



(12) 发明专利

(10) 授权公告号 CN 101341252 B

(45) 授权公告日 2011.08.17

(21) 申请号 200680040367. X
 (22) 申请日 2006.10.26
 (30) 优先权数据
 05023611.6 2005.10.28 EP
 (85) PCT申请进入国家阶段日
 2008.04.28
 (86) PCT申请的申请数据
 PCT/EP2006/010308 2006.10.26
 (87) PCT申请的公布数据
 W02007/048601 EN 2007.05.03
 (73) 专利权人 弗·哈夫曼-拉罗切有限公司
 地址 瑞士巴塞尔
 (72) 发明人 U·格普费特 E·科佩特茨基
 (74) 专利代理机构 北京市中咨律师事务所
 11247
 代理人 黄革生 林柏楠
 (51) Int. Cl.
 C12N 15/869 (2006.01)

(56) 对比文件
 WO 2005/024030 A1, 2005.03.17,
 H Mizuguchi等. Long-term replication of Epstein-Barr virus-derived episomal vectors in the rodent cells. 《FEBS Letters》. 2000, 第472卷 173-178.
 K-I Tomiyasu等. Gene transfer in vitro and in vivo with Epstein-Barr virus-based episomal vector results in markedly high transient expression in rodent cells. 《Biochemical and biophysical research communications》. 1998, 第253卷(第3期), 733-738.

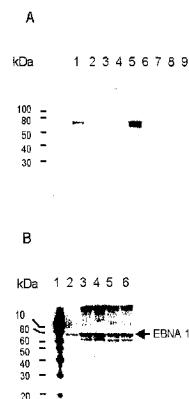
审查员 张丽华

权利要求书 3 页 说明书 18 页
序列表 7 页 附图 6 页

(54) 发明名称
啮齿动物细胞中的蛋白质表达

(57) 摘要

本发明描述了用于在啮齿动物细胞中表达异源多肽的方法。所述方法包括 EB 病毒 (EBV) 的 oriP/EBNA-1 游离型复制和保持系统。随着 EBNA-1 蛋白质表达盒在启动子控制下稳定整合进啮齿动物细胞的基因组, 在细胞中获得 EBNA-1 蛋白质的表达。异源蛋白质从包含 EBV 复制起点和所述异源蛋白质的功能性表达盒的附加体表达。本发明还包含经转化的啮齿动物细胞系、在所述细胞系中生产异源蛋白质的方法以及用于构建所述细胞系的试剂盒。



1. CHO 细胞,其特征在于所述 CHO 细胞:
 - a) 表达 EB 病毒核抗原 1,其中编码 EB 病毒核抗原 1 蛋白质的结构基因与强启动子有效连接,且细胞不含 EB 病毒复制起点 oriP 序列的功能性拷贝;
 - b) 含有附加体,其中,所述附加体包含:
 - i) 原核复制起点;
 - ii) 选择标记;
 - iii) EB 病毒复制起点 oriP,其作为单个源自 EB 病毒的元件,且所述附加体不含 EB 病毒核抗原 1 结构基因的功能性拷贝;
 - iv) 适于在所述 CHO 细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码所述异源多肽的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译区域,
 - c) 不包含多瘤病毒的大 T 抗原。
2. 根据权利要求 1 的 CHO 细胞,其特征在于编码 EB 病毒核抗原 1 蛋白质的结构基因与异源启动子有效相连。
3. 获得 CHO 细胞的方法,其特征在于,所述方法包括如下步骤:
 - a) 提供 CHO 细胞,其不含有多瘤病毒的大 T 抗原;
 - b) 提供下述质粒,所述质粒包含
 - (i) 原核复制起点;
 - (ii) 选择标记;
 - (iii) 针对 EB 病毒核抗原 1 的功能性表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码 EB 病毒核抗原 1 的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译序列;其中所述质粒不含 EB 病毒复制起点 oriP 序列的功能性拷贝;
 - c) 将所述质粒 b) 引入所述 CHO 细胞 a),以及选择经稳定转化的 CHO 细胞;
 - d) 提供一个至四个下述其它质粒,所述质粒包含
 - (i) 原核复制起点;
 - (ii) 选择标记;
 - (iii) EB 病毒复制起点 oriP,其作为单个源自 EB 病毒的元件;
 - (iv) 适于在所述经转化的 CHO 细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码异源多肽的核酸序列和具有聚腺苷酸化信号的 3' 非翻译区域;
 - (v) 无编码 EB 病毒核抗原 1 蛋白质的结构基因;
 - e) 将所述其它质粒 d) 引入所述经稳定转化的 CHO 细胞。
4. 生产异源多肽的方法,其特征在于,所述方法包括:
 - a) 提供根据权利要求 1 的 CHO 细胞;
 - b) 在适于表达所述异源多肽的条件下培养所述 CHO 细胞;
 - c) 从培养物中回收所述多肽。
5. 用于生产根据权利要求 1 的 CHO 细胞的试剂盒,其特征在于,所述试剂盒包含:
 - a) CHO 细胞,其不含有多瘤病毒的大 T 抗原;
 - b) 第一质粒,其包含:

- (i) 原核复制起点；
- (ii) 选择标记；
- (iii) 针对EB病毒核抗原1的功能性表达盒,其中,所述表达盒包含启动子序列、5'非翻译区域、编码EB病毒核抗原1的核酸序列以及包含聚腺苷酸化信号的3'非翻译区域；其中所述质粒不含EB病毒复制起点oriP序列的功能性拷贝；
- c) 第二质粒,其包含：
 - (i) 原核复制起点；
 - (ii) 选择标记；
 - (iii) EB病毒复制起点oriP,其作为单个源自EB病毒的元件；
 - (iv) 适于在所述啮齿动物细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5'非翻译区域、用于引入核酸序列的克隆位点和包含聚腺苷酸化信号的3'非翻译区域；
 - (v) 无编码EB病毒核抗原1蛋白质的结构基因。
- 6. 根据权利要求3至4中任意一项的方法,其特征在于,所述异源多肽为前药。
- 7. 根据权利要求3至4中任意一项的方法,其特征在于,所述异源多肽为生物活性多肽。
- 8. 根据权利要求3至4中任意一项的方法,其特征在于,所述异源多肽为生长因子。
- 9. 根据权利要求3至4中任意一项的方法,其特征在于,所述异源多肽选自酶、酶片段、酶抑制剂或酶激活剂。
- 10. 根据权利要求3至4中任意一项的方法,其特征在于,所述异源多肽选自刺猬蛋白质、骨形态发生蛋白、促红细胞生成素、血小板生成素、G-CSF、白细胞介素、干扰素、免疫球蛋白或免疫球蛋白片段。
- 11. 根据权利要求4的方法,其特征在于所述异源多肽是免疫球蛋白或免疫球蛋白片段。
- 12. 在CHO细胞中表达异源多肽的方法,其特征在于,所述方法包括：
 - a) 提供经编码EB病毒核抗原1蛋白质的结构基因稳定转染的CHO细胞,其不含有多瘤病毒的大T抗原,其中编码EB病毒核抗原1蛋白质的结构基因与强启动子有效连接,且细胞不含EB病毒复制起点oriP序列的功能性拷贝,
 - b) 用表达质粒转染所述CHO细胞,所述表达质粒包含：
 - (i) EB病毒复制起点oriP,其作为单个源自EB病毒的元件,
 - (ii) 无编码EB病毒核抗原1蛋白质的结构基因,
 - (iii) 原核复制起点,
 - (iv) 选择标记,
 - (v) 适合于在所述CHO细胞中表达异源多肽的表达盒,其中所述表达盒包含启动子序列、5'非翻译区域、编码所述异源蛋白质的核酸序列,以及包含聚腺苷酸化信号的3'非翻译区域；
 - c) 在适于表达所述异源多肽的条件下培养所述经转染的细胞,
 - d) 从培养物中回收所述异源多肽。
- 13. 根据权利要求3至4和12中任意一项的方法,其特征在于,所述异源多肽是分泌的

异源多肽。

14. 根据权利要求 4、11 和 12 中任意一项的方法,其特征在于,所述培养在瞬时转染下进行。

15. 根据权利要求 3 至 4 和 12 中任意一项的方法,其特征在于,所述 CHO 细胞选自 CHO-DBX11 细胞、CHO-K1 细胞、CHO-DG44 细胞和表达 EB 病毒核抗原 1 的 CHO 细胞。

16. 生产异源免疫球蛋白的方法,其特征在于所述方法包括下列步骤:

a) 提供 CHO 细胞,其

(i) 表达 EB 病毒核抗原 1;

(ii) 包含附加体,其中所述附加体包含

(i) 原核复制起点;

(ii) 选择标记;

(iii) EB 病毒复制起点 oriP,其作为单个源自 EB 病毒的元件,且所述附加体不含 EB 病毒核抗原 1 结构基因的功能性拷贝;

(iv) 适合于在所述 CHO 细胞中表达异源免疫球蛋白的表达盒,其中所述表达盒包含启动子序列、5' 非翻译区域、编码所述异源免疫球蛋白的核酸序列,以及包含聚腺苷酸化信号的 3' 非翻译区域,

(iii) 不包含多瘤病毒的大 T 抗原,

(iv) 其中编码 EB 病毒核抗原 1 蛋白质的结构基因与强启动子有效连接,且细胞不含 EB 病毒复制起点 oriP 序列的功能性拷贝;

b) 在适合于表达所述异源免疫球蛋白的条件下培养所述 CHO 细胞,

c) 从培养物中回收所述异源免疫球蛋白。

啮齿动物细胞中的蛋白质表达

[0001] 本发明涉及异源蛋白质在啮齿动物细胞中的表达。编码蛋白质的核酸被提供于附加体上,其处于 EB 病毒的 oriP/EBNA-1 系统的控制和保持下,其中,oriP 和 EBNA-1 的结构基因位于细胞中不同的元件上。

技术背景

[0002] 过去数年间生物技术生产工艺的作用和影响越发重要。随着生物技术工艺重要性日益增加,生产的产品的复杂性也稳步增长。

[0003] 可能用来表达异源蛋白质的宿主细胞是 HEK(人胚肾)、HeLa(Henrietta Lacks)、COS(经 SV40 转化的非洲绿猴肾细胞)和 CHO(中国仓鼠卵巢)细胞,因为在这些宿主中表达的异源蛋白质可按照人细胞中的情形被折叠、加工、由蛋白酶使其成熟、糖基化和硫酸化(见,例如 Watson, E., 等人, *Glycobiol.* 4(1994)227-37)。

[0004] EB 病毒 (EBV) 是人 B 淋巴细胞的病原体。其属于人类疱疹病毒类 (疱疹病毒科 (herpesviridae))。经 EBV 转化的 B 淋巴细胞可能不受调控地增殖 (propagate) (Miller, G., 由 Fields, B. 在 *Virology* 中编著, B., Raven Press, N. Y. (1985)563-590)。经 EBV 转化的细胞的特征性特点是所谓的 EB 病毒核抗原 (EBNA) 蛋白质的表达。其中,目前已鉴定出了六种不同的变体。

[0005] EBNA-1 蛋白质在 EBV 核酸在经转化细胞中的复制循环中发挥重要作用。与第二 EBV 元件 (顺式作用的复制起点 (oriP)) 组合,附加体得以在细胞中复制和保持 (Lupton, S., 和 Levine, A. J., *Mol. Cell. Biol.* 5(1985)2533-2542; Yates, J. L., 等人, *Nature (London)* 313(1985)812-815)。

[0006] oriP 片断由两个区域构成:二重对称元件和重复家族。第一种是 65bp 的序列片断。其在病毒的核酸中被大约 1000bp 与重复家族分离开。这一第二种片断由重复了 20 次的 30bp 的序列组成 (Hudson, G. S., 等人, *Virology* 147(1985)81-98; Reisman, D., 等人, *Mol. Cell. Biol.* 5(1985)1822-1832)。30bp 重复家族具有增强转录的能力,而二重对称元件则在复制中发挥作用。完整的 oriP- 片断协助对质粒的染色体外保持。

[0007] EBNA-1 蛋白质 (即,反式作用起始蛋白质)和 oriP 的组合允许质粒的构建,该质粒一旦被引入细胞,它们可作为附加体稳定保持,并且在细胞增殖期间稳定复制 (见,例如, Yates, J. L., 等人, *Nature (London)* 313(1985)812-815; Reisman, D., 等人, *Mol. Cell. Biol.* 5(1985)1822-1832)。这两种元件利用宿主细胞的复制机制来复制附加体 (Bode, J., 等人, *Gene Ther. Mol. Biol.* 6(2001)33-46)。

[0008] Krysan, P. J., 等人 (*Mol. Cell. Biol.* 9(1989)1026-1033) 表明,包含 EBV 复制起点重复家族的质粒可被永久保留于人类细胞的细胞核中。因此,EBNA-1 蛋白质的存在是足够的,不需要 EBV 的其它元件 (还见 Aiyar, A., 等人, *EMBO J.* 17(1998)6394-6403; Hung, S. C., 等人, *PNAS* 98(2001)1865-1870; Yates, J. L., in *DNA Replication in Eukaryotic Cells*, DePhamphilis 编著, M. L., Cold Spring Harbor Laboratory, N. Y. (1996) 第 751-774 页; Yates, J. L., 等人, *J. Vir.* 74(2000)4512-4522)。

[0009] EBNA-1 蛋白质在游离型复制期间将附加体与宿主细胞染色体相连,由此确保了细胞分裂期间附加体的增殖。

[0010] 此种相互关系对于其它一些病毒来说也是已知的,例如,来自猿猴病毒 40(SV 40)的大 T-抗原和来自牛乳头瘤病毒(BPV)的 E1/E2 蛋白质(见,例如 Gilbert, D. M., 等人, Cell 50(1987)59-68 ;DuBridge, R. B., 等人, Mutagen. 3(1988)1-9 ;Lebkowski, J. S., 等人, Mol. Cell. Biol. 4(1984)1951-1960)。

