(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101687920 B

(45)授权公告日 2013.06.19

(21)申请号 200880022771.3

(22)申请日 2008.06.25

(30)优先权数据

07012774.1 2007.06.29 EP

(85)PCT申请进入国家阶段日

2009.12.29

(86)PCT申请的申请数据

PCT/EP2008/005136 2008.06.25

(87)PCT申请的公布数据

WO2009/003623 EN 2009.01.08

(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 S·汉森 H·克内根 E·科佩茨基 O·普罗特奈尔

(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.

C07K 16/00(2006.01)

C12N 15/09(2006.01)

C12N 15/67(2006.01)

C12N 15/00(2006.01)

(56)对比文件

Rua Prof.et al.QV0-TT0010-091199-053-g03 TT0010 Homo sapiens cDNA, mRNA sequence..《EMBL-Bank》.2000, 全文.

审查员 姚进孝

权利要求书2页 说明书23页

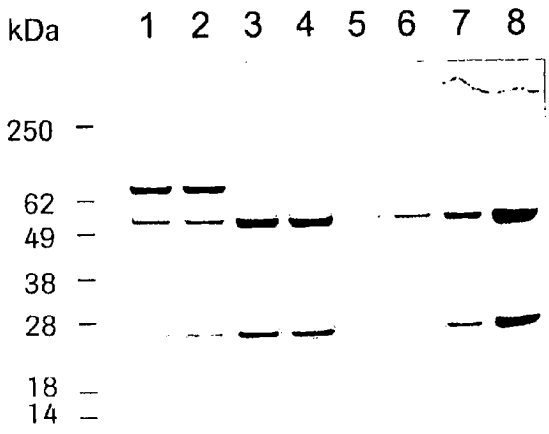
序列表12页 附图4页

(54)发明名称

导致免疫球蛋白生产改善的重链突变体

(57)摘要

本发明包括编码IgA或IgG类免疫球蛋白C_H3-结构域的C-末端部分、或编码IgE或IgM类免疫球蛋白C_H4-结构域的C-末端部分的氨基酸序列的核酸，其中包含在C_H3-或C_H4-结构域的C-末端部分的氨基酸序列中的甘氨酸-赖氨酸二肽，由核酸ggaaaa、或核酸ggcaaa、或核酸gggaaa、或核酸gggaag、或核酸ggcaag、或核酸ggaaag编码。



CN 101687920 B

1. 核酸，其编码 IgA 或 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分氨基酸序列、或编码 IgE 或 IgM 类免疫球蛋白 C_H4- 结构域的 C- 末端部分氨基酸序列，该核酸特征在于包含在 C_H3- 或 C_H4- 结构域的 C- 末端部分的所述氨基酸序列中的甘氨酸 - 赖氨酸二肽，由核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa、或核酸 ggaaag、或核酸 ggcaag、或核酸 gggaag 编码，其中所述 C- 末端部分氨基酸序列为 C- 末端的 20 个氨基酸。

2. 根据权利要求 1 所述的核酸，其特征在于所述核酸编码氨基酸序列，其选自于 SEQ ID NO :1、3、4、5、6、7、或 8 的氨基酸序列。

3. 根据权利要求 1 至 2 任一项所述的核酸，其特征在于编码所述甘氨酸 - 赖氨酸二肽的核酸前面是核苷酸 g 或 a。

4. 根据权利要求 1 至 2 任一项所述的核酸，其特征在于所述甘氨酸 - 赖氨酸二肽是由核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa 编码。

5. 根据权利要求 1 所述的核酸，其特征在于所述 C- 末端部分包括免疫球蛋白重链初级氨基酸序列的至少 20 个 C- 末端氨基酸。

6. 根据权利要求 1 所述的核酸，其特征在于所述核酸编码 IgA、IgE、IgM 或 IgG 类免疫球蛋白重链 C- 末端恒定区的一部分，并选自于 SEQ ID NO :17、18、19、20、21、22、23、30、或 31 的核酸。

