(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 102089432 A
(43)申请公布日 2011.06.08

(21)申请号 200980126670.5

(22)申请日 2009.07.07

(30)优先权数据
61/078,963 2008.07.08 US

(85)PCT申请进入国家阶段日
2011.01.07

(86)PCT申请的申请数据
PCT/CA2009/000941 2009.07.07

(87)PCT申请的公布数据
WO2010/003235 EN 2010.01.14

(71)申请人 麦迪卡格公司
地址 加拿大魁北克

(72)发明人 马农·科图雷 纳萨莉·兰德里
路易斯-菲利普·韦齐纳
米凯莱·达吉斯

(74)专利代理机构 北京集佳知识产权代理有限公司 11227
代理人 顾晋伟 彭鲲鹏

(51)Int.Cl.
C12N 15/62(2006.01)
A61K 39/145(2006.01)
A61K 39/385(2006.01)
C07K 14/11(2006.01)
C07K 19/00(2006.01)
C12N 15/44(2006.01)
C12N 15/82(2006.01)
C12P 21/02(2006.01)

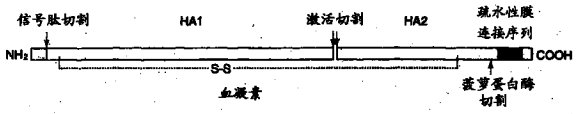
权利要求书 1 页 说明书 17 页
序列表 23 页 附图 10 页
按照条约第19条修改的权利要求书 1 页

(54)发明名称

可溶性重组流感抗原

(57)摘要

本发明提供了重组可溶性血凝素(rHA)三聚体蛋白,其包含血凝素胞外结构域和寡聚化结构域。所述rHA作为可溶性同三聚体产生,还可包含信号肽和/或内质网(ER)驻留信号。本发明还涉及编码本发明之rHA的核酸,以及包含所述核酸的载体和嵌合构建体。还提供了产生所述rHA的方法。本发明所述的rHA可用于配制流感疫苗,或者可用于强化已有疫苗。



CN 102089432 A

1. 一种重组血凝素 (rHA),其包含 :

a) 血凝素结构域 ;和

b) 寡聚化结构域

其中所述 rHA 作为嵌合可溶性同三聚体产生。

2. 权利要求 1 的 rHA,其中所述蛋白还包含信号肽。

3. 权利要求 1 或 2 的 rHA,其中所述蛋白还包含内质网 (ER) 驻留信号。

4. 编码权利要求 1 至 3 中任一项的 rHA 的核苷酸序列。

5. 核酸序列,其包含 :

a) 编码血凝素结构域的核苷酸序列 ;和

b) 编码寡聚化结构域的核苷酸序列 ;

其中所述核酸编码形成同三聚体的可溶性 rHA。

6. 权利要求 5 的核酸序列,还包含编码信号肽的核苷酸序列。

7. 权利要求 5 或 6 的核酸,还包含编码内质网 (ER) 驻留信号的核苷酸序列。

8. 一种载体,其包含权利要求 4 至 7 中任一项的核苷酸。

9. 表达权利要求 1 至 3 中任一项的 rHA 的宿主细胞。

10. 用权利要求 4 至 7 中任一项的核苷酸转化的宿主细胞。

11. 用权利要求 8 的载体转化的宿主细胞。

12. 一种产生重组 rHA 蛋白的方法,包括提供含有载体的宿主细胞,所述载体包含 :

a) 编码血凝素结构域的核苷酸序列 ;和

b) 编码寡聚化结构域的核苷酸序列 ;

其中所述核酸编码形成同三聚体的可溶性 rHA。

13. 一种在植物中表达重组血凝素 (rHA) 的方法,包括向植物中引入载体并表达 rHA,其中所述载体包含 :

a) 编码血凝素结构域的核苷酸序列,其中所述核酸编码形成同三聚体的可溶性 rHA ;和

b) 编码寡聚化结构域的核苷酸序列。

14. 权利要求 13 的方法,其中在引入步骤 (步骤 a) 中,所述核酸以瞬时方式被引入所述植物中。

15. 权利要求 13 的方法,其中在引入步骤 (步骤 a) 中,所述核酸以使其稳定的方式被引入所述植物中。

16. 一种在植物中产生重组血凝素 (rHA) 的方法,包括 :

a) 将核酸序列引入所述植物或其部分中,所述核酸序列包含与编码血凝素结构域和寡聚化结构域之核苷酸序列有效连接的调控区域 :其中所述核酸编码形成同三聚体的可溶性 rHA ;以及

b) 培养所述转基因植物,从而产生所述 rHA。

17. 权利要求 16 的方法,其中在引入步骤 (步骤 a) 中,所述核酸以瞬时方式被引入所述植物中。

18. 权利要求 16 的方法,其中在引入步骤 (步骤 a) 中,所述核酸以使其稳定的方式被引入所述植物中。

可溶性重组流感抗原

发明领域

[0001] 本发明涉及可溶性重组流感抗原的产生。更具体地，本发明涉及保持免疫原性的可溶性重组流感抗原的产生。

[0002] 发明背景

[0003] 流感是由呼吸系统病毒导致的人类死亡的首要原因。常见症状包括发烧、咽喉疼痛、呼吸短促以及肌肉酸痛等。在流感季节，流感病毒感染世界人口的 10-20％，每年导致 250-500,000 例死亡。

[0004] 流感病毒根据存在的核蛋白和基质蛋白抗原分为甲、乙或丙型。甲型流感病毒可根据存在的血凝素 (hemagglutinin，HA) 和神经氨酸酶 (neuraminidase，NA) 表面糖蛋白的组合进一步分成亚型。HA 决定着病毒结合和进入宿主细胞的能力。NA 从宿主细胞和病毒表面蛋白的多糖链上除去末端唾液酸残基，所述唾液酸残基防止病毒聚集并促进病毒运动。目前，已确认 16 种 HA(H1-H16) 和 9 种 NA(N1-N9) 亚型。每种甲型流感病毒存在一种类型的 HA 和一种类型的 NA 糖蛋白。一般来讲，每种亚型都显示物种特异性；例如，已知所有 HA 和 NA 亚型都感染鸟类，而只有 H1、H2、H3、H5、N1 和 N7 表明感染人。认为含 H5 和 H7 的流感病毒为甲型流感病毒的最强致病形式，并且最可能引起未来的大流行。

[0005] 流感大流行通常由高传播性和毒力的流感病毒引起，并且可导致全球更高水平的疾病和死亡。新甲型流感病毒亚型的出现在 20 世纪引发了 4 次主要大流行。1918-1919 年由 H1N1 病毒引起的西班牙流感在 1917 至 1920 年间导致全球超过五千万人死亡。出现新亚型或动物特有的亚型传播到人的风险始终存在。尤其关注的是高度毒力形式的鸟类流感（也称为“禽流感”），其爆发已在全世界几个国家中有报道。在很多情况下，该禽流感可在 48 小时内导致将近 100％的死亡率。已假定禽流感病毒 (H5N1，其在 1997 年在香港首次鉴定）在其它亚洲国家和欧洲的传播与野生鸟类的迁移模式相关。

[0006] 越来越得到关注的问题是，病毒可能变得对人具有高度感染性。人类健康的主要问题在于流感病毒抗原性不稳定这一事实，即它们快速变异。如果禽流感病毒与人病毒接触，则该鸟类病毒的遗传重排可产生高致病性的流感病毒，其可在人中引起严重的疾病或死亡。而且，这种突变还可导致流感病毒在人中顺利地传播。

[0007] 抗击人中流感的现行方法为每年接种。每年世界卫生组织选择 3 种病毒株纳入当年度的在受精卵中生产的流感疫苗中。然而，每年生产的疫苗剂量并不足以接种全世界人口。例如，加拿大和美国可获得足够免疫其约三分之一人口的疫苗剂量，而在欧盟只有 17％的人口可被接种。很显然，当面临全球流感大流行时全球的流感疫苗生产将不足。因此，政府和私人企业等已将注意力转向有效流感疫苗的生产上。

[0008] 如上所述，目前获得流感病毒疫苗的方法是通过在受精卵中生产。将病毒在受精卵中培养，而后灭活病毒并纯化病毒糖蛋白。尽管该方法保持了抗原表位和翻译后修饰，但有几个缺点，包括由于采用全病毒而导致的污染风险，以及产量根据病毒株不同而变化。不理想的防护水平可能由于病毒被引入卵中所导致的其遗传异质性而引起。其它缺点包括获得卵所需的大量工作、由于在纯化中使用的化学物导致的污染风险，以及生产时间长。还

有，对卵蛋白过敏的人可能不是接受疫苗的合适候选者。

[0009] 为了避免卵的使用，还在哺乳动物细胞培养中生产了流感病毒，例如，在 MDCK 或 PERC.6 等细胞中。另一个方法为反向遗传学，其中通过用病毒基因转化细胞产生病毒。然而这些方法也需要采用完整病毒，并且方法复杂，以及需要特定的培养环境。

[0010] 已经研究了病毒 DNA 作为疫苗的应用。在该技术中，通过在人细胞中表达病毒抗原获得保护；随后抗原被识别为外来抗原，这导致特异性的抗体应答。然而，存在将 DNA 引入人细胞基因组决定性部分而使癌基因激活的风险 - 这是重大的缺陷。

[0011] Dow Agroscience 也制备了包含用病毒 DNA 转化的昆虫或植物细胞表达的重组病毒抗原的疫苗（例如见 WO2004/098530）。尽管避免了与采用活病毒相关的风险，并且缩短了生产过程，但是蛋白构象和翻译后修饰仍受到影响。放大和纯化步骤也相对复杂，因为抗体与细胞膜相结合。此外，有效免疫动物所需的杆状病毒重组 HA 的剂量比受精卵中产生的天然 HA 的剂量高 10 倍。在两种情况下，病毒抗原表达水平都很低。

[0012] 为了避免与纯化膜蛋白相关的困难，Huang 等 (2001, Vaccine, 19 :2163-2171) 用 ER 驻留信号替换了麻疹 HA 的跨膜结构域和胞质尾区。所得的 HA 蛋白在烟草植物细胞中产生用于开发口服疫苗（可食用疫苗）。表达的 HA 在 ER 中的存留没有用跨膜结构域产生的存留强，因此简化了纯化过程。然而，在这些条件下不能形成 HA 的天然三聚体形式，这可影响重组蛋白的免疫原性。

[0013] Saelens 等 (1999, Eur. J. Biochm, 260 :166-175) 在酵母中 (Pichiapastoris) 表达了缺少跨膜结构域的 HA 基因，导致了 HA 单体的分泌。然而，该形式与三聚体 HA 相比免疫原性较低。

[0014] 为了保护世界人口免受流感侵袭并且避开未来的大流行，将需要疫苗生产者开发生产疫苗剂的有效快速方法。目前采用受精卵生产疫苗不能满足需要而且过程长。重组技术提供了下生产流感病毒抗原的有希望的方法。然而，血凝素的生产受限于涉及复杂提取过程而产量低的膜相关蛋白，或受限于低免疫原性的可溶性蛋白。

[0015] 发明概述

[0016] 本发明涉及可溶性重组流感抗原的生产。更具体来讲，本发明涉及保持三聚体组装和免疫原性的可溶性重组流感抗原。

[0017] 本发明的一个目的是提供可溶性重组流感抗原。

[0018] 本发明提供了包含血凝素结构域和寡聚化结构域的重组血凝素 (rHA)。rHA 作为可溶性同三聚体而产生。该蛋白可进一步包含信号肽和 / 或内质网 (ER) 驻留信号。

[0019] 本发明还提供了编码上述 rHA 的核苷酸序列。

[0020] 本发明进一步提供了核酸序列，其包含 :a) 编码血凝素结构域的核苷酸序列 ;和 b) 编码寡聚化结构域的核苷酸序列。该核酸编码形成同三聚体的可溶性 rHA。该核酸还可包含编码信号肽和 / 或内质网 (ER) 驻留信号的核苷酸序列。

[0021] 本发明还提供了包含上述核苷酸的载体。

[0022] 本发明进一步提供了表达上述 rHA 的宿主细胞，用上述核苷酸转化的宿主细胞，或用上述载体转化的宿主细胞。

[0023] 本发明还提供了一种生产重组 rHA 蛋白的方法。所述方法包括提供带有载体的宿主细胞，其中所述载体包含 :a) 编码血凝素结构域的核苷酸序列，其中所述核酸编码形成

同三聚体的可溶性 rHA，以及 b) 编码寡聚化结构域的核苷酸序列；然后表达该 rHA。

[0024] 本发明进一步提供了在植物中表达重组血凝素 (rHA) 的方法。在第一步中，将载体引入植物中，所述载体包含编码血凝素结构域的核苷酸序列（其中所述核酸编码形成同三聚体的可溶性 rHA) 以及编码寡聚化结构域的核苷酸序列。在所述引入步骤（步骤 a) 中，所述核酸可以瞬时方式被引入植物中，或者所述核酸可以使之稳定的方式被引入植物中。

[0025] 本发明还提供了在植物中生产重组血凝素 (rHA) 的方法，包括：a) 将核酸序列引入植物或其部分中，该核酸序列包含与编码血凝素结构域和寡聚化结构域之核苷酸序列有效连接的调控区域，其中所述核酸编码形成同三聚体的可溶性 rHA；以及 b) 培养所述转基因植物，从而生产 rHA。在引入步骤（步骤 a) 中，所述核酸可以瞬时的方式被引入植物中，或者所述核酸可以使之稳定的方式被引入植物中。

[0026] 对于生产而言，rHA 是非常复杂的分子。当前生产系统的重组 HA 的表达水平和产量低；因此生产成本高。这很大程度上是由于该蛋白质的复杂三聚体结构，其在其合成中必须经历复杂的组装过程。而且，HA 为具有跨膜结构域的大蛋白，并且高度糖基化。生产可溶形式的 HA 可使得能够以更高水平生产，并降低纯化过程的复杂性。这将对生产成本产生重要影响。已表明用可溶性的 α－螺旋或其它二级结构（其适于结构性稳定 HA 并与 HA 蛋白胞外结构域的卷曲螺旋核心相容）替代此跨膜结构域可产生稳定的可溶性 HA 三聚体。可用这种重组蛋白来强化目前的流感疫苗，或制备新疫苗。

