质体。

15 在其它实施方案中，上述任何类型的佐剂可相互组合或与其它佐剂组合使用。例如，可用来给予本发明 hSARS 相关抗原的组合佐剂制剂的非限制性实例包括包含免疫刺激蛋白质、细胞因子、T 细胞和/或 B 细胞肽的脂质体；或带有或不带包埋 IL-2 的微生物或含有肠毒素的微粒。本领域熟知的其它佐剂也包括在本发明的范围内(参见
20 *Vaccine Design: The Subunit and Adjuvant Approach*, 第 7 章, Michael F.Powell 和 Mark J.Newman (编辑), Plenum Press, New York, 1995, 其通过引用整体结合到本文中)。

佐剂的效力可通过测量诱导针对免疫原性多肽(含有 hSARS 多肽表位)的抗体来确定，所述抗体由在也包含各种佐剂的疫苗中给予
25 这种多肽而产生。

所述多肽可配制成中性形式或盐形式的疫苗。药物可接受的盐包括酸加成盐(通过肽的游离氨基形成)和与无机酸(例如盐酸或磷酸)或有机酸(例如乙酸、草酸、酒石酸、马来酸等)形成的盐。通过游离

羧基形成的盐也可来源于无机碱，例如氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙或氢氧化铁，及有机碱，例如异丙胺、三甲胺、2-乙氨基乙醇、组氨酸、普鲁卡因等。

5 5.4. 抗体的制备

抗体可从感染 SARS 的对象的血清中分离。特异性识别本发明多肽，例如但不限于包含 SEQ ID NO: 2、12 或 14 序列的多肽、或如图 11 (SEQ ID NO: 17-239、241-736 和 738-1107)和图 12 (SEQ ID NO: 1109-1589、1591-1964、1966-2470)所示的多肽、或 hSARS 表位的抗体或其抗原结合片段，可用来检测、筛选和分离本发明多肽或其片段或可能编码其它生物类似酶的类似序列。例如，在一个具体实施方案中，免疫特异性结合 hSARS 表位的抗体或其片段可用于各种体外检测试验中，包括酶联免疫吸附测定(ELISA)、放射免疫测定、蛋白质印迹等，以在样品例如生物材料中检测本发明多肽或优选 hSARS 病毒的多肽，所述生物材料包括细胞、细胞培养基(例如细菌细胞培养基、哺乳动物细胞培养基、昆虫细胞培养基、酵母细胞培养基等)、血液、血清、血浆、唾液、尿、粪便、组织、痰、鼻咽抽吸物等。

对本发明多肽或 hSARS 病毒的任何表位具有特异性的抗体可通过本领域熟知的任何适合方法产生。抗目的抗原(例如保藏号为 CCTCC- V200303 或包含 SEQ ID NO: 15 核酸序列的 hSARS 病毒)的多克隆抗体可通过本领域熟知的各种方法生产。例如，可将抗原给予各种宿主动物，包括但不限于兔、小鼠、大鼠等，以诱导产生含有抗原特异性多克隆抗体的抗血清。可使用多种佐剂来增强免疫应答，取决于宿主物种，佐剂包括但不限于弗氏(完全和不完全)佐剂、无机凝胶如氢氧化铝、表面活性物质如溶血卵磷脂、聚醚多元醇、聚阴离子、肽、油乳液、匙孔蛾血蓝蛋白、二硝基酚和对人类具有潜在用途的佐剂如 BCG (卡介苗)和小棒杆菌。这种佐剂在本领域也

是熟知的。

单克隆抗体可使用多种本领域熟知的技术制备，包括使用杂交瘤技术、重组体技术和噬菌体呈现技术或它们的组合。例如单克隆抗体可使用杂交瘤技术生产，包括本领域公知的杂交瘤技术和例如在 Harlow 等, *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 第 2 版, 1988); Hammerling 等, *Monoclonal Antibodies and T-Cell Hybridomas*, 第 563-681 页(Elsevier, N.Y., 1981)中讲授的杂交瘤技术(两者通过引用整体结合到本文中)。本文所用术语“单克隆抗体”不限于通过杂交瘤技术生产的抗体。术语“单克隆抗体”指来源于单个克隆的抗体，包括任何真核克隆、原核克隆和噬菌体克隆，而不是指抗体生产的方法。

使用杂交瘤技术生产和筛选具体抗体的方法是常规的和本领域熟知的。在非限制性实例中，小鼠可用目的抗原或表达这种抗原的细胞免疫。一旦检测出免疫应答，例如在小鼠血清中检测出对抗原特异的抗体，则收获小鼠脾脏，分离脾细胞。然后通过熟知的技术将脾细胞与任何适合的骨髓瘤细胞融合。通过有限稀释选择和克隆杂交瘤。然后用本领域熟知的方法检测杂交瘤克隆中分泌能结合抗原的抗体的细胞。通常含有高水平抗体的腹水可通过用阳性杂交瘤克隆腹膜内接种小鼠来产生。

识别特定表位的抗体片段可通过公知技术产生。例如，Fab 和 $F(ab')_2$ 片段可通过使用酶例如木瓜蛋白酶(用于产生 Fab 片段)或胃蛋白酶(用于产生 $F(ab')_2$ 片段)，对免疫球蛋白分子进行蛋白酶水解来生产。 $F(ab')_2$ 片段含有全部轻链以及重链的可变区、CHI 区和铰链区。

本发明抗体或其片段也可通过本领域熟知的任何抗体合成方法来生产，具体地说通过化学合成法，或优选通过重组表达技术来生产。

编码抗体的核苷酸序列可从本领域普通技术人员可获取的任何信息获得(即从 Genbank、文献获得或通过常规克隆和序列分析获得)。

如果含有编码特定抗体或其表位结合片段的核酸的克隆不可获得，但抗体分子或其表位结合片段的序列是已知的，则编码免疫球蛋白的核酸可通过化学法合成，或获得自合适的来源(例如抗体 cDNA 文库，或从表达抗体的任何组织或细胞(如选择用来表达抗体的杂交瘤细胞)产生的 cDNA 文库或从其中分离的核酸，优选 poly A+ RNA)，5 通过使用可与序列的 3'和 5'末端杂交的合成引物进行 PCR 扩增，或通过使用对特定基因序列具有特异性、例如用于从编码抗体的 cDNA 文库中鉴别 cDNA 克隆的寡核苷酸探针进行克隆来获得。通过 PCR 产生的扩增核酸随后可使用本领域熟知的任何方法克隆至可复制克隆载体中。10

一旦确定了抗体的核苷酸序列，抗体的核苷酸序列可使用本领域熟知的操作核苷酸序列的方法，例如重组 DNA 技术、定点诱变、PCR 等(参见例如上述 Sambrook 等和 Ausubel 等编，Current Protocols in Molecular Biology, John Wiley & Sons, NY，它们通过引用整体结合到本文中)进行操作，以通过例如在抗体的表位结合域中或者在可增强或降低抗体生物活性的抗体任何部分中引入氨基酸取代、缺失和/或插入，来产生具有不同氨基酸序列的抗体。15

抗体的重组表达需要构建含有编码抗体的核苷酸序列的表达载体。一旦获得编码抗体分子或抗体的重链或轻链或其部分的核苷酸序列，用于生产抗体分子的载体可通过重组 DNA 技术，使用前面各节讨论的本领域熟知的技术来生产。可使用本领域普通技术人员熟知的方法来构建含有抗体编码序列和合适转录和翻译控制信号的表达载体。这些方法包括例如体外重组 DNA 技术、合成技术和体内遗传重组。可将编码重链可变区、轻链可变区、重链和轻链可变区、重链和/或轻链可变区的表位结合片段、或抗体的一个或多个互补决定区(CDR)的核苷酸序列克隆到这种载体中进行表达。然后，如此制备的表达载体可被引入到合适的宿主细胞中以表达抗体。因此，本25 发明包括含有多核苷酸的宿主细胞，所述多核苷酸编码对本发明多

肽或其片段具有特异性的抗体。

5 宿主细胞可用本发明的两种表达载体共转染，第一种载体编码重链衍生的多肽，第二种载体编码轻链衍生的多肽。所述两种载体可含有相同的可选择标记，以使重链和轻链多肽的表达相等，或含有不同的可选择标记，以确保维持两种质粒。另外，也可使用编码和能够表达重链多肽和轻链多肽的单一载体。在这种情况下，轻链应位于重链的前面，以避免出现过量的有毒游离重链(Proudfoot, 1986, Nature, 322: 52 和 Kohler, 1980, Proc. Natl. Acad. Sci. USA, 77: 2 197)。重链和轻链的编码序列可包含 cDNA 或基因组 DNA。

10 在另一实施方案中，抗体也可使用本领域熟知的各种噬菌体呈现方法来产生。在噬菌体呈现方法中，功能抗体结构域被呈现在携带其编码多核苷酸序列的噬菌体颗粒的表面上。在一个具体实施方案中，这种噬菌体可被用来呈现从所有成分或组合抗体文库(例如人的或鼠的)表达的抗原结合域，如 Fab 和 Fv 或二硫键稳定的 Fv。表达结合目的抗原的抗原结合域的噬菌体可用抗原来选择或鉴定，例如使用标记抗原或被结合或捕捉到固体表面或珠粒的抗原。在这些方法中使用的噬菌体通常是丝状噬菌体，包括 fd 和 M13。抗原结合域被表达为与噬菌体基因 III 或基因 VIII 蛋白质的重组融合蛋白质。可用来制备本发明免疫球蛋白或其片段的噬菌体呈现方法的实例包括在以下文献中公开的方法：Brinkman 等, 1995, J. Immunol. Methods, 182: 41-50; Ames 等, 1995, J. Immunol. Methods, 184: 177-186; Kettleborough 等, 1994, Eur. J. Immunol., 24: 952-958; Persic 等, 1997, Gene, 187: 9-18; Burton 等, 1994, Advances in Immunology, 57: 191-280; PCT 申请号 PCT/GB91/01134; PCT 公开号 WO 90/02809; 15 WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; 及美国专利号 5,698,426、5,223,409、5,403,484、5,580,717、5,427,908、5,750,753、5,821,047、5,571,698、5,427,908、5,516,637、5,780,225、5,658,727、5,733,743 和 5,969,108;

20

25

以上各文献通过引用整体结合到本文中。

如以上参考文献所述，进行噬菌体选择后，可分离和使用来自噬菌体的抗体编码区以产生所有的抗体，包括人抗体或任何其它需要的片段，并可表达于任何需要的宿主中，包括哺乳动物细胞、昆虫细胞、植物细胞、酵母和细菌，例如以下所详细描述。例如，5 也可应用重组生产 Fab、Fab'和 F(ab')₂ 片段的技术，使用例如在以下文献中公开的本领域熟知的方法：PCT 公开号 WO 92/22324；Mullinax 等，1992，BioTechniques, 12(6): 864-869；及 Sawai 等，AJRI, 34: 26-34, 1995；及 Better 等，1988，Science, 240: 1041-1043 (以上各文献通过引用整体结合到本文中)。10 5 10 15 20 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995

一旦本发明抗体分子已通过上述任何方法生产出来，其随后可通过本领域熟知的任何纯化免疫球蛋白分子的方法进行纯化，所述方法例如色谱法(例如离子交换色谱，亲和色谱，尤其是进行蛋白质 A 和蛋白质 G 纯化后对特定抗原的亲和色谱，以及排阻柱色谱)、离心法、差示溶解度法、或任何其它蛋白质纯化的标准技术。此外，20 本发明抗体或其片段可与本文所述的或本领域熟知的异源多肽序列融合，以促进纯化。

对于某些应用，包括在人体内使用抗体和体外检测试验，优选使用嵌合抗体、人源化抗体或人抗体。嵌合抗体是其中抗体的不同部分来源于不同的动物物种的抗体分子，例如具有来源于鼠单克隆25 抗体的可变区和来源于人免疫球蛋白的恒定区的抗体。生产嵌合抗体的方法是本领域熟知的。参见例如 Morrison, 1985, Science, 229: 1202; Oi 等, 1986, BioTechniques, 4: 214; Gillies 等, 1989, J. Immunol. Methods, 125: 191-202; 美国专利号 5,807,715、4,816,567 和 4,816,397,

以上各文献通过引用整体结合到本文中。人源化抗体是来自非人类物种的能结合所需抗原的抗体分子，其具有一个或多个来自非人类物种的互补决定区(CDR)和来自人免疫球蛋白分子的构架区。通常，人构架区中的构架残基会被来自 CDR 供体抗体的相应残基所取代，以改变、优选改进抗原的结合。这些构架取代可通过本领域熟知的方法来鉴定，例如通过对 CDR 和构架残基的相互作用进行建模，以鉴定对抗原结合重要的构架残基，以及通过序列比较，以鉴定在特定位置的不寻常构架残基。参见例如 Queen 等，美国专利号 5,585,089；Riechmann 等，1988, Nature, 332: 323，以上文献通过引用整体结合到本文中。抗体可使用多种本领域熟知的各种技术进行人源化，包括例如 CDR 嫁接(EP 239,400；PCT 公开号 WO 91/09967；美国专利号 5,225,539；5,530,101 和 5,585,089)、表面胶合(veneering)或表面重修(resurfacing) (EP 592,106；EP 519,596；Padlan, 1991, Molecular Immunology, 28(4/5): 489-498；Studnicka 等，1994, Protein Engineering, 7(6): 805-814；Roguska 等，1994, Proc Natl. Acad. Sci. USA, 91: 969-973)和链改组(美国专利号 5,565,332)，以上各文献通过引用整体结合到本文中。

完全人抗体特别适合用于对人患者进行治疗处理。人抗体可通过多种本领域熟知的方法来制备，包括使用来源于人免疫球蛋白序列的抗体文库通过上述噬菌体呈现技术来制备。参见美国专利号 4,444,887 和 4,716,111；以及 PCT 公开号 WO 98/46645；WO 98/50433；WO 98/24893；WO 98/16654；WO 96/34096；WO 96/33735 和 WO 91/10741，以上各文献通过引用整体结合到本文中。

人抗体也可使用不能表达功能性内源免疫球蛋白、但能表达人免疫球蛋白基因的转基因小鼠来生产。有关这种生产人抗体的技术的综述，参见 Lonberg 和 Huszar, 1995, Int. Rev. Immunol., 13: 65-93。有关这种生产人抗体和人单克隆抗体的技术以及生产这种抗体的方案的详细讨论，参见例如 PCT 公开号 WO 98/24893；WO 92/01047；

WO 96/34096; WO 96/33735; 欧洲专利号 0 598 877; 美国专利号 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771 和 5,939,598, 以上各文献通过引用整体结合到本文中。另外, 可与例如 Abgenix, Inc. (Fremont, CA),
5 Medarex (NJ)和 Genpharm (San Jose, CA)等公司约定, 以提供使用类似上述技术生产的针对选定抗原的人抗体。

识别选定表位的完全人抗体可使用称为“导向选择”的技术来产生。在这种方法中, 选定的非人单克隆抗体如小鼠抗体被用来指导选择识别同一表位的完全人抗体。(Jespers 等, 1988, Bio/technology,
10 12: 899-903)。

与异源多肽融合或缀合的抗体可用于本领域熟知的体外免疫试验或纯化方法(例如亲和色谱)中。参见例如 PCT 公开号 WO 93/21232; EP 439,095; Naramura 等, 1994, Immunol. Lett., 39: 91-99; 美国专利号 5,474,981; Gillies 等, 1992, PNAS, 89: 1428-1432 和 Fell
15 等, 1991, J. Immunol., 146: 2446-2452, 以上各文献通过引用整体结合到本文中。

抗体也可吸附到固体支持体上, 这对免疫试验和纯化本发明多肽或其片段、衍生物、类似物或变体或具有与本发明多肽类似酶活性的类似分子来说是特别有用的。这种固体支持体包括但不限于玻
20 璃、纤维素、聚丙烯酰胺、尼龙、聚苯乙烯、聚氯乙烯或聚丙烯。

5.5 药物组合物和药盒

本发明包括包含本发明抗病毒剂的药物组合物。在一个具体实施方案中, 抗病毒剂是免疫特异性结合和中和 hSARS 病毒或其天然
25 或人工变体、类似物或衍生物或其任何衍生蛋白质的抗体。所述病毒中和抗体中和病毒的感染性, 并当随后将野生型病毒给予动物时保护动物抵抗疾病。

在另一个具体实施方案中, 所述抗病毒剂为本发明多肽或核酸

分子。药物组合物具有预防性抗病毒剂的效用，可给予已暴露于或预期暴露于病毒的对象。

已知有各种给药系统可用来给予本发明药物组合物，所述给药系统例如脂质体包囊化、微粒、微胶囊、能表达突变病毒的重组细胞、受体介导胞吞作用(参见例如 Wu 和 Wu, 1987, J. Biol. Chem. 262: 4429-4432)。引入的方法包括但不限于皮内、肌内、腹膜内、静脉内、皮下、鼻内、硬膜外、划痕和口服途径。化合物可通过任何方便的途径给药，例如通过灌输或大剂量注射、通过上皮或黏膜皮肤内层(例如口腔黏膜、直肠黏膜和肠黏膜等)吸收，也可与其它生物活性剂一起给药。给药可以是全身给药或局部给药。在优选的实施方案中，期望通过任何合适的途径将本发明药物组合物引入肺部。也可例如通过使用吸入器或喷雾器，与气雾化剂一起配制来实施肺部给药。

在一个具体实施方案中，期望将本发明药物组合物局部给予需要进行治疗的区域；这可例如(非限制性地)通过手术中的局部灌输、局部敷用(例如与手术后的伤口敷料一起)、通过注射、通过导管、通过栓剂或通过植入物来实现，所述植入物是多孔、无孔或凝胶状材料，包括薄膜如硅橡胶(silastic)膜或纤维。在一个实施方案中，可通过在受感染组织的部位(或从前部位)直接注射来给药。

在另一实施方案中，药物组合物可在小泡、尤其是脂质体中进行送递(参见 Langer, 1990, Science 249: 1527-1533; Treat 等, Liposomes in the Therapy of Infectious Disease and Cancer, Lopez Berestein 和 Fidler(编辑), Liss, New York, 第 53-365 页(1989); Lopez- Berestein, 出处同上, 第 317-327 页; 通常参见出处同上)。

在又一个实施方案中，药物组合物可以控释系统送递。在一个实施方案中，可使用泵(参见 Langer, 出处同上; Sefton, 1987, CRC Crit. Ref. Biomed. Eng. 14: 201; Buchwald 等, 1980, Surgery 88: 507; 和 Saudek 等, 1989, N. Engl. J. Med. 321: 574)。在另一实施方案中，可使用聚合物材料(参见 Medical Applications of Controlled Release,

Langer 和 Wise (编辑), CRC Pres., Boca Raton, Florida(1974);
Controlled Drug Bioavailability, Drug Product Design and Performance,
Smolen 和 Ball(编辑), Wiley, New York(1984); Ranger 和 Peppas, 1983,
J. Macromol. Sci. Rev. Macromol. Chem. 23: 61; 也参见 Levy 等, 1985,
5 Science 228: 190; During 等, 1989, Ann. Neturol. 25: 351; Howard
等, 1989, J. Neurosurg. 71: 105)。在又一个实施方案中, 控释系统可
置于组合物靶标即肺的附近, 从而只需要全身剂量的一部分即可(参
见例如 Goodson, Medical Applications of Controlled Release, 出处同
上, 第 2 卷, 第 115-138 页(1984))。

10 其它控释系统在 Langer 的综述(1990, Science 249: 1527-1533)
中有讨论。

本发明药物组合物包含治疗有效量的活但减毒的、灭活的或死的
的 hSARS 病毒, 或重组或嵌合的 hSARS 病毒以及药物可接受的载
体。在一个具体实施方案中, 术语“药物可接受的”指得到联邦政
15 府或州政府的管理机关批准, 或被列入美国药典或其它通常公认的
药典中, 以供在动物中、更具体在人类中使用。术语“载体”指与
药物组合物一起给予的稀释剂、佐剂、赋形剂或介质。这种药物载
体可以是无菌液体如水和油, 包括石油、动物、植物或合成来源的
液体, 如花生油、大豆油、矿物油、芝麻油等。当静脉内给予药物
20 组合物时, 水是优选的载体。盐水溶液和葡萄糖及甘油水溶液也可
用作液体载体, 尤其是用于可注射溶液。合适的药物赋形剂包括淀
粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、大米、面粉、石灰石、硅
胶、硬脂酸钠、一硬脂酸甘油酯、滑石、氯化钠、脱脂奶粉、甘油、
丙烯、乙二醇、水、乙醇等。如需要, 组合物也可含有少量的湿润
25 剂或乳化剂或 pH 缓冲剂。这些组合物可以溶液剂、混悬剂、乳剂、
片剂、丸剂、胶囊剂、散剂、缓释剂等形式出现。组合物可与传统
的粘合剂和载体如甘油三酯一起配制成栓剂。口服剂型可包含标准
的载体, 如药物级甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维

素、碳酸镁等。合适的药物载体的实例在 E. W. Martin 著的“Remington's Pharmaceutical Sciences”中有描述。剂型应与给药方式相适应。

5 在优选的实施方案中，组合物按照常规程序配制成适合供静脉内给予人类的药物组合物。典型的是，供静脉内给药的组合物是无菌等渗水缓冲液的溶液。必要时，组合物也可包括增溶剂和局部麻醉剂如利多卡因，以减轻注射部位的疼痛。通常，各成分以单位剂量形式单独或混和供应，例如作为在密封容器如安瓿或小袋中的冻干粉或无水浓缩物，容器上标明活性药物的量。在组合物通过输液
10 给药的情况下，其可用装有无菌药物级水或盐水的输液瓶来调剂。在组合物通过注射给药的情况下，可提供一安瓿注射用无菌水或盐水，以便在给药前混合各成分。

本发明药物组合物可配制成中性形式或盐形式。药物可接受的盐包括与游离氨基形成的盐，如来源于盐酸、磷酸、乙酸、草酸、
15 酒石酸等的盐，以及与游离羧基形成的盐，如来源于氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙、氢氧化铁、异丙胺、三乙胺、2-乙氨基乙醇、组氨酸、普鲁卡因等的盐。

本发明药物组合物能有效治疗具体病状或病症的量须取决于病状或病症的性质，可通过标准的临床技术来确定。另外，可任选采用体外试验来帮助确定最佳的剂量范围。在制剂中待采用的精确剂量
20 也须取决于给药途径和疾病或病状的严重程度，且应按照执业医师的判断和各个患者的情况来决定。但是，静脉内给药合适剂量范围通常是每公斤体重约 20-500 微克活性化合物。鼻内给药合适剂量范围通常是约 0.01 pg/kg 体重至 1 mg/kg 体重。有效剂量可从得
25 自体外或动物模型测试系统的量效曲线推知。

栓剂通常含有 0.5%-10% (重量)的活性成分；口服剂型优选含有 10%-95%的活性成分。

本发明还提供包含一个或多个容器的药物小包或药盒，容器中

装有一种或多种本发明药物组合物成分。可以任选伴随这种容器的是一份管理药品或生物制品的生产、使用或销售的政府机关的规定形式的通知书，该通知书反映了政府机关对生产、使用或销售人用药品或生物制品的批准。在一个优选实施方案中，药盒包含本发明
5 抗病毒剂，例如对 SEQ ID NO: 1、11、13、15、2471、2472、2473、2474、2475 或 2476 核酸序列编码的多肽、或图 11 (SEQ ID NO: 17-239、241-736 或 738-1107)和图 12 (SEQ ID NO: 1109-1589、1591-1964 或 1966-2470)所示的多肽、或任何 hSARS 表位、或本发明多肽或蛋白质、或本发明核酸分子具有特异性的抗体，其可单独存在，或与
10 佐剂、抗病毒药物、抗生素、止痛剂、支气管扩张药或其它药物可接受赋形剂组合。

本发明进一步包括这种药盒，其包含含有本发明药物组合物的容器和使用说明书。

15 5.6 检测试验

本发明提供在来自 SARS 患者的生物样品如细胞、血液、血清、血浆、唾液、尿、粪便、痰、鼻咽抽吸物等中检测免疫特异性结合 hSARS 病毒的抗体的方法。在一个具体实施方案中，所述方法包括将样品与直接固定化在底物上的 hSARS 病毒，例如保藏号为 CCTCC-
20 V200303 的 hSARS 病毒或具有 SEQ ID NO: 15 基因组核酸序列的 hSARS 病毒接触，然后直接检测或通过标记的异源抗同种型抗体间接检测与病毒结合的抗体。在另一具体实施方案中，样品与被 hSARS 病毒，例如保藏号为 CCTCC-V200303 的 hSARS 病毒或具有 SEQ ID NO: 15 基因组核酸序列的 hSARS 病毒感染的宿主细胞接触，然后
25 被结合的抗体可通过以下 6.5 节描述的免疫荧光试验来检测。

检测生物样品中本发明多肽或核酸是否存在的示例性方法包括，从各种来源获得生物样品，将样品与能够检测 hSARS 病毒的表位或核酸(例如 mRNA、基因组 RNA)的化合物或试剂接触，这样检

测样品中 hSARS 病毒的存在。检测本发明 hSARS mRNA 或基因组 RNA 的优选试剂是能够与编码本发明多肽的 mRNA 或基因组 RNA 杂交的标记核酸探针。该核酸探针可以是例如包含或由 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590、1965、2471、2472、
5 2473、2474、2475 或 2476 核酸序列或其互补物、类似物、衍生物、片段或部分组成的核酸分子，如长度为至少 15、20、25、30、50、100、250、500、750、1000 或更多连续核苷酸的寡核苷酸，其在严格条件下足以特异性与 hSARS mRNA 或基因组 RNA 杂交。

在另一优选的具体实施方案中，样品中 hSARS 病毒的存在是使用如下引物通过逆转录聚合酶链反应(RT-PCR)来检测的，所述引物
10 基于 hSARS 病毒(例如保藏检索号 CCTCC-V200303 的 hSARS 病毒)的基因组的部分核苷酸序列构建，或基于 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590 或 1965 核酸序列构建。在一个非限制性具体实施方案中，用于 RT-PCR 方法的优选引物是：5'-
15 TACACACCTCAGCGTTG-3' (SEQ ID NO : 3) 和 5'-CACGAACGTGACGAAT-3' (SEQ ID NO: 4)，在 2.5 mM MgCl₂ 存在下，热循环例如但不限于 94°C 8 分钟，随后 94°C 1 分钟、50°C 1 分钟、72°C 1 分钟的循环 40 次(另参见下文 6.7 节)。在更优选的具体实施方案中，本发明提供实时定量 PCR 试验以如下检测生物样品中
20 hSARS 病毒的存在：从样品抽提总 RNA，对抽提总 RNA 进行逆转录获得 cDNA，使用特定的引物如具有 SEQ ID NO: 3 和 4 核酸序列的引物以及荧光染料如 SYBR[®] Green I (其当非特异性与双链 DNA 结合时会发出荧光)，使 cDNA 经历 PCR 反应。在又一个优选具体实施方案中，本发明中使用的实时定量 PCR 为 TaqMan[®] 试验(参见上文 5
25 节)。明确地说，对于在生物样品中检测 hSARS 病毒存在的实时定量 PCR 试验，使用的引物优选为具有 SEQ ID NO: 2471 和 2472 核酸序列的引物。在这种情况下，通过 TaqMan[®] 探针检测扩增产物，优选所述探针具有 SEQ ID NO: 2473 核苷酸序列。TaqMan[®] 试验中使

用的另一种优选引物为具有 SEQ ID NO: 2474 和 2475 核酸序列的引物, 优选 TaqMan[®] 探针具有 SEQ ID NO: 2476 核苷酸序列。由于经一系列热循环后产生 PCR 产物, 在各延伸步骤结束时捕捉这些反应的荧光信号, 从而可以基于扩增图定量测定样品中的病毒负载量(参见下文 6.7 和 6.8 节)。

在另一个优选具体实施方案中, 样品中 hSARS 病毒的存在用荧光 cDNA 微阵列技术测定。一系列来源于 hSARS 病毒(例如保藏号 CCTCC-V200303 或具有 SEQ ID NO: 15 基因组核酸序列的病毒)的 cDNA 探针通过逆转录和使用合适引物的扩增制备, 所述引物基于例如所述 hSARS 病毒基因组的部分核苷酸序列或 SEQ ID NO: 1、11、13、15、16、240、737、1108、1590 或 1965 核酸序列构建。然后纯化这种扩增产物, 并将其固定化在芯片例如聚-L-赖氨酸包被的玻璃板上, 作为 cDNA 微阵列。从生物样品提取总 RNA, 在荧光标记的核苷酸存在下逆转录。然后表示样品中 mRNA 的标记 cDNA 与微阵列上的固定化 cDNA 探针接触, 检测并定量结合 cDNA 的荧光信号。已描述了各种 DNA 微阵列方法, 例如在 *Nucleic Acid Res.* 28(22): 4552-7 (Kane, M.D.等, 2000); *Science* 2000 Sep 8; 289(5485): 1757-60 (Taton, T.A.等, 2000); *Nature*, 405(6788): 827-836 (Lockhart, D.J.等, 2000)中。

优选检测 hSARS 病毒的另一种试剂是特异性结合本发明多肽或任何 hSARS 表位的抗体, 优选具有可检测标记的抗体。所述抗体可以是多克隆抗体, 或更优选为单克隆抗体。可使用完整的抗体或其片段(例如 Fab 或 F(ab')₂)。

涉及探针或抗体的术语“标记”包括通过将可检测物质偶联(即物理连接)到探针或抗体以对探针或抗体进行直接标记, 以及通过与被直接标记的另一试剂反应以对探针或抗体进行间接的标记。间接标记的实例包括用荧光标记的第二抗体检测第一抗体和用生物素对 DNA 探针进行末端标记, 这样就可用荧光标记的链霉抗生物素蛋白

来检测。本发明检测方法可用来体外或体内检测样品中的 mRNA、蛋白质(或任何表位)或基因组 RNA。例如体外检测 mRNA 的技术包括 Northern 杂交、原位杂交、RT-PCR 和 RNA 酶保护。体外检测 hSARS 病毒表位的技术包括酶联免疫吸附试验(ELISA)、蛋白质印迹、免疫沉淀法和免疫荧光法。体外检测基因组 RNA 的技术包括 Northern 杂交、RT-PCR 和 RNA 酶保护。此外,体内检测 hSARS 病毒的技术包括向对象生物体中引入针对多肽的标记抗体。例如,抗体可用放射性标志标记,所述放射性标志在对象生物体中的存在和位置可通过标准成像技术(包括放射自显影术)来检测。

10 在一个具体实施方案中,本发明方法进一步包括从对照对象获得对照样品,将对照样品与能检测 hSARS 病毒(例如本发明多肽或编码本发明多肽的 mRNA 或基因组 RNA)的化合物或试剂接触,从而检测样品中 hSARS 病毒或多肽或编码多肽的 mRNA 或基因组 RNA 的存在,并将对照样品中 hSARS 病毒或多肽或编码多肽的 mRNA 或基因组 RNA 的存在与待测样品中 hSARS 病毒或多肽或编码多肽的 mRNA 或基因组 RNA 的存在相比较。

20 在一个具体实施方案中,本发明提供包含核酸分子的诊断试剂盒,所述核酸分子适合用于检测 hSARS 病毒或其天然或人工变体、类似物或衍生物。在一个具体实施方案中,核酸分子具有 SEQ ID NO: 2471 和 2472 核酸序列。在具体实施方案中,核酸分子具有 SEQ ID NO: 2473 核酸序列。在另一个具体实施方案中,核酸分子具有 SEQ ID NO: 2474 和 2475 核酸序列。在具体实施方案中,核酸分子具有 SEQ ID NO: 2476 核酸序列。

25 本发明还包括检测待测样品中本发明 hSARS 病毒或多肽或核酸的存在的试剂盒。所述试剂盒可例如包含能检测待测样品中 hSARS 病毒或多肽或编码多肽的核酸分子的标记化合物或试剂,在某些实施方案中,还包含确定样品中多肽或 mRNA 的量的工具(例如结合多肽的抗体或与编码多肽的 DNA 或 mRNA 结合的寡核苷酸探针)。试

剂盒还可包含使用说明书。

对于基于抗体的试剂盒，其可包含例如：(1)第一抗体(例如连接到固体支持体上)，其结合本发明多肽或 hSARS 病毒表位；任选(2)不同的第二抗体，其结合多肽或第一抗体，且缀合到可检测试剂上。

- 5 对于基于寡核苷酸的试剂盒，其可包含例如：(1)寡核苷酸，例如可检测标记寡核苷酸，其可与编码本发明多肽的核酸序列或 hSARS 基因组内的序列杂交，或(2)一对引物，其可用于扩增含有 hSARS 序列的核酸分子。所述试剂盒还可包含例如缓冲剂、防腐剂或蛋白质稳定剂。所述试剂盒也可包含用于检测可检测的试剂所需的成分(例如酶或底物)。
- 10 所述试剂盒还可含有对照样品或一系列对照样品，所述对照样品可用来进行试验并与待测样品所含的相比较。试剂盒的各成分通常封闭在单独的容器中，所有的各种容器与使用说明书一起装在单一包装物中。

15 5.7. 筛选试验

本发明提供用于鉴定抑制 hSARS 病毒感染宿主或宿主细胞能力的化合物的方法。在某些实施方案中，本发明提供用于鉴定降低 hSARS 病毒在宿主或宿主细胞中复制能力的化合物的方法。可使用普通技术人员熟知的任何技术筛选破坏或降低 hSARS 病毒感染宿主

20 和/或在宿主或宿主细胞中复制的能力的化合物。

在某些实施方案中，本发明提供用于鉴定抑制 hSARS 病毒在哺乳动物或哺乳动物细胞中复制能力的化合物的方法。更具体地说，本发明提供用于鉴定抑制 hSARS 病毒感染哺乳动物或哺乳动物细胞能力的化合物的方法。在某些实施方案中，本发明提供用于鉴定抑制

25 抑制 hSARS 病毒在哺乳动物细胞中复制能力的化合物的方法。在一个具体实施方案中，哺乳动物细胞是人细胞。

在另一实施方案中，将细胞与测试化合物接触并用 hSARS 病毒感染。在某些实施方案中，对照培养物在测试化合物不存在下用

hSARS 病毒感染。细胞可在用 hSARS 病毒感染之前、同时或之后与测试化合物接触。在一个具体实施方案中，细胞是哺乳动物细胞。在一个甚至更具体实施方案中，细胞是人细胞。在某些实施方案中，细胞与测试化合物一起孵育至少 1 分钟、5 分钟、15 分钟、30 分钟、1 小时、2 小时、5 小时、12 小时或 1 天。病毒的滴度可在试验的任何时间进行测量。在某些实施方案中，测定培养物中病毒生长的时间进程。如果在测试化合物存在下病毒生长被抑制或降低，则测试化合物可鉴定为对抑制或降低 hSARS 病毒的生长或感染有效。在一个具体实施方案中，测试抑制或降低 hSARS 病毒生长的化合物抑制或降低其它病毒生长速率的能力，和/或测试其对 hSARS 病毒的特异性。

在一个实施方案中，将测试化合物给予模型动物，后者用 hSARS 病毒感染。在某些实施方案中，对照模型动物用 hSARS 病毒感染但不给予测试化合物。测试化合物可在用 hSARS 病毒感染之前、同时或之后给予。在一个具体实施方案中，模型动物是哺乳动物。在一个甚至更具体实施方案中，模型动物是可以是但不限于棉鼠、小鼠或猴。模型动物中的病毒滴度可在试验的任何时间进行测量。在某些实施方案中，测定培养物中病毒生长的时间进程。如果在测试化合物存在下病毒生长被抑制或降低，则测试化合物可鉴定为对抑制或降低 hSARS 病毒的生长或感染有效。在一个具体实施方案中，测试抑制或降低模型动物中 hSARS 病毒生长的化合物抑制或降低其它病毒生长速率的能力，以测试其对 hSARS 病毒的特异性。

6. 实施例

以下实施例说明了新型 hSARS 病毒的分离和鉴定。这些实施例不应被解释为对本发明的限制。

方法和结果

使用 Wiedbrauk DL 和 Johnston SLG 的 *Manual of Clinical Virology*, Raven Press, New York, 1993 作为一般参考文献。

6.1 临床对象

5 本研究包括共 50 名符合 SARS 的世界卫生组织(WHO)修正定义的患者, 他们于 2003 年 2 月 26 日至 3 月 26 日之间被香港特别行政区(HKSAR)的两个急性病地区医院收治(WHO. Severe acute respiratory syndrome (SARS) 2000, *Weekly Epidemiol Rec* 78: 81-83)。本研究还包括另外一名患者的肺部活组织检查, 该患者患典型的 SARS 并被
10 第三家医院收治。简单地说, SARS 的病例定义是: (i)发热 38°C 或更高; (ii)咳嗽或气短; (iii)胸部 X 光照片显示新的肺部浸润; 及(iv)具有与 SARS 患者的接触史或对典型和非典型肺炎的经验抗生素药物覆盖(β -内酰胺和大环内酯药物、氟喹诺酮或四环素)无反应。

 从所有患者收集鼻咽抽吸物和血清样品。从一些患者身上可同
15 时获得急性期和康复期的血清和粪便。对来自一个患者的肺部活组织进行处理, 供进行病毒培养、RT-PCR、常规组织病理学检查和电子显微镜检查。供进行其它疾病的微生物试验的鼻咽抽吸物、粪便和血清在盲试情况下加入本研究中, 作为对照物。

 由主治医师和临床微生物学者对医疗记录进行回顾审阅。进行
20 常规血液学、生物化学和微生物学检查, 包括血液和痰的细菌培养、血清学研究, 并收集鼻咽抽吸物进行病毒学试验。

6.2 细胞系

 FRhK-4 (胎猕猴肾)细胞在含 1%胎牛血清、1%链霉素和青霉素、
25 0.2%制霉菌素和 0.05%硫酸庆大霉素的极限必需培养基(MEM)中维持。

6.3 病毒感染

用来自两个患者(参见下文“结果”一节)、处于病毒运输培养基中的 200 μ l 临床(鼻咽抽吸物)样品感染 FRhk-4 细胞。将接种细胞在 37°C 下温育 1 小时。然后将含 1 μ g 胰蛋白酶的 1 ml MEM 加入到培养物中, 将感染细胞在提供 5%二氧化碳的 37°C 培养箱中温育。温育 2-4 天后, 观察感染细胞中出现的致细胞病变效应。使感染细胞传代成为新的 FRhK-4 细胞, 在接种后 1 天内观察致细胞病变效应。通过免疫荧光试验测试感染细胞的流感病毒 A、流感病毒 B、呼吸道合胞体病毒、副流感病毒 1 型、2 型和 3 型、腺病毒和人间质肺病毒(hMPV), 所有病例的试验结果均为阴性。还通过 RP-PCR 测试感染细胞的流感病毒 A 和人间质肺病毒, 结果为阴性。

6.4 病毒形态学

收集如上制备的感染细胞, 离心成颗粒状物, 处理细胞颗粒状物, 供进行薄切片透射电子显微镜检。在感染两个临床样本的细胞中鉴定出病毒颗粒, 但在不被病毒感染的对照细胞中没有。从感染细胞分离的病毒颗粒约有 70-100 纳米(图 2)。病毒衣壳主要发现于高尔基体和内质网的小泡中, 在细胞质中也有发现。在细胞膜中也发现病毒颗粒。

对一份病毒分离物进行超离心, 用磷钨酸对所得细胞颗粒状物进行负染色。这样显示了具有冠状病毒科特征的病毒颗粒。由于迄今公认的人类冠状病毒已知并不会导致类似疾病, 本发明人认为所述病毒分离物代表了感染人类的新型病毒。

6.5 抗体应答

为进一步确认此新型病毒在受感染患者中引起 SARS, 从 SARS 患者身上获取血清样品, 进行中和试验。将典型稀释的血清(x50、x200、x800 和 x1600)与用丙酮固定的感染 hSARS 病毒的 FRhK-4 细胞一起在 37°C 下温育 45 分钟。然后用磷酸缓冲盐水洗涤温育细胞,

用抗人 IgG-FITC 缀合抗体染色。然后洗涤细胞，在荧光显微镜下检查。在这些实验中，在 8 名 SARS 患者中发现阳性信号(图 3)，表明这些患者对此冠状病毒科新型人呼吸道病毒具有 IgG 抗体应答。与此对比，在 4 份阴性对照成对血清中没有信号检出。受试患者的抗 hSARS 抗体的血清滴度见表 1。

表 1

| 姓名 | 日期 | 实验编号 | 抗 SARS |
|------|---------------|------------|--------|
| 患者 A | 03 年 2 月 25 日 | S2728 | <50 |
| | 03 年 3 月 6 日 | S2728 | 1600 |
| 患者 B | 03 年 2 月 26 日 | S2441 | 50 |
| | 03 年 3 月 3 日 | S2441 | 200 |
| 患者 C | 03 年 3 月 4 日 | S3279 | 200 |
| | 03 年 3 月 14 日 | S3279 | 1600 |
| 患者 D | 03 年 3 月 6 日 | M41045 | <50 |
| | 03 年 3 月 11 日 | MB943703 | 800 |
| 患者 E | 03 年 3 月 4 日 | M38953 | <50 |
| | 03 年 3 月 18 日 | KWH03/3601 | 800 |
| 对照 F | 03 年 2 月 13 日 | M27124 | <50 |
| | 03 年 3 月 1 日 | MB942968 | <50 |
| 患者 G | 03 年 3 月 3 日 | M38685 | <50 |
| | 03 年 3 月 7 日 | KWH03/2900 | 可疑 |

盲试样品:

| | | |
|------|-----|------|
| 1a * | 急性期 | <50 |
| 1b | 康复期 | 1600 |
| 2a * | 急性期 | 50 |

| | | |
|------|-----|-------|
| 2b | 康复期 | >1600 |
| 3a * | 急性期 | 50 |
| 3b | 康复期 | >1600 |
| 4a * | 急性期 | <50 |
| 4b | 康复期 | <50 |
| 5a * | 急性期 | <50 |
| 5b | 康复期 | <50 |
| 6a * | 急性期 | <50 |
| 6b | 康复期 | <50 |

注: * SARS 患者

这些结果表明, 此冠状病毒科新成员是 SARS 的关键病原体。

5 6.6 hSARS 病毒的序列

感染后两天从感染或未感染 FrHK-4 细胞收获总 RNA。用 Superscript[®] II 逆转录酶(Invitrogen)按生产商的推荐在含 10 pg 简并引物(5'-GCCGGAGCTCTGCAGAATTCNNNNNNNN-3', SEQ ID NO: 5; N=A、T、G 或 C)的 20 μ l 反应混合物中对 100 ng 纯化 RNA 进行逆

10 转录。然后通过 QIAquick[®] PCR 纯化试剂盒按生产商的使用说明纯化逆转录产物, 洗脱到 30 μ l 的 10 mM Tris-HCl (pH 8.0)中。将 3 μ l 纯化 cDNA 产物加入到含以下成分的 25 μ l 反应混合物中: 2.5 μ l 10x PCR 缓冲液、4 μ l 25 mM MgCl₂、0.5 μ l 10 mM dNTP、0.25 μ l AmpliTaq Gold[®] DNA 聚合酶 (Applied Biosystems)、2.5 μ Ci [α -³²P]CTP

15 (Amersham)、2 μ l 10 μ M 引物(5'-GCCGGAGCTCTGCAGAATTC-3'; SEQ ID NO: 6)。按以下程序对反应进行热循环: 94 $^{\circ}$ C 8 分钟, 然后 94 $^{\circ}$ C 1 分钟、40 $^{\circ}$ C 1 分钟, 72 $^{\circ}$ C 2 分钟的循环 2 次。该热循环后, 进行 94 $^{\circ}$ C 1 分钟、60 $^{\circ}$ C 1 分钟、72 $^{\circ}$ C 1 分钟的循环 35 次。取 6 μ l PCR 产物进行 5%变性聚丙烯酰胺凝胶电泳分析。将凝胶对 X 光片曝光,

X 光片曝光过夜后显影。将只在感染细胞样品中鉴定出的独特 PCR 产物从凝胶中分离出来，用 50 μ l 1x TE 缓冲液洗脱。然后将洗脱的 PCR 产物在含有以下成分的 25 μ l 反应混合物中再扩增：2.5 μ l 10x PCR 缓冲液、4 μ l 25 mM $MgCl_2$ 、0.5 μ l 10 mM dNTP、0.25 μ l AmpliTaq Gold[®] DNA 聚合酶(Applied Biosystems)、1 μ l 10 μ M 引物(5'-GCCGGA GCTCTGCAGAATTC-3': SEQ ID NO: 6)。按以下程序对反应进行热循环：94 $^{\circ}$ C 8 分钟，然后 94 $^{\circ}$ C 1 分钟、60 $^{\circ}$ C 1 分钟、72 $^{\circ}$ C 1 分钟的循环 35 次。用 TOPO TA Cloning[®]试剂盒(Invitrogen)克隆 PCR 产物，将连接的质粒转化到 TOP 10 大肠杆菌(*E.coli*)感受态细胞 (Invitrogen)中。PCR 插入片段通过 BigDye[®]循环测序试剂盒按生产商 (Applied Biosystems)的建议进行测序，测序产物通过自动测序仪 (Applied Biosystems, 型号 3770)进行分析。获得的序列(SEQ ID NO: 1)见图 1。从获得的 DNA 序列推断的氨基酸序列(SEQ ID NO: 2)显示，其与已鉴定冠状病毒的聚合酶蛋白有 57%的同源性。

与此类似，从 hSARS 病毒获得另外两种部分序列(SEQ ID NO: 11 和 13)及推断的氨基酸序列(分别为 SEQ ID NO: 12 和 14)，见图 8 (SEQ ID NO: 11 和 12)和图 9 (SEQ ID NO: 13 和 14)。

hSARS 病毒的整个基因组序列见图 10 (SEQ ID NO:15)。以全部三种读框获得的 SEQ ID NO:15 的推断氨基酸序列见图 11 (DNA 序列见 SEQ ID NO: 16、240 和 737；氨基酸序列分别见 SEQ ID NO: 17-239、241-736 和 738-1107)。以全部三种读框获得的 SEQ ID NO:15 的互补物的推断氨基酸序列见图 12 (DNA 序列见 SEQ NO: 1108、1590 和 1965；氨基酸序列分别见 SEQ ID NO: 1109-1589、1591-1964 和 1966-2470)。

25

6.7 鼻咽抽吸物中 hSARS 病毒的检测

首先，通过快速免疫荧光抗原检测法检查鼻咽抽吸物(NPA)中的流感病毒 A 和 B、副流感病毒 1 型、2 型和 3 型、呼吸道合胞体病毒

和腺病毒(Chan KH, Maldeis N, Pope W, Yup A, Ozinskas A, Gill J, Seto WH, Shortridge KF, Peiris JSM. Evaluation of Directigen Fly A+B test for rapid diagnosis of influenza A and B virus infections. *J Clin Microbiol.* 2002; **40**: 1675-1680), 在 Mardin Darby 犬肾细胞、LLC-Mk2、RDE、
5 Hep-2 和 MRC-5 细胞中培养鼻咽抽吸物的常规呼吸道病原体 (Wiedbrauk DL, Johnston SLG. *Manual of clinical virology*. Raven Press, New York.1993)。随后, 将胎猕猴肾细胞(FRhk-4)和 A-549 细胞加入到所用的一系列细胞系中。直接对临床样本进行逆转录聚合酶链反
10 应(RT-PCR)以检测流感病毒 A (Fouchier RA, Bestebroer TM, Herfst S, Van Der Kemp L, Rimmelzwan GF, Osterhaus AD. Detection of influenza A virus from different species by PCR amplification of conserved sequences in the matrix gene. *J Clin Microbiol.* 2000; **38**: 4096-101)和
15 人间质肺病毒(HMPV)。用于 HMPV 的引物为: 第一轮, 5'-AARGTSAATGCATCAGC-3' (SEQ ID NO : 7) 和 5'-CAKATTYTGCTTATGCTTTC-3' (SEQ ID NO: 8); 及嵌套引物: 5'-ACACCTGTTACAATACCAGC-3' (SEQ ID NO : 9) 和 5'-GACTTGAGTCCCAGCTCCA-3' (SEQ ID NO: 10)。嵌套 PCR 产物的大小是 201 bp。用针对支原体的 ELISA 筛选细胞培养物(Roche Diagnostics GmbH, Roche, Indianapolis, USA)。

20 6.7.1. RT-PCR 试验

对来自两个患者的 hSARS 病毒进行培养和遗传测序后(参见上文 6.6 节), 开发出用于从 NPA 样品检测 hSARS 病毒的 RT-PCR。来自临床样品的总 RNA 用随机六聚体进行逆转录, cDNA 用引物 5'-
25 TACACACCTCAGCGTTG-3' (SEQ ID NO: 3)和 5'-CACGAACGTG-ACGAAT-3' (SEQ ID NO: 4)在 2.5 mM MgCl₂ 存在下扩增(94°C 8 分钟, 然后 94°C 1 分钟、50°C 1 分钟、72°C 1 分钟的循环 40 次), 所述两种引物是基于 hSARS 病毒基因组构建的。

典型的 RT-PCR 方案概述如下:

RNA 提取

通过 QIAquick[®]病毒 RNA 提取试剂盒从 140 μl NPA 样品中提取 RNA，洗脱在 50 μl 洗脱缓冲液中。

逆转录

| | |
|---|--------------------|
| RNA | 11.5 μl |
| 0.1 M DTT | 2 μl |
| 5x 缓冲液 | 4 μl |
| 10 mM dNTP | 1 μl |
| Superscript II, 200 U/ μl (Invitrogen) | 1 μl |
| 随机六聚体, 0.3 $\mu\text{g}/\mu\text{l}$ | 0.5 μl |

反应条件

42°C, 50 分钟

94°C, 3 分钟

4°C

PCR

5 如下在 50 μl 反应物中扩增通过随机引物产生的 cDNA:

| | |
|--|---------------------|
| cDNA | 2 μl |
| 10 mM dNTP | 0.5 μl |
| 10x 缓冲液 | 5 μl |
| 25 mM MgCl_2 | 5 μl |
| 25 μM 正向引物 | 0.5 μl |
| 25 μM 反向引物 | 0.5 μl |
| AmpliTaq Gold [®] 聚合酶, 5U/ μl (Applied Biosystems) | 0.25 μl |
| 水 | 36.25 μl |

热循环条件: 95°C 10 分钟, 然后 95°C 1 分钟、50°C 1 分钟、72°C 1 分钟的循环 40 次。

引物序列

10 引物基于 hSARS 病毒的 RNA 依赖性 RNA 聚合酶编码序列 (SEQ ID NO: 1) 设计。

正向引物: 5' TACACACCTCAGCGTTG 3' (SEQ ID NO: 3)

反向引物: 5' CACGAACGTGACGAAT 3' (SEQ ID NO: 4)

产物(扩增子)大小: 182 bp

实时定量 PCR 试验

通过 QIAamp[®]病毒 RNA 微型试剂盒(Qiagen)按生产商的说明从
5 140 µl 鼻咽抽吸物(NPA)提取总 RNA。在含有 0.15 µg 随机六聚体、10
mmol/L DTT 和 0.5 mmol/L dNTP 的 20 µl 反应混合物中, 将 10 µl 洗
脱的 RNA 样品用 200 U Superscript[®]II 逆转录酶(Invitrogen)按说明进
行逆转录。然后互补 DNA 在 SYBR[®] Green I 荧光反应(Roche)混合物
中进行扩增。简单地说, 含有 2 µl of cDNA、3.5 mmol/L MgCl₂、0.25
10 µmol/L 正向引物(5'-TACACACCTCAGCGTTG-3'; SEQ ID NO: 3)和
0.25 µmol/L 反向引物(5'-CACGAACGTGACGAAT-3'; SEQ ID NO: 4)
的 20 µl 反应混合物中, 用 Light-Cycler (Roche)按 PCR 程序[95°C 10
分钟, 然后 95°C 10 分钟; 57°C 5 秒; 72°C 9 秒的循环 50 次]进行热
循环。含有目标序列的质粒用作阳性对照。这些反应物的荧光信号
15 在每个循环的延伸步骤结束时捕捉。为确定试验的特异性, 在试验
结束时对 PCR 产物(184 个碱基对)进行解链曲线分析(65°C 至 95°C, 每
秒 0.1°C)。

临床结果

临床发现:

所有 50 个 SARS 患者均为中国种族。他们代表了 5 个不同的流
行病学相关群以及符合病例定义的其他偶发病例。他们平均在症状
开始后 5 天住院。他们的平均年龄是 42 岁(23 岁至 74 岁), 女性与男
性的比例是 1.3。其中 14 人(28%)是医护人员, 5 人(10%)有到严重
25 暴发 SARS 的医院的探望史。13 人(26%)在家庭接触 SARS 患者, 另
外 12 人(24%)在社会上接触过 SARS 患者。4 人(8%)有最近到中国大
陆的旅行史。

大部分患者的主诉是发热(90%)和气短。半数以上的患者出现咳

嗽和肌痛(表 2)。少数患者出现上呼吸道症状,如鼻溢(24%)和喉咙痛(20%)。腹泻(10%)和食欲减退(10%)也有报道。最初听诊检查发现,只有 38%的患者出现捻发音和进气减少。62%的患者报告有干咳。所有患者在接诊时经放射检查均发现实变迹象,包括 1 个区域(36 例)、
5 2 个区域(13 例)和 3 个区域(1 例)。

表 2

| 临床症状 | 患者数目(百分比) |
|-------|-----------|
| 发热 | 50 (100%) |
| 发冷或寒战 | 37 (74%) |
| 咳嗽 | 31 (62%) |
| 肌痛 | 27 (54%) |
| 身体不适 | 25 (50%) |
| 流鼻涕 | 12 (24%) |
| 喉咙痛 | 10 (20%) |
| 气短 | 10 (20%) |
| 食欲减退 | 10 (20%) |
| 腹泻 | 5 (10%) |
| 头痛 | 10 (20%) |
| 头晕 | 6 (12%) |

* 1 名患者出现躯干斑丘疹。

10 大部分患者(98%)尽管发高烧,却没有白细胞增多的迹象。外周血检查发现淋巴细胞减少(68%)、白细胞减少(26%)、血小板减少(40%)和贫血(18%)(表 3)。肝实质酶丙氨酸转氨酶(ALT)和肌肉酶肌酸酐激酶(CPK)的水平分别在 34%和 26%的病例中升高。

表 3

| 化验参数 | 平均值(范围) | 异常百分比 | 正常范围 |
|------|---------|-------|------|
|------|---------|-------|------|

| | | |
|---|----------------|------------------------------|
| 血红蛋白 | 12.9(8.9-15.9) | 11.5-16.5 g/dl |
| 贫血 | 9(18%) | |
| 白细胞计数 | 5.17(1.1-11.4) | 4-11 x 10 ⁹ /L |
| 白细胞减少 | 13(26%) | |
| 淋巴细胞计数 | 0.78(0.3-1.5) | 1.5-4.0 x 10 ⁹ /L |
| 显著淋巴细胞减少 (<1.0 x 10 ⁹ /L) | 34(68%) | |
| 血小板计数 | 174(88-351) | 150-400 x 10 ⁹ /L |
| 血小板减少 | 20(40%) | |
| 丙氨酸转氨酶(ALT) | 63(11-350) | 6-53 U/L |
| ALT 升高 | 17(34%) | |
| 白蛋白 | 37(26-50) | 42-54 g/dl |
| 低白蛋白 | 34(68%) | |
| 球蛋白 | 33(21-42) | 24-36 g/dl |
| 球蛋白升高 | 10(20%) | |
| 肌酸酐激酶 | 244(31-1379) | 34-138 U/L |
| 肌酸酐激酶升高 | 13(26%) | |

5 通过培养、抗原检测和 PCR 对已知病毒和细菌进行的常规微生物学检查在大部分病例中为阴性。一名接入重病监护室的 74 岁男性患者的血液培养发现大肠杆菌阳性，这是因为医院获得性尿道感染。另外两名患者入院时从其痰样本中分离出肺炎克雷伯氏菌(*Klebsiella pneumoniae*)和流感嗜血菌(*Hemophilus influenzae*)。

10 给 9 名患者每 24 小时口服 500 mg 左氧氟沙星，给另外 40 名患者静脉注射(每 8 小时 1.2 g)/口服(375 mg，每日三次)阿莫西林-克拉维酸盐并每 12 小时静脉注射/口服 500 mg 克拉霉素。给 4 名患者每日两次口服 75 mg 奥塞米韦。给 1 名患者每 24 小时静脉注射 2 gm 头孢曲松，每 24 小时口服 500 mg 阿奇霉素，每日两次口服 100 mg

金刚烷胺，以对典型和非典型肺炎进行经验覆盖。

5 19 名患者发展为带有氧去饱和的严重疾病，需要接受重病监护和通气支持。自症状开始起病情恶化的平均天数是 8.3 天。症状开始后平均 6.7 天给予 49 名患者每 8 小时静脉注射利巴韦林 8 mg/kg 以及类固醇。

10 与需要接受重病监护和通气支持的严重并发症有关的风险因素是年老、淋巴细胞减少、ALT 受损及延迟给予利巴韦林和类固醇(表 4)。所有并发病例在接入重病监护室后用利巴韦林和类固醇治疗，而所有无并发病例在普通病房中就用利巴韦林和类固醇治疗。正如所料，31 个无并发病例痊愈或好转，而 8 个并发病例病情恶化，其中 1 个在本说明书写作时死亡。所有 50 名患者在本说明书写作时平均接受监控 12 天。

表 4

| | 并发病例 (n= 19) | 无并发病例 (n= 31) | P 值 |
|---------------------|-----------------|------------------|----------|
| 平均(SD)年龄(范围) | 49.5 ± 12.7 | 39.0 ± 10.7 | P < 0.01 |
| 男性/女性比例 | 8/11 | 14/17 | N.S |
| 潜在疾病 | 5 † | 1 ‡ | P < 0.05 |
| 接触方式 | | | |
| 到中国旅行 | 1 | 3 | N.S. |
| 医护人员 | 5 | 9 | N.S. |
| 到医院探访 | 1 | 4 | N.S. |
| 家庭接触 | 8 | 5 | P < 0.05 |
| 社会接触 | 4 | 10 | N.S. |
| 至入院时症状平均(SD)持续时间(天) | 5.2 ± 2.0 | 4.7 ± 2.5 | N.S. |
| 平均(SD)入院温度(°C) | 38.8 ± 0.9 | 38.7 ± 0.8 | N.S. |

| | | | |
|--|----------------|----------------|-----------|
| 平均(SD)初始总外周 WBC 计数($\times 10^9/L$) | 5.1 \pm 2.4 | 5.2 \pm 1.8 | N.S. |
| 平均(SD)初始淋巴细胞计 数($\times 10^9/L$) | 0.66 \pm 0.3 | 0.85 \pm 0.3 | P < 0.05 |
| 血小板减少的存在 ($< 150 \times 10^9/L$) | 8 | 12 | N.S. |
| 肝脏功能受损检查 | 11 | 6 | P < 0.01 |
| CXR 改变(受影响区域的 数量) | 1.4 | 1.2 | N.S. |
| 自症状出现起病情恶化的 平均(SD)天数 § | 8.3 \pm 2.6 | 不适用 | |
| 自症状出现起至开始给予 利巴韦林和类固醇的平均 (SD)天数 | 7.7 \pm 2.9 | 5.7 \pm 2.6 | P < 0.05 |
| 病情恶化后开始给予利巴 韦林和类固醇 | 12 | 0 | P < 0.001 |
| 对利巴韦林和类固醇有反 应 结果 | 11 | 28 | P < 0.05 |
| 好转或痊愈 | 10 | 31 | P < 0.01 |
| 没有好转 | 8 | 0 | P < 0.01 |

* 由于病例数量少, 没有进行多变量分析;

† 2名患者有糖尿病, 1名有肥厚性梗阻性心肌症, 1名有慢性活动性乙型肝炎, 1名有脑肿瘤;

‡ 1名患者有原发性高血压;