[0011] EBV 的元件 EBNA-1 和 oriP 的组合已用于制备非整合的、染色体外的、自主复制的附加体。例如 Horlick 等人使用稳定表达 EBNA-1 蛋白质的 HEK(人胚肾)细胞来从含有 EBV oriP 的质粒表达 CRHR(促肾上腺皮质激素释放激素受体)(Horlick, R. A., 等人, Prot. Exp. Purif. 9(1997)301-308)。

[0012] 在美国专利 4,686,186 中,已报道了重组载体以及由此转化的真核宿主。EBV 元件 oriP 和 EBNA-1 在重组载体上组合。

[0013] 在 WO 2002/090533 中,报道了一种生产重组蛋白质的方法,该方法通过对悬浮培养的人胚肾细胞(293 细胞系及其遗传变体)进行瞬时转染来进行。提供 EBV 复制起点的质粒以染色体外方式保持。

[0014] 在 WO 2004/053137 中,报道了一种在宿主细胞中生产重组多肽和/或未翻译的 RNA 分子的方法。WO 2004/018506 报道了可用于被瞬时转染进哺乳动物宿主细胞的可控表达系统的组合物和方法。

[0015] 美国专利申请 2002/0086419 报道了用于在真核宿主细胞中稳定维持外源 DNA 的重组载体以及该重组载体用于在宿主细胞中长期稳定生产基因产物的用途。

[0016] 在美国专利 5,707,830 中报道了可用于转染选用的哺乳动物宿主细胞的表达载体。所述表达载体含有 EB 病毒(Epstein-Barr-Virus)重复家族、可在宿主细胞中功能性表达的 EBNA-1 基因的拷贝、真核 DNA 片段(其提供载体在宿主细胞中复制的能力)和表达盒。

[0017] WO 2002/027005 报道了增强的转染系统,其包含附加体保持系统、强启动子/增强子、蛋白质反式激活系统和编码异源蛋白质的 DNA。优选的细胞系应当是非啮齿动物的,因为含 oriP 的质粒不能高效复制,并且 CHO 细胞,例如,缺乏用于反式激活系统的细胞因子。

[0018] Wysokenski, D. A., 和 Yates, J. L., (J. Vir. 63(1989)2657-2666) 提到,EBV 质粒复制不能将物种系从灵长类动物细胞交叉到啮齿动物细胞。这基于缺乏与 DNA 复制所涉及的宿主蛋白质(其在啮齿动物中缺乏)的相互作用。

[0019] Tomiyasu, K-i., 等人, Biochem. Biophys. Res. Commun. 253(1998)733-738、Mizuguchi, H., 等人, FEBS Letters 472(2000)173-178 和 W02005/024030 报道了含有 oriP 和 EBNA-1 结构基因以及其它元件的质粒。

[0020] 美国专利 5,976,807 报道了生产表达多种感兴趣的蛋白质的重组真核细胞系的方法。获得表达 EBNA-1 蛋白质的经转染的细胞,其含有至少两种经转染的附加体。

发明内容

[0021] 本发明提供了表达系统,用于在啮齿动物细胞系中表达异源蛋白质。该系统包含

EB 病毒的 oriP/EBNA-1 游离型复制和保持系统。

[0022] 具体地,本发明提供了啮齿动物细胞,其中,所述啮齿动物细胞表达 EB 病毒的核抗原 1 (EBNA-1),并且含有附加体,其中,所述附加体包含:

[0023] a) 原核复制起点;

[0024] b) 选择标记;

[0025] c) EB 病毒 (EBV) 复制起点 (oriP);

[0026] d) 适于在所述啮齿动物细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码所述异源蛋白质的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译区域。

[0027] 本发明还提供了一种方法,用于获得根据本发明的啮齿动物细胞,其中,所述方法包括如下步骤:

[0028] a) 提供啮齿动物细胞;

[0029] b) 提供下述质粒,其包含原核复制起点、选择标记和针对 EB 病毒核抗原 1 (EBNA-1) 的功能性表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码 EBNA-1 蛋白质的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译区域;

[0030] c) 将所述质粒 b) 引入所述啮齿动物细胞 a);

[0031] d) 选择经稳定转化的啮齿动物细胞;

[0032] e) 提供一个或多个下述另一质粒,其包含原核复制起点、选择标记、EB 病毒 (EBV) 复制起点 (oriP)、适于在经转化的啮齿动物细胞中表达异源蛋白质的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码异源多肽的核酸序列和包含聚腺苷酸化信号的 3' 非翻译区域,

[0033] f) 将所述另一质粒 e) 引入啮齿动物细胞 d);

[0034] g) 重复步骤 e) 至 f) 多至五次。

[0035] 本发明还提供了用于生产异源多肽的方法,其中,所述方法包含:

[0036] a) 提供根据本发明的啮齿动物细胞;

[0037] b) 在适合表达所述异源多肽的条件下培养所述啮齿动物细胞;

[0038] c) 从培养物中回收所述异源多肽。

[0039] 本发明还提供了生产根据本发明的啮齿动物细胞的试剂盒,其中,所述试剂盒包含:

[0040] a) 啮齿动物细胞;

[0041] b) 第一质粒,其包含原核复制起点、选择标记和针对 EB 病毒核抗原 1 (EBNA-1) 的功能性表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码 EBNA-1 蛋白质的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译区域;

[0042] c) 第二质粒,其包含原核复制起点、选择标记、EB 病毒 (EBV) 复制起点 (oriP)、适于在啮齿动物细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、用于引入核酸序列的克隆位点和包含聚腺苷酸化信号的 3' 非翻译区域。

[0043] 在本发明的一种实施方案中,啮齿动物细胞是 CHO 细胞。

[0044] 在本发明的另一实施方案中,CHO 细胞选自 CHO-K1 细胞、CHO-DXB11 细胞、CHO-DG44 细胞和表达 EBNA-1 蛋白质的 CHO 细胞。

[0045] 在本发明的另一实施方案中,待表达的异源多肽选自前药、酶、酶片段、酶抑制剂、酶激活剂、生物活性多肽、刺猬(hedgehog)蛋白质、骨形态发生蛋白、生长因子、促红细胞生成素、血小板生成素、G-CSF、白细胞介素、干扰素、免疫球蛋白或免疫球蛋白片段的组。

[0046] 在一种实施方案中,所述异源多肽是免疫球蛋白或免疫球蛋白片段。

[0047] 在另一实施方案中,根据本发明的对异源蛋白质的生产在瞬时转染下进行。

[0048] 在另一实施方案中,所述异源多肽分泌进培养基。

[0049] 在本发明的又一实施方案中,引入啮齿动物细胞的质粒含有至少两种选择标记,优选至少一种原核选择标记以及至少一种真核选择标记。

[0050] 在本发明的又一实施方案中,引入一个啮齿动物细胞的质粒的数量为 1 至 5 个质粒之间。

[0051] 在本发明的另一实施方案中,在同一步骤中引入啮齿动物细胞的质粒的数量在 1 至 3 之间,优选,1 至 2 个质粒之间。

[0052] 发明详述

[0053] 本发明提供了啮齿动物细胞,其中,所述啮齿动物细胞表达 EB 病毒核抗原 1 (EBNA-1),并且含有附加体,其中,所述附加体包含:

[0054] a) 原核复制起点;

[0055] b) 选择标记;

[0056] c) EB 病毒 (EBV) 复制起点 (oriP);

[0057] d) 适于在所述啮齿动物细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码所述异源多肽的核酸序列以及包含聚腺苷酸化信号的 3' 非翻译区域。

[0058] 可用于开展本发明的方法和技术被描述于,例如, Ausubel, F. M., 编著, Current Protocols in Molecular Biology, 卷 I 至 III (1997); Glover, N. D., 和 Hames, B. D., 编著, DNA Cloning: A Practical Approach, 卷 I 和 II (1995), Oxford University Press; Freshney, R. I. (编著), Animal Cell Culture—a practical approach, IRL Press (1986); Watson, J. D., 等人, Recombinant DNA, 第二版, CHSL Press (1992); Winnacker, E. L., From Genes to Clones; N. Y., VCH Publishers (1987); Celis, J., 编著, Cell Biology, 第二版, Academic Press (1998); Freshney, R. I., Culture of Animal Cells: A Manual of Basic Techniques, 第二版, Alan R. Liss, Inc., N. Y. (1987) 中。

[0059] 用于本文中的“核酸”指编码可重组生产的多肽的、至少部分非天然存在的核酸。“非天然存在”可指各核苷酸的序列或所用的功能性元件的组合,所述功能性元件不限于启动子、3' 非翻译区域、增强子或复制起点。核酸可由核酸片段来构成,所述核酸片段优选是 DNA 片段,其是经分离的或通过化学手段合成的。核酸可被整合进另外的核酸,例如,整合进质粒或真核宿主细胞的基因组/染色体。质粒包括穿梭载体和表达载体等。通常,质粒还将包含原核增殖单元,其包含复制起点(例如 ColE1 复制起点)和选择标记(例如,氨苄青霉素或四环素抗性基因),分别用于在细菌中复制和选择载体。

[0060] 同样地,核酸可通过由各核苷酸构成的其核酸序列来表征。

[0061] 用于本申请中的术语“核酸序列”指核酸分子和其变体的核苷酸序列,所述核酸分子和其变体编码肽、多肽或蛋白质或其功能性变体,即,例如,具有不同氨基酸序列但具有

相同生物学功能 / 活性的蛋白质。这些改变是由于,例如,遗传密码子的简并、突变(例如点突变)、缺失、插入等导致的。蛋白质变体的氨基酸序列与亲本蛋白质的氨基酸序列有所不同,这是由于亲本蛋白质序列中一个或多个氨基酸残基的添加、缺失和 / 或取代导致的。普遍地,变体将具有与亲本蛋白质序列有至少 90% 氨基酸序列同一性的氨基酸序列,更优选地,至少 95%,最优选地,至少 99%。

[0062] “表达盒”指这样的核酸序列,其包含对于在细胞中表达和分泌至少其所含有的结构基因来说必要的元件。

[0063] “基因”指这样的片断,例如,位于染色体或质粒上,其对于表达肽、多肽或蛋白质来说是必要的。除了编码区域之外,基因还包含其它功能性元件,包括启动子、内含子和终止子。

[0064] “结构基因”指基因的编码区域,而没有信号序列。

[0065] 在本申请中使用的术语“启动子”指用于推动下游核酸序列转录的调控性核酸序列。启动子序列可选自以下的启动子序列:巨细胞病毒(CMV)、早期和晚期猿猴病毒 40(SV40)(Bernoist, C., 和 Chambon, P., Nature 290(1981)304-10)、劳斯肉瘤病毒(RSV)3' 长末端重复中含有的启动子(Yamamoto 等人, Cell 22 :787-97(1980))、甘油醛磷酸脱氢酶(GADPH)、逆转录病毒 LTRs、延伸因子 1 α (EF-1 α)、泛素、单纯疱疹病毒的胸苷激酶(HSVTK)(还见,例如, Lee, A., 等人, Mol. Cell. 7(1997)495-501, Wagner 等人, Proc. Natl. Acad. Sci. U. S. A. 78 :1441-45(1981))以及金属硫蛋白基因的调节序列(Brinster 等人, Nature 296(1982)39-42)。优选的是强启动子,例如,腺病毒启动子(例如,腺病毒主要晚期启动子),清蛋白启动子, ApoA1 启动子, β -肌动蛋白启动子,热激启动子,异源启动子(例如 CMV),人珠蛋白和生长激素启动子,可诱导启动子(例如 MMT),逆转录病毒 LTR 启动子、RSV 以及胸苷激酶启动子(例如单纯疱疹病毒的胸苷激酶启动子)。尤其优选的是强病毒启动子,例如,腺病毒启动子,即时早期和晚期巨细胞病毒启动子(CMV ;Boshart 等人, Cell 41(1985)521-30),小鼠乳肿瘤病毒启动子(MMTV)和来自长末端重复的劳斯肉瘤病毒启动子(LTR-RSV ;Gorman 等人, Proc. Natl. Acad. Sci. USA 79(1982)6777-81)。

[0066] 用于本申请中的术语“聚腺苷酸化信号”指用于诱导特定核酸序列片断的初级转录本的切割和聚腺苷酸化的核酸序列。包含聚腺苷酸化信号的 3' 非翻译区域可选自包含源自 SV40、牛生长素(BGH)的基因、免疫球蛋白基因和胸苷激酶基因(tk, 例如,单纯疱疹的胸苷激酶聚腺苷酸化信号)的聚腺苷酸化信号的 3' 非翻译区域构成。

[0067] “抗性基因”或“选择标记”可在本申请中互换使用,其是这样的基因,在存在相应选择剂时,其使得携带该基因的细胞被特异性选出或选掉。有用的阳性抗性基因是抗生素抗性基因。该选择标记使得经该基因转化的宿主细胞在存在相应抗生素时被阳性选出,而未经转化的宿主细胞则不能在选择性培养条件下生长或存活。选择标记可以是阳性的、阴性的或双功能的。阳性选择标记使得携带所述标记的细胞被选出,而阴性选择标记则使得携带所述标记的细胞被选择性清除。通常,选择标记将赋予对药物的抗性或对宿主细胞中代谢或异化缺陷的补偿。在原核细胞中,通常使用赋予抗氨苄青霉素、四环素、卡那霉素或氯霉素抗性的基因等。对于真核细胞有用的抗性基因包括,例如,针对氨基糖苷磷酸转移酶(APH)(例如潮霉素磷酸转移酶(hyg)、新霉素和 G418APH)、二氢叶酸还原酶(DHFR)、胸苷激酶(tk)、谷氨酰胺合成酶(GS)、天冬酰胺合成酶、色氨酸合成酶(吡啶)、组氨醇脱氢酶(组