[0027] 本发明的此概述并不一定描述本发明的所有特征。

[0028] 附图简述

[0029] 根据以下描述并参考附图，本发明的这些和其它特征将更加显而易见，其中：

[0030] 图 1 显示了天然血凝素 (HA) 蛋白的结构域示意图。

[0031] 图 2 显示了 GCN4-pII 肽的氨基酸序列 (SEQ ID NO：1)。

[0032] 图 3 显示了 PDI 的氨基酸和核苷酸序列 (SEQ ID NO：6 和 7；(Genbank 登记号 Z11499)，即苜蓿信号肽。该 PDI 信号肽与小鼠 ERp59 同源。BglII 限制性位点以粗体显示。

[0033] 图 4 显示了甲型流感病毒株 /New Caledonia/20/99(H1N1) 的 HA 的氨基酸序列 (SEQ ID NO：8)(Genbank 登记号 AY289929，初始登记 UniProt KB/TrEMBL：Q6WG00)；rHA 信号肽以斜体显示。HA0 切割位点以粗体显示，融合肽以下划线表示。跨膜结构域以灰色背景显示。

[0034] 图 5 显示了根据本发明的多种 rHA 构建体的氨基酸序列。氨基酸的标号已根据 HA 原始氨基酸的标号做了调整。PDI 信号肽以斜体表示，HA_0 切割位点以粗体表示，融合肽用下划线表示，终止密码子用*代表。图 5A 为全长 rHA 的氨基酸序列 (SEQ ID NO：9)，其包含 PDI 信号肽、跨膜结构域和胞质尾区 (cytoplasmic tail)。跨膜结构域以灰色背景表示。图 5B 为采用 SEKDEL 驻留信号的 ER 驻留 rHA 的氨基酸序列 (SEQ IDNO：10)。驻留信号以灰色背景显示。图 5C 为采用 HDEL 驻留信号的 ER 驻留 rHA 的氨基酸序列 (SEQ ID NO：11)。驻留信号以灰色背景显示。图 5D 为无跨膜结构域的可溶性 rHA 的氨基酸序列 (SEQ ID NO：12)。图 5E 为采用 GCN4-pII 三聚体肽的可溶三聚体 rHA 的氨基酸序列 (SEQ IDNO：13)。GCN4-pII 肽以灰色背景显示。图 5F 为采用 GCN4-pII 三聚体肽并驻留于 ER 的可溶性三聚体 rHA 的氨基酸序列 (SEQ ID NO：14)。GCN4-pII 肽以灰色背景显示，SKDEL 驻留信号以斜体显示。图 5G 为采用 PRD 三聚体肽的可溶性三聚体 rHA 的氨基酸序列 (SEQ ID NO：15)。

PRD 肽以灰色背景显示。图 5H 为采用 PRD 三聚体肽并驻留于 ER 的可溶性三聚体 rHA 的氨基酸序列 (SEQ ID NO :16)。PRD 肽以灰色背景显示，驻留信号以斜体显示。

[0035] 图 6 显示了本发明多个片段的核苷酸序列。非编码序列用小写字母表示，可用的限制位点用下划线表示。图 6A 显示了 HA0 基因片段的核苷酸序列 (SEQ ID NO :17)。图 6B 显示了跨膜结构域和胞质尾区基因片段的核苷酸序列 (SEQ ID NO :18)。图 6C 显示了 ER 驻留 SEKDL 基因片段的核苷酸序列 (SEQ ID NO :19)。图 6D 显示了 ER 驻留 HDEL 基因片段的核苷酸序列 (SEQ ID NO :20)。图 6E 显示了 GCN4-pII 基因片段的核苷酸序列 (SEQ ID NO :21)。图 6F 显示了 ER 驻留 GCN4-pII 基因片段的核苷酸序列 (SEQ ID NO :22)。图 6G 显示了 PRD 基因片段的核苷酸序列 (SEQ ID NO :23)。图 6H 显示了 ER 驻留 PRD 基因片段的核苷酸序列 (SEQ ID NO :24)。

[0036] 图 7 为根据本发明一个实施方案的 pCAMBIA 二元质粒中 rHA 转移 DNA(t-DNA) 的示意图。

[0037] 图 8 为 Western 印迹图，其显示了烟草中 rHA 表达的免疫检测。泳道 :1) 纯的 rHA 标准品 (1ng) ;2) 掺入 1ng 标准 rHA 的 10μg 植物提取物 ;3)10μg 植物提取物 ;4) 来自表达构建体 #540 之生物质的 10μg 蛋白提取物 ;5) 来自表达构建体 #541 之生物质的 10μg 蛋白提取物 ;6) 来自表达构建体 #542 之生物质的 10μg 蛋白提取物 ;7) 来自表达构建体 #544 之生物质的 10μg 蛋白提取物 ;8) 来自表达构建体 #545 之生物质的 10μg 蛋白提取物 ;9) 来自表达构建体 #546 之生物质的 10μg 蛋白提取物 ;10) 来自表达构建体 #547 之生物质的 10μg 蛋白提取物。

[0038] 图 9 为 Western 印迹图，其显示了用 5μg 提取物进行的烟草中 rHA 表达的免疫检测。图 9A 显示了从本塞姆氏烟草 (N.benthamiana) 获得的结果，图 9B 显示了从普通烟草 (N.tabacum) 获得的结果。泳道 :1) 分别为区域 B、D 和 A、C 的 5μg 植物提取物 ;2) 分别为区域 B、D 和 A、C 的掺入了 1ng 标准 rHA 的 2 和 5μg 植物提取物 ;3) 来自表达构建体 #540 之生物质的提取物 ;4) 来自表达构建体 #541 之生物质的提取物 ;5) 来自表达构建体 #542 之生物质的提取物 ;6) 来自表达构建体 #544 之生物质的提取物 ;7) 来自表达构建体 #545 之生物质的提取物 ;8) 来自表达构建体 #546 之生物质提取物 ;9) 来自表达构建体 #547 之生物质的提取物。

[0039] 图 10 为 Western 印迹图，其显示了用 5μg 提取物在还原条件下进行的烟草中 rHA 表达的免疫检测。图 10A 显示了从本塞姆氏烟草获得的结果，而图 10B 显示了从普通烟草获得的结果。泳道 :1)5μg 植物提取物 ;2) 掺入了 1ng 标准 rHA 的 5μg 植物提取物 ;3) 来自表达构建体 #540 之生物质的提取物 ;4) 来自表达构建体 #541 之生物质的提取物 ;5) 来自表达构建体 #542 之生物质的提取物 ;6) 来自表达构建体 #544 之生物质的提取物 ;7) 来自表达构建体 #545 之生物质的提取物 ;8) 来自表达构建体 #546 之生物质的提取物 ;9) 来自表达构建体 #547 之生物质的提取物。

[0040] 图 11 显示了有凝血反应测定结果的板。行 1 :PBS(阴性对照) ;行 2 :PBS+1000ng HA(PSC) ;行 3 :PBS+100ng HA(PSC) ;行 4 :PBS+10ngHA(PSC) ;行 5 :PBS+1ng HA(PSC) ;行 6 :未转化植物的提取物 ;行 7 :未转化的植物提取物 +1000ng HA(PSC) ;行 8 :未转化的植物提取物 +100ng HA(PSC) ;行 9 :未转化的植物提取物 +10ng HA(PSC) ;行 10 :未转化的植物提取物 +1ng HA(PSC) ;行 11 :表达构建体 540(跨膜 rHA) 的植物提取物 ;行 12 :表达构建体

544(与 GCN4 融合的可溶性 rHA)的植物提取物。

[0041] 优选实施方案的描述

[0042] 本发明涉及可溶性重组流感抗原的生产。更具体地,本发明涉及保留有免疫原性的可溶性重组流感抗原的生产。

[0043] 以下为对优选实施方案的描述。

[0044] 本发明提供了含血凝素结构域和寡聚化结构域的重组血凝素(rHA)。该重组蛋白作为可溶性同三聚体而产生。所述 rHA 还可包含信号肽和/或内质网(ER)驻留信号。

[0045] 流感由流感病毒引起,流感病毒可分为甲、乙或丙型。甲型和乙型为最经常与流行相关的类型。甲型流感病毒可根据所存在的血凝素(HA)和神经氨酸酶(NA)表面糖蛋白的组合进一步分成亚型。目前,已确认了 16 种 HA(H1-H16)和 9 种 NA(N1-N9)亚型。每种甲型流感病毒存在一种类型的 HA 和一种类型的 NA 糖蛋白。

[0046] 术语重组血凝素(也称为“重组 HA”或“rHA”)是指由本领域技术人员熟知的重组技术产生的血凝素蛋白。血凝素(HA)为在甲型流感病毒上发现的病毒表面蛋白。迄今为止,已经确定了 16 个 HA 亚型(H1-H16)。HA 负责使病毒与所感染宿主细胞表面碳水化合物部分的唾液酸残基结合。在细胞将病毒内吞后,HA 蛋白经历剧烈的构型变化,这引起病毒与细胞膜融合并使病毒进入细胞。

[0047] HA 为同三聚体的 I 型膜糖蛋白,通常包含信号肽、HA_0 结构域、C 末端的跨膜锚定位点和小胞质尾区(图 1)。术语“同三聚体”指由三个 HA 蛋白分子形成的寡聚体。HA 蛋白以 75kDa 单体前体蛋白(HA_0)形式合成,后者在表面组装成伸长的三聚体蛋白。在三聚体化之前,将前体蛋白 HA_0 在保守的活化切割位点(也称为融合肽)切割为通过二硫键连接的 2 个多肽链,HA1(328 个氨基酸)和 HA2(221 个氨基酸)。尽管该步骤对病毒感染很重要,但其对蛋白的三聚体化并不是必要的。HA 在宿主细胞内质网(ER)膜中的插入、信号肽切割和蛋白糖基化为共翻译事件。HA 的正确折叠需要蛋白的糖基化和形成 6 个链间二硫键。HA 三聚体在高尔基复合体顺面和反面组装,跨膜结构域在三聚化过程中起作用。经菠萝蛋白酶处理的 HA 蛋白(缺少跨膜结构域)的晶体结构在流感病毒株中显示了高度保守的结构(Russell 等,2004)。还已经确认了 HA 在感染过程中经历显著的构象变化,这需要前体 HA_0 被切割为 2 个多肽链 HA1 和 HA2。

[0048] 本发明的重组 HA 可以为任何亚型。例如,HA 可为亚型 H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15 或 H16。本发明的 rHA 可包含基于现有技术中已知任何血凝素序列的氨基酸序列。此外,所述 rHA 还可以基于从新出现的流感病毒中分离出的血凝素序列。

[0049] 本发明的 rHA 可为嵌合蛋白构建体,其包含血凝素结构域和寡聚化结构域。术语“血凝素结构域”指含 HA_0 结构域或者 HA1 和 HA2 结构域的氨基酸序列。也就是说,所述 rHA 蛋白可被加工(即包含 HA1 和 HA2 结构域),或者可以未被加工(即包含 HA_0 结构域)。血凝素结构域不包括在天然蛋白中发现的信号肽、跨膜结构域或胞质尾区。“寡聚化结构域”(也称为“三聚体肽”)指促进 rHA 蛋白寡聚化的结构域。所述寡聚化结构域可以是本领域已知的促进三聚体形成的任何氨基酸序列。例如,所述寡聚化结构域可为异源肽,例如亮氨酸拉链或采取卷曲螺旋结构的肽。所述寡聚化结构域可与其所替换的跨膜结构域(长度为 26 个氨基酸)具有相似的长度和/或结构。或者,所述寡聚化结构域可与其所替换的跨膜结

构域（26 个氨基酸）和胞质尾区（10 个氨基酸）具有相似的长度和 / 或结构。例如（但不希望有限制），所述寡聚化结构域的长度可为约 20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49 或 50 个氨基酸。在一个具体的非限定性实施例中，所述寡聚化结构域的长度为约 25-48 个氨基酸，或其间的任何量。合适的寡聚化结构域的非限定性实例包括 GCN4-pII 肽 (Harbury et al，1993，Science，262：1401-7)、γ 玉米醇溶蛋白的脯氨酸富含结构域 (PRD)、或噬菌体 T4 纤维蛋白 (fibritin) (Strelkov 等，1996，Virology 219：190-194)，或在纤连蛋白或胶原凝集素家族中鉴定的三聚化组件。

[0050] 在一个具体的非限定性实施例中，所述寡聚化结构域可为 GCN4-pII 肽，其为 GCN4 酵母亮氨酸拉链的变体。该 GCN4 突变体在 α 螺旋上每 7 个氨基酸出现的 a 和 d 位含 IIe 残基 (pII)，这导致三聚化的高发倾向。所述三聚体的熔解温度 (Tm) 大于 100℃，使得该寡聚体具有高度的内在稳定性 (Harbour 等）。图 2 中显示了 GCN4-pII 的氨基酸序列。GCN4-pII 非常适合用作寡聚化结构域；GCN4-pII 的该 29 个氨基酸序列置于 HA 结构域序列的 C 末端，基本上替换了 26 个氨基酸的所述跨膜结构域。如需要，可在 rHA 的 C 末端放置额外的氨基酸序列，以使重组结构不被 α 螺旋终止。例如（但不希望有限制），可在 GCN4-pII 的 C 末端加入 Ser-Ala-Ala 氨基酸残基。

