



(12) 发明专利

(10) 授权公告号 CN 1910293 B

(45) 授权公告日 2011.04.13

(21) 申请号 200480038606.9

代理人 王允方 刘国伟

(22) 申请日 2004.10.22

(51) Int. Cl.

(30) 优先权数据

C12Q 1/68 (2006.01)

60/513,888 2003.10.23 US

C12N 15/54 (2006.01)

60/542,373 2004.02.06 US

A61K 31/7088 (2006.01)

60/554,758 2004.03.19 US

C12N 15/11 (2006.01)

60/560,524 2004.04.08 US

C07K 16/40 (2006.01)

60/578,323 2004.06.09 US

60/605,243 2004.08.26 US

(56) 对比文件

WO 02081741 A3,2003.11.06,

(85) PCT申请进入国家阶段日

S Knapp1,et al.Polymorphisms in

2006.06.22

interferon-induced genes and the. 《Genes and Immunity》.2003,第4卷(第6期),

(86) PCT申请的申请数据

G.A. Hitman, et al.2'-5' oligoadenylate

PCT/US2004/035284 2004.10.22

synthetase and its relationship to HLA and.

(87) PCT申请的公布数据

《Immunogenetics》.1989,第30卷(第6期),

WO2005/040428 EN 2005.05.06

J. Justesen*, R. Hartmann and N. O.

(73) 专利权人 伊洛默格生物科学公司

Kjeldgaard.Gene structure and function of the

地址 美国马萨诸塞州

2'-5'-oligoadenylate. 《CELLULAR AND

(72) 发明人 肖恩·P·艾度纳托

查尔斯·L·麦格尼斯

盖瑞·罗森堡

克莉斯汀娜·A·席尔

蒂埃里·吉尔洛德约克斯

MOLCULAR LIFE SCIENCES》.2000,第57卷(第11期),

审查员 温庭江

(74) 专利代理机构 北京律盟知识产权代理有限

责任公司 11287

权利要求书 1 页 说明书 45 页 附图 33 页

(54) 发明名称

检测与抗病毒感染性有关的基因 OAS1 中的突变的方法

(57) 摘要

本发明涉及一种检测与编码 OAS1 的基因有关的突变的方法。这一突变和其他所揭示的突变与人类对包括 C 型肝炎的病毒感染的抗性有关。本发明也提供一种由所述突变基因编码的蛋白质或多肽或编码所述蛋白质或多肽的聚核苷酸组成的治疗剂。本发明也提供对人类 OAS1 呈特异性的包括反义寡核苷酸的人类 OAS1 抑制剂、方法和组合物。

1. 一种经分离的多肽，其是由选自 SEQ ID NO：48、SEQ ID NO：49 和 SEQ ID NO：50 组成的群组的氨基酸序列组成。
2. 根据权利要求 1 所述的经分离的多肽，其 N 末端仅存在单个甲硫氨酸残基。
3. 根据权利要求 1 或权利要求 2 所述的经分离的多肽，其与包含蛋白转导域的多肽共价连接。
4. 根据权利要求 3 所述的经分离的多肽，其中所述蛋白转导域包含选自 SEQ IDNO：85-94 组成的群组的的多肽。
5. 根据权利要求 3 所述的经分离的多肽，其中所述蛋白转导域因精氨酸、赖氨酸或组氨酸的加入或取代而不同于选自 SEQ ID NO：85-94 组成的群组的的多肽。
6. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其与聚乙二醇共价连接。
7. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其封装于脂质体中。
8. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其与内体破裂剂共价连接。
9. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其与内体破裂剂非共价连接。
10. 根据权利要求 8 所述的经分离的多肽，其中所述内体破裂剂具有 pH 敏感性。
11. 根据权利要求 9 所述的经分离的多肽，其中所述内体破裂剂具有 pH 敏感性。
12. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其与糖部分共价共轭。
13. 根据权利要求 12 所述的经分离的多肽，其中所述糖部分包含半乳糖或甘露糖。
14. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其是由重组宿主细胞产生。
15. 根据权利要求 1-5 中任一权利要求所述的经分离的多肽，其包含异源多肽序列。
16. 一种经分离的多肽，其是通过包含以下步骤的方法产生：
 - (a) 由细胞表达根据权利要求 1-5 中任一权利要求所述的多肽；和
 - (b) 回收所述多肽。
17. 一种组合物，其包含根据权利要求 1-16 中任一权利要求所述的经分离的多肽和医学上可接受的载剂。

检测与抗病毒感染性有关的基因 OAS1 中的突变的方法

技术领域

[0001] 本发明涉及一种用于检测人类寡腺苷酸合成酶基因中的突变的方法及所编码的蛋白质及其抗体的用途，其中突变赋予对黄病毒感染（包括 C 型肝炎感染）的抗性，且突变与包括前列腺癌和糖尿病的其他病状有关。

背景技术

[0002] 迄今为止已鉴别许多疾病，其中人类群体中存在天然的感染抗性。Alter 和 Moyer, *J.Acquir.Immune Defic.Syindr.Hum Retrovirol.*18 增刊 1 : S6-10 (1998) 报导在诸如注射药物使用者的高危群组中的 C 型肝炎病毒感染 (HCV) 高达 90%。然而，尚未有文献确认剩余的 10% 明显对抗感染的机制为何。在 HCV 感染中起作用的蛋白质包括 2- 端、5- 端寡腺苷酸合成酶。OAS 是干扰素诱导的蛋白质，其特征为能催化 2- 端、5- 端寡腺苷寡聚物 (2-5A) 的合成。Hovanessian 等人, *EMBO* 6 : 1273-1280 (1987) 发现经干扰素治疗的人类细胞含有数种与 40 (OAS1)、46 (OAS1)、69 及 100kD 的蛋白质对应的 OAS。Marie 等人, *Biochem.Biophys.Res.Comm.*160 : 580-587 (1989) 产生高度特异性的抗 p69 (69-kD OAS) 多克隆抗体。通过以抗 p69 抗体筛检经干扰素治疗的人类细胞表达文库, Marie 和 Hovanessian, *J.Biol.Chem.*267 : 9933-9939 (1992) 分离出部分 OAS2 cDNA。他们以部分 cDNA 筛检另外的文库并回收编码两个 OAS2 同功异型物 (isoform) 的 cDNA。较小的同功异型物由两个 3- 端未转译区长度不同的 mRNA 编码。

[0003] Northern 印迹分析揭示 OAS2 表达为人类细胞中四个由干扰素诱导的 mRNA。所预测的 OAS2 蛋白质具有共同的 683- 氨基酸序列及不同的 3- 末端。根据活体外转录 / 转译产物的 SDS-PAGE, 两个同功异型物的分子质量为 69 和 71kD。两个同功异型物在活体外均显示出 OAS 活性。序列分析表明 OAS2 含有两个由富含脯氨酸的假定连接子区所分隔的 OAS1 同源域。N 末端及 C 末端域分别为 41% 和 53% 相同于 OAS1。

[0004] 通过荧光原位杂交和通过包含于定位克隆之内, Hovanian 等人, *Genomics* 52 : 267-277 (1998) 确定 OAS1、OAS2 和 OAS3 基因是由 12q24.2 上的 130-kb 区丛集。2-5A 结合至降解病毒及细胞 RNA 的核糖核酸酶 I 并激活其, 从而导致细胞蛋白质合成的抑制和病毒复制的削弱。

[0005] 被称为 OASL 的第四种人类 OAS 基因与 OAS1、OAS2 和 OAS3 的不同在于 OASL 缺乏酶活性。OASL 基因编码双域蛋白质, 其是由融合至与泛素的串联重复同源的 164- 氨基酸 C 末端域的 OAS 单元组成。(Eskildsen 等人, *Nuc.Acids Res.*31 : 3166-3173, 2003 ; Kakuta 等人, *J.Interferon & Cytokine Res.*22 : 981-993, 2002)

[0006] 由于其在抑制病毒复制和病毒感染中的作用, 因此此项技术中需要抑制与 OAS1 活性有关的病毒复制的方法和组合物, 包括进一步需要抑制 HCV 复制的基于抑制剂的疗法。

发明内容

[0007] 本发明涉及检测与 C 型肝炎抗性有关的突变, 所述突变的特征在于其为寡腺苷

酸合成酶 1 基因中的点突变。

[0008] 在一实施例中，涵盖一种人类基因筛检方法。所述方法包含对分离自人类的核酸样本分析寡腺苷酸合成酶 1 基因点突变的存在，所述点突变的特征为在根据 Genbank 序列登记号第 NT_009775.13 号（连续核苷酸 2,130,000-2,157,999，其以 SEQ ID NO：19 显示于图 2 中）的寡腺苷酸合成酶 1 基因（OAS1）的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523 或 2156638 核苷酸位置处的碱基取代或在 2156595 核苷酸位置处的碱基缺失。

[0009] 在一优选实施例中，所述方法包含于扩增条件下以聚合酶链式反应（PCR）引物对处理来自人类的基因组 DNA 样本，以扩增含有寡腺苷酸合成酶 1 基因 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 核苷酸位置的人类基因组 DNA 区。PCR 处理产生含有所述区的扩增产物，然后分析其点突变的存在。

[0010] 在另一实施例中，本发明提供一种由在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处具有至少一个突变的基因所编码的蛋白质，和所述蛋白质用以制备抗病毒感染、优选为黄病毒感染、最优选为 C 型肝炎感染用的诊断剂的用途。在特定实施例中，所述诊断剂为抗体。

[0011] 在另一实施例中，本发明提供一种用于预防或抑制病毒感染、优选为黄病毒感染、最优选为 C 型肝炎病毒感染的治疗化合物，其中所述治疗化合物为由在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处具有至少一个突变的 OAS1 基因所编码的蛋白质。在其他实施例中，所述治疗化合物为编码蛋白质的聚核苷酸，诸如 DNA 或 RNA。

[0012] 在另一实施例中，本发明提供一种用于预防或抑制病毒感染、优选为黄病毒感染、最优选为 C 型肝炎病毒感染的治疗化合物，其中所述治疗化合物为具有以下序列的蛋白质：SEQ ID NO：20、SEQ ID NO：21、SEQ ID NO：22、SEQ ID NO：23、SEQ ID NO：24、SEQ ID NO：25、SEQ ID NO：26、SEQ ID NO：27、SEQ ID NO：28、SEQ ID NO：29、SEQ ID NO：32、SEQ ID NO：33、SEQ ID NO：34、SEQ ID NO：35、SEQ ID NO：46、SEQ ID NO：47 和 / 或 SEQ ID NO：48。

[0013] 在另一实施例中，本发明提供一种用于预防或抑制病毒感染、优选为黄病毒感染、最优选为 C 型肝炎病毒感染的治疗化合物，其中所述治疗化合物模拟在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处的至少一个突变的有利作用。所述治疗化合物可为小分子、蛋白质、肽、DNA 或 RNA 分子或抗体。

[0014] 在另一实施例中，本发明提供一种用于预防或治疗癌症、优选为前列腺癌的治疗化合物，其中所述治疗化合物为由在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、

2144985、2156523、2156595 或 2156638 位置处具有至少一个突变的 OAS1 基因所编码的蛋白质。在其他实施例中，所述治疗化合物为编码蛋白质的聚核苷酸，诸如 DNA 或 RNA。

[0015] 在另一实施例中，本发明提供一种用于预防或治疗癌症、优选为前列腺癌的治疗化合物，其中所述治疗化合物为具有以下序列的蛋白质：SEQ ID NO：20、SEQ ID NO：21、SEQ ID NO：22、SEQ ID NO：23、SEQ ID NO：24、SEQ ID NO：25、SEQ ID NO：26、SEQ ID NO：27、SEQ ID NO：28、SEQ ID NO：29、SEQ ID NO：32、SEQ ID NO：33、SEQ ID NO：34、SEQ ID NO：35、SEQ ID NO：46、SEQ ID NO：47 和 / 或 SEQ ID NO：48。

[0016] 在另一实施例中，本发明提供一种用于预防或治疗癌症、优选为前列腺癌的治疗化合物，其中所述治疗化合物模拟在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处的至少一个突变的有利作用。所述治疗化合物可为小分子、蛋白质、肽、DNA 或 RNA 分子或抗体。

[0017] 在其他实施例中，所述治疗化合物能够抑制 OAS1 或整个蛋白质的至少一个子区或子功能的活性，且所述化合物的代表为能特异性结合至 OAS1 聚核苷酸的反义分子、核酶和 RNAi 分子，以及能特异性结合至 OAS1 蛋白质和多肽的抗体及其片段。

[0018] 在另一实施例中，本发明提供 OAS1 的抑制剂。本发明的抑制剂包括（但不限于）反义分子、核酶、RNAi、抗体或抗体片段、蛋白质或多肽以及小分子。例示性反义分子包含 SEQ ID NO：19 的至少 10、15 或 20 个连续核苷酸，或是在严格条件下杂交至 SEQ ID NO：19 的聚核苷酸的反义分子。更优选的是包含 SEQ ID NO：19 的至少 25 个连续核苷酸或在严格条件下杂交至 SEQ ID NO：19 的序列的反义分子。

[0019] 在另一实施例中，设想 OAS1 的抑制剂特异性结合至由 SEQ ID NO：30 的多肽所界定的蛋白质区。本发明的抑制剂包括（但不限于）抗体、抗体片段、小分子、蛋白质或多肽。

[0020] 在另一实施例中，设想 OAS1 的抑制剂包含特异性结合或杂交至 SEQ ID NO：31 的聚核苷酸的反义或 RNAi 分子。

[0021] 在另一实施例中，提供包含在医药学上可接受的载剂中的一种或一种以上 OAS1 抑制剂的组合物。

[0022] 额外的实施例提供降低 OAS1 基因表达或生物学活性的方法。

[0023] 额外的实施例提供特异性增加或降低在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处具有至少一个突变的 OAS1 基因的某些形式的表达的方法。

[0024] 本发明提供包含至少一个经修饰的核苷间键合的反义寡核苷酸。

[0025] 本发明进一步提供具有硫代磷酸酯键合的反义寡核苷酸。

[0026] 本发明又进一步提供包含至少一个经修饰的糖部分的反义寡核苷酸。

[0027] 本发明也提供包含至少一个经修饰的糖部分的反义寡核苷酸，其中所述糖部分为 2' -O- 甲基糖部分。

- [0028] 本发明进一步提供包含至少一个经修饰的核苷酸碱基的反义寡核苷酸。
- [0029] 本发明又进一步提供具有经修饰的核苷酸碱基的反义寡核苷酸，其中所述经修饰的核苷酸碱基为 5- 甲基胞嘧啶。
- [0030] 本发明也提供一种反义化合物，其中所述反义化合物为嵌合寡核苷酸。
- [0031] 本发明提供一种抑制人类细胞或组织中人类 OAS1 的表达的方法，其包含使活体内细胞或组织与靶向于编码人类 OAS1 的核酸分子的长度为 8 至 35 个核苷酸的反义化合物或核酶接触，从而抑制人类 OAS1 的表达。
- [0032] 本发明进一步提供一种使用反义或 RNAi 化合物或核酶降低或增加活体内 OAS1 的特定形式的表达的方法，所述形式是由在 NT 009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处具有至少一个突变来界定。
- [0033] 本发明进一步提供一种调节癌细胞生长的方法，其包含使活体内癌细胞与靶向于编码人类 OAS1 的核酸分子的长度为 8 至 35 个核苷酸的反义化合物或核酶接触，从而抑制人类 OAS1 的表达。
- [0034] 本发明又进一步提供鉴别 OAS1 聚核苷酸的目标区。本发明也提供用于通过原位杂交鉴别 OAS1 聚核苷酸的经标记的探针。
- [0035] 本发明提供根据本发明的 OAS1 抑制剂用以制备预防或抑制 HCV 感染用的药剂的用途。
- [0036] 本发明进一步提供使 OAS1 抑制剂针对 OAS1 蛋白质的特定区或所述蛋白质的特定功能。
- [0037] 本发明也提供一种用于抑制 OAS1 表达的医药组合物，其包含本发明的反义寡核苷酸与生理学上可接受的载剂或稀释剂的混合物。
- [0038] 本发明进一步提供能特异性裂解 OAS1 RNA 的核酶和包含所述核酶的医药组合物。
- [0039] 本发明也提供 OAS1 的小分子抑制剂，其中所述抑制剂能降低 OAS1 的活性或能降低或阻碍 OAS1 mRNA 的表达。
- [0040] 本发明进一步提供除修饰 2' -5' 寡腺苷酸的合成外还修饰蛋白质的特定功能的 OAS1 抑制剂，所述功能包括与诸如 C 型肝炎病毒 NS5 A 蛋白的其他蛋白质的相互作用。
- [0041] 本发明进一步提供改变包括（但不限于）糖基化和磷酸化作用的 OAS1 转译后修饰的化合物。本发明进一步提供用于鉴别寡腺苷酸合成酶基因突变的人类基因筛检方法，其包含：(a) 于扩增条件下以聚合酶链式反应 (PCR) 引物对处理来自人类的基因组 DNA 样本，以扩增含有寡腺苷酸合成酶基因的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 核苷酸位置的人类基因组 DNA 区，所述处理产生含有所述区的扩增产物；及 (b) 检测步骤 (a) 的扩增产物中在 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 核苷酸位置处的核苷酸点突变的存在，从而鉴别所述突变。

[0042] 在此方法的某些实施例中，所述区包含以选自由 SEQ ID NO：1-7 及 SEQ IDNO：57-64 组成的群组的序列所代表的核苷酸序列。在其他实施例中，所述区基本上由选自由 SEQ ID NO：1-7 及 SEQ ID NO：57-64 组成的群组的核苷酸序列组成。本发明也提供一种检测方法，其中所述检测包含在杂交条件下以对点突变特异的寡核苷酸探针处理以上步骤 (a) 的扩增产物，并检测杂交产物的形成。在所述方法的某些实施例中，寡核苷酸探针包含选自由 SEQ ID NO：12-18 组成的群组的核苷酸序列。

[0043] 本发明也涉及在人类中检测含有点突变的 C 型肝炎感染抗性疾病等位基因的方法，所述点突变包含在寡腺苷酸合成酶基因 (OAS1) 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 核苷酸位置处以非野生型核苷酸取代野生型核苷酸，所述方法包含：(a) 通过将来自所述人类的基因组 DNA 样本与选自由 SEQ ID NO：8 及 9 和 SEQ ID NO：10 及 11 组成的群组的寡腺苷酸合成酶基因特异性 PCR 引物对在 PCR 缓冲液中组合而形成聚合酶链式反应 (PCR) 混合物；(b) 使 PCR 混合物经受多个 PCR 热循环以产生寡腺苷酸合成酶基因扩增产物；和 (c) 在杂交条件下以选自由 SEQ ID NO：12-18 组成的群组的探针处理步骤 (b) 中所产生的产物，从而检测所述突变。

[0044] 本发明也提供一种经分离的 OAS1 抑制剂，其是选自由下列各物组成的群组：反义寡核苷酸、核酶、抑制性小 RNA (RNAi)、蛋白质、多肽、抗体和小分子。所述经分离的抑制剂可为包含至少 15 个 SEQ ID NO：19 序列的连续核酸的反义分子或其互补序列。在其他实施例中，所述经分离的 OAS1 抑制剂（反义分子或其互补序列）在高度严格条件下与 SEQ ID NO：19 的序列杂交。

[0045] 所述经分离的 OAS1 抑制剂可选自由抗体和抗体片段组成的群组。本发明也提供一种组合物，其包含在医药学上可接受的载剂中的治疗有效量的至少一种 OAS1 抑制剂。

[0046] 本发明也涉及一种抑制哺乳动物细胞中 OAS1 表达的方法，其包含向所述细胞投与选自由下列各物组成的群组的 OAS1 抑制剂：反义寡核苷酸、核酶、蛋白质、RNAi、多肽、抗体和小分子。

[0047] 本发明进一步涉及一种抑制受检者中 OAS1 基因表达的方法，其包含向所述受检者投与在医药学上有效的媒剂中的反义寡核苷酸，所述反义寡核苷酸的量可有效地与衍生自所述 OAS1 基因的经选定目标核酸序列的全部或部分特异性杂交。

[0048] 本发明又进一步涉及一种预防易受感染的人类受检者中黄病毒感染的方法，其包含向所述人类受检者投与选自由下列各物组成的群组的 OAS1 抑制剂：反义寡核苷酸、核酶、RNAi、蛋白质、多肽、抗体和小分子，其中所述 OAS1 抑制剂预防所述黄病毒的感染。

[0049] 本发明又进一步涉及一种预防或治愈易受感染的人类受检者中黄病毒或其他病毒感染的的方法，其包含向所述人类受检者投与选自由下列各物组成的群组的 OAS1 抑制剂：反义寡核苷酸、核酶、RNAi、蛋白质、多肽、抗体和小分子，其中所述 OAS1 抑制剂预防所述黄病毒或其他病毒感染且其中所述 OAS1 抑制剂针对一种或一种以上由在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处的突变所界定的蛋白质的特定形式。

[0050] 本发明也进一步涉及一种通过投与具有下列序列的多肽中的一者来预防或治愈易受感染的人类受检者中黄病毒或任何其他病毒感染的方法：SEQ ID NO：20、SEQ ID NO：21、SEQ ID NO：22、SEQ ID NO：23、SEQ ID NO：24、SEQ ID NO：25、SEQ ID NO：26、SEQ ID NO：27、SEQ ID NO：28、SEQ ID NO：29、SEQ ID NO：32、SEQ ID NO：33、SEQ ID NO：34、SEQ ID NO：35、SEQ ID NO：46、SEQ ID NO：47 和 / 或 SEQ ID NO：48。

[0051] 本发明也涉及人类免疫缺陷病毒 (HIV) 感染的治疗。

[0052] 本发明又进一步涉及一种预防人类受检者中的胰岛素依赖型糖尿病 (IDDM) 的方法，其包含向所述人类受检者投与选自下列各物组成的群组的 OAS1 抑制剂：反义寡核苷酸、核酶、RNAi、蛋白质、多肽、抗体和小分子，其中所述 OAS1 抑制剂预防 IDDM。

[0053] 本发明又进一步涉及一种预防人类受检者中的 IDDM 的方法，其包含向所述人类受检者投与选自下列各物组成的群组的 OAS1 抑制剂：反义寡核苷酸、核酶、RNAi、蛋白质、多肽、抗体和小分子，其中所述 OAS1 抑制剂预防 IDDM 且其中所述 OAS1 抑制剂针对一种或一种以上由在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处的突变所界定的蛋白质的特定形式。

[0054] 本发明又进一步涉及一种通过增加 OAS1 基因表达或通过治疗性投与本文中所揭示的多肽来治疗癌症（诸如前列腺癌）的方法。

[0055] 本发明也提供一种抑制活体外细胞中 OAS1 目标基因表达的方法，其包含将足以抑制 OAS1 目标基因表达的量的核糖核酸 (RNA) 引入细胞中，其中所述 RNA 为双股分子，第一股基本上由与 OAS1 目标基因的核苷酸序列对应的核糖核苷酸序列组成，且第二股基本上由与 OAS1 目标基因的核苷酸序列互补的核糖核苷酸序列组成，其中第一与第二核糖核苷酸股为互相杂交而形成所述双股分子的独立互补股，且所述双股分子抑制目标基因的表达。

[0056] 在所述方法的某些实施例中，第一核糖核苷酸序列包含至少 20 个与 OAS1 目标基因对应的碱基且第二核糖核苷酸序列包含至少 20 个与 OAS1 目标基因的核苷酸序列互补的碱基。在其他实施例中，目标基因表达被抑制至少 10%。

[0057] 在所述方法的其他实施例中，双股核糖核酸结构为长度至少 20 个碱基且每一核糖核酸链都能特异性地与超过所述至少 20 个碱基的 OAS1 目标基因的脱氧核糖核酸链杂交。

[0058] 本发明提供能恢复可能由于基因突变而减弱或失去的 OAS1 功能的多肽或蛋白质。在某些实施例中，所述多肽或蛋白质具有由包含 SEQ ID NO：19 的基因所编码的野生型 OAS1 的氨基酸序列。在其他实施例中，其中 OAS1 基因中的突变赋予 OAS1 蛋白增加的活性、稳定性和 / 或半衰期或其他使 OAS1 蛋白更适用于抗病毒活性的变化，由经突变的 OAS1 基因所编码的蛋白质或多肽是优选的。

[0059] 任何上述蛋白质和多肽均可作为治疗组合物的组份来提供。

[0060] 本发明也提供由以下各序列组成的蛋白质中的任一者作为治疗组合物的组份的用途：SEQ ID NO：20、SEQ ID NO：21、SEQ ID NO：22、SEQ ID NO：23、SEQ ID

NO : 24、SEQ ID NO : 25、SEQ ID UNO : 26、SEQ ID NO : 27、SEQ ID NO : 28、SEQ ID NO : 29、SEQ ID NO : 32、SEQ ID NO : 33、SEQ ID NO : 34、SEQ ID NO : 35、SEQ ID NO : 46、SEQ ID NO : 47 和 / 或 SEQ ID NO : 48。

[0061] 在另一实施例中, 编码 OAS1 蛋白、OAS1 突变型蛋白或 OAS1 多肽的核酸可以基因疗法形式来投与。

附图说明

[0062] 图 1 是显示 OAS1 基因中本发明突变的位置、等位基因变异体 (碱基取代)、基因组序列上突变的座标和 (若有的话) NCBI dbSNP ID 的表。

[0063] 图 2 (SEQ ID NO : 19) 是由在 NCBI 登记号 NT_009775.13 的 2,130,000-2,157,999 位置处的连续核苷酸碱基组成的聚核苷酸序列 OAS1。

[0064] 图 3 显示 SEQ ID NO : 20-64。

[0065] 图 4 显示高加索群体中 OAS1 基因内的大致单倍体分布。

[0066] 图 5A 及 B 阐明本发明所揭示的转录物变异体的种类及其与本发明所揭示的突变的关系。图 5A 显示转录物变异体 (TV) 形式 1-6, 且图 5B 显示 TV 形式 7-10。图 5C 进一步阐明黑猩猩和大猩猩变异体相对于人类的推测结构。

[0067] 图 6 是描述非人类的灵长类相对于人类的 OAS1 基因的突变和在对应人类基因组序列上的突变座标的表。

[0068] 图 7 表明本发明的例示性多肽当自大肠杆菌表达和回收时的酶活性。

[0069] 图 8 表明由本发明的聚核苷酸产生的例示性多肽于 HeLa 细胞中表达时的酶活性。

[0070] 图 9 表明本发明的例示性多肽的抗体的显影。

[0071] 图 10 是详述例示性蛋白质转导域多肽的表。

具体实施方式

[0072] 引言和定义

[0073] 本发明涉及寡核苷酸合成酶基因的新颖突变、所述突变用于诊断对病毒感染的易感性或抗性的用途、由具有根据本发明的突变的基因所编码的蛋白质和使用所述蛋白质、抗体及相关核酸预防或抑制病毒感染。所述突变与携带者对黄病毒感染、尤其是 C 型肝炎病毒感染的抗性有关。

[0074] 许多当前的医学研究集中于鉴别引起或有助于疾病的突变和缺陷。设计所述研究以导致针对病状来治疗的化合物和方法。对于尽管暴露于感染剂及其他危险因素下但仍然使人们保持健康的基因影响的研究则较少关注。本发明表示由发明者所研制的方法的成功应用, 由此确定并分析人类受检者的特定群体以便揭示赋予对疾病的抗性的基因变异或突变。对具有对特定疾病或生物学病症的天然抗性的亚群体部分的鉴别使得能进一步鉴别对于药物介入、诊断评估或预防 (诸如预防性接种) 为适合目标的基因和蛋白质。

[0075] 本文中所鉴别的亚群体部分包含尽管反复暴露于 C 型肝炎病毒 (HCV) 下但是仍保持血清阴性而同类者已受感染 (血清阳性) 的个体。所研究的群体包括经受反复输血

的血友病患者，和通过共用的针及其他危险因素而暴露的静脉内药物使用者。

[0076] HCV 感染包括一组共同作用以达成一定感染水平的复杂蛋白质和免疫系统组份，当所述感染水平引起疾病时，其可于受感染细胞内发展成病毒的低稳态，从而显然允许 HCV 逃离宿主免疫监视系统，同时促成病毒持久感染 (Dansako 等人, *Virus Research*97 : 17-30, 2003)。本发明集中于这个系统中的一种组份，即可由干扰素诱导的 2' -5' -寡腺苷酸合成酶基因，尤其是 OAS1。OAS1 基因通过激活核糖核酸酶 L (RNase L) 以裂解病毒 RNA 而于人类宿主细胞的抗病毒活性方面起主要作用。HCV RNA 激活 2' -5' -OAS/RNase L 途径。如 Dansako 等人所指出，HCV RNA 激活导致病毒 RNA 裂解的途径可看似矛盾。然而，所述活性可用来保持宿主免疫防御与可杀死宿主的感染水平之间的平衡。

[0077] 鉴于 OAS1 基因的这种复杂作用，本发明已确定 OAS1 基因中的突变与所述突变的携带者中抗 HCV 感染性之间的强相关性具有显著重要性。现在所述个体的存在可允许阐明尽管反复暴露于感染水平的病毒下但是 OAS1 仍如何有助于对 HCV 感染产生抗性。这个信息随后将导致开发用于在缺少天然抗性的个体中复制抗性机制的方法和组合物。

[0078] 因此，本发明提供：本文中所鉴别的突变无论机制为何均适用于鉴别抗 HCV 感染的个体。抗性可通过 OAS1 蛋白的功能丧失来产生，在此情况下预测 HCV 病毒含量将足够高以阻止病毒自宿主免疫监视系统逃离，从而有利于病毒的破坏。抗性也可通过功能的获得而产生，因为 OAS1 蛋白含量增加，蛋白质半衰期增加和 / 或蛋白质结构以增强其激活核糖核酸酶 L 裂解病毒 RNA 的能力的方式受影响。抗性也可通过对 OAS1 蛋白的修饰而产生，所述修饰阻止 HCV 病毒蛋白或核苷酸对正常 OAS1 蛋白功能的抑制。抗性也可通过对 OAS1 蛋白的修饰而产生，所述修饰阻止蛋白与正常 HCV 病毒寿命所必需的 HCV 病毒蛋白或核苷酸之间的相互作用。本发明不受限于一种机制。此外，尽管本文中揭示几种不同的点突变，然而这并非想要表明每一突变对于 OAS1 蛋白结构或功能均具有相同的作用。

[0079] OAS1 在由 HCV 为其中一员的黄病毒家族的其他病毒引起的感染方面起作用。黄病毒家族也包括引起黄热病、登革热、圣路易斯脑炎 (St.Louis encephalitis)、日本脑炎 (Japanese encephalitis) 和本文中所揭示的其他病毒疾病的病毒。宿主对于这些病毒的防御包括可由病毒诱导的干扰素。所述干扰素诱导 2' -5' -寡腺苷酸合成酶，如上所讨论所述合成酶涉及于 RNaseL 的活化作用中。RNaseL 接着裂解病毒 RNA。其他病毒感染可通过本文中所揭示的方法得以预防和 / 或抑制，包括 RSV。

[0080] 关于详细描述和优选实施例，使用下列定义：

[0081] A：腺嘌呤；C：胞嘧啶；G：鸟嘌呤；T：胸腺嘧啶 (DNA 中)；和 U：尿嘧啶 (RNA 中)

[0082] 等位基因：特定基因的 DNA 序列的变异体。二倍体细胞中最多将存在两个等位基因，其各自位于染色体组的同源染色体的相同的相对位置或基因座上。当位于任一基因座上的等位基因相同时，认为就那个基因座而言个体是纯合的，而当其不同时，认为就那个基因座而言个体是杂合的。因为任一基因的不同等位基因可能仅有单个碱基改变，所以任一基因的等位基因的可能数目十分巨大。当等位基因不同时，一个相对于另一个是显性而另一个则被认为是隐性的。显性是表现型的性质且并不意味隐性等位基因

由显性基因导致失活。在许多实例中，正常发挥功能（野生型）的等位基因对于所有或多或少缺陷功能的突变型等位基因均呈显性。在所述情况下，一般的解释是两个功能等位基因中的一个足以产生足够活性的基因产物以支持有机体的正常发育（即基因产物的数量通常存在 2 倍安全限度）。

[0083] 单倍体：许多可能的多个等位基因中的一者，其通过染色体定位而连续排列，且呈现由染色体组的一个特定同源染色体所携带的等位基因组。

[0084] 核苷酸：DNA 或 RNA 的单体单元，由糖部分（戊糖）、磷酸酯和含氮杂环碱基组成。碱基与糖部分经由糖苷碳（戊糖的 1' 碳）连接，且碱基与糖的所述组合为核苷。当核苷含有键结于戊糖的 3' 或 5' 位置的磷酸酯基团时，称其为核苷酸。可经由操作连接的核苷酸的序列于本文中通常称为“碱基序列”或“核苷酸序列”及其合乎文法的等效物，且本文中由自左至右方向为 5' - 末端至 3' - 末端的常规定向的式来表示。

[0085] 碱基对 (bp)：双股 DNA 分子中腺嘌呤 (A) 与胸腺嘧啶 (T) 的配对或胞嘧啶 (C) 与鸟嘌呤 (G) 的配对。在 RNA 中，以尿嘧啶 (U) 替代胸腺嘧啶。当本文中提及 RNA 时，可互换使用符号 T 与 U 来表示处于 RNA 分子中的特定位置的尿嘧啶。

[0086] 核酸：单股或双股的核苷酸聚合物。

[0087] 聚核苷酸：单股或双股核苷酸的聚合物。本文中所用的“聚核苷酸”及其合乎文法的等效物包括全部范围的核酸。聚核苷酸通常意指包含两种或两种以上脱氧核糖核苷酸和 / 或核糖核苷酸的直链的核酸分子。如此项技术中所熟知的，精确尺寸将取决于许多因素，而这些因素又取决于最终使用条件。本发明的聚核苷酸包括引物、探针、RNA/DNA 节段、寡核苷酸（相对短的聚核苷酸）、基因、载体、质粒等等。

[0088] RNAi：RNA 干扰 (RNAi) 是一种藉以将干扰性小 RNA (siRNA)（通常约 21-23 个核苷酸长度的双螺旋）引入细胞内，最终导致含有相同或互补序列的目标基因的信使 RNA 降解并有效地使其沉默的方法。

[0089] 基因：其核苷酸序列编码 RNA 或多肽的核酸。基因可为 RNA 或 DNA。

[0090] 双螺旋 DNA：双股核酸分子，其包含两股由一个或一个以上存在于双螺旋碱基对中的每一互补碱基之间的氢键保持在一起的实质上互补的聚核苷酸。由于形成碱基对的核苷酸可为核糖核苷酸碱基或脱氧核糖核苷酸碱基，因此短语“双螺旋 DNA”意指包含两股 DNA (ds DNA) 的 DNA-DNA 双螺旋，或包含一股 DNA 和一股 RNA 的 RNA-DNA 双螺旋。

[0091] 互补碱基：当 DNA 或 RNA 采用双股构型时正常配对的核苷酸。

[0092] 互补核苷酸序列：DNA 或 RNA 的单股分子中的核苷酸序列，其与另一单股上的核苷酸序列充分互补，从而以所产生的氢键与其特异性地杂交。

[0093] 保守：若核苷酸序列与预选（参考）序列的精确互补序列非随机杂交，则所述核苷酸序列对所述预选序列保守。

[0094] 杂交：通过互补碱基对之间建立氢键的使实质上互补的核苷酸序列（核酸链）配对以形成双螺旋或异源双螺旋。其为两条互补聚核苷酸之间可竞争性抑制的特异性（即非随机的）相互作用。

[0095] 核苷酸类似物：结构上与 A、T、G、C 或 U 不同但足够类似以取代核酸分子中的正常核苷酸的嘌呤或嘧啶核苷酸。

[0096] DNA 同源体：具有预选保守核苷酸序列和编码能结合预选配位体的受体的序列的核酸。

[0097] 上游：在与 DNA 转录方向相反的方向上，且因此在非编码链上为自 5' 至 3' 行进，或在 mRNA 上为自 3' 至 5' 行进。

[0098] 下游：进一步在序列转录或读取的方向上沿着 DNA 序列，其沿 DNA 非编码链的 3' 至 5' 方向行进，或沿 RNA 转录物的 5' 至 3' 方向行进。

[0099] 终止密码子：不编码氨基酸而是导致蛋白质合成终止的三个密码子中的任一者。它们是 UAG、UAA 及 UGA 且也称为无义或终止密码子。

[0100] 阅读框：转译中使用的邻近核苷酸三联体（密码子）的特定序列。阅读框取决于转译起始密码子的位置。

[0101] 内含子：也称为插入序列，其为最初复制入 RNA 但由最终 RNA 转录物切除的 DNA 非编码序列。

[0102] 本发明的实施模式

[0103] 本发明提供一种用于筛检人类中与对黄病毒感染、尤其是 C 型肝炎感染的抗性有关的寡腺苷酸合成酶等位基因的新颖方法。本发明是基于所述抗性与寡腺苷酸合成酶基因 DNA 序列中在编码人类 OAS1 基因的 Genbank 登记号 NT_009775.13 (SEQ ID NO: 19) 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 核苷酸位置处的点突变（碱基取代）有关的发现。