§ 去饱和作用需要重病监护支持;

|| 1名患者死亡。

5

从两名患者身上分离了两个病毒分离物，后来鉴定为冠状病毒科成员(参见下文)。一个病毒分离物来自一名 53 岁的中国香港居民的切开肺活组织检查组织，另一个病毒分离物来自一名以往健康良好的 42 岁女性的鼻咽抽吸物。该 53 岁男性与一名来自广州、后来死于 SARS 的中国游客有 10 个小时的家庭接触史。接触两天后，他就出现发热、身体不适、肌痛和头痛症状。肺右下区有捻发音，胸部放射照片显示有相应的肺泡阴影。血液学检查显示淋巴细胞减少，为 $0.7 \times 10^9/L$ ，总白细胞和血小板计数正常。ALT (41 U/L)和 CPK (405 U/L)均受损。尽管他口服了阿奇霉素、金刚烷胺并静脉注射了头孢曲松，仍出现两侧肺部浸润物增加以及进行性氧气去饱和。因此，他入院 9 天后进行切开肺活组织检查。组织病理学检查显示中等间质炎症，分散的肺泡细胞呈现细胞巨大化、粒状两染性细胞质、细胞核增大、核仁突出。没有细胞显示出疱疹病毒或腺病毒感染的典型包涵体。该患者进行手术后需接受通气和重病监护。经验性给他静脉注射了利巴韦林和氢化可的松。但他还是在入院 20 天后死亡。在回顾时在其鼻咽抽吸物、肺活组织检查和死后的肺中发现冠状病毒样 RNA。他抗自身 hSARS 分离物的抗体滴度显著升高，从 1/200 升至 1/1600。

第二名分离出 hSARS 病毒的患者是一个以往健康良好的 42 岁女性。她曾到中国大陆的广州出行两天，回到香港五天后出现发热和腹泻症状。对她进行身体检查显示，肺右下区有捻发音，胸部放射照片显示有相应的肺泡阴影。检查还显示白细胞减少($2.7 \times 10^9/L$)、淋巴细胞减少($0.6 \times 10^9/L$)和血小板减少($104 \times 10^9/L$)。尽管用阿莫西林-克拉维酸盐、克拉霉素和奥塞米韦给她进行经验性抗微生物覆盖，她在入院后五天病情还是恶化，需要接受机械通气和重病监护达五天。随后她逐渐好转，不需接受利巴韦林或类固醇治疗。其鼻咽抽吸物在 RT-PCR 试验中对病毒为阳性，并且她实现血清转化，抗 hSARS 分离物抗体滴度从 $<1/50$ 至 $<1/1600$ 。

病毒学发现:

在 FRhk-4 细胞上分别从上述两名患者的肺活组织检查和鼻咽抽
吸物分离出病毒。接种后 2 天至 4 天出现初始致细胞病变效应，但
5 经随后的传代，致细胞病变效应在 24 小时内出现。两个病毒分离物
均不与用于鉴定病毒分离物的一系列常规试剂反应，包括用于鉴定
流感病毒 A、B、副流感病毒 1 型、2 型和 3 型、腺病毒和呼吸道合
胞体病毒的试剂(DAKO, Glostrup, 丹麦)。所述两个病毒分离物也不在
10 用于流感病毒 A 和 HMPV 的 RT-PCR 试验中反应，或在用于支原体的
PCR 试验中反应。病毒对乙醚敏感，表明它是有包膜病毒。对通
过超离心获得的负染色(2%磷钨酸钾, pH 7.0)细胞培养提取物进行电
子显微镜检查发现，存在多型有包膜病毒颗粒，直径约为 80-90 nm
(70-130 nm 的范围)，其表面形态学似乎可与冠状病毒科成员相比(图
5A)。对感染细胞进行薄切片电子显微镜检查显示，细胞质中的滑壁
15 小泡内有直径为 55-90 nm 的病毒颗粒(图 5A 和 5B)。在细胞表面也
可见病毒颗粒。总体发现与冠状病毒科病毒造成的细胞感染一致。

所述 53 岁男性的肺活组织检查薄切片电子显微照片显示，其脱
屑细胞的细胞质中含有 60-90 nm 病毒颗粒。这些病毒颗粒在大小和
形态学上与在来自两名患者的细胞培养病毒分离物中观察到的病毒
20 颗粒相似(图 4)。

对随机引物 RT-PCR 试验中产生的 RT-PCR 产物进行了分析，
对病毒感染样本中发现的特有条带进行了克隆和测序。在检查的 30
个克隆中，鉴定出含未知来源的 646 个碱基对(SEQ ID NO:1)的克隆。
对此 DNA 片段进行测序分析表明，此序列与冠状病毒科家族的病毒
25 有弱的同源性(数据未给出)。但是从此未知序列推断的氨基酸序列
(215 个氨基酸: SEQ ID NO:2)与牛冠状病毒和鼠肝炎病毒的 RNA 聚
合酶具有高度的同源性(57%)，确证此病毒属于冠状病毒科家族。蛋
白质序列的系统发生学分析显示，此病毒虽然与冠状病毒 II 类最紧

密相关，但却是截然不同的病毒(图 5A 和 5B)。

5 根据此分离物的 646 个碱基对序列，设计了用于检测新病毒的特异性引物，供在临床样本中对此 hSARS 病毒基因组进行 RT-PCR 检测。从 50 名 SARS 患者获得的 44 个鼻咽抽吸物样本中，22 个样本有 hSARS RNA 迹象。在 18 个受检的粪便样品中有 10 个可检测出病毒 RNA。RT-PCR 反应的特异性通过对选出的阳性 RT-PCR 扩增产物进行测序来确认的。来自无关疾病患者的 40 个鼻咽抽吸物和粪便样本中没有一个是 RT-PCR 试验中具有反应性。

10 为确定实时定量 PCR 的动态范围，制备含目标序列的质粒 DNA 的连续稀释物，对其进行实时定量 PCR 试验。如图 7A 所示，该试验能够检测少至 10 个拷贝的目标序列。与之相比，在水对照中没有观察到信号(图 7A)。在 29 名血清学确认的 SARS 患者中有 23 名观察到阳性信号。在所有这些阳性病例中，观察到与阳性对照的信号相对应的独特 PCR 产物($T_m = 82^\circ\text{C}$)(图 7B，数据未给出)。这些数据
15 表明，此试验对目标具有高度的特异性。这些反应中目标序列的拷贝数从 4539 至少于 10。因此，在 1 ml NPA 样品中可发现高达 6.48×10^5 个拷贝的此病毒序列。上述阳性病例中有 5 个可以在血清转化前收集 NPA 样品。在这些样品中的 3 个检测出病毒 RNA，表明此试验甚至在感染发作的早期就可检测出病毒。

20 为进一步确认此试验的特异性，征集健康人(n=11)和感染腺病毒(n=11)、呼吸道合胞体病毒(n=11)、人间质肺病毒(n=11)、流感病毒 A (n=13)或流感病毒 B (n=1)的患者的 NPA 样品作为阴性对照。所有这些样品除一个外，试验结果均为阴性。假阳性病例在随后的测试中为阴性。包括假阳性的病例在内，实时定量 PCR 试验的灵敏度是
25 79%，特异性是 98%。

流行病学数据表明，飞沫传播是此病毒的主要传播途径之一。本研究从 NPA 样品中检测出活病毒和高拷贝的病毒序列，明确支持 SARS 患者的咳嗽和喷嚏飞沫可能是此传染媒介的主要来源。有趣的

是, 本研究中来自 SARS 患者的 4 个粪便样品中有 2 个在试验中呈阳性(数据未给出)。粪便中检测出病毒表明可能存在其它传播途径。有关指出的是, 某些动物冠状病毒是通过粪-口途径传播的(McIntosh K.,1974, *Coronaviruses: a comparative review. Current Top Microbiol Immunol.* 63:85-112)。但是, 需要进一步的研究以检验粪便中的病毒是否有传染性。

除此 hSARS 病毒外, 目前还有两种已知的人冠状病毒血清群(229E 和 OC43)(Hruskova J.等, 1990, *Antibodies to human coronaviruses 229E and OC43 in the population of C.R., Acta Virol.* 34: 346-52)。用于本试验的引物对与 229E 毒株没有同源性。由于在 Genebank 中不能获得相应的 OC43 序列, 不知道这些引物是否能与此毒株发生交叉反应。但是, 对在 OC43 聚合酶基因其它区域中可获得的序列进行的序列分析表明, 与 SARS 有关的新型人病毒在遗传上与 OC43 截然不同。此外, 在本研究中使用的引物与已知冠状病毒的任何序列都没有同源性。因此, 这些引物不大可能会与 OC43 毒株发生交叉反应。

有报道说, 除所述新型病原体外, 在某些 SARS 患者身上还鉴定出间质肺病毒(Center for Disease Control and Prevention, 2003, *Morbidity and Mortality Weekly Report* 52:269-272)。在本研究的任何患者中没有检出间质肺病毒感染的任何迹象(数据未给出), 表明本发明新型 hSARS 病毒是 SARS 发病机理中的主要参与者。

免疫荧光抗体测定:

在来自 SARS 患者的 50 份最新血清样品中有 35 份具有抗 hSARS 抗体的迹象(参见图 3)。可获得成对急性期和康复期血清的 27 名患者都发生血清转化或其抗病毒抗体滴度均提高>4 倍。来自本研究组以外群体的其它 SARS 患者的其它 5 对血清也进行了检测, 以在社会上对 SARS 患者进行更广泛的取样, 他们全部都发生血清转化。来自呼吸道疾病患者或其它疾病患者的 80 份血清以及 200 个正常献血

者均没有可检测的抗体。

如果单一血清中对 HP-CV 血清阳性或在 NPA 或粪便中检出病毒 RNA 均被认为是感染 hSARS 病毒的证据, 那么 50 名患者中有 45 名具有感染迹象。在 5 名没有任何冠状病毒科病毒感染的病毒学证据的患者当中, 只有一名患者在临床症状发作>14 天后接受了血清检验。

6.8 用于检测 hSARS 病毒的定量 TaqMan[®] 试验

6.8.1. 材料和方法

10 患者和样品收集

在本研究中使用来自 50 个满足 SARS 临床 WHO 病例定义 (<http://www.who.int/csr/sars/casedefinition/en/>) 的患者的贮藏临床样品, 这些患者的诊断随后通过血清转变证实。从症状开始的 1-3 天入上述收集 NPA 样品(Poon 等, 2003, Clin. Chem. 49: 953-955)。使用来自无关疾病患者的 NPA 样品作为对照。

15 RNA 提取和逆转录

用 QIAamp[®] 病毒 RNA 微型试剂盒(Qiagen)按厂商说明从临床样品提取 RNA。在上述常规 RT-PCR 试验中, 140 μ l NPA 用于 RNA 提取。在改进的 RNA 提取方案中, 540 μ l NPA 用于 RNA 提取。提取的 RNA 最终洗脱到 30 μ L 无 RNA 酶水中, 储存在-20 $^{\circ}$ C。然后用随机六聚体逆转录来自临床样品的总 RNA。

20 用于 SARS-CoV 的常规 PCR

常规 PCR 试验如 6.7.1. 节中所述进行。

25 用于 SARS-CoV 的实时定量 PCR 试验

在本研究中使用特异性针对 SARS-CoV 1b 区的实时定量 PCR。通过 TaqMan[®] PCR Core Reagent 试剂盒在 7000 Sequence Detection System (Applied Biosystems) 中扩增互补 DNA。简单地说, 在含有下述成分的 25 μ l 反应物中扩增 4 μ l cDNA: 0.625 U AmpliTaq Gold[®] 聚

合酶(Applied Biosystems)、2.5 μ l 10x TaqMan[®]缓冲液 A、0.2 mM dNTP、5.5 mM MgCl₂、2.5 U AmpErase[®] UNG 和 1x 引物-探针混合物 (Assays by Design, Applied Biosystems)。引物序列为 5'-CAGAACGCTGTAGCTTCAAAAATCT-3' (SEQ ID NO: 2471)和 5'-TCAGAACCCTGTGATGAATCAACAG-3' (SEQ ID NO: 2472), 探针为 5'-(FAM)TCTGCGTAGGCAATCC(NFQ)-3' (SEQ ID NO: 2473; FAM, 6-羧基荧光素; NFQ, 无荧光猝灭剂)。反应最初在 50°C 孵育 2 分钟, 然后在 95°C 孵育 10 分钟。反应物然后进行热循环 45 次(95°C 15 秒、60°C 1 分钟)。含有目标序列的质粒用作阳性对照。

6.8.2. 结果

研究从血清学证实的 SARS 患者收集的其疾病最初 3 天的共 50 个分离 NPA 样品。它们之中, 11 个(22%)在我们前面报道的常规 RT-PCR 试验中为阳性(参见 6.7.1 节)(表 5)。

表 5

| 开始天数 | 样本大小 | 阳性数量 | | |
|------|------|--------------|------------------------------|-------------------------------|
| | | 常规 RT-PCR 试验 | 用改进 RNA 提取方案进行的常规 RT-PCR 试验* | 用改进 RNA 提取方案进行的实时 RT-PCR 试验** |
| 1 | 8 | 0 (0%) | 2 (25%) | 5 (63%) |
| 2 | 16 | 3 (19%) | 8 (50%) | 14 (88%) |
| 3 | 26 | 8 (31%) | 12 (46%) | 21 (81%) |

*试验的总检测率与常规 RT-PCR 试验有统计学差异(McNemar 检验, $P < 0.001$)

**试验的总检测率与用改进 RNA 提取方案进行的常规 RT-PCR 试验有统计学差异(McNemar 检验, $P < 0.0001$)

我们推论, SARS-CoV RT-PCR 检测在疾病早期的低敏感性可通过将最初 NPA 样品的提取体积从 140 μ l 增加到 560 μ l 而得到增强。

使用这种改进的 RNA 提取方案,传统 RT-PCR 试验的敏感性从 11/50 加倍至 22/50 (表 5)。改进 RT-PCR 方案的总检测率与我们的第一代 RT-PCR 方案有统计学差异(McNemar 检验, $P < 0.001$, 表 5)。对于 30 个阴性对照样品,观察到一个假阳性结果。采用 RNA 提取改进,常规 RT-PCR 对疾病头 3 天收集的样品的敏感性和特异性分别为 44.0% 和 96.6%。

为进一步改进发病早期样品中 SARS-CoV 的检测,我们采用了高度敏感的实时定量试验进行 SARS-CoV 检测(图 14)。采用改进的 RNA 提取方案,实时试验中 50 个 NPA 样品的 40 个为阳性(图 15 和表 5)。改进 RT-PCR 方案的总检测率与其它两种试验有统计学差异(McNemar 检验, $P < 0.0001$, 表 5)。具体地说,疾病开始第 1 天分离的 NPA 样品在实时定量 RT-PCR 试验中 63%为阳性。与之相比,第 1 天分离的样品在传统 RT-PCR 试验中无阳性。对于第 2-3 天分离的样品,超过 81%的样品在定量试验中为阳性(表 5)。采用改进的 RNA 提取方案和实时 PCR 技术,定量试验对早期 SARS 样品的敏感性和特异性分别为 80%和 100%。

实时试验也允许定量这些临床样品的病毒载量(1 拷贝/反应=27.8 拷贝/ml NPA 样品)。如图 16 中所示,疾病的进展导致 NPA 中病毒载量升高(空白条)。此外,我们进一步调查了我们第一代 RT-PCR 试验中的阴性临床样品($N = 39$)的病毒载量(图 16, 灰色条)。如期望的,这些临床样品的病毒载量(灰色条)比整体的全面病毒载量(空白条)低得多。

6.8.3. 讨论

此研究我们的目标是建立用于检测 SARS-CoV 的高敏感性 RT-PCR 试验。具体地说,我们集中于检测症状开始 1-3 天分离的样品中的 SARS-CoV RNA。使用我们第一代常规 RT-PCR 试验,仅有 22% 的这些样品表现出具有 SARS-CoV RNA。为了建立更敏感的试验,我们改进了 RNA 提取方法,并在我们当前的研究中采用了定量技术。

通过将用于 RNA 提取的最初体积从 140 μl 增加到 540 μl , 阳性病例的比例增加到 44%。此外, 通过在改进试验中进一步使用实时定量 PCR 技术, 80%的早期 SARS 样品成为阳性。更重要的是, 实时定量试验中 5'核酸酶探针的使用会最小化假阳性率, 这是因为信号特异性增加。结合起来, 本研究的结果表明, 我们改进的 RT-PCR 试验允许早期和准确地诊断 SARS。

我们改进 RT-PCR 试验的定量结果进一步提供了关于这些临床样品中 SARS-CoV 病毒载量的信息。我们的结果表明, 病毒载量随疾病的进展而增加。对于在第一代 RT-PCR 试验中为阴性的样品, 所有都含有极少量的病毒 RNA (图 15 和 16)。该现象解释了为什么大多数这些样品在使用我们的第一代 RT-PCR 试验时为阴性。有趣的是, 对于在第一代试验中为阳性的样品, 某些含有极高量的病毒 RNA (图 16)。

总之, 通过增加用于 RNA 提取的最初样品体积和使用实时定量 PCR 技术, 我们建立了敏感和准确的 RT-PCR 试验用于即时鉴定 SARS-CoV。期望的是, 采用这种快速诊断方法, 这种病原体的即时鉴定会促进疾病的控制和即时治疗的制度。

6.9. 临床观察和讨论

SARS 的暴发在许多方面是不寻常的, 尤其是在医护工作者和家庭接触中集中出现肺炎患者。在这一系列 SARS 患者当中, 对非典型肺炎的常规病原体的检查证实为阴性。但是从分别获自两名 SARS 患者的肺活组织检查和鼻咽抽吸物中分离出属于冠状病毒科家族的病毒。该病毒在系统发生学上与任何已知的人或动物冠状病毒或环状病毒都不紧密相关。本分析基于聚合酶基因的 646 碱基对片段 (SEQ ID NO:1), 表明该病毒与冠状病毒的抗原 2 类以及鼠肝炎病毒和牛冠状病毒相关。但是, 冠状病毒科病毒能够在病毒家族内部进行异源重组, 故有必要对新型病毒的基因组的其它部分进行遗传分析, 然

后更加确凿地定义此病毒的本质(Holmes KV. Coronaviruses. Eds Knipe DM, Howley PM Fields Virology, 第4版, Lippincott Williams & Wilkins, Philadelphia, 1187-1203 页)。将生物学、遗传学和临床数据综合在一起,表明此新型病毒不是两种已知人冠状病毒的任一种。

5 患上临床定义的 SARS 的大部分患者(90%)有感染此病毒的血清学和 RT-PCR 证据。与此对比,在健康人对照中没有抗体或病毒 RNA 可检测出。可获得急性期和康复期血清的所有 27 名患者均显示抗 hSARS 病毒的抗体滴度升高,这加强了以下论点,即最近感染此病毒是 SARS 发展中的必要因素。另外,来自香港其它医院的患者的
10 所有 5 对急性期和康复期血清经检测也显示对病毒的血清转化。没有显示 hSARS 病毒感染的血清学或病毒学证据的 5 名患者以后需要对康复期血清进行检测,以确定他们是否也发生血清转化。但是,假使临床病例的定义从未明确,hSARS 病毒与 SARS 临床定义的一致性表现得仍很显著。

15 无论是通过 RT-PCR 还是根据抗 HMPV 抗体滴度的升高,这些患者中没有一个检测到 HMPV 感染的迹象。在我们的 SARS 患者组中始终没有检测到其它病原体。因此,很可能此 hSARS 病毒是导致 SARS 的原因或是该疾病发展的必要前提。是否有其它微生物因素或另外辅助因素在该疾病的发展中发生作用的问题还有待调查。

20 冠状病毒科家族包括冠状病毒属和环状病毒属。它们是有包膜 RNA 病毒,可导致人类和动物患病。以前知道的人冠状病毒 229E 和 OC43 型是导致普通感冒的主要原因(Holmes KV. Coronaviruses. Eds Knipe DM, Howley PM Fields Virology, 第4版, Lippincott Williams & Wilkins, Philadelphia, 1187-1203)。但是,虽然冠状病毒有时能在老人、新生婴儿或免疫受损患者中导致肺炎(El-Sahly HM, Atmar RL, Glezen WP, Greenberg SB. Spectrum of clinical illness in
25 hospitalized patients with “common cold” virus infections. *Clin Infect Dis.* 2000; 31:96-100; 和 Foltz EJ, Elkordy MA. Coronavirus pneumonia following autologous bone marrow transplantation for breast cancer. *Chest*

1999;115:901-905), 有报道说它们是军队新兵中肺炎的重要原因, 在某些研究中占高达 30% 的病例(Wenzel RP, Hendley JO, Davies JA, Gwaltney JM, Coronavirus infections in military recruits: Three-year study with coronavirus strains OC43 and 229E. *Am Rev Respir Dis.* 1974; 5 109:621-624)。人冠状病毒能感染神经元, 已在多发性硬化症患者的大脑中检测到病毒 RNA (Talbot PJ, Cote G, Arbour N. Human coronavirus OC43 and 229E persistence in neural cell cultures and human brains. *Adv Exp Med Biol.* 发表中)。另一方面, 某些动物冠状病毒(例如猪传染性肠胃炎病毒、鼠肝炎病毒、禽传染性支气管炎病毒)会导致它们各自的宿主发生呼吸道疾病、胃肠疾病、神经疾病或肝病 10 (McIntosh K. Coronaviruses: a comparative review. *Current Top Microbiol Immunol.* 1974; 63: 85-112)。

我们第一次对 SARS 的临床表现和并发症进行了描述。不到 25% 的冠状病毒肺炎患者具有上呼吸道症状。正如对非典型肺炎所料到的, 15 呼吸道症状和阳性听诊结果与胸部放射照片结果很不相称。10% 的患者出现胃肠道症状。有关的是, 冠状病毒 RNA 可在某些患者的粪便样品中检出, 并且冠状病毒已知与动物和人类的腹泻有关(Caul EO, Egglestone SI. Further studies on human enteric coronaviruses *Arch Virol.* 1977; 54: 107-17)。肝脏功能紊乱、白细胞减少、显著淋巴细胞减少、血小板减少及随后发展为成人呼吸窘迫综合征的高发率表明, 20 此 hSARS 病毒导致了严重的全身性炎性损害。因此通过类固醇进行免疫调节以辅助利巴韦林的抗病毒疗法是很重要的。在这点上, 同样假定与禽流感 H5N1 亚型(近来从动物交叉传染给人类的另一种病毒)有关的严重人类疾病具有免疫病理学成分是恰当的(Cheung CY, 25 Poon LLM, Lau ASY 等, Induction of proinflammatory cytokines in human macrophages by influenza A (H5N1) viruses: a mechanism for the unusual severity of human disease. *Lancet* 2002; 360:1831-1837)。和 H5N1 疾病一样, 严重 SARS 患者也是成年人, 其淋巴细胞减少更显著, 且具有呼吸道以外器官功能障碍的特征(表 4) (Yuen KY, Chan PKS,

Peiris JSM 等, Clinical features and rapid viral diagnosis of human disease associated with avian influenza A H5N1 virus. *Lancet* 1998; **351**:467-471)。值得说明的是, 从疾病开始到呼吸衰竭有约 8 天的机会窗口。严重的并发症病例与潜在疾病及利巴韦林和类固醇疗法的延迟使用强烈相关。根据我们从最初病例得到的临床经验, 在入院时基本上没有并发症的后来病例中我们很早就实施上述组合疗法。采用这个治疗方案, 到本说明书写作时总死亡率只有 2%。在 19 个并发症病例中还有 8 个没有出现明显的反应。由于剂量及起始治疗时间不一致, 不可能详细分析对此组合方案的治疗反应。

10 与此严重疾病有关的其它因素是通过家庭接触而患病, 这可归因于高剂量或持续暴露于病毒以及存在潜在疾病。

在此所作的临床描述基本上是关于入院治疗的较严重病例。目前我们还没有任何关于社会和门诊中出现的冠状病毒科感染的完整临床范围数据。本文描述的诊断试验的有效性将有助于解决以上问题。另外, 这也允许解决关于康复中病毒脱落(及传染性)的时期、病毒在其它体液和排泄物中的存在以及潜伏期中病毒脱落的发生等问题。

目前的流行病学数据似乎表明病毒是通过飞沫或直接和间接接触而传播的, 虽然在某些情况下不能排除通过空气传播。在呼吸道中发现传染性病毒支持了此论点。初步的证据也暗示病毒可能脱落于粪便中。但是, 值得说明的是, 病毒 RNA 的检出并不能证明病毒有生存力或传染性。如果在粪便中检出活病毒, 这可能是需要加以考虑的另一潜在传播途径。可恰当指出的是, 某些动物冠状病毒是通过粪-口途径传播的 (McIntosh K., *Coronaviruses: a comparative review. Current Top Microbiol Immunol.* 1974, **63**:85-112)。

25 总之, 本报告提供了如下证据, 即冠状病毒科家族的病毒是 SARS 的病原因子。本发明公开了快速、敏感和特异性鉴定 hSARS 病毒的定量诊断试验。

7. 保藏

5 分离 hSARS 病毒的样品根据微生物保藏布达佩斯条约于 2003 年 4 月 2 日保藏在位于武汉大学(中国武汉 430072)的中国典型培养物中心(CCTCC), 给予的保藏检索号是 CCTCC-V200303, 其通过引用整体结合到本文中。

8. 市场潜力

10 现能大规模培养 hSARS 病毒, 这使得可以开发如上所述的各种诊断试验和开发能有效预防、改善或治疗 SARS 的疫苗和抗病毒药物。鉴于该疾病的严重性及其在全世界的迅速蔓延, 对用于对抗该疾病的诊断试验、疗法和疫苗的需求很可能在全世界范围内明显上升。另外, 此病毒含有对临床和科研应用极其重要和有价值的遗传信息。

15

9. 等价方案

本领域普通技术人员仅仅采用常规实验, 就将认识到或能够确定本文所述具体实施方案的许多等价方案。这种等价方案被以下权利要求书所包涵。

20 在本说明书中提到的所有出版物、专利和专利申请通过引用整体结合到本文说明书中, 其程度如同各个单独的出版物、专利或专利申请被明确和个别指出通过引用整体结合到本文中。

本文对参考文献的引用或讨论不应被解释为承认这就是本发明的现有技术。

```

a cag gac gct gta gct tca aaa atc tta gga ttg cct acg cag act gtt 49
  Gln Asp Ala Val Ala Ser Lys Ile Leu Gly Leu Pro Thr Gln Thr Val
  1           5           10          15
gat tca tca cag ggt tct gaa tat gac tat gtc ata ttc aca caa act 97
Asp Ser Ser Gln Gly Ser Glu Tyr Asp Tyr Val Ile Phe Thr Gln Thr
  20           25           30
act gaa aca gca cac tct tgt aat gtc aac cgc ttc aat gtg gct atc 145
Thr Glu Thr Ala His Ser Cys Asn Val Asn Arg Phe Asn Val Ala Ile
  35           40           45
aca agg gca aaa att ggc att ttg tgc ata atg tct gat aga gat ctt 193
Thr Arg Ala Lys Ile Gly Ile Leu Cys Ile Met Ser Asp Arg Asp Leu
  50           55           60
tat gac aaa ctg caa ttt aca agt cta gaa ata cca cgt cgc aat gtg 241
Tyr Asp Lys Leu Gln Phe Thr Ser Leu Glu Ile Pro Arg Arg Asn Val
  65           70           75           80
gct aca tta caa gca gaa aat gta act gga ctt ttt aag gac tgt agt 289
Ala Thr Leu Gln Ala Glu Asn Val Thr Gly Leu Phe Lys Asp Cys Ser
  85           90           95
aag atc att act ggt ctt cat cct aca cag gca cct aca cac ctc agc 337
Lys Ile Ile Thr Gly Leu His Pro Thr Gln Ala Pro Thr His Leu Ser
  100          105          110
gtt gat ata aaa ttc aag act gaa gga tta tgt gtt gac ata cca ggc 385
Val Asp Ile Lys Phe Lys Thr Glu Gly Leu Cys Val Asp Ile Pro Gly
  115          120          125
ata cca aag gac atg acc tac cgt aga ctc atc tct atg atg ggt ttc 433
Ile Pro Lys Asp Met Thr Tyr Arg Arg Leu Ile Ser Met Met Gly Phe
  130          135          140
aaa atg aat tac caa gtc aat ggt tac cct aat atg ttt atc acc cgc 481
Lys Met Asn Tyr Gln Val Asn Gly Tyr Pro Asn Met Phe Ile Thr Arg
  145          150          155          160
gaa gaa gct att cgt cac gtt cgt gcg tgg att ggc ttt gat gta gag 529
Glu Glu Ala Ile Arg His Val Arg Ala Trp Ile Gly Phe Asp Val Glu
  165          170          175
ggc tgt cat gca act aga gat gct gtg ggt act aac cta cct ctc cag 577
Gly Cys His Ala Thr Arg Asp Ala Val Gly Thr Asn Leu Pro Leu Gln
  180          185          190
cta gga ttt tct aca ggt gtt aac tta gta gct gta ccg act ggt tat 625
Leu Gly Phe Ser Thr Gly Val Asn Leu Val Ala Val Pro Thr Gly Tyr
  195          200          205
gtt gac act gaa aat aac cta 646
Val Asp Thr Glu Asn Asn Leu
  210          215

```

图 1

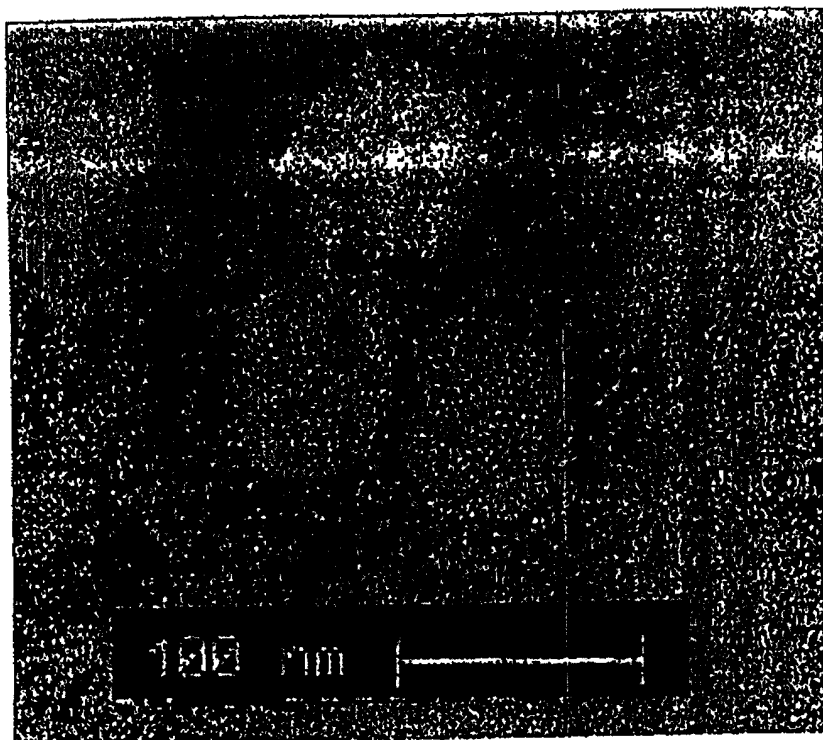


图 2

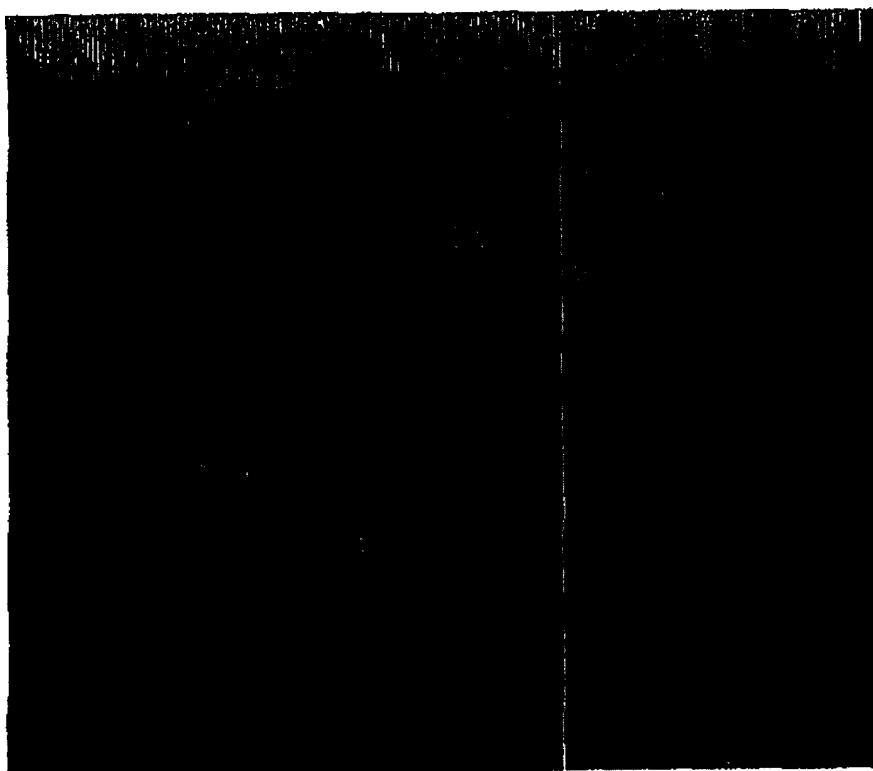


图 3

图 4

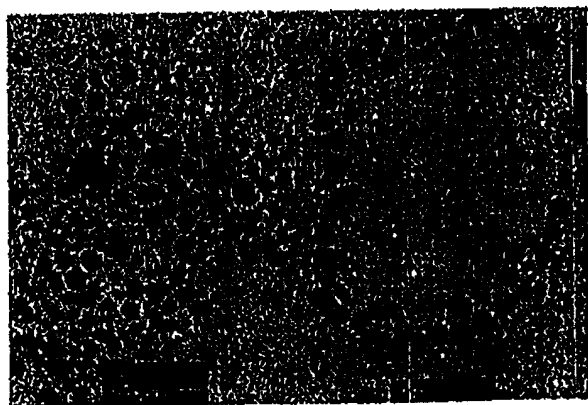
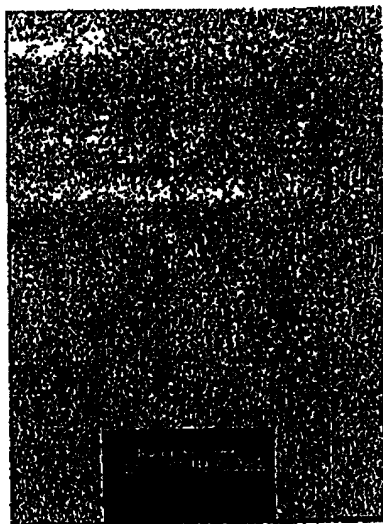


图 5A

图 5B

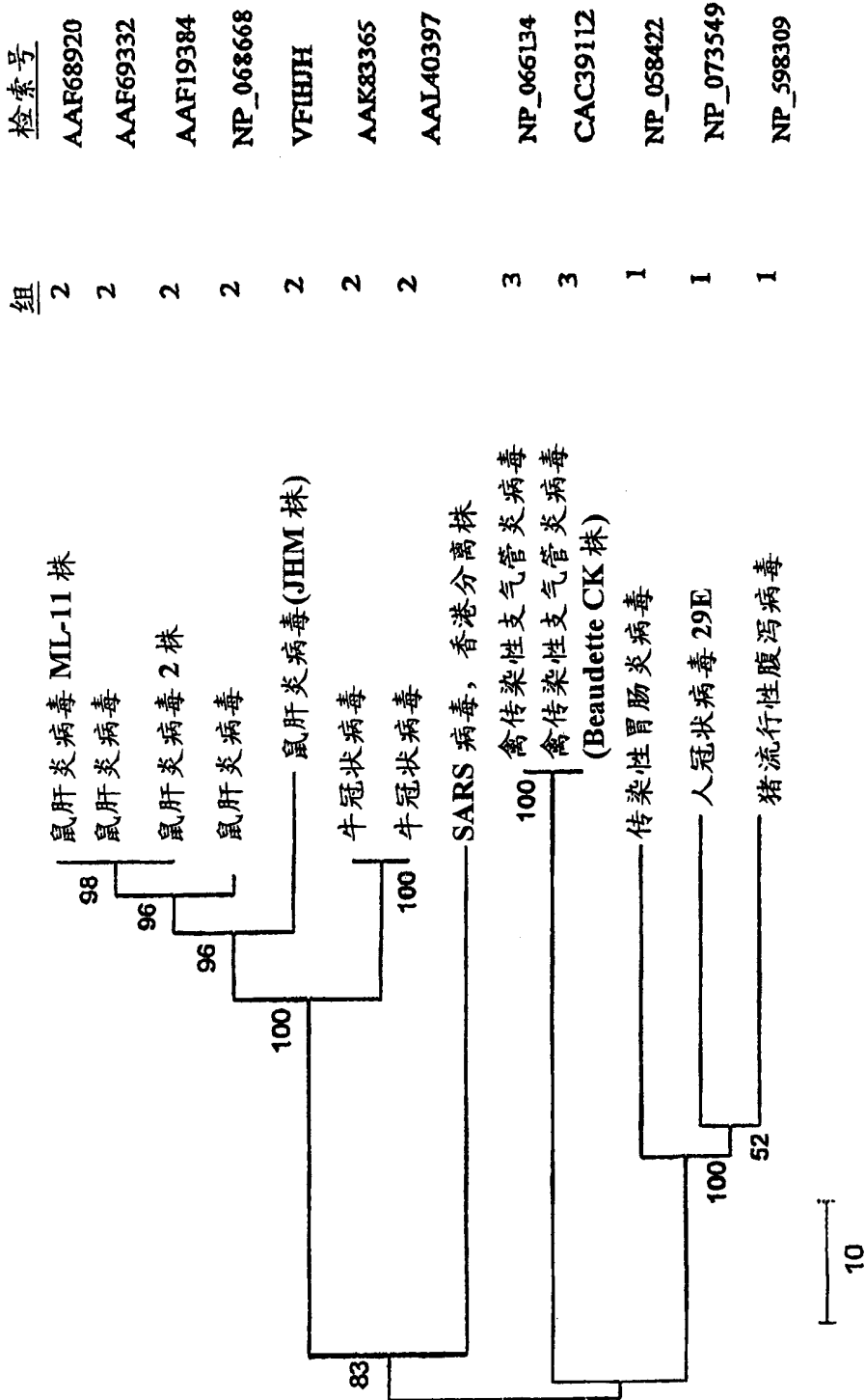


图 6

图 7A

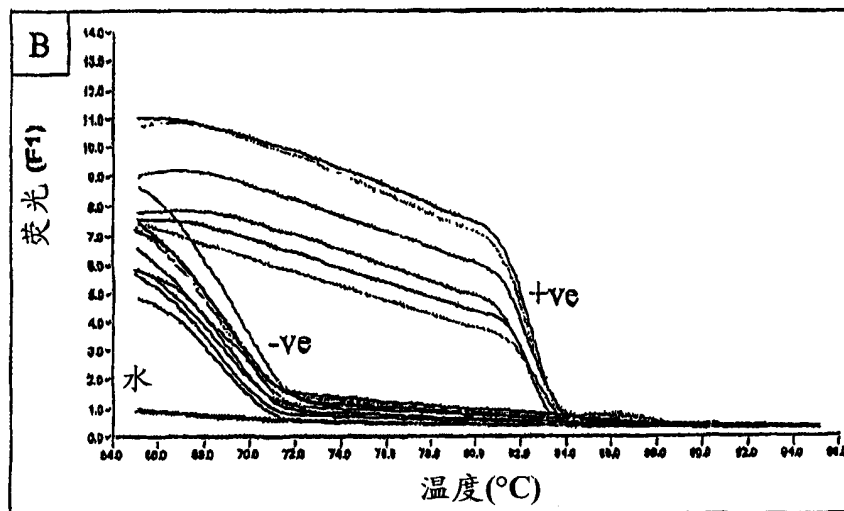
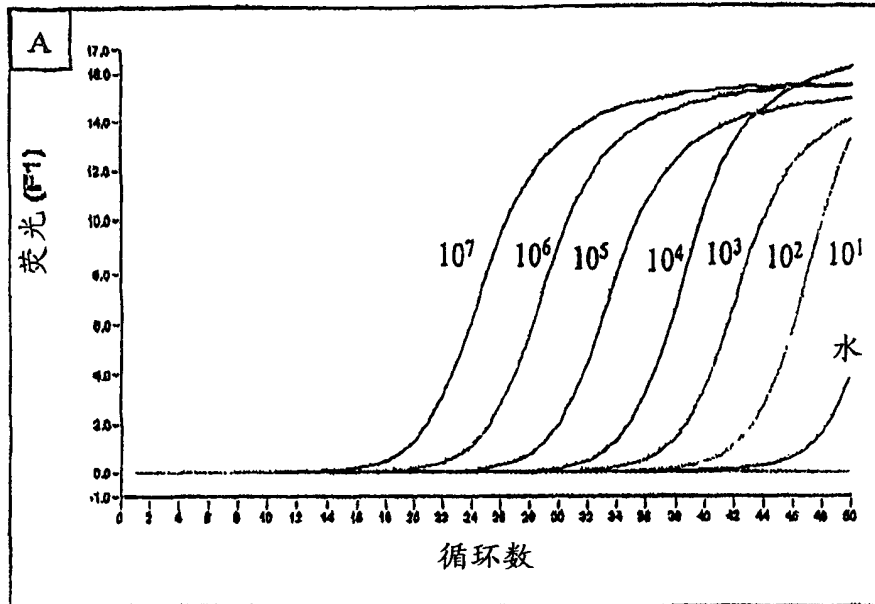


图 7B

```

t aaa tgt agt aga atc ata cct gcg cgt gcg cgc gta gag tgt ttt gat 49
  Lys Cys Ser Arg Ile Ile Pro Ala Arg Ala Arg Val Glu Cys Phe Asp
  1           5           10           15

aaa ttc aaa gtg aat tca aca cta gaa cag tat gtt ttc tgc act gta 97
  Lys Phe Lys Val Asn Ser Thr Leu Glu Gln Tyr Val Phe Cys Thr Val
           20           25           30

aat gca ttg cca gaa aca act gct gac att gta gtc ttt gat gaa atc 145
  Asn Ala Leu Pro Glu Thr Thr Ala Asp Ile Val Val Phe Asp Glu Ile
           35           40           45

tot atg gct act aat tat gac ttg agt gtt gtc aat gct aga ott cgt 193
  Ser Met Ala Thr Asn Tyr Asp Leu Ser Val Val Asn Ala Arg Leu Arg
           50           55           60

gca aaa cac tac gtc tat att ggc gat cct gct caa tta cca gcc ccc 241
  Ala Lys His Tyr Val Tyr Ile Gly Asp Pro Ala Gln Leu Pro Ala Pro
  65           70           75           80

cgc aca ttg ctg act aaa ggc aca cta gaa cca gaa tat ttt aat tca 289
  Arg Thr Leu Leu Thr Lys Gly Thr Leu Glu Pro Glu Tyr Phe Asn Ser
           85           90           95

gtg tgc aga ctt atg aaa aca ata ggt cca gac atg ttc ctt gga act 337
  Val Cys Arg Leu Met Lys Thr Ile Gly Pro Asp Met Phe Leu Gly Thr
           100          105          110

tgt cgc cgt tgt cct gct gaa att gtt gac act gtg agt got tta gtt 385
  Cys Arg Arg Cys Pro Ala Glu Ile Val Asp Thr Val Ser Ala Leu Val
           115          120          125

tat gac aat aag cta aaa gca cac aag gag aag tca gct caa tgc ttc 433
  Tyr Asp Asn Lys Leu Lys Ala His Lys Glu Lys Ser Ala Gln Cys Phe
           130          135          140

aaa atg ttc tac aaa ggt gtt att aca cat gat gtt tca tct gca atc 481
  Lys Met Phe Tyr Lys Gly Val Ile Thr His Asp Val Ser Ser Ala Ile
  145          150          155          160

aac aga cct caa ata ggc gtt gta aga gaa ttt ctt aca cgc aat cct 529
  Asn Arg Pro Gln Ile Gly Val Val Arg Glu Phe Leu Thr Arg Asn Pro
           165          170          175

gct tgg aga aaa gct gtt ttt atc tca cct tat aat tca cag aac gct 577
  Ala Trp Arg Lys Ala Val Phe Ile Ser Pro Tyr Asn Ser Gln Asn Ala
           180          185          190

gta gct tca aaa atc tta gga ttg cct acg cag act gtt gat tca tca 625
  Val Ala Ser Lys Ile Leu Gly Leu Pro Thr Gln Thr Val Asp Ser Ser
           195          200          205

cag ggt tct gaa tat gac tat gtc ata ttc aca caa act act gaa aca 673
  Gln Gly Ser Glu Tyr Asp Tyr Val Ile Phe Thr Gln Thr Thr Glu Thr
           210          215          220

```

图 8

| | |
|---|-----------------|
| gca cac tct tgt aat gtc aac cgc ttc aat gtg gct atc aca agg gca | 721 |
| Ala His Ser Cys Asn Val Asn Arg Phe Asn Val Ala Ile Thr Arg Ala | |
| 225 | 230 235 240 |
| aaa att ggc att ttg tgc ata atg tct gat aga gat ctt tat gac aaa | 769 |
| Lys Ile Gly Ile Leu Cys Ile Met Ser Asp Arg Asp Leu Tyr Asp Lys | |
| | 245 250 255 |
| ctg caa ttt aca agt cta gaa ata cca cgt cgc aat gtg gct aca tta | 817 |
| Leu Gln Phe Thr Ser Leu Glu Ile Pro Arg Arg Asn Val Ala Thr Leu | |
| | 260 265 270 |
| caa gca gaa aat gta act gga ctt ttt aag gac tgt agt aag atc att | 865 |
| Gln Ala Glu Asn Val Thr Gly Leu Phe Lys Asp Cys Ser Lys Ile Ile | |
| | 275 280 285 |
| act ggt ctt cat cct aca cag gca cct aca cac ctc agc gtt gat ata | 913 |
| Thr Gly Leu His Pro Thr Gln Ala Pro Thr His Leu Ser Val Asp Ile | |
| | 290 295 300 |
| aaa ttc aag act gaa gga tta tgt gtt gac ata cca ggc ata cca aag | 961 |
| Lys Phe Lys Thr Glu Gly Leu Cys Val Asp Ile Pro Gly Ile Pro Lys | |
| | 305 310 315 320 |
| gac atg acc tac cgt aga ctc atc tct atg atg ggt ttc aaa atg aat | 1009 |
| Asp Met Thr Tyr Arg Arg Leu Ile Ser Met Met Gly Phe Lys Met Asn | |
| | 325 330 335 |
| tac caa gtc aat ggt tac cct aat atg ttt atc acc cgc gaa gaa gct | 1057 |
| Tyr Gln Val Asn Gly Tyr Pro Asn Met Phe Ile Thr Arg Glu Glu Ala | |
| | 340 345 350 |
| att cgt cac gtt cgt gog tgg att ggo ttt gat gta gag ggc tgt cat | 1105 |
| Ile Arg His Val Arg Ala Trp Ile Gly Phe Asp Val Glu Gly Cys His | |
| | 355 360 365 |
| gca act aga gat gct gtg ggt act aac cta cct ctc cag cta gga ttt | 1153 |
| Ala Thr Arg Asp Ala Val Gly Thr Asn Leu Pro Leu Gln Leu Gly Phe | |
| | 370 375 380 |
| tct aca ggt gtt aac tta gta gct gta ccg act ggt tat gtt gac act | 1201 |
| Ser Thr Gly Val Asn Leu Val Ala Val Pro Thr Gly Tyr Val Asp Thr | |
| | 385 390 395 400 |
| gaa aat aac cta | 1213 |
| | Glu Asn Asn Leu |

图 8 续

```

c aga acc atg cct aac atg ctt agg ata atg gcc tct utt gtt ctt gct 49
  Arg Thr Met Pro Asn Met Leu Arg Ile Met Ala Ser Leu Val Leu Ala
  1          5          10          15

cgc aaa cat aac act tgc tgt aac tta tca cac cgt ttc tac agy tta 97
  Arg Lys His Asn Thr Cys Cys Asn Leu Ser His Arg Phe Tyr Arg Leu
          20          25          30

gct aac gag tgt gcg caa gta tta agt gag atg gtc atg tgt ggc ggc 145
  Ala Asn Glu Cys Ala Gln Val Leu Ser Glu Met Val Met Cys Gly Gly
          35          40          45

tca cta tat gtt aaa cca ggt gga aca tca toc ggt gat gct aca act 193
  Ser Leu Tyr Val Lys Pro Gly Gly Thr Ser Ser Gly Asp Ala Thr Thr
  50          55          60

gct tat gct aat agt gtc ttt aac att tgt caa gct gtt aca gcc aat 241
  Ala Tyr Ala Asn Ser Val Phe Asn Ile Cys Gln Ala Val Thr Ala Asn
  65          70          75          80

gta aat gca ctt ctt tca act gat ggt aat aag ata gct gac aag tat 289
  Val Asn Ala Leu Leu Ser Thr Asp Gly Asn Lys Ile Ala Asp Lys Tyr
          85          90          95

gtc cgc aat cta caa cac agg ctc tat gag tgt ctc tat aga aat agg 337
  Val Arg Asn Leu Gln His Arg Leu Tyr Glu Cys Leu Tyr Arg Asn Arg
          100          105          110

gat gtt gat cat gaa ttc gty gat gag ttt tac gct tac ctg cgt aaa 385
  Asp Val Asp His Glu Phe Val Asp Glu Phe Tyr Ala Tyr Leu Arg Lys
          115          120          125

cat ttc tcc atg atg att ctt tct gat gat gcc gtt gty tgc tat aac 433
  His Phe Ser Met Met Ile Leu Ser Asp Asp Ala Val Val Cys Tyr Asn
  130          135          140

agt aac tat gcg got caa ggt tta gta gct agc att aag aac ttt aag 481
  Ser Asn Tyr Ala Ala Gln Gly Leu Val Ala Ser Ile Lys Asn Phe Lys
  145          150          155          160

gca gtt ctt tat tat caa aat aat gty ttc atg tot gag gca aaa tgt 529
  Ala Val Leu Tyr Tyr Gln Asn Asn Val Phe Met Ser Glu Ala Lys Cys
          165          170          175

tgg act gag act gac ctt act aaa gga cot cac gaa ttt tgc tca oag 577
  Trp Thr Glu Thr Asp Leu Thr Lys Gly Pro His Glu Phe Cys Ser Gln
          180          185          190

cat aca atg cta gtt aaa caa gya gat gat tac gty tac ctg cct tac 625
  His Thr Met Leu Val Lys Gln Gly Asp Asp Tyr Val Tyr Leu Pro Tyr
          195          200          205

cca gat cca tca aga ata tta gcc gca gcc tgt ttt gtc gat gat att 673
  Pro Asp Pro Ser Arg Ile Leu Gly Ala Gly Cys Phe Val Asp Asp Ile
          210          215          220

gtc aaa cag atg gta cac tta tga ttg aaa ggt tcc gty tca ctg gct 721
  Val Lys Gln Met Val His Leu
  225          230

att gat gc 729

```

图 9

```

1 atattaggtt tttacctacc caggaaaagc caaccaacct cgatctcttg tagatctggt
61 ctctaacga actttaaat ctgtgtagct gtcgc:cggc tgcatgccta gtcacctac
121 gcagtataaa caataataaa ttttactgtc gttgacaaga aacgagtaac tcgtccctct
181 tctgcagact gcttaecggt tcgtccgtgt tgcagtogat catcagcata cctaggtttc
241 gtcggggtgt gaccgaaagg taagatggag agccttgttc ttggtgtcaa cgagaaaaca
301 cacgtccaac tcagtttgcc tgccttcag gtttagagacg tgcagtgcg tggcttcggg
361 gactctgtgy aagaggccct atcggaggca cgtgaacaoc tcaaaaatgg cacttgtggt
421 otagtagagc tggaaaaagg cgtactgccc cagcttgaac agccctatgt gttcattaaa
481 cgttctgatg ccttaagoac caatoacggo caoaaggctg ttgagctggt tgoagaaatg
541 gacggcattc agtaocgctg tagcgggata acaocgggag taocgctgoc acatgtgggc
601 gaaaccccaa ttgcataocg caatgttctt cttcgttaaga acggtataaa gggagccggg
661 ggtcatagot atggcatoga tctaaagtct tatgacttag gtgacgagct tggcactgat
721 cccattgaag attatgaaca aaactggaac actaagcatg gcagtggctg aocctgtgaa
781 ctactcgtg agctcaatgg aggtgcagtc actcgcctatg tcgacaacaa tttctgtggc
841 ocagatgggt accctottga ttgcatcaa gattttctog cacgcgaggg caagtcaatg
901 tgcactcttt ccgaacaact tgattacatc gactogaaga gagggtgcta tctgtcgcgt
961 gaccatgagc atgaaattgc tgggttcaot gagcgcctcg ataagagcta cgagcaccag
1021 acaccctctg aaattaagag tgcacaagaaa tttgacactt tcaaaaggga atgcccacag
1081 tttgtgtttc ctcttaactc aaaaagcaaa gtcattcaac cacgtgttga aaagaaaaag
1141 actgagggtt toatggggcg tatacgcctc gtgtaccctg ttgcatctoc acaggagtgt
1201 acaaatatgc acttgtctac cttgatgaaa tgyaaitcatt gcgatgaagt ttcatggcag
1261 acgtgogaot ttctgaaagc caottgtgaa cattgtggca otgaaaattt agttattgaa
1321 ggacotacta catgtgggta cctaocctac aatgctgtag tgaaaatgoc atgtcctgoc
1381 tgtcaagacg cagagattgg accgtgagcat agtgttgoag attatcaca ccaotcaaac
1441 attgaaactc gaotcogcaa gggaggtagg actagatggt ttggaggctg tgtgtttgoc
1501 tatgttggct gotataataa gcgtgcctac tgggttcctc gtgctagtgc tgatattggc
1561 tcaggccata ctggcattac tggtgacaat gtggagacct tgaatgagga tctccttgag
1621 aactgagtc gtgaacgtgt taacattaac attgttggcg attttcattt gaatgaagag
1681 gttgccatca ttttggcatc tttctotgot tctacaagtg cctttattga cactataaag
1741 agtcttgatt acaagtcttt caaaaccatt gttgagctct gcggtaacta taaagttacc
1801 aagggaagc ccgtaaaagg tgottggaac attggacaac agagatcagt tttaaocaca
1861 ctgtgtggtt ttccctcaaa ggtcgtggtt gttatocagat caatttttgc gcgcacactt
1921 gatgcagcaa accactcaat tctgtatttg caaagagcag ctgtcaccat acttgatggt
1981 atttctgaac agtcattacg tcttgtogac gccatgggtt ataactcaga cctgctcacc
2041 aacagtgtca ttattatggc atatgtaact ggtgtctctg tacaacagac ttctcagtg
2101 ttgtotaatc ttttgggcao taotgttgaa naaotcaggo otatctttga atggattgag
2161 gogaaactta gtgcaggagt tgaatttctc aaggatgott gggagattot caaatttctc
2221 attacaggty tttttgacat cgtcaagggt caaatcacgg ttgottcaga taacatcag
2281 gattgtgtaa aatgcttcat tgatgttgtt aacaaggcac tcgaaatgtg cattgatcaa
2341 gtcactatcg ctggcgcaaa gttcggatca ctcaacttag gtgaagtctt catcgtcaa
2401 agcaagggac tttacogtca gtgtatacgt ggcaaggagc agotgoaact actcatgctt
2461 cttaggcacc caaaaagaag aaoccttctt gaaggtgatt oaatgaoac agtaottacc
2521 tctgaggagg ttgtttotcaa gaacggtgaa otcgaagcac tcgagacgcc cgttgatagc
2581 ttcaaaaatg gagctatcgt cggcacacca gtctgtgtaa atggcctcat gctcttagag
2641 attaaggaca aagaacaata ctgcgcattg tctcctgggt tactggctac aaacaatgct
2701 tttcgtctaa aagggggtgc acaattaaa ggtgtaacct ttggagaaga tactgtttg
2761 gaagttcaag gttacaagaa tgtgagaatc acatttgagc ttgatgaaog tgttgacaaa
2821 gtgcttaatg aaaagtgtc tgtotacact gttgaaatcg gtaccgaagt tactgagttt
2881 goatgtgttg tagcagaggc tgttgtgaa actttacaac cagtttctga tctccttacc
2941 aacatgggta ttgatcttga tgagtggagt gtagotacat totaottatt tgatgatgct
3001 ggtgaagaaa aottttcatc aogtatgtat tgttctttt aocctocaga tgaggaagaa
3061 gaggacgatg cagagtgtga ggaagaagaa attgatgaaa cctgtgaca tgagtacggg
3121 acagaggatg attatoaagg tctccctctg gaatttggty cctcagctga aacagttoga
3181 gttgaggaag aagaagagga agactggctg gatgatacta ctgagcactc agagattgag
3241 caagaacag aacctacaoc tgaagaacaa gttaatcagt ttaactggta tttaaaactt
3301 actgaacatg ttgcaattaa atgtgttgac atcgttaagg aggcaaaaag tgetaatctt

```

图 10


```

3361 atggtgattg taaatgctgc taacatacac ctgaaacatg gtggtggtgt agcagggtgca
3421 ctcaacaagg caaccaatgg tgcocatgcaa aaggagagtg atgattacat taagctaaat
3481 ggcoctctta cagtaggagg gtcttgtttg ctttctggac ataantcttg taagaagtgt
3541 ctgcatgttg ttggaactaa cotaaatgca ggtgaggaoa tccagcttct taaggcagca
3601 tatgaaaatt tcaattcaca ggacatctta ottgcacoat tgttgtcagc aggeatattt
3661 ggtgctaaac cacttcagtc ttacaagtg tgcgtgcaga cgttctgtac acaggttat
3721 attgacagtca atgacaaagc tctttatgag caggttgtca tggattatct tgataacctg
3781 aagoctagag tggaaagoacc taacaagag gagcoaccaa acacagaaga ttocaaaact
3841 gaggagaaat ctgtogtaca gaagcctgtc gatgtgaagc caaaaattaa ggcoctgcat
3901 gatgaggtta ocacoaact ggagaaaact aagttctta ccaataagtt actottgttt
3961 gctgatataca atggttaagot ttacoatgat totcagaaoa tgcttagagg tgaagatag
4021 tctttccttg agaaggatgc acottaocat gtagg-gatg ttatcactag tggatgatac
4081 acttgtgttg taataocctc caaaaaggct ggtggcacta ctggagatgt ccaagagct
4141 ttgaaagaag tgcagttga tgagtatata accaogtacc ctggacaagg atgtgctggt
4201 tatacacttg aggaagctaa gactgctott aagaaatgca aatctgcatt ttatgtacta
4261 ccttcagaag cacctaagtc taaggaagag attctaggaa ctgtatctct gaatttgaga
4321 gaaatgcttg ctcatgtgca agagacaaga aaattaatgc ctatatgcat ggatgttaga
4381 gccataatgg caacoatoca acgtaagtat aaaggaatta aaattcaaga ggcacatggt
4441 gactatggtg tccgattott cttttatact agtaaagagc ctgtagcttc tattattacg
4501 aagctgaact ctctaaatga gccgcttgtc acaatgcaa ttggttatgt gacacatggt
4561 tttaatcttg aagaggotgc gcgctgtatg cgttctctta aagctcctgc cgtagtgtca
4621 gtabcatoac cagatgotgt taactaolat aatggatacc tcaactctgc atcaagaca
4681 tctgaggagc aotttgtaga aacagttctt ttggotggot cttacagaga ttggtcotat
4741 tcaggacagc gtacagagtt aggtgttgaa tttcttaagc gtggtgacaa aattgtgtac
4801 cacactctgg agagocccgt cgagtttcat ottgacgggt aggttcttct acttgacaaa
4861 ctaagagatc tottatccct gcgggaggtt aagactataa aagtgttcc acotgtggac
4921 aacactaatc tocaacoaca gottgtggat atgtotatga catatggaca gcagtttgg
4981 ccaacatact tggatggtgc tgatgttaca aaaattaac ctcatgtaa tcatgagggt
5041 aagaotttct ttgtaactacc tagtgatgac acactaogta gtgaagottt cgagtactac
5101 ctaactcttg atgagagttt tcttggtagg tacatgtctg otttaaccaa cacaagaaa
5161 tggaaatttc ctcaagttgg tggtttaact tcaattaat gggotgataa oaattgttat
5221 ttgtotagtg ttttattagc aottcaacag ottgaagtca aattcaatgc accagcaott
5281 caagaggctt attatagagc cogtgotggg gatgctgcta acttttgtgc actcactctc
5341 gottacagta ataaaactgt tggcgagott ggtgatgtca gagaaactat gaoccatctt
5401 ctacagcatg ctaatttggg atotgcaaag cgagttctta atgtggtgtg taaacattgt
5461 ggtocagaaa ctactacott aaoggtgtga gaagctgtga tgtatatggg tactctatct
5521 tatgataatc ttaagacagg tgtttccatt coatgtgtgt gtggtcgtga tgotacaaa
5581 tatctagtac aacaagagtc tcttttgtt atgatgtctg caccocgtgc tgagtataa
5641 ttacagcaag gtaoattctt atgtgcgaat gagtacaotg gtaactatca gtgtggtcac
5701 tacaotoata taactgotaa ggagaccoto tctcgtattg acggagotca cottacaaag
5761 atgtcagagt acaaaggacc agtgactgat gtttcttaca aggaaacatc ttacactaca
5821 accatcaagc ctgtgtcgtg taaactogat ggagtttaott acacagagat tgaaccaaaa
5881 ttggatgggt attataaaaa ggataatgct tactatacag agcagcctat agaocttcta
5941 ccaactcaac cattacaaaa tgogagtttt gataatttca aactcactat ttotaacaca
6001 aattttgtgt atgattttaa toaatgaca ggcttcaoa agccagotto aogagagota
6061 tctgtcacat tcttcccaga cttgaatggc gatgtagtgt ctattgacta tagacactat
6121 tcagogagtt tcaagaaaggt tgctaattaa ctgcataagc caattgtttg gcacattaac
6181 caggctacaa ocaagacaac gttcaaaoca acaacttggg gtttacgttg totttggagt
6241 acaagccag tagatacttc aaattcattt gaagtctctg cagtagaaga cacacaagga
6301 atggacaatc ttgottgtga aagtoacaaa cooacotctg aagaagtagt ggaanaatcct
6361 aocatacaga aggaagtoat agagtgtgac gtgaaaoota oogaagttgt aggoaatgtc
6421 ataactaac catcagatga aggtgtttaa gtaacacaag agttaggtca tgaggatctt
6481 atggctgctt atgtggaaaa cacaagcatt accattaaga aacctaatga gotttcaacta
6541 gcottaggtt taaaacaact tgccactcat ggtattgctg caattaatag tgttctctgg
6601 agtaaaattt tggottatgt caaacatto ttaggacaag cagcaattac aacatcaaat
6661 tgcgctaaga gattagcaoa aogtgtgttt aacaattata tgocttatgt gtttacatta

