

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁶

[12] 发明专利申请公开说明书

[21] 申请号 96198423.6

C12N 15/38
C12N 15/62 C12N 15/70
C12N 1/21 C12Q 1/68
C07K 14/03 C07K 16/08
C12P 21/08 A61K 39/245
A61K 39/42 G01N 33/569

[43]公开日 1999年4月14日

[11]公开号 CN 1214080A

[22]申请日 96.9.26 [21]申请号 96198423.6

[30]优先权

[32]95.9.26 [33]US [31]60/004,297

[86]国际申请 PCT/US96/15702 96.9.26

[87]国际公布 WO97/12042 英 97.4.3

[85]进入国家阶段日期 98.5.19

[71]申请人 华盛顿州大学

地址 美国华盛顿州

[72]发明人 T·M·罗瑟 M·L·波施

K·斯特兰德

[74]专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 齐曾度

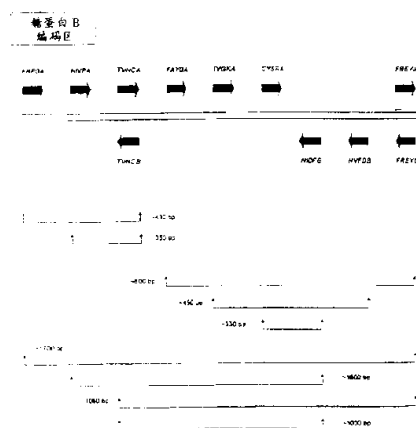
权利要求书 5 页 说明书 112 页 附图页数 26 页

[54]发明名称 RFHV/KSHV 亚族疱疹病毒的糖蛋白 B

[57]摘要

本发明涉及编码 γ 疱疹病毒 RFHV/KSHV 亚族糖蛋白 B 的多核苷酸,此亚族的 3 个成员的特征已被详细描述。DNA 提取物得自被腹膜后纤维瘤病(RF)感染的豚尾猴和食蟹猴,以及被卡波济氏肉瘤(KS)感染的 AIDS 人患者。使用由 γ 疱疹病毒已知的蛋白质和 DNA 序列设计的共有 - 简并寡核苷酸探针扩增提取物。RFHV1 和 KSHV 之间的 319 个碱基对的核苷酸序列片段约有 76% 相同,此片段与 RFHV/KSHV 亚族以外的最相关的 γ 疱疹病毒的相应片段约有 60 - 63% 相同。RFHV1 和 KSHV 之间由 此片段编码的蛋白质序列约 91% 相同,与其它 γ 疱疹病毒的相应序列 \sim 65% 相同。全长的 KSHV 糖蛋白 B 序列在 N-末端含有跨膜区,在细胞外区域含有大量潜在的抗原性位点。本发明还提供了鉴定 RFHV/KSHV 亚族 成员的糖蛋白 B 编码区的材料和方法,所述成员包括但不限于 RFHV1,

RFHV2 和 KSHV。本发明的肽,多核苷酸和抗体可用于诊断感染,并引发针对糖蛋白 B 的免疫应答。



ISSN 1008-4274



权 利 要 求 书

1. 含有编码疱疹病毒糖蛋白 B 之区域的分离的多核苷酸, 所述多核苷酸含有与 SEQ. ID NO: 1 或 SEQ. ID NO: 3 的核苷酸 36 - 354 至少 65% 相同的序列。

5 2. 含有权利要求 1 之多核苷酸的糖蛋白 B 编码区域中至少 50 个连串核苷酸的片段的分离的多核苷酸。

10 3. 含有编码疱疹病毒糖蛋白 B 之区域的分离的多核苷酸, 所述多核苷酸含有选自以下的序列: 与寡核苷酸 SHMDA 至少 74% 相同的 35 个核苷酸的序列 (SEQ. ID NO: 41); 与寡核苷酸 CFSSB 至少 73% 相同的 30 个核苷酸的序列 (SEQ. ID NO: 43); 与寡核苷酸 ENTFA 至少 72% 相同的 29 个核苷酸的序列 (SEQ. ID NO: 45); 和与寡核苷酸 DNIQB 至少 80% 相同的 35 个核苷酸的序列 (SEQ. ID NO: 46)。

4. 含有权利要求 3 之多核苷酸的糖蛋白 B 编码区域中至少 50 个连串核苷酸的片段的分离的多核苷酸。

15 5. 权利要求 1 或权利要求 2 的多核苷酸, 其中所述的疱疹病毒能感染灵长类动物。

6. 权利要求 1 或权利要求 2 的多核苷酸, 其中所述的疱疹病毒是 RFHV1, RFHV2 或 KSHV。

20 7. 含有与 SEQ. ID NOS: 1, 3, 或 92 所含的核苷酸 36 - 354 之间或 SEQ. ID NO: 96 中任何位置处的线性序列, 而非 SEQ. ID NO: 96 中的线性序列至少有 21 个核苷酸相同的线性序列的分离的多核苷酸。

8. 权利要求 7 的分离的多核苷酸, 含有与 SEQ. ID NO: 1, SEQ. ID NO: 3, 或 SEQ. ID NO: 96 的核苷酸 36 - 354, 或者 SEQ. ID NO: 92 中的至少 200 个连串核苷酸基本上相同的线性序列。

25 9. 由权利要求 2 的多核苷酸编码的分离的多肽。

10. 含有与 SEQ. ID NOS: 2, 4, 或 97 所含的氨基酸 13 - 118, 或 SEQ. ID NO: 94 中任何位置处的序列, 而非 SEQ. ID NO: 99 中的序列至少有 17 个氨基酸基本上相同的线性序列的分离的多肽。

30 11. 权利要求 10 的分离的多肽, 基本上由具有 SEQ. ID NO: 2, SEQ. ID NO: 4, SEQ. ID NO: 94 或 SEQ. ID NO: 97 所示序列的多肽组成。

12. 含有与根据权利要求 10 的多肽所含的序列至少有 17 个氨基酸相同的线性序列的融合多肽。



13. 权利要求 10 的分离的多肽, 涉及疱疹病毒与哺乳动物细胞的结合或融合。

14. 权利要求 10 的分离的多肽, 它是糖基化的。

15. 权利要求 10 的分离的多肽, 它是非糖基化的。

5 16. 权利要求 10 的分离的多肽, 它是免疫原性的。

17. 权利要求 10 的分离的多肽, 含有选自 SEQ. ID NOS: 67 - 76 的序列。

18. 编码权利要求 10 的多肽的分离的多核苷酸。

19. 编码权利要求 10 的多肽的非自然产生的多核苷酸。

10 20. 编码融合多肽的多核苷酸, 它含有与编码多肽的第二多核苷酸直接相连的权利要求 2 的多核苷酸。

21. 含有编码与根据权利要求 10 的多肽所含的序列至少有 17 个氨基酸相同的线性序列的多核苷酸序列的重组克隆载体。

15 22. 含有与调控多核苷酸序列有效地连接, 并编码与根据权利要求 10 的多肽所含的序列至少有 17 个氨基酸相同的线性序列的多核苷酸序列的重组表达载体。

23. 含有与权利要求 7 的多核苷酸所含的线性序列至少有 21 个核苷酸相同的线性序列的重组克隆载体。

20 24. 被权利要求 18 或 19 的多核苷酸, 或被权利要求 21, 22 或 23 的载体转化的宿主细胞。

25. 特异针对权利要求 1 多核苷酸的所述编码区所编码的糖蛋白 B 多肽的单克隆或分离的多克隆抗体。

26. 特异针对权利要求 11 之多肽, 但不针对具有 SEQ ID NOS: 30 - 41 中任何一个的氨基酸序列的多肽的单克隆或分离的多克隆抗体。

25 27. 权利要求 26 的抗体, 它是单克隆抗体。

28. 权利要求 26 的抗体, 它是分离的多克隆抗体。

29. 含有权利要求 9 或 10 中的多肽和药物可相容的赋形剂的疫苗。

30. 权利要求 29 的疫苗, 它也含有佐剂。

31. 治疗疱疹病毒感染的方法, 包括施用权利要求 29 的疫苗。

30 32. 含有权利要求 2 或 7 中的多核苷酸和药物可相容的赋形剂的疫苗。

33. 权利要求 32 的疫苗, 它是活病毒或病毒表达载体。



34. 治疗疱疹病毒感染的方法，包括施用权利要求 32 的疫苗。
35. 含有权利要求 25 或 26 中的抗体和药物可相容的赋形剂的疫苗。
36. 治疗疱疹病毒感染的方法，包括施用权利要求 35 的疫苗。
- 5 37. 与选自 SEQ ID NOS: 24 - 63, SEQ ID NOS: 77 - 78 和 SEQ ID NOS: 80 - 90 的寡核苷酸基本上相同的寡核苷酸。
38. 得到编码糖蛋白 B 的多核苷酸的扩增拷贝的方法，包括步骤：
a) 用权利要求 37 的寡核苷酸接触多核苷酸；并
b) 延伸已与多核苷酸形成双螺旋的寡核苷酸。
- 10 39. 权利要求 38 的方法，其中所述扩增反应是聚合酶链反应 (PCR)。
40. 权利要求 39 的方法，其中所述 PCR 包括退火和延伸的重复循环，其中退火在至少 50°C 的温度下进行。
41. 权利要求 38 的方法，其中被扩增的多核苷酸首先得自生物样品，所述生物样品取自患有以成纤维细胞增殖和胶原沉积为特征的疾病的个体。
- 15 42. 检测样品中病毒 DNA 或 RNA 的方法，包括步骤：
a) 将样品中的 DNA 或 RNA 与含有权利要求 2 或 7 的多核苷酸的探针接触，所处的条件应能允许探针与具有 SEQ. ID NO: 1, 3, 92 或 94 所示序列的至少一个多核苷酸，而不与具有 SEQ. ID NOS: 5 - 13 中任一个的序列的多核苷酸形成稳定的双螺旋；和
b) 如果有的话，检测步骤 a) 中形成的所述稳定双螺旋的存在。
- 20 43. 检测样品中病毒 DNA 或 RNA 的方法，包括步骤：
a) 将样品中的 DNA 或 RNA 与含有权利要求 2 或 7 的多核苷酸的探针接触，所处的条件应能允许探针与具有 SEQ. ID NO: 1 所示序列的多核苷酸，和具有 SEQ. ID NO: 34 所示序列的多核苷酸，而不与具有 SEQ. ID NOS: 5 - 13 中任一个的序列的多核苷酸形成稳定的双螺旋；和
b) 如果有的话，检测步骤 a) 中形成的所述稳定双螺旋的存在。
- 25 44. 检测来源于灵长类动物的样品中病毒 DNA 或 RNA 的方法，包括步骤：
a) 将样品中的 DNA 或 RNA 与含有权利要求 2 或 7 的多核苷酸的探针接触，所处的条件应能允许探针与具有 SEQ. ID NO: 1, 3, 92 或
- 30



94 所示序列的至少一个多核苷酸，而不与具有 SEQ. ID NOS: 5, 或 9 - 13 所示的任一个序列的多核苷酸形成稳定的双螺旋；和

b) 如果有的话，检测步骤 a) 中形成的所述稳定双螺旋的存在。

5 45. 权利要求 42 的方法进一步包括在与探针接触之前对样品中的 DNA 或 RNA 进行扩增反应。

46. 检测样品中病毒 DNA 或 RNA 的方法，包括步骤：

a) 在反应中使用权利要求 37 的寡核苷酸作为引物对样品中的多核苷酸进行扩增反应；和

b) 如果有的话，检测多核苷酸扩增拷贝的存在。

10 47. 能与含有选自 SEQ. ID NO: 1, 3, 92 和 94 和它们各自的互补序列的序列的第二多核苷酸形成稳定双螺旋的分离的多核苷酸，所处的条件中第二多核苷酸能与具有 SEQ. ID NO: 1, 3, 92 或 94 所示序列的至少一个多核苷酸，而不与具有 SEQ. ID NOS: 5 - 13 中任一个的序列的多核苷酸形成稳定的双螺旋。

15 48. 权利要求 47 的分离的多核苷酸，其核苷酸序列包含于自然产生的病毒的基因组中。

49. 含有由权利要求 47 的多核苷酸编码的至少 25 个氨基酸的线性序列的分离的多肽。

20 50. 检测疱疹病毒对个体的感染的方法，包括检测得自个体的生物样品中的病毒 DNA 或 RNA，其中通过权利要求 42, 44, 或 46 的方法检测病毒 DNA 或 RNA。

51. 检测生物样品中疱疹病毒多核苷酸的诊断试剂盒，它含有适当包装的试剂，其中试剂包括权利要求 2 或 7 的多核苷酸。

25 52. 检测生物样品中疱疹病毒多核苷酸的诊断试剂盒，它含有适当包装的试剂，其中试剂包括权利要求 37 的寡核苷酸。

53. 检测疱疹病毒对个体的感染的方法，包括下列步骤：

a) 将得自个体的样品中的抗体与权利要求 25 或 26 的多肽接触，所处的条件能允许稳定的抗原 - 抗体复合物的形成；和

b) 如果有的话，检测步骤 a) 中所形成的所述稳定复合物。

30 54. 检测生物样品中存在的抗 - 疱疹病毒抗体的诊断试剂盒，它含有适当包装的试剂，其中试剂包括权利要求 25 或 26 的多肽。

55. 检测疱疹病毒对个体的感染的方法，包括下列步骤：



a) 将得自个体的样品中的多肽与权利要求 9 或 10 的抗体接触，所处的条件能允许稳定的抗原 - 抗体复合物的形成；和

b) 如果有的话，检测步骤 a) 中形成的所述稳定复合物。

5 56. 检测生物样品中存在的疱疹病毒多肽的诊断试剂盒，它含有适当包装的试剂，其中试剂包括权利要求 9 或 10 的抗体。

57. 确定候选药物是否能用于治疗 γ 疱疹病毒感染的方法，包括下列步骤：

a) 将权利要求 9 或 10 的多肽与候选药物接触；和

b) 确定候选药物是否能改变多肽的生化功能。

10 58. 权利要求 57 的方法，其中步骤 b) 中确定的多肽生化功能是多肽与哺乳动物细胞表面的结合。

59. 产生糖蛋白 B 多肽的方法，包括在真核细胞中表达权利要求 2 或 7 的多核苷酸。

15 60. 产生含有疱疹病毒糖蛋白 B 编码区的多核苷酸的方法，包括复制权利要求 21 或 23 的重组克隆载体。

说明书

RFHV/KSHV 亚族疱疹病毒的糖蛋白 B

技术领域

5 本发明一般涉及病毒学领域，尤其是疱疹病毒家族的病毒。本发明更具体地涉及疱疹病毒糖蛋白 B 分子的鉴定和特征描述，所述分子与包括人的灵长类动物的纤维增殖性和肿瘤疾病相关。

背景技术

10 卡波济氏肉瘤是一种损坏外形并潜在致命的出血性肉瘤形式，其特征在于在皮肤上以深色斑或小结出现的多发性脉管肿瘤。在组织学水平上，其特征在于相对一致的纺锤形细胞的增殖，形成束和脉管裂缝。在炎症浸润中经常会出现浆细胞，T 细胞和单核细胞，胃肠损害处或相关的淋巴瘤处出血的最终结果可能是死亡（一般参见 Martin 等人，Finesmith 等人）。

15 一旦遇到相对而言起因不明的疾病，公众的注意力立刻会转向该疾病与 AIDS 的相关性。多达 20% 的某些感染了 AIDS 的群体在疾病过程中会患卡波济氏肉瘤，卡波济氏肉瘤也出现在与免疫缺损相关的其它疾病中，包括肾透析和治疗级免疫抑制，然而此疾病的流行病学暗示免疫缺损不是唯一的引发因素。具体地说，卡波济氏肉瘤与某些性服务高水平的相关性暗示非人类免疫缺损病毒这一病原因子的介入（Berel 等人）。

20 已鉴定出得自 AIDS 病人卡波济氏损害的组织样品中疱疹病毒样 DNA 序列（Chang 等人，由 Ambroziuk 等人进一步证实），通过表现度差异分析得到序列（Lisitsyn 等人），其中使用不相关的引物寡核苷酸扩增了已感染和未感染组织中的 DNA，然后使所述 DNA 杂交以凸现细胞间的差异。此序列与 Epstein Barr 病毒和松鼠猴疱疹病毒的已知序列部分相同，它编码两个被隔绝在病毒内部的结构组分—衣壳和外被蛋白。在概括研究多种来源的组织时发现不论病人的 HIV 状况如何，95% 的卡波济氏肉瘤损害中都有此序列（Moore 等人，1995a）。相同病人的 21% 未介入组织呈阳性，而对照群体样品中 5% 呈阳性，样品之间大约有 0.5% 的序列差异。

30 在体腔淋巴瘤中也检测到相同的序列，体腔淋巴瘤为具有 B 细胞特征的淋巴瘤渗出物，只在 AIDS 病人中出现（Cesarman 等人）。与卡波



济氏肉瘤相比,体腔淋巴瘤中拷贝数更高。其它与 AIDS 相关的淋巴瘤为阴性。在患有 Castleman 氏疾病的患者的外周血单核细胞中也发现了此序列 (Dupin 等人), 此病的特征在于血管滤泡增生的形态学特征, 并且与发烧, 腺病, 和脾大相关。已知推断的从中得到序列的病毒是与卡波济氏肉瘤相关的疱疹病毒 (KSHV)。

Boshoff 等人使用 PCR 原位杂交已检测出据认为代表卡波济氏肉瘤中致瘤性细胞的细胞类型的 KSHV 多核苷酸序列, 血清学证据支持 KSHV 在卡波济氏肉瘤病原学中的重要作用 (O' Leary)。Kedes 等人研究出一种免疫荧光血清学测定法, 该方法可检测出被 KSHV 潜伏感染的 B 细胞中与潜伏状态相关的核抗原的抗体, 他们还发现在卡波济氏肉瘤的患者中 KSHV 血清正性高。Gao 等人通过免疫印迹测定法发现: 在卡波济氏肉瘤的 40 个患者中, 32 个患者的抗 KSHV 抗原的抗体为阳性, 与之相比, 在 AIDS 突然爆发前一刻未患有卡波济氏肉瘤的 40 个男同性恋者只有 7 个呈现出抗 KSHV 抗原的抗体为阳性。Miller 等人从含有 KSHV 和 Epstein-Barr 病毒基因组的体腔淋巴瘤细胞系中制备出 KSHV 抗原。已在 48 个被 HIV-1 感染的卡波济氏肉瘤患者中的 32 个中鉴定出针对一个抗原的抗体, 该抗原被称为 p40, 与之相比, 在 54 个被 HIV-1 感染但未患有卡波济氏肉瘤的患者中, 只在 7 个患者中鉴定出上述抗体。

Zhong 等人在 mRNA 水平上分析了 KSHV 序列在被感染组织中的表达, 发现两个小的转录本表现出大部分由 KSHV 基因组转录的病毒特异性 RNA。他们推测一个转录本编码小的膜蛋白; 另一个是在细胞核中积累的异常的 poly-A RNA, 可能没有蛋白质编码序列。通过从基因组 DNA 的 λ 文库中克隆大量重叠的 KSHV 基因组片断可以分析 mRNA, 所述片断跨越了 ~120kb 的 KSHV 基因组, 将克隆用作探针以进行 Northern 分析, 但未得到或公开它们的序列。

Moore 等人已部分鉴定出得自体腔淋巴瘤的 KSHV 基因组片断, 尽管未公开基因组中 20.7kb 区域的序列, 但据报道已测定出此区域的序列。此片断中存在 17 个部分的或完整的开放阅读框, 除 1 个以外, 都与其它已知的 γ 疱疹病毒基因具有序列和位置的同源性, 所述基因包括衣壳成熟基因和胸苷激酶基因。系统发育分析表明相对于 KSHV 与 Epstein Barr 病毒的相关性而言, KSHV 与马疱疹病毒 2 型和松鼠猴病毒更加密切相关。20.7kb 的区域不含有编码糖蛋白 B 或 DNA 聚合酶的序

列。

疱疹病毒家族作为一个整体含有很多大小约为 100nm 的多包膜病毒，所述病毒能够感染脊椎动物（一般性评论例见 Emery 等人，Field 等人）。双链 DNA 基因组异乎寻常得大 - 长度约为 88 - 229 个千碱基，它在病毒生命周期的不同时期可产生 50 个以上不同的转录本。大量糖蛋白表达于病毒表面，它们对于病毒识别靶细胞，病毒侵入细胞内起作用。与内部病毒成分相比，在种之间这些表面蛋白有相对更多的变体（Karlin 等人）。在由携有病毒的细胞产生的缺损病毒颗粒上也出现了相同的表面蛋白，一个这种非感染性的形式是 L-颗粒，它含有外壳和病毒包膜，但缺乏核衣壳。

疱疹病毒家族已被分成几个亚族，对每一类目的划分起先是基于生物学特性，当基因组序列数据出现时才得以精简。 α 亚族含有的病毒具有宽宿主范围，短的复制周期，和对感觉神经节的亲和性，它们包括人单纯疱疹病毒和水痘 - 带状疱疹病毒。 β 亚族含有的病毒的宿主范围受到限制，它们包括巨细胞病毒和人疱疹病毒 6 型。 γ 亚族含有的病毒通常嗜淋巴细胞，其 DNA 的特征在于约 110 千碱基的低 GC 含量的片段旁侧有多个高 GC 含量的串联重复， γ 亚族包括 Epstein Barr 病毒 (EBV)，松鼠猴疱疹病毒，马疱疹病毒 2 和 5 型，和牛疱疹病毒 4 型。

疱疹病毒与具有复杂临床进程的疾病有关，很多疱疹病毒的特征是能够在宿主体内进入潜伏期达相当长的时间。 α 亚族的病毒在感觉和植物性神经节中保持潜伏状态，而 γ 亚族的病毒在例如淋巴细胞谱系的细胞中保持潜伏状态。潜伏期与某些病毒基因的转录有关，并可以持续几十年直至身体状况最适于病毒恢复活性复制，这种身体状况可包括免疫缺损。另外，一些 γ 亚族的疱疹病毒能够遗传转化它们感染的细胞，例如，EBV 与 B 细胞淋巴瘤，口毛白斑病，淋巴组织间质肺炎和鼻咽癌有关。

在人和其它脊椎动物中出现的许多其它疾病涉及纤维增殖和前 - 肿瘤细胞的产生，出现在人中的疾病的例子是腹膜后纤维变性，小结状纤维瘤病，假肉瘤纤维瘤病，和硬化肠系膜炎。在华盛顿大学的地区性灵长类动物研究中心的猕猴群体中已观察到已知为地方性腹膜后纤维瘤病 (RF) 的另一种疾病 (Giddens 等人)。此疾病晚期的特征在于纤维组织围绕肠系膜和腹膜腔的背面部分增殖，延伸至腹股沟内，穿过隔膜，

并进入腹壁。一旦在临床上出现此病的明显症状，病人一律在 1-2 个月内死亡。此病与由 D 型猿猴逆转录病毒 SRV-2 造成的猿猴免疫缺损 (SAIDS) 相关 (Tsai 等人)，但是在感染了 SAIDS 的其它猴群体中未显示出相同的 RF 频率，近年来华盛顿的 RF 频率已有所降低。

5 在非人灵长类动物中研究这种疾病是重要的，这不仅是由于它可作为人疾病的模型，也是因为一种灵长类动物可作为感染另一种灵长类动物的病毒的来源库，例如，松鼠猴疱疹病毒在其天然宿主松鼠猴 (*Saimiri sciureus*) 中似乎不会引起疾病，但它在其它灵长类动物，尤其是猫头鹰 (owl) 猴中会引起多克隆的 T-细胞淋巴瘤和急性白血病。

10 有必要开发试剂和方法以用于检测和治疗疱疹病毒感染。由血清学证据进一步证实的 KSHV 与卡波济氏肉瘤之间的病原学联系显示出这种需要的重要性。

例如，有必要开发能够用于诊断和评价卡波济氏肉瘤和类似疾病的试剂和方法。能够检测新病人的病原因子在鉴别诊断中会起作用；能够
15 评价正在发生的疾病病原因子的水平在临床处理时会起作用。合乎需要的标记物包括那些与乙肝病毒的 HBsAg 相似，对于活跃的和潜伏形式的病毒感染都可提供很敏感的指示的标记物。合乎需要的标记物也包括那些免疫原性的，可用于评估对在抗体应答中出现的病毒因子的免疫学暴露的标记物。得自病毒包膜的糖蛋白抗原特别适于作为具有这些特征的
20 标记物，它们可以不仅在病毒复制形式的表面附近，还在由病毒感染的细胞产生的 L-颗粒上以高丰度表达。

第二，有必要开发能够用于治疗病毒感染的试剂和方法—既是预防性的，也在被病毒攻击后起作用。这种试剂包括可赋予抗病毒的免疫力水平的疫苗。被动疫苗，如那些含有抗-病毒抗体的疫苗可用于为新近
25 被暴露于病毒的个体提供即时保护或防止病毒侵入细胞和复制。主动疫苗，如那些含有免疫原性的病毒成分的疫苗可用于在个体中产生主动的和不断发展的免疫应答。由主动疫苗产生的抗体有助于保护个体抗活病毒的后续攻击。由主动疫苗产生的细胞毒 T 细胞通过消除病毒复制所涉及的宿主细胞可有助于消灭同时发生的感染。保护性免疫应答，尤其是
30 抗体的适当的靶是暴露于病毒颗粒表面的蛋白质抗原，和那些牵涉到病毒与靶细胞融合的蛋白质抗原。

第三，有必要开发能够用于开发治疗卡波济氏肉瘤和类似疾病的新

药物的试剂和方法。对卡波济氏肉瘤的最新治疗方法是放射性疗法与传统的化学疗法，如长春新碱（Northfelt, Mitsuyasu）的联合，尽管损害会对这些物理疗法作出反应，但反应是暂时的，减弱的临床进程一般会恢复，甚至实验性的疗法，如用细胞因子治疗也是针对疾病综合征而不是病因。基于病原因子的药物筛选和合理的药物设计可直接针对对于具有长期效力的临床方案的长期需要。这种药物的适当的靶是涉及识别和侵入宿主细胞的病毒成分，这些靶包括病毒包膜的糖蛋白成分。

第四，有必要开发能够用于鉴定可能与其它纤维增殖性疾病相关的新的病毒因子的试剂和方法。Chang 等人使用的表现度差异分析技术非常复杂，可能不适于作为一般的筛选试验。在研究多种怀疑含有相关病原因子的组织样品的更常规的试验中能用作试剂的一套寡核苷酸探针，肽和抗体更加合乎需要。此试剂要有足够的特异性以避免鉴定出不相关的病毒和宿主的内源成分，并且此试剂可以充分地交叉反应以鉴定相关的但以前未描述过的病毒病原体。

发明内容

本发明的目的是提供分离的多核苷酸、多肽、和抗体，它们或得自新的编码疱疹病毒 RFHV/KSHV 亚族的糖蛋白 B 分子的基因，或能与所述基因的产物反应。此家族的两个成员是与腹膜后纤维瘤病相关的疱疹病毒（RFHV）和与卡波济氏肉瘤相关的疱疹病毒（KSHV）。这些材料和相关的方法可用于诊断和治疗灵长类，包括人类的疱疹病毒感染。分离的或重组的糖蛋白 B 片断或编码它们的多核苷酸可用作疱疹病毒主动疫苗的成分，而特异于糖蛋白 B 的抗体可用作被动疫苗的成分。

因此，本发明的一个实施方案是具有疱疹病毒 RFHV/KSHV 亚族的糖蛋白 B 编码区域的分离的多核苷酸，所述多核苷酸含有与 SEQ. ID NO: 1 或 SEQ. ID NO: 3 的核苷酸 36-354 至少有 65% 相同的 319 个核苷酸长的序列，它们分别是编码 RFHV 和 KSHV 的糖蛋白 B 的 319 个核苷酸长的片断。本发明还包括具有糖蛋白 B 编码区的分离的多核苷酸，此多核苷酸含有选自以下的序列：与寡核苷酸 SHMDA 至少有 74% 相同的 35 个核苷酸长的序列（SEQ. ID NO: 41）；与寡核苷酸 CFSSB 至少有 73% 相同的 30 个核苷酸长的序列（SEQ. ID NO: 43）；与寡核苷酸 ENTFA 至少有 72% 相同的 29 个核苷酸长的序列（SEQ. ID NO: 45）；和与寡核苷酸 DNIQB 至少有 80% 相同的 35 个核苷酸长的序列（SEQ. ID NO:

46)。

本发明的另一个实施方案是含有上述实施方案中的糖蛋白 B 多核苷酸编码区的至少 21 个，优选 35 个，更优选 50 个，再更优选 75 个，甚至更优选 100 个连串核苷酸的片断的分离的多核苷酸。此多核苷酸优选来自能感染灵长类动物的病毒，包括来自 RFHV 和 KSHV 的编码糖蛋白 B 的多核苷酸片断。本发明的另一个实施方案是含有与 SEQ. ID NO: 1, SEQ. ID NO: 3, 或 SEQ. ID NO: 92 所含的核苷酸 36 - 354 之间的糖蛋白 B 编码序列，或者 SEQ. ID NO: 96 中的任何位置处的序列，而非 SEQ. ID NO: 98 所含序列至少约有 21 个核苷酸相同的分离的多核苷酸。

本发明的另一个实施方案是由任一个前述实施方案编码的分离的多肽。本发明还包括一种分离的多肽，它含有与 SEQ. ID NO: 2, SEQ. ID NO: 4, 或 SEQ. ID NO: 97 所示的糖蛋白 B 的蛋白质序列，或者 SEQ. ID NO: 94 (KSHV) 中的任何位置处的序列，而非 SEQ. ID NO: 99 所含序列至少有 17 个氨基酸基本上相同的线性序列。它们包括融合多肽，免疫原性的多肽，和以糖基化和非糖基化形式存在的多肽，SEQ. ID NOS: 67 - 76 中列出了一些优选的抗原肽。本发明中还包括编码任一个前述多肽的分离的和非天然产生的多核苷酸，以及克隆载体，表达载体和由它们得到的被转染的宿主细胞。本发明的另一个实施方案是生产本发明的多核苷酸或多肽的方法，此方法包括复制本发明的载体或在适当的宿主细胞中表达多核苷酸。

本发明的另一个实施方案是特异针对本发明中包括的糖蛋白 B 多肽，或特异于在本发明所包含的多核苷酸编码区编码的糖蛋白 B 的单克隆或分离的多克隆抗体。所述抗体特异于 RFHV/KSHV 亚族的成员，而不与更远相关的糖蛋白 B 序列，特别是 SEQ. ID NOS: 30 - 41 交叉反应。

其它糖蛋白 B 抗体为特异针对权利要求 9 的多肽，而不是具有 SEQ. ID NOS: 30-41 任一氨基酸序列的多肽的特异性单克隆或分离的多克隆抗体。

本发明的另一个实施方案是含有本发明的多肽和药物可接受的赋形剂，也可非强制性地含有佐剂的疫苗。本发明的另一个实施方案是含有本发明的多核苷酸的疫苗，它可以是活病毒或病毒表达载体的形式。本发明的另一个实施方案是含有本发明的抗体和药物可接受的赋形剂的疫苗。其它实施方案是或预防性地或在正进行着的感染过程中治疗疱疹病

毒感染的方法，所述方法包括施用前述实施方案中的一个。

5 本发明的另一个实施方案是特异于 γ 疱疹病毒亚族，RFHV/KSHV 亚族，RFHV，和 KSHV 的糖蛋白 B 编码序列的寡核苷酸，尤其是 SEQ. ID NOS: 24-63 中所列的那些。本发明还包括得到编码糖蛋白 B 的多核苷酸的扩
增拷贝的方法，所述方法包括将多核苷酸与一个或多个前述的寡核苷酸
接触。欲被扩增的多核苷酸可以取自患有以成纤维细胞增殖和胶原沉积
为特征之疾病的个体，所述疾病包括但不限于腹膜后纤维瘤病或卡波济
氏肉瘤，或淋巴细胞谱系的恶性肿瘤。

10 本发明的另一个实施方案是检测样品中病毒 DNA 或 RNA 的方法，一个方法包括的步骤有：将样品中的 DNA 或 RNA 与含有本发明的多核苷酸
或寡核苷酸的探针接触，所处条件应能允许探针与具有 SEQ. ID NO: 1
或 SEQ. ID NO: 3，或这两者所示序列的多核苷酸，而不与具有除 RFHV/KSHV
亚族以外的疱疹病毒序列，尤其是 SEQ. ID NOS: 5-13 的多核苷酸形
成稳定的双螺旋；和检测由此形成的任何双螺旋的存在。涉及的条件是
15 一单套反应参数，如保温时间，温度，溶质浓度和洗涤步骤。在这些条
件下，多核苷酸如果与具有 SEQ. ID NO: 1 所示序列的多核苷酸，或具
有 SEQ. ID NO: 3 所示序列的多核苷酸，或同时具有这两者的多核苷酸
接触，可以形成稳定的双螺旋，如果与具有 SEQ. ID NO: 5-13 中任一
20 个的序列的多核苷酸接触，不能形成稳定的双螺旋。另一个方法包括的
步骤有：在扩增反应中使用本发明的寡核苷酸作为引物对样品中的 DNA
或 RNA 进行扩增反应，并检测任何经扩增的拷贝的存在。本发明还包括
由前述方法鉴定的分离的多核苷酸，所述多核苷酸可能存在于天然产生
的病毒的基因组或被感染的组织中。

25 本发明的另一个实施方案是检测生物样品中与疱疹病毒感染相关之
成分的诊断试剂盒，所述生物样品可得自怀疑具有这种感染的个体，所
述试剂盒中含有适当包装的本发明的多核苷酸，寡核苷酸，多肽或抗体。
本发明还包括检测个体感染的方法，所述方法包括对得自个体的生物样
品使用本发明的试剂，方法，或试剂盒。

30 本发明的另一个实施方案是用于治疗被 γ 疱疹病毒感染之个体的治
疗性的化合物和组合物，所包括的治疗剂含有以基因疗法为目的的本发
明的多核苷酸和载体。本发明还包括通过将本发明中所包含的多肽与所
述化合物接触，并测定多肽的生化功能是否改变而鉴定的药用化合物。

本发明还包括基于糖蛋白 B 分子的结构和生化特征，由合理的药物设计得到的药用化合物。

附图简述

图 1 列出了从 RFHV 和 KSHV 的糖蛋白 B 编码区扩增的多核苷酸序列。下划线部分是残基 36 - 354 之间的 319 - 碱基的多核苷酸区段，并表示用于扩增它的引物之间的各个病毒基因区段。与多核苷酸序列排在一起的是可用作杂交探针或 PCR 引物的寡核苷酸。Type1 寡核苷酸含有 γ 疱疹病毒共有序列，可被用于扩增 γ 疱疹病毒的糖蛋白 B 基因区段，所示例子是 NIVPA 和 TVNCB。Type2 寡核苷酸含有 RFHV/ KSHV 亚族的共有序列，可被用于扩增属于此亚族的病毒的糖蛋白 B 基因区段，所示例子是 SHMDA, CFSSB, ENTFA 和 DNIQB。所示的其它寡核苷酸是 Type3 寡核苷酸，这些寡核苷酸含有直接取自 RFHV 或 KSHV 序列的序列，并且特异于各个病毒的序列。在编码序列的方向上起始扩增的寡核苷酸（名称以“A”结束）以 5' \rightarrow 3' 的方向被列出。以与编码序列的方向相反的方向起始扩增的寡核苷酸（名称以“B”结束）以 3' \rightarrow 5' 的方向被列出。图 1 中还示出了由 RFHV 和 KSHV 多核苷酸序列编码的多肽。在两个序列中由核苷酸 238 - 240 编码的天冬酰胺是潜在的 N-键糖基化位点，此位点在这两种病毒和其它疱疹病毒中是保守的。

图 2 是据信包含在 KSHV 基因组，和 RFHV/KSHV 亚族其它成员的基因组中的编码糖蛋白 B 的 DNA 序列的图谱。所示出的是本文所述的 KSHV 糖蛋白 B 序列的大致位置。还示出了表示 Type1 共有/简并寡核苷酸之杂交位点的推断的保守区段，它可用于从 γ 疱疹病毒中探测和扩增出糖蛋白 B 序列。

图 3 列出了一些以前已知的疱疹病毒糖蛋白 B 的蛋白质序列，和它排在一起的是完整的 KSHV 糖蛋白 B 蛋白质序列和 RFHV1 及 RFHV2 的片断。被框着的区域表示推断的加工前的信号序列和跨膜区。下划线的是半胱氨酸残基，在疱疹病毒糖蛋白 B 序列中高度保守的残基下标星号 (*)。只在 KSHV 糖蛋白 B 中出现的半胱氨酸下标圆点 (·)。

图 4 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列，示出了保守区域，和由此设计出的 Type1 寡核苷酸 FRFDA。

图 5 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列，示出了保守区域，和由此设计出的 Type1 寡核苷酸 NIVPA 和 NIVPASQ。

图 6 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 TVNCA, TVNCB 和 TVNCBSQ.

图 7 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 FAYDA.

5 图 8 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 IYGKA 和 IYGKASQ.

图 9 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 CYSRA 和 CYSRASQ.

10 图 10 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 NIDFB 和 NIDFBSQ.

图 11 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 FREYA, FREYB 和 NVFDA.

图 12 列出了以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 示出了保守区域, 和由此设计出的 Type1 寡核苷酸 GGMA.

15 图 13 列出了得自 RFHV 和 KSHV 的糖蛋白 B 多核苷酸序列的部分, 排在一起的是以前已知的 γ 疱疹病毒糖蛋白 B 的多核苷酸序列, 每个共用的残基以小圆点表示.

20 图 14 列出了不同 γ 疱疹病毒的糖蛋白 B 多肽序列的比较, 所述蛋白质由多核苷酸序列中 NIVPA 和 TVNCB 的杂交位点之间的区域编码. 据推测下划线所示的 II 类序列片断是 RFHV/KSHV 交叉反应的抗原肽, 而以小写字母表示的 III 类序列是 RFHV 或 KSHV 病毒特异性的肽.

图 15 排列了较广范围的 α , β , 和 γ 亚族疱疹病毒中的糖蛋白 B 的多肽序列.

图 16 是以图 15 所示多肽序列为基础的糖蛋白 B 关系图谱.

25 图 17 列出了 Type2 (亚族-特异性) 寡核苷酸的例子, 排在一起的是得到它们的核苷酸序列.

图 18 是 KSHV 基因组中出现的糖蛋白 B 和 DNA 聚合酶的大致图谱, 显示了寡核苷酸引物的杂交位置.

30 图 19 列出了通过从图 1 所示序列的上游和下游扩增片断而得到的 KSHV DNA 序列. 显示了完整的 KSHV 糖蛋白 B 序列的开放阅读框, 其侧翼是衣壳成熟基因和 DNA 聚合酶的开放阅读框. 核苷酸序列中的下划线部分是推断的糖蛋白 B 启动子.

图 20 是 NIVPA 和 TVNCB 之间编码的 106 个核苷酸的 RFHV 糖蛋白 B 多肽片段的 Hopp - Woods 抗原性图，下面示出的是序列中疏水性和抗原性残基的跨度。

图 21 是 NIVPA 和 TVNCB 之间编码的 106 个核苷酸的 KSHV 糖蛋白 B 多肽片段的 Hopp-Woods 抗原性图。

图 22 是 KSHV 完整的糖蛋白 B 的 Hopp-Woods 抗原性图。

图 23 列出了被称为 RFHV2 的 RFHV/KSHV 亚族的第三个成员的糖蛋白 B 片段的 DNA 和蛋白质序列。下划线部分是残基 36 - 354 之间的 319 个碱基的多核苷酸区段，并表示用于扩增它的引物之间的糖蛋白 B 编码区段。

实施本发明的最佳方式

我们已从疱疹病毒 RFHV/KSHV 亚族中发现并鉴定了编码糖蛋白 B 的多核苷酸，本发明中包含的多核苷酸，寡核苷酸，多肽和抗体能用于诊断，临床监测和治疗疱疹病毒感染以及相关疾病。

RFHV 糖蛋白 B 的多核苷酸来源是取自患有腹膜后纤维瘤病（“RF”）的豚尾猴的被感染组织的样品，KSHV 糖蛋白 B 的多核苷酸得自取自患有卡波济氏肉瘤（“KS”）的人的被感染组织的样品。已知用于本发明的组织含有 RFHV 或 KSHV 的遗传物质，因为以前已成功地使用它们克隆了相应的编码 DNA 聚合酶的片断。在共有的美国专利申请 60/001, 148 中描述了 DNA 聚合酶区域的扩增。

为了从这些样品中扩增糖蛋白 B 序列，我们由其它疱疹病毒的序列设计了寡核苷酸。由于糖蛋白 B 暴露于病毒包膜的外部，从而处于它们所感染的宿主的免疫系统的选择压力之下，因此预计疱疹病毒之间的糖蛋白 B 较不保守。相应地，由据信与 RFHV 和 KSHV 最密切相关的疱疹病毒的序列设计了寡核苷酸。由 DNA 聚合酶序列可以知道这两个病毒是密切相关的 γ 型疱疹病毒。

主要从先前已知的四个 γ 疱疹病毒：sHV1, eHV2, bHV4, mHV68 和 hEBV 的糖蛋白 B 序列设计了寡核苷酸，这四个糖蛋白 B 分子的氨基酸序列的比较揭示了九个相对保守的区域。根据序列资料构建的寡核苷酸含有简并区段和共有区段，这一点将在以下段落中描述。已经将这些寡核苷酸中的三个用作扩增反应中的引物，所述反应已从 RF 和 KS 组织中产生了 RFHV 和 KSHV 糖蛋白 B 编码区段的片断。

最终的扩增反应步骤之后得到的 RFHV 和 KSHV 多核苷酸序列片段示于图 1 (分别为 SEQ. ID NO: 1 和 SEQ. ID NO: 3), 所包括的区段的每一末端对应于扩增反应中所用 NIVPA 和 TVNCB 引物的杂交区域。引物结合区段之间的片断长为 319 个碱基对(残基 36 - 354), 据信它是 RFHV 和 KSHV 基因组的各个糖蛋白 B 编码区域序列的准确反映。

得自 RFHV 的编码糖蛋白 B 的 319 个碱基对长的多核苷酸区段与 sHV1 和 bHV4 的相应部分仅有 60% 相同, 后两个序列是 RFHV/KSHV 亚族以外的最密切相关的序列。得自 KSHV 的 319 个碱基对长的多核苷酸区段与 sHV1 和 bHV4 的相应部分仅有 63% 相同。RFHV 和 KSHV 之间的此区段有 76% 相同。

还显示了相应的推测的氨基酸序列 (SEQ. ID NO: 2 和 SEQ. ID NO: 4)。多肽序列是新的, 并且与其它疱疹病毒的糖蛋白 B 序列部分同源, 推测所示片断约为整个糖蛋白 B 序列的 1/8, 它们约从加工前蛋白质的推测的 N-端甲硫氨酸下游 80 个氨基酸处开始。根据序列 Asn-Xaa- (Thr/Ser), 在氨基酸序列的第 80 位有一个潜在的 N-键糖基化位点, RFHV 和 KSHV 之间的此位点是保守的, 在其它已知的 γ 疱疹病毒中此位点也是保守的。在第 58 位也有一个半胱氨酸残基, 此残基在 γ , β , 和 α 亚族的疱疹病毒之间是保守的, 可能对于维持蛋白质的三维结构起作用。

RFHV 和 KSHV 之间由扩增引物之间的 319 个碱基对编码的 106 个氨基酸长的糖蛋白 B 区段有 91% 相同, 但在 KSHV 和 bHV4 的相应部分之间只有 65% 相同, bHV4 的所述序列是 RFHV/KSHV 亚族以外最接近的序列。

预期由 RFHV/KSHV 疱疹病毒亚族表达的糖蛋白 B 分子具有其它疱疹病毒糖蛋白 B 所具有的很多特性。糖蛋白 B 分子的大小一般约为 110kDa, 对应于约 800 - 900 个氨基酸或约 2400 - 2700 个碱基对。疏水性作图显示出从 N 末端至 C 末端以下列次序排列的区域: 对应于导向膜的前导序列的疏水区; 对应于细胞外区的混合极性区; 对应于跨膜区的疏水区; 对应于胞质区的另一个混合极性区。

图 19 所示的 KSHV 糖蛋白 B 的完整序列进一步证实了这些推测: 此基因编码了包括 C 末端附近的信号肽和跨膜区的约 845 个氨基酸。KSHV 和其它糖蛋白 B 序列之间的半胱氨酸残基是保守的, 另一个潜在的二硫键可有助于稳定三维结构。

糖蛋白 B 一般在感染性和缺损的病毒颗粒包膜, 和被感染的细胞表面表达。它一般是糖基化的, 可含有 5-20 或更多个糖基化位点。它一般被表达成蛋白质二聚体, 在病毒芽殖前转运至宿主细胞表面的过程中装配。负责二聚化的位点似乎位于约氨基酸 475 至跨膜区段之间(Navarro 等人)。

以往的研究已经绘出涉及糖蛋白 B 分子不同区域的感染性的几个生化功能图。糖蛋白 B 和糖蛋白 C 都涉及 HSV1 和牛疱疹病毒 1 型与靶细胞的最初的结合 (Herold 等人, Byme 等人)。由糖蛋白 B 识别的细胞上的组成成分似乎是硫酸乙酰肝素; 通过液相肝素可以抑制结合。缺乏糖蛋白 C 的突变体仍可结合靶细胞, 但对于既缺乏糖蛋白 C 又缺乏糖蛋白 B 的突变体而言, 其进入细胞的能力受到严重损害。

糖蛋白 B 的另一个明显重要的功能是促进膜融合和病毒进入细胞的能力。在人 CMV 中, 融合作用似乎由糖蛋白 B 的第一个疏水区来担当, 该作用可能与此区域内部保守的甘氨酸残基有关 (Reschke 等人)。在 HSV1 突变体中, 糖蛋白 B 促进合胞体形成的能力归功于蛋白质胞质区 C 末端附近的多个位点 (Kostal 等人)。

为了行使这些更复杂的功能中的一些功能, 糖蛋白 B 看上去似乎不仅与第二个糖蛋白 B 分子相关, 而且与病毒编码的其它成分相关, 例如, 糖蛋白 B 似乎需要 UL45 基因产物来诱导融合 (Haanes 等人)。已经有人假设糖蛋白 B 与其它表面蛋白合作以在靶细胞表面形成疏水融合微孔 (Pereira 等人)。已发现糖蛋白 B 能产生潜在的可中和完整病毒的抗体应答, 具有中和活性的单克隆抗体可针对糖蛋白 B 分子上的多个不同位点。

因而, 预期糖蛋白 B 分子携有的位点可与靶细胞相互作用, 可有助于促进融合, 并与其它病毒蛋白质相关。预测 RFHV/KSHV 亚族病毒的糖蛋白 B 分子会在其它种疱疹病毒中行使糖蛋白 B 的很多功能, 并且携有具有一些相同特性的活性区域。用药物, 抗体, 或通过突变干扰这些活性区域中的任何一个可能会有损于病毒的感染性或毒力。

在发现了 RFHV 和 KSHV 的糖蛋白 B 之后, 在猕猴的被感染的组织样品中鉴定出 RFHV/KSHV 亚族的第三个成员 (实施例 12), 该成员的糖蛋白 B 与 RFHV 的密切相关但不相同, 将该成员称为 RFHV2。预测 RFHV/KSHV 亚族的其它成员也会出现, 包括一些对人是病原体的成员。这一发现的教导是如何检测和鉴定亚族的新成员。

RFHV/KSHV 亚族内糖蛋白 B 序列之间的同源性意味着本发明中包含的多核苷酸和多肽是亚族的不同毒株中可靠的标记物。本发明中包含的多核苷酸，多肽和抗体在这种应用中可用于检测和治疗个体中因 RFHV, KSHV, 或相同亚族的其它疱疹病毒造成的病毒感染。在下文中将更详细地描述与糖蛋白 B 相关的多核苷酸，寡核苷酸探针，多肽，抗体和疫苗组合物，以及这些化合物的制备和用途。

缩写

本文使用的下列缩写指的是疱疹病毒的种类，以及从中得到的多核苷酸和多肽：

表 1: 疱疹病毒株的缩写		
名称	病毒	暂时的亚族排布
RFHV	与猕猴腹膜后纤维瘤病相关的疱疹病毒	γ -疱疹病毒
KSHV	与人卡波济氏肉瘤相关的疱疹病毒	
mHV68	鼠疱疹病毒 68 型	
bHV4	牛疱疹病毒 4 型	
eHV2	马疱疹病毒 2 型	
sHV1	松鼠猴疱疹病毒 1 型	
hEBV	人 Epstein-Barr 病毒	
hCMV	人巨细胞病毒	β -疱疹病毒
mCMV	鼠巨细胞病毒	
gpCMV	豚鼠巨细胞病毒	
hHV6	人疱疹病毒 6 型	α -疱疹病毒
hVZV	人水痘-带状疱疹病毒	
HSV1	人单纯疱疹病毒 1 型	
HSV2	人单纯疱疹病毒 2 型	
sHVS A8	猴疱疹病毒 A8 型	
eHV1	马疱疹病毒 1 型	
iHV1	ictalurid catfish 疱疹病毒	

10

一般性的定义

“糖蛋白 B” 是疱疹病毒的特定的蛋白质成分，它由病毒基因组编



码，据信表达于完整病毒的表面。对某些种类的疱疹病毒，尤其是 HSV1, hCMV 和牛疱疹病毒 1 型进行的功能性研究已在有关病毒感染性的多种生化功能中涉及到糖蛋白 B，这些功能包括与靶细胞表面的成分，如硫酸乙酰肝素结合，病毒膜与靶细胞的膜融合，病毒衣壳穿入细胞内，和多核化胞体细胞的形成。据观察糖蛋白 B 是同二聚体，它可以与其它病毒表面蛋白相互作用以行使它的一些生化功能。不同的生化功能，尤其是硫酸乙酰肝素结合和膜融合，似乎应归功于糖蛋白 B 分子的不同部分。包括 RFHV/KSHV 亚族成员的其它疱疹病毒的糖蛋白 B 分子可行使这些功能中的任何一部分或全部。本文所用的术语糖蛋白 B 包括非糖基化的，部分糖基化的，和完全糖基化的形式，既包括单体也包括多聚体。

本文所用的糖蛋白 B 片断，区域，或区段是糖蛋白 B 分子的片断，或编码糖蛋白 B 的多核苷酸亚区的转录本。完整的糖蛋白 B 分子，或全长的转录本将行使如上文所述的那些与病毒活性相关的生化功能。这些功能中的一些或全部可在此片断上被保留，或此片断可来自不能依靠自身行使这些功能的完整分子的一部分。

“糖蛋白 B 活性”指的是糖蛋白 B 的任何生化功能，或疱疹病毒的可归功于糖蛋白 B 的任何生物活性。这些可包括但不局限于蛋白质与细胞，如硫酸乙酰肝素的细胞受体，和受体类似物的结合；病毒结合或穿透细胞，或细胞融合。

术语“糖蛋白 B 基因”指的是含有编码如上文所定义的糖蛋白 B 分子之序列的基因。应理解糖蛋白 B 基因可产生经加工和经改变的翻译产物，包括但不局限于具有或不具有信号或前导序列的糖蛋白 B 形式，截短的或内部缺失的形式，多聚体的形式，和具有不同程度的糖基化的形式。

本文所用的“DNA 聚合酶”是蛋白质或蛋白质类似物，在适当条件下能够催化 DNA 多核苷酸与序列的装配，所述序列与用作模板的多核苷酸互补。DNA 聚合酶也可以含有其它催化活性，如 3'-5' 外切核酸酶活性；任何一种活性都可能占优势。DNA 聚合酶可能需要与附加的蛋白质或辅-因子连接以行使其催化功能。

“RFHV”是在患有腹膜后纤维瘤病 (RF) 的豚尾猴组织样品中检测到的疱疹病毒家族的病毒，“RFHV”是术语“RFHV1”，“RFHVM_n”和“RFM_n”的同义词。“KSHV”是在患有卡波济氏肉瘤 (KS) 的人的组织



样品中检测到的疱疹病毒家族的病毒。RFHV/KSHV 亚族的第三成员是在猕猴中鉴定出的病毒，本文中将此病毒称作“RFHV2”，“RFHV2”是术语“RFHVMm”和“RFMm”的同义词。

“RFHV/KSHV 亚族”这一术语用于本文是指能感染脊椎动物种类的疱疹病毒的总和，组成这一亚族的成员所具有的糖蛋白 B 序列与 RFHV 或 KSHV 的相应序列之间的相关性相对于与其它任何疱疹病毒之间的相关性而言，前者更加密切，所述其它疱疹病毒包括 sHV1，eHV2，bHV4，mHV68 和 hEBV。优选编码糖蛋白 B 的多核苷酸含有在残基 36 和 354 之间与 RFHV (SEQ. ID NO: 1) 或 KSHV (SEQ. ID NO: 3) 的相应部分至少有 65% 相同的区段；或与寡核苷酸 SHMDA 至少约有 74% 相同，或与寡核苷酸 CFSSB 至少约有 73% 相同，或与核苷酸 ENTFA 至少约有 72% 相同，或与核苷酸 DNIQB 至少约有 80% 相同。RFHV 和 KSHV 是 RFHV/KSHV 亚族成员的例子，RFHV/KSHV 亚族代表一亚组 γ 亚族的疱疹病毒。

术语“多核苷酸”和“寡核苷酸”可以互换使用，它们指的是任何长度的核苷酸的聚合形式，所述核苷酸或者是脱氧核糖核苷酸或者是核糖核苷酸，或者是它们的类似物。多核苷酸可以具有任何三维结构，并可行使已知或未知的任何功能。以下是多核苷酸非限制性的例子：基因或基因片段，外显子，内含子，信使 RNA (mRNA)，转运 RNA，核糖体 RNA，核酶，cDNA，重组多核苷酸，分支多核苷酸，质粒，载体，任何序列分离的 DNA，任何序列分离的 RNA，核酸探针和引物。多核苷酸可含有经修饰的核苷酸，如甲基化的核苷酸和核苷酸类似物。如果存在修饰，则可在聚合物装配之前或之后对核苷酸结构进行修饰。非核苷酸成分可打断核苷酸的序列，聚合之后可通过例如与标记成分结合进一步修饰多核苷酸。

本文使用的术语多核苷酸既可指双链也可指单链分子。除非另有特别说明或需要，本文所述发明的任何实施方案中的多核苷酸既包含双链形式，也包含已知或预测组成此双链形式的两个互补单链形式中的每一种。

在多核苷酸的上下文中，“线性序列”或“序列”是多核苷酸中 5' 至 3' 方向上的核苷酸顺序，其中在序列中互为邻居的残基在多核苷酸的一级结构中是邻接的。“部分序列”是已知在一个或两个方向上含有额外残基的多核苷酸之一部分的线性序列。



“杂交”指的是一种或多种多核苷酸反应形成复合物的反应，所述复合物通过核苷酸残基碱基之间的氢键得以稳定。Watson-Crick 碱基配对，Hoogsteen 结合或任何其它序列-特异性方式都可能产生氢键键合。复合物可含有形成双螺旋结构的两条链，形成多链复合物的三条或更多条链，单条自身杂交链或任何它们的组合。在更加外延的方法，如 PCR 的起始，或通过核酶酶促裂解多核苷酸中，杂交反应可构成一个步骤。

可在不同的“严紧”条件下进行杂交反应。增加杂交反应严紧性的条件已广为人知并已在本技术领域公开，例见 Sambrook Fritsch & Maniatis。相关条件的例子包括（以增加严紧性的次序）：保温温度为 25℃，37℃，50℃ 和 68℃；缓冲液浓度为 10×SSC，6×SSC，1×SSC，0.1×SSC（其中 SSC 是 0.15M NaCl 和 15mM 柠檬酸缓冲液）和使用其它缓冲液系统的它们的等同物；甲酰胺浓度为 0%，25%，50% 和 75%；保温时间为 5 分钟至 24 小时；1，2，或更多次的洗涤步骤；1，5，或 15 分钟的洗涤保温时间；洗涤溶液是 6×SSC，1×SSC，0.1×SSC 或去离子水。

“ T_m ”是实验条件下 50% 多核苷酸双螺旋解离成单链时的摄氏温度，所述双螺旋由通过 Watson-Crick 碱基配对在反平行方向以氢键键合的互补链组成。根据标准公式可预测 T_m ，例如：

$$T_m = 81.5 + 16.6 \log [Na^+] + 0.41 (\%G/C) - 0.61 (\%F) - 600/L$$

其中 Na^+ 是以 mol/L 表示的阳离子浓度（通常为钠离子）；(%G/C) 是以在双螺旋中占总残基的百分比表示的 G 和 C 残基的数目；(%F) 是溶液中甲酰胺的百分含量 (wt/vol)；L 是双螺旋的每一链中核苷酸的数目。