7. 根据权利要求 6 所述的核酸，其特征在于所述核酸选自于 SEQ ID NO :17、18、19、22、23、30、或 31 的核酸。

8. 根据权利要求 7 所述的核酸，其特征在于所述核酸选自于 SEQ ID NO :17、18、19、22、或 23 的核酸。

9. 根据权利要求 1 至 2 任一项所述的核酸，其特征在于所述核酸编码 IgG1 或 IgG4 类人免疫球蛋白重链 C- 末端恒定区的一部分。

10. 包含根据权利要求 1 所述核酸的质粒。

11. 包含根据权利要求 1 所述核酸的细胞。

12. 根据权利要求 11 所述的细胞，其特征在于所述细胞是哺乳动物细胞。

13. 根据权利要求 12 所述的细胞，其特征在于所述哺乳动物细胞选自于 CHO 细胞、HEK 细胞或 BHK 细胞。

14. 由 SEQ ID NO :17、18、19、22、23、30、或 31 的核苷酸序列组成的核酸。

15. 根据权利要求 14 所述的核酸，其由 SEQ ID NO :17、18、19、22、或 23 的核苷酸序列组成。

16. 在哺乳动物细胞中生产免疫球蛋白的方法，包括下列步骤 :

a) 用编码免疫球蛋白重链的核酸转染所述哺乳动物细胞，其中 SEQ ID NO :17、18、19、20、21、22、23、30、或 31 的核酸编码免疫球蛋白重链的 C- 末端部分，

b) 在适合免疫球蛋白表达的条件下培养转染的哺乳动物细胞，

c) 从培养物或细胞回收免疫球蛋白。

17. 根据权利要求 16 所述的方法，其特征在于所述哺乳动物细胞用一种核酸转染，该核酸包含编码免疫球蛋白重链的表达盒和编码免疫球蛋白轻链的表达盒。

18. 根据权利要求 16 所述的方法，其特征在于用两种核酸转染所述哺乳动物细胞，其中第一种核酸包含编码免疫球蛋白轻链的表达盒，第二种核酸包括编码免疫球蛋白重链的

表达盒。

19. 根据权利要求16至18中任一项所述的方法，其特征在于所述转染所述哺乳动物细胞是利用编码免疫球蛋白重链的核酸，其中SEQ ID NO：17、18、19、22、23、30、或31的核酸编码免疫球蛋白重链的C-末端部分。

20. 根据权利要求19所述的方法，其特征在于所述转染所述哺乳动物细胞是利用编码免疫球蛋白重链的核酸，其中SEQ ID NO：17、18、19、22、或23的核酸编码免疫球蛋白重链的C-末端部分。

21. 改善哺乳动物细胞中免疫球蛋白表达的方法，其中编码免疫球蛋白重链的核酸包括编码甘氨酸-赖氨酸二肽的核酸ggaaaa、或核酸ggcaaa、或核酸gggaaa、或核酸ggaaag、或核酸酸ggcaag、或核酸gggaag，该甘氨酸-赖氨酸二肽包含在免疫球蛋白重链的C_H3-或C_H4-结构域中，其中所述C-末端部分氨基酸序列为C-末端的20个氨基酸。

导致免疫球蛋白生产改善的重链突变体

[0001] 本申请描述了对哺乳动物细胞中免疫球蛋白的生产有用的方法和核酸。

[0002] 发明背景

[0003] 生产重组多肽的表达系统是众所周知的，并在本领域现有技术的文献中有报道。为了生产用于药物用途的多肽，使用哺乳动物细胞例如CHO细胞、BHK细胞、NSO细胞、Sp2/0细胞、COS细胞、HEK细胞、**PER.C6®**细胞，等等。编码多肽的核酸例如在质粒中被引入细胞中，例如，表达质粒。表达质粒的基本元件是原核质粒繁殖单元，如对于大肠杆菌，其包括复制起点和选择标记，真核选择标记，以及一个或多个用于目的核酸表达的表达盒，其中每一个包括启动子、结构基因和包括聚腺苷酸化信号的转录终止子。对于在哺乳动物细胞中的瞬时表达，可以包括哺乳动物复制起点，如SV40 Ori或OriP。作为启动子可以选择组成型启动子或诱导型启动子。对于优化的转录，可以在5′非编码区中包含Kozak序列。对于mRNA加工，还以包含多腺苷酸化信号。