```

图 10 续

```

6721 ttgttccaat tgtgtacttt tactaaaagt accaattota gaattagago ttcactaocct
6781 acaactattg ctaaaaatag tgtaagagt gttgctaaat tatgtttgga tgccggcatt
6841 aattatgtga agtoacccaa attttctaaa ttgttcacaa tcgctatgtg gctattgttg
6901 ttaagtattt gotttaggtc tctaactctgt gtaactgctg cttttggtgt actcttatct
6961 aattttgggt ctcottotta ttgtaatggc gttagagaat tgtatcctaa ttcgtctaac
7021 gttactacta tggatttotg tgaaggttot tttcottgoa gcatttgttt aagtggatta
7081 gactcccttg attcttatcc agctottgaa accattoagg tgacgattto atogtaoaag
7141 ctagacttga caattttagg tctggcogct gagtgggttt tggcataat gttgttcaca
7201 aaattctttt atttattagg tctttcagct ataatgcagg tgttctttgg ctattttgcct
7261 agtcaattca tcagcaatto ttggotoatg tggtttatca ttagtattgt acaaatggca
7321 ccogtttotg oaatggttag gatgtacatc ttctttgctt otttctacta cstatggaag
7381 agctatgttc atatcatgga tggttgcacc tottcgactt gcattgatgt ctataagcgc
7441 aatogtgcca cagcogttga gtgtacaact attgttaatg gcattgaagag atctttctat
7501 gctcatgcaa atggaggcog tggcttotgoc aagacccaca attggaattg totcaattgt
7561 gacacatttt gcaactgtag tacattcatt agtgaatgaa ttgctogtga tttgtoactc
7621 cagtttaaaa gaccaatcaa cctactgac cagtcactgt atattgttga tagtgttgcct
7681 gtgaaaaatg gogcogttca cctotaottt gaoaaggctg gtoaaaagac ctatgagaga
7741 catcogctct occattttgt caattttagac aatttgagag ctaacaacac taaaggttca
7801 ctgctatta atgtcatagt ttttgatggc aagtccaaat gcgacgagtc tgcctctaag
7861 tctgcttctg tgtactacag tcagotgatg tgccaaccta ttctgttgct tgaccaagot
7921 cttgtatcaa acgttggaga tagtactgaa gtttccogtta agatgtttga tgccttatgtc
7981 gacacotttt cagcaacttt tagtgttccct atggaaaaac ttaaggcaact tgttgcatac
8041 gctcacagcg agttagcaa ggggtgtagct ttagatgggt tcctttctac tctogtgoa
8101 gctgcccagc aaggtgttgt tgataccgat gttgacacaa aggatgttat tgaatgtctc
8161 aaactttcac atcaotctga cttagaagtg acaggtgaca gttgtaacaa tttoatgctc
8221 acctataata aggttgaaaa catgacgccc agagatottg gogcatgtat tgaotgtaat
8281 gcaaggcata tcaatgccc agtagcaaaa agtcaaatg tttcactcat ctggaatgta
8341 aaagactaoa tgtctttatc tgaacagotg cgtaaacaaa ttogtactgo tgccaagaag
8401 aacaaotac cttttacaot aacttgtgct acaactagac aggttgtcaa tgcataact
8461 actaaaactt cactcaaggg tggttaagatt gtttagtactt gttttaaact tatgtttaag
8521 gooacttat tgtgcttctc tgotgoattg gtttgttata tcgttatgcc agtacaatac
8581 ttgtcaatoc atgatggtta oacaaatgaa atcattgggtt acaaagccat tcaggatggt
8641 gtoactogtg acatcatttc tactgatgat tgttttgoaa ataaacatgc tggttttgac
8701 gcatggttta gccagcgtgg tggttcatac aaaaatgaca aaagctgccc tgtagtagct
8761 gctatcatta caagagagat tggttcatac gtgctggctc taccgggtao tgtgctgaga
8821 gcaatcaatg gtgacttctt goattttcta cctogtgttt ttagtgcctg tggcaactt
8881 tgetaocaco cttccaaact cattgagtat agtgattttg ctacctctgc ttgcttctt
8941 gctgctgagt gtacaatttt taaggatgct atgggcaaac ctgtgocata ttgttatgac
9001 actaatttgc tagagggttc tatttottat agtgaotto gtocagacao togttatgtg
9061 cttatggatg gttccatcat acagtttctc aacaacttaco tggagggttc tgttagagta
9121 gtaacaactt ttgatgctga gtactgtaga catggtacat gogaaaggtc agaagtaggt
9181 atttgcctat ctaccagtg tagatgggtt cttaataatg agcattacag agototatca
9241 ggagttttct gtggtgttga tgogatgaa otcatagota acatctttac tctcttgtg
9301 caacctgtgg gtgctttaga tgtgtctgot tcagttagtg ctgggtggtat tattgocata
9361 ttggtgactt gtgctgocca ctaotttatg aaattcagac gtgtttttgg ttagtacaac
9421 catgttbtgt ctgctaagtc acttttgttt ttgatgtott toactataot ctgtctggtg
9481 oagcttaca gctttctgcc gggagtctac tcagtctttt aottgtactt gacattctat
9541 ttaccaatg atgttttatt cttggctcac ottcaatggt ttgocattgt ttctcotatt
9601 gtgctttttt ggataacagc aatctatgta ttctgtattt ctctgaagca ctgcatggtg
9661 ttctttaaca actatcttag gaaaagagtc atgtttaatg gagttacatt tagtaccttc
9721 gaggaggctg otttgtgtac ctttttgctc aacaaggaaa tgtacctaaa attgogtagc
9781 gagacactgt tgcoacttac acagtataac aggtatottg ototataaa caagtacaag
9841 ttttccagtg gacacttaga tactaccagc tatcgtgaa gacgttgotg ccacttagca
9901 aaggotctaa atgactttag caactcaggt gctgatgttc tctaccaacc accacagaca
9961 tcaatcactt ctgctgttct gcagagtggc tttaggaaaa tggcattccc gtcaggcaaa
10021 gttgaagggt gcatggtaca agtaacotgt ggaactacaa ctottaatgg attgtggttg

```

图 10 续

```

10081 gatgaacag tatactgtcc aagacatgtc atttgcacag cagaagacat gcttaatcct
10141 aactatgaag atotgtcoat togcaaatcc aacctagot ttcttgctca ggctggcaat
10201 gttcaacttc gtgttattgg ccattctatg caaaattgtc tgottaggct taaagttgat
10261 aottotaacc otaagacaoc caagtataaa tttgtccgta tccaacctgg tcaaacattt
10321 tcagttctag catgctacaa tggttcacca tctggtggtt atcagtgctc catgagacct
10381 aatoatacca ttaaaggtto ttcocttaat ggatoatgtg gtagtggttg ttttaacatt
10441 gattatgatt gogtgtcttt ctgctatatg catcatatgg agcttocaac aggagtacac
10501 gotggtactg acttagaagg taaattctat ggtccatttg ttgacagaca aactgcacag
10561 gctgcaggta cagacacaac cataacatta aatgttttgg catggctgta tgctgctggt
10621 atcaatggty ataggtgggt tottaataga ttcaccacta ctttgaatga ctttaacctt
10681 gtyggaatga agtacaacta tgaacctttg acacaagatc atgttgacat attyggacct
10741 ctttctgctc aaacaggaat tgcgccttta gatatgtgtg ctgctttgaa agagotgctg
10801 cagaatggta tgaatggctg tactatcott ggtagcacta ttttgaaga tgagtttaca
10861 ccatttgatg ttgttagaca atgctctggt gttacottcc aaggtaagtt caagaaaatt
10921 gtttaagggca ctcatcattg gatgotttta actttcttga catcaactatt gattcttggt
10981 caaagtacac agtggctact gtttttcttt gtttacgaga atgctttctt gccatttact
11041 cttggtatta tggcaattgc tgcattgtgt atgctgcttg ttaagcataa gcacgcattc
11101 ttgtgottgt ttctgttacc ttctcttgca acagttgctt actttaatat ggtctacatg
11161 cctgctagot ggggtgatgc tatcatgaga tggcttgaat tggctgacac tagcttctct
11221 ggtgotttagc ttaaggattg tgttatgtat gottcagott tagttttgot tattotoatg
11281 aoagotcgcg ctgtttatga tgcattgtgt agacgtggtt ggaocactgat gaatgtcatt
11341 acacttggtt acaaagtcta ctatggtaat gotttagatc aagctatttc catgtgggctc
11401 ttagtatttt ctgtaacctc taactattct ggtgtcgtta cgactatcat gtttttagct
11461 agagctatag tgtttgtgtg tgttgagtat taccattgtt tatttattac tggcaaacac
11521 ttacagtgta tcatgottgt ttattgtttc ttaggctatt gttgctgctg ctactttggc
11581 cttttctggt tactcaaccg ttacttcagg cttactcttg gtgtttatga ctacttggct
11641 tctacacaag aatttaggta tatgaactcc caggggcttt tgcctcota gaagtattt
11701 gatgotttca agcttaacat taagttgctg ggtattggag gtaaacatg tatcaaggtt
11761 gctactgtac agtctaaaat gtctgacgta aagtgcacat ctgtggtact gctctcggtt
11821 ctttaacaac tttagagtaga gtoatottct aaattgtggg caaatgtgt acaaatocac
11881 aatgatattc ttcttgoaaa agacaacaot gaagcttctg agaagatggt ttctottttg
11941 tctgttttgc tatccatgoa ggggtcgtga gacattaata ggttgtgoga ggaatgctc
12001 gataaccgtg ctactcttca ggctattgct tcagaattta gttctttaoc atcatatgoc
12061 gottatgcca ctgcccagga ggcctatgag caggctgtag ctaatgggta ttctgaagtc
12121 gttctcaaaa agttaaagaa atctttgaaat gtggctaaat ctgagttga oogttagctc
12181 gcoatgcaac gcaagttgga aaagatggca gatcaggcta tgaccaaat gtacaacag
12241 gcaagatctg aggacaagag ggcaaaagta actagtgtca tgcaacaat gctcttcact
12301 atgcttagga agcttgataa tgatgcaott aacaacatta tcaacaatgc gogttagtgt
12361 tgtgttccac tcaacatcat accattgact acagcagcca aaotcaggt tgyttgocot
12421 gattatggta cctacaagaa cacttgtgat ggtaacacct ttacatatgc atctgcactc
12481 tgggaaatcc agcaagttgt tgatgcggat agcaagattg ttcaacttag tgaatttaac
12541 atggacaatt caccaaattt ggcttggcct cttattgtta cagctctaag agccaactca
12601 gctgttaaac tacagaataa tgaactgagt ccagtagoac taogacagat gtcctgtgog
12661 gctggtacca cacaaacagc ttgtaotgat gacaatgoac ttgcotacta taacaattcg
12721 aagggaggta ggtttgtgct ggcattacta tcagaccacc aagatoccaa atgggctaga
12781 ttccotaaga gtgatggtac aggtacaatt tacacagaac tggaaaccacc ttgtaggttt
12841 gttacagaca caocaaagg gcoataagty aaatacttgt acttoatcaa aggcttaaac
12901 aacctaaata gaggtatggt gctgggcagt ttagotgcta cagtaogtot tcaaggctgga
12961 aatgctacag aagtaootgo caattoaacot gtgotttoot totgtgcttt tgcagtagac
13021 cctgctaaag catataagga ttacctagca agtggaggac aaccaatcac caactgtgtg
13081 aagatgttgt gtacacacac tgggtacagga caggcaatta ctgtaacacc agaagctaac
13141 atggaccaag agtcttttgg tgggtcttoa tgyttgtctgt attgtagatg ccacattgac
13201 catocaaatc ctaaaggatt ctgtgacttg aaaggttaagt acgtccaaat acctaccact
13261 tgtgotaatg acocagtggg ttttcaactt agaaacacag tctgtacogt ctgcggaatg
13321 tggaaagggt atgyctgtag ttgtgaocaa ctocogaaac ccttgatgca gctgcggat
13381 gcatcaacgt ttttaaacgg gtttgcggty taagtgaagc cctgtttaca oogtgggca

```

图 10 续

```

13441 caggcactag taotgatgto gtctacaggg cttttgatat ttacaacgaa aaaagtgctg
13501 gttttgcaaa gttcctaana actaattgct gtcgcttoca ggagaaggat gaggaaggca
13561 atttattaga ctcttaoott gtatgtaaga ggcatactat gtctaactac caacatgaag
13621 agactattta taacttggtt aaagattgto cageggttg cgtocatgac tttttcaagt
13681 tttagataga tgggtgacatg gtaccacata taccacgtca cgtctactac aaatacaca
13741 tggotgattt agtctatgct ctacgtcatt ttgatgagg ttgattgatt acattaaaag
13801 aaataactct caoatacaat tgotgtgatg atgattattt caataagaag gattggtatg
13861 acttogtaga gaatoctgao atcttaogog tatatgctaa cttaggtgag ogtgtacgcc
13921 aatoattatt aaagaotgta caattctgog atgctatgog tgatgocagg attgtaggog
13981 tactgacatt agataatcag gatcttaatg ggaactggta cgtattcggg gatttogtac
14041 aagtagcacc aggtctcggg gtccctattg tggattcata ttactcattg ctgatgccca
14101 tectoacttt gactagggca ttggtgctg agtocoatat ggatgctgat ctgcgaaaac
14161 caactattaa gtgggatttg ctgaaatat attttaogga agagagactt tgtctcttcg
14221 accgtatttt taaatatgg gacpagaact accatcocoa ttgtattaac tgtttgattg
14281 ataggtgtat ccttcattgt gcaaacotta atgtgttatt ttctactgtg tttcoacta
14341 caagttttgg accactagta agaaaaatat ttgtgatgg tgttctttt gttgtttcaa
14401 ctggatacca ttttctgtag ttaggagtog tacataatca ggatgtaaac ttacatagct
14461 cgcgtctcag tttcaaggaa cttttagtgt atgctgctga tccagctatg catgacgctt
14521 ctggcaattt attgctagat aaacgcacta catgcttttc agtagctgca ctacaaaaca
14581 atgttgcttt toaaactgto aaaccoggta attttaataa agacttttat gactttgctg
14641 tgtotaagg ttcttttaag gaaggagtt ctgttgaact aaaacactc ttctttgctc
14701 aggatggcaa ogctgctatc atgtattatg aotattatog ttataatctg ccaacaatgt
14761 gtgatatcag acaactccta ttctgtagtg aagttgttga taataoott gattgttacg
14821 atggtggctg tattaatgcc aaccaagtaa tctgtaaaaa tctggataaa tcaactggtt
14881 tccattttaa taaatgggg aaggctagac tttattatga ctcaatgagt tatgaggatc
14941 aagatgoact tttcogctat aotaagogta atgtcoatoc tactataact caaatgaatc
15001 ttaagtatgo cattagtgoa aagaatagag ctgcacocgt agctgggtgc tctatctgta
15061 gtactatgac aaatagacag tttoatcaga aattattgaa gtcaatagcc gccoatagag
15121 cgtctactgt ggtaattgga acaagcaagt tttacoggtg ctggcataat atgttaaaaa
15181 ctgtttacag tgatgtagaa actocacacc ttatgggttg ggattatcca aaatgtgaca
15241 gagccatgoc taacatgott aggataatgg, cctctcttgt tcttgotcgc aaacataaca
15301 ottgotgtaa ottatcacac cgtttctaca ggtagctaa caggtgtgog caagtattaa
15361 gtgagatggt catgtgtggo ggtoactat atgttaaaac aggtggaaca tcatcoggtg
15421 atgctacaac tgottatgot aatagtgctt ttaacattg tcaagotgtt acagccaatg
15481 taaatgoact tctttcaact gatggtaata agatagotga caagtatgto ogoaatctac
15541 aacacaggt ctatgagtgt ctctatagaa atagggatgt tgatcatgaa ttogtgatg
15601 agttttaocg ttacttgcgt aaacattctt ccatgatgat tctttctgat gatgocgtg
15661 tgtgtataaa cagtaactat gcggctoag gtttagtagc tagcattaag aactttaagg
15721 oagttcttta ttatcaaaat aatgtgttca tgtctgaggg aaaatgttgg actgagactg
15781 acottaactaa aggacctcac gaattttgot cacagcatac aatgctagt aaacaaggag
15841 atgattaogt gtacctgcot taaccoagato catcaagaat attaggogca ggctgtttt
15901 tcatgatgat tgtcaaaaac gatggtaoac ttatgattga aaggtctgtg tcaactggcta
15961 ttgatgctta ccoacttaca aaacatccta atcaggagta tgctgatgto tttcacttgt
16021 atttaoata cattagaaag ttacatgatg agcttactgg ccacatgttg gacatgtatt
16081 ocgtaatgct aactaatgat aacaocctac ggtactggga aocctggtt tatgaggcta
16141 tgtacacac acatacagto ttgcaggotg taggtgottg tgtattgtgc aattcacaga
16201 ctctactctg ttgcggtgcc tgtattagga gaccattcct atgttgoaag tgotgotatg
16261 accatgtcat ttcaacatca cacaaatag tgtgtotgt taatootat gtttgcaatg
16321 cccaggttg tgatgtoact gatgtgacac aactgtatct aggaggtatg agotattatt
16381 gcaagtacaca taagcctccc attagtttto cattatgtgo taatggcag gtttttgggt
16441 tatacaaaaa cacatgtgta ggcagtgaca atgtoactga cttaactgg atagcaaat
16501 gtgattggac taatgctggc gattacatac ttgocacac ttgtactgag agactcaagc
16561 ttttcgcagc agaaacgcto aaagcoactg aggaaacatt taagctgtoa tatggtattg
16621 cactgtacg cgaagtactc tctgacagag aattgcatct ttoatgggag gttggaaaao
16681 ctgaccacc attgaacaga aactatgtct ttactggta cctgttaact aaaaatagta
16741 aagtacagat tggagagtac acotttgaag aaggtgacta tgggtgatgct gttgtgtaca

```

图 10 续

```

16801 gaggtactac gacatacaag ttgaatggtg gtgattaott tgtgttgaoa tctcaactg
16861 taatgccact tagtgcacct actctagtgc cacaagagca ctatgtgaga attactggct
16921 tgtacccaac actcaacatc tcagatgagt tttctagcaa tgttgcaaat tatcaaaagg
16981 toggcattgca aaagtactct acaactocaag gaocacotgg tactggtaag agtcattttg
17041 ccacgggact tgctctotat taccoatctg ctogootagt gtatacggca tgotctcatg
17101 cagctgttga tgcoctatgt gaaaaggcat taaaatattt gcccatagat aaatgtagta
17161 gaatcatacc tgccgctgcg cgcgtagagt gttttgataa attcaaagtg aattcaacac
17221 tagaacagta tgttttctgc actgtaaagt cattgccaga aacaactgct gacattgtag
17281 totttgtaga aatctctatg gctactaatt atgacttgag tgtgtcaat gctagacttc
17341 gtgcaaaaaca ctacgtctat attggcgatc otgctoatc aocagcccc cgcacattgc
17401 tgactaaagg cacactagaa ccagaatatt ttaattcagt gtgcagactt atgaaaacaa
17461 taggtccaga catgttcott ggaactgtgc googttgtoc tgcgaaatt gttgacactg
17521 tgagtgtttt agtttatgac aataagotaa aagcaacaaa ggataagtca gotcaatgct
17581 tcaaaatggt ctacaaagggt gttattacac atgatgtttc atctgoaatc aacagacctc
17641 aaataggogt tgtaagagaa tttcttaoac goaatctctc ttggagaaaa gotgttttta
17701 tctcacotta taattoacag aaogctgtag cttcaaaaat cttaggattg cctacgcaga
17761 ctgttgattc atcacagggt tctgaatatg actatgtcat attcacacaa actactgaaa
17821 cagocactc ttgtaatgtc aaccgcttca atgtggotat caoaagggca aaaattggca
17881 ttttgtgat aatgtctgat agagatcttt atgacaaact gcaatttaca agtctagaaa
17941 taccacgtcg caatgtggct aottacaag cagaaaatgt aactggactt ttaaggact
18001 gtagtaagat cattactggt cttcatccta cacaggcacc tacacacctc agcgttgata
18061 taaaattcaa gaotgaagga ttatgtgttg acataccagg cataccaay gacatgacct
18121 aocgtagaot catctotatg atgggttcca aaatgaatta ccaagtcaat ggttacotta
18181 atatgtttat oacccgcgaa gaagctattc gtcacgttcg tgcgtggatg ggcgttgatg
18241 tagagggctg tcatgcaact agagatgotg tgggtactaa cctacctctc cagctaggat
18301 tttctoaagg tgtaactta gtactgtac cgactggta tgttgacact gaaaataaca
18361 cagaattcac cagagttaat gcaaacctc caccagggtg ccagtttaa catcttatac
18421 cactcatgta taaggcttg cctggaatg tagtgggtat taagatagta caaatgctca
18481 gtgatacact gaaaggattg tcagacagag togtgttogt ootttggcg catggtttg
18541 agcttacatc aatgaagtac tttgtoaaga ttggaootga aagaacgtyt tgtctgtgtg
18601 acaaacgtgc aacttgcttt totacttcat cagatactta tgootgctgg aatcattctg
18661 tgggttttga ctatgtctat aaccattta tgattgatgt tcagcagtyg ggccttaccg
18721 gtaacottca gagtaacct gaccaacatt gccaggtaca tggaaatgca catgtggcta
18781 gttgtgatgc tatcatgact agatgtttag cagtccatga gtgctttgtt aagcggctg
18841 attggtctgt tgaatacoot attataggag atgaaotgag ggttaattct gottgcagaa
18901 aagtacaaca catggtgtg aagtotgoat tgottgotga taagtttoca gttottoatg
18961 acattggaat tccaaaggct atcaagtgtg tgccctcaggc tgaagtagaa tggaaagtct
19021 acgatgotca gccatgtagt gacaaagctt aczaaataga ggaactctc tattcttatg
19081 ctacacatca ogataaatto actgatggtg tttgtttgtt ttggaattgt aacgttgato
19141 gttacocagc oaatgcaatt ktgtgtaggt ttgacacaag agtcttgtca aacttgaact
19201 taocaggctg tgatggtggt agtttgtatg tgaataagoa tgcattccac actccagctt
19261 togataaaag tgoatttaot aatttaaago aattgocctt othttactat totgatagtc
19321 cttgtgagtc tcatggoaaa oaagttaggt cggatattga ttatgttoca ctcaaatotg
19381 ctacgtgtat tacaogatgo aatttaggtg gtgotgtttg cagacaccat gcaaatgagt
19441 aocgacagta ottggatgca tataatatga tgattctctg tggatttagc ctatggattt
19501 acaaaoaatt tgatacttat aacctgtgga atacatttac caggttaoag agtttagaaa
19561 atgtggetta taatgttgtt aataaaggac aotttgatgg acacgcggc gaagcacctg
19621 tttocatcat taataatgct gtttacacaa aggtagatgg tattgatgtg gagatctttg
19681 aaaataagac aacacttctt gttaatgttg oatttgagct ttgggctaag cgtaacatta
19741 aaccagtgcc agagattaag ataactcaata atttgggtgt tgatctcgot gctaatactg
19801 taactctgga ctacaaaaga gaagccccag cacatgtatc tacaatagggt gctcgcacaa
19861 tgaotgacat tgccaagaaa cctactgaga gtgotgttct ttcacttaot gctgtgtttg
19921 atggttagat ggaaggacag gttagcotht ttgaaaacgc cogtaatggt gttttaataa
19981 cagaaggctc agtcaagggt ctaaacottt caaagggacc agcacaagct agogtcaatg
20041 gagtcacatt aattggagaa tcagtaaaaa caoagtttaa ctactttaag aaagtagacg
20101 gcattattca acagttgoot gaaacctact ttaotcagag cagagactta gaggatttta

```

图 10 续

```

20161 agcccagato acaaatggaa actgactttc togagotogc tatggatgaa ttcatacagc
20221 gatataagct cgagggctat gocttogaac acatcgttta tggagatttc agtcatggac
20281 aacttggcgg tcttoattta atgataggct tagocaagcg ctoacaagat tcaccactta
20341 aatttagagga ttttatcoot atggacagca cagtgaanaa ttacttcata acagatgcgc
20401 aaacaggttc atoaaaatgt gtgtgttotg tgattgatct tttacttgat gactttgtcg
20461 agataataaa gtcadaagat ttgtoagtga tttcaaaagt ggtoaagggt acaattgact
20521 atgctgaaat ttcattcatg ctttgggtga aggatggaca tgttgaaac ttctaccaa
20581 aactacaagc aagtcaagcg tggcaaacag gtgttgcgat gcctaacttg tacaagatgc
20641 aaagaatgct tcttgaaaag tgtgaccttc agaattatgg tgaaaatgct gttataccaa
20701 aaggaataat gatgaatgtc gcaaagtata ctcaactgtg tcaactacta aatacactta
20761 ctttagctgt accotacaac atgagagtta ttoactttgg tgctggctct gataaaggag
20821 ttgcaocagg tacagctgtg ctoagacaat ggttgocaac tggcacacta cttgtcgatt
20881 cagatcttaa tgaactcgtc tccgaogcag attctacttt aattggagac tgtgcaacag
20941 taocatacggc taataaatgg gacottatta ttagcgatat gtatgaocot aggaccaaac
21001 atgtgacaaa agagaatgac tctaagaag ggtttttcac tbatotgtgt ggatttataa
21061 agcaaaaact agccctgggt ggttctatag ctgtaaagat aacagagcat tcttggaatg
21121 atgacactta caagcttatg ggccatttct catggtggac agcttttgtt acaaatgtaa
21181 atgcatcato atoggaagoa ttttaattg gggctaacta tcttggcaag ccgaaggaac
21241 aaattgatgg ctataccatg catgctaact acattttctg gaggaacaca aatcctatcc
21301 agttgtcttc ctattcactc tttgaoatga goaaatctcc tottaaatta agaggaaactg
21361 ctgtaatgto tcttaaggag aatcaaatca atgatatgat ttattctctt ctggaaaaag
21421 gtaggcttat cattagagaa aacaaoagag ttgtggtttc aagtgatatt cttgttaaca
21481 actaaacgaa catgtttatt ttcttattat ttcttactct cactagtggt agtgaocottg
21541 aocggtgca cacttttgat gatgttcaag ctcootaatta cactcaacat acttcatcta
21601 tgaggggggt ttactatcct gatgaaattt ttagatcaga cactctttat ttaactcagg
21661 atttatktot tccatrttat tctaagtta cagggtttca tactattaat catacgtttg
21721 gcaaccctgt catacctttt aaggatggta tttattttgc tgoaacagag aatcaaatg
21781 ttgtcogtgg ttgggttttt ggttotacca tgaacaacaa gtcacagtog gtgattatta
21841 ttaoaatct taactaatgtt gttatacagag catgtaactt tgaattgtgt gacaacocct
21901 tctttgctgt ttctaaacc c atgggtacac agaacacatac tatgatattc gataatgcat
21961 ttaattgca tttogagtao atatctgatg cctttctgct tgatgtttca gaaaagtcag
22021 gtaattttaa acaacttaoga gatgttgtgt ttaaaaataa agatgggttt ctctatgttt
22081 ataagggcta tcaacotata gatgtagtto gtgatctacc ttctggtttt aacactttga
22141 aacctatttt taagttgcct cttggtatta acattacaaa ttttagagoc attcttaoag
22201 ccttttcacc tgctcaagac atttggggca cgtcagctgc agcctatttt gttggtatt
22261 taaagocaa c taactttatg ctoaaagtat atgnaaatgg tacaatcaca gatgctgtt
22321 attgttotoa aaatccactt gctgaactoa aatgctctgt taagagcttt gagattgaca
22381 aaggaattta coagacotot aatctcaggg ttgttcooto aggagatggt gtgagattcc
22441 ctaattttac aaacttgtgt ootttggag aggtttttaa tgotactaaa ttccctctg
22501 totatgcatg ggagagaaaa aaaatttota attgtgttgc tgattactct gtgctotaos
22561 actcaacatt tttttcaacc ttttaagtgt atggcgtttc tgooaotaag ttgaatgato
22621 tttgototo oaatgtotat gcagattcct ttgtagtcaa gggagatgat gtaagacaaa
22681 tagcgcaggg acaaaotggt gttattgctg attataatta taaattgoca gatgatttoa
22741 tgggttgtgt cttgcttggt aatactagga acattgatgo tacttoaac ggtaattata
22801 attataata taggtatcct agaatggca agottaggcc ctttgagaga gacatatota
22861 atgtgccttt ctcccctgat ggcococctt gcacocococ tgotottaat tgttattggc
22921 cattaaatga ttatggtttt tacacocacta ctggoattgg ctaccaacot tacagagttg
22981 tagtactttc ttttgaactt ttaaatgcac oggccacggg ttgtggaoca aaattatcoa
23041 ctgacottat taagaaccag tgtgtcaatt ttaattttaa tggactcact ggtactggty
23101 tgttaacotc ttottoaaag agatttcaac catttcaaca atttggccgt gatgtttctg
23161 atttcactga ttcogttoga gatootaaaa catotgaaat attagacatt toacctgtot
23221 ootttggggg tytaagtgtta attacacotg gaocaaatgc ttcactctgaa gttgotgtto
23281 tatacaaga tgttaactgo actgatgttt ctacagoaat tcatgcagat caactocac
23341 cagcttggcg catatattct actggaacaa atgtattcca gactcaagca ggctgtotta
23401 taggagctga gcatgtogac acttcttatg agtgcgacat tccatttggg gctggcattt
23461 gtgctagtta ccatacagtt totttattac gtagtactag ccaaaaatct atttgggtt

```

图 10 续

```

23521 ataactatgtc tttaggtgct gatagttcoa ttgcttactc taataacaoc attgotatac
23581 ctactaactt ttoaattago attactacag aagtaatgcc tgtttctatg gctaaaacct
23641 ccgtagattg taatatgtac atctgocggag attotaotga atgtgctaata ttgcttctcc
23701 aatatggtag cttttgocaa caaotaaatc gtgcaactctc aggtattgct gctgaacagg
23761 atcgcaaac acgtgaagtg ttogctcaag toaaaocaaat gtacaaaaoc coaactttga
23821 aatatttttg tggttttaat ttttoacaaa tattaactga ccctctaaag coaactaaga
23881 ggtcttttat tgaggacttg ctctttaata aggtgaoact cgctgatgot ggottoatga
23941 agcaatatgg cgaatgccta ggtgatatta atgctagaga tctcatttgt gcgcagaagt
24001 tcaatggact tacagtgttg ccacctctgc tcactgatga tatgattgct goctacactg
24061 ctgctctagt tagtggctact gccactgctg gatggacatt tgggtgctggc gctgctcttc
24121 aaataccttt tgctatgcaa atggcatata ggttcaatgg cattggagtt acccaaaatg
24181 ttctctatga gaaccacaaa oaaatcgoca accaatttaa caaggcgatt agtcaaattc
24241 aagaatcact tacaacaaca tcaactgcat tgggcaagot gcaagacgtt gtaaccaga
24301 atgctcaagc attaaacaca ctgtttaaac aacttagctc taattttggt gaatttcaa
24361 gtgtgctaaa tgatacctt tcgcgacttg ataaagtcca ggcggaggtg caaattgaca
24421 ggttaattac aggcagactt caaagccttc aaacctatgt aacacacaa ctaatcaggg
24481 ctgctgaaat cagggcttct gctaactctg ctgctactaa aatgtctgag tgtgtcttg
24541 gacaatcaaa aagagttgac ttttgtggaa agggctacca ccttatgtcc ttcccacag
24601 cagocccgoa tgggtttgto ttootacatg toacgtatgt gccatcccag gagaggaact
24661 tcaccacagc gccagcaatt tgcctatgag gcaaagcata ottcoctogt gaagggtgtt
24721 ttgtgtttaa tggcacttct tggtttatta cacagaggaa ctctctttct ccaaaaata
24781 ttactacaga caatacattt gtctcaggaa atgtgtatgt cgttattggc atcattaaca
24841 acacagttta tgatcctctg caacctgagc ttgactcatt caaagaagag ctggacaagt
24901 acttcaaaaa tcatacatoa ccagatgttg atcttggcga catttcaggc attaaogctt
24961 ctgtcgtcaa cattcaaaaa gaaattgacc gcotcaatga ggtocgtaaa aatttaaatg
25021 aatcactcat tgacctcaa gaattgggaa aatatgagca atatatataa tggccttgg
25081 atgtttggct cggcttcatt gctggactaa ttgocatogt catggttaca atottgottt
25141 gttgcatgac tagttgttgc agttgoccca aggggtcatg ctcttbtggt tcttgtctga
25201 agtttgatga ggatgactot gagocagttc tcaaggggtg caaattacat tacacataaa
25261 cgaactatg gatttgttta tgagattttt taotcttggg tcaattactg cacagccagt
25321 aaaaattgac aatgcttctc ctgcaagtac tgttoatgct acagocaaog tacogctaca
25381 agcttaoact cctttoggat tggctgttat tggogttgca tttottgtga tttttcagag
25441 cgctaccaaa ataattgctc tcaataaaaag atggcagcta gccctttata agggottcca
25501 gttcatttgc aatttactgc tgctatttgt taccatctat tcacatcttt tgccttgcgc
25561 tgcaggtaag gaggcgcaat ttttgtacct ctatgcottg atatattttc tacaatgcat
25621 caacgcagtg agaattatta tgagatgttg gctttgttgg aagtgcfaat ccaagaacc
25681 attactttat gatgcaact actttgtttg ctggcacaca cataactatg actactgtat
25741 accatataac agtgtcaog ataoaattgt cgttactgaa ggtgacggca ttcaacacc
25801 aaaaactaaa gaagaotaco aaattgggtg ttattotgag gataggcaot oagggtgtaa
25861 agactatgto gttgtacatg gctatttoac cgaagtttac taocagctg agtotacaa
25921 aattactaca gacactggta ttgaaaatgc facattcttc atctttaaoc agcttgttaa
25981 agacccaccg aatgtgcaaa tacacocaaat cgacggctct tcaggagttg otaatccagc
26041 aatggatcca atttatgatg agccgaogac gaotactagc gtgootttgt aagcacaaga
26101 aagttagtac gaacttatgt actcattcgt ttccggaaga acaggtacgt taatagttaa
26161 tagcgtactt cttttctctg ctctcgtggt attottgcta gtcacaotag ccatccttac
26221 tgcgcttoga ttgtgtgctg actgctgcaa tattgttaac gtgagtttag taaaaccaac
26281 ggtttacgto tactogcgtg ttaaaaatct gaactcttct gaaggagttc ctgatcttct
26341 ggtotaaacg aactaactat tattattatt ctgtttgaa ctttaacatt gottatcatg
26401 gcagacaacg gtactattac cgttgaggag ottaaacaac tcctggaaca atggaacct
26461 gtaatagggt tcotattcct agcctggatt atgttactac aatttgcta ttctaactg
26521 aacaggtttt tgtacataat aaagottggt ttcoctctgg totttgtggc agtaacactt
26581 gcttgttttg tgottgctgt tgtctacaga ataaattggg tgactggggc gattgcgatt
26641 goaatggctt gtattgtagg ctgtatgttg cttagetact tegtgtctc cttcaggctg
26701 tttctcgtta cccgctcaat gttgtcattc aocccagaaa caaacattct tctcaatgtg
26761 cctctocggg ggaacattgt gaccagaccy ctcatgaaa gtgaactgtt cattggtgct
26821 gtgatcatto gtggtcactt gcgaatggcc ggaoactccc tagggcctg tgacattaag

```

图 10 续

```

26881 gacctgcca aagagatcac tgtggctaca tcacgaacgc tttcttatta caaattagga
26941 gcgtcgagc gtgtaggcac tgattcaggt tttgotgcat acaaccgcta ccgtattgga
27001 aactataaat taaatacaga ccacgccggt agcaacgaoa atattgcttt gotagtacag
27061 taagtgacaa cagatgtttc atcttgttga ottccagggt acaatagocag agatattgat
27121 tatcattatg aggacttcoa ggattgctat ttggaatctt gacgttataa taagltcaat
27181 agtgagacaa ttatttaagc ctctaactaa gaagaattat tccggagttag atgatgaaga
27241 acctatggag ttagattatc oataaaaoga acatgaaaat tattotcttc ctgacattga
27301 ttgtatttac atcttgcgag ctatatcact atcaggagtg tgttagaggt acgactgtac
27361 tactaaaaga accttgccca tcaggaocat acgagggcaa ttcaccattt caccctcttg
27421 ctgacaataa atttgcacta acttgcacta gcacacactt tgcttttctg tgtgctgaog
27481 gtactcgaca tacctatcag ctgogtgcaa gatcagtttc accaaaaott ttcacogao
27541 aagaggaggt tcaacaagag ctctactogc cactttttct cattgttctg gctctagtat
27601 ttttaatact ttgcttcacc attaagagaa agacagaatg aatgagctca ctttaattga
27661 cttctatttg tgctttttag ctttctctgt attccttgtt ttaataatgc ttattatatt
27721 ttggttttca otcgaatcc aggatctaga agnaccttgt accaaagtct aaacgaacat
27781 gaaacttctc attgttttga ctgtatttbo tctatgcagt tgcatatgca ctgtagtaca
27841 gcgctgtgca tctaataaac ctcatgtctg tgaagatcct tghtaagttac aocactaggg
27901 gtaataacta tagcactgct tggctttgtg ototaggaaa ggttttacct tttcatagat
27961 ggcacactat ggttcaaaca tgcacacota atgttaactat oaactgtcaa gatccagctg
28021 gtgggtgcgct tatagctagg tgttggtaacc ttcattgaagg tcaccaaaact gctgcattta
28081 gagacgtact tgttgtttta aataaacgaa caaattaaaa tgtctgataa tggaccccaa
28141 tcaaaocaaac gtatgtcccc ccgcattaca tttggtggac ccacagattc aactgacaat
28201 aaccagaatg gaggacgcaa tgggggaaagg ccaaaacagc gcogaccoca aggtttacc
28261 aataatactg cgtcttgggt cacagctctc actcagcatg gcaaggagga acttagattc
28321 cctcagggcc agggogttcc aatcaacacc aatagtggtc cagatgacca ctgtgctac
28381 taccgaagag ctacocgaog agttctgggt ggtgaogga aatgaaaga gctcagcccc
28441 agatggtaact tctattaoot aggaactggc ccagaagctt caottccota cggcgctaac
28501 aaagaaggca tctgtatgggt tgcactgag ggagccttga atacacccaa agaccacatt
28561 ggcaccogca atcctaataa caatgctgcc accgtgctac aacttctctc aggaacaaca
28621 ttgocaaaag gottotacgc agagggaaagc agaggoggoa gtcaagootc tctctogotcc
28681 tcatcacgta gtcgoggtaa ttcaagaaat toaactoctg gcagcagtag gggaaattct
28741 cctgctcgaa tggctagcgg aggtggtgaa actgccctcg cgctattgct gctagacaga
28801 ttgaaccagc ttgagagcaa agttctgggt aaaggccaac aacaacaagg caaactgtc
28861 aotaagaat ctgctgctga ggcactctaa aagcctogcc aaaaacttac tgccacaaaa
28921 cagtaocaaog tcaotcaagc atttgggaga cgtggctccag aacaaacca aggaaatttc
28981 ggggacocaaog acotaatcag aocaggaact gattacaaac attggcogca aattgcacaa
29041 tttgctocaa gtgootctgc attctttgga atgtcacgca ttggcatgga agtcacaoot
29101 tcgggaacat ggctgactta tcatggagcc attaaattgg atgacaaaga tccacaattc
29161 aaagocaaog tcatactgct gaacaagcac attgacgcat acaaaacatt cccacocaaa
29221 gagcctaaaa aggacaaaaa gaanaagact gatgaagoto agccttggc goagagocaa
29281 aagaagcagc ccactgtgac tcttctctct cgggctgaca tggatgattt ctccagacaa
29341 cttcaaaatt ccatgagtgg agcttctgct gattcaactc aggcataaac actcatgatg
29401 accacacaag gcagatgggc tatgtaaacy ttttcgcaat tccgtttacg atacatagtc
29461 tactcttctg cagaatgaat tctcgttaot aaacagocaa agtaggttta gtttaactta
29521 atctcacata gcaatcttta atcaatgtgt aacattaggg aggaottgaa agagccacca
29581 cattttcato gaggccacgc ggagtacgat cgaggggtaca gtgaataatg ctagggagag
29641 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta tccccatgtg
29701 attttaatag cttcttagga gaatgacaaa aaaaaaaaaa aa

```

图 10 续


```

1 - ATATTAGGTTTTTACCTACCCAGGAAAAGCCAAACCAACCTCGATCTCTGTAGATCTGTT - 60
- I L G F Y L P R K S Q P T S I S C R S V
- Y * V F T Y P G K A N Q P R S L V D L F
- I R F L P T Q E K P T N L D L L * I C S
61 - CTCTAAACGAACCTTAAATCTGTGTAGCTGTGCTCGGCTGCATGCCTAGTGCACCTAC - 120
- L * T N F K I C V A V A R L H A * C T Y
- S K R T L K S V * L S L G C M P S A P T
- L N E L * N L C S C R S A A C L V H L R
121 - GCAGTATAACAATAATAAATTTACTGTCGTGACAAGAAACGAGTAACTCGTCCCTCT - 180
- A V * T I I N F T V V D K K R V T R P S
- Q Y K Q * * I L L S L T R N E * L V P L
- S I N N N K F Y C R * Q E T S N S S L F
181 - TCTGCAGACTGCTACGGTTTCGTCCGTGTCAGTGCATCAGCATACCTAGGTTTC - 240
- S A D C L R F R P C C S R S S A Y L G F
- L Q T A Y G F V R V A V D H Q H T * V S
- C R L L T V S S V L Q S I I S I P R F R
241 - GTCCGGTGTGACCGAAAGGTAAGATGGAGGCCTTGTCTTGGTGTCAACGAGAAAACA - 300
- V R V * P K G K M E S L V L G V N E K T
- S G C D R K V R W R A L F L V S T R K H
- P G V T E R * D G E P C S W C Q R E N T
301 - CACGTCGAACCTCAGTTTGCCTGTCCITCAGGTTAGAGACGTGCTAGTGCCTGGCTCGGG - 360
- H V Q L S L P V L Q V R D V L V R G F G
- T S N S V C L S F R L E T C * C V A S G
- R P T Q F A C P S G * R R A S A W L R G
361 - GACTCTGTGGAAGAGGCCCTATCGGAGGCACGTGAACACCTCAAAAATGGCACTTGTGGT - 420
- D S V E E A L S E A R E H L K N G T C G
- T L W K R P Y R R H V N T S K M A L V V
- L C G R G P I G G T * T P Q K W H L W S
421 - CTAGTAGAGCTGGAAAAGGCGTACTGCCAGCTTGAACAGCCCTATGTGTTCAATAAA - 480
- L V E L E K G V L P Q L E Q P Y V F I K
- * * S W K K A Y C P S L N S P M C S L N
- S R A G K R R T A P A * T A L C V H * T
481 - CGTTCGTGATGCCCTAAGCACCAATCACGGCCACAAGGTCGTGAGCTGGTTCAGAAATG - 540
- R S D A L S T N R G H K V V E L V A E M
- V L M P * A P I T A T R S L S W L Q K W
- F * C L K H Q S R P Q G R * A G C R N G
541 - GACGGATTACAGTACGGTCTAGCGGTATAACACTGGGAGTACTCGTCCACATGTGGGC - 600
- D G I Q Y G R S G I T L G V L V P H V G
- T A F S T V V A V * H W E Y S C H M W A
- R H S V R S * R Y N T G S T R A T C G R
601 - GAAACCCCAATTCATACCGCAATGTTCTTCTTCTGTAAGAACGGTAATAAGGGAGCCGGT - 660
- E T P I A Y R N V L L R K N G N K G A G
- K P Q L H T A M F F F V R T V I R E P V
- N P N C I P Q C S S S * E R * * G S R W
661 - GGTCATAGCTATGGCATCGATCTAAAGTCTTATGACTTAGGTGACGAGCTTGGCACTGAT - 720
- G H S Y G I D L K S Y D L G D E L G T D
- V I A M A S I * S L M T * V T S L A L I
- S * L W H R S K V L * L R * R A W H * S
721 - CCCATTGAAGATTATGAACAAAACCTGGAACACTAAGCATGGCAGTGGTGCCTCCGIGAA - 780
- P I E D Y E Q N W N T K H G S G A L R E
- P L K I M N K T G T L S M A V V H S V N
- H * R L * T K L E H * A W Q W C T P * T
781 - CTCACTCGTGAGCTCAATGGAGGTGCAGTCACTCGCTATGTGACAACAATTTCTGTGGC - 840
- L T R E L N G G A V T R Y V D N N F C G
- S L V S S M E V Q S L A M S T T I S V A
- H S * A Q W R C S H S L C R Q Q F L W P

```

图 11

```

841 - CCAGATGGGTACCCTCTTGATTGCATCAAAGATTTTCTCGCACGGCGGGCAAGTCAATG - 900
- P D G Y P L D C I K D F L A R A G K S M
- Q M G T L L I A S K I F S H A R A S Q C
- R W V P S * L H Q R F S R T R G Q V N V
901 - TGCACICTTTCCGAACAACCTTGATTACATCGAGTCGAAGAGAGGTTGCTACTGCTGCCGT - 960
- C T L S E Q L D Y I E S K R G V Y C C R
- A L F P N N L I T S S R R E V S T A A V
- H S F R T T * L H R V E E R C L L P *
961 - GACCATGAGCATGAAATGCGTTCCTGACTGAGCGCTCTGATAAGAGCTACGAGCACCAG - 1020
- D H E H E I A W F T E R S D K S Y E H Q
- T M S M K L P G S L S A L I R A T S T R
- P * A * N C L V H * A L * * E L R A P D
1021 - ACACCCCTCGAAATTAAGAGTGCCAAGAAATTTGACACTTTCAAAGGGGAATGCCCAAAG - 1080
- T P F E I K S A K K F D T F K G E C P K
- H P S K L R V P R N L T L S K G N A Q S
- T L R N * E C Q E I * H F Q R G M P K V
1081 - TTTGTGTTTCTTAAGTCAAAGTCAAAGTCAATCAACCACGTTGTTGAAAAGAAAAG - 1140
- F V F P L N S K V K V I Q P R V E K K K
- L C F L L T Q K S K S F N H V L K R K R
- C V S S * L K S Q S H S T T C * K E K D
1141 - ACTGAGGGTTTCATGGGGGTATACGCTCTGTGTACCCTGTGCATCTCCACAGGAGTGT - 1200
- T E G F M G R I R S V Y P V A S P Q E C
- L R V S W G V Y A L C T L L H L H R S V
- * G F H G A Y T L C V P C C I S T G V *
1201 - AACAAATGCACTTGTCTACCTTGATGAAATGTAATCATTGCGATGAAGTTTCATGGCAG - 1260
- N N M H L S T L M K C N H C D E V S W Q
- T I C T C L P * * N V I I A M K P H G R
- Q Y A L V Y L D E M * S L R * S F M A D
1261 - ACGTGCCTACTTTCTGAAAGCCACTTGTGAACATGTGGCACTGAAAATTTAGTTATTGAA - 1320
- T C D F L K A T C E H C G T E N L V I E
- R A T F * K P L V N I V A L K I * L L K
- V R L S E S H L * T L W H * K F S Y * R
1321 - GGACCTACTACATGTGGGTACCTACCTACTAATGCTGTAGTGAATAATGCCATGCTCTGCC - 1380
- G P T T C G Y L P T N A V V K M P C P A
- D L L H V G T Y L L M L * * K C H V L P
- T Y Y M W V P T Y * C C S E N A M S C L
1381 - TGTCAAAGACCCAGAGATTGGACCTGAGCATAGTGTGCAGATTATCACAACCACTCAAAC - 1440
- C Q D P E I G P E H S V A D Y H N H S N
- V K T Q R L D L S I V L Q I I T T T Q T
- S R P R D W T * A * C C R L S Q P L K H
1441 - ATTGAAACTCGACTCCGCAAGGGAGGTAGGACTAGATGTTTTGGAGGCTGTGTGTTTGGC - 1500
- I E T R L R K G G R T R C F G G C V F A
- L K L D S A R E V G L D V L E A V C L P
- * N S T P Q G R * D * M F W R L C V C L
1501 - TATGTTGGCTGCTATAATAAGCGTGCCTACTGGGTTCTCGTGTAGTGTGATATTGGC - 1560
- Y V G C Y N K R A Y W V P R A S A D I G
- M L A A I I S V P T G F L V L V L I L A
- C W L L * * A C L G S S C * C * Y W L
1561 - TCAGGCATACTGSCATTACTGGTGACAATGTGGAGACCTTGAATGAGGATCTCCTTGAG - 1620
- S G H T G I T G D N V E T L N E D L L E
- Q A I L A L L V T M W R P * M R I S L R
- R P Y W H Y W * Q C G D L E * G S P * D
1621 - ATACTGAGTCGTGAACGTGTTAACATTAACATTGTTGGCGATTTTCATTGAATGAAGAG - 1680
- I L S R E R V N I N I V G D F H L N E E
- Y * V V N V L T L L L A I F I * M K R
- T E S * T C * H * H C W R F S F E * R G

```

图 11 续

```

1681 - GTTGCATCATTTTGGCATCTTCTCTGCTTCTACAAGTGCCTTTATTGACACTATAAAG - 1740
- V A I I L A S F S A S T S A F I D T I K
- L P S F W H L S L L L Q V P L L T L * R
- C H H F G I F L C F Y K C L Y * H Y K E
1741 - AGTCTTGATTACAAGTCTTTCAAACCATTGTTGAGTCCTGCGGTAACATAAAGTTACC - 1800
- S L D Y K S F K T I V E S C G N Y K V T
- V L I T S L S K P L L S P A V T I K L P
- S * L Q V F Q N H C * V L R * L * S Y Q
1801 - AAGGGAAGCCCGTAAAGGTGCTTGGACATTGGACAACAGAGATCAGTTTAAACCCA - 1860
- K G K P V K G A W N I G Q Q R S V L T P
- R E S P * K V L G T L D N R D Q F * H H
- G K A R K R C L E H W T T E I S F N T T
1861 - CTGTGTGTTTTCCCTCACAGGCTGCTGGTGTATCAGATCAATTTTTGCGGCACACTT - 1920
- L C B F P S Q A A G V I R S I F A R T L
- C V V F P H R L L V L S D Q F L R A H L
- V W F S L T G C W C Y Q I N F C A H T *
1921 - GATGCAGCAAACCACTCAATTCCTGATTTGCAAAGAGCAGCTGCACCATACTTGATGGT - 1980
- D A A N H S I P D L Q R A A V T I L D G
- M Q Q T T Q F L I C K E Q L S P Y L M V
- C S K P L N S * F A K S S C H H T * W Y
1981 - ATTTCTGAACAGTCATTACGCTTGTGCGAGCCATGGTTTATACTTCAGACCTGCTCACC - 2040
- I S E Q S L R L V D A M V Y T S D L L T
- F L N S H Y V L S T P W F I L Q T C S P
- F * T V I T S C R R H G L Y F R P A H Q
2041 - AACAGTGTCAATTATTATGGCATATGTAAGTGGTGGTCTTGTACAACAGACTTCTCAGTGG - 2100
- N S V I L M A Y V T G G L V Q Q T S Q W
- T V S L L W H M * L V V L Y N R L L S G
- Q C H Y Y G I C N W W S C T T D F S V V
2101 - TTGTCTAATCTTTTGGGCACTACTGTTGAAAACTCAGGCCTATCTTTGAATGGATTGAG - 2160
- L S N L L G T T V E K L R P I F E W I E
- C L I F W A L L L K N S G L S L N G L R
- V * S F G H Y C * K T Q A Y L * M D * G
2161 - GCGAAACTTAGTGCAGGAGTTGAATTTCTCAAGGATGCTTGGGAGATTCTCAAATTTCTC - 2220
- A K L S A G V E F L K D A W E I L K F L
- R N L V Q E L N F S R M L G R F S N F S
- E T * C R S * I S Q G C L G D S Q I S H
2221 - ATTACAGGTGTTTTGACATCGTCAAGGGTCAAATACAGGTTGCTTCAGATAACATCAAG - 2280
- I T G V F D I V K G Q I Q V A S D N I K
- L Q V F L T S S R V K Y R L L Q I T S R
- Y R C F * H R Q G S N T G C F R * H Q G
2281 - GATTGTGTAAAATGCTTCATTGATGTTGTTAACAAGGCACTCGAAATGTGCATTGATCAA - 2340
- D C V K C F I D V V N K A L E M C I D Q
- I V * N A S L M L L T R H S K C A L I K
- L C K M L H * C C * Q G T R N V H * S S
2341 - GTCATATCGCTGGCGCAAAGTTGCGATCACTCACTTAGGTGAAGTCTTCATCGCTCAA - 2400
- V T I A G A K L R S L N L G E V F I A Q
- S L S L A Q S C D H S T * V K S S S L K
- H Y R W R K V A I T Q L R * S L H R S K
2401 - AGCAAGGGACTTTACCGTCAGTGTATACGTGGCAAGGAGCAGCTGCAACTACTCATGCTT - 2460
- S K G L Y R Q C I R G K E Q L Q L L M P
- A R D F T V S V Y V A R S S C N Y S C L
- Q G T L P S V Y T W Q G A A A T T H A S
2461 - CTTAAGGCACCAAAGAAGTAACCTTTCTTGAAGGTGATTCACATGACACAGTACTTACC - 2520
- L K A P K E V T F L E G D S H D T V L T
- L R H Q K K * P F L K V I H M T Q Y L P
- * G T K R S N L S * R * F T * H S T Y L

```

图 11 续

```

2521 - TCTGAGGAGGTTGTTCTCAAGAACGGTGAAGCACTCGAGACGCCCGTTGATAGC - 2580
- S E E V V L K N G E L E A L E T P V D S
- L R R L F S R T V N S K H S R R P L I A
- * G G C S Q E R * T R S T R D A R * * L
2581 - TTCACAAATGGAGCTATCGTCGGCACACCAGTCTGTGTAATGGCCCTCATGCTCTTAGAG - 2640
- F T N G A I V G T P V C V N G L M L L E
- S Q M E L S S A H Q S V * M A S C S * R
- H K W S Y R R H T S L C K W P H A L R D
2641 - ATTAAGACAAAGAACAATACTGCGCATTGCTCCTGGTTTACTGGCTACAACAATGTC - 2700
- I K D K E Q Y C A L S P G L L A T N N V
- L R T K N N T A H C L L V Y W L Q T M S
- * G Q R T I L R I V S W F T G Y K Q C L
2701 - TTTGCGTTAAAAGGGGGTGCACCAATTAAGGTGTAACCTTGGAGAAGATACTGTTTGG - 2760
- F R L K G G A P I K G V T F G E D T V W
- F A * K G V H Q L K V * P L E K I L F G
- S L K R G C T N * R C N L W R R Y C L G
2761 - GAAGTTCAAGGTTACAAGAATGTGAGAATCACATTTGAGCTTGATGAACGTGTTGACAAA - 2820
- E V Q G Y K N V R I T F E J D E R V D K
- K F K V T R M * E S H L S L M N V L T K
- S S R L Q E C E N H I * A * * T C * Q S
2821 - GTGCTTAATGAAAAGTGTCTGTCTACACTGTTGAATCCGGTACCGAAGTACTGAGTTT - 2880
- V L N E K C S V Y T V E S G T E V T E F
- C L M K S A L S T L L N P V P K L L S L
- A * * K V L C L H C * I R Y R S Y * V C
2881 - GCATGTGTGTAGCAGAGGCTGTTGTGAAGACTTTACAACCAAGTTTCTGATCTCCTTACC - 2940
- A C V V A E A V V K T L Q P V S D L L T
- H V L * Q R L L * R L Y N Q F L I S L P
- M C C S R G C C E D F T T S F * S P Y Q
2941 - AACATGGGTATTGATCTTGATGAGTGGAGTGTAGCTACATTCTACTTATTTGATGATGCT - 3000
- N M S I D L D E W S V A T F Y L F D D A
- T W V L I L M S G V * L H S T Y L M M L
- H G Y * S * * V E C S Y I L L I * * C W
3001 - GGTGAAGAAAACTTTTCATCAGTATGTTTCCTTTTACCCCTCCAGATGAGGAAGAA - 3060
- G E B N F S S R M Y C S F Y P P D E E E
- V K K T F H H V C I V P F T L Q M R K K
- * R K L F I T Y V L F L L P S R * G R R
3061 - GAGGACGATGCAGAGTGTGAGGAAGAAGAAATGATGAAACCTGTGAACATGAGTACGGT - 3120
- E D D A E C E E E E I D E T C E H E Y G
- R T M Q S V R K K L M K P V N M S T V
- G R C R V * G R R N * * N L * T * V R Y
3121 - ACAGAGGATGATTATCAAGTCTCCCTCTGGAATTTGGTGCCTCAGCTGAAACAGTTCGA - 3180
- T E D D Y Q G L P L E F G A S A E T V R
- Q R M I I K V S L W N L V P Q L K Q F E
- R G * L S R S P S G I W C L S * N S S S
3181 - GTTGAGGAAGAAGAAGAGGAAGACTGGCTGGATGATACTACTGAGCAATCAGAGATTGAG - 3240
- V E E E E E E D W L D D T T E Q S E I E
- L R K K K R K T G W M I L L S N Q R L S
- * G R R R G R L A G * Y Y * A I R D * A
3241 - CCAGAACCAGAACCTACACCTGAAGACCAGTAAATCAGTTTACTGGTTATTTAAAACCT - 3300
- P E P E P T P E E P V N Q F T G Y L K L
- Q N Q N L H L K N Q L I S L L V I * N L
- R T R T Y T * R T S * S V Y W L F K T Y
3301 - ACTGACAAATGTTGCCATTAATGTGTTGACATCGTTAAGGAGGCACAAAGTGCATATCCT - 3360
- T D N V A I K C V D I V K E A Q S A N P
- L T M L P L N V L T S L R R H K V L I L
- * Q C C H * M C * H R * G G T K C * S Y