多核苷酸“稳定的双螺旋”或生化反应中任何两个或多个成分之间形成的“稳定的复合物”指的是在双螺旋或复合物的形成和其随后的检测之间能维持足够长时间的双螺旋或复合物。双螺旋或复合物无论在何种条件存在下都应该能维持，或在形成时刻和检测时刻之间的某一时刻被导入，这些条件是所进行的检测或反应的参数。非强制性存在的能破坏双螺旋或复合物的有关条件包括洗涤，加热，在反应混合物中加入额外的溶质或溶剂（如变性剂），和与多余的反应类别竞争。稳定的双螺旋或复合物可以是不可逆的，也可以是可逆的，但必须满足此定义的其



它必需条件。因此，反应混合物中可能形成瞬时复合物，但是如果它自动解离或者由于在检测前引入新增加的条件或操作而发生解离，就不能构成稳定的复合物。

5 当在两个单链多核苷酸之间以反平行的构型形成稳定的双螺旋时，特别是在高度严紧的条件下，链基本上是“互补的”。双链多核苷酸可以与另一个多核苷酸“互补”，条件是第一多核苷酸中的一条链和第二多核苷酸中的一条链之间能形成稳定的双螺旋。从单链多核苷酸序列推测的互补序列是根据广泛为人接受的碱基配对规则预期与单链多核苷酸形成氢键键合的标准核苷酸的最适序列。

10 “有义”链和“反义”链当用于相同的上下文中时指的是相互之间互补的单链多核苷酸。它们可以是双链多核苷酸的相反链，或者根据广泛为人接受的碱基配对规则由一条链推测另一条链。如果不是特指，可任意地将一条链或其它链指定为“有义”或“反义”链，至于编码肽的多核苷酸区段，“有义”链一般指的是含有编码区段的链。

15 当在多核苷酸之间进行相同程度比较时，只能理解为容易产生互补链，选择有义或反义链，或者预测它们能使被比较的多核苷酸之间的相同程度最大化。例如，当被比较的多核苷酸中的一个或两个是双链时，如果第一多核苷酸的一条链与第二多核苷酸的一条链相同，则序列相同。类似地，当多核苷酸探针被描述成与其靶相同时，应理解它是靶的
20 互补链，所述互补链参与了探针和靶之间的杂交反应。

如果两个序列都能与相同的互补多核苷酸杂交形成双螺旋，则其中一个核苷酸的线性序列与另一个线性序列“基本上相同”，更优选能在较严紧的条件下杂交的序列。应理解杂交反应可容纳核苷酸序列中的插入，缺失和取代。因此即使一些核苷酸残基不能精确地对应或匹配，核苷酸的线性序列也基本上相同，相比之下更优选与本发明公开的更加对应或匹配的序列。通常约为 25 个残基的多核苷酸区域与另一个区域基本上相同，条件是序列至少约 85% 相同；更优选序列至少约 90% 相同；更优选序列至少约 95% 相同；再更优选序列为 100% 相同。在同源部分
25 相互匹配之后，40 或更多个残基的多核苷酸区域与另一个区域基本上相
30 同，条件是序列至少约 75% 相同；更优选序列至少约 80% 相同；更优选序列至少约 85% 相同；甚至更优选序列至少约 90% 相同；再更优选序列为 100% 相同。



在测定多核苷酸序列是否基本上相同时，特别优选保留了多核苷酸的功能性并以此进行比较的序列。可通过不同的参数测定功能性，例如，如果多核苷酸用于涉及与另一个多核苷酸杂交的反应中，则优选的序列为那些在类似条件下与相同靶杂交的序列。通常，对于 200 或更多个残基的双螺旋而言，序列的同一性每减少 1%，DNA 双螺旋的 T_m 约降低 1℃；或对于少于 40 个残基的双螺旋而言约降低 5℃，这取决于错配残基的位置（例见 Meinkoth 等人）。一般而言，约 100 个残基的基本上相同的序列在低于 T_m 约 20℃ 时能与彼此各自互补的序列形成稳定的双螺旋；优选它们能在低于 T_m 约 15℃ 时形成稳定的双螺旋；更优选它们能在低于 T_m 约 10℃ 时形成稳定的双螺旋，甚至更优选它们能在低于 T_m 约 5℃ 时形成稳定的双螺旋；再更优选它们能在约 T_m 时形成稳定的双螺旋。在另一个实施例中，如果由多核苷酸编码的多肽是其功能性的重要部分，则优选的序列为那些编码相同或基本上相同的多肽的序列。因此能导致保守的氨基酸取代的核苷酸差异比那些导致非保守取代的核苷酸差异更加优选，更优选不会改变氨基酸序列的核苷酸差异，而甚至更优选相同的核苷酸。导致多肽插入或缺失的多核苷酸插入或缺失比那些导致下游编码区移码的多核苷酸插入或缺失更加优选，甚至更优选不含插入或缺失的多核苷酸序列。杂交特性和多核苷酸编码的多肽序列的相对重要性取决于本发明的应用。

多核苷酸具有与另一个多核苷酸相同的“特性”，条件是这两个多核苷酸在类似的最高严紧度的条件下，都能与特殊的第三个多核苷酸形成稳定的双螺旋。优选除了类似的杂交特性外，多核苷酸也编码基本上相同的多肽。

多核苷酸序列的“保守”残基是那些不变地出现在被比较的两个或多个相关序列的相同位置上的残基。相对保守的残基是那些在更相关的序列中保守或相对于在序列别处出现的残基而言具有更程度的相同性的残基。

“相关的”多核苷酸是共用的相同残基占相当比例的多核苷酸。

本文所用的“简并”寡核苷酸序列是按下述衍生自至少两个相关的来源多核苷酸序列的经设计的序列：在来源序列中保守的残基被保留在简并序列中，而在来源序列中不保守的残基可以简并序列中的几个可替代物提供。例如，可从来源序列 ATACA 和 ACAGA 出发设计简并序列



AYASA, 其中 Y 是 C 或 T, S 是 C 或 G。Y 和 S 是“多义”残基的例子。
简并区段是含有简并序列的多核苷酸区段。

5 应理解含有简并序列的合成寡核苷酸实际上是除了在多义位置以外
共用相同序列的密切相关的寡核苷酸混合物。这种寡核苷酸通常被合成
为多义位置处的核苷酸的所有可能的组合的混合物, 混合物中每种寡核
苷酸被称作“可替换的形式”, 混合物中该形式的数目等于

$$\prod_{i=1}^n k_i$$

10 其中 k_i 是每个位置处允许的可替换核苷酸的数目。

本文所用的“共有”寡核苷酸序列是按下述衍生自至少两个相关的
来源多核苷酸序列的经设计序列: 在所有来源序列中保守的残基被保留
在共有序列中; 而在残基不保守的位置处, 从来源序列中选择一个替
15 代物。通常选择的核苷酸是在来源序列中以最高频率出现的核苷酸。例
如, 可由来源序列 CAAAA, AAGAA 和 AAAAT 出发设计共有序列 AAAAA,
共有区段是含有共有序列的多核苷酸区段。

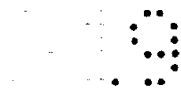
本文所用的多核苷酸“片段”或“插入物”通常表示全长形式的次
分区域, 但也可包括完整的全长多核苷酸。

20 如果据信多核苷酸可以互相衍生或来自共同的祖先, 则多核苷酸相
互“对应”。例如, 如果不同病毒基因的编码区有显著水平的同一性,
在基因组图谱上处于相同的位置, 或编码行使类似生化功能的蛋白质,
则可认为它们相互对应。信使 RNA 对应于它所转录的基因, cDNA 对应于
产生它的 RNA, 以及编码 RNA 的基因。蛋白质对应于编码它的多核苷酸,
和能与它特异性结合的抗体。

25 在多核苷酸操作上下文中所用的“探针”指的是以试剂的形式提供
以通过与靶杂交检测受试样品中潜在存在的靶的寡核苷酸。通常探针会
含有标记物或者使标记物在杂交反应之前或之后能被结合的物质。适当
的标记物包括但不限于放射性同位素, 荧光染料, 化学发光化合物, 染
料和包括酶的蛋白质。

30 “引物”是通常具有游离的 3'-OH 基团的寡核苷酸, 它通过与靶杂
交来结合受试样品中潜在存在的靶, 然后促使与靶互补的多核苷酸聚合。

产生相同多核苷酸的复制拷贝的方法, 如 PCR 或基因克隆在本文中



被统称为“扩增”或“复制”。例如，可以复制单链或双链 DNA 以形成具有相同序列的另一个 DNA。例如可通过 RNA 指导的 RNA 聚合酶，或通过逆转录 DNA 然后进行 PCR 来复制 RNA。在后一种情况下，RNA 的扩增拷贝是具有相同序列的 DNA。

5 “聚合酶链反应”（“PCR”）是使用一种或多种引物和聚合作用催化剂，如逆转录酶或 DNA 聚合酶，特别是热稳定性的聚合酶由靶多核苷酸制成复制拷贝的反应。通常 PCR 涉及反复形成的三个步骤：“退火”，其中温度被调节，以使寡核苷酸引物能与欲被扩增的多核苷酸形成双螺旋；“延伸”，其中温度被调节，以使通过 DNA 聚合酶，使用它们已与之形成双螺旋的多核苷酸为模板，使得已形成双螺旋的寡核苷酸被延
10 伸；和“解链”，其中温度被调节，以使多核苷酸和被延伸的寡核苷酸解离。然后重复此循环直至得到所需量的经扩增的多核苷酸。PCR 方法参见美国专利号 4,683,195 (Mullis) 和 4,683,202 (Mullis 等人)。

“调控元件”或“调控序列”是涉及对多核苷酸的功能性调节有贡献的分子相互作用的核苷酸序列，所述作用包括多核苷酸的复制，重复，
15 转录，剪接，翻译，或降解。调节可以影响到进程的频率，速度或特异性，也可以在性质上逐渐增强或被抑制。调控元件是本技术领域已知。例如，“启动子”就是一例调控元件。启动子是能在某些条件下结合 RNA 聚合酶并启动位于启动子下游（3'方向）的编码区转录的 DNA 区域。
20 “有效连接”指的是遗传元件的并列，其中元件的排列关系应能允许它们以预期的方式起作用。例如，如果启动子帮助启动编码序列的转录，则启动子与编码区域有效地连接，在启动子和编码区域之间应存在有相关的残基，只要这种功能关系能得以维持。

术语“多肽”，“肽”和“蛋白质”在本文中可互换使用，意指任何长度的氨基酸的聚合物。聚合物可以是线性的或分支的，它可含有经
25 修饰的氨基酸，可被非氨基酸打断。此术语也包含已被自然修饰或通过间插修饰过的氨基酸聚合物；修饰可包括例如，二硫键的形成，糖基化，脂化，乙酰化，磷酸化，或任何其它操作，如与标记物成分缀合。

在多肽的上下文中，“线性序列”或“序列”是多肽中 N 末端至 C
30 末端方向上的氨基酸顺序，其中序列中互相邻接的残基在多肽的一级结构中是邻接的。“部分序列”是已知在一个或两个方向上含有额外残基的多肽的一部分的线性序列。

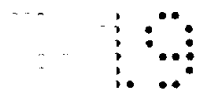


如果氨基酸的线性序列与另一个序列具有相当程度上的序列同一性，则这两个序列“基本上相同”。应理解蛋白质的折叠和生化功能可容纳氨基酸序列中的插入，缺失和取代。因此，即使一些残基不能精确地对应或匹配，氨基酸的线性序列也基本上相同，更加优选与本发明公开的序列更紧密对应或匹配的序列。还应理解一些氨基酸的取代更容易忍受。例如由具有类似特性侧链的另一个氨基酸取代具有疏水侧链，芳香侧链，极性侧链，具有正或负电荷的侧链，或含有两个或较少碳原子的侧链的氨基酸不会影响两个序列基本上相同的特性。测定同源区域和评价同源性的方法是本技术领域熟知的；例见 Altschul 等人和 Henikoff 等人。较能忍受的序列差异指的是“保守的取代”，因此具有保守取代的序列比那些在相同位置具有其它取代的序列更加优选；更优选在相同位置处具有相同残基的序列。通常，在同源部分排列之后，多肽区域与另一个区域基本上相同的条件是：序列至少约 92% 相同；更优选至少约 95% 相同；更优选它们至少约 95% 相同，并含有至少另一个 2% 的相同或保守取代；更优选它们至少约 97% 相同；更优选它们至少约 97% 相同，并含有至少另一个 2% 的相同或保守取代；更优选它们至少约 99% 相同；再更优选序列 100% 相同。

在测定是否多肽序列基本上相同时，特别优选保留了与其进行比较的多肽的功能性的序列。可通过不同的参数确定功能性，如酶活性，底物-酶或受体-配体相互作用中的结合速率或亲和性，与抗体的结合亲和性和 X-射线晶体结构。

如果多肽与另一个多肽显示出相同的生化功能，如酶活性，配体结合或抗体反应性，则这两个多肽具有相同的“特性”。与糖蛋白 B 或糖蛋白 B 片段相关之多肽的优选特性是：结合被其它疱疹病毒种类的糖蛋白 B 结合的细胞表面受体类似物的能力，促进膜与靶细胞融合的能力，促进病毒穿透宿主细胞的能力。也优选与所比较的多肽显示出相同生化功能的多肽，另外还优选据信通过计算机模拟预测或通过如 X-射线晶体学这类技术测定，与所比较的多肽具有类似三维构象的多肽。

多肽的“生化功能”，“生物学功能”或“生物活性”包括通过适当的实验研究可检测到的多肽的任何特征。“改变的”生化功能指的是通过例如分子量测定，圆二色性，抗体结合，差异光谱学或核磁共振可以检测到的多肽一级，二级，三级或四级结构的改变。它也可以指反应



性的改变，如催化某一反应的能力，或结合辅因子，底物，抑制物，药物，半抗原或其它多肽的能力。如果某物质以这些方式中的任一种改变了多肽的生化功能，就可以说此物质“干扰”了多肽的生化功能。

“融合多肽”是与天然出现的多肽相比，含有序列中不同位置处的区域的多肽。此区域可以在不同的蛋白质中正常地存在，并在融合多肽中被连在一起；或者它们也可以正常地存在于相同的蛋白质中，但在融合多肽中以新的安排被放置。可通过例如化学合成，或通过以所需的关系产生和翻译编码肽区域的多核苷酸来产生融合多肽。

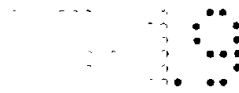
“抗体”是能够通过位于免疫球蛋白分子可变区上的至少一个抗原识别位点与靶（如多肽）特异性结合的免疫球蛋白分子。本文所用的术语不仅包含完整的抗体，也包含其片段，其突变体，融合蛋白，人源化抗体，和含有所需特异性的抗原识别位点的免疫球蛋白分子的任何其它经修饰的构型。

“免疫识别”或“免疫反应性”指的是通过免疫球蛋白或相关分子上的至少一个抗原识别位点，如B细胞受体或T细胞受体与靶特异性地结合。

术语“抗原”指的是抗体通过其抗原识别位点特异性结合的靶分子。抗原可以但不必需在化学上与刺激抗体生成的免疫原相关。抗原可以是多价的或单价的半抗原。可被抗体识别的多种抗原的例子包括多肽，多核苷酸，其它抗体分子，寡糖，复合脂，药物和化合物。

“免疫原”是当注射到适当宿主，通常为哺乳动物体内时能够刺激抗体生成的抗原。具有此特性的化合物被描述成“免疫原性的”。通过本技术领域已知的各种技术可使化合物变成免疫原性的化合物，所述技术包括与载体交联或缀合以增加价位，与促细胞分裂原混合以增强免疫应答，和与佐剂联合以增强呈递作用。

“疫苗”是供人或动物使用的药物制品，施用疫苗的目的是赋予受体抗特殊靶或靶组一定程度的特异性免疫反应性。免疫反应性可以是对靶具有免疫反应性的抗体或细胞（尤其是B细胞，浆细胞，T辅助细胞和细胞毒T淋巴细胞和它们的前体），或它们的任何组合。可能的靶包括外来的或病理性的化合物，如外源蛋白质，病原性病毒或由癌细胞表达的抗原。免疫反应性对于实验目的，治疗特殊疾病，消除特殊物质或预防特殊疾病或物质是合乎需要的。除非特别地指出，本文所指的疫苗



可以是被动疫苗，也可以是主动疫苗，或者同时具有两者的特性。

“被动疫苗”是不需要受体免疫应答参与以产生其效果的疫苗。通常它由对靶反应的抗体分子组成。抗体可得自供体受试者，并被充分纯化以施用给受体，或者它们也可在体外产生，例如来自杂交瘤细胞培养物，或者通过经遗传工程改造的编码抗体分子的多核苷酸产生。

“主动疫苗”是经施用意欲在受体体内产生特异性免疫应答的疫苗，所述免疫应答则具有合乎需要的针对靶的免疫反应性。主动疫苗含有适当的免疫原。合乎需要的免疫应答可以是体液的或细胞的，全身性的或分泌性的，或它们的任何组合。

“试剂”多核苷酸，多肽或抗体是为反应提供的物质，此物质具有一些已知的和合乎反应需要的参数。

反应混合物中也可含有试剂能与之反应的“靶”，如多核苷酸，抗体或多肽。例如在一些类型的诊断试验中，通过加入试剂，使试剂和靶反应，并测量反应产物的量可测定样品中靶的量。在临床处理的上下文中，“靶”也可以是被施用物质，如药物化合物之目标的细胞、细胞集合、组织或器官。身为病毒感染靶的细胞是病毒为了复制或转化成潜伏形式的目的而优先定位的细胞。

“分离的”多核苷酸，多肽，蛋白质，抗体，或其它物质指的是该物质制品缺乏至少一些该物质或类似物质天然存在或最初取得时也可存在的其它成分。因此例如可通过使用纯化技术从来源混合物中富集以制备分离的物质。可在绝对的基础（如每体积溶液的重量）之上测量富集，或者相对于来源混合物中存在的潜在干扰的第二物质来测量富集。本发明实施方案中富集的增加更为优选。因此，例如优选 2-倍富集，更优选 10-倍富集，更优选 100-倍富集，甚至更优选 1000-倍富集。通过人工装配的方法，如通过化学合成或重组表达也可以分离的状态提供物质。

反应中使用的多核苷酸，如杂交反应中使用的探针，PCR 中使用的引物，或药物制品中存在的多核苷酸为“特异性的”或“选择性的”的条件是，相对于它与可替换的物质杂交或反应而言，它与想要的靶的杂交或反应更加频繁，更加快速，或者持续的时间更长。类似地，多肽为“特异性的”或“选择性的”的条件是：相对于它与可替换的物质的结合而言，它与想要的靶，如配体，半抗原，底物，抗体或其它多肽的结合更加频繁，更加快速，或持续的时间更长。抗体为“特异性的”或“选



择性的”的条件是：相对于它与可替换的物质的结合而言，它通过至少一个抗原识别位点与想要的靶的结合更加频繁，更加快速，或持续的时间更长。多核苷酸，多肽或抗体被说成“选择性地抑制”或“选择性地干扰”反应的条件是：相对于它抑制或干扰可替换的底物之间的反应而言，它抑制或干扰特殊底物之间的反应的程度更高或持续时间更长。

5 “候选药物”或“候选药剂”是据信具有治疗潜力的化合物，其效力有待检测。候选药物的“筛选”指的是进行能够评价候选药物的效力和/或特异性的试验。在这一上下文中，“效力”指的是候选药物以有利的方式影响细胞或它所施用的生物体的能力：例如，限制因侵入病毒
10 导致的病理学。

药物制品的“效应物成分”是当施用给提供细胞的受试者时，通过以合乎需要的方式改变靶细胞的功能以修饰靶细胞的成分。一些先进的药物制品也具有“靶向成分”，如抗体，它可协助将效应物成分更有效地传递到靶位点。依据所需要的作用，效应物成分可具有多种作用模式
15 中的任何一种。例如，它可以恢复或增强细胞的正常功能，它可消除或抑制细胞不正常的功能，或者它可改变细胞的表型。或者它可使具有病理学特征的细胞，如被病毒感染的细胞被杀死或变成休眠状态。以后的章节中会提供效应物成分的例子。

“细胞系”或“细胞培养物”表示在体外生长或维持的较高级的真核生物的细胞。应理解细胞的后代与母细胞不完全相同（无论是形态学，
20 基因型还是表型）。

“宿主细胞”是通过施用外源多核苷酸已经被转化，或者能够被转化的细胞。“宿主细胞”包括原始转化子的后代。

“遗传改变”指的是不通过自然的细胞分裂而使遗传因子导入细胞
25 的过程。所述遗传因子对于细胞而言可以是异源的，它也可以是细胞中已存在的遗传因子的多余的拷贝或改良的形式。通过例如本领域中已知的任何方法用重组质粒或其它多核苷酸转染细胞，所述方法如电穿孔，磷酸钙沉淀，与多核苷酸-脂质体复合物接触，或者通过用 DNA 或 RNA 病毒或病毒载体转导或感染能够实现遗传改变。这种改变优选、但并不
30 必需由经改变细胞的后代遗传下去。

“个体”指的是脊椎动物，特别是哺乳动物种的成员，它包括但不限于家养动物，野生动物和包括人的灵长类动物。



本文所用的术语“灵长类动物”指的是哺乳动物中进化地位最高的任何成员。它们包括（但不限于）原猴类，如狐猴和懒猴；跗猴类，如跗猴；新世界猴，如松鼠猴（*Saimiri sciureus*）和狨猴；旧世界猴，如猕猴（包括豚尾猴，食蟹猴，和日本猴）；长臂猿类，如长臂猿和合趾猴；猩猩类，如马来猩猩，大猩猩，和黑猩猩；和包括人的人科动物。

由疱疹病毒感染引起的“病理学”是损害宿主健康或正常生理学的一切事情。它可包括（但不限于）病毒破坏性地侵入先前未被感染的细胞，病毒在损害细胞正常代谢的情况下复制，病毒产生毒素或其它非天然的分子，细胞或细胞间结构的不规则生长（包括纤维变性），被感染细胞的不寻常或受抑制的生物活性，恶性转化，干扰邻近细胞的正常功能，炎症或免疫反应的加重或抑制，对其它病原生物体和疾病的易感性增加。

对个体或细胞的“治疗”是试图改变个体或细胞的自然进程的任何形式的介入。例如，进行对个体的治疗可以降低或限制由感染个体的疱疹病毒引起的病理学。治疗包括（但不限于）施用如药物组合物的组合物，可以在病理事件开始或与病原因子接触之后预防性地，也可以治疗性地进行。

应理解临床或生物“样品”包含得自受试者并可用于体外步骤，如诊断试验的各种样品类型。此定义包含以外科手术切除的方式得到的固体组织样品，病理学样本，或活体解剖样本，从中得到的组织培养物或细胞及其后代，从这些来源中的任一个制备的切片或涂片。非限制性的例子是得自被感染的部位，纤维化的部位，未受影响的部位和肿瘤的样品。此定义也包含血液，脊髓液和生物来源的其它液体样品，也可以指其中悬浮的细胞或细胞片段，或指液体介质及其溶质。此定义也包括已经被溶解或已经富集了某些成分，如 DNA，RNA，蛋白质或抗体的样品。

本文所述的寡核苷酸引物和探针已经按下述被命名：名称的第一部分是它们所根据的多肽保守区部分的单个氨基酸密码子，通常为 4 个残基长。紧随其后的是字母 A 或 B，分别表示寡核苷酸与 DNA 编码区的有义或反义链互补。用于测序和标记反应的次级共有寡核苷酸在名称末端具有字母 SQ。

一般技术

除非另有说明，本发明的操作将使用本技术领域熟知的分子生物



学，微生物学，重组 DNA 和免疫学的常规技术。这类技术在文献中有全面的解释，例见“Molecular Cloning: A laboratory Manual”，第二版（Sambrook, Fritsch & Maniatis, 1989），“Oligonucleotide Synthesis”（M. J. Gait, 编，1984），“Animal Cell Culture”（R. I. Freshney, 编，1987）；系列“Methods in Enzymology”（Academic Press, Inc.）；“Handbook of Experimental Immunology”（D. M. Weir & C. C. Blackwell, 编），“Gene Transfer Vectors for Mammalian Cells”（J. M. Miller & M. P. Calos, 编，1987），“Current Protocols in Molecular Biology”（F. M. Ausubel 等，编，1987）；和“Current Protocols in Immunology”（J. E. Coligan 等，编，1991）。

上文和下文提到的所有专利，专利申请，文章和出版物都已列入本文参考文献。

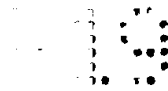
编码疱疹病毒 RFHV/KSHV 亚族糖蛋白 B 的多核苷酸

本发明包括衍生自疱疹病毒中存在的糖蛋白 B 基因的分离的多核苷酸区段，所述多核苷酸区段编码糖蛋白 B 多肽的片段。多核苷酸与疱疹病毒 RFHV/KSHV 亚族有关。列举的多核苷酸编码 RFHV 或 KSHV 的糖蛋白 B 片段。优选的片段包括那些按下文实施例所述得到的示于图 1 的片断及其亚片断。特别优选的多核苷酸含有 SEQ. ID NO: 1, SEQ. ID NO: 3, 或 SEQ. ID NO: 96 的残基 36 和 354 之间的序列，或 SEQ. ID NO: 92 中所含的多核苷酸。

RFHV 和 KSHV 的残基 36 和 354 之间的多核苷酸区段 76% 相同，共用残基在图 1 中由“*”表示，此区段内 RFHV 和 KSHV 之间共用的最长亚区的长度为 15, 17 和 20 个核苷酸。

相对于它们中的任一个与以前已经测序的任何疱疹病毒的相应区段相比，RFHV 和 KSHV 在扩增引物结合位点之间的 319 个碱基对的片段相互之间更加一致。下一个最密切相关的序列是 sHV1 和 bHV4 的有关序列，所述序列与 KSHV 的相应序列 63% 相同，与 RFHV 的相应序列 60% 相同。糖蛋白 B 片断和先前已测序的任何病毒之间共用的最长数目的连串碱基是 14。据信 RFHV 或 KSHV 序列 16 个碱基对或更长的任何亚片段对于 RFHV/KSHV 亚族而言，或者对于此亚族内特定的疱疹病毒种和变体而言是独特的。

本发明包括 RFHV/KSHV 亚族的糖蛋白 B 基因中所含的亚片段，优选

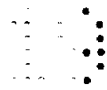


包含对应于 SEQ. ID NO: 1, SEQ. ID NO: 3, 或 SEQ. ID NO: 96 的残基 36 和 354 之间, 或 SEQ. ID NO: 92 中的任何位置处的 319 个碱基对的区域。优选亚片段至少约 16 个核苷酸长; 更优选它们至少长为 18 个核苷酸; 更优选它们至少长为 21 个核苷酸; 更优选它们至少长约为 25 个核苷酸; 更优选它们至少长约 35 个核苷酸; 再更优选它们至少长约 50 个核苷酸; 更加优选它们至少长约 75 个核苷酸, 甚至更优选它们为 100 个核苷酸长或更长。本发明中还包括含有每种独自の疱疹病毒糖蛋白 B 的整个开放阅读框的多核苷酸。

相对于 RFHV 或 KSHV 与表 1 所列的任何其它病毒的比较而言, 组成 RFHV/KSHV 亚族的成员所具有的序列与 RFHV 或 KSHV 的相应序列更加一致。根据图 1 相应区域中糖蛋白 B 基因的序列可鉴定出此家族中优选的成员。表 2 提供了此区域内序列同一性的程度:

表 2: KSHV 和其它疱疹病毒的糖蛋白 B 之间的序列同一性				
糖蛋白 B 序列	SEQ. ID NO:	与多核苷酸片断的同一性		
		RFHV (SEQ. ID NO: 1) 碱基 36 - 354	KSHV (SEQ. ID NO: 3) 碱基 36 - 354	
RFHV	1	(100%)	76%	
RFHV/KSHV 亚族	KSHV 3	76%	(100%)	
其它 γ 疱疹 病毒	sHV1 5	60%	63%	
	bHV4 6	60%	63%	
	eHV2 7	52%	54%	
	mHV68 8	56%	54%	
	hEBV 9	<50%	52%	
α 和 β 疱疹 病毒	hCMV 10	<50%	<50%	
	hHV6 11	<50%	<50%	
	hVZV 12	<50%	<50%	
	HSV1 13	<50%	<50%	

通过首先排列被编码的氨基酸序列, 进行编码多核苷酸的相应对比, 然后计算所比较的序列之间每个排列位置处共用的残基数目即可计算出序列同一性的百分比。插入或缺失的存在不会产生不利的后果, 但是只有在被排列的序列之一中需要容纳数目明显增加的氨基酸残基的地

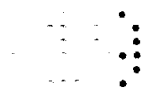


方才允许插入或缺失。无论在多肽水平上还是在多核苷酸水平上都不允许仅为了改善序列排列的补偿性插入。因此，多核苷酸序列中的任何插入的长度均为3的倍数。给出了受试序列中与比较或参照序列中的残基相同的残基的百分比。

本发明优选的编码糖蛋白B的多核苷酸序列是那些衍生自RFHV/KSHV 疱疹病毒亚族的序列。它们包括那些与RFHV或KSHV序列的碱基36和354之间至少65%相同的序列；更优选序列至少67%相同；更优选序列至少约70%相同；更优选序列至少约75%相同；更优选序列至少约80%相同；更优选序列至少约85%相同；更优选序列至少约90%相同；甚至更优选序列95%以上相同。也包括位于能满足所述同一性标准之区域的上游或下游的糖蛋白B编码区域。

下一章节将更详细地描述，通过与RFHV/KSHV亚族特异性的寡核苷酸（Type2寡核苷酸）相比之下的同一性百分比可鉴定出其它优选的编码糖蛋白B的多核苷酸序列。RFHV和KSHV糖蛋白B与Type2寡核苷酸实例之间的同一性百分比示于表3：

糖蛋白B序列	SEQ. ID NO:	与 SHMDA 的同一性 (SEQ. IDNO: 41)	与 CFSSB 的同一性 (SEQ. IDNO: 43)	与 ENTFA 的同一性 (SEQ. ID NO: 45)	与 DNIQB 的同一性 (SEQ. IDNO: 46)
RFHV	1	91%	91%	89%	91%
KSHV	3	100%	85%	89%	97%
sHV1	5	71%	70%	66%	66%
bHV4	6	57%	64%	69%	74%
eHV2	7	57%	61%	54%	60%
mHV68	8	<50%	55%	54%	77%
hEBV	9	57%	55%	60%	51%
hCMV	10	57%	55%	60%	51%
hHV6	11	< 50%	52%	60%	57%
hVZV	12	54%	58%	66%	57%
hSV1	13	57%	60%	54%	54%



为了序列排列的目的，可通过不允许寡核苷酸或多肽中有缺口计算出这一长度的寡核苷酸的同源性百分率。在本发明的所有内容中，每当被比较的两个序列中至少有一个是含有多义残基的简并寡核苷酸时，如果简并寡核苷酸的至少一个可替换形式与与之相比较的序列相同，则这两个序列相同。例如，由于AYAAA是ATAAA和ACAAA的混合物，所以AYAAA与ATAAA 100%相同。

优选的糖蛋白 B 编码序列是那些所覆盖的相应区域与 SHMDA 至少约 72% 相同的序列；更优选它们至少约 74% 相同；更优选它们至少约 77% 相同；更优选它们至少约 80% 相同；更优选它们至少 85% 相同，更加优选它们至少约 91% 相同。其它优选的糖蛋白 B 编码序列是那些所覆盖的相应区域与 CFSSB 至少约 71% 相同的序列；更优选它们至少约 73% 相同；更优选它们至少约 77% 相同；更优选它们至少约 80% 相同；更优选它们至少约 85% 相同。其它优选的糖蛋白 B 编码序列是那些所覆盖的相应区域与 ENTFA 至少约 70% 相同的序列，更优选它们至少约 72% 相同；更优选它们至少约 75% 相同；更优选它们至少约 80% 相同；更优选它们至少约 85% 相同，甚至更优选它们至少 89% 相同。其它优选的糖蛋白 B 编码序列是那些所覆盖的相应区域与 DNIQB 至少约 78% 相同的序列；更优选它们至少约 80% 相同；更优选它们至少约 85% 相同；更优选它们至少约 91% 相同。也包括位于能满足所述同一性标准之区域的上游或下游的糖蛋白 B 编码区。

通过使用 RFHV 或 KSHV 序列，或者亚族特异性寡核苷酸的前述序列比较中的任一种鉴定出的 RFHV/KSHV 亚族成员的糖蛋白 B 编码序列同样是优选的。例举的序列是 RFHV 和 KSHV 的糖蛋白 B 编码序列。本发明中还包括亚族的任何糖蛋白 B 编码序列的片段，和含有这种多核苷酸片段的较长的多核苷酸。

本章节描述的多核苷酸序列提供了得到以后章节中概述的合成的寡核苷酸，蛋白质和抗体的基础。使用本领域普通技术人员已知的一般性技术可制备这些化合物，如下文所述，所述化合物可用于各种研究，诊断和治疗目的。

30 多核苷酸的制备

可通过本领域中已知的任何适当的方法制备本发明的多核苷酸和寡核苷酸。例如，如下文实施例 3 和实施例 11 所述，在 PCR 扩增得自被



疱疹病毒感染之组织的 DNA 时，可以使用寡核苷酸引物。或者如下文实施例 8 所述，可以使用寡核苷酸鉴定 DNA 文库中适当的细菌克隆。

也可以通过化学合成由本文提供的序列出发直接制备多核苷酸。本领域中已知几种合成方法，包括三酯法和亚磷酸酯法。在优选的方法中，
5 通过使用单核苷氨基亚磷酸酯偶联单位的固相合成法制备多核苷酸，例见 Horise 等, Beaucage 等, Kumar 等, 和 U.S. 专利号 4,415,732。

典型的固相合成包括重复进行的四个步骤：脱保护，偶联，加帽和氧化。这可以导致在 3' 至 5' 方向上逐步合成寡核苷酸。

在第一步中，其 3' 末端通过 (-O-) 基团与固相支持物结合的逐渐
10 增长的寡核苷酸在其 5' 末端脱保护。例如，通过 -ODMT 基团可以保护 5' 末端，所述 -ODMT 基团是通过与吡啶中的 4,4'-二甲氧基三苯甲基氯 (DMT - Cl) 反应而形成的。此基团在碱性条件下是稳定的，但是在酸性条件下容易被除去，例如在二氯乙酸 (DCA) 或三氯乙酸 (TCA) 的存在下易被除去。脱保护提供了 5'-OH 反应基团。

15 在第二步中，寡核苷酸与所需的核苷酸单体反应，所述单体自身已先转变成 5'-保护的 3'-氨基亚磷酸酯。单体的 5'-OH 可以例如 -ODMT 基团的形式被保护，3'-OH 基团可以转变成氨基亚磷酸酯，例如 -
OP(OR')NR₂；其中 R 是异丙基 -CH(CH₃)₂；R' 是例如 -H (产生氨基亚磷酸酯二酯)，或 -CH₃，-CH₂CH₃，或 β-氰乙基 -CH₂CH₂CN (产生氨基亚磷酸
20 酯三酯)。单体的 3'-氨基亚磷酸酯基团与逐渐增长的寡核苷酸的 5'-OH 基团反应可产生亚磷酸酯连键 5'-OP(OR')O-3'。

在第三步中，未与单体偶联的寡核苷酸退出进一步合成以防止不完全聚合物的形成。通过与乙酸酐 (CH₃C(O)-O-C(O)CH₃) 反应，以乙酸酯 (-OC(O)CH₃) 的形式给保持原状的 5'-OH 基团加帽可以做到这一点。

25 在第四步中，新形成的亚磷酸酯基团 (即 5'-OP(OR')O-3') 被氧化成磷酸酯基团 (即 5'-OP(=O)(OR')O-3')；例如可通过与含水的碘和吡啶反应做到这一点。

由于在所述方法最后得到的寡核苷酸的 5' 端得到保护，并可用于第一步中，因此可重复进行四步法。当已得到合乎要求的全长寡核苷酸时，
30 可通过例如碱和热处理从固相支持物上将它裂解下来。此步骤也可用于将磷酸三酯 (即当 R' 不是 -H 时) 转变成磷酸二酯 (-OP(=O)₂O-)，并使核苷酸碱基的碱基 - 不稳定地被保护的氨基基团脱保护。

可以通过任何一般性技术，如 PCR 扩增或基因克隆来复制通过这些方法中的任一种而制备的多核苷酸，以提供较丰富的多核苷酸来源。

含有编码糖蛋白 B 之多核苷酸的克隆和表达载体

本发明中提供的克隆载体和表达载体含有编码疱疹病毒糖蛋白 B 的序列或其变异体或其片段。可根据标准技术构建适当的克隆载体，或者也可从本领域中适用的大量克隆载体中选择。尽管根据欲被使用的宿主细胞而选择的克隆载体会有所不同，但有用的克隆载体一般应具有自我复制的能力，对于特定的限制性核酸内切酶具有单个作用位点，并携有能用于选择被转染之克隆的标记物的基因。适当的例子包括质粒和细菌的病毒；如 pUC 18, mp 18, mp 19, pBR 322, pMB 9, ColE 1, pCR 1, RP 4, 噬菌体 DNA 和如 pSA 3 和 pAT 28 的穿梭载体。

表达载体通常为与适当的转录和翻译调控元件有效连接的编码多肽的可复制多核苷酸构建体。转录调控元件的例子是启动子，增强子，转录起始位点和转录终止位点。翻译调控元件的例子是核糖体结合位点，翻译起始位点和终止密码子。也可包括蛋白质加工元件：例如编码前导或信号肽以及将多肽跨膜转运或从细胞中分泌出来所需的蛋白酶裂解位点的区域。所用元件在用于表达的宿主细胞中可能是功能性的。调控元件可以得自与载体中所用相同的糖蛋白 B 基因，或者它们也可以是异源的（即衍生自其它基因和/或其它生物体）。

可以通过本领域中任何已知方法将多核苷酸插入宿主细胞。适当的宿主细胞包括如大肠杆菌，分枝杆菌的细菌细胞，其它原核微生物和真核细胞（包括真菌细胞，昆虫细胞，植物细胞和动物细胞）。通过直接摄入，胞吞作用，转染，*f*-交配或电穿孔插入外源多核苷酸可以转化细胞。随后，外源多核苷酸在细胞内以未-整合的载体，如质粒的形式被维持，或者也可以被整合到宿主细胞基因组中。

克隆载体可被用于得到它们所含的多核苷酸的复制拷贝，或者作为将多核苷酸贮存到仓库中以便将来回收的工具。表达载体和宿主细胞可被用于得到由它们所含的多核苷酸转录的多肽。它们也可以用于如药物筛选试验的试验中，其中具有能合成多肽的完整细胞才合乎需要。

γ 疱疹病毒糖蛋白 B 的合成的 Type1 寡核苷酸

本发明中包括的由疱疹病毒糖蛋白 B 序列设计的寡核苷酸可被用作鉴定相关序列的探针，或者在如 PCR 的扩增反应中用作引物。



在以下章节中描述了具有不同特性的不同寡核苷酸。被称作 Type 1 的寡核苷酸是从先前已知的 γ 疱疹病毒糖蛋白 B 多核苷酸序列中被设计出的。它们被设计成能与编码任何 γ 疱疹病毒糖蛋白 B 的多核苷酸杂交, 并可用于检测先前已知的 γ 疱疹病毒种。它们也可用于检测和鉴定 γ 疱疹病毒新种。被称作 Type 2 的寡核苷酸是从 RFHV 和 KSHV 糖蛋白 B 多核苷酸序列这两者中被设计出的, 它们被设计成能与 RFHV/KSHV 亚族(包括但不限于 RFHV 和 KSHV)的编码糖蛋白 B 的多核苷酸杂交。被称作 Type 3 的寡核苷酸是从 RFHV 或 KSHV 糖蛋白序列中被设计出的, 所述序列对于各个病毒而言是相对独特的, 这类寡核苷酸被设计成只能与 RFHV 或 KSHV 和密切相关的病毒株的编码糖蛋白 B 的多核苷酸特异性杂交。

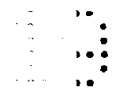
一些优选的 Type 1 寡核苷酸的例子列于表 4, 这些寡核苷酸对于广范围疱疹病毒的编码糖蛋白 B 的多核苷酸具有特异性。

表 4: 用于检测、扩增或鉴定疱疹病毒编码糖蛋白 B 的多核苷酸的 Type 1 寡核苷酸

靶: 疱疹病毒, 尤其是 γ 疱疹病毒的糖蛋白 B

名称	序列 (5'至 3')	长度	类型数目	方向	SEQ ID:
FRFDA	GCTG TTCAGATTTGACTTAGAYMANM CNTGYCC	33	256	有义	13
NIVPA	GTGTACAAGAAGAACATCGTGCCNTA YATNTTYAA	32	64	有义	14
NIVPASQ	GTGTACAAGAAGAACATCGTGCC	23	1		15
TVNCB	AACATGTCTACAATCTCACARTTNAC NGTNGT	32	128	反义	16
TVNCBSQ	AACATGTCTACAATCTCACA	20	1		17
FAYDA	AATAACCTCTTTACGGCCCAAATTCA RTWYGCNTAYGA	38	64	有义	18
IYGKA	CCAACGAGTGTGATGTCAGCCATTTA YGGNAARCCNGT	38	64	有义	19
IYGKASQ	CCAACGAGTGTGATGTCAGCC	21	1		20
CYSRA	TGCTACTCGCGACCTCTAGTCACCTT YAARTTYRTNAA	38	64	有义	21
CYSRASQ	TGCTACTCGCGACCTCTAGTCACC	24	1		22
NIDFB	ACCGGAGTACAGTTCCACTGTYTTRA ARTCDATRTT	36	48	反义	23
NIDFBSQ	TGTCACCTTGACATGAGGCCA	21	1		24
FREYA	TTTGACCTGGAGACTATGTTYMGNGA RTAYAA	32	64	有义	25
FREYB	GCTCTGGGTGTAGTAGTTRTAYTCYC TRAACAT	33	16	反义	26
NVFDB	TCTCGGAACATGCTCTCCAGRTCRAA MACRTT	32	32	反义	27
GGMA	ACCTTCATCAAAAATCCCTTNGGNGG NATGYT	32	128	有义	28
TVNCA	TGGACTTACAGGACTCGAACNACNGT NAAYTG	32	128	有义	29

表 4 所示方向是相对于多核苷酸的编码区。“有义”方向的寡聚体与编码链的反义链杂交并以编码序列的方向起动扩增。“反义”方向的



寡聚体与编码链杂交并以与编码序列相反的方向起扩扩增。

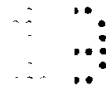
5 这些寡核苷酸已被设计成具有几个希望有的特性：1) 即使当来源材料中存在的靶 DNA 拷贝数很低时，对靶 DNA 仍有敏感性；2) 有足够的特异性以避免与不需要的序列，如具有有限的类似性的宿主序列杂交；3) 有足够的交叉反应性以使未知靶和用于设计它的序列之间的差异不会阻止寡核苷酸与靶形成稳定的双螺旋。

10 对于一些应用而言，特别有效的设计是在 3' 末端具有简并区段的寡核苷酸，所述寡核苷酸是从至少 2 个据信在某种程度上与多核苷酸靶有保守部分的已知多核苷酸的区域设计成的。多义残基的各种变换有助于确保寡核苷酸的至少一个可替换的形式能与靶杂交。在寡核苷酸的 5' 末端与简并区段邻接的是加强任何可能形成的双螺旋或允许在较高温度下进行杂交或扩增反应的共有区段。简并区段位于分子的 3' 末端以增加寡核苷酸和靶之间在聚合酶链反应过程中延伸开始的位点处紧密匹配的可能性。

15 根据下列密码子表示序列简并部分中的多义残基：

表 5: 多义位置的单字母密码子	
密码子	代表物
R	A 或 G (嘌呤)
Y	C 或 T (嘧啶)
W	A 或 T
S	C 或 G
M	A 或 C
K	G 或 T
B	C 或 G 或 T (非 A)
D	A 或 G 或 T (非 C)
H	A 或 C 或 T (非 G)
V	A 或 C 或 G (非 T)
N	A 或 C 或 G 或 T

20 表 4 所示的 Type 1 寡核苷酸通常用于与编码糖蛋白 B 的多核苷酸区段杂交。可以进行这种杂交以检测多核苷酸的存在，或者引发扩增反应以使多核苷酸可以被进一步鉴定。适当的靶包括编码广谱 γ 疱疹病毒



糖蛋白 B 区域的多核苷酸，所述病毒包括 RFHV/KSHV 亚族中的成员。适当的病毒是那些感染任何脊椎动物，包括人和非人灵长类动物的病毒，不论糖蛋白 B 或病毒是否已知或已经被描述。非限制性的例子包括编码表 1 所列任何 γ 疱疹病毒的糖蛋白 B 的多核苷酸。

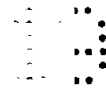
- 5 寡核苷酸可以特别地用于引发从寡核苷酸杂交的位点起向 3' 方向扩增靶多核苷酸区域的反应。FRFDA, HIVPA, TVNCB, FAYDA, IYGKA, CYSRA, NIDFB, FREYA, FREYB, NVFDB, GGMA, 和 TVNCA 是在邻近简并区段处具有共有区段的寡核苷酸，可以用于此目的。

10 图 2 显示了 RFHV/KSHV 亚族的整个糖蛋白 B 多核苷酸序列上前述寡核苷酸引物有望杂交的位置。此图没有按比例绘制，但准确地显示出了沿着糖蛋白 B 编码链的 5' - 3' 方向上的推测的杂交位点的顺序。此图还显示出了通过在扩增反应中使用多套引物而产生的扩增产物的大致长度。所示长度包括每个末端的引物结合位点，和它们之间包含的多核苷酸。

- 15 用作表 4 寡核苷酸的靶的优选 DNA 来源是任何生物样品（包括固形组织和组织培养物），特别是已知或怀疑携带疱疹病毒的脊椎动物来源的样品。通过本领域中已知的任何方法，包括用有机溶剂提取或高盐浓度下沉淀可从样品来源中提取 DNA。

20 优选的扩增方法是聚合酶链反应：一般可参见美国专利号 4,683,195 (Mullis) 和 4,683,202 (Mullis 等人)；对于病毒多核苷酸的应用参见美国专利 5,176,995 (Sninsky 等人)。通过将欲被扩增的靶多核苷酸与能同靶杂交的短寡核苷酸合并，并用作聚合反应的引物即可进行扩增反应。还需加入底物单核苷酸和热稳定性的依赖于 DNA 的糖蛋白 B，如 Taq。用于扩增反应的条件一般为本领域中已知的条件，使用已知病毒源，如 RFHV, KSHV, hEBV 或 HSV1 凭经验可使条件最优化。条件也可以改变，例如可改变扩增循环，尤其是杂交时相的时间和温度；改变寡核苷酸引物的重量摩尔成分；改变缓冲液浓度；改变扩增循环的数目。25 细微调节扩增的条件对于本领域技术人员而言是常规技术。

30 在一个方法中，扩增中使用了本发明的单个引物，可选择使用第二个引物，如随机引物以起动相反方向上的第一引物下游的复制。在优选的方法中，相同反应中至少使用了两个本发明的引物以起动相反方向上的复制。至少两个特异性引物的使用增强了扩增反应的特异性，并限定



了样品之间比较片段的大小。例如，使用引物 NIVPA 和 TVNCB 可以进行扩增。更优选以嵌套方式使用几套引物以增强扩增。通过使用产生中间体产物的引物进行首次扩增可实现嵌套，所述引物含有另一个引物的一个或多个内部结合位点。接着使用与上述套中的一个引物联合的新引物，或两个新引物进行第二次扩增，因此第二次扩增的产物是第一次产物的亚片断。如有必要，可使用另一个引物进行另一轮扩增。

因此，使用三个或更多个寡核苷酸引物的任何组合可进行嵌套扩增反应，所述引物组合中含有至少一个有义方向上的引物和一个反义方向上的引物。优选选定的引物能使中间体扩增产物不超过约 2000 个碱基对；更优选它们不超过约 1500 个碱基对；甚至更优选它们不超过约 750 个碱基对。优选最靠内的引物提供的最终扩增产物不超过约 1200 个碱基对；更优选它们不超过约 750 个碱基对；甚至更优选它们不超过约 500 个碱基对。因此，优选的组合是选自 FAYDA, IYGKA, CYSRA, NIDFB, NVFDB, 和 FREYB 的至少三个引物。另一个优选的组合是选自 FRFDA, NIVPA, TVNCA, NIDFB, NVFDB 和 FREYB 的至少三个引物。

特别优选使用引物 FRFDA 和 TVNCB 第一次扩增，再使用引物 NIVPA 和 TVNCB 第二次扩增。当对 KSHV 的糖蛋白 B 基因的多核苷酸进行扩增时，包括引物结合区域的最终片断的大小约为 386 个碱基。

在扩增反应的任何阶段，通过例如大小测定即可鉴定出经扩增的多核苷酸。优选通过在约 1-2% 的琼脂糖凝胶上电泳多核苷酸来进行鉴定。如果凝胶中的多核苷酸以充足的量存在，即可用溴化乙锭染色并在紫外光下检测。或者也可以在上样到约 6% 的聚丙烯酰胺凝胶之前，用如 ^{32}P 或 ^{35}S 的放射性同位素标记多核苷酸，随后可利用凝胶产生放射自显影图。优选的标记经扩增之多核苷酸的方法是使用多核苷酸激酶和 γ - $[\text{}^{32}\text{P}]\text{-ATP}$ ，连续扩增约 5-15 个循环以用 ^{32}P 末端标记如 NIVPA 的寡核苷酸引物。

如果需要，在制备经扩增的多核苷酸时也可将大小分离用作一个步骤。当发现扩增混合物含有不同大小的假象多核苷酸，如可能是由于与不需要的靶的交叉反应性而引起这一现象时，所述方法特别有用。如上一段落中描述的分离的凝胶被放在其背后的纸上晾干并用于产生放射自显影图。切下凝胶中对应于放射自显影图上所需带的位置，并通过标准技术提取。提取出的多核苷酸能被直接鉴定，克隆或用于下一个扩增循

环。

寡核苷酸 NIVPASQ, TVNCBSQ, IYGKASQ, CYSRASQ, 和 NIDFBSQ 中的每一种都得自共有 - 简并的 Type1 寡核苷酸, 它们保留了共有区段, 但缺乏简并区段。它们对于测定使用共有 - 简并寡核苷酸成功扩增的糖蛋白 B 多核苷酸片断的序列特别有用。

通过采用这种类型的引物也可以降低扩增反应混合物中不需要的多核苷酸的比例。例如, 使用 1/5 至 1/25 正常量的引物 NIVPA 和 TVNCB 进行最初的 3 - 5 个循环扩增, 然后加入摩尔过量 (例如 50pmol) 的 NIVPASQ 和/或 TVNCBSQ, 再继续扩增 30 - 35 个循环。这可以降低扩增混合物中存在的寡核苷酸的复杂度, 并允许反应温度增高以减少不需要的多核苷酸的扩增。

RFHV/KSHV 亚族糖蛋白 B 的 Type2 型寡核苷酸引物

Type2 寡核苷酸用于 RFHV/KSHV 亚族任何病毒的糖蛋白 B 的检测或扩增反应。它们是由 RFHV 和 KSHV, 而不是其它先前已测序的 γ 疱疹病毒之间相对很保守的糖蛋白 B 编码区的区段设计成的。优选的例子示于表 6:

表 6: 用于检测, 扩增, 或鉴定编码疱疹病毒糖蛋白 B 的多核苷酸的 Type 2 型寡核苷酸					
靶: 疱疹病毒 RFHV/KSHV 亚族的糖蛋白 B					
名称	序列 (5'至 3')	长度	类型数目	方向	SEQ ID:
SHMDA	AGACCCGTGCCACTCTATGARATHA GYCAYATGGA	35	24	有义	41
SHMDASQ	AGACCCGTGCCACTCTATGA	20	1		42
CFSSB	GTTCACAACAATCTTCATNGARCTR AARCA	30	32	反义	43
CFSSBSQ	GTTCACAACAATCTTCAT	18	1		44
ENTFA	GTCAACGGAGTAGARAAYACNTTYA CNGA	29	128	有义	45
DNIQB	ACTGGCTGGCTAAAGTACCTTTGAA TRTTRTCNGT	35	16	反义	46
DNIQBSQ	ACTGGCTGGCTAAAGTACCTTTG	23	1		47



Type2寡核苷酸可用于需要RFHV/KSHV亚族特异性的多种目的，包括检测或扩增得自己知的RFHV/KSHV亚族病毒的糖蛋白B，或鉴定此亚族新成员的糖蛋白B。

5 SHMDA, CFSSB, ENTFA 和 DNIQB 是具有简并的 3'末端的共有-简并寡核苷酸，可用作 PCR 扩增的最初的引物，它们包括与 RFHV 或 KSHV 中的任一个都不相同的 RFHV/KSHV 亚族的多核苷酸。SHMDASQ, CFSSBSQ 和 DNIQBSQ 仅含有共有区段，例如可用于标记使用共有-简并寡核苷酸已经扩增的多核苷酸，或测定所述多核苷酸的序列。

10 在一个应用中，可单个使用或者组合使用这些 Type 2 寡核苷酸作为扩增引物。在此应用的一个实施例中，对得自组织样品中的 DNA 直接使用寡核苷酸以得到衍生自 RFHV/KSHV 亚族，而不是存在于相同组织中的如 hEBV, hCMV 或 HSV1 的更不相关之病毒的糖蛋白 B。因此，SHMDA 和 DNIQB 可用作 PCR 中的引物，任选使用如 NIVPA 和 TVNCB 的 Type1 寡核苷酸预扩增所述引物。其它组合也适用。在另一个实施例中，将表 6
15 的 Type2 寡核苷酸中的一个与前文所列的适当的 Type1 寡核苷酸联合使用。因此，NIVPA 可以与 DNIQB 联合，或 SHMDA 可以与 TVNCB 联合以用作 PCR 引物。可任选使用 NIVPA 和 TVNCB 预扩增 DNA 源。其它组合也适用。

20 在另一个应用中，Type 2 寡核苷酸，或含有这些序列的寡核苷酸或其片段被用作检测试验中的探针。例如，它们可被提供成具有适当的标记物，如 ^{32}P ，然后用于具有适当靶，如使用 FRFDA 和/或 NIVPA, 和 TVNCB 一起扩增的 DNA 的杂交试验中。

RFHV 或 KSHV 糖蛋白 B 特异性的 Type3 寡核苷酸引物

25 Type3 寡核苷酸用于特殊病毒特异性的检测或扩增反应，它们是 RFHV 或 KSHV 糖蛋白 B 编码区的非简并区段，相对于此区域的其它区段而言，此区段在这两个病毒之间比其它疱疹病毒相对更易变。优选的例子示于表 7 和实施例部分：



表 7: 用于检测, 扩增, 或鉴定编码疱疹病毒糖蛋白 B 的多核苷酸的 Type3 型寡核苷酸

靶: RFHV 的糖蛋白 B					
名称	序列 (5'至 3')	长度	类型数目	方向	SEQ ID:
GMTEB	TGCTGCTTCTGTCATACCGCG	21	1	反义	48
AAITB	TATTTGTTTGTGATTGCTGCT	21	1	反义	49
GMTEA	GCGGTATGACAGAAGCAGCAA	21	1	有义	50
KYEIA	AACAAATATGAGATCCCCAGG	21	1	有义	51
TDRDB	TCATCCCGATCGGTGAACGTA	21	1	反义	52
VEGLB	TTGTCAGTTAGACCTTCGACG	21	1	反义	53
VEGLA	CCCGTCGAAGGTCTAACTGAC	21	1	有义	54
PVLYA	AGCCAACCAGTACTGTACTCT	21	1	有义	55
靶: KSHV 的糖蛋白 B					
名称	序列 (5'至 3')	长度	类型数目	方向	SEQ ID:
GLTEB	TGATGGCGGACTCTGTCAAGC	21	1	反义	56
TNKYB	GTTTCATACTTGTGGTGATGG	21	1	反义	57
GLTEA	GGGCTTGACAGAGTCCGCCAT	21	1	有义	58
YELPA	ACAAGTATGAACTCCCGAGAC	21	1	有义	59
VNVNB	ACCCCGTTGACATTTACCTTC	21	1	反义	60
TFTDV	TCGTCTCTGTCAGTAAATGTG	21	1	反义	61
TVFLA	CCACAGTATTCCTCCAACCAG	21	1	有义	62
SQPVA	GGTACTTTAGCCAGCCGGTCA	21	1	有义	63