[0051] 在另一个实施例中，所述寡聚化结构域可为 γ 玉米醇溶蛋白的 PRD（在本文中也称为“PRD”）。已知 γ 玉米醇溶蛋白一旦进入 ER 就大量贮存在蛋白质体中。合成的 PRD 肽采取两亲性的聚脯氨酸 II 构象，其可将自身组装为三聚体 (Kogan 等，2002，Biophysical J.，83：1194-1204)。在其天然形式中，PRD 包含 8 个重复的肽 PPPVHL(SEQ ID NO：2)。γ 玉米醇溶蛋白的 PRD 肽也置于 HA 结构域的 C 末端，替换跨膜锚定区。由于天然形式的 PRD 相当长，因此提供不同肽长度 (4、6 或 8 肽重复，即 24、36 或 48 氨基酸长度）的 PRD 也在本发明的范围内。如果需要，可在 HA 结构域和 PRD 之间放置氨基酸连接子，以使 PRD 肽链的取向为朝向聚脯氨酸左螺旋。可使用现有的任何合适的肽连接子。例如（但不希望有限制），可使用诸如 Gly-Gly-Ala-Gly(SEQ ID NO：3) 的四肽。如果需要，还可在 rHA 的 C 末端放置额外的氨基酸序列，以使重组结构不被 α 螺旋终止。例如（但不希望有限制），可在 PRD 的 C 末端加入 Ser-Ala-Ala 氨基酸残基。

[0052] 在另一个非限制性实施例中，所述寡聚化结构域可为噬菌体 T4 纤维蛋白 (Strelkov 等，1996，Virology 219：190-194)。该结构域包含纤维蛋白 C 末端的最后 29 个氨基酸残基。

[0053] 在另一个非限制性实施例中，所述寡聚化结构域还可为在 WO98/56906（其通过引用并入本文）中公开的三聚化组件，其为在四联蛋白 (tetranectin) 家族中鉴定的三聚化组件。该四联蛋白三聚化组件也显示出稳定性，体现在该三聚体可在 60℃下、或甚至在 700℃下存在。该三聚化组件可与 rHA 共价连接，并能与其它两个三聚化组件形成稳定的复合物。寡聚化结构域的另外一个实例为在 WO 95/31540（其通过引用并入本文）中公开的三聚化肽，其在胶原凝集素家族中确定。该肽长度为约 25 至 40 个氨基酸，其源自胶原凝集素家族蛋白的颈部区域。

[0054] 本发明的 rHA 可进一步包含信号肽。所述信号肽可为任何现有技术中已知的任何合适的肽，以将重组蛋白引导到目标细胞区室或膜。例如（但不希望有限制），可采用在天

然HA中发现的信号肽，其将HA引导到ER。在另一个非限制性实施例中，所述信号肽可以是PDI，即苜蓿信号肽。图3中显示了PDI的氨基酸和核苷酸序列。有利地，所述PDI信号肽包含bglII限制性位点，其可以对克隆有帮助。

[0055] 上述rHA还可包含内质网(ER)驻留信号。可以使用本领域技术人员已知的任何合适的ER驻留信号。例如（但不希望以任何形式成为限制），可以使用Ser-Glu-Lys-Asp-Glu-Leu(SEKDEL；SEQ ID NO：4)或His-Asp-Glu-Leu(HDEL；SEQ ID NO：5)ER驻留信号。所选择的ER驻留信号可在rHA蛋白序列的C末端。有利地，在几种情况下已显示植物中重组蛋白的ER驻留可将表达水平提高2至10倍(Schillberg等，2003，Cell MoI.Life Sci.60：443-445)。不希望受理论局限，ER驻留信号可使蛋白在ER和高尔基复合体之间来回运动，这允许三聚化发生。

[0056] 术语“可溶性”指rHA在宿主细胞中以可溶形式产生。如上所述，重组HA向可溶形式的转化产生于以下过程：将疏水性的跨膜结构域替换为可溶性的α螺旋，后者与HA结构域在结构上相容。以可溶形式表达rHA可提高产量（更高的表达水平），并降低纯化的复杂性，因此降低生产成本。

[0057] 本发明还提供了编码上述rHA的核酸。该核酸为包含编码血凝素结构域(HA_0)之核苷酸序列和编码寡聚化结构域之核苷酸序列的嵌合构建体。编码rHA的核酸还可包含编码信号肽的核苷酸序列和/或编码ER驻留信号的核苷酸序列。

[0058] 本发明进一步涉及包含编码上述rHA之核酸的嵌合基因构建体，所述核酸与调控元件有效连接。“调控元件”或“调控区域”指一部分核酸，其通常（但不总是）为基因的上游，并可包含DNA或RNA，或者同时包含DNA和RNA。调控元件可包括能够调节器官特异性或控制发育或时序基因激活的元件。进一步讲，“调控元件”包括启动子元件、核心启动子元件、对外界刺激产生响应的可诱导元件、组成性激活的元件，或者降低或升高启动子活性的元件（分别如负向调控元件或转录增强子）。显示调控元件活性的核苷酸序列是指当与目的编码序列有效连接时，起到启动子、核心启动子、组成型调控元件、负向元件或沉默子（即降低启动子活性的元件）、或者转录或翻译增强子作用的核苷酸序列。

[0059] “有效连接”是指特定的序列（例如调控元件和目的编码区域）直接或间接地相互作用以执行目的功能，如介导或调解基因表达。有效连接序列的相互作用可例如由以有效连接序列相互作用的蛋白质所介导。

[0060] 本文所用的调控元件还包括在转录起始或转录后有活性的元件，例如调节基因表达的调节性元件，如翻译和转录增强子，翻译和转录抑制子，以及mRNA稳定性或不稳定性决定子。在本发明中，术语“调控元件”还指DNA序列，其通常但不总是为结构基因编码序列的上游(5')，其包含通过提供识别RNA多聚酶和/或其它转录所需因子的特定位点来控制编码区域之表达的序列。为RNA多聚酶和/或其它转录因子提供识别以保证在特定位点起始的调控元件的一个实例为启动子元件。启动子元件包含核心启动子元件（负责转录的起始）及调节基因表达的其它调控元件（如上面列举的）。应该理解，位于内含子中或编码区域序列3'的核苷酸序列也可能有助于调节目标编码区域的表达。调控元件还可包括那些位于转录起始位点下游(3')或在转录区域中或者在以上两者中的元件。在本发明中，转录后调节元件可包括在转录起始后有活性的元件，例如翻译和转录增强子，翻译和转录抑制子，以及mRNA稳定性决定子。

[0061] 调控元件或其片段可与异源调控元件或启动子有效关联（有效连接），以调节所述异源调控元件的活性。该调节包括增强或抑制异源调控元件的转录活性、调节转录后事件，或增强／抑制异源调控元件的转录活性和调节转录后事件这两者。例如，一个或多个调控元件或其片段可与组成型、可诱导型、组织特异性启动子或其片段，或者调控元件片段有效关联，例如但不限于，TATA 或 GC 序列可与本发明的调控元件有效关联，以调节这些启动子在植物、昆虫、真菌、细菌、酵母或动物细胞中的活性。

[0062] 有几种类型的调控元件，包括发育调控的、可诱导的和组成型的元件。受发育调控的调控元件，或控制在其控制下基因的差异表达的调控元件，在器官或组织发育期间在特定时间在特定器官或组织中被激活。然而，一些受发育调控的调控元件可优选地在特定发育阶段在某些器官或组织中被激活，它们还可以发育调节方式被激活，或在植物的其它器官或组织中为基础水平。

[0063] “启动子”指在编码区域或其片段的 5’端的核苷酸序列，其包含转录起始和转录速率调节的所有必需信号。通常有两种类型的启动子，可诱导启动子和组成型启动子。如果需要组织特异性基因表达，例如种子或叶的特异性表达，则还可采用对这些组织特异性的启动子。

[0064] 可诱导启动子为能够响应于诱导物而直接或间接激活一个或多个 DNA 序列或基因转录的启动子。在没有诱导物时 DNA 序列或基因不被转录。通常，与可诱导启动子特异性结合以激活转录的蛋白因子以非活性形式存在，其随后被诱导物直接或间接转化为活性形式。诱导物可为化学物质如蛋白质、代谢产物、生长调节剂、除草剂或酚类化合物，或者通过热、冷、盐或毒性元素所直接施加的或者通过病原体或疾病物质（如病毒）的作用而间接施加的生理胁迫。含有可诱导启动子的植物细胞可通过对所述细胞或植物外部施加诱导物（例如通过喷洒、灌溉、加热或类似方法）来使其暴露于诱导物。可诱导启动子的实例包括但不限于植物启动子，如：光调控的苜蓿质体蓝素 (plastocyanine) 启动子（例如见 WO01/025455）；苜蓿亚硝酸还原酶启动子 (NiR；例如见 WO01/025454，其可通过使用硝酸盐类施肥而诱导）(3)；苜蓿脱水蛋白 (dehydrine) 启动子（美国申请 60/757,486），其可由环境胁迫（如寒冷）所诱导。

[0065] 组成型启动子在植物的各个部分以及持续地在植物发育过程中指导基因的表达。任何合适的组成型启动子都可用于在宿主有机体的转化细胞、或者所有器官或组织、或者所有的器官和组织中驱动 rHA 的表达。已知的组成型启动子的实例包括与 CaMV 35S 转录本相关的那些启动子 (Odell 等，1985，Nature，313：810-812）、稻米肌动蛋白 1(Zhang 等，1991，Plant Cell，3：1155-1165) 和磷酸丙糖异构酶 1(Xu 等，1994，Plant Physiol. 106：459-467) 基因、玉米泛素 1 基因 (Cornejo 等，1993，Plant Mol. Biol. 29：637-646）、拟南芥 (Arabidopsis) 泛素 1 至 6 基因 (Holtorf 等 1995，PlantMol. Biol. 29：637-646) 以及烟草转录起始因子 4A 基因 (Mandel 等，1995Plant Mol. Biol. 29：995-1004）。

[0066] 本文所用术语“组成型”不一定表示基因在所有细胞类型中都以相同的水平表达，但是该基因在各种各样的细胞类型中表达，尽管经常可观察到在丰度方面的一些变化。

[0067] 本发明的嵌合基因构建体还可包含 3’非翻译区。3’非翻译区指含 DNA 片段的基因部分，其包含多聚腺苷酸化信号或能够实现 mRNA 加工或基因表达的任何其它调控信号。多聚腺苷酸化信号通常特征在于，在 mRNA 前体的 3Y 末端添加多聚腺苷酸尾。多聚腺苷酸

化信号通常通过存在典型形式 5’-AATAAA-3’ 的同源物而进行识别，尽管变体并非不常见。

[0068] 合适的 3’区域的实例为包含农杆菌属 (Agrobacterium) 肿瘤诱导 (Ti) 质粒基因之多聚腺苷酸化信号的 3’转录非翻译区，所述基因例如胭脂碱合成基因 (Nos 基因) 和植物基因 (例如大豆储存蛋白基因和) 以及核酮糖 -1,5- 二磷酸羧化酶小亚基 (ssRUBISCO) 基因。因此，本构建体结构基因的 3’非翻译区可用于构建在植物中表达的嵌合基因。其它合适的 3’区域的实例为终止子，其可包括但不限于质体蓝素或亚硝酸还原酶或脱水蛋白苜蓿基因序列的 3’非编码区。

[0069] 如果需要，本发明的嵌合基因构建体还可进一步包含增强子、翻译或转录增强子。本领域技术人员熟知这些增强子区域，其可包括 ATG 起始编码子和邻近序列。起始编码子必须与编码序列的阅读框相同以确保全部序列的翻译。翻译控制信号和起始密码子可来自多种天然和合成来源。可从转录起始区域或从结构基因中获取翻译起始区域。该序列还可源自所选的用于表达基因的调控元件，并可进行特别的调节以增加 mRNA 的转录。

[0070] 包含含有编码本发明 rHA 之核酸序列的嵌合基因构建体的植物、植物部分或组织、植物细胞、树木、树木部分、树木细胞、酵母、细菌、真菌、昆虫和动物细胞也被认为是本发明的一部分。然而，应当理解，本发明的调控元件也可与目的编码区域结合以在一系列易于转化的宿主有机体中表达。这些宿主有机体包括但不限于 ：

[0071] 植物 (单子叶和双子叶植物)，例如，玉米、谷类植物、小麦、大麦、燕麦、烟草、芸苔、大豆、豆、豌豆、苜蓿、马铃薯、番茄、人参、拟南芥 ；这些植物的部分或组织，例如，叶、根、茎、分生组织、花结构，这些植物的细胞，以及

[0072] - 酵母、真菌、昆虫、动物和细菌细胞。

[0073] 这些有机体的转化和再生的方法已在现有技术中建立，并且为本领域技术人员所知，获得转化和再生植物的方法对本发明而言并不关键。

[0074] “转化”指在种间转移遗传信息 (核苷酸序列)，其可表现为基因型、表型或两者皆有。遗传信息从嵌合构建体到宿主的种间转移可为可遗传的，并且遗传信息的转移被认为是稳定的 ；或者所述转移可为瞬时的，并且遗传信息的转移不可遗传。

[0075] 由植物细胞再生整个植物的方法也是现有技术中已知的。一般来讲，将转化的植物细胞在合适的培养基中培养，所述培养基可含有选择试剂 (如抗生素)，其中使用可选择标记物以有助于对转化植物细胞的鉴定。一旦愈伤组织形成，就可按照已知方法采用合适的植物激素来促进芽的形成，并将芽转移到生根培养基来进行植物的再生。然后可采用该植物来建立来自用种子或利用营养繁殖技术的可繁殖世代。

[0076] [01] 可利用 Ti 质粒、Ri 质粒、植物病毒载体、DNA 直接转化、微注射、电穿孔等将本发明的构建体引入植物细胞中。关于这些技术的综述可见例如 Weissbach 和 Weissbach, Methods for Plant Molecular Biology, Academy Press, New York VIII,421-463 页 (1988) ;Geierson 和 Corey, Plant Molecular Biology, 第二版 .(1988) ; 以及 Miki 和 Iyer,Fundamentalsof Gene Transfer in Plants.Plant Metabolism,第二版.DT.Dennis, DHTurpin,DD Lefebrve,DB Layzell(编),Addison Wesly,Langmans Ltd.London,561-579 页 (1997)。其它方法包括 DNA 直接摄取、采用脂质体、电穿孔，例如采用原生质体、微注射、微弹 (microprojectile) 或颈须 (whisker)，以及真空渗入。参见例如 Bilang 等 (Gene 100 :247-250(1991), Scheid 等 (Mol.Gen.Genet.228 :104-112,1991), Guerche 等 (Plant