氨醇 D) 的基因和编码对嘌呤霉素、博来霉素、腐草霉素、氯霉素、Zeocin 和霉酚酸的抗性的基因。其它标记基因被描述于 W092/08796 和 W0 94/28143 中。

[0068] 除了在存在相应选择物质时进行选择之外,选择标记还可提供编码正常情况下不存在于细胞中的分子(例如,绿色荧光蛋白(GFP))的基因。带有这种编码 GFP 的基因的细胞可容易地与不带该基因的细胞区分开,这仅通过探测 GFP 发出的荧光即可实现。

[0069] 真核表达载体/质粒可在原核细胞中增殖。因此,真核表达载体/质粒可以并且通常将携带不止一种抗性基因,即,可用于原核选择的一种抗性基因以及可用于真核选择的一种抗性基因。

[0070] 本文中使用的“调控元件”指编码感兴趣的肽、多肽或蛋白质的结构基因的转录和/或翻译必需的、以顺式存在的核苷酸序列。转录调控元件通常包含处于待表达的结构基因上游的启动子、转录起始和终止位点以及聚腺苷酸化信号序列。术语“转录起始位点”指对应于基因中将掺入初级转录本(即,mRNA 前体)的第一个核酸的核酸碱基;转录起始位点可与启动子序列重叠。术语“转录终止位点”指通常呈现于待转录的感兴趣基因 3' 末端的核苷酸序列,其导致 RNA 聚合酶终止转录。聚腺苷酸化信号序列,或多聚 A 添加信号,提供了在真核 mRNA 的 3' 末端的特定位置进行切割以及在核中向经切割的 3' 末端翻译后添加大约 100-200 个腺苷酸序列(多聚 A 尾巴)的信号。聚腺苷酸化信号序列可包括位于切割位点上游大约 10-30 个核苷酸处的共有序列——AATAAA。

[0071] 为生产分泌的多肽,感兴趣的结构基因额外包含编码信号序列/前导肽的 DNA 片段。信号序列将新合成的肽、多肽或蛋白质指导到 ER 膜上以及穿过 ER 膜,在 ER 膜处,所述多肽可进入分泌途径。在蛋白质穿过 ER 膜期间,信号序列可被信号肽酶切下。对于信号序列的功能而言,宿主细胞分泌机制的识别是必要的。因此,所用的信号序列必须被宿主细胞分泌机制的蛋白质和酶识别。

[0072] 翻译调控元件包括翻译起始(AUG)和停止密码子(TAA、TAG 或 TGA)。内部核糖体进入位点(IRES)可包括在一些构建体中。

[0073] “启动子”指控制基因/结构基因或与其有效相连的核酸序列转录的多核苷酸序列。启动子包括针对 RNA 聚合酶结合和转录起始的信号。所使用的启动子将在其中选用的序列预期被表达的宿主细胞的细胞类型中具有功能。来自多种不同来源的大量的启动子,包括组成型、诱导型和阻遏型启动子是本领域公知的(并被鉴定于数据库,例如 GenBank 中),并可(从,例如,保藏机构,例如 ATCC 以及其它商业或个体来源)作为克隆的多核苷酸获得,或在克隆的多核苷酸中获得。“启动子”包含指导结构基因转录的核苷酸序列。通常,启动子位于基因的上游区域(5'),邻近结构基因的转录开始位点。启动子中作用于起始转录的序列元件通常通过共有的核苷酸序列来表征。这些启动子元件包括 RNA 聚合酶结合位点,TATA 序列,CAAT 序列,分化特异性元件(DSE;McGehee, R. E., 等人, Mol. Endocrinol. 7(1993)551-60),环化 AMP 应答元件(CRE),血清应答元件(SRE;Treisman, R., Seminars in Cancer Biol. 1(1990)47-58),糖皮质激素应答元件(GRE)和针对其它转录因子的结合位点,例如,CRE/ATF(0' Reilly, M. A., 等人, J. Biol. Chem. 267(1992)19938-43)、AP2(Ye, J., 等人, J. Biol. Chem. 269(1994)25728)、SP1、cAMP 应答元件结合蛋白质(CREB; Loeken, M. R., Gene Expr. 3(1993)253-64)和八聚体因子(一般而言,见, Watson 等人, 编著, Molecular Biology of the Gene, 第 4 版, The Benjamin/Cummings Publishing

Company, Inc. 1987 和 Lemaigre, F. P. 和 Rousseau, G. G., *Biochem. J.* 303(1994)1-14)。如果启动子是诱导型启动子,那么转录速率应答于诱导剂增加。相反,如果启动子是组成型启动子,转录速率则不受诱导剂调控。阻遏型启动子也是已知的。例如,当生长因素与其细胞表面上的受体结合时,c-fos 启动子被特异性激活。可通过由,例如,CMV 启动子以及接着的两个 Tet- 操纵子位点构成的人工杂合启动子来实现由四环素 (tet) 调控的表达。Tet- 阻遏子与两个 Tet- 操纵子位点结合,阻止转录。加入诱导剂四环素之后,Tet- 阻遏子从 Tet- 操纵子位点释放,转录得以进行 (Gossen, M. 和 Bujard, H. *PNAS* 89(1992)5547-5551)。关于其它诱导型启动子,包括金属硫蛋白和热激启动子,见,例如, Sambrook 等人 (见前文) 和 Gossen, M., 等人, *Curr. Opin. Biotech.* 5(1994)516-520。已被鉴定为用于高水平表达的强启动子的真核启动子包括 SV40 早期启动子、腺病毒主要晚期启动子、小鼠金属硫蛋白 -I 启动子、劳斯肉瘤病毒长末端重复、中国仓鼠延伸因子 1 α (CHEF-1, 见,例如 US 5,888,809)、人 EF-1 α 、泛素和人巨细胞病毒即刻早期启动子 (CMV IE)。

[0074] “启动子”可以是组成型或诱导型的。增强子 (即,作用于启动子上以增加转录的顺式作用 DNA 元件) 可能是必要的,以与启动子一起发挥功能以增加表达水平 (相对于单独用启动子获得的),其可作为转录调控元件被包括进去。通常,含有启动子的多核苷酸片段将还包括增强子 (例如,CMV 或 SV40)。

[0075] 本文中使用的“增强子”指增强与其有效相连的编码序列或基因的转录的多核苷酸序列。与启动子不同,增强子的定向和位置相对独立,其被发现对于转录单元来说是 5' 或 3' 的 (Lusky, M., 等人, *Mol. Cell Bio.*, 3(1983)1108-22),位于内含子中 (Banerji, J., 等人, *Cell*, 33(1983)729-40) 以及位于编码序列本身中 (Osborne, T. F., 等人, *Mol. Cell Bio.*, 4(1984)1293-305)。因此,增强子可被放置于转录起始位点的上游或下游,或者与启动子有相当大的距离,虽然实践中,增强子可能与启动子物理和功能重叠。来自多种不同来源的大量增强子是本领域公知的 (并被鉴定于数据库,例如 GenBank 中),并可 (从,例如,保藏机构,例如 ATCC 以及其它商业或个体来源) 作为克隆的多核苷酸序列获得,或在克隆的多核苷酸序列中获得。多种包含启动子序列的多核苷酸 (例如,商业使用的 CMV 启动子) 也包含增强子序列。例如,上文列出的所有强启动子也可含有强增强子 (见,例如, Bendig, M. M., *Genetic Engineering*, 7(1988)91-127)。

[0076] “有效相连”指两种或多种组分的并列,其中,所描述的组分处于允许它们以其期望的方式发挥功能的关系中。例如,如果启动子和 / 或增强子顺式作用,以控制或调节相连序列的转录的话,那么启动子和 / 或增强子则是与编码序列有效相连的。通常但并非必须,“有效相连”的 DNA 序列是邻接的 (contiguous), 根据需要将两个蛋白质编码区域 (例如分泌前导区和多肽) 连接成邻接并且符合读码框的。但是,虽然有效相连的启动子通常位于编码序列的上游,但其并不一定要与编码序列邻接。增强子不必要邻接。如果增强子增强了编码序列的转录,那么增强子即是与编码序列有效相连的。有效相连的增强子可位于编码序列上游、之中或下游,以及与启动子有相当大的距离。如果聚腺苷酸化位点位于编码序列的下游末端,使得转录通过编码序列进行至聚腺苷酸化序列,那么聚腺苷酸化位点即是与编码序列 / 结构基因有效相连的。相连通过本领域已知的重组方法来完成,例如,使用 PCR 方法和 / 或通过方便的限制性位点进行连接。如果不存在方便的限制性位点,那么可根据常规实践使用合成的寡核苷酸衔接子或接头。

[0077] 本文中使用的术语“表达”指在宿主细胞内发生的转录和 / 或翻译。宿主细胞中想要的产物的转录水平可基于细胞中存在的相应 mRNA 的量来测定。例如,可通过 PCR 或 Northern 杂交(见, Sambrook 等人, *Molecular Cloning :A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York (1989)) 来对从选定的序列转录的 mRNA 加以定量。可通过多种方法来对选定的序列编码的蛋白质加以定量,例如,通过 ELISA, 通过针对蛋白质生物学活性加以检测, 或者通过使用独立于此类活性的检测, 例如, Western 印迹或放射性免疫测定(其中使用识别并结合蛋白质的抗体)(见, 前文的 Sambrook 等人, 1989)。

[0078] “宿主细胞”指将向其中引入编码本发明多肽的基因的细胞。宿主细胞包括用于质粒 / 载体的增殖的原核细胞, 以及用于表达结构基因的真核细胞。通常, 真核细胞是哺乳动物细胞。

[0079] “多肽”是通过肽键连接的氨基酸残基的聚合物, 无论其是天然生产或合成生产的。少于大约 20 个氨基酸残基的多肽可被称为“肽”。

[0080] “蛋白质”是包含一条或多条多肽链的大分子, 其中, 至少一条链具有长度为 100 个氨基酸或更大的氨基酸长度。蛋白质还可包含非肽性组分, 例如, 糖类基团。糖类和其非肽性取代基可通过在其中生产蛋白质的细胞加入到蛋白质中, 其可随着细胞类型变动。本文中将根据其氨基酸主链结构来限定蛋白质; 添加(例如, 糖类基团)通常并不特别指出, 但也可以存在。

[0081] “异源 DNA”或“异源多肽”指天然不存在于给定的宿主细胞中的 DNA 分子或多肽或 DNA 分子的群或多肽的群。对于特定宿主细胞来说异源的 DNA 分子可含有源自宿主细胞物种的 DNA(即, 内源 DNA), 只要该宿主 DNA 与非宿主 DNA(即, 外源 DNA)组合即可。例如, 含有与包含启动子的宿主 DNA 片断有效相连的、编码多肽的非宿主 DNA 片断的 DNA 分子被认为是异源 DNA 分子。反过来, 异源 DNA 分子可包含与外源启动子有效相连的内源结构基因。

[0082] 非宿主 DNA 分子编码的肽或多肽是“异源”肽或多肽。

[0083] 用于本申请中的术语“质粒”和“载体”指用于将遗传物质转入细胞的载体。该物质通常(但并不排它)是环状核酸分子。根据质粒的意欲用途, 其可额外含有复制起点以及必要的复制控制片断, 例如, 以使得质粒可在宿主中复制或转录。

[0084] “克隆载体”是具有在宿主细胞中自主复制的能力的核酸分子, 例如质粒、黏粒、噬菌粒(phageimid)或细菌人工染色体(BAC)。通常, 克隆载体含有一处或少量限制性内切酶识别位点(这使得核酸分子可以以可决定的方式插入, 而不会损失载体的必要生物学功能)以及编码抗性基因的核苷酸序列, 其适用于对经克隆载体转化的细胞进行鉴定和选择。通常, 抗性基因包括提供四环素抗性或氨苄青霉素抗性的基因。

[0085] “表达质粒”是编码将在宿主细胞中表达的蛋白质的核酸分子。通常, 表达质粒包含原核质粒增殖单元(例如, 对于大肠杆菌(*E. coli*)而言, 包含复制起点和选择标记)、真核选择标记以及用于表达感兴趣的结构基因的一个或多个表达盒。通常, “表达盒”包含启动子、5' 非翻译区域、结构基因以及包括聚腺苷酸化信号的 3' 非翻译区域。基因表达通常处于启动子控制下, 此类结构基因被说成是与启动子“有效相连”的。类似地, 如果调控元件调节核心启动子的活性, 那么调控元件和核心启动子是有效相连的。

[0086] “经分离的多肽”是这样的多肽, 其基本不含被细胞组分所污染, 例如糖类、脂类或

与多肽天然相关联的其它蛋白质性杂质。通常,经分离的多肽的制品含有高度纯化形式的多肽,即,至少大约 80% 纯,至少大约 90% 纯,至少大约 95% 纯,超过 95% 纯或超过 99% 纯。显示出特定蛋白质制品含有经分离的多肽的一种途径是通过蛋白质制品进行十二烷基硫酸钠 (SDS)-聚丙烯酰胺凝胶电泳以及对凝胶进行 Coomassie 亮蓝染色之后的单一条带的出现来判断的。但是,术语“经分离得”并不排除以另外的物理形式(例如二体、突变形式(例如点突变)或者可选地糖基化或衍生化的形式)存在的同样的多肽的存在。