GMTEB, AAITB, GMTEA, KYEIA, TDRDB, VEGLB, VEGLA 和 PVLYA 是 RFHV 糖蛋白 B 的特异性非简并寡核苷酸, 可用于扩增或检测 RFHV 来源的编码糖蛋白 B 的多核苷酸。优选使用嵌套形式的寡核苷酸进行扩增: 例如使用 GMTEA 和 VEGLB 作为引物进行首次扩增; 然后使用 KYEIA 和 TDRDB 作为引物进行第二次扩增。这可以提供特异于 RFHV 糖蛋白 B 的极度敏感的扩增试验。GMTEB 和 AAITB 在片断的 5' 末端附近杂交, 并可与在上游杂交的 Type1 寡核苷酸联合使用以扩增或检测 5' 方向上的序列。VEGLA 和 PVLYA 在片断的 3' 末端附近杂交, 可与在下游杂交的 Type1 寡



核苷酸联合使用以扩增或检测 3' 方向上的序列。

类似地，GLTEB，TNKYB，GLTEA，YELPA，VNVNB，ENTFB，SQPVA 和 TVFLA 是 KSHV 糖蛋白 B 的特异性非简并寡核苷酸，可以类似的方式被使用，包括作为扩增反应的引物。优选使用嵌套形式的寡核苷酸进行扩增：
5 例如使用 GLTEA 和 ENTFB 作为引物进行首次扩增；然后使用 YELPA 和 VNVNB 作为引物进行第二次扩增。这可以提供特异于 KSHV 糖蛋白 B 的极度敏感的扩增试验。GLTEB 和 TNKYB 在片断的 5' 末端附近杂交，可与在上游杂交的 Type1 寡核苷酸联合使用以扩增或检测 5' 方向上的序列。

10 SQPVA 和 TVFLA 在片断的 3' 末端附近杂交，可与在下游杂交的 Type1 寡核苷酸联合使用以扩增或检测 3' 方向上的序列。

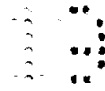
本领域技术人员会很快认识到 Type 1, 2 和 3 寡核苷酸（尤其是那些表 4, 6 和 7 中所示的）可以互相联合用于 PCR 以扩增编码糖蛋白 B 之多核苷酸的不同部分。扩增反应的特异性一般由具有最少量交叉反应性的引物决定。扩增片段的大小和位置由扩增的最后一次循环所用的引物决定。例如，NIVPA 与 SQPVB 联合使用会扩增约 310 个碱基的编码与
15 KSHV 密切相关之病毒的糖蛋白 B 的多核苷酸。寡核苷酸的适当联合可被用作嵌套形式的扩增引物。

使用合成的寡核苷酸鉴定多核苷酸靶

如前面章节所述，本发明中包含的寡核苷酸可被用作引物以扩增编
20 码疱疹病毒糖蛋白 B 的多核苷酸，尤其在聚合酶链反应中更是如此。

进行 PCR 的条件取决于所使用寡核苷酸的特性。具体地说，当所用寡核苷酸含有的简并区段或共有区段与靶的相应区段只部分相同时，和当靶多核苷酸含有未知序列时，条件的选择对于扩增的成功会很重要。对于新引物或新多核苷酸靶的条件的最优化是本领域技术人员的常规技
25 术。下文是有助于此目的的指导性方案。

首先，使 PCR 退火步骤的温度最优化以增加被扩增的靶多核苷酸的量，使之大于被扩增的不相关多核苷酸的量。理想状态下，所述温度允许引物与靶序列杂交，而不与其它序列杂交。对于含有共有区段的引物（Type1 型）而言，PCR 重复循环中退火步骤的温度一般至少约为 45℃；
30 优选至少约 50℃。优选首先在如 55 或甚至 60℃ 的较高温度下进行几轮 PCR 循环。较高的温度可以在循环过程中强使退火更具有序列特异性，并可降低不相关序列的背景扩增。可以在较不严紧的条件下进行随后循



环的退火步骤以提高扩增率。在特别优选的方法中，最初的 PCR 扩增循环含有在 60℃ 下进行约 1 分钟的退火步骤。在每个循环的温度递减 2℃ 的条件下进行随后的循环中的退火步骤，直至温度达到 50℃。然后进行退火温度为 50℃ 的后续循环，直至达到所需程度的扩增。

5 病毒特异性并不含共有区段的引物更具选择性，可在较宽的温度范围内有效。PCR 扩增循环中退火步骤的优选温度介于 50℃ 到 65℃ 之间。

第二，使缓冲液条件最优化。我们已发现由 Taq 聚合酶的商购制品提供的缓冲液有时难以使用，这部分是由于对镁离子浓度的严格的依赖性。使用如 M. Wigler (Lisitsyn 等) 推荐的缓冲液一般更容易进行使用本
10 发明的寡核苷酸进行的 PCR。优选最终的 PCR 反应混合物含有 $(\text{NH}_4)_2\text{SO}_4$ 代替 KCl 作为主要的离子源。优选最终反应混合物中 $(\text{NH}_4)_2\text{SO}_4$ 的浓度约为 5 - 50mM，更优选约 10 - 30mM，甚至更优选 16mM。缓冲成分优选为 Tris，优选其终浓度约 67mM，pH 约 8.8。在这些条件下， MgCl_2 浓度不必严格限制。优选其终浓度约 1 - 10mM，更优选约 3 - 6mM，最佳约为 4mM。反
15 应混合物也含有约 10mM β - 巯基乙醇和 0.05 - 1mg/ml 牛血清白蛋白。特别优选的缓冲液是 WB4 缓冲液 (67mM Tris 缓冲液，pH 8.8，4mM MgCl_2 ，16mM $(\text{NH}_4)_2\text{SO}_4$ ，10mM β - 巯基乙醇和 0.1 mg/ml 白蛋白)。下文实施例 3 中提供了进行此反应的优选条件。

为了进行 PCR 反应，需制备含有寡核苷酸引物，四种脱氧核苷酸，
20 适当缓冲液，欲扩增的 DNA，和热稳定性的依赖于 DNA 的 DNA 聚合酶的混合物。然后通过退火，延伸，和解链步骤的温度循环加工混合物，直至达到所需程度的扩增。例如，通过溴化乙锭染色可以测定产生的 DNA 的量，染色可在琼脂糖凝胶上分离扩增片段后进行。

扩增反应可能的复杂情况是寡核苷酸引物自身的二聚体化和扩增，
25 在琼脂糖凝胶上可以轻易检测出低分子量 (<100 个碱基对) 的此片断。通过琼脂糖或聚丙烯酰胺凝胶分离可除去扩增引物，通过稍微调节一下反应条件，尤其是退火步骤的温度可以降低扩增的二聚体的量。也优选预混合引物，脱氧核苷酸和缓冲液，并在加入欲扩增的 DNA 前，将混合物加热到 80℃。

30 使用本发明的任何寡核苷酸作为引物的扩增反应产生了编码糖蛋白 B 部分的多核苷酸片段。通过本领域技术人员已知的大量技术可以鉴定这些片段。一些鉴定片段的非限制性方法如下所述：

在一个方法中，可根据本领域已知的任何序列测定法测定片段的序列，所述方法包括 Maxam & Gilbert 法，或者 Sanger & Nicholson 法。或者也可以将片段提交给任何提供多核苷酸测序服务的商业机构。在测序前可任选将片段克隆和/或扩增。核苷酸序列可被用于预测由此片段编码的氨基酸序列。可使用序列资料在多核苷酸水平上或者在氨基酸水平上与其它已测序的糖蛋白 B 比较，以鉴定原始来源材料中存在的疱疹病毒种类。在制作模型的算法规则中也可使用序列资料预测抗原区或三维结构。

在第二个鉴定方法中，可通过任何适当的方法测定片段的大小，所述方法如在聚丙烯酰胺或琼脂糖凝胶上电泳，或通过适当的密度梯度离心。例如，对于 RFHV 和 KSHV 而言，NIVPA 和 TVNCB 之间的片段约为 319 个碱基。因此包括引物结合区域的整个扩增片段的长度约为 386 个碱基。相应的 sHV1 片段含有额外的 6 个碱基对。因此，例如通过将每个片段扩增的多核苷酸片段置于分离凝胶的相邻泳道上电泳，或者通过将 sHV1 片段置于适当的分子量标准旁边电泳即可将 sHV1 片段与 RFHV 或 KSHV 的片段区分开来。在大小上与 RFHV 和 KSHV 相同的多核苷酸片段可衍生自相同或相关的病毒种类。大小基本上不同的片段更可能衍生自不同的疱疹病毒。

在第三个鉴定方法中，通过尝试将片段与寡核苷酸探针杂交来检测片段。在优选的实施例中，检测了片段与 RFHV 或 KSHV 的糖蛋白 B 编码区的相关性。可使用含有糖蛋白 B 编码区序列或其遗传互补物的探针进行试验。适当的探针是含有 RFHV 或 KSHV 的序列的多核苷酸，如表 7 所列的 Type3 寡核苷酸。

探针的长度和特性以及杂交条件的选择取决于试验的目的。如果目的是仅仅检测得自 RFHV 或 KSHV 的多核苷酸，包括次要的株变体，那么在高严谨性的条件下进行杂交。使用了来自各个 RFHV 或 KSHV 糖蛋白 B 的序列。较长的序列改善了试验的特异性，可以在较高严谨性的条件下使用。优选探针含有至少约 30 个核苷酸的糖蛋白 B 序列；更优选此序列至少约 50 个核苷酸；甚至更优选此序列至少约 75 个核苷酸长。

如果目的是检测与 RFHV 或 KSHV 密切相关但不相同的多核苷酸，如在筛选试验或征集先前未曾描述过的 RFHV/KSHV 亚族病毒的试验中，可选择不同的条件。得自 RFHV 或 KSHV 的序列可以被使用，但一般优选两

个的混合物或简并探针。选择序列的长度和杂交反应的条件以提供足够的特异性，从而排除不需要的序列，但要么提供与潜在靶的最大程度的交叉反应性。使用上文给出的公式，通过计算 T_m ，再计算能忍受的最大程度错配的相应温度即可预测适当的条件。通过检测探针与含有编码疱疹病毒糖蛋白 B 的已知靶多核苷酸样品的交叉反应性，即可凭经验检测出条件的适当性。

在此试验条件下形成稳定双螺旋所需的最小程度的互补性将决定何种糖蛋白 B 序列会与探针杂交。可以考虑，例如扩增得自人或非人灵长类动物的靶产生对应于 SEQ. ID NO: 3 的碱基 36 - 354 的片段，再用 KSHV 多核苷酸的相应片段来作探针。根据表 2 中的资料，如果杂交反应在仅需要约 50% 相同即可形成稳定双螺旋的条件下进行，探针会与得自任何已被测序的 γ 疱疹病毒（包括 hEBV 和 sHV1）糖蛋白 B 基因的靶杂交。如果反应在需要探针和靶之间至少约有 65% 相同，优选至少约 67% 相同，更优选至少约 70% 相同，甚至更优选至少约 75% 相同才能形成稳定双螺旋的条件下进行，此试验将检测到来自 RFHV/KSHV 亚族即或者是 RFHV，KSHV，或者是具有仍未被测序的糖蛋白 B 多核苷酸的密切相关疱疹病毒的靶多核苷酸。甚至在仅需要约 50 - 55% 的同一性以形成稳定的双螺旋的杂交条件下，由于据信 bHV4，eHV2 或 mHV68 不能感染灵长类动物，因此阳性反应不能显示这些病毒的存在。

可以将通过大小的鉴定与通过杂交的鉴定联合。例如，可在丙烯酰胺或琼脂糖凝胶上分离被扩增的多核苷酸，将所述多核苷酸印迹到适当材料的膜上，如硝酸纤维素膜上，然后与具有适当标记物，如 ^{32}P 的探针杂交。洗涤后标记物的存在反映了样品中可杂交材料的存在，而与适当分子量标准比较的迁移距离反映了材料的大小。在高严紧度条件下与前述探针之一杂交但具有非预期大小的片段序列表示与探针有高水平的同一性，但与 RFHV 或 KSHV 有所不同的糖蛋白 B 序列。

使用多核苷酸和寡核苷酸检测疱疹病毒感染

本发明中包含的编码疱疹病毒糖蛋白 B 的多核苷酸，和根据此多核苷酸的合成寡核苷酸可用于诊断与疱疹病毒感染相关的临床疾病。例如，临床样品中可检测到疱疹病毒糖蛋白 B 的存在可暗示各个疱疹病毒作为病原因子参与了疾病的发展。特殊组织而不是周围组织中糖蛋白 B 的存在可用于定位感染的损害。临床样品中 γ 疱疹病毒和其它疱疹病毒

之间的鉴别可用于预测感染的临床进程或选择适于治疗的药物。由于糖蛋白 B 是由复制病毒，L-颗粒，和感染细胞表达的，因此我们预测它可用作与任何这些形式的蛋白质表达有关之疾病的活动期和潜伏期的有用的标记物。

- 5 进行诊断试验的方法是本领域中广为人知的，对于本领域普通技术人员而言是常规技术。通常为了进行本发明的诊断方法，本发明的一个组合物被提供成试剂以检测临床样品中与之反应的靶。例如，本发明的多核苷酸可被用作试剂以检测如可能存在于被疱疹病毒感染的细胞中的 DNA 或 RNA 靶。本发明的多肽可被用作试剂以检测能与之形成特异性复
- 10 合物的靶。如抗体分子或（如果多肽是受体）相应配体。本发明的抗体可被用作试剂以检测它特异性识别的靶，如由病毒感染细胞表达的多肽。

通过从被测量诊断参数的个体处得到适当的组织样品即可提供靶。相关的试验样品是那些得自怀疑携带疱疹病毒的个体的样品。多种类型的

15 样品适于这一目的，包括那些通过活组织检查或外科解剖得自感染或病理学可疑部位附近的样品，衍生自它们的体外细胞培养物，溶解的提取物，血液和血液成分。如果需要，在进行检测前可从样品中将靶部分纯化或扩增。通过在允许试剂与靶之间形成复合物的条件下将试剂与样品接触来进行反应。可在溶液中进行反应，或者例如使用组织学切片在

20 固体组织样品上进行反应。通过各种本领域中已知的技术可检测复合物的形成。例如可提供具有标记物的试剂，从复合物中除去未反应的试剂；从而剩下的标记物的量可表示所形成复合物的量。在说明书下文中将提供复合物检测的进一步的细节和替换物。

为了测定所形成的复合物的量是否代表了感染或未感染疱疹病毒的

25 细胞，优选将试验结果与对对照样品进行的类似试验结果相比较。一般优选使用得自未感染来源的对照样品，要么对照样品在组成上类似于被检测的临床样品。然而，如果已知对照样品中靶的相对量，或者此相对量能用于比较的目的，则任何对照样品都适合。通常优选同时对试验样品和对照样品进行检测。然而，如果所形成复合物的量可以定量并且有

30 足够的一致性，则在不同日期或不同实验室检测试验样品和对照样品是可以接受的。

因此，本发明所包含的编码 RFHV/KSHV 亚族糖蛋白 B 的多核苷酸和

合成的寡核苷酸可被用于检测生物样品中可能存在的 γ 疱疹病毒的多核苷酸。在具体的诊断试验中使用多核苷酸的一般方法是本领域技术人员所熟知的，例见日本专利申请 5309000 (latron)。

例如：1) 通过与特异性探针进行杂交反应；2) 通过与特异性引物进行扩增，或 3) 通过上述两种方法的联合可使使用多核苷酸试剂的检测变成特异性的。

为了进行因与特异性探针杂交而被赋予特异性的试验，选定的多核苷酸应具有所需程度的与想要的靶的互补性。优选的探针包括长度至少约为 16 个核苷酸的编码 RFHV, KSHV 或 RFHV/KSHV 亚族其它成员糖蛋白 B 的一部分的多核苷酸。更加优选的探针含有至少约 18, 21, 25, 30, 50 或 100 个核苷酸长的糖蛋白 B 编码区。还优选在所用条件下能与 RFHV/KSHV 亚族的多核苷酸而不能与其它疱疹病毒的多核苷酸形成稳定双螺旋的简并探针。

一般情况下提供的探针具有标记物。经常用于这种类型检测的一些标记物包括如 ^{32}P 和 ^{33}P 的放射性同位素，化学发光剂或如荧光素的荧光剂，能产生有色溶质或沉淀物的酶，如碱性磷酸酶。标记物对于试剂而言可以是内源性的，它可通过直接的化学键被连接，或者通过一系列中间反应分子，如生物素-亲和素复合物或一系列相互反应的多核苷酸被连接。可在与靶多核苷酸杂交之前或之后在试剂中加入标记物。为了改善试验的敏感性，经常需要增加因杂交引起的信号，通过以复合标记物成分掺入每个复合物的方式使用系列杂交的多核苷酸或分支的多核苷酸的联合可以实现这一目的，参见美国专利号 5,124,246 (Urdea 等人)。

如果需要，可从样品中提取靶多核苷酸，也可以将靶多核苷酸部分纯化。为了测量病毒颗粒，优选富集 DNA 的制备；为了测量糖蛋白 B 的活性转录，优选富集 RNA 的制备。一般预期临床样品中疱疹病毒的多核苷酸水平较低，潜伏有病毒的每个细胞中仅有几拷贝编码糖蛋白 B 的 DNA，或正在复制病毒的每个细胞中有多至几百拷贝的 DNA。与那些基因在其中无活性的细胞相比，活跃表达蛋白质的细胞中 mRNA 的水平较高，因此需要通过扩增 DNA 或 RNA 来增强样品中靶的水平。适当的扩增方法是 PCR，优选使用本发明包含的一个或多个寡核苷酸引物进行 PCR。通过使用逆转录酶制备 cDNA 拷贝，然后使用上述引物进行 PCR 可以扩增 RNA。

靶多核苷酸可任选经受任何额外处理的联合，包括用限制性核酸内

切酶消化，例如通过在琼脂糖或聚丙烯酰胺中电泳以进行大小分离，和贴在反应基质，如印迹材料上。

5 在适当的反应条件下，通过将试剂多核苷酸与怀疑含有靶多核苷酸的样品混合，可允许杂交发生。接着通过洗涤或分离以除去未反应的试剂。一般情况下，靶多核苷酸和试剂必须都至少部分地被平衡成单链形式以便互补序列能有效杂交。因此，通过本领域技术人员已知的标准变性技术制备样品是有用的（特别是在 DNA 检测中）。

10 选定的杂交条件的严紧性水平取决于检测的目的。如果需要此检测对于 RFHV 或 KSHV 是特异性的，则可使用含有各个糖蛋白 B 一个区段的探针，在高严紧性的条件下进行反应。例如对于优选的 50 个或更多个核苷酸的探针所用的一套优选条件是 37℃，50% 甲酰胺，6×SSC，然后在低离子强度下洗涤。一般需要靶与多核苷酸探针至少约 90% 相同才形成稳定的双螺旋。通过增加所用探针的长度也可以增加对特定可疑病毒的反应的特异性。因此对于本发明的这一应用特别优选较长的探针。或者，如果需要试验能检测与 KSHV 相关的其它疱疹病毒，则可使用较低的严紧性。适当的探针包括得自 KSHV 糖蛋白 B 多核苷酸的片段，其混合物，或如表 7 所列的那些寡核苷酸。

20 适当的杂交条件被确定成仅允许探针与糖蛋白 B 序列杂交，所述序列具有所需程度的与探针的同一性。所需的严紧性取决于多核苷酸探针的长度和探针与所需靶序列之间同一性的程度。可以考虑例如探针由 NIVPA 和 TVNCB 杂交位点之间的 KSHV 多核苷酸片段组成，需要最少 60% 同一性的条件会导致与 KSHV 和如 sHV1 的其它 γ 疱疹病毒的相应多核苷酸形成稳定的双螺旋；需要最少 90% 同一性的条件会导致仅与 KSHV 和密切相关的变体的多核苷酸形成稳定的双螺旋；需要最少 65-70% 同一性的中间严紧度的条件会允许与 KSHV，和 RFHV/KSHV 亚族的一些其它成员的糖蛋白 B 多核苷酸形成双螺旋，而不与其它已知的疱疹病毒，包括 γ 疱疹病毒 eHV2，sHV1，mHV68，bHV4，EBV，和如 hCMV，hHV6，hVZV 和 HSV1 的其它人病原体的相应多核苷酸形成双螺旋。

30 使用上文给出的公式可预先估计条件。优选通过用已知含有多核苷酸的分离样品检测探针以进一步确证精确的条件，所述样品包括那些在试验中需被检测的和那些不需被检测的多核苷酸。或者通过由公开序列合成多核苷酸，或者通过从据信被各个疱疹病毒感染的组织中提取和扩

增 DNA 可以提供这种样品。确定杂交条件对于本领域技术人员而言是一种常规的调整，不需要太多的实验。由于 eHV2, sHV1, mHV68, bHV4 和 EBV 相对于 α 和 β 亚族的成员而言，与 RFHV/KSHV 亚族更加密切相同，排除了 RFHV/KSHV 亚族以外的 γ 疱疹病毒的条件一般也排除了表 1 所列的其它疱疹病毒。另外，如果据信在最终测定中，某些病毒不会存在于受试样品中（如 eHV2, mHV68 或 bHV4 在人组织样品中），那么当确定了试验条件后即可任选忽略相应的靶序列。因此条件可被确定为允许如表 6 所列的那些 Type2 寡核苷酸探针能与含有 SEQ. ID NO: 1 或 SEQ. ID NO: 3 的多核苷酸序列，而不与选自 SEQ. ID NO: 5-13 的序列形成稳定的双螺旋。条件也可被确定为允许含有 SEQ. ID NO: 1 或 SEQ. ID NO: 3 的至少 21 个或更多个连串碱基的适当片段能与含有 SEQ. ID NO: 1 和 SEQ. ID NO: 3 的多核苷酸，而不与含有 SEQ. ID NO: 5-13 中任何一个的多核苷酸形成稳定的双螺旋。

或者为了进行因用特异性引物扩增而被赋予特异性的试验，可如前所述从生物样品中制备 DNA 或 RNA。任选在 PCR 中使用非种特异性的引物，如表 4 或 6 中所列的那些，预扩增靶多核苷酸。然后使用特异性引物，如表 7 中所列的那些，或得自表 4, 6 和 7 的引物的联合扩增靶。在优选的实施方案中，以嵌套形式使用寡核苷酸引物进行了两轮扩增：第一轮中为病毒特异性或非特异性；第二轮中为病毒特异性。这提供了既敏感又特异的试验。

扩增过程中特异性 Type3 引物的使用足以提供所需的特异性。扩增系列的最后充足的反应产物的存在表示阳性试验。通过例如将反应混合物印迹到如硝酸纤维素的介质上并用溴化乙锭染色可以检测被扩增的多核苷酸。或者可以在最终的扩增循环期间在混合物中加入经放射性标记的底物；将掺入的标记物与未掺入的标记物分开（例如通过印迹或通过大小分离），检测标记物（例如通过计数或放射性自显影）。如果在琼脂糖或聚丙烯酰胺凝胶上电泳，产物的大小将有助于进一步证实被扩增片段的本质。通过使用得自前述方法的扩增混合物作为前文概述的杂交反应的靶源，在特异性扩增之后可进行特异性杂交。

多核苷酸用于基因治疗

本发明中包含的药物含有病毒特异性多核苷酸，多肽或抗体作为活性组分。这种组合物会降低病毒或感染细胞本身的病理学，或者使得病

毒或被感染的细胞对非特异性药物化合物的治疗更加敏感。

可以使用本发明编码疱疹病毒糖蛋白 B 部分的多核苷酸，例如给被感染的个体用药以进行基因治疗（一般参见美国专利号 5, 399, 346; Anderson 等人）。一般的原理是以或者促进或者减弱其编码的多肽表达的方式施用多核苷酸。

5 优选的基因治疗模式是以能在细胞内部复制，增强和延长效果的方式提供多核苷酸。因此，可使多核苷酸与适当的启动子有效连接，所述启动子如相应基因的天然启动子，在靶组织类型的细胞中有内在活性的异源启动子，或能够通过适当试剂被诱导的异源启动子。通过本领域中已知的适当技术，如以适当的载体，如病毒表达载体的形式提供多核苷酸，或将多核苷酸包裹在脂质体内，便于多核苷酸进入细胞。多核苷酸可以全身性注射，或者通过抗原特异性寻靶机制，或通过直接注射将多核苷酸提供给感染部位。

15 在一个有所变化的方案中，多核苷酸含有的启动子与具有相同方向的多核苷酸链连接，所述链在疱疹病毒感染过程中被转录。优选被编码的糖蛋白 B 包括外部成分，跨膜成分和转运至表面的信号序列。被这种多核苷酸转染的感染了病毒的细胞有望在表面表达增强水平的糖蛋白 B。这种形式的增强的糖蛋白 B 表达可增强通过免疫系统因子对这些细胞的识别，所述因子包括抗体（和依赖于抗体的效应物，如 ADCC）和病毒特异性的细胞毒 T 细胞。

20 在另一个有所变化的方案中，多核苷酸含有的启动子与具有相反方向的多核苷酸链连接，所述链在疱疹病毒感染过程中被转录。被这种多核苷酸转染的感染了病毒的细胞有望表达降低水平的糖蛋白 B。转录本有望与由病毒基因转录的互补链杂交，并防止它被翻译。已知这种方法为反义疗法。

RFHV/KSHV 亚族具有糖蛋白 B 活性的多肽及其片段

30 图 1 所示的 RFHV 和 KSHV 多核苷酸都具有开放阅读框。被它编码的多肽分别示于 SEQ. ID NO: 2 和 SEQ. ID NO: 4。用于得到多核苷酸序列的引物 NIVPA 和 TVNCB 的杂交区域之间编码的是在 RFHV 和 KSHV 之间有 91% 相同的 106 个氨基酸的糖蛋白 B 分子的片断。KSHV 糖蛋白 B 的整个蛋白质序列示于 SEQ. ID NO: 94，RFHV/KSHV 亚族的第三成员 RFHV2 的糖蛋白 B 片断示于 SEQ. ID NO: 97。

有很多已被测序的其它疱疹病毒的糖蛋白 B 分子的同源性残基, SEQ. ID NO: 2 或 SEQ. ID NO: 4 而不是其它疱疹病毒已知的序列中所含的最长序列为 9 个氨基酸长, 所述其它已知序列中有两个除外 (SEQ. ID NOS: 64 和 65)。在糖蛋白 B 氨基酸序列的别处发现了较长的匹配部分, 5 最长的是 SEQ. ID NO: 99 所示的得自 bHV4 的 21 个氨基酸长的序列, 其余的都是 16 个氨基酸长或更短。除 SEQ. ID NO: 99 以外, RFHV 和 KSHV 糖蛋白 B 蛋白质序列的 17 个氨基酸或更长的任何片段据信分别对于 RFHV 或 KSHV, 或密切相关的株而言是特异性的。由于据信 bHV4 和具有匹配区段的其它病毒不能感染灵长类动物, 因此 SEQ. ID NO: 4 中所含的在 10 灵长类动物中发现的约 10 个或更多个氨基酸的任何片断会显示与 KSHV 密切相关的感染因子的存在。

本发明既包含 RFHV/KSHV 亚族疱疹病毒完整的糖蛋白 B, 也包含对此亚族有特异性的其任何片段。本发明的优选糖蛋白 B 片段至少 10 个氨基酸长; 更优选它们至少 13 个氨基酸长; 更优选它们至少至少 17 个 15 氨基酸长; 更优选它们至少 20 个氨基酸长; 甚至更优选它们至少约 25 个氨基酸长, 再更优选它们至少约 30 个氨基酸长。

SEQ. ID NOS: 2, 4, 94 和 96 所示的 RFHV 和 KSHV 糖蛋白 B 片段的氨基酸序列可用于鉴定病毒特异性和交叉反应性的抗原区域。

原则上, 特异性的抗体会识别序列间非抗体来源物种共有的任何氨基酸 20 差异。抗体结合位点一般应足够大以包含抗原的 5 至 9 个氨基酸残基, 所述位点应足能识别单个氨基酸的差异。特异性抗体可能是多克隆应答的一部分, 所述应答在被表达糖蛋白 B 的病毒感染的动物中自发地产生。也可以通过给实验动物注射完整的糖蛋白 B 或糖蛋白 B 片段来诱导特异性的抗体。

25 因此, KSHV 独特的 5 个或更多个氨基酸的任何肽是潜在的病毒特异性抗原, 可以被 KSHV - 特异性抗体识别。类似地, 在 RFHV/KSHV 亚族内部, 而不是与其它疱疹病毒之间共用的足够长度的任何肽是潜在的亚族特异性抗原。

30 优选肽的一些例子示于表 8。本领域技术人员很快会认识到通过细微改变所列肽的长度和/或在任一方向上将肽的框架移动几个残基即可设计出具有类似特异性的其它肽。

在 γ 疱疹病毒亚族的某些其它成员的糖蛋白 B 和 KSHV 的糖蛋白 B 之

间，表 8 所示的 I 类肽是保守的。针对此区域中的一个这种糖蛋白 B 的抗体因此能与一些其它抗体交叉反应。在 RFHV 和 KSHV 的糖蛋白 B 之间 II 类肽是保守的，但在其它 γ 疱疹病毒之间 II 类肽却不保守。针对此区域的抗体预期在 RFHV, KSHV 和 RFHV/KSHV 亚族的其它病毒之间能交叉反应；但不能与此亚族以外的疱疹病毒交叉反应。RFHV, KSHV 和其它已知的 γ 疱疹病毒之间的 III 类肽有所不同。与此区域，特别是与其中所含的不相同的残基结合的抗体预期可将 RFHV 与 KSHV 糖蛋白 B 互相区分，并与更远相关的疱疹病毒的糖蛋白 B 区分开来。

表 8: 抗原肽				
特异性		序列	长度	SEQ. ID NO:
I 类: 与下列病毒 共用 在 RFHV/ KSHV 亚族 和一些其它 γ 疱疹病毒 中共用	bHV4	YRKIATSVTVYRG	13	64
	bHV4, mHV68	RYFSQP	6	66
	bHV4	IYAEPGWFPGIYRVR IYAEPGWFPGIYRVRTTVNCE	15 21	65 99
	mHV68	VLEELSRWCREQVRD	16	100
	II 类: 在 RFHV/ KSHV 亚族中 共用	VTVYRG	6	67
	AITNKYE	7	68	
	SHMDSTY	7	69	
	VENTFTD	7	70	
	TVFLQPV	7	71	
	TDNIQRY	7	72	
III 类: 病毒特异 性的 ¹	特异于			
	RFHV	RGMTEAA	7	73
	KSHV	RGLTESA	7	75
	RFHV	PVLYSEP	7	74
	KSHV	PVIYAEP	7	76

¹-不与任何其它已测序的疱疹病毒共用; 可能存在于一些未测序的 RFHV/KSHV 亚族的病毒中

特别优选的 III 类肽是那些如实施例中所所述的包含糖蛋白 B 区域的肽, 所述区域具有适于抗原表位的极性特征。给定 KSHV 和 RFHV/KSHV 亚族其它成员的糖蛋白 B 的完整序列, 通过类似分析可预测分子其它区域的病毒或亚族特异性肽。

多肽的制备

可通过本领域技术人员已知的几个不同方法制备本发明的多肽。

例如, 通过化学合成可由序列资料方便地制备长度约为 5-50 个氨基酸的短多肽。优选的方法是固相 Merrifield 技术。或者可以根据前

述方法中的一个分离或合成编码所需多肽的 mRNA，并使用体外翻译系统，如兔网织红细胞系统翻译，例见 Dorsky 等人。

5 使用适当的表达系统可以方便地制备大到并包括整个糖蛋白 B 的较长的多肽。例如，将全长 cDNA 的编码链与适当的启动子有效连接，插入表达载体，并转染适当的宿主细胞。再在允许转录和翻译发生的条件下培养宿主细胞，随后回收多肽。例如从其它种疱疹病毒中表达和回收糖蛋白 B，参见美国专利号 4,642,333 (Person)；5,244,792 (Burke 等人)；Manservigi 等人。

10 对于多种目的而言，使用重组的糖蛋白 B 多核苷酸特别方便，所述多核苷酸包括编码转运至细胞表面的信号的区域，但缺乏编码蛋白质跨膜区的区域。可在跨膜编码区的 5' 截短多核苷酸，或者多核苷酸可含有细胞外和胞质这两个编码区，但缺乏跨膜区。预期这种特性的构建体可从细胞中以可溶的形式分泌出来。在需要糖蛋白 B 片断为单体时，可将重组体设计成能限制蛋白质的约前 475 个氨基酸的翻译。

15 例如，为了在酵母中表达任何这些形式的糖蛋白 B，可使用甘油醛-3-磷酸-脱氢酶 (GAPDH) 启动子区和终止区制备盒 (cassette)。在酵母文库中鉴定 GAPDH 基因片断，分离之并以适当的构型连接。将所述的盒克隆到 pBR322 中，分离之并通过 DNA 测序进一步证实。构建出的 pCI/I 质粒含有糖蛋白 B 插入物和 GAPDH 启动子区和终止区。用此质粒转化酵母菌株啤酒糖酵母 (*S. cerevisiae*)，培养后离心沉淀酵母细胞，并将细胞沉淀物重新悬浮于含有蛋白酶抑制剂的缓冲液中，所述抑制剂如 1mM 的苯甲基磺酰氟和 0.1 μ g/ml 的胃酶抑制剂。用玻璃珠旋涡破碎经洗涤的细胞并重新离心。通过例如使用按下文所述制备的抗糖蛋白 B 的抗体进行 Western 印迹可进一步证实上清液中正确大小的糖蛋白 B 的存在。通过标准的蛋白质化学技术的联合可从上清液中纯化糖蛋白 B，所述技术包括离子交换层析，使用抗体或底物的亲和层析和高压液相层析。

25 为了在哺乳动物细胞中表达糖蛋白 B，例如可使用如 pSV1/dhfr 的哺乳动物表达载体。此质粒具有氨苄青霉素-抗性的 β 内酰胺酶基因，和与 SV40 早期启动子连接的可选择的哺乳动物细胞标记物二氢叶酸还原酶。将糖蛋白 B 多核苷酸的钝端片断连接到 pSV1/dhfr 载体中，并用核酸内切酶消化以提供包括 SV40 启动子，糖蛋白编码区，和 SV40 剪接

和聚腺苷酸化位点的盒。例如使用质粒转化 dhfr 有缺陷的 CHO 细胞，并选择转化子，通过例如使用抗糖蛋白 B 的抗体作为第一抗体的免疫荧光可鉴定表达糖蛋白 B 的细胞。

5 在另一个实施例中，表达糖蛋白 B 的重组质料被克隆使之在附加型复制载体 pRP-RSV 中的 Rous 肉瘤病毒长的末端重复的控制之下，所述质粒含有人乳多空病毒 BK 的复制起点和早期区域，以及 dhfr 抗性标记物。然后可使用此载体例如转化人 293 细胞。通过使用缺乏跨膜区的糖蛋白 B 编码区，糖蛋白 B 多肽能以 0.15-0.25pg/细胞/天的速度组成型分泌到培养基中。在 0.6-6 μ M 氨甲蝶呤的存在下，由于附加型重组体的扩
10 增，产量可增加 10-100 倍。以这种方式制备的糖蛋白 B 特别适用于诊断新的和复发的疱疹病毒感染，并制备抗这种感染的保护性疫苗 (Manservigi 等)。

使用多肽评估疱疹病毒感染

15 在几个不同的应用中，本发明所包含的多肽也可被用来检测或评估个体中疱疹病毒感染的现状。

在一个应用中，编码疱疹病毒糖蛋白 B 一部分的多肽被提供成检测用的试剂以检测可特异性地识别它的抗体的存在。这种抗体可能存在于例如目前或继往暴露于疱疹病毒中的个体的循环系统中。

20 循环系统中抗糖蛋白 B 之抗体的存在能提供病毒感染的敏感的早期征兆。由于糖蛋白 B 是病毒包膜的功能性成分，因此它比病毒颗粒内被隔绝的其它转录本产量更高，比似乎只在病毒生命循环中短暂停留的转录本分布更广。另外，它不仅可被完整的病毒表达，也可被病毒可感染的细胞的非感染性产物，如 L-颗粒表达。已知多种疱疹病毒的糖蛋白 B 具有强烈的免疫原性，因此，个体中抗糖蛋白 B 的抗体的检测可以
25 是正在发展中的活跃的疱疹病毒感染，潜伏的感染，先前的暴露，或用糖蛋白 B 疫苗处理的征兆。

被测量其中的抗体水平的适当临床样品包括得自怀疑有 γ 疱疹病毒感染
的个体的血清或血浆。通过例如免疫测定法可测定抗体的存在。

本领域中已确立了大量免疫测定法以使用病毒肽进行抗体的定量
30 (例见美国专利号 5,350,671: Houghton 等人)。例如，可以将潜在含有特异性抗体的试验样品与预先确定的不限定量的试剂多肽混合。此试剂可含有直接结合的标记物，如酶或放射性同位素。对于液相检测而言，

通过如过滤或层析的分离技术除去未反应的试剂。或者，通过固相上的试剂首先捕获样品中的抗体。所述试剂可以是例如特异性的多肽，抗-免疫球蛋白，或 A 蛋白，然后用如结合有标记物的特异性多肽，抗-免疫球蛋白或 A 蛋白的第二试剂检测被捕获的抗体。至少一种捕获试剂或检测试剂必须是特异性多肽。在第三种变化中，含有多肽的细胞或组织切片可首先被含有抗体的试验样品覆盖，然后被如经标记的抗-免疫球蛋白的检测试剂覆盖。在所有这些例子中，复合物中捕获的标记物的量与试验样品中存在的特异性抗体的量正相关。可设计类似的试验，其中试验样品中的抗体与经标记的抗体竞争同有限量的特异肽结合，那么复合物中标记物的量与试验样品中特异性抗体的量负相关。比较试验样品和得自未被感染之个体的对照样品之间使用这些试验中的任一个得到的结果。

通过适当选择试剂多肽，可检测所需特异性的抗体。例如，如果使用完整的糖蛋白 B，或者使用含有疱疹病毒保守区域的片段，则在试验样品中检测到的抗体可能是病毒特异性的，交叉反应性的，或两者皆是。优选将多-表位的试剂用于与疱疹病毒感染有关的抗体的一般性筛选试验中。为了使试验特异针对抗 RFHV 或抗 KSHV 的抗体，可选择含有适当病毒糖蛋白 B 的非保守区的抗原肽，如表 8 的 III 类所列的那些肽，优选使用这种肽的混合物。为了同时检测抗 RFHV，KSHV，和密切相关的 γ 疱疹病毒家族的病毒，而不抗 sHV1 和 EBV 的抗体，可选择具有表 8 II 类所列肽之特性的抗原肽，优选使用这种肽的混合物。

一旦感染消退，在疱疹病毒感染期间刺激产生的抗体也会消退，或者它们也可以作为宿主免疫记忆的一部分被保存。在后一种情况下，通过测定抗体的类别可将因目前的感染产生的抗体与因免疫记忆产生的抗体区分开来。例如，可进行这样一种试验，其中试验样品中的抗体被特异性的多肽捕获，再用经标记的抗-IgM 或抗-IgG 使之显现出来。试验样品中 IgM 类特异性抗体的存在表明感染正在发生，而只存在 IgG 抗体表明活性是由于先前感染或免疫接种的免疫记忆而产生的。

使用多肽设计或筛选抗病毒药物

干扰糖蛋白 B 基因或基因产物会修饰感染的进程或疾病的进展。本发明的目的之一是提供一种方法，通过此方法可开发和试验有用的药物组合物和利用这种化合物治疗 γ 疱疹病毒感染的方法。特别优选可用于

治疗 RFHV, KSHV 和 RFHV/KSHV 亚族的其它成员感染的药物化合物。适当的药物可以干扰糖蛋白 B 基因的转录或翻译, 并能干扰由此基因编码的多肽的生物学功能。不需要知道干扰的机理; 只需要对于与感染过程相关的反应而言, 干扰是优先的。

- 5 优选的药物包括那些竞争干扰糖蛋白 B 与靶细胞上其底物, 如硫酸乙酰肝素和其类似物结合的药物, 另外优选的药物竞争干扰糖蛋白 B 与其它病毒包膜成分的任何相互作用, 所述成分可能是病毒行使一种生物学功能, 如侵入靶细胞所必需的。另外优选的分子能够交联或要不然的话固定化糖蛋白 B, 从而防止糖蛋白 B 结合其底物或行使在病毒感染性
- 10 中起作用的任何生物学功能。

本发明提供了筛选候选药物以决定哪种适于临床使用的方法。此方法会瞄准目前已知的抗病毒化合物和那些将来设计的抗病毒化合物。

- 此方法包括将有活性的糖蛋白 B 与候选药物联合, 并测定候选药物是否改变了生化功能。糖蛋白 B 可以由 RFHV/KSHV 亚族的糖蛋白 B 基因编码的具有糖蛋白 B 活性的任何片段。通过表达编码分子活性位点的
- 15 经遗传工程改造的多肽, 或通过用蛋白酶裂解糖蛋白 B 并纯化活性片段可得到适当的片段。在优选的实施方案中, 提供了完整的糖蛋白 B。反应混合物也可含有测量蛋白质生物活性所必需的其它成分。例如, 在测量底物结合的试验中, 可以提供硫酸乙酰肝素或其类似物, 可以将它们
- 20 与固相支持物连接以便于测量结合反应。

- 筛选方法的一个实施方案是测量候选药物直接与分离的糖蛋白 B 或其片断的结合。预期与分子的活性位点结合的化合物可以干扰糖蛋白 B 活性, 因此, 将含有活性位点的整个糖蛋白 B 或片断与候选药物混合。通过例如以放射性标记或稳定的 γ -同位素标记的形式提供候选药物可直接
- 25 测量候选药物的结合。通过例如用适当抗体沉淀糖蛋白 B, 或通过提供与固相结合分子并在反应后洗涤固相可以检测与糖蛋白 B 结合的标记物的存在。通过例如用示差光谱学, 核磁共振或圆二色性检测构象的变化也可显示出候选药物与糖蛋白 B 的结合。或者, 也可以在竞争试验中检测结合: 例如, 将糖蛋白 B 与候选药物混合, 然后加入调节性亚单
- 30 位的经标记的核苷酸或片断。候选药物与生化相关位点的结合应能抑制经标记的化合物随后的结合。

筛选方法的第二个实施方案是测量候选药物抑制糖蛋白 B 与底物或

底物类似物结合的能力。优选的类似物是与如 Sepharose™ 珠的固相支持物偶联的肝素。通过例如提供经标记的糖蛋白 B，将它与候选药物一起保温，加入亲和性树脂，然后洗涤并计数树脂以测定是否候选药物已降低了被结合的放射性的量，即可测量抑制作用。也可以检测候选药物竞争干扰糖蛋白 B 与其它疱疹病毒蛋白质之间的相互作用的能力。

筛选方法的第三个实施方案是测量候选药物抑制如病毒颗粒的活性颗粒由糖蛋白 B 介导的活性的能力。此颗粒经改造可以表达糖蛋白 B 而不是能介导相同功能的其它成分。然后在存在和不存在候选药物的情况下，通过提供适当的靶以检测颗粒抑制生物学功能的能力，所述功能如底物结合或膜融合。

本发明也通过合理的药物设计提供了治疗疱疹病毒感染之药物的开发，一般可参见 Hodgson 和 Erickson 等人。在此实施方案中，或者通过以氨基酸序列为基础的预测模型，或者优选通过实验测定法可测定糖蛋白 B 的三维结构。实验的方法包括抗体作图，突变分析和抗-独特型的形成。特别优选 X-射线晶体学。知道糖蛋白的三维结构，特别是底物结合位点附近的重要氨基酸基团的方向，就可以从头设计化合物，或者可适当修饰现有的化合物。经设计的化合物要具有适当的电荷平衡，疏水性，和/或形状以使得它能在糖蛋白 B 活性位点的附近结合，并在空间上干扰此位点的正常生化功能。优选随后在上文所述的药物筛选试验中检测通过此方法设计的化合物。

抗糖蛋白 B 的抗体及其制备

本发明中包含的糖蛋白 B 分子的氨基酸序列对于所述病毒感染的宿主而言是外源的。已知其它种类疱疹病毒的糖蛋白 B 对于哺乳动物具有强烈的免疫原性。例如，人被 hCMV 感染的通常结果是形成抗-糖蛋白 B 的抗体。与之类似，预期 RFHV，KSHV 和 RFHV/KSHV 亚族的其它成员的糖蛋白 B 对于包括人的哺乳动物而言是免疫原性的。这些预期得到下文实施例所述观察结果的支持。

一般可通过本领域已知的任何方法制备抗多肽的抗体。为了在实验动物体内刺激抗体的产生，经常优选通过如与戊二醛聚合，或与如弗氏佐剂等佐剂联合这类技术增强多肽的免疫原性。将免疫原注射到适当的实验动物体内：优选注射到啮齿动物体内以制备单克隆抗体；优选注射到如兔或绵羊的较大动物中以制备多克隆抗体。优选在约 4 周后提供第

二次或加强注射，并在不少于约 1 周后开始收获抗体源。

从经免疫的动物体内收获的血清提供了多克隆抗体的来源。从来源材料中纯化特异性抗体活性的详细方法是本领域中已知的。如有必要，可通过如 A 蛋白层析，硫酸铵沉淀，离子交换层析，高效液相层析和在偶联至固体支持物上的免疫多肽柱上进行的免疫亲和层析之类的技术进一步纯化特异性抗体活性。

通过完整的糖蛋白 B 或含有保守序列的片段免疫产生的多克隆抗体在疱疹病毒之间可能有交叉反应。通过用适当的特异性抗原，如上文表 8 中所列的那些抗原免疫可产生病毒或亚族特异性的抗体。或者，通过例如使抗体穿过由糖蛋白 B 制成的吸附剂并收集未结合的级分以除去不需要的抗其它病毒糖蛋白 B 的活性，可以使得抗较大片段产生的多克隆抗体为特异性的。

或者，可以从免疫动物体内回收如脾细胞等免疫细胞并用于制备能产生单克隆抗体的细胞系，例见 Harrow & Lane (1988)，美国专利号 4,472,500 (Milstein 等人)，和 U. S. 4,444,887 (Hoffman 等人)。

简单地说，可以特别地通过细胞融合，或者通过用 Epstein Barr 病毒转化产生抗体的细胞，或者用致癌 DNA 转化来产生能生产抗体的细胞系。克隆并培养经处理的细胞，选择能产生所需特异性抗体的克隆。可通过多种技术对培养物上清液进行特异性检测，所述技术如在标准的免疫测定法中使用免疫多肽作为检测试剂，或者在免疫组化法中使用表达多肽的细胞。可从大体积的组织培养物上清液，或者从注射有克隆的适当准备的宿主动物腹水液中纯化选定克隆的单克隆抗体供应。

此方法的有效变动包括对分离的细胞进行多肽的免疫接种。通过标准的蛋白质化学方法，如用蛋白酶使抗体裂解可制备抗体片段和其它衍生物。通过得到编码抗体的多核苷酸，并应用一般的分子生物学方法以引入突变并翻译变异体可以产生抗体的经遗传工程改造的变异体。

通过注射完整的糖蛋白 B 或含有保守序列的片段而产生的单克隆抗体在疱疹病毒之间可能有交叉反应。通过用适当的特异性抗原，如选自表 8 的那些抗原免疫可产生病毒或亚族特异性的抗体。或者，通过在筛选方法中使用适当的抗原，如选自表 8 中的 III 类抗原可从经克隆的杂交瘤中选择病毒特异性的克隆。

抗疱疹病毒糖蛋白 B 的特异性抗体在开发，诊断和治疗工作中作用

很多。例如，抗体可用于筛选药物（见美国专利号 5,120,639），如下文所述，它们也可用作被动疫苗的成分或用作检测生物样品中的疱疹病毒和靶向药物。

也可制备与糖蛋白 B 相关的抗-独特型，通过根据上述方法学首先制备糖蛋白 B 抗体，通常是单克隆抗体可做到这一点。然后将抗体用作自愿者或实验动物的免疫原以产生抗-独特型。此抗-独特型可以是单克隆的也可以是多克隆的，它的形成一般根据第一抗体所用的方法学。使用免疫原抗体作为阳性选择子，使用不相关特异性的抗体作为阴性选择子来选择抗-独特型或表达抗-独特型的杂交瘤克隆。通常阴性选择子抗体是多克隆的免疫球蛋白制品或含有相同免疫球蛋白类和亚类并与免疫原抗体种类相同的单克隆抗体的集合体。抗-独特型可用作抗糖蛋白 B 的主动疫苗的可替换成分。

使用抗体检测生物样品中的糖蛋白 B

特异于糖蛋白 B 的抗体可被用于检测可能存在于，例如，固形组织样品和经培养的细胞中病毒来源的糖蛋白 B 多肽和片段。进行这种测定的免疫组织学技术对于本领域技术人员而言是显而易见的。一般情况下通过下述技术的联合可保护组织，所述技术可包括冷冻，交换成不同的溶剂，用如低聚甲醛的试剂固定，用如醇的试剂干燥，或在如石蜡或 OCT 的商购介质中包埋。适当制备样品的切片，并用特异针对蛋白质的第一抗体覆盖。

可直接提供具有适当标记物的第一抗体。更多情况下使用多种易制备或可商购的显色剂中的一种来检测第一抗体。典型地，这些显色剂是抗-免疫球蛋白或 A 蛋白，它们一般具有的标记物包括但不限于：如荧光素的荧光标记物，如过氧化物酶的能沉淀适当化合物的酶，如胶体金的电子密度高的标记物或如 ^{125}I 的放射性同位素。再使用适当的显微技术观察切片，比较怀疑受病毒感染的细胞和对照细胞之间的标记水平，所述对照细胞如感染区域周围或得自较远部位的细胞。

在标准的定量免疫测定法中，也可检测由糖蛋白 B 基因编码的蛋白质。如果蛋白质由受感染的细胞以任何可估计的量分泌或排出，则可在血浆或血清样品中检测到蛋白质。或者，可从固形组织样品中溶解或提取靶蛋白质。定量前可通过例如印迹技术或使用捕获抗体将蛋白质粘附于固相上。



本领域中已建立了大量进行定量的免疫测定法，例如，可将蛋白质与预定的非限制性量的特异针对蛋白质的试剂抗体相混合。试剂抗体可含有直接结合的标记物，如酶或放射性同位素，或者也可加入第二个经标记的试剂，如抗免疫球蛋白或 A 蛋白。对于固相测定法而言，通过洗
5 涤除去未反应的试剂。对于液相测定法而言，通过一些其它的分离技术，如过滤或层析除去未反应的试剂。复合物中捕获的标记物量与试验样品中存在的靶蛋白质量正相关。此技术的变化是竞争测定法，其中靶蛋白质与经标记的类似物竞争特异性抗体的结合位点，此时被捕获的标记物量与试验样品中存在的靶蛋白质的量负相关。比较试验样品和得自未受
10 感染之来源的对照样品之间使用任何这类检定法得到的结果。

使用抗体使药物靶向作用

抗体如何能被用于治疗疱疹病毒感染的例子是效应物成分的特异性靶向作用，被病毒感染的细胞一般呈现病毒的肽，尤其是在病毒包膜外部表达的蛋白质，因此肽为被感染的细胞提供了可与特异性抗体结合的
15 标记物。而与抗体结合的效应物成分在被感染的细胞附近变得密集；改善了对感染细胞的效应并降低了对未受感染之细胞的效应。并且，如果抗体能诱导胞吞作用，则可促进效应物进入细胞内部。

为了靶向作用这一目的，优选通过共价或高亲和性键使特异针对病毒多肽（此时为糖蛋白 B 的区域）的抗体与适当的效应物成分缀合。这种组合物中适当的效应物成分包括如 ^{131}I 等放射性核素，有毒的化合物
20 和如白喉毒素的有毒肽。另一个适当的效应物成分是可包裹于脂质体中的反义多核苷酸。

诊断试剂盒

诊断实验室，实验性实验室，技术人员或私人都可进行使用本发明的多核苷酸，寡核苷酸，肽或抗体的诊断方法。本发明提供了能用于这些装置的诊断试剂盒。在得自受感染个体的临床样品中，个体中存在疱疹病毒表现为样品中所含 DNA，RNA，蛋白质或抗体的改变。与得自健康
25 个体的样品中的成分相比，因疱疹病毒的存在而导致的这些成分之一的变化采取的形式是成分水平的增加或降低，或成分形式的变化。任选预处理临床样品以富集被检测的靶。然后使用者可使用试剂盒中所含的试剂以检测诊断成分改变的水平或形式变化。
30

每个试剂盒必须含有赋予方法特异性的试剂：用于检测靶 DNA 或 RNA



的试剂多核苷酸；用于检测靶蛋白质的试剂抗体；用于检测欲被分析的样品中可能存在的靶抗体的试剂多肽。可以适于贮存的固体形式或液体缓冲液形式提供试剂，然后当进行试验时，用于交换或加入反应介质中。还提供了适当的包装。此试剂盒可提供对此方法有用的另外的成分，这些可选择的成分包括缓冲剂，捕获试剂，显色剂，标记物，反应表面，检测的工具，对照样品，说明书和解释性的资料。

RFHV/KSHV 亚族的其它成员

RFHV 和 KSHV 是 RFHV/KSHV 亚族成员的例子。本发明包含编码如本文所限定的亚族其它成员糖蛋白 B 的多核苷酸序列。本文所述的共有 - 简并 γ 疱疹病毒寡核苷酸 Type1 和 Type2 引物以及方法被设计成适于鉴定 RFHV/KSHV 亚族其它成员的相应多核苷酸片断。一个这类成员是感染猴的另一个病毒，它被称为 RFHV2。如实施例 12 所述，从得自猕猴的 RF 组织中克隆了此病毒糖蛋白编码序列的区段。

为了鉴定和描述此家族中的其它成员，对提取自怀疑被这种病毒感染

的组织样品的 DNA 使用本发明的试剂和方法。

用于此目的适当来源的 DNA 包括得自在人和其它脊椎动物中出现的广范围疾病的生物样品。优选疾病中的因子与 γ 疱疹病毒亚族的其它成员相似，被怀疑是嗜淋巴细胞的；例如非 EBV 来源的感染性单核细胞增多。更优选的疾病在至少一个临床或组织学特征上类似于与 RFHV 或 KSHV 相关之疾病。这些疾病包括：a) 纤维增殖是疾病病理学一部分的疾病，尤其是与胶原沉积相关的疾病，和尤其是纤维组织被打乱

的疾病；b) 涉及血管发育异常的疾病；c) 涉及恶性转化的疾病，特别是但不限于淋巴细胞谱系的细胞；d) 隐伏的免疫缺损影响疾病的频率或严重性的疾病；e) 在器官的各个位点或在全身自发产生的疾病；f) 其流行病学的资料暗示与传染性

或环境因子相关的疾病。相对而言更优选满足不止一个上述标准的疾病。特别优选的一些疾病的例子包括腹膜后纤维变性，小结状纤维瘤病，假肉瘤纤维瘤病，纤维肉瘤，硬化肠系膜炎，急性呼吸病综合征，自发性肺纤维变性，不同类型的扩散增殖性肾小球肾炎，神经胶质瘤，成胶质细胞瘤，神经胶质增生和所有类型的白血病和淋巴瘤。

所用组织样品的类型取决于疾病的临床表现。更可能含有与疾病相关之病毒



它迹象的部位。被感染的个体的外周血单核细胞也可用作 RFHV/KSHV 亚族病毒的载体。在卡波济氏肉瘤 (Moore 等人, 1995b) 和 Castleman 病 (Dupin 等人) 患者的 PBMC 中已经检测到 KSHV。其它适当的来源是由这种来源形成的细胞培养物, 和得自这种来源的浓缩或分离的病毒制品。对于阴性对照样品而言, 组织可得自相同个体明显未受感染的部位, 或得自明显不会患有此病的匹配个体。

鉴定 RFHV/KSHV 亚族之成员的方法优选包括或单独地或联合地使用本发明提供的方法和试剂。

一个方法涉及扩增编码疱疹病毒糖蛋白 B 的多核苷酸, 所述多核苷酸得自从样品中提取的 DNA。例如可通过在如 PCR 的反应中扩增多核苷酸来实施此方法。在一个有所变动的方法中, 使用如表 4 所示的那些有广泛特异性的共有 - 简并 Type 1 寡核苷酸引发扩增反应, 此方法主要扩增 γ 型疱疹病毒。由于 RFHV/KSHV 亚族是一亚组 γ 疱疹病毒, 因此由此经变动的方法检测的糖蛋白 B 序列需被鉴定以进一步测定它们是否属于 RFHV/KSHV 亚族。在第二个有所变动的方法中, 使用如表 7 所列的那些 RFHV 或 KSHV 特异性的 Type 3 寡核苷酸, 或取自这些病毒的其它糖蛋白 B 多核苷酸区段引发扩增反应。扩增在低严谨性介质的条件下进行, 以使寡核苷酸与相关种类的病毒交叉 - 杂交。在更优选的有所变动的方法中, 使用如表 6 所列的那些 RFHV/KSHV 亚族特异性的 Type 2 寡核苷酸引发扩增反应。在适当的杂交条件下, 这些引物会优选扩增此亚族疱疹病毒的糖蛋白 B。