Science52 :111-116,1987), Neuhause 等 (Theor. Appl Genet. 75 :30-36,1987), Klein 等, Nature 327 :70-73(1987) ;Howell 等 (Science 208 :1265,1980), Horsch 等, (Science 227 :1229-1231,1985), DeBlock 等, Plant Physiology 91 :694-701,1989), Methods for Plant Molecular Biology(Weissbach 和 Weissbach 编, Academic Press Inc.,1988), Methods in Plant MolecularBiology(Schuler 和 Zielinski 编, Academic Press Inc., 1989), Liu 和 Lomonossoff(J VirolMeth,105 :343-348,2002,), 美国专利 4,945,050 ; 5,036,006 和 5,100,792,1995 年 5 月 10 日提交的美国专利申请 08/438,666 以及 1992 年 9 月 25 日提交的美国专利申请 07/951,715(所有这些参考文献均通过引用并入本文)。

[0077] 如下所述, 可采用瞬时表达方法来表达本发明的构建体 (参见 Liu 和 Lomonossoff,2002, Journal of Virological Methods,105 :343-348 ;其通过引用并入本文)。或者, 可采用基于真空的瞬时表达方法 (如 Kapila 等所述,1997, 其通过引用并入本文)。这些方法可包括例如但不限于农杆菌接种 (Agro-inoculation) 或农杆菌渗入 (Agro-infiltration)、注射浸润 (synringe infiltration) 的方法, 但是也可使用上述其它瞬时方法。利用农杆菌接种、农杆菌渗入或注射浸润, 可使含有所需核酸的农杆菌的混合物进入组织的细胞间空隙内, 所述组织例如叶、植物的气生部分 (包括茎、叶和花)、植物的其它部分 (茎、根、花) 或整个植物。在通过表皮后, 农杆菌感染 t-DNA 并将 t-DNA 拷贝转移进细胞。t-DNA 以附加体转录, mRNA 翻译, 使得在感染细胞中产生目的蛋白, 但是, t-DNA 在核中的传代为瞬时的。

[0078] 为了帮助鉴定转化的植物细胞, 可进一步处理本发明的构建体使其包含植物可选择标记物。可用的可选择标记物包括提供对抗生素之抗性的酶, 所述抗生素例如庆大霉素、潮霉素、卡那霉素等。类似地, 提供可通过颜色变化鉴定化合物之生成的酶 (如 GUS(β - 葡糖苷酸酶) 或荧光的酶 (例如萤光素酶) 也是可用的。

[0079] 当在本发明中涉及具体序列时, 应该理解, 这些序列在其范围内包括与所述具体序列“基本同源”的序列, 或与本文定义的一个或多个核苷酸序列在严格杂交条件下杂交的互补序列。只要这些同源序列表现出本文公开的一个或多于一个调控元件的活性, 当在确定长度的核苷酸序列中有至少 70%, 或更优选 75% 的核苷酸序列匹配时, 这些序列就是“基本同源”的。

[0080] 可利用核苷酸序列比较程序来确定这些序列的相似度, 所述程序例如在 DNASIS 中提供的 (采用例如但不限定于以下的参数 :GAP penalty 5,#0f top diagonals 5,fixed GAP penalty 10,k-tuple 2,floating gap 10 andwindow size 5)。但是, 在现有技术中还熟知其它序列比较的比对方法, 例如, Smith & Waterman 算法 (1981, Adv. Appl. Math. 2 : 482)、Needleman & Wunsch(J. Mol. Biol. 48 :443,1970)、Pearson & Lipman(1988, Proc. Nat′ l. Acad. Sci. USA 85 :2444), 以及这些算法的计算机执行 (GAP、BESTFIT、FASTA 和 BLAST, 可通过 NIH 获得), 或者通过人工比对并目检的方法 (参见例如 Current Protocols in Molecular Biology, Ausubel 等编,1995 年增刊), 或在严格条件下用 Southern 或 Northern 杂交 (参见 Maniatis 等, Molecular Cloning(A Laboratory Manual), ColdSpring Harbor Laboratory,1982)。优选地, 基本同源的序列在确定长度的该分子中显示出至少 80%, 以及最优选至少 90% 的序列相似性。

[0081] 该严格杂交条件的一个实例可为在 65℃下在 4X SSC 中过夜杂交 (约为 16-20 小

时），而后在65℃下在0.1X SSC中清洗1小时，或在65℃下在0.1X SSC中清洗2次，每次20或30分钟。或者，示例性的严格杂交条件可为在42℃下50％甲酰胺、4X SSC中过夜(16-20小时），而后在65℃下在0.1X SSC中清洗1小时，或在65℃下在0.1X SSC中清洗2次，每次20或30分钟，或过夜（16-20小时）；或者对于独特的序列区域而言，在65℃下在Church磷酸盐缓冲液（7％ SDS；0.5M $NaPO_4$缓冲液pH 7.2；10mM EDTA）中杂交，在50℃下在0.1X SSC、0.1％ SDS中清洗2次，每次20或30分钟，或在65℃下在2X SSC、0.1％ SDS中清洗2次，每次20或30分钟。

[0082] 本发明的rHA可与已有的流感疫苗联合使用，以补充疫苗、使它们更有效，以及减少所需的施用剂量。本领域技术人员已知，疫苗可靶向一种或多于一种流感病毒。合适的疫苗的实例包括但不限于那些可从Sanofi-Pasteur、ID Biomedical、Merial、Sinovac、Chiron、Roche、Medlmmune、GlaxoSmithKline等购得的。

[0083] 现在将参考以下非限制性实施例对本发明进行详细描述。

[0084] 实施例1：rHA策略

[0085] 在本发明中选择作为示例的HA获取自流感病毒株A/NewCaledonia/20/99(H1N1)。A/New Caledonia/20/99 HA已被充分表征，并且可购得免疫检测工具。由于所有HA的三维结构都相当保守，因此本表达策略可直接用于任何其它HA。

[0086] rHA的表达采用苜蓿PDI信号肽（SEQ ID NO：3）靶向ER和分泌途径。该信号肽直接与成熟rHA的N末端一级序列（SEQ ID NO：8）融合。用SignalP server 3.0验证了有PDI信号肽的成熟rHA之信号肽的理论性切割。

[0087] 可使用SEKDEL(SEQ ID NO：4)或HDEL(SEQ ID NO：5)驻留信号（两者均与rHA蛋白序列的C末端融合）将rHA驻留在ER中。

[0088] 当采用GCN4-pII肽作为寡聚化结构域时，其直接与HA0最后的Tyr残基的C末端融合。将GCN4-pII第1位的Met残基改变为Leu，因为Leu比Met更具惰性，并且其与起初的Met相比有相似的包装体积。在GCN4-pII的C末端添加氨基酸Ser-Ala-Ala。

[0089] 当采用γ玉米醇溶蛋白的PRD时，将含有8个重复的PPPVHL(SEQID NO：2)的肽与HA融合，以替换HA的跨膜结构域。在融合中还引入了前4个脯氨酸残基。去除了C末端半胱氨酸。为了适应与HA C末端聚脯氨酸螺旋的融合，在PRD的N末端添加了肽Gly-Gly-Ala-Gly(EQ IDNO：3)。在PRD的C末端添加了Ser-Ala-Ala肽。

[0090] 总共制备了8种不同的rHA基因构建体以检测它们在植物中的表达：

[0091] 1. 全长rHA，包括含PDI信号肽（SEQ ID NO：9）的跨膜结构域；

[0092] 2. 从项目1的rHA中去除跨膜结构域和胞质尾区，并用驻留信号SEKDEL(SEQ ID NO：10）替换；

[0093] 3. 从项目1的rHA中去除跨膜结构域和胞质尾区，并用驻留信号HDEL(SEQ ID NO：11）替换；

[0094] 4. 从项目1的rHA中去除跨膜结构域和胞质尾区（SEQ ID NO：12）；

[0095] 5. 从项目1的rHA中去除跨膜结构域和胞质尾区，并用GCN4-pII(SEQID NO：13）替换；

[0096] 6. 从项目1的rHA中去除跨膜结构域和胞质尾区，并用在C末端有驻留信号SEKDEL(SEQ ID NO：14）的GCN4-pII替换；

[0097] 7. 从项目 1 的 rHA 中去除跨膜结构域和胞质尾区，并用 PRD 结构域 (SEQID NO :15) 替换 ;以及

[0098] 8. 从项目 1 的 rHA 中去除跨膜结构域和胞质尾区，并用在 C 末端有驻留信号 SEKDEL(SEQ ID NO :16) 的 PRD 结构域替换。

[0099] 实施例 2 :基因合成

[0100] 合成了在实施例 1 中表述的 8 种不同的基因构建体。对流感病毒株 A/New Caledonia/20/99 的野生型 HA 核苷酸序列只进行修饰以插入或去除限制性位点，以便于克隆步骤。这些对基因序列的微小修饰不会改变所得蛋白质的氨基酸序列。合成的 HA_0 基因在 5'端有 ApaI 限制位点，与 PDI 信号肽的 ATG 同相。其它 7 个基因在其 5'端有 KpnI 限制位点。每个基因在 3'端以终止密码子 (TAA) 终止，其后是 SacI 和 StuI 限制位点。按照设计好的克隆策略，以下限制位点在最终的质粒中被设计为唯一的 :ApaI、BstZ17 I、Sac I、Kpn I、BgI II 和 Stu I。在植物生物技术研究所 (Plant Biotechnology Institute) 合成了不同的基因片段，并转运克隆进 pUC18 中。在图 6 中给出了所选的核苷酸序列信息 (SEQ ID NO :17-24)。

[0101] 实施例 3 :表达 rHA 融合蛋白的载体的遗传组装

[0102] 采用已知的 DNA 重组技术 (Sambrook 等 (1989)Molecular cloning :Alaboratory Manual，第二版，Cold Spring Harbor Laboratory press，ColdSpring Harbor，NY) 组装了多种 DNA 构建体。除非另外说明，否则采用大肠杆菌 (E.coli)DH5α 菌株作为扩增二元载体的宿主。利用 bglII 和 SacI 限制位点直接将 HA0 基因片段连接进 plastocynine 盒 (已经含有 pdi 信号肽) 中，或者采用 KpnI、bglII 和 SacI 限制位点与其它 7 个 C- 末端共连接。在每种情况下，都将基因插入到二元载体上的转移 DNA(t-DNA) 所含的 plastocynine 表达盒中。t-DNA 为通过转化整合入宿主植物细胞基因组 (从右边缘到左边缘) 的二元载体上的一部分。在本实施例中所用的二元载体来自商品化质粒 pCAMBIA 2300(Cambia，Canberra，Au)，并含以下元件 (从 t-DNA 的右边缘到左边缘) :

[0103] - 其中插入 rHA 基因的多个专有表达盒 (启动子和终止子)，其可获自 Medicago ;

[0104] - 多克隆位点以便于克隆 ;以及

[0105] - 编码新霉素磷酸转移酶 II(nptII) 的基因，可在胭脂碱合成酶 (NOS) 启动子和花椰菜 35S 花叶病毒 (35S) 终止子的控制下提供对卡那霉素抗性的可选择标记物。

[0106] 将每个克隆进行 DNA 测序以测定该克隆的核苷酸序列。如果所得的氨基酸序列仍编码完全相同的蛋白序列 (SEQ ID NO :9-16)，则可允许 rHA 基因的序列有核苷酸的变化。在表 1 中列出了 8 种不同的 DNA 构建体，并且在图 7 中示意性显示了所得到的 rHA t-DNA 的实例。

[0107] 表 1 :rHA 的 DNA 构建体

[0108]

编号	表达策略	亚细胞靶向	ER 驻留信号	SEQ ID NO :
540	完全跨膜分泌的	质外体		9
541	ER 驻留的	ER	SEKDEL	10

542	ER 驻留的	ER	HDEL	11
543	可溶的,分泌的	质外体		12
544	可溶性三聚体 PRD	质外体		13
545	可溶性三聚体 PRD	ER	SEKDEL	14
546	可溶性三聚体 PRD	质外体		15
547	可溶性三聚体 PRD	ER	SEKDEL	16

[0109] 最终将从大肠杆菌分离的 DNA 插入根瘤农根菌 (Agrobacteriumtumefaciens) Ag11 或其它菌株中。表 1 中显示了含有多种本发明所用的 t-DNA 的农杆菌 (Agrobacterium) 克隆列表。在植物瞬时表达试验的准备中,将质粒 DNA 从农杆菌克隆中分离出来,并进行限制消化作图以确保 DNA 的完整性。

[0110] 实施例 3 :在烟草中瞬时表达 rHA

[0111] 农杆菌培养在 28℃下于 2ml 分别含有 25μg/mL 羧苄青霉素和 50μg/mL 卡纳霉素的 YEB 或 LB 培养基中培养包含含有 rHA DNA 构建体 (表 1) 之二元载体的农杆菌克隆 24 小时。用 10μL 该培养物作为生成 25mL YEB 诱导培养基 (YEB 培养基,10mM 2(N- 吗啉基) 乙磺酸 (MES),pH 调节为 5.6,25mg/L 羧苄青霉素,50mg/L 卡那霉素,20μM 乙酰丁香酮) 之培养物的起始接种物。将后者在 28℃下在旋转震荡 (220rpm) 培养箱中培养 18 小时,或者直至达到 600nm 下的光密度 (OD600) 为 0.8-1。

[0112] 非转基因烟草的培养 在温室中以泥炭为主的基质 (AgroMix) 中以种子培育本塞姆氏烟草和普通烟草。幼苗起初在培育室中生长,而后移植到盆中。植株每天灌溉两次,每次施用时接受 180ppm 的氮。温室条件保持在白天 25℃,夜间 21℃,采用长的日光周期方案 (16 小时光照 /8 小时黑暗循环),采用 20 瓦每平方米植株水平的人工照明。