[0104] 本发明揭示鉴别对 C 型肝炎病毒 (HCV) 感染具有抗性或部分抗性的受检者群体和进一步鉴别赋予这种有益作用的基因突变的研究结果。鉴别 2' -5' -寡腺苷酸合成酶基因中的几种基因突变，其与对 HCV 感染的抗性显著相关。所使用的研究设计为病例 - 对照型等位基因关联分析。指定为受检者的病例已经连续证实或推测暴露于 HCV，但由于发展对病毒的抗体而经证实并未发展感染（即 HCV 血清阴性）。对照受检者为连续暴露且血清转变为 HCV 阳性的受检者。病例和对照受检者募集自三种群体：来自加拿大不列颠哥伦比亚省的范库弗峰 (Vancouver) 的血友病患者；来自法国西北部的血友病患者；和来自西雅图 (Seattle) 大都会地区的注射药物使用者。

[0105] 病例和对照的定义在血友病组与注射药物使用者 (IDU) 组之间不同且是基于文献中所发表的感染风险的流行病学模型和如本文中所述由发明者提出的其他模型。对于血友病群体，使用商业诊断实验室测试证实对照受检者对于 HCV 抗体呈血清阳性。证实病例受检者为 HCV 血清阴性、具有低于 5% 的正常凝血因子且于 1987 年 1 月前已经接收浓缩凝血因子。将对照注射药物使用者定义为经证实的 HCV 血清阳性。将病例注射药物使用者定义为经证实的 HCV 血清阴性，其注射药物超过 10 年以上，且据报导其参与一种或一种以上额外的危险行为。额外的危险行为包括与另一名 IDU 共用注射器、蒸煮锅或棉球。在一特定高加索 (Caucasian) 群体中有 20 名病例和 42 名对照包括于本研究中。病例和对照受检者的选择基本上如美国专利申请案 09/707,576 中所述使用受感染的有风险（“对照”）和未受感染的有风险（“病例”）群体群组进行。

[0106] 鉴别与 HCV 感染抗性有关的基因突变的本发明方法包括选择候选基因。调查约 50 种在病毒结合至细胞表面、细胞内病毒繁殖、干扰素反应和先天免疫系统方面和抗病

毒反应中所涉及的候选基因。通过使用聚合酶链式反应以扩增来自每一受检者的基因组 DNA 的目标序列而在病例和对照中为候选基因测序。使用基于荧光的自动化 DNA 测序法和 ABI3730 自动化测序仪来直接测序来自候选基因的 PCR 产物。

[0107] 鉴别寡腺苷酸合成酶基因 (OAS1) 中的基因突变, 其单独或以联合形式与 HCV 感染抗性显著相关 ($p < 0.05$)。图 1 中显示构成这些突变的碱基取代和缺失。OAS1 基因的变异形式 (“OAS1R”) 是因本发明突变中的一种或一种以上的存在而产生 (确定为 SEQ ID NO : 1-7 和 SEQ ID NO : 57-64)。据信 OAS1 基因的这些变异 OAS1R 形式赋予抗病毒感染性。

[0108] 自通过所属领域技术人员已知的期望最大化 (Expectation Maximization) 方法建立的病例-对照基因型数据计算推测包含多个 OAS1 突变 (SEQ ID NO : 1-7 和 SEQ ID NO : 57-64) 的群体和受检者亲本单倍体 (Excoffier 和 Slatkin, Mol.Biol.Evol.12 : 921-927, 1995)。本发明涵盖使用用于单倍体分析的全部范围的 OAS1 突变以及使用出于计算便利的 OAS1 突变的子集。对单倍体和单倍体子集的分离样式与病例和对照组的比较分析鉴别关于病毒抗性或易感性的特定功效的突变。

[0109] 在一说明性实例中, 计算包含多个由 SEQ ID NO : 2 至 SEQ ID NO : 6 确定的 OAS1 突变的单倍体。通过此分析鉴别高加索人病例和对照群体中的数个单倍体。图 4 中显示这些单倍体的定义。确定两种常见单倍体 (确定为 HAP1 和 HAP2), 其占所推测的单倍体约 85% 且处于哈迪-韦伯格 (Hardy-Weinberg) 平衡中, 所述平衡尤其是关于群体中单倍体纯合子的出现。各种人类群体和灵长类中的 OAS1 的进一步分析表明 HAP2 是遗传的灵长类单倍体, 其预先表明旧世纪猴与原始人类的趋异。一种额外的单倍体 (确定为 HAP3) 与这一特定群体中的持久 HCV- 抗性病例组有关。因此携带 HAP3 单倍体的受检者处于对 HCV 感染的实质上较低风险中。HAP3 单倍体似乎是经由一系列源自遗传单倍体的复杂重组和突变而出现。与单倍体 HAP3 高发生率相结合的所述事件的结合少见性表明作出在群体中发展和保持单倍体 HAP3 的积极选择, 其可能随时间推移作为对复发病毒攻击的反应。在这个实例中, 单倍体 HAP3 是以可评估的群体频率发生的仅有单倍体, 其将单一前体 RNA 中的突变 SEQ ID NO : 2 的 G 核苷酸的作用与突变 SEQ ID NO : 4 的 A 核苷酸的作用相结合。

[0110] 本发明不受限于上述说明性实例。本发明也不受限于提供对特定 OAS1 突变的相关性和效用的了解的其他说明性实例。在另一说明性实例中, SEQ ID NO : 2 中 G 核苷酸取代 A 核苷酸引起所预测的丝氨酸对甘氨酸的氨基酸取代。所属领域技术人员已知的计算预测高度表明: 丝氨酸是磷酸化作用位点而甘氨酸不会经磷酸化。

[0111] 在另一说明性实例中, 突变 SEQ ID NO : 4 中的 A 核苷酸取代 G 核苷酸发生于 OAS1 中的野生型第六外显子的一致剪接受体位点。此取代替剪接受体辨别信号中所必需的 G, 但在这个过程中创造新的剪接识别位点, 即一个下游碱基对。因此经突变形式于转译蛋白中创造移码。经突变位点也为较少有效的剪接信号, 且因而促进除移码外显子 6 剪接之外的前体 RNA 的额外替代性剪接。图 5A 和 5B 中提供这些替代性剪接形式的优选实施例。以下提供基因分析的其他说明性实例。

[0112] OAS1 转录物变异体的混乱剪接通过来源于自新鲜人血清分离而得的淋巴细胞系和周边血液单核细胞 (PBMC) 的 RNA 的逆转录而独立地确定。对来自携带各种单倍体的

细胞系和 PBMC 的逆转录 RNA 产物的 PCR 分析表明携带突变 SEQ ID NO: 4 的 A 核苷酸的 RNA 形式导致众多转录物变异体 OAS1R 形式。这些 OAS1R 转录物变异体于图 5A 及 5B 中图示描述且包含如图 3 中所提供的 SEQ ID NO: 36 至 SEQ ID NO: 47。

[0113] 据信 OAS1 基因的这些变异 OAS1R 形式和相应的转录物变异体编码由 SEQ ID NO: 20、SEQ ID NO: 21、SEQ ID NO: 22、SEQ ID NO: 23、SEQ ID NO: 24、SEQ ID NO: 25、SEQ ID NO: 26、SEQ ID NO: 27、SEQ ID NO: 28、SEQ ID NO: 29、SEQ ID NO: 32、SEQ ID NO: 35、SEQ ID NO: 46 和 / 或 SEQ ID NO: 47 组成的多肽中的一种或一种以上。单个或多个上述多肽于本文中可互换地称作 OAS1R 多肽或 OAS1R 蛋白。许多上述多肽的共同特征在于其主要在其羧基末端不同而保留氨基末端部分。

[0114] 除其自身替代性转录物的产生之外，基因的 OAS1R 形式也可含有或除去特定序列背景（诸如外显子剪接增强子），其改变对于特定转录物变异体的选择性优选。此举又引起大量所得蛋白质的相对含量不同。OAS1 基因的这些变异 OAS1R 形式也可改变所得蛋白质的定位或转译后修饰。所属领域技术人员应了解特定 OAS1 蛋白形式的丰度增加或其他改良其活性、稳定性或可用性的修饰可改良 2' -5' -OAS/RNase L 途径的总体抗病毒性能。所属领域技术人员同样可了解在特定 OAS1 形式与其他特定 OAS1 形式相比并不有利或甚至不利的情况下，抑制所述特定蛋白的活性或可用性也可改良 2' -5' -OAS/RNaseL 途径的总体抗病毒性能。不利的 OAS1 蛋白的一实施例为（非局限于）以消除所述特定 OAS1 蛋白的酶活性的方式由病毒蛋白特异性靶向的蛋白。非有利的 OAS1 蛋白的另一实施例为具有与其他活性形式聚合的较低酶活性因此降低或消除聚合蛋白的总酶活性（且因而降低总体抗病毒效果）的蛋白。上述机制中的一种或一种以上可有助于对病毒感染产生抗性。然而，本发明不受限于所揭示的变异聚核苷酸或多肽的特定作用机制。

[0115] 因此，本发明提供人类 2' -5' -寡腺苷酸合成酶基因的新颖形式、新颖 mRNA 转录物和相关蛋白质。本发明也揭示新颖 mRNA 转录物和新颖蛋白质的效用。这些新颖形式的特征在于基因中存在公共数据库中未揭示的数种罕见基因突变或单倍体中的一种或一种以上。OAS1 的这些新颖形式 OAS1R 赋予携带者对 C 型肝炎病毒和相关黄病毒一定程度的抗性，所述相关黄病毒包括（但不限于）西尼罗河病毒（West Nile virus）、登革热病毒、黄热病病毒、蜱传脑炎病毒（tick-borne encephalitis virus）、日本脑炎病毒、圣路易斯脑炎病毒、墨累谷病毒（Murray Valley virus）、玻瓦桑病毒（Powassan virus）、罗西奥病毒（Rocio virus）、跳跃病病毒（louping-ill virus）、班奇病毒（Banzi virus）、伊列乌斯病毒（Ilheus virus）、科科贝拉病毒（Kokobera virus）、昆金病毒（Kunjin virus）、沃父病毒（Alfuy virus）、牛痢疾病毒和卡萨诺尔森林病病毒（Kyasanur forest disease virus）。OAS 蛋白也已显示其在实验性呼吸融合病毒和小核糖核酸病毒细胞培养物感染系统的感染减弱中是重要的。已将受人类免疫缺陷病毒-1（HIV-1）感染的细胞无法释放病毒与高浓度的 OAS 和 / 或 2-5A 联系起来。此外，HIV-1 反式激活蛋白（tat）已显示阻断 OAS 的活化（Muller 等人，J Biol Chem.1990 Mar 5；265(7)：3803-8），因而表明 OAS 的新颖形式可避开 HIV-1 防御机制且提供有效的治疗。因此本文中所揭示的这些 OAS1 的 OAS1R 形式也可赋予对这些非黄病毒感染剂的抗性。

[0116] 本发明也提供对 OAS1 的黑猩猩 (*Pan troglodytes*) 和大猩猩 (*Gorilla gorilla*) 形式的新颖描述, 每一形式均导致具有有效用的新颖 mRNA 和多肽。虽然基因在关系密切的灵长类 (诸如人类、黑猩猩和大猩猩) 中通常高度保守, 但是观察到三种物种之间 OAS1 的重要差异。最亲的人类亲属黑猩猩在 OAS1 外显子 5 中具有单个碱基取代 (位于相当于人类 NT_009775.13 中的 2,142,351 的位点且由 SEQ ID NO: 53 所界定), 其导致生成截短型蛋白质产物。黑猩猩 OAS1 多肽和 mRNA 序列分别由 SEQ ID NO: 51 和 SEQ ID NO: 55 提供。大猩猩也在靠近受体剪接位点的外显子 6 中具有两碱基对缺失 (位于相当于人类 NT_009775.13 中的 2,144,089-2,144,090 的位点且由 SEQ ID NO: 54 所界定), 其导致翻译提前终止。大猩猩部分多肽和部分 mRNA 序列分别由 SEQ ID NO: 52 和 SEQ ID NO: 56 提供。图 5C 中提供相对于人类的黑猩猩和大猩猩转录物的推测结构。这些情况中的每一种 (类似于人类 OAS1R 多肽) 均具有高度保守的多肽序列, 其在羧基末端尾部的结构和内容物方面最显著不同。这些多肽的共有氨基末端部分含有先前证实为 OAS1 酶活性所必需的所有元件。由于黑猩猩和大猩猩已经受与非洲人的病毒攻毒类似的病毒攻毒, 因此这些不同但功能上类似的灵长类变异体的流行提供进一步证据表明较长 OAS1 形式的羧基末端部分对于病毒攻毒后的继续存活为非必要或甚至不利的。黑猩猩具有 OAS1 多肽的完全截短 (其与人类转录物变异体的异质性相对) 的事实与以下观察相一致: 尽管黑猩猩为已知的易受 HCV 感染的仅有其他灵长类, 但黑猩猩确实具有相对于人类的非典型性感染, 其特征为病毒清除频率增加和不导致纤维化或肝细胞癌。

[0117] 每一新颖 OAS1R cDNA 都是克隆自携带这些突变的人类受检者。克隆是通过标准 cDNA 克隆方法进行, 这些方法包括自细胞或组织分离 RNA、将 RNA 转化为 cDNA 和将 cDNA 转化为适于克隆的双股 DNA。正如所属领域技术人员将认识到的, 所有这些步骤都是常规分子生物学分析。其他方法包括使用逆转录酶 PCR、5' RACE (cDNA 末端快速扩增) 或传统的 cDNA 文库构建和 Southern 杂交筛选。本文中所述的所有 OAS1R 等位基因都是自患者携带者回收。对每一新近克隆的 OAS1R cDNA 测序以确定其身份且鉴别相对于野生型的任何额外序列差异。

[0118] 新颖 OAS1R 基因突变可通过改变所得 OAS1 mRNA 的性质而影响对病毒感染的抗性。因此, 评估 OAS1R 等位基因携带者与纯合野生型受检者之间的 mRNA 稳定性差异。使用包括 Taqman[®] 和简单 Northern 杂交法的已知分析来评估并比较 RNA 稳定性。这些构成分子生物学中的常规方法。

[0119] OAS1R 突变可通过改变 OAS1 基因的调控来影响感染抗性。已知 OAS 基因的表达是由干扰素治疗诱导且于病毒感染期间诱导。OAS1R 等位基因可经由组成性表达、过度表达或其他不受调控型表达来赋予对病毒感染的抗性。使用数种方法评估具有或不具有干扰素或病毒刺激的基因表达。这些方法包括表达微阵列分析、Northern 杂交法、Taqman[®] 及其他方法。自己知表达 OAS 基因的组织 (诸如周边血液单核细胞) 采集样本。比较来自 OAS1R 携带者和非携带者的组织之间的基因表达。在一实施例中, 自携带者和非携带者采集周边血液单核细胞, 在培养物中繁殖且由干扰素刺激。将干扰素诱导期间 OAS1R 等位基因的表达水平与野生型等位基因进行比较。在另一实施例中, 以干扰素治疗人类受检者并评估 OAS1R 突变的携带者中相对于非携带者中的 OAS1 基因的诱导程度。正如所属领域技术人员可了解, 将组织、实验设计和分析方法的许多组合用于

评估 OAS1R 基因调控。

[0120] 一旦对每一 OAS1R 的新颖 cDNA 进行克隆, 即可使用众多不同的已知表达克隆系统中的任一种将其用于制备重组 OAS1R 蛋白。在此方法的一实施例中, OAS1R cDNA 是通过标准分子生物学方法克隆至大肠杆菌 (*Escherichia coli*) 表达载体中邻近于含有编码聚组氨酸多肽的 DNA 序列的附加表位。随后使用固定化金属亲和色谱法或类似方法自大肠杆菌溶解产物纯化重组蛋白。所属领域技术人员将认识到有许多可用于纯化重组蛋白的不同表达载体和宿主细胞, 包括(但不限于)酵母表达系统、杆状病毒表达系统、中国仓鼠卵巢细胞等等。

[0121] 使用计算方法鉴别来自 OAS1R 蛋白的短肽序列, 其独特地将这些蛋白与野生型 OAS1 蛋白区别开来。各种计算方法和市售软件包可用于肽选择。可使用 FMOC 肽合成化学或类似方法制备这些经计算选择的肽序列。所属领域技术人员将认识到有大量用于根据所供应序列合成短多肽的化学方法。

[0122] 可使用来自 OAS1R 基因的肽片段和重组蛋白开发对此基因产物呈特异性的抗体。正如所属领域技术人员将认识到的, 存在大量用于抗体开发的方法, 包括使用多个不同的宿主有机体、佐剂等等。在一经典实施例中, 将少量(150 微克)纯化重组蛋白皮下注射入新西兰白兔 (*New Zealand White Rabbit*) 背部, 随后每几个月注射类似数量作为辅助剂。随后通过静脉穿刺收集兔血清并可收集血清、纯化 IgG 或对免疫蛋白呈特异性的亲和纯化抗体。正如所属领域技术人员将认识到的, 可使用类似方法开发大鼠、小鼠、山羊及其他有机体中的抗体。也可使用如上所述的肽片段开发对 OAS1R 蛋白呈特异性的抗体。单克隆抗体和多克隆抗体的开发都适于实施本发明。可通过标准分子技术产生分泌 OAS1R 蛋白的特异性单克隆抗体的小鼠杂交瘤细胞系。

[0123] 可使用如上所述制备的抗体开发用于评估细胞、组织和有机体中 OAS1R 蛋白的存在或不存在的诊断方法。在此方法的一实施例中, 可使用经纯化的重组 OAS1R 蛋白和特异性抗体开发酶联免疫吸附分析以检测人类血清中的这些蛋白。可使用这些诊断方法验证上述基因突变的携带者和非携带者的组织中 OAS1R 蛋白的存在或不存在。

[0124] 也可使用如上所述制备的抗体自那些携带这些突变的患者纯化天然 OAS1R 蛋白。对于使用抗体自人类细胞和组织纯化天然蛋白有大量方法可利用。在一实施例中, 抗体可用于包括均质化人类组织和使用蛋白 A 捕捉抗体的免疫沉淀实验中。这个方法能够浓缩且进一步评估突变型 OAS1R 蛋白。可利用许多其他用于分离 OAS1R 的天然形式的方法, 包括柱色谱法、亲和色谱法、高压液相色谱法、盐析法、透析法、电泳法、等电聚焦法、差速离心法等等。

[0125] 使用蛋白质体方法评估 OAS1R 突变对二级、三级和四级蛋白质结构的影响。也使用蛋白质体方法评估 OAS1R 突变对 OAS 蛋白的转译后修饰的影响。存在许多已知可能的蛋白质转译后修饰, 包括蛋白酶裂解、糖基化、磷酸化、硫酸化、加入化学基团或复合分子等等。评估二级和三级蛋白质结构的一般方法为核磁共振 (NMR) 光谱法。使用 NMR 探测野生型 OAS1 蛋白与 OAS1R 蛋白之间二级和三级结构的差异。传统 NMR 的修改方案也是适合的, 包括评估功能位点活性的方法, 包括核转移欧佛豪瑟光谱法 (TrNOESY) 等等。正如所属领域技术人员将认识到的, 可采用此方法和用于结果的数据解释的方法的许多微小修改方案。所有这些方法都意欲包括于本发明的实施之中。使用

通过结晶和 X 光衍射来确定蛋白质结构的其他方法。

[0126] 也可使用质谱法评估突变型与野生型 OAS 蛋白之间的差异。可使用此方法评估结构差异以及蛋白质转译后修饰的差异。在此方法的一典型实施例中，使用上述方法中的一种自人类周边血液单核细胞纯化野生型 OAS1 蛋白和突变型 OAS1R 蛋白。这些细胞可如上所述经干扰素刺激以增加 OAS 蛋白的表达。以特异性蛋白酶（诸如胰蛋白酶）消化经纯化蛋白并使用质谱法评估。正如所属领域技术人员将认识到的，也可使用许多替代方法。本发明涵盖这些额外的替代方法。例如，可使用基质辅助激光解吸附 / 电离法 (MALDI) 或电喷雾电离 (ESI) 质谱法。此外，质谱法可与细胞蛋白的二维凝胶电泳分离法联用以作为对综合预纯化法的替代。质谱法也可与肽指纹数据库和各种检索算法联用。也可通过质谱法与各种预处理（诸如以糖基化酶及磷酸酶预处理）联用来探测转译后修饰（诸如磷酸化或糖基化作用）的差异。将所有这些方法都视为本申请案的一部分。

[0127] OAS1 为四聚物形式时有活性，且干扰自身缔合的突变影响酶活性。使用已知方法评估 OAS1R 突变对四聚物形成的影响。例如，以 OAS1 和 OAS1R 特异性抗体进行免疫沉淀以自患者细胞、细胞培养物或过度表达 OAS1 或 OAS1R 的经转染细胞分离 OAS1/OAS1R 复合物。随后可通过凝胶电泳法或所属领域技术人员熟知的其他色谱方法来评估这些复合物。

[0128] OAS1 可通过与其他蛋白相互作用而赋予抗病毒性。酶含有与 bcl-2 家族的 BH3 域在结构上同源的区。此域可能对 OAS1 的功能是关键性的。根据本发明，可使用 OAS1 特异性抗体分离蛋白复合物，包括来自如上讨论的各种来源的 OAS1 蛋白。随后可通过凝胶电泳法评估这些复合物以分离相互作用的复合物的成员。可使用包括 Western 印迹法的许多方法探测凝胶，且可使用肽测序法分离并鉴别新颖的相互作用的蛋白质。也将评估野生型与 OAS1R 提取物中 OAS 复合物含量的差异。正如所属领域技术人员将认识到的，所述方法仅为许多可用于纯化、鉴别和评估 OAS 复合物中相互作用的蛋白质的不同方法中的一些。额外方法包括（但不限于）噬菌体展示技术和使用酵母双杂交体方法。

[0129] 已知 OAS1 与 C 型肝炎病毒 NS5A 蛋白相互作用 (Taguchi, T. 等人, J.Gen. Virol.85 : 959-969, 2004)。不受机制的约束，本发明因此涉及不与 NS5A 蛋白相互作用的 OAS1 蛋白，其中所述蛋白为本发明的多肽，其由本发明的聚核苷酸表达、由 OAS1 的剪接变异体所编码的 mRNA 表达、由含有至少一个本发明的突变的 OAS1 聚核苷酸表达、由编码区内具有至少一个突变的 OAS1 聚核苷酸表达和 / 或由具有至少一个碱基取代、缺失或添加的 OAS1 聚核苷酸表达，其中与 NS5A 蛋白的结合被改变或阻碍。

[0130] NS5A 蛋白可通过抑制干扰素的抗病毒活性而发挥生物学活性。当干扰素在诸如双股 RNA 的辅因子的存在下刺激 OAS1 活性时，此抗病毒活性处于一正常实施的模型中。OAS1 使 ATP 聚合成 2' -5' -连接的寡腺苷酸，其激活 RNase L 以裂解包括 mRNA 的单股 RNA。NS5A 与 OAS1 的结合可抑制其活性，因此抑制或阻碍将另外导致病毒 RNA 破坏的活性级联。

[0131] 尽管本发明并非取决于此模型，但是 NS5A 与 OAS1 的结合与其中 OAS1 的突变形式避免 NS5A 结合和抑制且因而能够实现聚合 ATP 的正常功能的模型相一致。在与本文中所述的临床结果相一致的这些情况下，携带这一突变的人对 C 型肝炎病毒感染存在

抗性。类似地，如上所揭示黑猩猩所具有的 OAS1 的截短形式可避开与 NS5A 或其他病毒蛋白结合且因此可观察到更高的黑猩猩病毒清除频率。在某些情况下突变可直接影响 OAS1 用于 NS5A 的结合位点。在其他情况下，突变可位于与实际结合位点分开的位点，但引起构象改变使得 OAS1 与 NS5A 的结合受到抑制、减慢或阻止。

[0132] 因此 OAS1 与 NS5A 以生理学和病理学上相关方式的结合提供用于分析 OAS1 聚核苷酸中碱基突变、缺失或添加对所编码的 OAS1 蛋白的生物学功能的影响的客观测试。以此项技术中已知的方式分析所述结合。在一诸如 Taguchi, T. 等人 (J.Gen.Virol.85 : 959-969, 2004) 所述之例示性而非限制性方法中，HeLa 细胞经编码 GST 标记的 NS5A 和 HA 标记的 OAS1 的表达质粒瞬时转染。此实例中的 OAS1 意指根据本发明的 OAS1，包括由 OAS1 的剪接变异体、由在编码区内具有至少一个突变的 OAS1 聚核苷酸和 / 或由具有至少一个碱基取代、缺失或添加的 OAS1 聚核苷酸所编码的 OAS1。在诸如 12-16 小时的适当培育时间之后，将细胞洗涤、溶解并离心，并将所得上清液与结合谷胱甘肽的琼脂糖 (Sephrose) 珠粒混合，此举有助于分离经 GST 标记的蛋白。通过使用并成像抗 NS5A 的抗体和抗 HA 抗体来鉴别经 GST 标记的 NS5A 与经 HA 标记的 OAS1 蛋白的复合物。此方法的变型包括对个别蛋白使用其他标记，诸如 FLAG- 标记。在本发明上下文中，主要变数为 OAS1 蛋白或多肽。使用这些分析法客观地测试 OAS1 蛋白或多肽进行一个生物学相关活动 (即结合至已知对于病毒在宿主中复制的能力具有保护性的 C 型肝炎蛋白) 的能力。未与 NS5A 结合的 OAS1 蛋白和多肽是作为治疗性蛋白的合适候选者。

[0133] OAS1 蛋白是催化 ATP 转化为寡腺苷酸分子的酶。可利用数种方法来评估 OAS1 酶的活性。使用这些方法来确定 OAS1R 突变对突变型蛋白相对于野生型酶的活性的影响。例如，可通过量化并入聚腺苷酸中的经 ³²P 放射性同位素标记的 ATP 来测量寡腺苷酸合成活性。可通过包括电泳法或离子交换色谱法的许多色谱方法对经放射性同位素标记的聚腺苷酸的长度进行量化及表征。这些分析也能够表征底物 (ATP) 结合和酶动力学。OAS1 是由 dsRNA 激活。使用本文中所述的活性分析和此项技术中所述的合成 dsRNA 来分析 OAS1 中此活化作用的动力学并与 OAS1R 作比较。

[0134] 通过这些和此项技术中已知的其他方法证实本发明的多肽具有寡腺苷酸合成活性。图 7 和图 8 证实本发明的数种例示性多肽的活性。不管其活性的量化水平为何，这种产生 2' -5' -寡腺苷酸的能力经由 RNaseL 的活化而产生抗病毒效果都为所属领域技术人员所十分了解。同样，本发明的多肽具有寡腺苷酸合成活性的纯粹事实表明所述多肽尤其是因为其于以下揭示的治疗用途而具有效用。

[0135] 执行生物学研究以评估 OAS1R 突变基因使得免受病毒感染的程度。这些生物学研究一般采用以下形式：将突变型 OAS1R 基因或蛋白引入细胞或整个有机体中，并相对于野生型对照物评估其生物学及抗病毒活性。在此方法的一典型实施例中，通过将如本文中所述经分离的 cDNA 克隆到哺乳动物表达载体 (所述载体驱动所克隆的 cDNA 自 SV40 启动子序列表达) 中，从而将 OAS1R 基因引入培养物中的非洲绿猴肾 (Vero) 细胞中。此载体也含有 SV40 和细胞巨化病毒增强子元件，其允许 OAS1R 基因和用于在培养物中选择的新霉素抗性基因的有效表达。随后可评估在感染登革热病毒的 Vero 细胞中 OAS1R 表达的生物学作用。在 OAS1R 赋予对多种黄病毒的广泛抗性的情况下，将预期到表达这些 OAS1 突变形式的细胞系中的病毒繁殖相对于野生型有所减少。正如所属领

域技术人员将认识到的，有多种不同实验方法可用于评估细胞和有机体中对不同感染剂有反应的 OAS1R 基因及蛋白的生物学效应。例如，在以上实例中，不同的表达载体、细胞类型和病毒物种可用于评估 OAS1R 抗性效应。相对于细胞系评估培养物中的初级人类细胞。在病毒引入前，细胞可由双股 RNA 或干扰素刺激。可评估含有替代启动子和增强子序列的表达载体。评估除黄病毒之外的病毒（例如呼吸融合病毒和小核糖核酸病毒）。

[0136] 开发转基因动物模型以评定 OAS1 的突变形式在保护整个有机体不受病毒感染方面的适用性。在一实施例中，将 OAS1R 基因引入易受黄病毒感染的小鼠的基因组中（例如 C3H/He 近交系实验室菌株）。评估这些 OAS1R 基因于易受感染的小鼠中调节感染或赋予感染抗性的能力。正如所属领域技术人员将了解的，可使用许多标准方法将转基因人类 OAS1R 基因引入小鼠中。这些方法可与影响组织特异性表达样式或允许经由引入内源性化学品、使用可诱导性或组织特异性启动子等来调控转基因的其他方法联合。

[0137] 作为 C 型肝炎感染的模型，评估表达 OAS1R 基因的细胞系对牛痢疾病毒 (BVDV) 感染的易感性、抗性或调节。BVDV 是用于测试潜在抗 HCV 抗病毒药物的功效的常用模型 (Buckwold 等人, *Antiviral Research* 60 : 1-15, 2003)。在一实施例中，可使用基本上如上所述的表达载体将 OAS1R 基因引入 KL (小牛肺) 细胞中并测试其于此细胞系中调节 BVDV 感染的能力。此外，也可对 HCV 感染的小鼠模型（例如，人肝脏移植入小鼠中、人肝细胞注入小鼠肝脏中等等）评估 OAS1R 基因的转基因组中 HCV 感染的改变。可进行实验由此在 HCV 病毒培养系统中评定 OAS1R 基因的表达效果。

[0138] 也可使用细胞培养系统来评定突变型 OAS1R 基因在不同条件下在促进细胞凋亡方面的影响。在一实施例中，可相对于野生型 OAS1 序列评定 OAS1R 的细胞培养突变形式在受包括 BVDV、HCV 及其他黄病毒的许多病毒所感染的细胞中促进细胞凋亡的能力。正如所属领域技术人员将认识到的，可利用许多方法测量细胞凋亡。最常用的方法包括联合使用荧光结合方法与琼脂糖凝胶电泳法检测凋亡细胞中的特征性基因组“DNA 阶梯 (DNA laddering)”效应。

[0139] 可在细胞培养物和小动物模型中测试缺陷型干扰病毒增强 OAS1R 突变形式的响应的能力。

[0140] 也可检测 OAS1R 基因型的存在或不存在影响其他人类表现型的程度。例如，评估 OAS1R 突变在感染 HCV 的受检者中与病毒滴度和自发病毒清除的相关性。也可采取将宿主 OAS1 基因型与其他黄病毒感染的过程相联系的类似方法。也检测在感染 HCV 的患者中干扰素治疗或干扰素与病毒唑 (ribavirin) 治疗期间 OAS1R 突变在促进成功结果方面的影响。这些突变不仅可赋予一定程度的感染抗性，而且也促进经或未经干扰素 - 病毒唑治疗的受感染受检者中的自发病毒清除。此外，已报导精神分裂症在黄病毒感染具地域性的地理区域发生频率较高，表明黄病毒抗性等位基因与易患精神分裂症的体质之间存在联系。通过进行额外的涉及精神分裂症表现型和 OAS1R 突变的基因关联性研究来评估此联系。也将评估 OAS1R 突变对 IDDM、前列腺癌及其他癌症及精神分裂症的易感性的影响。

[0141] 本发明揭示新颖且具有效用的 OAS1R 变异型 mRNA (确定为 SEQ ID NO : 36 至 SEQ ID NO : 43)。本发明不受限于所揭示的变异型 mRNA 的使用模式。在一优选实施

例中，这些变异型 mRNA 是用于差别筛检人类受检者增加或降低的病毒（包括 HCV）易感性。在其他优选实施例中，这些变异型 mRNA 适用于筛检对 IDDM、前列腺癌及其他癌症和 / 或精神分裂症的易感性。通过所属领域技术人员已知的表达分析执行所述差别筛检以测定一种或一种以上存在于来源于给定人类受检者的样本中的变异型 OAS1R mRNA 的相对量。相对于对照样本而言，人类受检者样本中一种或一种以上 OAS1R mRNA 变异体的量增加或降低表明受检者易感染病毒、IDDM、前列腺癌及其他癌症和 / 或精神分裂症的程度，此适于所考虑的测试。

[0142] 如本文中所述讨论，2'，5' -寡腺苷酸合成酶 (OAS) 为 IFN- α 诱导性、RNA 依赖性效应分子酶家族，其自 ATP 合成短的 2' 至 5' 连接的寡腺苷酸 (2-5A) 分子。OAS 酶构成对病毒感染的非特异性免疫防御的重要部分且已用作病毒感染的细胞标记物。除在本文中所述讨论的 C 型肝炎感染中的作用外，在其他病状、尤其是那些其中病毒感染起作用的病状中也涉及 OAS 活性。

[0143] 虽然当前分析的主题为特异性致病机制，但是据信病毒感染在诸如糖尿病的疾病的发展中起作用。1 型糖尿病患者中淋巴细胞 OAS 活性显著提高，表明 OAS 可能为病毒感染与疾病发展之间的重要联系。在涉及来自单卵双生对的糖尿病双胞胎的研究中，Bonnie Nielsen 等人 (Clin Immunol.2000 Jul ; 96(1) : 11-8) 显示 OAS 在新近发作和长期存在的 1 型糖尿病中均得以持久激活。1 型糖尿病中连续升高的 OAS 活性与正常抗病毒反应明显不同且可能表示酶的慢性刺激、下调机制的失效或对内源性或外源性病毒或其产物的异常反应。

[0144] 病毒感染与糖尿病发展之间更直接的联系由许多研究加以例证，这些研究显示介于 13% 与 33% 之间的慢性 C 型肝炎患者患有糖尿病 (2 型糖尿病)，此水平与匹配的健康对照或慢性 B 型肝炎患者相比显著增加 (Knobler 等人, Am J Gastroenterol.2003 Dec ; 98(12) : 2751-6)。虽然迄今为止尚未报导 OAS 对在 C 型肝炎感染之后发展成糖尿病起作用，但是其可为用于抗病毒反应系统的有用标记物。此外，根据本发明所报导的结果说明，若 C 型肝炎感染与糖尿病有因果关系，则使用本文中所揭示的组合物和方法抑制或消除 C 型肝炎感染可有利于预防或减轻糖尿病的发展。

[0145] 另一公开研究已显示 OAS 在伤口痊愈及其病理学病症方面、尤其是在静脉溃疡和与糖尿病相关的不易痊愈伤口的情况下起本质作用 (WO 02/090552)。在不易痊愈伤口的情况下，受感染组织中 OAS mRNA 含量减少，而不是象在来源于 1 型糖尿病患者的淋巴细胞中那样升高。这些发现针对 OAS 作为糖尿病和与糖尿病相关的伤口痊愈中的免疫反应的病因学上的重要标记物。

[0146] OAS 也可在前列腺癌中所涉及的细胞过程中起中间物作用。OAS 的主要生物学功能为促进 RNaseL 的活性，RNase L 是一种由 2-5A 分子进行酶刺激的受独特调控的核糖核酸内切酶。RNaseL 在调节 IFN 的抗病毒效应上具有非常确实的作用，且为遗传性前列腺癌 1 等位基因 (HPC1) 的强力候选物。已显示 RNaseL 中的突变倾向于增加男性中的前列腺癌发病率，在某些情况中其表现为与非 RNaseL 相关病例相比更具侵袭性的疾病和 / 或发作年龄降低。Xiang 等人 (Cancer Res.2003 Oct 15 ; 63(20) : 6795-801) 证实 2-5A 的生物稳定性硫代磷酸酯类似物诱导 RNaseL 活性并引起晚期转移性人类前列腺癌细胞系培养物中的细胞凋亡。他们的发现表明同时伴随 2-5A 含量增加的 OAS 活性提高可经由