使用糖蛋白 B 多核苷酸探针检测的此亚族的优选成员是那些在 SEQ. ID NO: 1 或 SEQ. ID NO: 3 的残基 36 - 354 之间与 RFHV 或 KSHV 之糖蛋白 B 核苷酸序列至少有 65% 相同的成员。更优选那些至少约 67% 相同的成员; 更优选那些至少约 70% 相同的成员; 更优选那些至少约 80% 相同的成员; 甚至更优选那些至少约 90% 或更多相同的成员。

通过使用适当的探针对样品中的多核苷酸进行杂交试验也可鉴定亚族的成员。在例如通过使用 Type 1 或 Type 2 寡核苷酸作为引物, 在进行杂交试验以前, 任选扩增欲被检测的多核苷酸。然后在杂交反应中用适当的经标记的探针检测靶。优选探针含有 SEQ. ID NOS: 1 和 3 中 RFHV 或 KSHV 的糖蛋白 B 序列所含的至少 21 个核苷酸, 优选至少约 25 个核苷酸, 更优选至少约 50 个核苷酸。甚至更优选探针含有 SEQ. ID NOS:



1 或 3 的核苷酸 36 - 354。其它优选的探针包括如表 6 所示的那些 Type 2 寡核苷酸。选择的杂交条件允许探针与 RFHV/KSHV 亚族的糖蛋白 B 多核苷酸序列杂交, 而不与先前已测序的疱疹病毒, 尤其是 sHV1, bHV4, eHV2, mHV68, hEBV, hCMV, hHV6, hVZV 和 HSV1 的相应序列杂交。在这些条件下与受试多核苷酸形成稳定的双螺旋表明在样品中有得自 RFHV/KSHV 亚族之成员的多核苷酸存在。

也可以使用 II 类抗体鉴定亚族的成员, 所述抗体的制备上文已有概述。II 类抗体能与由 RFHV/KSHV 亚族成员产生的抗原交叉反应, 但不与其它抗原, 包括由非亚族成员的疱疹病毒产生的那些抗原交叉反应。通过例如在对制备自患有上文所列疾病之个体的组织切片进行免疫组化研究中, 使用抗体可检测新亚族成员。组织切片被抗体正染色表明样品中存在 RFHV/KSHV 亚族之成员的糖蛋白 B, 这可能是由于组织已被病毒感染。另外, 如果组织切片不能与 RFHV 和 KSHV 特异性 III 类抗体反应, 则组织中的糖蛋白 B 可能来自此亚族的另一个成员。类似地, 如果在个体的循环系统中发现 II 类抗体, 这表明个体已经被 RFHV/KSHV 亚族的成员现时或既往感染。

一旦通过上述方法中的任何一种鉴定出了推断的新病毒, 通过得到和测序此病毒糖蛋白 B 基因区域, 根据 RFHV/KSHV 亚族的定义将此区域与 RFHV 或 KSHV 的相应区域进行比较, 即可证实其为 RFHV/KSHV 亚族的成员。对于 RFHV/KSHV 亚族的新成员而言, 本发明的其它实施方案将致力于检测、诊断和药物开发的目的。使它们适于新亚族成员的变化如果是必需的, 则希望此变化较小, 并且根据新的序列资料是显而易见的, 或者可以常规调节。

RFHV/KSHV 亚族之糖蛋白 B 的变化形式

25 本发明也包括 RFHV/KSHV 亚族之糖蛋白 B 的变化形式。

对于 HSV1 和 hCMV 糖蛋白 B 的自然产生的和经诱导的突变的大量研究集中于分子的某些区域对于其多种生化功能的作用。例见 Reschke 等人和 Baghian 等人的羧基端氨基酸对融合的作用; Shiu 等人和 Pellett 等人的表位对中和抗体的作用; Gage 等人的涉及合胞体形成的分子区域; Navarro 等人 (1992) 的涉及病毒侵入和细胞 - 细胞传播的区域; Quadri 等人和 Novarro 等人 (1991) 的涉及生物合成过程中糖蛋白 B 细胞内转运的区域。



以前已研究过的病毒的糖蛋白 B 分子和本文所述的糖蛋白 B 分子之间的一些所述残基可能是保守的。类似地，预期 RFHV/KSHV 亚族糖蛋白 B 中相同残基的突变具有与所述的其它病毒的相应效应类似的效应。或者，可以联合不同糖蛋白 B 分子的功能性区域以产生具有经改变的功能的糖蛋白 B 重组体。例如，用非病原体病毒的糖蛋白 B 基因置换病原体病毒的相应基因可降低重组体的病原性 (Kostal 等人)。RFHV/KSHV 疱疹病毒亚族糖蛋白 B 的突变和重组都可导致减毒株，所述减毒株中感染性，复制活性或病原性有所降低。本发明中考虑了具有这些效应的糖蛋白 B 序列改变。

10 疱疹病毒的减毒株可用于例如开发多价疫苗。特别是在发展中国家需要提供能刺激免疫系统同时抗几种潜在的病原体的预防性疫苗。经遗传工程改造可表达几种不同病原体的免疫原性肽的病毒可实现此目的。由于疱疹病毒大的基因组可轻易容纳几千个碱基的编码肽的额外 DNA，因此疱疹病毒是特别合适的载体。理想的情况是病毒载体足够完整以表现出一些生物活性并吸引宿主免疫系统的注意，而同时又被充分减毒以免引起显著的病理状况。因此，RFHV/KSHV 亚族的减毒病毒可用作抗类似毒性形式的疫苗，并且可以被修饰以表达另外的肽和拓宽免疫保护的范

15 围。

疱疹病毒减毒形式的另一用途是作为基因疗法的传递载体 (Latchman 等人, Glorioso 等人)。为了有效，基因疗法中的多核苷酸必须被传递到靶组织部位。在治疗纤维性疾病，恶性肿瘤和相关疾病的过程中，由于 RFHV/KSHV 亚族的减毒病毒载体具有靶向受感染组织的能力，故优于其它靶向作用机理，包括其它疱疹病毒。在此实施方案中，病毒先被减毒，然后被修饰以含有基因疗法所需的多核苷酸，如前文所

20 概述的那些。

RFHV/KSHV 亚族疫苗中的糖蛋白 B

由于糖蛋白 B 在感染性病毒的包膜和被感染的细胞上较为显著，因此预期糖蛋白 B 能成为免疫效应物的有用的靶。疱疹病毒糖蛋白 B 一般是免疫原性的，可引起能中和病毒的抗体并预防此病毒进入复制期。另

30 外，糖蛋白 B 能产生 T-细胞反应，所述反应通过进攻宿主细胞中病毒复制的位点可有助于消灭发展中的病毒感染。

本发明包括疫苗组合物，和使用它们预防和处理 RFHV/KSHV 亚族病



毒感染的方法。

一个系列的实施方案与主动疫苗有关。这些组合物被设计成能刺激被治疗的个体抗糖蛋白 B 的免疫应答。它们一般含有糖蛋白 B 分子，其免疫原性的片断或变体，或者能表达糖蛋白 B 分子的细胞或颗粒。或者它们可含有编码免疫原性的糖蛋白 B 片断的多核苷酸（Horn 等人），优选的多核苷酸形式是表达载体。多核苷酸疫苗可任选含有如脂质体或病毒载体颗粒的传递载体，或者可以作为裸露的 DNA 被施用。

以呈递免疫原性的片断以刺激适当免疫细胞类型的增殖和/或生物学功能的方式设计本发明的疫苗组合物。目的在于引发抗体反应的组合物含有或编码 B 细胞表位，也可含有或编码能增强抗原 - 呈递细胞的摄取和展示，或能征集 T 细胞帮助的其他因子。目的在于产生 T 辅助细胞，尤其是 CD4⁺细胞的组合物一般含有在 II 类组织相容性分子的上下文中能被呈递的 T 细胞表位。目的在于刺激细胞毒 T 细胞和它们的前体，尤其是 CD8⁺细胞的组合物一般含有在 I 类组织相容性分子的上下文中能被呈递的 T 细胞表位。

在保护将来可能接触疱疹病毒的个体时，抗体应答可能已足够。预防性的组合物优选含有可引起 B 细胞应答的成分。成功地消灭处于发展中的疱疹病毒感染可能涉及细胞毒 T 细胞，T 辅助 - 诱导细胞，或这两者的参与。治疗发展中的感染的组合物优选含有能产生 T 辅助细胞和细胞毒 T 细胞的成分。甚至更优选能优先刺激 Type1 辅助（T_{H1}）细胞而不是 Type2 辅助（T_{H2}）细胞的组合物。有关主动疫苗的适当组合物的制备和检测在下文中有所概述。

另一系列的实施方案涉及被动疫苗和过继转移的其他材料。这些组合物一般含有立即准备参与病毒中和或消灭的抗糖蛋白 B 的特异性免疫成分。优选使用这些组合物的治疗方法可防止最近暴露于病毒的病理学结果。它们在不能对主动疫苗发起足够的免疫应答的免疫妥协的个体中也是优选的。这种个体包括那些具有先天性的免疫缺损，获得性免疫缺损（如那些被 HIV 感染或进行肾透析的个体），和那些接受免疫抑制治疗，如皮质类固醇的个体。

过继转移的适当材料包括如下文所述的针对糖蛋白 B 的特异性抗体，也包括免疫细胞的过继转移。例如，针对糖蛋白 B 反应的 T 细胞可取自供体，任选在体外克隆或培养，然后转移到组织相容性受体中。更



优选被转移的细胞是受体自身的细胞，并在体外被刺激。因此，从需要治疗的个体中纯化 T 细胞，在糖蛋白 B 免疫原性成分和适当的刺激因子的存在下培养 T 细胞以产生病毒 - 特异性的细胞，然后重新施用之。

5 本文所包含的某些组合物可能具有主动疫苗和被动疫苗这两者的特性，例如，由过继转移得到的糖蛋白 B 抗体可赋予即时的抗疱疹病毒的保护，也可通过抗 - 独特型网络，或通过增强病毒抗原的免疫呈递来刺激发展中的应答。

含有糖蛋白 B 多肽的疫苗

10 刺激抗糖蛋白 B 的免疫应答的疫苗的特异性成分包括完整的糖蛋白 B 分子和对于宿主而言为免疫原性的糖蛋白 B 片断。

通过上文所述的任何方法，尤其是由含有编码糖蛋白 B 的多核苷酸的适当表达载体纯化可制备完整的糖蛋白 B 及其较长的片断。得自其它病毒株的分离的糖蛋白 B 刺激保护性免疫应答（见美国专利号 15 5, 171, 568; Burke 等人）。优选的片断含有完整的病毒包膜外部暴露的分子区域；位于成熟蛋白质 N - 末端的约 650 个氨基酸内。

糖蛋白 B 的糖基化并不是免疫原性所必需的（O' Donnell 等人），因此，同样优选糖基化和非糖基化形式的分子。由标准技术可测定糖基化；例如，在用商购的 F 或 H 型内切糖苷酶处理之前和之后在 SDS 聚丙烯酰胺凝胶电泳中比较蛋白质的迁移率。

20 含有糖蛋白 B 特殊表位的 5 - 50 个氨基酸的较小片断也是适当的疫苗成分。通过上文所述的任何方法可制备这些片断；最方便的方法是化学合成。优选的片断是那些免疫原性的和在病毒包膜外部表达的片断。甚至更优选涉及糖蛋白 B 生物学功能的片断，所述功能如与细胞表面的受体结合或病毒穿透靶细胞。

25 可由本技术中已知的规则推测多种表位的免疫原性。通过例如鉴定可变的极性区可测定 B 细胞受体的抗原性区域（Hopp 等人，见实施例 9）。通过例如鉴定能在组织相容性分子的呈递沟中形成两亲性螺旋的区域可测定 T 细胞受体的抗原性区域。通过与其它病毒种类的糖蛋白 B 分子的类比也可鉴定抗原性区域。例见 Sanchez-Pescador 等人和 Mester 等人的 HSV1 之 B 细胞表位；Liu 等人的 hCMV 之 HLA - 限制的 T 辅助细胞表位；和 Hanke 等人的 HSV1 之细胞毒 T 淋巴细胞表位。

30 通过多种不同的技术可实验性地测量多种表位的免疫原性。这些技



术一般包括制备含有潜在的抗原性区域的长为 5-20 个氨基酸的蛋白质片断，并在特异性的生物试验中检测它们。通过 CNBr 和/或蛋白酶裂解降解糖蛋白 B 的较大区段，并通过例如凝胶电泳和印迹到硝酸纤维素上纯化可制备片断 (Demotz 等人)。通过标准的肽合成也可制备片断

5 (Schumacher 等人, Liu 等人)。在优选的方法中，使用尼龙膜支持物上的 F-Moc 化学，根据成熟的糖蛋白 B 分子的整个细胞外区域合成有 8 个残基重叠的 12 个氨基酸的连串肽 (见实施例 11)。

10 然后测定暴露于完整病毒或糖蛋白 B 成分的个体的样品中针对经制备的片断的反应性。通过故意施用已将此个体实验性地暴露于糖蛋白 B 成分，或者，此个体已自然发生病毒感染，此自然感染优选通过阳性扩增反应来证实，所述扩增反应使用糖蛋白 B 或 DNA 聚合酶的病毒-特异性寡核苷酸探针。通过标准技术从个体中取得血液样品，并用于制备血清，T 细胞，和外周血单核细胞 (PBMC)。

15 可在酶联免疫吸附测定法中检测血清中糖蛋白 B 特异性抗体的存在，例如，将与如尼龙膜的固相支持物结合的肽与血清一起保温，洗涤，与酶联抗-免疫球蛋白一起保温，和使用酶底物显色。试验孔中反应产物水平比含有不相关肽的孔中反应产物的水平高显示针对特殊的糖蛋白 B 肽的抗体的存在 (实施例 11)。

20 可在增殖试验中检测淋巴细胞制品中糖蛋白 B 特异性 T 辅助细胞的存在。在自身扩散或扩散的作为抗原呈递细胞的 10^5 个 PBMC 存在的条件下，将约 2×10^4 个 T 辅助细胞与 $10^{-4} - 10^{-6}$ M 的肽一起保温约 3 天。在培养的约最后 16 小时加入 [3 H] 胸苷，然后收获细胞并洗涤，经洗涤的细胞的放射性水平比那些在缺乏肽的条件下培养的细胞高约 10 倍反映了特异于此肽的 T 细胞的增殖 (Liu 等人)。如有必要，可克隆具有 CD3⁺
25 4⁺ 8⁻ 表型的细胞以进一步鉴定 T 辅助细胞反应。

30 在 51 Cr 释放试验中检测淋巴细胞制品中糖蛋白 B 特异性细胞毒 T 细胞的存在。通过用含有可表达的糖蛋白 B 基因的疱疹病毒感染同种异型细胞来制备靶，或者，可使用被糖蛋白 B 表达载体转染的同种异型细胞。在 37°C 将靶与 51 Cr 保温约 90 分钟，然后洗涤。在 37°C 将约 5×10^4 个靶细胞与 $10^{-4} - 10^{-5}$ M 的肽和 $0.1 - 2 \times 10^4$ 个试验 T 细胞保温约 30 分钟。释放至上清液中的放射性水平比自发裂解产生的放射性高反映了 CTL 活性。如有必要，可克隆具有 CD3⁺ 4⁻ 8⁺ 表型的细胞以进一步鉴定 CTL 反



应。

在疫苗中，糖蛋白 B 肽可任选与相同病毒的其它肽联合，适当的肽包括疱疹病毒任何其它成分的肽，如糖蛋白 C, D, H, E, I, J 和 G。糖蛋白 B 肽也可任选与不同病毒的免疫原性的肽联合以提供抗不止一种病原性有机体的多价疫苗。通过制备肽溶液的混合物，或通过合成多种肽成分相连的融合蛋白可联合肽。

含有适当表位的糖蛋白 B 形式可任选被化学处理以增强它们的免疫原性，尤其是如果它们含有 100 个或更少的氨基酸时。这种处理可包括例如与戊二醛交联；与如匙孔血蓝蛋白 (KLH) 或破伤风类毒素的蛋白质载体连接。

肽或肽混合物可单独使用，但正常情况下将与生理学和药理学可接受的赋形剂联合，所述赋形剂如水，盐水，生理缓冲盐水或糖溶液。

在优选的实施方案中，主动疫苗也含有佐剂，佐剂可增强免疫原的呈递或要不然的话强调针对免疫原的免疫应答，所述佐剂如明矾，氢氧化铝， β -2 微球蛋白 (WO 91/16924: Rock 等人)，胞壁酰二肽，胞壁酰三肽 (美国专利号 5,171,568: Burke 等人) 和单磷脂 A (美国专利号 4,436,728: Ribi 等人；和 WO 92/16231: Francotte 等人)。也可存在如白介素 2 的免疫调节剂。肽和其它成分 (如果有的话) 任选被包裹于脂质体或微球体中。多种佐剂的实验性检测的概述见美国专利号 5,171,568 (Burke 等人)。多种佐剂有效。佐剂的选择至少部分取决于在佐剂存在下疫苗的稳定性，施药的途径，和尤其是当给人使用时佐剂的调节接受性。

多肽疫苗一般具有宽的有效范围，通常的施药途径是肌内，但也要开发通过其它途径给药也有效的制品，所述其它途径包括静脉内，腹膜内，口腔，鼻内和通过吸入给药。肌内给药的每剂量疫苗中糖蛋白 B 多肽的总量一般约为 $10 \mu\text{g} - 5\text{mg}$ ；通常约为 $50 \mu\text{g} - 2\text{mg}$ ；更通常约为 $100 - 500 \mu\text{g}$ 。优选首先以引发剂量施用疫苗，然后通常在至少 4 周后，再以加强剂量施用疫苗，可在定期的基础上以进一步加强的剂量给药以增强或复壮反应。

含有表达糖蛋白 B 的病毒颗粒的疫苗

主动疫苗也可被制备成表达糖蛋白 B 免疫原性表位的颗粒。

一个这种疫苗含有重组疱疹病毒的 L-颗粒 (见美国专利号



5, 284, 122: Cunningham 等人)。重组病毒基因组中的衣壳成分有缺陷，或要不然的话不能形成完整的病毒；然而，它保留了制备 L-颗粒的能力。经遗传工程改造的基因组包括与重组病毒的控制元件有效相连的本发明编码糖蛋白 B 的多核苷酸。然后例如在经培养的细胞中培养病毒，
5 通过在适当的梯度，如 FICOLL™ 中离心以纯化此颗粒。这种制品不含有感染性的病毒，并能表达大量不同的合乎需要的表位的肽成分。

另一个这种疫苗含有将本发明的糖蛋白 B 作为异源抗原表达的活病毒，这种病毒包括 HIV, SIV, FIV, 马感染性贫血病毒，维斯纳病毒和其它种类的疱疹病毒。被处理的病毒种应该是天生非病原性的；或者
10 应该通过遗传修饰例如降低复制或毒力以减毒。通过突变涉及复制的基因，如 DNA 聚合酶基因也可使疱疹病毒减毒。通过缺失必需的晚期成分，如糖蛋白 H 也可使疱疹病毒减毒 (WO 92/05263: Inglis 等人)。活疫苗可以在宿主中低水平地复制，这样做的特殊条件是能增强蛋白质的表达，但不至于在被治疗的受试者中引起任何病理学表现。

15 优选的制备活疫苗的病毒种类是腺病毒。对于治疗人而言，人腺病毒 4 型和 7 型已显示出没有副作用，适合将它们用作载体。因此，本发明的糖蛋白 B 多核苷酸可以被遗传工程改造，例如插入病毒基因组的 E1 或 E3 区域。已知表达 HSV1 或 HSV2 的糖蛋白 B 的腺病毒载体能刺激高滴度的可中和病毒的抗体产生 (McDermott 等人)。此反应可保护实验
20 动物抵抗各个活病毒的致死量攻击。

活重组疫苗中的病毒也优选重组痘病毒，尤其是痘苗。甚至更优选已经被修饰以灭活非必需的毒力因子的痘苗病毒株，所述修饰为缺失或插入与所述因子相关的开放阅读框 (美国专利号 5, 364, 773: Paoletti 等人)。为了制备疫苗，将本发明编码糖蛋白 B 的多核苷酸经遗传工程
25 改造插入病毒基因组，并在痘苗病毒启动子的控制下表达。这种类型的重组体可直接用于以每剂量约 $10^7 - 10^8$ 个噬斑形成单位来免疫接种。单次剂量足以刺激抗体应答。含有 HSV1 的糖蛋白 B 的痘苗病毒重组体可有效保护小鼠抵抗致死量的 HSV1 感染 (Cantin 等人)。

这一类型的另一个疫苗是自我装配的复制缺陷型杂合病毒。例见 WO
30 92/05263 (Inglis 等人)。此颗粒可含有例如衣壳和包膜糖蛋白，但不含有完整的病毒基因组。病毒包膜中的一个糖蛋白是本发明所包含的糖蛋白 B。



在优选的实施方案中，通过第一类病毒载体以及异源种类的衣壳和包膜编码区生产颗粒，所述第一类病毒载体具有足够的该种类基因组区段以进行复制（见美国专利号 5,420,026: Payne）。选择第一种类的遗传元件以使载体对真核细胞的感染产生能自我装配成复制缺陷型颗粒的衣壳和包膜糖蛋白。在此实施方案的变体中，由衍生自第一病毒种类的两个分离的载体提供编码衣壳和包膜糖蛋白的多核苷酸。衣壳编码区可得自慢病毒属，如 HIV, SIV, FIV, 马感染性贫血病毒或维斯纳病毒。包膜编码区含有本发明编码糖蛋白 B 的多核苷酸。优选所有的包膜成分由疱疹病毒，尤其是 RFHV/KSHV 亚族的疱疹病毒编码。通过用载体感染敏感的真核细胞系，如 BSC-40 并在约 18 小时后收获上清液能得到有缺陷的病毒颗粒。如有必要，可通过穿过蔗糖垫层离心进一步纯化病毒颗粒。在作为疫苗施用前也可以在 40℃ 下用 0.8% 的福尔马林将颗粒处理 24 小时。

含有活的减毒病毒或病毒类似物的疫苗可被冻干以冷藏。稀释剂可任选包括组织培养基，山梨醇，明胶，碳酸氢钠，白蛋白，盐水溶液，磷酸盐缓冲液和无菌水。可任选加入其它活性成分，如麻疹，腮腺炎和风疹病毒的减毒株以产生多价疫苗。通过例如气注技术可使悬浮液冻干。按下述进行冻干：将疫苗小管放入预冷至约 -45℃ 并充有 10-18 Pa 干燥无菌氮气的冻干室中，每小时将温度升高约 5-25℃ 直至温度达 +30℃，在完全真空条件下进行第二次冻干循环，然后在氮气下以一般性的方式密封小管（见 EP0290197B1: McAleer 等人）。对于含有活疱疹病毒的疫苗，最终的冻干制品优选含有 2-8% 的水分。

据认为主动疫苗的大量可替换的组合物不局限于本文详细描述的那些，都能有效引发特异性 B- 和 T- 细胞免疫力。所有这种组合物都包括在本发明的精神中，条件是它们包括作为活性成分的 RFHV/KSHV 亚族糖蛋白 B 多核苷酸或多肽。

含有糖蛋白 B 抗体的疫苗

通过过继转移施用针对 RFHV/KSHV 亚族糖蛋白 B 的抗体以直接赋予被治疗的受试者一定的体液免疫力。被动施用抗糖蛋白 B 的抗体甚至可实验性地保护 T- 细胞免疫妥协的受试者以抵抗其它疱疹病毒的致死量攻击（Eis-Hubinger 等人）。

需要抗糖蛋白 B 的保护作用，所用的抗体分子应该针对糖蛋白 B。



所述抗体应不与其它抗原，尤其是被治疗的受试者的内源性抗原交叉反应。抗体可特异针对整个 RFHV/KSHV 亚族（II 类抗体），或者针对特殊的病毒种类（III 类抗体），这取决于治疗的目的。优选抗体对多价抗原的全部亲和性至少约为 10^8 M^{-1} ；更优选至少约为 10^{10} M^{-1} ；更优选至少约为 10^{12} M^{-1} ；甚至更优选至少约为 10^{13} M^{-1} 或更高。可使用完整的抗体分子，重组体，融合蛋白或抗体片断；然而，优选能表达天然抗体效应物功能的完整的抗体分子或重组体。相关的效应物功能包括但不限于病毒聚集；依赖于抗体的细胞的细胞毒性；补体激活；和调理作用。

5 根据上文提供的描述可制备抗体。对于全身性保护而言，优选抗体是单体，并优选抗体为 IgG 类单体。对于粘膜保护而言，抗体可以是多聚体，优选是 IgA 类多聚体。抗体可以是单克隆的也可以是多克隆的；典型地优选单克隆抗体的混合物。也优选得自最初抗体来源的其它生物学成分基本上纯净的制品。纯化后可加入具有所需反应性的其它抗体分子，和载体或稳定剂。

15 在有些情况下，需要抗体与需被治疗的种的抗体尽可能地密切相似。这可预防被施用的抗体将其自身变成受体免疫应答的靶。当受试者具有主动免疫系统，或当抗体以重复的剂量被施用这种类型的抗体尤其合乎需要。

20 因此，本发明包括人或人源化的抗-糖蛋白 B 的抗体。可从先前已被各个 RFHV/KSHV 亚族疱疹病毒感染的人个体血清中，或施用了主动疫苗的志愿者中纯化多克隆的人抗体。可由得自这种个体的例如外周血的淋巴细胞产生单克隆的人抗体。一般可根据上文概述的方法产生人杂交瘤。通常，稳定的人杂交瘤的产生需要多种技术的联合，如既与人骨髓瘤细胞系融合又用例如 EBV 转化。

25 在优选的方法中，由具有功能性人免疫球蛋白基因座的嵌合的非灵长类动物产生人抗体（WO 91/10741: Jakobovits 等人）。非灵长类动物株（典型地为小鼠）不能表达内源性免疫球蛋白的重链，并且最理想的情况下也不表达至少一个内源性免疫球蛋白的轻链。此动物经遗传工程改造可以表达人重链，并且最理想的情况下也能表达人轻链。用

30 RFHV/KSHV 亚族疱疹病毒的糖蛋白 B 接种这些动物，然后用它们的血清制备多克隆抗体，用它们的淋巴细胞以一般性的方式制备杂交瘤。适当选择和纯化后，所得抗体是具有所需特异性的人抗体。

在另一个优选的方法中，首先在如小鼠等的另一种类中开发针对糖蛋白 B 的具有所需特异性的单克隆抗体，然后再人源化。为了使抗体人源化，可分离编码特异性抗体的多核苷酸，得到抗原结合区，然后与编码不相关特异性的人免疫球蛋白元件的多核苷酸重组。或者，得到特异性抗体的核苷酸序列，并将此序列用于设计具有人特性的相关序列，通过例如化学合成可制备所述序列。用所需类型的人免疫球蛋白的恒定区取代特异性抗体的重链恒定区或轻链恒定区，优选同时取代这两者。优选互补决定区（CDR）外部的这两条链的可变区区段也被它们的人等价物取代（EP 0329400: Winter）。

甚至更优选可变区区段被它们的人等价物取代，条件是它们不涉及抗原结合或结合位点结构的维持。例如按 Padlan 所述可鉴定重要的氨基酸。在一个特定的技术中（WO 94/11509: Couto 等人），使用已知免疫球蛋白的序列和晶体学资料得到位置共有序列。将糖蛋白 B 特异性抗体的氨基酸序列与模式序列相比较，鉴定出涉及抗原结合，与 CDR 接触，或与相反链接触的氨基酸。在需要时可改变其它氨基酸使它们与人免疫球蛋白序列的共有序列一致。然后制备编码人源化序列的多核苷酸，将所述多核苷酸转染至宿主细胞，并用于产生具有与最初得到的抗体克隆相同的糖蛋白 B 特异性的人源化抗体。

使用这些方法中的任何一种得到的特异性抗体一般是无菌的，并与药物可相容的赋形剂混合。也可存在如 0.3 摩尔甘氨酸等的稳定剂和如 1: 10,000 乙基汞硫代水杨酸钠等的防腐剂。通过例如碳酸钠可将悬浮液缓冲至中性 pH（~ 7.0）。通过加入正常的人 IgG 可任选地调节效力，所述 IgG 是通过例如 Cohn 冷乙醇分级方法从正常血浆大库中得到的。如白蛋白溶液等的其它稀释剂可用作替换物。调节浓度以使单剂量施用为 0.005 - 0.2mg/kg，优选约 0.06mg/kg。优选单剂量导致循环水平的抗-糖蛋白 B，这可由 ELISA 或其它适当的技术来测定，所述水平可与在已接受过主动的糖蛋白 B 疫苗或已从相应病毒的急性感染中恢复的个体，或从实验工作中已知具有抗病理剂量的病毒攻击的保护作用的个体中观察到的水平相比。

一般应通过肌肉而不是静脉内注射进行施药，应小心操作以确保针头不进入血管。对于施用了人白蛋白后有全身性变态反应病史的个体应格外小心。如上文所述，对于预防性应用而言，抗体制品可任选与针对

糖蛋白 B 的主动疫苗联合施用。对于暴露后的施用而言，优选在暴露 1 周内施用抗体制品，更优选在 24 小时内，或暴露后尽可能地马上施用抗体制品。任选以约 3 个月的间隔给予下一剂量。

5 至于本文所述的所有治疗仪器，所用组合物的量，和施药的适当途径和时间表取决于临床状况和被治疗的特殊个体的需要。特殊方案的选择最终是开药的医生或兽医的责任。

前文描述特别地提供了有关疱疹病毒，尤其是 RFHV/KSHV 亚族的那些病毒的糖蛋白 B 编码区如何被鉴定和它们的序列如何被得到的详细解释。提供了 RFHV 和 KSHV 糖蛋白 B 编码区域的多核苷酸序列。

10 据信本文所列的 RFHV 和 KSHV 的多核苷酸序列是用于此研究之组织样品中疱疹病毒多核苷酸所含序列的精确复本，它们表示得自多个克隆的共有序列资料。然而据认为由如 PCR 的扩增方法得到的序列可能由于扩增而在序列中含有偶然的错误。对于单次测定而言，估计错误率在约 0.44% 至 0.75% 之间；对于总共 n 个不同的测定而言，错误率约为相同的频率被 $\sqrt{(n-1)}$ 除。不过，错误率可能高至 1% 或更高。通过制备疱疹病毒多核苷酸序列文库，使用如表 7 所提供的那些寡核苷酸选择相关克隆，并测定选定克隆中 DNA 的序列即可得到没有扩增错误的序列。相关的方法学是本领域普通技术人员所熟知的，他们也许希望参考下文实施例中的描述。

20 据认为出现了疱疹病毒的等位变体和逃逸变体，在不违背本发明精神的前提下，可分离或得到掺入突变的多核苷酸和多肽，所述突变或者为自然产生，或者为偶然地或故意地诱导产生。

提供下文所示实施例是为了进一步指导本领域的普通技术人员，而不是以任何方式限制本发明。

25

实施例

实施例 1: 疱疹病毒糖蛋白 B 的寡核苷酸引物

已知的疱疹病毒糖蛋白 B 的氨基酸序列得自 PIR 蛋白质数据库，或衍生自得自 GenBank 数据库的 DNA 序列，序列是通过计算机辅助的排列程序和通过手工排列的。

30

结果示于图 3，使用 sHV1，bHV4，mHV68，EBV 和 hHV6 序列鉴定尤其在 γ 疱疹病毒中相对较保守的区域。选择九个区域设计扩增引物。然后将这些区域的 DNA 序列用于设计寡核苷酸引物，引物被设计成在 3'

末端具有 8-14 碱基对的简并区段，在 5' 末端具有 18-30 碱基的共有区段。这可提供具有最适敏感性和特异性的引物。

5 简并区段延伸跨越了疱疹病毒糖蛋白 B 序列的高度保守的区域，该区域包含最少数目的可替换密码子。因此引物可以在简并位置用可替换的核苷酸残基合成并产生最少数目的组合。所得到的每个引物有不超过 256 个的可替换形式。

10 共有区段衍生自糖蛋白 B 序列的相应侧翼区域。通常，通过选择被分析的所有糖蛋白 B 序列的每个位置处最经常出现的核苷酸可得到共有区段。然而，更倾向于选择 C 或 G 核苷酸以使引物形成稳定双螺旋的能力最大化。

15 结果示于图 4-12，并概括于表 4。在 PCR 中，通过与编码链的反义链杂交，并在与糖蛋白 B 编码序列的相同方向上起聚合可将表 4 中所列的具有“有义”方向的寡核苷酸用作引物。表 4 中所列的具有“反义”方向的寡核苷酸会与编码链杂交并在与糖蛋白 B 编码序列相反的方向上起聚合。

根据设计好的序列的合成寡核苷酸订购自 Oligo Etc, Inc.

实施例 2: DNA 提取

20 活组织检查的样本得自被诊断为患 AIDS 的人受试者的卡波济氏肉瘤损害，样本被固定于低聚甲醛中并包埋于石蜡中，如此进行以供正常的组织学检查。

25 在 1.5ml 的 EPPENDORF™ 圆锥形离心管中用 500 μ l 的二甲苯提取石蜡样品的片段。在室温下将样品缓慢摇晃 5 分钟，将试管置于 EPPENDORF™ 台式离心机中，以 14,000rpm 的转速离心 5 分钟。用巴氏吸管除去二甲苯后，加入 500 μ l 95% 的乙醇，重新悬浮样品，然后再离心。除去乙醇，重复洗涤步骤。再将样品自然干燥约 1 小时，加入 500 μ l 蛋白酶-K 缓冲液（0.5% TWEEN™ 20，一种去污剂；50mM Tris 缓冲液，pH 7.5，50mM NaCl）和 5 μ l 蛋白酶 K（20mg/ml），将样品在 55 $^{\circ}$ C 下保温 3 小时。通过在 95 $^{\circ}$ C 下保温 10 分钟使蛋白酶 K 失活。

30 收集得自 KS 组织的 DNA 样品以提供用于扩增反应的多核苷酸的一致来源。如共同拥有的美国专利申请 60/001,148 中所述，通过 KSHV DNA 聚合酶序列的扩增可检测出此收集物中含有得自 KSHV 的 DNA。

实施例 3: 得到 KSHV 糖蛋白 B 的经扩增区段



根据下述方法，使用实施例 1 中得到的寡核苷酸扩增实施例 2 中从 KSHV 组织提取的 DNA 区段。

使用 2 μ l 收集的 DNA 模板，1 μ l 寡核苷酸 FRFDA (50pmol/ μ l)，1 μ l 寡核苷酸 TVNCB (50pmol/ μ l)，10 μ l 10 \times 缓冲液，1 μ l 含有 2.5mM 的各种脱氧核糖核苷酸三磷酸 (dNTPs)，65 μ l 蒸馏水和 65 μ l 矿物油进行第一次 PCR 反应。在 Perkin - Elmer (Model 480) PCR 仪中将混合物加热至 80 $^{\circ}$ C，再加入 0.5 μ l Taq 聚合酶 (BRL, 5U/ μ l) 和 19.5 μ l 水。以下列顺序进行 35 次扩增循环：94 $^{\circ}$ C 1 分钟，退火温度 1 分钟和 72 $^{\circ}$ C 1 分钟。第一次循环中退火温度是 60 $^{\circ}$ C，每次循环降低 2 $^{\circ}$ C 直至达到 50 $^{\circ}$ C，使用 50 $^{\circ}$ C 作为退火温度进行其余的循环。

按下述进行第二次 PCR 反应：在前一步骤的 1 μ l 反应混合物中加入 1 μ l 寡核苷酸 NIVPA (50 pmol/ μ l)，1 μ l 寡核苷酸 TVNCB (50pmol/ μ l)，10 μ l 10 \times 缓冲液，1 μ l dNTPs，66 μ l 水和 65 μ l 矿物油。将混合物加热至 80 $^{\circ}$ C，加入于 19.5 μ l 水中的 0.5 μ l Taq 聚合酶。使用与上相同的温度逐步降低的方法进行 35 次扩增循环。通过在 2% 的琼脂糖凝胶上电泳和用溴化乙锭染色来分析 PCR 产物。

使用 14 个试验缓冲液进行两轮扩增，5 个缓冲液产生的 PCR 产物的大致的大小由与其它疱疹病毒序列类似的方法推测。这些缓冲液包括 WB4 缓冲液 (10 \times WB4 缓冲液是：0.67M Tris 缓冲液，pH8.8，40 mM MgCl₂，0.16M (NH₄)₂SO₄，0.1M β - 巯基乙醇，1mg/ml 牛血清白蛋白，此缓冲液在反应时被稀释 10 倍)。也试验了 WB2 缓冲液 (与 WB4 缓冲液相同，除了在 10 倍浓缩液中为 20 mM MgCl₂)。另外还试验的缓冲液当稀释至最终的反应体积时含有 10mM Tris, pH8.3, 3.5 mM MgCl₂ 和 25mM KCl; 或 10mM Tris, pH8.3, 3.5 mM MgCl₂ 和 75mM KCl; 或 10mM Tris, pH8.8, 3.5 mM MgCl₂ 和 75mM KCl。WB4 缓冲液显示了最强的带，和一些另外的较微弱的带。这可能是由于 WB4 样品中经标记的扩增的多核苷酸的总量较多的原因。

选择用 WB2 缓冲液扩增的产物以进一步研究。进行第三轮循环以引入放射性标记。用 γ ³²P - ATP 末端标记最后使用的寡核苷酸 (TVNCB)，在得自前一扩增步骤的 20 μ l 反应混合物中加入 γ ³²P - ATP 和 1 μ l 2.5mM dNTP。将混合物加热至 80 $^{\circ}$ C，加入 0.5 μ l Taq 聚合酶。经过 94 $^{\circ}$ C，60 $^{\circ}$ C 和 72 $^{\circ}$ C 的 5 轮循环进行扩增，使用 8.8 μ l 得自 Circumvent 测序凝胶

试剂盒的上样缓冲液终止反应。

于 51°C 在 6% 的聚丙烯酰胺测序凝胶上将经放射性标记的反应产物的 ~4 μ l 等分试样电泳 1.5 小时，将凝胶干燥 1.5 小时，通过暴露 12 小时产生放射性自显影，鉴定两个带，较大的带具有类似于其它 γ 疱疹病毒序列的片断的预期大小。

较大的带被标明并切下，通过在 40 μ l TE 缓冲液 (10mM Tris 和 1mM EDTA, pH8.0) 中保温洗脱 DNA。使用 1 μ l 洗脱液, 10 μ l 10 \times WB2 缓冲液, 1 μ l 2.5mM dNTP, 1 μ l 第二套寡核苷酸引物中的每一种 (NIVPA 和 TVNCB) 和 65 μ l 水对提取的 DNA 进行进一步的扩增反应。将混合物加热至 80°C, 加入于 19.5 μ l 水中的 0.5 μ l Taq 聚合酶。使用与上相同的温度逐步降低的方法进行 35 次扩增循环。

实施例 4: KSHV 糖蛋白 B 的 386 个碱基之片段的序列

根据下述方法纯化和克隆 KSHV 糖蛋白 B 基因的经扩增的多核苷酸片断。

在 2% 的琼脂糖凝胶上电泳 40 μ l 扩增产物，并使用 0.125 μ g/ml 溴化乙锭染色。切下约 400 个碱基对的单带，并使用 QIAGEN™ II 凝胶提取试剂盒，根据厂商的说明纯化此带。

将经纯化的 PCR 产物连接到 pGEM™-t 克隆载体上，使用载体转化感受态细菌 (大肠杆菌 JM-109)。挑选并培养含有经扩增之 DNA 的细菌克隆。裂解细菌，并接着使用酚-氯仿通过乙醇沉淀来提取 DNA。使用含有正确大小的插入物的菌落得到 DNA 以测序。根据厂商的说明，借助于 SEQUENASE™ 7-脱氮 dGTP 试剂盒，使用载体-特异性寡核苷酸 (正向和反向引物) 从两端测定克隆插入物的序列。通过将得自一个 KSHV 糖蛋白 B 扩增产物的 5 个克隆的序列资料联合可得到新片断的共有序列。

引物杂交区域之间的片段长度为 319 个碱基对。核苷酸序列列为 SEQ. ID NO: 3 并示于图 1。被编码的多肽序列列为 SEQ. ID NO: 4。

图 13 比较了糖蛋白 B 基因片断的序列与其它 γ 疱疹病毒的相应序列。单点 (.) 表示其它 γ 疱疹病毒中与 KSHV 序列相同的残基。短横 (-) 表示加入缺口以提供被编码的蛋白质的最佳排列的位置。在 KSHV 序列和任何其它已列出的序列之间相同的最长一段连串核苷酸是 14 个。短的保守序列散布于整个片断。总的来说，KSHV 和两个最接近的疱疹病毒 sHV1 和 bHV4 的序列之间的多核苷酸片断有 63% 相同。



使用序列资料设计长度为 20 - 40 个碱基对的 Type3 寡核苷酸引物，引物被设计成优先与 KSHV 糖蛋白 B 多核苷酸杂交，而不与其它已测序的编码糖蛋白 B 的多核苷酸杂交。这种引物的例子已列入上文的表 7。

图 14 比较了由相同的糖蛋白 B 基因片断编码的推测的氨基酸序列。

5 在氨基酸水平上，KSHV 和先前已知的 γ 疱疹病毒 bHV4 共用两个短区段，第一个 (SEQ. ID NO: 64) 长度为 13 个氨基酸，位于片断的 N-末端附近，第二个 (SEQ. ID NO: 65) 长度为 15 个氨基酸，位于片断的 C-末端附近。KSHV 和其它 γ 疱疹病毒共用的所有其它区段为 9 个氨基酸或更短。

10 实施例 5: RFHV 糖蛋白 B 的 386 个碱基之片段的序列

组织样本得自华盛顿大学地方灵长类动物研究中心的豚尾猴肿瘤。样本被固定于低聚甲醛中并包埋于石蜡中，根据实施例 2 中的方法从样本中提取 DNA。

15 通过使用与 DNA 聚合酶基因杂交的寡核苷酸引物进行 PCR 扩增反应来测定 DNA 制品中 RFHV 多核苷酸的存在。此方法的细节见共同拥有的美国专利申请 60/001, 148。收集以这种方式测定的含有 RFHV 多核苷酸的 DNA 提取物以用于本研究。

20 含有 RFHV 多核苷酸的 DNA 制品用作 PCR 扩增反应中的模板，所述扩增反应的第一轮使用糖蛋白 B 共有 - 简并寡核苷酸 FRFDA 和 TVNCB，第二轮扩增使用寡核苷酸 NIVPA 和 TVNCB。条件基本上与实施例 3 中的相同，不同的是用最初来源的 DNA 量和所用的条件只有 WB4 缓冲液产生很亮的带。与上文相同，用 ^{32}P 末端标记的 NIVPA 标记经扩增的 DNA；在 6% 的聚丙烯酰胺凝胶上电泳产物，得到放射性自显影。观察到对应于约 386 个碱基对和约 10 个较高分子量的多联体的带的阶梯。从凝胶上切下 386 个碱基对的带 (与同时电泳的 KSHV 片断的迁移率相同) 并提取之。

25 为了测定此提取物中的 DNA 是否得自特殊的扩增反应，仅使用 NIVPASQ，仅使用 TVNCBSQ 或同时使用这两个引物建立 PCR。缓冲液的条件与最初的扩增反应相同，混合物被加热至 80 $^{\circ}\text{C}$ ，加入 Taq 聚合酶，使用温度逐步降低的方法进行 35 轮扩增循环。理论上说，当使用一个引物时，特异性扩增反应按线性方式积累产物，当使用方向相反的两个引物时，则按指数方式积累产物。因此，通过使用两个引物的反应中产



物更多可以显示特异性，而在所有三个混合物中产物相等则暗示了非-特异性扩增。在被溴化乙锭染色的琼脂糖凝胶上分析这些检测反应中得到的扩增产物，RF 提取物未显示出有 NIVPASQ 反应的产物，它在适当的迁移率处，有 TVNCBSQ 反应的中度染色的带，对于同时使用两个引物的反应而言有强染色的带。对于平行检测的 KSHV 片断而言，有 NIVPASQ 反应的微弱的带，没有 TVNCBSQ 反应的带，对于同时使用两个引物的反应而言有强染色的带。我们推断出 RF 提取物中的 386 个碱基对的带表示特异性的扩增产物。

因此，在制备性的 2% 琼脂糖凝胶上电泳 40 μ l 已用这两种引物扩增的 RF 提取物，切下 ~ 386 个碱基对的带，使用 QIAGEN™ 试剂盒除去琼脂糖，如实施例 4 中所述将产物克隆到 E. coli 中并测定产物的序列。测定得自相同的经扩增的 RFHV 产物的 3 个不同克隆的共有序列。

将 RFHV 糖蛋白 B 片断的多核苷酸序列 (SEQ. ID NO: 1) 与 KSHV 的相应序列一起排列于图 1 中，还显示了被编码的 RFHV 氨基酸序列 (SEQ. ID NO: 2)。在引物杂交区域之间 (核苷酸 36 - 354)，多核苷酸序列有 76% 相同；氨基酸序列有 91% 相同。在这两种病毒之间，内在的半胱氨酸残基和潜在的 N-糖基化位点都是保守的。

使用序列资料设计长度为 20 - 40 个碱基对的 Type3 寡核苷酸引物，引物被设计成优先与 RFHV 糖蛋白 B 多核苷酸杂交，而不与其它已测定的编码糖蛋白 B 的多核苷酸杂交。这种引物的例子已列入上文的表 7。

图 15 比较了由糖蛋白 B 基因片断的核苷酸 36 - 354 编码的推测的氨基酸序列。至于 KSHV 序列，RFHV 和先前已知的 γ 疱疹病毒 bHV4 共用两个短区段。RFHV 和其它 γ 疱疹病毒共用的所有其它区段的长度小于 9 个氨基酸。

图 16 是可得到序列资料的疱疹病毒范围中相同糖蛋白 B 片断的序列资料排列。图 17 根据图 16 所示的糖蛋白 B 部分氨基酸序列的同一性程度，显示了疱疹病毒之间系统发育的关系。根据所示病毒中氨基酸的同源性，RFHV 和 KSHV 与 bHV4, eHV2 和 sHV1 最密切相关。

实施例 6: RFHV/KSHV 亚族的寡核苷酸引物和探针

以针对 RFHV 和 KSHV 得到的多核苷酸片断为基础，设计了 7 个 Type2 寡核苷酸，所述寡核苷酸可用作 RFHV/KSHV 亚族成员的 PCR 引物或杂交探针。



图 17 显示了四个共有 - 简并 Type2 寡核苷酸, 即 SHMDA, CFSSB, ENTFA 和 DNIQB, 排在一起的是衍生它们的序列。与实施例 1 中的寡核苷酸类似, 在 5' 末端具有共有区段, 在 3' 末端具有简并区段。然而, 这些寡核苷酸仅以 RFHV 和 KSHV 的序列为基础, 因此优先与 RFHV/KSHV 亚族的糖蛋白 B 形成稳定的双螺旋。上文表 6 显示了一系列 Type2 寡核苷酸的例子。

不同的 Type2 寡核苷酸具有有义或无义方向。在 PCR 扩增中可同时使用具有相反方向的引物, 或者, Type2 寡核苷酸可与具有相反方向的 Type1 寡核苷酸联合使用。

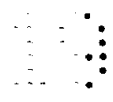
10 实施例 7: 上游和下游的糖蛋白 B 序列

进行进一步的扩增反应以得到另外的序列资料。KSHV DNA 的来源是根据实施例 2 制备的或者是冷冻的组织块或者是石蜡 - 包埋组织的卡波济氏肉瘤组织, 或由具有 KSHV 病原学的癌症, 如体腔淋巴瘤形成的细胞系。经培养增殖的 KSHV 也适合 (Weiss 等人)。

15 得到编码链 5' 方向上进一步的序列资料的一般策略是将与片断上游杂交的共有 - 简并 (Type1) 寡核苷酸作为 5' 引物, 并与作为 3' 引物的最接近的病毒 - 特异性 (Type3) 寡核苷酸联合使用以进行扩增反应。因此, 例如使用 FRFDA 和 TNKYB 作为第一套引物进行第一系列的扩增循环, 任选接着使用例如 FRFDA 和 GLTEB 作为第二套引物进行第二系列的
20 扩增循环。

所用的条件类似于实施例 3 和 4 中所述的那些。使用实施例 3 中所述的温度逐步降低法在 WB4 缓冲液中进行反应, 两轮扩增后, 使用经 γ ³²P - ATP 末端标记的最后使用的病毒 - 特异性寡核苷酸 (此时为 GLTEB) 标记产物。在 6% 聚丙烯酰胺上电泳经标记的产物, 切下对应于通过与其它疱疹病毒的类比预测的适当大小的带。重新扩增之后, 如前文所述
25 纯化, 克隆和测序产物, 通过联合约 3 个测定的结果得到了新片断的共有序列。

为了得到编码链 3' 方向上进一步的序列资料, 可将与片断下游杂交的共有 - 简并 (Type1) 寡核苷酸作为 3' 引物, 并与作为 5' 引物的最接近的病毒 - 特异性 (Type3) 寡核苷酸联合使用以进行扩增反应。在一个实施例中, 使用 NVFDB 和 TVFLA 进行第一系列的扩增循环, 任选接着使用 NVFDB 和 SQPVA 进行第二系列的扩增循环。如前文所述进行扩增
30



和测序。使用新序列设计具有有义方向的另一个 Type3 寡核苷酸，它与其它能与下游杂交的 Type1 寡核苷酸（如 FREYB 和 NVFDB）一起使用能得到进一步的序列资料。或者，使用具有相反方向的 Type1 寡核苷酸能得到 3' 方向上进一步的序列资料：例如，两个引物选自 FRFDA, NIVPA, TVNCA, NIDFB, NVFDB 和 FREYB；可选择另外的引物以进行嵌套扩增。

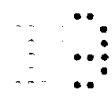
为了得到最下游的寡核苷酸引物的 3' 端的序列资料，可将如 CYSRA 的 Type1 引物，或如 TVFLA 的 Type3 引物与同 DNA 聚合酶基因 5' 末端杂交的引物联合使用。在共同拥有的美国专利申请 60/001,148 中描述了能与疱疹病毒的 DNA 聚合酶基因杂交的寡核苷酸引物，所述疱疹病毒与 RFHV 和 KSHV 相关。DNA 聚合酶编码区位于糖蛋白 B 编码区的 3' 端。预期使用这种联合引物进行的 PCR 扩增出的多核苷酸含有糖蛋白 B 编码区的 3' 末端，可能有的任何插入序列，和 DNA 聚合酶编码区的 5' 末端。

此策略按下述实施：

从被称为 RiGr 的冷冻的卡波济氏肉瘤样品，和被称为 BC-1 的从体腔淋巴瘤中衍生的细胞系中制备含有 KSHV 的糖蛋白 B 编码序列的 DNA。

为了得到整个 5' 序列，设计了怀疑是糖蛋白 B 上游的编码序列即衣壳成熟基因 (CAPMAT) 的 Type1 寡核苷酸探针。使用其它病毒的 CAPMAT 的已知序列鉴定相关的保守区，并设计了被称为 FENSA 的共有-简并引物以与糖蛋白 B 有义方向上的 CAPMAT 杂交。设计了怀疑是糖蛋白 B 下游的编码序列即 DNA 聚合酶的 Type 1 寡核苷酸探针。这些寡核苷酸列于表 9：

表 9: 检测, 扩增或鉴定疱疹病毒多核苷酸所用的另外的 Type1 寡核苷酸					
靶: 疱疹病毒, 尤其是 γ 疱疹病毒的衣壳/成熟基因					
名称	序列 (5' - 3')	长度	类型数目	方向	SEQ. ID:
FENSAC	GCCTTTGAGAATTCYAARTAYATHAAR	27	48	有义	77
FENSAG	GGGTTTGAGAATTCYAARTAYATHAAR	27	48	有义	78
靶: 疱疹病毒, 尤其是 γ 疱疹病毒的 DNA 聚合酶基因					
名称	序列 (5' - 3')	长度	类型数目	方向	SEQ. ID:
CVNVB	TAAAAGTACAGCTCCTGCCCGAANACRTTNACRCA	35	64	反义	79



使用覆盖整个糖蛋白 B 编码区的有义和反义引物对进行扩增，得到的片断包括表 10 所列的那些。

表 10: 所得的 KSHV 糖蛋白 B 片断

	片段	长度	位置
1	NIVPA→TVNCB	0.39kb	原始的片断
2	FENSA→VNVNB	0.9kb	横跨至 CAPMAT 的片断 1 的 5' 端
3	TVNCA→FREYB	2.3kb	片断 1 的 3' 端
4	FAYDA→FREYB	0.65kb	片断 1 的 3' 端
5	SQPVA→HVLQB	2.5kb	横跨至 DNA 聚合酶的片断 1 的 3' 端
6	FREYA→SCGFB	1.1kb	横跨至 DNA 聚合酶的片断 2 的 3' 端

5

用于扩增和测序的方法如下所述：使用具有引物对（如 FREYA 和 SCGFB）的 DNA 模板进行 PCR 扩增，以 94℃ 45 秒，60℃ 45 秒和 72℃ 45 秒进行 35 次循环；然后接着在 72℃ 最后延伸 10 分钟。使用 Quiagen 的 QIAQUICK™ PCR 纯化试剂盒在琼脂糖凝胶上纯化推测长度的 PCR 产物。在第二轮扩增中重新扩增经纯化的 PCR 产物，或者以嵌套或非嵌套的形式进行第二轮扩增。在给出的实施例中，使用 FREYA 和 SCGFB，或使用 FREYA 和 HVLQB 进行第二轮扩增。以 94℃ 45 秒，65℃ 45 秒和 72℃ 45 秒进行 35 次扩增循环；然后接着在 72℃ 最后延伸 60 分钟。

10

将 PCR 产物连接到 Novagen PT7 BLUE™ 载体上，并转化到 Novablue 感受态的 E. coli 中，使用 Novagen 方案进行连接和转化。通过使用 M13 正向和反向寡核苷酸的 PCR 筛选菌落。使用 Quiaquick 质粒分离试剂盒，从已在 37℃ 下，5ml 的 LB 肉汤中培养过夜的 PCR 阳性菌落中分离质粒。按照 USB Sequenase Kit 方案（USB）使用 M13 正向和反向测序引物对质粒手工测序。通过 ABI 方法进行自动测序。

15

当 KSHV 序列出现时，设计了另外的 KSHV - 特异性的 Type3 寡核苷酸。Type3 寡核苷酸可用于多对联合或与 Type1 寡核苷酸一起使用以 PCR 扩增，克隆和测序 KSHV DNA 片断。所用的 Type3 寡核苷酸列于表 11:

20

表 11: 检测, 扩增或鉴定疱疹病毒编码糖蛋白 B 的多核苷酸所用的另外的 Type3 寡核苷酸

靶: KSHV 的糖蛋白 B					
名称	序列 (5' - 3')	长度	类型 数目	方向	SEQ ID:
GAYTA	TGTGGAAACGGGAGCGTACAC	21	1	有义	80
DTYSB	TCAGACAAGAGTACGTGTCGG	21	1	反义	81
AIYGB	TACAGGTCGACCGTAGATGGC	21	1	反义	82
VTECA	CGCCATTTCCGTGACCGAGTG	21	1	有义	83
CEHYB	TGATGAAGTAGTGTTCGCAGG	21	1	反义	84
DLGGB	GATGCCACCCAGGTCCGCCAC	21	1	反义	85
DLGGA	GTGGCGGACCTGGGTGGCATC	21	1	有义	86
RAPPA	CGTAGATCGCAGGGCACCTCC	21	1	有义	87
靶: KSHV 的 DNA 聚合酶					
名称	序列 (5' - 3')	长度	类型 数目	方向	SEQ ID:
GEVFB	GTCTCTCCCGGAATACTTCT	21	1	反义	88
HVLQB	GAGGGCCTGCTGGAGGACGTG	21	1	反义	89
SCGFB	CGGTGGAGAAGCCGCAGGATG	21	1	反义	90