[0113] rHA 烟草的瞬时表达将按照如前所述制备的农杆菌培养物以 10 000g 离心 8 分钟,重悬于相同体积的接种培养基 (10mM MgCl2,10mMMES, pH 调节为 5.6,并补充 100μM 乙酰丁香酮) 中,并在接种前在室温 (RT,23℃) 下放置 1 小时。或者,可在接种前使悬浮液在 4℃保持 24 小时。本塞姆氏烟草和普通烟草的瞬时转化基本上按照 Liu 和 Lomonossoff 描述的 (2002, Journal of Virological Methods,105 :343-348) 进行,并做以下修改。对于 rHA 的表达而言,接种了两种农杆菌株的混合物。第一种菌株含有表 1 中所描述的克隆之一,第二种菌株含有在 35S 启动子控制下的马铃薯病毒 Y 的 HcPro 沉默抑制子。接种后,将植物在温室中孵育。温度保持在白天最低 25℃,夜间最低 21℃。植物一天灌溉两次,每次施用时接受 180ppm 的氮。在 4-8 天后进行生物质的收获。

[0114] 实施例 5 :证实烟草叶中 rHA 的表达

[0115] 从接种的叶中制备可溶性蛋白提取物在收获后直接分析叶或在 -80℃冷冻生物质后分析叶。称量约 0.1g 的农杆菌接种叶的植物生物质,并用于分析 rHA 瞬时表达。

[0116] 采用以下两种提取方法之一来制备总蛋白提取物 :用研钵和研杵研磨植物组织,或通过在 Retsch 的 MixerMill300(MM300) 中将其粉碎。将 0.1g 的植物生物质转移到洁净

且预冷的研钵中。加入 0.3mL 冷的提取缓冲液 (Tris 50mM pH 7.4,含 NaCl 150mM,0.1% Triton X-100 和 5% 甘油) 及 PMSF(终浓度为 1mM) 和抑凝乳蛋白酶素 (chymostatin)(终浓度为 10μM)。用研杵磨碎叶直到获得均质制备物。然后将植物提取物转移到 1.5mL 微管 (microtube) 中并以 20,000g 在 4℃下离心 20 分钟。或者,将含有 0.3mL 提取缓冲液的 0.1g 植物组织引入非灭菌的 1.5 微管中。向每个管中加入钨珠。将盒子在 30Hz 下进行 3 分钟一循环的搅拌。重复该循环 2 次。然后如上所述将植物提取物离心。

[0117] 离心后,将上清液转移进洁净的微管中并置于冰上。最后,通过 Bradford 方法使用 BSA 作为参比蛋白来测量每种蛋白提取物的总蛋白含量。

[0118] 在烟草叶材料中瞬时表达的 rHA 的免疫检测如上所述制备了接种了农杆菌的烟草叶蛋白提取物。还制备了以下 3 个对照 :1) 加入有 1μg BSA(Pierce) 的 1ng 纯 rHA(由 CHO 细胞产生的重组 rHA,Protein SciencesCorporation,catalog# 3006),2) 从接种有经空 pCAMBIA 质粒转化之农杆菌的农杆菌生物质中获得的 2-10μg 烟草蛋白提取物,以及 3) 掺入 1ng 纯 rHA 的 2-10μg#2 样品蛋白提取物。

[0119] 通常,将来自每种生物质的 2-10μg 蛋白提取物 (用 8 种不同的 DNA 构建体之一转化,参见表 1) 上样到 SDS-PAGE 上进行 Western 印迹分析。在上样到变性 SDS-PAGE(浓缩胶为 6% 丙烯酰胺 ;分离胶为 10% 丙烯酰胺,Tris- 甘氨酸缓冲液) 之前,将样品在 100℃下煮 5 分钟,基本上如 Thomas 所述进行 (1975, Proc.Nat.Acad.Sci.USA,72 :2626-2630)。用 Mini Protean III 装置 (BioRad) 在 110 伏下进行电泳 90 分钟。然后在 Tris 25mM,甘氨酸 (Glyine)192mM,甲醇 10% (v/v)pH 8.3 缓冲液中在 100V 下将分离的蛋白印迹在 PVDF 膜上,采用 Criterion blotter 或 Mini-trans blot Unit(BioRad),时间分别为 30 分钟或 60 分钟。

[0120] 用以下抗体实现 rHA 的免疫检测。使用甲型流感病毒的小鼠单克隆抗体 (Fitzgerald Industries International Inc. 目录号 10-150) 作为 1/1,000 的一抗。二抗为以 1/10,000 的过氧化物酶缀合的山羊抗小鼠 IgG(H+L)(Jackson ImmunoResearch #111-035-146)。用 Immobilon Western 化学发光 HRP 底物 (Milipore,目录号 WBRL50100) 或 BM 化学发光印迹底物 (POD)(Roche,目录号 11500 694 001) 检测过氧化物酶活性。

[0121] 优化了免疫检测的方法以获得对每个 rHA 的特异性信号,而对内源性植物蛋白则没有交叉反应活性。检测限很低 ;其允许检测到以低至大于 0.004% 的总可溶性蛋白 (TSP) 的水平 (25μg 提取物中的 1ng) 表达的蛋白。

[0122] 如图 8 中所显示的,rHA 在每个检测条件下和两个烟草种中都有表达。采用非还原条件允许检测血凝素的不同寡聚形式 :单体、二聚体和三聚体。如在图 9 中示例的,随着在凝胶上上样的蛋白质的量的降低,可分辨出 rHA 的寡聚形式。后者在所检测的每种构建体和两个烟草种中都相同 :在三聚体和单体水平上观察到免疫检测信号。还可观察到了第三个免疫反应条带,其迁移率在二聚体和三聚体预计分子量之间。

[0123] 将跨膜结构域替换为不同的三聚体的卷曲螺旋肽使 rHA 表达的表达水平等于或高于 rHA 的跨膜形式。在与 GCN4-pII 融合时尤其如此。与 GCN4-pII 融合相比,rHA 与 PRD 的融合允许降低水平的表达。GCN4-pII-rHA 融合蛋白的表达水平为至少 0.01% TSP(在 10μg TSP 中大于 1ng rHA)。与普通烟草相比,rHA 融合蛋白的积累似乎在本塞姆氏烟草中更多。最终,对于所检测的每种构建体而言,rHA 的内质网驻留显著提高了 rHA 的表达水

平。在很多其它蛋白中也观察到了该趋势。与其它区室相比，通常公认 ER 亚细胞区域的蛋白水解活性较低，这导致重组蛋白积累更高。

[0124] 对同样的蛋白提取物在还原条件下进行了分析并按照之前所述进行了免疫检测，以研究 HA_0 链是否被蛋白酶解为 rHA 的两链形式，即 HA1 和 HA2。后者的预计分子量分别为 47 和 28kDa。如图 10 中所示，HA_0 并不容易转化为两链形式。应当指出，在普通烟草提取物中 rHA 的蛋白酶解或切割似乎程度更大。

[0125] 实施例 6 ：证实 rHA 的三聚体组装

[0126] 血凝反应测试基本按照 Nayak 等的描述 (2004，J. Virol Methods，122 ：9-15) 进行血凝反应测定，进行了以下调整。在含有 1MM PMSV 和 10 μ M 抑凝乳蛋白酶素 (作为蛋白酶抑制剂) 的 50mM Tris pH7.4，150mM NaCl，0.1％ Triton X-100 中，通过组织匀浆作用制备提取物。将可溶性片段在 PBS 中稀释至 0.2mg/mL。在含有 100 μ L PBS 的圆底 96 微孔板中进行样品 (tample)(100 μ L) 的系列倍比稀释。向每孔中加入 100 μ L 鸡血红细胞 ($2X10^7$ rbc/mL)(Innovative Research，South Field，MI，USA)。然后将板在室温下孵育 60-120 分钟。测定包括以下阳性对照 ：来自 PSC 的 rHA(从 1 μ g 至 7.8pg)，在 PBS 中稀释或掺入到未转化提取物中。还测定了采用未转化提取物的阴性对照。

[0127] 使用血凝反应测定来证实与卷曲螺旋肽融合的 rHA 的三聚体组装 ：该测试尽管非常简单，但它是基于定量流感病毒滴度的常规基础而使用的。该检测的原理是 ：游离的血红细胞在 ELISA 孔底部滚动，形成容易检测到的血红细胞堆。如果发生血凝反应，则血细胞将形成板块状物 (matrix)，从而血细胞不能在 ELISA 孔底部滚动，并且不能观察到点。图 11 表明，鸡血红细胞的凝血反应可直接在表达于植物提取物中的 rHA 上进行。此外，利用表达与 GCN4-pII 融合之 rHA 的植物提取物，可发生血凝反应。由于 rHA 单体无法进行血红细胞的凝血反应，所以该结果最终证明所述嵌合 rHA 在植物中自身组装为三聚体。

[0128] 表 2 归纳了本申请中列出的序列。

[0129] 表 2 ：对序列识别符号的序列描述。

[0130]

SEQ ID NO:	序列描述
1	GCN4-pII 肽的氨基酸序列
2	γ 玉米醇溶蛋白的 PRD 肽的肽重复
3	四肽连接子
4	SEKDEL ER 驻留信号
5	HDEL ER 驻留信号
6	苜蓿 PDI 信号肽的氨基酸序列
7	苜蓿 PDI 信号肽的核苷酸序列
8	成熟 HA 的氨基酸序列
9	全长 rHA，包括跨膜结构域，且用 PDI 信号肽替换了 HA 信号肽
10	SEQ ID NO:9，其中用驻留信号 SEKDEL 替换了跨膜结构域和胞质尾区
11	SEQ ID NO:9，其中用驻留信号 HDEL 替换了跨膜结构域和胞质尾区
12	SEQ ID NO:9，无跨膜结构域和胞质尾区
13	SEQ ID NO:9，其中用 GCN4-pII 肽替换了跨膜结构域和胞质尾区

[0131]

14	SEQ ID NO:13，在 C 末端有 SEKDEL 驻留信号
15	SEQ ID NO:9，其中用 PRD 结构域替换了跨膜结构域和胞质尾区
16	SEQ ID NO:15，在 C 末端有 SEKDEL 驻留信号
17	HA_0 基因片段的核苷酸序列
18	跨膜结构域和胞质尾区基因片段的核苷酸序列
19	ER 驻留的 SEKDEL 基因片段的核苷酸序列
20	ER 驻留的 HDEL 基因片段的核苷酸序列
21	GCN4-pII 基因片段的核苷酸序列
22	ER 驻留的 GCN4-pII 基因片段的核苷酸序列
23	PRD 基因片段的核苷酸序列
24	ER 驻留的 PRD 基因片段的核苷酸序列

[0132] 本发明已通过一个或更多个实施方案进行了描述。然而，对本领域技术人员而言，显然可进行多种变化和调整而不背离权利要求中所定义的本发明的范围。

[0133] 在本申请中引用的所有参考文献和专利的内容均通过引用全文并入本文。

[0001]

序 列 表

<110> Medicago Inc.
COUTURE, Manon
VEZINA, Louis-Philippe
DARGIS, Michele
LANDRY, Nathalie

<120> 可溶性重组流感抗原

<130> V81901WO

<150> 61/078,963
<151> 2008-07-08

<160> 24

<170> PatentIn version 3.4

<210> 1
<211> 34
<212> PRT
<213> 人工

<220>
<223> GCN4-pII 肽的氨基酸序列

<400> 1

Leu Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr
1 5 10 15

His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Ser
20 25 30

Ala Ala

<210> 2
<211> 6
<212> PRT
<213> 人工

<220>
<223> 肽 PPPVHL 的重复序列

<400> 2

Pro Pro Pro Val His Leu
1 5

<210> 3
<211> 4
<212> PRT
<213> 人工

<220>
<223> 四肽连接子

[0002]

<400> 3

Gly Gly Ala Gly
1

<210> 4
<211> 6
<212> PRT
<213> 人工

<220>
<223> SEKDEL 驻留信号

<400> 4

Ser Glu Lys Asp Glu Leu
1 5

<210> 5
<211> 4
<212> PRT
<213> 人工

<220>
<223> HDEL 驻留信号

<400> 5

His Asp Glu Leu
1

<210> 6
<211> 24
<212> PRT
<213> 苜蓿 (Medicago sativa)

<400> 6

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala
20

<210> 7
<211> 72
<212> DNA
<213> 苜蓿

<400> 7
atggcgaaaa acgttgcgat tttcggttta ttgttttctc ttcttctgtt ggttccttct 60

cagatcttcg ct 72

<210> 8
<211> 565

[0003]

<212> PRT
<213> 流感病毒

<400> 8

```
Met Lys Ala Lys Leu Leu Val Leu Leu Cys Thr Phe Thr Ala Thr Tyr
1               5                  10                  15

Ala Asp Thr Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Asp Thr
            20                  25                  30

Val Asp Thr Val Leu Glu Lys Asn Val Thr Val Thr His Ser Val Asn
        35                  40                  45

Leu Leu Glu Asp Ser His Asn Gly Lys Leu Cys Leu Leu Lys Gly Ile
    50                  55                  60

Ala Pro Leu Gln Leu Gly Asn Cys Ser Val Ala Gly Trp Ile Leu Gly
65                  70                  75                  80

Asn Pro Glu Cys Glu Leu Leu Ile Ser Lys Glu Ser Trp Ser Tyr Ile
                85                  90                  95

Val Glu Thr Pro Asn Pro Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe
            100                 105                 110

Ala Asp Tyr Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe
        115                 120                 125

Glu Arg Phe Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr
    130                 135                 140

Val Thr Gly Val Ser Ala Ser Cys Ser His Asn Gly Lys Ser Ser Phe
145                 150                 155                 160

Tyr Arg Asn Leu Leu Trp Leu Thr Gly Lys Asn Gly Leu Tyr Pro Asn
                165                 170                 175

Leu Ser Lys Ser Tyr Val Asn Asn Lys Glu Lys Glu Val Leu Val Leu
            180                 185                 190

Trp Gly Val His His Pro Pro Asn Ile Gly Asn Gln Arg Ala Leu Tyr
        195                 200                 205

His Thr Glu Asn Ala Tyr Val Ser Val Val Ser Ser His Tyr Ser Arg
    210                 215                 220

Arg Phe Thr Pro Glu Ile Ala Lys Arg Pro Lys Val Arg Asp Gln Glu
225                 230                 235                 240

Gly Arg Ile Asn Tyr Tyr Trp Thr Leu Leu Glu Pro Gly Asp Thr Ile
                245                 250                 255

Ile Phe Glu Ala Asn Gly Asn Leu Ile Ala Pro Trp Tyr Ala Phe Ala
            260                 265                 270

Leu Ser Arg Gly Phe Gly Ser Gly Ile Ile Thr Ser Asn Ala Pro Met
        275                 280                 285

Asp Glu Cys Asp Ala Lys Cys Gln Thr Pro Gln Gly Ala Ile Asn Ser
    290                 295                 300
```