[0087] 术语“免疫球蛋白”指由基本由免疫球蛋白基因编码的一种或多种多肽构成的蛋白质。已知的免疫球蛋白基因包括不同的恒定区基因以及无数的免疫球蛋白可变区基因。免疫球蛋白可以以多种形式存在,包括,例如, Fv、Fab 和 F(ab)₂ 以及单链 (scFv) (例如, Huston, J. S., 等人, PNAS USA 85(1988)5879-5883 ; Bird 等人, Science 242(1988)423-426 以及,一般性地, Hood, L. E., Weissman, I., Wood, W. B., 和 Wilson, J. H., Immunology, Benjamin/Cummings, Menlo Park, California. (1983) 和 Hunkapiller 和 Hood, Nature 323(1986)15-16)。

[0088] 免疫球蛋白通常包含至少两条轻链多肽和两条重链多肽。每条重链和轻链多肽链可含有可变区(通常是多肽链的氨基末端部分),其含有能与抗原相互作用的结合结构域。每条重链和轻链多肽链包含恒定区(通常是羧基末端部分)。重链的恒定区介导抗体与 i) 带有 Fc γ 受体 (Fc γ R) 的细胞,例如,吞噬细胞,的结合, ii) 带有新生儿 Fc 受体 (FcRn) (也被称为 Brambell 受体) 的细胞的结合。其还介导与一些因子的结合,所述因子包括经典补体系统的因子,例如,补体成分 (component) (C1q)。

[0089] 免疫球蛋白轻链或重链的可变区进而包含不同片断,即,四个构架区 (FR) 和三个高变区 (CDR)。

[0090] “免疫球蛋白片段”指包含选自下述的一种或多种片断的多肽: C_H1 结构域、铰链区、C_H2 结构域、C_H3 结构域、C_H4 结构域、C_L 结构域、V_H 结构域、V_L 结构域、构架区 1、构架区 2、构架区 3、构架区 4、高变区 1、高变区 2 和高变区 3 (有或没有插入、缺失和 / 或突变)。

[0091] 在本申请中指出的“包含聚腺苷酸化信号的 3' 非翻译区域”是长度为 50-750 个碱基对的 DNA 序列,其提供了用于在真核 mRNA 的 3' 末端的特异性位点进行切割以及在细胞核中向经切割的 3' 末端翻译后添加大约 100-200 个腺苷酸序列 (多聚 A 尾巴) 的信号。建议采用非常高效的聚腺苷酸化信号,因为效率低的切割和聚腺苷酸化可以导致类似操纵子的 mRNA 形成,这可能是不想要的,例如,质粒编码的基因表达的原因。

[0092] 本申请中使用的术语“宿主”和“细胞”指适于接收根据本发明的质粒的细胞。优选地,该细胞选自 CHO (中国仓鼠卵巢) 细胞、BHK (仓鼠肾) 细胞和其它啮齿动物细胞。优选的是 CHO 或源自 CHO 的细胞,其包含 CHO-K1 细胞、CHO-DXB11 细胞、CHO-DG44 细胞和表达 EBNA-1 蛋白质的 CHO 细胞。根据本发明提供的细胞的后代也包括在内。还包括在连续世代具有修饰的、根据本发明提供的细胞的后代,所述修饰可能是由于突变、环境影响或遗传密码的简并性导致的。术语“细胞系”尤其指可体外增殖一段特定时间段的细胞群。在该时间段内,细胞生长并通过细胞分裂增殖。

[0093] 在本申请中使用的术语“瞬时转染”指这样的过程,其中,引入细胞的核酸没有整合进细胞的基因组或染色体 DNA。其实际上作为染色体外元件,例如作为附加体保持于细胞中。附加体的核酸的转录过程不受影响,例如,产生附加体的核酸编码的蛋白质。

[0094] 本申请中使用的术语“稳定转化”指外源核酸可遗传地稳定整合进宿主细胞基因组/染色体。

[0095] 本文中使用的术语“生物活性多肽”指当被给予人工生物学系统（例如使用细胞系和病毒进行的生物检测）或体内给予动物（包括但不限于鸟和哺乳动物，包括人）时导致生物学效果的有机分子，例如生物大分子，例如，肽、蛋白质、核蛋白、黏蛋白、脂蛋白、合成多肽或蛋白质。生物学效果可以是（但不限于）酶抑制、激活或变构修饰，与受体的结合（在结合位点或周围），阻碍或激活受体，信号触发或抗原结合。

[0096] EBV 的 oriP 复制起点、EBV 二重对称元件和 EBV 重复家族的核酸序列表示 SEQ ID NO :01 至 SEQ ID NO :03, 其已从 GenBank 条目 V01555 获得。编码 EBNA-1 蛋白质的序列表示在 SEQ ID NO :04(核酸序列)和 SEQ ID NO :05(氨基酸序列)。这些序列已分别从 V01555(GenBank) 和 P03211(SwissProt) 获得。

[0097] 已经很好地建立了不同的方法，它们被广泛用于蛋白质纯化，例如用微生物蛋白质进行的亲和色谱（例如，A 蛋白或 G 蛋白亲和色谱）、离子交换色谱（例如，阳离子交换（羧甲基树脂）、阴离子交换（氨基乙基树脂）和混合模式交换）、亲疏吸附（例如，用 β 巯基乙醇和其它 SH 配体）、疏水相互作用或芳香族吸附色谱（例如，用苯基-琼脂糖凝胶、氮杂-arenophilic 树脂或间氨基苯硼酸）、金属螯合亲和色谱（例如，用 Ni(II)-和 Cu(II)-亲和材料）、大小排除色谱和电泳方法（例如凝胶电泳、毛细管电泳）(Vijayalakshmi, M. A., Appl. Biochem. Biotech. 75(1998)93-102)。

[0098] 本发明提供了方法，用于在啮齿动物细胞中表达目标肽、多肽和蛋白质。本发明还提供了啮齿动物细胞，用于生产异源肽、多肽或蛋白质，其中，编码所述异源肽、多肽或蛋白质的结构基因是啮齿动物细胞内的附加体提供的。由两个或多个亚基组成的多肽或蛋白质，即，例如免疫球蛋白也可按照本发明的方法来生产。例如，在免疫球蛋白（由两对轻链和重链构成）的情况下，可以使用含有编码免疫球蛋白两种链的两个结构基因的一个附加体。或者，可以使用两个附加体，例如，其中一个含有编码轻链的结构基因，一个含有编码重链的结构基因。

[0099] 在本发明中使用指导 EBNA-1 蛋白质表达的启动子。优选地，使用强启动子，优选地，使用异源强启动子，例如，CMV。EBNA-1 蛋白质的表达盒稳定整合进染色体 DNA 导致产生经修饰的啮齿动物细胞，其表达 EBNA-1 蛋白质，即，编码 EBNA-1 蛋白质的结构基因，并且将和启动子有效连接地整合进染色体 DNA。

[0100] 为阐述本发明的主题，将在下文中对本发明加以描述。作为第一个步骤，构建表达 EBNA-1 蛋白质的基础啮齿动物细胞系。优选地，构建这样的基础啮齿动物细胞系，其在培养物中稳定表达 EBNA-1 蛋白质。之后，设计含有针对一种或多种异源多肽的表达盒的质粒，将其引入基础啮齿动物细胞系。在适于表达异源多肽的条件下，培养获得的表达 EBNA-1 蛋白质并且携带有用于表达异源多肽的质粒的细胞系（见实施例 1-4）。

[0101] 提供下述描述和实施例是为了阐述本发明的目的，而非限制本发明范围的目的。

[0102] 构建宿主细胞系

[0103] 本发明的宿主细胞提供编码 EBNA-1 蛋白质的核酸。

[0104] 在实施例 2a) 中描述了对含有下述表达盒的质粒的构建，所述表达盒具有编码 EBNA-1 蛋白质的结构基因。质粒的带注释图谱示于图 1。

[0105] 用该质粒转染啮齿动物细胞。在选择性培养条件（用 G418 作为选择剂）下对经转染的细胞进行选择后，挑出转化体，扩展，并针对 EBNA-1 蛋白质的生产加以测试。

[0106] 构建的用于形成基础啮齿动物细胞系的质粒不含 EBV oriP 序列的功能性拷贝。这包括 oriP 的两个元件——二重对称元件和重复家族。

[0107] 构建表达质粒

[0108] 为表达异源肽、多肽或蛋白质，必须构建含有针对相应结构基因的表达盒的表达质粒。

[0109] 对蛋白质举例而言，选用单克隆人抗 IGF-1R 抗体（见 US2005/0008642）。为生产 HuMab 抗 IGF-1R，使用本发明的基础啮齿动物细胞系，构建出表达质粒（见实施例 3 和图 3-5）。构建的用于表达异源多肽的质粒不含 EBNA-1 结构基因的功能性拷贝。

[0110] 一般而言，通过亚克隆和重叠 PCR，将编码 HuMab 抗 IGF-1R 轻链可变区 (V_L) 和人 κ -轻链恒定区 (C_L) 的结构基因连接起来，针对 HuMab 抗 IGF-1R 重链可变区 (V_H) 和人 $\gamma 1$ 重链恒定区 (C_{H1} -铰链- C_{H2} - C_{H3}) 也一样。随后将这些构建体插入缺乏或携带 EBV oriP 的哺乳动物细胞表达载体。

[0111] 在基础啮齿动物细胞系中表达多肽

[0112] 用一种或多种构建的表达质粒来转染根据本发明的基础啮齿动物细胞系。在适于表达异源多肽的条件下，例如，在瞬时转染下，进行对经转染细胞的培养。

[0113] 一般而言，在瞬时转染下的培养允许：经转染的表达质粒不整合进宿主细胞的染色体，这是由于不存在选择剂施加的选择压力。由于这些非选择性的培养条件，表达质粒不能在所有培养细胞中以 100% 的频率复制。这导致总的蛋白质表达率缓慢降低。培养通常进行至直到培养物中表达的蛋白质浓度达到最大值时。优选地，培养进行多至 20 天，优选地，多至 10 天，更优选地，5 至 10 天。

[0114] 在该时间段后，分出上清液，按照本领域技术人员已知的标准技术，对产生的分泌的免疫球蛋白进行分离和纯化。

[0115] 除了上文所述的分批培养之外，还可使用分流分批 (split-batch) 工艺来培养细胞。在分流分批培养工艺期间，在培养期一半之后更换营养培养基。

[0116] 因此，本发明的目的是提供表达异源多肽的方法，该方法允许在短时间内生产出具有与哺乳动物细胞（优选地，人细胞）相似的糖基化模式的所述异源多肽。现在已经吃惊地发现，根据本发明的方法满足此需求。

[0117] 该方法基于的细胞系是啮齿动物细胞，优选地，CHO 细胞，其遵从针对生产治疗性多肽和蛋白质的细胞系建立的规则，即，不含原癌基因，例如多瘤病毒的大 T 抗原。

[0118] 已经吃惊地发现，在用编码 EBNA-1 蛋白质的结构基因（优选地，与启动子有效相连）稳定转染的细胞中，染色体外元件，例如，表达质粒（具有 EB 病毒 (EBV) 复制起点并且没有，即，不含编码 EBNA-1 蛋白质的结构基因）在对异源多肽的表达中发挥功能。一般而言，在根据本发明的方法中，源自 EB 病毒的两个元件是采用复制起点 (oriP) 和编码 EBNA-1 蛋白质的结构基因。

[0119] 本发明的一个方面是在啮齿动物细胞中表达异源多肽的方法，其特征在于，所述方法包括：

[0120] a) 提供经编码 EBNA-1 蛋白质的结构基因稳定转染的啮齿动物细胞，

- [0121] b) 用包含 EB 病毒 (EBV) 复制起点 (oriP) 的表达质粒来转染所述啮齿动物细胞,
- [0122] c) 在适于表达所述异源多肽的条件下培养所述经转染的细胞,
- [0123] d) 从培养物中回收所述异源多肽。
- [0124] 在一种实施方案中,所述异源多肽是免疫球蛋白或免疫球蛋白片段,优选地,所述免疫球蛋白是免疫球 G 蛋白或免疫球蛋白 E。
- [0125] 在一种实施方案中,所述异源多肽是分泌的异源多肽,并且从培养基中回收所述异源多肽。
- [0126] 在一种实施方案中,所述啮齿动物细胞是 CHO 细胞。
- [0127] 在一种实施方案中,所述编码 EBNA-1 蛋白质的结构基因与启动子,优选地,强启动子,尤其优选地,源自 CMV 的启动子有效相连。
- [0128] 在一种实施方案中,所述表达质粒包含单个源自 EBV 的元件——EB 病毒 (EBV) 复制起点 (oriP)。
- [0129] 在一种实施方案中,所述表达质粒不包含编码 EBNA-1 蛋白质的结构基因。
- [0130] 在一种实施方案中,所述表达质粒额外包含
- [0131] - 原核复制起点
- [0132] - 选择标记
- [0133] - 适于在所述啮齿动物细胞中表达异源多肽的表达盒,其中,所述表达盒包含启动子序列、5' 非翻译区域、编码所述异源蛋白质的核酸序列 (结构基因) 和包含聚腺苷酸化信号的 3' 非翻译区域。
- [0134] 在一种实施方案中,如果异源多肽是分泌的异源多肽,那么所述表达盒额外饱含与编码所述异源多肽的结构基因有效相连的信号序列。
- [0135] 在一种实施方案中,表达异源多肽的方法作为分批工艺、分流分批工艺或作为连续工艺进行。在优选的实施方案中,所述方法作为分批或分流分批工艺进行。
- [0136] 在一种实施方案中,表达异源多肽的方法额外含有步骤 e) 纯化所述异源多肽。
- [0137] 一般性的色谱方法及其用途是本领域技术人员已知的。见,例如, Chromatography, 第五版,部分 A: Fundamentals and Techniques, Heftmann, E. (编著), Elsevier Science Publishing Company, New York (1992); Advanced Chromatographic and Electromigration Methods in Biosciences, Deyl, Z. (编著), Elsevier Science BV, Amsterdam, The Netherlands (1998); Chromatography Today, Poole, C. F., 和 Poole, S. K., Elsevier Science Publishing Company, New York (1991); Scopes, Protein Purification: Principles and Practice (1982); Sambrook, J., 等人 (编著), Molecular Cloning: A Laboratory Manual, 第二版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989), 或 Current Protocols in Molecular Biology, Ausubel, F. M., 等人 (编著), John Wiley & Sons, Inc., New York。
- [0138] 一般而言,对免疫球蛋白的纯化工艺包含多步骤的色谱部分。在第一个步骤中,通过亲和色谱,例如用 A 蛋白将非免疫球蛋白多肽和蛋白质与免疫球蛋白级分分开。之后,可进行离子交换色谱来分离各免疫球蛋白的类以及除去从第一柱上共洗脱下来的痕量 A 蛋白。最后,必须进行第三个色谱步骤,来将同一类的免疫球蛋白单体与多体和片段分开。聚集物的量有时较高 (5% 或更高),在第三个纯化步骤中不可能对它们高效分离,因此还需