图 18 显示了与 KSHV DNA 杂交的寡核苷酸的位置图。所用缩写如下所述: d 或 h = 与疱疹病毒序列杂交的共有 - 简并探针 (Type 1); sq = 5 可使用的另外的测序尾; g = 与 γ 疱疹病毒杂交的探针 (Type 1); f = 与 RFHV/KSHV 亚族疱疹病毒杂交的探针 (Type 2); ks = KSHV 特异性的探针 (Type 3)。

图 19 列出了通过汇编每个已鉴定的片断的序列资料得到的共有序列, 显示了多核苷酸序列 (SEQ. ID NO: 91)。本发明的实施方案中包括在 DNA 聚合酶编码序列之前掺入区域的核苷酸 1 - 3056 (SEQ. ID NO: 10 92), 此共有序列表示得自卡波济氏肉瘤样品 RiGr 和淋巴瘤细胞系 BC - 1 的共有序列资料, 对于上述每个样品和每个基因区段而言, 都有多个克隆已被测序。在每个位置处进行约 3 - 9 次测定。

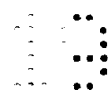


图 19 中还显示了三个开放阅读框 (SEQ. ID NOS: 93-95) 的氨基酸翻译, 被编码的 CAPMAT 蛋白质片断 (SEQ. ID NO: 93) 在不同时相上与糖蛋白 B 编码序列 (SEQ. ID NO: 94) 的 5' 末端重叠。在更加上游的位置处, CAPMAT 编码序列由于与 Epstein Barr 病毒的 RNA 聚合酶 2 的结合位点具有同源性, 因而也被怀疑含有糖蛋白 B 转录的调控元件。图中的下划线部分是这种推断的启动子区域。在糖蛋白 B 编码序列的 3' 末端, 有一个包括聚腺苷酸信号的非翻译序列, 其更下游的位置处是 DNA 聚合酶片断的编码序列 (SEQ. ID NO: 95)。

当将糖蛋白 B 编码序列与 GenBank 中的其它序列相比较时, 发现仅与其它疱疹病毒的糖蛋白 B 序列具有同源性。几个疱疹病毒只共用 20 个或更少的核苷酸的偶然序列。与 eHV2 共用序列 ATGTTTCAGGGAGTACA ACTACTACAC (SEQ. ID NO: 98)。除了此序列以外, 与其它先前已知的序列相比, KSHV 编码区域的 21 个或更多个核苷酸的区段显然是独特的。

在糖蛋白 B 编码序列内, 在使用卡波济氏肉瘤样品得到的序列资料和使用体腔淋巴瘤细胞系得到的序列资料之间标明了多核苷酸水平上的四个等位基因变体, 它们在图中由箭头表示。除了一个以外, 所有的变体都是沉默的。第四个变体导致了基因产物中脯氨酸变成亮氨酸的差异。

由 KSHV 糖蛋白 B 基因编码的蛋白质产物具有下列特征: 在 N-末端有一个对应于其它疱疹病毒糖蛋白 B 的信号肽区域 (“前导序列”) 的区域。图 3 提供了完整的 KSHV 糖蛋白 B 氨基酸序列和其它疱疹病毒已知的相应序列, 并显示了同源性区域。疱疹病毒糖蛋白 B 序列中高度保守的残基以星号 (*) 标明。在其它疱疹病毒糖蛋白 B 序列中保守的半胱氨酸残基也出现在 KSHV 的相应序列中。另外, 有两个可形成额外的内部二硫键并稳定三维结构的额外的半胱氨酸 (以 “·” 表示)。KSHV 糖蛋白 B 序列也具有对应于其它疱疹病毒糖蛋白 B 跨膜区的推测的跨膜区。

通过类似的策略得到 RFHV 的整个糖蛋白 B 序列。RFHV DNA 的来源类似地制备自华盛顿大学地方灵长类动物研究中心的被感染的猴组织。按实施例 5 所述提取 DNA。

为了得到编码链 5' 方向上进一步的序列资料, 将与片断上游杂交的共有-简并 (Type1) 寡核苷酸作为 5' 引物, 并与作为 3' 引物的最接近的病毒-特异性 (Type3) 寡核苷酸联合使用以进行扩增。因此, 例

如使用 FRFDA 和 AAITB 作为第一套引物进行第一系列的扩增循环，接着使用相同引物，或使用 FRFDA 和 GMTEB 嵌套引物进行第二系列的扩增循环。扩增条件与所描述的 KSHV 的相类似。

5 为了得到编码链 3' 方向上进一步的序列资料，可将与片断下游杂交的共有 - 简并 (Type1) 寡核苷酸作为 3' 引物，并与作为 5' 引物的最接近的病毒 - 特异性 (Type3) 寡核苷酸联合使用以进行扩增。因此，使用 NVFDB 和 VEGLA 进行第一系列的扩增循环，接着使用 NVFDB 和 PVLVA 进行第二系列的扩增循环。如前文所述进行扩增和测序。使用新序列设计具有有义方向的另一个 Type3 寡核苷酸，它与其它能与下游杂交的
10 Type1 寡核苷酸 (即 FREYB 和 NVFDB) 一起使用能得到进一步的序列资料。

使用多核苷酸和氨基酸序列资料将 RFHV 和 KSHV 的糖蛋白 B 互相比较，并与其它疱疹病毒的糖蛋白 B 比较。如实施例 6 所述，可使用 RFHV 和 KSHV 序列设计另一个亚族 - 特异性的 Type2 寡核苷酸。

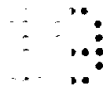
15 实施例 8: 得自 DNA 文库的糖蛋白 B 序列

通过从被感染的组织中建立 DNA 文库可得到或进一步证实整个糖蛋白 B 序列。用于此研究的 DNA 来源同实施例 7。

用蛋白酶 K 消化 DNA 裂解物，使用酚 - 氯仿提取 DNA。充分透析后，用 Sau3AI 限制性核酸内切酶部分消化制品。在蔗糖梯度中离心消化物，
20 回收约 10 - 23 千碱基的片段。通过用 BamHI 切割制备 λ DASH - 2™ 载体噬菌体 (Stratagene)。然后将选定大小的片段与载体混合，并使用 DNA 连接酶连接。

根据厂商的指导，用 Stratagene 的包装提取物制备经连接的载体，将所述载体用于感染 XL1 - BLUE™ MRA 细菌。在琼脂上以每个平板约
25 20,000 个细菌的密度涂布约 200,000 个被噬菌体感染的细菌。培养后，用硝酸纤维素覆盖平板上的菌落，将硝酸纤维素切成片段，从片段上洗脱噬菌体，使用适当的病毒特异性引物对它们的 DNA 进行扩增反应。在琼脂糖凝胶上电泳反应产物，并用溴化乙锭染色。从显示出所期望之大小的经扩增的 DNA 的平板区域回收噬菌体。将回收的噬菌体用于感染新的
30 XL1 细菌并重新涂布于新鲜培养基上，重复此过程直至得到有限稀释的单个克隆。

然后使用限制性核酸酶作出由此方法选择的每个克隆的图谱以确证



所掺入的片段大小。使用载体特异性引物由两个末端测定大于完整 DNA 聚合酶序列的插入物的序列。将此序列与完整的 EBV 基因组的已知多核苷酸序列相比较以确定此片段是否跨越了完整的糖蛋白 B 序列。从适当的克隆中得到 DNA，剪切，并通过根据标准技术的鸟枪克隆测定 DNA 的序列。

5 实施例 9: 糖蛋白 B 的抗原性区域

糖蛋白 B 基因中 NIVPA 和 TVNCB 杂交位点之间的多核苷酸片断具有图 14 所示的推测的氨基酸序列。根据这些序列，可鉴定出 RFHV 或 KSHV 独特的肽，或种之间共用的肽。

10 图 14 显示了长度为 6 或 7 个氨基酸的肽的例子。与相应的 γ 疱疹病毒肽相比，一些肽含有对 RFHV 或 KSHV 而言 (III 类) 或者对 RFHV/KSHV 亚族而言 (II 类) 为独特的一个或多个残基。

15 为了进一步证实糖蛋白 B 的这一 106 个氨基酸的区域内所含区域可以被抗体识别，可进行计算机分析以产生 Hopp 和 Woods 抗原性图。Hopp 和 Woods 测定法部分地基于连串氨基酸残基的相对亲水性和疏水性 (Hopp 等人)。结果示于图 20, 21 和 22。图标: \sim = 抗原性的; \wedge = 疏水性的; # = 潜在的 N-键糖基化位点。图 20 表示 RFHV 的 106 个氨基酸的糖蛋白 B 片断的分析; 图 21 表示 KSHV 片断的分析, 图 22 表示整个 KSHV 序列的分析。

20 RFHV 和 KSHV 都含有经预测可能是抗体靶位点的几个区域。尤其是 KSHV 序列在此片断的 N-末端附近和潜在的 N-键糖基化位点附近显示了抗原性区域。全长的 KSHV 序列显示出既对应于信号肽 (残基 ~ 25) 又对应于跨膜区 (残基 ~ 750) 的疏水性的小段序列。观察到评分 >1.0 或 >1.5 的推断的大量抗原性区域。特别值得注意的是在约残基 440 - 460
25 处出现了评分达到 ~ 2.5 的区域。

实施例 10: 病毒特异性的糖蛋白 B 扩增试验

在嵌套病毒特异性扩增反应中使用 Type3 寡核苷酸检测得自潜在感染的受试者的一系列组织样品中 RFHV 或 KSHV 的存在。

30 对于 KSHV 而言，从被怀疑携有此病毒的组织; 尤其是患有卡波济氏肉瘤损害和体腔 B-细胞淋巴瘤的人受试者的活检样品中提取 DNA。使用了大量不同的组织样品，包括得自 KS 损害，得自相同个体的明显未受感染的组织，得自没有明显的 KS 损害的 HIV 阳性个体和得自 HIV

阴性个体的一些样品。每种类型得到 5 个样品，按实施例 2 所述制备 DNA。

寡核苷酸引物 GLTEA, YELPA, VNVNB 和 ENTFB 订购自 Oligos Etc., Inc. 分两个阶段扩增 DNA，在第一阶段中使用引物 GLTEA 和 ENTFB，在第二阶段中使用引物 YELPA 和 VNVNB。扩增的条件与实施例 3 中的类似。在 2% 琼脂糖凝胶上电泳反应产物，用溴化乙锭染色并在紫外灯下检查。通过溴化乙锭染色检测出反应产物中高丰度的多核苷酸的存在表示阳性结果，这反映了样品中得自 KSHV 的 DNA 的存在；尤其是从 YELPA 至 VNVNB 的糖蛋白 B 编码片断的存在。将结果与病人的病史和样品的组织病理学相比较以确定阳性的试验结果是否与对 KS 的怀疑一致。

对于 RFHV 而言，从取自生长于华盛顿大学地方灵长类动物研究中心的灵长类群体的豚尾猴和食蟹猴的冷冻组织样品中提取 DNA。从具有明显的纤维瘤病症状的组织部位，相同猴的明显未受感染的部位，和没有显示症状的猴的相应部位各取 10 个样品。先使用 GMTEA 和 VEGLB，再使用 KYEIA 和 TDRDB 进行嵌套 PCR 扩增。如上文所述电泳和染色扩增产物以测定样品中是否存在 RFHV 多核苷酸。

实施例 11: 糖蛋白 B 的免疫原性区域

为了鉴定在 KSHV 感染的自然过程中会产生何种抗体，从 10-20 个具有卡波济氏肉瘤损害的 AIDS 受试者，10-20 个 HIV-阳性症状-阴性的受试者和 10-20 个 HIV-阴性的对照者中得到血清样品。在最初的研究中，收集每一群体的血清以分析抗体。

根据成熟的 KSHV 糖蛋白 B 分子的推测的整个细胞外区域合成 12 个残基长的肽。制备出的顺序肽覆盖了整个序列，并重叠了 8 个残基。通过标准的 F-Moc 化学法，根据厂商的指导使用得自 Genosys 的 SPOTS™ 试剂盒在尼龙膜支持物上制备肽。用血清覆盖制备出的膜，洗涤之并用与 β -半乳糖偶联的抗-人 IgG 覆盖此膜。通过加入底物 X-gal 显色此试验，阳性染色表示血清中有抗相应肽的 IgG 抗体的反应性。

类似地，为了鉴定 RFHV 感染的自然过程中形成的抗体，从 10 只豚尾猴和 10 只食蟹猴中收集血样，这些猴中的一部分显示了明显的纤维瘤病症状。通过如实施例 10 中所述，使用 RFHV-特异性寡核苷酸进行 PCR 扩增试验进一步证实正在发生的 RFHV 感染存在与否。通过离心分离血浆和血细胞，在类似于 KSHV 所用的方法中，使用这些血清检测抗体，不同之处在于根据 RFHV 糖蛋白 B 序列合成 12-聚体。



也要在动物模型中测试选定的 RFHV 和 KSHV 肽以测定它们与合乎需要的佐剂，如明矾和 DETOX™ 等联合使用时的免疫原性。适当的肽包括那些在上述实验中被鉴定出能在病毒感染的自然过程中产生抗体的肽。其它候选者包括那些据信能参与糖蛋白 B 的生物学功能的肽，和那些与已知能产生中和病毒之抗体的其它疱疹病毒肽相对应的肽。将此肽偶联到作为载体的匙孔血蓝蛋白 (KLH) 上，根据标准方法与佐剂联合，给兔肌肉注射于 1-2ml 接种物中的相当于 100 μg 的肽。4 周后，用第二剂量加强免疫动物，过 2 周后取血化验。

通过用免疫原或不相关的肽 - KLH 对照物包被来制备用于 ELISA 的微量滴定板孔。用得自试验血样的系列稀释血浆覆盖孔，洗涤，使用 β - 半乳糖抗 - 人 IgG 和 X-gal 显色。在试验孔而不是对照孔中的阳性染色表示肽在所用的条件下是免疫原性的。

实施例 12: RFHV/KSHV 亚族其它成员的糖蛋白 B 的鉴定和特征描述
冷冻保存怀疑含有上文未曾描述过的 γ 疱疹病毒，尤其是纤维增殖性疾病，淋巴细胞恶性肿瘤和与免疫缺损和免疫抑制相关之疾病，如急性呼吸道疾病综合症 (ARDS) 的组织样品，按实施例 2 所述提取 DNA。在第一轮中使用 Type1 寡核苷酸 FRFDA 和 TVNCB，然后在第二轮中使用嵌套的 Type1 或 Type2 寡核苷酸进行两轮 PCR 扩增。

任选地，通过用适当的探针探测扩增产物来证实 RFHV/KSHV 亚族糖蛋白 B 多核苷酸的存在。在琼脂糖上电泳经扩增的多核苷酸并印迹到尼龙膜上。将印迹与经 ³²P 标记的，含有得自编码糖蛋白 B 之 KSHV 多核苷酸的多核苷酸片段探针 (SEQ. ID NO: 3 的残基 36 - 354) 杂交。在允许探针和疱疹病毒糖蛋白 B 之间形成稳定的双螺旋，但不允许探针与 RFHV/KSHV 亚族以外来源的编码糖蛋白 B 的多核苷酸之间形成稳定双螺旋的条件下进行杂交反应。杂交条件需要探针和靶的杂交区段之间有大体 70% 的同一性以形成稳定的复合物。根据探针的长度和序列和靶的相应序列，使用上文给出的公式计算这些条件。估计条件如下：a) 允许探针在室温下，缺乏甲酰胺时，于 6 × SSC (0.15M NaCl, 15mM 柠檬酸钠缓冲液) 中与靶杂交；和 b) 在室温下用 2 × SSC 将新形成的双螺旋洗涤较短时间 (5 - 10 分钟)。

选择在这些条件下与经标记的探针杂交的经扩增的多核苷酸以进一步鉴定。或者，怀疑与从 KSHV 推测的大小大致相同的 PCR 扩增产物具

有相关序列。如果样品也已被用于得到包含疱疹病毒基因组的其它区域，如 DNA 聚合酶的多核苷酸，则可怀疑此样品具有相关序列。按实施例 4 中所述，横跨此片断测定含有由于大小差异或者来源不同而与 RFHV 或 KSHV 潜在不同之片断的样品的序列。通过与实施例 7 或 8 中类似的方法，使用具有新序列的这些片断测定整个糖蛋白 B 基因序列。

按下述得到 RFHV/KSHV 疱疹病毒亚族第三成员的糖蛋白 B 编码序列。

从两个冷冻的得自患有腹膜后纤维瘤病的猕猴的组织样品中提取 DNA，根据实施例 1 进行提取，在 40 μ g 糖原作为载体存在时，用乙醇沉淀提取的 DNA，用 70% 的乙醇洗涤，并重新悬浮于 10mM 的 Tris, pH8.0 缓冲液中。使用提取的 DNA 得到 151 个碱基对的疱疹病毒 DNA 聚合酶基因片断，此片断与 KSHV, RFHV 或任何其它先前已鉴定的 DNA 聚合酶的相应片断不相同。这导致了下列怀疑：样品含有不同疱疹病毒的基因组 DNA，可用于对新的糖蛋白 B 基因进行鉴定和特征描述。

使用半-嵌套的 PCR 从样品中扩增 386 个碱基对的糖蛋白 B 编码序列的片断。方法类似于实施例 4 和 5 中所用的方法，使用 FRFDA 和 TVNCB 进行第一轮扩增，接着使用 NIVPA 和 TVNCB 进行第二轮扩增，如上文所述测定最终的 PCR 产物的序列。

图 23 列出了多核苷酸序列 (SEQ. ID NO: 96) 以及相应的氨基酸翻译 (SEQ. ID NO: 97)。下划线的是两个引物杂交位点之间的 319 个碱基对的序列，此序列与 KSHV 和 RFHV 的相应序列不同。糖蛋白 B 得自疱疹病毒 RFHV/KSHV 亚族的被称为 RFHV2 的新成员。

参考文献

Altschul 等, (1986)。数学生物学通报, 48: 603-616。

Ambroziuk 等, (1995)。科学, 268: 582-583。

A. M. Eis-Hubinger 等, (1993)。普通病毒学杂志, 74: 379-385。

Baghian A. 等, (1993)。病毒学杂志, 67: 2396-2401。

Basco 等, (1992)。生物与化学杂志, 267: 19427-19434。

Basco 等, (1993)。染色体, 102: 32-38。

Beaucage 等, (1981)。四面体通讯, 22: 1859-1862。

Berel V. 等, (1990)。柳叶刀, 335: 123-128。

- Bernard 等, (1989)。细胞, 59: 219 - 228.
- Bernard 等, (1990)。美国国家科学院院刊, 87: 4610 - 4614.
- Boshoff 等, (1995)。天然药物学, 1: 1274 - 1278.
- Byrne K.M. 等, (1995)。病毒学, 290: 230 - 235.
- 5 Cantin E.M. 等, (1987)。美国国家科学院院刊, 84: 5908 - 5912.
- Cesarman E. 等, (1995)。新英格兰医学杂志, 332: 1186 - 1191.
- Chang Y. 等, (1994)。科学, 266: 1865 - 1869.
- Demotz S. 等, (1989)。免疫学方法杂志, 122: 67 - 72.
- Derbyshire 等, (1991)。EMBO J., 10: 17 - 24.
- 10 Digard P. 等, (1995)。美国国家科学院院刊, 92: 1456 - 1460.
- Dorsky D. I. 等, (1990)。病毒学杂志, 64: 1394 - 1397.
- Dorsky D. I. 等, (1988)。病毒学杂志, 62: 3224 - 3232.
- Dupin N. 等, (1995)。新英格兰医学杂志, 333: 798.
- Emery V. C. 等, (1992)。pp. 257 - 277 见, AIDS 机会感染的分子
15 和细胞生物学; S. Myint & A. Cann, eds, Chapman & Hall.
- Erickson 等, (1990)。科学 249: 527 - 533.
- Fields B. N. & Knipe D. M., eds. (1991)。基础病毒学, 第 2
版, Raven Press.
- Firesmith T.H. 等, (1994)。Int. J. Dermatol. 33: 755 - 762.
- 20 Gage P. J. 等, (1993)。病毒学杂志, 67: 2191 - 2201.
- Gao S-J. 等, (1996)。新英格兰医学杂志, 335: 233 - 241.
- Gibbs J. S. 等, (1988a)。美国国家科学院院刊, 85: 6672 - 6676.
- Gibbs J. S. 等, (1988b)。美国国家科学院院刊, 85: 7969 - 7973.
- Giddens W. E. Jr. 等, (1983)。pp. 249 - 253 见, 非人灵长类的
25 病毒和免疫性疾病; Alan R. Liss Inc.
- Glorioso J. C. 等, (1994)。Dev. Biol. Stand 82: 79 - 87.
- Haanes E. J. 等, (1994)。病毒学杂志, 68: 5825 - 5834.
- Haffey M. L. 等, (1988)。病毒学杂志, 62: 4493 - 4498.
- Hall J. D. 等, (1989)。核酸研究, 17: 9231 - 9244.
- 30 Hanke T. 等, (1991)。病毒学杂志, , 65: 1177 - 1186.
- Henikoff 等, (1992)。美国国家科学院院刊, 89: 10915 - 10919.
- Herold B. C. 等, (1994)。普通病毒学杂志, 75: 1211 - 1222.

- Hirose 等, (1978)。四面体通讯, (1978) 19: 2449 - 2452.
- Hodgson (1991)。生物/技术学, 9: 19 - 21.
- Horn 等, (1995)。人基因疗法, 6: 565 - 573.
- Hopp T.P. 等, (1981)。美国国家科学院院刊, 78: 3824 - 3828.
- 5 Johnson P.A. 等, (1994)。细胞生物学方法, 43A: 191 - 210.
- Karlin S. 等, (1994)。病毒学杂志, 68: 1886 - 1902.
- Kedes D.H. 等, (1996)。天然药物学, 2: 918 - 924.
- Knopf C.W. 等, (1988)。生物化学与生物物理学报, 951: 298 - 314.
- 10 Kostal M. 等, (1994)。Acta Virologica, 38: 77 - 88.
- Kumar 等, (1984)。有机化学杂志, 49: 4905 - 4912.
- Larder B.A. 等, (1987)。EMBO J. 6: 169 - 175.
- Latchman D.S. 等, (1994)。分子生物技术, 2: 179 - 195.
- Lin L.S. 等, (1995)。医学病毒学杂志, 45: 99 - 105.
- 15 Lisitsyn N. 等, (1993)。科学, 259: 946 - 。
- Liu M.Y. 等, (1989)。医学病毒学杂志, 28: 101 - 105.
- Liu Y.-N.C. 等, (1993)。普通病毒学杂志, 74: 2207 - 2214.
- Manservigi R. 等, (1990)。病毒学杂志, 64: 431 - 436.
- Marcy A.I. 等, (1990)。病毒学杂志, 64: 5883 - 5890.
- 20 Martin R.W. 等, (1993)。医学, 72: 245 - 26.
- McDermott M.R. 等, (1989)。病毒学, 169: 244 - 247.
- Meier J.L. 等, (1993), 病毒学杂志, 67: 7573 - 7581.
- Meinkoth J. 等, (1984)。分析生物化学, 138: 267 - 。
- Mester J.C. 等, (1990)。病毒学杂志, 64: 5277 - 5283.
- 25 Miles S.A. (1994)。当代肿瘤学论点, 6: 497 - 502.
- Miller G. (1996)。新英格兰医学杂志, 334: 1292 - 1297.
- Mitsuyasu R.T. (1993)。当代肿瘤学论点, 5: 835 - 844.
- Moore P.S. 等, (1995a)。新英格兰医学杂志, 332: 1181 - 1185.
- Moore P.S. 等, (1995b)。新英格兰医学杂志, 333: 798-799.
- 30 Moore P.S. 等, (1996)。病毒学杂志, 70: 549 - 558.
- Navarro D. 等, (1991)。病毒学, 184: 253 - 264.
- Navarro D. 等, (1992)。病毒学, 186: 99 - 112.

- Northfelt. D.W. (1994). 药学(新西兰), 48: 569-582.
- Nugent C.T.等, (1994). 病毒学杂志, 68: 7644-7648.
- O'Donnell C.A.等, (1991). 临床实验免疫学, 86: 30-36.
- O'Donnell M.E.等, (1987). 生物与化学杂志, 262: 4252-4259.
- 5 O'Leary J.J., (1996). 天然药物学, 2: 862-863.
- Padlan E.A., (1991). 分子免疫学, 28: 489-494.
- Pellett P.E.等, (1985). 病毒学杂志, 53: 243-253.
- Pereira L., (1994). 疾病传染因子, 3: 9-28.
- Qadri I.等, (1991). 病毒学, 180: 135-152.
- 10 Reardon J.E.等, (1989). 生物与化学杂志, 264: 7405-7411.
- Reschke M.等, (1995). 普通病毒学杂志, 76: 113-122.
- Sanchez-Pescador L.等, (1992). 传染病杂志, 166: 623-627.
- Schumacher T.N.等, (1992). 欧洲免疫学杂志, 22: 1405-1412.
- Shiu S.Y.W.等, (1994). Arch. Virol., 137: 133-138.
- 15 Simon等, (1991). EMBO J. 10: 2165-2171.
- Soengas等, (1992). EMBO J. 11: 4227-4237.
- Stow N.D. (1993). 核酸研究, 21: 87-92.
- Tsai C.-C.等, (1986). 实验动物科学, 36: 119-124.
- VanDevanter等, (1996). 临床微生物学杂志, 34: 1666-1671.
- 20 Wang T.S.-F.等, (1989). FASEB J. 3: 14-21.
- Ward P.L.等, (1994). 遗传学趋势, 10: 267-274.
- Weiss R.A.等, (1996). 天然药物学, 2: 277-278.
- Yeung K.C.等, (1991). 当代眼科研究, 10 (Suppl.) 31-37.
- Zhong W.等, (1996). 美国国家科学院院刊, 93: 6641-6646.
- 25 专利和专利申请:
- | | | |
|------------|-------------------|-------------|
| US 4762708 | Cohen 等 | (Gd 疫苗) |
| US 4415732 | Caruthers M. H. 等 | (多核苷酸合成) |
| US 4444887 | Hoffman M. K. | (mAb 方法) |
| US 4472500 | Milstein C. 等 | (mAb 细胞) |
| US 4642333 | Person S. | (HSV Gb 表达) |
| US 4683195 | Mullis K. B. | (PCR) |
| US 4683202 | Mullis K. B. 等 | (PCR) |

US 5124246	Urdea M. S. 等	(分支的 DNA)
US 5171568	Burke R. L. 等	(HSV Gb/Gd 疫苗)
US 5176995	Sninsky J. J. 等	(病毒的 PCR 方法)
US 5244792	Burke R. L. 等	(HSV Gb 表达)
US 5350671	Houghton M. 等	(HCV 诊断)
US 5354653	Matsumoto T. 等	(HSV 菌株探针检测)
US 5364773	Paoletti 等	(痘苗疫苗)
US 5384122	Cunningham 等	(疱疹病毒 L-颗粒疫苗)
US 5399346	Anderson W. F 等	(基因疗法)
US 5420026	Payne	(装配缺损颗粒)
WO 91/16420	Blum 等	(聚合酶突变)
WO 92/05263	Inglis 等	(减毒的疱疹病毒)
WO 92/16231	Francotte 等	(Gd/MPL-A 疫苗)
WO 94/11509	Couto 等	(人源化 ab)
EP 0239400	Winter	(人源化 ab)
EP 0290197	Mcaleer 等	(疱疹病毒活疫苗)
JP 5309000	Iatron Lab Inc	(EBV POL 的 PCR 检测)

美国专利申请 60/001, 148; 和 1996 年 7 月 11 日提交的部分继续申请[流水号待定; 代理人案号 299938 - 20001.00]: T. M. Rose, M. Bosch, K. Strand & G. Todaro. “与卡波济氏肉瘤和腹膜后纤维瘤病相关的 γ 疱疹病毒的 DNA 聚合酶”。

序列表

SEQ. ID	名称	描述	类型	来源
1	RFHV	糖蛋白 B 的 PCR 区段	dsDNA	图 1
2	RFHV	糖蛋白 B 的 PCR 区段	蛋白质	图 1
3	KSHV	糖蛋白 B 的 PCR 区段	dsDNA	图 1
4	KSHV	糖蛋白 B 的 PCR 区段	蛋白质	图 1
5	sHV1	糖蛋白 B 序列	dsDNA	GenBank HSVSPOLGBP
6	bHV4	糖蛋白 B 序列	dsDNA	GenBank BHT4GLYB
7	eHV2	糖蛋白 B 序列	dsDNA	GenBank EHVU20824
8	mHV68	糖蛋白 B 序列	dsDNA	GenBank MVU08990
9	hEBV	糖蛋白 B 序列	dsDNA	Genbank EBV
10	hCMV	糖蛋白 B 序列	dsDNA	Genbank HEHCMVGB
11	hHV6	糖蛋白 B 序列	dsDNA	Genbank HH6GBXA
12	hVZV	糖蛋白 B 序列	dsDNA	Genbank HEVZVXX
13	HSV1	糖蛋白 B 序列	dsDNA	Genbank HS1GLYB
14	sHV1	糖蛋白 B 序列	蛋白质	翻译
15	bHV4	糖蛋白 B 序列	蛋白质	翻译
16	eHV2	糖蛋白 B 序列	蛋白质	翻译
17	mHV68	糖蛋白 B 序列	蛋白质	翻译
18	hEBV	糖蛋白 B 序列	蛋白质	翻译
19	hCMV	糖蛋白 B 序列	蛋白质	翻译
20	hHV6	糖蛋白 B 序列	蛋白质	翻译
21	hVZV	糖蛋白 B 序列	蛋白质	翻译
22	HSV1	糖蛋白 B 序列	蛋白质	翻译
23	sHVS A8	糖蛋白 B 序列	蛋白质	翻译
24 - 40		Type1 寡核苷酸 (γ 疱疹病毒糖蛋白 B)	ssDNA (IUPAC)	表 4
41 - 47		Type2 寡核苷酸 (RFHV/KSHV 亚族糖蛋白 B)	ssDNA (IUPAC)	表 6
48 - 55		Type3 寡核苷酸 - RFHV 特异性糖蛋白 B	ssDNA	表 7
56 - 63		Type3 寡核苷酸 - KSHV 特异性糖蛋白 B	ssDNA	表 7
64 - 66		I 类抗原肽 (γ 疱疹病毒糖蛋白 B)	蛋白质	表 8
67 - 72		II 类抗原肽		

		(RFHV/KSHV 亚族糖蛋白 B)	蛋白质	表 8
73-74		III 类抗原肽 - RFHV 特异性糖蛋白 B	蛋白质	表 8
75-76		III 类抗原肽 - KSHV 特异性糖蛋白 B	蛋白质	表 8
77-78		Type1 寡核苷酸 (γ 疱疹病毒衣壳成熟基因)	SsDNA (IUPAC)	表 9
79		Type1 寡核苷酸 (γ 疱疹病毒 DNA 聚合酶)	ssDNA (IUPAC)	表 9
80-87		Type3 寡核苷酸 - KSHV 特异性糖蛋白 B		表 11
88-90		Type3 寡核苷酸 - KSHV 特异性 DNA 聚合酶		表 11
91	KSHV	含有衣壳成熟片断, 糖蛋白 B 和 DNA 聚合酶片断的编码区的 DNA 序列	DsDNA	图 18
92	KSHV	含有衣壳成熟片断和糖蛋白 B 编 码区的 DNA 序列(残基 1-3056)	DsDNA	实施例 7
93	KSHV	衣壳成熟序列	蛋白质	图 18
94	KSHV	糖蛋白 B 序列	蛋白质	图 18
95	KSHV	DNA 聚合酶序列	蛋白质	图 18
96	RFHV2	糖蛋白 B PCR 区段	dsDNA	图 22
97	RFHV2	糖蛋白 B PCR 区段	蛋白质	图 22
98		共用的序列	dsDNA	实施例 7
99-100		糖蛋白 B 的 I 类抗原肽	蛋白质	表 8

部分序列表

(1) 一般性资料:

(i) 申请人: Rose, Timothy M.

Bosch, Marnix

Strand, Kurt

(ii) 发明题目: RFHV/KSHV 亚族疱疹病毒的糖蛋白 B

(iii) 序列的数目: 40

(iv) 联系地址:

(A) 联系人: Morrison & Foerster

(B) 街道: 755 Page Mill Road

(C) 城市: Palo Alto

(D) 州: CA

(E) 国家: 美国

(F) 邮编: 94304 - 1018

(viii) 代理人/代理资料:

5

(A) 姓名: Schiff, J. Michael

(B) 登记号: 40,253

(C) 参考文献/概要号: 29938 - 20002.00

(ix) 电讯资料:

10

(A) 电话: (415) 813 - 5600

(B) 传真: (415) 494 - 0792

(C) 电传: 706141

SEQ ID NO:1:

	GTGTACAAGA AGAACATCGT GCCGTACATT TTCAAGGTAC GCAGGTACAT AAAAATAGCA	60
30	ACATCTGTCA CGGTCTACCG CGGTATGACA GAAGCAGCAA TCACAAACAA ATATGAGATC	120
	CCCAGGCCCG TGCCCTCTTA CGAGATCAGT CACATGGACA GCACCTACCA GTGCTTTAGT	180
	TCCATGAAAA TTGTAGTGAA CGGAGTCGAA AATACGTTCA CCGATCGGGA TGACGTAAAC	240
	AAAACCGTAT TTCTCCAGCC CGTCGAAGGT CTAAGTACA ACATACAAAG ATACTTTAGC	300
35	CAACCAGTAC TGTAAGTCTGA ACCCGGATGG TTCCAGGTA TCTACAGGGT TGGACAACA	360
	GTAACCTGTG AGATTGTAGA CATGTT	386

SEQ ID NO:2:

	Val Tyr Lys Lys Asn Ile Val Pro Tyr Ile Phe Lys Val Arg Arg Tyr
	Ile Lys Ile Ala Thr Ser Val Thr Val Tyr Arg Gly Met Thr Glu Ala
40	Ala Ile Thr Asn Lys Tyr Glu Ile Pro Arg Pro Val Pro Leu Tyr Glu
	Ile Ser His Met Asp Ser Thr Tyr Gln Cys Phe Ser Ser Met Lys Ile
	Val Val Asn Gly Val Glu Asn Thr Phe Thr Asp Arg Asp Asp Val Asn
	Lys Thr Val Phe Leu Gln Pro Val Glu Gly Leu Thr Asp Asn Ile Gln
	Arg Tyr Phe Ser Gln Pro Val Leu Tyr Ser Glu Pro Gly Trp Phe Pro
45	Gly Ile Tyr Arg Val Gly Thr Thr Val Asn Cys Glu Ile Val Asp Met

SEQ ID NO:3:

	GTGTACAAGA AGAACATCGT GCCGTATATT TTTAAGGTGC GGGCTATAG GAAAATTGCC	60
	ACCTCTGTCA CGGTCTACAG GGGCTTGACA GAGTCCGCCA TCACCAACAA GTATGAACTC	120
50	CCGAGACCCG TGCCACTCTA TGAGATAAGC CACATGGACA GCACCTATCA GTGCTTTAGT	180
	TCCATGAAGG TAAATGTCAA CGGGGTAGAA AACACATTTA CTGACAGAGA CGATGTTAAC	240
	ACCACAGTAT TCCTCCAACC AGTAGAGGGG CTTACGGATA ACATTCAAAG GACTTTAGC	300
	CAGCCGGTCA TCTACGGGA ACCCGGCTGG TTTCCGGCA TATACAGAGT TAGGACAACA	360
55	GTCAACTGTG AGATTGTAGA CATGTT	386

SEQ ID NO:4:

	Val Tyr Lys Lys Asn Ile Val Pro Tyr Ile Phe Lys Val Arg Arg Tyr
	Arg Lys Ile Ala Thr Ser Val Thr Val Tyr Arg Gly Leu Thr Glu Ser
60	Ala Ile Thr Asn Lys Tyr Glu Leu Pro Arg Pro Val Pro Leu Tyr Glu
	Ile Ser His Met Asp Ser Thr Tyr Gln Cys Phe Ser Ser Met Lys Val
	Asn Val Asn Gly Val Glu Asn Thr Phe Thr Asp Arg Asp Asp Val Asn
	Thr Thr Val Phe Leu Gln Pro Val Glu Gly Leu Thr Asp Asn Ile Gln

Arg Tyr Phe Ser Gln Pro Val Ile Tyr Ala Glu Pro Gly Trp Phe Pro
 Gly Ile Tyr Arg Val Arg Thr Thr Val Asn Cys Glu Ile Val Asp Met

SEQ ID NO:5:

5	ATGGTACCTA	ATAAACACTT	ACTGCTTATA	ATTTTGTCTG	TTTCTACTGC	ATGTGGACAA	60
	ACGACACCTA	CTACAGCTGT	TGAAAAAAT	AAAACCTCAAG	CTATATACCA	AGAGTATTTT	120
	AAATATCGTG	TATGTAGTGC	ATCAACTACT	GGAGAATTGT	TTAGATTTGA	TTTAGACAGA	180
	ACTTGTCCAA	GTAAGTGAAG	CAAAGTTCAT	AAGGAAGGCA	TTCTTTTGTG	GTACAAAAAA	240
	AATATAGTTC	CATATATCTT	TAAAGTCAGA	AGATACAAAA	AAATCACAAC	ATCAGTCCGT	300
10	ATTTTTAATG	GCTGGACTAG	AGAAGGTGTT	GCTATTACAA	ACAAATGGGA	ACTTTCTAGA	360
	GCTGTTCCAA	AATATGAGAT	AGATATTATG	GATAAGACTT	ACCAATGTCA	TAATTGCATG	420
	CAGATAGAAG	TAAACGGAAT	GTTAAATTCT	TACTATGACA	GAGATGGAAA	TAACAAAACT	480
	GTAGACTTAA	AGCCTGTAGA	TGGTCTAACG	GGTGCAATTA	CAAGATACAT	TAGCCAACCT	540
	AAAGTTTTTG	CTGATCTGG	CTGGCTATGG	GGAACCTACA	GSACTCGAAC	TACCGTTAAC	600
15	TGTGAAATTG	TAGACATGTT	TGCTAGGTCT	GCTGACCCTT	ACACATACTT	TGTGACTGCG	660
	CTTGGCGACA	CAGTAGAAGT	GTCTCCTTTC	TGTGATGTAG	ATAATTCATG	CCCAAATGCA	720
	ACTGACGTGT	TGTCAGTACA	AATAGACTTA	AATCACACTG	TTGTTGACTA	TGGAAATAGA	780
	GCTACATCAC	AGCAGCATAA	AAAAAGAATA	TTTGCTCATA	CTTTAGATTA	TTCTGTTTCT	840
	TGGGAAGCTG	TAAACAAATC	CGCGTCAGTA	TGCTCAATGG	TTTTTTGGAA	GAGTTTTTCAA	900
20	CGAGCTATCC	AAACTGAACA	TGACTTAACT	TATCATTTCA	TTGCTAATGA	AATAACAGCA	960
	GGATTCTCTA	CAGTGAAGA	ACCCTTAGCA	AATTTTACAA	GTGATTACAA	TTGTCTTATG	1020
	ACTCATATCA	ACACTACTTT	AGAGGATAAG	ATAGCAAGAG	TCAACAATAC	TCACACTCCA	1080
	AATGGTACAG	CAGAATATTA	TCAAACAGAA	GGTGGAAATG	TTTTAGTGTG	GCAGCCATTA	1140
	ATAGCAATAG	AATTAGAAGA	AGCAATGTTG	GAAGCAACTA	CATCTCCAGT	AACTCCTAGT	1200
25	GCACCAACTA	GCTCATCTAG	AAGTAAGCGA	GCAATAAGAA	GCATAAGAGA	TGTGAGTGCA	1260
	GGTTCAGAAA	ATAATGTGTT	TCTATCACAA	ATACAATATG	CATATGATAA	GCTACGTCAA	1320
	AGTATCAACA	ACGTGCTAGA	AGAGTTAGCT	ATAACATGGT	GTAGAGAACA	AGTGAGACAA	1380
	ACAAATGGTG	GGTATGAGAT	AGCAAAAAAT	AATCCAACAA	GTGTTATGAC	AGCAATATAT	1440
	GGAAAACTGT	TCTCTCGTAA	AGCTTTAGGA	GATGTAATCT	CTGTTACAGA	ATGTATAAAT	1500
30	GTTGACCAAT	CTAGTGTGAG	CATACACAAG	AGTCTTAAAA	CAGAAAATAA	TGACATATGC	1560
	TATTCACGGC	CTCCAGTTAC	ATTTAAATTT	GTTAACAGTA	GTCAGCTGTT	TAAAGGACAG	1620
	TTAGGGGCTA	GAAATGAAAT	TCTTCTGTCA	GAAAGTCTTG	TAGAAAATG	CCACCAAAT	1680
	GCAGAGACTT	TTTTACAGC	TAAAAATGAA	ACTTACCACT	TTAAAAATTA	TGTGCATGTA	1740
	GAAACTTTGC	CAGTGAATAA	CATTTCAACT	TTAGACACTT	TTTTAGCTCT	TAACCTAACT	1800
35	TTCATAGAAA	ATATTGACTT	TAAAGCTGTT	GAATTGTATT	CAAGTGGAGA	GAGAAAGTTA	1860
	GCAAACTGTG	TTGATTTAGA	GACTATGTTT	AGAGAATATA	ACTATTACGC	TCAGAGTATA	1920
	TCTGGCTTAA	GAAGAGATTT	TGATAACTCT	CAAAGAAAACA	ACAGAGACAG	AATCATTCAA	1980
	GATTTTTTAC	AAATTCTAGC	AGACTTAGGC	TCTATCGGCA	AAGTTATTGT	TAATGTGGCA	2040
	AGCGGCGCAT	TTTTCTTTTT	TGGAGGTATT	GTAACAGGCA	TATTAATTTT	TATTAATAAT	2100
40	CCTTTAGTGG	GCATGTTTAC	ATTTCTATTA	ATAGGAGCAG	TTATAATCTT	AGTAATTCTA	2160
	CTAGTACGGC	GCACAAATAA	TATGTCTCAA	GCTCCAATTA	GAATGATTTA	CCCAGATGTT	2220
	GAGAAATCTA	AATCTACTGT	GACGCTATG	GAGCCTGAAA	CAATTAACA	AATTTTGCTT	2280
	GGAAATGCATA	ACATGCAGCA	AGAAGCATAT	AAGAAAAAAG	AAGAACAAG	AGCTGCTAGA	2340
	CCGTCTATTT	TTAGACAAGC	TGCTGAGACA	TTTTTGCGTA	AGCGATCTGG	TTACAAAACAG	2400
45	ATTTCAACCG	AAGACAAAAT	AGTAT				2425

SEQ ID NO:6:

	ATGTATTATA	AGACTATCTT	ATTCTTCGCT	CTAATTAAGG	TATGCAAGTTT	CAACCAGACC	60
	ACTACACACT	CAACCACAAC	CTCACCAAGT	ATTTTCATCAA	CCACCTCTTC	CACAACAACA	120
50	TCAACAAGCA	AGCCATCAAA	CACAACCTCA	ACAAATAGTT	CATTAGCTGC	CTCTCCCAG	180
	AACACGTCAA	CAAGCAAGCC	ATCCACTGAT	AATCAGGGTA	CCAGTACCCC	CACTATTCCA	240
	ACTGTTACTG	ATGACACAGC	CAGTAAAAAT	TTTTATAAAT	ACAGAGTATG	CAGTGCATCA	300
	TCTTCTCTG	GAGAACTATT	CAGATTTGAC	CTTGATCAGA	CATGTCCAGA	TACAAAAGAT	360
	AAAAAACATG	TGGAAGGCAT	CCTGCTGGTA	CTAAAAAAGA	ATATTGTCCC	ATACATCTTC	420
55	AAAGTGAGGA	AATATAGAAA	AATTGCCACC	TCAGTGACAG	TTTACAGAGG	GTGGTCCCAG	480
	GCAGCTGTTA	CCAATAGGGA	TGATATCAGC	AGAGCCATAC	CCTATAATGA	AATTTCAATG	540
	ATAGATAGGA	CCTATCATTG	TTTCTCTGCT	ATGGCAACAG	TCATTAATGG	GATTCTGAAC	600
	ACCTATATAG	ACAGGGATTC	TGAAAAAAG	TCTGTTCCCC	TCCAGCCAGT	GGCCGGACTG	660
	ACTGAGAAC	TAAACAGATA	CTTTAGTCAA	CCTCTCATAT	ATGCAGAACC	TGGCTGGTTT	720
60	CCAGGGATTT	ATAGAGTGAG	AACAACCTGTT	AATTGTGAGG	TTGTTGACAT	GTATGCCCGC	780
	TCTGTGGAAC	CATATACTCA	CTTTATTACA	GCTCTGGGGG	ACACTATTGA	AATCTCCCCA	840
	TTCTGTGACA	ACAATTTCTCA	ATGCACCACT	GGTAATTCCA	CCTCAAGGGA	TGCCACAAG	900
	GTATGGATAG	AAGAAAATCA	CCAAACTGTT	GACTATGAAA	GACGGGGGCA	TCCCACTAAA	960

	GATAAAAGAA TCTTTCTAAA AGATGAGGAA TATACCATCT CCTGGAAAGC AGAAGATAGA	1020
	GAGAGAGCTA TTTGTGATTT TGTGATATGG AAAACCTTTC CCAGGGCCAT ACAAACAATC	1080
	CATAATGAGA GCTTTCACCTT TGTGGCAAAT GAAGTCACAG CCAGCTTTTT AACATCCAAC	1140
	CAAGAAGAAA CGGAGCTACG TGGAAATACC GAGATATTGA ATTGCATGAA TAGTACCATA	1200
5	AATGAAACTC TAGAAGAGAC AGTCAAAAAA TTTAACAAAT CCCATATCAG AGATGGGGAG	1260
	GTAAGTACT ATAAACAAA TGGGGGACTA TTCCTTATCT GGCAGGCAAT GAAACCCCTT	1320
	AATCTGTGAG AACACACAAA CTACACTATT GAAAGGAATA ACAAGACTGG AAATAAATCA	1380
	AGACAAAAAA GGTCTGTAGA TACAAAGACC TTCCAAGGCG CCAAGGGCCT GTCCACTGCC	1440
10	CAGGTTCAAT ATGCCTATGA CCATTTAAGA ACAAGCATGA ATCACATCCT AGAGGAATTA	1500
	ACCAAAACAT GGTGCCGGGA ACAA AAAAAG GACAATCTAA TGTGGTATGA GCTGAGTAAA	1560
	ATTAACCCAG TGAGTGTGAT GGCAGCCATT TATGGGAAAC CTGTGGCAGT GAAAGCCATG	1620
	GGAGATGCAT TCATGGTTTC TGAGTGCATC AATGTTGACC AGGCAAGTGT CAATATCCAT	1680
	AAAAGTATGA GAACGGATGA TCCCAAGGTA TGTTACTCCA GACCCTGGT CACATTTAAA	1740
	TTTGTGAATA GACTGCCAC CTTGAGGGT CAGCTTGGAA CAAGGAATGA AATCTTGCTC	1800
15	ACAAACACAC ACGTGGAAAC TTGTAGACCA ACAGCAGATC ATTATTTTTT TGTAAAGAAC	1860
	ATGACACACT ATTTTAAGGA CTATAAATTT GTGAAGACAA TGGATACCAA TAACATATCC	1920
	ACCCTGGATA CATTTTTAAC TCTCAATTTA ACTTTTATAG ACAATATAGA TTTCAAGACA	1980
	GTGGAACCTT ACAGTGAGAC TGAAAGAAAG ATGGCCAGTG CCCTCGACCT GGAGACGATG	2040
20	TTTAGAGAGT ATAATTAATA CACACAGAAG CTTGCAAGTC TGAGAGAAGA TCTAGACAAC	2100
	ACCATTGACC TGAACAGGGA CAGACTAGTT AAAGATCTCT CTGAAATGAT GGCAGACCTT	2160
	GGAGACATTG GAAAAGTGGT GGTCAACACA TTCAGTGGCA TTGTCAGTGT TTTTGGGTCT	2220
	ATAGTTGGTG GATTTGTGAG TTTTTCACA AACCCCATG GGGGGCTGAC GATCATCCTC	2280
	CTTCTCATAG TTGTGGTTTT TGTTGTTTTT ATAGTCTCCA GGAGAACCAA TAACATGAAC	2340
	GAGGCCCCCA TAAAATGAT CTATCCAAC ATTGACAAAG CCTCTGAGCA GGAGAACATT	2400
25	CAGCCCTAC CCGGAGAGGA GATTAAGCGC ATCCTCCTTG GAATGCACCA GCTCCAGCAA	2460
	AGTGAGCAGC GCAAATCTGA GGAAGAGGCT AGCCATAAAC CAGGGTTGTT CCAACTATTG	2520
	GGGGATGGCC TACAATTGCT GCGCAGGCGC GGGTATACTA GGTTACCAAC TTTTGACCCC	2580
	AGTCCAGGCA ATGACACATC TGAGACACAC CAAAAATATG TTT	2623
30	SEQ ID NO:7:	
	ATGGGGGTGCG GGGGCGGGCC TCGCGTCGTC CTCTGTCTAT GGTGCGTGC TGCGCTTCTC	60
	TGCCAGGGGG TGGCGCAAGA AGTTGTGGCT GAAACGACCA CCCCGTTGCG AACCCACAGA	120
	CCAGAAGTGG TGGCCGAGGA GAACCCGGCC AACCCCTTTC TGCCGTTGAG GGTATGCGGG	180
	GCCTCGCCTA CGGGCGGAGA GATATTCAGG TCCCCCTGG AGGAGAGCTG CCCC AACACG	240
35	GAAGACAAGG ACCACATAGA GGGCATAGCT CTCATCTACA AGACCAACAT AGTGCCTTAT	300
	GTTTTTAATG TCAGAAAGTA TAGGAAGATC ATGACCTCGA CCACCATCTA CAAGGGTTGG	360
	AGCGAGGATG CCATAACAAA CCAGCACAGC AGGAGCTACG CCGTCCCCT GTACGAGGTC	420
	CAGATGATGG ACCACTATTA TCAGTGCTTT AGCGCGTAC AGGTCAACGA GGGGGGCAC	480
40	GTCAACACCT ACTATGACAG GGACGGTGG AACGAGACCG CCTTCCTCAA ACCGGCCGAT	540
	GGTCTCACT CTAGCATAAC GCGTATCAG AGTCAACCAG AGGTGTACGC CACCCCCAGA	600
	AACCTGTTGT GGTCTTACAC AACAGAACC ACAGTCAACT GCGAGGTGAC AGAGATGTCT	660
	GCGAGATCCA TGA AACATT TGAGTTCTTT GTGACGCTG TTGGTGACAC TATAGAGATG	720
	TCGCCCTTTT TAAAAGAAA TGGCACAGAG CCAGAGAAAA TCTTGAAGAG ACCCACTCT	780
45	ATTC AACTGC TGA AAAACTA TGCTGT CACA AAGTACGGTG TGGGTTGGG GCAGGCTGAT	840
	AACGCTACCA GATTC TTG C AATATTTGG GACTATTCCC TGCTTTGGAA AGCCCACT	900
	GAAAACAGCT CCTACTGTGA TTTAATTTTA TGAAGGGGT TTTCCAATGC CATTCAAAT	960
	CAACACAATA GCAGTCTCCA TTTTATTGCC AATGATATAA CAGCCTCCT CTCTACTCCT	1020
	TTAGAAGAAG AGGCTAATTT TAACGAGACA TTTAAGTGTA TATGSAACAA CACCCAAGAA	1080
	GAAATTCAAA AAAAGTTAAA AGAGTTGAA AAAACTCACA GACCTAACGG TACTGCGAAG	1140
50	GTCTATAAAA CAACAGGCAA TCTGTACATT GTTTGGCAAC CGCTTATACA GATAGACCTG	1200
	CTAGATACTC ATGCCAAGCT GTACAATCTC ACAAACGCTA CAGCTTCACC TACATCAACA	1260
	CCCACAACAT CTCCCAGGAG AAGACGAGG GATACTTCAA GTGTTAGTGG CGGTGGAAAT	1320
	AATGGAGACA ACTCAACTAA GGAAGAGAGT GTGGCGGCT CCCAGGTTCA GTTTGCCTAT	1380
	GACAATCTCA GAAAGAGCAT CAACAGGGTG TTGGGAGAGC TGTCAGGGC ATGGTGCAGG	1440
55	GAACAGTACA GGGCCTCGCT CATGTGGTAC GAGCTGAGCA AGATCAACCC CACCAGCGTC	1500
	ATGAGCGCCA TCTATGGCAG GCCAGTGTCT GCCAAGTTGA TAGGGGACGT GGTGTGAGTG	1560
	TCAGATTGTA TCAGTGTGA CCAAAGAGC GTGTTTGTGC ACAA AATAT GAAGGTGCCT	1620
	GGCAAAGAAG ACCTGTGTTA CACCAGGCT GTGGTGGCT TCAAGTTTAT CAATGGGAGC	1680
	GAAGTGTG CTGGCCAGCT GGGTCCAGG AACGAGATTG TGCTGTCCAC CTCTCAGGTG	1740
60	GAGGTCTGCC AGCACAGCTG CGAGCACTAC TTCCAGGCG GGAACAGAT GTACAAGTAC	1800
	AAGGACTACT ACTATGTCAG TACCCTCAAC CTGACTGACA TACCACCT ACACACCATG	1860
	ATTACCCTGA ACCTGTCTCT GGTAGAGAAT ATAGATTTTA AGGTGATTGA GCTCTATTCT	1920
	AAAACAGAGA AAAGGCTGTC CAACGTGTTT GACATCGAGA CCATGTTGAG GGAGTACAAC	1980

	TACTACACTC	AGAACCTCAA	CGGGCTGAGG	AAGSACCTGG	ATGACAGCAT	AGATCATGGC	2040
	AGGGACAGCT	TCATCCAGAC	CCTGGGTGAC	ATCATGCAGG	ACCTGGGCAC	CATAGGCAAG	2100
	GTGTGGTCA	ATGTGGCCAG	CGGAGTGTTC	TCCCTCTTTG	GGAGCATAGT	CTCGGGGGTG	2160
	ATAAGCTTTT	TCAAAAAATCC	CTTTGGGGGG	ATGCTGTCTA	TAGTCTCTCAT	CATAGCCGGG	2220
5	GTAGTGGTGG	TGTACTCTGT	TATGACCAGG	TCCAGGAGCA	TATACTCTGC	CCCCATTAGA	2280
	ATGCTCTACC	CGGGGTGGA	GAGGGCGGCC	CAGGAGCCGG	GCGGCGACCC	GGTGTCTAGAA	2340
	GACCAAAATCA	GGAACATCCT	GATGGGAATG	CACCAATTTT	AGCAGCGGCA	GCGGGCGGAA	2400
	GAGGAGGCCC	GACGAGAGGA	AGAAGTAAAA	GGAAAAAGAA	CTCTCTTTGA	AGTGATAAGA	2460
	GACTCTGCGA	CCAGCGTTCT	GAGGAGGAGA	AGAGGGGGTG	GTGGGTACCA	GCGCCTACAG	2520
10	CGAGACGGGA	GCGACGATGA	GGGGGATTAT	GAGCCATTGA	GGCGACAAGA	TGGAGGCTAC	2580
	GACGACGTGG	ACGTGGAGGC	AGGCACGGCG	GATACCGGTG	TGTAA		2625

SEQ ID NO:8:

	ATGTACCCTA	CAGTGAAAA	TATGAGAGTC	GCCACCTAA	CCAATCTCCT	AACCCTTCTG	60
15	TGCTGTCTGT	GCCACACGCA	TCTCTACGTA	TGTCAGCCAA	CCACTCTGAG	GCAGCCATCA	120
	GACATGACCC	CAGCCCAGGA	CGCTCCANCA	GAGACTCCCC	CACCCCTCTC	AACTAACACT	180
	AACAGAGGAT	TTGAGTACTT	TCGCGTGTGT	GGGGTGGCTG	CCACGGGGGA	GACCTTCAGG	240
	TTTGATTAG	ACAAAACATG	CCCCAGTACA	CAAGATAAGA	AGCATGTGGA	GGGCATCTTG	300
	CTCGTGTATA	AGATCAACAT	CGTGCCCTAC	ATCTTCAAAA	TCAGGAGATA	TAGAAAAATA	360
20	ATTACTCAAC	TGACCATCTG	GCGAGGCCTA	ACCACTAGTT	CAGTCACTGG	TAAATTTGAA	420
	ATGGCCACTC	AGGCCACGGA	GTGGGAATG	GGCGACTTTG	ACAGCATCTA	TCAGTGTCTAC	480
	AATAGCGCCA	CCATGGTGGT	AAACAACGTC	AGACAGGTGT	ATGTGACAG	AGATGGGGTC	540
	AATAAACTG	TGAACATACG	CCCTGTTGAT	GGTCTAACAG	GGAAATATCA	AAGATACTTT	600
	AGTCAGCCCA	CCCTTTATTC	AGAACCTGGT	TGGATGCCTG	GCTTTTATCG	TGTTGGAACC	660
25	ACCGTTAACT	GTGAAATGTT	AGACATGGTG	GCACGCTCCA	TGGATCCCTA	TAACACATC	720
	GCTACCGCCC	TGGGAGACAG	CCTGGAGCTC	TCCCGTTTC	AAACCTTTGA	CAACACCAGC	780
	CAGTGTACTG	CGCCTAAGAG	AGCTGATATG	AGGGTCAGGG	AGGTCAAGAA	TTACAAGTTT	840
	GTAGATTATA	ATAACAGGGG	AACTGCCCCC	GCTGGACAAA	GCAGGACCTT	TCTAGAGACT	900
	CCCTCTGCCA	CTTACTCCTG	GAAAACAGCC	ACCAGACAAA	CTGCCACGTG	CGACCTGGTG	960
30	CACTGGA AAA	CATTCCCTCG	CGCCATCCAA	ACTGCTCATG	AACATAGCTA	CCATTTTGTG	1020
	GCCAATGAAG	TCACCGCCAC	CTTCAATACA	CCCCTGACTG	AGGTAGAAAA	TTTCACCAGC	1080
	ACGTATAGCT	GCGTCAGTGA	CCAGATCAAT	AAGACCATCT	CTGAATATAT	CCAAAAGTTG	1140
	AACAACTCC	ACGTGGCCAG	TGGGAAAACA	CAGTATTTCA	AGACTGATGG	TAACCTGTAC	1200
	CTCATCTGGC	AACCACCTCGA	ACATCCAGAG	ATTGAAGACA	TAGACGAGGA	CAGCGACCCA	1260
35	GAACCAACCC	CCGCCCCACC	AAAGTCCACA	AGGAGAAAAA	GAGAGGCAGC	TGACAATGGA	1320
	AACTCAACAT	CTGAGGTCTC	AAAGGGCTCA	GAAAATCCGC	TCATTACGGC	CCAAAATCAA	1380
	TTTGCTATG	ACAAGCTGAC	CACCAGGCTC	AACAACGTGC	TTGAGGAGTT	GTCCAGGGCG	1440
	TGGTGTAGAG	AACAGGTCAG	AGACACCCTC	ATGTGGTATG	AGCTTAGCAA	GGTCAACCCT	1500
	ACGAGTGTGA	TGCTCTGCCAT	TTATGGAAAG	CCTGTGCTG	CCAGGTACGT	GGGCGACGCC	1560
40	ATATCTGTGA	CAGACTGTAT	CTATGTGGAC	CAAAGTTCAG	TCAACATCCA	CCAGAGCTTG	1620
	CGGCTGCAGC	ATGATAAAAC	CACCTGTCTAC	TCGAGACCTA	GAGTCACTT	CAAATTTATA	1680
	AACAGTACAG	ACCGCTAAC	TGGCCAGTTG	GGTCTTAGAA	AAGAAATTAT	CCTCTCCAAC	1740
	ACAAACATAG	AAACATGCAA	GGATGAGAGT	GAACACTACT	TCATTGTGGG	GGAATACATT	1800
	TACTATTATA	AAAATTACAT	TTTTGAAGAA	AAGCTAAACC	TCTCAAGCAT	CGCTACCCTA	1860
45	GACACATTTA	TAGCCCTCAA	TATCTCATTT	ATTGAAAATA	TCGACTTCAA	AACAGTAGAA	1920
	CTGTACTCCT	CTACTGAAAG	GAAACTCGCA	TCGAGCGTCT	TTGATATAGA	ATCCATGTTT	1980
	AGGGAAATATA	ACTATTACAC	CTACAGCCTC	GCGGGCATT	AGAAGGACCT	AGACAACACC	2040
	ATCGACTACA	ATAGAGACAG	ACTGGTTCAG	GACCTGTCTG	ACATGATGGC	TGATCTGGGA	2100
	GACATTGGAA	GATCTGTGGT	GAATGTGGTC	AGCTCGGTAG	TCACATTTTT	CAGTAGTATT	2160
50	GTGACAGGGT	TCATTA AAT	CTTTACCAAC	CCTCTAGGGG	GAATATTCAT	TCTCCTAATT	2220
	ATTGTTGGAA	TAATCTTCTT	GGTGGTAGTC	CTAAATAGAA	GAAACTCACA	GTTTCACGAT	2280
	GCACCCATCA	AAATGCTGTA	CCCTTCTGTT	GAAAACACTG	CTGCCAGACA	GGCGCCACCT	2340
	CCCTATAGCG	CATCACCTCC	AGCTATAGAC	AAAGAGGAAA	TTAAGCGCAT	ACTTTTGGGC	2400
	ATGCATCAGG	TACACCAGGA	AGAAAAGGAA	GCACAGAAAC	AACTAACCAA	CTCTGGCCCT	2460
55	ACTTTGTGGC	AGAAAAGCCAC	AGGATTCCTT	AGAAATCGCC	GGAAGGGATA	CAGCCAACCT	2520
	CCTCTGGAAG	ATGAATCAAC	TTCCCTCT				2548

SEQ ID NO:9:

	ATGACTCGGC	GTAGGGTGTCT	AAGCGTGGTC	GTGCTGCTAG	CCGCCCTGGC	GTGCCGTCTC	60
60	GGTGGCAGCA	CCCCAGAGCA	GCCCGCACCC	CCC GCCACCA	CGGTGCAGCC	TACCGCCACG	120
	CGTCAGCAAA	CCAGCTTTCC	TTTCCGAGTC	TGCGAGCTCT	CCAGCCACGG	CGACCTGTTC	180
	CGCTTCTCCT	CGGACATCCA	GTGTCCCTCG	TTTGGCACGC	GGGAGAAATCA	CACGGAGGGC	240
	CTGTGTGATGG	TGTTTAAAGA	CAACATTATT	CCCTACTCGT	TTAAGGTCCG	CTCCTACACC	300

	AAGATAGTGA CCAACATTCT CATCTACAAT GGCTGGTACG CGGACTCCGT GACCAACCGG	360
	CACGAGGAGA AGTTCTCCGT TGACAGCTAC GAAACTGACC AGATGGATAC CATCTACCAG	420
	TGCTACAACG CGGTCAAGAT GACAAAAGAT GGGCTGACGC GCGTGTATGT AGACCGCGAC	480
	GGAGTTAACA TCACCGTCAA CCTAAAGCCC ACCGGGGGCC TGGCCAACGG GGTGGGCGC	540
5	TACGCCAGCC AGACGGAGCT CTATGACGCC CCCGGTGGT TGATATGGAC TTACAGAACA	600
	AGAACTACCG TCAACTGCCT GATAACTGAC ATGATGGCCA AGTCCAACAG CCCCTTCGAC	660
	TTCTTTGTGA CCACCACCGG GCAGACTGTG GAAATGTCC CTTTCTATGA CGGAAAAAT	720
	AAGGAAACCT TCCATGAGCG GGCAGACTCC TTCCACGTGA GAACTAACTA CAAGATAGTG	780
	GACTACGACA ACCGAGGGAC GAACCCGCAA GCGCAACGCC GAGCCTTCT GGACAAGGGC	840
10	ACTTACACGC TATCTTGAA GCTCGAGAAC AGGACAGCCT ACTGCCGCT TCAACTCTGG	900
	CAAACCTTTG CTAGCTTCGT GACCAACACA ACCGTGGGCA TAGAGCTCCC GGACGCCTC	960
	GAGGGCACCT ACTCGACCAT CGCCACAGAA ACAGGGAAGT CAATACATT TTGTACTGAC	1020
	AAGTGCATCG AAGAGCAGGT GAACAAGACC ATGCATGAGA AGTACGAGC CGTCCAGGAT	1080
	CGTTACAGCA AGGCCAGGA AGCCATTACA TATTTATAA CGAGCGGAGG ATTGTTATTA	1140
15	GCTTGGCTAC CTCTGACCCC GCGCTCGTTG GCCACCGTCA AGAACCTGAC GGAGCTTACC	1200
	ACTCCGACTT CCTCACCCCC CAGCAGTCCA TCGCCCCAG CCCCATCCGC GCCCGCGGG	1260
	AGCACCCCGG CCGCCGTTCT GAGGCGTCGG AGGCGGGATG CGGGGAACGC CACCACCCG	1320
	GTGCCCCCA CCGCCCCCGG GAAGTCCCTG GGCACCCTCA ACAATCCGC CACCGTCCAG	1380
	ATCCAATTTG CCTACGACTC CCTGCCCCG CAGATCAACC GCATGCTGGG AGACCTTGGC	1440
20	CGGGCCTGGT GCCTGGAGCA GAAGAGGCAG AACATGGTGC TGAGAGAACT AACCAAGATT	1500
	AATCCAACCA CCGTCATGTC CAGCATCTAC GGTAAGGGCG TGGCGGCCAA GCGCCTGGGG	1560
	GATGTCATCT CAGTCTCCCA GTGGGTGCC GTTAACCAGG CCACCGTCA CCTGCGCAAG	1620
	AGCATGAGG TCCCTGGCTC CGAGACCATG TGCTACTCGC GCCCCTGGT GTCTTCAGC	1680
	TTTATCAACG ACACCAAGAC CTACGAGGGA CAGCTGGGCA CCGACAACGA GATCTTCTC	1740
25	ACAAAAAGA TGACGGAGGT GTGCCAGCG ACCAGCCAGT ACTACTTCCA GTCCGGCAAC	1800
	GAGATCCAG TCTACAACGA CTACCACCAC TTTAAACCA TCGAGCTGGA CGGCATTGCC	1860
	ACCCTGCAGA CCTTCATCTC ACTAAACACC TCCCTCATG AGAACATTGA CTTTGCCTCC	1920
	CTGGAGCTGT ACTCACGGGA CGAACAGCGT GCCTCCAAG TCTTTGACCT GGAGGGCATC	1980
	TTCCGGGAGT ACAACTTCCA GCGCAAAAC ATCGCCGGCC TGGCGAAGGA TTTGGACAAT	2040
30	GCAAGTCAA ACGGAAGAAA TCAATTCGTG GACGGCCTGG GGGAACTTAT GGACAGTCTG	2100
	GGTAGCGTG GTACAGTCCAT CACCAACCTA GTCAGCACGG TGGGGGTTT GTTTAGCAGC	2160
	CTGGTCTCTG GTTTCATCTC CTCTTCAA AACCCTTCG GCGGCATGCT CATTCTGGTC	2220
	CTGGTGGCG GGGTGGTGAT CCTGGTTATT TCCCTCACGA GGGCACGCG CCAGATGTCG	2280
	CAGCAGCCGG TGCAGATGCT CTACCCGGG ATCGAGGAG TCGCTCAGCA ACATGCTCT	2340
35	GGTAGGGTC CAGGCATTA TCCATTAGT AAGACAGAA TACAAGCCAT CATGTTAGCG	2400
	CTGCATGAG AAAACCAAGGA GCAAAAGAGA GCAGCTCAGA GGGCGGCCG ACCCTCAGTG	2460
	GCCAGCAGAG CATTGCAGGC AGCCAGGGAC CGTTTTCCAG GCCTACGCAG AAGACGCTAT	2520
	CACGATCCAG AGACCGCCG CGCACTGCTT GGGGAGGCAG AGACTGAGTT TT	2572
40	SEQ ID NO:10:	
	ATGGAATCCA GGATCTGGTG CCTGGTAGTC TGCGTAAAC TGTGTATCGT CTGTCTGGGT	60
	GCTGCGGTTT CTTCTCTAG TACTTCCCAT GCAACTTCTT CTACTCACAA TGGGAAGCCAT	120
	ACTTCTCGTA CGACGTCTGC TCAAACCCGG TCAGTCTATT CTCAACAGT AACGTCTTCT	180
	GAAGCCGTCA GTCATAGAC CAACGAGACT ATCTACAACA CTACCCTCAA GTACGGAGAT	240
45	GTGGTGGGAG TCAACACTAC CAAGTACCCC TATCGCGTGT GTTCTATGGC CCAGGGTACG	300
	GATCTTATTC GCTTTGAACG TAATATCATC TGCACCTCGA TGAAGCCTAT CAATGAAGAC	360
	TTGGATGAGG GCATCATGGT GGTCTACAAG CGCAACATCG TGGCGCACAC CTTTAAGGTA	420
	CGGGTCTACC AAAAGTTTT GACGTTTTGT CGTAGCTACG CTTACATCTA CACCATTAT	480
	CTGCTGGGCA GCAATACGGA ATACGTGGCG CCTCCTATGT GGGAGATTCA TCACATCAAC	540
50	AAGTTTGCTC AATGTACAG TTCCTACAG CCGCTTATAG GAGGCACGGT TTTGTTGGCA	600
	TATCATAGGG ACAGTTATGA AAACAAAACC ATGCAATTA TTTCCGACGA TTATTCCAAC	660
	ACCCACAGTA CCCGTTACGT GACGGTCAAG GATCAGTGGC ACAGCCGCGG CAGCACCTGG	720
	CTCTATCGTG AGACCTGTAA TCTGAACTGT ATGCTGACCA TCACTACTGC GCGCTCCAAG	780
	TATCCTTATC ATTTTTTGC AACTTCCAG GGTGATGTGG TTTACATTC TCCTTTCTAC	840
55	AACGGAACCA ATCGCAATGC CAGCTACTTT GGAGAAAACG CCGACAAGTT TTTCATTTTC	900
	CCGAACTACA CCATCGTTTC CGACTTTGGA AGACCCAACG CTGCGCCAGA AACCCATAGG	960
	TTGGTGGCTT TTCTCGAACG TGCCGACTCG GTGATCTCT GGGATATACA GGACGAGAAG	1020
	AATGTACCT GCCAGCTCAC CTTCTGGGAA GCCTCGAAC GTACTATCCG TTCCGAAGCC	1080
	GAAGACTCGT ACCACTTTTC TTCTGCCAAA ATGACTGCAA CTTTTCTGTC TAAGAAACAA	1140
60	GAAGTGAACA TGTCCGACTC CCGCTGGAC TCGGTACGTG ATGAGGCTAT AAATAAGTTA	1200
	CAGCAGATTT TCAATACTTC ATACAATCAA ACATATGAAA AATACGGAAA CGTGTCCGTC	1260
	TTCGAAACCA GCGCGGTCT GGTGGTGTTC TGGCAAGGCA TCAAGCAAAA ATCTTTGGTG	1320
	GAATTGGAAC GTTTGGCCAA TCGATCCAGT CTGAATATCA CTCATAGGAC CAGAAGAAGT	1380

	ACGAGTGACA	ATAATACAAC	TCATTTGTCC	AGCATGGAAT	CGGTGCACAA	TCTGGTCTAC	1440
	GCCCAGCTGC	AGTTCACCTA	TGACACGTTG	CGGGTTACA	TCAACCGGGC	GCTGGCGCAA	1500
	ATCGCAGAAG	CCTGGTGTGT	GGATCAACGG	CGCACCCCTAG	AGGTCTTCAA	GGAACCTCAGC	1560
	AAGATCAACC	CGTCAGCCAT	TCTCTCGGCC	ATTTACAACA	AACCGATTGC	CGCGGTTTC	1620
5	ATGGGTGATG	TCTTGGCCCT	GGCCAGCTGC	GTGACCATCA	ACCAAACCAG	CGTCAAGGTG	1680
	CTGCGTGATA	TGAACGTGAA	GGAAATCGCCA	GGACGCTGCT	ACTCACGACC	CGTGGTCATC	1740
	TTTAATTTCC	CCAACAGCTC	GTACGTGCAG	TAOGGTCAAC	TGGCGGAGGA	CAACGAAATC	1800
	CTGTTGGGCA	ACCACCGCAC	TGAGGAATGT	CAGCTTCCCA	GCCTCAAGAT	CTTCATCGCC	1860
	GGGAACTCGG	CCTACGAGTA	CGTGGACTAC	CTCTTCAAAC	GCATGATTGA	CCTCAGCAGT	1920
10	ATCTCCACCG	TCGACAGCAT	GATCGCCCTG	GATATCGACC	CGCTGGAAAA	TACCGACTTC	1980
	AGGGTACTGG	AACTTTACTC	GCAGAAAGAG	CTGCGTTCCA	GCAAGTTTTT	TGACCTCGAA	2040
	GAGATCATGC	GCGAATTCOA	CTCGTACAAG	CAGCGGGTAA	AGTACGTGGA	GGACAAGGTA	2100
	GTGCAACCCG	TACCGCCCTA	CCTCAAGGTT	CTGGACGACC	TCATGAGCGG	CCTGGGCGCC	2160
	GCGGGAAGG	CCGTTGGCCT	AGCCATTGGG	GCGGTGGGTG	GCGCGGTGGC	CTCGTGGTC	2220
15	GAAGCGGTTG	CCACCTTCTC	CAAAAACCC	TTGGAGCCT	TCACCATCAT	CCTCGTGGCC	2280
	ATAGCGGTAG	TCATTATCAC	TTATTTGATC	TATACTCGAC	AGCGGCGTCT	GTGCACGCAG	2340
	CCGCTGCAGA	ACCTCTTTCC	CTATCTGGTG	TCGCCGACG	GGACCACCGT	GACGTCGGGC	2400
	AGCACCAAAG	ACACGTCGTT	ACAGGCTCCG	CCTTCTACG	AGGAAAGTGT	TTATAATTCT	2460
	GGTCGCAAAG	GACCGGGACC	ACCGTCGTCT	GATGCATCCA	CGGCGGCTCC	GCCTTACACC	2520
20	AACGAGCAGG	CTTACCAGAT	GCTTCTGGCC	CTGGCCGTC	TGGACGCAGA	GCAGCGAGCG	2580
	CAGCAGAACG	GTACAGATTC	TTTGACGGA	CAGACTGGCA	CGCAGGACAA	GGACAGAAG	2640
	CCTAACCTGC	TAGACCGGCT	GCGACATCGC	AAAAACGGCT	ACAGACACTT	GAAAGACTCC	2700
	GACGAAGAAG	AGAACGTCTG	AA				2722
25	SEQ ID NO:11:						
	ATGAGCAAGA	TGAGAGTATT	ATTCCTGGCT	GTCTTTTTGA	TGAATAGTGT	TTTAAATGATA	60
	TATTGCGATT	CGGATGATTA	TATCAGAGCG	GGCTATAATC	ACAAATATCC	TTTTCGGATT	120
	TGTTGCGATT	CCAAAGGCAC	TGATTTGATG	CGTTTCGACA	GAGATATTTT	GTGTTCCGCA	180
	TATAAGTCTA	ATGCAAAGAT	GTCCGAGGGT	TTTTTCATCA	TTTACAAAAC	AAATATCGAG	240
30	ACCTACACTT	TTCCAGTGAG	AACATATAAA	AACGAGCTGA	CGTTCCAAAC	CAGTTACCGT	300
	GATGTGGGTG	TGTTTTATTT	TCTGGATCGG	ACGGTGATGG	GTTTGGCCAT	GCCGGTGTAC	360
	GAAGCAAATT	TAGTTAATTC	TCGTGCGCAG	TGTTATTAG	CCGTAGCGAT	AAAACGACCC	420
	GATGGTACGG	TGTTTAGTGC	CTATCATGAG	GATAATAATA	AAAACGAAAC	TCTAGAATTA	480
	TTTCTCTGTA	ATTTCAAGTC	TGTTACTAAT	AAAAGATTTA	TCACTACGAA	AGAACCCTAC	540
35	TTTGCAAGGG	GTCCTTTGTG	GCTCTATTCT	ACATCGACGT	CTCTCAATTG	TATTGTGACG	600
	GAGGCTACGG	CTAAGGCGAA	ATATCGGTTT	AGTTACTTTG	CTTTGACGAC	TGGTGAATC	660
	GTGGAAGGGT	CTCCGTTCTT	CGACGGTTCA	AACGGTAAAC	ATTTTGCAGA	GCCGTTAGAA	720
	AAATTGACAA	TCCTTGAAAA	CTATACTATG	ATAGAAGATC	TAATGAATGG	TATGAATGGG	780
	GCTACTACGT	TAGTAAGGAA	GATCGCTTTT	CTGGAGAAAG	GGGACTTTT	GTTTTCTTGG	840
40	GAAATCAAGG	AAGAGATGA	ATCGGTGTGT	ATGCTAAAGC	ACTGGACTAC	GGTGACTCAC	900
	GGGCTTCGAG	CGGAGACGGA	TGAGACTTAT	CACTTTATTT	CTAAGGAGTT	GACAGCCGCT	960
	TTGTCGCCT	CCAAGGAGTC	TTTAAATCTT	ACCGATCCCA	AACAAACGTG	TATTAAGAAT	1020
	GAATTTGAGA	AGATAATTAC	AGATGTCTAT	ATGTCAGATT	ATAATGATGA	CTACAGCATG	1080
	AACGGTAGTT	ATCAAATTTT	TAAGACTACG	GGAGATCTGA	TTTTGATTTG	GCAGCCTCTT	1140
45	GTGCAAAAAT	CTCTTATGTT	TCTTGAGCAG	GGTTCAAGTAA	ACTTACGTAG	GAGGCGAGAT	1200
	TTGGTGGATG	TCAAGTCTAG	ACATGATATT	CTTTATGTGC	AATTACAGTA	CCTCTATTGAT	1260
	ACTTTGAAAG	ATTATATCAA	CGATGCCTTG	GGGAATTTGG	CAGAACTTTG	GTGCCTCGAT	1320
	CAAAAACGAA	CGATAACGAT	GTTGCACGAA	CTTAGTAAGA	TCAGTCCATC	GAGTATCGTG	1380
	TCTGAGGTTT	ACGGTCGTCC	GATATCTGCA	CAGTTGCATG	GTGATGTGTT	AGCTATCTCG	1440
50	AAATGCATAG	AAGTTAATCA	ATCATCCGTT	CAGCTTTATA	AGAGTATGCG	GGTCGTCGAT	1500
	GCGAAGGGAG	TAAGGAGTGA	AACGATGTGT	TATAATCGGC	CCTTGGTGAC	GTTTAGCTTT	1560
	GTGAACTCCA	CGCCTGAGGT	TGTCCTTGGT	CAGCTAGGGT	TAGATAATGA	GATTCTGTTG	1620
	GGTGATCATA	GGACAGAGGA	ATGTGAGATA	CCTAGTACAA	AGATATTTCT	ATCTGGAAAT	1680
	CATGCACACG	TGTATACCGA	TTATACGCAT	ACGAATTCGA	CGCCCATAGA	AGACATTGAG	1740
55	GTATTGGATG	CTTTTATTAG	ACTAAAGATC	GACCTCTCG	AAAATGCTGA	TTTTAAACTA	1800
	CTTGATTTAT	ATTCGCGCGA	CGAATTTGAGT	AGAGCAAACG	TTTTCGATTT	AGAGAATATT	1860
	CTTCGTGAAT	ATAACTCATA	TAAGAGCGCA	CTATATACTA	TAGAAGCTAA	AATTGCTACT	1920
	AATACGCCGT	CGTATGTCAA	TGGGATTAAT	TCTTTTTTAC	AAGGGCTTGG	GGCTATAGGC	1980
	ACTGGATTGG	GCTCGGTTAT	AAGTGTACG	GCAGGAGCAC	TTGGGGATAT	TGTGGGTGGA	2040
60	GTGGTGTCTT	TTTTAAAAAA	TCCATTCGGG	GGTGGTCTCA	TGTTGATTTT	AGCGATAGTA	2100
	GTTGTGTTA	TAATAATTTG	GGTTTTGTTG	AGACAAAAAC	ATGTGCTTAG	TAAGCCTATT	2160
	GACATGATGT	TTCTTATATG	CACCAATCCG	GTGACTACTG	TGTCCAGTGT	TACGGGGACC	2220
	ACTGTGCTCA	AGACGCCTAG	TGTTAAAGAT	GCTGACGGGG	GCACATCTGT	TGCGGTTTCG	2280

GAAAAAGAGG	AGGGTATGGC	TGACGTCACT	GGACAAATAA	GTGGTGATGA	ATATTCACAA	2340
GAAGATGCTT	TAAAAATGCT	CAAGGCCATA	AAGTCTTTAG	ACGAGTCTTA	CAGAAGAAAA	2400
CCTTCGTCTT	CTGAGTCTCA	TGCTCAAAA	CCTAGTTTGA	TAGACAGGAT	CAGGTATAGA	2460
GGTTATAAGA	GTGTAATGT	AGAAGAAGCG	TGA			2493

5

SEQ ID NO:12:

ATGTTTGTTA	CGGCGTGT	GTGCTCTCT	CCAAGCTCGT	TTTATGAGAG	TTTACAAGTA	60
GAGCCACAC	AATCAGAAGA	TATAACCCGG	TCTGCTCCTC	TGGGCGATGG	TGATGAAATC	120
AGAGAAGCTA	TACACAAGTC	CCAGGACGCC	GAACAAAAAC	CCACGTTTTA	CGTCTGCCCA	180
10	CCGCCAACAG	GCTCCACAAT	CGTACGATTA	GAACCAACTC	GGACATGTCC	240
	CTTGGTAAAA	ACTTTACAGA	GGGTATTGCT	GTTGTTTATA	AAGAAAAACAT	300
	AAGTTTAAAG	CGACGGTATA	TTACAAAGAT	GTTATCGTTA	GCACGGCGTG	360
	TCTTATACGC	AAATTACTAA	TAGATATGCG	GATAGGGTAC	CAATTCCCGT	420
	ACGGACACCA	TTGATAAGTT	TGCAAGTGT	TCTTCTAAAG	CAACGTACGT	480
15	CACAAAGTTG	AAGCCTTTAA	TGAGGATAAA	AATCCACAGG	ATATGCCTCT	540
	AAATATAATT	CTGTGGGATC	CAAAGCATGG	CATACTACCA	ATGACACGTA	600
	GGAACCCCGG	GAACATATAG	GACGGGCACG	TCGGTGAATT	GCATCATTGA	660
	GCCAGATCAA	TATCCCTTA	TGATAGTTTT	GGACTTTCCA	CGGGAGATAT	720
	TCCCGTTTTT	TTGGCCTACG	GGATGGTGCA	TACAGAGAAC	ATTCCAATTA	780
20	CGTTTTACC	AGTTTGAGGG	TTATAGACAA	AGGGATCTTG	ACACTAGAGC	840
	CCTGCAGCGC	GGAACTTTTT	AGTCACGCCT	CATTTAACCG	TTGGTTGGAA	900
	AAACGAACGG	AAGTTTGTTC	GCTTGTCAAG	TGGCGTGAGG	TTGAAGACGT	960
	GAGTATGCAC	ACAATTTTCG	CTTTACAATG	AAAACACTTT	CTACCACGTT	1020
	ACAACCGAGT	TTAATCTTAA	CCAAATCCAT	CTCAGTCAAT	GTGTAAGGA	1080
25	GCTATTATTA	ACCGGATCTA	TACAACCAGA	TACAACCTCAT	CTCATGTTAG	1140
	ATCCAGACCT	ACCTTGCCAG	AGGGGGTTTT	GTGTGGTGT	TTCAACCCCT	1200
	TCCCTCGCCC	GTCTCTATCT	CCAAGAATTG	GTCCGTGAAA	ACACTAATCA	1260
	AAACACCCGA	CTCGAAATAC	CAGATCCCGA	CGAAGCGTGC	CAGTTGAGTT	1320
	AGAACAATAA	CAACCACCTC	ATCGGTGGAA	TTTGCTATGC	TCCAGTTTAC	1380
30	ATTCAAGAGC	ATGTTAATGA	AATGTTGGCA	CGTATCTCCT	CGTCTGGTG	1440
	AATCGCGAAC	GCGCCCTTTG	GAGCGGACTA	TTTCCAATTA	ACCCAAGTGC	1500
	ACCATTTTGG	ATCAACGTGT	TAAAGCTCGT	ATTCTCGGG	ACGTTATCTC	1560
	TGTCAGAAC	TGGGATCAGA	TACACGCATT	ATACTTCAAA	ACTCTATGAG	1620
	AGTACTACGC	GTTGTTATAG	CCGTCTTTA	ATTTCAATAG	TTAGTTTAAA	1680
35	ACGGTGGAGG	GCCAGCTTGG	AACAGATAAC	GAGTTAATTA	TGTCCAGAGA	1740
	CCATGCGTGG	CTAATCACAA	GCGATATTTT	CTATTTGGGC	ATCACTACGT	1800
	GATTATCGTT	TAAGCTGTA	AATCGCAGTC	CATGATGTGG	GAATGATTAG	1860
	GATTTAAACT	TAACACTTCT	TAAAGATAGA	GAGTTTATGC	CGCTGCAAGT	1920
	GACGAGCTGC	GGGATACAGG	ATTACTAGAC	TACAGTAAAA	TTCAACGCCG	1980
40	CATTGCTGTC	GTTTTTATGA	CATAGACAAG	GTTGTGCAAT	ATGATAGCGG	2040
	ATGCAGGGCA	TGGCTCAGTT	TTCCAGGGA	CTTGGGACCG	CGGGCCAGGC	2100
	GTGTTCTTGG	GGGCCACGGG	AGCGCTGCTT	TCCACCGTAC	ACGGATTTAC	2160
	TCTAACCCAT	TTGGGGCATT	GGCCGTGGGA	TTATTGGTTT	TGGCGGGACT	2220
	TTTTTTGCGT	ACCGGTACGT	GCTTAAACTT	AAAACAAGCC	CGATGAAGGC	2280
45	CTCACAAACA	AGGGGTTAAA	ACAGTTACCG	GAAGGAATGG	ATCCCTTTGC	2340
	AACGCTACTG	ATACCCCAAT	AGAAGAAATT	GGCGACTCAC	AAAACACTGA	2400
	AATAGCGGGT	TTGATCCCGA	TAAATTTTGA	GAAGCCAGG	AAATGATTAA	2460
	TTAGTATCTG	CGGCTGAGCG	CCAAGAATCT	AAAGCCCGCA	AAAAAATAA	2520
	CTTTTAACTT	CAGCTCTTAC	CGGCCTTGT	TTACGAAATC	GCCGAGGATA	2580
50	CGCACCCAGA	ATGTAACGGG	GGTGTAATA			2608

SEQ ID NO:13:

ATGCGCCAGG	GCGCCGCGG	GGGGTGCCGG	TGTTTCGTGG	TATGGGCGCT	CTTGGGGTTG	60
ACGCTGGGGG	TCCTGGTGGC	GTGCGCGGCT	CCGAGTTCCC	CCGGCACGCC	TGGGGTGGCG	120
55	GCCCGGACCC	AGGCGGGGAA	CGGGGACCT	GCCACTCCGG	CGCGGCCCGC	180
	GCCCAACCGG	GGGACACGAA	ACCGAAGAAG	AACAAAAAAC	CGAAAAACCC	240
	CGCCCCCGCG	GCGACAACGC	GACCGTCGCC	GCGGGCCACG	CCACCCTGCG	300
	CGGGACATCA	AGGCGGAGAA	CACCGATGCA	AACTTTTACG	TGTGCCACC	360
	GCCACGGTGG	TGCAGTTTGA	GCAGCCGCGC	CGCTGCCCGA	CCCGGCCCGA	420
60	TACACGGAGG	GCATCGCGGT	GGTCTTCAAG	GAGAACATCG	CCCCGTACAA	480
	ACCATGTACT	ACAAGACGCT	CACCGTTTCC	CAGGTGTGGT	TEGGCCACCG	540
	TTTATGGGGA	TCTTTGAGGA	CCGCGCCCCC	GTCCCCCTCG	AGGAGGTGAT	600
	AACGCCAAGG	GGGTCTGTGG	GTCCACGGCC	AAGTACGTGC	GCAACAACCT	660

	GCCTTTCACC	GGGACGACCA	CGAGACCGAC	ATGGAGCTGA	AACCGGCCAA	CGCCGCGACC	720
	CGCACGAGCC	GGGGCTGGCA	CACCACCGAC	CTCAAGTACA	ACCCCTCGCG	GGTGGAGGCG	780
	TTCCACCGGT	ACGGGACGAC	GGTAAACTGC	ATGTGCGAGG	AGGTGGACGC	GCGCTCGGTG	840
	TACCCGTACG	ACGAGTTTGT	GCTGGCGACT	GGCGACTTTG	TGTACATGTC	CCCGTTTTAC	900
5	GGCTACCGGG	AGGGGTGCGA	CACCGAACAC	ACCAGCTACG	CGCCGACCG	CTTCAAGCAG	960
	GTTGACGGCT	TCTACGCGCG	CGACCTCACC	ACCAAGGCC	GGCCACGGC	GCCGACCACC	1020
	CGAACCTGCG	TCACGACCCC	CAAGTTCACC	GTGGCCTGGG	ACTGGGTGCC	AAAGCGCCCG	1080
	TCGGTCTGCA	CCATGACCAA	GTGGCAGGAG	GTGGACGAGA	TGCTGCGCTC	CGAGTACGGC	1140
	GGCTCCTTCC	GATTCTCCTC	CGACGCCATA	TCCACCACCT	TCACCACCAA	CCTGACCGAG	1200
10	TACCCGCTCT	CGCGGTGGA	CCTGGGGGAC	TGCATCGGCA	AGGACGCCCG	CGACGCCATG	1260
	GACCGCATCT	TCGCCCGCAT	GTACAACGCG	ACGCACATCA	AGGTGGGCCA	GCCGCGATAC	1320
	TACCTGGCCA	ATGGGGGCTT	TCTGATCGCG	TACCAGCCCC	TTCTCAGCAA	CACGCTCGCG	1380
	GAGCTGTACG	TGCGGGAACA	CCTCCGAGAG	CAGAGCCGCA	AGCCCCAAA	CCCCACGCC	1440
	CCGCCGCCCG	GGGCCAGCGC	CAACGCGTCC	GTGGAGCGCA	TCAAGACCAC	CTCCTCCATC	1500
15	GAGTTCGCCC	GGCTGCAGTT	TACGTACAAC	CACATACAGC	GCCATGTCAA	CGATATGTTG	1560
	GGCCGCGTTG	CCATCGCGTG	GTGGGAGCTG	CAGAATCACG	AGCTGACCC	GTGGAACGAG	1620
	GCCCGCAAGC	TGAACCCCAA	CGCCATCGCC	TCGGCCACCG	TGGCCGGCG	GGTGAAGCGG	1680
	CGGATGCTCG	GCGACGTGAT	GGCCGTCTCC	ACGTGCGTGC	CGGTGCGCG	GGACAACGTG	1740
	ATCGTCCAAA	ACTCGATGCG	CATCAGCTCG	CGGCCCGGGG	CCTGCTACAG	CCGCCCCCTG	1800
20	GTCAGCTTTC	GGTACGAAGA	CCAGGGCCCG	TTGGTGGAGG	GGCAGTGGG	GGAGAACAAC	1860
	GAGCTGCGGC	TGACGCGCGA	TGCGATCGAG	CCGTGCACCG	TGGACACCG	GCGCTACTTC	1920
	ACCTTCGGTG	GGGGTACGT	GACTTTCGAG	GAGTACGCGT	ACTCCACCA	GCTGAGCCGC	1980
	GCCGACATCA	CCACCGTCAG	CACCTTCATC	GACCTCAACA	TCACCATGCT	GGAGGATCAC	2040
	GAGTTTGTCC	CCCTGGAGGT	GTACACCCGC	CACGAGATCA	AGGACAGCGG	CCTGTGGAC	2100
25	TACACGGAGG	TCCAGCGCCG	CAACCAGCTG	CACGACCTGC	GCTTCGCCGA	CATCGACACG	2160
	GTCATCCACG	CCGACGCCAA	CGCCGCCATG	TTCCGGGGCC	TGGCGCGT	CTTCGAGGGG	2220
	ATGGGCGACC	TGGGGCGCGC	GGTGGGCAAG	GTGGTGTGG	GCATCGTGGG	CGCGTGGTA	2280
	TCGGCCGTGT	CGGGCGTGT	CTCCTTCATG	TCCAACCCCT	TTGGGGCGCT	GGCCGTGGGT	2340
	CTGTTGGTCC	TGGCCGGCCT	GGCGGCGCT	TTCTTCGCCT	TTGCTACGT	CATGCGGCTG	2400
30	CAGAGCAACC	CCATGAAGGC	CCTGTACCCG	CTAACCACCA	AGGAGCTCAA	GAACCCACC	2460
	AACCCGGACG	CGTCCGGGA	GGCGGAGGAG	GGCGGCGACT	TTGACGAGGC	CAAGCTAGCC	2520
	GAGGCCCGGG	AGATGATACG	GTACATGGCC	CTGGTGTCTG	CCATGGAGCG	CACGGAACAC	2580
	AAGGCCAAGA	AGAAGGGCAC	GAGCGCGCTG	CTCAGCGCCA	AGGTACCCGA	CATGCTCATG	2640
	CGAAGCGCC	GCAACACCAA	CTACACCCAA	GTTCCCAACA	AAGACGGTGA	CGCCGACGAG	2700
35	GACGACCTGT	GAC					2713

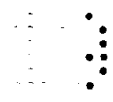
SEQ ID NO:14:

	Met	Val	Pro	Asn	Lys	His	Leu	Leu	Leu	Ile	Ile	Leu	Ser	Phe	Ser	Thr
	Ala	Cys	Gly	Gln	Thr	Thr	Pro	Thr	Thr	Ala	Val	Glu	Lys	Asn	Lys	Thr
40	Gln	Ala	Ile	Tyr	Gln	Glu	Tyr	Phe	Lys	Tyr	Arg	Val	Cys	Ser	Ala	Ser
	Thr	Thr	Gly	Glu	Leu	Phe	Arg	Phe	Asp	Leu	Asp	Arg	Thr	Cys	Pro	Ser
	Thr	Glu	Asp	Lys	Val	His	Lys	Glu	Gly	Ile	Leu	Leu	Val	Tyr	Lys	Lys
	Asn	Ile	Val	Pro	Tyr	Ile	Phe	Lys	Val	Arg	Arg	Tyr	Lys	Lys	Ile	Thr
	Thr	Ser	Val	Arg	Ile	Phe	Asn	Gly	Trp	Thr	Arg	Glu	Gly	Val	Ala	Ile
45	Thr	Asn	Lys	Trp	Glu	Leu	Ser	Arg	Ala	Val	Pro	Lys	Tyr	Glu	Ile	Asp
	Ile	Met	Asp	Lys	Thr	Tyr	Gln	Cys	His	Asn	Cys	Met	Gln	Ile	Glu	Val
	Asn	Gly	Met	Leu	Asn	Ser	Tyr	Tyr	Asp	Arg	Asp	Gly	Asn	Asn	Lys	Thr
	Val	Asp	Leu	Lys	Pro	Val	Asp	Gly	Leu	Thr	Gly	Ala	Ile	Thr	Arg	Tyr
	Ile	Ser	Gln	Pro	Lys	Val	Phe	Ala	Asp	Pro	Gly	Trp	Leu	Trp	Gly	Thr
50	Tyr	Arg	Thr	Arg	Thr	Thr	Val	Asn	Cys	Glu	Ile	Val	Asp	Met	Phe	Ala
	Arg	Ser	Ala	Asp	Pro	Tyr	Thr	Tyr	Phe	Val	Thr	Ala	Leu	Gly	Asp	Thr
	Val	Glu	Val	Ser	Pro	Phe	Cys	Asp	Val	Asp	Asn	Ser	Cys	Pro	Asn	Ala
	Thr	Asp	Val	Leu	Ser	Val	Gln	Ile	Asp	Leu	Asn	His	Thr	Val	Val	Asp
	Tyr	Gly	Asn	Arg	Ala	Thr	Ser	Gln	Gln	His	Lys	Lys	Arg	Ile	Phe	Ala
55	His	Thr	Leu	Asp	Tyr	Ser	Val	Ser	Trp	Glu	Ala	Val	Asn	Lys	Ser	Ala
	Ser	Val	Cys	Ser	Met	Val	Phe	Trp	Lys	Ser	Phe	Gln	Arg	Ala	Ile	Gln
	Thr	Glu	His	Asp	Leu	Thr	Tyr	His	Phe	Ile	Ala	Asn	Glu	Ile	Thr	Ala
	Gly	Phe	Ser	Thr	Val	Lys	Glu	Pro	Leu	Ala	Asn	Phe	Thr	Ser	Asp	Tyr
	Asn	Cys	Leu	Met	Thr	His	Ile	Asn	Thr	Thr	Leu	Glu	Asp	Lys	Ile	Ala
60	Arg	Val	Asn	Asn	Thr	His	Thr	Pro	Asn	Gly	Thr	Ala	Glu	Tyr	Tyr	Gln
	Thr	Glu	Gly	Gly	Met	Ile	Leu	Val	Trp	Gln	Pro	Leu	Ile	Ala	Ile	Glu
	Leu	Glu	Glu	Ala	Met	Leu	Glu	Ala	Thr	Thr	Ser	Pro	Val	Thr	Pro	Ser
	Ala	Pro	Thr	Ser	Ser	Ser	Arg	Ser	Lys	Arg	Ala	Ile	Arg	Ser	Ile	Arg

Asp Val Ser Ala Gly Ser Glu Asn Asn Val Phe Leu Ser Gln Ile Gln
 Tyr Ala Tyr Asp Lys Leu Arg Gln Ser Ile Asn Asn Val Leu Glu Glu
 Leu Ala Ile Thr Trp Cys Arg Glu Gln Val Arg Gln Thr Met Val Trp
 Tyr Glu Ile Ala Lys Ile Asn Pro Thr Ser Val Met Thr Ala Ile Tyr
 5 Gly Lys Pro Val Ser Arg Lys Ala Leu Gly Asp Val Ile Ser Val Thr
 Glu Cys Ile Asn Val Asp Gln Ser Ser Val Ser Ile His Lys Ser Leu
 Lys Thr Glu Asn Asn Asp Ile Cys Tyr Ser Arg Pro Pro Val Thr Phe
 Lys Phe Val Asn Ser Ser Gln Leu Phe Lys Gly Gln Leu Gly Ala Arg
 10 Asn Glu Ile Leu Leu Ser Glu Ser Leu Val Glu Asn Cys His Gln Asn
 Ala Glu Thr Phe Phe Thr Ala Lys Asn Glu Thr Tyr His Phe Lys Asn
 Tyr Val His Val Glu Thr Leu Pro Val Asn Asn Ile Ser Thr Leu Asp
 Thr Phe Leu Ala Leu Asn Leu Thr Phe Ile Glu Asn Ile Asp Phe Lys
 Ala Val Glu Leu Tyr Ser Ser Gly Glu Arg Lys Leu Ala Asn Val Phe
 Asp Leu Glu Thr Met Phe Arg Glu Tyr Asn Tyr Tyr Ala Gln Ser Ile
 15 Ser Gly Leu Arg Lys Asp Phe Asp Asn Ser Gln Arg Asn Asn Arg Asp
 Arg Ile Ile Gln Asp Phe Ser Glu Ile Leu Ala Asp Leu Gly Ser Ile
 Gly Lys Val Ile Val Asn Val Ala Ser Gly Ala Phe Ser Leu Phe Gly
 Gly Ile Val Thr Gly Ile Leu Asn Phe Ile Lys Asn Pro Leu Gly Gly
 Met Phe Thr Phe Leu Leu Ile Gly Ala Val Ile Ile Leu Val Ile Leu
 20 Leu Val Arg Arg Thr Asn Asn Met Ser Gln Ala Pro Ile Arg Met Ile
 Tyr Pro Asp Val Glu Lys Ser Lys Ser Thr Val Thr Pro Met Glu Pro
 Glu Thr Ile Lys Gln Ile Leu Leu Gly Met His Asn Met Gln Gln Glu
 Ala Tyr Lys Lys Lys Glu Glu Gln Arg Ala Ala Arg Pro Ser Ile Phe
 Arg Gln Ala Ala Glu Thr Phe Leu Arg Lys Arg Ser Gly Tyr Lys Gln
 25 Ile Ser Thr Glu Asp Lys Ile Val

SEQ ID NO:15:

Met Tyr Tyr Lys Thr Ile Leu Phe Phe Ala Leu Ile Lys Val Cys Ser
 Phe Asn Gln Thr Thr Thr His Ser Thr Thr Thr Ser Pro Ser Ile Ser
 30 Ser Thr Thr Ser Ser Thr Thr Thr Ser Thr Ser Lys Pro Ser Asn Thr
 Thr Ser Thr Asn Ser Ser Leu Ala Ala Ser Pro Gln Asn Thr Ser Thr
 Ser Lys Pro Ser Thr Asp Asn Gln Gly Thr Ser Thr Pro Thr Ile Pro
 Thr Val Thr Asp Asp Thr Ala Ser Lys Asn Phe Tyr Lys Tyr Arg Val
 Cys Ser Ala Ser Ser Ser Ser Gly Glu Leu Phe Arg Phe Asp Leu Asp
 35 Gln Thr Cys Pro Asp Thr Lys Asp Lys Lys His Val Glu Gly Ile Leu
 Leu Val Leu Lys Lys Asn Ile Val Pro Tyr Ile Phe Lys Val Arg Lys
 Tyr Arg Lys Ile Ala Thr Ser Val Thr Val Tyr Arg Gly Trp Ser Gln
 Ala Ala Val Thr Asn Arg Asp Asp Ile Ser Arg Ala Ile Pro Tyr Asn
 Glu Ile Ser Met Ile Asp Arg Thr Tyr His Cys Phe Ser Ala Met Ala
 40 Thr Val Ile Asn Gly Ile Leu Asn Thr Tyr Ile Asp Arg Asp Ser Glu
 Asn Lys Ser Val Pro Leu Gln Pro Val Ala Gly Leu Thr Glu Asn Ile
 Asn Arg Tyr Phe Ser Gln Pro Leu Ile Tyr Ala Glu Pro Gly Trp Phe
 Pro Gly Ile Tyr Arg Val Arg Thr Thr Val Asn Cys Glu Val Val Asp
 Met Tyr Ala Arg Ser Val Glu Pro Tyr Thr His Phe Ile Thr Ala Leu
 45 Gly Asp Thr Ile Glu Ile Ser Pro Phe Cys His Asn Asn Ser Gln Cys
 Thr Thr Gly Asn Ser Thr Ser Arg Asp Ala Thr Lys Val Trp Ile Glu
 Glu Asn His Gln Thr Val Asp Tyr Glu Arg Arg Gly His Pro Thr Lys
 Asp Lys Arg Ile Phe Leu Lys Asp Glu Glu Tyr Thr Ile Ser Trp Lys
 Ala Glu Asp Arg Glu Arg Ala Ile Cys Asp Phe Val Ile Trp Lys Thr
 50 Phe Pro Arg Ala Ile Gln Thr Ile His Asn Glu Ser Phe His Phe Val
 Ala Asn Glu Val Thr Ala Ser Phe Leu Thr Ser Asn Gln Glu Glu Thr
 Glu Leu Arg Gly Asn Thr Glu Ile Leu Asn Cys Met Asn Ser Thr Ile
 Asn Glu Thr Leu Glu Glu Thr Val Lys Lys Phe Asn Lys Ser His Ile
 Arg Asp Gly Glu Val Lys Tyr Tyr Lys Thr Asn Gly Gly Leu Phe Leu
 55 Ile Trp Gln Ala Met Lys Pro Leu Asn Leu Ser Glu His Thr Asn Tyr
 Thr Ile Glu Arg Asn Asn Lys Thr Gly Asn Lys Ser Arg Gln Lys Arg
 Ser Val Asp Thr Lys Thr Phe Gln Gly Ala Lys Gly Leu Ser Thr Ala
 Gln Val Gln Tyr Ala Tyr Asp His Leu Arg Thr Ser Met Asn His Ile
 Leu Glu Glu Leu Thr Lys Thr Trp Cys Arg Glu Gln Lys Lys Asp Asn
 60 Leu Met Trp Tyr Glu Leu Ser Lys Ile Asn Pro Val Ser Val Met Ala
 Ala Ile Tyr Gly Lys Pro Val Ala Val Lys Ala Met Gly Asp Ala Phe
 Met Val Ser Glu Cys Ile Asn Val Asp Gln Ala Ser Val Asn Ile His
 Lys Ser Met Arg Thr Asp Asp Pro Lys Val Cys Tyr Ser Arg Pro Leu



Val Thr Phe Lys Phe Val Asn Ser Thr Ala Thr Phe Arg Gly Gln Leu
 Gly Thr Arg Asn Glu Ile Leu Leu Thr Asn Thr His Val Glu Thr Cys
 Arg Pro Thr Ala Asp His Tyr Phe Phe Val Lys Asn Met Thr His Tyr
 Phe Lys Asp Tyr Lys Phe Val Lys Thr Met Asp Thr Asn Asn Ile Ser
 5 Thr Leu Asp Thr Phe Leu Thr Leu Asn Leu Thr Phe Ile Asp Asn Ile
 Asp Phe Lys Thr Val Glu Leu Tyr Ser Glu Thr Glu Arg Lys Met Ala
 Ser Ala Leu Asp Leu Glu Thr Met Phe Arg Glu Tyr Asn Tyr Tyr Thr
 Gln Lys Leu Ala Ser Leu Arg Glu Asp Leu Asp Asn Thr Ile Asp Leu
 10 Asn Arg Asp Arg Leu Val Lys Asp Leu Ser Glu Met Met Ala Asp Leu
 Gly Asp Ile Gly Lys Val Val Val Asn Thr Phe Ser Gly Ile Val Thr
 Val Phe Gly Ser Ile Val Gly Phe Val Ser Phe Phe Thr Asn Pro
 Ile Gly Gly Val Thr Ile Ile Leu Leu Leu Ile Val Val Val Phe Val
 Val Phe Ile Val Ser Arg Arg Thr Asn Asn Met Asn Glu Ala Pro Ile
 15 Lys Met Ile Tyr Pro Asn Ile Asp Lys Ala Ser Glu Gln Glu Asn Ile
 Gln Pro Leu Pro Gly Glu Glu Ile Lys Arg Ile Leu Leu Gly Met His
 Gln Leu Gln Gln Ser Glu His Gly Lys Ser Glu Glu Glu Ala Ser His
 Lys Pro Gly Leu Phe Gln Leu Leu Gly Asp Gly Leu Gln Leu Leu Arg
 Arg Arg Gly Tyr Thr Arg Leu Pro Thr Phe Asp Pro Ser Pro Gly Asn
 20 Asp Thr Ser Glu Thr His Gln Lys Tyr Val

SEQ ID NO:16:

Met Gly Val Gly Gly Gly Pro Arg Val Val Leu Cys Leu Trp Cys Val
 Ala Ala Leu Leu Cys Gln Gly Val Ala Gln Glu Val Val Ala Glu Thr
 Thr Thr Pro Phe Ala Thr His Arg Pro Glu Val Val Ala Glu Glu Asn
 25 Pro Ala Asn Pro Phe Leu Pro Phe Arg Val Cys Gly Ala Ser Pro Thr
 Gly Gly Glu Ile Phe Arg Phe Pro Leu Glu Glu Ser Cys Pro Asn Thr
 Glu Asp Lys Asp His Ile Glu Gly Ile Ala Leu Ile Tyr Lys Thr Asn
 Ile Val Pro Tyr Val Phe Asn Val Arg Lys Tyr Arg Lys Ile Met Thr
 Ser Thr Thr Ile Tyr Lys Gly Trp Ser Glu Asp Ala Ile Thr Asn Gln
 30 His Thr Arg Ser Tyr Ala Val Pro Leu Tyr Glu Val Gln Met Met Asp
 His Tyr Tyr Gln Cys Phe Ser Ala Val Gln Val Asn Glu Gly Gly His
 Val Asn Thr Tyr Tyr Asp Arg Asp Gly Trp Asn Glu Thr Ala Phe Leu
 Lys Pro Ala Asp Gly Leu Thr Ser Ser Ile Thr Arg Tyr Gln Ser Gln
 Pro Glu Val Tyr Ala Thr Pro Arg Asn Leu Leu Trp Ser Tyr Thr Thr
 35 Arg Thr Thr Val Asn Cys Glu Val Thr Glu Met Ser Ala Arg Ser Met
 Lys Pro Phe Glu Phe Phe Val Thr Ser Val Gly Asp Thr Ile Glu Met
 Ser Pro Phe Leu Lys Glu Asn Gly Thr Glu Pro Glu Lys Ile Leu Lys
 Arg Pro His Ser Ile Gln Leu Leu Lys Asn Tyr Ala Val Thr Lys Tyr
 Gly Val Gly Leu Gly Gln Ala Asp Asn Ala Thr Arg Phe Phe Ala Ile
 40 Phe Gly Asp Tyr Ser Leu Ser Trp Lys Ala Thr Thr Glu Asn Ser Ser
 Tyr Cys Asp Leu Ile Leu Trp Lys Gly Phe Ser Asn Ala Ile Gln Thr
 Gln His Asn Ser Ser Leu His Phe Ile Ala Asn Asp Ile Thr Ala Ser
 Phe Ser Thr Pro Leu Glu Glu Glu Ala Asn Phe Asn Glu Thr Phe Lys
 Cys Ile Trp Asn Asn Thr Gln Glu Glu Ile Gln Lys Lys Leu Lys Glu
 45 Val Glu Lys Thr His Arg Pro Asn Gly Thr Ala Lys Val Tyr Lys Thr
 Thr Gly Asn Leu Tyr Ile Val Trp Gln Pro Leu Ile Gln Ile Asp Leu
 Leu Asp Thr His Ala Lys Leu Tyr Asn Leu Thr Asn Ala Thr Ala Ser
 Pro Thr Ser Thr Pro Thr Thr Ser Pro Arg Arg Arg Arg Arg Asp Thr
 Ser Ser Val Ser Gly Gly Gly Asn Asn Gly Asp Asn Ser Thr Lys Glu
 50 Glu Ser Val Ala Ala Ser Gln Val Gln Phe Ala Tyr Asp Asn Leu Arg
 Lys Ser Ile Asn Arg Val Leu Gly Glu Leu Ser Arg Ala Trp Cys Arg
 Glu Gln Tyr Arg Ala Ser Leu Met Trp Tyr Glu Leu Ser Lys Ile Asn
 Pro Thr Ser Val Met Ser Ala Ile Tyr Gly Arg Pro Val Ser Ala Lys
 Leu Ile Gly Asp Val Val Ser Val Ser Asp Cys Ile Ser Val Asp Gln
 55 Lys Ser Val Phe Val His Lys Asn Met Lys Val Pro Gly Lys Glu Asp
 Leu Cys Tyr Thr Arg Pro Val Val Gly Phe Lys Phe Ile Asn Gly Ser
 Glu Leu Phe Ala Gly Gln Leu Gly Pro Arg Asn Glu Ile Val Leu Ser
 Thr Ser Gln Val Glu Val Cys Gln His Ser Cys Glu His Tyr Phe Gln
 Ala Gly Asn Gln Met Tyr Lys Tyr Lys Asp Tyr Tyr Tyr Val Ser Thr
 60 Leu Asn Leu Thr Asp Ile Pro Thr Leu His Thr Met Ile Thr Asn
 Leu Ser Leu Val Glu Asn Ile Asp Phe Lys Val Ile Glu Leu Tyr Ser
 Lys Thr Glu Lys Arg Leu Ser Asn Val Phe Asp Ile Glu Thr Met Phe
 Arg Glu Tyr Asn Tyr Tyr Thr Gln Asn Leu Asn Gly Leu Arg Lys Asp



Leu Asp Asp Ser Ile Asp His Gly Arg Asp Ser Phe Ile Gln Thr Leu
 Gly Asp Ile Met Gln Asp Leu Gly Thr Ile Gly Lys Val Val Val Asn
 Val Ala Ser Gly Val Phe Ser Leu Phe Gly Ser Ile Val Ser Gly Val
 Ile Ser Phe Phe Lys Asn Pro Phe Gly Gly Met Leu Leu Ile Val Leu
 5 Ile Ile Ala Gly Val Val Val Val Tyr Leu Phe Met Thr Arg Ser Arg
 Ser Ile Tyr Ser Ala Pro Ile Arg Met Leu Tyr Pro Gly Val Glu Arg
 Ala Ala Gln Glu Pro Gly Ala His Pro Val Ser Glu Asp Gln Ile Arg
 Asn Ile Leu Met Gly Met His Gln Phe Gln Gln Arg Gln Arg Ala Glu
 10 Glu Glu Ala Arg Arg Glu Glu Glu Val Lys Gly Lys Arg Thr Leu Phe
 Glu Val Ile Arg Asp Ser Ala Thr Ser Val Leu Arg Arg Arg Arg Gly
 Gly Gly Gly Tyr Gln Arg Leu Gln Arg Asp Gly Ser Asp Asp Glu Gly
 Asp Tyr Glu Pro Leu Arg Arg Gln Asp Gly Gly Tyr Asp Asp Val Asp
 Val Glu Ala Gly Thr Ala Asp Thr Gly Val