[0004] 蛋白质和特别是免疫球蛋白，在今天的医疗产品中发挥了重要作用。对于人类应用，每一药物物质必须满足不同的标准。为了确保用于人类的生物制药因子的安全，必须特别去除会造成严重危害的物质。

[0005] mRNA剪接受剪接供体位点与剪接受体位点的组合调控，这两个位点分别位于内含子的5′末端和3′末端。根据Watson等(Watson等(编)，Recombinant DNA:A Short course，Scientific American Books，distributed byW.H.Freeman and Company，New York，New York，U SA(1983))是5′剪接供体位点ag|gtragt(外显子|内含子)和3′剪接受体位点$(y)_nNcag|g$(内含子|外显子)的共有序列(r＝嘌呤碱基；y＝嘧啶碱基；n＝整数；N＝任何自然碱基)。

[0006] 1980年发表了第一篇论述分泌型和膜结合型免疫球蛋白起源的文章。重链前体mRNA的选择性剪接导致分泌型(slg)和膜结合型(mlg)同种型的形成。在mlg同种型中，编码分泌型的C-末端区域的外显子中的剪接供体位点(即，分别C_H3或C_H4结构域)和位于其下游一段距离处的剪接受体位点被用于连接恒定区和编码跨膜区的下游外显子。

[0007] 在WO 2002/016944中报道了一种制备合成的核酸分子的方法，当其在特定宿主细胞中表达时具有减少的不适当或不期望的转录特征。WO2006/042158中报道了经过修饰从而增强重组蛋白表达和/或减少或消除错误剪接和/或内含子通读副产物的核酸分子。

[0008] 因此，存在对于具有减少的副产物的免疫球蛋白的重组生产方法的需要。

[0009] 发明简述

[0010] 本发明第一个方面包括编码IgA或IgG类免疫球蛋白C_H3-结构域的C-末端部分，或编码IgE或IgM类免疫球蛋白CH4-结构域的C-末端部分氨基酸序列的核酸，其中包含在C_H3-或C_H4-结构域的C-末端部分的主要氨基酸序列中所包含的甘氨酸-赖氨酸二肽，由核酸ggaaaa、或核酸ggcaaa、或核酸gggaaa、或核酸ggaaag、或核酸ggcaag、或核酸gggaag编码。

[0011] 在一个实施方案中，根据发明的核酸编码的氨基酸序列选自于SEQ IDNO:1、3、4、5、6、7或8的氨基酸序列。在另一个实施方案中编码甘氨酸-赖氨酸二肽的核酸前面是核

苷酸 g 或 a。在另一个实施方案中编码甘氨酸 - 赖氨酸二肽的核酸是核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa。

[0012] 本发明的第二方面是包含根据发明核酸的质粒，本发明的第三方面是包含根据本发明的核酸的细胞。

[0013] 本发明的另一个方面是具有 SEQ ID NO :17、18、19、20、21、22、23、30、或 31 的核苷酸序列的核酸。

[0014] 本发明的第五方面是在哺乳动物细胞中生产免疫球蛋白的方法，包括下列步骤 :

[0015] a) 用编码免疫球蛋白重链的核酸转染哺乳动物细胞，该免疫球蛋白重链包含 SEQ ID NO :17、18、19、20、21、22、23、30、或 31 的核酸，其编码免疫球蛋白重链的 C- 末端部分，

[0016] b) 在适合免疫球蛋白表达的条件下培养转染的哺乳动物细胞，

[0017] c) 从培养物或细胞再生免疫球蛋白。

[0018] 在一个实施方案中，哺乳动物细胞是 CHO 细胞、BHK 细胞、NSO 细胞、Sp2/0 细胞、COS 细胞、HEK 细胞或**PER.C6®**细胞。优选地哺乳动物细胞是 CHO 细胞、或 BHK 细胞、或**PER.C6®**细胞。在另一个实施方案中用两种核酸转染哺乳动物细胞，其中第一种核酸包含编码免疫球蛋白轻链的表达盒，且其中第二种核酸包含编码免疫球蛋白重链的表达盒，该免疫球蛋白重链包括编码免疫球蛋白重链 C- 末端部分的 SEQ ID NO :17、18、19、20、21、22、23、30、或 31 的核酸。