```

图 11 续

```

3361 - ATGGTGATTGTAAATGCTGCTAACATACACCTGAAACATGGTGGTGGTGTAGCAGGTGCA - 3420
- M V I V N A A N I H L K H G G G V A G A
- W * L * M L L T Y T * N M V V V * Q V H
- G D C K C C * H T P E T W W W C S R C T
3421 - CTCACCAAGGCAACCAATGGTGCCATGCAAAAGGAGAGTGATGATTACATTAAGCTAAT - 3480
- L N K A T N G A M Q K E S J D Y I K L N
- S T R Q P M V P C K R R V M I T L S * M
- Q Q G N Q W C H A K G E * * L H * A K W
3481 - GGCCCTCTTACAGTAGGAGGTCTTGTGTTGCTTCTGGACATAATCTTGCTAAGAAGTGT - 3540
- G P L T V G G S C L L S G H N L A K K C
- A L L Q * E G L V C F L D I I L L R S V
- P S Y S R R V L F A F W T * S C * E V S
3541 - CTGCATGTTGTTGGACCTAACCTAAATGCAGGTGAGGACATCCAGCTTCTTAAGGCAGCA - 3600
- L H V V G P N L N A G E D I Q L L K A A
- C M L L D L T * M Q V R T S S F L R Q H
- A C C W T * P K C R * G H P A S * G S I
3601 - TATGAAAATTTCAATTCACAGGACATCTTACTTGCACCATTGTTGTGACAGGCATATTT - 3660
- Y E N F N S Q D I L L A P L L S A G I F
- M K I S I H R T S Y L H H C C Q Q A Y L
- * K F Q F T G H L T C T I V V S R H I W
3661 - GGTGCTAAACCACTTCAGTCTTTACAAGTGTGCGTGCAGACGGTTCGTACACAGGTTAT - 3720
- G A K P L Q S L Q V C V Q T V R T Q V Y
- V L N H F S L Y K C A C R R F V H R F I
- C * T T S V F T S V R A D G S Y T G L Y
3721 - ATTGCAGTCAATGACAAAGCTCTTTATGAGCAGGTTGTCATGGATTATCTTGATAACCTG - 3780
- I A V N D K A L Y E Q V V M D Y L D N L
- L Q S M T K L F M S R L S W I I L I T *
- C S Q * Q S S L * A G C H G L S * * P E
3781 - AAGCCTAGAGTGAAGCACCTAAACAAGAGGAGCCACCAACACAGAAGATTCAAAACCT - 3840
- K P R V E A P K Q E E P P N T E D S K T
- S L E W K H L N K R S H Q T Q K I P K L
- A * S G S T * T R G A T K H R R F Q N *
3841 - GAGGAGAAATCTGTCGTACAGAAGCCTGTCGATGTGAAGCCAAAATTAAGGCCCTGCATT - 3900
- E E K S V V Q K P V D V K P K I K A C I
- R R N L S Y R S L S M * S Q K L R P A L
- G E I C R T E A C R C E A K N * G L H *
3901 - GATGAGGTTACCAACAACCTGGAAGAACTAAGTTTCTTACCAATAAGTTACTCTTGTTT - 3960
- D E V T T T L E E T K F L T N K L L L F
- M R L P Q H W K K L S F L P I S Y S C L
- * G Y H N T G R N * V S Y Q * V T L V C
3961 - GCTGATATCAATGGTAAGCTTTACCATGATTCTCAGAACATGCTTAGAGGTGAAGATATG - 4020
- A D I N G K L Y H D S Q N K L R G E D M
- L I S M V S F T M I L R T C L E V K I C
- * Y Q W * A L P * F S E H A * R * R Y V
4021 - TCTTTCCTTGAGAAGGATGCACCTTACATGGTAGGTGATGTTATCACTAGTGGTGATATC - 4080
- S F L E K D A P Y M V G D V I T S G D I
- L S L R R M H L T W * V M L S L V V I S
- F P * E G C T L H G R * C Y H * W * Y H
4081 - ACTTGTGTTGTAATACCCCTCAAAAAGGCTGGTGGCACTACTGAGATGCTCTCAAGAGCT - 4140
- T C V V I P S K K A G G T T E M L S R A
- L V L * Y P P K R L V A L L R C S Q E L
- L C C N T L Q K G W W H Y * D A L K S F
4141 - TTGAAGAAAGTGCCAGTTGATGAGTATATAACCACGTACCTGGACAAGGATGTGCTGGT - 4200
- L K K V P V D E Y I T T Y P G Q G C A G
- * R K C Q L M S I * P R T L D K D V L V
- E E S A S * * V Y N H V P W T R M C W L

```

图 11 续

```

4201 - TATACACTTGAGGAAGCTAAGACTGCTCTTAAGAAATGCAAATCTGCATTTTATGTACTA - 4260
- Y T L E E A K T A L K K C K S A F Y V L
- I H L R K L R L L L R N A N L H F M Y Y
- Y T * G S * D C S * E M Q I C I L C T T
4261 - CCTTCAGAAGCACCTAATGCTAAGGAAGAGATTCTAGGAAGCTATCCTGGAATTTGAGA - 4320
- P S E A P N A K E E I L G T V S W N L R
- L Q K H L M L R K R F * E L Y P G I * E
- F R S T * C * G R D S R N C I L E F E R
4321 - GAAATGCTTGCTCATGCTGAAGAGACAAGAAAATTAATGCCTATATGCATGGATGTTAGA - 4380
- E M L A H A E E T R K L M P I C M D V R
- K C L L M L K R Q E N * C L Y A W M L E
- N A C S C * R D K K I N A Y M H G C * S
4381 - GCCATAATGGCAACCATCCAACGTAAGTATAAAGGAATTAATCAAGAGGGCATCGTT - 4440
- A I M A T I Q R K Y K G I K I Q E G I V
- P * W Q P S N V S I K E L K F K R A S L
- H N G N H P T * V * R N * N S R G H R *
4441 - GACTATGGTCCGATTCCTTTTATACTAGTAAAGAGCCTGTAGCTTCTATTATTACG - 4500
- D Y E V R F F F Y T S K E P V A S I I T
- T M V S D S S F I L V K S L * L L L L R
- L W C P I L L L Y * * R A C S F Y Y Y E
4501 - AAGCTGAAGCTCTAATGAGCCGCTGTGACAAATGCCAATGGTTATGTGACACATGGT - 4560
- K L N S L N E P L V T M P I G Y V T H G
- S * T L * M S R L S Q C Q L V M * H M V
- A E L S K * A A C H N A N W L C D T W F
4561 - TTTAATCTTGAAGAGGCTGCGCGCTGTATGCGTTCTCTTAAAGCTCCTGCCGTAGTGCA - 4620
- F N L E E A A R C M R S L K A P A V V S
- L I L K R L R A V C V L L K L L P * C Q
- * S * R G C A L Y A F S * S S C R S V S
4621 - GTATCATCACCAGATGCTGTTACTACATATAATGGATACCTCACTTCGTCATCAAAGACA - 4680
- V S S P D A V T T Y N G Y L T S S S K T
- Y H H Q M L L L H I M D T S L R H Q R H
- I I T R C C Y Y I * W I P H F V I K D I
4681 - TCTGAGGACACTTTGTAGAAACAGTTTCTTTGGCTGGCTTACAGAGATTGGTCTAT - 4740
- S E H F V E T V S L A G S Y R D W S Y
- L R S T L * K Q F L W L A L T E I G P I
- * G A L C R N S F F G W L L Q R L V L F
4741 - TCAGGACAGCGTACAGAGTTAGGTGTTGAATTTCTTAAGCGTGGTGACAAAATGTGTAC - 4800
- S G Q R T E L G V E F L K R G D K I V Y
- Q D S V Q S * V L N F L S V V T K L C T
- R T A Y R V R C * I S * A W * Q N C V P
4801 - CACACTGTGAGAGCCCGCTCGAGTTTCATCTTGACGGTGAGGTTCTTTCACCTGACAAA - 4860
- H T L E S P V E F H L D G E V L S L D K
- T L W R A P S S F I L T V R F F H L T N
- H S G E P R R V S S * R * G S F T * Q T
4861 - CTAAGAGTCTCTTATCCCTGCGGGAGGTTAAGACTATAAAGTGTTCACAAGTGTGGAC - 4920
- L K S L L S L R E V K T I K V F T T V D
- * R V S Y P C G R L R L * K C S Q L W T
- K E S L I P A G G * D Y K S V H N C G Q
4921 - AACACTAATCTCCACACAGCTTGTGGATATGTCTATGACATATGGACAGCAGTTTGGT - 4980
- N T N L H T Q L V D M S M T Y G Q Q F G
- T L I S T H S L W I C L * H M D S S L V
- H * S P H T A C G Y V Y D I W T A V W S
4981 - CCAACATACTTGGATGGTCTGATGTTACAAAATTAACCTCATGTAATCATGAGGGT - 5040
- P T Y L D G A D V T K I K P H V N H E G
- Q H T W M V L M L Q K L N L M * I M R V
- N I L G W C * C Y K N * T S C K S * G *

```

图 11 续

```

5041 - AAGACTTCTTTGTACTACCTAGTGATGACACACTACGTAGTGAAGCTTCGAGTACTAC - 5100
- K T F F V L P S D D T L R S E A F E Y Y
- R L S L Y Y L V M T H Y V V K L S S T T
- D F L C T T * * * H T T * * S F R V L P
5101 - CATACTCTTGATGAGAGTTTTCTTGGTAGGTACATGTCTGCTTTAAACCACACAAGAAA - 5160
- H T L D E S F L G R Y M S A L N H T K K
- I L L M R V F L V G T C L L * T T Q R N
- Y S * * E F S W * V H V C F K P H K E M
5161 - TGGAAATTCCTCAAGTTGGTGGTTAACTTCAATTAATGGGCTGATAACAATTGTTAT - 5220
- W K F P Q V G G L T S I K W A D N N C Y
- G N F L K L V V * L Q L N G L I T I V I
- E I S S S W W F N F N * M G * * Q L L F
5221 - TTGTCTAGTGTATTATTAGCACTTCAACAGCTTGAAGTCAATTCATGCACCAGCACTT - 5280
- L S S V L L A L Q Q L E V K F N A P A L
- C L V F Y * H F N S L K S N S M H Q H F
- V * C F I S T S T A * S Q I Q C T S T S
5281 - CAAGAGCTTATTATAGAGCCGTCGGTGATGCTGCTAACTTTTGTGCACTCATACTC - 5340
- Q E A Y Y R A R A G D A A N F C A L I L
- K R L I I E P V L V M L L T F V H S Y S
- R G L L * S P C W * C C * L L C T H T R
5341 - GCTTACGTAATAAACTGTGGCGAGCTTGGTGATGTCAGAGAACTATGACCCATCTT - 5400
- A Y S N K T V G E L G D V R E T M T H F
- L T V I K L L A S L V M S E K L * P I F
- L Q * * N C W R A W * C Q R N Y D P S S
5401 - CTACAGCATGCTAATTTGGAATCTGCAAAGCGAGTTCTTAATGTGGTGTGTAACATTGT - 5460
- L Q H A N L E S A K R V L N V V C K H C
- Y S M L I W N L Q S E F L M W C V N I V
- T A C * F G I C K A S S * C G V * T L W
5461 - GGTCAGAAACTACTACCTTAACGGGTGTAGAAGCTGTGATGTATATGGGTACTCTATCT - 5520
- G Q K T T T L T G V E A V M Y M G T L S
- V R K L L P * R V * K L * C I W V L Y L
- S E N Y Y L N G C R S C D V Y G Y S L
5521 - TATGATAATCTTAAGACAGGTGTTCCATTCATGTGTGTGGTCGTGATGCTACACAA - 5580
- Y D N L K T G V S I P C V C G R D A T Q
- M I I L R Q V F P F H V C V V V M L H N
- * * S * D R C F H S M C V W S * C Y T I
5581 - TATCTAGTACAACAAGAGTCTTCTTTGTTATGATGTCTGCACCACCTGCTGAGTATAAA - 5640
- Y L V Q Q E S S F V M M S A P P A E Y K
- I * Y N K S L L L L * C L H H L L S I N
- S S T T R V F F C Y D V C T T C * V * I
5641 - TTACAGGAAGGTACATTCTTATGTGCGAATGAGTACACTGGTAACTATCAGTGTGGTCAT - 5700
- L Q Q G T F L C A N E Y T G N Y Q C G H
- Y S K V H S Y V R M S T L V T I S V V I
- T A R Y I L M C E * V H W * L S V W S L
5701 - TACTGATATAACTGCTAAGGAGACCCCTCTATCGTATTGACGGAGCTCACCTTACAAG - 5760
- Y T H I T A K E T L Y R I D G A H L T K
- T L I * L L R R P S I V L T E L T L Q R
- H S Y N C * G D P L S Y * R S S P Y K D
5761 - ATGTCAGAGTACAAGGACCAGTACTGATGTTTTCTACAAGGAACATCTTACTACTACA - 5820
- M S E Y K G P V T D V F Y K E T S Y T T
- C Q S T K D Q * L M F S T R K H L T L Q
- V R V Q R T S D * C F L Q G N I L H Y N
5821 - ACCATCAAGCCTGTGTCGTATAAATCGATGGAGTTACTTACACAGAGATTGACCAAAA - 5880
- T I K P V S Y K L D G V T Y T E I E P K
- P S S L C R I N S M E L L T Q R L N Q N
- H Q A C V V * T R W S Y L H R D * T K I

```

图 11 续

```

5881 - TTGGATGGGTATTATAAAAAGGATAATGCTTACTATACAGAGCAGCCTATAGACCTTGTA - 5940
- L D G Y Y K K D N A Y Y T E Q P I D L V
- W M G I I K R I M L T I Q S S L * T L Y
- G W V L * K G * C L L Y R A A Y R P C T
5941 - CCAACTCAACCATTACCAAATGCGAGTTTTGATAATTCAAACTCACATGTTCTAACACA - 6000
- P T Q P L P N A S P D N F K L T C S N T
- Q L N H Y Q M R V L I I S N S H V L T Q
- N S T I T K C E F * * F Q T H M F * H K
6001 - AAATTTGCTGATGATTTAAATCAAATGACAGGCTTCACAAAGCCAGCTCAGGAGACTA - 6060
- K F A D D L N Q M T G F T K P A S R E L
- N L L M I * I K * Q A S Q S Q L H E S Y
- I C * * F K S N D R L H K A S F T R A I
6061 - TCTGTCACATTCTCCAGACTTGAATGGCGATGTAGTGGCTATTGACTATAGACACTAT - 6120
- S V T F F P D L N G D V V A I D Y R H Y
- L S H S S Q T * M A M * W L L T I D T I
- C H I L P R L E W R C S G Y * L * T L F
6121 - TCAGCGAGTTTCAAGAAAGGTGCTAAATTACTGCATAAGCCAATTGTTGGCACATTAAC - 6180
- S A S F K K G A K L L H K P I V W H I N
- Q R V S R K V L N Y C I S Q L F G T L T
- S E F Q E R C * I T A * A N C L A H * P
6181 - CAGGCTACAACCAAGACAACGTTCAAACCAACACTTGGTGTTTACGTGTCTTTGGAGT - 6240
- Q A T T K T T F K P N T W C L R C L W S
- R L Q P R Q R S N Q T L G V Y V V F G V
- G Y N Q D N V Q T K H L V F T L S L E Y
6241 - ACAAGCCAGTAGATACTTCAAATTCATTGAGTTCCTGGCAGTAGAAGACACACAAGGA - 6300
- T K P V D T S N S F E V L A V E D T Q G
- Q S Q * I L Q I H L K F W Q * K T H K E
- K A S R Y F K F I * S S G S R R H T R N
6301 - ATGGACAATCTTGCTTGTGAAAGTCAACAACCACCTCTGAAGAAGTAGTGGAAAATCCT - 6360
- M D N L A C E S Q Q P T S E E V V E N P
- W T I L L V K V N N P P L K K * W K I L
- G Q S C L * K S T H L * R S S G K S Y
6361 - ACCATACAGAAGGATCATAGAGTGTGACGTGAAAACCTACCGAAGTGTAGGCAATGTC - 6420
- T I Q K E V I E C D V K T T E V V G N Y
- P Y R R K S * S V T * K L P K L * A M S
- H T E G S H R V * R E N Y R S C R Q C H
6421 - ATRACTAACCATCAGATGAAGGTGTTAAAGTAAACACAGAGTTAGGTCATGAGGATCTT - 6480
- I L K P S D E G V K V T Q E L G H E D L
- Y L N H Q M K V L K * H K S * V M R I L
- T * T I R * R C * S N T R V R S * G S Y
6481 - ATGGCTGCTTATGTGGAAAACACAAGCATTACCATTAAGAAACCTAATGAGCTTCACTA - 6540
- M A A Y V E N T S I T I K K P N E L S L
- W L L M W K T Q A L P L R N L M S F H *
- G C L C G K H K H Y H * E T * * A F T S
6541 - GCCTTAGGTTAAAACAATGGCCACTCATGGTATTGCTGCAATTAATAGTGTTCCTTGG - 6600
- A L G L K T I A T H G I A A I N S V P W
- P * V * K Q L P L M V L L Q L I V F L G
- L R F K N N C H S W Y C C N * * C S L E
6601 - AGTAAAATTTGGCTTATGTCAAACCATTCTTAGGACAAGCAGCAATTACAACATCAAAAT - 6660
- S K I L A Y V K P F L G Q A A I T T S N
- V K F W L M S N H S * D K Q Q L Q H Q I
- * N F G L C Q T I L R T S S N Y N I K L
6661 - TGGCTAAGAGATTAGCACAACSTGTGTTAAACAATTATGCTTATGTGTTACATA - 6720
- C A K R L A Q R V F N N Y M P Y V F T L
- A L R D * H N V C L T I I C L M C L H Y
- R * E I S T T C V * Q L Y A L C V Y I I

```

图 11 续


```

6721 - TTGTTCCAATTGTGTACTTTTACTAAAAGTACCAATTCTAGAATTAGAGCTTCACTACCT - 6780
- L F Q L C T F T K S T N S R I R A S L P
- C S N C V L L L K V P I L E L E L H Y L
- V P I V Y F Y * K Y Q F * N * S F T T Y
6781 - ACAACTATTGCTAAAAATAGTGTAAAGAGTGTGCTAAATTATGTTGGATGCCGGCATT - 6840
- T T T A K N S V K S V A K L C L D A G I
- Q L L L K I V L R V L L N Y V W M P A L
- N Y C * K * C * E C C * I M F G C R H *
6841 - AATTATGTGAAGTCACCCAAATTTCTAAATTGTTTACCAATCGCTATGTGGCTATTGTTG - 6900
- N Y V K S P K F S K L F T I A M W L L L
- I M * S H P N F L N C S Q S L C G Y C C
- L C E V T Q I F * I V H N R Y V A I V V
6901 - TTAAGTATTTGCTTAGGTTCTCTAATCTGTGTAAGTCTGCTGCTTTTGGTGTACTCTTATCT - 6960
- L S I C L G S L I C V T A A F G V L L S
- * V F A * V L * S V * L L L L V Y S Y L
- K Y L L R F S N L C N C C F W C T L I *
6961 - AATTTGGTGCCTTCTTATTGTAATGGCGTTAGAGAATTGTATCTTAATTCGTCTAAC - 7020
- N F G A P S Y C N G V R E L Y L N S S N
- I L V L L L I V M A L E N C I L I R L T
- F W C S F L L * W R * R I V S * F V * R
7021 - GTTACTACTATGGATTCTGTGAAGGTTCTTTTCTGTCAGCATTGTTTAAAGTGGATTA - 7080
- V T T M D F C E G S F P C S I C L S G L
- L L L W I S V K V L F L A A F V * V D *
- Y Y Y G F L * R F F S L Q H L F K W I R
7081 - GACTCCCTTGATTCTTATCCAGCTCTTGAACCATTTCAGGTGACGATTTTCATCGTACAAG - 7140
- D S L D S Y P A L E T I Q V T I S S Y K
- T P L I L I Q L L K P F R * R F H R T S
- L P * F L S S S * N H S G D D F I V Q A
7141 - CTAGACTTGACAATTTTAGGCTGCGCCGCTGAGTGGGTTTGGCATATATGTTGTTTACA - 7200
- L D L T I L G L A A E W V L A Y M L F T
- * T * Q F * V W P L S G F W H I C C S Q
- R L D N F R S G R * V G F G I Y V V H K
7201 - AAATCTTTTATTATTAGGCTTTTTCAGCTATAATGCAGGTGTTCTTTGGCTATTTTGGCT - 7260
- K F Y L L G L S A I M Q V F P G Y F A
- N S F I Y * V F Q L * C R C S L A I L L
- I L L F I R S F S Y N A G V L W L F C *
7261 - AGTCATTTTCATCAGCAATTCTTGGCTCATGTGGTTTATCATTAGTATTGTACAAATGGCA - 7320
- S H F I S N S W L M W F I I S I V Q M A
- V I S S A I L G S C G L S L V L Y K W H
- S F H Q Q F L A H V V Y H * Y C T N G T
7321 - CCCGTTTCTGCAATGGTTAGGATGTACATCTTCTTTGCTTCTTCTACTACATATGGAG - 7380
- P V S A M V R M Y I F F A S F Y Y I W K
- P F L Q W L G C T S S L L L S T T Y G R
- R F C N G * D V H L L C F F L L H M E E
7381 - AGCTATGTTTCATATCATGGATGGTTGCACCTCTTCGACTTGCATGATGTGCTATAAGCGC - 7440
- S Y V H I M D G C T S S T C M M C Y K R
- A M F I S W M V A P L R L A * C A I S A
- L C S Y H G W L H L F D L H D V L + A Q
7441 - AATCGTGCCACACGCGTTGAGTGTACAACATTGTTAATGGCATGAAGAGATCTTTCTAT - 7500
- N R A T R V E C T T I V N G M K R S F Y
- I V P H A L S V Q L L L M A * R D L S M
- S C H T R * V Y N Y C * W H E E I F L C
7501 - GTCTATGCAAAATGGAGGCCGTTGCTTCTGCAAGACTCACAATGGAATTGTCTCAATTGT - 7560
- V Y A N G G R G F C K T H N W N C L N C
- S M Q M E A V A S A R L T I G I V S I V
- L C K W R P W L L Q D S Q L E L S Q L *

```

图 11 续

```

7561 - GACACATTTTGCCTGGTAGTACATTCATTAGTGTGAAGTTGCTCGTGATTTGTCCTC - 7620
- D T F C T G S T F I S D E V A R D L S L
- T H F A L V V H S L V M K L L V I C H S
- H I L H W * Y I H * * * S C S * F V T P
7621 - CAGTTTAAAAGACCAATCAACCCTACTGACCAGTCATCGTATATTGTTGATAGTGTGCT - 7680
- Q F K R P I N P T D Q S S Y I V D S V A
- S L K D Q S T L L T S H R I L L I V L L
- V * K T N Q P Y * P V I V Y C * * C C C
7681 - GTGAAAATGGCGGCTTACCTCTACTTTGACAAGGCTGGTCAAAAGACCTATGAGAGA - 7740
- V K N G A L H L Y F D K A G Q K T Y E R
- * K M A R F T S T L T R L V K R P M R D
- E K W R A S P L L * Q G W S K D L * E T
7741 - CATCCGCTCTCCCATTTTGTCAATTTAGACAATTTGAGAGTAAACAACACTAAAGGTTCA - 7800
- H P L S H F V N L D N L R A N N T K G S
- I R S P I L S I * T I * E L T T L K V H
- S A L P F C Q F R Q F E S * Q H * R F T
7801 - CTGCCATTAAATGTCATAGTTTTGATGGCAAGTCCAAATGCGACGAGTCTGCTTCTAAG - 7860
- L P I N V I V F D G K S K C D E S A S K
- C L L M S * F L M A S P N A T S L L S
- A Y * C H S F * W Q V Q M R R V C F * V
7861 - TCTGCTTCTGTACTACAGTCAGCTGATGTGCCAACCTATTCTGTTGCTTGACCAAGCT - 7920
- S A S V Y Y S Q L M C Q P I L L L D Q A
- L L L C T T V S * C A N L F C C L T K L
- C F C V L Q S A D V P T Y S V A * P S S
7921 - CTTGTATCAACGTTGGAGATAGTACTGAAGTTCCGTTAAGATGTTTGTGCTTATGTC - 7980
- L V S N V G D S T E V S V K M F D A Y V
- L Y Q T L E I V L K F P L R C L M L M S
- C I K R W R * Y * S F R * D V * C L C R
7981 - GACACCTTTTCAGCAACTTTTAGTGTTCCTATGAAAACTTAAGGCATTGTTGCTACA - 8040
- D T F S A T F S V P M E K L K A L V A T
- T P F Q Q L L V F L W K N L R H L L L Q
- H L F S N F * C S Y G K T * G T C C Y S
8041 - GCTCACAGCGAGTTAGCAAAGGGTGTAGCTTTAGATGGTGTCCCTTCTACATTCGTGTCA - 8100
- A H S E L A K G V A L D G V L S T F V S
- L T A S * Q R V * L * M V S F L H S C Q
- S Q R V S K G C S F R W C P F Y I R V S
8101 - GCTGCCGACAAGGTGTTGTTGATACCGATGTTGACACAAAGGATGTTATTGAATGTCTC - 8160
- A A R Q G V V D T D V D T K D V I E C L
- L P D K V L L I P M L T Q R M L L N V S
- C P T R C C * Y R C * H K G C Y * M S Q
8161 - AAACCTTCACATCACTCTGACTTAGAAGTGACAGGTGACAGTTGTAACAATTCATGCTC - 8220
- K L S H H S D L E V T G D S C N N F M L
- N F H I T L T * K * Q V T V V T I S C S
- T F T S L * L R S D R * Q L * Q F H A H
8221 - ACCATAATAAGGTTGAAAACATGACGCCAGAGATCTGGCGCATGTATTGACTGTAAT - 8280
- T Y N K V E N M T P R D L G A C I D C N
- P I I R L K T * R P E I L A H V L T V M
- L * * G * K H D A Q R S W R M Y * L * C
8281 - GCAAGGCATATCAATGCCAAGTAGCAAAAAGTCACAATGTTTCACTCATCTGGAATGTA - 8340
- A R H I N A Q V A K S H N V S L I W N V
- Q G I S M P K * Q K V T M F H S S G M *
- K A Y Q C P S S K K S Q C F T H L E C K
8341 - AAAGACTACATGTCTTATCTGACAGCTGCGTAAACAAATTCGTACTGCTGCCAAGAAG - 8400
- K D Y M S L S E Q L R K Q I R T A A K K
- K T T C L Y L N S C V N K F V L L P R R
- R L H V F I * T A A * T N S Y C C Q E E

```

图 11 续

```

8401 - AACAACTACCTTTTACACTAACTTGTGCTACAACCTAGACAGGTTGTCAATGTCATAACT - 8460
- N N I P F T L T C A T T R Q V V N V I T
- T T Y L L H * L V L Q L D R L S M S * L
- Q H T F Y T N L C Y N * T G C Q C H N Y
8461 - ACTAAATCTCACTCAAGGGTGGTAAGATTGTTAGTACTTGTTTAACTTATGCTAAG - 8520
- T K I S L K G G K I V S T C F K L M L K
- L K S H S R V V R L L V L V L N L C L R
- * N L T Q G W * D C * Y L F * T Y A * G
8521 - GCCACATATTGTCGTTCTTGTGCTGCATGGTTTGTATATCGTTATGCCAGTACATACA - 8580
- A T L L C V L A A L V C Y I V M P V H T
- P H Y C A F L L H W F V I S L C Q Y I H
- H I I V R S C C I G L L Y R Y A S T Y I
8581 - TTGTCAATCCATGATGGTTACACAAATGAAATCATTGGTTACAAAGCCATTCAGGATGGT - 8640
- L S I H D G Y T N E I I G Y K A I Q D G
- C Q S M M V T Q M K S L V T K P F R M V
- V N P * W L H K * N H W L Q S H S G W C
8641 - GTCACCTCGTGACATCATTTCTACTGATGATGTTTTCGCAAAATAACATGCTGGTTTTCGAC - 8700
- V T R D I I S T D D C F A N K H A G F D
- S L V T S F L L M I V L Q I N M L V L T
- H S * H H F Y * * L F C K * T C W F * R
8701 - GCATGGTTAGCCAGCGTGGTGGTTCATACAAAATGACAAAAGCTGCCCTGTAGTAGCT - 8760
- A W F S Q R G G S Y K N D K S C P V V A
- H G L A S V V V H T K M T K A A L * * L
- M V * P A W W F I Q K * Q K L P C S S C
8761 - GCTATCATTACAAGAGAGATTGGTTTCATAGTGCCTGGCTTACCGGGTACTGTGCTGAGA - 8820
- A I I T R E I G F I V P G L P G T V L R
- L S L Q E R L V S * C L A Y R V L C * E
- Y H Y K R D W F H S A W L T G Y C A E S
8821 - GCAATCAATGGTGACTTCTTGCATTTTCTACCTCGTGTTCCTAGTGTGTGGCAACATT - 8880
- A I N G D F L H F L P R V F S A V G N I
- Q S M V T S C I F Y L V F L V L L A T F
- N Q W * L L A F S T S C F * C C W Q H L
8881 - TGCTACACACCTTCAAACCTCATTGAGTATAGTATTTGCTACCTCTGCTTGCCTTCTT - 8940
- C Y T P S K L I E Y S D F A T S A C V L
- A T H L P N S L S I V I L L P L L A F L
- L H T F Q T H * V * * F C Y L C L R S C
8941 - GCTGCTGAGTGTACAATTTTAAGGATGCTATGGGCAAACCTGTGCCATATTGTTATGAC - 9000
- A A E C T I F K D A M G K P V P Y C Y D
- L L S V Q F L R M L W A N L C H I V M T
- C * V Y N F * G C Y G Q T C A I L L * H
9001 - ACTAATTGCTAGAGGGTCTATTCTTATAGTGAGCTTCGTCAGACACTCGTTATGIG - 9060
- T N L L E G S I S Y S E L R P D T R Y V
- L I C * R V L F L I V S F V Q T L V M C
- * F A R G F Y F L * * A S S R H S L C A
9061 - CTTATGGATGGTTCCATCATACAGTTTCTAACACTTACCTGGAGGGTCTGTTAGAGTA - 9120
- L M D G S I I Q F P N T Y L E G S V R V
- L W M V P S Y S F L T L T W R V L L E *
- Y G W F H H T V S * H L P G G F C * S S
9121 - GTAACAATTTTGTGCTGAGTACTGTAGACATGGTACATGCCAAAGGTCAGAAGTAGGT - 9180
- V T T F D A E Y C R H G T C E R S E V G
- * Q L L M L S T V D M V H A K G Q K * V
- N N F * C * V L * T W Y M R K V R S R Y
9181 - ATTTGCCATCTACCAGTGTAGATGGGTTCTTAATAATGACATTACAGAGCTCTATCA - 9240
- I C L S T S G R W V L N N E H Y R A L S
- F A Y L P V V D G F L I M S I T E L Y Q
- L P I Y Q W * M G S * * * A L Q S S I R

```

图 11 续

```

9241 - GGAGTTTCTGTGGTGGTGGATCCGATGAATCTCATAGCTAACATCTTTACTCCTCTGTG - 9300
- G V F C G V D A M N L I A N I F T P L V
- E F S V V L M R * I S * L T S L L L L C
- S F L W C * C D E S H S * H L Y S S C A
9301 - CAACCTGTGGGTGCTTAGATGTGTCTGCTTCAGTAGTGGCTGGTGGTATTATTGCCATA - 9360
- Q P V G A L D V S A S V V A G G I Y A I
- N L W V L * M C L L Q * W L V V L L P Y
- T C G C F R C V C F S S G W W Y Y C H I
9361 - TTGGTACTTGTGCTGCTACTACTTTATGAAATCAGACGTTGTTTTGGTGAGTACAAC - 9420
- L V T C A A Y Y F M K F R R V F G E Y N
- W * L V L P T T L * N S D V F L V S T T
- G D L C C L L L Y E I Q T C F W * V Q P
9421 - CATGTTGTGCTGCTAATGCACTTTTGTGTTTGGATGCTTTTCACTATACTCTGTCTGGTA - 9480
- H V V A A N A L L F L M S F T I L C L V
- M L L L L M H F C F * C L S L Y S V W Y
- C C C C * C T F V F D V F R Y T L S G T
9481 - CCAGCTTACAGCTTCTGCGGGAGTCTACTCAGTCTTTTACTGTACTTGACATTCAT - 9540
- P A Y S F L P G V Y S V F Y L Y L T F Y
- Q L T A F C R E S T Q S F T C T * H S I
- S L Q L S A G S L L S L L L V L D I L F
9541 - TTCACCAATGATGTTTCATTCTTGGCTCACCTTCAATGGTTGCCATGTTTCTCCTATT - 9600
- F T N D V S F L A H L Q W F A M F S P I
- S P : M M F H S W L T F N G L P C F L L L
- H Q * C F I L G S P S M V C H V F S Y C
9601 - GTGCCTTTTGGATAACAGCAATCTATGTATTCTGTATTCTCTGAAGCACTGCCATTGG - 9660
- V P F W I T A I Y V F C I S L K H C H W
- C L F G * Q Q S M Y S V F L * S T A I G
- A F L D N S N L C I L Y F S E A L P L V
9661 - TTCTTTAACTATCTTAGGAAAAGATCATGTTAATGGAGTTACATTTAGTACCTTC - 9720
- F F N N Y L R K R V M F N G V T F S T F
- S L T T I L G K E S C L M E L H L V P S
- L * Q L S * E K S H V * W S Y I * Y L R
9721 - GAGGAGCTGCTTGTGTACTTTTTGCTCAACAAGGAAATGTACCTAAAATTCGCTAGC - 9780
- E E A A L C T F L L N K E M Y L K L R S
- R R L L C V P F C S T R K C T * N C V A
- G G C F V Y L F A Q Q G N V P K I A * R
9781 - GAGACCTGTGCCACTTACACAGTATAACAGGTATCTTGTCTATATAACAAGTACAAG - 9840
- E T L L P L T Q Y N R Y L A L Y N K Y K
- R H C C H L H S I T G I L L Y I T S T S
- D T V A T Y T V * Q V S C S I * Q V Q V
9841 - TATTCAGTGGAGCCTTAGATACTACCAGCTATCGTGAAGCAGCTTGCTGCCACTTAGCA - 9900
- Y F S G A L D T T S Y R E A A C C H L A
- I S V E P * I L P A I V K Q L A A T * Q
- F Q W S L R Y Y Q L S * S S L L P L S K
9901 - AAGGCTCTAATGACTTTAGCAACTCAGGTGCTGATGTTCTCTACCAACCACCACAGACA - 9960
- K A L N D F S N S G A D V L Y Q P P Q T
- R L * M T L A T Q V L M F S T N H H R H
- G S K * L * Q L R C * C S L P T T T D I
9961 - TCAATCACTTCTGCTGTTCTGCAGAGTGGTTTTAGGAAAATGGCATTCCCGTCAGGCAA - 10020
- S I T S A V L Q S G F R K M A F P S G K
- Q S L L L F C R V V L G K W H S R Q A K
- N H F C C S A E W F * E N G I P V R Q S
10021 - GTTGAAGGGTGCATGGTACAAGTAACTGTGGAACAACAATCTTAATGGATTGTGGTTG - 10080
- V E G C M V Q V T C G T T T L N G L W L
- L K G A W Y K * P V E L Q L L M D C G W
- * R V H G T S N L W N Y N S * W I V V G

```

图 11 续

```

10081 - GATGACACAGTATACTGTCCAAGACATGTCATTTCACACAGCAGAAGACATGCTTAATCCT - 10140
- D D T V Y C P R H V I C T A E D M L N P
- M T Q Y T V Q D M S F A Q Q K T C L I L
- * H S I L S K T C H L H S R R H A * S *
10141 - AACTATGARGATCTGCTCATTGCGAAATCCAACCATAGCTTCTTGTTCAGGCTGGCAAT - 10200
- N Y E D L L I R K S N H S F L V Q A G N
- T M K I C S F A N P T I A F L F R L A M
- L * R S A H S Q I Q P * L S C S G W Q C
10201 - GTTCAACTTCGTTATTGGCCATTCTATGCAAAATGTCTGCTTAGGCTTAAAGTTGAT - 10260
- V Q L R V I G H S M Q N C L L R L K V D
- F N F V L L A I L C K I V C L G L K L I
- S T S C Y W P F Y A K L S A * A * S * Y
10261 - ACTTCTAACCTAAGACACCCAAGTATAAATTTGTCGGTATCCAACCTGGTCAAACATTT - 10320
- T S N P K T P K Y K F V R I Q P G Q T F
- L L T L R H P S I N L S V S N L V K H F
- F * P * D T Q V * I C P Y P T W S N I F
10321 - TCAGTTCTAGCATGCTACAATGGTTCACCATCTGGTGTATCAGTGTGCCATGAGACCT - 10380
- S V L A C Y N G S P S G V Y Q C A M R P
- Q F * H A T M V H H L V F I S V P * D L
- S S S M L Q W F T I W C L S V C H E T *
10381 - AATCATACCATTAAAGGTTCTTTCCITATGGATCATGTGGTAGTGTGGTTTTAACATT - 10440
- N H T I K G S F L N G S C G S V G F N I
- I I P L K V L S L M D H V V V L V L T L
- S Y H * R F F P * W I M W * C W F * H *
10441 - GATTATGATGCGTGTCTTTCTGCTATATGCATCATATGGAGCTTCCAACAGGAGTACAC - 10500
- D Y D C V S F C Y M H H M E L P T G V H
- I M I A C L S A I C I I W S F Q Q E Y T
- L * L R V F L L Y A S Y G A S N R S T R
10501 - GCTGGTACTGACTTAGAAGGTAATTTCTATGGTCCATTGTTGACAGACAACTGCACAG - 10560
- A G T D L E G K F Y G P F V D R Q T A Q
- L V L T * K V N S M V H L L T D K L H R
- W Y * L R R * I L W S I C * Q T N C T G
10561 - GCTGCAGGTACAGACACAACCATAACATTAATGTTTGGCATGGCTGTATGCTGCTGTT - 10620
- A A G T D T T I T L N V L A W L Y A A V
- L Q V Q T Q P * R * M F W H G C M L L L
- C R Y R H N H N I K C F G M A V C C C Y
10621 - ATCAATGGTATAGGTGGTTTCTTAATAGATTACCCTACTTTGAATGACTTTAACCTT - 10680
- I N G D R W F L N R F T T T L N D F N L
- S M V I G G F L I D S P L L * M T L T L
- Q W * * V V S * * I H H Y F E * L * P C
10681 - GTGGCAATGAAGTACAACATGAACTTTGACACAAGATCATGTTGACATATTGGGACCT - 10740
- V A M K Y N Y E P L T Q D H V D I L G P
- W Q * S T T M N L * H K I M L T Y W D L
- G N E V Q L * T F D T R S C * H I G T S
10741 - CTTTCTGCTCAAACAGGAATTGCCGCTTAGATATGTGTGCTGCTTTGAAAGAGCTGCTG - 10800
- L S A Q T G I A V L D M C A A L K E L L
- F L L K Q E L P S * I C V L L * K S C C
- F C S N R N C R L R Y V C C F E R A A A
10801 - CAGATGGTATGAATGGTCGTAATCTATCCTGGTAGCACTATTTAGAAGATGAGTTTACA - 10860
- Q N G M N G R T I L G S T I L E D E F T
- R M V * M V V L S L V A L F * K M S L H
- E W Y E W S Y Y P W * H Y F R R * V Y T
10861 - CCATTGATGTTGTAGACAAATGCTCTGTTTACCTTCCAAGGTAAGTCAAGAAAT - 10920
- P F D V V R Q C S G V T F Q G K F K K I
- H L M L L D N A L V L P S K V S S R K L
- I * C C * T M L W C Y L P R * V Q E N C

```

图 11 续

10921 - GTTAAGGGCACTCATCATTGGATGCTTTAACTTTCTTGACATCACTATTGATTCTTGTT - 10980
 - V K S T H H W M L L T F L T S L L I L V
 - L R A L I I G C F * L S * H H Y * F L F
 - * G H S S L D A F N F L D I T I D S C S

10981 - CAAAGTACACAGTGGTCACTGTTTTCTTTGTTACGAGAATGCTTTCTTGCCATTTACT - 11040
 - Q S T Q W S L F F F V Y E N A F L P F T
 - K V H S G H C F S L F T R M L S C H L L
 - K Y T V V T V F L C L R E C F L A I Y S

11041 - CTTGGTATTATGGCAATTGCTGCATGCTGCTATGCTGCTGTTAAGCATAAGCAGCATT - 11100
 - L G I M A I A A C A M L L V K H K H A F
 - L V L W Q L L H V L C C L L S I S T H S
 - W Y Y G N C C M C Y A A C * A * A R I L

11101 - TTGTGCTTGTCTGTTACCTTCTCTTGCAACAGTTGCTTACTTTAATATGGTCTACATG - 11160
 - L C L F L L P S L A T V A Y F N M V Y M
 - C A C F C Y L L L Q Q L L T L I W S T C
 - V L V S V T F S C N S C L L * Y G L H A

11161 - CCTGCTAGCTGGGTGATGCGTATCATGACATGGCTTGAATTGGCTGACACTAGCTTGTCT - 11220
 - P A S W V M R I M T W L E L A D T S L S
 - L L A G * C V S * H G L N W L T L A C L
 - C * L G D A Y H D M A * I G * H * L V W

11221 - GGTATAGGCTTAAGGATTGTGTATGATGCTTCAGCTTTAGTTTTGCTTATTCTCATG - 11280
 - G Y R L K D C V M Y A S A L V L L I L M
 - V I G L R I V L C M L Q L * F C L F S *
 - L * A * G L C Y V C F S F S F A Y S H D

11281 - ACAGCTCGCACTGTTTATGATGATGCTGCTAGACGTTTGGACACTGATGAATGTCATT - 11340
 - T A R T V Y D D A A R R V W T L M N V I
 - Q L A L F M M M L L D V F G H * * M S L
 - S S H C L * * C C * T C L D T D E C H Y

11341 - ACACTGTTTACAAAGTCTACTATGGTAATGCTTTAGATCAAGCTATTTCCATGTGGGCC - 11400
 - T L V Y K V Y Y G N A L D Q A I S M W A
 - H L F T K S T M V M L * I K L F P C G P
 - T C L Q S L L W * C F R S S Y F H V G L

11401 - TTAGTTATTCTGTAACCTCTAACTATTCTGGTGTGCTTACGACTATCATGTTTTAGCT - 11460
 - L V I S V T S N Y S G V V T T I M F L A
 - * L F L * P L T I L V S L R L S C F * L
 - S Y F C N L * L F W C R Y D Y H V F S *

11461 - AGAGCTATAGTGTGTGTGTTGAGTATTACCCATTGTTATTTACTGGCAACACC - 11520
 - R A I V F V C V E Y Y P L L F I T G N T
 - E L * C L C V L S I T H C Y L L L A T P
 - S Y S V C V C * V L P I V I Y Y W Q H L

11521 - TTACAGTGTATCATGCTTGTATTGTTTCTTAGGCTATTGTGCTGCTGCTACTTTGGC - 11580
 - L Q C I M L V Y C F L G Y C C C C Y F G
 - Y S V S C L F I V S * A I V A A A T L A
 - T V Y H A C L L F L R L L L L L L L W P

11581 - CTTTCTGTTTACTCAACCGTTACTTCAGGCTTACTCTTGGTGTGTTAAGTACTTGGTC - 11640
 - L F C L L N R Y F R L T L G V Y D Y L V
 - F S V Y S T V T S G L L L V F M T T W S
 - F L F T O P L L Q A Y S W C L * L L G L

11641 - TCTACAAGAATTTAGGTATATGAACCTCCAGGGCTTTGCTCCTAAGAGTAGTATT - 11700
 - S T Q E F R Y M N S Q G L L P P K S S I
 - L H K N L G I * T P R G F C L L R V V L
 - Y T R I * V Y E L P G A F A S * E * Y *

11701 - GATGCTTCAAGCTTAACATTAAGTTGTTGGGTATTGAGGTAACCATGTATCAAGGTT - 11760
 - D A F K L N I K L L G I G G K P C I K V
 - M L S S L T L S C W V L E V N H V S R L
 - C F Q A * H * V V G Y W R * T M Y Q G C

图 11 续

```

11761 - GCTACTGTACAGTCTAAAATGTCTGACGTAAAGTGCACATCTGTGGTACTGCTCTCGGTT - 11820
- A T V Q S K M S D V K C T S V V L L S V
- L L Y S L K C L T * S A H L W Y C S R F
- Y C T V * N V * R K V H I C G T A L G S
11821 - CTTCAACAACCTAGAGTAGAGTCATCTTCTAAATTGTGGGCACAATGTGTACAACCTCCAC - 11880
- L Q Q L R V E S S B K L W A Q C V Q L H
- F N N L E * S H L L N C G H N V Y N S T
- S T T * S R V I F * I V G T M C T T P Q
11881 - AATGATATTCTTCTTGCAAAAGACACAACCTGAAGCTTTCGAGAAGATGGTTTCTCTTTTG - 11940
- N D I L L A K D T T E A F E K M V S L L
- M I F F L Q K T Q L K L S R R W F L F C
- * Y S S C K R H N * S F R E D G F S F V
11941 - TCTGTTTGGCTATCCATGCAGGGTCTGTAGACATTAATAGGTTGTGCGAGGAAATGCTC - 12000
- S V L L S M Q G A V D I N R L C E E M L
- L F C Y P C R V L * T L I G C A R K C S
- C F A I H A G C C R H * * V V R G N A R
12001 - GATAACCGTGTACTCTTCAGGCTATTGCTTCAGAATTTAGTTCTTTACCATCATATGCC - 12060
- D N R A T L Q A I A S E F S S L P S Y A
- I T V L L F R L L L O N L V L Y H H M P
- * P C Y S S G Y C F R I * F F T I I C R
12061 - GCTTATGCCACTGCCAGGAGGCCTATGAGCAGGCTGTAGCTAATGGTATTCTGAAGTC - 12120
- A Y A T A Q E A Y E Q A V A N G D S E V
- L M P L P R R P M S R L * L M V I L K S
- L C H C P G G L * A G C S * W * F * S R
12121 - GTTCTCAAAAAGTTAAAGAAATCTTTGAATGTGGCTRAATCTGAGTTTGACCGTGATGCT - 12180
- V L K K L K K S L N V A K S E F D R D A
- F S K S * R N L * M W L N L S L T V M L
- S Q K V K E I F E C G * I * V * P * C C
12181 - GCCATGCAACGCAAGTTGAAAAGATGGCAGATCAGGCTATGACCCAAATGTACAAACAG - 12240
- A M Q R K L E K M A D Q A M T Q M Y K Q
- P C N A S W K R W Q I R L * P K C T N R
- H A T Q V G K D G R S G Y D P N V Q T G
12241 - GCAAGATCTGAGGACAAGAGGGCAAAAGTAACTAGTGTATGCAACAATGCTCTTCACT - 12300
- A R S E D K R A K V T S A M Q T M L F T
- Q D L R T R G Q K * L V L C K Q C S S L
- K I * G Q E G K S N * C Y A N N A L H Y
12301 - ATGCTTAGGAAGCTTGATAATGATGCACTTAACRACATTATGAACAATGGCGTGATGGT - 12360
- M L R K L D N D A L N N I I N N A R D G
- C L G S L I M M H L T T L S T M R V M V
- A * E A * * * C T * Q H Y Q Q C A * W L
12361 - TGTGTTCCACTCAACATCATACCATTGACTACAGCAGCCAACTCATGGTTGTTGTCOOT - 12420
- C V F L N I I P L T T A A K L M V V V P
- V F H S T S Y H * L Q Q P N S W L L S L
- C S T Q H H T I D Y S S Q T H G C C P *
12421 - GATTATGGTACCTACAAGAACACTTGTGATGGTAACACCTTTACATATGCATCTGCACTC - 12480
- D Y G T Y K N T C D G N T F T Y A S A L
- I M V P T R T L V M V T P L H M H L S
- L W Y L Q E H L * W * H L Y I C I C T L
12481 - TGGGAAATCCAGCAAGTTGTTGATGCGGATAGCAAGATTGTTCAACTTAGTGAATTAAC - 12540
- W E I Q Q V V D A D S K I V Q L S E I N
- G K S S K L L M R I A R L F N L V K L T
- G N P A S C * C G * Q D C S T * * N * H
12541 - ATGGACAATTACCAAAATTTGGCTTGGCCTCTTATTGTTACAGCTCTAAGAGCCAACTCA - 12600
- M D N S P N L A W P L I V T A L R A N S
- W T I H Q I W L G L L L L Q L * E P T Q
- G Q F T K F G L A S Y C Y S S K S Q L S

```

图 11 续

```

12601 - GCTGTTAAACTACAGAATAATGAACTGAGTCCAGTAGCACTACGACAGATGTCCTGTGCG - 12660
- A V K L Q N N E L S P V A D R Q M S C A
- L L N Y R I M N * V Q * H Y D R C P V R
- C * T T E * * T E S S S T T T D V L C G
12661 - GCTGGTACCACACAAACAGCTTGTACTGATGACAATGCACTTGCCTACTATAACAATTG - 12720
- A G T T Q T A C T D D N A L A Y Y N N S
- L V P H K Q L V L M T M H L P T I T I R
- W Y H T N S L Y * * Q C T C L L * Q F E
12721 - AAGGGAGGTAGGTTGTGCTGGCATTACTATCAGACCACCAAGATCTCAAATGGGCTAGA - 12780
- K G G R F V L A L L S D H Q D L K W A R
- R E V G L C W H Y Y Q T T K I S N G L D
- G R * V C A G I T I R P P R S Q M G * I
12781 - TTCCTAAGAGTGTGGTACAGGTACAATTTACACAGAACGGAAACCACCTTGTAGGTTT - 12840
- F P K S D G T G T I Y T E L E P P C R F
- S L R V M V Q V Q F T Q N W N H L V G L
- P * E * W Y R Y N L H R T G T T L * V C
12841 - GTTACAGACACACAAAAGGGCCTAAAGTGAATACTTGTACTTCATCAAAGGCTTAAAC - 12900
- V T D T P K G P K V K Y L Y F I K G L N
- L Q T H Q K G L K * N T C T S S K A * T
- Y R H T K R A * S E I L V L H Q R L K Q
12901 - AACCTAATAGAGTATGGTGTGGGAGTTTGTGCTACAGTACGTCTTCAGGCTGGA - 12960
- N L N R G M V L G S L A A T V R L Q A G
- T * I E V W C W A V * L L Q Y V F R L E
- P K * R Y G A G Q F S C Y S T S S G W K
12961 - AATGCTACAGAAGTACCTGCCAATTCACACTGTGCTTCCCTCTGTGCTTTGCAGTAGAC - 13020
- N A T E V P A N S T V L S F C A F A V D
- M L Q K Y L P I Q L C F P S V L L Q * T
- C Y R S T C Q F N C A F L L C F C S R P
13021 - CCTGCTAAAGCATATAAGGATTACCTAGCAAGTGGAGGACAACCAATCACCAACTGTGTG - 13080
- P A K A Y K D Y L A S G G Q P I T N C V
- L L K H I R I T * Q V E D N Q S P T V *
- C * S I * G L P S K W R T T N H Q L C E
13081 - AAGATGTGTGTACACACACTGGTACAGGACAGGCAATTACTGTAACACCAGAAGCTAAC - 13140
- K M L C T H T G T G Q A I T V T P E A N
- R C C V H T L V Q D R Q L L * H Q K L T
- D V V Y T H W Y R T G N Y C N T R S * H
13141 - ATGGACCAAGAGTCCCTTGGTGGTGTTCATGTTGTCTGTATTGTAGATGCCACATTGAC - 13200
- M D Q E S F G G A S C C L Y C R C H I D
- W T K S P L V V L H V V C I V D A T L T
- G P R V L W W C F M L S V L * M P H * P
13201 - CATCCAAATCCTAAGGATTCTGTGACTGAAAGGTAAGTAGGTCCAAATACCTACCACT - 13260
- H P N P K G F C D L K G K Y V Q I P T T
- I Q I L K D S V T * K V S T S K Y L P L
- S K S * R I L * L E R * V R P N T Y H L
13261 - TGTGCTAATGACCCAGTGGGTTTACACTTAGAAACACAGTCTGTACCGTCTGCGGAATG - 13320
- C A N D P V G F T L R N T V C T V C G M
- V L M T Q W V L H L E T Q S V P S A E C
- C * * P S G F Y T * K H S L Y R L R N V
13321 - TGGAAAGGTTATGGCTGTAGTTGTGACCACTCCGCGAACCTTGATGCAGTCTGCGGAT - 13380
- W K G Y G C S C D Q L R E P L M Q S A D
- G K V M A V V V T N S A N P * C S L R M
- E R L W L * L * P T P R T L D A V C G C
13381 - GCATCAACGTTTTTAAACGGGTTGCGGTGAAGTGCAGCCCGTCTTACACCGTGGCGCA - 13440
- A S T F L N G F A V * V Q P V L H R A A
- H Q R F * T G L R C K C S P S Y T V R H
- I N V F K R V C G V S A A R L T P C G T

```

图 11 续


```

13441 - CAGGCAC TAGACTGATGTCGCTACAGGGCTTTTGATATTTACAACGAAAAAGTGCTG - 13500
- Q A L V L M S S T G L L I F T T K K V L
- R H * Y * C R L O G F * Y L Q R K K C W
- G T S T D V V Y R A F D I Y N E K S A G
13501 - GTTTTGCAAAGTTCCTAAAACTAATTGCTGTCGCTCCAGGAGAAGGATGAGGAAGGCA - 13560
- V L Q S S * K L I A V A S R R R M R K A
- F C K V P K N * L L S L P G E G * G R Q
- F A K F L K T N C A C R F Q E K D E E G N
13561 - ATTTATFAGACTCTTACTTTGTAAGGAGGCATACTATGTCTAACTACCAACATGAAG - 13620
- I Y * T L T L * L R G I L C L T T N M K
- F I R L L L C S * E A Y Y V * L P T * R
- L L D S Y F V V K R H T M S N Y Q H E E
13621 - AGACTATTTATAACTTGGTTAAAGATTGCCAGCGTGTGCTGTCCATGACTTTTTCAAGT - 13680
- R L F I T W L K I V Q R L L S M T F S S
- D Y L * L G * R L S S G C C P * L F Q V
- T I Y N L V K D C P A V A V H D F F K F
13681 - TTAGATYAGATGGTGACATGGTACCACATATATCACGTCAGCGTCTAACTAATACACAA - 13740
- L E * M V T W Y H I Y H V S V * L N T Q
- * S R W * H G T T Y I T S A S N * I H N
- R V D G D M V P H I S R Q R L T K Y T M
13741 - TGGCTGATTTAGTCTATGCTCTACGTCATTTTGATGAGGGTAATGTGATACATAAAG - 13800
- W L I * S M L Y V I L M R V I V I H * K
- G * F S L C S T S F * * G * L * Y I K R
- A D L V Y A L R H F F D E G N C D T L K E
13801 - AAATACTCGTCACATACAATTGCTGTGATGATGATTATTTCAATAAGAAGGATTGGTATG - 13860
- K Y S S H T I A V M M I I S I R R I G M
- N T R H I Q L L * * * L F Q * E G L V *
- I L V T Y N C C D D D Y F N K K D W Y D
13861 - ACTTCGTAGAGAATCCTGACATCTACCGGTATATGCTAACTTAGGTGAGCGGTACGCC - 13920
- T S * R I L T S Y A Y M L T * V S V Y A
- L R R E S * H L T R I C * L R * A C T P
- F V E N P D I L R V Y A N L G E R V R Q
13921 - AATCATTATTAAGACTGTACAATTCTGCGATGCTATGCGTGATGCAGGCATGTAGGGG - 13980
- N H Y * R L Y N S A M L C V M Q A L * A
- I I I K D C T I L R C Y A * C R H C R R
- S L L K T V Q F C D A M R D A G I V G V
13981 - TACTGACATTAGATAATCAGGATCTTAATGGGAAGTACGATTTCGGTGATTTCGTAC - 14040
- Y * H * I I R I L M G T G T I S V I S Y
- T D I R * S G S * W E L V R F R * F R T
- L T L D N Q D L N G N W Y D F G D F V Q
14041 - AAGTAGCACCAGGCTGCGGAGTTCCTATTGTGGATTCAATTACTCATTGCTGATGCCCA - 14100
- K * H Q A A E F L L W I H I T H C * C P
- S S T R L R S S Y C G F I L L I A D A H
- V A P G C G V P I V D S Y Y S L L M P I
14101 - TCCTCACTTTGACTAGGGCATTGGCTGCTGAGTCCCATATGGATGCTGATCTCGAAAAC - 14160
- S S L * L G H W L L S P I W M L I S Q N
- P H F D * G I G C * V P Y G C * S R K T
- L T L T R A L A E S H M D A D L A K P
14161 - CACTTATTAAGTGGGATTGCTGAATATGATTTTACGGAAGAGACTTTGTCTCTCG - 14220
- H L L S G I C * N M I L R K R D F V S S
- T Y * V G F A E I * F Y G R E T L S L R
- L I K W D L L K Y D F T E E R L C L F D
14221 - ACCGTTATTTAAATATTGGGACCAGACATACCATCCCAATTGTATTAAGTGTGGATG - 14280
- T V I L N I G T R H T I P I V L T V W M
- P L F * I L G P D I P S Q L Y * L F G *
- R Y F K Y W D Q T Y H P N C I N C L D D