有效的细胞凋亡途径而有利于癌细胞的毁灭。因此，本文中所揭示的组合物和方法的使用可在前列腺癌的检测、治疗和 / 或预防中具有效用。

[0147] 另外 OAS 可通过其对 RNaseL 的调控或经由另一尚未发现的途径而在正常细胞生长调控中起作用。有大量证据支持 OAS 在负调控细胞生长中的重要性。Rysiecki 等人 (J.Interferon Res.1989 Dec ; 9(6) : 649-57) 证实人类 OAS 至神经胶质瘤细胞系中的稳定转染导致细胞增殖减少。在几项将静止细胞系与增殖细胞系相比的研究中 OAS 含量也已显示为可测量的 (例如 Hassel 和 Ts' O, Mol Carcinog.1992 ; 5(1) : 41-51 及 Kimchi 等人, Eur J Biochem.1981 ; 114(1) : 5-10) 且在每一情况下 OAS 含量均为在静止细胞中最高。其他研究已显示 OAS 含量与细胞周期阶段之间的相关性, 其中 OAS 含量在后 S 阶段急剧上升而接着在 G2 阶段突然下降 (Wells 和 Mallucci, Exp Cell Res.1985 Jul ; 159(1) : 27-36)。几项研究已显示在经各种形式的干扰素刺激后 OAS 诱导作用与抗增殖效应发作之间的相关性 (参见 Player 和 Torrence, Pharmacol Ther.1998 May ; 78(2) : 55-113)。细胞分化期间也显示出 OAS 的诱导作用 (例如 Salzberg 等人, J Cell Sci.1996 Jun ; 109(Pt6) : 1517-26 及 Schwartz 和 Nilson, Mol Cell Biol.1989 Sep ; 9(9) : 3897-903)。关于 OAS 由血小板衍生生长因子 (PDGF) (Zullo 等人, Cell.1985 Dec ; 43(3 Pt 2) : 793-800) 在热休克诱导生长条件下 (Chousterman 等人, J Biol Chem.1987 Apr 5 ; 262(10) : 4806-11) 诱导的其他报导导致以下假设: OAS 的诱导作用为正常细胞生长控制机制。因此, 本文中所揭示的组合物和方法的使用可在癌症的检测、治疗和 / 或预防中具有广泛效用。

[0148] 聚核苷酸分析

[0149] 寡腺苷酸合成酶基因是一种核酸, 其核苷酸序列编码寡腺苷酸合成酶、突变型寡腺苷酸合成酶或寡腺苷酸合成酶假基因。其可呈基因组 DNA、mRNA 或 cDNA 形式, 且呈单股或双股形式。优选使用基因组 DNA, 因为其相对于 mRNA 而言在生物学样本中具有相对稳定性。在序列表中以 SEQ ID NO : 19 提供由参考寡腺苷酸合成酶基因的完整基因组序列的连续核苷酸 2,130,000-2,157,999 组成的聚核苷酸序列, 且对应于 Genbank 登记号 NT_009775.13。

[0150] 自细胞 (通常为周边血液白细胞) 获得核酸样本。在使用 mRNA 时, 于 RNase 抑制条件下溶解细胞。在一实施例中, 第一步为分离总细胞 mRNA。随后可通过与寡聚 dT 纤维素柱杂交来选择 Poly A+mRNA。

[0151] 在一优选实施例中, 富集核酸样本以获得寡腺苷酸合成酶等位基因物质。通常通过使用如本文中所述的聚核苷酸合成引物使基因组 DNA 或 mRNA 经受引物延伸反应来实现富集。用于产生待分析样本的尤其优选方法为在聚合酶链式反应 (PCR) 中使用预选的聚核苷酸作为引物以形成扩增 (PCR) 产物。

[0152] 聚核苷酸引物的制备

[0153] 将本文中关于欲通过引物延伸而合成的引物、探针和核酸片段或节段所使用的术语 “聚核苷酸” 定义为包含两种或两种以上、优选为三种以上脱氧核糖核苷酸或核糖核苷酸的分子。其精确大小将取决于许多因素, 这些因素又取决于最终使用条件。

[0154] 本文中所用的术语 “引物” 意指自核酸限制性消化纯化或由合成产生的聚核苷酸, 当置于诱导与核酸链互补的引物延伸产物的合成的条件下 (即在核苷酸和诸如 DNA

聚合酶、逆转录酶等等的聚合剂存在下且在合适温度和 pH 下) 时, 其能够充当核酸合成的起始点。引物优选为单股以达成最大效率, 但或者可呈双股形式。若引物为双股, 则在用于制备延伸产物之前首先对其加以处理以与其互补股分离。引物优选为聚脱氧核糖核苷酸。引物必须足够长以在聚合剂的存在下引发延伸产物的合成。引物的精确长度将取决于许多因素, 包括温度和引物的来源。例如, 视目标序列的复杂性而定, 聚核苷酸引物通常含有 15 至 25 个或更多核苷酸, 虽然其也可含有较少的核苷酸。短引物分子一般需要较低温度以与模板形成足够稳定的杂交体复合物。

[0155] 选择本文中所述的引物以与待合成或扩增的每一特定序列的不同链“实质上”互补。此意谓引物必须充分互补以与其各自的模板链非随机杂交。因此, 引物序列可能反映或可能不反映模板的精确序列。例如, 非互补核苷酸片段可连接于引物的 5' 末端, 而引物序列的剩余部分与所述链实质上互补。这些非互补片段通常编码核酸内切酶限制性位点。或者, 非互补碱基或更长序列可引入引物中, 条件是引物序列具有与待合成或扩增的链序列的足够互补性以与其非随机杂交且因而在聚核苷酸合成条件下形成延伸产物。

[0156] 本发明的引物也可含有 DNA 依赖性 RNA 聚合酶启动子序列或其互补序列。例如参见 Krieg 等人, Nucl.Acids Res., 12 : 7057-70 (1984); Studier 等人, J.Mol.Biol., 189 : 113-130 (1986); 和 Molecular Cloning : A Laboratory Manual, 第二版, Maniatis 等人编, Cold SpringHarbor, N.Y. (1989)。

[0157] 当使用含 DNA 依赖性 RNA 聚合酶启动子的引物时, 所述引物与待扩增的聚核苷酸链杂交且使用诸如大肠杆菌 DNA 聚合酶 I 或大肠杆菌 DNA 聚合酶的 Klenow 片段的诱导剂完成 DNA 依赖性 RNA 聚合酶启动子的第二聚核苷酸链。起始聚核苷酸通过 RNA 聚核苷酸产生与 DNA 聚核苷酸产生之间交替来扩增。

[0158] 引物也可含有用于 RNA 指导的 RNA 聚合酶的模板序列或复制起始位点。典型的 RNA 指导的 RNA 聚合酶包括由 Lizardi 等人, Biotechnology, 6 : 1197-1202 (1988) 描述的 QB 复制酶。RNA 指导的聚合酶自少量含有模板序列或复制起始位点的模板 RNA 链生成大量 RNA 链。如 Kramer 等人, J.Mol.Biol., 89 : 719-736 (1974) 已描述, 这些聚合酶通常使模板链扩增一百万倍。

[0159] 可使用任何合适方法制备聚核苷酸引物, 诸如磷酸三酯法或磷酸二酯法, 参见 Narang 等人, Meth.Enzymol., 68 : 90, (1979); 美国专利第 4,356,270 号、第 4,458,066 号、第 4,416,988 号、第 4,293,652 号; 和 Brown 等人, Meth.Enzymol., 68 : 109 (1979)。

[0160] 引物的核苷酸序列的选择取决于下列因素: 诸如自杂交点至编码待检测突变的区的核酸距离、核酸上相对于任何待使用的第二引物的杂交位点等等。若通过 PCR 扩增来富集核酸样本以获得寡腺苷酸合成酶基因物质, 则必须使用两个引物(即 PCR 引物对)以用于待扩增核酸的每一编码链。第一引物成为非编码(反义或负或互补)链的一部分且与正链或编码链上的核苷酸序列杂交。第二引物成为编码(有义或正)链的一部分且与负链或非编码链上的核苷酸序列杂交。第一和第二引物中的一者或两者可含有界定核酸内切酶识别位点的核苷酸序列。所述位点可与正在扩增的寡腺苷酸合成酶基因异源。

[0161] 在一实施例中, 本发明使用一组形成具有位于引物 3' - 末端的引发区的引物的聚核苷酸。引发区通常为 3' - 端(3' - 末端)15 至 30 个核苷酸碱基。每一引物的

3' - 末端引发部分均能作为引物以催化核酸合成, 即引发自其 3' 末端起始的引物延伸反应。引物中的一者或两者可另外含有 5' - 末端 (5' - 端) 非引发部分, 即不参与与优选模板杂交的区。

[0162] 在 PCR 中, 每一引物与第二引物一起发挥作用以扩增目标核酸序列。用于 PCR 中的 PCR 引物对的选择是由如本文中对于生成寡腺苷酸合成酶基因区所讨论的考虑因素来决定。当选择引物序列以与寡腺苷酸合成酶基因等位基因内含子内的目标序列杂交 (退火) 时, 目标序列在等位基因中应为保守的以确保产生待分析的目标序列。

[0163] 聚合酶链式反应

[0164] 寡腺苷酸合成酶基因包含聚核苷酸编码链, 诸如 mRNA 和 / 或基因组 DNA 的有义链。若待分析的遗传物质呈双股基因组 DNA 的形式, 则通常通过熔融首先将其变性成为单股。通过以 PCR 引物对处理 (接触) 样本而使核酸经受 PCR 反应, 引物对中的每一成员具有经预选的核苷酸序列。PCR 引物对能够通过与其长度优选为至少约 10 个核苷酸、更优选为至少约 20 个核苷酸且在寡腺苷酸合成酶等位基因内为保守的核苷酸序列杂交来引发引物延伸反应。PCR 引物对的第一引物在本文中有时称为“反义引物”, 因为其与核酸的非编码链或反义链 (即编码链的互补链) 杂交。PCR 引物对的第二引物在本文中有时称为“有义引物”, 因为其与核酸的编码链或有义链杂交。

[0165] 通过将 PCR 引物对 (优选为其预定量) 与样本核酸 (优选为其预定量) 在 PCR 缓冲液中混合以形成 PCR 反应混合物来进行 PCR 反应。将所述混合物热循环足以形成 PCR 反应产物的若干周期 (通常为预定的), 从而富集待分析样本以获得寡腺苷酸合成酶遗传物质。

[0166] PCR 通常通过热循环进行, 即在下限为约 30 摄氏度 (30°C) 至约 55°C 和上限为约 90°C 至约 100°C 的温度范围内重复增加及降低 PCR 反应混合物的温度。所述增加及降低可是连续的, 但优选为在每一有利于聚核苷酸合成、变性和杂交的温度下具有相对温度稳定性时期的时相。

[0167] 在每一扩增中均可使用多个第一引物和 / 或多个第二引物, 例如, 一类第一引物可与许多不同的第二引物配对以形成几种不同的引物对。或者, 可使用第一与第二引物的个别对。在任何情况下, 均可将使用第一引物与第二引物的相同或不同组合所扩增的扩增产物组合用于分析突变。

[0168] 使用任何合适方法进行 PCR 反应。其一般于 pH 值优选为 7-9、最优选为约 8 的缓冲水溶液 (即 PCR 缓冲液) 中发生。优选地, 将摩尔过量 (对于基因组核酸而言, 通常引物: 模板约为 10^6 : 1) 的引物与含有模板链的缓冲液混合。优选为大量摩尔过量以提高过程的效率。

[0169] PCR 缓冲液也含有三磷酸脱氧核糖核苷酸 (聚核苷酸合成底物) dATP、dCTP、dGTP 和 dTTP 和通常热稳定的聚合酶, 所有物质均为用于引物延伸 (聚核苷酸合成) 反应的足够量。将所得溶液 (PCR 混合物) 加热至约 90°C -100°C 历时约 1 至 10 分钟, 优选为 1 至 4 分钟。在此加热时期之后, 将溶液冷却至 54°C, 此温度对于引物杂交是优选的。合成反应可在室温至超过聚合酶 (诱导剂) 不再有效发挥作用的温度下发生。重复进行热循环直至生成所要量的 PCR 产物。例示性 PCR 缓冲液包含下列物质: 每 100 微升缓冲液中有 50mM KCl; 10mM pH 值为 8.3 的 Tris-HCl; 1.5mM MgCl₂; 0.001% (重量 /

体积)明胶; 200 μ M dATP; 200 μ M dTTP; 200 μ M dCTP; 200² μ M dGTP; 和 2.5 单位水生栖热菌 (*Thermus aquaticus*) (Taq) DNA 聚合酶 I (美国专利第 4,889,818 号)。

[0170] 诱导剂可为将发挥作用以实现引物延伸产物的合成的任何化合物或系统, 包括酶。为达成此目的合适酶例如包括大肠杆菌 DNA 聚合酶 I、大肠杆菌 DNA 聚合酶 I 的 Klenow 片段、T4 DNA 聚合酶、其他可利用的 DNA 聚合酶、逆转录酶及其他酶(包括热稳定性酶), 其将有利于核苷酸以适当方式组合以形成与每一核酸链互补的引物延伸产物。一般来说, 合成将于每一引物的 3' 末端引发且沿模板链以 5' 方向行进直至合成终止, 产生不同长度的分子。然而, 可存在使用如上所述的相同过程而于 5' 末端引发合成且以上述方向行进的诱导剂。

[0171] 诱导剂也可为将发挥作用以实现 RNA 引物延伸产物的合成的化合物或系统, 包括酶。在优选实施例中, 诱导剂可为 DNA 依赖性 RNA 聚合酶, 诸如 T7 RNA 聚合酶、T3RNA 聚合酶或 SP6 RNA 聚合酶。这些聚合酶生成互补 RNA 聚核苷酸。如 Chamberlin 等人, *The Enzymes*, P.Boyer 编, 第 87-108 页, Academic Press, New York (1982) 已描述, RNA 聚合酶的高转换率 (turn-over rate) 扩增起始聚核苷酸。基于转录的扩增系统已由 Gingeras 等人在 *PCR Protocols: A Guide to Methods and Applications*, 第 245-252 页, Innis 等人编, Academic Press, Inc., San Diego, Calif. (1990) 中描述。

[0172] 若诱导剂为 DNA 依赖性 RNA 聚合酶且因而合并三磷酸核糖核苷酸, 则将足够量的 ATP、CTP、GTP 和 UTP 与引物延伸反应混合物混合并如上所述处理所得溶液。

[0173] 新合成链及其互补核酸链形成可用于所述过程的随后步骤中的双股分子。

[0174] PCR 反应可有利地用于将适用于检测寡腺苷酸合成酶基因中的突变的预选限制性位点并入所述产物中。

[0175] PCR 扩增方法在美国专利第 4,683,192 号、第 4,683,202 号、第 4,800,159 号和第 4,965,188 号中和至少在包括 *PCR Technology: Principles and Applications for DNA Amplification*, H.Erlich 编, Stockton Press, New York (1989) 和 *PCR Protocols: A Guide to Methods and Applications*, Innis 等人编, Academic Press, San Diego, Calif. (1990) 的若干教科书中详细描述。

[0176] 在某些实施例中, 每个扩增反应使用两对第一和第二引物。自多个不同扩增反应(每个反应使用多个不同引物对)获得的扩增反应产物可一起或单独进行分析。

[0177] 然而, 本发明涵盖使用仅一对第一和第二引物的扩增反应。以下表 1 中显示用于扩增含有本文中所揭示的突变的 DNA 部分的例示性引物。扩增子 A 对应于含有 SEQ ID NO: 1-3 中所指的突变的聚核苷酸序列。扩增子 B 对应于含有 SEQ ID NO: 4-7 和 SEQ ID NO: 60 中所指的突变的聚核苷酸序列。扩增子 C 对应于含有 SEQ ID NO: 57 中所指的突变的聚核苷酸序列。扩增子 D 对应于含有 SEQ ID NO: 58 中所指的突变的聚核苷酸序列。扩增子 E 对应于含有 SEQ ID NO: 59 中所指的突变的聚核苷酸序列。扩增子 F 对应于含有 SEQ ID NO: 61 中所指的突变的聚核苷酸序列。扩增子 G 对应于含有 SEQ ID NO: 62-64 中所指的突变的聚核苷酸序列。

[0178] 表 1

[0179] 含有本发明的突变的扩增子

[0180]

扩增子	引物 A	引物 B	产物大小 (bp)
扩增子 A	5' -AATGGACCTCAAGACTTCCC-3' (SEQ ID NO : 8)	5' -ATTCTCCCTTCTGTTGCAGG-3' (SEQ ID NO : 9)	509
扩增子 B	5' -TCCAGATGGCATGTCACAGT-3' (SEQ ID NO : 10)	5' -GAGCTATGCTTGGCACATAG-3' (SEQ ID NO : 11)	747
扩增子 C	5' - CACAAAGAGTGAACCTTAATGT-3' (SEQ ID NO : 65)	5' - CCAGGAAGTGGAAAGATCAT-3' (SEQ ID NO : 66)	603
扩增子 D	5' - ATCTCCCACAGTTTGAGAGC-3' (SEQ ID NO : 67)	5' - TCAGCCTCCAAAAGTGTGG-3' (SEQ ID NO : 68)	553
扩增子 E	5' - GGGTACATGTGCACAATGTG-3' (SEQ ID NO : 69)	5' - CCCTTATACAAAATTCAACTC-3' (SEQ ID NO : 70)	532
扩增子 F	5' - GAGCCAAGAAGTACAGATGC-3' (SEQ ID NO : 71)	5' - AGGACAGAGCTGTCCAATAG-3' (SEQ ID NO : 72)	648
扩增子 G	5' - GGCTCAGAGAAGCTAAGTGA-3' (SEQ ID NO : 73)	5' - CCACAGCATCCTTTTCAGTC-3' (SEQ ID NO : 74)	581

[0181] 表 2 揭示以上扩增子中本发明突变的位置

[0182] 表 2

[0183] 扩增子中本发明的突变的位置

[0184]

突变	扩增子	扩增子中的位置 (相对于扩增子的引物 A 侧的 5' 末端)
1 (SEQ ID NO : 1)	扩增子 A	134
2 (SEQ ID NO : 2)	扩增子 A	155
3 (SEQ ID NO : 3)	扩增子 A	384
4 (SEQ ID NO : 4)	扩增子 B	98

5 (SEQ ID NO : 5)	扩增子 B	114
6 (SEQ ID NO : 6)	扩增子 B	142
7 (SEQ ID NO : 7)	扩增子 B	347
8 (SEQ ID NO : 57)	扩增子 C	319
9 (SEQ ID NO : 58)	扩增子 D	404
10 (SEQ ID NO : 59)	扩增子 E	133
11 (SEQ ID NO : 60)	扩增子 B	320
12 (SEQ ID NO : 61)	扩增子 F	367
13 (SEQ ID NO : 62)	扩增子 G	138
14 (SEQ ID NO : 63)	扩增子 G	210
15 (SEQ ID NO : 64)	扩增子 G	253

[0185] 核酸序列分析

[0186] 通过 (a) 基于探针链及其互补目标的杂交或变性的物理化学技术与 (b) 与核酸内切酶、连接酶和聚合酶的酶促反应的组合来进行核酸序列分析。可在 DNA 或 RNA 水平上分析核酸。前者分析个体人类的遗传潜力而后者分析特定细胞的表达信息。

[0187] 在使用核酸杂交的分析中，可通过多种手段实现检测在本发明的过程中 DNA 双螺旋的存在。

[0188] 在一用于检测 DNA 双螺旋存在的方法中，在 DNA 双螺旋中经杂交的寡核苷酸包括使双螺旋可检测的标记或指示基团。所述标记通常包括放射性原子、经化学修饰的核苷酸碱基等等。

[0189] 可对寡核苷酸进行标记，即可经由操作与指示工具或基团连接，且可将其用于检测目标模板中特定核苷酸序列的存在。可经由操作连接至寡核苷酸探针（经标记的寡核苷酸）或作为寡核苷酸探针的一部分存在的放射性元素提供有利于检测 DNA 双螺旋的有用方法。典型的放射性元素为产生 β 射线发射的元素。诸如 ^3H 、 ^{12}C 、 ^{32}P 和 ^{35}S 等发射 β 射线的元素代表一类产生 β 射线发射的放射性元素标记。通常通过使用 DNA 激酶将经放射性标记的核苷酸以酶促法并入核酸中来制备放射性聚核苷酸探针。经放射性标记的寡核苷酸的替代物为经化学修饰以含有金属络合剂、含生物素基团、荧光化合物等等的寡核苷酸。

[0190] 一种有用的金属络合剂为由镧系元素与芳族 β -二酮形成的镧系元素螯合物，镧系元素经由形成螯合物的化合物（诸如 EDTA 类似物）结合至核酸或寡核苷酸，从而形

成荧光性镧系元素络合物。参见美国专利第 4,374,120 号、第 4,569,790 号和公开专利申请案 EP0139675 及 W087/02708。

[0191] 已描述经生物素或吡啶酯标记的寡核苷酸及其在标记聚核苷酸上的用途。参见美国专利第 4,707,404 号、公开专利申请案 EP0212951 和欧洲专利第 0087636 号。有用的荧光标记化合物包括荧光素、若丹明 (rhodamine)、德克萨斯红 (Texas Red)、NBD 等等。

[0192] 存在于 DNA 双螺旋中的经标记的寡核苷酸使双螺旋自身得以标记且因此可区别于待分析样本中所存在的其他核酸。检测双螺旋中标记的存在且藉此检测双螺旋的存在，通常包括将 DNA 双螺旋自任何经标记的未与 DNA 双螺旋杂交的寡核苷酸探针中分离。

[0193] 用于自 DNA 双螺旋中分离单股寡核苷酸（诸如未经杂交的经标记寡核苷酸探针）的技术为人所熟知，且通常包括基于其化学性质将单股核酸自双股核酸中分离。更常用的分离技术包括使用异质杂交型式，其中未经杂交的探针通常通过洗涤而自与不溶性基质结合的 DNA 双螺旋中分离。例示性实例为 Southern 印迹技术，其中基质为硝化纤维素薄片且标记为 ^{32}P 。Southern, J.Mol.Biol., 98 : 503(1975)。

[0194] 也可通常于其 5' - 末端或附近处有利地将寡核苷酸连接至固体基质，即非水溶性固体支撑物。有用的固体基质为此项技术中所熟知且包括诸如以 SEPHADEX 为商品名购自 Pharmacia Fine Chemicals (Piscataway, N.J.) 的交联葡聚糖、琼脂糖、直径为约 1 微米至约 5 毫米的聚苯乙烯或乳胶珠粒、聚氯乙烯、聚苯乙烯、交联聚丙烯酰胺、基于硝化纤维素或尼龙的织物，诸如薄片、条带、浆叶、平板、微滴定板孔等等。

[0195] 也有可能将“连接”核苷酸加入成员寡核苷酸的 5' 或 3' 末端，并使用所述连接寡核苷酸经由操作将此成员连接到固体支撑物。在核苷酸杂交分析中，在所涵盖方法中杂交反应混合物于杂交条件下维持一段足够长时间以使得与模板上的预定序列互补的寡核苷酸与存在于模板中的互补核酸序列杂交以形成杂交产物，即含有寡核苷酸和目标核酸的复合物。当使用维持时期时，短语“杂交条件”及其合乎文法的均等物表示使杂交反应混合物在反应物和伴随试剂于混合物中的浓度范围内经受足以使一种或一种以上寡核苷酸与目标序列退火以形成核酸双螺旋的时间、温度和 pH 条件。如此项技术中所熟知，这些为实现杂交所需要的时间、温度和 pH 条件取决于以下因素：待杂交寡核苷酸的长度、寡核苷酸与目标之间的互补性程度、寡核苷酸的鸟嘌呤及胞嘧啶含量、所要的杂交严格度和在杂交反应混合物中可影响杂交动力学的盐或额外试剂的存在。用于优化给定杂交反应混合物的杂交条件的方法为此项技术中所熟知。典型杂交条件包括使用缓冲至介于 4 与 9 之间的 pH 值的溶液，并在 4°C 至 37°C、优选为约 12°C 至约 30°C、更优选为约 22°C 的温度下进行，且持续 0.5 秒至 24 小时、优选为 2 分钟至 1 小时的时期。

[0196] 如所熟知的，杂交可以同质或异质型式进行。同质杂交反应完全于溶液中进行，其中寡核苷酸和待杂交的核酸序列（目标）均以可溶形式存在于溶液中。异质反应包括使用不溶于反应介质且寡核苷酸、聚核苷酸探针或目标核酸与其相结合的基质。

[0197] 当含有目标序列的核酸呈双股 (ds) 形式时，则最好在杂交反应之前首先通过加热或碱处理来使双股 DNA 变性。双股 DNA 的变性可在与待杂交的寡核苷酸混合之前进行，或可在双股 DNA 与寡核苷酸混合之后进行。

[0198] 寡核苷酸与模板之间的预定互补性是以两种替代方式达成。模板 DNA 中的序列可为已知，诸如其中待形成的引物可与已知寡核苷酸合成酶序列杂交并可引发引物延伸至用于测序目的以及随后的上述分析目的的 DNA 区中，或其中先前测序已确定核苷酸序列的区且引物经设计为自最近经测序区向未知序列区延伸。已将后一种方法称为“定向测序”，因为每一轮测序都是由基于先前确定的序列而设计的引物来指导。

[0199] 存在于杂交反应混合物中的寡核苷酸的有效量为一般熟知的且通常以待杂交的寡核苷酸与模板之间的摩尔比表示。优选比率为含有等摩尔量的目标序列与寡核苷酸的杂交反应混合物。如人们所熟知，偏离等摩尔浓度时仍将产生杂交反应产物，只是效率较低。因此，尽管比率中一种组份可相对于另一组份摩尔过量多达 100 倍，但在实施本发明中所想要的是过量小于 50 倍、优选为小于 10 倍且更优选为小于 2 倍。

[0200] 经薄膜固定化的目标序列的检测

[0201] 在 DNA (Southern) 印迹技术中，通过先前讨论的 PCR 扩增来制备 DNA。在琼脂糖凝胶中根据尺寸分离 PCR 产物 (DNA 片段) 并将其转移 (转渍) 至硝化纤维素或尼龙薄膜上。常规电泳法分离长度范围为 100 至 30,000 个碱基对的片段而脉冲场凝胶电泳法分离长达 20,000,000 个碱基对的片段。在薄膜上含有特定 PCR 产物的位置是通过与特异性经标记的核酸探针杂交来确定。

[0202] 在优选实施例中，使用斑点印迹 (dot-blot) (细长印迹 (slot-blot)) 装置将 PCR 产物直接固定至固体基质 (硝化纤维素膜) 上，且通过探针杂交来分析。参见美国专利第 4,582,789 号和第 4,617,261 号。

[0203] 可通过以等位基因特异性寡核苷酸 (ASO) 探针探测来分析经固定的 DNA 序列，这些探针为长度约 15、17、20、25 或长达约 30 个核苷酸的合成 DNA 寡聚物。这些探针足够长以呈现基因组中的独特序列，但足够短以便由于其与目标分子杂交中的内部错配而失去稳定。因此，可通过在小心控制的杂交条件下 ASO 探针与正常或突变型目标之间的杂交体的不同变性行为来区分在单个核苷酸上不同的任何序列。例示性探针在本文中揭示为 SEQ ID NO: 1-7 和 SEQ ID NO: 57-64 (表 3)，但只要探针与携带所选择的点突变的 OAS1 基因区特异性杂交且能够特异性区分携带点突变的聚核苷酸与野生型聚核苷酸，则任何探针均为适合的。

[0204] 溶液中目标序列的检测

[0205] 已开发出数种不需要核酸纯化或固定化的快速技术。例如，可将探针 / 目标杂交体于优选结合双股核酸的固体基质 (诸如羟基磷灰石) 上选择性分离。或者，可将探针核酸固定于固体支撑物上并用于自溶液中俘获目标序列。目标序列的检测可在第二经标记探针的辅助下实现，所述探针在竞争型分析法中由目标序列自支撑物中置换或经由目标序列的桥接作用以夹层型式连接至支撑物。

[0206] 在寡核苷酸连接分析法 (OLA) 中，酶 DNA 连接酶是用于共价连接两条所选择的合成寡核苷酸序列，使得两者可以精确的头尾 (head-to-tail) 并置方式与目标序列碱基配对。接合区的错配核苷酸的存在妨碍两种寡聚物连接。此程序考虑到在不需 DNA 纯化的情况下细胞样本中已知序列变异体之间的差别。可通过固定两种寡核苷酸中的一种并观察第二经标记寡核苷酸是否也被俘获来监控两种寡核苷酸的连接。

[0207] 用于检测碱基取代的扫描技术

[0208] 三种技术允许分析长度为几百个碱基对的探针 / 目标双螺旋中未知的单一核苷酸取代或其他序列差异。在核糖核酸酶 (RNase) A 技术中, 所述酶在经标记的 RNA 探针与目标 RNA 或 DNA 序列错配的位置处裂解经标记的 RNA 探针。考虑到对突变大致位置的确定, 可按照尺寸分离片段。参见美国专利第 4,946,773 号。

[0209] 在变性梯度凝胶技术中, 用电泳法以渐增强度的变性梯度来分析探针 - 目标 DNA 双螺旋。变性伴随着迁移率降低。带有错配碱基对的双螺旋比完美配对的双螺旋变性更快。

[0210] 第三种方法依赖于错配碱基对的化学裂解。可在异源双螺旋中检测 T 与 C、G 或 T 之间的错配以及 C 与 T、A 或 C 之间的错配。与四氧化钨 (T 与 C 错配) 或羟基胺 (C 错配) 反应之后以哌啶处理可在适当错配处裂解探针。

[0211] 用于恢复和 / 或增强 OAS1 功能的治疗剂

[0212] 无论是经由 OAS1 蛋白的含量降低、影响蛋白功能的蛋白突变还是经由其他机制, 当 OAS1 基因中的突变导致缺陷性 OAS1 功能且此缺陷性功能与患者对病原体感染的易感性增加相关时, 用野生型 OAS1 蛋白治疗患者均可能是有利的。此外, 若突变在抗感染携带者中产生与野生型蛋白不同且在抑制 HCV 感染方面具有优势的蛋白形式, 则投与由突变基因编码的蛋白可为有利的。如先前所述, 在包括 (但不限于) 癌症、糖尿病和伤口痊愈的其他病症的治疗中, 投与天然或突变形式的 OAS1 蛋白或多肽也可为有利的。以下讨论是关于任何上述蛋白或多肽的投与。

[0213] 本发明的多肽 (包括那些由 OAS1R 编码者) 可为天然纯化产物或化学合成程序的产物, 或通过重组技术自本发明的聚核苷酸序列的原核或真核宿主 (例如培养物中的细菌、酵母、高等植物、昆虫和哺乳动物细胞) 产生。视重组生产程序中所用的宿主而定, 本发明的多肽可经哺乳动物或其他真核生物碳水化合物进行糖基化或可未经糖基化。本发明的多肽也可包括初始甲硫氨酸氨基酸残基 (于负 1 位置处)。

[0214] 本发明的多肽也包括 SEQ ID NO : 20、SEQ ID NO : 21、SEQ ID NO : 22、SEQ ID NO : 23、SEQ ID NO : 24 及 SEQ ID NO : 25、SEQ ID NO : 26、SEQ ID NO : 27、SEQ ID NO : 28、SEQ ID NO : 29、SEQ ID NO : 32、SEQ ID NO : 33、SEQ ID NO : 34、SEQ ID NO : 35、SEQ ID NO : 46 和 SEQ ID NO : 47 及其衍生物中所界定的蛋白质序列。

[0215] 除天然产生的等位基因形式的多肽之外, 本发明也包含其类似物和片段, 其功能与天然产生的等位基因形式相似。因此, 举例而言, 只要保持 OAS1R 蛋白的功能, 则多肽的氨基酸残基中的一种或一种以上可由保守氨基酸残基置换。以下实例 8-10 提供关于本发明多肽的合适氨基酸置换的代表性例证。作为另一实例, 本发明的多肽特定地包括 SEQ ID NO : 48、SEQ ID NO : 49、SEQ ID NO : 50、SEQ ID NO : 51 和 SEQ ID NO : 52 中所界定的 OAS1R 的截短形式或类似物形式。如先前所讨论, SEQ ID NO : 51 表示黑猩猩所具有的 OAS1 的截短形式而 SEQ ID NO : 52 表示大猩猩所具有的稍长但仍为截短形式的羧基末端片段。SEQ ID NO : 49 和 SEQ ID NO : 50 分别表示按照对应的黑猩猩和大猩猩截短位点而截短的合成人 OAS1R 构筑体。SEQ ID NO : 48 表示按照黑猩猩与大猩猩形式的长度中间值而截短的合成人 OAS1R 多肽。通过如本文中其他处所揭示的此项技术中已知的方法进一步证明 SEQ ID NO : 48 具有酶活性。因此, 也可证明剩

余高度类似的截短形式具有酶活性。正如所属领域技术人员所了解的，OAS1R 多肽的截短形式而非功能性形式的治疗用途可阻止另外阻碍多肽治疗功效的抗体反应的发展。上述截短型多肽和所属领域技术人员所设想的其他形式多肽保持功能但却移除多肽的非普遍存在部分，所述部分能诱导不具有内源性全长 OAS1R 多肽的个体中的抗体反应。所属领域技术人员也将了解较小多肽一般来说更经受得住治疗进程中通常所遭遇的制造、传递和清除的复杂性。另外，所属领域技术人员将了解黑猩猩和大猩猩中不同纯合截短变异体的出现也高度表明目前所揭示的截短型 OAS1 形式的广泛抗病毒效力。尽管以上特定揭示的截短型多肽形式表示对多肽羧基末端的截短，但是本发明不受限于所述片段的形式且特定包括保留酶促功能的氨基末端截短和内部氨基酸缺失。

[0216] 保留特定揭示的多肽的至少一种活性但与所揭示的氨基酸序列不同的多肽也包括在本发明的范畴内。如使用标准比对方法所计算的，所述多肽优选地具有与相应所揭示的 SEQ ID NO：相比至少 80% 序列同源性、优选为 85% 序列同源性、更优选为 90% 序列同源性、最优选为 95% 以上序列同源性。

[0217] 多肽也可通过所述多肽在活体内的表达根据本发明来使用，此举通常称为基因治疗。因此，举例而言，可以编码多肽的聚核苷酸 (DNA 或 RNA) 在活体外转导细胞，随后将这些经转导细胞提供给待用所述多肽治疗的患者。所述方法在此项技术中为熟知的。例如，可通过使用含有编码本发明多肽的 RNA 的逆转录病毒粒子经由此项技术中已知的程序来转导细胞。

[0218] 类似地，例如可通过此项技术中已知的程序在活体内完成细胞转导以用于在活体内表达多肽。如此项技术中已知，可将产生含有编码本发明多肽的 RNA 的逆转录病毒粒子的生产细胞投与患者以用于活体内转导和多肽在活体内的表达。

[0219] 对于所属领域技术人员而言，这些方法及其他通过所述方法投与本发明多肽的方法由于本发明的教示应变得显而易见。例如，用于转导细胞的表达媒介可为除逆转录病毒以外的媒介，例如与合适传递媒介组合之后可用于活体内转导细胞的腺病毒。

[0220] 在多肽制备成液体调配物并通过注射投与的情况下，溶液优选为含有 140 毫摩尔氯化钠和 10 毫摩尔钙的 pH 值为 7.4 的等渗盐溶液。举例而言，可将注射液以治疗有效量投与，考虑到投与途径、患者健康状况等因素，优选的每日剂量为约 1 μ g/kg 体重至约 5mg/kg 体重。

[0221] 本发明的多肽可与合适的医药载剂联合使用。所述组合物包含治疗有效量的蛋白质和医药学上可接受的载剂或赋形剂。所述载剂包括 (但不限于) 盐水、缓冲盐水、右旋糖、水、甘油、乙醇及其组合。调配物应适于投与模式。

[0222] 本发明的多肽也可通过多肽与一种或一种以上部分或共轭物化学连接而进行修饰以增强多肽的活性、细胞分布或细胞摄取。所述部分或共轭物包括脂质，诸如胆固醇、胆酸、硫醚、脂族链、磷脂及其衍生物；多元胺；聚乙二醇 (PEG)；棕榈基部分；和例如于美国专利第 5,514,758 号、第 5,565,552 号、第 5,567,810 号、第 5,574,142 号、第 5,585,481 号、第 5,587,371 号、第 5,597,696 号及第 5,958,773 号中所揭示的其他物质。

[0223] 也可修饰本发明的多肽以靶向于特定病症的特定细胞类型，包括 (但不限于) C 型肝炎感染情况中的肝细胞。如所属领域技术人员所能了解，已描述达成所述靶向目标的适当方法且这些方法包括 (但不限于) 脂质体靶向、受体调节的内吞作用和抗体 - 抗

原结合。在一实施例中，可通过将半乳糖部分加入多肽中而使脱唾液酸糖蛋白受体用于靶向于肝细胞。在另一实施例中，可将甘露糖部分共轭至多肽以便靶向于巨噬细胞和肝细胞上所发现的甘露糖受体。也可通过所属领域技术人员已知的方法修饰本发明的多肽以用于胞质传递，这些方法包括（但不限于）内体逃离机制或蛋白转导域（PTD）系统。例如，Vives E 等人（1997）*J.Biol.Chem.*272：16010-16017、Derossi 等人（1994）*J.Biol.Chem.*269：10444-10450、Elliott 等人（1997）*Cell* 88：223-233、Wadia, JS 等人（2004）*Nat.Med.*10：310-315 和 Kabouridis, PS.（2003）*Trends Biotech.*, 21：498-503 中揭示 PTD 系统。已知的内体逃离系统包括使用 pH 值响应性聚合载体，诸如聚（丙基丙烯酸）。已知的 PTD 系统的范围为诸如 HIV-1 TAT、HSV-1 VP22、果蝇触角足（*Drosophila Antennapedia*）或白喉毒素的天然肽至合成肽载体（Wadia 和 Dowdy, *Cur.Opin. Biotech.*13：52-56, 2002；Becker-Hapak 等人, *Methods* 24：247-256, 2001）。图 10 提供这些例示性 PTD 中的几种的详细描述。正如所属领域技术人员认识到的，可将多重传递与靶向方法相联合。例如，可通过在脂质体内封装使本发明的多肽靶向于肝细胞，其中所述脂质体共轭至半乳糖以用于靶向于脱唾液酸糖蛋白受体。