[0019] 本发明最后的方面是提高哺乳动物细胞中免疫球蛋白表达的方法，其中编码免疫球蛋白重链的核酸包括编码甘氨酸 - 赖氨酸二肽的核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa、或核酸 ggaaag、或核酸酸 ggcaag、或核酸 gggaag，该甘氨酸 - 赖氨酸二肽包含在 C_H3- 或 C_H4- 结构域。

[0020] 发明详述

[0021] 本发明包括编码 IgA 或 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分，或编码 IgE 或 IgM 类免疫球蛋白 C_H4- 结构域的 C- 末端部分氨基酸序列的核酸，其中包含在 C_H3- 或 C_H4- 结构域的 C- 末端部分的氨基酸序列中的甘氨酸 - 赖氨酸二肽，由核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa、或核酸 ggaaag、或核酸 ggcaag、或核酸 gggaag 编码，且其中编码甘氨酸 - 赖氨酸二肽的核酸前面任选为核苷酸 g 或核苷酸 a。

[0022] 用来实施本发明的方法和技术是本领域技术人员众所周知的，并被描述于例如在 Ausubel, F.M., 编, Current Protocols in Molecular Biology, 第一至第三卷 (1997)，和 Sambrook 等, Molecular Cloning :A LaboratoryManual, 第二版, Cold Spring Harbor Laboratory Press, Cold SpringHarbor, N.Y. (1989)。如本领域技术人员众所周知，使用重组 DNA 技术能够生产许多核酸和 / 或多肽的衍生物。这些衍生物可以，例如，在单个或若干位置通过取代、改变、交换、缺失或插入而进行修饰。这种修饰或衍生可以，例如，通过定点诱变的方法实现。本领域技术人员可以很容易地实现这种修饰 (见例如, Sambrook, J., 等, Molecular Cloning :A laboratorymanual(1999)Cold Spring Harbor Laboratory Press, New York, USA)。重组技术的应用，使本领域技术人员能够用异源核酸转化不同宿主细胞。虽然不同细胞的转录和翻译，即表达体系使用相同的元件，但属于不同物种的细胞可能具有与其他生物中不同的所谓的密码子使用。因此相同的多肽 (关于氨基酸序列) 可能由不同的核酸编码。此外，由于遗传密码的简并性，不同的核酸可能编码相同的多肽。

[0023] 本文中所使用的“核酸”，是指由单个核苷酸（又称碱基）a、c、g 和 t（或 u 在 RNA 中）组成的聚合物分子，例如指 DNA、RNA 或其修饰。该多核苷酸分子可以是自然存在的多核苷酸分子或合成的多核苷酸分子、或一个或多个自然存在的多核苷酸分子与一个或多个合成的多核苷酸分子的组合物。这一定义还包含有自然存在的多核苷酸分子，其中的一个或多个核苷酸被改变（如通过诱变）、缺失或添加。核酸可以被分离、或整合到另一核酸，例如在表达盒、质粒、或宿主细胞染色体中。同样，核酸的特征在于其由单个核苷酸组成的核酸序列。

[0024] 对于本领域的技术人员来说，将氨基酸序列，例如一个多肽，转换为相应的编码这一氨基酸序列的核酸序列的程序和方法是众所周知的。因此，核酸特征在于其由单个苷酸组成的核酸序列，还在于由此编码的多肽的氨基酸序列。

[0025] 术语“质粒”包括如穿梭和表达质粒以及转染质粒。在本申请中术语“载体”与“质粒”同义使用。通常，“质粒”还将包括复制起点（例如复制起点 ColE1 或 oriP）和选择标记（例如，氨苄青霉素、卡那霉素、四环素或氯霉素选择标记），分别用于细菌中质粒的复制和筛选。