```

图 11 续

```

14281 - ATAGGTGTATCCTTCATTGTGCAAACTTTAATGTGTTATTTTCTACTGTGTTTCCACCTA - 14340
- I G V S F I V Q T L M C Y F L L C F H L
- * V Y P S L C K L * C V I F Y C V S T Y
- R C I L H C A N F N V L F S T V F P P T
14341 - CAAGTTTTGGACCRCTAGTAAGAAAAATATTGTAGATGGTGTTCCTTTTGTGTTTCAA - 14400
- Q V L D H * * E K Y L * M V F L L L F Q
- K F W T T S K K N I C R W C S F C C F N
- S F G P L V R K I F V D G V P F V V S T
14401 - CTGGATACCATTTCGTGAGTTAGGAGTCGTACATAATCAGGATGTAAACTTACATAGCT - 14460
- L D T I F V S * E S Y I I R M * T Y I A
- W I P F S * V R S R T * S G C K L T * L
- G Y H F R E L G V V H N Q D V N L H S S
14461 - CGCGTCTCAGTTCAAGGAAGCTTTTAGTGTATGCTGCTGATCCAGCTATGCATGCAGCTT - 14520
- R V S V S R N F * C M L L I Q L C M Q L
- A S Q F Q G T F S V C C * S S Y A C S F
- R L S F K E L L V Y A A D P A M H A A S
14521 - CTGGCAATTTATGCTAGATAAACGCCACTACATGCTTTTTCAGTAGCTGCACTAACAAACA - 14580
- L A I Y C * I N A L H A F Q * L H * Q T
- W Q F I A R * T H Y M L F S S C T N K Q
- G N L L L D K R T T C F S V A A L T N N
14581 - ATGTTGTTTTTCAAAGCTGTCAAACCCGTAATTTAATAAAGACTTTTATGACTTTGCTG - 14640
- M L L F K L S N P V I L I K T F M T L L
- C C F S N C Q T R * F * * R L L * L C C
- V A F Q T V K P G N F N K D F Y D F A V
14641 - TGTCTAAAGGTTTCTTTAAGGAAGGAAGTCTGTTGAAGTAAACACTTCTTCTTTGCTC - 14700
- C L K V S L R K E V L L N * N T S S L L
- V * R F L * G R K F C * T K T L L L C S
- S K G F F K E G S S V E L K H F F F A Q
14701 - AGGATGGCAACGCTGCTATCAGTGATTATGACTATTATCGTTATAATCTGCCAACCAATGT - 14760
- R M A T L L S V I M T I I V I I C Q Q C
- G W Q R C Y Q * L * L L S L * S A N N V
- D G N A A I S D Y D Y R Y N L P T M C
14761 - GTGATATCAGACAAGTCTTATTCGTAGTTGAAGTTGTTGATAAATACTTTGATTGTTACG - 14820
- V I S D N S Y S * L K L L I N T L I V T
- * Y Q T T P I R S * S C * * I L * L L R
- D I R Q L L F V V E V V D K Y F D C Y D
14821 - ATGGTGGCTGTATTAATGCCAACCAAGTAATCGTTAACAATCTGGATAAATCAGCTGGTT - 14880
- M V A V L M P T K * S L T I W I N Q L V
- W W L Y * C Q P S N R * Q S G * I S W F
- G G C I N A N Q V I V N N L D K S A G F
14881 - TCCCATTTAATAAATGGGGTAAGGCTAGACTTTTATATGACTCAATGAGTTATGAGGATC - 14940
- S H L I N G V R L D F I M T Q * V M R I
- P I * * M G * G * T L L * L N E L * G S
- P F N K W G K A R L Y Y D S M S Y E D Q
14941 - AAGATGGACTTTTCGGTATACTAAGCGTAATGTCATCCCTACTATAACTCAAATGAATC - 15000
- K M H F S R I L S V M S S L L * L K * I
- R C T F R V Y * A * C H P Y Y N S N E S
- D A L F A Y T K R N V I P T I T Q M N L
15001 - TTAAGTATGCCATTAGTGC AAGAATAGAGCTCGCACCGTAGCTGGTGTCTCTATCTGTA - 15060
- L S M P L V Q R I E L A P * L V S L S V
- * V C H * C K E * S S H R S W C L Y L *
- K Y A I S A K N R A R T V A G V S I C S
15061 - GTACTATGACAAATAGACAGTTTCATCAGAATTATTGAAGTCAATAGCCGCCACTAGAG - 15120
- V L * Q I D S F I R N Y * S Q * P P L E
- Y Y D K * T V S S E I I E V N S R H * R
- T M T N R Q F H Q K L L K S I A A T R G

```

图 11 续

```

15121 - GAGCTACTGTGGTAATTGGAACAAGCAAGTTTTACGGTGGCTGGCATAATATGTTAAAAA - 15180
- E L L W * L E Q A S F T V A G I I C * K
- S Y C G N W N K Q V L R W L A * Y V K N
- A T V V I G T S K F Y G G W H N M L K T
15181 - CTGTTTACAGTGATGTAGAACTCCACACCTTATGGGTGGGATTATCCAAAATGTGACA - 15240
- L F T V M * K L R T L W V G I Y Q N V T
- C L Q * C R N S T P Y G L G L S K M * Q
- V Y S D V E T P H L M G W D Y P K C D R
15241 - GAGCCATGCCTAACATGCTTAGGATAATGGCCTCTCTGTCTTGCTCGCAACATAACA - 15300
- E P C L T C L G * W P L L F L L A N I T
- S H A * H A * D N G L S C S C S Q T * H
- A M P N M L R I M A S L V L A R K H N T
15301 - CTGCTGTAACTTATCACACCGTTTCTACAGTTAGCTAACGAGTGTGCGCAAGTATTAA - 15360
- L A V T Y H T V S T G * L T S V R K Y *
- L L * L I T P F L Q V S * R V C A S I K
- C C N L S H R F Y R L A N E C A Q V L S
15361 - GTGAGATGGTCATGTGTGGCGGCTCACTATATGTTAAACCAGGTGAACATCATCCGGTG - 15420
- V R W S C V A A H Y M L N Q V E H H P V
- * D G H V W R L T I C * T R W N I I R *
- E M V M C G G S L Y V K P G G T S S G D
15421 - ATGCTACAACCTGCTTATGCTAATAGTGTCTTTAACATTTGTCAAGCTGTTACAGCCAATG - 15480
- M L Q L L M L I V S L T F V K L L Q P M
- C Y N C L C * * C L * H L S S C Y S Q C
- A T T A Y A N S V F N I C Q A V T A N V
15481 - TAAATGCACTTCTTTCAACTGATGGTAATAAGATAGCTGACAAGTATGTCCGCAATCTAC - 15540
- * M H F F Q L M V I R * L T S M S A I Y
- K C T S F N * W * * D S * Q V C P Q S T
- N A L L S T D G N K I A D K Y V R N L Q
15541 - AACACAGGCTCTATGAGTGTCTCTATAGAAATAGGGATGTTGATCATGAATTCGTGGATG - 15600
- N T G S M S V S I E I G M L I M N S W M
- T Q A L * V S L * K * G C * S * I R G *
- H R L Y E C L Y R N R D V D H E F V D E
15601 - AGTTTTACGCTTACCTGCGTAAACATTTCTCCATGATGATCTTTCTGATGATGCCGTTG - 15660
- S F T L T C V N I S P * * F F L M M P L
- V L R L P A * T F L H D D S F * * C R C
- F Y A Y L R K H F S M M I L S D D A V V
15661 - TGTGCTATAACAGTAACTATGCGGCTCAAGTTTAGTAGCTAGCATTAAAGAACTTTAAGG - 15720
- C A I T V T M R L K V * * L A L R T L R
- V L * Q * L C G S R F S S * H * E L * G
- C Y N S N Y A A Q G L V A S I K N F K A
15721 - CAGTTCTTTATTATCAAATAATGTGTTTCATGTCTGAGGCAAAATGTTGGACTGAGACTG - 15780
- Q F F I I K I M C S C L R Q N V G L R L
- S S L L S K * C V H V * G K M L D * D *
- V L Y Y Q N N V F M S E A K C W T E T D
15781 - ACCTTACTAAAGGACCTCAGGAATTTGCTCACAGCATAAATGCTAGTTAAACAAGGAG - 15840
- T L L K D L T N F A H S I Q C * L N K E
- P Y * R T S R I L L T A Y N A S * T R R
- L T K G P H E F C S Q H T M L V K Q G D
15841 - ATGATTACGTGTACCTGACCTTACCCAGATCCATCAAGAATATTAGGCGCAGGCTGTTTTG - 15900
- M I T C T C L T Q I H Q E Y * A Q A V L
- * L R V P A L P R S I K N I R R R L F C
- D Y V Y L P Y P D P S R I L G A G C F V
15901 - TCGATGATATTGTCAAACAGATGGTACACTTATGATTGAAGGTTGCTGCTACTGGCTA - 15960
- S M I L S K Q M V H L * L K G S C H W L
- R * Y C Q N R W Y T Y D * K V R V T G Y
- D D I V K T D G T L M I E R F V S L A I

```

图 11 续

```

15961 - TTGATGCTTACCCACTTACAAAACATCCTAATCAGGAGTATGCTGATGTCTTTCACTTGT - 16020
- L M L T H L O N I L I R S M L M S F T C
- * C L P T Y K T S * S G V C * C L S L V
- D A Y P L T K H P N Q E Y A D V F H L Y
16021 - ATTTACAATACATTAGAAAAGTTACATGATGAGCTTACTGGCCACATGTGGACATGTATT - 16080
- I Y N T L E S Y M M S L L A T C W T C I
- F T I H * K V T * * A Y W P H V G H V F
- L Q Y I R K L H D E L T G H M L D M Y S
16081 - CCGTAAAGCTAACTAATGATAACACCTCACGGTACTGGGAACCTGAGTTTATGAGGCTA - 16140
- P * C * L M I T P H G T G N L S F M R L
- R N A N * * * H L T V L G T * V L * G Y
- V M L T N D N T S R Y W E P E F Y E A M
16141 - TGTACACACCACATACAGTCTTGCAGGCTGTAGGTGCTTGTGTATTGTGCAATTCACAGA - 16200
- C T H H I Q S C R L * V L V Y C A I H R
- V H T T Y S L A G C R C L C I V Q F T D
- Y T P H T V L Q A V G A C V L C N S Q T
16201 - CTTCACTTCGTTGCGGTGCCTGTATTAGGAGACCATTCCTATGTTGCAAGTGCTGCTATG - 16260
- L H F V A V P V L G D H S Y V A S A A M
- F T S L R C L Y * E T I P M L Q V L L *
- S L R C G A C I R R P F L C C K C C Y D
16261 - ACCATGTCATTTCAACATCACAAAATTAGTGTGTCTGTAAATCCCTATGTTTGAATG - 16320
- T M S F Q H H T N * C C L L I P M F A M
- P C H F N I T Q I S V V C * S L C L Q C
- H V I S T S H K L V L S V N P Y V C N A
16321 - CCCCAGGTTGTGATGTCAGTGTGACACAACCTGTATCTAGGAGGTATGAGCTATTATT - 16380
- P Q V V M S L M * H N C I * E V * A I I
- P R L * C H * C D T T V S R R Y E L L L
- P G C D V T D V T Q L Y L G G M S Y C
16381 - GCAAGTCACATAAGCCTCCATTAGTTTTCCATTATGTGCTAATGGTCAGGTTTTGGTT - 16440
- A S H I S L P L V F H Y V L M V R F L V
- Q V T * A S H * F S I M C * W S G F W F
- K S H K P P I S F P L C A N G Q V F G L
16441 - TATACA AAAACACATGTGTAGGCAGTGACAATGTCAGTACTTCAATGCGATAGCAACAT - 16500
- Y T K T H V * A V T M S L T S M R * Q H
- I Q K H M C R Q * Q C H * L Q C D S N M
- Y K N T C V G S D N V T D F N A I A T C
16501 - GTGATGGACTAATGCTGGGATTACATACTTGCCAACACTGTACTGAGAGACTCAAGC - 16560
- V I G L M L A I T Y L P T L V L R D S S
- * L D * C W R L H T C Q H L Y * E T Q A
- D W T N A G D Y I L A N T C T E R L K L
16561 - TTTTCGGAGCAGAAACGCTCAAAGCCACTGAGGAAACATTAAGCTGCATATGGTATTG - 16620
- F S Q Q K R S K P L R K H L S C H M V L
- F R S R N A Q S H * G N I * A V I W Y C
- F A A E T L K A T E E T F K L S Y G I A
16621 - CCACTGTACGCGAAGTACTCTCTGACAGAGAATTGCATCTTTCATGGGAGGTTGAAAAC - 16680
- P L Y A K Y S L T E N C I F H G R L E N
- H C T R S T L * Q R I A S F M G G W K T
- T V R E V L S D R E L H L S W E V G K P
16681 - CTAGACCACCATGAACAGAACTATGTCTTTACTGGTTACCGTGTAACATAAAATAGTA - 16740
- L D H H * T E T M S L L V T V * L K I V
- * T T I E Q K L C L Y W L P C N * K * *
- R P P L N R N Y V F T G Y R V T K N S K
16741 - AAGTACAGATTGGAGAGTACACCTTTGAAAAGGTGACTATGGTGATGCTGTTGTGTACA - 16800
- K Y R L E S T P L K K V T M V M L L C T
- S T D W R V H L * K R * L W * C C C V Q
- V Q I G E Y T F E K G D Y G D A V V Y R

```

图 11 续

```

16801 - GAGGTACTACGACATACAAGTTGAATGTTGGTGATTACTTTGTGTTGACATCTCACACTG - 16860
- E V L R H T S * M L V I T L C * H L T L
- R Y Y D I O V E C W * L L C V D I S H C
- G T T T Y K L N V G D Y F V L T S H T V
16861 - TAATGCCACTTAGTGACCTACTCTAGTGCCACAAGGACACTATGTGAGAATTACTGGCT - 16920
- * C H L V H L L * C H K S T M * E L L A
- N A T * C T Y S S A T R A L C E N Y W L
- M P L S A P T L V P Q E H Y V R I T G L
16921 - TGTACCCAACTCAACATCTCAGATGAGTTTTCTAGCAATGTTGCAAATTATCAAAGG - 16980
- C T Q H S T S Q M S F L A M L Q I I K R
- V P N T Q H L R * V F * Q C C K L S K G
- Y P T L N I S D E F S S N V A N Y Q K V
16981 - TCGGCATGCAAAGTACTCTACACTCCAAGGACCACCTGGTACTGGTAAGAGTCATTTG - 17040
- S A C K S T L H S K D H L V L V R V I L
- R H A K V L Y T P R T T W Y W * E S F C
- G M Q K Y S T L Q G P P G T G K S H F A
17041 - CCATCGGACTTGCTCTCTATTACCCATCTGCTCGCATAGTGATACGGCATGCTCTCATG - 17100
- P S D L L S I T H L L A * C I R H A L M
- H R T C S L L P I C S H S V Y G M L S C
- I G L A L Y Y P S A R I V Y T A C S H A
17101 - CAGCTGTTGATGCCCTATGTGAAAAGGCATTAATAATTTGCCCATAGATAAATGTAGTA - 17160
- Q L L M P Y V K R H * N I C P * I N V V
- S C * C P M * K G I K I F A H R * M * *
- A V D A L C E K A L K Y L P I D K C S R
17161 - GAATCATACCTGCGCGTGGCGCGTAGAGTGTTTTGATAAATTCAAAGTGAATTC AACAC - 17220
- E S Y L R V R A * S V L I N S K * I Q H
- N H T C A C A R R V F * * I O S E F N T
- I I P A R A R V E C F D K F K V N S T L
17221 - TAGAACAGTATGTTTTCTGCACTGTAATGCATTGCCAGAAACAAGTCTGCTGACTGTAG - 17280
- * N S M F S A L * M H C Q K Q L L T L *
- R T V C F L H C K C I A R N N C * H C S
- E Q Y V F C T V N A L P E T T A D I V V
17281 - TCTTTGATGAATCTCTATGGCTACTAATTAGACTTGAGTGTGTCAATGCTAGACTTC - 17340
- S L M K S L W L L I M T * V L S M L D F
- L * * N L Y G Y * L * L E C C Q C * T S
- F D E I S M A T N Y D L S V V N A R L R
17341 - GTGCAAACACTACGCTATATTGGCGATCCTGCTCAATTACCAGCCCCCGCACATTGC - 17400
- V Q N T T S I L A I L L N Y Q P P A R C
- C K T L R L Y W R S C S I T S P P H I A
- A K H Y V Y I G D P A Q L P A P R T L L
17401 - TGACTAAAGGCACACTAGAACCAGATATTTAATTCAGTGTGCAGACTTATGAAAACAA - 17460
- * L K A H * N Q N I L I Q C A D L * K Q
- D * R H T R T R I F * F S V Q T Y E N N
- T K G T L E P E Y F N S V C R L M K T I
17461 - TAGGTCQAGACATGTTCCCTGGAACCTGTCGCGTGTGCTGCTGAAATGTTGACACTG - 17520
- * V Q T C S L E L V A V V L L K L L T L
- R S R H V P W N L S P L S C * N C * H C
- G P D M F L G T C R R C P A E I V D T V
17521 - TGAGTGGTTTTAGTTTATGACAATAAGCTAAAAGCACACAGGATAAGTCAGCTCAATGCT - 17580
- * V L * F M T I S * K H T R I S Q L N A
- E C F S L * Q * A K S T Q G * V S S M L
- S A L V Y D N K L K A H K D K S A Q C F
17581 - TCAAAATGTTTACAAAGGTATTACACATGATGTTTCATCTGCAATCAACAGACCTC - 17640
- S K C S T K V L L H M M F H L Q S T D L
- Q N V L Q R C Y Y T * C F I C N Q Q T S
- K M F Y K G V I T H D V S S A I N R P Q

```

图 11 续

```

17641 - AAATAGCGTTGTAAGAGAATTTCTTACACGCAATCCTGCTGGAGAAAAGCTGTTTTTA - 17700
- K * A L * E N F L H A I L L G E K L F L
- N R R C K R I S Y T Q S C L E K S C F Y
- I G V V R E F L T R N P A W R K A V F I
17701 - TCTCACCTATAATTCACAGAACGCTGTAGCTTCAAAAATCTTAGGATTGCCTACGCAGA - 17760
- S H L I I H R T L * L Q K S * D C L R R
- L T L * F T E R C S F K N L R I A Y A D
- S P Y N S Q N A V A S K I L G L P T Q T
17761 - CTGTTGATTCATCACAGGGTTCTGAATATGACTATGTCATATTCACACAACTACTGAAA - 17820
- L L I H H R V L N M T M S Y S H K L L K
- C * F I T G F * I * L C H I H T N Y * N
- V D S S Q G S E Y D Y V I F T Q T T E T
17821 - CAGCACACTCTTGTATGTCAACCGCTTCAATGTGGCTATCACAGGGCAAAAATTGGCA - 17880
- Q H T L V M S T A S M W L S Q G Q K L A
- S T L L * C Q P L Q C G Y H K G K N W H
- A H S C N V N R F N V A I T R A K I G I
17881 - TTTTGTGCATAATGTCTGATAGAGATCTTTATGACAACTGCAATTTACAAGTCTAGAAA - 17940
- F C A * C L I E I F M T N C N L Q V * K
- F V H N V * * R S L * Q T A I Y K S R N
- L C I M S D R D L Y D K L Q F T S L E I
17941 - TACCACGTCGCAATGTGGCTACATTACAAGCAGAAAATGTAAGTGGACTTTTAAGGACT - 18000
- Y H V A M W L H Y K Q K M * L D F L R T
- T T S Q C G Y I T S R K C N W T F * G L
- P R R N V A T L Q A E N V T G L F K D C
18001 - GTAGTAAGATCATTACTGGTCTTCATCCTACACAGGCACCTACACACCTCAGCGTTGATA - 18060
- V V R S L L V F I L H R H L H T S A L I
- * * D H Y W S S S Y T G T Y T P Q R * Y
- S K I I T G L H P T Q A P T H L S V D I
18061 - TAAAATCAAGACTGAAGGATTATGTGTTGACATACCAGGCATACCAAAGGACATGACCT - 18120
- * N S R L K D Y V L T Y Q A Y Q R T * P
- K I Q D * R I M C * H T R H T K G H D L
- K F K T E G L C V D I P G I P K D M T Y
18121 - ACCGTAGACTCATCTCTATGATGGGTTTCAAATGAATTACCAAGTCAATGGTTACCCCTA - 18180
- T V D S S L * W V S K * I T K S M V T L
- P * T H L Y D G F O N E L P S Q W L P *
- R R L I S M M G F K M N Y Q V N G Y P N
18181 - ATATGTTTATCACCCGCGAAGAAGCTATTGCTCACGTTGCTGCGTGGATTGGCTTTGATG - 18240
- I C L S P A K K L F V T F V R G L A L M
- Y V Y H P R R S Y S S R S C V D W L * C
- M F I T R E E A I R H V R A W I G F D V
18241 - TAGAGGGCTGTCATGCAACTAGAGATGCTGTGGGTACTAACCTACCTCTCCAGCTAGGAT - 18300
- * R A V M Q L E M L W V L T Y L S S * D
- R G L S C N * R C C G Y * P T S P A R I
- E G C H A T R D A V G T N L P L Q L G F
18301 - TTTCTACAGGTGTTAACTTAGTAGCTGTACCGACTGGTTATGTTGACACTGAAAATAACA - 18360
- F L Q V L T * * L Y R L V M L T L K I T
- F Y R C * L S S C T D W L C * H * K * H
- S T G V N L V A V P T G Y V D T E N N T
18361 - CAGAATTCACCAGAGTTAATGCAAACCTCCACCAGGTGACCAGTTAAACATCTTATAC - 18420
- Q N S P E L M Q N L H Q V T S L N I L Y
- R I H Q S * C K T S T R * P V * T S Y T
- E F T R V N A K P P P G D Q F K H L I P
18421 - CACTCATGTATAAAGGCTTGCCCTGGAATGTAGTGCCTATTAAGATAGTACAAATGCTCA - 18480
- H S G I K A C P G M * C V L R * Y K C S
- T H V * R L A L E C S A Y * D S T N A Q
- L M Y K G L P W N V V R I K I V Q M L S

```

图 11 续

18481 - GTGATACACTGAAAGGATTGTCAGACAGAGTCGTGTTCCCTTTGGGCGCATGGCTTTG - 18540
 - V I H * K D C Q T E S C S S F G R M A L
 - * Y T E R I V R Q S R V R P L G A W L *
 - D T L K G L S D R V V F V L W A H G F E
 18541 - AGCTTACATCAATGAAGTACTTTGTCAAGATTGGACCTGAAAGAACGTGTGTCTGTGTG - 18600
 - S L H Q * S T L S R L D L K E R V V C V
 - A Y I N E V L C Q D W T * K N V L S V *
 - L T S M K Y F V K I G P E R T C C L C D
 18601 - ACAAACGTGCAACTGCTTTTCTACTTCATCAGATACTTATGCCTGCTGGAATCATTCTG - 18660
 - T N V Q L A F L L H Q I L M P A G I I L
 - Q T C N L L F Y F I R Y L C L L E S F C
 - K R A T C F S T S S D T Y A C W N H S V
 18661 - TGGGTTTGGACTATGTCTATAACCCATTTATGATTGATGTTTCAGCAGTGGGGCTTTACGG - 18720
 - W V L T M S I T H L * L M F S S G A L R
 - G F * L C L * P I Y D * C S A V G L Y G
 - G F D Y V Y N P F M I D V Q Q W G F T G
 18721 - GTAACCTTCAGAGTAACCATGACCAACATTGCCAGGTACATGGAATGCACATGTGGCTA - 18780
 - V T F R V T M T N I A R Y M E M H M W L
 - * P S E * P * P T L P G T W K C T C G *
 - N L Q S N H D Q H C Q V H G N A H V A S
 18781 - GTTGTGATGCTATCATGACTAGATGTTTAGCAGTCCATGAGTGCTTTGTTAAGCGCGTTG - 18840
 - V V M L S * L D V * Q S M S A L L S A L
 - L * C Y H D * M F S S P * V L C * A R *
 - C D A I M T R C L A V H E C F V K R V D
 18841 - ATTGGTCTGTGAATACCCATTATAGGAGATGAACTGAGGGTTAATTCTGCTGCAGAA - 18900
 - I G L L N T L L * B M N * G L I L L A E
 - L V C * I P Y Y R R * T E G * F C L Q K
 - W S V E Y P I I G D E L R V N S A C R K
 18901 - AAGTACAACACATGGTTGTAAGTCTGCATTGCTTGCTGATAAGTTCCAGTTCTTCATG - 18960
 - K Y N T W L * S L H C L L I S F Q F F M
 - S T T H G C E V C I A C * * V S S S S *
 - V Q H M V V K S A L L A D K F P V L R D
 18961 - ACATTGGAATCCAAAGGCTATCAAGTGTGTGCCCTCAGGCTGAAGTAGAATGGAAGTTCT - 19020
 - T L E I Q R L S S V C L R L K * N G S S
 - H W K S K G Y Q V C A S G * S R M E V L
 - I G N P K A I K C V P Q A E V E W K F Y
 19021 - ACGATGCTCAGCCATGTAGTGACAAAGCTTACAAAATAGAGGAACCTCTTCTATTCTTATG - 19080
 - T M L S H V V T K L T K * R N S S I L M
 - R C S A M * * Q S L Q N R G T L L F L C
 - D A Q P C S D K A Y K I E E L F Y S Y A
 19081 - CTACACATCAGATAAATTCACIGATGGTGTGTTGTTTGGATTGTAACGTTGATC - 19140
 - L H I T I N S L M V F V C F G I V T L I
 - Y T S R * I H * W C L F V L E L * R * S
 - T H H D K F T D G V C L F W N C N V D R
 19141 - GTTACCAGCCAATGCAATTGTGTGTAGGTTGACACAAGAGTCTTGTCAAACTGAACT - 19200
 - V T Q P M Q L C V G L T Q E S C Q T * T
 - L P S Q C N C V * V * H K S L V K L E L
 - Y P A N A I V C R F D T R V L S N L N L
 19201 - TACCAGGCTGTGATGGTGGTGTGATGTAATAAGCATGCATCCACACTCCAGCTT - 19260
 - Y Q A V M V V V C M * I S M H S T L Q L
 - T R L * W W * F V C E * A C I P H S S F
 - P G C D G G S L Y V N K H A F H T P A F
 19261 - TCGATAAAAGTCATTTACTAATTTAAAGCAATTGCCTTTCTTTTACTATTCTGATAGTC - 19320
 - S I K V H L L I * S N C L S F T I L I V
 - R * K C I Y * F K A I A F L L L F * * S
 - D K S A F T N L K O L P F F Y Y S D S P

图 11 续

19321 - CTTGTGAGTCTCATGGCAAACAAGTAGTGTCCGGATATTGATTATGTTCCACTCAAATCTG - 19380
 - L V S L M A N K * C R I L I M F H S N L
 - L * V S W Q T S S V G Y * L C S T Q I C
 - C E S H G K Q V V S D I D Y V P L K S A
 19381 - CTACGTGATTACACGATGCAATTTAGGTGGTGTGTTGCAGACACCATGCAATGAGT - 19440
 - L R V L H D A I * V V L F A D T M Q M S
 - Y V Y Y T M Q F R W C C L Q T P C K * V
 - T C I T R C N L G G A V C R H H A N E Y
 19441 - ACCGACGACTTGGATGCATATAATATGATGATTTCTGCTGGATTAGCCTATGGATT - 19500
 - T D S T W M H I I * * F L L D L A Y G F
 - P T V L G C I * Y D D F C W I * P M D L
 - R Q Y L D A Y N M M I S A G F S L W I Y
 19501 - ACAACAATTTGATACTTATAACCTGTGGAATACATTTACCAGGTTACAGAGTTAGAAA - 19560
 - T N N L I L I T C G I H L P G Y R V * K
 - Q T I * Y L * P V E Y I Y Q V T E F R K
 - K Q F D T Y N L W N T F T R L Q S L E N
 19561 - ATGTGGCTTATAATGTTGTTAATAAAGGACACTTTGATGGACACGCCGCCGAGCACCTG - 19620
 - M W L I M L L I K D T L M D T P A K H L
 - C G L * C C * * R T L * W T R R R S T C
 - V A Y N V V N K G H F D G H A G E A P V
 19621 - TTTCCATCATTATAATGCTGTTTACAAAGGTAGATGGTATTGATGTTGGAGATCTTTG - 19680
 - F P S L I M L F T Q R * M V L M W R S L
 - F H H * * C C L H K G R W Y * C G D L *
 - S I I N N A V Y T K V D G I D V E I F E
 19681 - AAAATAAGACAACACTTCTGTTAATGTTGCATTTGAGCTTTGGGCTAAGCGTAACATTA - 19740
 - K I R Q H F L L M L H L S F G L S V T L
 - K * D N T S C * C C I * A L G * A * H *
 - N K T T L P V N V A F E L W A K R N I K
 19741 - AACCACTGCCAGAGATTAAGATACTCAATAATTTGGGTGTGATATCGCTGCTAATACTG - 19800
 - N Q C Q R L R Y S I I W V L I S L L I L
 - T S A R D * D T Q * F G C * Y R C * Y C
 - P V P E I K I L N N L G V D I A A N T V
 19801 - TAATCTGGGACTACAAAAGAGAAGCCCCAGCACATGTATCTACAATAGGTGTCTGCACAA - 19860
 - * S G T T K E K P Q H M Y L Q * V S A Q
 - N L G L Q K R S P S T C I Y N R C L H N
 - I W D Y K R E A P A H V S T I G V C T M
 19861 - TGACTGACATTGCCAAGAAACCTACTGAGAGTGTCTTCTTCACTACTGTCTGTTTG - 19920
 - * L T L P R N L R V L V L H L L S C L
 - D * H C Q E T Y * E C L F F T Y C L V *
 - T D I A K K P T E S A C S S L T V L F D
 19921 - ATGGTAGAGTGGAGGACAGGTAGACCTTTTTAGAAACGCCCGTAATGGTGTTTAATAA - 19980
 - M V E W K D R * T P L E T P V M V F * *
 - W * S G R T G R P F * K R P * W C F N N
 - G R V E G Q V D L F R N A R N G V L I T
 19981 - CAGAAGTTCAGTCAAAGGTCTAACACCTTCAAAGGGACCAGCACAAGCTAGCGTCAATG - 20040
 - Q K Y Q S K V * H L Q R D Q H K L A S M
 - R R F S Q R S N T F K G T S T S * R Q W
 - E G S V K G L T P S K G P A Q A S V N G
 20041 - GAGTCAATTAATTGGAGATCAGTAAAACACAGTTTAACTACTTTAAGAAAGTAGACG - 20100
 - E S H * L E N Q * K H S L T T L R K * T
 - S H I N W R I S K N T V * L L * E S R R
 - V T L I G E S V K T Q F N Y F . K K V D G
 20101 - GCATTATTCACAGTTGCCTGAAACCTACTTTACTCAGAGCAGAGACTTAGAGGATTTA - 20160
 - A L F N S C L K P T L L R A E T * R I L
 - H Y S T V A * N L L Y S E Q R L R G F *
 - I I Q Q L P E T Y F T Q S R D L E D F K

图 11 续

20161 - AGCCAGATCACAAATGGAACTGACTTTCTCGAGCTCGCTATGGATGAATTCATACAGC - 20220
 - S P D H K W K L T F S S S L W M N S Y S
 - A Q I T N G N * L S R A R Y G * I H T A
 - P R S Q M E T D F L E L A M D E F I Q R
 20221 - GATATAAGCTCGAGGGCTATGCCTTCGAACACATCGTTTATGGAGATTCAGTCATGGAC - 20280
 - D I S S R A M P S N T S F M E I S V M D
 - I * A R G L C L R T H R L W R F Q S W T
 - Y K L E G Y A F E H I V Y G D F S H G Q
 20281 - AACTGGCGGCTTCATTTAATGATAGGCTTAGCCAAGCGCTCACAAAGATTCACCACTTA - 20340
 - N L A V F I * * * A * P S A H K I H H L
 - T W R S S F N D R L S Q A L T R F T T *
 - L G G L H L M I G L A K R S Q D S P L K
 20341 - AATTAGAGGATTTATCCCTATGGACAGCACAGTGAATAACTTACTCATAACAGATGCGC - 20400
 - N * R I L S L W T A Q * K I T S * Q M R
 - I R G F Y P Y G Q H S E K L L H N R C A
 - L E D F I P M D S T V K N Y F I T D A Q
 20401 - AAACAGTTTCACAAATGTGTGTCTGTGATTGATCTTTACTTGATGACTTTGTGCG - 20460
 - K Q V H Q N V C V L * L I F Y L M T L S
 - N R F I K M C V F C D * S F T * * L C R
 - T G S S K C V C S V I D L L L D D F V E
 20461 - AGATAATAAGTCACAAGATTTGTCAGTGATTTCAAAGTGGTCAAGGTTACAATTGACT - 20520
 - R * * S H K I C Q * F Q K W S R L Q L T
 - D N K V T R F V S D F K S G Q G Y N * L
 - Y I K S Q D L S V I S K V V K V T I D Y
 20521 - ATGCTGAAATTTTCATTCATGCTTTGGTGAAGGATGGACATGTTGAAACCTTCTACCCAA - 20580
 - M L K F H S C F G V R M D M L K P S T Q
 - C * N F I H A L V * G W T C * N L L P K
 - A E I S F M L W C K D G H V E T F Y P K
 20581 - AACTACAAGCAAGTCAAGCGTGGCAACCAGGTGTGGGATGCCTAAGTGTACAAGATGC - 20640
 - N Y K Q V K R G N Q V L R C L T C T R C
 - T T S K S S V A T R C C D A * L V Q D A
 - L Q A S Q A W Q P G V A M P N L Y K M Q
 20641 - AAAGAATGCTTCTGAAAGTGTGACCTTCAGAATTATGGTGAATAATGCTGTTATACCAA - 20700
 - K E E F L K S V T F R I M V K M L L Y Q
 - K N A S * K V * P S E L W * K C C Y T K
 - R M L L E K C D L Q N Y G E N A V I P K
 20701 - AAGGAATAATGATGAATGTGCGAAAGTATACTCAACTGTGTCAATACTTAAATACACTTA - 20760
 - K E * * * M S Q S I L N C V N T * I H L
 - R N N D E C R K V Y S T V S I L K Y T Y
 - G I M M N V A K Y T Q L C Q Y L N T L T
 20761 - CTTTAGCTGTACCCTACAACATGAGAGTTATTCACTTTGGTGTGGCTCTGATAAAGGAG - 20820
 - L * L Y P T T * E L F T L V L A L I K E
 - F S C T L Q H E S Y S L W C W L * * R S
 - L A V P Y N M R V I H F G A G S D K G V
 20821 - TTGCACCAGGTACAGCTGTGCTCAGACAATGGTGGCAACTGGCACACTACTTGTGCGATT - 20880
 - L H Q V Q L C S D N G C Q L A H Y L S I
 - C T R Y S C A Q T M V A N W H T T C R F
 - A P G T A V L R Q W L P T G T L L V D S
 20881 - CAGATCTTAATGACTTCGTCTCCGACGCAGATTCTACTTTAATTGGAGACTGTGCAACAG - 20940
 - Q I L M T S S P T Q I L L * L E T V Q Q
 - R S * * L R L R R F Y F N W R L C N S
 - D L N D F V S D A D S T L I G D C A T V
 20941 - TACATACGGCTAATAAATGGGACCTTATTATTAGCGATATGTATGACCCTAGGACCAAC - 21000
 - Y I R L I N G T L L L A I C M T L G P N
 - T Y G * * M G P Y Y * R Y V * P * D Q T
 - H T A N K W D L I I S D M Y D P R T K H

图 11 续

21001 - ATGTGACAAAAGAGAATGACTCTAAAGAAGGGTTTTCACTTATCTGTGTGGATTATAA - 21060
 - M * Q K R M T L K K G F S L I C V D L *
 - C D K R E * L * R R V F H L S V W I Y K
 - V T K E N D S K E G F F T Y L C G F I K
 21061 - AGCAAAACTAGCCCTGGGTGTTCTATAGCTGTRAAGATAACAGAGCATTCTTGGAAATG - 21120
 - S K N * P W V V L * L * R * Q S I L G M
 - A K T S P G W F Y S C K D N R A F L E C
 - Q K L A L G G S I A V K I T E H S W N A
 21121 - CTGACCTTACAAGCTTATGGGCCATTTCTCATGGTGACAGCTTTTGTACAAATGTAA - 21180
 - L T F T S L W A I S H G G Q L L L Q M *
 - * P L Q A Y G P F L M V D S F C Y K C K
 - D L Y K L M G H F S W W T A F V T N V N
 21181 - ATGCATCATCATCGGAAGCATTTTAATGGGGCTAACTATCTTGGCAAGCCGAAGAAC - 21240
 - M H H H R K H F * L G L T I L A S R R N
 - C I I I G S I F N W G * L S W Q A E G T
 - A S S S E A F L I G A N Y L G K P K E Q
 21241 - AAATTGATGGCTATACCATGCGTAAGTACTACATTTTCTGGAGAACACAAATCCTATCC - 21300
 - K L M A I P C M L T T F S G G T Q I L S
 - N * W L Y H A C * L H F L E E H K S Y P
 - I D G Y T M H A N Y I F W R N T N P I Q
 21301 - AGTTGTCTTCTTATTCATCTTTGACATGAGCAAATTCCTCTTAAATTAAGAGGAAGT - 21360
 - S C L P I H S L T * A N F L L N * E E L
 - V V F L F T L * H E Q I S S * I K R N C
 - L S S Y S L F D M S K F P L K L R G T A
 21361 - CTGTAATGTCTCTTAAGGAGAATCAATCAATGATATGATTATTCTTCTTCTGGAAAAG - 21420
 - L * C L L R R I K S M I * F I L F W K K
 - C N V S * G E S N Q * Y D L F S S G K R
 - V M S L K E N Q I N D M I Y S L L E K G
 21421 - GTAGCTTATCATTAGAGAAAACAACAGAGTTGTGGTTCAAGTGATATTCTTGTAAACA - 21480
 - V G L S L E K T T E L W F Q V I F L L T
 - * A Y H * R K Q Q S C G F K * Y S C * Q
 - R L I I R E N N R V V V S S D I L V N N
 21481 - ACTAAGGAACATGTTTATTTCTTATTATTTCTACTCTACTAGTGGTAGTGACCTTG - 21540
 - T K R T C L F S Y Y F L L S L V V V T L
 - L N E H V Y F L I I S Y S H * W * * P *
 - * T N M F I F L L F L T L T S G S D L D
 21541 - ACCGGTGCACCACTTTTATGATGTTCAAGCTCCTAATTACACTCAACATACTTCATCTA - 21600
 - T G A P L L M M F K L L I T L N I L H L
 - P V H H F * * C S S S * L H S T Y F I Y
 - R C T T F D D V Q A P N Y T Q H T S S M
 21601 - TGAGGGGGTTTACTATCCTGATGAAATTTTATGATCAGACACTCTTTATTTAAGTACAGG - 21660
 - * G G F T I L M K F L D Q T L F I * L R
 - E G G L L S * * N F * I R H S L F N S G
 - R G V Y Y P D E I F R S D T L Y L T Q D
 21661 - AITTTTCTTCCATTTTATTCTAATGTTACAGGGTTTCATACTATTAATCATAAGTTG - 21720
 - I Y F F H F I L M L Q G F I L L I R L
 - F I S S I L F * C Y R V S Y Y * S Y V W
 - L F L P F Y S N V T G F H T I N H T F G
 21721 - GCAACCTGTCATACCTTTAAGGATGGTATTTATTTGCTGCCACAGAGAAATCAATG - 21780
 - A T L S Y L L R M V F I L L P Q R N Q M
 - Q P C H T F * G W Y L F C C H R E I K C
 - N P V I P F K D G I Y F A A T E K S N V
 21781 - TTGTCCTGGTTGGGTTTTGGTCTACCATGAACAAGTCACAGTCGGTGATTATTA - 21840
 - L S V V G F L V L P * T T S H S R * L L
 - C P W L G F W F Y H E Q Q V T V G D Y Y
 - V R G W V F G S T M N N K S O S V I I I

图 11 续

```

21841 - TTAACAATTCTACTAATGTTGTTATACGAGCATGTAACCTTGAATTGTGTGACAACCCTT - 21900
- L T I L L M L L Y E H V T L N C V T T L
- * Q F Y * C C Y T S M * L * I V * Q P F
- N N S T N V V I R A C N F E L C D N P F
21901 - TCTTTGCTGTTTCTAAACCCATGGGTACACAGACACATACTATGATATTGATAATGCAT - 21960
- S L L F L N P W V H R H I L * Y S I M H
- L C C F * T H G Y T D T Y Y D I R * C I
- F A V S K P M G T Q T H T M I F D N A F
21961 - TTAATTGCACCTTCGAGTACATATCTGATGCCTTTTCGCTTGATGTTTCAGAAAAGTCAG - 22020
- L I A L S S T Y L M P F R L M F Q K S Q
- * L H F R V H I * C L F A * C F R K V R
- N C T F E Y I S D A F S L D V S E K S G
22021 - GTAATTTTAAACACTTACGAGAGTTTGTGTTTAAAAATAAGATGGGTTTCTCTATGTTT - 22080
- V I L N T Y E S L C L K I K M G F S M F
- * F * T L T R V L C V * K * R W V S L C L
- N F K H L R E F V F K N K D G F L Y V Y
22081 - ATAAGGGCTATCAACCTATAGATGTAGTTCGTGATCTACCTTCTGGTTTTAACACTTTGA - 22140
- I R A I N L * M * F V I Y L L V L T L *
- * G L S T Y R C S S * S T F W F * H F E
- K G Y Q P I D V V R D L P S G F N T L K
22141 - AACCTATTTTTAAGTTGCCCTTGGTATTAACATTACAAATTTAGAGCCATTCTTACAG - 22200
- N L F L S C L L V L T L Q I L E P F L Q
- T Y F * V A S W Y * H Y K F * S H S Y S
- P I F K L P L G I N I T N F R A I L T A
22201 - CCTTTTCACTGCTCAAGACATTTGGGGCACGTCAGCTGCAGCCTATTTTGTGGCTATT - 22260
- P F H L L K T F G A R Q L Q P I L L A I
- L F T C S R H L G H V S C S L F C W L F
- F S P A Q D I W G T S A A A Y F V G Y L
22261 - TAAAGCCAACTACATTTATGCTCAAGTATGATGAAAATGGTACAATCACAGATGCTGTTG - 22320
- * S Q L H L C S S M M K M V Q S Q M L L
- K A N Y I Y A Q V * * K W Y N H R C C *
- K P T T F M L K Y D E N G T I T D A V D
22321 - ATTGTTCTCAAAATCCACTTGTGTAACCTCAATGCTCTGTTAAGAGCTTTGAGATTGACA - 22380
- I V L K I H L L N S N A L L R A L R L T
- L F S K S T C * T Q M L C * E L * D * Q
- C S Q N P L A E L K C S V K S F E I D K
22381 - AAGGAATTTACCAGACCTCTAATTTCCAGGGTGTTCCTCAGGAGATGTTGTGAGATTCC - 22440
- K E F T R P L I S G L F P Q E M L * D S
- R N L P D L * F Q G C S L R R C C E I P
- G I Y Q T S N F R V V P S G D V V R F P
22441 - CTAATATTACAACTTGTGTCCTTTTGGAGAGGTTTTAATGCTACTAAATCCCTTCTG - 22500
- L I L Q T C V L L E R F L M L L N S L L
- * Y Y K L V S F W R G F * C Y * I P F C
- N I T N L C P F G E V F N A T K F P S V
22501 - TCTATGCATGGGAGAGAAAAAATTTCTAATTTGTTGCTGATTACTCTGTGCTCTACA - 22560
- S M H G R E K K F L I V L L I T L C S T
- L C M G E K K N F * L C C * L L C A L Q
- Y A W E R K K I S N C V A D Y S V L Y N
22561 - ACTCAAGATTTTTTCAACCTTAAAGTGTATGGCGTTTCTGCCACTAAGTTGAATGATC - 22620
- T Q H F F Q P L S A M A F L P L S * M I
- L N I F F N L * V L W R F C H * V E * S
- S T F F S T F K C Y G V S A T K L N D L
22621 - TTTGCTTCTCCAATGTCTATGCAGATTCTTTGTAGTCAAGGGAGATGATGTAAGACAAA - 22680
- F A S P M S M Q I L L * S R E M M * D K
- L L L Q C L C R F F C S Q G R * C K T N
- C F S N V Y A D S F V V K G D D V R Q I

```

图 11 续

22681 - TAGCGCCAGGACAACTGGTGTATTGCTGATTATAATTATAAATTGCCAGATGATTTC - 22740
 - * R Q D K L V L L L I I I I N C Q M I S
 - S A R T N W C Y C * L * L * I A R * F H
 - A P G Q T G V I A D Y N Y K L P D D F M
 22741 - TGGTTGTGTCCTTGCTGGAATACTAGGAACATTGATGCTACTTCAACTGGTAATTATA - 22800
 - W V V S L L G I L G T L M L L Q L V I I
 - G L C P C L E Y * E H * C Y F N W * L *
 - G C V L A W N T R N I D A T S T G N Y N
 22801 - ATTATAAATAGGTATCTTAGACATGGCAAGCTTAGGCCCTTGAGAGAGACATATCTA - 22860
 - I I N I G I L D M A S L G P L R E T Y L
 - L * I * V S * T W Q A * A L * E R H I *
 - Y K Y R Y L R H G K L R P F F E R D I S N
 22861 - ATGTGCCTTTCTCCCCTGATGGCAACCTTGACCCACCTGCTCTAATTGTTATTGGC - 22920
 - M C L S P L M A N L A P H L L L I V I G
 - C A F L P * W Q T L H P T C S * L L L A
 - V P F S P D G K P C T P P A L N C Y W P
 22921 - CATTAAATGATTATGGTTTTACACCACTACTGGCATTGGCTACCAACCTTACAGAGTTG - 22980
 - H * M I M V F T P L L A L A T N L T E L
 - I K * L W F L H H Y W H W L P T L Q S C
 - L N D Y G F Y T T T G I G Y Q P Y R V V
 22981 - TAGTACTTTCTTTGAACTTTTAAATGCACCGCCACGGTTTGTGGACCAAATATCCA - 23040
 - * Y F L L N F * M H R P R F V D Q N Y P
 - S T P F * T F K C T G H G L W T K I I H
 - V L S F E L L N A P A T V C G P K L S T
 23041 - CTGACCTTATAAGAACCAGTGTGTCAATTTTAAATTTTATGGACTCACTGGTACTGGTG - 23100
 - L T L L R T S V S I L I L M D S L V L V
 - * P Y * E P V C Q F * F * W T H W Y W C
 - D L I K N Q C V N F N F N G L T G T G V
 23101 - TGTTAACTCCTTCTCAAGAGATTCAACCATTCAACAATTTGGCCGTGTTTCTG - 23160
 - C * L L L Q R D F N H F N N L A V M F L
 - V N S F F K E I S T I S T I W P * C F *
 - L T P S S K R F Q P F Q Q F G R D V S D
 23161 - ATTTCACTGATCCGTTGAGATCCTAAAACATCTGAAATATTAGACATTTACCTTGCT - 23220
 - I S L I P F E I L K H L K Y * T F H L A
 - F H * F R S R S * N I * N I R H F T L L
 - F T D S V R D P K T S E I L D I S P C S
 23221 - CTTTTGGGGTGAAGTGAATTACACCTGGAACAAATGCTTCACTGAAGTTGCTGTTT - 23280
 - L L G V * V * L H L E Q M L H L K L L F
 - F W G C K C N Y T W N K C F I * S C C S
 - F G G V S V I T P G T N A S S E V A V L
 23281 - TATATCAAGATGTTAACTGCACTGATTTCTACAGCAATTCATGCAGATCACTCACAC - 23340
 - Y I K M L T A L M F L Q Q F M Q I N S H
 - I S R C * L H * C F Y S N S C R S T H T
 - Y Q D V N C T D V S T A I H A D Q L T P
 23341 - CAGCTTGGCGCATATATTCTACTGGAACAATGTTCCAGACTCAAGCAGGCTGTCTTA - 23400
 - Q L G A Y I L L E T M Y S R L K Q A V L
 - S L A H I F Y W K Q C I P D S S R L S Y
 - A W R I Y S T G N N V F Q T Q A G C L I
 23401 - TAGGAGCTGAGCATGTCGACACTTCTTATGAGTGGACATTCTATTGGAGCTGGCATT - 23460
 - * E L S M S T L L M S A T F L L E L A F
 - R S * A C R H F L * V R H S Y W S W H L
 - G A E H V D T S Y E C D I P I G A G I C
 23461 - GTGCTAGTTACCATACAGTTTCTTTATTACGTAGTACTAGCCAAAATCTATTGTGGCTT - 23520
 - V L V T I Q F L Y Y V V L A K N L L W L
 - C * L P Y S F F I T * Y * P K I Y C G L
 - A S Y H T V S L L R S T S Q K S I V A Y

图 11 续

23521 - ATACTAAGTCTTTAGGTGCTGATAGTTC AATTGCTTACTCTAATAACACCATTGCTATAC - 23580
 - I L C L * V L I V Q L L T L I T P L L Y
 - Y Y V F R C * * F N C L L * * H H C Y T
 - T M S L G A D S S I A Y S N N T I A I P
 23581 - C T A C T A A C T T T T C A A T T A G C A T T A C T A C A G A A G T A A T G C C T G T T T C T A T G G C T A A A A C C T - 23640
 - L L T F Q L A L L Q K * C L F L W L K P
 - Y * L F N * H Y Y R S N A C F Y G * N L
 - T N F S I S I T T E V M P V S M A K T S
 23641 - C C G T A G A T T G T A A T A T G T A C A T C T G C G G A G A T T C T A C T G A A T G T G C T A A T T T G C T T C T C C - 23700
 - P * I V I C T S A E I L L N V L I C F S
 - R R L * Y V H L R R F Y * M C * F A S P
 - V D C N M Y I C G D S T E C A N L L L Q
 23701 - A A T A T G G T A G C T T T T G C A C A C A A C T A A A T C G T G C A C T C T C A G G T A T T G C T G C T G A A C A G G - 23760
 - N M V A F A H N * I V H S Q V L L L N R
 - I W * L L H T T K S C T L R Y C C * T G
 - Y G S F C T Q L N R A L S G I A A E Q D
 23761 - A T C G C A A C A C A C G T G A A G T G T T C G C T C A A G T C A A C A A A T G T A C A A A C C C C A A C T T T G A - 23820
 - I A T H V K C S L K S N K C T K P Q L *
 - S Q H T * S V R S S Q T N V Q N P N F E
 - R N T R E V F A Q V K Q M Y K T P T L K
 23821 - A A T A T T T T G G T G G T T T T A A T T T T C A C A A T A T T A C C T G A C C C T C T A A G C C A A C T A A G A - 23880
 - N I L V V L I F H K Y Y L T L * S Q L R
 - I F W W F * F F T N I T * P S K A N * E
 - Y F G G F N F S Q I L P D P L K P T K R
 23881 - G G T C T T T T A T T G A G G A C T T G C T C T T T A A A A G G T G A C A C T C G C T G A T G C T G G C T T C A T G A - 23940
 - G L L L R T C S L I R * H S L M L A S *
 - V F Y * G L A L * * G D T R * C W L H E
 - S F I E D L L F N K V T L A D A G F M K
 23941 - A G C A A T A T G G C G A A T G C C T A G G T G A T A T A A T G C T A G A G A T C T C A T T T G T G C G C A G A G T - 24000
 - S N M A N A * V I L M L E I S F V R R S
 - A I W R M P R * Y * C * R S H L C A E V
 - Q Y G E C L G D I N A R D L I C A Q K F
 24001 - T C A A T G G A C T T A C A G T G T T G C C A C C T C T G C T C A C T G A T G A T A T G A T T G C T G C C T A C A C T G - 24060
 - S M D L Q C C H L C S L M I * L L P T L
 - Q W T Y S V A T S A H * * Y D C C L H C
 - N G L T V L P P L L T D D M I A A Y T A
 24061 - C T G C T C T A G T T A G T G G T A C T G C C A C T G C T G G A T G G A C A T T T G G T G C T G G C G C T G C T C T T C - 24120
 - L L * L V V L P L L D G H L V L A L L F
 - C S S * W Y C H C W M D I W C W R C S S
 - A L V S G T A T A G W T F G A G A A L Q
 24121 - A A A T A C C T T T T G C T A T G C A A A T G G C A T A T A G G T T C A A T G G C A T T G G A G T T A C C C A A A T G - 24180
 - K Y L L L C K W H I G S M A L E L P K M
 - N T F C Y A N G I * V Q W H W S Y P K C
 - I P F A M Q M A Y R F N G I G V T Q N V
 24181 - T T C T C T A T G A G A C C A A A A C A A A T C G C C A C C A A T T T A C A A G C G A T T A G T C A A A T T C - 24240
 - F S M R T K N K S P T N L T R R L V K F
 - S L * E P K T N R Q P I * Q G D * S N S
 - L Y E N Q K Q I A N Q F N K A I S Q I Q
 24241 - A R G A A T C A C T T A C A C A C A T C A A C T G C A T T G G G C A A G C T G C A A G C G T T G T T A A C C A G A - 24300
 - K N H L Q Q H Q L H W A S C K T L L T R
 - R I T Y N N I N C I G Q A A R R C * P E
 - E S L T T T S T A L G K L Q D V V N Q N
 24301 - A T G C T C A A G C A T T A A C A C A C T T G T T A A C A A C T T A G C T C T A A T T T T G G T G C A A T T T C A A - 24360
 - M L K H * T H L L N N L A L I L V Q F Q
 - C S S I K H T C * T T * L * F W C N F K
 - A Q A L N T L V K Q L S S N F G A I S S

图 11 续

24361 - GTGTGCTAAATGATATCCTTTTCGCGACTTGATAAAGTCGAGGCGGAGGTACAAATTGACA - 24420
 - V C * M I S F R D L I K S R R R Y K L T
 - C A K * Y P F A T * * S R G G G T N * Q
 - V L N D I L S R L D K V E A E V Q I D R
 24421 - GGTAAATTACAGGCAGACTTCAAAGCCTTCAAACCTATGTAACACAACAATAATCAGGG - 24480
 - G * L Q A D F K A F K P M * H N N * S G
 - V N Y R Q T S K P S N L C N T T T N Q G
 - L I T G R L Q S L Q T Y V T Q Q L I R A
 24481 - CTGCTGAAATCAGGGCTTCTGCTAATCTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTG - 24540
 - L L K S G L L L I L L L L L K C L S V F L
 - C * N Q G F C * S C C Y * N V * V C S W
 - A E I R A S A N L A A T K M S E C V L G
 24541 - GACAATCAAAAAGAGTTGACTTTTGTTGAAAGGGCTACCACCTTATGTCCTTCCCACAAG - 24600
 - D N Q K E L T F V E R A T T L C P S H K
 - T I K K S * L L W K G L P P Y V L P T S
 - Q S K R V D F C G K G Y H L M S F P Q A
 24601 - CAGCCCCGCATGGTGTGCTTCTTACATGTACGTATGTGCCATCCCAGGAGAGGAAGT - 24660
 - Q P R M V L S S Y M S R M C H P R R G T
 - S P A W C C L P T C H V C A I P G E E L
 - A P H G V V F L H V T Y V P S Q E R N F
 24661 - TCACCAAGCGCCAGCAATTTGTCTGAAGGCAAAGCATACTCCCTCGTGAAGGTGTT - 24720
 - S P Q R Q Q F V M K A K H T S L V K V F
 - H H S A S N L S * R Q S I L P S * R C F
 - T T A P A I C H E G K A Y F P R E G V F
 24721 - TTGTGTTAATGGCACTTCTGGTTTATTACACAGAGGAAGTCTTTTCTCCACAATAA - 24780
 - L C L M A L L G L L H R G T S F L H K *
 - C V * W H F L V Y Y T E E L L F S T N N
 - V F N G T S W F I T Q R N F F S P Q I I
 24781 - TTACTACAGACAATACATTTGTCTCAGGAAATTTGATGTCTGTTATTTGGCATCATTAA - 24840
 - L L Q T I H L S Q E I V M S L L A S L T
 - Y Y R Q Y I C L R K L * C R Y W H H * Q
 - T T D N T F V S G N C D V V I G I I N N
 24841 - ACACAGTTTATGATCCTCTGCAACCTGAGCTTGACTCATTCAAAGAAGAGCTGGACAAGT - 24900
 - T Q F M I L C N L S L T H S K K S W T S
 - H S L * S S A T * A * L I Q R R A G Q V
 - T V Y D P L Q P E L D S F K E E L D K Y
 24901 - ACTTCAAAAATCATAATCACCAGATGTTGATCTTGGCGACATTTCAGGCATTAACGCTT - 24960
 - T S K I I H H Q M L I L A T F Q A L T L
 - L Q K S Y I T R C * S W R H F R H * R F
 - F K N H T S P D V D L G D I S G I N A S
 24961 - CTGTGCTCAACATTCAAAAGAAATGACCGCTCAATGAGGTCGCTAAAATTTAAATG - 25020
 - L S S T F K K K L T A S M R S L K I * M
 - C R Q H S K R N * P P Q * G R * K F K *
 - V V N I Q K E I D R L N E V A K N L N E
 25021 - AATCACTCATTGACCTTCAAGAATTGGGAAAATATGAGCAATATATTAATGGCCTTGGT - 25080
 - N H S L T F K N W E N M S N I L N G L G
 - I T H * P S R I G K I * A I Y * M A L V
 - S L I D L Q E L G K Y E Q Y I K W P W Y
 25081 - ATGTTTGGCTCGGCTTCAATGCTGACTAATGCCATCGTCATGGTTACAATCTTGCTTT - 25140
 - M F G S A S L L D * L P S S W L Q S C F
 - C L A R L H C W T N C H R H G Y N L A L
 - V W L G F I A G L I A I V M V T I L L C
 25141 - GTTGCATGACTAGTTGTTGCAGTGCCTCAAGGGTGCATGCTCTTGTGGTTCTTGCTGCA - 25200
 - V A * L V V A V A S R V H A L V V L A A
 - L H D * L L Q L P Q G C M L L W F L L Q
 - C M T S C C S C L K G A C S C G S C C K

图 11 续

25201 - AGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACATTACACATAAA - 25260
 - S L M R M T L S Q F S R V S N Y I T H K
 - V * + G * L * A S S Q G C Q I T L H I N
 - F D E D D S E P V L K G V K L H Y T * T
 25261 - CGAACTTATGGATTGTTTATGAGATTTTTACTCTGGATCAATTACTGCACAGCCAGT - 25320
 - R T Y G F V Y E I F Y S W I N Y C T A S
 - E L M D L F M R F F T L G S I T A Q P V
 - N L W I C L * D F L L L D Q L L H S Q *
 25321 - AAAAATTGACAATGCTTCTCCTGCAAGTACTGTTTCATGCTACAGCAACGATACCGCTACA - 25380
 - K N * Q C F S C K Y C S C Y S N D T A T
 - K I D N A S P A S T V H A T A T I P L Q
 - K L T M L L L Q V L F M L Q Q R Y R Y K
 25381 - AGCCTCACTCCCTTTTCGGATGGCTTGTATTGGCGTTGCATTCTTGCTGTTTTTCAGAG - 25440
 - S L T P F R M A C Y W R C I S C C F S E
 - A S L P F G W L V I G V A F L A V F Q S
 - P H S L S D G L L L A L H F L L F F R A
 25441 - CGCTACCAAATAATGGCGTCAATAAAGATGGCAGCTAGCCCTTATAAGGGCTTCCA - 25500
 - R Y Q N N C A Q * K M A A S P L * G L P
 - A T K I I A L N K R W Q L A L Y K G F Q
 - L P K * L R S I K D G S * P F I R A S S
 25501 - GTTCATTGCAATTTACTGCTGCTATTGTTACCATCTATTCACATCTTTTGCTTGCTGCG - 25560
 - V H L Q F T A A I C Y H L F T S F A C R
 - F I C N L L L F V T I Y S H L L L V A
 - S F A I Y C C Y L L P S I H I F C L S L
 25561 - TGCAGGTAAGGAGGCGCAATTTTTGTACCTCTATGCCTGATATATTTCTACAATGCAT - 25620
 - C R * G G A I F V P L C L D I F S T M H
 - A G K E A Q F L Y L Y A L I Y F L Q C I
 - Q V R R R N F C T S M P * Y I F Y N A S
 25621 - CAACCGATGTAGAATTATTATGAGATGTTGGCTTGTGGGAAGTGCAAATCCAAGAACCC - 25680
 - Q R M * N Y Y E M L A L L E V Q I Q E P
 - N A C R I I M R C W L C W K C K S K N P
 - T H V E L L * D V G F V G S A N P R T H
 25681 - ATTACTTTATGATGCCAACTACTTTGTTTGGCTGGCACACATAACTATGACTACTGTAT - 25740
 - I T L * C Q L L C L L A H T * L * L L Y
 - L L Y D A N Y F V C W H T H N Y D Y C I
 - Y F M M P T T L F A G T H I T M T T Y Y
 25741 - ACCATATAACAGTGTACAGATACAATTGTCGTTACTGAAGGTGACGGCATTTCACACC - 25800
 - T I * Q C H R Y N C R Y * R * R H F N T
 - P Y N S V T D T I V V T E G D G I S T P
 - H I T V S Q I Q L S L L K V T A F Q H Q
 25801 - AAAACTCAAAGAAGACTACCAATTTGGTGTATTCTGAGGATAGGCACACTCAGGTGTAA - 25860
 - K T Q R R L P N W W L F * G * A L R C *
 - K L K E D Y Q I G G Y S E D R H S G V K
 - N S K K T T K L V V I L R I G T Q V L K
 25861 - AGACTATGTCGTTGTACATGGCTATTTACCCGAAGTTTACTACCAGCTTGAGTCTACACA - 25920
 - R L C R C T W L F H R S L L P A * V Y T
 - D Y V V V H G Y F T E V Y Y Q L E S T Q
 - T M S L Y M A I S P K F T T S L S L H K
 25921 - AATTACTACAGACACTGGTATTGAAAATGCTACATTTCTTCATCTTTAACAAGCTTGTAA - 25980
 - N Y Y R H W Y * K C Y I L H L * Q A C *
 - I T T D T G I E N A T F F I F N K L V K
 - L L Q T L V L K M L H S S S L T S L L K
 25981 - AGACCCACCGAATGTGCAAATACACAAATCGACGGCTCTTCAGGAGTTGCTAATCCAGC - 26040
 - R P T E C A N T H N R R L F R S C * S S
 - D P P N V Q I H T I D G S S G V A N P A
 - T H R M C K Y T Q S T A L Q E L L I Q O