[0004]

Ser Leu Pro Phe Gln Asn Val His Pro Val Thr Ile Gly Glu Cys Pro
305 310 315 320

Lys Tyr Val Arg Ser Ala Lys Leu Arg Met Val Thr Gly Leu Arg Asn
325 330 335

Ile Pro Ser Ile Gln Ser Arg Gly Leu Phe Gly Ala Ile Ala Gly Phe
340 345 350

Ile Glu Gly Gly Trp Thr Gly Met Val Asp Gly Trp Tyr Gly Tyr His
355 360 365

His Gln Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Gln Lys Ser Thr
370 375 380

Gln Asn Ala Ile Asn Gly Ile Thr Asn Lys Val Asn Ser Val Ile Glu
385 390 395 400

Lys Met Asn Thr Gln Phe Thr Ala Val Gly Lys Glu Phe Asn Lys Leu
405 410 415

Glu Arg Arg Met Glu Asn Leu Asn Lys Lys Val Asp Asp Gly Phe Leu
420 425 430

Asp Ile Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Leu Glu Asn Glu
435 440 445

Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr Glu Lys
450 455 460

Val Lys Ser Gln Leu Lys Asn Asn Ala Lys Glu Ile Gly Asn Gly Cys
465 470 475 480

Phe Glu Phe Tyr His Lys Cys Asn Asn Glu Cys Met Glu Ser Val Lys
485 490 495

Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr Ser Glu Glu Ser Lys Leu Asn
500 505 510

Arg Glu Lys Ile Asp Gly Val Lys Leu Glu Ser Met Gly Val Tyr Gln
515 520 525

Ile Leu Ala Ile Tyr Ser Thr Val Ala Ser Ser Leu Val Leu Leu Val
530 535 540

Ser Leu Gly Ala Ile Ser Phe Trp Met Cys Ser Asn Gly Ser Leu Gln
545 550 555 560

Cys Arg Ile Cys Ile
565

<210> 9
<211> 572
<212> PRT
<213> 人工

<220>
<223> 含有 PDI 信号肽以及跨膜结构域和胞质尾区的全长 rHA

[0005]

<400> 9

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300
```

[0006]

```
Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
                485                 490                 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
            500                 505                 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
        515                 520                 525

Glu Ser Met Gly Val Tyr Gln Ile Leu Ala Ile Tyr Ser Thr Val Ala
    530                 535                 540

Ser Ser Leu Val Leu Leu Val Ser Leu Gly Ala Ile Ser Phe Trp Met
545                 550                 555                 560

Cys Ser Asn Gly Ser Leu Gln Cys Arg Ile Cys Ile
                565                 570
```

<210> 10
<211> 540
<212> PRT
<213> 人工

<220>
<223> 利用 SEKDEL 驻留信号的 ER 驻留 rHA

<400> 10

[0007]

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                  10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320
```

[0008]

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
325 330 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
340 345 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
355 360 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
370 375 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385 390 395 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
405 410 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
420 425 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
435 440 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
450 455 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465 470 475 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
485 490 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
515 520 525

Glu Ser Met Gly Val Tyr Ser Glu Lys Asp Glu Leu
530 535 540

<210> 11
<211> 538
<212> PRT
<213> 人工

<220>
<223> 利用 HDEL 驻留信号的 ER 驻留 rHA

<400> 11

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

[0009]

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
50 55 60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65 70 75 80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
85 90 95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
100 105 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
115 120 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
130 135 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145 150 155 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
165 170 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
180 185 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
195 200 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
210 215 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225 230 235 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
245 250 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
260 265 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
275 280 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
290 295 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305 310 315 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
325 330 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
340 345 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
355 360 365

[0010]

```
Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
                485                 490                 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
            500                 505                 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
        515                 520                 525

Glu Ser Met Gly Val Tyr His Asp Glu Leu
    530                 535


<210>  12
<211>  534
<212>  PRT
<213>  人工

<220>
<223>  无跨膜结构域的可溶性 rHA

<400>  12

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                   10                  15
Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95
```

[0011]

```
Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415
```

[0012]

```
Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
                485                 490                 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
            500                 505                 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
        515                 520                 525

Glu Ser Met Gly Val Tyr
    530
```

<210> 13
<211> 568
<212> PRT
<213> 人工

<220>
<223> 利用 GCN4-pII 三聚体肽的可溶性三聚体 rHA

<400> 13

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                   10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140
```

[0013]

```
Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460
```

[0014]

```
Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
                485                 490                 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
            500                 505                 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
        515                 520                 525

Glu Ser Met Gly Val Tyr Leu Lys Gln Ile Glu Asp Lys Ile Glu Glu
    530                 535                 540

Ile Leu Ser Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys
545                 550                 555                 560

Lys Leu Ile Gly Glu Ser Ala Ala
                565
```

<210> 14
<211> 574
<212> PRT
<213> 人工

<220>
<223> 利用 GCN4-pII 三聚体肽并驻留在 ER 中的可溶性三聚体 rHA

<400> 14

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                   10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140
```

[0015]

```
Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460
```

[0016]

```
Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
                485                 490                 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
            500                 505                 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
        515                 520                 525

Glu Ser Met Gly Val Tyr Leu Lys Gln Ile Glu Asp Lys Ile Glu Glu
    530                 535                 540

Ile Leu Ser Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys
545                 550                 555                 560

Lys Leu Ile Gly Glu Ser Ala Ala Ser Glu Asp Lys Glu Leu
                565                 570
```

<210> 15
<211> 593
<212> PRT
<213> 人工

<220>
<223> 利用 PRD 三聚体肽的可溶性三聚体 rHA

<400> 15

```
Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1               5                   10                  15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
            20                  25                  30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
        35                  40                  45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
    50                  55                  60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65                  70                  75                  80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
                85                  90                  95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
            100                 105                 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
        115                 120                 125

Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160
```

[0017]

```
His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
    450                 455                 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465                 470                 475                 480
```

[0018]

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
485 490 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
515 520 525

Glu Ser Met Gly Val Tyr Gly Gly Ala Gly Pro Pro Pro Pro Val His
530 535 540

Leu Pro Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro
545 550 555 560

Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro
565 570 575

Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro Ser Ala
580 585 590

Ala

<210> 16
<211> 599
<212> PRT
<213> 人工

<220>
<223> 利用 RDP 三聚体肽并驻留在 ER 中的可溶性三聚体 rHA

<400> 16

Met Ala Lys Asn Val Ala Ile Phe Gly Leu Leu Phe Ser Leu Leu Val
1 5 10 15

Leu Val Pro Ser Gln Ile Phe Ala Asp Thr Ile Cys Ile Gly Tyr His
20 25 30

Ala Asn Asn Ser Thr Asp Thr Val Asp Thr Val Leu Glu Lys Asn Val
35 40 45

Thr Val Thr His Ser Val Asn Leu Leu Glu Asp Ser His Asn Gly Lys
50 55 60

Leu Cys Leu Leu Lys Gly Ile Ala Pro Leu Gln Leu Gly Asn Cys Ser
65 70 75 80

Val Ala Gly Trp Ile Leu Gly Asn Pro Glu Cys Glu Leu Leu Ile Ser
85 90 95

Lys Glu Ser Trp Ser Tyr Ile Val Glu Thr Pro Asn Pro Glu Asn Gly
100 105 110

Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr Glu Glu Leu Arg Glu Gln
115 120 125

[0019]

```
Leu Ser Ser Val Ser Ser Phe Glu Arg Phe Glu Ile Phe Pro Lys Glu
    130                 135                 140

Ser Ser Trp Pro Asn His Thr Val Thr Gly Val Ser Ala Ser Cys Ser
145                 150                 155                 160

His Asn Gly Lys Ser Ser Phe Tyr Arg Asn Leu Leu Trp Leu Thr Gly
                165                 170                 175

Lys Asn Gly Leu Tyr Pro Asn Leu Ser Lys Ser Tyr Val Asn Asn Lys
            180                 185                 190

Glu Lys Glu Val Leu Val Leu Trp Gly Val His His Pro Pro Asn Ile
        195                 200                 205

Gly Asn Gln Arg Ala Leu Tyr His Thr Glu Asn Ala Tyr Val Ser Val
    210                 215                 220

Val Ser Ser His Tyr Ser Arg Arg Phe Thr Pro Glu Ile Ala Lys Arg
225                 230                 235                 240

Pro Lys Val Arg Asp Gln Glu Gly Arg Ile Asn Tyr Tyr Trp Thr Leu
                245                 250                 255

Leu Glu Pro Gly Asp Thr Ile Ile Phe Glu Ala Asn Gly Asn Leu Ile
            260                 265                 270

Ala Pro Trp Tyr Ala Phe Ala Leu Ser Arg Gly Phe Gly Ser Gly Ile
        275                 280                 285

Ile Thr Ser Asn Ala Pro Met Asp Glu Cys Asp Ala Lys Cys Gln Thr
    290                 295                 300

Pro Gln Gly Ala Ile Asn Ser Ser Leu Pro Phe Gln Asn Val His Pro
305                 310                 315                 320

Val Thr Ile Gly Glu Cys Pro Lys Tyr Val Arg Ser Ala Lys Leu Arg
                325                 330                 335

Met Val Thr Gly Leu Arg Asn Ile Pro Ser Ile Gln Ser Arg Gly Leu
            340                 345                 350

Phe Gly Ala Ile Ala Gly Phe Ile Glu Gly Gly Trp Thr Gly Met Val
        355                 360                 365

Asp Gly Trp Tyr Gly Tyr His His Gln Asn Glu Gln Gly Ser Gly Tyr
    370                 375                 380

Ala Ala Asp Gln Lys Ser Thr Gln Asn Ala Ile Asn Gly Ile Thr Asn
385                 390                 395                 400

Lys Val Asn Ser Val Ile Glu Lys Met Asn Thr Gln Phe Thr Ala Val
                405                 410                 415

Gly Lys Glu Phe Asn Lys Leu Glu Arg Arg Met Glu Asn Leu Asn Lys
            420                 425                 430

Lys Val Asp Asp Gly Phe Leu Asp Ile Trp Thr Tyr Asn Ala Glu Leu
        435                 440                 445
```

[0020]

Leu Val Leu Leu Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn
450 455 460

Val Lys Asn Leu Tyr Glu Lys Val Lys Ser Gln Leu Lys Asn Asn Ala
465 470 475 480

Lys Glu Ile Gly Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asn Asn
485 490 495

Glu Cys Met Glu Ser Val Lys Asn Gly Thr Tyr Asp Tyr Pro Lys Tyr
500 505 510

Ser Glu Glu Ser Lys Leu Asn Arg Glu Lys Ile Asp Gly Val Lys Leu
515 520 525

Glu Ser Met Gly Val Tyr Gly Gly Ala Gly Pro Pro Pro Pro Val His
530 535 540

Leu Pro Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro
545 550 555 560

Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro
565 570 575

Pro Pro Val His Leu Pro Pro Pro Val His Leu Pro Pro Pro Ser Ala
580 585 590

Ala Ser Glu Lys Asp Glu Leu
595

<210> 17
<211> 1623
<212> DNA
<213> 流感病毒

<400> 17
gggcccatgg cgaaaaacgt tgcgattttc ggtttattgt tttctcttct tctgttggtt 60

ccttctcaga tcttcgctga cacaatatgt ataggctacc atgccaacaa ctcaaccgac 120

actgttgaca cagtacttga gaagaatgtg acagtgacac actctgtcaa cctacttgag 180

gacagtcaca atggaaaact atgtctacta aaaggaatag ccccactaca attgggtaat 240

tgcagcgttg ccggatggat cttaggaaac ccagaatgcg aattactgat ttccaaggaa 300

tcatggtcct acattgtaga aacaccaaat cctgagaatg gaacatgtta cccagggtat 360

ttcgccgact atgaggaact gagggagcaa ttgagttcag tatcttcatt tgagagattc 420

gaaatattcc ccaaagaaag ctcatggccc aaccacaccg taaccggagt atcagcatca 480

tgctcccata atgggaaaag cagtttttac agaaatttgc tatggctgac ggggaagaat 540

ggtttgtacc caaacctgag caagtcctat gtaaacaaca aagagaaaga agtccttgta 600

ctatggggtg ttcatcaccc gcctaacata gggaaccaaa gggcactcta tcatacagaa 660

aatgcttatg tctctgtagt gtcttcacat tatagcagaa gattcacccc agaaatagcc 720

[0021]

aaaagaccca aagtaagaga tcaggaagga agaatcaact actactggac tctgctggaa 780

cctggggata caataatatt tgaggcaaat ggaaatctaa tagcgccatg gtatgctttt 840

gcactgagta gaggctttgg atcaggaatc atcacctcaa atgcaccaat ggatgaatgt 900

gatgcgaagt gtcaaacacc tcagggagct ataaacagca gtcttccttt ccagaatgta 960

cacccagtca caataggaga gtgtccaaag tatgtcagga gtgcaaaatt aaggatggtt 1020

acaggactaa ggaacatccc atccattcaa tccagaggtt tgtttggagc cattgccggt 1080

ttcattgaag gggggtggac tggaatggta gatgggtggt atggttatca tcatcagaat 1140

gagcaaggat ctggctatgc tgcagatcaa aaaagtacac aaaatgccat taacgggatt 1200

acaaacaagg tcaattctgt aattgagaaa atgaacactc aattcacagc tgtgggcaaa 1260

gagttcaaca aattggaaag aaggatggaa aacttaaata aaaaagttga tgatgggttt 1320

ctagacattt ggacatataa tgcagaattg ttggttctac tggaaaatga aaggactttg 1380

gatttccatg actccaatgt gaagaatctg tatgagaaag taaaaagcca attaaagaat 1440

aatgccaaag aaataggaaa cgggtgtttt gagttctatc acaagtgtaa caatgaatgc 1500

atggagagtg tgaaaaatgg tacctatgac tatccaaaat attccgaaga atcaaagtta 1560

aacagggaga aaattgatgg agtgaaattg gaatcaatgg gagtatacta agagctcagg 1620

cct 1623

<210> 18
<211> 219
<212> DNA
<213> 流感病毒

<400> 18
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac cagattctgg cgatctactc aactgtcgcc 120

agttccctgg ttcttttggt ctccctgggg gcaatcagct tctggatgtg ttccaatggg 180

tctttgcagt gtagaatatg catctaagag ctcaggcct 219

<210> 19
<211> 123
<212> DNA
<213> 人工

<220>
<223> ER 驻留 SEKDEL 基因片段

<400> 19
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac tctgagaagg atgaacttta agagctcagg 120