[0224] 本发明也提供医药包或试剂盒，其包含一种或一种以上填充有本发明医药组合物的成份中的一种或一种以上的容器。与所述容器相关的是呈管理医药品或生物产品的制备、使用或销售的政府机构所规定形式的公告，此公告反映该机构对于人投与的所述产品的制备、使用或销售的批准。另外，本发明的多肽可与其他治疗性化合物联合使用。

[0225] 当本发明的 OAS1 变异体用作医药品时，其可在适当媒介中给予哺乳动物。当本发明的多肽用作如上所述的医药品时，考虑到投与途径、患者健康状况等因素，其例如以每日约 10 μ g/kg 体重至约 4mg/kg 体重的治疗有效剂量来给药。给药量优选为足以达成预防或抑制病毒感染、优选为黄病毒感染、最优选为 HCV 感染，从而复制携带本文中所述所揭示的 OAS1R 等位基因的人中所发现的天然抗性。

[0226] 也预见到模拟在 NT_009775.13 的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处至少一个突变的有益作用的基于抑制剂的药物疗法，如以下详细讨论。如先前所讨论，开发所述抑制剂的一例示性原理为有益突变减弱或消除一种或一种以上 OAS1 特定同功异型物的表达、转译或功能的情况。本发明既不受限于有益突变的精确形式或效应也不受限于因此受影响的特定同功异型物的生物学活性。在所述情况下，所属领域技术人员将了解治疗学上抑制 OAS1 的所述特定同功异型物的效用。这些基于抑制剂的疗法可采用以下形式：化学个体、肽或蛋白质、反义寡核苷酸、干扰性小 RNA 和抗体。

[0227] 蛋白质、其片段或其他衍生物或其类似物或表达其的细胞可用作免疫原以产生其抗体。举例而言，这些抗体可为多克隆、单克隆、嵌合、单链、Fab 片段或 Fab 表达文库的产物。此项技术中已知的各种程序可用于产生多克隆抗体。

[0228] 可通过将多肽直接注射至动物或通过多肽投与动物（优选为非人类）来获得针对由本发明的 OAS1R 编码的多肽而产生的抗体。由此获得的抗体随后将结合多肽本身。以此方式，甚至仅编码多肽片段的序列也可用于产生结合整个天然多肽的抗体。此外，

对大量多肽具有特异性的一组所述抗体可用于鉴别和区分所述组织。 举例而言，图 9 演示对本发明的特定例示性多肽呈特异性的抗体的显影。

[0229] 提供由连续细胞系培养物产生的抗体的任何技术均可用于制备单克隆抗体。 实例包括杂交瘤技术 (Kohler 和 Milstein, 1975, *Nature*, 256 : 495-597)、Trioma 技术、人类 B 细胞杂交瘤技术 (Kozbor 等人, 1983, *Immunology Today* 4 : 72) 和用以产生人类单克隆抗体的 EBV 杂交瘤技术 (Coe 等人, 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R.Liss, Inc. 第 77-96 页)。

[0230] 用于产生单链抗体的所述技术 (美国专利第 4,946,778 号) 可适于产生抗本发明的免疫原性多肽产物的单链抗体。

[0231] 所述抗体可用于与本发明的蛋白序列定位和活性有关的方法中，例如用于为这些蛋白成像、测量其在适当生理学样本中的含量等等。

[0232] 本发明提供用于为细胞内的 OAS1 聚核苷酸成像的可检测性经标记的寡核苷酸。所述寡核苷酸适用于确定是否发生基因扩增，且适用于例如使用此项技术中已知的原位杂交技术分析细胞或组织中的表达水平。

[0233] 抑制 OAS1 功能的治疗剂

[0234] 本发明也涉及为干扰 OAS1 聚核苷酸正常功能而设计的反义寡核苷酸。 此项技术中已知有待广泛应用于反义技术的反义分子的任何修饰或变体均包括于本发明的范畴内。 所述修饰包括如美国专利第 5,536,821 号、第 5,541,306 号、第 5,550,111 号、第 5,563,253 号、第 5,571,799 号、第 5,587,361 号、第 5,625,050 号和第 5,958,773 号中所揭示的含磷键的制备。

[0235] 本发明的反义化合物可包括如 5,958,773 和本文中所揭示的专利中所揭示的经修饰碱基。 本发明的反义寡核苷酸也可通过将寡核苷酸与一种或一种以上部分或共轭物化学连接来加以修饰以增强反义寡核苷酸的活性、细胞分布或细胞摄取。 所述部分或共轭物包括脂质，诸如胆固醇、胆酸、硫醚、脂族链、磷脂；多元胺；聚乙二醇 (PEG)；棕榈基部分；和例如于美国专利第 5,514,758 号、第 5,565,552 号、第 5,567,810 号、第 5,574,142 号、第 5,585,481 号、第 5,587,371 号、第 5,597,696 号及第 5,958,773 号中所揭示的其他物质。

[0236] 嵌合反义寡核苷酸也属于本发明范畴内，且可使用例如美国专利第 5,013,830 号、第 5,149,797 号、第 5,403,711 号、第 5,491,133 号、第 5,565,350 号、第 5,652,355 号、第 5,700,922 号和第 5,958,773 号中所述的方法自本发明的寡核苷酸制备。

[0237] 优选的反义寡核苷酸可通过常规实验使用例如实例中所述的分析来选择。 尽管发明者不受特定作用机制约束，但是据信反义寡核苷酸通过使用 Watson-Crick 碱基配对来与目标聚核苷酸的互补区在细胞内结合而达成抑制效果。 当目标聚核苷酸为 RNA 时，实验证据表明杂交体的 RNA 组份是由 RNase H 裂解 (Giles 等人, *Nuc.Acids Res.*23 : 954-61, 1995 ; 美国专利第 6,001,653 号)。 一般来说，含有 10 个碱基对的杂交体具有足够长度以用作 RNase H 的底物。 然而，为达成结合的特异性，最好使用具有至少 17 个核苷酸的反义分子，因为此长度的序列在人类基因的中有可能是独特的。

[0238] 如以引用方式并入本文中的美国专利第 5,998,383 号中所揭示，选择寡核苷酸使得序列展现对与其互补模板形成寡核苷酸双螺旋来说为重要的适当能量相关特

征，并显示自身二聚化或自身互补的低可能性 (Anazodo 等人, *Biochem.Biophys.Res. Commun.*229 : 305-09, 1996)。将计算机程序 OLIGO (引物分析软件, 3.4 版) 用于确定反义序列熔融温度、自由能性质且用于评估潜在的自身二聚体形成及自身互补性质。所述程序允许确定这两项参数 (潜在的自身二聚体形成及自身互补性) 的定性评估且提供 “不可能”、“有一些可能” 或 “基本完全可能” 的指示。一般选择在这些参数中具有不可能评估的 OAS1 聚核苷酸节段。然而, 可使用在这些范畴的一者中具有 “一些可能” 的节段。在选择中使用这些参数的平衡。

[0239] 在反义技术中需要一定程度的常规实验以选择用于特定目标的最佳反义分子。为了有效起见, 优选将反义分子靶向于目标 RNA 分子的易接近或已暴露部分。尽管在某些情况下可利用关于目标 mRNA 分子结构的信息, 但目前使用反义进行抑制的方法是经由实验法达成。根据本发明, 此实验法可通过使用实例中所述的方法以反义寡核苷酸转染细胞来常规地进行。细胞中 mRNA 含量可通过逆转录 mRNA 和检定 cDNA 含量而在经处理细胞及对照细胞中常规地测量。如此项技术中已知, 可通过测量细胞生长或生存能力来常规地测定生物学效应。

[0240] 通过检定和分析 cDNA 含量来测量反义活性的特异性是验证反义结果的技术认可方法。已表明来自经处理细胞及对照细胞的 RNA 应经逆转录且应分析所得 cDNA 群体。(Branch, A.D., *T.I.B.S.*23 : 45-50, 1998.) 根据本发明, 以为了靶向于 OAS1 而设计的两条不同的反义寡核苷酸转染细胞培养物。在经处理细胞和对照细胞中测量对应于 OAS1 的 mRNA 含量。

[0241] 额外的抑制剂包括核酶、蛋白质或多肽、抗体或其片段以及小分子。这些 OAS1 抑制剂中的每一种的共同特征在于其减少 OAS1 的表达和 / 或生物学活性。除了本文中所揭示的例示性 OAS1 抑制剂之外, 可利用本文中特定揭示或另外不难获得且在熟练技工的专业技术之内的方法经由常规实验法获得替代性抑制剂。

[0242] **核酶**

[0243] OAS1 抑制剂可为核酶。核酶是特异性地裂解诸如 mRNA 的 RNA 底物、从而导致特异性抑制或干扰细胞基因表达的 RNA 分子。如本文中所用, 术语核酶包括含有用于特异性辨别的反义序列且含有 RNA 裂解酶活性的 RNA 分子。催化链于高于化学计量浓度时裂解目标 RNA 中的特异性位点。

[0244] 在本发明的上下文中可利用多种核酶, 例如包括锤头状核酶 (例如, 如由 Forster 和 Symons, *Cell* 48 : 211-20, 1987 ; Haseloff 和 Gerlach, *Nature* 328 : 596-600, 1988 ; Walbot 和 Bruening, *Nature* 334 : 196, 1988 ; Haseloff 和 Gerlach, *Nature* 334 : 585, 1988 所述) ; 发夹状核酶 (例如, 如 1993 年 10 月 19 日颁发的 Haseloff 等人的美国专利第 5,254,678 号和 1990 年 3 月 26 日公开的 Hempel 等人的欧洲专利公告第 0 360 257 号) ; 和基于四膜虫 (*Tetrahymena*) 核糖体 RNA 的核酶 (参见 Cech 等人, 美国专利第 4,987,071 号)。本发明的核酶通常是由 RNA 组成, 但也可由 DNA、核酸类似物 (例如硫代磷酸酯) 或其嵌合体 (例如 DNA/RNA/RNA) 构成。

[0245] 核酶可靶向于任何 RNA 转录物且可催化性地裂解所述转录物 (例如参见 Cech 等人的美国专利第 5,272,262 号 ; 美国专利第 5,144,019 号 ; 和美国专利第 5,168,053 号、第 5,180,818 号、第 5,116,742 号及第 5,093,246 号)。根据本发明的某些实施例, 可将任何

所述 OAS1 mRNA- 特异性核酶或编码所述核酶的核酸传递至宿主细胞中以实现 OAS1 基因表达的抑制。因此可通过编码与真核生物启动子（诸如真核病毒启动子）相连接的核酶的 DNA 将核酶等等传递至宿主细胞，使得在引入核内之后核酶将直接得以转录。

[0246] RNAi

[0247] 本发明也提供将带有部分或全部双股特征的 RNA 引入细胞或细胞外环境中。抑制对 OAS1 表达为特异性的，因为来自目标 OAS1 基因的一部分的核苷酸序列经选择以用于产生抑制性 RNA。此过程为 (1) 有效产生对基因表达的抑制和 (2) 对目标 OAS1 基因为特异性。所述程序可提供目标 OAS1 基因的部分或全部功能丧失。使用与其他目标基因相当的技术已显示至少 99% 的目标细胞中基因表达减少或损失。较低剂量的注射物质和双股 RNA 投与后的较长时间可导致较低百分率的细胞的抑制。细胞中基因表达的量化可显示在目标 mRNA 积累水平或目标蛋白转译水平上类似的抑制量。制备和使用 RNAi 的方法一般揭示于美国专利 6,506,559 中，此专利以引用的方式并入本文中。

[0248] RNA 可包含一条或一条以上聚合的核糖核苷酸链；其可包括对磷酸酯-糖骨架或核苷的修饰。双股结构可由单条自身互补 RNA 链或两条互补 RNA 链形成。可在细胞内或细胞外引发 RNA 双螺旋形成。可以允许每个细胞传递至少一个拷贝的量将 RNA 引入。更高剂量的双股物质可导致更高效的抑制。抑制具有序列特异性，因为对应于 RNA 双螺旋区的核苷酸序列是以基因抑制为目标。含有与一部分 OAS1 目标基因相同的核苷酸序列的 RNA 对于抑制是优选的。也已发现相对于目标序列具有插入、缺失和单点突变的 RNA 序列对抑制有效。因此，可通过此项技术中已知的比对算法和计算核苷酸序列之间的差异百分比来优化序列一致性。或者，RNA 双螺旋区可在功能上定义为能与一部分目标基因转录物杂交的核苷酸序列。

[0249] 可在活体内或活体外合成 RNA。细胞的内源性 RNA 聚合酶可调节活体内转录，或经克隆的 RNA 聚合酶可用于活体内或活体外转录。对于自活体内转基因或表达构筑体转录而言，调节区可用于转录 RNA 链。

[0250] 对于 RNAi，可将 RNA 直接引入细胞中（即细胞内地），或细胞外地引入空腔（胞间隙）中，引入有机体循环中，经口引入，或可通过将有机体浸泡在含有 RNA 的溶液中而引入。经口引入方法包括将 RNA 与有机体的食物直接混合，以及基因工程学方法，其中对用作食物的物质进行基因工程改造以表达 RNA，随后喂入待感染的有机体中。引入核酸的物理方法包括将 RNA 溶液直接注射至细胞中或细胞外注射至有机体中。

[0251] 所述方法的优势包括易于将双股 RNA 引入细胞内、可使用低浓度的 RNA、双股 RNA 的稳定性和抑制的有效性。

[0252] 抑制基因表达意指不存在（或明显降低）来自 OAS1 目标基因的蛋白质和 / 或 mRNA 产物的含量。特异性意指在不对细胞其他基因产生明显影响的情况下抑制目标基因的能力。抑制结果可通过检测细胞或有机体的外部性质或通过生物化学技术来确定，这些生物化学技术例如为 RNA 溶液杂交法、核酸酶保护法、Northern 杂交法、逆转录、以微阵列监控基因表达、抗体结合法、酶联免疫吸附分析 (ELISA)、Western 印迹法、放射性免疫分析 (RIA)、其他免疫分析和荧光活化细胞分析法 (FACS)。对于细胞系或整个有机体中由 RNA 调节的抑制而言，基因表达是通过使用其蛋白产物易于检定的报道基因或药物抗性基因而方便地检定。所述报道基因包括乙酰羟基酸合酶 (AHAS)、碱

性磷酸酶 (AP)、 β 半乳糖苷酶 (LacZ)、 β 葡萄糖醛酸苷酶 (GUS)、氯霉素乙酰转移酶 (CAT)、绿色荧光蛋白 (GFP)、辣根过氧化物酶 (HRP)、荧光素酶 (Luc)、胭脂氨酸 (nopaline) 合酶 (NOS)、章鱼氨酸 (octopine) 合酶 (OCS) 及其衍生物。可利用多种可选择的赋予对胺苄青霉素 (ampicillin)、博莱霉素 (bleomycin)、氯霉素 (chloramphenicol)、庆大霉素 (gentamycin)、潮霉素 (hygromycin)、卡那霉素 (kanamycin)、林可霉素 (lincomycin)、甲胺喋呤 (methotrexate)、草铵膦 (phosphinothricin)、嘌呤霉素 (puromycin) 和四环素 (tetracyclin) 的抗性的标记物。

[0253] 视分析而定, 基因表达量的量化允许确定与未根据本发明处理的细胞相比超过 10%、33%、50%、90%、95% 或 99% 的抑制程度。较低剂量的注射物质和双股 RNA 投与后的较长时间可引起较低百分率的细胞的抑制 (例如目标细胞的至少 10%、20%、50%、75%、90% 或 95%)。细胞中 OAS1 基因表达的量化可显示在 OAS1 目标 mRNA 积累水平或 OAS1 目标蛋白转译水平上类似的抑制量。举例而言, 抑制效率可通过分析细胞中基因产物的量来确定: 可以具有在用于抑制性双股 RNA 的区以外的核苷酸序列的杂交探针来检测 mRNA, 或可以针对所述区的多肽序列而产生的抗体来检测经转译多肽。

[0254] RNA 可包含一条或一条以上聚合的核糖核苷酸链。其可包括对磷酸酯-糖骨架或核苷的修饰。举例而言, 可修饰天然 RNA 的磷酸二酯键以包括氮或硫杂原子中的至少一个。RNA 结构中的修饰可适于允许特异性基因抑制同时避免在某些有机体中由双股 RNA 产生的普遍惊恐反应。同样地, 可修饰碱基以阻断腺苷脱氨酶的活性。RNA 可由酶促法产生或通过部分/全部有机合成产生, 可通过活体外酶促合成或有机合成将任何经修饰的核糖核苷酸引入。

[0255] 双股结构可由单条自身互补 RNA 链或两条互补 RNA 链形成。可在细胞内或细胞外引发 RNA 双螺旋形成。可以允许每个细胞传递至少一个拷贝的量将 RNA 引入。更高剂量 (例如每个细胞至少 5、10、100、500 或 1000 个拷贝) 的双股物质可导致更高效的抑制; 较低剂量也可适用于特殊应用。抑制具有序列特异性, 因为对应于 RNA 双螺旋区的核苷酸序列是以基因抑制为目标。

[0256] 含有与一部分 OAS1 目标基因相同的核苷酸序列的 RNA 对抑制是优选的。相对于目标序列具有插入、缺失和单点突变的 RNA 序列可对抑制有效。因此, 可通过序列对比和此项技术中已知的比对算法 (参见 Gribskov 和 Devereux, *Sequence Analysis Primer*, Stockton Press, 1991 及其中所引用的文献) 及通过例如如 BESTFIT 软件程序中所实施的 Smith-Waterman 演算法使用默认参数 (例如 University of Wisconsin Genetic Computing Group) 计算核苷酸序列之间的差异百分比来优化序列一致性。在抑制性 RNA 与部分 OAS1 目标基因之间超过 90% 序列一致性或甚至 100% 序列一致性是优选的。或者, RNA 双螺旋区可在功能上定义为能与一部分 OAS1 目标基因转录物杂交 (例如, 400 mM NaCl, pH 值为 6.4 的 40 mM PIPES, 1 mM EDTA, 50°C 或 70°C 下杂交 12-16 小时; 随后洗涤) 的核苷酸序列。相同核苷酸序列的长度可为至少 25、50、100、200、300 或 400 个碱基。

[0257] 实施本发明并不需要 RNA 与 OAS1 目标基因之间 100% 的序列一致性。因此, 所述方法具有能接受可能由于基因突变、菌株多形态或进化趋异而预期的序列变异的优势。

[0258] 可在活体内或活体外合成 OAS1 RNA。细胞的内源性 RNA 聚合酶可调节活体内转录，或经克隆的 RNA 聚合酶可用于活体内或活体外转录。对于自活体内转基因或表达构筑体转录而言，调节区（例如启动子、增强子、沉默子、剪接供体及受体、多聚腺苷酸化作用）可用于转录 RNA 链。有机体、组织或细胞类型中的特异性转录；环境条件刺激（例如感染、应激、温度、化学诱导剂）；和 / 或在发育阶段或时期的遗传工程转录均可以抑制为目标。RNA 链可经或未经多聚腺苷酸化；RNA 链能或不能通过细胞转译装置转译为多肽。可通过手动或自动反应而化学合成或酶促合成 RNA。可通过细胞 RNA 聚合酶或噬菌体 RNA 聚合酶（例如 T3、T7、SP6）合成 RNA。表达构筑体的使用和产生在（参见 WO 第 97/32016 号；美国专利第 5,593,874 号、第 5,698,425 号、第 5,712,135 号、第 5,789,214 号和第 5,804,693 号；及其中所引用的文献）中已知。若化学合成或通过活体外酶促法合成，则 RNA 可在引入细胞之前进行纯化。例如，RNA 可通过以溶剂或树脂萃取、沉淀、电泳、色谱分析或其组合而自混合物中纯化。或者，可使用未经纯化或最小限度纯化的 RNA 以避免因样本处理而引起的损失。可将 RNA 干燥储存或溶解于水溶液中。所述溶液可含有缓冲液或盐以促进双螺旋链的退火和 / 或稳定化。

[0259] 可将 RNA 直接引入细胞中（即细胞内地）；或细胞外地引入空腔（胞间隙）中，引入有机体的循环中，经口引入，或可通过将有机体浸泡在含有 RNA 的溶液中而引入。经口引入方法包括将 RNA 与有机体食物直接混合，以及遗传工程学方法，其中对用作食物的物质进行遗传工程操作以表达 RNA，随后喂入待感染的有机体中。举例而言，可将 RNA 喷射在植物上或可以基因工程学方法改造植物使其表达足够量的 RNA 以杀灭一些或全部已知会感染植物的病原体。也可使用引入核酸的物理方法，例如直接注射至细胞内或细胞外注射至有机体内。RNA 可引入的位点为血管或血管外循环、血液或淋巴系统和脑脊髓液。表达来自重组构筑体的 RNA 的转基因有机体可通过将所述构筑体引入来源于适当有机体的接合子、胚胎干细胞或另一多潜能细胞而产生。

[0260] 引入核酸的物理方法包括注射含有 RNA 的溶液、通过经 RNA 覆盖的粒子轰击、将细胞或有机体浸泡在 RNA 溶液中或在 RNA 存在下电穿孔细胞膜。封装入病毒粒子中的病毒构筑体将实现表达构筑体至细胞内的有效引入和由所述表达构筑体所编码的 RNA 的转录。可使用此项技术中已知的将核酸引入细胞内的其他方法，诸如由脂质调节的载体输送、由化学品（诸如磷酸钙）调节的输送等等。因此可将 RNA 连同执行下列活动中的一种或一种以上的组份一起引入：增强 RNA 被细胞摄取、促进双螺旋链退火、稳定退火链或另外增加目标基因的抑制。

[0261] 本发明可单独使用或作为试剂盒的组份使用，所述试剂盒含有为进行活体外或活体内将 RNA 引入测试样本或受检者所必需的试剂中的至少一种。优选的组份为双股 RNA 和促进双股 RNA 引入的媒介。所述试剂盒也可包括允许试剂盒使用者实施本发明的说明书。

[0262] 建构适当的注射混合物使得动物接收平均 0.5×10^6 至 1.0×10^6 个 RNA 分子。为比较有义、反义和双股 RNA 活性，以相等质量 RNA（即双股 RNA 摩尔浓度为单股摩尔浓度的一半）来比较注射。基于注射物质中的 RNA 浓度（由溴化乙锭染色评估）和注射体积（由注射部位的可见位移评估）粗略大致地给出每个成人所注射的分子数。个别动物之间注射体积上的数倍变化性是可能的。

[0263] 蛋白质和多肽

[0264] 除本文中所揭示的反义分子和核酶以外，本发明的 OAS1 抑制剂也包括有效减少 OAS1 基因表达或有效降低 OAS1 生物学活性中的一种或一种以上的蛋白质或多肽，所述生物学活性包括（但不限于）酶活性；与单股 RNA 构型相互作用；和与诸如 C 型肝炎病毒 NS5A 或其片段的其他蛋白结合。熟练技工可通过此项技术中易于利用的多种方法经由常规实验法快速鉴别所述 OAS1 抑制剂。本发明不受限于以下例示性方法。

[0265] 熟练技工可利用描述检测和分析蛋白与蛋白之间相互作用的文献。于 Phizicky 等人，*Microbiological Reviews* 59 : 94-123, 1995 中评述，此文献以引用的方式并入本文中。所述方法包括（但不限于）：物理方法，例如蛋白亲和色谱法、亲和印迹法、免疫沉淀法和交联法；以及基于文库的方法，例如蛋白探测法、噬菌体展示法和双杂交体筛选法。其他可用于鉴别蛋白与蛋白之间相互作用的方法包括基因方法，诸如使用基因外抑制基因、合成致死效应和未连接型非互补。例示性方法在以下进一步详细描述。

[0266] 可经由依赖于 OAS1 蛋白与潜在抑制剂蛋白组或文库之间的直接相互作用的生物学筛选分析来鉴别本发明的 OAS1 抑制剂。包括各种“n 杂交体技术”的生物学筛选方法例如描述于 Vidal 等人，*Nucl.Acids Res.*27(4) : 919-29, 1999 ; Frederickson, R.M., *Curr.Opin.Biotechnol.*9(1) : 90-96, 1998 ; Brachmann 等人，*Curr.Opin.Biotechnol.*8(5) : 561-68, 1997 ; 和 White, M.A., *Proc.Natl.Acad.Sci.U.S.A.*93 : 10001-03, 1996 中，各文献均以引用的方式并入本文中。

[0267] 双杂交体筛选方法可用于寻找具有抑制性质的 OAS1 结合蛋白的新颖或现有目标 cDNA 文库。双杂交体系统是借助于增加报道基因的转录来检测蛋白与蛋白之间相互作用的基因方法。所述系统依赖于位点特异性转录活化子具有 DNA 结合域和转录活化域的事实。DNA 结合域使活化域靶向于待表达的特异基因。由于转录活化子的模块式性质，因此 DNA 结合域可自转录活化域共价分开而不会损失任一域的活性。此外，这两个域可因两种与转录机构无关的蛋白之间的蛋白与蛋白接触而毗邻。因此，建构双杂交体以创建功能系统。第一杂交体（即诱饵）是由融合至相关蛋白的转录活化子 DNA 结合域组成。第二杂交体（目标）是由转录活化域与蛋白或多肽文库融合而产生。诱饵蛋白与目标文库的成员之间的相互作用导致 DNA 结合域与转录活化域毗邻和随之发生的报道基因表达的上调。

[0268] 熟练技工可利用多种基于双杂交体的系统，最常用的是酵母 Gal4 或大肠杆菌 LexADNA 结合域 (BD) 和酵母 Gal4 或单纯疱疹病毒 VP16 转录活化域。Chien 等人，*Proc. Natl.Acad.Sci.U.S.A.*88 : 9578-82, 1991 ; Dalton 等人，*Cell* 68 : 597-612, 1992 ; Durfee 等人，*Genes Dev.*7 : 555-69, 1993 ; Vojtek 等人，*Cell* 74 : 205-14, 1993 ; 和 Zervos 等人，*Cell*72 : 223-32, 1993。常用报道基因包括大肠杆菌 lacZ 基因以及可选择的酵母基因，诸如 HIS3 和 LEU2。Fields 等人，*Nature (London)* 340 : 245-46, 1989 ; Durfee, T.K., 同上 ; 和 Zervos, A.S., 同上。此项技术中有多种活化域文库易于利用，使得可通过常规实验法进行相互作用的蛋白的筛选。

[0269] 用于鉴别 OAS 1 相互作用蛋白的适当诱饵蛋白可基于本文中作为 SEQ ID NO : 19 呈现的 OAS1 DNA 序列而设计。所述诱饵蛋白包括全长 OAS1 蛋白或其片段。

[0270] 用于制备 OAS1 诱饵构筑体和目标文库的质粒载体（诸如 pBTM116 和 pAS2-1）

为技工易于利用的且可自诸如 Clontech (Palo Alto, CA)、Invitrogen (Carlsbad, CA) 和 Stratagene (La Jolla, CA) 的商业来源获得。这些质粒载体允许 cDNA 与分别作为 LexA 或 Gal4BD 的 DNA 结合域框内融合。

[0271] 或者可通过此项技术中可利用的检测蛋白与蛋白相互作用的物理或生化方法中的一种来鉴别本发明的 OAS1 抑制剂。

[0272] 经由蛋白亲和色谱方法, 作为潜在 OAS1 抑制剂的待测试的先导化合物可借助其在共价或非共价偶合至固体基质(诸如琼脂糖珠粒)时对 OAS1 的特异性滞留而加以鉴别。蛋白亲和柱的制备在例如 Beeckmans 等人, Eur.J.Biochem.117: 527-35, 1981 和 Formosa 等人, Methods Enzymol.208: 24-45, 1991 中描述。含有细胞蛋白的完全补体的细胞溶解产物可穿过 OAS1 亲和柱。对 OAS1 具有高亲和性的蛋白在低盐条件下将特异性滞留, 而大多数细胞蛋白将穿过柱。可在高盐条件下以离液溶剂或以十二烷基硫酸钠(SDS)将所述高亲和性蛋白自固定化 OAS1 洗提。在某些实施例中, 最好在制备溶解产物之前对细胞进行放射性同位素标记以有助于鉴别 OAS1 特异性结合蛋白。对哺乳动物细胞进行放射性同位素标记的方法在此项技术中为熟知的且例如在 Sopta 等人, J.Biol.Chem.260: 10353-60, 1985 中提供。

[0273] 可将用于亲和色谱法的适当 OAS1 蛋白融合至蛋白或多肽以允许在适当亲和树脂上快速纯化。例如, 可将 OAS1 cDNA 融合至有利于融合蛋白吸附至谷胱甘肽-琼脂糖柱上的谷胱甘肽 S-转移酶(GST)的编码区。Smith 等人, Gene 67: 31-40, 1988。或者, 融合蛋白可包括: 蛋白 A, 其可在具有免疫球蛋白 G 的柱上纯化; 含有寡组氨酸的肽, 其可在具有 Ni²⁺ 的柱上纯化; 麦芽糖结合蛋白, 其可在含有直链淀粉的树脂上纯化; 和二氢叶酸还原酶, 其可在甲胺喋呤柱上纯化。一种适于制备本文中所提供的 OAS1 融合蛋白的例示性标记为流感病毒红血球凝聚素(HA)的抗原决定基, 可易于获得针对其的单克隆抗体且可自所述抗体制备亲和柱。

[0274] 在经受 SDS 聚丙烯酰胺凝胶电泳法(SDS-PAGE)后可鉴别特异性滞留在 OAS1 亲和柱上的蛋白。因此, 在制备细胞溶解产物且穿过 OAS1 亲和柱之前对细胞进行放射性同位素标记时, 可通过放射自显影法检测对 OAS1 具有高亲和性的蛋白。可通过熟练技工易于使用的蛋白测序技术确定 OAS1 特异性结合蛋白的一致性, 诸如 Mathews, C.K. 等人, Biochemistry, The Benjamin/Cummings Publishing Company, Inc., 1990, 第 166-70 页。

[0275] 小分子

[0276] 本发明也提供可经由常规应用高通量筛检(HTS)方法而易于鉴别的小分子 OAS1 抑制剂。由 Persidis, A., Nature Biotechnology 16: 488-89, 1998 评述。HTS 方法一般意指允许快速分析先导化合物(诸如小分子)的治疗潜力的那些技术。HTS 方法采用测试物质的自动机械处理、正信号检测和数据解译。举例而言, 所述方法包括使用可溶性分子以及基于细胞的系统(诸如以上详述的双杂交体系统)的自动机械筛检技术。

[0277] 可利用多种基于细胞系的 HTS 方法, 其优点为操作简便和相对于溶液形式在细胞背景中所发生的相互作用的临床关联性。先导化合物可经由并入放射性或经由依赖于所读取的吸光度、荧光性或发光性的光学分析来鉴别。例如参见 Gonzalez 等人, Curr.Opin.Biotechnol.9(6): 624-31, 1998, 其以引用的方式并入本文中。

[0278] 举例而言，HTS 方法可用于筛检阻断 OAS1 的一种生物学活性的先导化合物。通过此方法，OAS1 蛋白可自表达所述蛋白的细胞中免疫沉淀且施加于适于自动机械筛检的分析板上的孔中。随后个别测试化合物可与免疫沉淀的蛋白接触且分析每个测试化合物对 OAS1 的效应。

[0279] 评定 OAS1 抑制剂功效的方法

[0280] 通过一种本文中所述的方法或经由此项技术中另外可利用的技术鉴别的先导分子或化合物，不论其为反义分子或核酶、蛋白和 / 或肽、抗体和 / 或抗体片段或小分子，均可在多种试管内、活体外和活体内动物模型分析系统中进一步表征其抑制 OAS1 基因表达的能力或生物学活性。如下提供的实例中进一步详细讨论，本发明的 OAS1 抑制剂可有效降低 OAS1 表达水平。因此，本发明进一步揭示允许熟练技工评定候选抑制剂的作用的方法。

[0281] 候选 OAS1 抑制剂可通过投与表达内源性 OAS1 或通过以重组 OAS1 质粒构筑体转染哺乳动物细胞来表达 OAS1 的细胞而进行测试。

[0282] 有效 OAS1 抑制性分子将有效减少 OAS1 的酶活性或 OAS1 对 IFN 诱导的反应能力。测量 OAS1 酶活性和 IFN 诱导的方法在此项技术中为已知的，例如如 Eskildsen 等人，Nuc.Acids Res.31 : 3166-3173, 2003 和 Justesen 等人，Nuc.Acids Res.8 : 3073-3085, 1980 中所述，所述文献以引用的方式并入本文中。可通过与已知在投与哺乳动物细胞时对 OAS1 表达无实质影响的对照“反义”分子相比来评定给定候选反义分子的有效性。

[0283] 进一步在活体外表征通过如上所讨论的方法中的一种或一种以上有效减少 OAS1 基因表达的 OAS1 抑制剂在一种如本文中所述讨论的易于建立的细胞培养物或原始细胞培养模型系统中的功效，参考通过诸如登革热病毒的黄病毒感染而攻毒的 Vero 细胞的用途。

[0284] 医药组合物

[0285] 可通过此项技术中已知的任何用于核糖核酸核苷酸或脱氧核糖核酸核苷酸的方法来合成本发明的反义寡核苷酸和核酶。例如，可使用固相合成诸如在 Applied Biosystems380B DNA 合成仪中制备寡核苷酸。如此项技术中所知来测定寡核苷酸的最终纯度。

[0286] 使用本发明方法所鉴别的反义寡核苷酸调节肿瘤细胞增殖。因此，提供用于干扰病毒感染、优选为黄病毒感染、最优选为 HCV 感染的医药组合物和方法，其包含将组织或细胞与使用本发明方法所鉴别的反义寡核苷酸中的一种或一种以上接触。

[0287] 本发明提供与 OAS1 mRNA 基因序列互补的反义寡核苷酸和核酶的医药组合物作为活性成份用于治疗应用。这些组合物也可用于本发明的方法中。当需要时，所述化合物具有核酸酶抗性。一般来说，用于抑制哺乳动物中病毒感染的医药组合物包括有效量的至少一种如上所述为实施本发明所需要的反义寡核苷酸，或显示具有相同作用的其片段，和医药学及生理学上可接受的载剂或稀释剂。

[0288] 组合物可经口、皮下或非经肠（包括静脉内、动脉内、肌肉内、腹膜内和鼻内投与）以及视需要的鞘内和输注技术投与。医药学上可接受的载剂、稀释剂、佐剂和媒剂以及植入式载体一般意指不与本发明的活性成份反应的惰性、无毒固体或液体填料、稀释剂或封装物质。组合物中也可包括阳离子脂质以有利于寡核苷酸摄取。化合物的植入也适用。一般来说，医药组合物是无菌的。

[0289] 生物活性（可表达的）意指当寡核苷酸直接传递至细胞时其在细胞内具有生物学活性和 / 或当在如下描述的载体中传递至细胞时其由适当启动子来表达并呈活性。核酸酶抗性是由此项技术中已知的不实质上干扰如本文中所述的生物学活性的任何方法提供。

[0290] “接触细胞”意指将反义寡核苷酸直接或通过病毒或非病毒载体暴露或传递至细胞且其中在传递后反义寡核苷酸具有生物活性的方法。

[0291] 可直接或以病毒或非病毒载体传递本发明的核苷酸序列。当直接传递时所述序列一般赋予核酸酶抗性。或者，可将序列并入表达框或构筑体中使得所述序列在细胞内得以表达。一般来说，所述构筑体含有适当的调节序列或启动子以使序列将在目标细胞中得以表达。

[0292] 一旦寡核苷酸序列准备开始传递，则如此项技术中已知可将其引入细胞中。转染、电穿孔、融合、脂质体、胶状聚合粒子和病毒载体以及此项技术中已知的其他方法可用于将寡核苷酸序列传递至细胞。所选方法将至少取决于待处理的细胞和所述细胞的位置且为所属领域技术人员所知。定位可由表面具有用于指导脂质体的特异性标记物的脂质体、通过直接注入含有目标细胞的组织中、通过具有与目标细胞在空间邻接性上相关的储槽、由特异性受体调节的摄取、病毒载体等等来达成。

[0293] 本发明提供包含可经由操作连接至本发明的寡核苷酸序列的表达控制序列的载体。本发明进一步提供选自适当的真核和原核生物细胞的宿主细胞，其必要时以这些载体转化。

[0294] 载体为已知的或可由所属领域技术人员建构且应含有为达成所要序列转录所必需的所有表达元件。载体内也可含有其他有益特征，诸如回收不同形式的寡核苷酸的机制。噬菌粒为所述有益载体的特例，因为其可用作质粒或噬菌体载体。其他载体的实例包括诸如噬菌体、杆状病毒和逆转录病毒的病毒、DNA 病毒、脂质体及其他重组载体。所述载体也可含有用于原核或真核生物宿主系统中的元件。一般技术者将了解何种宿主系统与特定载体相容。