图 11 续

26041 - AATGGATCCAATTTATGATGAGCCGACGACGACTACTAGCGTGCCTTTGTAAGCACAAGA - 26100
 - N G S N L * * A D D D Y * R A F V S T R
 - M D P I Y D E P T T T T S V P L * A Q E
 - W I Q F M M S R R R L L A C L C K H K K
 26101 - AAGTGAGTACGAACCTTATGTACTCATTGCTTCGGAAGAACAGGTACGTTAATAGTTAA - 26160
 - K * V R T Y V L I R F G R N R Y V N S *
 - S E Y E L M Y S F V S E E T G T L I V N
 - V S T N L C T H S F R K K Q V R * * L I
 26161 - TAGCGTACTTCTTTTCTGCTTCGTTGCTGTTATCTTCTAGTCACACTAGCCATCCTTAC - 26220
 - * R T S F S C F R G I L A S R T S H P Y
 - S V L L F L A F V V F L L V T L A I L T
 - A Y F F F L L S W Y S C * S H * P S L L
 26221 - TGCGCTTCGATTGTGTGCTACTGCTGCAATATTGTTAACGTGAGTTAGTAAACCAAC - 26280
 - C A S I V C V L L Q Y C * R E F S K T N
 - A L R L C A Y C C N I V N V S L V K P T
 - R E D C V R T A A I L L T * V * * N Q R
 26281 - GGTTCACGTCTACTCGCGTGTAAAAATCTGAACTCTTCTGAAGGAGTTCCTGATCTTCT - 26340
 - G L R L L A C * K S E L F * R S S * S S
 - V Y V Y S R V K N L N S S E G V P D L L
 - F T S T R V L K I * T L L K E F L I F W
 26341 - GGTCTAAACGAACCTAATATTATTATTCTGTTGGAACCTTAACATTGCTTATCATG - 26400
 - G L N E L T I I I I L F G T L T L L I M
 - V * T N * L L L L F C L E L * H C L S W
 - S K R T N Y Y Y Y S V W N F N I A Y H G
 26401 - GCAGACAACGGTACTATTACCGTTGAGGAGCTTAAACAACCTCCTGGAACAATGGAACCTA - 26460
 - A D N G T I T V E E L K Q L L E Q W N L
 - Q T T V L L P L R S L N N S W N N G T *
 - R Q R Y Y R * G A * T T P G T M E P S
 26461 - GTAATAGGTTTCTTATTCCTAGCCTGATTATGTACTACAATTTGCCTATTCTAATCGG - 26520
 - V I G F L F L A W I M L L Q F A Y S N R
 - * * V S Y S * P G L C Y Y N L P I L I G
 - N R F P I P S L D Y V T T I C L F * S E
 26521 - AACAGGTTTTGTACATAATAAAGCTTGTTCCTCTGGCTCTTGTGGCCAGTAACACT - 26580
 - N R F L Y I I K L V F L W L L W P V T L
 - T G F C T * * S L F S S G S C G Q * H L
 - Q V F V H N K A C F P L A L V A S N T C
 26581 - GCTTGTGTTGCTGCTGTGTCTACAGAATTAATGGGTGACTGGCGGGATTGCGATT - 26640
 - A C F V L A V V Y R I N W V T G G I A I
 - L V L C L L L S T E L I G * L A G L R L
 - L F C A C C C L Q N * L G D W R D C D C
 26641 - GCAATGGCTTGTATTGTAGGCTTGATGTGGCTTAGCTACTTCGTTGCTTCCTCAGGCTG - 26700
 - A M A C I V G L M W L S Y F V A S F R L
 - Q W L V L * A * C G L A T S L L P S G C
 - N G L Y C R L D V A * L L R C F L Q A V
 26701 - TTTGCTGTTACCGCTCAATGTGGTCATCAACCCAGAAACAACATTCTTCTCAATGTG - 26760
 - F A R T R S M W S F N P E T N I L L N V
 - L L V P A Q C G H S T Q K Q T F F S M C
 - C S Y P L N V V I Q P R N K H S S O C A
 26761 - CCTCTCGGGGACAAATGTGACCAGACCGCTCATGGAAAGTGAACCTTGTATTGGTGTGCT - 26820
 - P L R G T I V T R P L M E S E L V I G A
 - L S G G Q L * P D R S W K V N L S L V L
 - S P G D N C D Q T A H G K * T C H W C C
 26821 - GTGATCATTCGTGGTCACTTGCGAATGGCCGACACTCCCTAGGGCGCTGTGACATTAG - 26880
 - V I I R G H L R M A G H S L G R C D I K
 - * S F V V T C E W P D T P * G A V T L R
 - D H S W S L A N G R T L P R A L * H * G

图 11 续

26881 - GACCTGCCAAAAGAGATCACTGTGGCTACATCACGAACGCTTCTTATTACAAATTAGGA - 26940
 - D L P K E I T V A T S R T L S Y Y K L G
 - T C Q K R S L W L H H E R F L I T N * E
 - P A K R D H C G Y I T N A F L L Q I R S
 26941 - GCGTCGCAGCGTGTAGGCACTGATTCAGGTTTTGCTGCATACAACCGCTACCGTATTGGA - 27000
 - A S Q R V G T D S G F A A Y N R Y R I G
 - R R S V * A L I Q V L L H T T A T V L E
 - V A A C R H * F R F C C I Q P L P Y W K
 27001 - AACTATAAATTAAATACAGACCACCGCGGTAGCAACGACAATATTGCTTTGCTAGTACAG - 27060
 - N Y K L N T D R A G S N D N I A L L V Q
 - T I N * I Q T T P V A T T I L L C * Y S
 - L * I K Y R P R R * Q R Q Y C F A S T V
 27061 - TAAGTGACAACAGATGTTTCATCTTGTGACTCCAGGTTACAATAGCAGAGATATTGAT - 27120
 - * V T T D V S S C * L P G Y N S R D I D
 - K * Q Q M F H L V D F Q V T I A E I L I
 - S D N R C F I L L T S R L Q * Q R Y * L
 27121 - TATCATTATGAGGACTTTCAGGATTGCTATTTGGAATCTTGACGTTATAATAAGTTCAAT - 27180
 - Y H Y E D F Q D C Y L E S * R Y N K F N
 - I I M R T F R I A I W N L D V I I S S I
 - S L * G L S G L L F G I L T L * * V Q *
 27181 - AGTGAGACAATTATTTAAGCCTCTAACTAAGAAGAATTATTCGGAGTTAGATGATGAAGA - 27240
 - S E T I I * A S N * E E L F G V R * * R
 - V R Q L F K P L T K K N Y S E L D D E E
 - * D N Y L S L * L R R I I R S * M M K N
 27241 - ACCTATGGAGTTAGATTATCCATAAACAACATGAAAATTATTCICTCTCCTGACATTGA - 27300
 - T Y G V R L S I K R T * K L F S S * H *
 - P M E L D Y P * N E H E N Y S L P D I D
 - L W S * I I H K T N M K I I L F L T L I
 27301 - TTGTATTACATCTTCCGAGCTATATCACTATCAGGAGTGTGTAGAGGTACGACTGTAC - 27360
 - L Y L H L A S Y I T I R S V L E V R L Y
 - C I Y I L R A I S L S G V C * R Y D C T
 - V F T S C E L Y H Y Q E C V R G T T V L
 27361 - TACTAAAGAACCTTGCCCATCAGGAACATACGAGGGCAATTCACCATTTCACCCCTTG - 27420
 - Y * K N L A H Q E H T R A I H H F T L L
 - T K R T L P I R N I R G Q F T I S P S C
 - L K E P C P S G T Y E G N S P F H P L A
 27421 - CTGACATAAATTGCACTAAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG - 27480
 - L T I N L H * L A L A H T L L L L V L T
 - * Q * I C T N L H * R T L C F C L C * R
 - D N K F A L T C T S T H F A F A C A D G
 27481 - GTACTCGACATACCTATCAGCTGCGTGAAGATCAGTTTCACCAAACITTCATCAGAC - 27540
 - V L D I P I S C V Q D Q F H Q N F S S D
 - Y S T Y L S A A C K I S F T K T F H Q T
 - T R H T Y Q L R A R S V S P K L F I R Q
 27541 - AAGAGGAGGTTCAACAAGAGCTCTACTCGCCACTTTTTCTCATGTGTGCTGCTCTAGTAT - 27600
 - K R R F N K S S T R H F F S L L L * Y
 - R G G S T R A L L A T F S H C C C S S I
 - E E V Q Q E L Y S P L F L I V A A L V F
 27601 - TTTTAACTTTGCTTCACCATTAAGAGAAGACAGAATGAATGAGCTCACTTTAATGA - 27660
 - F * Y F A S P L R E R Q N E * A H F N *
 - F N T L L H H * E K D R M N E L T L I D
 - L I L C F T I K R K T E * M S S L * L T
 27661 - CTTCTATTGTGCTTTTAGCCTTCTGCTTCTGCTTTTATAATGCTTATTATATT - 27720
 - L L F V L F S L S A I P C F N N A Y Y I
 - F Y L C F L A F L L F L V L I M L I I F
 - S I C A F * P F C Y S L F * * C L L Y F

图 11 续

27721 - TTGGTTTTCACCTCGAAATCCAGGATCTAGAAGAACCTTGACCAAAGTCTAAACGAACAT - 27780
 - L V F T R N P G S R R T L Y Q S L N E H
 - W F S L E I Q D L E E P C T K V * T N M
 - G F H S K S R I * K N L V P K S K R T *
 27781 - GAAACTTCTCATTGTTTTGACTTGTATTCTCTATGCAGTTGCATATGCACCTGTAGTACA - 27840
 - E T S H C F D L Y F S M Q L H M H C S T
 - K L L I V L T C I S L C S C I C T V V Q
 - N F S L F * L V F L Y A V A Y A L * Y S
 27841 - GCGCTGTGCATCTAATAAACCTCATGTGCTGAAGATCCCTTGTAAGGTACAACACTAGGG - 27900
 - A L C I * * T S C A * R S L * G T T L G
 - R C A S N K P H V L E D P C K V Q H * G
 - A V H L I N L M C L K I L V R Y N T R G
 27901 - GTAATACTTATAGCACTGCTTGGCTTGTGCTCTAGGAAAGGTTTTACCTTTTCATAGAT - 27960
 - V I L I A L L G F V L * E R F Y L F I D
 - * Y L * H C L A L C S R K G F T F S * M
 - N T Y S T A W L C A L G K V L P F H R W
 27961 - GGCACACTATGGTTCAAACATGCACACCTAATGTTACTATCAACTGTCAAGATCCAGCTG - 28020
 - G T L W F K H A H L M L L S T V K I Q L
 - A H Y G S N M H T * C Y Y Q L S R S S W
 - H T M V Q T C T P N V T I N C Q D P A G
 28021 - GTGGTGGCCTTATAGCTAGGTGTTGGTACCTTCATGAAGGTCACCAAAGTGTGCTGATT - 28080
 - V V R L * L G V G T F M K V T K L L H L
 - W C A Y S * V L V P S * R S P N C C I *
 - G A L I A R C W Y L H E G H Q T A A F R
 28081 - GAGACGTACTTGTGTTTTAAATAACGAACAATTAATGTTCTGATAATGGACCCCAA - 28140
 - E T Y L L F * I N E Q I K M S D N G P Q
 - R R T C C F K * T N K L K C L I M D P N
 - D V L V V L N K R T N * N V * * W T P I
 28141 - TCAAACCAACGTAGTGCACCCCGCATTACATTTGGTGGACCCACAGATTCAACTGACAAT - 28200
 - S N Q R S A P P R I T F G G P T D S T D N
 - Q T N V V P P A L H L V D P Q I Q L T I
 - K P T * C P P H Y I W W T H R F N * Q *
 28201 - AACAGAAATGGAGGACGCAATGGGGCAAGGCCAAAACAGCGCCGACCCCAAGGTTTACCC - 28260
 - N Q N G G R N G A R P K Q R R P Q G L P
 - T R M E D A M G Q G Q N S A D P K V Y P
 - P E W R T Q W G K A K T A P T P R F T Q
 28261 - AATAAFACTGCGTCTTGGTTTCACAGCTCTCACTCAGCATGGCAAGGAGGAAGTATGATTC - 28320
 - N N T A S W F T A L T Q H G K E E L R F
 - I I L R L G S Q L S L S M A R R N L D S
 - * Y C V L V H S S H S A W Q G G T * I P
 28321 - CCTCGAGGCCAGGGCGTTCCAATCAACACCAATAGTGGTCCAGATGACCAAATTTGGCTAC - 28380
 - P R G Q G V P I N T N S G P D D Q I G Y
 - L E A R A F Q S T P I V V Q M T K L A T
 - S R P G R S N Q H Q * W S R * P N W L L
 28381 - TACCGAAGAGCTACCCGACGAGTTCGTGGTGGTACGGCAAATGAAAGAGCTCAGCCCC - 28440
 - Y R R A T R R V R G G D G K M K E L S P
 - T E E L P D E F V V V T A K * K S S A P
 - P K S Y P T S S W W * R Q N E R A Q P Q
 28441 - AGATGGTACTTCTATTACCTAGGAACTGGCCAGAAGCTTCACTCCCTACGGGCTAAC - 28500
 - R W Y F Y Y L G T G P E A S L P Y G A N
 - D G T S I T * E L A Q K L H F P T A L T
 - M V L L L P R N W P R S F T S L R * Q
 28501 - AAAGAAGGCATCGTATGGGTTGCAACTGAGGAGCCTGAATACCCCAAAGACCACATT - 28560
 - K E G I V W V A T E G A L N T P K D H I
 - K K A S Y G L Q L R E P * I H P K T T L
 - R R H R M G C N * G S L E Y T Q R P H W

图 11 续

28561 - GGCACCCGCAATCCTAATAACAATGCTGCCACCGTGCTACAACCTCCTCAAGGAACAACA - 28620
 - G T R N P N N N A A T V L Q L P Q G T T
 - A P A I L I T M L P P C Y N F L K E Q H
 - H P Q S * * Q C C H R A T T S S R N N I
 28621 - TTGCCAAAAGGCTTCTACGCAGAGGGAAGCAGAGGCGGCAGTCAAGCCTTCTTCGCTCC - 28680
 - L P K G F Y A E G S R G G S Q A S S R S
 - C Q K A S T Q R E A E A A V K P L L A P
 - A K R L L R R G K Q R R Q S S L F S L L
 28681 - TCATCACGTAGTCGGGTAATTCAGAAATCAACTCCTGGCAGCAGTAGGGGAAATTCT - 28740
 - S S R S R G N S R N S T P G S S R G N S
 - H H V V A V I Q E I Q L L A A V G E I L
 - I T * S R * F K K F N S W Q Q * G K F S
 28741 - CCTGCTGAATGGCTAGCGGAGGTGGTGAAGTGCCTCGCGCTATTGCTGCTAGACAGA - 28800
 - P A R M A S G G G E T A L A L L L L D R
 - L L E W L A E V V K L P S R Y C C * T D
 - C S N G * R R W * N C P R A I A A R Q I
 28801 - TTGAACCGCTTGAGAGCAAAGTTTCTGGTAAAGGCCAACAAACAAGGCCAAAAGTGC - 28860
 - L N Q L E S K V S G K G Q Q Q Q G Q T V
 - * T S L R A K F L V K A N N N K A K L S
 - E P A * E Q S F W * R P T T T R P N C H
 28861 - ACTAAGAAATCTGCTGCTGAGGCATCTAAAAGCCTCGCCAAAACGTACTGCCACAAA - 28920
 - T K K S A A E A S K K P R Q K R T A T Q
 - L R N L L L R H L K S L A K N V L P Q N
 - * E I C C * G I * K A S P K T Y C H K T
 28921 - CAGTACAACGCTCACTCAAGCATTGGGAGACGTGGTCCAGAACAACCCCAAGGAAATTC - 28980
 - Q Y N V T Q A F G R R G P E Q T Q G N F
 - S T T S L K H L G D V V Q N K P K E I S
 - V Q R H S S I W E T W S R T N P R K F R
 28981 - GGGGACCAAGACCTAATCAGACAAGGAAGTATTACAACATTGGCCGCAATTCGCACAA - 29040
 - G D Q D L I R Q G T D Y K H W P Q I A Q
 - G T K T * S D K E L I T N I G R K L H N
 - G P R P N Q T R N * L Q T L A A N C T I
 29041 - TTTGCTCCAAGTGCCTCTGCATTCTTTGGAATGTCACGCATTGGCATGGAAGTCCACCT - 29100
 - F A P S A S A F F G M S R I G M E V T P
 - L L Q V P L H S L E C H A L A W K S H L
 - C S K C L C I L W N V T H W H G S H T F
 29101 - TCGGGAACATGGCTGACTTATCATGGAGCCATTAATTTGGATGACAAAGATCCACAATTC - 29160
 - S G T W L T Y H G A I K L D D K D P Q F
 - R E H G * L I M E P L N W M T K I H N S
 - G N M A D L S W S H * I G * Q R S T I Q
 29161 - AAAGACAACGTCATACTGCTGAACAAGCACATTGACGCATACAAAACATTCACCAACA - 29220
 - K D N V I L L N K H I D A Y K T F P P T
 - K T T S Y C * T S T L T H T K H S H Q Q
 - R Q R H T A E Q A H * R I Q N I P T N R
 29221 - GAGCCTAAAAGGACAAAAGAAAAGACTGATGAAGCTCAGCCTTTGCCGAGAGACAA - 29280
 - E P K K D K K K K T D E A Q P L P Q R Q
 - S L K R T K R K R L M K L S L C R R D K
 - A * K G Q K E K D * * S S A F A A E T K
 29281 - AAGAAGCAGCCCACTGTGACTCTTCTCCGCGGCTGACATGGATGATTTCTCCAGACAA - 29340
 - K K Q P T V T L L P A A D M D D F S R Q
 - R S S P L * L F L R L T W M I S P D N
 - E A A H C D S S S C G * H G * F L Q T T
 29341 - CTTCAAATTCATGAGTGGAGCTTCTGCTGATTCAACTCAGGCATAAACACTCATGATG - 29400
 - L Q N S M S G A S A D S T Q A * T L M M
 - F K I P * V E L L L I Q L R H K H S * *
 - S K F H E W S F C * F N S G I N T H D D

图 11 续

```

29401 - ACCACACAAGGCAGATGGGCTATGTRAACGTTTTCGCAATCCGTTTACGATACATAGTC - 29460
- T T Q G R W A M * T F S Q F R L R Y I V
- P H K A D G L C K R F R N S V Y D T * S
- H T R Q M G Y V N V F A I P F T I H S L
29461 - TACTCTTGTGCAGAATGAATTCTCGTAACTAAACAGCACAAGTAGGTTTAGTAACTTTA - 29520
- Y S C A E * I L V T K Q H K * V * L T L
- T L I V Q N E F S * L N S T S R F S * L *
- L L C R M N S R N * T A Q V G L V N F N
29521 - ATCTCACATAGCAATCTTTAATCAATGTGTAACATTAGGGAGGACTTGAAGAGCCACCA - 29580
- I S H S N L * S M C N I R E D L K E P P
- S H I A I F N Q C V T L G R T * K S H H
- L T * Q S L I N V * H * G G L E R A T T
29581 - CATTTCATCGAGGCCACGGGAGTACGATCGAGGGTACAGTGAATAATGCTAGGGAGAG - 29640
- R F H R G H A E Y D R G Y S E * C * G E
- I P I E A T R S T I E G T V N N A R E S
- F S S R P R G V R S R V Q * I M L G R A
29641 - CTGCCTATATGGAAGAGCCCTAATGTGTAATAATTTTAGTAGTGCTATCCCCATGTG - 29700
- L P I W K S P N V * N * F * * C Y P H V
- C L Y G R A L M C K I N F S S A I P M *
- A Y M E E P * C V K L I L V V L S P C D
29701 - ATTTTAATAGCTTCTTAGGAGAATGACAAAAAAAAAAAAAAAA - 29742
- I L I A S * E N D K K K K K X
- F * * L L R R M T K K K K X
- F N S F L G E * Q K K K K X

```

图 11 续

```

1 - TTTTTTTTTTTTTTGTTCATTCTCCTAAGAAGCTATTAAAATCACATGGGGATAGCACTA - 60
- F F F F F V I L L R S Y * N H M G I A L
- F F F F L S F S * E A I K I T W G * H Y
- F F F F C H S P K K L L K S H G D S T T
61 - CTAAAATTAATTTTACACATTAGGGCTCTCCATATAGGCAGCTCTCCCTAGCATTATC - 120
- L K L I L H I R A L P Y R Q L S L A L F
- * N * F Y T L G L F H I G S S P * H Y S
- K I N E T H * G S S I * A A L P S I I H
121 - ACTGTACCCCTCGATCGTACTCCGCGTGGCCCTCGATGAAAATGTGGTGGCTCTTTCAAGTC - 180
- T V P S I V L R V A S M K M W W L F Q V
- L Y P R S Y S A W P R * K C G G S F K S
- C T L D R T P R G L D E N V V A L S S P
181 - CTCCCTAATGTTACACATTGATTAAGATTGCTATGTGAGATTAAGTTAACTAACCTA - 240
- L P N V T H * L K I A M * D * S * L N L
- S L M L H I D * R L L C E I K V N * T Y
- P * C Y T L I K D C Y V R L K L T K P T
241 - CTGTGCTGTTTAGTTACGAGAATTCATTCTGCACAAGAGTAGACTATGTATCGTAAACG - 300
- L V L F S Y E N S F C T R V D Y V S * T
- L C C L V T R I H S A Q E * T M Y R K R
- C A V * L R E F I L H K S R L C I V N G
301 - GAATTGCGAAAACGTTTACATAGCCCATCGCCTTGTGGTTCATCATGAGTGTATTATGC - 360
- E L R K R L H S P S A L C G H H E C L C
- N C E N V Y I A H L P C V V I M S V Y A
- I A K T F T * P I C L V W S S * V F M P
361 - CTGAGTTGAATCAGCAGAAGCTCCACTCATGGAATTTGAAGTTGTCTGGAGAAATCATC - 420
- L S * I S R S S T H G I L K L S G E I I
- * V E S A E A P L M E F * S C L E K S S
- E L N Q Q K L H S W N F E V V W R N H P
421 - CATGTCAGCCGCAAGAAGAGTACAGTGGGCTGCTTCTTTGTCTCTGCGGCAAGG - 480
- H V S R R K K S H S G L L L L S L R Q R
- M S A A G R R V T V G C F F C L C G K G
- C Q P Q E E E S Q W A A S F V S A A K A
481 - CTGAGCTTCATCAGTCTTTTTCTTTTTGTCCTTTTTAGGCTCTGTTGGTGGGAATGTTTT - 540
- L S F I S L F L F V L F R L C W W E C F
- * A S S V F F F L S F L G S V G G N V L
- E L R Q S F S F C P F * A L L V G M F C
541 - GTATGCGTCAATGTGCTTGTTCAGCAGTATGACGTTGTCTTTGAATTGTGGATCTTTGTC - 600
- V C Y N V L V Q Q Y D V V F E L W I F V
- Y A S M C L F S S M T L S L N C G S L S
- M R Q C A C S A V * R C L * I V D L C H
601 - ATCCAATTTAATGGCTCCATGATAAGTCAGCCATGTTCCCGAAGGTGTGACTTCCATGCC - 660
- I Q F N G S M I S Q P C S R R C D F H A
- S N L M A P * * V S H V P E G V T S M P
- P I * W L H D K S A M F P K V * L P C Q
661 - AATGCGTGCATTCCAAGAATGCAGAGGCTTGGAGCAAATTGTGCAATTTGCGGCCA - 720
- N A * H S K E C R G T W S K L C N L R P
- M R D I P K N A E A L G A N C A I C G Q
- C V T F Q R M Q R H L E Q I V Q F A A N
721 - ATGTTTGAATCAGTTCCTTGTCTGATTAGGTCTTGGTCCCGAAATTCCTTGGGTTTG - 780
- M F V I S S L S D * V L V P E I S L G L
- C L * S V P C L I R S W S P K F P W V C
- V C N Q F L V * L G L G P R N F L G F V
781 - TTCTGGACCACGCTCCCAATGCTTGAGTGACGTTGTACTGTTTTGTGGCAGTACGTTT - 840
- F W T T S P K C L S D V V L F C G S T F
- S G P R L P N A * V T L Y C F V A V R F
- L D H V S Q M L E * R C T V L W Q Y V F

```

图 12

```

841 - TTGGCGAGGCTTTTTAGATGCCTCAGCAGCAGATTTCTTAGTGACAGTTTGGCCTTGTG - 900
- L A R L F R C L S S R F L S D S L A L L
- W R G F L D A S A A D F L V T V W P C C
- G E A F * M P Q Q Q I S * * Q F G L V V
901 - TTGTTGGCCTTTACCAGAACTTTGCTCTCAAGCTGGTTCAATCTGTCTAGCAGCAATAG - 960
- L L A F T R N F A L K L V Q S V * Q Q *
- C W P L P E T L L S S W F N L S S S N S
- V G L Y Q K L C S Q A G S I C L A A I A
961 - CGCGAGGGCAGTTTACCACCTCCGCTAGCCATTGAGCAGGAGAATTTCCCTACTGCT - 1020
- R E G S F T T S A S H S S R R I S P T A
- A R A V S P P P L A I R A G E F P L L L
- R G Q F H H L R * P F E Q E N F P Y C C
1021 - GCCAGGAGTTGAATTTCTGAATTACCGGACTACGTGATGAGGAGCGAGAAGAGCCTG - 1080
- A R S * I S * I T A T T * * G A R R G L
- P G V E F L E L P R L R D E E R E E A *
- Q E L N F L N Y R D Y V M R S E K R L D
1081 - ACTGCCGCCTCTGCTTCCCTCTCGGTAGAAGCCTTTTGGCAATGTTGTTCCCTTGAGGAAG - 1140
- T A A S A S L C V E A F W Q C C S L R K
- L P P L L P S A * K P F G N V V P * G S
- C R L C F P L R R S L L A M L F L E V
1141 - TTGTAGCACGGTGGCAGCATTGTTATTAGGATTGCGGGTGCCAATGTGGTCTTTGGGTG - 1200
- L * H G G S I V I R I A G A N V V F G C
- C S T V A A L L L G L R V P M W S L G V
- V A R W Q H C Y * D C G C Q C G L W V Y
1201 - ATTCAAGGCTCCCTCAGTTGCAACCCATACGATGCCTTCTTTGTTAGCGCGTAGGGAAG - 1260
- I Q G S L S C N P Y D A F F V S A V G K
- F K A P S V A T H T M P S L L A P * G S
- S R L P Q L Q P I R C L L C * R R R E V
1261 - TGAAGCTTCTGGGCCAGTTCCCTAGGTAATAGAAGTACCATCTGGGGCTGAGCTTTTCAT - 1320
- * S F W A S S * V I E V P S G A E L F H
- E A S G P V P R * * K Y H L G L S S F I
- K L L G Q F L G N R S T I W G * A L S F
1321 - TTTGCCGTCACCACCAGAACTCGTGGGTAGCTCTTCGGTAGTAGCCAATTTGGTCATC - 1380
- F A V T T T N S S G S S S V V A N L V I
- L P S P P R T R V A L R * * P I W S S
- C R H H H E L V G * L F G S S Q F G H L
1381 - TGACCACCTATTGGTGTGATTGGAACGCCCTGGCCTCGAGGGAATCTAAGTTCCTCCTT - 1440
- W T E I G V D W N A L A S R E S K F L L
- G P L L V L I G T P W P R G N L S S S L
- D H Y W C * L E R P G L E G I * V P P C
1441 - GCCATGGTGAGTGAGAGCTGTGAACCAAGACGAGTATTATTGGGTAAACCTTGGGGTGC - 1500
- A M L S E S C E P R R S I I G * T L G S
- P C * V R A V N Q D A V L L G K P W G R
- H A E * E L * T K T Q Y Y W V N L G V G
1501 - GCGCTGTTTGGCCTTGCCCATTTGCTCCTCCATTCTGGTTATTGTCAGTTGAATCTGT - 1560
- A L F W P C P I A S S I L V I V S * I C
- R C F G L A P L R P P F W L L S V E S V
- A V L A L P H C V L H S G Y C Q L N L W
1561 - GGGTCCACCAATGTAATCGGGGGGCACTACGTTGGTTGATTGGGGTCCATTATCAGA - 1620
- G S T K C N A G G T T L V * L G S I I R
- G P P N V M R G A L R W F D W G P L S D
- V H Q M * C G G H Y V G L I G V H Y Q T
1621 - CATTTTAATTTGTTGTTTTATTTAAACAACAAGTACGCTCTAATGCAGCAGTTTGGT - 1680
- H F N L F V Y L K Q Q V R L * M Q Q F G
- I L I C S F I * N N K Y V S K C S S L V
- F * F V R L F K T T S T S L N A A V W *

```

图 12 续

```

1681 - GACCTTCATGAAGGTACCAACACCTAGCTATAAGCGCACCACCAGCTGGATCTTGACAGT - 1740
- D L H E G T N T * L * A H H Q L D L D S
- T F M K V P T P S Y K R T T S W I L T V
- P S * R Y Q H L A I S A P P A G S * Q L
1741 - TGATAGTAACATTAGGTGTCATGTTTGAACCATAGTGTGCCATCTATGAAAAGGTAAAA - 1800
- * * * H * V C M F E P * C A I Y E K V K
- D S : N I R C A C L N H S V P S M K R * N
- I V T L G V H V * T I V C H L * K G K T
1801 - CCTTTCCTAGAGCACAAAGCCAAGCAGTGCCTATAAGTATTACCCCTAGTGTGTACCTTA - 1860
- P F L E H K A K Q C Y K Y Y P * C C T L
- L S * S T K P S S A I S I T P S V V P Y
- F P R A O S Q A V L * V L P L V L Y L T
1861 - CAAGGATCTTCAAGCACATGAGGTTTATTAGATGCACAGCGCTGTACTACAGTGCATATG - 1920
- Q G S S S T * G L L D A Q R C T T V H M
- K D : L Q A H E V Y * M H S A V L Q C I C
- R I F K H M R F I R C T A L Y Y S A Y A
1921 - CCACTGCATAGAGAAATACAAGTCAAACAATGAGAAGTTTCATGTTTCGTTTACACTTTG - 1980
- Q L H R E I Q V K T M R S F M F V * T L
- N C I E K Y K S K Q * E V S C S F R L W
- T A : * R N T S Q N N E K F H V R L D F G
1981 - GTACAAGGTTCTTCTAGATCCTGGATTCGAGTGAAAACCAAATATAATAAGCATTATT - 2040
- V Q G S S R S W I S S E N Q N I I S I I
- Y K V L L D P G F R V K T K I * * A L L
- T R F F * I L D F E * K P K Y N K H Y *
2041 - AAAACAAGGAATAGCAGAAAGGCTAAAAAGCACAAATAGAAGTCAATTAAGTGAGCTCA - 2100
- K T R N S R K A K K H K * K S I K V S S
- K Q G I A E R L K S T N R S Q L K * A H
- N K : E * Q K G * K A Q I E V N * S E L I
2101 - TTCAATCTGTCTTCTTCTTAATGGTGAAGCAAAGTATTAATAACTAGAGCAGCAACAA - 2160
- F I L S F S * W * S K V L K I L E Q Q Q
- S F C L S L N G E A K Y * K Y * S S N N
- H S : V F L L M V K Q S I K N T R A A T M
2161 - TGAGAAAAGTGGCGAGTAGAGCTCTGTGTAACCTCCTTGTCTGATGAAAAGTTTTG - 2220
- * E K V A S R A L V E P P L V * * K V L
- E K : K W R V E L L L N L L L S D E K F W
- R K : S G E * S S C * T S S C L M K S F G
2221 - GTGAAACTGATCTTGCACGCAGCTGATAGGTATGTCGAGTACCGTCAGCACAGCAAAG - 2280
- V K L I L H A A D R Y V E Y R Q H K Q K
- * N : * S C T Q L I G M S S T V S T S K S
- E T : D L A R S * * V C R V P S A Q A K A
2281 - CAAAGTGTGTGCTAGTGCAAGTTAGTGCAAAATTTATTGTCAGCAAGAGGGTGAATGGTG - 2340
- Q S V C * C K L V Q I Y C Q Q E G E M V
- K V C A S A S * C K F I V S K R V K W *
- K C V L V Q V S A N L L S A R G * N G E
2341 - AATTGCCCTCGTATGTTCCCTGATGGCAAGGTTCTTTTATAGTAGTACAGTCGTACCTCTAA - 2400
- N C P R M F L M G K V L L V V Q S Y L *
- I A : L V C S * W A R F F * * Y S R T S N
- L P S Y V P D G Q G S F S S T V V P L T
2401 - CACACTCCTGATAGTGATATAGCTCGCAAGATGTAATACAATCAATGTCAGGAAGAGAA - 2460
- H T P D S D I A R K M * I Q S M S G R E
- T L : L I V I * L A R C K Y N Q C Q E N
- H S : * * * Y S S Q D V N T I N V R K R I
2461 - TAATTTTCATGTTTCGTTTATGGAATAATCTAACTCCATAGGTTCTTCATCATCTAACTCC - 2520
- * F S C S F Y G * S N S I G S S S S S N S
- N F : H V R F M D N L T P * V L H H L T P
- I F M F V L W I I * L H R F F I I * L R

```

图 12 续

```

2521 - GAATAATTCTTCTTAGTTAGAGGCTTAAATAATTGTCTCACTATTGAACTTATTATAACG - 2580
- E * F F L V R G L N N C L T I E L I I T
- N N S S * L E A * I I V S L L N L L * R
- I I L L S * R L K * L S H Y * T Y Y N V
2581 - TCAAGATTCCAAATAGCAATCCTGAAAGTCCTCATAATGATAATCAATATCTCTGCTATT - 2640
- S R F Q I A I L K V L I M C I N I S A I
- Q D S K * Q S * K S S * * * S I S L L L
- K I P N S N P E S P H N D N Q Y L C Y C
2641 - GTAACCTGGAAGTCAACAAGATGAACATCTGTTGTCACTTACTGTACTAGCAAAGCAAT - 2700
- V T W K S T R * N I C C H L L Y * Q S N
- * P G S Q Q D E T S V V T Y C T S K A I
- N L E V N K M K H L L S L T V L A K Q Y
2701 - ATTGTCGTTGCTACCGCGTGGTCTGTATTAAATTTATAGTTTCCAATACGGTAGCGGTT - 2760
- I V V A T G V V C I * F I V S N T V A V
- L S L L P A W S V F N L * F P I R * R L
- C R C Y R R G L Y L I Y S F Q Y G S G C
2761 - GTATGCAGCAAACCTGAATCAGTGCCACAGCTGCGACGCTCCTAATTTGTAATAAGA - 2820
- V C S K T * I S A Y T L R R S * F V I R
- Y A A K P E S V P T R C D A P N L * * E
- M Q Q N L N Q C L H A A T L L I C N K K
2821 - AAGCGTTCCGTGATGTAGCCACAGTGATCTCTTTTGGCAGGTCCTTAATGTCACAGCGCCC - 2880
- K R S * C S H S D L F W Q V L N V T A P
- S V R D V A T V I S F G R S L M S Q R P
- A F V M * P Q * S L L A G P * C H S R L
2881 - TAGGGAGTGTCCGGCCATTCCGAAGTGACCACGAATGATCACAGCACC AATGACAAGTTC - 2940
- * G Y S G H S Q V T T N D H S T N D K F
- R E C P A I R K * P R M I T A P M T S S
- G S V R P F A S D H E * S Q H Q * Q V H
2941 - ACTTCCATGAGCGGTCTGGTCACAAATGTCACCGGAGAGGCACATGAGAAGAATGTT - 3000
- T F H E R S G H N C P P E R H I E K N V
- L S M S G L V T I V P R R G T L R R M F
- F P * A V W S Q L S P G E A H * E E C L
3001 - TGTTTCTGGGTGAATGACCACATTGAGCGGGTACGAGCAAACAGCCTGAAGGAAGCAAC - 3060
- C F W V E * P H * A G T S K Q P E G S N
- V S G L N D H I E R V R A N S L K E A T
- F L G * M T T L S G Y E Q T A * R K Q R
3061 - GAAGTAGCTAAGCCACATCAAGCCTACAATAACAAGCCATTGCAATCGCAATCCCGCCAGT - 3120
- E V A K P H Q A Y N T S H C N R N P A S
- K * L S H I K P T I Q A I A I A I P P V
- S S * A T S S L Q Y K P L Q S Q S R Q S
3121 - CACCCAATTAATTCTGTAGACAACAGCAAGCAGCAAAAACAAGCAAGTGTACTGGCCACAA - 3180
- H P I N S V D N S K H K T S K C Y W P Q
- T Q L I L * T T A S T K Q A S V T G H K
- P N * F C R Q Q A Q N K Q V L L A T R
3181 - GAGCCAGAGGAAAACAAGCTTTATTATGTACAAAACCTGTTCCGATTAGAATAGGCAAA - 3240
- E P E E N K L Y Y V Q K P V P I R I G K
- S Q R K T S F I M Y K N L F R L E * A N
- A R G K Q A L L C T K T C S D * N R Q I
3241 - TTGTAGTACATAATCCAGGCTAGGAATAGGAAACCTATACTAGGTTCCATTGTTCCAG - 3300
- L * * H N P G * E * E T Y Y * V P L F Q
- C S N I I Q A R N R K P I T R F H C S R
- V V T * S R L G I G N L L L G S I V P G
3301 - GAGTTGTTTAAAGCTCCTCAACGSTAATAGTACCGTTGTCTGCCATGATAAGCAATGTAA - 3360
- E L F K L L N G N S T V V C H D K Q C *
- S C L S S S T V I V P L S A M I S N V K
- V V * A P Q R * * Y R C L P * * A M L K

```

图 12 续


```

3361 - ACTTCCAACAGAATAATAATAATAGTTAGTTTCGTTTAGACCAGAAGATCAGGAACTCCT - 3420
- S S K Q N N N N S * F V * T R R S G T P
- V P I N R I I I I V S S F R P E D Q E L L
- F Q T E * * * * L V R L D Q K I R N S F
3421 - TCAGAAGAGTTTCAGATTTTAAACACGCGAGTAGACGTAACCCTGGTTTACTAAATC - 3480
- S E E F R F L T R E * T * T V G F T K L
- Q K S S D F * H A S R R K P L V L L N S
- R R V Q I F N T R V D V N R W F Y * T H
3481 - ACGTTAACAAATATTGCAGCAGTACGCACACAATCGAAGCGCAGTAAGGATGGCTAGTGTG - 3540
- T L T I L Q Q Y A H N R S A V R M A S V
- R * Q Y C S S T H T I E A Q * G W L V *
- V N N I A A V R T Q S K R S K D G * C D
3541 - ACTAGCAAGAAATACCAGAAAGCAAGAAAAGAAGTACGCTATTAATTAACGTACCT - 3600
- T S K N T T K A R K R S T L L T I N V P
- L A R I P R K Q E K E V R Y * L L T Y L
- * Q E Y H E S K K K Y A I N Y * R T C
3601 - GTTCTTCCGAAACGAATGAGTACATAAGTTTCGTACTCACTTCTTGTGCTTACAAAGGC - 3660
- V S S E T N E Y I S S Y S L S C A Y K G
- F L P K R M S T * V R T H F L V L T K A
- F F R N E * V H K F V L T F L C L Q R H
3661 - ACGCTAGTAGTCGTCGTCGGCTCATATAATGGATCCATTGCTGGATTAGCAACTCCT - 3720
- T L V V V V G S S * I G S I A G L A T P .
- R * S S S A H H K L D P L L D * Q L L
- A S S R R R L I I N W I H C W I S N S *
3721 - GAAGAGCCGTCGATTGTGTATTGACATTGGTGGGTCTTTAACAAGCTGTAAAG - 3780
- E E P S I V C I C T F G G S L T S L L K
- K S R R L C V F A H S V G L * Q A C * R
- R A V D C V Y L H I R W V F N K L V K D
3781 - ATGAAGAATGTAGCATTTTCAATACCAGTGTCTGTAGTAATTTGTGTAGACTCAAGCTGG - 3840
- M K N V A F S I P V S V V I C V D S S W
- * R M * H F Q Y Q C L * * F V * T Q A G
- E E C S I F N T S V C S N L C R L K L V
3841 - TAGTAAACTTCGGTGAATAGCCATGTACAACGACATAGTCTTTAACACCTGAGTGCCTA - 3900
- * * T S V K * P C T T T * S L T P E C L
- S K L R * N S H V Q R H S L * H L S A Y
- V N F G E I A M Y N D I V F N T * V P I
3901 - TCCTCAGAATAACCACCAATTTGGTAGTCTTCTTTGAGTTTGGTGTGAAATGCCGTCA - 3960
- S S E * P P I W * S S L S F G V E M P S
- P Q N N H Q F G S L L * V L V L K C R H
- L R I T T N L V V F F E F W C * N A V T
3961 - CCTTCAGTAACGACAATTGTATCTGTGACACTGTTATATGGTATACAGTAGTCATAGTTA - 4020
- P S V T T I V S V T L L Y G I Q * S * L
- L Q * R Q L Y L * H C Y M V Y S S H S Y
- F S N D N C I C D T V I W Y T V V I V M
4021 - TGTGTGTGCCAGCAAACAAGTAGTTGGCATCATAAAGTAATGGGTCTTGGATTGTCAC - 4080
- C V C Q Q T K * L A S * S N G F L D L H
- V C A S K Q S S W H H K V M G S W I C T
- C V P A N K V V G I I K * W V L G F A L
4081 - TTCCAACAAGCCAACATCTCATAATAATTCTACATGCGTTGATGCATTGTAGAAAATAT - 4140
- F Q Q S Q H L I I I L H A L M H C R K Y
- S N K A N I S * * F Y M R * C I V E N I
- P T K P T S H N S T C V D A L * K I Y
4141 - ATCAAGGCATAGAGGTACAAAATTCGCCTTACCTGCAGCGACAAGCAAAGATGT - 4200
- I K A * R Y K N C A S L P A A T S K R C
- S R H R G T K I A P P Y L Q R Q A K D V
- Q G I E V Q K L R L L T C S D K Q K M *

```

图 12 续

```

4201 - GAATAGATGGTRACAAATAGCAGCAGTAAATTGCAAATGAACTGGAAGCCCTTATAAAGG - 4260
- E * M V T N S S S K L Q M N W K P L * R
- N R W * Q I A A V N C K * T G S P Y K G
- I D G N K * Q Q * I A N E L E A L I K G
4261 - GCTAGCTGCCATCTTTTATTGAGCGCAATTATTTGGTAGCGCTCTGAAAAACAGCAAGA - 4320
- A S C H L L L S A I I L V A L * K T A R
- L A A I F Y * A Q L F W * R S E K Q Q E
- * L P S F I E R N Y F G S A L K N S K K
4321 - AATGCAACGCCAATAACAAGCCATCCGAAAGGGAGTGAGGCTTGTAGCGGTATCGTTGCT - 4380
- N A T P I T S H P K G S E A C S G I V A
- M Q R Q * Q A I R K G V R L V A V S L L
- C N A N N K P S E R E * G L * R Y R C C
4381 - GTAGCATGAACAGTACTTGCAGGAGAAGCATTGTCAATTTTACTGGCTGTGCAGTAATT - 4440
- V A * T V L A G E A L S I F T G C A V I
- * H E Q Y L Q E K H C Q F L L A V Q * L
- S M N S T C R R S I V N F Y W L C S N *
4441 - GATCCAAGAGTAAAAATCTCATAAACAATCCATAAGTTCGTTTATGTGTAATGTAATT - 4500
- D P R V K N L I N K S I S S F M C N V I
- I Q E * K I S * T N P * V R L C V M * F
- S K S K K S H K Q I H K F V Y V * C N L
4501 - TGACACCCTTGAGAACTGGCTCAGAGTCATCCTCATCAACTTGCAGCAAGAACCACAAG - 4560
- * H P * E L A Q S H P H Q T C S K N H K
- D T L E N W L R V I L I K L A A R T R
- T P L R T G S E S S S S N L Q Q E P Q E
4561 - AGCATGCACCCTTGAGGCAACTGCAACAAGTTCATGCAACAAGCAAGATTGTAACCA - 4620
- S M H P * G N C N N * S C N K A R L * P
- A C T L E A T A T T S H A T K Q D C N H
- H A P L R Q L Q Q L V M Q Q S K I V T M
4621 - TGACGATGGCAATTAGTCCAGCAATGAAGCCGAGCCAACATACCAAGGCCATTAATAT - 4680
- * R W Q L V Q Q * S R A K H T K A I * Y
- D D G N * S S N E A E P N I P R P F N I
- T M A I S P A M K P S Q T Y Q G H L I Y
4681 - ATTGCTCATATTTCCCAATTCTTGAAGGTCATGAGTGATTCAATTAATTTTAGCGA - 4740
- I A H I F P I L E G Q * V I H L N F * R
- L L I F S Q F L K V N E * F I * I F S D
- C S Y F P N S * R S M S D S F K F L A T
4741 - CCTCATTGAGCGGTCATTTCTTTTGAATGTTGACGACAGAAGCGTTAATGCCTGAAA - 4800
- P H * G G Q F L F E C * R Q K R * C L K
- L I E A V N F F L N V D D R S V N A * N
- S L R R S I S F * M L T T E A L M P E M
4801 - TGTCGCCAAGATCAACATCTGGTGATGTATGATTTTGAAGTACTGTCCAGCTCTTCT - 4860
- C R Q D Q H L V M Y D F * S T C P A L L
- V A K I N I W * C M I F E V L V Q L S F F
- S P R S T S G D V * F L K Y L S S S L
4861 - TGAATGAGTCAAGCTCAGGTTGCAGAGGATCAFAACTGTGTTAATGATGCCAATAA - 4920
- * M S Q A Q V A E D H K L C C * * C Q *
- E * V K L R L Q R I I N C V V N D A N N
- N E S S S G C R G S * T V L L M M P I T
4921 - CGACATCACAATTCCTGAGACAAATGTATTGCTGTAGTAATTATTGTTGGAGAAAAGA - 4980
- R H H N F L R Q M Y C L * * L F V E K R
- D I T I S * D K C I V C S N Y L W R K E
- T S Q F P E T N V L S V V I I C G E K K
4981 - AGTTCCTCTGTGTAATAAACAAGAGTGCCATTAACACAAAACACCTTCACGAGGGA - 5040
- S S S V * * T K K C H * T Q K H L H E G
- V P L C N K P R S A I K H K N T F T R E
- F L C V I N Q E V P L N T K T P S R G K

```

图 12 续

5041 - AGTATGCTTTGCCTTCATGACAAAATTGCTGGCGCTGTGGTGAAGTTCCTCTCCTGGGATG - 5100
 - S M L C L H D K L L A L W * S S S P G M
 - V C F A F M T N C W R C G E V P L L G W
 - Y A L P S * Q I A G A V V K F L S W D G
 5101 - GCACATACGTGACATGTAGGAAGACAACCCATGCGGGGCTGCTTGTGGGAAGGACATAA - 5160
 - A H T * H V G R Q H H A G L L V G R T *
 - H I R D M * E D N T M R G C L W E G H K
 - T Y V T C R K T T P C G A A C G K D I R
 5161 - GGTGGTAGCCCTTTCCACAAAAGTCAACTCTTTTTGATTGTCCAAGAACACACTCAGACA - 5220
 - G G S P F H K S Q L F L I V Q E H T Q T
 - V V A L S T K V N S F * L S K N T L R H
 - W * P P P Q K S T L F D C P R T H S D I
 5221 - TTTAGTAGCAGCAAGATTAGCAGAAGCCCTGATTTTCAGCAGCCCTGATTAGTTGTTGTG - 5280
 - F * * Q Q D * Q K P * F Q Q P * L V V V
 - F S S S K I S R S P D F S S P D * L V L C
 - L V A A R L A E A L I S A A L I S C C V
 5281 - TTACATAGGTTGAAGGCTTTGAAGTCTGCCTGTAATTAACCTGTCAATTTGTACCTCCG - 5340
 - L H R F E G F E V C L * L T C Q F V P P
 - Y I G L K A L K S A C N * P V N L Y L R
 - T * V * R L * S L P V I N L S I C T S A
 5341 - CCTCGACTTTATCAAGTCGCGAAGGATATCATTAGCACACTTGAATTCACCAAAT - 5400
 - P R L Y Q V A K G Y H L A H L K L H Q N
 - L D F I K S R K D I I * H T * N C T K I
 - S T L S S R E R I S F S T L E I A P K L
 5401 - TAGAGCTAAGTTGTTTAAACAAGTGTGTTAATGCTTGAGCATTCTGGTTAACAACGTCTT - 5460
 - * S * V V * Q V C L M L E H S G * Q R L
 - R A K L F N K C V * C L S I L V N N V L
 - E L S C L T S V F N A * A F W L T T S C
 5461 - GCAGCTTGCCCAATGCAGTTGATGTTGTTGTAAGTGAATTCTTGAATTTGACTAATCGCCT - 5520
 - A A C P M Q L M L L * V I L E F D * S P
 - Q L A Q C S * C C C K * F L N L T N R L
 - S L P N A V D V V V S D S * I * L I A L
 5521 - TGTTAAATTGTTGGCGATTGTTTTGGTTCTCATAGAGAACATTTGGGTAACCTCAA - 5580
 - C * I G W R F V F G S H R E H F G * L Q
 - V K L V G D L F L V L I E N I L G N S N
 - L N W L A I C F W F S * R T F W V T P M
 5581 - TGCCATTGAACCTATATGCCATTTGCATAGCAAAAGGTATTTGAAGAGCAGCGCCAGCAC - 5640
 - C H * T Y M P F A * Q K V F E E Q R Q H
 - A I E P I C H L H S K R Y L K S S A S T
 - P L N L Y A I C I A K G I * R A A P A P
 5641 - CAAATGTCCATCCAGCAGTGGCAGTACCCTAAGTAGAGCAGCAGTGTAGGCAGCAATCA - 5700
 - Q M S I Q Q W Q Y H * L E Q Q C R Q Q S
 - K C P S S S G S T T N * S S S V G S N H
 - N V H P A V A V P L T R A A V * A A I I
 5701 - TATCATCAGTGAGCAGAGGTGGCAACACTGTAAGTCCATTGAACTTCTGCGCACAAATGA - 5760
 - Y H Q * A E V A T L * V H * T S A H K *
 - I I S E Q R W Q H C K S I E L L R T N E
 - S S V S R G G N T V S P L N F C A Q M R
 5761 - GATCTTAGCATTAAATACCTTAGGCATTGCCATATTGCTTCATGAAGCCAGCATCAG - 5820
 - D L * H * Y H L G I R H I A S * S Q H Q
 - I S S I N I T * A F A I L L H E A S I S
 - S L A L I S P R H S P Y C F M K P A S A
 5821 - CGAGTGTACCTTATTAAGAGCAAGTCCCTCAATAAAAGACCTCTTAGTTGGCTTAGAG - 5880
 - R V S P Y * R A S P Q * K T S * L A L E
 - E C H L I K E Q V L N K R P L S W L * R
 - S V T L L K S K S S I K D L L V G F R G

图 12 续

```

5881 - GGTCAGGTAATATTTGTGAAAAATAAAACCACCAAAATATTTCAAAGTTGGGGTTTGT - 5940
- G Q V I F V K N * N H Q N I S K L G F C
- V R . * Y L * K I K T T K I F Q S W G F V
- S G N I C E K L K P P K Y F K V G V L Y
5941 - ACATTTGTTTACTTGTAGCGAACACTTCACGTGTGTTGCGATCCTGTTGAGCAGCAATAC - 6000
- T F V * L E R T L H V C C D P V Q Q Q Y
- H L F D L S E H F T C V A I L F S S N T
- I C L T * A N T S R V L R S C S A A I P
6001 - CTGAGAGTGCAGATTTAGTTGTGTGCAAAAGCTACCATATTGGAGAAGCAAATTAGCAC - 6060
- L R V H D L V V C K S Y H I G E A N * H
- * E C T I * L C A K A T I L E K Q I S T
- E S A R F S C V Q K L P Y W R S K L A H
6061 - ATTCAGTAGAATCTCCGCAGATGTACATATTCAATCTACGGAGGTTTTAGCCATAGAAA - 6120
- I Q * N L R R C T Y Y N L R R F * P * K
- F S R I S A D V H I T I Y G G F S H R N
- S V E S P Q M Y I L Q S T E V L A I E T
6121 - CAGGCATTACTTCTGTAGTAATGCTAATTGAAAAGTTAGTAGGTATAGCAATGGTGTAT - 6180
- Q A L L L * * C * L K S * * V * Q W C Y
- R H Y F C S N A N * K V S R Y S N G V I
- G I . T S V V M L I E K L V G I A M V L L
6181 - TAGAGTAAGCAATTGAACTATCAGCACCTAAAGACATAGTATAAGCCACAATAGATTTTT - 6240
- * S K Q L N Y Q H L K T * Y K P Q * I F
- R V S N * T I S T * R H S I S H N R F L
- E * A I E L S A P K D I V * A T I D F W
6241 - GGCTAGTACTCGTAATAAGAACTGTATGGTAACTAGCACAAATGCCAGCTCCAATAG - 6300
- G * Y Y V I K K L Y G N * H K C Q L Q *
- A S T T * * R N C M V T S T N A S S N R
- L V L R N K E T V W * L A Q M P A P I G
6301 - GAATGTCGCACTCATAAGAAGTGTGCGACATGTCAGCTCCTATAAGACAGCCTGCTTGAG - 6360
- E C R T H K K C R H A Q L L * D S L L E
- N V A L I R S V D M L S S Y K T A C L S
- M S H S * E V S T C S A P I R Q P A * V
6361 - TCTGGAATACATTGTTTCCAGTAGAATATATGCGCCAAGCTGGTGTGAGTTGATCTGCAT - 6420
- S G I H C F Q * N I C A K L V * V D L H
- L E Y I V S S R I Y A P S W C E L I C M
- W N T L F P V E Y M R Q A G V S * S A *
6421 - GAATGCTGTAGAAACATCAGTGCAGTTAACATCTTGATATAGAACAGCAACTTCAGATG - 6480
- E L L * K H Q C S * H L D I E Q Q L Q M
- N C C R N I S A V N I L I * N S N F R *
- I A V E T S V Q L T S * Y R T A T S D E
6481 - AAGCATTGTTCCAGGTGTAATTACACTTACACCCCAAGAGCAAGGTGAAATGTCTA - 6540
- K H L F Q V * L H L H P Q K S K V K C L
- S I C S R C N Y T Y T P K R A R * N V *
- A F V P G V I T L T P P K E Q G E M S N
6541 - ATATTTAGATGTTTGGATCTCGAACGGAATCAGTGAATCAGAAACATCAGGCCAA - 6600
- I F Q M F * D L E R N Q * N Q K H H G Q
- Y F R C F R I S N G I S E I R N I T A K
- I S D V L G S R T E S V K S E T S R P N
6601 - ATGTTGAAATGGTTGAATCTCTTTGAAGAAGGAGTTAACACACCAGTACCAGTGAGTC - 6660
- I V E M V E I S L K K E L T H Q Y Q * V
- L L K W L K S L * R R S * H T S T S E S
- C * N G * N L F E E G V N T P V P V S P
6661 - CATTAAAATTAATGACACACTGTTCTTAATAAGGTCAGTGGATAATTTGGTCCAC - 6720
- H * N * N * H T G S * * G Q W I I L V H
- I K I K I D T L V L N K V S G * F W S T
- L K L K L T H W F L I R S V D N F G P Q

```

图 12 续

```

6721 - AAACCGTGGCCGGTGCATTTAAAAGTTCAAAAGAAAGTACTACAACCTCTGTAAGGTGGT - 6780
- K P W P V H L K V Q K K V L Q L C K V G
- N R G R C I * K F K R K Y Y N S V R L V
- T V A G A F K S S K E S T T T L * G W *
6781 - AGCCAATGCCAGTAGTGGTGTAAAAACCATAATCATTTAATGGCCAATAACAATTAAGAG - 6840
- S Q C Q * W C K N H N H L M A N N N * E
- A N A S S G V K T I I I * W P I T I K S
- P M P V V V * K P * S F N G Q * Q L R A
6841 - CAGGTGGGGTGCAAGGTTTGCCATCAGGGGAGAAAGGCACATTAGATATGTCTCTCTCAA - 6900
- Q V G C K V C H Q G R K A H * I C L S Q
- R W G A R F A I R G E R H I R Y V S L K
- G G V Q G L P S G E K G T L D M S L S K
6901 - AGGGCCTAAGCTTGCCATGTCTAAGATACCTATATTTATAATTATAATTACCAGTTGAAG - 6960
- R A * A C H V * D T Y I Y N Y N Y Q L K
- G P . K L A M S K I P I F I I I I T S * S
- G L S L P C L R Y L Y L * L * L P V E V
6961 - TAGCATCAATGTTCTTAGTATTCCAAGCAAGGACACAACCCATGAAATCATCTGGCAATT - 7020
- * H Q C S * Y S K Q G H N P * N H L A I
- S I N V P S I P S K D T T H E I I W Q F
- A S M F L V F Q A R T Q P M K S S G N L
7021 - TATAATTATAATCAGCAATAACCCAGTTGTCTCTGGCGCTATTGTCTTACATCATCTC - 7080
- Y N Y N Q Q * H Q F V L A L F V L H H L
- I I I I S N N T S L S W R Y L S Y I S
- * L * S A I T P V C P G A I C L T S S P
7081 - CCTTGACTACAAAAGAATCTGCATAGACATTGGAGAAGCAAAGATCATTCAACTTAGTGG - 7140
- P * L Q K N L H R H W R S K D H S T * W
- L D Y K R I C I D I G E A K I I Q L S G
- L T T K E S A * T L E K Q R S F N L V A
7141 - CAGAAACGCCATAGCACTTAAAGGTTGAAAAAATGTTGAGTTGTAGAGCACAGAGTAAT - 7200
- Q K R H S T * R L K K M L S C R A Q S N
- R N A I A L K G * K K C * V V E H R V I
- E T P * H L K V E K N V E L * S T E * S
7201 - CAGCAACACAATTAGAAATTTTTTCTCTCCCATGCATAGACAGGGAATTTAGTAG - 7260
- Q Q H N * K F F F S P M H R Q K G I * *
- S N T I R N F F S L P C I D R R E F S S
- A T Q L E I F F L S H A * T E G N L V A
7261 - CATTAAAACCTCTCCAAAAGGACACAAGTTTGTAAATATTAGGGAATCTCACACATCTC - 7320
- H * K P L Q K D T S L * Y * G I S Q H L
- I K N L S K R T Q V C N I R E S H N I S
- L K T S P K G H K F V I L G N L T T S P
7321 - CTGAGGGGAACAACCTGAAATTAGAGGTCTGGTAAATTCCTTTGTCAATCTCAAAGCTCT - 7380
- L R E Q P * N * R S G K F L C Q S Q S S
- * G N N P E I R G L V N S F V N L K A L
- E G T T L K L E V W * I P L S I S K L L
7381 - TAACAGAGCATTGAGTTCAGCAAGTGGATTTGAGAACAATCAACAGCATCTCTGATTG - 7440
- * Q S I * V Q Q V D F E N N Q Q H L * L
- N R A F E F S K W I L R T I N S I C D C
- T E H L S S A S G F * E Q S T A S V I V
7441 - TACCATTTTCATCACTTGAGCATAAATGTAGTTGGCTTTAAATAGCCAACAAAATAGG - 7500
- Y H F H H T * A * M * L A L N S Q Q N R
- T I F I I L E H K C S W L * I A N K I G
- P F S S Y L S I N V V G F K * P T K * A
7501 - CTGCAGCTGACGTGCCCAATGTCTTGAGCAGGTGAAAAGGCTGTAGAATGGCTCTAA - 7560
- L Q L T C P K C L E Q V K R L * E W L *
- C S * R A P N V L S R * K G C K N G S K
- A A D V P Q M S * A G E K A V R M A L K