[0022]

cct 123

<210> 20
<211> 117
<212> DNA
<213> 人工

<220>
<223> ER 驻留 HDEL 基因片段

<400> 20
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac cacgatgaac tttaagagct caggcct 117

<210> 21
<211> 207
<212> DNA
<213> 人工

<220>
<223> GCN4-pII 基因片段

<400> 21
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac ctcaaacaaa tcgaagacaa gatcgaagag 120

atcctctcga aaatctacca catcgaaaac gagattgcca ggatcaagaa gctcataggc 180

gagtcagcag cctaagagct caggcct 207

<210> 22
<211> 225
<212> DNA
<213> 人工

<220>
<223> ER 驻留 GCN4-pII 基因片断

<400> 22
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac ctcaaacaaa tcgaagacaa gatcgaagag 120

atcctctcga aaatctacca catcgaaaac gagattgcca ggatcaagaa gctcataggc 180

gagtcagcag cctctgagaa ggatgaactt taagagctca ggcct 225

<210> 23
<211> 279
<212> DNA
<213> 玉米 (Zea maize)

<400> 23
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

[0023]

ggagtgaaat tggaatcaat gggagtatac ggaggcgctg gcccaccgcc agttcatcta 120

ccgcctccgg tgcatctgcc acctccggtt cacctgccac ctccggtgca tctcccaccg 180

ccagtccacc tgccgcctcc ggtccacctg ccaccgcctg tccatgtgcc gccaccggtt 240

catctgccgc ctccatcagc agcctaagag ctcaggcct 279

<210> 24
<211> 297
<212> DNA
<213> 人工

<220>
<223> ER 驻留基因片段

<400> 24
ggtacctatg actatccaaa atattccgaa gaatcaaagt taaacaggga gaaaattgat 60

ggagtgaaat tggaatcaat gggagtatac ggaggcgctg gcccaccgcc agttcatcta 120

ccgcctccgg tgcatctgcc acctccggtt cacctgccac ctccggtgca tctcccaccg 180

ccagtccacc tgccgcctcc ggtccacctg ccaccgcctg tccatgtgcc gccaccggtt 240

catctgccgc ctccatcagc agcctctgag aaggatgaac tttaagagct caggcct 297

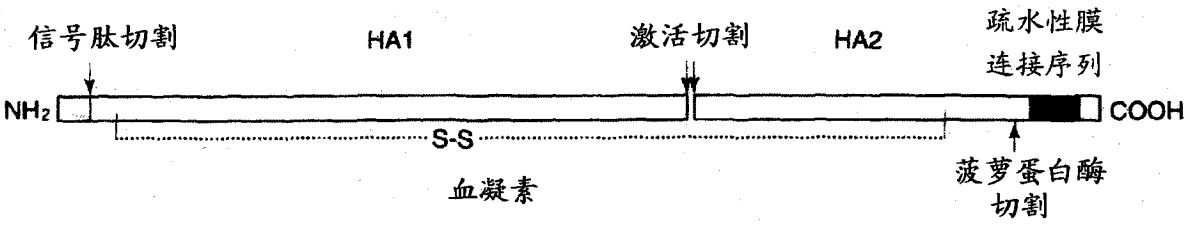


图 1

LKQIEDKIEEILSKIYHIENEIARIKKLIGESAA (SEQ ID NO :1)

图 2

```
1       M   A   K   N   V   A   I   F   G   L   L   F   S   L
1   5'-ATG GCG AAA AAC GTT GCG ATT TTC GGT TTA TTG TTT TCT CTT

15      L   V   L   V   P   S   Q   I   F   A       SEQ ID NO:6
43  CTT CTG TTG GTT CCT TCT CAG ATC TTC GCT -3'   SEQ ID NO:7
```

图 3

```
1   MKAKLLVLLC TFTATYADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL EDSHNGKLCL
61  LKGIAPLQLG NCSVAGWILG NPECELLISK ESWSYIVETP NPENGTCYPG YFADYEELRE
121 QLSSVSSFER FEIFPKESSW PNHTVTGVSA SCSHNGKSSF YRNLLWLTGK NGLYPNLSKS
181 YVNNKEKEVL VLWGVHHPPN IGNQRALYHT ENAYVSVVSS HYSRRFTPEI AKRPKVRDQE
241 GRINYYWTLL EPGDTIIFEA NGNLIAPWYA FALSRGFGSG IITSNAPMDE CDAKCQTPQG
301 AINSSLPFQN VHPVTIGECP KYVRSAKLRM VTGLRNIPSI QSRGLFGAIA GFIEGGWTGM
361 VDGWYGYHHQ NEQGSGYAAD QKSTQNAING ITNKVNSVIE KMNTQFTAVG KEFNKLERRM
421 ENLNKKVDDG FLDIWTYNAE LLVLLENERT LDFHDSNVKN LYEKVKSQLK NNAKEIGNGC
481 FEFYHKCNNE CMESVKNGTY DYPKYSEESK LNREKIDGVK LESMGVYQIL AIYSTVASSL
541 VLLVSLGAIS FWMCSNGSLQ CRICI          SEQ ID NO:8
```

图 4

```
-7    MAKNVAIFGL LFSLL VLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54    GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114   EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174   NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234   VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294   QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354   GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414   LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474   IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYQIL AIYS
534   TVASSLVLLV SLGAISFWMC SNGSLQCRIC I*          SEQ ID NO:9
```

图 5A

```
-7    MAKNVAIFGL LFSLL VLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54    GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114   EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174   NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234   VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294   QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354   GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414   LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474   IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYSEKDE*
      SEQ ID NO:10
```

图 5B

```
-7      MAKNVAIPGL LFSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54      GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114     EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174     NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234     VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294     QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354     GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414     LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474     IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYHDQ*
        SEQ ID NO:11
```

图 5C

```
-7      MAKNVAIPGL LFSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54      GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114     EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174     NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234     VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294     QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354     GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414     LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474     IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VY*
        SEQ ID NO:12
```

图 5D

```
-7      MAKNVAIPGL LPSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54      GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114     EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174     NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234     VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294     QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354     GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414     LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474     IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYLKQIEDKI
527     EEILSKIYHI ENEIARIKKL IGESAMSEDK EL*          SEQ ID NO:13
```

图 5E

```
-7      MAKNVAIPGL LPSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54      GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114     EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174     NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234     VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294     QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354     GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414     LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474     IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYLKQIEDKI
527     EEILSKIYHI ENEIARIKKL IGESAMSEDK EL*          SEQ ID NO:14
```

图 5F

```
-7    MAKNVAIFGL LFSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54    GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114   EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174   NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234   VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294   QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354   GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414   LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474   IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYGGAGPPPA
527   VHLPPPVHLP PPVHLPPPVH LPPPVHLPPP VHLPPPVHLP PPVHLPPPSA A*
      SEQ ID NO:15
```

图 5G

```
-7    MAKNVAIFGL LFSLLVLVPS QIFADTICIG YHANNSTDTVDTVLEKNVTVTHSVNLLEDSHN
54    GKLCLLKGIA PLQLGNCSVA GWILGNPECE LLISKESWSY IVETPNPENG TCYPGYFADY
114   EELREQLSSV SSFERFEIFP KESSWPNHTV TGVSASCSHN GKSSFYRNLL WLTGKNGLYP
174   NLSKSYVNNK EKEVLVLWGV HHPPNIGNQR ALYHTENAYV SVVSSHYSRR FTPEIAKRPK
234   VRDQEGRINY YWTLLEPGDT IIFEANGNLI APWYAFALSR GFGSGIITSN APMDECDAKC
294   QTPQGAINSS LPFQNVHPVT IGECPKYVRS AKLRMVTGLR NIPSIQSRGL FGAIAGFIEG
354   GWTGMVDGWY GYHHQNEQGS GYAADQKSTQ NAINGITNKV NSVIEKMNTQ FTAVGKEFNK
414   LERRMENLNK KVDDGFLDIW TYNAELLVLL ENERTLDFHD SNVKNLYEKV KSQLKNNAKE
474   IGNGCFEFYH KCNNECMESV KNGTYDYPKY SEESKLNREK IDGVKLESMG VYGGAGPPPA
527   VHLPPPVHLP PPVHLPPPVH LPPPVHLPPP VHLPPPVHLP PPVHLPPPSA ASEKDEL*
      SEQ ID NO:16
```

图 5H

```
1       gggcccATGG CGAAAAACGT TGCGATTTTC GGTTTATTGT TTTCTCTTCT TCTGTTGGTT
61      CCTTCTCAGA TCTTCGCTGA CACAATATGT ATAGGCTACC ATGCCAACAA CTCAACCGAC
121     ACTGTTGACA CAGTACTTGA GAAGAATGTG ACAGTGACAC ACTCTGTCAA CCTACTTGAG
181     GACAGTCACA ATGGAAAACT ATGTCTACTA AAAGGAATAG CCCCACTACA ATTGGGTAAT
241     TGCAGCGTTG CCGGATGGAT CTTAGGAAAC CCAGAATGCG AATTACTGAT TTCCAAGGAA
301     TCATGGTCCT ACATTGTAGA AACACCAAAT CCTGAGAATG GAACATGTTA CCCAGGGTAT
361     TTCGCCGACT ATGAGGAACT GAGGGAGCAA TTGAGTTCAG TATCTTCATT TGAGAGATTC
421     GAAATATTCC CCAAAGAAAG CTCATGGCCC AACCACACCG TAACCGGAGT ATCAGCATCA
481     TGCTCCCATA ATGGGAAAAG CAGTTTTTAC AGAAATTTGC TATGGCTGAC GGGGAAGAAT
541     GGTTTGTACC CAAACCTGAG CAAGTCCTAT GTAAACAACA AAGAGAAAGA AGTCCTTGTA
601     CTATGGGGTG TTCATCACCC GCCTAACATA GGGAACCAAA GGGCACTCTA TCATACAGAA
661     AATGCTTATG TCTCTGTAGT GTCTTCACAT TATAGCAGAA GATTCACCCC AGAAATAGCC
721     AAAAGACCCA AAGTAAGAGA TCAGGAAGGA AGAATCAACT ACTACTGGAC TCTGCTGGAA
781     CCTGGGGATA CAATAATATT TGAGGCAAAT GGAAATCTAA TAGCGCCATG GTATGCTTTT
841     GCACTGAGTA GAGGCTTTGG ATCAGGAATC ATCACCTCAA ATGCACCAAT GGATGAATGT
901     GATGCGAAGT GTCAAACACC TCAGGGAGCT ATAAACAGCA GTCTTCCTTT CCAGAATGTA
961     CACCCAGTCA CAATAGGAGA GTGTCCAAAG TATGTCAGGA GTGCAAAATT AAGGATGGTT
1021    ACAGGACTAA GGAACATCCC ATCCATTCAA TCCAGAGGTT TGTTTGGAGC CATTGCCGGT
1081    TTCATTGAAG GGGGGTGGAC TGGAATGGTA GATGGGTGGT ATGGTTATCA TCATCAGAAT
1141    GAGCAAGGAT CTGGCTATGC TGCAGATCAA AAAAGTACAC AAAATGCCAT TAACGGGATT
1201    ACAAACAAGG TCAATTCTGT AATTGAGAAA ATGAACACTC AATTCACAGC TGTGGGCAAA
1261    GAGTTCAACA AATTGGAAAG AAGGATGGAA AACTTAAATA AAAAAGTTGA TGATGGGTTT
1321    CTAGACATTT GGACATATAA TGCAGAATTG TTGGTTCTAC TGGAAAATGA AAGGACTTTG
1381    GATTTCCATG ACTCCAATGT GAAGAATCTG TATGAGAAAG TAAAAAGCCA ATTAAAGAAT
1441    AATGCCAAAG AAATAGGAAA CGGGTGTTTT GAGTTCTATC ACAAGTGTAA CAATGAATGC
1501    ATGGAGAGTG TGAAAAATGG TACCTATGAC TATCCAAAAT ATTCCGAAGA ATCAAAGTTA
1561    AACAGGGAGA AAATTGATGG AGTGAAATTG GAATCAATGG GAGTATACta agagctcagg
1621    cct                SEQ ID NO:17
```

图 6A

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC CAGATTCTGG CGATCTACTC AACTGTCGCC
121     AGTTCCCTGG TTCTTTTGGT CTCCCTGGGG GCAATCAGCT TCTGGATGTG TTCCAATGGG
181     TCTTTGCAGT GTAGAATATG CATCtaagag ctcaggcct          SEQ ID NO:18
```

图 6B

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC TCTGAGAAGG ATGAACTTta agagctcagg
121     cct        SEQ ID NO:19
```