15 SEQ ID NO:17:

Met Tyr Pro Thr Val Lys Ser Met Arg Val Ala His Leu Thr Asn Leu
 Leu Thr Leu Leu Cys Leu Leu Cys His Thr His Leu Tyr Val Cys Gln
 Pro Thr Thr Leu Arg Gln Pro Ser Asp Met Thr Pro Ala Gln Asp Ala
 20 Pro Thr Glu Thr Pro Pro Pro Leu Ser Thr Asn Thr Asn Arg Gly Phe
 Glu Tyr Phe Arg Val Cys Gly Val Ala Ala Thr Gly Glu Thr Phe Arg
 Phe Asp Leu Asp Lys Thr Cys Pro Ser Thr Gln Asp Lys Lys His Val
 Glu Gly Ile Leu Leu Val Tyr Lys Ile Asn Ile Val Pro Tyr Ile Phe
 Lys Ile Arg Arg Tyr Arg Lys Ile Ile Thr Gln Leu Thr Ile Trp Arg
 Gly Leu Thr Thr Ser Ser Val Thr Gly Lys Phe Glu Met Ala Thr Gln
 25 Ala His Glu Trp Glu Val Gly Asp Phe Asp Ser Ile Tyr Gln Cys Tyr
 Asn Ser Ala Thr Met Val Val Asn Asn Val Arg Gln Val Tyr Val Asp
 Arg Asp Gly Val Asn Lys Thr Val Asn Ile Arg Pro Val Asp Gly Leu
 Thr Gly Asn Ile Gln Arg Tyr Phe Ser Gln Pro Thr Leu Tyr Ser Glu
 Pro Gly Trp Met Pro Gly Phe Tyr Arg Val Arg Thr Thr Val Asn Cys
 30 Glu Ile Val Asp Met Val Ala Arg Ser Met Asp Pro Tyr Asn Tyr Ile
 Ala Thr Ala Leu Gly Asp Ser Leu Glu Leu Ser Pro Phe Gln Thr Phe
 Asp Asn Thr Ser Gln Cys Thr Ala Pro Lys Arg Ala Asp Met Arg Val
 Arg Glu Val Lys Asn Tyr Lys Phe Val Asp Tyr Asn Asn Arg Gly Thr
 Ala Pro Ala Gly Gln Ser Arg Thr Phe Leu Glu Thr Pro Ser Ala Thr
 35 Tyr Ser Trp Lys Thr Ala Thr Arg Gln Thr Ala Thr Cys Asp Leu Val
 His Trp Lys Thr Phe Pro Arg Ala Ile Gln Thr Ala His Glu His Ser
 Tyr His Phe Val Ala Asn Glu Val Thr Ala Thr Phe Asn Thr Pro Leu
 Thr Glu Val Glu Asn Phe Thr Ser Thr Tyr Ser Cys Val Ser Asp Gln
 Ile Asn Lys Thr Ile Ser Glu Tyr Ile Gln Lys Leu Asn Asn Ser Tyr
 40 Val Ala Ser Gly Lys Thr Gln Tyr Phe Lys Thr Asp Gly Asn Leu Tyr
 Leu Ile Trp Gln Pro Leu Glu His Pro Glu Ile Glu Asp Ile Asp Glu
 Asp Ser Asp Pro Glu Pro Thr Pro Ala Pro Pro Lys Ser Thr Arg Arg
 Lys Arg Glu Ala Ala Asp Asn Gly Asn Ser Thr Ser Glu Val Ser Lys
 Gly Ser Glu Asn Pro Leu Ile Thr Ala Gln Ile Gln Phe Ala Tyr Asp
 45 Lys Leu Thr Thr Ser Val Asn Asn Val Leu Glu Glu Leu Ser Arg Ala
 Trp Cys Arg Glu Gln Val Arg Asp Thr Leu Met Trp Tyr Glu Leu Ser
 Lys Val Asn Pro Thr Ser Val Met Ser Ala Ile Tyr Gly Lys Pro Val
 Ala Ala Arg Tyr Val Gly Asp Ala Ile Ser Val Thr Asp Cys Ile Tyr
 Val Asp Gln Ser Ser Val Asn Ile His Gln Ser Leu Arg Leu Gln His
 50 Asp Lys Thr Thr Cys Tyr Ser Arg Pro Arg Val Thr Phe Lys Phe Ile
 Asn Ser Thr Asp Pro Leu Thr Gly Gln Leu Gly Pro Arg Lys Glu Ile
 Ile Leu Ser Asn Thr Asn Ile Glu Thr Cys Lys Asp Glu Ser Glu His
 Tyr Phe Ile Val Gly Glu Tyr Ile Tyr Tyr Tyr Lys Asn Tyr Ile Phe
 Glu Glu Lys Leu Asn Leu Ser Ser Ile Ala Thr Leu Asp Thr Phe Ile
 55 Ala Leu Asn Ile Ser Phe Ile Glu Asn Ile Asp Phe Lys Thr Val Glu
 Leu Tyr Ser Ser Thr Glu Arg Lys Leu Ala Ser Ser Val Phe Asp Ile
 Glu Ser Met Phe Arg Glu Tyr Asn Tyr Tyr Thr Tyr Ser Leu Ala Gly
 Ile Lys Lys Asp Leu Asp Asn Thr Ile Asp Tyr Asn Arg Asp Arg Leu
 Val Gln Asp Leu Ser Asp Met Met Ala Asp Leu Gly Asp Ile Gly Arg
 60 Ser Val Val Asn Val Val Ser Ser Val Val Thr Phe Phe Ser Ser Ile
 Val Thr Gly Phe Ile Lys Phe Phe Thr Asn Pro Leu Gly Gly Ile Phe
 Ile Leu Leu Ile Ile Gly Gly Ile Ile Phe Leu Val Val Leu Asn
 Arg Arg Asn Ser Gln Phe His Asp Ala Pro Ile Lys Met Leu Tyr Pro

Ser Val Glu Asn Tyr Ala Ala Arg Gln Ala Pro Pro Pro Tyr Ser Ala
 Ser Pro Pro Ala Ile Asp Lys Glu Glu Ile Lys Arg Ile Leu Leu Gly
 Met His Gln Val His Gln Glu Glu Lys Glu Ala Gln Lys Gln Leu Thr
 5 Asn Ser Gly Pro Thr Leu Trp Gln Lys Ala Thr Gly Phe Leu Arg Asn
 Arg Arg Lys Gly Tyr Ser Gln Leu Pro Leu Glu Asp Glu Ser Thr Ser
 Leu

SEQ ID NO:18:

10 Met Thr Arg Arg Arg Val Leu Ser Val Val Val Leu Leu Ala Ala Leu
 Ala Cys Arg Leu Gly Ala Gln Thr Pro Glu Gln Pro Ala Pro Pro Ala
 Thr Thr Val Gln Pro Thr Ala Thr Arg Gln Gln Thr Ser Phe Pro Phe
 Arg Val Cys Glu Leu Ser Ser His Gly Asp Leu Phe Arg Phe Ser Ser
 Asp Ile Gln Cys Pro Ser Phe Gly Thr Arg Glu Asn His Thr Glu Gly
 15 Leu Leu Met Val Phe Lys Asp Asn Ile Ile Pro Tyr Ser Phe Lys Val
 Arg Ser Tyr Thr Lys Ile Val Thr Asn Ile Leu Ile Tyr Asn Gly Trp
 Tyr Ala Asp Ser Val Thr Asn Arg His Glu Glu Lys Phe Ser Val Asp
 Ser Tyr Glu Thr Asp Gln Met Asp Thr Ile Tyr Gln Cys Tyr Asn Ala
 Val Lys Met Thr Lys Asp Gly Leu Thr Arg Val Tyr Val Asp Arg Asp
 20 Gly Val Asn Ile Thr Val Asn Leu Lys Pro Thr Gly Gly Leu Ala Asn
 Gly Val Arg Arg Tyr Ala Ser Gln Thr Glu Leu Tyr Asp Ala Pro Gly
 Trp Leu Ile Trp Thr Tyr Arg Thr Arg Thr Thr Val Asn Cys Leu Ile
 Thr Asp Met Met Ala Lys Ser Asn Ser Pro Phe Asp Phe Phe Val Thr
 Thr Thr Gly Gln Thr Val Glu Met Ser Pro Phe Tyr Asp Gly Lys Asn
 25 Lys Glu Thr Phe His Glu Arg Ala Asp Ser Phe His Val Arg Thr Asn
 Tyr Lys Ile Val Asp Tyr Asp Asn Arg Gly Thr Asn Pro Gln Gly Glu
 Arg Arg Ala Phe Leu Asp Lys Gly Thr Tyr Thr Leu Ser Trp Lys Leu
 Glu Asn Arg Thr Ala Tyr Cys Pro Leu Gln His Trp Gln Thr Phe Asp
 Ser Thr Ile Ala Thr Glu Thr Gly Lys Ser Ile His Phe Val Thr Asp
 30 Glu Gly Thr Ser Ser Phe Val Thr Asn Thr Thr Val Gly Ile Glu Leu
 Pro Asp Ala Phe Lys Cys Ile Glu Glu Gln Val Asn Lys Thr Met His
 Glu Lys Tyr Glu Ala Val Gln Asp Arg Tyr Thr Lys Gly Gln Glu Ala
 Ile Thr Tyr Phe Ile Thr Ser Gly Gly Leu Leu Leu Ala Trp Leu Pro
 Leu Thr Pro Arg Ser Leu Ala Thr Val Lys Asn Leu Thr Glu Leu Thr
 Thr Pro Thr Ser Ser Pro Pro Ser Ser Pro Pro Ala Pro Ser
 35 Ala Ala Arg Gly Ser Thr Pro Ala Ala Val Leu Arg Arg Arg Arg Arg
 Asp Ala Gly Asn Ala Thr Thr Pro Val Pro Pro Thr Ala Pro Gly Lys
 Ser Leu Gly Thr Leu Asn Asn Pro Ala Thr Val Gln Ile Gln Phe Ala
 Tyr Asp Ser Leu Arg Arg Gln Ile Asn Arg Met Leu Gly Asp Leu Ala
 Arg Ala Trp Cys Leu Glu Gln Lys Arg Gln Asn Met Val Leu Arg Glu
 40 Leu Thr Lys Ile Asn Pro Thr Thr Val Met Ser Ser Ile Tyr Gly Lys
 Ala Val Ala Ala Lys Arg Leu Gly Asp Val Ile Ser Val Ser Gln Cys
 Val Pro Val Asn Gln Ala Thr Val Thr Leu Arg Lys Ser Met Arg Val
 Pro Gly Ser Glu Thr Met Cys Tyr Ser Arg Pro Leu Val Ser Phe Ser
 Phe Ile Asn Asp Thr Lys Thr Tyr Glu Gly Gln Leu Gly Thr Asp Asn
 45 Glu Ile Phe Leu Thr Lys Lys Met Thr Glu Val Cys Gln Ala Thr Ser
 Gln Tyr Tyr Phe Gln Ser Gly Asn Glu Ile His Val Tyr Asn Asp Tyr
 His His Phe Lys Thr Ile Glu Leu Asp Gly Ile Ala Thr Leu Gln Thr
 Phe Ile Ser Leu Asn Thr Ser Leu Ile Glu Asn Ile Asp Phe Ala Ser
 Leu Glu Leu Tyr Ser Arg Asp Glu Gln Arg Ala Ser Asn Val Phe Asp
 50 Leu Glu Gly Ile Phe Arg Glu Tyr Asn Phe Gln Ala Gln Asn Ile Ala
 Gly Leu Arg Lys Asp Leu Asp Asn Ala Val Ser Asn Gly Arg Asn Gln
 Phe Val Asp Gly Leu Gly Glu Leu Met Asp Ser Leu Gly Ser Val Gly
 Gln Ser Ile Thr Asn Leu Val Ser Thr Val Gly Gly Leu Phe Ser Ser
 Leu Val Ser Gly Phe Ile Ser Phe Phe Lys Asn Pro Phe Gly Gly Met
 55 Leu Ile Leu Val Leu Val Ala Gly Val Val Ile Leu Val Ile Ser Leu
 Thr Arg Arg Thr Arg Gln Met Ser Gln Gln Pro Val Gln Met Leu Tyr
 Pro Gly Ile Asp Glu Leu Ala Gln Gln His Ala Ser Gly Glu Gly Pro
 Gly Ile Asn Pro Ile Ser Lys Thr Glu Leu Gln Ala Ile Met Leu Ala
 Leu His Glu Gln Asn Gln Glu Gln Lys Arg Ala Ala Gln Arg Ala Ala
 60 Gly Pro Ser Val Ala Ser Arg Ala Leu Gln Ala Ala Arg Asp Arg Phe
 Pro Gly Leu Arg Arg Arg Tyr His Asp Pro Glu Thr Ala Ala Ala
 Leu Leu Gly Glu Ala Glu Thr Glu Phe

SEQ ID NO:19:

Met Glu Ser Arg Ile Trp Cys Leu Val Val Cys Val Asn Leu Cys Ile
Val Cys Leu Gly Ala Ala Val Ser Ser Ser Ser Thr Arg Gly Thr Ser
Ala Thr His Ser His His Ser Ser His Thr Thr Ser Ala Ala His Ser
5 Arg Ser Gly Ser Val Ser Gln Arg Val Thr Ser Ser Gln Thr Val Ser
His Gly Val Asn Glu Thr Ile Tyr Asn Thr Thr Leu Lys Tyr Gly Asp
Val Val Gly Val Asn Thr Thr Lys Tyr Pro Tyr Arg Val Cys Ser Met
Ala Gln Gly Thr Asp Leu Ile Arg Phe Glu Arg Asn Ile Val Cys Thr
Ser Met Lys Pro Ile Asn Glu Asp Leu Asp Glu Gly Ile Met Val Val
10 Tyr Lys Arg Asn Ile Val Ala His Thr Phe Lys Val Arg Val Tyr Gln
Lys Val Leu Thr Phe Arg Arg Ser Tyr Ala Tyr Ile His Thr Thr Tyr
Leu Leu Gly Ser Asn Thr Glu Tyr Val Ala Pro Pro Met Trp Glu Ile
His His Ile Asn Ser His Ser Gln Cys Tyr Ser Ser Tyr Ser Arg Val
Ile Ala Gly Thr Val Phe Val Ala Tyr His Arg Asp Ser Tyr Glu Asn
15 Lys Thr Met Gln Leu Met Pro Asp Asp Tyr Ser Asn Thr His Ser Thr
Arg Tyr Val Thr Val Lys Asp Gln Trp His Ser Arg Gly Ser Thr Trp
Leu Tyr Arg Glu Thr Cys Asn Leu Asn Cys Met Val Thr Ile Thr Thr
Ala Arg Ser Lys Tyr Pro Tyr His Phe Phe Ala Thr Ser Thr Gly Asp
Val Val Asp Ile Ser Pro Phe Tyr Asn Gly Thr Asn Arg Asn Ala Ser
20 Tyr Phe Gly Glu Asn Ala Asp Lys Phe Phe Ile Phe Pro Asn Tyr Thr
Ile Val Ser Asp Phe Gly Arg Pro Asn Ser Ala Leu Glu Thr His Arg
Leu Val Ala Phe Leu Glu Arg Ala Asp Ser Val Ile Ser Trp Asp Ile
Gln Asp Glu Lys Asn Val Thr Cys Gln Leu Thr Phe Trp Glu Ala Ser
Glu Arg Thr Ile Arg Ser Glu Ala Glu Asp Ser Tyr His Phe Ser Ser
25 Ala Lys Met Thr Ala Thr Phe Leu Ser Lys Lys Gln Glu Val Asn Met
Ser Asp Ser Ala Leu Asp Cys Val Arg Asp Glu Ala Ile Asn Lys Leu
Gln Gln Ile Phe Asn Thr Ser Tyr Asn Gln Thr Tyr Glu Lys Tyr Gly
Asn Val Ser Val Phe Glu Thr Thr Gly Gly Leu Val Val Phe Trp Gln
Gly Ile Lys Gln Lys Ser Leu Val Glu Leu Glu Arg Leu Ala Asn Arg
30 Ser Ser Leu Asn Leu Thr His Asn Arg Thr Lys Arg Ser Thr Asp Gly
Asn Asn Ala Thr His Leu Ser Asn Met Glu Ser Val His Asn Leu Val
Tyr Ala Gln Leu Gln Phe Thr Tyr Asp Thr Leu Arg Gly Tyr Ile Asn
Arg Ala Leu Ala Gln Ile Ala Glu Ala Trp Cys Val Asp Gln Arg Arg
Thr Leu Glu Val Phe Lys Glu Leu Ser Lys Ile Asn Pro Ser Ala Ile
35 Leu Ser Ala Ile Tyr Asn Lys Pro Ile Ala Ala Arg Phe Met Gly Asp
Val Leu Gly Leu Ala Ser Cys Val Thr Ile Asn Gln Thr Ser Val Lys
Val Leu Arg Asp Met Asn Val Lys Glu Ser Pro Gly Arg Cys Tyr Ser
Arg Pro Val Val Ile Phe Asn Phe Ala Asn Ser Ser Tyr Val Gln Tyr
Gly Gln Leu Gly Glu Asp Asn Glu Ile Leu Leu Gly Asn His Arg Thr
40 Glu Glu Cys Gln Leu Pro Ser Leu Lys Ile Phe Ile Ala Gly Asn Ser
Ala Tyr Glu Tyr Val Asp Tyr Leu Phe Lys Arg Met Ile Asp Leu Ser
Ser Ile Ser Thr Val Asp Ser Met Ile Ala Leu Asp Ile Asp Pro Leu
Glu Asn Thr Asp Phe Arg Val Leu Glu Leu Tyr Ser Gln Lys Glu Leu
Arg Ser Ser Asn Val Phe Asp Leu Glu Glu Ile Met Arg Glu Phe Asn
45 Ser Tyr Lys Gln Arg Val Lys Tyr Val Glu Asp Lys Val Val Asp Pro
Leu Pro Pro Tyr Leu Lys Gly Leu Asp Asp Leu Met Ser Gly Leu Gly
Ala Ala Gly Lys Ala Val Gly Val Ala Ile Gly Ala Val Gly Gly Ala
Val Ala Ser Val Val Glu Gly Val Ala Thr Phe Leu Lys Asn Pro Phe
Gly Ala Phe Thr Ile Ile Leu Val Ala Ile Ala Val Val Ile Ile Ile
50 Tyr Leu Ile Tyr Thr Arg Gln Arg Arg Leu Cys Met Gln Pro Leu Gln
Asn Leu Phe Pro Tyr Leu Val Ser Ala Asp Gly Thr Thr Val Thr Ser
Gly Asn Thr Lys Asp Thr Ser Leu Gln Ala Pro Pro Ser Tyr Glu Glu
Ser Val Tyr Asn Ser Gly Arg Lys Gly Pro Gly Pro Pro Ser Ser Asp
Ala Ser Thr Ala Ala Pro Pro Tyr Thr Asn Glu Gln Ala Tyr Gln Met
55 Leu Leu Ala Leu Val Arg Leu Asp Ala Glu Gln Arg Ala Gln Gln Asn
Gly Thr Asp Ser Leu Asp Gly Gln Thr Gly Thr Gln Asp Lys Gly Gln
Lys Pro Asn Leu Leu Asp Arg Leu Arg His Arg Lys Asn Gly Tyr Arg
His Leu Lys Asp Ser Asp Glu Glu Glu Asn Val

60 SEQ ID NO:20:

Met Ser Lys Met Val Val Leu Phe Leu Ala Val Phe Leu Met Asn Ser
Val Leu Met Ile Tyr Cys Asp Pro Asp His Tyr Ile Arg Ala Gly Tyr
Asn His Lys Tyr Pro Phe Arg Ile Cys Ser Ile Ala Lys Gly Thr Asp

```

Leu Met Arg Phe Asp Arg Asp Ile Ser Cys Ser Pro Tyr Lys Ser Asn
Ala Lys Met Ser Glu Gly Phe Phe Ile Ile Tyr Lys Thr Asn Ile Glu
Thr Tyr Thr Phe Pro Val Arg Thr Tyr Lys Lys Glu Leu Thr Phe Gln
Ser Ser Tyr Arg Asp Val Gly Val Val Tyr Phe Leu Asp Arg Thr Val
5 Met Gly Leu Ala Met Pro Val Tyr Glu Ala Asn Leu Val Asn Ser His
Ala Gln Cys Tyr Ser Ala Val Ala Met Lys Arg Pro Asp Gly Thr Val
Phe Ser Ala Phe His Glu Asp Asn Asn Lys Asn Asn Thr Leu Asn Leu
Phe Pro Leu Asn Phe Lys Ser Ile Thr Asn Lys Arg Phe Ile Thr Thr
10 Lys Glu Pro Tyr Phe Ala Arg Gly Pro Leu Trp Leu Tyr Ser Thr Ser
Thr Ser Leu Asn Cys Ile Val Thr Glu Ala Thr Ala Lys Ala Lys Tyr
Pro Phe Ser Tyr Phe Ala Leu Thr Thr Gly Glu Ile Val Glu Gly Ser
Pro Phe Phe Asn Gly Ser Asn Gly Lys His Phe Ala Glu Pro Leu Glu
Lys Leu Thr Ile Leu Glu Asn Tyr Thr Met Ile Glu Asp Leu Met Asn
15 Gly Met Asn Gly Ala Thr Thr Leu Val Arg Lys Ile Ala Phe Leu Glu
Lys Ala Asp Thr Leu Phe Ser Trp Glu Ile Lys Glu Glu Asn Glu Ser
Val Cys Met Leu Lys His Trp Thr Thr Val Thr His Gly Leu Arg Ala
Glu Thr Asp Glu Thr Tyr His Phe Ile Ser Lys Glu Leu Thr Ala Ala
Phe Val Ala Pro Lys Glu Ser Leu Asn Leu Thr Asp Pro Lys Gln Thr
20 Cys Ile Lys Asp Glu Phe Glu Lys Ile Ile Asn Glu Val Tyr Met Ser
Asp Tyr Asn Asp Thr Tyr Ser Met Asn Gly Ser Tyr Gln Ile Phe Lys
Thr Thr Gly Asp Leu Ile Leu Ile Trp Gln Pro Leu Val Gln Lys Ser
Leu Met Phe Leu Glu Gln Gly Ser Glu Lys Ile Arg Arg Arg Arg Asp
Val Val Asp Val Lys Ser Arg His Asp Ile Leu Tyr Val Gln Leu Gln
25 Tyr Leu Tyr Asp Thr Leu Lys Asp Tyr Ile Asn Asp Ala Leu Gly Asn
Leu Ala Glu Ser Trp Cys Leu Asp Gln Lys Arg Thr Ile Thr Met Leu
His Glu Leu Ser Lys Ile Ser Pro Ser Ser Ile Val Ser Glu Val Tyr
Gly Arg Pro Ile Ser Ala Gln Leu His Gly Asp Val Leu Ala Ile Ser
Lys Cys Ile Glu Val Asn Gln Ser Ser Val Gln Leu His Lys Ser Met
30 Arg Val Val Asp Ala Lys Gly Val Arg Ser Glu Thr Met Cys Tyr Asn
Arg Pro Leu Val Thr Phe Ser Phe Val Asn Ser Thr Pro Glu Val Val
Pro Gly Gln Leu Gly Leu Asp Asn Glu Ile Leu Leu Gly Asp His Arg
Thr Glu Glu Cys Glu Ile Pro Ser Thr Lys Ile Phe Leu Ser Gly Asn
His Ala His Val Tyr Thr Asp Tyr Thr His Thr Asn Ser Thr Pro Ile
35 Glu Asp Ile Glu Val Leu Asp Ala Phe Ile Arg Leu Lys Ile Asp Pro
Leu Glu Asn Ala Asp Phe Lys Val Leu Asp Leu Tyr Ser Pro Asp Glu
Leu Ser Arg Ala Asn Val Phe Asp Leu Glu Asn Ile Leu Arg Glu Tyr
Asn Ser Tyr Lys Ser Ala Leu Tyr Thr Ile Glu Ala Lys Ile Ala Thr
Asn Thr Pro Ser Tyr Val Asn Gly Ile Asn Ser Phe Leu Gln Gly Leu
Gly Ala Ile Gly Thr Gly Leu Gly Ser Val Ile Ser Val Thr Ala Gly
40 Ala Leu Gly Asp Ile Val Gly Gly Val Val Ser Phe Leu Lys Asn Pro
Phe Gly Gly Gly Leu Met Leu Ile Leu Ala Ile Val Val Val Val Ile
Ile Ile Val Val Phe Val Arg Gln Arg His Val Leu Ser Lys Pro Ile
Asp Met Met Phe Pro Tyr Ala Thr Asn Pro Val Thr Thr Val Ser Ser
Val Thr Gly Thr Thr Val Val Lys Thr Pro Ser Val Lys Asp Val Asp
45 Gly Gly Thr Ser Val Ala Val Ser Glu Lys Glu Glu Gly Met Ala Asp
Val Ser Gly Gln Val Ser Asp Asp Glu Tyr Ser Gln Glu Ala Ala Leu
Lys Met Leu Lys Ala Ile Lys Ser Leu Asp Glu Ser Tyr Arg Arg Lys
Pro Ser Ser Ser Glu Ser His Ala Ser Lys Pro Ser Leu Ile Asp Arg
Ile Arg Tyr Arg Gly Tyr Lys Ser Val Asn Val Glu Glu Ala
50

```

SEQ ID NO:21:

```

Met Phe Val Thr Ala Val Val Ser Val Ser Pro Ser Ser Phe Tyr Glu
Ser Leu Gln Val Glu Pro Thr Gln Ser Glu Asp Ile Thr Arg Ser Ala
His Leu Gly Asp Gly Asp Glu Ile Arg Glu Ala Ile His Lys Ser Gln
55 Asp Ala Glu Thr Lys Pro Thr Phe Tyr Val Cys Pro Pro Pro Thr Gly
Ser Thr Ile Val Arg Leu Glu Pro Thr Arg Thr Cys Pro Asp Tyr His
Leu Gly Lys Asn Phe Thr Glu Gly Ile Ala Val Val Tyr Lys Glu Asn
Ile Ala Ala Tyr Lys Phe Lys Ala Thr Val Tyr Tyr Lys Asp Val Ile
Val Ser Thr Ala Trp Ala Gly Ser Ser Tyr Thr Gln Ile Thr Asn Arg
60 Tyr Ala Asp Arg Val Pro Ile Pro Val Ser Glu Ile Thr Asp Thr Ile
Asp Lys Phe Gly Lys Cys Ser Ser Lys Ala Thr Tyr Val Arg Asn Asn
His Lys Val Glu Ala Phe Asn Glu Asp Lys Asn Pro Gln Asp Met Pro
Leu Ile Ala Ser Lys Tyr Asn Ser Val Gly Ser Lys Ala Trp His Thr

```

Thr Asn Asp Thr Tyr Met Val Ala Gly Thr Pro Gly Thr Tyr Arg Thr
 Gly Thr Ser Val Asn Cys Ile Ile Glu Glu Val Glu Ala Arg Ser Ile
 Phe Pro Tyr Asp Ser Phe Gly Leu Ser Thr Gly Asp Ile Ile Tyr Met
 Ser Pro Phe Phe Gly Leu Arg Asp Gly Ala Tyr Arg Glu His Ser Asn
 5 Tyr Ala Met Asp Arg Phe His Gln Phe Glu Gly Tyr Arg Gln Arg Asp
 Leu Asp Thr Arg Ala Leu Leu Glu Pro Ala Ala Arg Asn Phe Leu Val
 Thr Pro His Leu Thr Val Gly Trp Asn Trp Lys Pro Lys Arg Thr Glu
 Val Cys Ser Leu Val Lys Trp Arg Glu Val Glu Asp Val Val Arg Asp
 Glu Tyr Ala His Asn Phe Arg Phe Thr Met Lys Thr Leu Ser Thr Thr
 10 Phe Ile Ser Glu Thr Asn Glu Phe Asn Leu Asn Gln Ile His Leu Ser
 Gln Cys Val Lys Glu Glu Ala Arg Ala Ile Ile Asn Arg Ile Tyr Thr
 Thr Arg Tyr Asn Ser Ser His Val Arg Thr Gly Asp Ile Gln Thr Tyr
 Leu Ala Arg Gly Gly Phe Val Val Val Phe Gln Pro Leu Leu Ser Asn
 Ser Leu Ala Arg Leu Tyr Leu Gln Glu Leu Val Arg Glu Asn Thr Asn
 15 His Ser Pro Gln Lys His Pro Thr Arg Asn Thr Arg Ser Arg Arg Ser
 Val Pro Val Glu Leu Arg Ala Asn Arg Thr Ile Thr Thr Thr Ser Ser
 Val Glu Phe Ala Met Leu Gln Phe Thr Tyr Asp His Ile Gln Glu His
 Val Asn Glu Met Leu Ala Arg Ile Ser Ser Ser Trp Cys Gln Leu Gln
 Asn Arg Glu Arg Ala Leu Trp Ser Gly Leu Phe Pro Ile Asn Pro Ser
 20 Ala Leu Ala Ser Thr Ile Leu Asp Gln Arg Val Lys Ala Arg Ile Leu
 Gly Asp Val Ile Ser Val Ser Asn Cys Pro Glu Leu Gly Ser Asp Thr
 Arg Ile Ile Leu Gln Asn Ser Met Arg Val Ser Gly Ser Thr Thr Arg
 Cys Tyr Ser Arg Pro Leu Ile Ser Ile Val Ser Leu Asn Gly Ser Gly
 Thr Val Glu Gly Gln Leu Gly Thr Asp Asn Glu Leu Ile Met Ser Arg
 25 Asp Leu Leu Glu Pro Cys Val Ala Asn His Lys Arg Tyr Phe Leu Phe
 Gly His His Tyr Val Tyr Tyr Glu Asp Tyr Arg Tyr Val Arg Glu Ile
 Ala Val His Asp Val Gly Met Ile Ser Thr Tyr Val Asp Leu Asn Leu
 Thr Leu Leu Lys Asp Arg Glu Phe Met Pro Leu Gln Val Tyr Thr Arg
 Asp Glu Leu Arg Asp Thr Gly Leu Leu Asp Tyr Ser Glu Ile Gln Arg
 30 Arg Asn Gln Met His Ser Leu Arg Phe Tyr Asp Ile Asp Lys Val Val
 Gln Tyr Asp Ser Gly Thr Ala Ile Met Gln Gly Met Ala Gln Phe Phe
 Gln Gly Leu Gly Thr Ala Gly Gln Ala Val Gly His Val Val Leu Gly
 Ala Thr Gly Ala Leu Leu Ser Thr Val His Gly Phe Thr Thr Phe Leu
 Ser Asn Pro Phe Gly Ala Leu Ala Val Gly Leu Leu Val Leu Ala Gly
 35 Leu Val Ala Ala Phe Phe Ala Tyr Arg Tyr Val Leu Lys Leu Lys Thr
 Ser Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys Gly Leu Lys Gln
 Leu Pro Glu Gly Met Asp Pro Phe Ala Glu Lys Pro Asn Ala Thr Asp
 Thr Pro Ile Glu Glu Ile Gly Asp Ser Gln Asn Thr Glu Pro Ser Val
 Asn Ser Gly Phe Asp Pro Asp Lys Phe Arg Glu Ala Gln Glu Met Ile
 40 Lys Tyr Met Thr Leu Val Ser Ala Ala Glu Arg Gln Glu Ser Lys Ala
 Arg Lys Lys Asn Lys Thr Ser Ala Leu Leu Thr Ser Arg Leu Thr Gly
 Leu Ala Leu Arg Asn Arg Arg Gly Tyr Ser Arg Val Arg Thr Glu Asn
 Val Thr Gly Val

45 SEQ ID NO:22:

Met Arg Gln Gly Ala Ala Arg Gly Cys Arg Trp Phe Val Val Trp Ala
 Leu Leu Gly Leu Thr Leu Gly Val Leu Val Ala Ser Ala Ala Pro Ser
 Ser Pro Gly Thr Pro Gly Val Ala Ala Ala Thr Gln Ala Ala Asn Gly
 Gly Pro Ala Thr Pro Ala Pro Pro Ala Pro Gly Pro Ala Pro Thr Gly
 50 Asp Thr Lys Pro Lys Lys Asn Lys Lys Pro Lys Asn Pro Pro Pro Pro
 Arg Pro Ala Gly Asp Asn Ala Thr Val Ala Ala Gly His Ala Thr Leu
 Arg Glu His Leu Arg Asp Ile Lys Ala Glu Asn Thr Asp Ala Asn Phe
 Tyr Val Cys Pro Pro Pro Thr Gly Ala Thr Val Val Gln Phe Glu Gln
 Pro Arg Arg Cys Pro Thr Arg Pro Glu Gly Gln Asn Tyr Thr Glu Gly
 55 Ile Ala Val Val Phe Lys Glu Asn Ile Ala Pro Tyr Lys Phe Lys Ala
 Thr Met Tyr Tyr Lys Asp Val Thr Val Ser Gln Val Trp Phe Gly His
 Arg Tyr Ser Gln Phe Met Gly Ile Phe Glu Asp Arg Ala Pro Val Pro
 Phe Glu Glu Val Ile Asp Lys Ile Asn Ala Lys Gly Val Cys Arg Ser
 Thr Ala Lys Tyr Val Arg Asn Asn Leu Glu Thr Thr Ala Phe His Arg
 60 Asp Asp His Glu Thr Asp Met Glu Leu Lys Pro Ala Asn Ala Ala Thr
 Arg Thr Ser Arg Gly Trp His Thr Thr Asp Leu Lys Tyr Asn Pro Ser
 Arg Val Glu Ala Phe His Arg Tyr Gly Thr Thr Val Asn Cys Ile Val
 Glu Glu Val Asp Ala Arg Ser Val Tyr Pro Tyr Asp Glu Phe Val Leu



Ala Thr Gly Asp Phe Val Tyr Met Ser Pro Phe Tyr Gly Tyr Arg Glu
 Gly Ser His Thr Glu His Thr Ser Tyr Ala Ala Asp Arg Phe Lys Gln
 Val Asp Gly Phe Tyr Ala Arg Asp Leu Thr Thr Lys Ala Arg Ala Thr
 5 Ala Pro Thr Thr Arg Asn Leu Leu Thr Thr Pro Lys Phe Thr Val Ala
 Trp Asp Trp Val Pro Lys Arg Pro Ser Val Cys Thr Met Thr Lys Trp
 Gln Glu Val Asp Glu Met Leu Arg Ser Glu Tyr Gly Gly Ser Phe Arg
 Phe Ser Ser Asp Ala Ile Ser Thr Thr Phe Thr Thr Asn Leu Thr Glu
 Tyr Pro Leu Ser Arg Val Asp Leu Gly Asp Cys Ile Gly Lys Asp Ala
 Arg Asp Ala Met Asp Arg Ile Phe Ala Arg Arg Tyr Asn Ala Thr His
 10 Ile Lys Val Gly Gln Pro Gln Tyr Tyr Leu Ala Asn Gly Gly Phe Leu
 Ile Ala Tyr Gln Pro Leu Leu Ser Asn Thr Leu Ala Glu Leu Tyr Val
 Arg Glu His Leu Arg Glu Gln Ser Arg Lys Pro Pro Asn Pro Thr Pro
 Pro Pro Pro Gly Ala Ser Ala Asn Ala Ser Val Glu Arg Ile Lys Thr
 Thr Ser Ser Ile Glu Phe Ala Arg Leu Gln Phe Thr Tyr Asn His Ile
 15 Gln Arg His Val Asn Asp Met Leu Gly Arg Val Ala Ile Ala Trp Cys
 Glu Leu Gln Asn His Glu Leu Thr Leu Trp Asn Glu Ala Arg Lys Leu
 Asn Pro Asn Ala Ile Ala Ser Ala Thr Val Gly Arg Arg Val Ser Ala
 Arg Met Leu Gly Asp Val Met Ala Val Ser Thr Cys Val Pro Val Ala
 Ala Asp Asn Val Ile Val Gln Asn Ser Met Arg Ile Ser Ser Arg Pro
 20 Gly Ala Cys Tyr Ser Arg Pro Leu Val Ser Phe Arg Tyr Glu Asp Gln
 Gly Pro Leu Val Glu Gly Gln Val Gly Glu Asn Asn Glu Leu Arg Leu
 Thr Arg Asp Ala Ile Glu Pro Cys Thr Val Gly His Arg Arg Tyr Phe
 Thr Phe Gly Gly Tyr Val Tyr Phe Glu Glu Tyr Ala Tyr Ser His
 Gln Leu Ser Arg Ala Asp Ile Thr Thr Val Ser Thr Phe Ile Asp Leu
 25 Asn Ile Thr Met Leu Glu Asp His Glu Phe Val Pro Leu Glu Val Tyr
 Thr Arg His Glu Ile Lys Asp Ser Gly Leu Leu Asp Tyr Thr Glu Val
 Gln Arg Arg Asn Gln Leu His Asp Leu Arg Phe Ala Asp Ile Asp Thr
 Val Ile His Ala Asp Ala Asn Ala Ala Met Phe Ala Gly Leu Gly Ala
 Phe Phe Glu Gly Met Gly Asp Leu Gly Arg Ala Val Gly Lys Val Val
 30 Met Gly Ile Val Gly Gly Val Val Ser Ala Val Ser Gly Val Ser Ser
 Phe Met Ser Asn Pro Phe Gly Ala Leu Ala Val Gly Leu Leu Val Leu
 Ala Gly Leu Ala Ala Ala Phe Phe Ala Phe Arg Tyr Val Met Arg Leu
 Gln Ser Asn Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys Glu Leu
 Lys Asn Pro Thr Asn Pro Asp Ala Ser Gly Glu Gly Glu Glu Gly Gly
 35 Asp Phe Asp Glu Ala Lys Leu Ala Glu Ala Arg Glu Met Ile Arg Tyr
 Met Ala Leu Val Ser Ala Met Glu Arg Thr Glu His Lys Ala Lys Lys
 Lys Gly Thr Ser Ala Leu Leu Ser Ala Lys Val Thr Asp Met Val Met
 Arg Lys Arg Arg Asn Thr Asn Tyr Thr Gln Val Pro Asn Lys Asp Gly
 40 Asp Ala Asp Glu Asp Asp Leu

SEQ ID NO:23:

Met Arg Pro Arg Gly Thr Pro Pro Ser Phe Leu Pro Leu Pro Val Leu
 Leu Ala Leu Ala Val Ile Ala Ala Ala Gly Arg Ala Ala Pro Ala Ala
 45 Ala Ala Ala Pro Thr Ala Asp Pro Ala Ala Thr Pro Ala Leu Pro Glu
 Asp Glu Glu Val Pro Asp Glu Asp Gly Glu Gly Val Ala Thr Pro Ala
 Pro Ala Ala Asn Ala Ser Val Glu Ala Gly Arg Ala Thr Leu Arg Glu
 Asp Leu Arg Glu Ile Lys Ala Arg Asp Gly Asp Ala Thr Phe Tyr Val
 Cys Pro Pro Pro Thr Gly Ala Thr Val Val Gln Phe Glu Gln Pro Arg
 Pro Cys Pro Arg Ala Pro Asp Gly Gln Asn Tyr Thr Glu Gly Ile Ala
 50 Val Val Phe Lys Glu Asn Ile Ala Pro Tyr Lys Phe Lys Ala Thr Met
 Tyr Tyr Lys Asp Val Thr Val Ser Gln Val Trp Phe Gly His Arg Tyr
 Ser Gln Phe Met Gly Ile Phe Glu Asp Arg Ala Pro Val Pro Phe Glu
 Glu Val Met Asp Lys Ile Asn Ala Lys Gly Val Cys Arg Ser Thr Ala
 Lys Tyr Val Arg Asn Asn Met Glu Ser Thr Ala Phe His Arg Asp Asp
 55 His Glu Ser Asp Met Ala Leu Lys Pro Ala Lys Ala Ala Thr Arg Thr
 Ser Arg Gly Trp His Thr Thr Asp Leu Lys Tyr Asn Pro Ala Arg Val
 Glu Ala Phe His Arg Tyr Gly Thr Thr Val Asn Cys Ile Val Glu Glu
 Val Glu Ala Arg Ser Val Tyr Pro Tyr Asp Glu Phe Val Leu Ala Thr
 Gly Asp Phe Val Tyr Met Ser Pro Phe Tyr Gly Tyr Arg Asp Gly Ser
 60 His Gly Glu His Thr Ala Tyr Ala Ala Asp Arg Phe Arg Gln Val Asp
 Gly Tyr Tyr Glu Arg Asp Leu Ser Thr Gly Arg Arg Ala Ala Ala Pro
 Val Thr Arg Asn Leu Leu Thr Thr Pro Lys Phe Thr Val Gly Trp Asp
 Trp Ala Pro Lys Arg Pro Ser Val Cys Thr Leu Thr Lys Trp Arg Glu

Val Asp Glu Met Leu Arg Ala Glu Tyr Gly Pro Ser Phe Arg Phe Ser
 Ser Ala Ala Leu Ser Thr Thr Phe Thr Ala Asn Arg Thr Glu Tyr Ala
 Leu Ser Arg Val Asp Leu Ala Asp Cys Val Gly Arg Glu Ala Arg Glu
 Ala Val Asp Arg Ile Phe Leu Arg Arg Tyr Asn Gly Thr His Val Lys
 5 Val Gly Gln Val Gln Tyr Tyr Leu Ala Thr Gly Gly Phe Leu Ile Ala
 Tyr Gln Pro Leu Leu Ser Asn Ala Leu Val Glu Leu Tyr Val Arg Glu
 Leu Val Arg Glu Gln Thr Arg Arg Pro Ala Gly Gly Asp Pro Gly Glu
 Ala Ala Thr Pro Gly Pro Ser Val Asp Pro Pro Ser Val Glu Arg Ile
 Lys Thr Thr Ser Ser Val Glu Phe Ala Arg Leu Gln Phe Thr Tyr Asp
 10 His Ile Gln Arg His Val Asn Asp Met Leu Gly Arg Ile Ala Thr Ala
 Trp Cys Glu Leu Gln Asn Arg Glu Leu Thr Leu Trp Asn Glu Ala Arg
 Arg Leu Asn Pro Gly Ala Ile Ala Ser Ala Thr Val Gly Arg Arg Val
 Ser Ala Arg Met Leu Gly Asp Val Met Ala Val Ser Thr Cys Val Pro
 Val Ala Pro Asp Asn Val Ile Met Gln Asn Ser Ile Gly Val Ala Ala
 15 Arg Pro Gly Thr Cys Tyr Ser Arg Pro Leu Val Ser Phe Arg Tyr Glu
 Ala Asp Gly Pro Leu Val Glu Gly Gln Leu Gly Glu Asp Asn Glu Ile
 Arg Leu Glu Arg Asp Ala Leu Glu Pro Cys Thr Val Gly His Arg Arg
 Tyr Phe Thr Phe Gly Ala Gly Tyr Val Tyr Phe Glu Glu Tyr Ala Tyr
 Ser His Gln Leu Gly Arg Ala Asp Val Thr Thr Val Ser Thr Phe Ile
 20 Asn Leu Asn Leu Thr Met Leu Glu Asp His Glu Phe Val Pro Leu Glu
 Val Tyr Thr Arg Gln Glu Ile Lys Asp Ser Gly Leu Leu Asp Tyr Thr
 Glu Val Gln Arg Arg Asn Gln Leu His Ala Leu Arg Phe Ala Asp Ile
 Asp Thr Val Ile Lys Ala Asp Ala His Ala Ala Leu Phe Ala Gly Leu
 Tyr Ser Phe Phe Glu Gly Leu Gly Asp Val Gly Arg Ala Val Gly Lys
 25 Val Val Met Gly Ile Val Gly Gly Val Val Ser Ala Val Ser Gly Val
 Ser Ser Phe Leu Ser Asn Pro Phe Gly Ala Leu Ala Val Gly Leu Leu
 Val Leu Ala Gly Leu Ala Ala Ala Phe Phe Ala Phe Arg Tyr Val Met
 Arg Leu Gln Arg Asn Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys
 Glu Leu Lys Ser Asp Gly Ala Pro Leu Ala Gly Gly Gly Glu Asp Gly
 30 Ala Glu Asp Phe Asp Glu Ala Lys Leu Ala Gln Ala Arg Glu Met Ile
 Arg Tyr Met Ala Leu Val Ser Ala Met Glu Arg Thr Glu His Lys Ala
 Arg Lys Lys Gly Thr Ser Ala Leu Leu Ser Ala Lys Val Thr Asp Ala
 Val Met Arg Lys Arg Ala Arg Pro Arg Tyr Ser Pro Leu Arg Asp Thr
 35 Asp Glu Glu Glu Leu

SEQ ID NO:24:
 GCTGTTCAAGA TTTGACTTAG AYMANNMONTG YCC 33

 SEQ ID NO:25:
 40 GTGTACAAGA AGAACATCGT GCCNTAYATN TTYAA 35

 SEQ ID NO:26:
 GTGTACAAGA AGAACATCGT GCC 23

 SEQ ID NO:27:
 45 AACATGTCTA CAATCTCACA RTTNACNGTN GT 32

 SEQ ID NO:28:
 AACATGTCTA CAATCTCACA 20

 SEQ ID NO:29:
 AATAACCTCT TTACGGCCCA AATTCARTWY GCNTAYGA 38

 SEQ ID NO:30:
 55 CCAACGAGTG TGATGTCAGC CATTAYYGN AARCCNGT 38

 SEQ ID NO:31:
 CCAACGAGTG TGATGTCAGC C 21

 SEQ ID NO:32:
 60 TGCTACTCGC GACCTCTAGT CACCTTYAAR TTYRTNAA 38

 SEQ ID NO:33:

	TGCTACTCGC GACCTCTAGT CACC	24
	SEQ ID NO:34: ACCGGAGTAC AGTTCCACTG TYTTRAARTC DATRTT	36
5	SEQ ID NO:35: TGTCACCTTG ACATGAGGCC A	21
10	SEQ ID NO:36: TTTGACCTGG AGACTATGTT YMGNGARTAY AA	32
	SEQ ID NO:37: GCTCTGGGTG TAGTAGTTRT AYTCYCTRAA CAT	33
15	SEQ ID NO:38: TCTCGGAACA TGCTCTCCAG RTCRAAMACR TT	32
	SEQ ID NO:39: ACCTTCATCA AAAATCCCTT NGGNGGNATG YT	32
20	SEQ ID NO:40: TGGACTTACA GGAICTCGAAC NACNGTNAAY TG	32

说明书附图

```

V Y K K N I V P Y I F K V R R Y I K I A
RFHV GTGTACAAGAAGAACATCGTGCCGTACATTTTCAAGGTACGCAGGTACATAAAAATAGCA 60
*****
KSHV GTGTACAAGAAGAACATCGTGCCGTATATTTTAAAGGTGCGGCGCTATAGGAAAATTGCC 60
V Y K K N I V P Y I F K V R R Y R K I A
(NIVPA>) gtgtacaagaagaacatcgtgcontayatnttyaa

(<GMTEB) gcgccatactgtcttcgctcgt
(<AAITB) tcgtcgttagtgtttgtttat
(GMTEA>) gcggtatgacagaagcagcaa
(KYEIA>) aacaaatatgagatc-
T S V T V Y R G M T E A A I T N K Y E I
RFHV ACATCTGTCACGGTCTACCGCGGTATGACAGAAGCAGCAATCACAACAAATATGAGATC 120
**
KSHV ACCTCTGTCACGGTCTACAGGGGCTTGACAGAGTCCGCCATCACCACAAAGTATGAACTC 120
T S V T V Y R G L T E S A I T N K Y E L
cgaactgtctcaggcggtagt (<GLTEB)
(<TNKYB) ggtagtggttgcatacttg
gggcttgacagagtcgcoat (GLTEA>)
(YELPA>) acaagtatgaactc-

-cccagg (KYEIA>)
P R P V P L Y E I S E M D S T Y Q C F S
RFHV CCCAGGCCCGTGCCTCTCTACGAGATCAGTCACATGGACAGCACCTACCAGTGCTTTAGT 180
**
KSHV CCGAGACCCGTGCCACTCTATGAGATAAGCCACATGGACAGCACCTATCAGTGCTTTAGT 180
P R P V P L Y E I S E M D S T Y Q C F S
-ccgagac (YELPA>)
agaccggtgccactctatgarathagycayatgga (SHMDA>)
(<CFSSB) acraarter-

(<TDRDB) atgcaagtggctagccctact
S M K I V V N G V E N T F T D R D D V N
RFHV TCCATGAAAATTGTAGTGAACGGAGTCGAAAATACGTTACCGATCGGGATGACGTAAC 240
*****
KSHV TCCATGAAGGTAAATGTCAACGGGGTAGAAAACACATTTACTGACAGAGACGATGTTAAC 240
S M K V N V N G V E N T F T D R D D V N
cttcatttacagttgcccc (VNVNB)
(<TFTDB) gtgtaaatgactgtctctgct
-agntacttctaacaacacttg (<CFSSB)
(ENTFA>) gtcaacggagtagaraayacnttyacnga

```

图 1a

```

(<VEGLB) gcagcttccagattgactgtt
(VEGLA) cccgtcgaaggtctaactgac (PVLIA) agc-
RFHV ATVFLQPVEGLTDNIQRYFS 300
AAAACCGTATTTCTCCAGCCCGTGAAGGTCTAACTGACAACATACAAAGATACTTTAGC
* * * * *
KSHV ACCACAGTATTCCTCCAACCAGTAGAGGGCTTACGGATAACATTCAAGGTACTTTAGC 300
ATVFLQPVEGLTDNIQRYFS
ccacagttattcctccaaccag (TVFLA) (SQPVA) ggtacttttagc-
(<DNIQB) tgnctrtrtaagtttccatgaaatcg-

-caaccagtactgtactct (PVLIA)
RFHV QPVLYS EPGWFPGIYRV GTT 360
CAACCAGTACTGTACTCTGAACCCGGATGGTCCCAGGTATCTACAGGGTTGGGACAACA
* * * * *
KSHV CAGCCGGTCATCTACGCGGAACCCGGCTGGTTCCCAGCATATACAGAGTTAGGACAACA 360
QPVLYS EPGWFPGIYRV RTT
-cagccgggtca (SQPVA) (<TVNCB) tgntgn-
-gtccgggtca (<DNIQB)

V N C E I V D M
RFHV GTAACTGTGAGATTGTAGACATGTT 386
* * * * *
KSHV GTCAACTGTGAGATTGTAGACATGTT 386
V N C E I V D M
-canttracactctaacaatctgtacaa (<TVNCB)

```

图 1b

糖蛋白 B
编码区

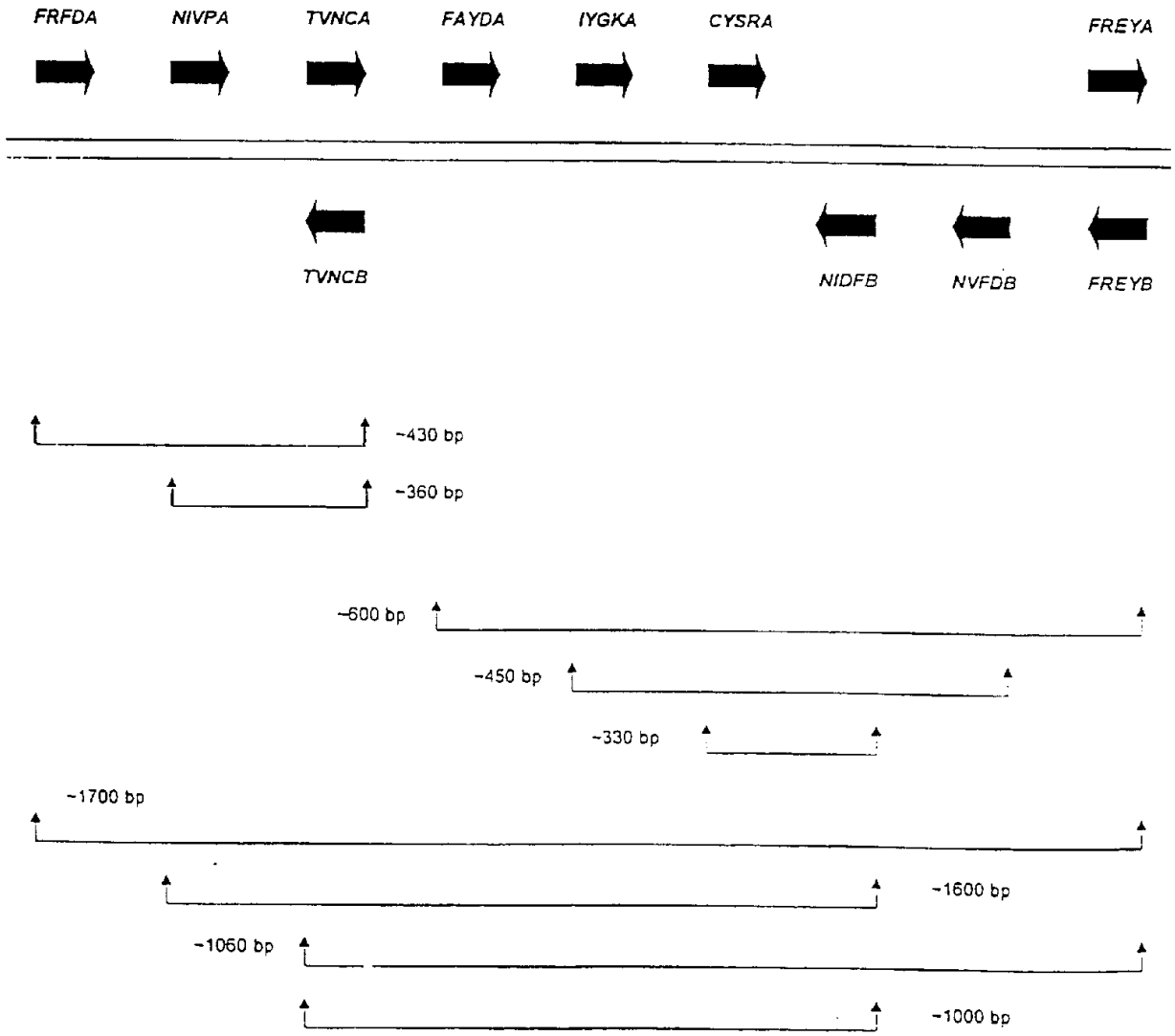


图 2

信号肽区

KSHV
HVS
bHV4
mHV68
EBV
hHV6
hHSV1

MTPRSRLA
MVRPKHL
MYKTIIE
MYPTVKSM
MTRRRVLS
MSKMWVEE
MROGAARG

KSHV **TLGTVIIILVCFCAGAAHS**RGDTFQTSSSPTPGSSSKAPTKEE-----
HVS **LTILSESTACGQITPI**TAVEKNKTOAI-----
bHV4 **EALIKVCSFNQITHTS**ITSPSISSTTSSTTSTSKPSNTTSTN-----
mHV68 **RVAHILTNIITLCCILCH**ILYVCOPTTLROPSDMTPAQDAPTETPPPLSTNTN-----
EBV **VVVLAAALACRLGAQTPEQP**APPATTVQPTATRO-----
hHV6 **LAVFLMNSVLMIYC**DPDHYIRAGYN-----
hHSV1 **GRWFVWALLGLGLGLV**ASAAPSSPGTGVAAATQAANGPATPAPPAGPAPTGDTPKPKNK

KSHV -----ASGPKSVDFYQFRVCSASIT-GELFRFNLEQTC
HVS -----YQEYFKYRVCSASTT-GELFRFDLDRTC
bHV4 SSLAASPONTSTSKPSTDNQGTSTPTIPTVTDOTASK-NFYKYRVCSASSSSGELFRFDLDRTC
mHV68 RTHLYVCOPTTLROPSDMTPAQDAPTETPPPLSTNTNRGFYFRVCGVAAT-GETFRFDLDRTC
EBV QLSVVVLLAALACRLGAQTPEQPAPPATTVQPTATROQTSFPFRVCELSH-GDLFRFSSDIQC
hHV6 -----HKYPPFRICSIAGK-TDLMRFRDRISC
hHSV1 KPKNPPPPRPAGDNATVAAGHATLREHLRDIKAENTDANFY---VCPPTG-ATVVQFEQPRRC
* * *