```

图 12 续

```

7561 - AATTTGTAATGTTAATAACCAAGAGGCAACTTAAAAATAGGTTTCAAAGTGTTAAAACCAG - 7620
- N L * C * Y Q E A T * K * V S K C * N Q
- I C N V N T K R Q L K N R F Q S V K T R
- F V M L I P R G N L K I G F K V L K P E
7621 - AAGGTAGATCACGAACTACATCTATAGGTTGATAGCCCTTATAAACATAGAGAACCAT - 7680
- K V D H E L H L * V D S P Y K H R E T H
- R * I T N Y I Y R L I A L I N I E K P I
- G R S R T T S I G * * P L * T * R N P S
7681 - CTTTATTTTAAACACAAACTCTCGTAAGTGTTAAAATTACCTGACTTTTCTGAACAT - 7740
- L Y F * T Q T L V S V * N Y L T F L K H
- F I F K H K L S * V F K I T * L F * N I
- L F L N T N S R K C L K L P D F S E T S
7741 - CAAGCGAAAAGGCATCAGATATGTAAGTGCATTAATGCATTATCGAATATCA - 7800
- Q A K R H Q I C T R K C N * M H Y R I S
- K R K G I R Y V L E S A I K C I I E Y H
- S E K A S D M Y S K V Q L N A L S N I I
7801 - TAGTATGTGTCTGTGTACCCATGGGTTTAGAAACAGCAAAGAAAGGGTTGTACACAAAT - 7860
- * Y V S V Y P W V * K Q Q R K G C H T I
- S M C L C T H G F R N S K E R V V T Q F
- V C V C V P M G L E T A K K G L S H N S
7861 - CAAAGTTACATGCTCGTATAACAACATTAGTAGAATTGTTAATAATAATCACCGACTGTG - 7920
- Q S Y M L V * Q H * * N C * * * S P T V
- K V T C S Y N N I S R I V N N N H R L *
- K L H A R I T T L V E L L I I I T D C D
7921 - ACTTGTGTTCATGGTAGAACCAAAAACCCAAACCACGGACAACATTTGATTTCTCTGTGG - 7980
- T C C S W * N Q K P N H G Q H L I S L W
- L V V H G R T K N P T T D N I * F L C G
- L L F M V E P K T Q P R T T F D F S V A
7981 - CAGCAAAATAAATACCATCCTTAAAGGTATGACAGGGTTGCCAAACGTATGATTAATAG - 8040
- Q Q N K Y H P * K V * Q G C Q T Y D * *
- S K I N T I L K R Y D R V A K R M I N S
- A K * I P S L K G M T G L P N V * L I V
8041 - TATGAAACCTGTAACATTAGAATAAATGGAAGAAATAAATCCTGAGTTAAATAAGAG - 8100
- Y E T L * H * N K M E E I N P E L N K E
- M K P C N I R I K W K K * I L S * I K S
- * N P V T L E * N G R N K S * V K * R V
8101 - TGCTGATCTAAAAATTTTCATCAGGATAGTAAACCCCTCATAGATGAAGTATGTTGAG - 8160
- C L I * K F H Q D S K P P S * M K Y V E
- V * S K N F I R I V N P P H R * S M L S
- S D L K I S S G * * T P L I D E V C * V
8161 - TGTAATTAGGAGCTTGAACATCATCAAAGTGGTGCACCGGTCAAGGTCCTACCCTAG - 8220
- C N * E L E H H Q K W C T G Q G H Y H *
- V I R S L N I I K S G A P V K V T T T S
- * L G A * T S S K V V H R S R S L P L V
8221 - TGAGAGTAAGAAATAAAGAAAATAAACATGTTTCGTTTGTGTTAACAAGAATATCAC - 8280
- * E * E I I R K * T C S F S C * Q E Y H
- E S K K * * E N K H V R L V V N K N I T
- R V R N N K K I N M F V * L L T R I S L
8281 - TTGAAACCACAACCTCTGTTGTTTCTCTAATGATAAGCCTACCTTTTCCAGAAGAGAAT - 8340
- L K P Q L C C F L * * * A Y L F P E E N
- * N H N S V V F S N D K P T F F Q K R I
- E T T T L L F S L M I S L P F S R R E *
8341 - AAATCATATCATTGATTCTCCTTAAGAGACATTACAGCAGTTCCTCTTAATTTAA - 8400
- K S Y H * F D S P * E T L Q Q F L L I *
- N H I I D L I L L K R H Y S S S S * F K
- I I S L I * F S L R D I T A V P L N L R

```

图 12 续

```

8401 - GAGGAAATTTGCTCATGTCAAAGAGTGAATAGGAAGCAACTGGATAGGATTTGTGTTCC - 8460
- E E I C S C Q R V N R K T T G * D L C S
- R K F A H V K E * I G R Q L D R I C V P
- G N L L M S K S E * E D N W I G F V F L
8461 - TCCAGAAAATGTAGTTAGCATGCATGGTATAGCCATCAATTTGTTCCCTCGGCTGCCAA - 8520
- S R K C S * H A W Y S H Q F V P S A C Q
- P E N V V S M H G I A I N L F L R L A K
- Q K M * L A C M V * P S I C S F G L P R
8521 - GATAGTTAGCCCCAATTA AAAATGCTTCCGATGATGATGCATTTACATTTGTAACAAAAG - 8580
- D S * P Q L K M L P M M M H L H L * Q K
- I V S P N * K C F R * * C I Y I C N K S
- * L A P I K N A S D D D A F T F V T K A
8581 - CTGTCCACCATGAGAAATGGCCATAAGCTTAAAGGTGAGCATTCCAAGAAGTCTCTG - 8640
- L S T M R N G P * A C K G Q H S K N A L
- C P P * E M A H K L V K V S I P R M L C
- V H H E K W P I S L * R S A F Q E C S V
8641 - TTATCTTTACAGCTATAGAACCACCCAGGGCTAGTTTTTGGCTTTATAAATCCACACAGAT - 8700
- L S L Q L * N H P G L V F A L * I H T D
- Y L Y S Y R T T Q G * F L L Y K S T Q I
- I F T A I E P P R A S F C F I N P H R *
8701 - AAGTGAAAACCCCTTCTTTAGAGTCATTCTCTTTTGTACATGTTTGGTCTAGGGTCAT - 8760
- K * K T L L * S H S L L S H V W S * G H
- S E K P F F R V I L F C H M F G P R V I
- V K N P S L E S F S F V T C L V L G S Y
8761 - ACATATCGCTAATAATAAGGTCCCATTATTAGCCGATGTACTGTGCACAGTCTCCAA - 8820
- T Y R * * * G P I Y * P Y V L L H S L Q
- H I A N N K V P F I S R M Y C C T V S N
- I S L I I R S H L L A V C T V A Q S P I
8821 - TTAAGTAGAATCTGCGTCGGAGACGAAGTCATTAAGATCTGAATCGACAAGTAGTGTGC - 8880
- L K * N L R R R' R S H * D L N R Q V V C
- * S R I C V G D E V I K I * I D K * C A
- K V E S A S E T K S L R S E S T S S V P
8881 - CAGTTGGCAACCATTGTCTGAGCACAGCTGTACCTGGTGAACCTCCTTTATCAGAGCCAG - 8940
- Q L A T I V * A Q L Y L V Q L L Y Q S Q
- S W Q P L S E H S C T W C N S F I R A S
- V G N H C L S T A V P G A T P L S E P A
8941 - CACCAAAGTGAATAACTCTCATGTTGTAGGGTACAGCTAAAGTAAAGTATTAAAGTATT - 9000
- H Q S E * L S C C R V Q L K * V Y L S I
- T K V N N S H V V G Y S * S K C I * V L
- P K * I T L M L * G T A K V S V F K Y *
9001 - GACACAGTTGAGTATACTTTGCGACATTCATCATTATTCCTTTTGGTATAACAGCATTTT - 9060
- D T V E Y T L R H S S L F L L V * Q H F
- T Q L S I L C D I H H Y S F W Y N S I F
- H S * V Y F A T F I I I P F G I T A F S
9061 - CACCATAATCTGAAGGTCACACTTTCAAGAAGCATTCTTTGCATCTTGTAAGTTAG - 9120
- H H N S E G H T F Q E A F F A S C T S *
- T I I L K V T L F K K H S L H L V Q V R
- P * F * R S H F S R S I L C I L Y K L G
9121 - GCATCGCAACACCTGGTTGCCACGCTTGACTTGTAGTTTTGGGTAGAAGGTTTCAA - 9180
- A S Q H L V A T L D L L V V L G R R F Q
- H R N T W L P R L T C L * F W V E G F N
- I A T P G C H A * L A C S F G * K V S T
9181 - CATGTCCATCCTTACACCAAAGCATGAATGAATTTAGCATAGTCAATTGTAACCTTGA - 9240
- H V H P Y T K A * M K F Q H S Q L * P *
- M S I L T P K H E * N F S I V N C N L D
- C P S L H Q S M N E I S A * S I V T L T

```

图 12 续

9241 - CCACTTTTGAAATCACTGACAAATCTTGTGACTTTATTATCTCGACAAAGTCATCAAGTA - 9300
 - P L L K S L T N L V T L L S R Q S H Q V
 - H F * N H * Q I L * L Y Y L D K V I K *
 - T F E I T D K S C D F I I S T K S S S K
 9301 - AAAGATCAATCACAGAACACACACATTTTGATGACCTGTTTGGCATCTGTTATGAAGT - 9360
 - K D Q S Q N T H I L M N L F A H L L * S
 - K I N H R T H T F * * T C L R I C Y E V
 - R S I T E H T H F D E P V C A S V M K *
 9361 - AATTTTTCACTGTGCTGTCCATAGGGATAAAATCCTTAATTTAAGTGGTGAATCTTGTG - 9420
 - N F S L C C P * G * N P L I * V V N L V
 - I F H C A V H R D K I L * F K W * I L *
 - F F T V L S I G I K S S N L S G E S C E
 9421 - AGCGCTTGGCTAAGCCTATCATTAATGAAGACCGCCAAGTTGTCCATGACTGAAATCTC - 9480
 - S A W L S L S L N E D R Q V V H D * N L
 - A L G * A Y H * M K T A K L S M T E I S
 - R L A K P I I K * R P P S C P * L K S P
 9481 - CATAACGATGTGTTTGAAGGCATAGCCCTCGAGCTTATATCGCTGTATGAATTCATCCA - 9540
 - H K R C V R R H S P R A Y * A V * I H P
 - I N D V F E G I A L E L I S L Y E F I H
 - * T M C S K A * P S S L Y R C M N S S I
 9541 - TAGCGAGCTCGAGAAAGTCAGTTTCCATTTGTGATCTGGGCTTAAAATCCTCTAAGTCTC - 9600
 - * R A R E S Q F P F V I W A * N P L S L
 - S E L E K V S F H L * S G L K I L * V S
 - A S S R K S V S I C D L G L K S S K S L
 9601 - TGCTCTGAGTAAAGTAGGTTTTCAGGCACTGTTGAATAATGCCGCTACTTTCTTAAAGT - 9660
 - C S E * S R F Q A T V E * C R L L S * S
 - A L S K V G F R Q L L N N A V Y F L K V
 - L * V K * V S G N C * I M P S T F L K *
 9661 - AGTTAAACTGTGTTTTTACTGATTCTCCAATTAATGTGACTCCATGACGCTAGCTTGTG - 9720
 - S * T V F L L I L Q L M * L H * R * L V
 - V K L C F Y * F S N * C D S I D A S L C
 - L N C V F T D S P I N V T P L T L A C A
 9721 - CTGGTCCCTTTGAAGGTGTTAGACCTTTGACTGAACCTTCTGTTATTAACACCATTAC - 9780
 - L V P L K V L D L * L N L L L L K H H Y
 - W S L * R C * T F D * T F C Y * N T I T
 - G P F E G V R P L T E P S V I K T P L R
 9781 - GGGCGTTTCTAAAAGGTCTACCTGTCTTCCACTCTACCATCAAACAAGACAGTAAGT - 9840
 - G R F * K G L P V L P L Y H Q T R Q * V
 - G V S K K V Y L S F H S T I K Q D S K *
 - A F L K R S T C P S T L P S N K T V S E
 9841 - AAGAACAAGCACTCTCAGTAGGTTTCTTGGCAATGTCAGTCATTGTGCAGACACCTATTG - 9900
 - K N K H S Q * V S W Q C Q S L C R H L L
 - R T S T L S R F L G N V S H C A D T Y C
 - E Q A L S V G F L A M S V I V Q T P I V
 9901 - TAGATACATGTGCTGGGGCTTCTCTTTGTAGTCCCAGATTACAGTATTAGCAGCGATAT - 9960
 - * I H V L G L L F C S P R L Q Y * Q R Y
 - R Y M C W G F S F V V P D Y S I S S D I
 - D T C A G A S L L * S Q I T V L A A I S
 9961 - CAACACCCAAATTATTGAGTATCTTAATCTCTGGCACTGGTTAATGTTACGCTTAGCCC - 10020
 - Q H P N Y * V S * S L A L V * C Y A * P
 - N T Q I I E Y L N L W H W F N V T L S P
 - T P K L L S I L I S G T G L M L R L A Q
 10021 - AAAGCTCAAATGCAACATTACAGGAAGTGTGTCTTATTTTCAAAGATCTCCACATCAA - 10080
 - K A Q M Q H * Q E V L S Y F Q R S P H Q
 - K L K C N I N R K C C L I F K D L H I N
 - S S N A T L T G S V V L F S K I S T S I

图 12 续


```

10081 - TACCATCTACCTTTGTGTAACAGCATTATTAATGATGGAAACAGGTGCTTCGCCGGCGT - 10140
- Y H L P L C K Q H Y * * W K Q V L R R R
- T I Y L C V N S I I N D G N R C F A G V
- P S T F V * T A L L M M E T G A S P A C
10141 - GTCCATCAAAGTGCCTTTATTAAACAACATTATAAGCCACATTTTCTAAACTCTGTAACC - 10200
- V H Q S V L Y * Q H Y K P H F L N S V T
- S I K V S F I N N I I S H I F * T L * P
- P S K C P L L T T L * A T F S K L C N L
10201 - TGGTAAATGTATCCACAGGTTATAAGTATCAAATTGTTGTAATCCATAGGCTAAATC - 10260
- W * M Y S T G Y K Y Q I V C K S I G * I
- G K C I P Q V I S I K L F V N P * A K S
- V N V F H R L * V S N C L * I H R L N P
10261 - CAGCAGAAATCATCATATTATATGCATCCAAGTACTGTCGGTACTCATTGCAATGGTGC - 10320
- Q Q K S S Y Y M H P S T V G T H L H G V
- S R N H H I I C I Q V L S V L I C M V S
- A E I I I L Y A S K Y C R Y S F A W C L
10321 - TGCAACAGCACCCACCTAAATTGCATCGTGTAAATACACGTAGCAGATTTGAGTGAACAT - 10380
- C K Q H H L N C I V * Y T * Q I * V E H
- A N S T T * I A S C N T R S R F E W N I
- Q T A P P K L H R V I H V A D L S G T *
10381 - AATCAATATCCGACACTACTTGTGGCCATGAGACTCACAAGGACTATCAGAATAGTAA - 10440
- N Q Y P T L L V C H E T H K D Y Q N S K
- I N I R H Y L F A M R L T R T I R I V K
- S I S D T T C L P * D S Q G L S E * * K
10441 - AGAAAGGCAATTGCTTTAAATTAGTAAATGCACCTTTTATCGAAAGCTGGAGTGTGGAATG - 10500
- R K A I A L N * * M H F Y R K L E C G M
- E R Q L L * I S K C T F I E S W S V E C
- K G N C F K L V N A L L S K A G V W N A
10501 - CATGCTTATTCACATACAACTACCACCATCACAGCCTGGTAAGTTCAAGTTTGACAAGA - 10560
- H A Y S H T N Y H H H S L V S S S L T R
- M L L H I Q T T T I T A W * V Q V * Q D
- C L F T Y K L P P S Q P G K F K F D K T
10561 - CTCTGTGTCAAACCTACACAAATGCATTGGCTGGGTAACGATCAACGTTACAATTCC - 10620
- L L C Q T Y T Q L H W L G N D Q R Y N S
- S C V K P T H N C I G W V T I N V T I P
- L V S N L H T I A L A G * R S T L Q F Q
10621 - AAAACAAACAACACCATCAGTGAATTTATCGTGATGTGTAGCATAGAATAGAAGAGTT - 10680
- K T N K H H Q * I Y R D V * H K N R R V
- K Q T N T I S E F I V M C S I R I E F
- N K Q T P S V N L S * C V A * E * K S S
10681 - CCTCTATTTTGTAAAGCTTTGTCACACTACATGGCTGAGCATCGTAGAACTTCCATCTACTT - 10740
- P L F C K L C H Y M A E H R R T S I L L
- L Y F V S F V T T W L S I V E L P F Y F
- S I L * A L S L H G * A S * N F H S T S
10741 - CAGCCTGAGGCACACACTTGATAGCCTTTGGATTCCAATGTCATGAAGAAGCTGGAACT - 10800
- Q P E A H T * * P L D F Q C H E E L E T
- S L R H T L D S L W I S N V M K N W K L
- A * G T H L I A F G F P M S * R T G N L
10801 - TATCAGCAAGCAATGCAGACTTCACAACCATGTGTTGTACTTTTCTGCAAGCAGAATTAA - 10860
- Y Q Q A M Q T S Q P C V V L F C K Q N *
- I S K Q C R L H N H V L Y F S A S R I N
- S A S N A D F T T M C C T F L Q A E L T
10861 - CCCTCAGTTTCTCCTATAATAGGTATTCAACAGACCAATCAACGCGCTTAACAAAGC - 10920
- P S V H L L * * G I Q Q T N Q R A * Q S
- P Q F I S Y N R V F N R P I N A L N K A
- L S S S P I I G Y S T D Q S T R L T K H

```

图 12 续

```

10921 - ACTCATGGACTGCTAAACATCTAGTCATGATAGCATCACAACCTAGCCACATGTGCATTC - 10980
- T H G L L N I * S * * H H N * P H V H F
- L M D C * T S S H D S I T T S H M C I S
- S W T A K H L V M I A S Q L A T C A F P
10981 - CATGTACCTGGCAATGTTGGTCATGGTTACTCTGAAGGTTACCCGTAAGCCCCACTGCT - 11040
- H V P G N V G H G Y S E G Y P * S P T A
- M Y L A M L V M V T L K V T R K A P L L
- C T W Q C W S W L L * R L P V K P H C *
11041 - GACATCAATCATAAATGGGTTATAGACATAGTCAAAACCCACAGAATGATTCCAGCAGG - 11100
- E H Q S * M G Y R H S Q N P Q N D S S R
- N I N H K W V I D I V K T H R M I P A G
- T S I I N G L * T * S K P T E * F Q Q A
11101 - CATAAGTATCTGATGAAGTAGAAAAGCAAGTTGCACGTTTGTACACAGACAACAGCTTC - 11160
- H K Y L M K * K S K L H V C H T D N T F
- I S I * * S R K A S C T F V T Q T T R S
- * V S D E V E K Q V A R L S H R Q H V L
11161 - TTTGAGGTCCAATCTTGACAAAGTACTTCATTGATGTRAGCTCAAAGCCATGCGCCAAA - 11220
- F Q V Q S * Q S T S L M * A Q S H A P K
- F R S N L D K V L H * C K L K A M R P K
- S G P I L T K Y F I D V S S K P C A Q R
11221 - GGACGAACACGACTCTGTCTGACAATCCTTTTCAGTGTATCACTGAGCATTTGTACTATCT - 11280
- G R T R L C L T I L S V Y H * A F V L S
- D E H D S V * Q S F Q C I T E H L Y I L
- T N T T L S D N P F S V S L S I C T I L
11281 - TAATACGCACTACATTCAGGGCAAGCCTTTATACATGAGTGGTATAAGATGTTTAAACT - 11340
- * Y A L H S R A S L Y T * V V * D V * T
- N T H Y I P G Q A F I H E W Y K M F K L
- I R T T F Q G K P L Y M S G I R C L N W
11341 - GGTCACCTGGTGGAGGTTTTCGATTAACTCTGGTGAATCTGTGTTATTTTCAGTGTCAA - 11400
- G H L V E V L H * L W * I L C Y F Q C Q
- V T W W R F C I N S G E F C V I F S V N
- S P G G F A L T L V N S V L F S V S T
11401 - CATAACCAGTCGGTACAGCTACTAAGTTAACCTGTAGAAAATCCTAGCTGGAGAGGTA - 11460
- H N Q S V Q L L S * H L * K I L A G E V
- I T S R Y S Y * V N T C R K S * L E R *
- * P V G T A T K L T P V E N P S W R G R
11461 - GGTTAGTACCCACAGCATCTCTAGTTGCATGACAGCCCTTACATCAAAGCCAATCCAGC - 11520
- G * Y P Q H L * L H D S P L H Q S Q S T
- V S T H S I S C M T A L Y I K A N P R
- L V P T A S L V A * Q P S T S K P I H A
11521 - CACGAACGTGACGAATAGCTTCTTCGCGGGTATAAACATATTAGGGTAACCATTGACTT - 11580
- H E R D E * L L R G * * T Y * G N R * L
- T N V T N S F F A G D K H I R V T I D L
- R T * R I A S S R V I N I L G * P L T W
11581 - GGTAATTCATTTGAAACCCATCATAGAGATGAGTCTACGGTAGGTCATGTCCTTTGGTA - 11640
- G N S F * N P S * R * V Y G R S C P L V
- V I H F E T H H R D E S T V G H V L W Y
- * F I L K P I I E M S L R * V M S F G M
11641 - TGCCTGGTATGTC AACACATAATCCTTCAGTCTTGAATTTATATCACGCTGAGGTTG - 11700
- C L V C Q H I I L Q S * I L Y Q R * G V
- A W Y V N T * S F S L E F Y I N A E V C
- P G M S T H N P S V L N F I S T L R C V
11701 - TAGTGCCTGTGTAGGATGAAGACCAGTAATGATCTTACTACAGTCCTTAAAAGTCCAG - 11760
- * V P V * D E D Q * * S Y Y S P * K V Q
- R C L C R M K T S N D L T T V L K K S S
- G A C V G * R P V M I L L Q S L K S P V

```

图 12 续

```

11761 - TTACATTTTCTGCTTGTAAATGTAGCCACATTGCGACGTGGTATTTCTAGACTTGTAAATT - 11820
- L H F L L V M * P H C D V V F L D L * I
- Y I F C L * C S H I A T W Y F * T C K L
- T F S A C N V A T L R R G I S R L V N C
11821 - GCAGTTTGTCAAAAGATCTCTATCAGACATTATGCACAAAATGCCAATTTTGGCCCTTG - 11880
- A V C H K D L Y Q T L C T K C Q F L P L
- Q F V I K I S I R H Y A Q N A N F C P C
- S L S * R S L S D I M H K M P I F A L V
11881 - TGATAGCCACATTGAAGCGTTGACATTACAAGAGTGTGCTGTTTCAGTAGTTTGTGTGA - 11940
- * * P H * S G * H Y K S V L F Q * F V *
- D S H I E A V D I T R V C C F S S L C E
- I A T L K R L T L Q E C A V S V V C V N
11941 - ATATGACATAGTCATATTCAGAACCCTGTGATGAATCAACAGTCTGCGTAGGCAATCCTA - 12000
- I * H S H I Q N P V M N Q Q S A * A I L
- Y D I V I F R T L * * I N S L R R Q S *
- M T * S Y S E P C D E S T V C V G N P K
12001 - AGATTTTGAAGCTACAGCGTCTGTGAATTATAAGGTGAGATAAAAAACAGCTTTTCTCC - 12060
- R F L K L Q R S V N Y K V R * K Q L F S
- D F * S Y S V L * I I R * D K N S F S P
- I F E A T A F C E L * G E I K T A F L Q
12061 - AAGCAGGATTGCGTGTAAAGAAATCTCTTACAACGCCTATTGAGGTCTGTTGATTGCAG - 12120
- K Q D C V * E I L L Q R L F E V C * L Q
- S R I A C K K F S Y N A Y L R S V D C R
- A G L R V R N S L T T P I * G L L I A D
12121 - ATGAAACATCATGTGTAATAACACCTTGTAGAACATTTGAAGCATTGAGCTGACTTAT - 12180
- M K H H V * * H L C R T F * S I E L T Y
- * N I M C N N T F V E H F E A L S * L I
- E T S C V I T P L * N I L K H * A D L S
12181 - CCTTGTGTGCTTTTAGCTTATTGTCAAAACTAAAGCACTCACAGTGTCAACATTTTCAG - 12240
- P C V L L A Y C H K L K H S Q C Q Q F Q
- L V C F * L I V I N * S T H S V N N F S
- L C A F S L L S * T K A L T V S T I S A
12241 - CAGGACAACGGCGACAAGTCCAAGGAACATGTCTGGACCTATTGTTTTCATAAGTCTGC - 12300
- Q D N G D K F Q G T C L D L L F S * V C
- R T T A T S S K E H V W T Y C F H K S A
- G Q R R Q V P R N M S G P I V F I S L H
12301 - AACTGAATTAATAATTTCTGGTTCTAGTGTGCCITTTAGTCAGCAATGTGCGGGGGGCTG - 12360
- T L N * N I L V L V C L * S A M C G G L
- H * I K I F W F * C A F S Q Q C A G G W
- T E L K Y S G S S V P L V S N V R G A G
12361 - GTAATTGAGCAGGATCGCCAATATAGACGTAGTGTTTTGCACGAAGTCTAGCATTGACAA - 12420
- V I E Q D R Q Y R R S V L H E V * H * Q
- * L S R I A N I D V V F C T K S S I D N
- N * A G S P I * T * C F A R S L A L T T
12421 - CACTCAAGTCATAATTAGTAGCCATAGAGATTTTCATCAAGACTACAATGTCAGCAGTTG - 12480
- H S S H N * * P * R F H Q R L Q C Q Q L
- T Q V I I S S H R D F I K D Y N V S S C
- L K S * L V A I E I S S K T T M S A V V
12481 - TTTCTGGCAATGCATTTACAGTGCAGAAAACATACTGTTCTAGTGTGAATTCACTTTGA - 12540
- F L A M H L Q C R K H T V L V L N S L *
- F W Q C I Y S A E N I L F * C * I H F E
- S G N A F T V Q K T Y C S S V E F T L N
12541 - ATTTATCAAAACTCTACGCGCGCACGCGCAGSTATGATTCTACTACATTTATCTATGG - 12600
- I Y Q N T L R A H A Q V * F Y Y I Y L W
- F I K T L Y A R T R R Y D S T T F I Y G
- L S K H S T R A R A G M I L L H L S M G

```

图 12 续

```

12601 - GCAAATATTTAATGCCTTTTCACATAGGGCATCAACAGCTGCATGAGAGCATGCCGTAT - 12660
- A N I L M P F H I G H Q Q L H E S M P Y
- Q I F * C L F T * G I N S C M R A C R I
- K Y F N A F S H R A S T A A * E H A V Y
12661 - ACACTATGCGAGCAGATGGGTAATAGAGGCAAGTCCGATGGCAAATGACTCTTACCAG - 12720
- T L C E Q M G N R E Q V R W Q N D S Y Q
- H Y A S R W V I E S K S D G K M T L T S
- T M R A D G * * R A S P M A K * L L P V
12721 - TACCAGGTGGTCCTTGGAGTGTAGAGTACTTTTGCATGCCGACCTTTTGATAATTTGCAA - 12780
- Y Q V V L G V * S T F A C R P F D N L Q
- T R W S L E C R V L L H A D L L I I C N
- P G G P W S V E Y F C M P T F * * F A T
12781 - CATTGCTAGAAAACCTCATCTGAGATGTTGAGTGTGGGTACAAGCCAGTAATCTCACAT - 12840
- H C * K T H L R C * V L G T S Q * F S H
- I A R K L I * D V E C W V Q A S N S H I
- L L E N S S E M L S V G Y K P V I L T *
12841 - AGTGCTCTTGGGCACTAGAGTAGGTCACCTAAGTGGCATTACAGTGTGAGATGTCAACA - 12900
- S A L V A L E * V H * V A L Q C E M S T
- V L L W H * S R C T K W H Y S V R C Q H
- C S C G T R V G A L S G I T V * D V N T
12901 - CAAAGTAATCACCACATTCAACTGTATGTCGTAGTACCTCTGTACACAACAGCATCAC - 12960
- Q S N H Q H S T C M S * Y L C T Q Q H H
- K V I T N I Q L V C R S T S V H N S I T
- K * S P T F N L Y V V V P L Y T T A S P
12961 - CATAGTCACCTTTTTCAAGGTGTACTCTCCAATCTGTACTTTACTATTTTGTAGTTACAC - 13020
- H S H L F Q R C T L Q S V L Y Y F * L H
- I V T F F K G V L S N L Y F T I F S Y T
- * S P F S K V Y S P I C T L L F L V T R
13021 - GGTAACCAGTAAAGACATAGTTTCTGTTCAATGGTGGTCTAGGTTTTCCAACCTCCCATG - 13080
- G N Q * R H S F C S M V V * V F Q P P M
- V T S K D I V S V Q W W S R F S N L P *
- * P V K T * F L F N G G L G F P T S H E
13081 - AAAGATGCAATCTCTGTACAGAGTACTTCGCGTACAGTGGCAATACCATATGACAGCT - 13140
- K D A I L C Q R V L R V Q W Q Y H M T A
- K M Q F S V R E Y F A Y S G N T I * Q L
- R C N S L S E S T S R T V A I P Y D S L
13141 - TAAATGTTTCTCAGTGGCTTTGAGCGTTTCTGCTGCCAAAAGCTTGAGTCTCTCAGTAC - 13200
- * M F P Q W L * A F L L R K A * V S Q Y
- K C F L S G F E R F C C E K L E S L S T
- N V S S V A L S V S A A K S L S L S V Q
13201 - AAGTGTGGCAAGTATGTAATCGCCAGCATTAGTCCAATCACATGTTGCTATCGCATTGA - 13260
- K C W Q V C N R Q H * S N H M L L S H *
- S V G K Y V I A S I S P I T C C Y R I E
- V L A S M * S P A L V Q S H V A I A L K
13261 - AGTCAGTGACATTGTCACTGCCTACACATGTGTTTTGTATAAACCAAAAACCTGACCAT - 13320
- S Q * H C H C L H M C F C I N Q K P D H
- V S D I V T A Y T C V F V * T K N L T I
- S V T L S L P T H V F L Y K P K T * P L
13321 - TAGCACATAATGGAAAACCTAATGGGAGGCTTATGTGACTTGCAATAATAGCTCATACCTC - 13380
- * H I M E N * W E A Y V T C N N S S Y L
- S T * W K T N G R L M * L A I I A H T S
- A H N G K L M G G L C D L Q * * L I P P
13381 - CTAGATACAGTTGTGTACATCAGTGACATCACAACTGGGGCATTGCAAACATAGGGAT - 13440
- L D T V V S H Q * H H N L G H C K H R D
- * I Q L C H I S D I T T W G I A N I G I
- R Y S C V T S V T S Q P G A L Q T * G L

```

图 12 续

13441 - TAACAGACAACACTAATTTGTGTGATGTTGAAATGACATGGTCATAGCAGCACTTGCAAC - 13500
 - * Q T T L I C V M L K * H G H S S T C N
 - N R Q H * F V * C * N D M V I A A L A T
 - T D N T N L C D V E M T W S * Q H L Q H
 13501 - ATAGGAATGGTCTOCTAATACAGGCACCGCAACGAAGTGAAGTCTGTGAATTGCACAATA - 13560
 - I G M V S * Y R H R N E V K S V N C T I
 - * E W S P N T G T A T K * S L * I A Q Y
 - R N G L L I Q A P Q R S E V C E L H N T
 13561 - CACAAGCACCTACAGCCTGCAAGACTGTATGTGGTGTGTACATAGCCTCATAAACTCAG - 13620
 - H K H L Q P A R L Y V V C T * P H K T Q
 - T S T Y S L Q D C M W C V H S L I K L R
 - Q A P T A C K T V C G V Y I A S * N S G
 13621 - GTTCCCAGTACCGTGAGGTGTTATCATTAGTTAGCATTACGGAATACATGTCCAACATGT - 13680
 - V P S T V R C Y H * L A L R N T C P T C
 - F P V P * G V I I S * H Y G I H V Q H V
 - S Q Y R E V L S L V S I T E Y M S N M W
 13681 - GGCCAGTAAGCTCATCATGTAACCTTTCTAATGTATTGTAATAACAAGTGAAGACATCAG - 13740
 - G Q * A H H V T F * C I V N T S E R H Q
 - A S K L I M * L S N V L * I Q V K D I S
 - P V S S S C N F L M Y C K Y K * K T S A
 13741 - CATACTCCTGATTAGGATGTTTGTAGTGGGTAAAGCATCAATAGCCAGTGACACGAACC - 13800
 - H T P D * D V L * V G K H Q * P V T R T
 - I L L I R M F C K W V S I N S Q * H E P
 - Y S * L G C F V S G * A S I A S D T N L
 13801 - TTTCAATCATAAGTGTACCATCTGTTTGTGACAATATCATCGACAAAACAGCCTGCGCCTA - 13860
 - F Q S * V Y H L F * Q Y H R Q N S L R L
 - F N H K C T I C F D N I I D K T A C A *
 - S I I S V P S V L T I S S T K Q P A P N
 13861 - ATATTCTTGATGGATCTGGGTAAGGCAGGTACACGTAATCATCTCCTTGTTAACTAGCA - 13920
 - I F L M D L G K A G T R N H L L V * L A
 - Y S * W I W V R Q V H V I I S L F N * H
 - I L D G S G * G R Y T * S S P C L T S I
 13921 - TTGTATGCTGTGAGCAAAATTCGTGAGGTCCTTTAGTAAGGTCAGTCTCAGTCCAACATT - 13980
 - L Y A V S K I R E V L * * G Q S Q S N I
 - C M L * A K F V R S F S K V S L S P T F
 - V C C E Q N S * G P L V R S V S V Q H F
 13981 - TTGCCTCAGACATGAACACATTATTTGATAATAAAGAAGTGCCTTAAAGTTCTTAATGC - 14040
 - L P Q T * T H Y F D N K E L P * S S * C
 - C L R H E H I I L I I K N C L K V L N A
 - A S D M N T L F * * R T A L K F L M L
 14041 - TAGCTACTAAACCTTGAGCCGCATAGTTACTGTTATAGCACACAACGGCATCATCAGAAA - 14100
 - * L L N L E P H S Y C Y S T Q R H H Q K
 - S Y * T L S R I V T V I A H N G I I R K
 - A T K P * A A * L L L * H T T A S S E R
 14101 - GAATCATCATGGAGAAATGTTTACGCAGGTAAGCGTAAACTCATCCCGAATTCATGAT - 14160
 - E S S W R N V Y A G K R K T H P R I H D
 - N H H G E M F T Q V S V K L I H E F M I
 - I I M E K C L R R * A * N S S T N S * S
 14161 - CAACATCCCTATTCTATAGAGACACTCATAGAGCCTGTGTGTAGATTGCGGACATACT - 14220
 - Q H P Y F Y R D T H R A C V V D C G H T
 - N I P I S I E T L I E P V L * I A D I L
 - T S L F L * R H S * S L C C R L R T Y L
 14221 - TGTCAGCTATCTTATTACCATCAGTTGAAAGAAGTGCATTACATTGGCTGTAACAGCTT - 14280
 - C Q L S Y Y H Q L K E V H L H W L * Q L
 - V S Y L I T I S * K K C I Y I G C N S L
 - S A I L L P S V E R S A F T L A V T A *

图 12 续

14281 - GACAAATGTTAAAGACACTATTAGCATAAGCAGTTGTAGCATCACCGGATGATGTTCCAC - 14340
 - D K C * R H Y * H K Q L * H H R M M F H
 - T N V K D T I S I S S C S I T G * C S T
 - Q M L K T L L A * A V V A S P D D V P P
 14341 - CTGGTTTAAACATATAGTGAGCCGCCACACATGACCATCTCACTTAATACTGGCGCACCT - 14400
 - L V * H I V S R H T * P S H L I L A H T
 - W F N I * * A A T H D H L T * Y L R T L
 - G L T Y S E P P H M T I S L N T C A H S
 14401 - CGTTAGCTAACCTGTAGAACGGTGTGATAAGTTACAGCAAGTGTATGTTTGGGAGCAA - 14460
 - R * L T C R N G V I S Y S K C Y V C E Q
 - V S * P V E T V * * V T A S V M F A S K
 - L A N L * K R C D K L Q Q V L C L R A R
 14461 - GAACAAGAGAGGCCATTATCCTAAGCATGTTAGGCATGGCTCTGTACATTTTGATAAT - 14520
 - E Q E R P L S * A C * A W L C H I L D N
 - N K R G H Y P K H V R H G S V T F W I I
 - T R E A I I L S M L G M A L S H F G * S
 14521 - CCCAACCCATAAGGTGTGGAGTTTCTACATCACTGTAACAGTTTTTAACATATTATGCC - 14580
 - P N P * G V E F L H H C K Q F L T Y Y A
 - P T H K V W S F Y I T V N S F * H I M P
 - Q P I R C G V S T S L * T V F N I L C Q
 14581 - AGCCACCGTAAACTTGCTGTTCCAATTACCACAGTAGCTCCTCTAGTGGCGGCTATTG - 14640
 - S H R K T C L F Q L P Q * L L * W R L L
 - A T V K L A C S N Y H S S S S S G G Y *
 - P P * N L L V P I T T V A P L V A A I D
 14641 - ACTTCAATAATTTCTGATGAACTGTCTATTTGTCATAGTACTACAGATAGAGACACCAG - 14700
 - T S I I S D E T V Y L S * Y Y R * R H Q
 - L Q * F L M K L S I C H S T T D R D T S
 - F N N F * * N C L F V I V L Q I E T P A
 14701 - CTACGGTGCAGCTCTATTCTTGGCACTAATGGCATACTTAAGATTCATTGAGTTATAG - 14760
 - L R C E L Y S L H * W H T * D S F E L *
 - Y G A S S I L C T N G I L K I H L S Y S
 - T V R A L F F A L M A Y L R F I * V I V
 14761 - TAGGGATGACATTACGCTTAGTATACGCGAAAAGTGCATCTTGATCCTCATAACTCATTG - 14820
 - * G * H Y A * Y T R K V H L D P H N S L
 - R D D I T L S I R E K C I L I L I T H *
 - G M T L R L V Y A K S A S * S S * L I E
 14821 - AGTCATAATAAGTCTAGCCTTACCCATTATTAATGGGAAACCAGCTGATTTATCCA - 14880
 - S H N K V * P Y P I Y * M G N Q L I Y P
 - V I I K S S L T P F I K W E T S * F I Q
 - S * * S L A L P H L L N G K P A D L S R
 14881 - GATTGTTAACGATTACTTGGTTGGCATTAAATACAGCCACCATCGTAACATCAAAGTATT - 14940
 - D C * R L L G W H * Y S H H R N N Q S I
 - I V N D Y L V G I N T A T I V T I K V F
 - L L T I T W L A L I Q P P S * Q S K Y L
 14941 - TATCAACAACCTCACTACGAATAGGAGTTGTCTGATATCACACATTGTTGGCAGATTAT - 15000
 - Y Q Q L Q L R I G V V * Y H T L L A D Y
 - I N N F N Y E * E L S D I T H C W Q I I
 - S T T S T T N R S C L I S H I V G R L *
 15001 - AACGATAATAGTCATAATCACTGATAGCAGCGTTGCCATCCTGAGCAAAGAAGAAGTGT - 15060
 - N D N S H N H * * Q R C H P E Q R R S V
 - T I I V I I T D S S V A I L S K E V F
 - R * * S * S L I A A L P S * A K K K C F
 15061 - TTAGTTCAACAGAACTTCCTTCTTAAAGAAACCTTAGACACAGCAAAGTCATAAAAGT - 15120
 - L V Q Q N F L P * R N L * T Q Q S H K S
 - * F N R T S F L K E T F R H S K V I K V
 - S S T E L P S L K K P L D T A K S * K S

图 12 续

```

15121 - CTTATTAATAATTACCGGGTTGACAGTTTGAAAAGCAACATTGTTTGTAGTGCAGCTA - 15180
- L Y * N Y R V * Q F E K Q H C L L V Q L
- F I K I T G F D S L K S N I V C * C S Y
- L L K L P G L T V * K A T L F V S A A T
15181 - CTGAAAAGCATGTAGTGCCTTTATCTAGCAATAAATTGCCAGAAGCTGCATGCATAGCTG - 15240
- L K S M * C V Y L A I N C Q K L H A * L
- * K A C S A F I * Q * I A R S C M H S W
- E K H V V R L S S N K L P E A A C I A G
15241 - GATCAGCAGCATACACTAAAAGTTCCTTGAAACTGAGACGCGAGCTATGTAAGTTTACAT - 15300
- D Q Q H T L K V P * N * D A S Y V S L H
- I S S I H * K F L E T E T R A M * V Y I
- S A A Y T K S S L K L R R E L C K F T S
15301 - CCTGATTATGTACGACTCCTAAGTTCAGAAAATGGTATCCAGTTGAAACAACAAAAGGAA - 15360
- P D Y V R L L T H E N G I Q L K Q Q K E
- L I M Y D S * L T K M V S S * N N K R N
- * L C T T P N S R K K W Y P V E T T K G T
15361 - CACCATCTACAAATATTTTTCTTACTAGTGGTCCAAAACCTGTAGGTGAAAACACAGTAG - 15420
- H H L Q I F F L L V V Q N L * V E T Q *
- T I Y K Y F S Y * W S K T C R W K H S R
- P S T N I F L T S G P K L V G G N T V E
15421 - AAAATAACACATTAAGTTTGACAATGAAGGATACACCTATCATCCAACAGTTAATAC - 15480
- K I T H * S L H N E G Y T Y H P N S * Y
- K * H I K V C T M K D T P I I Q T V N T
- N N T L K F A Q * R I H L S S K Q L I Q
15481 - AATTGGGATGGTATGTCTGGTCCCAATATTTAAATAACGGTCGAAGAGACAAAGTCTCT - 15540
- N W D G M S G P N I * N N G R R D K V S
- I G M V C L V P I F K I T V E E T K S L
- L G W Y V W S Q Y L K * R S K R Q S L S
15541 - CTTCCGTAAAATCATATTTTCAGCAAATCCCACTTAATAAGTGGTTTTGCGAGATCAGCAT - 15600
- L P * N H I S A N P T * * V V L R D Q H
- F R K I I F Q Q I P L N K W F C E I S I
- S V K S Y F S K S H L I S G F A R S A S
15601 - CCATATGGGACTCAGCAGCCATGCCCTAGTCAAAGTGAGGATGGGCATCAGCAATGAGT - 15660
- P Y G T Q Q P M P * S K * G W A S A M S
- H M G L S S Q C P S Q S E D G H Q Q * V
- I W D S A A N A L V K V R M G I S N E *
15661 - AATATGAATCCACAATAGGAATCCGCAGCCTGGTCTACTTGTACGAAATCACCGAAT - 15720
- N M N P Q * E L R S L V L L V R N H R N
- I * I H N R N S A A W C Y L Y E I T E I
- Y E S T I G T P Q P G A T C T K S P K S
15721 - CGTACCAGTTCCCATTAAGATCCTGATTATCTAATGTCAGTACGCCTACAATGCCTGCAT - 15780
- R T S S H * D P D Y L M S V R L Q C L H
- V P V P I K I L I I * C Q Y A Y N A C I
- Y Q F P L R S * L S N V S T P T M P A S
15781 - CACGCATAGCATCGCAGAATTGTACAGTCTTAATAATGATTGGCGTACACGCTCACCTA - 15840
- H A * H R R I V Q S L I M I G V H A H L
- T H S I A E L Y S L * * * L A Y T L T *
- R I A S Q N C T V F N N D W R T R S P K
15841 - AGTTAGCATATACGGTAAGATGTCAGGATCTCTACGAAGTCATACCAATCCTTCTTAT - 15900
- S * H I R V R C Q D S L R S H T N P S Y
- V S I Y A * D V R I L Y E V I P I L L I
- L A Y T R K M S G F S T K S Y Q S F L L
15901 - TGAATAATCATCATCAGCAATGTATGTGACGAGTATTTCTTTAATGTATCACAAT - 15960
- * N N H H H S N C M * R V F L L M Y H N
- E I I I I T A I V C D E Y F F * C I T I
- K * S S S Q Q L Y V T S I S F N V S Q L

```

图 12 续

15961 - TACCCTCATCAAATGACGTAGAGCATAGACTAAATCAGCCATTGTGTATTTAGTTAGAC - 16020
 - Y P H Q N D V E H R L N Q P L C I * L D
 - T L I K M T * S I D * I S H C V F S * T
 - P S S K * R R A * T K S A I V Y L V R R
 16021 - GCTGACGTGATATATGTGGTACCATGTCACCATCTACTCTAAACTTGAAAAAGTCATGGA - 16080
 - A D V I Y V V P C H H L L * T * K S H G
 - L T * Y M W Y H V T I Y S K L E K V M D
 - * R D I C G T M S P S T L N L K K S W T
 16081 - CAGCAACCGCTGGACAATCTTTAACCAAGTTATAAATAGTCTCTTCATGTTGGTAGTTAG - 16140
 - Q Q P L D N L * P S Y K * S L H V G S *
 - S N R W T I F N Q V I N S L F M L V V R
 - A T A G Q S L T K L * I V S S C W * L D
 16141 - ACATAGTATGCCTCTTAACTACAAGTAAGAGTCTAATAAATGCCTTCCTCATCCTTCT - 16200
 - T * Y A S * L Q S K S L I N C L P H P S
 - H S M P L N Y K V R V * * I A F L I L L
 - I V C L L T T K * E S N K L P S S S F S
 16201 - CCTGGAAGCGACAGCAATTAGTTTTAGGAACTTTGCAAACCAGCACTTTTTTCGTTGT - 16260
 - P G S D S N * F L G T L Q N Q H F F R C
 - L E A T A I S F * E L C K T S T F F V V
 - W K R Q Q L V F R N F A K P A L F S L *
 16261 - AAATATCAAAGCCCTGTAGACGACATCAGTACTAGTGCCTGTGCCGCACGGTGAACAC - 16320
 - K Y Q K P C R R H Q Y * C L C R T V * D
 - N I K S P V D D I S T S A C A A R C K T
 - I S K A L * T T S V L V P V P H G V R R
 16321 - GGGCTGCACCTACACCGCAAACCCGTTAAAAACGTTGATGCATCCGCAGACTGCATCAA - 16380
 - G L H L H R K P V * K R * C I R R L H Q
 - G C T Y T A N P F K N V D A S A D C I K
 - A A L T P Q T R L K T L M H P Q T A S R
 16381 - GGGTTCCGGGAGTTGGTCACAACACTACGCCATAACCTTCCACATTCGCCAGCAGGTACA - 16440
 - G F A E L V T T T A I T F P H S A D G T
 - G S R S W S Q L Q P * P F R I P Q T V Q
 - V R G V G H N Y S H N L S T F R R R Y R
 16441 - GACTGTGTTTCTAAGTGTAAAACCCACTGGGTCAATTAGCACAAAGTGGTAGGTATTTGGAC - 16500
 - D C V S K C K T H W V I S T S G R Y L D
 - T V F L S V K P T G S L A Q V V G I W T
 - L C F * V * N P L G H * H K W * V F G R
 16501 - GACTTACCTTCAAGTCACAGAATCCTTTAGGATTTGGATGGTCAATGTGGCATCTACA - 16560
 - V L T F Q V T E S F R I W M V N V A S T
 - Y L P F K S Q N P L G F G W S M W H L Q
 - T Y L S S H R I L * D L D G Q C G I Y N
 16561 - ATACAGACAACATGAAGCACCACCAAGGACTCTTGGTCCATGTTAGCTTCTGGTGTAC - 16620
 - I Q T T * S T T K G L L V H V S F W C Y
 - Y R Q H E A P P K D S W S M L A S G V T
 - T D N M K H H Q R T L G P C * L L V L Q
 16621 - AGTAATTGCCTGTCTGTACCAGTGTGTACACAACATCTTACACAGTTGGTGTATGG - 16680
 - S N C L S C T S V C T Q H L H T V G D W
 - V I A C P V P V C V H N I F T Q L V I G
 - * L P V L Y Q C V Y T T S S H S W * L V
 16681 - TTGTCTCCACTTGCTAGGTAATCCTTATATGCTTAGCAGGGTCTACTGCAAAAGCACA - 16740
 - L S S T C * V I L I C F S R V Y C K S T
 - C P P L A R * S L Y A L A G S T A K A Q
 - V L H L L G N P Y M L * Q G L L Q K H R
 16741 - GAAGGAAAGCACAGTTGAATTGGCAGTACTTCTGTAGCATTCCAGCCTGAAGACGTAC - 16800
 - E G K H S * I G R Y F C S I S S L K T Y
 - K E S T V E L A G T S V A F P A * R R T
 - R K A Q L N W Q V L L * R F Q P E D V L

图 12 续


```

16801 - TGTAGCAGCTAAACTGCCAGCACCATACCTCTATTTAGGTTGTTTAAGCCTTTGATGAA - 16860
- C S S * T A Q H H T S I * V V * A F D E
- V A A K L P S T I P L F R L F K P L M K
- * Q L N C P A P Y L Y L G C L S L * * S
16861 - GTACAAGTATTTCACTTAGGCCCTTTGGTGTGTCTGTAACAAACCTACAAGGTGGTTC - 16920
- V Q V F H F R P F W C V C N K P T R W F
- Y K Y F T L G P F G V S V T N L Q G G S
- T S I S L * A L L V C L * Q T Y K V V P
16921 - CAGTCTGTGTAATTTACCTGTACCATCACTCTTAGGGAATCTAGCCATTGAGATC - 16980
- Q F C V N C T C T I T L R E S S P F E I
- S S V * I V P V P S L L G N L A H L R S
- V L C K L Y L Y H H S * G I * P I * D L
16981 - TTGGTGGTCTGATAGTAATGCCAGCACAAACCTACCTCCCTCGAATTGTTATAGTAGGC - 17040
- L V V * * * C Q H K P T S L R I V I V G
- W W S D S N A S T N L P P F E L L * * A
- G G L I V M P A Q T Y L P S N C Y S R Q
17041 - AAGTGCATTGTCATCAGTACAAGCTGTTGTGTGGTACCAGCCGCACAGGACATCTGTGC - 17100
- K C I V I S T S C L C G T S R T G H L S
- S A L S S V Q A V C V V P A A Q D I C R
- V H C H Q Y K L F V W Y Q P H R T S V V
17101 - TAGTCTACTGGACTCAGTTCATTATTCTGTAGTTTAACAGCTGAGTGGCTCTTAGAGC - 17160
- * C Y W T Q F I I L * F N S * V G S * S
- S A T G L S S L F C S L T A E L A L R A
- V L L D S V H Y S V V * Q L S W L L E L
17161 - TGTAACAATAAGAGGCCAAGCCAAATTTGGTGAATTGTCCATGTTAATTTCACTAAGTTG - 17220
- C N N K R P S Q I W * I V H V N F T K L
- V T I R G Q A K F G E L S M L I S L S *
- * Q * E A K P N L V N C P C * F H * V E
17221 - AACAACTTTGCTATCCGCATCAACAACCTGCTGGATTCCAGAGTGAGATGCATATGT - 17280
- N N L A I R I N N L L D F P E C R C I C
- T I L L S A S T T C W I S Q S A D A Y V
- Q S C Y P H Q Q L A G F P R V Q M H M *
17281 - AAAGTGTACCATCACAAGTGTCTGTAGGTACCATAATCAGGGACAACAACCATGAG - 17340
- K G V T I T S V L V G T I I R D N N H E
- K V L P S Q V F L * V P * S G T T T M S
- R C Y H H K C S C R Y H N Q G Q Q P * V
17341 - TTTGGCTGCTGTAGTCAATGGTATGATGTTGAGTGGAAACCAACCATCAGCGCATTGTT - 17400
- F G C C S Q W Y D V E W N T T I T R I V
- L A A V V N G M M L S G T Q P S R A L L
- W L L * S M V * C * V E H N H H A H C *
17401 - GATAATGTTGTTAAGTGCATCATTATCAAGCTTCCTAAGCATAGTGAAGAGCATTGTTG - 17460
- D N V V K C I I I K L P K H S E E H C L
- I M L L S A S L S S F L S I V K S I V C
- * C C * V H H Y Q A S * A * * R A L F A
17461 - CATAGCACTAGTTACTTTTGCCTCTTGTCTCAGATCTTGCCTGTTTGTACATTTGGGT - 17520
- H S T S Y F C P L V L R S C L F V H L G
- I A L V T F A L L S S D L A C L Y I W V
- * H * L L L P S C P Q I L P V C T F G S
17521 - CATAGCCTGATCTGCCATCTTTCCAACCTGCGTTGCATGGCAGCATCACGGTCAAACCTC - 17580
- H S L I C H L F Q L A L H G S I T V K L
- I A * S A I F S N L R C M A A S R S N S
- * P D L P S F P T C V A W Q H H G Q T Q
17581 - AGATTTAGCCACATTCAAAGATTTCTTTAACTTTTTGAGAACGACTTCAGATCACCATT - 17640
- R F S H I Q R F L * L F E N D F R I T I
- D L A T F K D F F N F L R T T S E S P L
- I * P H S K I S L T F * E R L Q N H H *

```

图 12 续

```

17641 - AGCTACAGCCTGCTCATAGGCTCCTGGGCAGTGGCATAAGCGGCATATGATGGTAAAGA - 17700
- S Y S L L I G L L G S G I S G I * W * R
- A T A C S * A S W A V A * A A Y D G K E
- L Q P A H R P P G Q W H K R H M M V K N
17701 - ACTAAATTCTGAAGCAATAGCCTGAAGAGTAGCACGGTTATCGAGCATTTCCTCGCACAA - 17760
- T K F * S N S L K S S T V I E H F L A Q
- L N S E A I A * R V A R L S S I S S H N
- * I L K Q * P E E * H G Y R A F P R T
17761 - CCTATTAATGTCTACAGCACCTGCATGGATAGCAAACAGACAAAAGAGAAACCATCTT - 17820
- P I N V Y S T L H G * Q N R Q K R N H L
- L L M S T A P C M D S K T D K R E T I F
- Y * C L Q H P A W I A K Q T K E K P S S
17821 - CTCGAAAGCTTCAGTTGTGCTTTTTCGAAGAAGAATATCATTGTGGAGTTGTACACATTG - 17880
- L E S F S C V F C K K N I I V E L Y T L
- S K A S V V S F A R R I S L W S C T H C
- R K L Q L C L L Q E E Y H C G V V H I V
17881 - TGCCCACAATTTAGAAGATGACTCTACTCTAAGTTGTTGAAGAACCAGAGCAGTACCAC - 17940
- C P Q F R R * L Y S K L L K N R E Q Y H
- A H N L E D D S T L S C * R T E S S T T
- P T I * K M T L L * V V E E P R A V P Q
17941 - AGATGTGCACTTTAGCTCAGACATTTTAGACTGTACAGTAGCAACCTTGATACATGGTTT - 18000
- R C A L Y V R H F R L Y S S N L D T W F
- D V H F T S D I L D C T V A T L I H G L
- M C T L R Q T F * T V Q * Q P * Y M V Y
18001 - ACCTCCAATACCCAACAACCTAATGTTAAGCTTGAAGCATCAATACTACTCTTAGGAGG - 18060
- T S N T Q Q L N V K L E S I N T T L R R
- P P I P N N L M L S L K A S I L L L G G
- L Q Y P T T * C * A * K H Q Y Y S * E A
18061 - CAAAAGCCCCTGGGAGTTCATATACCTAAATTCCTGTGTAGAGACCAAGTAGTCATAAAC - 18120
- Q K P L G V H I P K F L C R D Q V V I N
- K S P W E F I Y L N S C V E T K * S * T
- K A P G S S Y T * I L V * R P S S H K H
18121 - ACCAAGAGTAAGCCTGAAGTAACGGTTAGTAACAGAAAAGGCCAAAGTAGCAGCAGCA - 18180
- T K S K P E V T V E * T E K A K V A A A
- P R V S L K * R L S K Q K R P K * Q Q Q
- Q E * A * S N G * V N R K G Q S S S S N
18181 - ACAATAGCCTAAGAAACAATAACAAGCATGATACACTGTAAGGTGTTGCCAGTATAAAA - 18240
- T I A * E T I N K H D T L * G V A S N K
- Q * P K K Q * T S M I H C K V L P V I N
- N S L R N N K Q A * Y T V R C C Q * * I
18241 - TAACAATGGGTAATACTCAACACACAAAACACTATAGCTCTAGCTAAAACATGATAGT - 18300
- * Q W V I L N T H K H Y S S S * K H D S
- N N G * Y S T H T N T I A L A K N M I V
- T M G N T Q H T Q T L * L * L K T * * S
18301 - CGTAACGACACCAGAATAGTTAGAGGTTACAGAAATAACTAAGGCCACATGGAATAGC - 18360
- R N D T R I V R G Y R N N * G P H G N S
- V T T P E * L E V T E I T K A H M E I A
- * R H Q N S * R L Q K * L R P T W K * L
18361 - TTGATCTAAAGCATTACCATAGTAGACTTTGTAACAAGTGTAAATGACATTCATCAGTGT - 18420
- L I * S I T I V D F V N K C N D I H Q C
- * S K A L P * * T L * T S V M T F I S V
- D L K H Y H S R L C K Q V * * H S S V S
18421 - CCAACAGCTCTAGCAGCATCATATAACAGTGCAGCTGTGCATGAGAATAAGCAAAC - 18480
- P N T S S S I I I N S A S C H E N K Q N
- Q T R L A A S S * T V R A V M R I S K T
- K H V * Q R H H K Q C E L S * E * A K L