要后续纯化步骤。

[0139] 随着对特定免疫球蛋白的重组生产,用于分离不同的免疫球蛋白类的分离步骤不再是必需的。因此,对重组产生的免疫球蛋白的整个纯化工序可减至两个色谱步骤。

[0140] 通常,在低于各免疫球蛋白等电点的 pH 值下,在阳离子交换材料上对 A 蛋白洗脱物进行色谱加工。

[0141] 在一种实施方案中,表达异源多肽的方法在瞬时转染下进行。

[0142] 在所述用于表达异源多肽的方法的一种实施方案中,所述编码 EBNA-1 蛋白质的结构基因是编码 EBNA-1 蛋白质的全长结构基因 (SEQID NO :4),所述表达质粒不包含编码 EBNA-1 蛋白质的全长结构基因 (SEQID NO :4)。

[0143] 下述实施例、序列和附图被提供来辅助理解本发明,本发明的实际保护范围示于所附的权利要求中。应当理解,可在不偏离本发明宗旨的情况下对示出的流程加以改变。

附图说明

[0144] 图 1 :质粒 pcDNA3 EBNA-1 的图谱,该质粒含有 EBNA-1 蛋白质的表达盒。

[0145] 图 2 :对 EBNA-1 表达的 Western 印迹分析。

[0146] 图 2a) 的图例 :道 1 :293EBNA,65 μ g 蛋白质 ;道 2 :CHO-DG44,100 μ g 蛋白质 ;道 3 :DG-700-IIH7,100 μ g 蛋白质 ;道 4 :DG-700-IIID9,100 μ g 蛋白质 ;道 5 :DG-700-IIIE3,100 μ g 蛋白质 ;道 6 :DG-700-IIIE8,100 μ g 蛋白质 ;道 7 :DG-700-IIIF1,100 μ g 蛋白质 ;道 8 :DG-700-IIIG10,100 μ g 蛋白质 ;道 9 :DG-700-IIIH8,100 μ g 蛋白质。

[0147] 图 2b) 的图例 :道 1 :蛋白标准 ;道 2 :293EBNA,100 μ g 蛋白质 ;道 3 :DG-700-IIIE3sub1,第 4 代,第 6 天,100 μ g 蛋白质 ;道 4 :DG-700-IIIE3sub1,第 15 代,第 37 天,100 μ g 蛋白质 ;道 5 :DG-700-IIIE3sub1,第 18 代,第 48 天,100 μ g 蛋白质 ;道 6 :DG-700-IIIE3sub1,第 20 代,第 54 天,100 μ g 蛋白质。

[0148] 图 3a :质粒 p4816-pUC-L-IR18- κ BsmI 的图谱

[0149] 图 3b :质粒 p4817-pUC-H-IR18- γ 1BsmI 的图谱

[0150] 图 4a :质粒 p4818-pUC-Hyg-OriP-重-IR18-BsmI 的图谱

[0151] 图 4b :质粒 p4819-pUC-Hyg-OriP-轻-IR18-BsmI 的图谱

[0152] 图 5 :质粒 p4821-pUC-Hyg-OriP-IR18-重-轻-BsmI 的图谱

[0153] 图 6 :对抗体在不同细胞系中以及通过不同表达质粒的表达进行的比较

[0154] 序列描述

[0155] SEQ ID NO :01 :EBV oriP 的核酸序列 ;V01555 (GenBank) ;Yates, J. L. , 等人, Proc Natl Acad Sci USA 81 (1984) 3806-3806。

[0156] SEQ ID NO :02 :EBV 二重对称元件的核酸序列 ;V01555 (GenBank) ;Reisman, D. , 等人, Mol. Cell Biol. 5 (1985) 1822-1832。

[0157] SEQ ID NO :03 :EBV 重复家族的核酸序列 ;V01555 (GenBank) ;Reisman 等人, 1985。

[0158] SEQ ID NO :04 :编码 EBNA-1 蛋白质的核酸序列 ;V01555 (GenBank)。

[0159] SEQ ID NO :05 :EBNA-1 蛋白质的氨基酸序列 ;P03211 (SwissProt)。

[0160] SEQ ID NO :06 :引物寡核苷酸 1。

[0161] SEQ ID NO :07 :引物寡核苷酸 2。

- [0162] SEQ ID NO :08 :引物寡核苷酸 3。
- [0163] SEQ ID NO :09 :引物寡核苷酸 4。
- [0164] 实施例描述
- [0165] 实施例 1 :一般性技术。
- [0166] 实施例 2 :构建表达 EBNA-1 蛋白质的 CHO 细胞系。
- [0167] 实施例 3 :构建携带有 EBV oriP 和针对人单克隆抗 IGF-1R 抗体的轻链及重链的表达盒的质粒。
- [0168] 实施例 4 :通过 EBNA-1 阳性的 DG-700-IIIIE3sub1 细胞从携带 oriP 的质粒生产抗体。
- [0169] 实施例 1
- [0170] 一般性技术
- [0171] a) 重组 DNA 技术
- [0172] 按照 Sambrook, J., Fritsch, E. F. 和 Maniatis, T. (1989), *Molecular cloning :A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York 所述,用标准方法来操作 DNA。按照厂商说明书来使用分子生物学试剂。
- [0173] b) DNA 序列测定
- [0174] 通过在麦蒂基因公司 (MediGenomix GmbH) (Martinsried, 德国) 上进行双链测序来测定 DNA 序列。
- [0175] c) DNA 和蛋白质序列分析和序列数据管理
- [0176] 用 GCG(麦迪逊遗传计算机公司 (Genetics Computer Group), Madison, Wisconsin) 的软件包版本 10.2 和 Infomax's Vector NTI Advance 套装版本 8.0 来产生序列、进行图谱定位、分析、注释和阐述。
- [0177] d) 细胞培养技术
- [0178] 按照 *Current Protocols in Cell Biology* (2000), Bonifacino, J. S., Dasso, M., Harford, J. B., Lippincott-Schwartz, J. 和 Yamada, K. M. (编著), John Wiley & Sons, Inc 所述,使用标准的细胞培养技术。
- [0179] e) 对 EBNA-1 表达进行 Western 印迹分析
- [0180] 通过在 200×g 离心收获细胞,用 PBS(磷酸盐缓冲盐水)洗,在裂解缓冲液 (50mM Tris·HCl(三(羟甲基)氨基甲烷盐酸盐), pH 8.0, 120mM NaCl, 0.5% (v/v) **Nonidet[®] P40**, 10% (v/v) 甘油, 5mM DTT(二硫苏糖醇), 1mM EGTA(亚乙基-双-(氧亚乙基次氨基)四乙酸), 1% (v/v) **Trasyol[®]**, 2mM PMSF(苯甲基磺酰氟)、50 μg/ml 亮抑酶肽 (Leupeptin)) 中于冰上孵育 30 分钟。在 13,000×g 离心之后,收获可溶的上清液,按照厂商方案,使用 Bio-Rad 蛋白质测定 (Cat-Nr. :5000-0001) 针对蛋白质浓度加以测试。
- [0181] 按照厂商推荐,使用 **NuPAGE[®]** 凝胶系统 (英杰生命技术公司 (Invitrogen)), 对蛋白质进行 SDS/聚丙烯酰胺凝胶电泳 (十二烷基硫酸钠, SDS-PAGE) 和电印迹 (electro-blotting)。简言之,将蛋白质裂解物 (100 μg 蛋白质) 与 4 倍体积的还原 LDS(十二烷基硫酸锂)样品缓冲液合并,在 70℃ 孵育 10 分钟,上样到 10% **NuPAGE[®] Novex Bis** /Tris 凝胶 (英杰生命技术公司 (Invitrogen), cat-Nr. :

NP0301) 上。对蛋白质的分离发生在还原**NuPAGE[®] MES SDS** (4- 吗啉代乙磺酸 / 十二烷基硫酸钠) 运行缓冲液中。为将蛋白质从 SDS/ 聚丙烯酰胺凝胶电转移, 使用标准尼龙膜。在电转移之后, 在 50mM Tris · HCl, pH 7.5, 150mM NaCl (TBS, tris 缓冲盐水) 中洗膜, 在 TBS, 1% (w/v) Western 封闭试剂 (罗氏公司 (Roche), Cat. Nr. :11921673001) 中于 4°C 对非特异性结合位点封闭过夜。针对 EBNA-1 的小鼠单克隆抗体 E8.26 (昂科基因公司 (Oncogene), Cat. Nr. :DP15L) 用作为一抗, 其是用 TBS, 1% (w/v) Western 封闭溶液进行了 1 : 1,000 稀释的。在 TBS 中洗两次以及在补充有 0.05% (v/v) Tween-20 的 TBS (TBST) 中洗两次之后, 用偶联过氧化物酶的抗小鼠 / 抗兔 IgG 抗体 (罗氏公司 (Roche), Cat. No. 1520709) 作为二抗, 其是用具有 1% (w/v) Western 封闭溶液的 TBS 进行了 1 : 10,000 稀释的。在用 TBST 洗两次以及用 TBS 洗三次之后, 使用 LumiLightPlus 底物溶液 (罗氏公司 (Roche), Cat. Nr. 12015196001) 和 Lumi-Imager F1 分析仪 (罗氏分子生物化学公司 (Roche MolecularBiochemicals)), 通过化学发光来检测结合的过氧化物酶缀合物。

[0182] f) 对细胞培养物上清液中的重组抗体进行定量

[0183] 使用人 Fc 检测试剂盒 (CIS 生物公司 (Cis Bio), Cat-Nr. :62HFCPEB), 通过竞争性免疫测定来对细胞培养物上清液中的抗体加以定量。该测定基于**HTRF[®]**技术 (均相时间分辨荧光)。简言之, 在稀释缓冲液中对样品进行 1 : 10 至 1 : 100 的稀释。将 50 μ l 经稀释的样品与 25 μ l 抗人 IgG Fc- 穴状化合物 (cryptate) 和 25 μ l 人 IgG-XL665 在 96 孔板 OptiPlates (普金艾尔莫公司 (Perkin Elmer), Cat-Nr. :6005279) 中合并。在室温下对测定混合物孵育过夜。在 320nm 激发后, 使用 Victor 1420 分析仪 (普金艾尔莫公司 (Perkin-Elmer)), 在 620nm 和 665nm 处测量荧光发射。通过与校准曲线比较, 从 665nm/620nm 之比推断出抗体浓度。

[0184] 实施例 2

[0185] 构建表达 EBNA-1 蛋白质的 CHO 细胞系

[0186] a) 构建用于在 CHO 细胞中表达 EBNA-1 蛋白质的质粒 pcDNA3_EBNA-1

[0187] 用限制性核酶 SacI 和 DraIII 切质粒 pcDNA3.1 (英杰生命技术公司 (Invitrogen), Cat-Nr. :V790-20), 将得到的 4715bp 的片段连接到通过退火寡核苷酸 1 (SEQ ID NO :06) 和寡核苷酸 2 (SEQ ID NO :07) 获得的 DNA 接头上。用 HindIII 和 DraIII 切得到的质粒。将 4748bp 的载体片段连接到编码 EBNA-1 的 HindIII/DraIII cDNA 片段上, 该片段是用 pCEP4 (英杰生命技术公司 (Invitrogen), Cat-Nr. :V044-50) 作为模板, 寡核苷酸 3 (SEQ ID NO :08) 和寡核苷酸 4 (SEQ ID NO :09) 作为引物, 通过聚合酶链式反应获得的。将针对 EBNA-1 的真核表达质粒命名为 pcDNA3 EBNA-1 (图 1)。

[0188] b) 用 pcDNA3 EBNA-1 转染 CHO 细胞以及选择稳定的转染子

[0189] 在湿润的温育箱中, 于 37°C 和 5% pCO₂ 下, 在 DHI 培养基 (Schlaeger, E. J., J. Immunol. Methods 194(1996)191-199) 中, 在旋转瓶 (spinner flask) 中, 保持对不含血清的悬浮培养物预适应的 CHO-DG44 细胞。

[0190] 在转染之前, 将细胞以 1×10^6 个细胞 /ml 接种进 24 孔板。按照厂商方案, 使用 LipofectAmine 2000 (英杰生命技术公司 (Invitrogen), Cat-Nr. :11668-027) 来进行转染。简言之, 在 OptiMEM I-培养基 (英杰生命技术公司 (Invitrogen), Cat-Nr. :31985-047) 中稀释 LipofectAmine 2000 和 DNA, 以 1 : 3 至 1 : 6 的 μ g DNA 对 μ l LipofectAmine