RFHVMn VRRYIKIATSVTVYRGM--TEAAITNKYETPRPVPLY
RFHVMm VRRYRKVATPVTLYRGM--TDAITNKYEIPRPVPLY
KSHV PDKDKY-HQEGILLVYKKNIVPHIFKVRRYRKIATSVTVYRGL--TESAITNKYELPRPVPLY
HVS PSTEDKV-HKEGILLVYKKNIVPYIFKVRRYKITTSVRIFNGWTREGVAITNKWELSRAPVKY
bHV4 PDKDKK-HVEGILLVYKKNIVPYIFKVRRYKATSVTVYRGM--SQAAVTNRDDISRAIPYN
mHV68 PSTQDKK-HVEGILLVYKKNIVPYIFKIRRYRKIITQLTIWRGL--TTSSVTGKFEMATOAEHW
EBV PSFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIVTNILYNGW--YADSVNRRHEEKFSVDSY
hHV6 SPYKSNAMSEGGFFIYKTNIEYTFPVRTYKKELTFOSSYRDV--GVVYFLDRVTMGLAMPVY
hHSV1 PTRPEGQNYTEGIAVVFKENIAPYKFKATMYKDVTVSQVWFGH-RYSQFMGIFEDRAPVPFEE
** * * * * *

RFHVMn EISHMDSTYQCFSSMKIVVNGVENTFTDRDDVNVKTVFLQPVEGLTDNIQRYFSQ--PVLYSEPG
RFHVMm EISHMDSTYQCFSSMKIVVNGVENTFTGRDDVNVKSVFLQPVEGLTDNIKRYFSQ--PLIYAEPG
KSHV EISHMDSTYQCFSSMKVNVNGVENTFTDRDDVNTTVFLQPVEGLTDNIQRYFSQ--PVIYAEPG
HV1 EIDIMDKTYOCHNCMQIEVNGMLNSYYDRDGNKTVDLKPV DGLTGAI TRYISQ--PKVFADPG
bHV4 EISMIDRTYHCFSAMATVINGILNTYIDRDNENKSVPLQPVAGLTENINRYFSQ--PLIYAEPG
mHV68 EVGDFDSIYOCYNSATMVVNNVRQVYVDRDGVNKT VNI R PV DGLTGNIQRYFSQ--PTLYSEPG
EBV ETDQMDTIYOCYNAVKMTKDGLTRVYVDRDGVNITVNLKPTGGLANGVRRYASQ--TELYDAPG
hHV6 EANLVNSHAQCYSAVAMKRPDGTVYVSAFHEDNNKNNTLNLFP LNFKSI TNKRFITTKEPYFARG
hHSV1 VIDKINAKGVC RSTAKYVRNLETTAFHRDDHETDMELK PANAATRTSRGWH--TTDLKYN-PS
*

图 3a

RFHVMn WFPGIYRVG
 RFHVMm WFPGIYRVR
 KSHV WFPGIYRVRTTVNCEIVDMIARSAEPYNYFVTSLGDTVEVSPFCYNES--CSTTPSNKN-GLS

HVS WLWGTYRTRTTVNCEIVDMFARSADPYTYFVTALGDTVEVSPFCVDNS--CPNAT---DVLS
 bHV4 WFPGIYRVRTTVNCEIVDMYARSVPEYTHFITALGDTIEISPFCH--NNSQCTTGNSTSRDATK
 mHV68 WMPGFYRVRTTVNCEIVDMVARSMOPYNYIATALGDSLELSPFQTFDNTSQCTAPKRA--DMRV
 EBV WLIWTYRTRTTVNCLITDMMAKSNPFDFFVTTTGQTVEMSPFYDGKNKETFHE-----RADS
 hHV6 P-LWLYSTSLNCEIVTEATAKAKYPFYFALTTGEIVEGSPFFNGSNGKHFAEPLK--LTIL
 hHSV1 RVEAFHRYGTTVNCEIVEVDARSVYYPDEFVLATGDFVYMSPFYGYREGSHTETSAAADRFRKQ
 * ** * * * **

KSHV VQVVLNHTVVITYSDRGTSPTQNRIFVETGAYTLWSASESKTTAVQPLALWKTFRPSIQTTHE

HVS VQIDLNHTVVVDYGNRATSQHKKRIFAHTLDYSVSWEAVNKSASVCSMVFWKSFQRAIQTEHDL
 bHV4 VWIEENHQTVDY-ERRGHPTKDKRIFLKDDEEYTI SWKAEDRERAI CDFVIWKTFRPRAIQTIHNE
 mHV68 REVK-NYKFVDYNNRGTAPAGQSRTFLETSPATYSWKATROTATCDLVHWKTFRPRAIQTAHEH
 EBV FHVRTNYKIVDYDNRGTNPOGERRAFLDKGTYYLSWKLENRTA-YCPLOHWQTFDSTIATETGK
 hHV6 ENYTMIEDLMNG-MNGATTLVRKIAFLEKADTLFSWEIKEENESVCMKHWTTVTHGLRAETDE
 hHSV1 -VDGFYARDLTKARATAPTRN--LLTPKFTVAWDVWPKRPSVCTMTKWQEVDEMLRSEGGG
 * * *

KSHV SFHFVANEITATFTAPLTP---VANFTDTYSCLTSDINTTLNA-SKAKLASTHVPNGTVQYFHT
 HVS TYHFIANEITAGFSTVKEP---LANFTSDYNC LMTHINTTLED-KIARVNNHTPNGTAEYYQT
 bHV4 SFHFVANEVTASFLTSNOEETELRGNTIELNCMNSTINETLEE-TVKKFNKSHIRDGEVKKYKT
 mHV68 SYHFVANEVTATFNTPLTE---VENFTSYSCVSDQINKTISE-YIQKLNNSYVASGKTOYFKT
 EBV SIHVFTEGTSFVTNTTV---GIELPDAFKCIEEQVNTMHEKYEAVQDRYTKGQEAITYFIT
 hHV6 TYHFISKELTAAAFVAPKES---LNLTDPKQTCIKDEFEKIINEVYMSDYNDTYSMNGSYQIFKT
 hHSV1 SFRFSSDAISTFTTNLTE--YPLSRVDLGDGIGKDARDAMDRIFARRYNATHIKVGGPQYYLA
 * *

KSHV TGGLYLWQPMSAINLTHAQGDSGNPTSSPPSPASPMTTASRRK-----

HVS EGGMILVWQPLIAIELEEAMLEATTSPVTPSAPTSSSRKRAI-----
 bHV4 NGGLFLIWQAMKPLNLSE-----HTNYTIERNNKTGNKSROK-----
 mHV68 DGNLYLIWQPLEHPEIEDIDEDSDPEP-TPAPPKSTRRKREAA-----
 EBV SGGLLLAWLPLTPRSLATVKNLTELTPPTSPPSSPSPAPSAARGSTPAAVLRRRRRDAGNAT
 hHV6 TGDLILIWQPLVQKSL--MFLEQGSEKIRRRRDVVDVKSROI-----
 hHSV1 NGGFLIAYQPLLSNTLAEIYVREHLREQSRKPPNPPTPPPGAS-----
 *

FAYDA>

KSHV RRSASTAAAGGGGSTDNLSYTLQFAYDKLRDGINQVLEELSRAWCREQVRDNLMWYELSKINP

HVS ----RSIRDVSAGSENNVFLSQIQYAYDKLRQSIINNVLLELAITWCREQVROTMMWYEIAKINP
 bHV4 ----RSVDTKTFQGAAGLSTAQVQYAYDHLRTSMNHILEELTKTWCREQKKNLMMWYELSKINP
 mHV68 -DNGNSTSEVSKGSENPLITAQIQFAYDKLTTSVNNVLEELSRAWCREQVRDNLMMWYELSKVNP
 EBV TPVPPTAPGKSLGTLNPNATVQIQFAYDSLRRQINRMLGDLARAWCLEQKRONMVLRELTKINP
 hHV6 -----LYVQLQLYLDTLKDYINDALGNLAESWCLDOKRTITMLHELKISIP
 hHSV1 ----ANASVERIKTSSIEFARLQTYNHIQRHVNDMLGRVAIAWCELONHETLWNEARKLNP
 * * * * *

图 3b

IYGKA> *CYSRA>*
 KSHV TSVMTAIYGRPVS AKFVGD AISVTEÇINVDQSSVNIHKSLRTNSK-----DVÇYARPLVTFKF
 HVS TSVMTAIYGKPVSRKALGDVISVTEÇINVDQSSVSIHKSLKTENN-----DICYSRPPVTFKF
 bHV4 VSVMAAIYGKPVAVKAMGDAFMVSEÇINVDQASVNIHKSMRTDDP-----KVCYSRPLVTFKF
 mHV68 TSVMSAIYGKPVAAARYVGD AISVTDÇIYVDQSSVNIHQSLRLQHD-----KTTÇYSRPRVTFKF
 EBV TTVMSSIYGKAVAAKRLGDVISVSOÇVPVNOATVTLRKS MRVPGS-----ETMÇYSRPLVSFSF
 hHV6 SSIYSEVYGRPI SAQLHGDVLAISKÇIEVNQSSVOLHKSMRVVDAKGVRS ETCYNRPLVTFSF
 hHSV1 NAIASATVGRRVSARMLGDVMAVSTÇVPVAADNVIQNSMRISSR-----PGACYSRPLVSFRY
 * * * * *

KSHV LNSS-NLFTGQLGARNE IILTNNQVETÇKDTÇEHYF ITRNETLVYKDYAYLRTINTD I STLNT
 HVS VNSS-QLFKGQLGARNE ILLSESLVENCHQNAETFF TAKNETYHFKNYVHVETLPVNNISTLDT
 bHV4 VNST-ATFRGQLGTRNE ILLTNTHVETÇRPTADHYFFVKNMTHYFKDYKFVKTMOTNNISTLDT
 mHV68 INST-DPLTGQLGPRKE IILSNTNIETÇKDESEHYFIVGEYIYYYKNYIFEELNLS SIATLDT
 EBV INDT-KTYEGQLGTDNE IFLTKKMTVEÇQATSQYYFQSGNEIHVYNDYHHFKTIELDGIATLOT
 hHV6 VNSTPEVVPGQLGLDNE ILLGDHRTEEÇEIPSTKIFLSGNHAHVYTDYTHNSTPIEDIEVLDA
 hHSV1 EDQG-PLVEGGVGENNELRLTRDAIEPÇTVGHRRYFTFGGGYVYFEEYAYSHQLSRADITTVST
 * * * * *

<NIDFB *<NVFDB* *<FREYB* *FREYA>*
 KSHV FIALNLSFIQNI DFKAI ELYSSAEKRLASSVFDLETMFREYNYYTHRLAGLREDLNTIDMNKE
 HVS FLALNLTFIENIDFKAVELYSSGERKLA-NVFDLETMFREYNYYAQSISGLRKDFD NSQRNDR
 bHV4 FLTLNLTFIENIDFKTVELYSETERKMA-SALDLETMFREYNYYTQKLASLREDLNTIDLNRD
 mHV68 FIALNISFIENIDFKTVELYSSTERKLASSVFDIESMFREYNYYTSLAGIKKDLDNATIDYNRD
 EBV FISLNTSLIENIDFASLELYSRDEQRAS-NVFDLEGIFREYNFQAQNIAGLRKDLDN AVSNGRN
 hHV6 FIRLKI DLENADFKVLDLYSPDELSRA-NVFDLENILREYNSYKSALYTI EAKI ---AANTP
 hHSV1 FIDLNITMLEDHEFVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDTVIHADANAAMV-
 * * * * *

跨膜区

KSHV RFVRDLSEIVADLGGIGKTVVNVASSVVTLCGSLVTGFIN-----FIKHP **GGMLMIIIVIA**
 HVS RIIQDFSEILADLGSIGKVI VNVASGAFSLFGGIVTGILN-----FIKNP **GGMFTFLLIGA**
 bHV4 RLVKDLSEMADLGDIGKVVNTFSGIVTVFGSIVGGFVS-----FFTNP **IGGVTIILLIV**
 mHV68 RLVDLSDMMADLGDIGRSVVNVSSVVTFFSSIVTGFIK-----FFTNP **GGIFILLIIGG**
 EBV QFVDGLGELMDSLGSVQSITNLVSTVGGLFSSLVSGFIS-----FFKNP **FGGMLILVLVAG**
 hHV6 SYVNGINSFLOGLGAIGTGLGSVISVTAGALGDIVGGVVS-----FLKNP **FGGMLLILAIIV**
 hHSV1 IHADANAAMFAGLGAF FEGMGDLGRAVGKVVMGIVGGVVS AVSGVSSFMSNP **FALAVGLLVLA**
 * * * * *

KSHV **IILII**FMLSRRTNTIAQAPVKMIYDPVDRRAPPSSGAP-----TREEIKNILLGMHQLQQE
 HVS **VII**LIVLLVRRTNMNSQAPIRMIYDPVEKSKSTVTP-----MEPETIKOILLGMHNMQQE
 bHV4 **VVE**VFI VSRRTNNMNEAPIKMIYPNIDKASEQENIQP-----LPGEEIKRILLGMHQLOQS
 mHV68 **TI**FLVVV LNRNSQFHDAPIKMLYPSVENYAARQAPPYSASPPAIDKEEIKRILLGMHQVHQE
 EBV **VV**ILVISLITRRTRQMSQOPVOMLYPGIDELAQHASGEGPGINP-ISKTELQAIMLALHEQNQE
 hHV6 **VVV**IIIVV FVRQRHVL SKPIDMMFPYATNPVTTVSSVTGTTVVKTPSVKDVDGGT SVAVSEKEE
 hHSV1 **GL**AAAFFAFRYVMRLQSNPMKALYPLTTKELKNPTNP DASGE GEGDFDEAKLAEAREMIRYM
 * *

图 3c

KSHV ERQKADDLKKSTPSVFQRTANG-LRQRLRGYKPLTQSLDISPETGE
HVS AYKKKEEQRAARPSIFRQAAETFLR-KRSGYKQISTEDKIV
bHV4 EHGKSEEEASHKPGFLQLLGDGLQLLRRRGYTR-LPTFDPSGNDTSETHQYV
mHV68 EKEAQQLTNSGPTLWQK-ATGFLRNRKGYSG-LPLEDESTSL
EBV OKRAAQRAAGPSVASRALQAARDRFPGLRRRYHDPETAAALLGEAETF
hHV6 ADVSGQVSDDEYSQEAALKMLKAIKSLDESRYR-KPSSSESHASKPSLIDRIRYRGYKSVNVEEA
hHSV1 ALVSAMERTEHKAKKKGTSALLSAKVTDMMVRKRRNTNYTQVPNKDGDADDDL

图 3d

EBV	GGCGACCTGTTCCGCTTCTCCTCGGACATCCAGTGTCCC
sHV1	GGAGAATTGTTTAGATTTGATTTAGACAGAACTTGTC
mHV68	GGGGAGACCTTCAGGTTTGATTTAGACAAAACATGCC
bHV4	GGAGAACTATTCAGATTTGACCTTGATCAGACATGTCCA

5'-gctgttcagatttgacttagaymanmctgycc-3'
FRFDA 256 倍 33 聚体

图 4

EBV	GAGGGCCTGTTGATGGTGTTTAAAGACAACATTATCCCTACTCGTTTAAAG
sHV1	GAAGGCATTCTTTTAGTGTACAAAAAATATAGTTCATATATCTTTAAA
mHV68	GAGGGCATCTTGCTCGTGTATAAGATCAACATCGTGCCCTACATCTTCAAA
bHV4	GAAGGCATCCTGCTGGTACTAAAAAAGAATATTGTCCCATACATCTTCAAA

5'-gtgtacaagaagaacatcgctgcentayatnttyaa-3'
NIVPA 64 倍 32 聚体

图 5

5'-gtgtacaagaagaacatcgctgcc-3'
NIVPASQ 23 聚体

EBV	ATATGGACTTACAGAACAAGAACTACCGTCAACTGCCTGATAACTGACATGATGGCCAAG
sHV1	TGGGGAACCTTACAGGACTCGAACTACCGTTAACTGTGAAATTGTAGACATGTTTGCTAGG
mHV68	CCTGGCTTTTATCGTGTTCGAACCACCGTTAACTGTGAAATTGTAGACATGGTGGCACGC
bHV4	CCAGGGATTTATAGAGTGAGAACAACGTAAATTGTGAGGTTGTTGACATGTATGCCCGC

5'-tggacttacaggactcgaacnacngtnaaytg-3'
TVNCA 128 倍 32 聚体

3'-tgntgncanttracactctaacatctgtacaa-5'
<TVNCB 128 倍 32 聚体

图 6

3'-acactctaacatctgtacaa-5'
<TVNCBSQ 20 聚体

EBV	AACAATCCCGCCACCGTCCAGATCCAATTTGCCTACGAC
sHV1	AATAATGTGTTTCTATCACAAATACAATATGCATATGAT
mHV68	AATCCGCTCATTACGGCCCAAATTC AATTTGCCTATGAC
bHV4	AAGGGCCTGTCCACTGCCAGGTTCAATATGCCTATGAC

5'-aataacctctttacggcccaaattcartwygcntayga-3'
 FAYDA 64 倍 38 聚体>

图 7

AATCCAACCACCGTCATGTCCAGCATCTACGGTAAGGCGGTG
 AATCCAACAAGTGTATGACAGCAATATATGGAAAACCTGTC
 AACCCTACGAGTGTGATGTCTGCCATTTATGGAAAGCCTGTC
 AACCAGTGAGTGTGATGGCAGCCATTTATGGGAAAACCTGTG

5'-ccaacgagtgatgtcagccatttayggnaarcngt-3'
 IYGKA 64 倍 38 聚体>

5'-ccaacgagtgatgtcagcc-3'
 IYGKASQ 21 聚体>

图 8

EBV	TGCTACTCGGCCCCCTGGTGTCTTCAGCTTTATCAACGAC
sHV1	TGCTATTCACGGCCTCCAGTIACATTTAAATTTGTAAACAGT
mHV68	TGCTACTCGAGACCTAGAGTCACCTTCAAATTTATAAACAGT
bHV4	TGTTACTCCAGACCCCTGGTCACATTTAAATTTGTGAATAGT

5'-tgctactcggacctctagtcaccttyaarttyrtnaa-3'
 CYSRA 64 倍 38 聚体>

5'-tgctactcggacctctagtcacc-3'
 CYSRASQ 24 聚体>

图 9

EBV
sHV1
mHV68
bHV4

'AACATTGACTTTGCCTCCCTGGAGCTGTACTCACGGGACGAAACAGCGT
AATATTGACTTTAAAGCTGTTGAATTGTATTCAAGTGGAGAGAGAAAAG
AATATCGACTTCAAACAGTAGAACTGTACTCCTCTACTGAAAGGAAA
AATATAGATTTCAAGACAGTGGAACTTTACAGTGAGACTGAAAGAAAAG

3'-ttrtadctraarttytgcaccttgacatgaggcca-5'
<NIDFB 48倍 36聚体>

3'-tgtcaccttgacatgaggcca-5'
<21聚体>

图 10

EBV
sHV1
mHV68
bHV4

AACGTCTTTGACCTGGAGGGCATCTCCGGGAGTACAACTTCCAGGCGCAAAAC
AACGTGTTTGATTTAGAGACTATGTTTAGAGAATATAACTATTACGCTCAGAGT
AGCGTCTTTGATATAGAATCCATGTTTAGGGAATATAACTATTACACCTACAGC
AGTGCCCTCGACCTGGAGACGATGTTTAGAGAGTATAATTACTACACACAGAAG

5'-tttgacctggagactatgttymngartayaa-3'
FREYA 64倍 32聚体>

3'-tacaartcyctyatrrttgatgatgtgggtctcg-5'
<FREYB 16倍 33聚体>

3'-ttrcanaarctracctctctgtacaaggetct-5'
<NVFDB 32倍 32聚体>

图 11

EBV
sHV1
mHV68
bHV4

TTCATCTCCTTCTTCAAAAACCCCTTCGGCGGCATGCTC
ATATTAAATTTTATTAAAAATCCTTTAGGTGGCATGTTC
TTCATTAAATCTTTACCAACCCCTCTAGGGGGAATATTC
TTTGTCAATTTTTTCACAAAACCCATTGGGGGCGTGACC

5'-accttcatcaaaaateccttngnggnatgyt-3'
GGMA 128倍 32聚体

图 12

V R R Y R K I A T S V T V Y R G L T E
 KSHV GGTGCGGCGCTATAGGAAAATTGCCACCTCTGTACGGTCTACAGGGGCTTGACAGAG--
 RFHV ...A..CA.G..C.TA.....A..A..A.....C.C..TA.....A--
 bHV4 A...A..AAA.....A.....A..G..A..T.....A..G.G.T.CC...--
 sHV1 A..CA.AA.A..C.AA.....CA.A..A..A...CGTA.T.TT.AT....G...TAGAGA
 eHV2 T..CA.AAAG.....G..CATG.....GAC...CA.....A...T.G..GC...--
 mHV68 AA.CA..A.A.....A..A..AATT..TCAAC.G..CA...GGC.A...C.A..CACT--
 hEBV ...C..CTC...C.CC..G..A.TG...AACA.TCTCA.....AT....G.TAC.C.--

S A I T N K Y E L P R P V P L Y E I S
 KSHV ----TCCGCCATCACCAACAAGTATGAACTCCCGAGACCCGTGCCACTCTATGAGATAAG
 RFHV ----G.A..A.....A.....A.....GA...C..G.....T.....C.....C..
 bHV4 ----G.A..TG.T.....T.G.G...TA..AGC...G..A.A..CTATA...A..TTC
 sHV1 AGGTGTT..T..T..A.....A.GG.....TT.T...G.T..T...AAA.....GA
 eHV2 ----GAT.....A..A...C..C.CACGAGGAGCTACG...C..C..G..C...G.CCA
 mHV68 ----AGTT.AG...TGGT..A.T...A.GG.C.CT.AG.CC.ACGAG.GG..AG.GG..
 hEBV ----GA.T..G.G.....CG.C.C..GGAGAA.TTCT....TGACAG...C..A.CTGA

H M D S T Y Q C F S S M K V N V N G V E
 KSHV CCACATGGACAGCACCTATCAGTGCTTTAGTTCCATGAAGGTAAATGTCAACGGGGTAGA
 RFHV T.....C.....AA.TGTA..G..C..A..C..
 bHV4 AATG..A..T..G.....T..T..CTC.G.T...GCAAC.GTCA.T..T...A.TCT
 sHV1 TATT.....T.AG..T..C..A..TCAT.A..G...C..A..G.A..A...AA.GTT
 eHV2 GATG.....CA.TAT.....CG..G.AC...C..C.AGGGG...CAC.T
 mHV68 .G..T.T.....T.....AC.A.AG.GCC.CCA.GGTG..A...AAC..CAG
 hEBV ...G.....T.C..T...C.....AC.ACG.GG.C...A.G.CAAAAG.T...C.GAC

图 13a

N T F T D R D D V N T T V F L Q P V E G
 KSHV AAACACATTTACTGACAGAGACGATGTTAACACCACAGTATTCCTCCAACCAGTAGAGGG
 RFHV ...T..G..C..C..TC.G..T..C..A...AA..C.....T.....G..C..C..A..
 bHV4 G.....C.A..TA.....G..TTC..AA..T.AGT.T..TCC.....G.....G.CC..
 sHV1 ...TT.T.ACTA.....T.GAAA.....AA..T...GA.T.AA.G..T.....T..
 eHV2 C.....C.ACTA.....G.....GGTGG...GAG..C.CC.....A...G.CC..T..
 mHV68 .C.GGTG.A.GTG.....T.GG..C..T.AA..T..GAA.A.A.GC..T..T..T..
 hEBV GCG.GTG.A.GTA...C.C....GA.....T...C..CAA...AA.G..CACC.G...

L T D N I Q R Y F S Q P V I Y A E P G W
 KSHV GCTTACGGATAACATTCAAAGGTACTTTAGCCAGCCGGTCATCTACGCGGAACCCGGCTG
 RFHV T..A..T..C.....A.....A.....A..A..AC.G...T.T.....A..
 bHV4 A..G..T..G....TAA.C..A.....T..T..A..TC...A..T..A.....T.....
 sHV1 T..A....G.GCA.T.AC...A...A.T.....A..TAAAG.T.TT..T..T..T.....
 eHV2 T..C..CTC..G..TAACGC.C..TCAG..T..A..A.AGG.G....CACC...A.AAA
 mHV68 T..A..A.GG..T..C.....A.....T.....CAC.C.T..TT.A.....T..T..
 hEBV C..GG.CA.CGGGG.G.GCC.C...GCC.....A...AGC...T.AC.CC.....G..

F P G I Y R V R
 KSHV GTTTCCTCCGGCATATACAGAGTTAGG
 RFHV ...C..A..T..C.....G...G..
 bHV4A..G..T..T.....G..A
 sHV1 .C.ATGG..A.CT....GAC.C.A
 eHV2 CC.GTTGT.GTCT....C.ACA..A
 mHV68 .A.G..T...T.T..TC.T...C.A
 hEBV ...GATAT.G.CT.....ACA..A

图 13b

KSHV VRRYRKIATSVTVYRGLTES--AITNKYELPRPVPLYEISHMDSTYQCFSSMKVN
 RFHV VRRYIKIATSVTVYRGMTEA--AITNKYEIPRPVPLYEISHMDSTYQCFSSMKIV
 bHV4 VRKYRKIATSVTVYRGWSQA--AVTNRDDISRAIPYNEISMDRTYHCFSAMATV
 sHV1 VRRYKKITTSVRI FNGWTREGVAITNKWELSRVAVPKYEIDIMDKTYQCHNCMQIE
 eHV2 VRKYRKIMTSTTIYKGWSED--AITNQHTRSAYVPLYEVQMMDHYYQCFSAVQVN
 mHV68 IRRYRKIIITQLTIWRGLTTS--SVTGKFEMATQAHEWEVGFDSIYQCYNSATMV
 hEBV VRSYTKIVTNILIYNGWYAD--SVTNRHEEKFSVDSYETDQMDTIYQCYNAVKMT
 hCMV VRVYQKVLTFRRSYAYIHTT--YLLGSNTEYVAPPMWEIHHINSHSQCYSYSRV
 hHHV6 VRTYKKELTFOSSYRDVGVV--YFLDRTVMGLAMPVYEANLVNSHAQCYSAVAMK
 hVZV ATVYYKDVIVSTAWAGSSYT-QITNRYADRVPIPVSEITDTIDKFGKCSSKATYV
 sHVS A8 ATMYKDVTVSQVWFGHRYS-QFMGIFEDRAPVPFEEVMDKINAKGVC RSTAKYV
 hHSV1 ATMYKDVTVSQVWFGHRYS-QFMGIFEDRAPVPFEEVIDKINAKGVC RSTAKYV

KSHV V-NGVENTFTDRDDV-NTTVFLQPVEGLTDNIQRYFSQPVIYAE PGWFPGIYRVR
 RFHV V-NGVENTFTDRDDV-NKTVFLQPVEGLTDNIQRYFSQPVL YSEPGWFPGIYRVG
 bHV4 I-NGILNTYIDRDSE-NKSVPLQPVAGLTENINRYFSQPLIYAE PGWFPGIYRVR
 sHV1 V-NGMLNSYYDRDGN-NKTVDLKPV DGLTGAITRYISQPKVFADPGWLWGTYRTR
 eHV2 E-GGHVNTYYDRDGV-NETAFLKPADGLTSSITRYQSQPEVYATPRNLLWSYTTR
 mHV68 V-NNVRQVYVDRDGV-NKTVNIRPVDGLTGNIQRYFSOPTLYSEPGWMPGFYRVR
 hEBV K-DGLTRVYVDRDGV-NITVNLKPTGGLANGVRRYASQTELYDAPGWL IWTYRTR
 hCMV I-AGTVFVAYHRDSYENKTMQLMPDDYSNTHSTRYVTVKDQWHSRGSTW-LYRET
 hHHV6 RPDGTVFSAFHEDNNKNNTLNL FPLNFKSITNKRFITTKEPYFARGPLW-LYSTS
 hVZV R-NNHKVEAFNEDKN-PQDMPLIASKYNSVGSKAWHTTNDTYMVAG-TPGTYRTG
 sHVS A8 R-NNMESTAFHRDDH-ESDMALKPAKAATRTSRGWHTTDLKYNPARVEAFHRYGT
 sHSV1 R-NNLETTAFHRDDH-ETDMELKPANAATRTSRGWHTTDLKYNPSRVEAFHRYGT

图 15

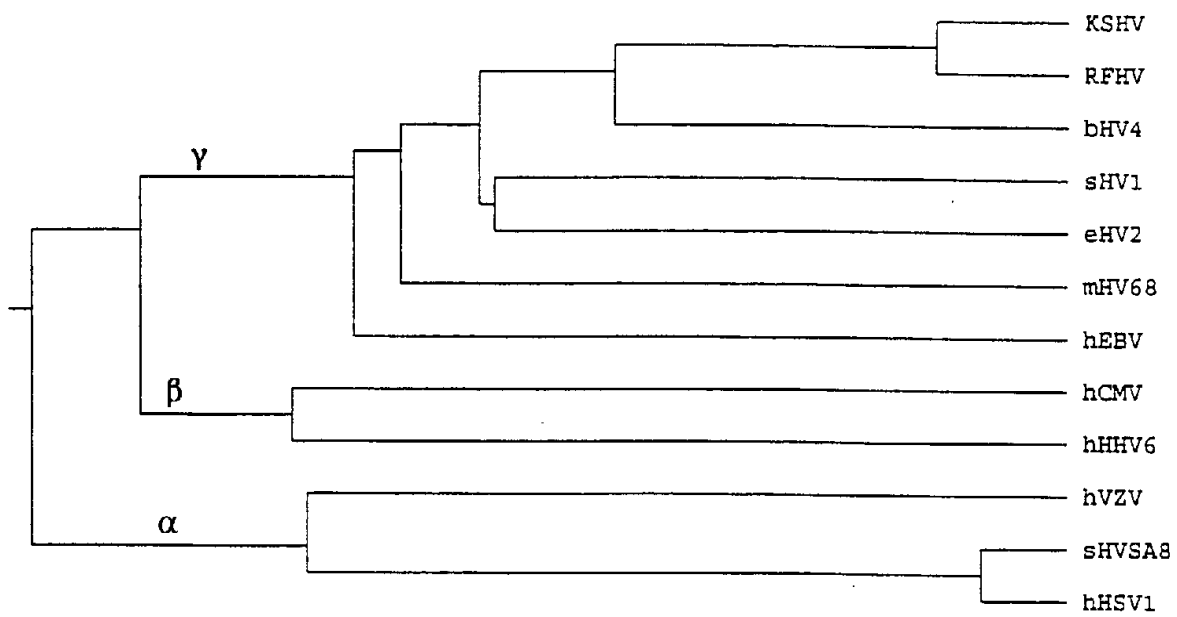


图 16

KSHV GGTGCGGCGCTATAGGAAAATTGCCACCTCTGTACGGTCTACAGGGGCTTGACAGAG--
 RFHV ...A..CA.G..C.TA.....A..A..A.....C.C..TA.....A--
 bHV4 A...A..AAA.....A.....A..G..A..T.....A..G.G.T.CC...--
 sHV1 A..CA.AA.A..C.AA.....CA.A..A..A...CGTA.T.TT.AT...G...TAGAGA
 eHV2 T..CA.AAAG.....G..CATG.....GAC...CA.....A...T.G..GC...--
 mHV68 AA.CA..A.A.....A..A..AATT..TCAAC.G..CA...GGC.A...C.A..CACT--
 hEBV ...C..CTC...C.CC..G..A.TG...AACA.TCTCA.....AT...G.TAC.C.--

KSHV ----TCCGCCATCACCAACAAGTATGAACTCCCGAGACCCGTGCCACTCTATGAGATAAG
 RFHV ----G.A..A.....A.....A.....GA....C..G.....T.....C.....C..
 bHV4 ----G.A..TG.T.....T.G.G....TA..AGC...G..A.A..CTATA...A..TTC
 sHV1 AGGTGTT..T..T..A.....A.GG.....TT.T..G.T..T...AAA.....GA
 eHV2 ----GAT.....A..A...C..C.CACGAGGAGCTACG....C..C..G..C...G.CCA
 mHV68 ----AGTT.AG...TGGT..A.T...A.GG.C.CT.AG.CC.ACGAG.GG..AG.GG.
 hEBV ----GA.T..G.G.....CG.C.C..GGAGAA.TTCT...TGACAG...C..A.CTGA

(SHMDA>) agaccctgcccactctatgarathag-

KSHV CCACATGGACAGCACCTATCAGTGCTTTAGTTCCATGAAGGTAAATGTCAACGGGGTAGA
 RFHV T.....C.....AA.TGTA..G..C..A..C..
 bHV4 AATG..A..T..G.....T..T..CTC.G.T...GCAAC.GTCA.T..T...A.TCT
 sHV1 TATT.....T.AG..T..C..A..TCAT.A..G....C..A..G.A..A.....AA.GTT
 eHV2 GATG.....CA.TAT.....CG..G.AC....C..C.AGGGG...CAC.T
 mHV68 .G..T.T.....T.....AC.A.AG.GCC.CCA.GGTG..A...AAC..CAG
 hEBV ...G.....T.C..T...C.....AC.ACG.GG.C...A.G.CAAAAG.T...C.GAC

(ENTFA> gtcaacggagtaga-
 (<CFSSB) acraartcragntacttctaacaacacttg
 ycayatgga (SHMDA>)

图 17a

KSHV AAACACATTTACTGACAGAGACGATGTTAACACCACAGTATTCCTCCAACCAGTAGAGGG
 RFHV ...T..G..C..C..TC.G..T..C..A....AA..C.....T.....G..C..C..A..
 bHV4 G.....C.A..TA.....G..TTC..AA..T.AGT.T..TCC.....G.....G.CC..
 sHV1 ...TT.T.ACTA.....T.GAAA.....AA..T...GA.T.AA.G..T.....T..
 eHV2 C.....C.ACTA.....G...GGTGG...GAG..C.CC.....A....G.CC..T..
 mHV68 .C.GGTG.A.GTG.....T.GG..C..T.AA..T..GAA.A.A.GC..T..T..T..
 hEBV GCG.GTG.A.GTA...C.C....GA.....T...C..CAA...AA.G..CACC.G...

-raayacnttyacnga (ENTFA>)

KSHV GCTTACGGATAACATTCAAAGGTACTTTAGCCAGCCGGTCATCTACGCGGAACCCGGCTG
 RFHV T..A..T..C.....A.....A.....A..A..AC.G...T.T.....A..
 bHV4 A..G..T..G....TAA.C..A.....T..T..A..TC...A..T..A.....T.....
 sHV1 T..A....G.GCA.T.AC...A...A.T....A..TAAAG.T.TT..T..T..T.....
 eHV2 T..C..CTC..G..TAACGC.C..TCAG..T..A..A.AGG.G....CACC...A.AAA
 mHV68 T..A..A.GG..T..C.....A.....T.....CAC.C.T..TT.A....T..T..
 hEBV C..GG.CA.CGGGG.G.GCC.C...GCC.....A...AGC...T.AC.CC.....G..

tgnctrtrtrtaagtttccatgaaatcggtcggtca (<DNIQB)

KSHV GTTTCCCGGCATATACAGAGTTAGG
 RFHV ...C..A..T..C.....G...G..
 bHV4A..G..T..T.....G..A
 sHV1 .C.ATGG..A.CT.....GAC.C.A
 eHV2 CC.GTTGT.GTCT....C.ACA..A
 mHV68 .A.G..T...T.T..TC.T...C.A
 hEBV ...GATAT.G.CT.....ACA..A

图 17 b

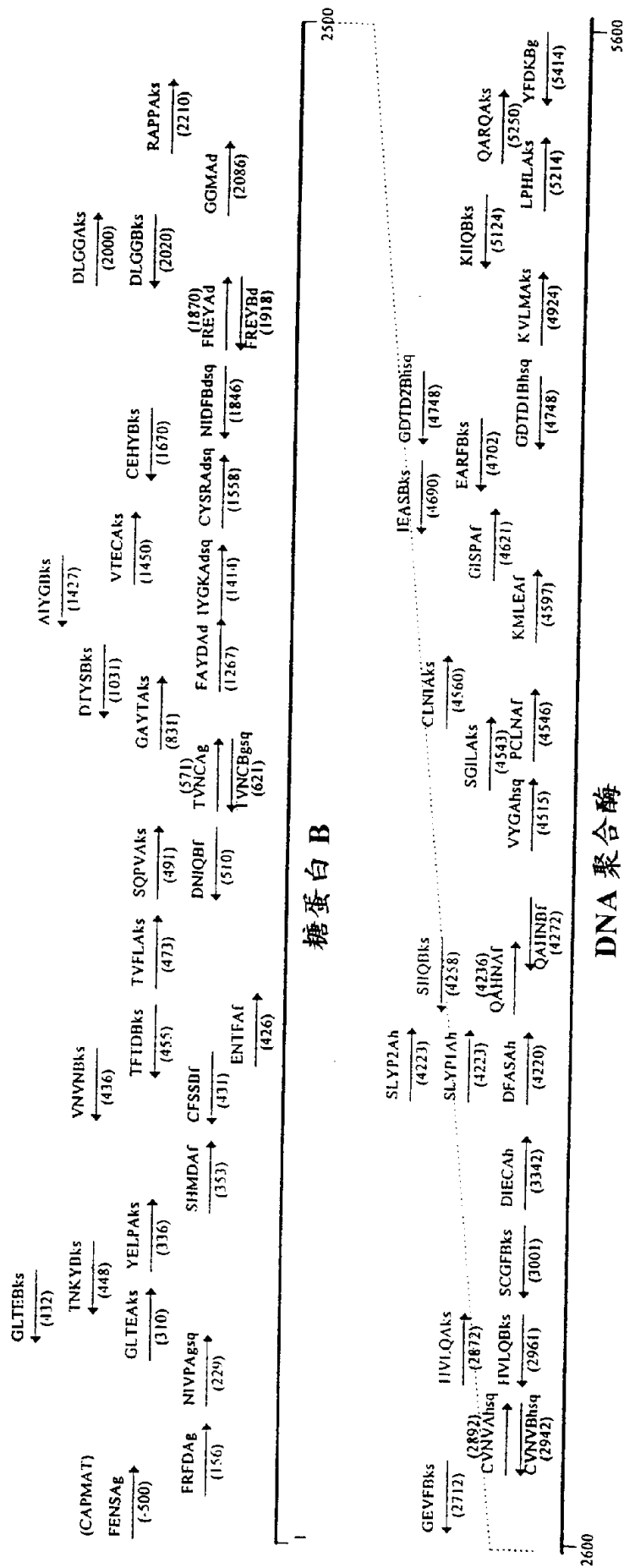
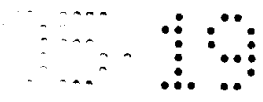


图 18



衣壳/成熟/转运基因

G G M F P I Q K M M V S E M I W P S I E
 TGGGGGCATGTTTCCCATTCAAAAGATGATGGTATCAGAGATGATCTGGCCCAGCATAGA 60

R K D W I E P N F N Q F Y S F E N Q D I
 GCGGAAGGACTGGATAGAGCCCAACTTCAACCAGTTCTATAGCTTTGAGAATCAAGACAT 120

N H L Q K R A W E Y I R E L V L S V S L
 AAACCATCTGCAAAAGAGAGCTTGGGAATATATCAGAGAGCTGGTATTATCGGTTTCTCT 180

N N R T W E R E L K I L L T P Q G S P G
 GAACAACAGAACTTGGGAGAGGGAGCTAAAAATACTTCTCACGCCTCAGGGCTCACCGGG 240

F E E P K P A G L T T G L Y L T F E I S
 GTTTGAGGAACCGAAACCCGAGGACTCACAACGGGGCTGTACCTAACATTTGAGATAIC 300

A P L V L V D K K Y G W I F K D L Y A L
 TGGCCCTTGGTGTGGTGGATAAAAAATATGGCTGGATATTAAAGACCTGTACGCCCT 360

L Y H H L Q L S N H N D S Q V * 糖蛋白 B 基因
 M T P R S R L A T L
 TCTGTACCACCACCTGCAACTGAGCAACCACAATGACTCCCAGGTCTAGATTGGCCACCC 420

G T V I L L V C F C A G A A H S R G D T
 TGGGACTGTCATCCTGTTGGTCTGCTTTTGCAGGCGCGGCGCACTCGAGGGGTGACA 480

F Q T S S S P T P P G S S S K A P T K P
 CCTTTCAGACGTCCAGTTCACCCACACCCCAAGGATCTTCTCTAAGGCCCCACCAAAC 540

G E E A S G P K S V D F Y Q F R V C S A
 CTGGTGAGGAAGCATCTGGTCCCTAAGAGTGTGGACTTTTACCAGTTCAGAGTGTGTAGTG 600

S I T G E L F R F N L E Q T C P D T K D
 CATCGATCACCAGGGAGCTTTTTTCGGTTCAACCTGGAGCAGCGTGCCAGACACCAAAG 660

K Y H Q E G I L L V Y K K N I V P H I F
 ACAAGTACCACCAAGAAGGAATTTTACTGGTGTACAAAAAAACATAGTGCCCTCATATCT 720

K V R R Y R K I A T S V T V Y R G L T E
 TTAAGGTGCGGCGCTATAGGAAAATTGCCACCTCTGTACGGTCTACAGGGGCTTGACAG 780

S A I T N K Y E L P R P V P L Y E I S H
 AGTCCGCCATCACCAACAAGTATGAACTCCCAGACCCGTGCCACTCTATGAGATAAGCC 840

M D S T Y Q C F S S M K V N V N G V E N
 ACATGGACAGCACCTATCAGTGCTTTAGTTCCATGAAGGTAAATGTCAACGGGGTAGAAA 900

图 19a

T F T D R D D V N T T V F L Q P V E G L
 ACACATTTACTGACAGAGACGATGTAAACACCACAGTATTCTCCAACCAGTAGAGGGGC 960

T D N I Q R Y F S Q P V I Y A E P G W F
 TTACGGATAACATTCAAAGGTACTTTAGCCAGCCGGTCATCTACGGGAACCCGGCTGGT 1020

P G I Y R V R T T V N C E I V D M I A R
 TTCCGGCATATACAGAGTTAGGACCACTGTCAATTGCGAGATAGTGGACATGATAGCCA
 > C 1080

S A E P Y N Y F V T S L G D T V E V S P
 GGTCTGCTGAACCATAACAATTACTTTGTACGTCACCTGGGTGACACGGTGAAGTCTCCC 1140

F C Y N E S S C S T T P S N K N G L S V
 CTTTTGCTATAACGAATCCTCATGCAGCAACCCCGCAACAAAATGGCCTTAGCG 1200

Q V V L N H T V V T Y S D R G T S P T P
 TCCAAGTAGTTCTCAACCACACTGTGGTCACGTA CTCTGACAGAGGAACCAGTCCCCTC 1260

Q N R I F V E T G A Y T L S W A S E S K
 CCCAAAACAGGATCTTTGTGGAAACGGGAGCGTACAGCTTTCGTGGGCCTCCGAGAGCA 1320

T T A V C P L A L W K T F P R S I Q T T
 AGACCACGGCCGTGTGTCCGCTGGCACTGTGGAAAACCTTCCCGCGCTCCATCCAGACTA 1380

H E D S F H F V A N E I T A T F T A P L
 CCCACGAGGACAGCTTCCACTTTGTGGCCAACGAGATCACGGCCACCTTCACGGCTCCTC 1440

T P V A N F T D T Y S C L T S D I N T T
 TAAGCCAGTGGCCAACTTTACCGACACGTA CTCTGTCTGACCTCGGATATCAACACCA 1500

L N A S K A K L A S T H V P N G T V Q Y
 CGCTTAACGCCAGCAAGGCCAAACTGGCGAGCACTCACGTCCCTAACGGGACGGTCCAGT 1560

F H T T G G L Y L V W Q P M S A I N L T
 ACTTCCACACAACAGGCGGACTCTATTTGGTCTGBCAGCCCATGTCCGCGATTAACCTGA 1620

H A Q G D S G N P T S S P P P S A S P M
 CTCACGCTCAGGGGACAGCGGGAACCCACGTCATCGCCGCCCCCTCCGCATCCCCCA 1680

T T S A S R R K R R S A S T A A A G G G
 TGACCACCTCTGCCAGCCGCAAGAGACGGTCAGCCAGTACCGCTGCTGCCGGCGGGC 1740

G S T D N L S Y T Q L Q F A Y D K L R D
 GGGGGTCCACGGACAACCTGTCTTACACGCAGCTGCAGTTTGCCTACGACAACTGCGGG 1800

G I N Q V L E E L S R A W C R E Q V R D
 ATGGCATTAAATCAGGTGTTAGAAGAACTCTCCAGGGCATGGTGTCCGAGCAGGTGAGGG 1860

图 19 b



N L M W Y E L S K I N P T S V M T A I Y
ACAACCTAATGTGGTACGAGCTCAGTAAATCAACCCACCAGCGTTATGACAGCCATCT 1920

G R P V S A K F V G D A I S V T E C I N
ACGGTCGACCTGTATCCGCCAAGTTCGTAGGAGACGCCATTTCCGTGACCGAGTGCATTA 1980

V D Q S S V N I H K S L R T N S K D V C
ACGTGGACCAGAGCTCCGTAACATCCACAAGAGCCTCAGAACCAATAGTAAGGACGTGT 2040

Y A R P L V T F K F L N S S N L F T G Q
GTTACGCGCGCCCTGGTGCAGTTTAAAGTTTTGAACAGTTC AACCTATTCACCGGCC 2100

L G A R N E I I L T N N Q V E T C K D T
AGCTGGGCGCGCAATGAGATAATACTGACCAACAACCAGGTGGAACCTGCAAAGACA 2160

C E H Y F I T R N E T L V Y K D Y A Y L
CCTGCGAACACTACTTCATCACCCGCAACGAGACTCTGGTGTATAAGGACTACGCGTACC 2220

R T I N T T D I S T L N T F I A L N L S
TGCGCACTATAAACACCACTGACATATCCACCCTGAACACTTTTATCGCCCTGAATCTAT 2280

F I Q N I D F K A I E L Y S S A E K R L
CCTTTATTCAAACATAGACTTCAAGGCCATCGAGCTGTACAGCAGTGCAGAGAAACGAC 2340

A S S V F D L E T M F R E Y N Y Y T H R
TCGCGAGTAGCGTGTGGTACCTGGAGACGATGTTTCAGGGAGTACAACACTACTACACATC 2400

L A G L R E D L D N T I D M N K E R F V
GTCTCGCGGGTTTGGCGGAGGATCTGGACAACCCATAGATATGAACAAGGAGCGCTTCG 2460

R D L S E I V A D L G G I G K T V V N V
TAAGGGACTTGTGCGGAGATAGTGGCGGACCTGGGTGGCATCGGAAAAACGGTGGTGAACG
> T 2520

A S S V V T L C G S L V T G F I N F I K
TGGCCAGCAGCGTGGTCACTCTATGTGGCTCATTGGTTACCGGATTCATAAATTTATTA 2580

H P L G G M L M I I I V I A I I L I I F
AACACCCCTAGGTGGCATGCTGATGATCATTATCGTTATAGCAATCATCCTGATCATT 2640

M L S R R T N T I A Q A P V K M I Y P D
TTATGCTCAGTCGCCGACCAATACCATAGCCCAGGCGCCGGTGAAGATGATCTACCCCG 2700

V D R R A P P S G G A P T R E E I K N I
ACGTAGATCGAGGGCACCTCCTAGCGGCGGAGCCCAACACGGGAGGAAATCAAAAACA 2760

图 19 c



L L G M H Q L Q Q E E R Q K A D D L K K
 TCCTGCTGGGAATGCACCAGCTACAACAAGAGGAGAGGCAGAAGGCGGATGATCTGAAAA 2820
 > T

S T P S V F Q R T A N G L R Q R L R G Y
 AAAGTACACCCTCGGTGTTTCAGCGTACCGCAAACGGCCTTCGTACGCGTCTGAGAGGAT 2880
 > L

K P L T Q S L D I S P E T G E *
 ATAAACCTCTGACTCAATCGCTAGACATCAGTCCGGAACGGGGAGTGACAGTGGATTC 2940
 > T

GAGGTTATTGTTTGATGTAAATTTAGGAAACACGGCCCGCCTCTGAAGCACCACATACAG 3000

DNA 聚合酶基因

M D

ACTGCAGTTATCAACCCTACTCGTTGCACACAGACACAAATTACCGTCCGCAGATCATGG 3060

F F N P F I D P T R G G P R N T V R Q P
 ATTTTTCAATCCATTTATCGACCCAACCTCGCGGAGGCCGAGAAACACTGTGAGGCAAC 3120

T P S Q S P T V P S E T R V C R L I P A
 CCACGCCGTACAGTCGCCAACTGTCCCCTCGGAGACAAGAGTATGCAGGCTTATACCGG 3180

C F Q T P G R P G V V A V D T T F P P T
 CCTGTTTCAAACCCGGGGCGACCCGGCGTGGTTGCCGTGGACACCACATTTCCACCCA 3240

Y F Q G P K R G E V F A G E T G S I W K
 CCTACTTCAGGGCCCAAGCGGGGAGAAGTATTCGCGGGAGAGACTGGGTCTATCTGGA 3300

T R R G Q A R N A P M S H L I F H V Y D
 AAACAAGGCGCGGACAGGCACGCAATGCTCCTATGTCGCACCTCATATTCCACGTATACG 3360

I V E T T Y T A D R C E D V P F S F Q T
 ACATCGTGGAGACCCTACACGGCCGACCGCTGCGAGGACGTGCCATTTAGCTTCCAGA 3420

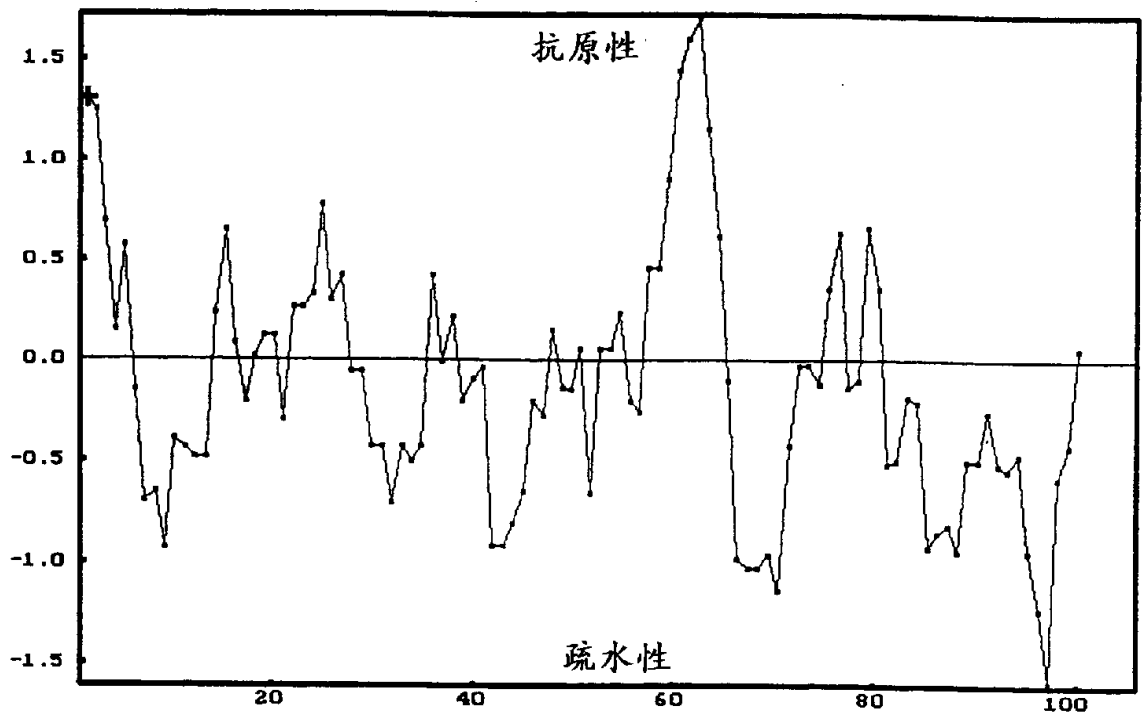
D I I P S G T V L K L L G R T L D G A S
 CTGATATCATTCCCAGCGGCACCGTCCCTCAAGCTGCTCGGCAGAACACTAGATGGCGCCA 3480

V C V N V F R Q R C Y F Y T L A P Q G V
 GTGTCTGCGTGAACGTTTTTCAGGCAGCGCTGCTACTTCTACACACTAGCACCCAGGGGG 3540

N L T H V L Q Q A L Q A G F G R A S C G
 TAAACCTGACCCACGTCCTCCAGCAGGCCCTCCAGGCTGGCTTCGGTTCGCGCATCCTGCG 3600

F S T
 GCTTCTCCACCG 3612

图 19 d



VRRYRKIATS VTVYRGLTES AITNKYELPR PVPLYEISHM DSTYQCFSSM KVNNGVENT 60
 ~~~~~  
 #  
 FTDRDDVNTT VFLQPVEGLT DNIQRYFSQP VIYAEPGWFP GIYRVR      106  
 ~~~~~^ ~~~~~^      ^^^^ ^^^

图 21

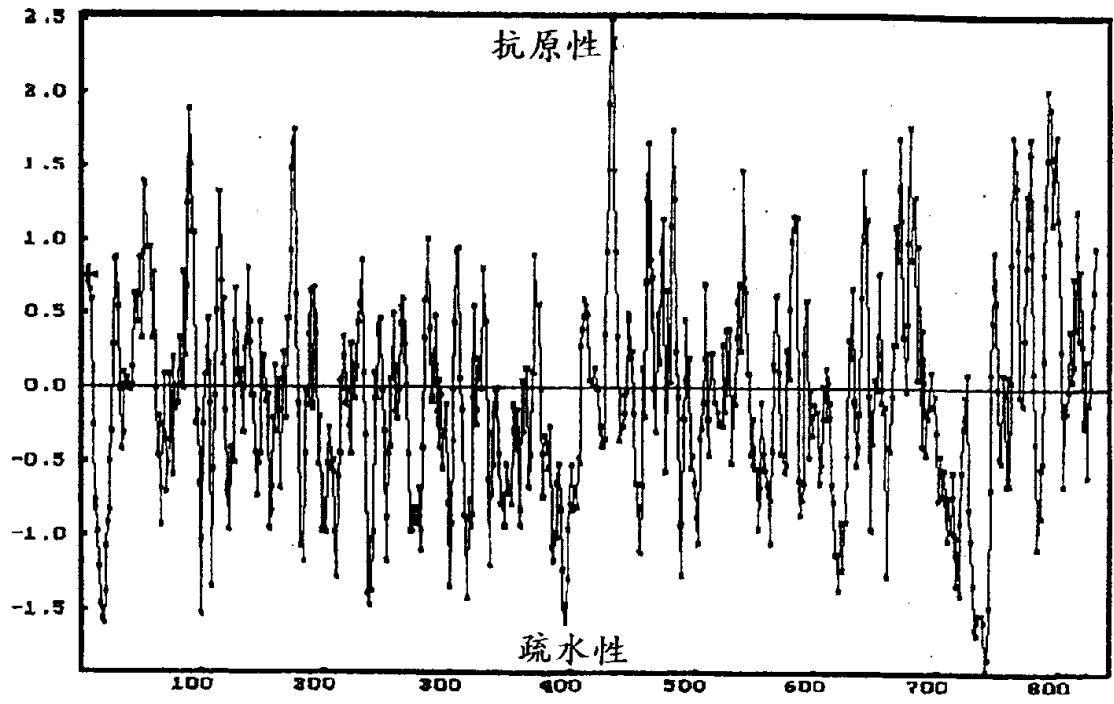
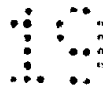


图 22

V Y K K N I V P N M F K V R R Y R K V A
 GTGTACAAGAAGAACATCGTGCCTAACATGTTCAAGGTACGCAGGTACAGAAAAGTAGCA 60

T P V T L Y R G M T D A A I T N K Y E I
 ACGCCTGTCACACTCTACCGGGTATGACAGACGCAGCAATAACTAACAAATATGAAATT 120

P R P V P L Y E I S H M D S T Y Q C F S
 CCCAGACCCGTACCACTATACGAGATCAGTCACATGGACAGCACCTACCAGTGCTTTAGT 180

S M K I V V N G V E N T F T G R D D V N
 TCCATGAAAATTGTAGTGAACGGAGTCGAAAACACGTTCCACCGGTCGGGATGACGTAAC 240

K S V F L Q P V E G L T D N I K R Y F S
 AAAAGCGTATTTCTCCAGCCAGTCGAAGGTCTAACTGACAACATAAAGAGATACTTTAGC 300

Q P V L Y S E P G W F P G I Y R V R T T
 CAGCCAGTGCTATATTCTGAACCCGGATGGTTTCCAGGTATCTACAGGGTTAGGACAACA 360

V N C E I V D M
 GTTAATTGTGAGATTGTAGACATGTT 380

图 23