[0295] 可通过多种此项技术中已知的方法中的任一种将载体引入细胞或组织中。可发现所述方法概述于 Sambrook 等人，*Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory, New York, 1989, 1992; Ausubel 等人，*Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD, 1989; Chang 等人，*Somatic Gene Therapy*, CRC Press, Ann Arbor, Mich., 1995; Vega 等人，*Gene Targeting*, CRC Press, Ann Arbor, Mich., 1995; Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Butterworths, Boston, Mass., 1988; 和 Gilboa 等人，*BioTechniques* 4: 504-12, 1986 中，且例如包括稳定或瞬时转染、脂质转染 (lipofection)、电穿孔和以重组病毒载体感染。

[0296] 此项技术中已知的重组方法也可用于达成目标核酸的反义抑制。例如，含有反义核酸的载体可用于表达反义信息以减少目标核酸的表达且因此降低其活性。

[0297] 本发明也提供一种评估化合物是否抑制 OAS1 基因转录或转译且因此调节（即削弱）细胞活化 RNaseL 的能力的方法，其包含以包含编码 OAS1 的核酸序列、核酸转录或转译的必要元件的表达载体来转染细胞；投与测试化合物；和将 OAS1 表达水平与不存

在所述测试化合物的对照所获得的水平相比较。

[0298] 优选实施例

[0299] 利用上述方法和此项技术中已知的其他方法，本发明涵盖一种筛检方法，其包含在扩增条件下以 PCR 引物对处理分离自人类的基因组 DNA 样本，以用于扩增含有寡核苷酸合成酶 (OAS1, SEQ ID NO : 19) 的核苷酸 (nt) 位置 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 的任一者的人类基因组 DNA 区。扩增条件包括存在对 DNA 合成有效量的 PCR 缓冲液和热循环温度。分析如此产生的 PCR 产物在相应核苷酸位置处点突变的存在。在一实施例中，PCR 产物含有包含自 5' 至 3' 方向书写的约 358 个碱基对 (bp) 且包括位置 2135728 (突变 1)、2135749 (突变 2)、2135978 (突变 3)、2144072 (突变 4)、2144088 (突变 5)、2144116 (突变 6) 或 2144321 (突变 7) 和在每一侧与所述位置侧接的约 175 个碱基的连续核苷酸序列。在另一实施例中，表 1 和表 2 中如上所述的扩增子为 PCR 产物和相应引物的例证。

[0300] 在一优选实施例中，通过在杂交条件下以对相应突变呈特异性的寡核苷酸探针处理扩增产物且检测任何杂交产物的形成来检定 PCR 产物的相应突变。优选的寡核苷酸探针包含以下表 3 中所指明的核苷酸序列。寡核苷酸与目标核酸的杂交描述于美国专利第 4,530,901 号中。

[0301] 表 3

[0302]

突变 SEQ ID NO	探针序列
SEQ ID NO : 1	GTAGATTTGCCYGAACAGGTCAGT (SEQ ID NO : 12)
SEQ ID NO : 2	CAGTTGACTGGCRGCTATAAACCTA (SEQ ID NO : 13)
SEQ ID NO : 3	CAGAGGAGGGGTRGGGGGAGGAGA (SEQ ID NO : 14)
SEQ ID NO : 4	TCTCACCTTTTCARGCTGAAAGCAAC (SEQ ID NO : 15)
SEQ ID NO : 5	GAAAGCAACAGTRCAGACGATGAGA (SEQ ID NO : 16)
SEQ ID NO : 6	ACGATCCCAGGASGTATCAGAAATAT (SEQ ID NO : 17)
SEQ ID NO : 7	TTGATCCAGAGARGACAAAGCTCCTC (SEQ ID NO : 18)

[0303] 其中 R = A/G、S = C/G 且 Y = C/T。

[0304] 使如此形成的 PCR 混合物经受多个 PCR 热循环以产生 OAS1 和 OAS1R 基因扩增产物。随后在杂交条件下以对每一突变均呈特异性的寡核苷酸探针处理扩增产物。随后检测任何杂交产物。

[0305] 以下实例旨在阐明本发明而并非理解为以任何方式限制本说明书和权利要求。

[0306] 实例

[0307] 实例 1

[0308] 基因组 DNA 的制备和初步筛检

[0309] 此实例涉及自两类特定患者群体筛检 DNA，但同样应用于其他经证实反复暴露于 HCV 的患者组，其中所述暴露并未导致感染。所述实例也涉及筛检已暴露于如上所述的其他黄病毒的患者，其中所述暴露并未导致感染。

[0310] 此处，研究两类群体：(1) 血友病群体，选择标准为中度至严重血友病且在 1987 年 1 月之前接收浓缩凝血因子；和 (2) 静脉药物使用者群体，其注射史超过 10 年且证明有其他危险行为，诸如共用注射针。所述研究包括经暴露但 HCV 阴性的患者和经暴露且 HCV 阳性的患者。

[0311] 自来自静脉药物使用者、血友病患者及其他具有 C 型肝炎感染或其他黄病毒感染风险的群体的白细胞中萃取高分子量 DNA。为进行基因组 DNA 的初始筛检，在得到如上所述的组中的患者同意之后采集血液并用 0.14M 柠檬酸、0.2M 柠檬酸三钠和 0.22M 右旋糖的混合物防止凝血。将未凝固血液在室温下以 $800\times g$ 离心 15 分钟并丢弃富含血小板的血浆上清液。将粒状红细胞、单核及多核细胞再悬浮并以与初始血液体积相同体积的 pH 值为 7.4 的冰冷 0.14 M 磷酸盐缓冲盐水 (PBS) 稀释。通过在低内毒素 Ficoll-Hypaque (Sigma Chem. Corp. St. Louis, Mo.) 上于 $18^{\circ}C$ 下以 $400\times g$ 离心 10 分钟而自经稀释的细胞悬浮液回收周边血液白细胞。随后将粒状白细胞再悬浮并用于高分子量 DNA 的来源。

[0312] 使用所属领域技术人员熟知和由 Maniatis 等人，Molecular Cloning: A Laboratory Manual, 第 2 版, Cold Spring Harbor Laboratory, Sections 9.16-9.23, (1989) 及美国专利号第 4,683,195 号所描述的方法自经分离的白细胞纯化高分子量 DNA。

[0313] 随后检查每一 DNA 样本在对应于寡腺苷酸合成酶 1 基因 (OAS1) 的 Genbank 登记号第 NT_009775.13 号的核苷酸位置的 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 位置处的任一核苷酸的点突变。

[0314] 实例 2

[0315] 与 HCV 感染抗性相关的 OAS1 基因中的突变

[0316] 使用实例 1 中所述的方法，研究一群非相关血友病患者和静脉药物使用者，并确定是否存在如 SEQ ID NO: 1-7 和 SEQ ID NO: 57-64 中所揭示的 OAS1 中的突变。

[0317] 在对高加索群体中 20 个病例和 42 个对照的研究中，在 C 型肝炎感染抗性的情况下发现了这些突变。HCV 感染抗性与 OAS1 中点突变的存在之间存在统计学上的显著相关性。

[0318] 实例 3

[0319] cDNA 的制备和测序

[0320] 自来自具有 C 型肝炎抗性表现型的患者的经培养的淋巴母细胞或成纤维细胞纯化总细胞 RNA。如 Chomczynski 等人，Anal. Biochem., 162: 156-159 (1987) 所述执行纯化程序。简言之，如实例 1 中所述制备细胞。随后在 10 毫升 (ml) 含有 4.0M 硫氰酸胍、pH 值为 7.5 的 0.1M Tris-HCl 和 0.1M β -巯基乙醇的变性溶液中将所述细胞均质化以形成细胞溶解产物。随后将十二烷基肌氨酸钠混合至细胞溶解产物中直至最终浓度为 0.5%，其后将混合物在室温下以 $5000\times g$ 离心 10 分钟。将所得含有总 RNA 的上清液在

pH 值为 7.5 的 5.7M 氯化铯和 0.01M EDTA 缓冲垫上分层并通过离心颗粒化。将所得 RNA 颗粒溶解于含有 0.1% 十二烷基硫酸钠 (SDS) 的 pH 值为 7.6 的 10mM Tris-HCl 与 1mM EDTA (TE) 的溶液中。在苯酚氯仿萃取和乙醇沉淀之后, 通过测量 260nm 下的光密度来估计纯化的细胞 RNA 总浓度。

[0321] 以上制备的总 RNA 是用作 cDNA 合成的模板, cDNA 合成中使用逆转录酶用于第一链合成并以所设计的寡核苷酸引物进行 PCR 以致在命名为 5' 和 3' 片段的两个重叠片段中扩增 cDNA。按照厂商说明书在 Applied Biosystems 381A DNA 合成仪上合成用于实施本发明的寡核苷酸。使用此项技术中已知的方法进行 PCR。PCR 扩增方法详细描述于美国专利第 4,683,192 号、第 4,683,202 号、第 4,800,159 号和第 4,965,188 号中且至少描述于包括 PCR Technology: Principles and Applications for DNA Amplification, H.Erlich 编, Stockton Press, New York (1989) 和 PCR Protocols: A Guide to Methods and Applications, Innis 等人编, Academic Press, San Diego, Calif.(1990) 及本文表 1 所述的引物的若干本教科书中。

[0322] 分析来自感染 HCV 或未感染 HCV 患者的经 PCR 扩增的 DNA 所直接测定的序列。可检测尽管反复暴露于病毒下但仍呈 HCV 血清阴性的患者中 OAS 基因编码区上游中突变的存在。

[0323] 实例 4

[0324] 经 PCR 扩增的含有点突变的基因组 DNA 的制备和通过等位基因特异性寡核苷酸杂交的检测

[0325] 寡腺苷酸合成酶 (OAS1) 基因中在核苷酸位置 2135728、2135749、2135978、2144072、2144088、2144116、2144321、2131025、2133961、2139587、2144294、2144985、2156523、2156595 或 2156638 中的一个位置的点突变可通过以下方法确定: 其中经由与所述区杂交的寡核苷酸探针杂交来检测经 PCR 扩增的含有突变的基因组 DNA。为扩增含有点突变的区以用于与寡核苷酸特异性探针杂交, 例如用 180 ng 表 1 中所示的每一个引物基本上如实例 3 中所述执行 PCR 扩增。

[0326] PCR 扩增之后, 将 2 μ l 经扩增的寡腺苷酸合成酶 DNA 产物转渍于硝化纤维素的单独薄片上。在干燥经转渍的扩增 DNA 之后, 以 0.5N NaOH 处理硝化纤维素 2 分钟, 以 pH 值为 7.5 的 1M Tris-HCl 处理 2 分钟, 随后以含有 1.5M NaCl 的 pH 值为 7.5 的 0.5M Tris-HCl 处理 2 分钟以使 DNA 变性并随后中和。将所得滤器在 80°C 下于真空中烘焙 1 小时, 并以由 6 \times SSC (1 \times = 0.15M NaCl、0.15M 柠檬酸钠)、5 \times Denhardt 溶液 (5 \times = 0.1% 聚乙烯吡咯烷酮、0.1% 菲可 (ficoll) 和 0.1% 牛血清白蛋白)、5mM pH 值为 7.0 的磷酸钠缓冲液、0.5mg/ml 鲑鱼睾丸 DNA 和 1% SDS 组成的预杂交溶液在 42°C 下预杂交至少 20 分钟。

[0327] 预杂交步骤之后, 将硝化纤维素滤器分别暴露于 32 P 标记的于预杂交缓冲液中稀释的寡核苷酸探针。通过将 2.5 μ l 激酶缓冲液的 10 \times 浓缩液 (10 \times = pH 值为 7.6 的 0.5M Tris[羟甲基]氨基甲烷盐酸盐 (Tris-HCl)、0.1M MgCl₂、50mM 二硫苏糖醇 (DTT)、1mM 亚精胺-HCl 和 1mM 乙二胺四乙酸 (EDTA))、1.1 μ l 60 μ g/ μ l 的所选寡核苷酸、18.4 μ l 水、浓度为 150mCi/ μ l 的 2 μ l 6000Ci/mM 的 γ 32 P ATP 和 1 μ l 10U/ μ l 的聚核苷酸激酶混合来进行 32 P 对探针的标记。将标记混合物在 37°C 下保持 20 分钟, 随后在 68°C

下保持 2 分钟。随后将经保持混合物施用于 Sephadex G50 (Pharmacia, Inc., Piscataway, NJ) 旋转柱以移除未并入的经 ^{32}P 标记的 ATP。

[0328] 以上表 3 中显示用于与含有突变的区杂交的寡核苷酸探针。 有下划线的核苷酸对应于突变核苷酸。 在用于检测野生型（正常）的探针中，以野生型核苷酸置换有下划线的核苷酸。

[0329] 将 10×10^6 cpm 的正常和突变型经标记探针分别与每一滤器混合。 随后在 42°C 下保持硝化纤维素滤器整夜以使得形成杂交产物。 将暴露于正常探针的硝化纤维素滤器以含有 0.1% SDS 的 $6 \times \text{SSC}$ 于 46°C 下洗涤，而将暴露于突变型探针的滤器以相同溶液在更严格温度 52°C 下洗涤。 随后将硝化纤维素滤器干燥并经放射自显影术。

[0330] 仅那些具有点突变的产物与突变型探针杂交。 每一分析中均包括阳性和阴性对照以确定 PCR 扩增是否成功。 因此，通过此方法确定实例 1 中所制备的患者基因组 DNA 在指定位置具有由非野生型核苷酸取代野生型核苷酸的独特点突变。

[0331] 实例 5

[0332] 目标 RNA 的反义抑制

[0333] A. 用于转染的寡核苷酸的制备

[0334] 通过在水中稀释至 0.5mM 、随后通过超声波处理以产生均匀溶液并滤经 $0.45 \mu\text{m}$ PVDF 薄膜来制备包含类脂质或胆固醇的载剂分子以用于转染。 随后将类脂质或胆固醇稀释至适当体积的 OptiMEMTM (Gibco/BRL) 中使得最终浓度为约 $1.5\text{--}2\text{nmol}$ 类脂质 / μg 寡核苷酸。

[0335] 通过首先在无菌 Millipore 水中稀释至 $100 \mu\text{M}$ 的工作浓度、随后在 OptiMEMTM 中稀释至 $2 \mu\text{M}$ (约 $20\text{mg}/\text{mL}$) 来制备反义和对照寡核苷酸。 随后立即将经稀释的寡核苷酸加入经稀释的类脂质中并通过上下吹吸而混合。

[0336] B. 转染

[0337] 使用易受 HCV 感染且支持 HCV 复制的人类 PH5CH8 肝细胞 (Dansako 等人, Virus Res.97 : 17-30, 2003 ; Ikeda 等人, Virus Res.56 : 157-167, 1998 ; Noguchi 和 Hirohashi, In Vitro Cell Dev. Biol Anim.32 : 135-137, 1996)。 所述细胞是通过在混合后立即加入寡核苷酸 / 类脂质混合物直到寡核苷酸最终浓度为 300nM 而转染。 随后将细胞与转染混合物在 37°C 、 $5\% \text{CO}_2$ 下培育整夜并将转染混合物在细胞上保持 3-4 天。

[0338] C. 总 RNA 萃取和逆转录

[0339] 按照生产商提供的方案使用 RNeasyTM 试剂盒 (Qiagen Corporation, Chatsworth, CA) 自经转染细胞萃取总 RNA。 萃取后将 RNA 逆转录以用作 PCR 模板。 一般将 $0.2\text{--}1 \mu\text{g}$ 经萃取的总 RNA 置放于无菌微离心管中，并加水使总体积达到 $3 \mu\text{L}$ 。 向每一管中加入 $7 \mu\text{L}$ 缓冲液 / 酶混合物。 通过按所列次序混合来制备缓冲液 / 酶混合物：

[0340] $4 \mu\text{L}$ 25mM MgCl_2

[0341] $2 \mu\text{L}$ $10 \times$ 反应缓冲液

[0342] $8 \mu\text{L}$ 2.5mM dNTP

[0343] $1 \mu\text{L}$ MuLV 逆转录酶 (50u) (Applied Biosystems)

[0344] $1 \mu\text{L}$ RNase 抑制剂 (20u)

[0345] $1 \mu\text{L}$ 寡聚 dT (50pmol)

[0346] 通过上下吹吸来混合微离心管中的内容物，并在 42°C 下培育反应物 1 小时。

[0347] **D. 目标序列的 PCR 扩增和量化**

[0348] 逆转录之后，使用 Roche Light Cycler™ 即时 PCR 仪扩增目标基因。通过将以下组份按所列次序混合来制备 PCR 扩增混合物的 20 μL 等份试样：2 μL 10×PCR 缓冲液 II (含有 10mM pH 值为 8.3 的 Tris 和 50mM KCl, Perkin-Elmer, Norwalk, CT)、3mM MgCl₂、140 μM 每一 dNTP、0.175pmol 每一 OAS1 寡聚物、SYBR® Green 的 1 : 50,000 稀释液、0.25mg/mL BSA、1 单位 Taq 聚合酶和为达到 20 μL 的 H₂O。SYBR® Green (Molecular Probes, Eugene, OR) 是一种当与双股 DNA 结合时发出荧光的染料，由此可直接测量每一反应中产生的 PCR 产物量。将 2 μL 完全逆转录反应物加入 PCR 扩增混合物的每一 20 μL 等份试样中，并根据标准实验方案进行扩增。

[0349] 实例 6

[0350] 以 OAS1 RNAi 处理细胞

[0351] 使用实例 5 的用于反义处理的方法以基于 OAS1 序列 (SEQ ID NO : 19) 的寡核苷酸处理细胞。合成并退火两个在 3' 末端具有脱氧 -TT 延伸的互补核糖核苷酸单体。以 50-200nM 具有 1 : 3 L2 类脂质的 RNAi 处理 PH3CH8 肝细胞系的细胞。在第 1、2、3 和 4 天收集细胞并通过 Western 分析法分析 OAS1 蛋白，如 Dansako 等人，Virus Res.97 : 17-30, 2003 所述。

[0352] 实例 7

[0353] OAS1 与 C 型肝炎病毒 NS5A 蛋白的相互作用

[0354] 使用 Taguchi, T. 等人，J.Gen.Virol.85 : 959-969, 2004 中所述的方法分析本发明的 OAS1 蛋白或多肽与 C 型肝炎病毒 NS5A 蛋白相互作用的能力。如上所述制备编码 OAS1 蛋白和多肽的聚核苷酸，并使用诸如 Taguchi, T. 等人所述的常规方法建构质粒。一种质粒含有编码 OAS1 蛋白或多肽的聚核苷酸，而另一质粒含有编码 NS5A 的聚核苷酸。所述质粒也编码各自蛋白的适当标记，诸如 FLAG- 标记、HA 或 GST。以编码标记和 NS5A 蛋白的质粒和编码不同标记和 OAS1 蛋白或多肽的质粒瞬时转染适当的细胞，诸如 HeLa 细胞。如 (Taguchi, T. 等人) 所述培育并制备上清液之后，可使用多种分析技术检测和量化 NS5A 与 OAS1 蛋白或多肽的结合。所述技术为此项技术中已知且包括共沉淀分析法、免疫荧光分析法和免疫印迹分析法。未显示与 NS5A 结合的 OAS1 蛋白和聚核苷酸适于作为 C 型肝炎感染的抑制剂而进一步分析。

[0355] 实例 8

[0356] OAS1 的化学和空间保守区

[0357] 如所属领域技术人员所认为的，当修饰 OAS1 的结构以提高酶活性或治疗潜力时，某些残基或残基区必定在化学及结构上为保守的。以实例说明，以下描述几种保守域。如所属领域技术人员所认为的，可容许对一些保持蛋白质结构和功能的氨基酸进行化学保守性改变。例如，Asp75 和 Asp77 均与对 OAS1 功能为必要的催化性二价金属离子配位。当容许对这些碱基进行修饰时 (例如对天冬酰胺或谷氨酸)，这些残基的本质上的极性和酸性性质必须受到保持。

[0358] 举例而言，关于 SEQ ID NO : 26-29、SEQ ID NO : 33、SEQ ID NO : 34 和 SEQ ID NO : 50，以下多肽片段表示保守域：

- [0359] 氨基酸 40-47 : FLKERCFR (SEQ ID NO : 75)
 [0360] 氨基酸 55-82 : VSKVVKGGSSGKGTTLRGRSDADLVVFL (SEQ ID NO : 76)
 [0361] 氨基酸 94-112 : RRGFIQEIIRRLQLEACQRE (SEQ ID NO : 77)
 [0362] 氨基酸 128-138 : NPRALSFVLSS (SEQ ID NO : 78)
 [0363] 氨基酸 145-158 : VEFDVLPALFDALGQ (SEQ ID NO : 79)
 [0364] 氨基酸 182-198 : KEGEFSTCFTELQRDFL (SEQ ID NO : 80)
 [0365] 氨基酸 201-217 : RPTKCLKSLIRLVKHWYQ (SEQ ID NO : 81)
 [0366] 氨基酸 225-241 : KLPPQYALELLTVYAW (SEQ ID NO : 82)
 [0367] 氨基酸 296-307 : PVILDPADPTGN (SEQ ID NO : 83)
 [0368] 氨基酸 337-343 : GSPVSSW (SEQ ID NO : 84)

[0369] 实例 9

[0370] 改良酶活性位点的氨基酸变化

[0371] 可设想改良蛋白质酶活性的 OAS 氨基酸序列变化。在一优选实施例中，可对酶的活性位点内的氨基酸进行修饰以改良 ATP 或金属离子结合、酶效率和酶的可反应性。所述改变的一实例为在 SEQ ID NO : 26、SEQ ID NO : 27、SEQ ID NO : 33、SEQ ID NO : 34、SEQ ID NO : 35 或 SEQ ID NO : 48 的氨基酸位置 61 处以酪氨酸取代甘氨酸。以化学上无害的甘氨酸取代极性酪氨酸将有利于 ATP 的 N3 原子与此氨基酸位置之间的氢键合，从而改良此相互作用的解离常数和能量学。举例而言，在可反应性更高的 poly-A 聚合酶中于此位置上发现了酪氨酸。如所属领域技术人员所了解的，可设想其他修饰。

[0372] 实例 10

[0373] 改良双股 RNA 结合的氨基酸变化

[0374] 对 OAS 进行的改良酶活性的氨基酸修饰的第二实例可为那些稳定此蛋白与双股病毒 RNA 之间相互作用的修饰。下表列出那些处在蛋白的 RNA 结合槽中的氨基酸和几种为稳定碱性、带正电荷的氨基酸侧链与带负电荷的核糖核酸之间的相互作用而设计的建议变化。设想增加蛋白的 RNA 结合槽中的正电荷密度的变化。如所属领域技术人员所认为的，可设想对 RNA 结合槽的相似类型的修饰。

[0375] 表 4

[0376] 对 RNA 结合槽中氨基酸的建议变化

[0377]

氨基酸	位置	建议修饰
甘氨酸	39	精氨酸或赖氨酸
赖氨酸	42	精氨酸
赖氨酸	60	精氨酸
精氨酸	195	赖氨酸

赖氨酸	199	精氨酸
赖氨酸	204	精氨酸

[0378] 实例 11

[0379] 基因突变分析

[0380] 所属领域技术人员认为，存在大量其他分析方法以评定基因分析中特定突变的进化重要性。一个实例为熟知的通常称为 D' 的连锁不平衡估计值的计算 (Lewontin, *Genetics* 49 : 49-67, 1964)。其他尤其相关方法试图通过比较基因座中高、中或低频率突变的相对丰度来评估基因座内选择压力和 / 或近代进化事件 (例如选择性清洗)。这些测试中最熟悉的是 Tajima D 统计法 (Tajima, *Genetics* 123 : 585-595, 1989)。Fu 和 Li, *Genetics* 133 : 693-709 (1993) 已开发出 Tajima 的变型及其他也利用关于每一突变的遗传等位基因的知识的统计法。将这些及其他方法应用于本发明的突变以评定对所观察到的效应的相对贡献。

[0381] 上述包括特定实施例和实例的说明书是用以阐明而非限制本发明。在不偏离本发明的真实精神和范畴的情况下可进行许多其他改变和修改。所引用的所有专利、专利公开案和非专利公开案均以引用的方式并入本文中。

<p>SEQ ID NO:1 等位基因变异体 :C/T Genbank基因组序列号 :NT_009775.13 基因组序列上突变的坐标 (起点-终点) :2135728-2135728 NCBI dbSNP ID (若有的话): 7955146</p> <p>CCCTCAGAGTGACTGAAGGAAATTCAGAGAAGAGCTGACACCTAAGTTGTAG ATTTTGCCYGAACAGGTCAGTTGACTGGCRGCTATAAACCTAACCCCAAAT CTATGTCAAGCTCATCG</p>
<p>SEQ ID NO:2 等位基因变异体 :G/A Genbank基因组序列号 :NT_009775.13 基因组序列上突变的坐标 (起点-终点) :2135749-2135749 NCBI dbSNP ID (若有的话): 3741981</p> <p>ATTCAGAGAAGAGCTGACACCTAAGTTGTAGATTTTGCCYGAACAGGTCAGT TGACTGGCRGCTATAAACCTAACCCCAAATCTATGTCAAGCTCATCGAGGA GTGCACCGACCTGCAGA</p>
<p>SEQ ID NO:3 等位基因变异体 :G/A Genbank基因组序列号 :NT_009775.13 基因组序列上突变的坐标 (起点-终点) :2135978-2135978 NCBI dbSNP ID (若有的话): 无</p> <p>ATGTATGGCCCTCCCACCAGGCCTGGTGGGTCCTGTCTCGACTGGGAGCAGA GGAGGGGTRGGGGGAGGAGAGAAAGAAGGGAGTGAAGGGAAGAGGGGGG GAGTGGTGGAGGGAAATAG</p>
<p>SEQ ID NO:4 等位基因变异体 :G/A Genbank基因组序列号 :NT_009775.13 基因组序列上突变的坐标 (起点-终点) :2144072-2144072 NCBI dbSNP ID (若有的话): 无</p> <p>ACTGAATCCAGCTGCAATGCAGGAAGACTCCCTGATGTGATCATGTGTCTCA CCCTTTCARGCTGAAAGCAACAGTRCAGACGATGAGACCGACGATCCCAGGA SGTATCAGAAATATGGT</p>

图1

<p>SEQ ID NO:5 等位基因变异体 : G/A Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2144088-2144088 NCBI dbSNP ID (若有的话): 2177979</p> <p>ATGCAGGAAGACTCCCTGATGTGATCATGTGTCTCACCCTTTCARGCTGAAAG CAACAGTRCAGACGATGAGACCGACGATCCCAGGASGTATCAGAAATATGGT TACATTGGAACACATG</p>
<p>SEQ ID NO:6 等位基因变异体 : G/C Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2144116-2144116 NCBI dbSNP ID (若有的话): 1051042</p> <p>GTGTCTCACCCTTTCARGCTGAAAGCAACAGTRCAGACGATGAGACCGACGA TCCCAGGASGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCTCATT TCTTCATAGACCCAG</p>
<p>SEQ ID NO:7 等位基因变异体 : G/A Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2144321-2144321 NCBI dbSNP ID (若有的话): 2660</p> <p>GGGCTCCAGTGTTATCTGGACCAGTTCCTTCATKTTTCAGGTGGGACTCTTGAT CCAGAGARGACAAAGCTCCTCAGTGAGCTGGTGTATAATCCAGGACAGAACC CAGGTCTCCTGACTCC</p>
<p>SEQ ID NO:57 等位基因变异体 : C/G Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2131025-2131025 NCBI dbSNP ID (若有的话): 1015542</p> <p>TAA CGCATGCCTGTAGTCCCAGGTATTCAGGAGGCTGGGGCAGGAGGATC SCTTGAACCCAGGAAGTTGAGGTTGCACGAGTCATGATCATGCCCCTGCAC</p>

图1

<p>SEQ ID NO:58 等位基因变异体 : G/A Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2133961-2133961 NCBI dbSNP ID (若有的话): 757398</p> <p>GACAGGAAGTGTAACCTCTCAGAGGCTCCCTTGCCACATCAGGAGAATTGR TAAAACCACTACCTGTATCATATCATTATTTAAGTGATAAATGATCA</p>
<p>SEQ ID NO:59 等位基因变异体 : C/A Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2139587-2139587 NCBI dbSNP ID (若有的话): (未知)</p> <p>TAGCATTAGGTATATCTCCTAATGCTATCCCTCCCAATTCCCCCACCCMGC TTGTTGGTATTTGTATATCTTCATTTGAGAATTCTCTGTTCATGTCCT</p>
<p>SEQ ID NO:60 等位基因变异体 : T/G Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2144294-2144294 NCBI dbSNP ID (若有的话): (未知)</p> <p>GTGCATCTTGGGGGAAAGGGCTCCAGTGTTATCTGGACCAGTTCCTTCATK TTCAGGTGGGACTCTTGATCCAGAGARGACAAAGCTCCTCAGTGAGCTGG</p>
<p>SEQ ID NO:61 等位基因变异体 : A/G Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2144985-2144985 NCBI dbSNP ID (若有的话): 7135577</p> <p>GAAAAATTATAGAACCTCCCTGTGTGACACAGCAGCCACTAGCCACATGTR TCAAATGCTTAAAATGTAGCTAGTCTAAATCTACATGTGCTGTGAGTGCA</p>
<p>SEQ ID NO:62 等位基因变异体 : C/T Genbank基因组序列号 : NT_009775.13 基因组序列上突变的坐标 (起点-终点) : 2156523-2156523 NCBI dbSNP ID (若有的话): 7968145</p> <p>ATGCTTCTATAGGCTTTTCTCACTGATGCTCTCTGGGCAGACAGGCTCCTY AATATGAGAGTGACACACTCCTTTCTTCATTTTCAGGTAAACCTCACA</p>

图1

<p>SEQ ID NO:63 等位基因变异体: A/- (A 或缺失 A) Genbank基因组序列号: NT_009775.13 基因组序列上突变的坐标(起点-终点) : 2156595-2156595 NCBI dbSNP ID (若有的话): (未知)</p> <p>CCTTTCTTCATTTTCAGGTAAACCTCACACTGGTTGGCAGAAGGAACTAT(A-) CCAATAATTAGTGAACATGCGGTGAATTTGCAACAGACAAGASGAGCCTC</p>
<p>SEQ ID NO:64 等位基因变异体: G/C Genbank基因组序列号: NT_009775.13 基因组序列上突变的坐标(起点-终点) : 2156638-2156638 NCBI dbSNP ID (若有的话): 7967461</p> <p>GAACTAT(A-)CCAATAATTAGTGAACATGCGGTGAATTTGCAACAGACAAGAS GAGCCTCATTATCCTATAGTTTCCAGGTTGCTTAGGGAGGCAGAAATCAC</p>

黑体单下划线碱基表示由各SEQ ID NO. 所鉴别的特定等位基因变异体。
 SEQ ID NO: 1、2、4、5、6、60、63和64内的双下划线碱基表示界定序列内的
 重叠突变，其由所提供的另一SEQ ID NO所鉴别。举例而言，
 为SEQ ID NO:1和SEQ ID NO:2提供的界定序列在基因组序列内重叠
 且因此其序列各自含有由SEQ ID NO:1和SEQ ID NO:2所鉴别的两种突变。

在SEQ ID NO:63中，据说突变或是A或是在
 指定位置上A的缺失，其表示为(A-)。

简并核酸密码:

R=A/G
 Y=C/T
 S=C/G
 K=G/T
 M=A/C
 W=A/T

图1

来自2, 130, 000-2, 157, 999的NCBI登记号NT_009775.13的序列片段

```
2130000 GGCTGCAGTGAGCTAAGATTGTGCCACTGCACCCAGCCTGGGCAACAGAGTGAGACCCT
2130060 GTCTCAACAAAATAAATAAATATATTTTGAATTAAAATAAAATGAAAAACAGCCTATCA
2130120 AAAAGTGTGGGATGTAGTGAAAGTAGTGTTTAGAGGGAAAGTTATAGCATTGAATGCATA
2130180 TATTAGGAGAAGAAAGATCTAACTCAGAGAGGGATGAACAGGTGAGGCATAGGGGATTTT
2130240 AAGGTAGTGAACCTCTTGCCAGGCAGGGGAGCCCAAACTGGGACACCGTCCGGGAAGG
2130300 CTCTTGGCTTCGTCCAGGAAGGAATTCAAGGGTGAGCTAGAGGAGGAAGAAAAGGTTTA
2130360 TTGAGGCAGCAGGGTTACAGTTCGTGTACTGTCCCTGCAGAGCAGGGCCACCCCATAGGC
2130420 AGTGCCTGGAGAGCAGCAGCTCAGGGCACTTCTATAGTCACATTATAACCCACTTTTAAA
2130480 TACGTGCAAATTAAGGGCAGGTTATTCAGAAATTTCTAGAAGAAGGGTGGTAACTGGGTC
2130540 ATTGCCAGGGAATGAGTAAACTGTTTCATGGTGTCTGGTGTCTATGCCAGCCAGTCTTCAAT
2130600 CTGGCCCTGAGTCAAGCCCCACCTCCTATCTCAAACTAATCTGCATGGTGTCTGTAATGG
2130660 TGGATACATGACATGTTATGTTTGGCAAAATCCATAGAACTGTAGGACACAAGAGTGAAC
2130720 CTTAATGTAAACCTTAATGTAAATGGACTTTTGTAAATATGATGTATTAATATCAATTC
2130780 ATCAATTGTAACAAATGTATCACAGTACTGTTAATAATAGAGGAAGTATTATGGAGAGA
2130840 GAGAGCTTATGGAACCTCTCTGCACATTCAAGTCAATATTTCTGTAAGCCTAAAAGTGGCTG
2130900 TGAGAAAATAAATCCAACCTGGGCAACATAGCAAGACCTTGTCTCTACAAAAAATAAAAA
2130960 ATGAGCTGGGTGCAGTAACGCATGCCTGTAGTCCAGGTATTCAAGAGGCTGGGGCAGGA
2131020 GGATCCCTTGAACCCAGGAAGTTGAGGTTGCACGAGTCAATGATCATGCCCTGCACCTCA
2131080 GCCTGGATAACAAAGCAAGATCCTGTCTCAAAAAATAAATAAATAAATAAATAAATACTAC
2131140 TAATTGAAAGGGAAAAAAGCATAGTATAATACCATTCTTAACAAAAAGAAAAGAGACCTG
2131200 TGTGTGTGTGTGTGTTAACATTTGAAAAAATCTGGAAGCTCTATATCAAAACGTTTAT
2131260 AGAGGCAATTTGTAGTGTAGAAATCATAGATGATCTTCCACTTCTGGTTTTTCTGCAC
2131320 TTTTTTCTTTTTGTCAGTGGGCATGTATTGCTGGAAAAACACAGCAACTGTGAAAGG
2131380 ATTTCATCAACAACAAAAAAGATAAAGAAGGAAACACAAAATCTGTTAATAAGATTT
2131440 ATGTTGGCTGGAGGTTAAAATGCATTTCCAGAGCAGAGTTCAGAGAAAGGCTGGGCTGCT
2131500 TGTGCTGGCTAAAGGACAAAGGGTAAGTTTCAGGAAGCAGAAGAGTGAACAGATGAAAT
2131560 TCAGCACTGGGATCAGGGGAGTGTCTGATTTGCAAAGGAAAGTGCAAAGACAGCTCCTC
2131620 CCTTCTGAGGAAACGAAACCAACAGCAGTCCAAGCTCAGTCAGCAGAAGAGATAAAAGCA
2131680 AACAGGTCTGGGAGGCAGTTCGTGGCCACTCTCTCTCCTGTCATGATGGATCTCAGAA
2131740 ATACCCAGCCAAATCTCTGGACAAGTTCATTGAAGACTATCTCTTGGCCAGACAGTGT
2131800 TCCGCATGCAAATCAACCATGCCATTTGACATCATCTGTGGGTTCTTGAAGGAAAGGCT
2131860 TCCGAGGTAGCTCCTACCTGTGTGTGTGTCCTCAAGGTGGTAAAGGTGAGTCCAGGCCTGC
2131920 CTGGCCAGGGGAGGGGTGGCTGAATGTGCAAGAGTTGAGATTGAGAATGAGAGAGAGAGA
2131980 GAGAGAGAAGCAAAAACCTAGAACCCAGGGTGCAAATGTGAGTACAGAGAGCTGAGATCT
2132040 TCTGGGATGGTGGTTTTCTTATTATTCACACAGCATGTTAAAATAGATTTCTGGGGTGAAA
2132100 TCCTACATCCCTATTATTAACAAGTGACCCTCCCCCTACTTCCCGCTGAAGTTTATGAA
2132160 CCACTGTCTGGGCGATGCCATTTTCAGAAATAGGGAAGTGAATCCCAGCTGTGTAAC
2132220 AGTTTGTCTAATTCGTGGCCAGGCTAGGGGCTCACCATTCTGCAGTGAAGAATCATATGT
2132280 TTTGAAAGCAAATAGCACCTGCTGGCTGCAAGACCTTGAGCAAGTCACTTAACACTCTG
2132340 TGTTCCAATTTCTCAGCCATAATCCCAATACTGTTGCAGTCTTGCCAGTGCACCTTAA
2132400 TGTAGCAGCTTCTCACTGAATTAGTACCCAAGGTTCTTTGTCTGCATCCAAGAAAATTA
2132460 AGGAACATGGACACAAACGTGAGCTTGGAGCAAAGTTCAAGTAAAGAAAAGAAAGCT
2132520 GTCTCCACTGTGGAGAGGGAAGTCTGAGTGGATTGCCAGATTGCAGCTGAATGCAAAAA
2132580 CTTTTATAAGAAAACCACTCTCTCCCTGTAAGTGTGAGAACTTTTTATCAGTAAAGC
2132640 TGTGCAACTTCCCTTACCTTATGCAGCTGTGGGTATATCTCTAGGCAAGCATAAAGCGCT
2132700 GCTTCTCTGTATGTATAACTGTGGATTTGTTTTAGGTAAGTCCCACTCCCTGCGCCAGT
2132760 TTCAGGCAGGCCGCTCCTCCAGGGCCAGCCTTGACCATTACCTAAGTATTTTCTCTC
2132820 TACTTTCCCTCAATACCTCATAGGGCGTGTAGATTAGTAAAATAGTAAAGTGTGAACCA
2132880 CCCAGCATAAGCTAGTCTGGCATCGTAAAGGACAATGGGAAAAGAACACAGATCCTGG
2132940 AAGAAGGCCCCAGGTTTGAATGTATTTGCCACCTACTAGCTGGGTGATGGGGCTGATA
2133000 TATTATCTCACTGAGCATCCATTTCCCATCTGTAAAATGGGAACATAATGATAATGGCAT
2133060 CCAAATCATAGCATATTGTGAGCATATAGGAGTTTAAAGACATGCAATGCCTTCAGAAC
2133120 AGTGGCTAGTGCTCCATAATGTAGTGTGCTCCTGTCTATTTATTTAGGGAGGTTTGC
2133180 CTCACTAAGCATCAATTAATTTTGTGCTCTTTTTCAGGGTGGCTCCTCAGGCAAGGG
2133240 CACCACCTCAGAGGCCGATCTGACGCTGACCTGGTTGTCTTCTCAGTCTCTCACCAC
```