图 6C

```
1       GGTACCTAT GACTATCCA AAATATTCC GAAGAATCA AAGTTAAAC AGGGAGAAA
61      ATTGATGGA GTGAAATTG GAATCAATG GGAGTATAC CACGATGAA CTTtaagagc
121     tcaggcct          SEQ ID NO:20
```

图 6D

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC CTCAAACAAA TCGAAGACAA GATCGAAGAG
121     ATCCTCTCGA AAATCTACCA CATCGAAAAC GAGATTGCCA GGATCAAGAA GCTCATAGGC
181     GAGTCAGCAG CCtaagagct caggcct          SEQ ID NO:21
```

图 6E

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC CTCAAACAAA TCGAAGACAA GATCGAAGAG
121     ATCCTCTCGA AAATCTACCA CATCGAAAAC GAGATTGCCA GGATCAAGAA GCTCATAGGC
181     GAGTCAGCAG CCTCTGAGAA GGATGAACTT taagagctca ggcct    SEQ ID NO:22
```

图 6F

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC GGAGGCGCTG GCCCACCGCC AGTTCATCTA
121     CCGCCTCCGG TGCATCTGCC ACCTCCGGTT CACCTGCCAC CTCCGGTGCA TCTCCCACCG
181     CCAGTCCACC TGCCGCCTCC GGTCCACCTG CCACCGCCTG TCCATGTGCC GCCACCGGTT
241     CATCTGCCGC CTCCATCAGC AGCCtaagag ctcaggcct          SEQ ID NO:23
```

图 6G

```
1       GGTACCTATG ACTATCCAAA ATATTCCGAA GAATCAAAGT TAAACAGGGA GAAAATTGAT
61      GGAGTGAAAT TGGAATCAAT GGGAGTATAC GGAGGCGCTG GCCCACCGCC AGTTCATCTA
121     CCGCCTCCGG TGCATCTGCC ACCTCCGGTT CACCTGCCAC CTCCGGTGCA TCTCCCACCG
181     CCAGTCCACC TGCCGCCTCC GGTCCACCTG CCACCGCCTG TCCATGTGCC GCCACCGGTT
241     CATCTGCCGC CTCCATCAGC AGCCTCTGAG AAGGATGAAC TTtaagagct caggcct
        SEQ ID NO:24
```

图 6H

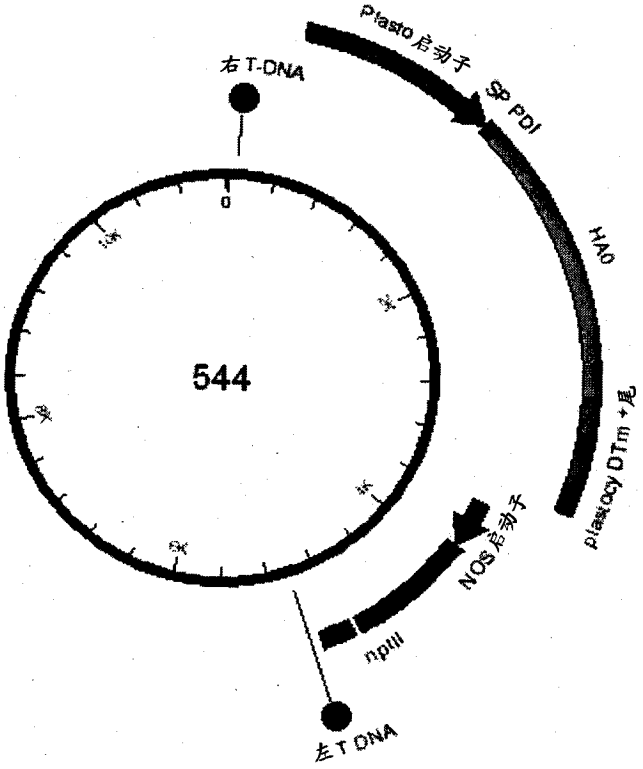

右 T-DNA
Plastd 启动子
SP PDI
HA0
plastocy DTm +尾
544
NOS 启动子
nptII
左 T DNA

图 7

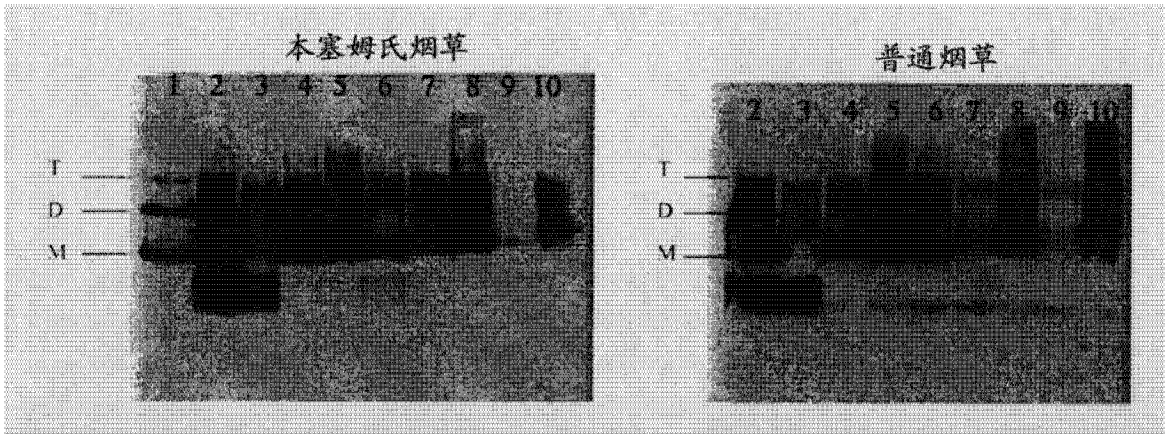

本塞姆氏烟草
1 2 3 4 5 6 7 8 9 10
T
D
M
普通烟草
2 3 4 5 6 7 8 9 10
T
D
M

图 8

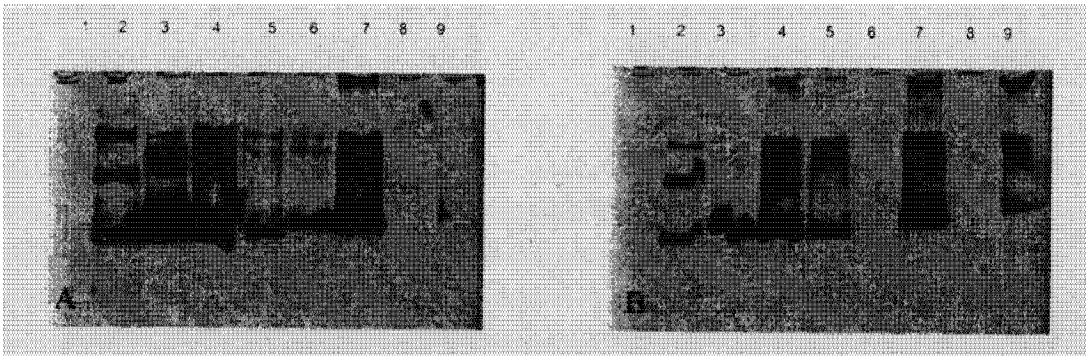


图 9

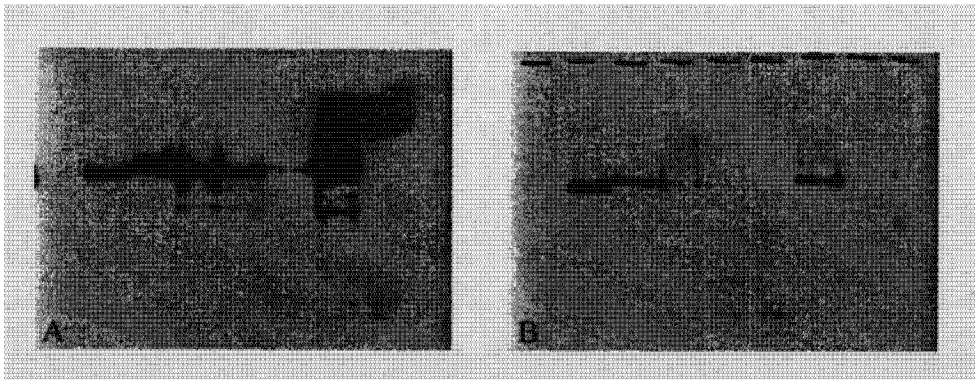


图 10

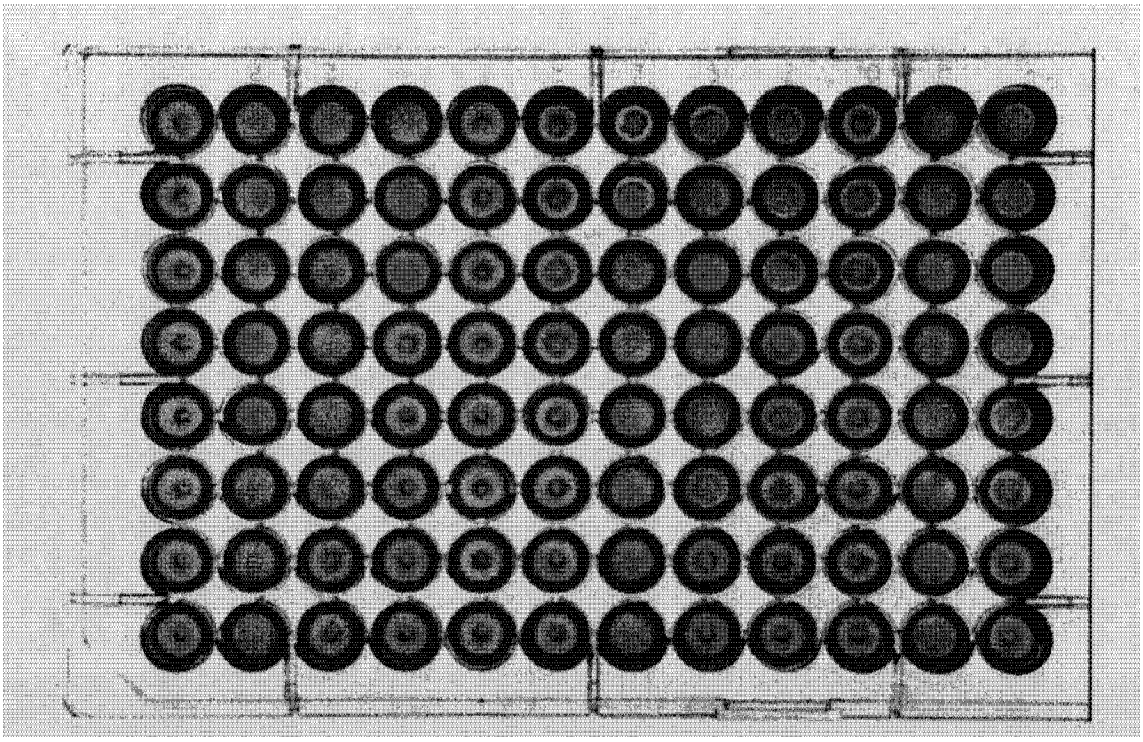


图 11

1. 一种重组血凝素 (rHA)，其包含 ：

a) 血凝素结构域，选自血凝素亚型 H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15 或 H16 ；和

b) 寡聚化结构域，选自 GCN4-pII 肽、γ 玉米醇溶蛋白的脯氨酸富含结构域 (PRD)，以及噬菌体 T4 纤维蛋白 ；

其中所述 rHA 作为嵌合式可溶性同三聚体产生。

2. 权利要求 1 的 rHA，其中所述 rHA 还包含信号肽。

3. 权利要求 1 或 2 的 rHA，其中所述 rHA 还包含内质网 (ER) 驻留信号。

4. 编码权利要求 1 至 3 中任一项的 rHA 的核苷酸。

5. 核酸，其包含 ：

a) 编码血凝素结构域的核苷酸序列，所述血凝素结构域选自血凝素亚型 H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15 或 H16 ；和

b) 编码寡聚化结构域的核苷酸序列，所述寡聚化结构域选自 GCN4-pII 肽、γ 玉米醇溶蛋白的脯氨酸富含结构域 (PRD)，以及噬菌体 T4 纤维蛋白 ；

其中所述核酸编码形成同三聚体的可溶性重组血凝素 (rHA)。

6. 权利要求 5 的核酸，还包含编码信号肽的核苷酸序列。

7. 权利要求 5 或 6 的核酸，还包含编码内质网 (ER) 驻留信号的核苷酸序列。

8. 一种载体，其包含权利要求 4 至 7 中任一项的核酸。

9. 表达权利要求 1 至 3 中任一项的 rHA 的宿主细胞。

10. 用权利要求 4 至 7 中任一项的核酸转化的宿主细胞。

11. 用权利要求 8 的载体转化的宿主细胞。

12. 一种产生 rHA 的方法，包括提供含有权利要求 5 所述核酸的宿主细胞，以及在所述宿主细胞中表达所述 rHA。

13. 一种在植物中表达 rHA 的方法，包括向植物中引入权利要求 8 的载体并表达所述 rHA。

14. 权利要求 13 的方法，其中在引入步骤 (步骤 a) 中，所述核酸以瞬时方式被引入所述植物中。

15. 权利要求 13 的方法，其中在引入步骤 (步骤 a) 中，所述核酸以使其稳定的方式被引入所述植物中。

16. 一种在植物中产生重组血凝素 (rHA) 的方法，包括 ：

a) 将核酸序列引入所述植物或其部分中，所述核酸序列包含与权利要求 5 所述核酸有效连接的调控区域 ；以及

b) 培养所述转基因植物，从而产生所述 rHA。

17. 权利要求 16 的方法，其中在引入步骤 (步骤 a) 中，所述核酸以瞬时方式被引入所述植物中。

18. 权利要求 16 的方法，其中在引入步骤 (步骤 a) 中，所述核酸以稳定的方式被引入所述植物中。