2000 之比对它们进行合并。在室温下孵育 20 分钟后,将混合物加入到细胞中。24 小时后,在 DHI 培养基中稀释细胞,以 300 个细胞 / 孔接种进 96 孔板。再 24 小时后,向培养基中以 700 μ g/ml 加入 G418。转染后 10 天,扩展 G418 抗性菌落,通过 Western 印迹杂交(图 2a) 针对 EBNA-1 蛋白质的表达加以分析。将稳定表达 EBNA-1 蛋白质的 HEK 293 细胞(HEK 293 EBNA 或 HEK 293 E) 和未经转染的 CHO-DG44 细胞用作为参照。克隆 DG-700-IIIIE3 对于 EBNA-1 来说是强阳性的,通过有限稀释对其亚克隆。连续培养 EBNA-1 阳性亚克隆 DG-700-IIIIE3sub1,通过 Western 印迹杂交检查 EBNA-1 表达(图 2b)。如图 2b 所示,EBNA-1 蛋白质水平在 54 天的培养中保持不变。

[0191] 实施例 3

[0192] 构建携带 EBV oriP 和针对人单克隆抗 IGF-1R 抗体的轻链和重链的表达盒的质粒

[0193] 通过亚克隆和重叠 PCR,将编码抗 IGF-1R 抗体轻链可变 (V_L) 区和人 κ -轻链恒定 (C_L) 区的基因片断精确连接,对针对抗 IGF-1R 抗体可变重链 (V_H) 区和人 $\gamma 1$ -重链恒定 ($C_H1-C_H2-C_H3$) 区的基因也一样。通过 DNA 测序验证编码抗 IGF-1R 抗体结构基因(κ -轻链和 $\gamma 1$ -重链)的 DNA 序列,随后将其插入缺乏或携带有 EBV oriP 的哺乳动物细胞表达载体中。

[0194] HuMab 抗-IR18 κ -轻链表达载体 p4816-pUC-L-IR-18- κ -BsmI (p4816) 由下述元件组成:

[0195] - 来自载体 pUC18 的复制起点,其使得该质粒能在大肠杆菌中复制 (pUC ori)

[0196] - β -内酰胺酶基因,其在大肠杆菌中赋予氨苄青霉素抗性 (Amp)

[0197] - 用于表达抗 IGF-1R 抗体 κ -轻链的转录单元,其由下述元件组成:

[0198] - 来自人巨细胞病毒的主要即时早期启动子和增强子 (hCMVIE1)

[0199] - 合成的 5' -UTR,其包括 Kozak 序列

[0200] - 鼠免疫球蛋白重链信号序列,其包括信号序列内含子 (LI_内含子_L2)

[0201] - 克隆的抗 IGF-1R 抗体可变轻链 cDNA,其在 5' 末端安排有唯一的 Bsm I 限制性位点,以及在 3' 末端安排有剪接供体位点和唯一的 NotI 限制性位点 (V_L)

[0202] - 基因组性的人 κ -轻链基因恒定区,其包括内含子性的小鼠 Ig- κ 增强子 [NotI_小鼠-Ig- κ -增强子-内含子-2_人-内含子-2_C- κ]

[0203] - 人 κ -免疫球蛋白 3' UTR,其包括聚腺苷酸化信号序列 (3' UTRC- κ)

[0204] - 分别处于 5' 和 3' 末端的唯一限制性位点 Sse8371I 和 FseI,以使得能将表达盒转入备选的表达质粒。

[0205] p4816-pUC-L-IR-18- κ -BsmI 的质粒图谱如图 3a) 所示。

[0206] 抗 IGF-1R 抗体 $\gamma 1$ -重链表达载体 p4817-pUC-H-IR-18- $\gamma 1$ -BsmI (p4817) 由下述元件组成:

[0207] - 来自载体 pUC18 的复制起点,其使得该质粒能在大肠杆菌中复制 (pUC ori)

[0208] - β -内酰胺酶基因,其在大肠杆菌中赋予氨苄青霉素抗性 (Amp)

[0209] - 用于表达抗 IGF-1R 抗体 $\gamma 1$ -重链的转录单元,其由下述元件组成:

[0210] - 来自人巨细胞病毒的主要即时早期启动子和增强子 (hCMVIE1)

[0211] - 合成的 5' -UTR,其包括 Kozak 序列

[0212] - 鼠免疫球蛋白重链信号序列,其包括信号序列内含子 (LI_内含子_L2)

[0213] - 克隆的抗 IGF-1R 抗体可变重链 cDNA,其在 5' 末端安排有唯一的 Bsm I 限制性位点,以及在 3' 末端安排有剪接供体位点和唯一的 NotI 限制性位点

[0214] - 基因组性的人 γ 1- 重链基因恒定区,其包括小鼠 Ig- μ (μ)- 增强子 [NotI_ 小鼠 -Ig- μ - 增强子 _human- 内含子 -2_C_H1-C_H2-C_H3,其具有间插内含子]

[0215] - 人 γ 1- 免疫球蛋白 3' UTR,其包括聚腺苷酸化信号序列

[0216] - 分别处于 5' 和 3' 末端的唯一限制性位点 SgrAI 和 AscI,以使得能将表达盒转入备选的表达质粒。

[0217] p4817-pUC-L-IR-18- γ 1-BsmI 的质粒图谱如图 3b) 所示。

[0218] 通过引入下述两个元件,分别从质粒 p4817-pUC-H-IR18- γ 1-BsmI 和 p4816-pUC-L-IR18- κ -BsmI 获得如图 4 所示的质粒 p4818-pUC-Hyg-OriP-重-IR18-BsmI (p4818) 和 p4819-pUC-Hyg-OriP-轻-IR18-BsmI (p4819) ;

[0219] - 潮霉素抗性基因,其适合用作为真核细胞中的选择标记

[0220] -EB 病毒的复制起点 oriP

[0221] p4818-pUC-Hyg-OriP-重-IR18-BsmI 和 p4819-pUC-Hyg-OriP-轻-IR18-BsmI 的质粒图谱如图 4a) 和图 4b) 所示。

[0222] 通过将 p4818 的 SgrAI/AscI 片段(包含 γ 1-重链转录单元)引入 p4819 的 SgrAI 和 AscI 限制性位点,来构建质粒 p4821-pUC-Hyg-OriP-IR18-重-轻-BsmI。得到的质粒含有与 p4819 相同的元件,包括 EBV oriP 和针对 κ -轻链的转录单元,加上 γ 1-重链转录单元。

[0223] p4821-pUC-Hyg-OriP-IR18-重-轻-BsmI 的质粒图谱如图 5 所示。

[0224] 实施例 4

[0225] 通过 EBNA-1 阳性的 DG-700-IIIIE3sub1 细胞从携带 oriP 的质粒生产抗体

[0226] 在 37°C 和 5% CO₂ 下,在静置培养中,在 ProCHO4-CDM(康伯斯公司(Cambrex), Cat-Nr. :BE12-029Q)、2mM 谷氨酰胺、2% (v/v) 50×HT 补充物(英杰生命技术公司(Invitrogen), Cat-Nr. :41065-012) 和 300 μ g/ml G418 中,对 DG-700-IIIIE3sub1 和未经修饰的 CHO-DG44 型细胞加以保持。在转染前 1 小时,通过离心收获细胞,将其以 0.5×10⁶ 个细胞/ml 重新悬浮于 DHI 培养基 (Schlaeger, E. J., J. Immunol. Methods 194(1996) 191-199) 中。

[0227] 使用 LipofectAmine 2000(英杰生命技术公司(Invitrogen), Cat-Nr. : 11668-027) 来进行转染。为转染 1ml 体积的培养物,合并下述组分:200 μ l OptiMEM I(英杰生命技术公司(Invitrogen), Cat-Nr. :31985-047)、2 μ g DNA 和 6 μ l LipofectAmine。在室温下孵育 5 至 30 分钟后,将混合物加入细胞。转染后 5-10 天,收集细胞培养物上清液,按照实施例 1f) 所述针对抗体浓度对其加以测试。

[0228] 已将三组不同的质粒转染进了 CHO-DG44 细胞系或表达 EBNA-1 的源自 CHO-DG44 的细胞系 DG-700-IIIIE3sub1。将用于表达抗体轻链的质粒 p4816 与用于表达抗体重链的 p4817 共转染。两种质粒都缺乏 EBV oriP。将用于表达抗体轻链的质粒 p4819 与用于表达抗体重链的 p4818 共转染。两种质粒都携带有 EBV oriP。最后,单独转染用于表达抗体轻链和重链的质粒 p4821。质粒 p4821 携带有 EBV oriP。

[0229] 如图 6 所示,在 EBNA-1 阳性的 DG-700-IIIIE3sub1 细胞中转染携带有 oriP 的质粒

p4818 和 p4819 导致抗体表达较之向相同细胞系或向 EBNA-1 阴性的 CHO-DG44 细胞中转染没有 oriP 的 p4816 和 p4817 而言增加 2 至 3 倍。当针对抗体轻链和重链的表达盒组合于携带 oriP 的单个质粒（例如 p4821）中时，DG-700-IIIIE3sub1 细胞中的抗体表达甚至更强。相反，携带 oriP 的质粒 p4817/p4818 和 p4821 在 EBNA-1 阴性的 CHO-DG44 细胞中不产生增加的抗体表达水平。结果，当携带 oriP 的质粒转染进 EBNA-1 阳性的 CHO 细胞中时，获得了免疫球蛋白的最大表达。

序列表

<110> 弗·哈夫曼-拉罗切有限公司 (F.Hoffmann-La Roche AG)

<120> 啮齿动物细胞中的蛋白质表达

<130>Case 23391 WO

<150>EP 05023611.6

<151>2005-10-28

<160>9

<170>PatentIn 版本 3.2

<210>1

<211>1998

<212>DNA

<213>EB 病毒 (Epstein Barr Virus)

<400>1

gaattctatc attaaacggc atgcaggaaa aggacaagca gcgaaaattc acgccccctt	60
gggaggtggc ggcatatgca aaggatagca ctcccactct actactgggt atcatatget	120
gactgtatat gcatgaggat agcatatgct acccggatac agattaggat agcatatact	180
accagatat agattaggat agcatatgct acccagatat agattaggat agcctatget	240
accagatat aaattaggat agcatatact acccagatat agattaggat agcatatget	300
accagatat agattaggat agcctatget acccagatat agattaggat agcatatget	360
accagatat agattaggat agcatatgct atccagatat ttgggtagta tatgetacce	420
agatataaat taggatagca tatactacce taatctctat taggatagca tatgetacce	480
ggatacagat taggatagca tatactacce agatatagat taggatagca tatgetacce	540
agatatagat taggatagcc tatgetacce agatataaat taggatagca tatactacce	600
agatatagat taggatagca tatgetacce agatatagat taggatagcc tatgetacce	660
agatatagat taggatagca tatgetatcc agatatttgg gtagtatatg ctaccatgg	720
caacattagc ccaccgtgct ctcagegacc tcgtgaatat gaggaccaac aaccctgtgc	780
ttggcgctca ggcgcaagtg tgtgtaattt gtccctcaga tcgcagcaat cgcgccceta	840
tcttggcccc cccacctact tatgcaggta ttccccgggg tgccattagt ggttttgtgg	900
gcaagtggtt tgaccgcagt ggtttagcggg gttacaatca gccaaagtat tacaccetta	960
ttttacagtc caaaaccgca gggcggcgtg tgggggctga cgcgtgcccc cactccacaa	1020
tttcaaaaaa aagagtggcc acttgtcttt gtttatgggc ccatttggcg tggagccccg	1080

tttaattttc	gggggtgta	gagacaacca	gtggagtccg	ctgctgtcgg	cgteccactct	1140
ctttcccctt	gttaciaaata	gagtgttaaca	acatggttca	cctgtcttgg	tcctcgctg	1200
ggacacatct	taataacccc	agtatcatat	tgactagga	ttatgtgttg	cccatagcca	1260
taaattcgtg	tgagatggac	atccagtctt	tacggcttgt	ccccaccca	tggatttcta	1320
ttgttaaaga	tattcagaat	gtttcattcc	tacactagta	tttattgccc	aaggggttg	1380
tgagggttat	attgggtgca	tagcacaatg	ccaccactga	acccccgctc	caaattttat	1440
tctgggggcg	tcacctgaaa	ccttgttttc	gagcacctca	catacacctt	actgttcaca	1500
actcagcagt	tattctatta	gctaaacgaa	ggagaatgaa	gaagcaggcg	aagattcagg	1560
agagttcact	gcccgtcct	tgatcttcag	ccactgcctt	tgtgactaaa	atggttcact	1620
accctcgtgg	aatcctgacc	ccatgtaaat	aaaaccgtga	cagctcatgg	ggtgggagat	1680
atcgtgttc	cttaggacce	ttttactaac	cctaattcga	tagcatatgc	ttcccgttgg	1740
gtaacatatg	ctattgaatt	agggttagtc	tggatagtat	atactactac	ccgggaagca	1800
tatgctacce	gtttagggtt	aacaaggggg	ccttataaac	actattgcta	atgccctctt	1860
gagggtccgc	ttatcggtag	ctacacagge	ccctctgatt	gacgttgggtg	tagcctcccg	1920
tagtcttctt	gggcccctgg	gaggtacatg	tccccagca	ttgggtgtaag	agcttcagcc	1980
aagagttaca	cataaagg					1998

<210>2

<211>146

<212>DNA

<213>EB 病毒

<400>2

gatategctg	ttccttagga	cccttttact	aaccctaatt	cgatagcata	tgcttcccgt	60
tgggtaacat	atgctattga	attagggtta	gtctggatag	tatatactac	taccgggaa	120
gcatatgcta	cccgtttagg	gttaac				146