图2

2133300 TTTTCAGGATCAGTTAAATCGCCGGGAGAGTTCATCCAGGAAATTAGGAGACAGCTGGA
 2133360 AGCCTGTCAAAGAGAGAGAGCATTTCGTTGAGGTCCAGGCTCCACGCTGGGG
 2133420 CAACCCCGTGCCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAGGGGTGAGTT
 2133480 CGATGTGCTGCCTGCCTTTGATGCCCTGGGTGAGAGCTCCAGCTTCTTTTTCTCCCTCT
 2133540 TCCCATTCTGAGCAGAAATCTCCACAGTTTGGAGAGCTTTTGGCCCAACAGGGCATCT
 2133600 CTCTAAAGCAGGTGGGAGGAGATCTTAGGATCTGTCCCGGGCAAGAATGAATACGGTC
 2133660 ATGATCTATCACAGGAGAGACATTAACAGCAAATTGGCATAATGTGGGGACAAAGACAT
 2133720 TTCTTACAGAACATCTGCAAGGCTTACTGGTCTGTTTAAAGGCAAAATGTGTGAATTTTA
 2133780 TCTTCTAAAATCAGGCAGCAAAGATGTGGCTTAAAGTTTATGTTACTCTCATCTTTGTC
 2133840 CCAACATGAGATCTCATCAAACGTATGCAGCACGTTGGGAGATAGATATTTATAATTTGC
 2133900 AGGAACATTTGGACAGGAAGTGTAACTCTCAGAGGCTCCCTTGCCACATCAGGAGAAT
 2133960 GGTAAAACACACTACCTGTATCATATCATTTATTTAAAGTGATAAATGATCATCTACAT
 2134020 CAGCTCTGATGAGTAATAGGTGTTCAAAAATAGGAACTTCCAGCCAAGTGTGGTGGCTCA
 2134080 TGCTTGTAATCCAACACTTTTGGAGGCTGAGGCAGGAGGGTCGCTTGAGCCCAGGAGTT
 2134140 CAAGACCAGCCTGGGCAGCAAAGTGAACCTCATCTCTACTAAAAATTTAAAACATTAG
 2134200 CCAAGTGTGGTGGTACATGCCTGTGGTGCAGTTATTGAGGACGCTGAGACTGAACGATC
 2134260 ACATGAGGCCAGCCAAGGATTCGAGGTGTCAGTGAGCCAGCAATGTACCCTGCCTCCA
 2134320 TCCTAGGCACAGAGCAAGAGCAAGACCCTGTCTCAATCAATCAGTCAATCAGTCAAACT
 2134380 ATGAATTTCCAGCTGTATATGAAGGCACCTCAAACACCACAGTGAACCTCACAGAGGGA
 2134440 CACGGAATAGTTTAGATTTTAAATTTTTTGGAGGAAATGCGATGACATCTGTACACACC
 2134500 CACAACCGGCTACTATTAACCTGAACCTACTGATTAGTGGCTACTAATTAATAGTTGGTC
 2134560 ATTAAGCAGTAATTAGTGATTAATTAATCAAGTAATTAGGACTTAATTAAGGAGGTGCA
 2134620 CAGTTTCCCTTAGTCTAGGGCAGCCATGAAAAAAAAAATGCTGACTCTCAAAGACACC
 2134680 AGGGTATGAGAAAGTTTGGATTCTCTCCTTTGTGCCATCTCCTGTGTGGGGGCTGAAG
 2134740 TACAATGGTTGTAAAAGACAAGAGGGGAGAAGGCTGGTACAGTGGCTCACGCCTGTAATC
 2134800 TCAGCACTTTGGGAGGCCAAAGTGGGGGATCACTTGAAGTCAGGAATCAAGACCAGCC
 2134860 TGGCCAACATGGTGAAATCTCACATCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGG
 2134920 TGTGTGCCTGTAATCACAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCCAGG
 2134980 AGATGGAGGTTGCAATGAGCCAAGATCATGCCATTGCACTCCAGCCTGGGCAACAGAGTG
 2135040 AGACTCCATCTCGAAAAAAGAAAAAGAAAAAATATAAGGAGTGATTAAAAAAGAAAA
 2135100 GAAAAGAAAATAAGTAGGGTGAACAATAGATAGCCATGGGGGTTAGGGAGCTTTTTTA
 2135160 GCAGGGTCTGAGGGAGGGTCCCTGAGCCTGAGTGGCGAGAAGGAGTGAGCCTTGGGGA
 2135220 GATCTGGAGGTTCTGGGAAGAGGAATGGCAAGTGCAGAGGCCCTGAAGCAGCAATGACCA
 2135280 TGCCACATTTGAGGAAGAGAGAAAAAGTGCAGAGAAGTAGAAAGTGGGCAAAGGAAGCAAG
 2135340 ACAGGAGGTGAGGTGGGAGAGGTTCCAGAGACCAGATCACACCAGACATCATTGGCCACC
 2135400 ATAAGATCTTTGGGTTTAAAAATCCAGATGTTATGGGATGCAGGAAGCAGCATGATCAG
 2135460 CAGCATTCTCTAGGTGCCAGGTTGAGAACAGGCTGTGGGGAACTGTAAGGAGGTGCT
 2135520 GCCATAGTTCCGGCAGTGACGGTGGTGGCTTGGATGGGGTGTGGCAGTGGAGAGGGCA
 2135580 GGAGGGAGGATCAGGAATGGACCTCAAGACTTCCAGCCCTGGGTCTGCTGCACTTTTCA
 2135640 ATCAAACCCCATGGCCAGGAGATTGTCCCTCAGAGTGACTGAAGGAAATTCAGAGAAG
 2135700 AGCTGACACCTAAGTTGTAGATTTTGGCCGACAGGTCAGTTGACTGGCGGCTATAAAC
 2135760 TAACCCCAAATCTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGA
 2135820 GTTCTCCACCTGCTTACAGAACTACAGAGACTTCTGAAAGCAGCGCCCAACCAAGCT
 2135880 CAAGAGCCTCATCCGCCTAGTCAAGCACTGGTACCAAATGTATGGCCCTCCCACCAGGC
 2135940 CTGGTGGGTCTGTCTCGACTGGGAGCAGAGGAGGGGTGGGGGAGGAGAGAAAGAGGG
 2136000 AGTGAAGGGAAGAGGAGGGGAGTGGTGGAGGGAAATAGAGGGATGGAAAAAGGAGAGAA
 2136060 AGGAAAAAGAGGTGGAGAGAGGAGCCTGCAACAGAAGGGAGAATGAAAGGGAAGGAAGAG
 2136120 AGAAAGGAAGGGATTTTGGTGTCTGTTCACTGCTGTATCCCCAGAACTTAAAACAGAGC
 2136180 CTGGTGCATAATAGGTGTAATAAATGTTGAATAAATGAATCAATGCTACATACACACAC
 2136240 GCACGCACACACACAGAGAGAGAGTCAACCACACTCTTCAGAAGGTGGATAAGTTAAA
 2136300 ACAAGAGTTTCAAACAAATATATGTTGATGATGCCCTTCTCTCCACTTACTGGCTGGCTG
 2136360 GCCTTAAGTAAGCAACTTAACTTCTGTTCTTCTGCTTTCTTATCTGCAACGAGTAGC
 2136420 ATGCCATAGCTAGAGTAACACGGCATATAGTTGGTCTGATAAATGTAGCATAATTTAGC
 2136480 CACCATAGGAGTACACATAATAAAGCTAACATGTAGTATGTGCTTAGCTTATCTATGTT
 2136540 TTGTGGATGTGATACAAATTTCTGTTCACTTTTAAATGCCCTGCATCTTAGTCAATTTTA
 2136600 ACAGTGATCTGTAAAGTTAGATAAGGTTAGGCATTATATTAATCCATTTACACCAAG
 2136660 AGAACTTGGGTCAAAAAGAGAACTCTGGGTACATGGCTCATTCGGCCAATAAGTAG

图2

2136720 CAGAAGTAAAAATTGAATTTGGCTGGGCGCGGTAGCTCACACCAGTAATCCCAGCACTTT
 2136780 GGGAAAGCCAAGGCAGGTAGATTGCTTGAGCCCAGGAGTTCAAGACTAGCCTGAGCAACAT
 2136840 GGCAAAACCTCGTCTCTACAAAATAAACTAAAAATTTAGCCAGGTGTGATGGTGAGCACC
 2136900 TGTAGCCCCAGCTACTGGGTAGGCTGAGGTGGGAGGATCGCTTGAGCCTGGGAGGAGGAG
 2136960 GTTGCAAGTAAAGTCAGGATTGCACTACTGCCCTCCAGCCTGTGAGACAGAGCAAGATCTTC
 2137020 TCTCAAACAAACAAACAAACAAACAAACAAAACTCGAATTTGGGTCTATTGACTTAAAGA
 2137080 GTTTGCCTGATAATAATAGGCATTCAATGTATATTTCTTGAATGAACGAATGAATGAAAA
 2137140 TAATCAGGAATAAACTTTCCAATTTAAAGTAACACCTCTAGGTAAGAAAAAGACAATCA
 2137200 TTTAGTTGCCAGACTTCTAAGTGTGGCTGTTCTATGAATTGTAATCATGGAGCCTGAGC
 2137260 ATTGTAGAAATTTACAAAAGCAGTTCCTGACAAAAGCAGCACTGCCCCAGGGACATATTG
 2137320 AAAATTAATGAGGGTGTTTTTGGTAAACCATGGTGATGGGAGGACATGGGTGCTACTTATA
 2137380 TTTAGTGGAAAGAAGACAAGAATGCTAGTTATTTGTACAATGATCAAGAGAGTCTGCACA
 2137440 GCCAAGAATTGTCTTTTTCTTTCTTTCTTGATGCTGTTCTCCTTTAAAACAAGACAAGAT
 2137500 TAACAATAATTTAACTCCACTAACCCATCATCACCACCTCCAACCTATATGCTACATTT
 2137560 TCTTGTATATTTCAAGTCTGTTTATATTTCAAGTGCCTCGAAGTATATTGTTTTATAG
 2137620 CCAAATGTTTAGTTAATCTGCTCACAGATTTACCCTTTCTTCACTATTCTTCTGCTTT
 2137680 ACACCTCTAACATTCATCTGGGGTAATTTCCCTAAATGATCATGCATCCTTTTGGGATTT
 2137740 CTTTTGATGATGGTCTATTGGTAGTAACTCTCTCAGTTATTGTTTGTCTGAAAAATGCA
 2137800 TGCTTTTGCCCTCATTGTTGAAGGGTGCTTTTGCTGGGTGGTCATTTCACTATATGAAT
 2137860 ATATCATTCCATCTTCCAGTGTCTATTAATAAAGTCAAGTGGCAGTCTAACTGCAGCTC
 2137920 TTTTATAAGTAACCTGTCTTATTCTTCTGGCTGCATGTAAGTATTTCTCTTTGCTTTT
 2137980 ATTTTGTTTAGCTTCAATCTGCTGTGCTTAAATGATGGGTTCCATTTGTTTGTCTGATTT
 2138040 GGGATTCGGTTAAGATTCCTGAATCTGTGGGTAGATATCTTTAATCAGTTTGGAACTTC
 2138100 TCAGCCATCTTCTAAAATATTGATTCTCCTTCATCTCTCCTCACCTTCTAGAATCCA
 2138160 ATTAATGTATGTTAGACCCTGCTCTATCTTCATATCTCTATACTCTCTCTGCTGTTTT
 2138220 TCATCCTTTTGTCTATTTTTCCATGCTTTATTTCTGAATAGTTCCTTCTAATCTACCTTCC
 2138280 AATTAACATAATTTCTCTTTAGCTATATCTAATTTGCTGTAATTAATTACAGTTGCCATT
 2138340 TTTATCCTAAAATTTCTATTTCAATATTTTTGTATCTGCCATGGTACTCTTTAGGCTTTT
 2138400 AATTCCTGCTAACTATTTAAAGTTCCTTATTTTATCCTGTGAATATGATATTCCTAGTTA
 2138460 TTTTATTTTAAATTTTTATTATTTGTAACTCTTATGTTTTATTTTACACTTCTTTCTGTG
 2138520 ACATGAGCACACAGATTCAATGTGTATACATATATGGCTCTGATACCTCTCCCTTCTG
 2138580 TCCTCATCAAACCACTGATCACAGAGAGAGGACTATTTTTTTTTATTTTTAAATTTTTCT
 2138640 ATTTCAATAGGTTTTTGGGGGAACAGGTGGTGTGGTTACATGAATAAGTCTTTAGTG
 2138700 GTGATTTTGGTGCACCCATCACCAACAGTGTACATGTACCCAATGTAACTCTTTA
 2138760 ACCCTTGCCACACCCACCCCTTTCCCGCAGTCCGCAAAGTCCCATGTATCATTCTTATG
 2138820 CCTTTGCTTCTCATAGCTTAGCTCCACATATGAGTGAGAACATACAATGTTTGGTTTT
 2138880 CCATTCCCTGAGTTATTTAATTAATAATAGTATCCAATTCATCCAGGTTGCTGTGAAT
 2138940 GCCATTATTTGTTCTTTTTATGGTTGAGTAGTATTCATGGTGTGTTTGTGTGTGAT
 2139000 AACATTTTTCTTTATCCACTCATGATTGATGGGCATTTGGGCTGGTTCATATTTTGC
 2139060 AATTGCAAATTTGTGCTGTTATAAACATGTGTGTGCAAGTATCTTTTTTGTATAATGACTT
 2139120 CTTTTCTCTGGGTAGATACCTAGTAGTGGGATGCTGGATCAAATGGTAGATCTACTTT
 2139180 TAGTTCTATAAGGAATCTCCACACTGTTTTCCATAGTGGTTGTATGAGTTTACATCCCA
 2139240 CCAATGGTGTAAAAGTGTTCCTTTTTACCACATCCACACCAACATCTATTATGTTTGA
 2139300 TTTTTTATTATGACCATTCTTGCAGGAGTGAGGTGGTATCACATTGTGGTTTTGATTTGC
 2139360 ATTTCCCTGATAATTAGGGATGTGAGCATTTTCCATATGCTTGTGTTGTTTGTGTTTT
 2139420 TTTTTTTTTTTTTTATTATATACTTTAAGTTTTAGGGTACATGTGCACAATGTGCAGGT
 2139480 TAGTTACATATGTATACATGTGCCATGCTGGTGTGCTGCACCCATTAAACCCGTCATTTAG
 2139540 CATTAGGTATATCTCCTAATGCTATCCCTCCCAATTCACCCACCCCGCTTGTGGTAT
 2139600 TTGTATATCTTCATTTGAGAATTCCTGTTTCATGTCCTTAGCCCACTTTTTGATGAGATT
 2139660 TTTTTTTCTTGTCTGATTGCTTTGAGTTCTTTGTAGATTCTGGATATTAGTTGGATGTAT
 2139720 AGATTGTGAAGATTTCTCCCATCTGTGGTGTGCTGTTAACTCTGCTAAATATTCTTT
 2139780 TTGCTTTGCAGAAGCTTTTTAGTTTAAATTAAGTCCCATCTATTTATCTTTGTTTTGTTG
 2139840 CATTGCTTTTGGGTCTTTGGTCATGAAGTCTTTGCCTAAGCCAATGTGTAGGAGGGTTT
 2139900 TTCCAATATATCTTCTAGAATCTTTATGGTTTCCAGGTCTTAGATTTAAGTATTTGATCG
 2139960 ATTTTGAGTTGAATTTTGTATAAGGGGAGAGACAAGGATTCAGTTTCATCTTCTACATG
 2140020 CAACTTGCCAATTATCCTAGGACCAATTTGTTGAATAGGGTGTCTTTCCCATTTTATGT
 2140080 TTTTGTGTTGTTGTCAAAGATCAGTTGGCTGTAAGTGTGTTGGCTTTATTTCTGGGTTAT

图2

2140140 CTATTCTGTTCCATTTGTCTACGTGACTATTTTTATACCAGTACCATGTTGTTTTGGTGA
 2140200 CSTATGGCCTTACAGTATAGTTTGAAGTCTGATAATGTAATGCCCTCCAGATTTGTTCTTTT
 2140260 TACTTAGTCTTGCTTTGGCTATGTGAGCTCTTTTTTGGTGCCATATGAAATTTAGGATTG
 2140320 TTTTTCTAGTCTGTGAAGAATGATGGTGGTATTTTGATGGGAATTGCATTGAATTTGT
 2140380 AGATTGTTTTTGGGAGTATGGTCATTTTCACAATATTGATTCTACCATTTCATGAGCATG
 2140440 GGATGTGTTTTCCATTTGTTTGTGTCTATGATTTTTCTTTCAGCAATGTTTTGTAGTTT
 2140500 TCCTTGTAGAGTTCCTAGTTATTTTAAAGTCTGTGTTCCGGTCTTTCAGCATTTAAAGTTT
 2140560 GTAGGTTTTATTACTATTTCTCTTCTTCTGTTGGTCATAACTCTTAGTGTGTTTTGTTCCCT
 2140620 TGTGTGCCCTGGTTACATATGTGCTGGTCATTGTATTTGAAAATTATGTGTGAAATAATT
 2140680 GAGGTTTTGGATTATGTATATTCCTCCAGAAAGAATTTCAFTTGCCTCTGTGCATTTCTT
 2140740 AGGAACATTTACAAGTCCCTCTTCTCAGTTAATTTTCGTAGTATCTTTATCAGATAGGTGC
 2140800 TATTACAACCACCTCACTTAGCAGATGAAAATCATGAGGCTCTGAGAGTCTAAGTCATCTA
 2140860 CTTAGAATTGGACAATGGTGAAGCCAGGATTCAAACCCACATCAATAAGAATCCAGCGCT
 2140920 CTTAAACAAGGGGCCAGTACACTTTTTTAAAAAATAAAAGGCTAGATAGTAAATATTTTAG
 2140980 ACTTTGTGGACTGCACAGCTGTGCAACTACTCAACCCCTGCCCTTGTAGCATGAATG
 2141040 CAGTCATAAACTATACATAAATGAATGAGCCTGGATTTCGTTCCAAGGAACTTTATAAAA
 2141100 ACAGGTGGCAGGCTGGATTTGGCCATGAGAAGTGTAGTTTACACAAAAGTTGAGCAAAAC
 2141160 CAATTTTTTTCTGATTGTTTTTCTCTTCTCAGTGTAAAGAAGAAGCTTGGGAAGCTGCCA
 2141220 CCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAAAACA
 2141280 CATTTC AACACAGCCAGGGATTTTCGACGGTCTTGGAAATTAGTCATAAACTACCAGCAA
 2141340 CTCTGCATCTACTGGACAAAGTATATGACTTTAAAAACCCATTATTGAAAAGTACCTG
 2141400 AGAAGGCAGCTCACGAAACCCAGGTATGCTATCCCCACATGGCTTAGCTCCCTATGTAA
 2141460 ATGAACACCTGGATACAGGTACAGTGCCTTGGAAATGGAGGAGGTGGGAGGGCTCCCCAC
 2141520 TTAGTGAGAATCTCCTGTTGCCCATCATTGTACTGGGCATTTTACTACTGCCATCTGTTT
 2141580 TAAACACCTACCTCCAACCTGTGAGGCAGGCACTATGCCAATTATTTTACAGGTGAGTA
 2141640 AACTGAGGTTCTGAGAGGTAAGGAGCTTGTCCAACCCCTAACAGAAAATGAGTAAAATAG
 2141700 CTGCAGTTGAACTGAAATAAGAACAGCAGCAACAACAATGATAGTAATTGCTCCCAGGT
 2141760 AATTGAAAGCTTTGTTGTAAGACTAACACATGCTAATAATAATAGTAAAAAATTATTAGCAATA
 2141820 TTACTGATATGTATGTTATGTTCTAGTCGCTGTGCTGAGCATTTTATATAACTGGGCTTT
 2141880 TTCTATCCTCACAGCATAGCCTTTGAGATAGGTATGTGGAACATTTCCCATTTTACAGAT
 2141940 AAGAATCCTGAGGCTTAGAGAGTTCAAGTGACCTACCCAAGGGCACATCACTGATAAAGG
 2142000 GCAGAGGTGGGATTCAAACCCACATCTGTGAGGTGCAAGTGCAAGGCTCCTTCTCCTCAT
 2142060 GCTCACTGCCCTGCTGGGGAATAGGGCACTGGGGACATACCCAGGGAGCCCTTCTCATG
 2142120 TTCTGAGTCCCAGTTCATCCCAGTCTGCTATTTTGGCTCTCCAGGAGCATCTAGCATCCC
 2142180 TAGACAGAGCCCCAGCTTCTCACCTGTCCCTCTCTAAATGCTGCTCTGCAGGCCTGTGAT
 2142240 CCTGGACCCGGCGGACCCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCA
 2142300 GCTGGCACAAGAGGCTGAGGCTTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGGTC
 2142360 CCCAGTGAGCTCCTGGATTCTGCTGGTGAGACCTCCTGCTTCCCTCCCTGCCATTCATCCC
 2142420 TGCCCCCTCTCCATGAAGCTTGTAGACATATAGCTGGAGACCATTCTTTCAAAGAACTTAC
 2142480 CTCTTGCCAAAGGCCATTTATATTATATAGTACAGGCTGTGCTCCATATTTTACAGTC
 2142540 ATTTTGGTCACAATCGAGGTTTTCTGGAATTTTACATCCCTTGTCCAGAATTCATCCC
 2142600 CTAAGAGTAATAATAAATAAATCTCTAACACCATTTATTGACTGTCTGCTTCGGGCTCAGG
 2142660 TTCTGTCTTAAGCCCTTTAATATGCACTCTCTCATTAAATAGTCACAACAATCCCATGAG
 2142720 GCATTTTTTAAAAATTTTTTATTATTTTAGATTAGAGGGCACATGTGCCATTTGTTACAC
 2142780 AGCTATATTGTGTAATGGTGGGTTTTGGGCTCTATTGATCCTGTGCGCCAAAATAGTGAA
 2142840 CAGAGTACCCAAAAGAATTTTTTCAACCTTTGCCTTTCTCCCTCCCTCCCTGTTGG
 2142900 AGTCCCTAGTGTCTATTGTTCCCATCTTAGCAGATGTTAAGTATTTGATTTTCTGTTTT
 2142960 TGGGTTAATTCACTTCGGATAATGGCCTCCAGCTGCAACCATGATTTTATTCTTTCTTAT
 2143020 GGCTGCATAATACTCCATGGTGTAGATATACCACACTTTCTTTATCCAGTTCACACTGAT
 2143080 GGGCACTTAAGTTGATTCCATGACTTTGCTATTGTGAATCGTACTGCGATAAACATACGA
 2143140 GTGCCGGTGTCTTTGATAGAATGATTTCTTTACCTTTGGGTAGATACCGAGTAGTGGGA
 2143200 TTGCTGGGTTGAATGGACATTTACTTTTAGTTATTTGAAAAGTCCCATGAGGCATGTTT
 2143260 TCTATCATTCCCATCTTACAGATGAGACAAAGGCTCAGAGAGGTGAGGTCACCTGCTCAA
 2143320 GGACATCAGCTAACAGTGGTGGAAATGGAATTCAGCTCAGTGGACTCTAAAGCCAGTG
 2143380 CTCATGTCACTGTGCTAAACAGCCTGCCTTGTACATCCCCACCTCTCATCTGACCAATG
 2143440 GGAGACTCTGAGCAGCTGAGTACTTTGGGTTGTCAACAGCTAAACAGGGGCAAGGACC
 2143500 CAGTCTTGGATCTTCCACCTCCAAGCAGGAATCTGTCTGATTCCAGGGGATTGATGATG

图2

2143560 TTGCAGATGGCTAGGAAGCAGACTCCAGGATGGAATTTAGTATGCAGGATGTTCTGGGGG
 2143620 AGAGCCACTGGAACCAGCAGCTCAGGGAAAGGGGGGAAGAAAGGATAGGAAGGAAGCATGA
 2143680 AAGAGAATAGGGAGAAGTGAACAGGGATGCAGAGCGAATGCCAGTTTCAGCCAACCTCAA
 2143740 GGACAGCCCTGGAGCTGGAATGGCCTTTAGAGCTGCCCCATGGTGAACAGAGGTGGCCAGG
 2143800 CTTCTATACCCCTACGTGGATCACTCACTGTGCTTGGGCACCTTGGGAAAGGGCATGGCT
 2143860 TTGAGCAAAGGCTCTCTGCAGCTGAGGCAACCCCTAAAAGGGCTGACGGCTGAAGTCTG
 2143920 TCTGCTGACCACGTGCCAGCAGCTGGGGCTTGTTAGTCCTTCTCAAAGGGGGATCCAG
 2143980 ATGGCATGTACAGTGTCTACCGTAAATGCTCACTGAATCCAGCTGCAATGCAGGAAGAC
 2144040 TCCCTGATGTGATCATGTGTCTCACCTTTCAGGCTGAAAGCAACAGTGCAGACGATGAG
 2144100 ACCGACGATCCCAGGAGGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCCTCAT
 2144160 TTCTCTCATAGACCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAGAGGACTGG
 2144220 ACCTGCACCATCCTCTGAATGCCAGTGCATCTTGGGGGAAAGGGCTCCAGTGTATCTGG
 2144280 ACCAGTTCCTTCATTTTCAGGTGGGACTCTTGATCCAGAGAGGACAAAGCTCCTCAGTGA
 2144340 GCTGGTGTATAATCCAGGACAGAACCAGGTCTCCTGACTCCTGGCCTTCTATGCCCTCT
 2144400 ATCCTATCATAGATAACATTTCCACAGCCTCACTTCATTCCACCTATTCTCTGAAAATA
 2144460 TTCCCTGAGAGAGAACAGAGAGATTTGATAAGAGAATGAAATTCAGCCTTGACTTTCT
 2144520 TCTGTGCACCTGATGGGAGGGTAATGTCTAATGTATTATCAATAACAATAAAAAATAAAGC
 2144580 AAATACCATTTATTTGGGTGTTTATTAACCTCAAGGCACAGAGCCAAGAAGTACAGATGCA
 2144640 TATCTAGGGGTATTGTGTGTGATATACATTGATTCAACAAGAAATATTTATTGAGCAGT
 2144700 TACTATGTGCCAAGCATAGCTCTGGGCACTGGGAATATAGCAATGCACAAAAGCAGACAG
 2144760 AAATCCCTGTCTCATGACCCTGCAGAGCCAAGACTTCCAGAAATTTTAAATAAAAAAAA
 2144820 TCCCTGTCTCATGGAGTTGACATTTGTGCAAAAACATCTTAATGTTAGATGGTTTTCTTA
 2144880 TTACTAATAATTTCTGAAATAAGCATCCTTGATTTATCCTTTCTCCATATCTCTGAGAAAA
 2144940 ATTATAGAACCCTCCCTGTGTGACACAGCAGCCACTAGCCACATGTATCAAATGCTTAAAA
 2145000 TGTAGCTAGTCTAAATCTACATGTGCTGTGAGTGAAGGTATATACTTGGTTTTCAAAGAC
 2145060 TTAGTACAAATGAAAAGAATGCCAAGTCTTGCCAACTGATAATTTTTTAAATTGTGTGC
 2145120 TCAAATGACAATTTTTAAATATATTTGAGTTAAATCAAATGAACTTCATCTCTTTCTTTT
 2145180 CCCTTTTTAAATTTGTGGCTACTAGAAAATGTGAAATCATAACATGTGGCTTGTGTTATATTA
 2145240 TGTATTTCTATTGGACAGCTCTGTCTCCAAGGTAAATCACTGGATTAAGATTTCGACTA
 2145300 TACTGACTTACATTGCCACATTGTCACTGTCTTGGGACCAAGAATCAACATATCATT
 2145360 CATAAGACTCTAAAATATAAACTCTCATAAATACTCACAAAAGAACCTAGCATGCTCTG
 2145420 ATCACCTGAGTTGCTGGTCACTTTTTGGTGGCTGGTAAGCAGCCTTTGGTCCGTCCAGATT
 2145480 ATATTTCTCCATTTAGTCCCCCACAATCCCTGTGAGATGGGTTTTGTTGTTATTCTCATA
 2145540 ATATTAAGTGAATAACTTTGAGGTTCTAAGAGGTTACAGTGTGCCCCAGGGTCAACCCAG
 2145600 CTGGTCAAGGGCAGAGTCTGAACTTGAAACCTAATCCTTCTCTCTCTAAAGCTCATGTTCT
 2145660 TTAATCACTGTAGCATGGTCTTAATGTGTCTCATTCATTGAAAGCTTATGTTTTCTCTAC
 2145720 TCTGGCCCATGAGAACCAGAAGCATCAATGTCCAGGGGCAGGAAAAGATGAATGTCCCA
 2145780 GCTCAACCTGACACAGATTCAACCTTCCCTGGTCTTTTGTCTTATTTGATAGCTGGAT
 2145840 TAGATGATGCCAGTCTACTGATTCAAATGTGAATCTCTTTCAGAAAAACCTCACAGATA
 2145900 CACCTAGAAATGATGTTTACCAGCTATCTTGGCACCTTAGCCAGTCAACTTGTCCACA
 2145960 TAAAATCAATCATCACACACTCCATGCTGATAGGCAAGTGTGGACATCCAATGTAATGG
 2146020 CTTCAATGTATTTACTGTGTGAAAAATGCACTTGTGTTGCCCTTTGAGAGTGTTCATT
 2146080 TTATAGCAATGCCACAACCAACAGTAGATTAATGGAATCAGTAATTAGTTGCTTGATCGA
 2146140 AGAGCCACATGGCCACGTGATCAGCCTTCCATCTACAACAGGACCCAGGAGTATACCGGG
 2146200 ATTGTTTTCAAAGGGCATAGACATTTTGTCTGCAAATGACATGGGCTTACTCCAGAGTCT
 2146260 CTGGAGGGTCTGTGTTATAATTTCTCAAATAGATATTGCCATAATCTCTGAATGACACC
 2146320 TTTTCCCATGACTAACACTTTGAACACCATGGGGTCTGCCAGGCTGGTGTGGGCCAAGTA
 2146380 GAGGGGCGACTTGCACCACAGCCTATACCAGCTGCAGAGCCCTTAGGACTTAATAAAAG
 2146440 GGTGCTAATTTCTGTACTTCCCTGGCTCTGAGATGTAATAGTGGTTTTAATTTACTATCC
 2146500 TGGCCAGGGAGGTGGCAGTTTCAAAGGTATCCCATGACCTTCCCTCACTGTGATAGCCCT
 2146560 CACTTAACCCTCAGGCCAATGTGGGTAGTGTCTTTTACCAAGCATGTCCTTTTTGAATA
 2146620 TACATTCAGGGGACAGGGAATGATCACCAGGTGGTCCATAGACACAGTGGGCAAATGA
 2146680 CAAGCTGACTTGGCCAGGGCTCCATTTTCACTCTTGGCTCCTATATGCCCCCTACTCTG
 2146740 ATGGGGACAGAATCTGATGACGTTTTCTGTTATCAGTGTGATCCTCTGCCCAACAGTCT
 2146800 TAGAAATGTTGAGTACTCCCTTTTCCAGTGTATCATCCTCTGAATAAATGGCCATAGG
 2146860 TCCTTTGGGGGAAACCATTACTATATATGCTATGGTGTACAGCATCCTTTCCAAGGGG
 2146920 ACTCTGCCTCCTTTGTTAATGGGCTAGTCTGAAAACCTGGGTTGGTTCAGTTCCAAAA