```

图 12 续

```

18481 - TAAAGCTGAAGCATAACACAAATCCTTAAGCCTATAACCAGACAAGCTAGTGTGAGC - 18540
- * S * S I H N T I L K P I T R Q A S V S
- K A E A Y I T Q S L S L * P D K L V S A
- K L K H T * H N P * A Y N Q T S * C Q P
18541 - CAATCAAGCCATGTCATGATACGCATCAGCCAGCTAGCAGGCATGTAGACCATATAAA - 18600
- Q F K P C H D T H H P A S R H V D H I K
- N S S H V M I R I T Q L A G M * T I L K
- I Q A M S * Y A S P S * Q A C R P Y * S
18601 - GTAAGCAACTGTTGCAAGAGAAGGTAACAGAAACAAGCACAGAATGCGTGCTTATGCTT - 18660
- V S N C C K R R * Q K Q A Q E C V L M L
- * A T V A R E G N R N K H K N A C L C L
- K Q L L Q E K V T E T S T R M R A Y A *
18661 - AACAGCAGCATAGCACATGCAGCAATTGCCATAATACCAAGAGTAAATGGCAAGAAAGC - 18720
- N K Q H S T C S N C H N T K S K W Q E S
- T S S I A H A A I A I I P R V N G K K A
- Q A A * H M Q Q L P * Y Q E * M A R K H
18721 - ATTCTCGTAAACAAAGAAAAACAGTGACCACTGTGTACTTTGAACAAGAATCAATAGTGA - 18780
- I L V N K E K Q * P L C T L N K N Q * *
- F S * T K K N S D H C V L * T R I N S D
- S R K Q R K T V T T V Y F E Q E S I V M
18781 - TGTCAGAAAGTTAAAGCATCCAATGATGAGTGCCCTTAACAATTTTCTTGAACCTACC - 18840
- C Q E S * K H P M M S A L N N F L E L T
- V K K V K S I Q * * V P L T I F L N L P
- S R K L K A S N D E C P * Q F S * T Y L
18841 - TTGGAAGGTAACACCAGAGCATTGTCTAACAAACATCAAATGGTGTAAACTCATCTTCTAA - 18900
- L E G N T R A L S N N I K W C K L I F *
- W K V T P E H C L T T S N G V N S S S K
- G R * H Q S I V * Q H Q M V * T H L L K
18901 - AATAGTGCTACCAAGGATAGTACGACCATTACATACCATTCTGCAGCAGCTCTTTCAAAGC - 18960
- N S A T K D S T T I H T I L Q Q L F O S
- I V L P R I V R P F I P F C S S S F K A
- * C Y Q G * Y D H S Y H S A A A L S K Q
18961 - AGCACACATATCTAAGACGGCAATTCCTGTTTGAGCAGAAAGAGGTCCARTATGTCAAC - 19020
- S T H I * D G N S C L S R K R S Q Y V N
- A H I S K T A I P V * A E R G P N M S T
- H T Y L R R Q F L F E Q K E V P I C Q H
19021 - ATGATCTTGTGTCAAAGGTTATAGTTGTACTTCATTGCCACAAGGTTAAAGTCATTCAA - 19080
- M I L C Q R F I V V L H C H K V K V I Q
- * S C V K G S * L Y F I A T R L K S F K
- D L V S K V H S C T S L P Q G * S H S K
19081 - AGTAGTGGTGAATCTATTAAGAAACCACCTATCACCATTGATAACAGCAGCATAACAGCCA - 19140
- S S G E S I K K P P I T I D N S S I Q P
- V V V N L L R N H L S P L I T A A Y S H
- * W * I Y * E T T Y H H * * Q Q H T A M
19141 - TGCCAAAACATTTAATGTTATGGTTGTGTCTGTACCTGCAGCCTGTGCAGTTTGTCTGTC - 19200
- C Q N I * C Y G C V C T C S L C S L S V
- A K T E N V M V V S V P A A C A V C L S
- P K H L M L W L C L Y L Q P V Q F V C Q
19201 - AACAAATGGACCATAGAATTTACCTTCTAAGTCAGTACCAGCGTGTACTCCTGTTGGAAG - 19260
- N K W T I E F T F * V S T S V Y S C W K
- T N G P * N L P S K S V P A C T P V G S
- Q M D H R I Y L L S Q Y Q R V L L E A
19261 - CTCCATGATGCATATAGCAGAAAGACAGCAATCATAATCAATGTTAAAACCAACACT - 19320
- L H M M H I A E R H A I I I N V K T N T
- S I * C I * Q K D T Q S * S M L K P T L
- P Y D A Y S R K T R N H N Q C * N Q H Y

```

图 12 续

```

19321 - ACCACATGATCCATTAAGGAAGAACCTTTAATGGTATGATTAGGTCATGGCACACTG - 19380
- T T * S I K E R T F N G M I R S H G T L
- P H D P L R K E P L M V * L G L M A H *
- H M I H * G K N L * W Y D * V S W H T D
19381 - ATAAACACCAGATGGTGAACCATTTGATGATGCTAGAACTGAAATGTTTGACCAGGTTG - 19440
- I N T R W * T I V A C * N * K C L T R L
- * T P D G E P L * H A R T E N V * P G W
- K H Q M V N H C S M L E L K M F D Q V G
19441 - GATACGGACAAATTTATACTGGGTGCTTAGGGTTAGAAGTATCACTTTAAGCCTAAG - 19500
- D T D K F I L G C L R V R S I N F K P K
- I R T N L Y L G V L G L E V S T L S L S
- Y G Q I Y T W V S * G * K Y Q L * A * A
19501 - CAGACATTTTGCATAGAATGCCAATAACACGAAGTTGAACATTGCCAGCCTGAACAAG - 19560
- Q T I L H R M A N N T K L M I A S L N K
- R O F C I E W P I T R S * T L P A * T R
- D N F A * N G Q * H E V B H C Q P E Q E
19561 - AAAGCTATGGTTGGATTGGCAATGAGCAGATCTTCATAGTTAGGATTAGCATGTCTTC - 19620
- K A M V G F A N E Q I F I V R I K H V F
- K L W L D L R M S R S S * L G L S M S S
- S Y G W I C E * A D L H S * D * A C L L
19621 - TGCTGTGCAAATGACATGTCTTGGACAGTATACTGTGTCCATCCAACCAATCCATTAAG - 19680
- C C A N D M S W T V Y C V I Q P Q S I K
- A V Q M T C L G Q Y T V S S N R N P L R
- L C K * H V L D S I L C H P T T I H * E
19681 - AGTTGTAGTTCCACAGGTTACTGTACCATGCACCCTTCACTTTGCCTGACGGGAATGC - 19740
- S C S S T G Y L Y H A P F N F A * R E C
- V V V P Q V T C T M H P S T L P D G N A
- L * F H R L L V P C T L Q L C L T G M P
19741 - CATTTCCTAAAACCACTCTGCAGAACAGCAGAAGTGATGTCTGTGGTGGTTGGTA - 19800
- H F P K T T L Q N S R S D * C L W W L V
- I F L K P L C R T A E V I D V C G G W *
- F S * N H S A E Q Q K * L M S V V V G R
19801 - GAGAACATCAGCACCTGAGTTGCTAAAGTCATTTAGAGCCTTTGCTAAGTGGCAGCAAGC - 19860
- E N I S T * V A K V I * S L C * V A A S
- R T S A P E L L K S F R A F A K W Q Q A
- E H Q H L S C * S H L E P L L S G S K L
19861 - TGCTTCAGATAGCTGGTAGTATCTAAGGCTCCACTGAAATACTTGTACTTGTATATAG - 19920
- C F T I A G S I * G S T E I L V L V I *
- A S R * L V V S K A P L K Y L Y L L Y R
- L H D S W * Y L R L H * N T C T C Y I E
19921 - AGCAAGATACCTGTTATACTGTGTAAGTGGCAACAGTGTCTCGCTACGCAATTTAGGTA - 19980
- S K I P V I L C K W Q Q C L A T Q F * V
- A R Y L L Y C V S G N S V S L R N F R Y
- Q D T C Y T V * V A T V S R Y A I L G T
19981 - CATTTCCTTGTGAGCAAAAAGGTACACAAAGCAGCCTCCTCGAAGGTACTAAATGTAAAC - 20040
- H F L V E Q K G T Q S S L L E G T K C N
- I S L L S K K V H K A A S S K V L N V T
- F P C * A K R Y T K O P P R R Y * M * L
20041 - TCCATTAAACATGACTCTTTTCCTAAGATAGTTGTTAAAGAACCAATGGCAGTGCCTCAG - 20100
- S I K H D S F P K I V V K E P M A V L Q
- P L N M T L F L R * L L K N Q W Q C F R
- H * T * L F S * D S C * R T N G S A S E
20101 - AGAAATACAGAATACATAGATTGCTGTTATCCAAAAAGGCACAAATAGGAGAAAACATGGC - 20160
- R N T E Y I D C C Y P K R H N R R K H G
- E I Q N T * I A V I Q K G T I G E N M A
- K Y R I H R L L L S K K A Q * E K T W Q

```

图 12 续

20161 - AAACCATTGAAGGTGAGCCAAGAATGAAACATCATTGGTGAATAGAATGTCAAGTACAA - 20220
 - K P L K V S Q E * N I I G E I E C Q V Q
 - N H * R * A K N E T S L V K * N V K Y K
 - T I E G E P R M K H H W * N R M S S T S
 20221 - GTAAAAGACTGAGTAGACTCCCAGCAGAAAGCTGTAAGCTGGTACCAGACAGAGTATAGT - 20280
 - V K D * V D S R Q K A V S W Y Q T E Y S
 - * K T E * T P G R K L * A G T R Q S I V
 - K R L S R L P A E S C K L V P D R V * *
 20281 - GAAAGACATCAAAAACAAAAGTGCATTAGCAGCAACAACATGGTTGTACTCACCAAAAC - 20340
 - E R H Q K Q K C I S S N N M V V L T K N
 - K D I K N K S A L A A T T W L Y S P K T
 - K T S K T K V H * Q Q Q H G C T H Q K H
 20341 - ACGTCTGAATTCATAAAGTAGTAGGCAGCACAAGTCACCAATATGGCAATAATACCACC - 20400
 - T S E F H K V V G S T S H Q Y G N N T T
 - R L N F I K * * A A Q V T N M A I I P P
 - V * I S * S S R Q H K S P I W Q * Y H Q
 20401 - AGCCACTACTGAAGCAGACACATCTAAGCACCACAGGTGCACAAGAGGAGTAAAGAT - 20460
 - S H Y * S R H I * S T H R L H K R S K D
 - A T T E A D T S K A P T G C T R G V K M
 - P L L K Q T H L K H P Q V A Q E E * R C
 20461 - GTTAGCTATGAGATTCATCGCATCAACACCACAGAAAACCTCGATAGAGCTCTGTAATG - 20520
 - V S Y E I H R I N T T E N S * * S S V M
 - L A M R F I A S T P Q K T P D R A L * C
 - * L * D S S H Q H H R K L L I E L C N A
 20521 - CTCATTATTAAGAACCCTACTACCCTGGTAGATAGGCAAATACCTACTTCTGACCTTTC - 20580
 - L I I K N P S T T G R * A N T Y F * P F
 - S L L R T H L P L V D R Q I P T S D L S
 - H Y * E P I Y H W * I G K Y L L L T F R
 20581 - GCATGTACCATGTCTACAGTACTCAGCATCAAAAGTTGTTACTACTCTAACAGAACCCCTC - 20640
 - A C T M S T V L S I K S C Y Y S N R T L
 - H V P C L Q Y S A S K V V T T L T E P S
 - M Y H V Y S T Q H Q K L L L L * Q N P P
 20641 - CAGGTAAGTGTAGGAAACTGTATGATGGAACCATCCATAAGCACATAACGAGTGTCTGG - 20700
 - Q V S V R K L Y D G T I H K H I T S V W
 - R * V L G N C M M E P S I S T * R V S G
 - G K C * E T V * W N H P * A H N E C L D
 20701 - ACGAAGCTCACTATAAGAAATAGAACCCCTCTAGCAAATTAGTGTCAACAAATATGGCAC - 20760
 - T K L T I R N R T L * Q I S V I T I W H
 - R S S L * E I E P S S K L V S * Q Y G T
 - E A H Y K K * N P L A N * C H N N M A Q
 20761 - AGGTTGCCCATAGCATCCTTAAAAATTGTACTACTCAGCAGCAAGAACGCAAGCAGAGGT - 20820
 - R F A H S I L K N C T L S S K N A S R G
 - G L P I A S L K I V H S A A R T Q A E V
 - V C P * H P * K L Y T Q Q Q E R K Q R *
 20821 - AGCAAATCACTATACTCAATGAGTTGGAAGGTGTGTAGCAAATGTTGCCAACAGCACT - 20880
 - S K I T I L N E F G R C V A N V A N S T
 - A K S L Y S M S L E G V * Q M L P T A L
 - Q N H Y T Q * V W K V C S K C C Q Q H *
 20881 - AAAAACACGAGGTAGAAAATGCAAGAAGTCACCATTGATTGCTCTCAGCACAGTACCCGG - 20940
 - K N T R * K M Q E V T I D C S Q H S T R
 - K T R G R K C K K S P L I A L S T V P G
 - K H E V E N A R S H H * L L S A Q Y P V
 20941 - TAAGCCAGGCACTATGAAACCAATCTCTTGTATGATAGCAGCTACTACAGGGCAGCT - 21000
 - * A R H Y E T N L S C N D S S Y Y R A A
 - K P G T M K P I S L V M I A A T T G Q L
 - S Q A L * N Q S L L * * * Q L L Q G S F

图 12 续

21001 - TTTGTCATTTTGTATGAACCACCACGCTGGCTAAACCATGCGTCAAACCAGCATGTTT - 21060
 - F V I F V * T T T L A K P C V K T S M F
 - L S F L Y E P P R W L N H A S K P A C L
 - C H F C M N H H A G * T M R Q N Q H V Y

21061 - ATTTGCAAAACAATCATCAGTAGAAATGATGTCACGAGTGACACCATCCTGAATGGCTTT - 21120
 - I C K T I I S R N D V T S D T I L N G F
 - F A K Q S S V E M M S R V T P S * M A L
 - L Q N N H Q * K * C H E * H H P E W L C

21121 - GTAACCAATGATTTTCATTTGTGTAACCATCATGGATTGACAATGTATGTACTGGCATAAC - 21180
 - V T N D F I C V T I M D * Q C M Y W H N
 - * P M I S F V * P S W I D N V C T G I T
 - N Q * F H L C N H H G L T M Y V L A * R

21181 - GATATAACAAACCAATGCAGCAAGAACGCACAATAATGTGGCCTTAAGCATAAGTTTAA - 21240
 - D I T N Q C S K N A O * C G L K H K F K
 - I * Q T N A A R T H N N V A L S I S L K
 - Y N K P M Q Q E R T I M W P * A * V * N

21241 - ACAAGTACTRACAATCTTACCACCCTTGAGTGAGATTTTAGTAGTTATGACATTGACAAC - 21300
 - T S T N N L T T L E * D F S S Y D I D N
 - Q V L T I L P P L S E I L V V M T L T T
 - K Y * Q S Y H P * V R E * * L * H * Q P

21301 - CTGCTAGTTGTAGCACAAGTTAGTGTAAAAGGTATGTTGTTCTTCTGGCAGCAGTACG - 21360
 - L S S C S T S * C K R Y V V L L G S S T
 - C L V V A Q V S V K G M L F F L A A V R
 - V * L * H K L V * K V C C S S W Q Q Y E

21361 - AATTGTTTACGCAGCTGTTTCAGATAAAGACATGTAGTCTTTTACATTCCAGATGAGTGA - 21420
 - N L F T Q L F R * R H V V F Y I P D E *
 - I C L R S C S D K D M * S F T F Q M S E
 - F V Y A A V Q I K T C S L L H S R * V K

21421 - AACATTGTGACTTTTTGCTACTTGGGCATTGATATGCCTTGCATTACAGTCAATACATGC - 21480
 - N I V T F C Y L G I D M P C I T V N T C
 - T L * L F A L T W A L I C L A L Q S I H A
 - H C D F L L L G H * Y A L H Y S Q Y M R

21481 - GCCAAGATCTCTGGGCGTCATGTTTTCAACCTTATTATAGGTGAGCATGAAATGTTACA - 21540
 - A K I S G R H V F N L I I G E H E I V T
 - P R S L G V M F S T L L * V S M K L L Q
 - Q D L W A S C F Q P Y Y R * A * N C Y N

21541 - ACTGTCACCTGTCACCTCTAAGTCAGAGTGATGTGAAAGTTTGAGACATTCAATAACATC - 21600
 - T V T C H F * V R V M * K F E T F N N I
 - L S P V T S K S E * C E S L R H S I T S
 - C H L S L L S Q S D V K V * D I Q * H P

21601 - CTTTGTGTCACATCGGTATCAACAACACCTTGTGGGCGCTGACACGAATGTAGAAAG - 21660
 - L C V N I G I N N T L S G S * H E C R K
 - F V S T S V S T T P C R A A D T N V E R
 - L C Q H R Y Q Q H L V G Q L T R M * K G

21661 - GACACCATCTAAAGCTACACCCTTGTGCTAACTCGCTGTGAGCTGTAGCAACAAGTGCCTT - 21720
 - D T I * S Y T L C * L A V S C S N K C L
 - T P S K A T P F A N S L * A V A T S A L
 - H H L K L H P L L T R C E L * Q Q V P *

21721 - AAGTTTTCCATAGGAACACTAAAAGTTGCTGAAAAGGTGTCGACATAAGCATCAAACAT - 21780
 - K F F H R N T K S C * K G V D I S I K H
 - S F S I G T L K V A E K V S T * A S N I
 - V F P * E H * K L L K R C R H K H Q T S

21781 - CTTAACGGAACCTTCAGTACTCTCCAACGTTTGATACAGAGCTTGGTCAAGCAACAG - 21840
 - L N G N F S T I S N V * Y K S L V K Q Q
 - L T E T S V L S P T F D T R A W S S N R
 - * R K L Q Y Y L Q R L I Q E L G Q A T E

图 12 续

```

21841 - AATAGGTTGGCACATCAGCTGACTGTAGTACACAGAAGCAGACTTAGAAGCAGACTCGTC - 21900
- N R L A H Q L T V V H R S R L R S R L V
- I G W H I S * L * Y T E A D L E A D S S
- * V G T S A D C S T Q K Q T * K Q T R R
21901 - GCATTGGACTTGCCATCAAAAATGACATTAATAGGCAGTGAACCTTTAGTGTGT - 21960
- A F G L A I K N Y D I N R Q * T F S V V
- H L D L P S K T M T L I G S E P L V L L
- I W T C H Q K L * H * * A V N L * C C *
21961 - AGCTCTCAAATTGTCTAAATGACAAAATGGGAGAGCGGATGTCTCTCATAGGTCCTTG - 22020
- S S Q I V * I D K M G E R M S L I G L L
- A L K L S K L T K W E S G C L S * V F *
- L S N C L N * Q N G R A D V S H R S F D
22021 - ACCAGCCTTGTCAAAGTAGAGGTGAAGCGGCCATTTTTCACAGCAACTATCAACAAT - 22080
- T S L V K V E V K R A I F H S N T I N N
- P A L S K * R * S A P F F T A T L S T I
- Q P C Q S R G E A R H F S Q Q H Y Q Y
22081 - ATACGATGACTGGTCAGTAGGTTGATGGTCTTTAAACTGGAGTGACAAATCAGGAGC - 22140
- I R * L V S R V D W S F K L E * Q I T S
- Y D D W S V G L I G L L N W S D K S R A
- T M T G Q * G * L V F * T G V T N H E Q
22141 - AACTTCATCACTAATGAATGTACTACCAGTGCAAATGTGTCAATGAGACAATCCA - 22200
- N F I T N E C T T S A K C V T I E T I P
- T S S L M N V L P V Q N V S Q L R Q F Q
- L H H * * M Y Y Q C K M C H N * D N S N
22201 - ATTGTGAGTCTTGACAGAAGCCAGGCCCTCCATTTGCATAGACATAGAAAGATCTCTCAT - 22260
- I V S L A E A T A S I C I D I E R S L H
- L * V L Q K P R P P F A * T * K D L F M
- C E S C R S H G L H L H R H R K I S S C
22261 - GCCATTAACAATAGTTGTACTCAACGGTGTGGCAGGATGCGCTTATAGCACATCAT - 22320
- A I N N S C T L N A C G T I A L I A H H
- P L T I V V H S T R V A R L R L * H I M
- H * Q * L Y T Q * R V W H D C A Y S T S C
22321 - GCAAGTCGAAGAGGTGCAACCATCCATGATATGAACATAGCTCTCCATATGTAGTAGAA - 22380
- A S R R G A T I H D M N I A L P Y V V E
- Q V E E V Q P S M I * T * L F H M * * K
- K S K R C N H P * Y E H S S S I C S R K
22381 - AGAAGCAAAGAAGATGTACATCCTAACCATTCGAGAAACGGGTGCCATTTGTACAATACT - 22440
- R S K E D V H P N H C R N G C H L Y N T
- E A K K M Y I L T I A E T G A I C T I L
- K Q R R C T S * P L Q K R V P F V Q Y *
22441 - AATGATAAACCACATGAGCCAGAATTGCTGATGAAATGACTAGCAAATAGCCAAAGAA - 22500
- N D K P H E P R I A D E M T S K I A K E
- M I N H M S Q E L L M K * L A K * P K N
- * * T T * A K N C * * N D * Q N S Q R T
22501 - CACCTGCATTATAGCTGAAAGACCTAATAAATAAAGAATTTGTGAACAACATATATGC - 22560
- R L H Y S * K T * * I K E F C E Q H I C
- T C I I A E R P N K * K N F V N N I Y A
- P A L * L K D L I N K R I L * T T Y M P
22561 - CAAACCCACTCAGCGGCCAGACCTAAAATTGTCAAGTCTAGCTTGTACGATGAAATCGT - 22620
- Q N P L S G Q T * N C Q V * L V R * N R
- K T H S A A R P K I V K S S L Y D E I V
- K P T Q R P D L K L S S L A C T M K S S
22621 - CACCTGAATGGTTCAAGAGCTGGATAAGAATCAAGGGAGTCTAATCCACTTAAACAAAT - 22680
- H L N G F K S W I R I K G V * S T * T N
- T * M V S R A G * E S R E S N P L K Q M
- P E W F Q E L D K N Q G S L I H L N K C

```

图 12 续

22681 - GCTGCAAGGAAAGAACCTTCACAGAAATCCATAGTAGTAACGTTAGACGAATTAAGATA - 22740
 - A A R K R T F T E I H S S N V R R I K I
 - L Q G K E P S Q K S I V V T L D E L R Y
 - C K E K N L H R N P * * * R * T N * D T
 22741 - CAATTCTTAACGCCATTACAATAAGAAGGAGCACCAAATTAGATAAGAGTACACCAA - 22800
 - Q F S N A I T I R R S T K I R * E Y T K
 - N S L T P L Q * E G A P K L D K S T P K
 - I L * R H Y N K K E H Q N * L R V H Q K
 22801 - AGCAGCAGTTACACAGATTAGAGAACCTAAGCAACTTAACAACAATAGCCACATAGC - 22860
 - S S S Y T D * R T * A N T * Q Q * P H S
 - A A V T Q I R E P K Q I L N N N S H I A
 - Q Q L H R L E N L S K Y L T T I A T * R
 22861 - GATTGTGAACAATTTAGAAAATTTGGGTGACTTCACATAATTAATGCCGGCATCAAACA - 22920
 - D C E Q F R K F G * L H I I N A G I Q T
 - I V N N L E N L G D F T * L M P A S K H
 - L * T I * K I W V T S H N * C R H P N I
 22921 - TAATTTAGCAACACTCTAACACTATTTTTAGCAATAGTTGTAGGTAGTGAAGCTCTAAT - 22980
 - * F S N T L N T I F S N S C R * * S S N
 - N L A T L L T L F L A I V V G S E A L I
 - I * Q H S * H Y F * Q * L * V V K L * F
 22981 - TCTAGAATGGTACTTTTAGTAAAAGTACACAATTGGAACAATAATGTAACACATAAGG - 23040
 - S R I G T F S K S T Q L E Q * C K H I R
 - L E L V L L V K V H N W N N N V N T * G
 - * N W Y F * * K Y T I G T I M * T H K A
 23041 - CATATAATGTTAAACACACGTTGTGCTAATCTCTTAGCGCAATTTGATGTTGTAATTGC - 23100
 - H I I V K H T L C * S L S A I * C C N C
 - I * L L N T R C A N L L A Q F D V V I A
 - Y N C * T H V V L I S * R N L M L * L L
 23101 - TGCTTGCTAAGAATGGTTTGACATAAGCCAAAATTTACTCCAAGGAACACTATAAT - 23160
 - C L S * E W F D I S Q N F T P R N T I N
 - A C P K N G L T * A K I L L Q G T L L I
 - L V L R M V * H K P K F Y S K E H Y * L
 23161 - TGCAGCAATACCATGAGTGGCAATGTTTTAAACCTAAGGCTAGTGAAGCTCATTAGG - 23220
 - C S N T M S G N C F * T * G * * K L I R
 - A A I P * V A I V F K P K A S E S S L G
 - Q Q Y H E W Q L F L N L R L V K A H * V
 23221 - TTTCTAATGGTAATGCTTGTGTTTTCCACATAAGCAGCCATAAGATCCTCATGACCTAA - 23280
 - F L N G N A C V F H I S S H K I L M T *
 - F L M V M L V F S T * A A I R S S * P N
 - S * W * C L C F P H K Q P * D P H D L T
 23281 - CTCTTGTGTTACTTTAACACCTTCATCTGATGGTTAAGTATGACATTGCCTACAACCTC - 23340
 - L L C Y F N T F I * W F K Y D I A Y N F
 - S C V T L T P S S D G L S M T L P T T S
 - L V L L * H L H L M V * V * H C L Q L R
 23341 - GGTAGTTTTACGTCACACTCTATGACTTCTTCTGTATGGTAGGATTTCCACTACTTC - 23400
 - G S F H V T L Y D F L L Y G R I F H Y F
 - V V F T S H S M T S F C M V G F S T T S
 - * F S R H T L * L P S V W * D F P L L L
 23401 - TTCAGAGGTGGTGTGTTGACTTTCACAAGCAAGATTGTCCATTCCTTGTGTCTTCTAC - 23460
 - F R G G L L T F T S K I V H S L C V F Y
 - S E V G C * L S Q A R L S I P C V S S T
 - Q R W V V D E H K Q D C P F L V C L L L
 23461 - TGCCAGAACTTCAATGAATTTGAAGTATCTACTGGCTTTGTACTCCAAGACACGTTAA - 23520
 - C Q N F K * I * S I Y W L C T P K T T *
 - A R T S N E F E V S T G F V L Q R Q R K
 - P E L Q M N L K Y L L A L Y S K D N V N

图 12 续

23521 - ACACCAAGTGTGGTTTGAACGTTGCTTGGTTGTAGCCTGGTTAATGTGCCAAACAAT - 23580
 - T P S V W F E R C L G C S L V N V P N N
 - H Q V F G L N V V L V V A W L M C Q T I
 - T K C L V * T L S W L * P G * C A K Q L
 23581 - TGGCTTATGCAGTAATTTAGCACCTTCTTGAARACTCGCTGAATAGTGTCTATAGTCAAT - 23640
 - W L M Q * F S T F L E T R * I V S I V N
 - G L C S N L A P F L K L A E * C L * S I
 - A Y A V I * H L S * N S L N S V Y S Q *
 23641 - AGCCACTACATCGCCATTCAAGTCTGGGAAGATGTGACAGATAGCTCTCGTGAAGCTGG - 23700
 - S H Y I A I Q V W E E C D R * L S * S W
 - A T T S P F K S G K N V T D S S R E A G
 - P L H R H S S L G R M * Q I A L V K L A
 23701 - CTTTGTGAAGCCTGTCAATTGATTTAATCATCAGCAAATTTGTGTTAGAACATGTGAG - 23760
 - L C E A C H L I * I I S K F C V R T C E
 - F V K P V I * F K S S A N F V L E H V S
 - L * S L S F D L N H Q Q I L C * N M * V
 23761 - TTTGAAATTATCAAACTCGCATTGGTAATGGTTGAGTTGGTACAAGGTCTATAGGCTG - 23820
 - F E I I K T R I W * W L S W Y K V Y R L
 - L K L S K L A F G N G * V G T R S I G C
 - * N Y Q N S H L V M V E L V Q G L * A A
 23821 - CTCTGTATAGTAAGCATTATCCTTTTTATAATACCCATCCAATTTGGTTCATCTCTGT - 23880
 - L C I V S I I L F I I P I Q F W F N L C
 - S V * * A L S F L * Y P S N F G S I S V
 - L Y S K H Y P F Y N T H P I L V Q S L C
 23881 - GTAAGTAACTCCATCGAGTTTATACGACACAGGCTTGATGGTTGTAGTGAAGATGTTTC - 23940
 - V S N S I E F I R H R L D G C S V R C F
 - * V T P S S L Y D T G L M V V V * D V S
 - K * L H R V Y T T Q A * W L * C K M F P
 23941 - CTTGTAGAAAACATCAGTCACTGGTCCTTTGTAAGGTGAGCTCC - 24000
 - L V E N I S H W S F V L * H L C K V S S
 - L * K T S V T G P L Y S D I F V R * A P
 - C R K H Q S L V L C T L T S L * G E L R
 24001 - GTCAATACGATAGAGGGTCTCCTTAGCAGTTATATGAGTGAATGACCACACTGATAGTT - 24060
 - V N T I E G L L S S Y M S V M T T L I V
 - S I R * R V S L A V I * V * * P H * * L
 - Q Y D R G S P * Q L Y E C N D H T D S Y
 24061 - ACCAGTGTACTCATTGCGACATAAGAATGTACCTTGCTGTAATTTATACTCAGCAGGTGG - 24120
 - T S V L I R T * E C T L L * F I L S R W
 - P V Y S F A H K N V P C C N L Y S A G G
 - Q C T H S H I R M Y L A V I Y T Q Q V V
 24121 - TGCAGACATCATAACAAAGAGACTCTTGTGTACTAGATATTGGTAGCATCAGGACC - 24180
 - C R H H N K R R L L L Y * I L C S I T T
 - A D I I T K E D S C C T R Y C V A S R P
 - Q T S * Q K K T L V V L D I V * H H D H
 24181 - ACACACACATGGAATGGAACACCTGTCTAAGATTATCATAAGATAGAGTACCCATATA - 24240
 - T H T W N G N T C L K I I I R * S T H I
 - H T H G M E T P V L R L S * D R V P I Y
 - T H M E W K H L S * D Y H K I E Y P Y T
 24241 - CATCAGCTTCTACACCCGTTAAGGTAGTATTTCTGACCACAATGTTACACACCAC - 24300
 - H H S F Y T R * G S S F L T T M F T H H
 - I T A S T P V K V V V F * P Q C L H T T
 - S Q L L H P L R * * F S D H N V Y T P H
 24301 - ATTAAGAACTCGCTTGCAGATCCAAATAGCATGCTGTAGAAGATGGTTCATAGTTTC - 24360
 - I K N S L C R F Q I S M L * K M G H S F
 - L R T R F A D S K L A C C R R W V I V S
 - * E L A L Q I P N * H A V E D G S * F L

图 12 续

24361 - TCTGACATCACCAAGCTCGCCAACAGTTTTATTACTGTAAGCGAGTATGAGTGCACAAAA - 24420
 - S D I T K L A N S F I T V S E Y E C T K
 - L T S P S S P T V L L L * A S M S A Q K
 - * H H Q A R Q Q F Y Y C K R V * V H K S
 24421 - GTTAGCAGCATCACCAGCAGGGCTCTATAATAAGCCTCTTGAAGTGTGGTGCATTGAA - 24480
 - V S S I T S T G S I I S L L K C W C I E
 - L A A S P A R A L * * A S * S A G A L N
 - * Q H H Q H G L Y N K P L E V L V H * I
 24481 - TTTGACTTCAAGCTGTTGAAGTGTCTATAAAACACTAGACAAATAACAATTGTTATCAGC - 24540
 - F D F K L L K C * * N T R Q I T I V I S
 - L T S S C * S A N K T L D K * Q L L S A
 - * L Q A V E V L I K H * T N N N C Y Q P
 24541 - CCATTTAATGAAGTTAAACCACCAACTTGAGGAAATTTCCATTCTTTGTGTGGTTTAA - 24600
 - P F N * S * T T N L R K F P F L C V V *
 - H L I E V K P P T * G N F H F F V W F K
 - I * L K L N H Q L E E I S I S L C G L K
 24601 - AGCAGACATGTACCTACCAAGAAAACCTCTCATCAAGAGTATGGTAGTACTCGAAAGCTTC - 24660
 - S R H V P T K K T L I K S M V V L E S F
 - A D M Y L P R K L S S R V W * Y S K A S
 - Q T C T Y Q E N S H Q E Y G S T R K L H
 24661 - ACTACGTAGTGTGCATCACTAGGTAGTACAAAGAAAGTCTTACCCTCATGATTTACATG - 24720
 - T T * C V I T R * Y K E S L T L M I Y M
 - L R S V S L G S T K K V L P S * F T *
 - Y V V C H H * V V Q R K S Y P H D L H E
 24721 - AGSTTTAATTTTGTAAACATCAGCACCATCCAAGTATGTTGGACCAAACCTGCTGTCCATA - 24780
 - R F N F C N I S T I Q V C W T K L L S I
 - G L I F V T S A P S K Y V G P N C C P Y
 - V * F L * H Q H H P S M L D Q T A V H M
 24781 - TGTCATAGACATATCCACAAGCTGTGTGTGGAGATTAGTGTGTGCCACAGTTGTGAACAC - 24840
 - C H R H I H K L C V E I S V V H S C E H
 - V I D I S T S C V W R L V L S T V V N T
 - S * T Y P Q A V C G D * C C P Q L * T L
 24841 - TTTTATAGTCTTAACCTCCCGCAGGGATAAGAGACTCTTTAGTTTGTCAAGTGAAGAAC - 24900
 - F Y S L N L P Q G * E T L * F V K * K N
 - F I V L T S R R D K R L F S L S S E R T
 - L * S * P P A G I R D S L V C Q V K E P
 24901 - CTCACCGTCAAGATGAAACTCGACGGGGCTCTCCAGAGTGTGGTACACAATTTGTCCACC - 24960
 - L T V K M K L D G A L Q S V V H N F V T
 - S P S R * N S T G L S R V W Y T I L S P
 - H R Q D E T R R G S P E C G T Q F C H H
 24961 - ACGCTTAAGAATTCACACCTAAGTGTGTACGGTGTCTGATAGGACCAATCTCTGTA - 25020
 - T L K K F N T * L C T L S * I G P I S V
 - R L R N S T P N S V R C P E * D Q S L *
 - A * E I Q H L T L Y A V L N R T N L C K
 25021 - AGAGCCAGCCAAAGAAACTGTTTCTACAAAGTGTCTCCTCAGATGTCTTTGATGACGAAGT - 25080
 - R A S Q R N C F Y K V L L R C L * * R S
 - E P A K E T V S T K C S S D V F D D E V
 - S Q P K K L F L Q S A P Q M S L M T K *
 25081 - GAGGTATCCATTATATGTAGTAACAGCATCTGGTGTGATACTGACACTACGGCAGGAGC - 25140
 - E V S I I C S N S I W * * Y * H Y G R S
 - R Y P L Y V V T A S G D D T D T T A G A
 - G I H Y M * * Q H L V M I L T L R Q E L
 25141 - TTTAAGAGAACGCATACAGCGCAGCCTCTTCAAGATTAAACCATGTGTACATAACC - 25200
 - F K R T H T A R S L F K I K T M C H I T
 - L R E R I Q R A A S S R L K P C V T * P
 - * E N A Y S A Q P L Q D * N H V S H N O

图 12 续

25201 - AATTGGCATTGTGACAAGCGGCTCATTAGAGAGTTCAGCTTCGTAATAATAGAAGCTAC - 25260
 - N W H C D K R L I * R V Q L R N N R S Y
 - I G I V T S G S F R E F S F V I I E A T
 - L A L * Q A A H L E S S A S * * * K L Q

25261 - AGGCTCTTTACTAGTATAAAGAAGAATCGGACACCATAGTCAACGATGCCCTTGAAT - 25320
 - R L F T S I K E E S D T I V N D A L L N
 - G S L L V * K K N R T P * S T M P S * I
 - A L Y * Y K R R I G H H S Q R C P L E F

25321 - TTTAATTCCTTATACTTACGTTGGATGGTTGCCATTATGGCTCTAACATCCATGCATAT - 25380
 - F N S F I L T L D G C H Y G S N I H A Y
 - L I P L Y L R W M V A I M A L T S M H I
 - * F L Y T Y V G W L P L W L * H P C I *

25381 - AGGCATTAATTTCTTGTCTCTTCAGCATGAGCAAGCATTCTCTCAAATCCAGGATAC - 25440
 - R H * F S C L F S M S K H F S Q I P G Y
 - G I N F L V S S A * A S I S L K F Q D T
 - A L I F L S L Q H E Q A F L S N S R I Q

25441 - AGTTCCTAGAACTCTTCTTAGCATTAGGTGCTTCTGAAGGTAGTACATAAATGCAGA - 25500
 - S S * N L F L S I R C F * R * Y I K C R
 - V P R I S S L A L G A S E G S T * N A D
 - F L E S L P * H * V L L K V V H K M Q I

25501 - TTTCATTTCTTAAGAGCAGTCTTAGCTTCTCAAGTGATAACCAGCACATCCTTGTCC - 25560
 - F A F L K S S L S F L K C I T S T S L S
 - L H F L R A V L A S S S V * P A H P C P
 - C I S * E Q S * L P Q V Y N Q H I L V Q

25561 - AGGTACGTGGTTATATACTCATCAACTGGCACTTTCTTCAAAGCTCTTGAGAGCATCTC - 25620
 - R V R G Y I L I N W H F L Q S S * E H L
 - G Y V V I Y S S T G T F F K A L E S I S
 - G T W L Y T H Q L A L S S K L L R A S Q

25621 - AGTAGTGCCACCAGCCTTTTGGAGGGTATTACAACAAGTGATATCACCACCTAGTGAT - 25680
 - S S A T S L F G G Y Y N T S D I T T S D
 - V V P P A F L E G I T T Q V I S P L V I
 - * C H Q P F W R V L Q H K * Y H H * * *

25681 - AACATCACCTACCATGTAAAGGTGCATCCTTCTCAAGGAAAGACATATCTTCACCTCTAAG - 25740
 - N I T Y H V R C I L L K E R H I F T S K
 - T S P T M * G A S F S R K D I S S P L S
 - H H L P C K V H P S Q G K T Y L H L * A

25741 - CATGTTCTGAGAATCATGGTAAAGCTTACCATTGATATCAGCAAACAAGAGTAACTTATT - 25800
 - H V L R I M V K L T I D I S K Q E * L I
 - M F * E S W * S L P L I S A N K S N L L
 - C S E N H G K A Y H * Y Q Q T R V T Y W

25801 - GGTAAGAACTTAGTTTTCTCCAGTGTGTGGTAACCTCATCAATGCAGGCCCTTAATTTT - 25860
 - G K K L S F F Q C C G N L I N A G L N F
 - V R N L V S S S V V V T S S M Q A L I F
 - * E T * F L P V L W * P H Q C R P * F L

25861 - TGGCTTCACATCGACAGGCTTCTGTACGACAGATTTCTCCTCAGTTTTGGAATCTTCTGT - 25920
 - W L H I D R L L Y D R F L L S F G I F C
 - G F T S T G F C T T D F S S V L E S S V
 - A S H R Q A S V R Q I S P Q F W N L L C

25921 - GTTTGGTGGCTCCTCTTGTAGGTGCTTCCACTCTAGGCTTCAGGTATCAAGATAATC - 25980
 - V W W L L L F R C F H S R L Q V I K I I
 - F G G S S C L G A S T L G F R L S R * S
 - L V A P L V * V L P L * A S G Y Q D N P

25981 - CATGACAACCTGCTCATAAAGAGCTTGTCTTACTGCAATATAAACCTGTGTACGAAC - 26040
 - H D N L L I K S F V I D C N I N L C T N
 - M T T C S * R A L S L T A I * T C V R T
 - * Q P A H K E L C H * L Q Y K P V Y E P

图 12 续

26881 - ACCGGATTCAACAGTGTAGACAGAGCACTTTTCATTAAAGCACTTTGTCAACACGTTTCATC - 26940
 - T G F N S V D R A L F I K H F V N T F I
 - P D S T V * T E H F S L S T L S T R S S
 - R I Q Q C R Q S T F H * A L C Q H V H Q
 26941 - AAGCTCAAATGTGATTCTCACATCTTGTAACTTGAACCTTCCCAAACAGTATCTTCTCC - 27000
 - K L K C D S H I L V T L N F P N S I F S
 - S S N V I L T F L * P * T S Q T V S S P
 - A Q M * F S H S C N L E L P K Q Y L L Q
 27001 - AAAGGTTACACCTTTAATGGTGCACCCCTTTTAAGCGAAAGACATTGTTTGTAGCCAG - 27060
 - K G Y T F N W C T P F * A K D I V C S Q
 - K V T P L I G A P P F K R K T L F V A S
 - R L H L * L V H P L L S E R H C L * P V
 27061 - TAAACCAGGAGACAATGCCAGTATTGTTCTTGTCCCTTAATCTCTAAGAGCATGAGGCC - 27120
 - * T R R Q C A V L F F V L N L * E H E A
 - K P G D N A Q Y C S L S L I S K S M R P
 - N Q E T M R S I V L C P * S L R A * G H
 27121 - ATTTACACAGACTGGTGTGCCGACGATAGCTCCATTGTGAAGCTATCAACGGGCGTCTC - 27180
 - I Y T D W C A D D S S I C E A I N G R L
 - F T Q T G V P T I A P F V K L S T G V S
 - L H R L V C R R * L H L * S Y Q R A S R
 27181 - GAGTGCTTCGAGTTCACCGTCTTGAGAACAACCTCCTCAGAGGTAAGTACTGTGTCATG - 27240
 - E C F E F T V L E N N L L R G K Y C V M
 - S A S S P F L R T T S S E V S T V S C
 - V L R V H R S * E Q P P Q R * V L C H V
 27241 - TGAATCACCTCAAGAAAGGTTACTTCTTTGGTGCCTTAAGAGGCATGAGTACTGTCAG - 27300
 - * I T F K K G Y F F W C L K R H E * L Q
 - E S P S R K V T S F G A L R G M S S C S
 - N H L Q E R L L L L V P * E A * V V A A
 27301 - CTGCTCCTTGCCACGTATACACTGACGGTAAAGTCCCTTGCTTTGAGCGATGAAGACTTC - 27360
 - L L L A T Y T L T V K S L A L S D E D F
 - C S L P R I H * R * S F L L * A M K T S
 - A P C H V Y T D G K V P C F E R * R L H
 27361 - ACCTAAGTTGAGTGATCGCAACTTTGGCCAGCGATAGTGACTTGATCAATGCACATTTTC - 27420
 - T * V E * S Q L C A S D S D L I N A H F
 - P K L S D R N F A P A I V T * S M H I S
 - L S * V I A T L R Q R * * L D Q C T F R
 27421 - GAGTGCCTTGTTAACAACATCAATGAAGCATTACACAATCCTTGATGTTATCTGAAGC - 27480
 - E C L V N N I N E A F Y T I L D V I * S
 - S A L L T T S M K H F T Q S L M L S E A
 - V P C * Q H Q * S I L H N P * C Y L K Q
 27481 - AACCTGTATTTGACCCTTGACGATGTCAAAAACCTGTATGAGAAATTTGAGAATCTC - 27540
 - N L Y L T L D D V K N T C N E K F E N L
 - T C I * P L T M S K T P V M R N L R I S
 - P V F D P * R C Q K H L * * E I * E S P
 27541 - CCAAGCATCCTTGAGAAATCAACTCCTGCACTAAGTTTCGCCTCAATCCATTCAAAGAT - 27600
 - P S I L E K F N S C T K F R L N P F K D
 - Q A S L R N S T P A L S F A S I H S K I
 - K H P * E I Q L L H * V S P Q S I Q R *
 27601 - AGGCCTGAGTTTTCAACAGTAGTGCCAAAAGATTAGACAACCACTGAGAAGTCTGTTG - 27660
 - R P E F F N S S A Q K I R Q P L R S L L
 - G L S F S T V V P K R L D N H * E V C C
 - A * V F Q Q * C P K D * T T T E K S V V
 27661 - TACAAGACCACCAGTTACATATGCCATAATAAGACTGTTGGTGTGAGCAGGTCTGAAGT - 27720
 - Y K T T S Y I C H N N D T V G E Q V * S
 - T R P F V T Y A I I M T L L V S R S E V
 - Q D H Q L H M P * * * H C W * A G L K Y

图 12 续

27721 - ATAAACCATGGCGTCGACAAGACGTAATGACTGTTTCAGAAATACCATCAAGTATGGTGAC - 27780
 - I N H G V D K T * * L F R N T I K Y G D
 - * T M A S T R R N D C S E I P S S M V T
 - K P W R R Q D V M T V Q K Y H Q V W * Q
 27781 - AGCTGCTCTTTGCAAATCAGGAATTGAGTGGTTTGTGCATCAAGTGTGCGCGCAAAAAT - 27840
 - S C S L Q I R N * V V C C I K C A R K N
 - A A L C K S G I E W F A A S S V R A K I
 - L L F A N Q E L S G L L H Q V C A Q K L
 27841 - TGATCTGATAACACCAGCAGCCTGTGAGGGAAAACCACACAGTGGTGTAAAACCTGATCT - 27900
 - * S D N T S S L * G K T T Q W C * N * S
 - D L I T P A A C E G K P H S G V K T D L
 - I * * H Q Q P V R E N H T V V L K L I S
 27901 - CTGTTGTCCAATGTTCCAAGCACCTTTTACGGGCTTTCCCTTGGTAACTTTATAGTTACC - 27960
 - L L S N V P S T F Y G L S L G N F I V T
 - C C P M F Q A P F T G F P L V T L * L P
 - V V Q C S K H L L R A F P W * L Y S Y R
 27961 - GCAGGACTCAACAATGGTTTTGAAAGACTTGAATCAAGACTCTTATAGTGTCAATAAA - 28020
 - A G L N N G F E R L V I K T L Y S V N K
 - Q D S T M V L K D L * S R L F I V S I K
 - R T Q Q W F * K T C N Q D S L * C Q * R
 28021 - GGCACTTGTAGAAGCAGAGAAAGATGCCAAATGATGGCAACCTCTTCAATCAATGAAA - 28080
 - G T C R S R E R C Q N D G N L F I Q M K
 - A L V E A E K D A K M M A T S S F K * K
 - H L * K Q R K M P K * W Q P L H S N E N
 28081 - ATCGCCAACAATGTTAATGTTAACAGTTCAGGACTCAGTATCTCAAGGAGATCCTCATT - 28140
 - I A N N V N V N T F T T Q Y L K E I L I
 - S P T M L M L T R S R L S I S R R S S F
 - R Q Q C * C * H V H D S V S Q G D P H S
 28141 - CRAAGTCTCCACATTGTCCAGTAATGCCAGTATGGCCTGAGCCAATATCAGCACTAGC - 28200
 - Q G L H I V T S N A S M A * A N I S T S
 - K V S T L S P V M P V W P E P I S A L A
 - R S P H C H Q * C Q Y G L S Q Y Q H * H
 28201 - ACGAGGAACCCAGTAGGCACGCTTATTATAGCAGCCAACATAGGCAACACACAGCCTCC - 28260
 - T R N P V G T L I I A A N I G K H T A S
 - R G T Q * A R L L * Q P T * A N T Q P P
 - E E P S R H A Y Y S S Q H R Q T H S L Q
 28261 - AAAACATCTAGTCTACCTCCCTTGGCGGAGTCGAGTTTCAATGTTTGGTGGTTGTGATA - 28320
 - K T S S P T S L A E S S F N V * V V V I
 - K H L V L P P L R S R V S M F E W L * *
 - N I * S Y L P C G V E F Q C L S G C D N
 28321 - ATCTGCAACACTATGCTCAGGTCCAATCTCTGGGTCTTGACAGGCAGGACATGGCATT - 28380
 - I C N T M L R S N L W V L T G R T W H F
 - S A T L C S G P I S G S * Q A G H G I F
 - L Q H Y A Q V Q S L G L D R Q D M A F S
 28381 - CACTACAGCATTAGTAGGTAGGTACCCACATGTAGTAGGTCCTTCAATACTAAATTTTC - 28440
 - H Y S I S R * V P T C S R S F N N * I F
 - T T A L V G R Y P H V V G P S I T K F S
 - L Q H * * V G T H M * * V L Q * L N F Q
 28441 - AGTGCCACAATGTTCAAGTGGCTTTTCAGAAAGTCGCAGGCTGCCATGAAACTTCATC - 28500
 - S A T M F T S G F Q K V A R L P * N F I
 - V P Q C S Q V A F R K S H V C H E T S S
 - C H N V H K W L S E S R T S A M K L H R
 28501 - GCAATGATTACATTTTCATCAAGGTAGACAAGTGCATATTGTTACTCCTGTGGAGATGC - 28560
 - A M I T F H Q G R Q V H I V T L L W R C
 - Q * L H F I K V D K C I L L H S C G D A
 - N D Y I S S R * T S A Y C Y T P V E M Q

图 12 续

28561 - AACAGGTTACACAGAGCGTATACGCCCATGAAACCCCTCAGTCTTTTTCTTTTCAACACG - 28620
 - N R V H R A Y T P H E T L S L F L F N T
 - T G Y T E R I R P M K P S V F F F S T R
 - Q G T Q S V Y A P * N P Q S F S F Q H V
 28621 - TGGTTGAATGACTTTTGACTTTTGAGTTAAGAGGAAACACAAACTTTGGGCATTCCCCTT - 28680
 - W L N D F D F * V K R K H K L W A F P F
 - G * M T L T F E L R G N T N F G H S P L
 - V E * L * L L S * E E T Q T L G I P L *
 28681 - GAAAGTGTCAAATTTCTTGGCACTCTTAATTTGGAAGGGTGTCTGGTGCTCGTAGCTCTT - 28740
 - E S V K F L G T L N F E G C L V L V A L
 - K V S N F L A L L I S K G V W C S * L L
 - K C Q I S W H S * F R R V S G A R S S Y
 28741 - ATCAGAGCGCTCAGTGAACCAGGCAATTCATGCTCATGGTCACGGCAGTAGACACC - 28800
 - I R A L S E P G N F M L M V T A A V D T
 - S E R S V N Q A I S C S W S R Q Q * T P
 - Q S A Q * T R Q E H A H G H G S S R H L
 28801 - TCTCTTCGACTCGATGTAATCAAGTTGTTGCGAAAGAGTGCACATTGACTTGCCCCGCCG - 28860
 - S L R L D V I K L F G K S A H * L A R A
 - L F D S M * S S C S E R V H I D L P A R
 - S S T R C N Q V V R K E C T L T C P R V
 28861 - TGCGAGAAAATCTTTGATGCAATCAGAGGGTACCCATCTGGGCCACAGAAATGTTGTC - 28920
 - C E K I F D A I K R V P I W A T E I V V
 - A R K S L M Q S R G Y P S G P Q K L L S
 - R E N L * C N Q E G T H L G H R N C C R
 28921 - GACATAGCGAGTACTGCACCTCCATTGAGCTCAGGAGTTCACGGAGTGCACCACT - 28980
 - D I A S D C T S I E L T S E F T E C T T
 - T * R V T A P P L S S R V S S R S A P L
 - H S E * L H L H * A H E * V H G V H C
 28981 - GCCATGCTTAGTGTCCAGTTTTGTTTCATAATCTCAATGGGATCAGTGCCAAAGCTCGTC - 29040
 - A M L S V P V L F I I F N G I S A K L V
 - P C L V F Q F C S * S S M G S V P S S S
 - H A * C S S F V H N L Q W D Q C Q A R H
 29041 - ACCTAAGTCATAAGACTTTAGATCGATGCCATAGCTATGACCACGGCTCCCTTATTACC - 29100
 - T * V I R L * I D A I A M T T G S L I T
 - P K S * D F R S M P * L * P P A P L L P
 - L S H K T L D R C H S Y D H R L P Y R
 29101 - GTTCTTACGAAGAAGAACATTGGCGTATGCAATGGGGTTTGGCCACATGTGGCAGCAG - 29160
 - V L T K K N I A V C N W G F A H M W H E
 - F L R R R T L R Y A I G V S P T C G T S
 - S Y E E E H C G M Q L G F R P H V A R V
 29161 - TACTCCAGTGTATACCGTACGACCGTACTGAAATGCCGTCCATTTCTGCAACCAGCTC - 29220
 - Y S Q C Y T A T T V L N A V H F C N Q L
 - T P S V I P L R P Y * M P S I S A T S S
 - L P V L Y R Y D R T E C R P F L Q P A Q
 29221 - AACGACCTTGTGGCCGTGATTGGTGCCTAAGGCATCAGAACGTTAATGAACACATAGGG - 29280
 - N D L V A V I G A * G I R T F N E H I G
 - T T L W P * L V L K A S E R L M N T * G
 - R P C G R D W C L R H Q N V * * T H R A
 29281 - CTGTTCAAGCTGGGGCAGTACGCCCTTTTCCAGCTCTACTAGACCACAAGTGCCATTTTT - 29340
 - L F K L G Q Y A F F Q L Y * T T S A I F
 - C S W G S T P F S S S T R P Q V P F L
 - V Q A G A V R L F P A L L D H K C H F *
 29341 - GAGGTGTTACCGTGCCTCCGATAGGGCCTCTCCACAGAGTCCCCGAGCCAGCAGTAG - 29400
 - E V F T C L R * G L F H R V P E A T H *
 - R C S R A S D R A S S T E S P K P R T S
 - G V H V P P I G P L P Q S P R S H A L A

图 12 续

```

29401 - CACGTCTCTAACCTGAAGGACAGGCAAACCTGAGTTGGACGTGTGTTTTCTCGTTGACACC - 29460
- H V S N L K D R Q T E L D V C F L V D T
- T S L T * R T G K L S W T C V F S L T P
- R L * P E G Q A N * V G R V F S R * H Q
29461 - AAGAACAAGGCTCTCCATCTTACCTTTCCGGTCACACCCGGACGAAACCTAGGTATGCTGA - 29520
- K N K A L H L T F R S H P D E T * V C *
- R T R L S I L P F G H T R T K P R Y A D
- E Q G S P S Y L S V T P G R N L G M L M
29521 - TGATCGACTGCAACACGGACGAAACCGTAAGCAGTCTGCAGAAGAGGGACGAGTTACTCG - 29580
- * S T A T R T K P * A V C R R G T S Y S
- D R L Q H G R N R K Q S A E E G R V T R
- I D C N T D E T V S S L Q K R D E L L V
29581 - TTTCTGTCAACGACAGTAAAATTTATTATTGTTTATACTGCGTAGGTGCACTAGGCATG - 29640
- F L V N D S K I Y Y C L Y C V G A L G M
- F L S T T V K F I I V Y T A * V H * A C
- S C Q R Q * N L L L F I L R R C T R H A
29641 - CAGCCGAGCGACAGCTACACAGATTTAAAGTTCGTTTAGAGAACGATCTACAAGAGAT - 29700
- Q P S D S Y T D F K V R L E N R S T R D
- S R A T A T Q I L K F V * R T D L Q E I
- A E R Q L H R F * S S F R E Q I Y K R S
29701 - CGAGGTTGGTTGGCTTTTCTGGGTAGGTAACCTAATAT - 29742
- R G W L A F P G * V K T * Y X
- E V G W L F L G R * K P N X
- R L V G F S W V G K N L I X

```

图 12 续

引物和探针序列

| | |
|-------|---|
| 正向引物: | 5'-CAGAACGCTGTAGCTTCAAAAATCT -3' (SEQ ID NO:2471) |
| 反向引物: | 5'-TCAGAACCCTGTGATGAATCAACAG -3' (SEQ ID NO:2472) |
| 探针: | 5'-TCTGCGTAGGCAATCC-3' (SEQ ID NO:2473) (5'用 FAM 标记; 3'用 NFQ-MGB 标记) |
| 正向引物: | 5'-ACCAGAATGGAGGACGCAATG-3' (SEQ ID NO:2474) |
| 反向引物: | 5'-GCTGTGAACCAAGACGCAGTATTAT -3' (SEQ ID NO:2475) |
| 探针: | 5'-ACCCCAAGGTTTACCC-3' (SEQ ID NO:2476) (5'用 FAM 标记; 3'用 NFQ-MGB 标记) |

图 13

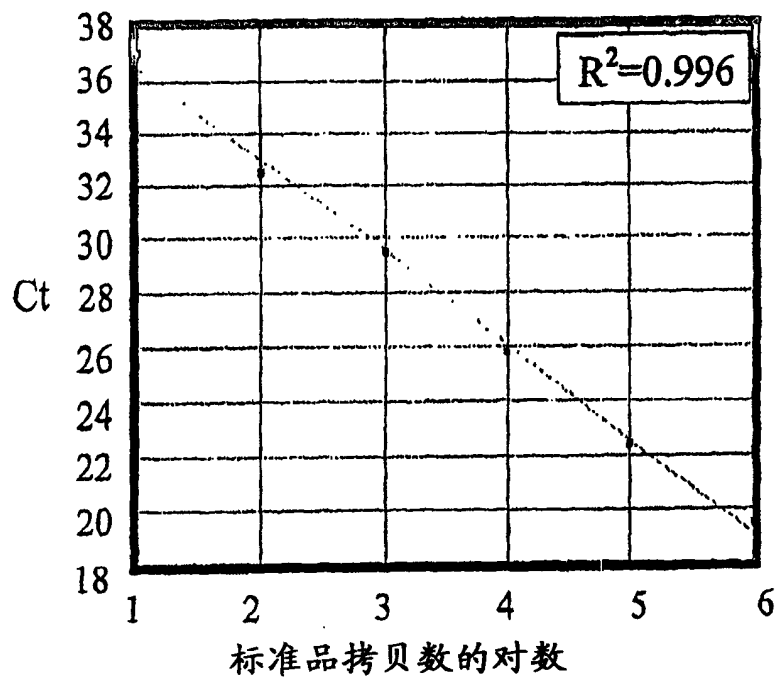


图 14

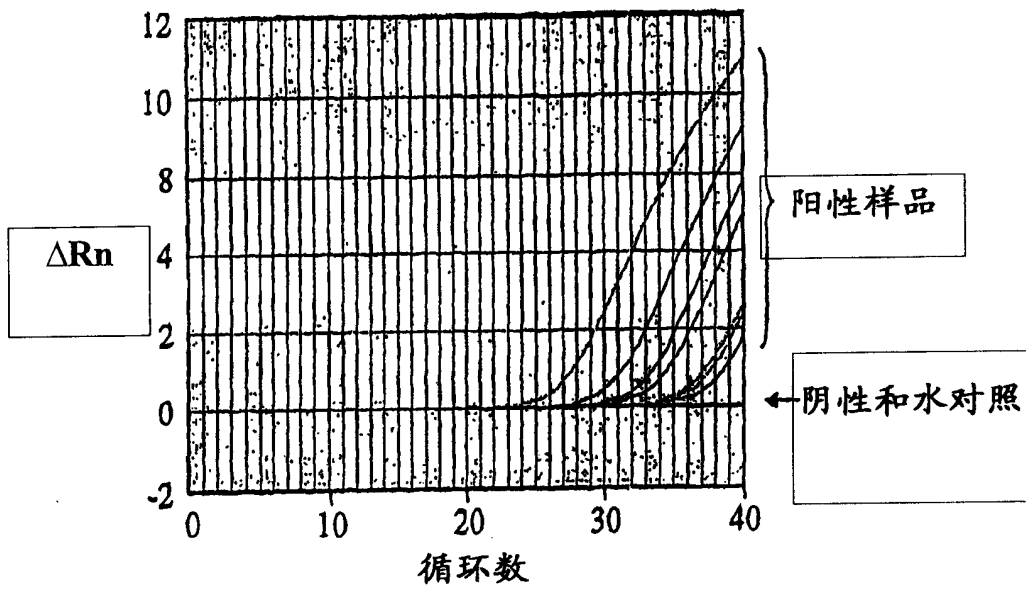


图 15

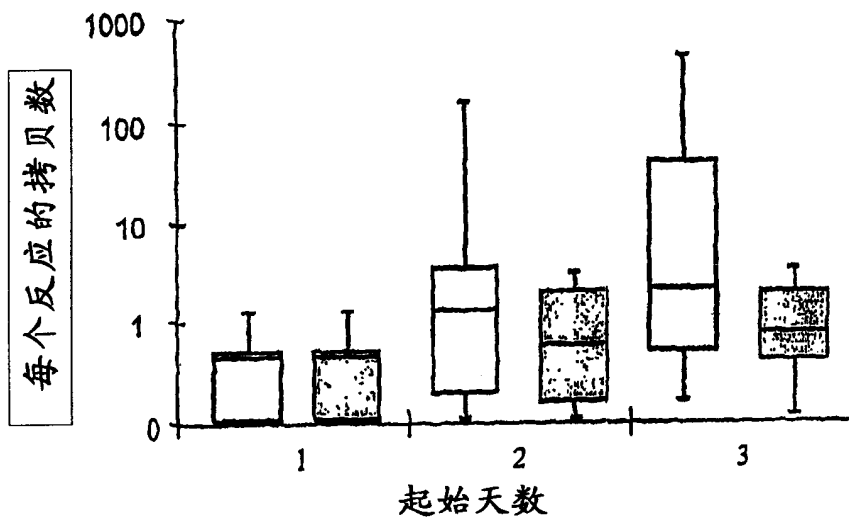


图 16