[0026] “表达盒”是指这样的构建体，其中包含必要的调控元件如启动子和多腺苷酸化位点，用于至少在细胞中包含的核酸的表达。

[0027] “选择标记”是核酸，其在相应的选择试剂存在的情况下，使携带选择标记的细胞被特异性地选择或不被选择。通常情况下，选择标记将赋予对药物的抗性或补偿宿主细胞中新陈代谢和分解代谢的缺陷。筛选标记可是正的、负的、或双功能的。一种有用的正选择标记是抗生素抗性基因。这一选择标记使以其转化的宿主细胞在相应选择试剂，例如选择抗生素存在的情况下被正向选择。在培养物中，未转化宿主细胞在选择性的培养条件下，即在选择试剂存在的情况在不能够生长或生存。正选择标记使得能够选择携带选择标记的细胞，而负选择标记使携带标记的细胞被有选择地消除。用于真核细胞的选择标记包括，例如对氨基糖甙磷酸转移酶 (APH) 的基因，例如，潮霉素 (hyg)、新霉素 (neo)、和 G418 筛选标记、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（选择剂吲哚）、组氨醇脱氢酶（选择剂组氨醇 D)、和赋予抗嘌呤霉素、博莱霉素、腐草霉素、氯霉素、Zeocin、和霉酚酸抗性的核酸。另外的标记基因在，例如 WO 92/08796 和 WO94/28143 中报道。

[0028] 本文中所使用的术语“表达”是指在细胞内发生的转录和 / 或翻译过程。目的核酸序列在细胞中的转录水平可以根据在细胞中存在的相应 mRNA 的数量决定。例如，从目的序列转录的 mRNA 可以通过 RT-PCR 或 Northern 杂交（见 Sambrook 等，1989，supra）定量。由目的核酸编码的多肽可以通过各种方法定量，例如通过 ELISA 法、通过测定多肽的生物活性、或通过采用不依赖这种活性的分析，如蛋白质印迹、或放射性免疫测定法，使用识别和结合于多肽的免疫球蛋白（见 Sambrook 等，1989，同上）。

[0029] 术语“细胞”或“宿主细胞”是指一种细胞，其中例如编码异源多肽的核酸可以或者已引入 / 转染至该细胞。术语“细胞”包括被用于质粒繁殖的原核细胞，和被用于表达核酸的真核细胞。优选地，该真核细胞是哺乳动物细胞。优选地该哺乳动物细胞选自于哺乳动物细胞包括 CHO 细胞（如 CHOK1、CHO DG44)、BHK 细胞、NSO 细胞、SP2/0 细胞、HEK 293 细胞、HEK 293EBNA 细胞、**PER.C6®**细胞和 COS 细胞。本文所使用的术语“细胞”，包括受试

细胞及其后代。因此，术语“转化体”和“转化细胞”包括原代受试细胞和由此得到的培养细胞而无需考虑传代次数。还应该理解，由于有意或无意突变，不是所有后代的 DNA 含量都精确地一致。具有相同的在最初转化细胞中所筛选的功能和生物活性的变体后代被包括在内。

[0030] “多肽”是一种聚合体，其由通过肽键连接的氨基酸组成，无论是自然生成还是人工合成的。少于约 20 个氨基酸残基的多肽，可被称之为“肽”，而由两个或更多的多肽组成的分子，或者包含一个超过 100 个氨基酸残基的多肽的分子，可被称为“蛋白质”。多肽还可以包含非氨基酸成分，如碳水化合物基团、金属离子、或羧酸酯。非氨基酸成分可以通过在其中表达多肽的细胞添加，并可能随细胞类型而有所不同。本文所定义的多肽是就其氨基酸主链结构或编码该氨基酸的核酸而言的。附加物，如碳水化合物基团通常是未指明的，但可能仍然存在。