图2

2146980 ATCAGGAAGGGATGATGACTTTTTATTGGAATGGCTCCCCTTAACCCCTGGTGGTCCGG
 2147040 CCTTGATTTCTTCTAATGGTATAAATGAGTAGTGCCCTTTCAAGATACTCATCCACTAA
 2147100 GTCACATAAGAGGCCACATTTGTGTTAACAGGTCCCTAACCTGTATCAGTCACGACCCCT
 2147160 GGCTGCCCCCTCAAACCTCACCATTCAATATAGTCATTGCACCACCTGGCTTCTATTGCT
 2147220 CTGGGAGTCTGCCTGCCCTATTGCCCTTATCAATCCAAGCTCTGTAACAGCCCCCTCTC
 2147280 CCAGGAGGCCACCACCAGAACACCTCATGATGCAGGTGTCCCTCTCTCCACACATTCCTT
 2147340 ATGGCCTCGGGGAATGTTTTGTCCCTGGGACTCTATTTCCAGACTTACATAGTAAACAT
 2147400 ATCCACTCCTCTTCACTGTGCCATTTCCCTTGGCCTCTAATACCTTCTTCCATCAGCT
 2147460 CCACAGCAATTTCTGGAATTTCTACTTTTCTGCATTTCTAGGAGCCATCCTAGGAGTCAG
 2147520 GGAAAACCTCGGCTTCACAAGTAAATGGGACTTCAGTTGTGTGCCAACATGCCTGGCTAA
 2147580 TTTTTAAATTTTTTTGTAGAGATGGGGTCTTGCTGTGTGTGCCAGGCTGGTCTTGAACCT
 2147640 CTGGCCTCAAACAATCTTTCCTCTTGGCCACTCAAAGTGTGGGATCACAGCTACCATT
 2147700 CCCACACACATTTTCTTATTTTTAAAGGCTGTATGTGCACGTATACATTAATGTGTGC
 2147760 ACTGTATACATTAACCTGTGTGCACGTATACATTAATTTCTTAAACGAATTTATCCATT
 2147820 TATAGTTGCTTTGGTTGTTTCCACTTGATTAATTGTGAATAGTGTGCAGTAAACATGGGA
 2147880 ATGCAGTTATCTCTTTGATATCCTGATTTCAATCTTTTGGATACTCAGAAGTGGGATTG
 2147940 CTGGAACATATCGTAGTTCATTTTTTAATTTTTTGGAGAACCTCCATACTGTTTTTACA
 2148000 GTGGCTCACCAACAGTGTGCAAGAGTTCCTATCCCTCCACATCCTCACACTTGTATCTT
 2148060 TTGTTCAATCTTTAAAAAATGATAGCCATCCTACCAGGAGTAAGGTGATATTGCATCGTG
 2148120 ATTTTGATTTGCACTTCTCTGATAATTAGTGATATTGAGTATATTTTCATAGACCTGTTA
 2148180 TCCATTTGTGTGTCTTCTTTGGAGAAAGATCTATTCCATCCTTAGCCATTTTTAAATC
 2148240 AAGTTATTAATTTTTTGTAGTAGTGGTAGGAGTTCCCTTACATATTGTGGAGATTAAAC
 2148300 CCTTATTAGATGTATGGTTTTGTGAATGTTTTCTTCAATCCATAGATTGTCTTTTCAGCCC
 2148360 GTTGATTGTTTTCTTTGCTATGCAGAGCTTTTTAGTCTGATGTAGTCCAAGTTGCCTAT
 2148420 TTTTGCTTTTGTGCTGTGCTTTGCATACGTGGCCACCTGATCTTTGACAAGATTGCCA
 2148480 AGAATACACAATGGGAAAGGACAGTGTCTTCAACAAATGGTGTTTGGAAAGCTGAATGT
 2148540 CCACATGCAAAAGAATAAAAATTGGACCTTACCTTACAGCATACACAAAAATCAACTCAA
 2148600 AATGGATTAAGACTTAAACGTAAGACCTGAAACTGAACTACTAGAAGAAAACTTAGGG
 2148660 GAAAACCTCATGACATTGGTCTTCCAGTGATTTCAGGGATGTGACACCAAAAGCACAGA
 2148720 CAACAAAAGGCATTTATTTTTTATATGGCATGTGAGGAAGGGTTTCAGTTCCAGTTCTTCC
 2148780 AATGTGGATGCTCAATTATCCAGCAGCATTATTGAACGGATCATGTTCTCTCCAATCTC
 2148840 TTTGCAAAGCCACCTCTTAAATATTCCAGAGCCCATCTATGTGGGAGTCTGTTCTGGA
 2148900 CTCTGCTCTTTCCATTGGTCTATTTTTTGTGTCTTTGAGTTAACACAACCTTGTCTTAA
 2148960 TTAATAACCTTATAATTTTAGTATCTTTGGGAGAACTCTGTCTCTTTTTATCAAGT
 2149020 CATTGGTTCCCTTTGGCCCTTTTTATTTCTACATTAATTTTATACTCAATTTGTGAAGCT
 2149080 CCTCCCAAAATATGGGGAGGCATTTGATTAGAATTACACATATTAGCTTGGGAAGAATAA
 2149140 CATCTATATTTATATATATATATATATAGATTTGATGATCTGATCCATGAACTTGGAA
 2149200 GTTTCTTTTCATTAATTTTTTGTCTCTTTAGTGACAGTGACAATGTTTTATCATTTTCCC
 2149260 TGCAGAGGCTTTTCATGGTTTTTGTATATGTATTCCCAAGTATTTGTTTCAGTACTATC
 2149320 GTACTTGGCATGCTTTCTTTAATTTCAATTTTGCATGGTTGCTTTCATGGGAGACTTAAAC
 2149380 ATTTTCAGTGGTGATATTTGGCTATAGCATTATAGGTGATCTTTATACTTTTCTAAATTT
 2149440 TCTGTGGCCACGGGAATAATAAGACACTTTTCTTGCACAGAAAAAAAACTTTGGG
 2149500 AAGTATTTCTCACAGCTAAGATCTGATAGTTTACGCAAAGTTGGCAGGCACAGGCTACAG
 2149560 AAAGCTCTGGGTGCTGTTGTTTGGAGCTGCTGGTTCAAGGACAAATTCACAAGATTTGG
 2149620 AAACAGAGGACCAAGTGTGTAAGGACGAGGGAACATGTTATAACATTGAAGCACCTGA
 2149680 GCTGGAAATTTCTGAGCCCTCAGAGATAAATTTCTCAGCTCCTCCCTGCCGAGAAAAACA
 2149740 AAATAAAAAGAGTTAATGTTTAGCCAACAGAAATGAGAGTGAAGTTACAGAAGAAAT
 2149800 AGGGCACAGCTGGAAGTGTCAAATGAGGAAACGTTCAATGTCAAGTTTGGATCCAGG
 2149860 GTTGATGGGTGAACTCTGCACGCTCAGCTTCATGTTAGGTCTCCAGCTCAACGAGGGGG
 2149920 TGAATTTTGGTGTACTGCTGTCTTTGAGGCCCTGCGTGAATCTCCACACCTCCCTCT
 2149980 TTTCTCATACAAAATCCCCTCCTTGCACACCCCTCACGCTGTCTGTGAGGTGCCAGGGC
 2150040 CCCCTCCAGCCAACCGCAGGCCAGTATTGCCCTCCCAAACTTCCCTCAGGCAGAT
 2150100 CAAACCCAGGCTCTGGAGTCACACTGCCTGGGCTCAAATCCTGCTTCTGAATCTTATGA
 2150160 GACATCAAGTCACTTACCTAACCTCTTGGTGAACAGGCTCTCTCTGTCAAATAGGTGT
 2150220 GCATGCTTTGGGAGGCCAAGGCAGGAGATTGCTTGGCCAGGAGTTCAAGACCAGCCT
 2150280 GGGTAACATAGTGAGACCCTGTCTCTATAAAAATAAAAAAATTAGCCAGGTGTGGTGGCT
 2150340 CATGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGAGGAAGAATTCCTTGAGCCCACGAA

图2

2150400 GTCAAGGCTGCAGTGAGCCATGATCAAGCCGCTGCACTCCAGCCTGGGCAACAGAATGCG
 2150460 ACCCTGTCTCAGAAATAAGTAAATAAAATAACATAAAATAAATTTGGGTGCAATTTGTGG
 2150520 CTCTTAGTGTACAGAGAGGACTGAAGGAGCTAATGGATGATGGATTTATGGCAGTACCT
 2150580 CACATATAGCTTATCCTAAAGGAAGTTAGAGCTTATTATGATGATTATTCAAAAATTTT
 2150640 ATCAAAGGTCCTGCCCTGGCCATGTTCTGAGCTAAGTGCTGGGGATGCAAAGATGAGCAA
 2150700 GAGACTCCTCAGGGACAATTTGCTGATGAGATAACAGGCCTATTTATGAGAGGTCCAAT
 2150760 CAATACAGTTCTATTTATCTTATAATTATCCAATAAATGATATAATAATTTATTAGAGGG
 2150820 CCAATAAATCTGATGGCAGGAGCCTGTGGGGGGTAATGGCCAGGTCTCACTATTTGTGCC
 2150880 AGGCTGGTCTTGAACCTCTGGCCTCAAGCATTCTTCTGCCTCCACCTCCCAGCATGCTG
 2150940 CAATTACAGAGGCATGAACAACCTGCACCTGGCCATAAAATTTTATGTTAATAAAAAATGCG
 2151000 ATGTATTTGAGGACTACGACATGATGCTTGAATATCATACTGTATCGGGGAACCCAGCC
 2151060 CCCGATATTTCAATGTAGGTTCTTTTCTATTTTCCCAAGTGTCGGCTGGTCTGAGAAAT
 2151120 AAAGGGAAAGAGTACAAAAGAGATAAATTTTAAAGCTGGGTGTCCAGGGCAGACATCACA
 2151180 TGTCCGGCAGGTTCTGTGTGCCCCCTGAGCCATAAAACCAGCAAGTTTTTTATTAGCAATC
 2151240 TTCAAAGGGAGGAAATGTACATATAGGGTGTGGGTACAGAGAACACATGATTCAAAGGCC
 2151300 GACAAAAGATCACAAGGCAGAAGGTGAGATCACAAGGTGAGGGCAAAACTAGAA
 2151360 TTACTAAGGAAGTTTCATGTTCCACTGTGCATGCATTGTTCATTGATAAACATCTTAACAG
 2151420 TGTTCAGAGCAGAGAACCAGTCTGACTAGAATTCGCCAGGCTGGAATTTCTAATCCTA
 2151480 GCAAGCCTGGGGGTGCTGCAGGAGACCAGGGCGTGTTCATCCCTTATCTGCAACTGGAT
 2151540 AAGGCAGACACCCCCAGAGCGGCCATTTTAGAGGCCCCCGGGAAATGCATTTCTTTCCCA
 2151600 GGGCTGTTAATTATTAATATTCCTTACTGGGGAAGAATTCAGGGATATTTCTTTACCT
 2151660 GTTTTTGGTAATAAGAGAAATATGGCTCTGTCTTGCCTGGCTCCCAGGCAGTCAGACCTA
 2151720 ATGGTTATCTCCCTTGTTCCTTCAACATCGCTATTATCTGTCTTCTTTCAAGGTGCC
 2151780 AGATTTCAATTTGTTCAAACACACATGCTTTACGAACAATTTGTGCAGTTAACGCAATCA
 2151840 TCACAGGGTCTGAGGCAACATACATCCTCAGCTTATGAAGATGACAGGATTAAGAGATT
 2151900 AAAGACAGACATAGGAAATTTAGAGAGTATTGATTGAGGAAGTGATAAATGTCATGAAA
 2151960 TCTTCACAAATTTATGTTCTTCTGTGATGGCTTCAGTAGGTCCCTCCGTTCCGGGTCCCTG
 2152020 ACTTCCCAACAACATCACTGTATACCTGAAATTAGCATTGATTCTAATTTCTTGGTTCACAC
 2152080 GTCATTCAGACATAGGATCTTTCGGTGGATTTAAGAAGTGCCCTCCCTCCCTATTCTCAGC
 2152140 CATGTGACTCCCAGAATCCTATGAAATTAAGATCTTGTGTTTGCCATAGGAGACTTCT
 2152200 ATTCACCATCTTTGTTCTCTCCCACAAATGGCGAGGCCCTCCAGTCTCCACATGACAGCT
 2152260 TTGTACTAAAATCAACCTTACTCTATAGAACATGCATGATTGCAGCAGGACTACTATGAT
 2152320 CTTGGTTTGATGAATTAGTTAGGATAACATTAGCTGCTGTAACAAACAGACCCCAAGCT
 2152380 GCAGCATGACTCAGATGCAATATAAGTGTCTTTTTCACTTATATCAAGCAAGAATGACCA
 2152440 GATTTTTCTGATTTTTTTTTCCGCTGTGCTAATGTAGGGAGAAGTTGTTGGAGGTACAGTC
 2152500 ACAGTTACAGCAACCATCTATGTTTGGGAGCAAGGATGCTGGAAATAGAATCCAGCATA
 2152560 CTTGTAGCTTGTCCATAATTACAGACACTTTGCATTTACTGAACCTGAATCCTTACTGAT
 2152620 GAAGACCAAAGACTGTAGCATGCCTGAAGGGACAGCCTCAGAACTGTGATGCCTGTCCC
 2152680 TCTCCTGGTTTGGGTTTGTGCTGACCACAGGCAAACCCACTGAACTCAGGATCACTGCAT
 2152740 AAAGTGACGTATTAAGCCTAGTGCCAGGATCTTTAGTGTGTTGCAGGAAGGTCTCATGATT
 2152800 TTTCTGCTAACTCAGCTAATAGGTAGTCTCTGTCCCTTCAAGTTACAAAACACATCCTTC
 2152860 ATTCAGGAATTTGAATGTTTCACTCCTTGATATTTTATCAACCCCTTCACTCTGTGGTCAA
 2152920 GGGCAGACGTCTCTCCAGTTAAGACTGGAAGTTAGCAGCCTGCCTCCAGGGATGGGGTT
 2152980 GTGGTTGCCTTCTGCTCTCTCTGTTTCTCTGGGAAGCAGCAGAAATCATTCCATGGGAGGA
 2153040 CTAGAGCAGTTCTTTCTTGAGAGAAGAGATTTACTCCTTCCAAGTGTATTGGTTAGTGAT
 2153100 TGCTACATAACAAACTACCCCAAACCTCTCAGTAGCTTAAAACAACCTGTGAAGTGATTGA
 2153160 TGCTCATGTGAGCATGGGTTGGTTGATCCAGGCTGGGCTCAGCTGGGCACCTCTGTATAT
 2153220 GCTGTGGGTTCTCCTGAGCTCAACTCCTTCCCTGCCAGTTGCATTTAAGGCTGTTCTGTGC
 2153280 GTGTCTTCTGGAGCCAGGCTGAATGGGCTTTGGAGATGGCAGAAGCACAAGCGAGTAAA
 2153340 CAGAGACATGAAGCCATTTAAGGCCCTCAGCCAGCACTGACGTGCCGTCACTTCTCAC
 2153400 ATTCCACTGGCCATGCAAGTCACATGGCTGAGCAAACCTCAAGGACTTGGGAAGTAACCAT
 2153460 CGTCTTTAGTGGGAGGAACTACAACATCCCATGGCAAAGCATGGATCCAGGAAGCAGTGA
 2153520 AATGGGGGCCAGTGACTCAGTTTACCACACTGAGGTCTGGCAGATGGCTAGAAGTGGCGC
 2153580 TTTCTCTTGAGGATTGGGGGAGAGGGTGTGTTTATGGATTCTACAGCAATCCCAGGCCTG
 2153640 GGAACCTCTGTAAGTCCCTTTCCAGGGCCTTACATCTCTCTCTACATGGTCCCCTCT
 2153700 AACTCCTGCCTCATCTAGATTTCGTACCACACCCAGCTTCTTGTGAGCATCTCTTGTCT
 2153760 GCCAGAGGGCCCTGTGAGACAAGCCCAAGATGACCCCATGCCAGAATTTGTCAACCATG

图2

2153820 TGATTACATCTGGACCAGAGGAAATGCCTCCCAAATGAAGCCCATCCTGCCCCCATCAG
 2153880 GCAACTACAGGCCACTTCAGCTTTCTTGGGTGTAAGGCAGACCTCAGAATCTCTGTGTCT
 2153940 CCCAGCTAGATGGAAAGCTTTCCAAGGGGTCTTGGGAAGCCAGCTGGATTGAGGCAAGG
 2154000 AATATCACACCCCATCCATCTCCAAAGGGGAAGCAACACATCACCTGACACAGTTCTC
 2154060 TCCAGGGCAATCTCTTGGCCAAACATTTGCTCCTCTCCACACTCCAACCCCTTTATGTATT
 2154120 CTTAACATGACTGAGGAGCCCTTTATAAAATTCTGCATTTGGGAGTTTGTTCATATTCT
 2154180 GTTTGGTTCCTGGAGTTACCTACACAAGAGCCTCAGGCAATGAGATTTTATTTATCAAAC
 2154240 CATTACCAGATGTCAGGTGCTGTCCCAAACACTTTGTAATGTTAACCCATTTAATCCTC
 2154300 ATATAAATCCTATGAGGTAGGTGCTATTAACATCTTTATTTTAGAGATAGGGAAGCTGAG
 2154360 GCACAGAGAGGTTAAGTAATAGCCCAAAGTCACACAGAAGGCAACTGCTTCTCAACCA
 2154420 GGCAAGAAGAGTCTTTTTTATGATAAATAAGGTGGAAGAAGGCAAGGGGGAGATAAGCAA
 2154480 GAAGATGATAATGATGATGGTGGCTTCCCTGAGTTACTGTGCTTAGCACTTAGTGTGCG
 2154540 TGGCCTTGCCCTGCTGCTTTTGGACAGGTATGTCAGACTCTCCCATTTCCGACATG
 2154600 AATAAACTGAGGCTGATATAGTTTGAATGTATGTCACCACCAAATCTCATGTTGAAATG
 2154660 TAATCCCAGTGTGGAGGTGGGGCCTGTTGGGAGGTGATTGGATCATGGGGGTGGATTT
 2154720 CTCATGGGTAGTTTAAACACCATCGCCTTGGTGTATCCTTATGATAGTGAGTACATTCTC
 2154780 ATGAGACCTGGTGTTTAAACTGGGTGGCACCTCCTCCCACTTCTCTCACTCCTGCT
 2154840 TTCGCCACATGATATGCCTGCTCCCTTTGCCCTTCTGCTATGATTGTAAGCCTCCTGAG
 2154900 GCCTCCCCAGAAGCCAAGCAGACGTCAGTACCATGCTTCTGTAAAGCCTGCAGAACCAT
 2154960 GAGCCGATTAACCTCTTTTTCTTTATAAAATTACCCAGTCTCAGATATCCCTTAATAGCA
 2155020 ATGCAAGAATGGCCTGATACAGAGGCTTTGAGAGGTCAAGTGACCTGCCAAAGGGCACAC
 2155080 CACTGATAAAGGACAGATGTGGGATTTGAACCCACTTTGTTAGCCCCAGTGCAAGGCC
 2155140 CCTTCTCCTCATGCTCACTGCCACTGCGGAGCTGGGCACCTTGGGACACACCCCTAGGGAG
 2155200 CTGTGCTCTGCTGGGGCCTTCTCATGCCCAAGTCCCTGCTGCCAAGGCCGGTGTCTC
 2155260 TGAGTCTCCTACAGCCCCTCCTCAGCCTCACTGGCCTCAGTCATCTTGGTTCAGGGAAG
 2155320 GACTAAGGTCCCCTTTGGTCCCTGCCAATCTGCACCCCAACCCAGTGTGACCTCAAGAG
 2155380 CCTGACTCTGGCCTTTTACAGAATAAATCTGAACAAAATCAGGGTTCATTTAATAGCAA
 2155440 CAGGCTGCTGATGCAGACCTTATCAACTCCATCAAACCTGTGTTCTTTCAAATGTTACGCT
 2155500 CCCCTGGGGGTGTCCACACCCCTGACGTCACACATTCACTCAGTGAAGCCCATATTCATT
 2155560 CGGGGAGCTCTTCTTTCTCTCTCTCTAACAACACACTCTCCTTGGTTAGCTGGCTGCTG
 2155620 ATCAACATTTCTGGATATACTGGTTTTACAGGAAAATATGATGGTTGGCTTCCAATCCCAG
 2155680 ATTTTTCACTGATGGGTTCCATATTTACACCATCTTGGCCACAGTCTCTGGGTCAACATT
 2155740 TCCACACATACCCACACACCATAAAGAGAGGCTTTTTCTGAGTCTTCTGTTTATGACGTC
 2155800 TGGAATTGTATTTGCTTTTGTGTTTGGGGCATCTGGGCAGCTCATTCCACTAATAGGCAT
 2155860 AACCATAACCATTGCAGTCTCCACTTACTGACATTTACAACCTTTCCAGGCACATGCTAGG
 2155920 GACCTTACATTCAATTTTCAATTTATTCTCATATCACAACTTGTGAGGTGGAGGAGCA
 2155980 TGATGGAAGGAGGAAGGAGCTAAAAGCAAAGAATCTGGAGTCCGACTGCCTGGGTTCAA
 2156040 TCCTAGCTCTACCAATTTCCAGCTCTGTAACATCGAGCTAATTTCTAACCCTCTATGTC
 2156100 CATTTCCCTATCTCTAAAAGGAAGCTGACAATAGCATCTATCTCATAGGATTTGTACGAA
 2156160 GATTAAATGAGTCAATATTTATAAAGTGTTCGGAATGATACCTGACATCTGGTAATGGTT
 2156220 TGATAAATAAATCCATTTAATGATGAGGAAACAGGCTCAGAAGAGGGCGCTCATTTGC
 2156280 TCATGTGGTACAGATAGGTTCCAGACTCAAACCTCAAGACCATCTGACTCTAAAACATCTA
 2156340 AAAGTGTGGCCCGCATCTCTTCAATTTGACAGATAATAAACTGAGGCTCAGAGAAGCT
 2156400 AAGTGACTCGCCTGGGACTGCACAGCAAATCAAGACAAATAAGACCTTAGGCTCCTGAC
 2156460 TGCCAGAGTGGAGATGCTTCTATAGGCTTTTCTCACTGATGCTCTCTGGGCAGACAGGCT
 2156520 CCTCAATATGAGAGTGACACACACTCCTTTCTTCAATTTTCCAGGTAACCTCACACTGGTT
 2156580 GGCAGAAGGAACATAACCAATAATGATGAACATGCGGTGAATTTGCAACAGACAAGAGG
 2156640 AGCCTCATTATCCTATAGTTTCCAGGTTGCTTAGGGAGGCAGAAATCACAGCAAGGAAAA
 2156700 CCTTCAATAATAAACAGACGCTCTCAAAAATTAATTGCAACCCAACTCTCTCTACTTT
 2156760 AAAATTAGCATCTATTTCCAGCTCTGCTTCAATGCCCCATATGAATACATGTGAATCC
 2156820 CTCCCTCTCTCCTCCCTGTCTCCTTCTCTCTCTCTCTGTCCTCATTAATAAATAAAT
 2156880 TTAAGAAAAAATAACAAGGTAGATTTACACAAATAGTGGGATCTCAGTCTTGAGTTAGCT
 2156940 GTGTATGACTGAAAAGGATGCTGTGGTTAATAATTATCATAAAAAACAATGACATGGCCGG
 2157000 GCACAGTGGCTCACGCCGTGTAATCCCAGAACTTTGGGAGGCCGAGGCAGGCAGATCACTT
 2157060 GAGGCCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAACCTGCATCTCTACTAAAAAT
 2157120 ACAAATAATAGCCGGCATCAATGGCCAGCCCCTGTAATCCAGCTAATCAGGAGTCTGA
 2157180 GGCAGGCGAATCACTTGAACCTGGGCTGGTGGTTGACAGTGAAGCCAGCTCACACCACT

图2

```
2157240   GCACTCCAGCCTGGGCGACAGAGTGAGACTACATCTCAAAAAACAAAAACAAACAAGCA
2157300   AAAAAAACCCACAGTAACACAAAAGTAATAAAACTGCTGCTATTTACTCAGTGCTTATC
2157360   TGATGCCAGCCACTTTGCTAAGCCTATGAATGCATTATTTCCCGTTGCTACAGATGAGA
2157420   GAATTGAGGTTGAGACAGGTTGAAATCATTGCTCCCAAAGTCACACAACTGGTGAGTGGC
2157480   AGAGCTGGGATGCAAACCCTAAACTGCCAGCCCTCAAAGCCTGTGCTCTTAATCTCCACC
2157540   CTGCTGTGCTTCCTTGTCCATTTAATTAAGCTCCACAGGCACACATTCACGCCCTCCTT
2157600   TGCTGTACAATCCCAGGCAAGTCGCTCAGCTTCTCTGAGCCTCAGTTTCATAATCTGTCA
2157660   AATGGAGGTAACACAAATAATTCCTAGTTGTGACCAAGAATCATCATAGAAATCTGCCAT
2157720   TTCCAGCCTATTGTGCAATTCCTCAAGCACTGTGACTCCAAGTGGCATCAGCTCCTGGAA
2157780   GAACACACTGTCTTACTGTTGTTTCCTCCTTTGTCAACTGATCCCCCTTGAACCTCACT
2157840   CTACCTCTGCTCTCAATGCCCATCTACTGCCACCTGATTAATAAAATCTTTTTGAAA
2157900   ATCATAAGTGTCATGAGTAAGGTTTCTTGGTGTGATGTAGAAGAACAAAACAGAATTGT
2157960   GAAATGAGAATCACTGCAGCTATCATGAAGTCCTGCCTAC
```

图2

SEQ ID NO:20

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQN

SEQ ID NO:21

1 SVSRRDKSKQ VWEAVLLPLS LLMMDLRNTP AKSLDKFIED YLLPDTCFRM QIXHAIDIIC
 61 GFLKERCFRG SSYPVCVSKV VKGGSSGKGT TLRGRSDADL VVFLSPLTTF QDQLNRRGEF
 121 IQEIRRQLEA CQRERAXSVK FEVQAPRWGN PRALSFVLSS LQLGEGVEFD VLPAPDALGQ
 181 LTGXYPNPQ IYVKLIEECT DLQKEGEFST CFTTELQRDFL KQRPTKLKSL IRLVKHWYQN

SEQ ID NO:22

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW PSHQAWWVLS RLGAEEG

SEQ ID NO:23

1 SVSRRDKSKQ VWEAVLLPLS LLMMDLRNTP AKSLDKFIED YLLPDTCFRM QIXHAIDIIC
 61 GFLKERCFRG SSYPVCVSKV VKGGSSGKGT TLRGRSDADL VVFLSPLTTF QDQLNRRGEF
 121 IQEIRRQLEA CQRERAXSVK FEVQAPRWGN PRALSFVLSS LQLGEGVEFD VLPAPDALGQ
 181 LTGXYPNPQ IYVKLIEECT DLQKEGEFST CFTTELQRDFL KQRPTKLKSL IRLVKHWYQN
 241 VWPSHQAWV LSRGAEEG

SEQ ID NO:24

1 SVSRRDKSKQ VWEAVLLPLS LLSMMDLRNT PAKSLDKFIE DYLLPDTCFR MQIXHAIDII
 61 CGFLKERCFR GSSYPVCVSK VVKGGSSGKG TTLRGRSDAD LVVFLSPLTT FQDQLNRRGE
 121 FIQEIRRQLE ACQRERAXSV KFEVQAPRWG NPRALSFVLS SLQLGEGVEF DVLPAFDALG
 181 QLTGXYPNP QIYVKLIEEC TDLQKEGEFS TCFTELQRDF LKQRPTKLKS LIRLVKHWYQ
 241 NVWPSHPACW YLYIFI

SEQ ID NO:25

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNVW PSHPACWYLY IFI

SEQ ID NO:26

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCK KKLGLPPOY ALELLTVYAW
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYDFKX PIIKYLRRQ LTKPRPVILD
 301 PADPTGNLGG GDPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLMRQ LREVRSLAQQ
 361 HQLTSGGNGI QAQWTLKPVL MSLC

图3

SEQ ID NO:27

```

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCK KKLGLKPPQY ALELLTVYAW
241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRO LTKPRPVILD
301 PADPTGNLGG GDPKGWRQLA QEAEAWLNXY CFKNWDGSPV SSWILLKAT VQTMRTIPG
361 XIRNMVTLFH MSTLISLIDP AHSRQHPHR QKRTGPAPSS ECQCILGERA PVLSGPVPSF
421 SGGTLDPEXT KLLSELVYNP GQNPGLLTPG LLCPLSYHR

```

SEQ ID NO:28

```

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCK KKLGLKPPQY ALELLTVYAW
241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRO LTKPRPVILD
301 PADPTGNLGG GDPKGWRQLA QEAEAWLNXY CFKNWDGSPV SSWILLIKLR LREAK

```

SEQ ID NO:29

```

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCK KKLGLKPPQY ALELLTVYAW
241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRO LTKPRPVILD
301 PADPTGNLGG GDPKGWRQLA QEAEAWLNXY CFKNWDGSPV SSWILLVNLV LVGRRNYP II
361 SEHAVNLQQT RXASLSYSFQ VA

```

SEQ ID NO:30

```

1 AESNSXDDDET DDPRXYQKYG YIGTHEYPHF SHRPSTLQAA STPQAEEDWT CTIL

```

SEQ ID NO:31

```

1 GCTGAAAGCA ACAGTRCAGA CGATGAGACC GACGATCCCA GGASGTATCA GAAATATGGT
61 TACATTGGAA CACATGAGTA CCCTCATTC TCTCATAGAC CCAGCACACT CCAGGCAGCA
121 TCCACCCAC AGGCAGAAGA GGACTGGACC TGCACCATCC TCTGAATGCC AGTGCATCTT
181 GGGGGAAAGG GCTCCAGTGT TATCTGGACC AGTTCCTTCA TTTTCAGGTG GGACTCTTGA
241 TCCAGAGARG ACAAAGCTCC TCAGTGAGCT GGTGTATAAT CCAGGACAGA ACCCAGGTCT
301 CCTGACTCCT GGCCTTCTAT GCCCTCTATC CTATCATAGA TAACATTCTC CACAGCCTCA
361 CTTTATTCCA CCTATTCTCT GAAAATATTC CCTGAGAGAG AACAGAGAGA TTTAGATAAG
421 AGAATGAAAT TCCAGCCTTG ACTTTCTTCT GTGCACCTGA TGGGAGGGTA ATGTCTAATG
481 TATTATCAAT AACAATAAAA ATAAAGCAA TACCATTTA

```

SEQ ID NO:32

```

1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
121 VQAPRWXNPR ALSFVLSSLQ LGEGVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCK KKLGLKPPQY ALELLTVYAW
241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIEKYLRRO LTKPR

```

图3

SEQ ID NO:33
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSLQ LGEVVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCCK KKLGLKLPPOY ALELLTVYAW
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIKYLRRQ LTKPRPVILD
 301 PADPTGNLGG GDPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLVRPP ASSLPFIPAP
 361 LHEA

SEQ ID NO:34
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSLQ LGEVVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCCK KKLGLKLPPOY ALELLTVYAW
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIKYLRRQ LTKPRPVILD
 301 PADPTGNLGG GDPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLAESN SXDDDETDDPR
 361 XYQKYGYIGT HEYPHFHRP STLQAASTPQ AEEDWTCTIL

SEQ ID NO:35
 1 MMDLRNTPAK SLDKFIEDYL LPDTCFRMQI XHAIDIICGF LKERCFRGSS YPVCVSKVVK
 61 GGSSGKGTTL RGRSDADLVV FLSPLTTFQD QLNRRGEFIQ EIRRQLEACQ RERAXSVKFE
 121 VQAPRWXNPR ALSFVLSLQ LGEVVEFDVL PAFDALGQLT GXYKPNPQIY VKLIEECTDL
 181 QKEGEFSTCF TELQRDFLKQ RPTKLKSLIR LVKHWYQNCCK KKLGLKLPPOY ALELLTVYAW
 241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFKX PIIKYLRRQ LTKPRPVILD
 301 PADPTGNLGG GDPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLTQHT PGSIHPTGRR
 361 GLDLHHPUNA SASWGKGLQC YLDQFLHFQV GLLIQRXQSS SVSWCIIQDR TQVS

SEQ ID NO:36
 GAGGCAGTTCTGTTGCCACTCTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACACGTGTTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTCC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTGTCCAAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACACCCCTCAGAGCCGATCTGACGCTGACCTGGTTGTCTTCCCTCAGTCCTCTCACCCTTTTCAGGATCAGTT
 AAATCGCCGGGGAGAGTTTCAATCCAGGAAATFAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTGCTCCAGCTCGGGGAG
 GGGGTGGAGTTCGATGTGCTGCCTGCCTTTGATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCTCCACCTGCTTCACAGAACTACAGA
 GAGACTTCTGAAAGCAGCGCCCAACAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAATTTGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA
 AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGAATTAGTCATAAACTACCAGCAACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAACCCCATFATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCTGTG
 ATCCTGGACCCGCGGACCCCTACAGGAAACTTGGGTGGTGGAGACCCAAAGGTTGGAGGCAGCTGCGACAAGAGGC
 TGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGGTCCCAGTGAGCTCCTGGATTCTGCTGGTGAGAC
 CTCTGCTTCTCCTGCCATTCATCCCTGCCCTCTCCATGAAGCTTGAGACATATAGCTGGAGACCATTCTTTCC
 AAAGAACTTACCTCTTGCCAAAGGCCATTTATATATAGTACAGGCTGTGCTCCATATTTTACAGTCATTTTGT
 GTCACAATCGAGGGTTTCTGGAATTTTACATCCCTTGTCCAGAATTCATTC CCTAAGAGTAATAATAAATAATCT
 CTAACACCATTTATTGACTGTCTGCTTCCGGCTCAGGTTCTGTCTTAAGCCCTTTAATATGACTCTCTCATTAAAT
 A

SEQ ID NO:37
 GAGGCAGTTCTGTTGCCACTCTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACACGTGTTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTCC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCCTGTGTGTGTGTCCAAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACACCCCTCAGAGCCGATCTGACGCTGACCTGGTTGTCTTCCCTCAGTCCTCTCACCCTTTTCAGGATCAGTT
 AAATCGCCGGGGAGAGTTTCAATCCAGGAAATFAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTACTGAGTTGCTCCAGCTCGGGGAG
 GGGGTGGAGTTCGATGTGCTGCCTGCCTTTGATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCTCCACCTGCTTCACAGAACTACAGA

图3

GAGACTTCCTGAAGCAGCGCCCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA
 AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCATAAACTACCAGCAACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAAACCCATTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCTGTG
 ATCCTGGACCCGGCGGACCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTGGCACAAAGGC
 TGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGGTCCCCAGTGAGCTCCTGGATTCTGCTGGCTGAAA
 GCAAAGTRCAGACGATGAGACCCGACGATCCAGGASGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCCT
 CATTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAGAGGACTGGACCTGCACCATCCT
 CTGAATGCCAGTCTTGGGGGAAAGGCTCCAGTGTATCTGGACCAGTTCCTTCATTTTCAGGTGGGACTCTT
 GATCCAGAGARGACAAAGCTCCTCAGTGAGCTGGTGTATAATCCAGGACAGAACCAGGTCTCCTGACTCCTGGCCT
 TCTATGCCCTCTATCCTATCATAGATAACATTTCTCCAAGCCTCACTTCATTCCACCTATTCTCTGAAAAATTTCC
 TGAGAGAGAAACAGAGAGATTTAGATAAGAGAAATGAAATTCAGCCTTGACTTTCTTCTGTGCACCTGATGGGAGGGT
 AATGTCTAATGTATTATCAATAACAATAAAAAATAAGCAAATACCATTTATTGGGTGTTTATTAACCTCAAGGCACA
 GAGCCAAGAAGTACAGATGCATATCTAGGGGTATTGTGTGTATATACATTGATTCAACAAGAAATATTTATTGAG
 CACTT

SEQ ID NO:38

GAGGCAGTTCTGTTGCCACTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACAGTGTTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTTC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCGTGTGTGTGTGTTCCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTTGTCTTCTCAGTCTCTCACCCTTTTCAGGATCAGTT
 AAATCGCCGGGAGAGTTTCATCCAGGAAATTAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTAAGTTCGCTCCAGCTCGGGGAG
 GGGTGGAGTTCGATGTGCTGCCTTGTATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACTGCAGAAAGAGGGGAGTTCTCCACTGCTTCACAGAACTACAGA
 GAGACTTCTGAAGCAGCGCCCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTTGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA
 AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCATAAACTACCAGCAACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAAACCCATTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCAGGCCTGTG
 ATCCTGGACCCGGCGGACCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTGGCACAAAGAGGC
 TGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGGTCCCCAGTGAGCTCCTGGATTCTGCTGACCAGC
 AACTCCAGGCAGCATCCACCCACAGGCAGAAGAGGACTGGACCTGCACCATCCTCTGAATGCCAGTGCATCTTGG
 GGGAAAGGGTCCAGTGTATCTGGACCAGTTCCTTCATTTTCAGGTGGGACTCTTGATCCAGAGARGACAAAGCTC
 CTCAGTGAGCTGGTGTATAATCCAGGACAGAACCAGGTCTCCTGACTCCTGGCCTCTATGCCCTCTATCCTATCA
 TAGATAACATTTCCACAGCCTCACTTCATCCACCTATTCTCTGAAAAATTTCCCTGAGAGAGAACAGAGAGATTT
 AGATAAGAGAATGAAATTCAGCCTTGACTTTCTTCTGTGCACCTGATGGGAGGGTAAATGTCTAATGTATTATCAAT
 AACAAATAAAAAATAAAGCAAATACCATTTATTGGGTGTTTATTAACCTCAAGGCACAGAGCCAAGAAGTACAGATGCA
 TATCTAGGGGTATTGTGTGTGTATATACATTGATTCAACAAGAAATATTTATTGAGCACTT

SEQ ID NO:39

GAGGCAGTTCTGTTGCCACTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACAGTGTTTCCGCATGCAAATCAACCATGCCATTGACATCATCTGTGGGTTC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCCGTGTGTGTGTGTTCCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTTGTCTTCTCAGTCTCTCACCCTTTTCAGGATCAGTT
 AAATCGCCGGGAGAGTTTCATCCAGGAAATTAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGTCCACGCTGGGGCAACCCCGTGCCTCAGCTTCGTAAGTTCGCTCCAGCTCGGGGAG
 GGGTGGAGTTCGATGTGCTGCCTTGTATGCCCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACTGCAGAAAGAGGGGAGTTCTCCACTGCTTCACAGAACTACAGA
 GAGACTTCTGAAGCAGCGCCCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTTGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA
 AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCATAAACTACCAGCAACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAAACCCATTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCTGTG
 ATCCTGGACCCGGCGGACCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTGGCACAAAGAGGC
 TGAGGCCTGGCTGAATACCCATGCTTTAAGAATTGGGATGGGTCCCCAGTGAGCTCCTGGATTCTGCTGTCTGAAAG
 CAACAGTRCAGACGATGAGACCCGACGATCCAGGASGTATCAGAAATATGGTTACATTGGAACACATGAGTACCCCTC
 ATTTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAGAGGACTGGACCTGCACCATCCTC
 TGAATGCCAGTGCATCTTGGGGGAAAGGCTCCAGTGTATCTGGACCAGTTCCTTCATTTTCAGGTGGGACTCTTG

图3

ATCCAGAGARGACAAAGCTCCTCAGTGAAGCTGGTGTATAATCCAGGACAGAACCCAGGTCTCCTGACTCCTGGCCTT
 CTATGCCCTCTATCCTATCATAGATAACATTCTCCACAGCCTCACTTCATTCCACCTATTCTCTGAAAATATTCCTT
 GAGAGAGAACAGAGAGATTAGATAAGAGAATGAAATCCAGCCTTGACTTCTCTGTGCACCTGATGGGAGGGTA
 ATGTCTAATGTATTATCAATAACAATAAAAAATAAGCAAATACCATTTATGGGTGTTTATTAACCTCAAGGCACAG
 AGCCAAGAAGTACAGATGCATATCTAGGGGTATTGTGTGTATATACATTGATTCAACAAGAAATATTTATTGAGC
 ACTT