<210>3

<211>622

<212>DNA

<213>EB 病毒

<400>3

gggtatcata	tgctgactgt	atatgcatga	ggatagcata	tgctaccg	atacagatta	60
ggatagcata	tactaccag	atatagatta	ggatagcata	tgctaccag	atatagatta	120
ggatagccta	tgctaccag	atataaatta	ggatagcata	tactaccag	atatagatta	180
ggatagcata	tgctaccag	atatagatta	ggatagccta	tgctaccag	atatagatta	240
ggatagcata	tgctaccag	atatagatta	ggatagcata	tgctatccag	atatttgggt	300
agtatatgct	accagatat	aaattaggat	agcatatact	accctaactt	ctattaggat	360

agcatatgct	acccgatac	agattaggat	agcatatact	accagatat	agattaggat	420
agcatatgct	accagatat	agattaggat	agcctatgct	accagatat	aaattaggat	480
agcatatact	accagatat	agattaggat	agcatatgct	accagatat	agattaggat	540
agcctatgct	accagatat	agattaggat	agcatatgct	atccagatat	ttgggtagta	600
tatgctaccc	atggcaacat	ta				622

<210>4

<211>1926

<212>DNA

<213>EB 病毒

<400>4

atgtctgacg	aggggccagg	tacaggacct	ggaaatggcc	taggagagaa	gggagacaca	60
tctggaccag	aaggctccgg	cggcagtgga	cctcaaagaa	gagggggtga	taaccatgga	120
cgaggacggg	gaagaggacg	aggacgagga	ggcgggaagac	caggagcccc	ggcgggtca	180
ggatcagggc	caagacatag	agatggtgtc	cggagacccc	aaaaacgtcc	aagttgcatt	240
ggctgcaaag	ggaccacg	tggaacagga	gcaggagcag	gagcgggagg	ggcaggagca	300
ggaggggcag	gagcaggagg	aggggcagga	gcaggaggag	gggcaggagg	ggcaggaggg	360
gcaggagggg	caggagcagg	aggaggggca	ggagcaggag	gaggggcagg	aggggcagga	420
ggggcaggag	caggagggag	ggcaggagca	ggaggagggg	caggaggggc	aggagcagga	480
ggaggggcag	gaggggcagg	aggggcagga	gcaggaggag	gggcaggagc	aggaggaggg	540
gcaggagggg	caggagcagg	aggaggggca	ggaggggcag	gaggggcagg	agcaggagga	600
ggggcaggag	caggaggggc	aggaggggca	ggaggggcag	gagcaggagg	ggcaggagca	660
ggaggagggg	caggaggggc	aggaggggca	ggagcaggag	gggcaggagc	aggaggggca	720
ggagcaggag	gggcaggagc	aggaggggca	ggaggggcag	gagcaggagg	ggcaggaggg	780
gcaggagcag	gaggggcagg	aggggcagga	gcaggaggag	gggcaggagg	ggcaggagca	840
ggaggagggg	caggaggggc	aggagcagga	ggggcaggag	gggcaggagc	aggaggggca	900
ggaggggcag	gagcaggagg	ggcaggaggg	gcaggagcag	gaggaggggc	aggagcagga	960
ggggcaggag	caggaggttg	aggccggggt	cgaggaggca	gtggaggccg	gggtcagga	1020
ggtagtggag	gccggggtcg	aggaggtagt	ggaggccgcc	gggtagagg	acgtgaaaga	1080
gccagggggg	gaagtcgtga	aagagccagg	gggagaggtc	gtggacgtgg	agaaaagagg	1140
cccaggagtc	ccagtagtca	gtcatcatca	tccgggtctc	caccgcgag	gccccctcca	1200
ggtagaagge	catttttcca	ccctgtaggg	gaagccgatt	attttgaata	ccaccaagaa	1260
ggtggcccag	atggtgagcc	tgacgtgccc	ccgggagcga	tagagcaggg	ccccgcagat	1320
gaccaggag	aaggccaag	cactggaccc	cggggtcagg	gtgatggagg	caggcgcaaa	1380
aaaggagggt	ggtttgaaa	gcatcgtggt	caaggagggt	ccaacccgaa	atttgagaac	1440
attgcagaag	gtttaagagc	tctcctggct	aggagtcacg	tagaaaaggac	taccgacgaa	1500
ggaacttggg	tcgccggtgt	gttcgtatat	ggaggtagta	agacctccct	ttacaaccta	1560
aggcgaggaa	ctgcccttgc	tattccacaa	tgctctctta	caccattgag	tcgtctcccc	1620

tttgaatgg cccctggacc eggcccacaa cctggcccgc taagggagtc cattgtctgt 1680
 tatttcatgg tctttttaca aactcatata tttgctgagg ttttgaagga tgcgattaag 1740
 gaccttgтта tgacaaagcc cgctectacc tgcaatatca gggtgactgt gtgcagcttt 1800
 gacgatggag tagattttgcc tcctggttt ccacctatgg tggaaagggc tgccgcggag 1860
 ggtgatgacg gagatgacgg agatgaagga ggtgatggag atgaggggtga ggaagggcag 1920
 gagtga 1926

<210>5

<211>641

<212>PRT

<213>EB 病毒

<400>5

Met Ser Asp Glu Gly Pro Gly Thr Gly Pro Gly Asn Gly Leu Gly Glu
 1 5 10 15
 Lys Gly Asp Thr Ser Gly Pro Glu Gly Ser Gly Gly Ser Gly Pro Gln
 20 25 30
 Arg Arg Gly Gly Asp Asn His Gly Arg Gly Arg Gly Arg Gly Arg Gly
 35 40 45
 Arg Gly Gly Gly Arg Pro Gly Ala Pro Gly Gly Ser Gly Ser Gly Pro
 50 55 60
 Arg His Arg Asp Gly Val Arg Arg Pro Gln Lys Arg Pro Ser Cys Ile
 65 70 75 80
 Gly Cys Lys Gly Thr His Gly Gly Thr Gly Ala Gly Ala Gly Ala Gly
 85 90 95
 Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Ala Gly
 100 105 110
 Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly
 115 120 125
 Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala
 130 135 140
 Gly Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly Ala Gly
 145 150 155 160
 Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly
 165 170 175
 Ala Gly Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly
 180 185 190
 Ala Gly Gly Ala Gly Ala Gly Gly Gly Gly Ala Gly Ala Gly Gly Ala Gly

195	200	205
Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala		
210	215	220
Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala		
225	230	235
Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly		
245	250	255
Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala		
260	265	270
Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Gly Ala Gly		
275	280	285
Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Ala Gly Gly Ala Gly		
290	295	300
Ala Gly Gly Ala Gly Gly Ala Gly Ala Gly Gly Gly Ala Gly Ala Gly		
305	310	315
Gly Ala Gly Ala Gly Gly Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly		
325	330	335
Arg Gly Arg Gly Gly Ser Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly		
340	345	350
Arg Arg Gly Arg Gly Arg Glu Arg Ala Arg Gly Gly Ser Arg Glu Arg		
355	360	365
Ala Arg Gly Arg Gly Arg Gly Arg Gly Glu Lys Arg Pro Arg Ser Pro		
370	375	380
Ser Ser Gln Ser Ser Ser Ser Gly Ser Pro Pro Arg Arg Pro Pro Pro		
385	390	395
Gly Arg Arg Pro Phe Phe His Pro Val Gly Glu Ala Asp Tyr Phe Glu		
405	410	415
Tyr His Gln Glu Gly Gly Pro Asp Gly Glu Pro Asp Val Pro Pro Gly		
420	425	430
Ala Ile Glu Gln Gly Pro Ala Asp Asp Pro Gly Glu Gly Pro Ser Thr		
435	440	445
Gly Pro Arg Gly Gln Gly Asp Gly Gly Arg Arg Lys Lys Gly Gly Trp		
450	455	460
Phe Gly Lys His Arg Gly Gln Gly Gly Ser Asn Pro Lys Phe Glu Asn		
465	470	475
Ile Ala Glu Gly Leu Arg Ala Leu Leu Ala Arg Ser His Val Glu Arg		
485	490	495
Thr Thr Asp Glu Gly Thr Trp Val Ala Gly Val Phe Val Tyr Gly Gly		
500	505	510

Ser Lys Thr Ser Leu Tyr Asn Leu Arg Arg Gly Thr Ala Leu Ala Ile
 515 520 525
 Pro Gln Cys Arg Leu Thr Pro Leu Ser Arg Leu Pro Phe Gly Met Ala
 530 535 540
 Pro Gly Pro Gly Pro Gln Pro Gly Pro Leu Arg Glu Ser Ile Val Cys
 545 550 555 560
 Tyr Phe Met Val Phe Leu Gln Thr His Ile Phe Ala Glu Val Leu Lys
 565 570 575
 Asp Ala Ile Lys Asp Leu Val Met Thr Lys Pro Ala Pro Thr Cys Asn
 580 585 590
 Ile Arg Val Thr Val Cys Ser Phe Asp Asp Gly Val Asp Leu Pro Pro
 595 600 605
 Trp Phe Pro Pro Met Val Glu Gly Ala Ala Ala Glu Gly Asp Asp Gly
 610 615 620
 Asp Asp Gly Asp Glu Gly Gly Asp Gly Asp Glu Gly Glu Glu Gly Gln
 625 630 635 640
 Glu

<210>6

<211>45

<212>DNA

<213> 人造

<220>

<223> 引物序列

<400>6

cgtttagtga accgtcagat cgcaaaaagc tttttttaac acgta 45

<210>7

<211>46

<212>DNA

<213> 人造

<220>

<223> 引物序列

<400>7

gtgttaaaaa aagcttttttg cgatctgacg gttcactaaa cgagct 46

<210>8

<211>44

<212>DNA

<213> 人造

<220>

<223> 引物序列

<400>8

ttttttaagc tttgccacca tgtctgacga ggggccaggt acag 44

<210>9

<211>42

<212>DNA

<213> 人造

<220>

<223> 引物序列

<400>9

ttttttcact acgtgggtgc tggttgctcc cattcttagg tg 42

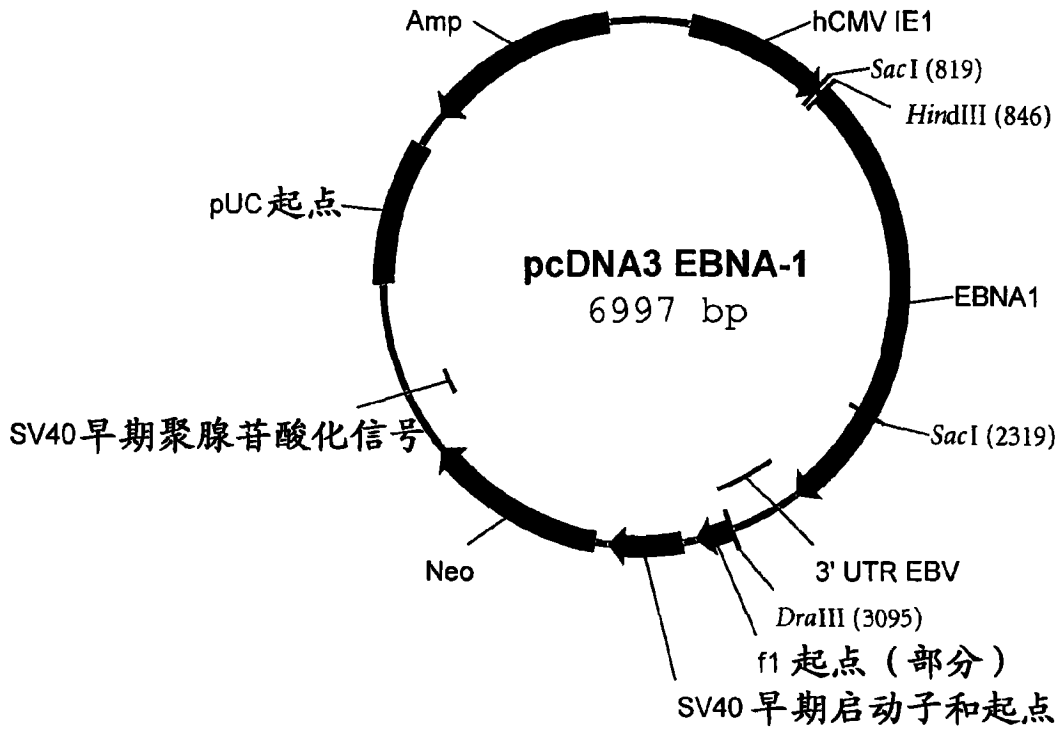
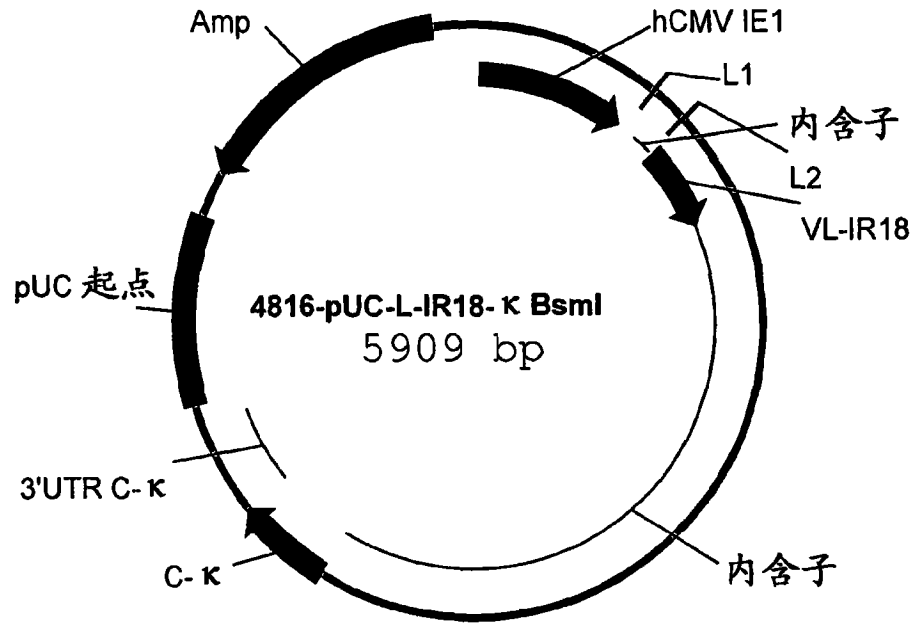


图 1



图 2

a)



b)

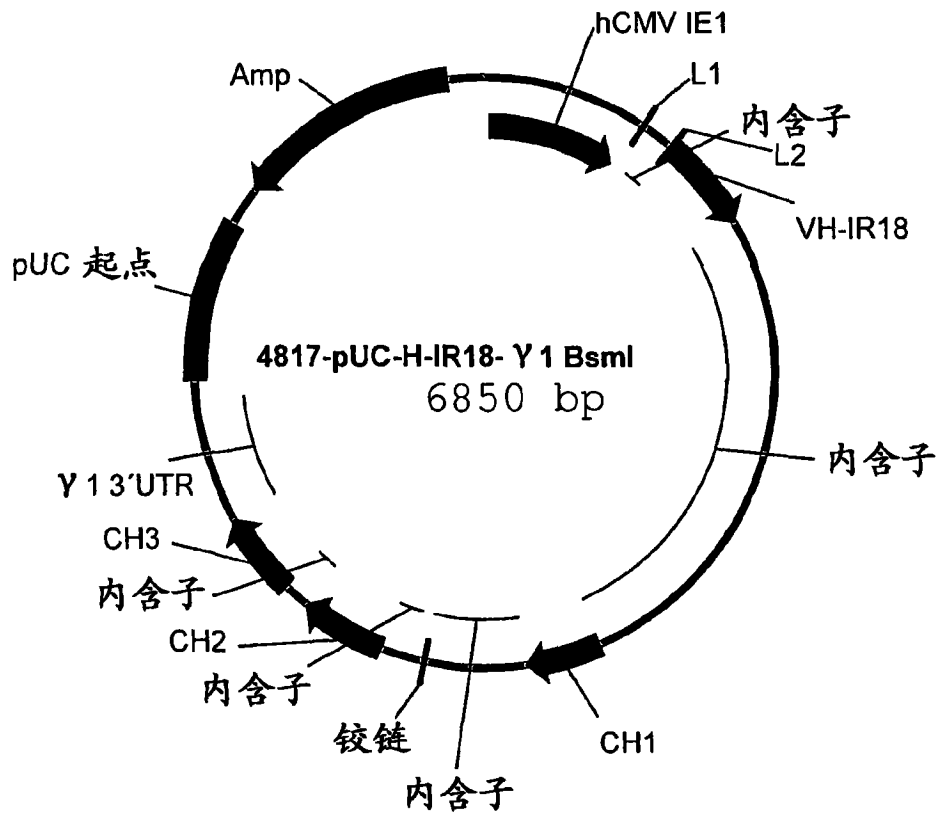
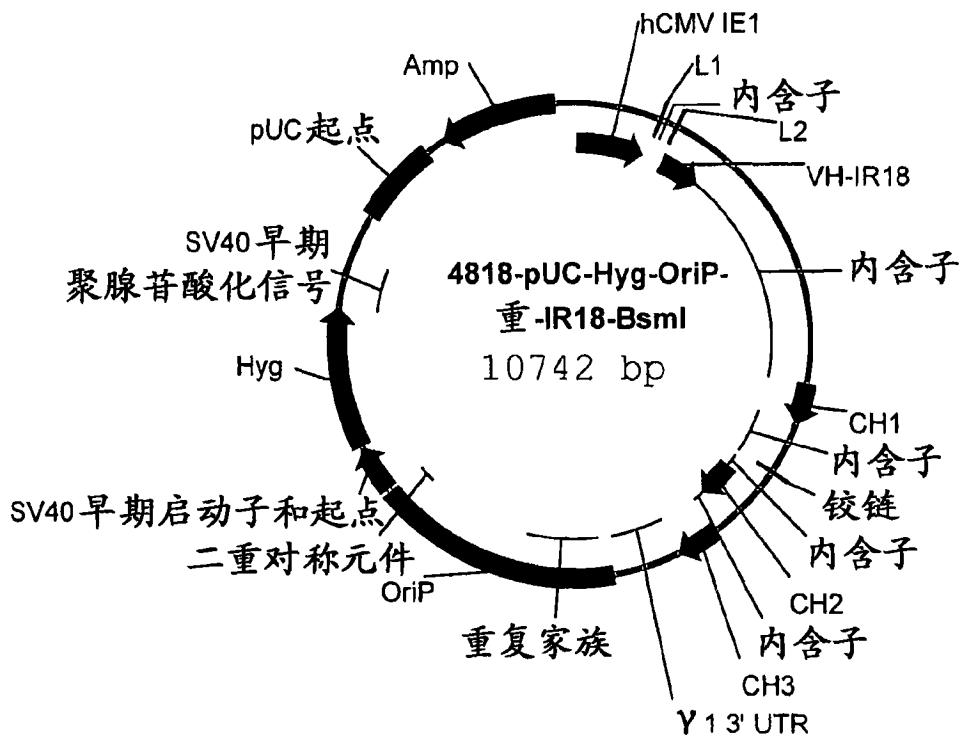


图 3

a)



b)

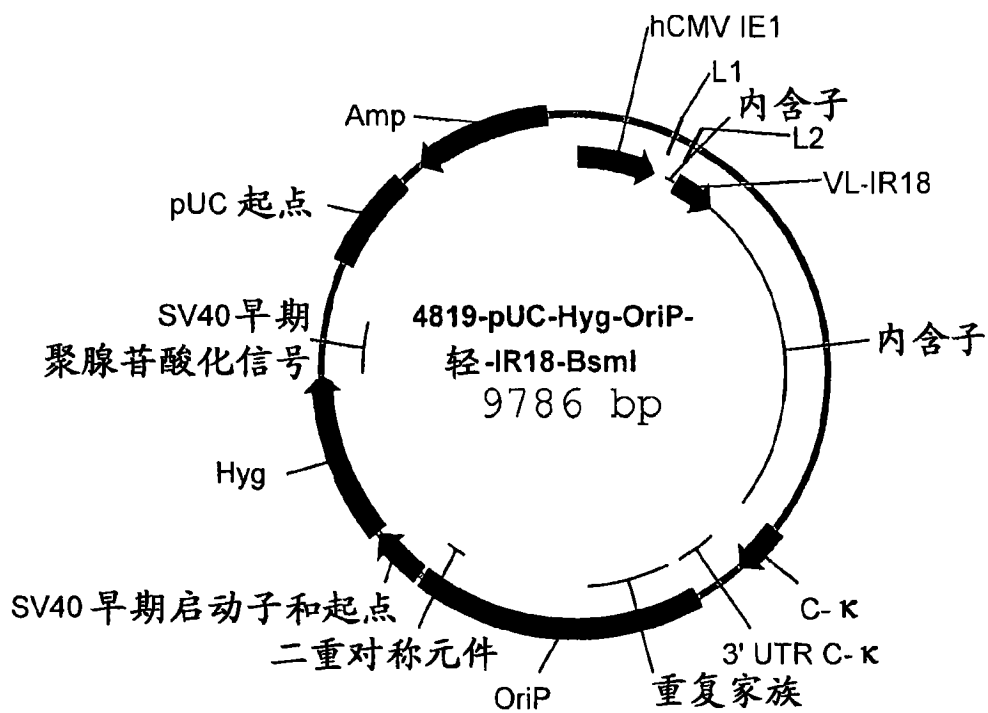


图 4

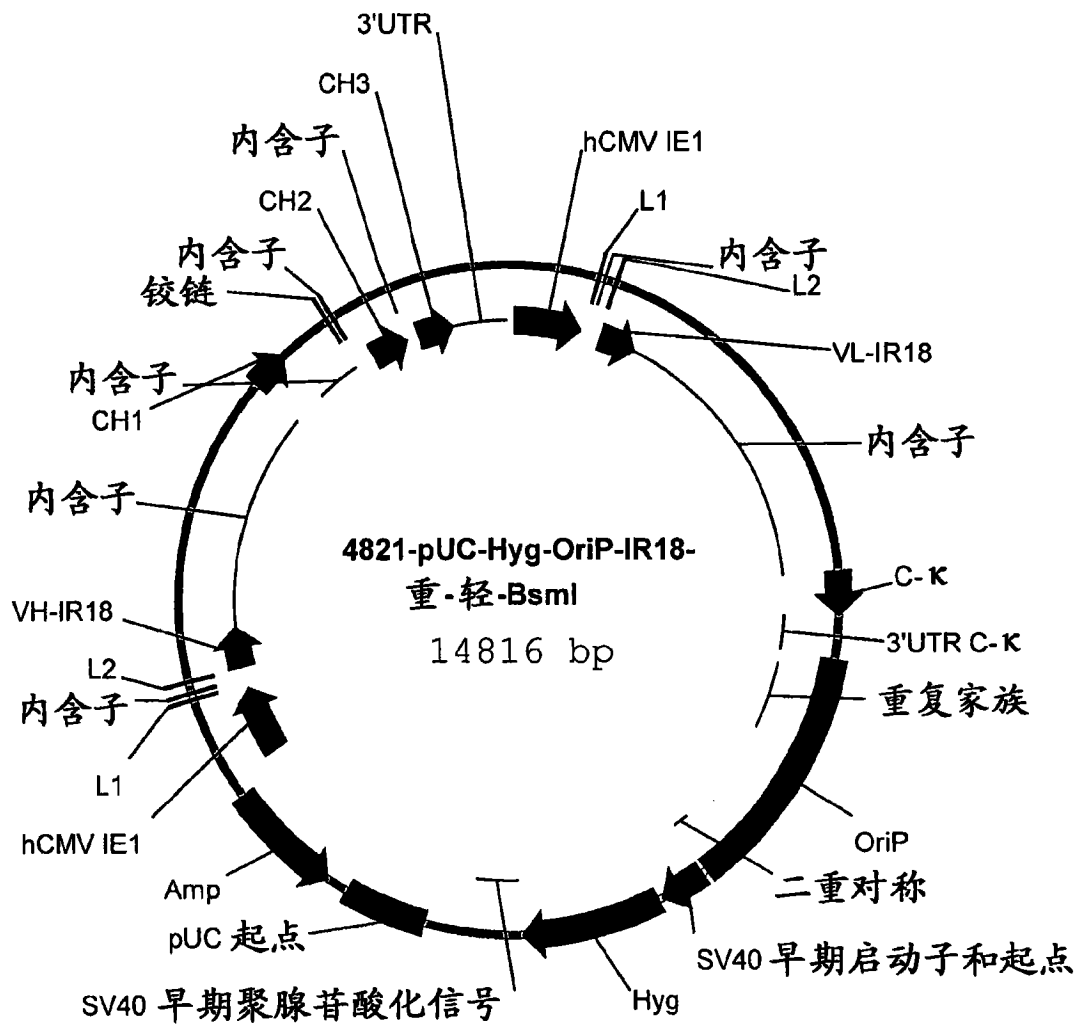
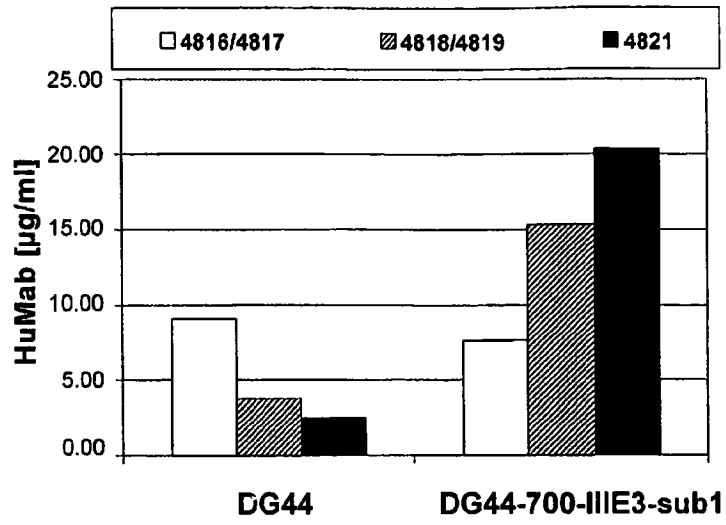


图 5

a)



b)

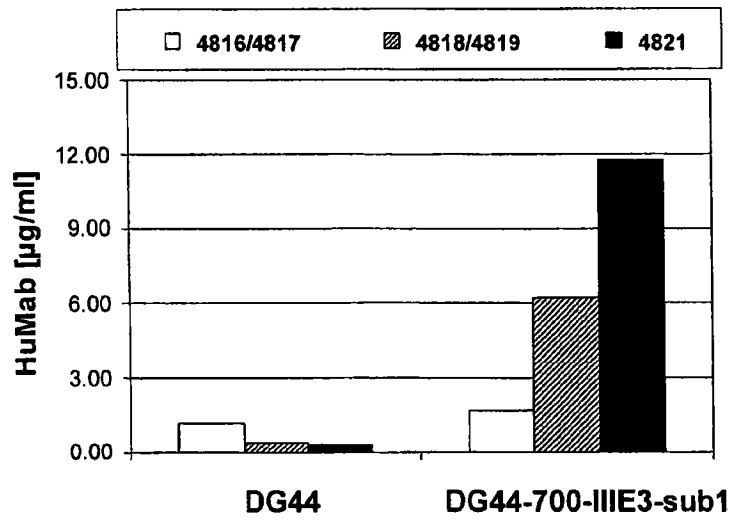


图 6