SEQ ID NO:40

GAGGCAGTTCGTGTGCCACTCTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACAGTGTTCGCGATGCAAATCAACCATGCCATTGACATCATCTGTGGGTTC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCTGTGTGTGTGTCCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTGTCTTCTCAGTCTCTCACCACCTTTTCAGGATCAGTT
 AAATCGCCCGGGAGATTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGGCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG
 GGGGTGGAGTTTCGATGTGTGCCTGCCTTTGATGCGCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCCTCCACTGCTTCACAGAACTACAGA
 GAGACTTCTGGAAGCAGCGCCCAACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGTAA
 AACACATTTCAACACAGCCAGGGATTCGAGACGCTTGGAAATAGTCAATAAACCTACCAGCACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAAACCCATTTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCTGTG
 ATCCTGGACCCGGCGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTGGAGGCAGCTGGCACAAGAGGC
 TGAGGCCTGGCTGAATTACCCATGCTTTAAGAATGGGATGGTCCCAAGTGGTCTTGGATTTCTGATTTCTGCTG
 AAAGGCTCAGAGAGGTGAGGTCACTTGCTCAAGGACATCAGTAAACAAGTGGTGGAAATGGAATCAAGCTCAGTGG
 ACTCTAAAGCCAGTGTCTATGCTACTGTGCTAAACAGCTTCCTTGTCAATCCCACTCTCATCTGACCAATGGG
 AGACTTCTGAGCAGCTGAGTACTGGTGTGTCACACAGCTTAAACAGGGGCAAAGGACCCAGTCTTGGATCTTCCAC
 CTCCAAGCAGGAATCTGTCTGATTCCAGGGGATGATGATGTTGCAGATGGCTAGGAAGCAGACTCCCAAGGATGA
 TTAGTATGCAGGATGTTCTGGGGGAGAGCCACTGGAAACGACACTCAGGAAAGGGGGGAAGAAAGGATAGGAAGGA
 AGCATGAAAGAGAATAGGGAGAAGTGAACAGGGATGCAGAGCGAATGCCAGTTTCAGCCAACTCAAGGACAGCCCT
 GGAGCTGGAATGGCCTTTAGAGCTGCCCATGGCTTGGAGCCAAAGGCTCTCTGCAGCTGAGGCAACCCCTAAAAGGC
 CTGTGCTTGGGCACCTTGGGAAAGGCATGGCTTTGAGCCAAAGGCTCTCTGCAGCTGAGGCAACCCCTAAAAGGC
 TGACGGCTGAAGTCTGTCTGCTGACCACTGTCCAGCAGCTGGGGCTTGTAGTCTTCTCTCAAAGGGGGATCCAGA
 TGGCATGTACAGTGTCTACCGTAAATGCTCACTGAATCCAGCTGCAATGCAGGAAGACTCCCTGATGTGATCATGT
 GTCTCACCTTTTCARGCTGAAAGCAACAGTRCAGACGATGAGACCCAGCATCCAGGASGTATCAGAAATATGGTTA
 CATTGGAACACATGAGTACCCTCATTTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCCAAGGCAGAA
 AGGACTGGACCTGCACCATCCTCTGAATGCCAGTGCATCTTGGGGGAAAGGGCTCCAGTGTATCTGGACCAGTTCC
 TTCATTTTCAGGTGGGACTCTTGTATCCAGAGARGACAAAGCTCCTCAGTGGAGCTGGTGTATAATCCAGGACAGAACC
 CAGTCTCCTGACTGCTGGCCCTCTATGCCCCTCTATCTATCATAGATAACATTTCTCCACAGCCTCACTTCATTCCA
 CCTATTCTGAAAATATTCCTTGAGAGAGAACAGAGAGATTTAGATAAGAGAATGAAATCCAGCCTTGACTTTCT
 TCTGTGCACCTGATGGGAGGGTAATGTCTAATGTATTATCAATAACAATAAAAAATAAGCAAATACCATTTATTGGG
 TGTATTAACTTCAAGGCACAGGCCAAGAAGTACAGATGCATATCTAGGGGTATTGTGTGTATATACATTGAT
 TCAACAAGAAATATTTATTGAGCACTT

SEQ ID NO:41

GAGGCAGTTCGTGTGCCACTCTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTGCCAGACAGTGTTCGCGATGCAAATCAACCATGCCATTGACATCATCTGTGGGTTC
 TGAAGGAAAGGTGCTTCCGAGGTAGCTCCTACCTGTGTGTGTCCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGGTGTCTTCTCAGTCTCTCACCACCTTTTCAGGATCAGTT
 AAATCGCCCGGGAGATTTCATCCAGGAAATAGGAGACAGCTGGAAGCCTGTCAAAGAGAGAGAGCATTTTCCGTGA
 AGTTTGAGGTCCAGGCTCCACGCTGGGGCAACCCCGTGGCTCAGCTTCGTACTGAGTTCGCTCCAGCTCGGGGAG
 GGGGTGGAGTTTCGATGTGTGCCTTTGATGCGCTGGGTGAGTTGACTGGCRGCTATAAACCTAACCCCAAAAT
 CTATGTCAAGCTCATCGAGGAGTGCACCGACCTGCAGAAAGAGGGCGAGTTCCTCCACTGCTTCACAGAACTACAGA
 GAGACTTCTGGAAGCAGCGCCCAACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATGTAAG
 AAGAAGCTTGGGAAGCTGCCACCTCAGTATGCCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCGAGGGAGCATGAA
 AACACATTTCAACACAGCCAGGGATTTCCGACGGTCTTGGAAATAGTCAATAAACCTACCAGCACTCTGCATCTACT
 GGACAAAGTATTATGACTTTAAAAACCCATTTATTGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCTGTG
 ATCCTGGACCCGGCGACCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTGGAGGCAGCTGGCACAAGAGGC
 TGAGGCCTGGCTGAATTACCCATGCTTTAAGAATGGGATGGGTCCCAAGTGGTCTTGGATTTCTGCTGGTAAACC

图3

241 ERGSMKTHFN TAQGFRTVLE LVINYQQLCI YWTKYYDFIX PIIEKYLRRQ LTKPRPVILD
 301 PADPTGNLGG GDPKGWRQLA QEAEAWLNYX CFKNWDGSPV SSWILLAESN SXD

SEQ ID NO:51

1 MMDLRNTPAK SLDKFIEDYL LPDKCFRKQI NHAIDIICGF LKERCFQSS YPVHVSQVVK
 61 GGSSGKGTTL RGRSDADLVV FLSEPLTTFQD QLNRRGEFIO EIRROLEACQ REERAFSVKF
 121 EVQAPRWDPN RALSFVLSL QLGEVEFDV LPAFDALGQL TDGYKDPQI YVKLIBECTY
 181 LQKEGEFSTC FTELQRDFLK QRPTKLKSLI RLVKHWYQNC KKKLGLPPQ YALELLTVYA
 241 WEQGSMTDF NTAQEFRTVL ELVINYQQLC IYWTKYYDFE NPIIEKYLRR QLTKPRPVIL
 301 DPADPTGNLG GDDPKGWRQL AQEAEAWLNY PCFKN

SEQ ID NO:52

1 PVILDPADPT GNLGGGDPKG WRQLAQEAEA WLNYPFCFKNW DGSPVSSWIL LAESDSGR

SEQ ID NO:55

GAGGCAGTTCTGTTGCCACTCTCTCTCCTGTCAATGATGGATCTCAGAAATACCCAGCCAAATCTCTGGACAAGTT
 CATTGAAGACTATCTCTTCCAGACAAGTGTTCCTCCGCAAGCAAATCAACCATGCCATTGACATCATCTGTGGGTCC
 TGAAGGAAAGGTGCTTCCAAGGTAGCTCCTACCTGTGCATGTCTCAAGGTGGTAAAGGGTGGCTCCTCAGGCAAG
 GGCACCACCTCAGAGGCCGATCTGACGCTGACCTGTTGTCTTCTCCTCAGTCTCTCACCACCTTTCAGGATCAGTT
 AAATCGCCGGGAGAGTTCATCCAGGAAATTAGGAGACAGCTGGAAGCCTGTCAAAGAGAGGAGAGAGCATTTCCG
 TGAAGTTTGAGGTCCAGGCTCCACGCTGGGACAACCCCGTGCCTCAGCTTCGTAAGTTCGCTCCAGCTCGGG
 GAGGGGTGGAGTTCGATGTGCTGCCTTGTATGCTGCTGGGTGAGTTCGACTGACGGCTATAAACCTGACCCCA
 AATCTATGTCAAGCTCATCGAGGAGTGCACCTACCTGCAGAAAGAGGGCGAGTTCCTCACCTGCTTACAGAACTAC
 AGAGAGACTTCCTGAAGCAGCGCCCCACCAAGCTCAAGAGCCTCATCCGCTAGTCAAGCACTGGTACCAAAATTGT
 AAGAAGAAGCTTGGAAGCTGCCACCTCAGTATGCCTGGAGCTCCTGACGGTCTATGCTTGGGAGCAAGGGAGCAT
 GGAAACAGATTTCAACACAGCCAGGAATTTCCGACGGTCTTGAATTAGTCATAAACTACCAGCAACTCTGCATCT
 ACTGGACAAAGTATTATGACTTTGAAAACCCCATTTATGAAAAGTACCTGAGAAGGCAGCTCACGAAACCCAGGCCT
 GTGATCCTGGACCCGGCGACCCTACAGGAAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTGGCACAAGA
 GGCTGAGGCTGGCTGAATTACCCATGCTTTAAGAATTGAGATGGGTCCCCAGTGGAGCTCCTGGATTCTGCTGGTGA
 GACCTCCTGCTTCTCCTGCCATTCATCCCTGCCCTCTCCATGAAGCTTGGAGACATATAGCTGGAGACCATTCTT
 TCCAAAGAACTTACCTCTTCCAAAGGCCATTTATATTATATAGTGACAGGCTGTGCTCCATATTTTACAGTTATT
 TTGGTCACAATCGAGGTTTCTGGAATTTTACATCCCTTGTCCAGAATTCATTCCCCTAAGAGTAATAATAATAA
 TCTCTAACAC

SEQ ID NO:56

GCCTGTGATCCTGGACCCGGCAGACCCTACAGGAACTTGGGTGGTGGAGACCCAAAGGGTTGGAGGCAGCTGGCAC
 AAGAGGCTGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTGGGATGGGTCCCCAGTGGAGCTCCTGGATTCTGCTG
 GCTGAAAGCGACAGTGGACGATGAGACCGACGATCCCAGGAGGTATCAGAAATATGGTTACATTGGAACACATGAGT
 ACCCTCAFTTCTCTCATAGACCCAGCACACTCCAGGCAGCATCCACCCACAGGCAGAAAGGACTGGACCTGCACC
 ATCCTCTGAATGCYAGTGCATCTTGGGGAAAGGCTCCAGTGTATCTGGAACAGTTCCTTCATTTTCAGGTGGGA
 CTCTTGATCCAGAGAGGACAAAGCTCCTCAGTGGTGGTGTATAATCCAGGACAGAACCCAGGTCTCCTGACTCCT
 GGCTTCTATGCCCTTATCCTATCATAGATAACATTCTCCACAGCCTCACTTCATTCCACCTATTCTCTGAAAATA
 TTCCCTGAGAGAGAACAGAGATTTAGATAAGAGAATGAAATCCAGCCTTGACTTTCTTCTGTGCACCTGATGGG
 AGGGTTATGTCTAATGTATTATCAATAACAGTAAAAATAAAGCAAATGCC

对于以上核苷酸序列 (SEQ ID NO:31、SEQ ID NO:36-45和SEQ ID NO:55-56) :

R表示A或G

S表示G或C

对于以上氨基酸序列 (SEQ ID NO:20-30、SEQ ID NO:32-35和SEQ ID NO:46-52) :

X表示根据下表的氨基酸变体

SEQ ID NO	氨基酸位置	氨基酸
-----------	-------	-----

图3

20	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	G 或 S
21	53	D 或 N
	137	L 或 F
	184	G 或 S
22	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	G 或 S
23	53	D 或 N
	137	L 或 F
	184	G 或 S
24	64	D 或 N
	138	L 或 F
	185	G 或 S
25	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	G 或 S
26	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	S 或 G
	280	N 或 T
	330	P 或 S
27	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	S 或 G
	280	N 或 T
	330	P 或 S
	361	G 或 R
28	429	K 或 R
	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	S 或 G
	280	N 或 T
29	330	P 或 S
	31	D 或 N
	115	L 或 F
	127	G 或 R

图3

	162	S或G
	280	N或T
	330	P或S
	372	R或G
30	6	A或T
	15	R或T
32	31	D或N
	115	L或F
	127	G或R
	162	S或G
	280	N或T
33	31	D或N
	115	L或F
	127	G或R
	162	S或G
	280	N或T
	330	P或S
34	31	D或N
	115	L或F
	127	G或R
	162	S或G
	280	N或T
	330	P或S
	352	A或T
	361	R或T
35	31	D或N
	115	L或F
	127	G或R
	162	S或G
	280	N或T
	330	P或S
	397	G或R
46	31	D或N
	122	N或T
	172	P或S
	214	R或G
47	31	D或N
	122	N或T
48	31	D或N
	115	L或F
	127	G或R
	162	S或G
	280	N或T
	330	P或S
49	31	D或N

图3

	115	L 或 F
	127	G 或 R
	162	S 或 G
	280	N 或 T
	330	P 或 S
50	31	D 或 N
	115	L 或 F
	127	G 或 R
	162	S 或 G
	280	N 或 T
	330	P 或 S
	352	A 或 T

图3

由病例和对照受检者所推断的群体单倍体

所推断的单倍体	大致的群体频率	SEQ ID NO:2 等位基因	SEQ ID NO:3 等位基因	SEQ ID NO:4 等位基因	SEQ ID NO:5 等位基因
HAP1	54%	A	G	A	A
HAP2	32%	G	G	G	G
HAP3	8%	G	A	A	A
HAP4	5%	G	G	G	A
HAP5	<1%	A	G	G	G
HAP6	<1%	G	G	A	A

图4

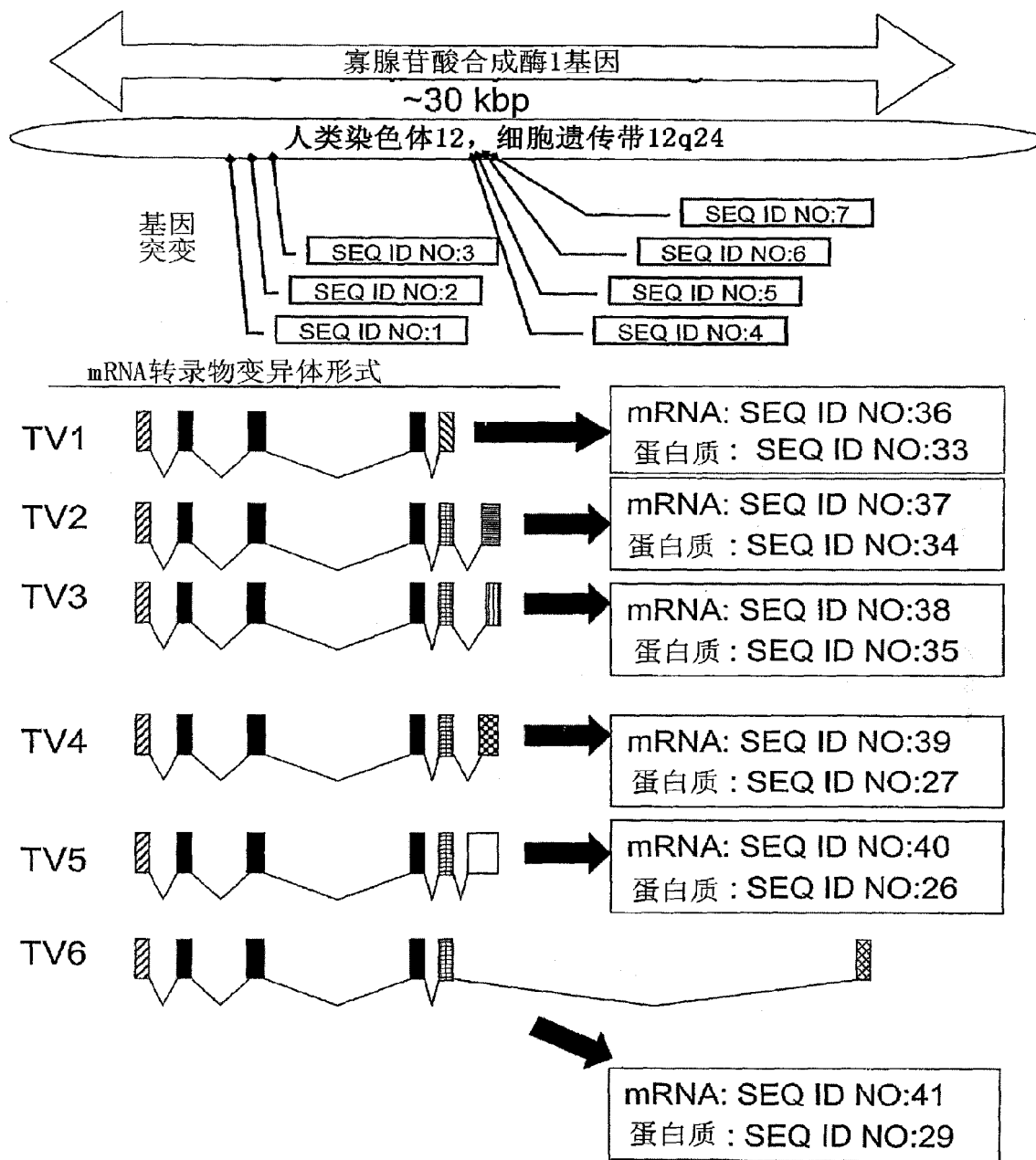


图5A

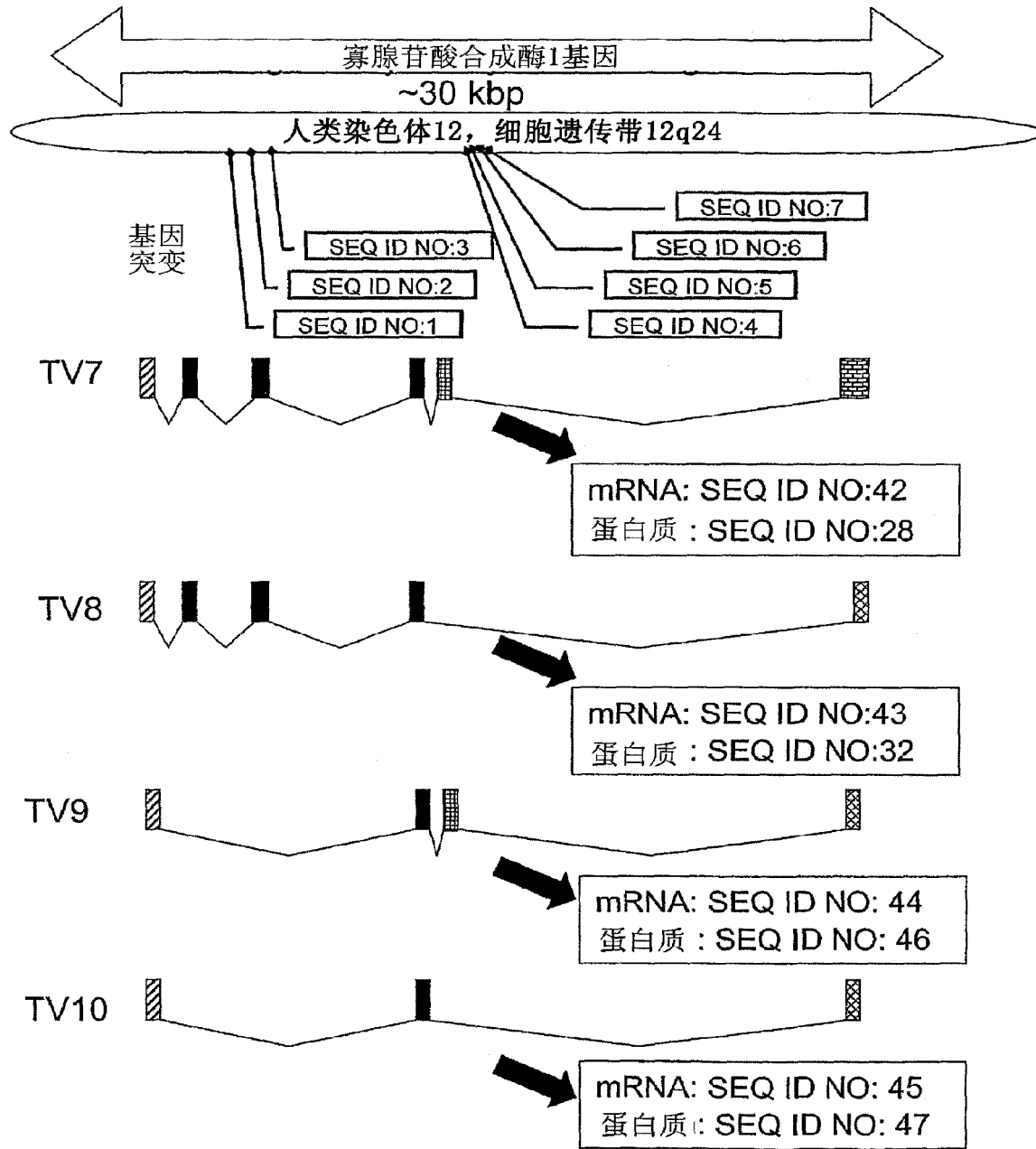
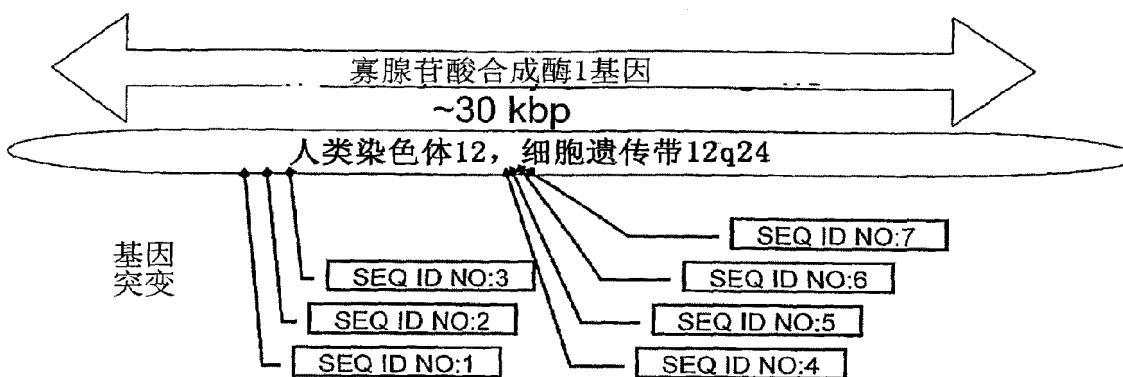


图5B



非人类OAS1基因的推断结构

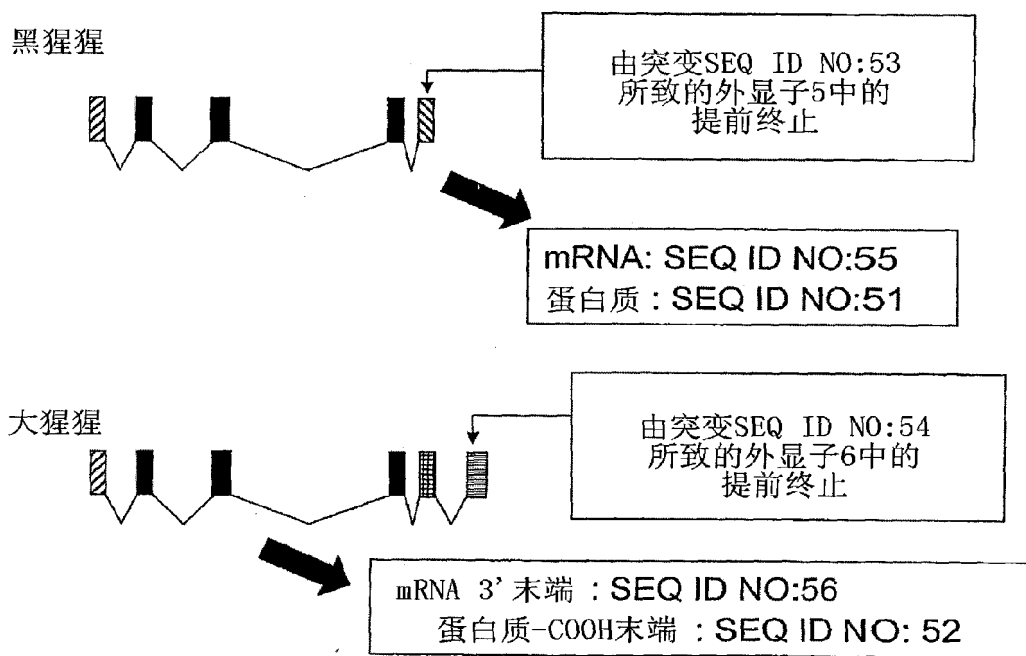


图5C

<p>SEQ ID NO: 53 有机体: 黑猩猩 相对于人类的变异体: A (A取代G) 同源人类基因组序列Genbank ID : NT 009775.13 人类基因组序列的突变坐标 (起点-终点) : 2,142,351-2,142,351 有机体的局部鉴别基因组序列背景: CTGGCACAAGAGGCTGAGGCCTGGCTGAATTACCCATGCTTTAAGAATTG <u>A</u> GATGGGTCCCCAGTGAGCTCCTGGATTCTGCTGGTGAGACCTCCTGCTTC</p>
<p>SEQ ID NO: 54 有机体: 大猩猩 相对于人类的变异体: -- (CA的两个碱基对缺失) 同源人类基因组序列Genbank ID : NT 009775.13 人类基因组序列的突变坐标 (起点-终点) : 2,144,089-2,144,090 有机体的局部鉴别基因组序列背景: CTCCCTGATGTGATCATGTGTCTCACCCCTTTCAGGCTGAAAGCG<u>ACAGTG</u> <u>--</u> GACGATGAGACCGACGATCCAGGAGGTATCAGAAATATGGTTACATTGG</p>

黑体单下划线的碱基表示由各SEQ ID NO所鉴别的相对于人类的特定有机体变异体。

双下划线碱基表示人类基因组序列与所指有机体之间的其他共定位变异体。

简并核酸密码:

R=A/G

Y=C/T

S=C/G

K=G/T

图6
非人的灵长类动物突变

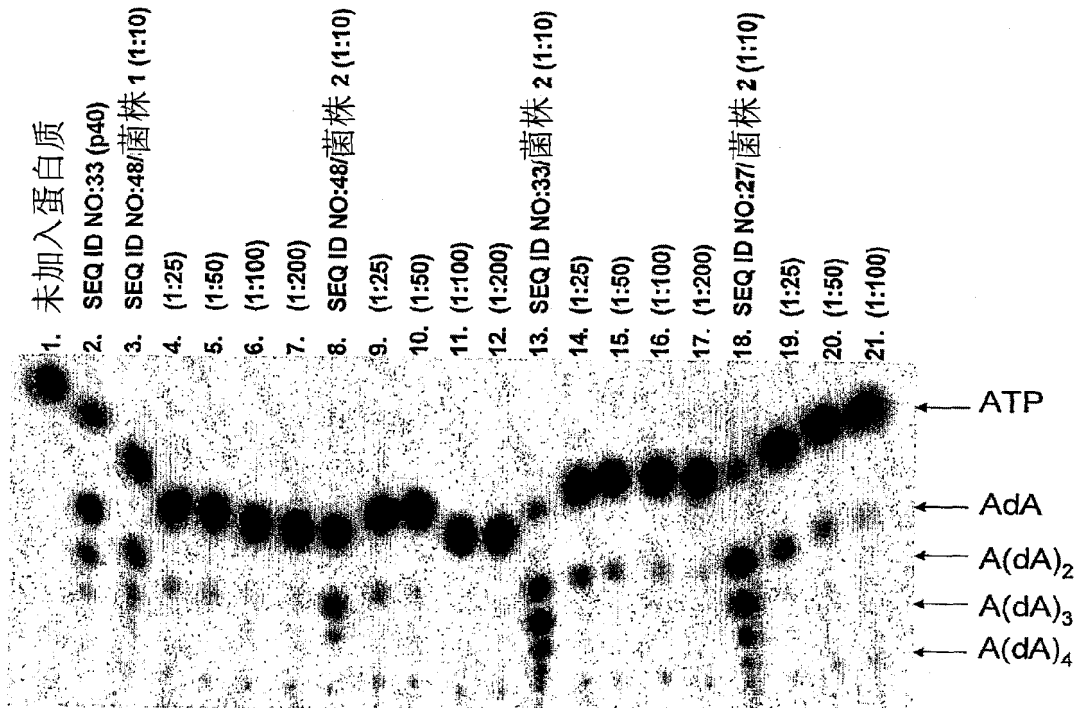


图7:大肠杆菌产生的蛋白质的寡腺苷酸合成酶1活性

将大肠杆菌中所产生的寡腺苷酸合成酶1的酶活性(条带3-21)与杆状病毒表达系统中所产生的酶活性(条带2)相比较。

大肠杆菌样本在两个菌株的任一者中产生且稀释为1:10(条带3、8、13和18)、1:25(条带4、9、14和19)、1:50等。在 $\alpha^{32}\text{P}$ dATP的存在下进行酶反应并将酶产物通过薄层色谱法在经聚乙亚胺纤维素涂覆的平板上离析。
2'-5'-寡腺苷酸的唯一已知功能是刺激细胞抗病毒反应。

图7 大肠杆菌产生的蛋白质的寡腺苷酸合成酶1活性

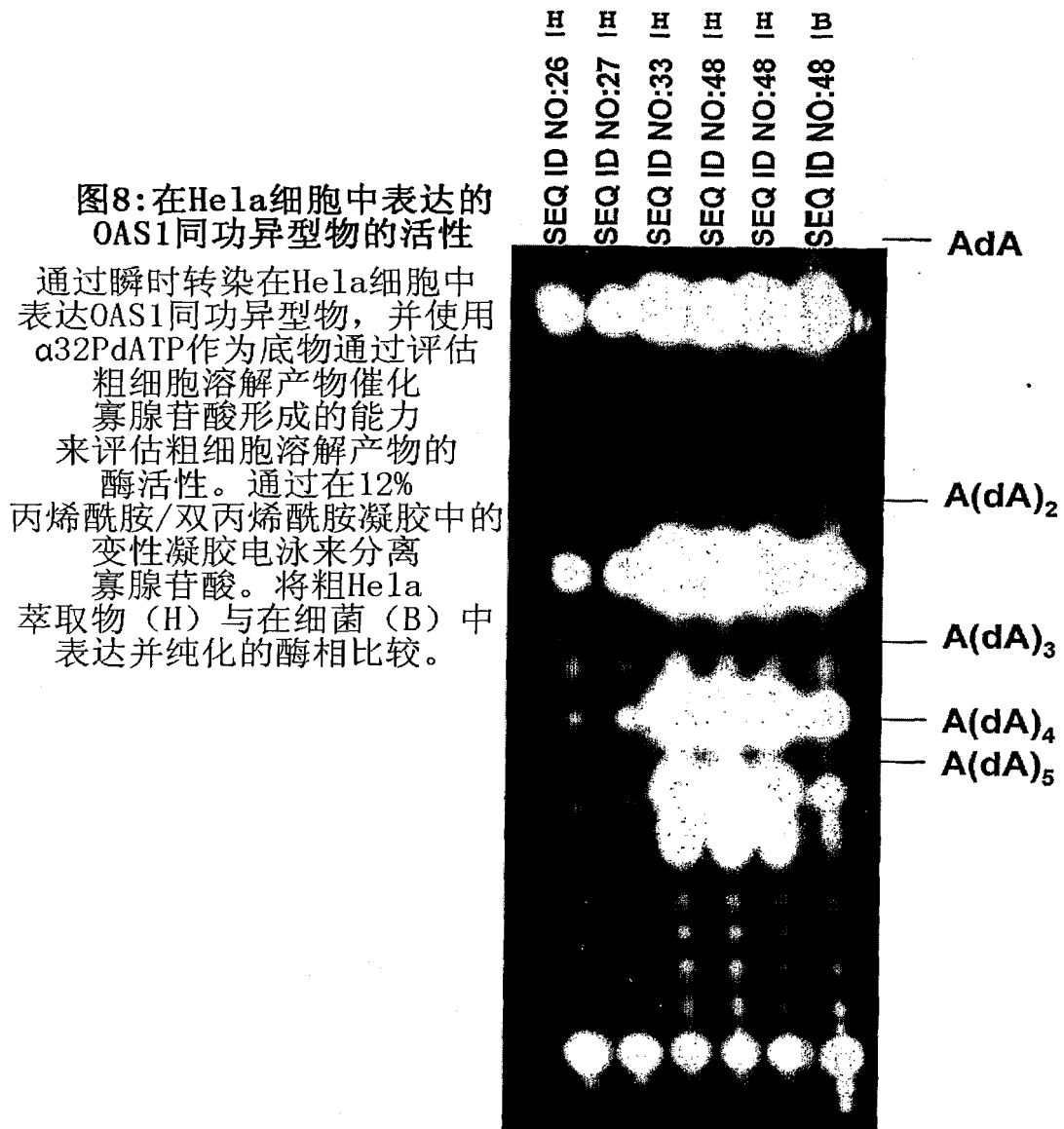


图8
HeLa细胞中的寡腺苷酸合成酶1活性

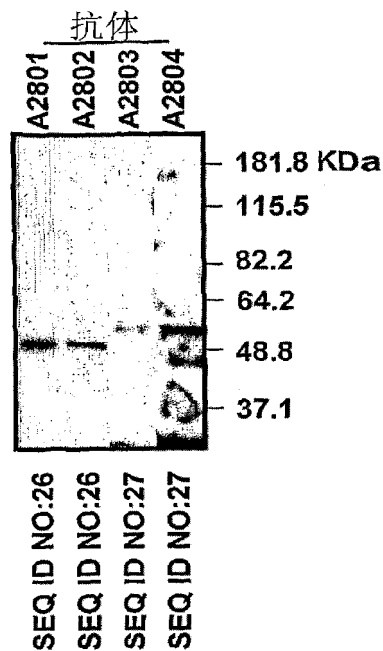


图9: 对新颖OAS1形式具特异性的抗体
 通过兔免疫将多克隆抗体培养成源自
 每个OAS1形式的独特部分的合成肽序列。
 针对在兔网织红细胞溶解产物
 偶合型转录-转译系统中
 活体外合成的蛋白来测试每个抗体。

图9

对新颖形式的OAS1具有特异性的抗体的显影

序列ID号	多肽序列
SEQ ID NO:85	YGRKKRRQRRR
SEQ ID NO:86	RQIKIWFQNRRMKWKK
SEQ ID NO:87	MTSRRSVKSGPREVPRDEYEDLYYIPSSGMASPDSPPD TSRRGALQTRSRQGEVRFVQYDESDYALYGGSSSEDD EHPEVPRTRRPVSGAVLSGPGPARAPPPAGSGGAGRTP TTAPRAPRTQRVATKAPAAPAAETTRGRKSAQPESAAL PDAPASTAPTRSKTPAQGLARKLHFSTAPPNPDPWPTR VAGFNKRVFCAAVGRLAAMHARMAAVQLWDMSPRPT DEDLNELLGITTRVTVCEGKNLLQRANELVNPVQDV DAATATRGRSAASRPTERPRAPARSASRPRRPVE
SEQ ID NO:88	MSRKL FASILIGALLGIGAPPSAHAGADDVVDSSKSFVM ENFSSYHGTKPGYVDSIQKGIQKPKSGTQGN YDDDWKGFYSTDNKYDAAGYSVDNENPLSGKAGGVV KVTYPGLTKVLALKVDNAETIKKELGLSLTEPLM EQVGTEEFIKRFGDGASRVVLSLPPFAEGSSSVEYINNWE QAKALSVELEINFETRGRKRGQDAMYEYMAQA CAGNRVRRSVGSSLSCLNDWDVIRDKTKTKIESLKEHG PIKNKMSSEPNKTVSEEKAKQYLEEFHQ TAL EHPSELKTVTGTNPVFAGANYAAWAVNVAQVIDSET ADNLEKTTAALSILPGIGSVMGIADGAVHHNT EEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIIN LFQVVHNSYNRPAYSPGHKTQPFLHDGYA VSWNTVEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLP TIPGKLDVNKSKTHISVNGRKRIRMRCRAID GDVTFCRPKSPVYVGNVGHANLHVAFHRSSSEKIHSNE ISSDSIGVLGYQKTVDH TKVNSKLSLFFEIKS
SEQ ID NO:89	GDIMGEWGNEIFGALAGFLG
SEQ ID NO:90	RRRRRRR
SEQ ID NO:91	RRQRRTSKLMKR
SEQ ID NO:92	GWTLNSAGYLLGKINLKALAALAKKIL
SEQ ID NO:93	WEAKLAKALAKALAKAHLAKALAKALKACEA
SEQ ID NO:94	KETWWETWWTEWSQPKKKRKV

图10
例示性蛋白转导域系统