[0031] 在本申请中使用的术语“氨基酸”指一组羧基 α - 氨基酸，其可以直接或以前体的形式由核酸编码。单独的氨基酸由核酸编码，其中每一个核酸由 3 个核苷酸，所谓的密码子或碱基三联体组成。每种氨基酸由至少一种密码子编码。例如氨基酸甘氨酸可以由四种核酸（密码子）gga、ggc、ggg、和 ggt 中的每一个编码，而氨基酸赖氨酸只能由两种核酸 aaa 和 aag 编码。这种现象被称为“遗传密码简并性”。氨基酸组包括丙氨酸 (3 字母代码 :ala，单字母代码 :A)、精氨酸 (arg，R)、天冬酰胺 (asn，N)、天门冬氨酸 (asp，D)、半胱氨酸 (cys，C)、谷氨酰胺 (gin，Q)、谷氨酸 (glu，E)、甘氨酸 (gly，G)、组氨酸 (his，H)、异亮氨酸 (ile，I)、亮氨酸 (leu，L)、赖氨酸 (lys，K)、蛋氨酸 (met，M)、苯丙氨酸 (phe，F)、脯氨酸 (pro，P)、丝氨酸 (ser，S)、苏氨酸 (thr，T)、色氨酸 (trp，W)、酪氨酸 (tyr，Y)、及缬氨酸 (val，V)。

[0032] 本文中所使用的术语“免疫球蛋白”，是指一种由一个或多个多肽组成的蛋白质，其中该多肽基本上由免疫球蛋白基因编码。这个定义包括变体如突变形式，即具有取代、缺失、和插入一个或多个氨基酸的形式、N- 末端截短的形式、融合的形式、嵌合体的形式、以及人源化的形式。公认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，其源自例如，灵长类动物和啮齿动物。单克隆免疫球蛋白是优选的。免疫球蛋白重多肽链和轻多肽链中的每一个均可包括一个恒定区（通常为羧基末端部分）。

[0033] 本文中所使用的术语“单克隆免疫球蛋白”是指从基本上同种的免疫球蛋白种群中获得的免疫球蛋白，即单独的免疫球蛋白包含的群是一致的，除了可能存在的自然发生的突变，其可能以较少的数量存在。单克隆免疫球蛋白是高度特异的，其针对单一抗原位点。此外，与在其中包括的不同免疫球蛋白针对不同抗原位点（决定簇或表位）的多克隆免疫球蛋白制剂不同，每一个单克隆免疫球蛋白针对抗原上一个单一的抗原位点。除了其特异性，单克隆免疫球蛋白是有利的，因为它们可以被合成而没有其他免疫球蛋白的污染。修饰语“单克隆”表示免疫球蛋白的作为源自于基本上同类种群的免疫球蛋白的特性，而不应解释为需要通过任何特定方法生产免疫球蛋。

[0034] 非人类（如啮齿动物）免疫球蛋白的“人源化”形式是嵌合免疫球蛋白，其含有源于非人类免疫球蛋白和源于人免疫球蛋白的部分序列。对于大部分，人源化的免疫球蛋白具有想要的特异性和亲和力，其是源于人类免疫球蛋白（受体免疫球蛋白），其中源自高变区的残基被源自非人类物种，如小鼠、大鼠、兔、或非人类的灵长类动物的高变区的残基（供体免疫球蛋白）取代（见例如，Morrison，S.L.，等，Proc.Natal.Acad.Sci.USA

81(1984)6851-6855 ;US 5,202,238 ;US 5,204,244)。在一些情况下,人类免疫球蛋白的框架区 (FR) 残基被相应的非人类残基取代。此外,人源化的免疫球蛋白可能包括其它修饰,如未在受体免疫球蛋白中或供体免疫球蛋白中发现的氨基酸残基。该修饰导致这种受体或供体免疫球蛋白的变体,其与相应的亲本序列是同源的,但又不一致。这些修饰是为了进一步改善免疫球蛋白的性能。总的来说,人源化的免疫球蛋白可能基本上包括全部至少一个,并通常两个可变域,其中所有或基本上所有高变环都与非人供体免疫球蛋白相对应,且所有或几乎所有的 FR 均是人受体免疫球蛋白的。该人源化免疫球蛋白任选地还包括至少一部分免疫球蛋白恒定区,通常是人免疫球蛋白的。本领域已经叙述了非人免疫球蛋白人源化的方法。优选地,人源化免疫球蛋白具有一个或多个氨基酸残基,其中该残基是从非人来源被引入人源化免疫球蛋白中。这些非人类氨基酸残基常常被称为“输入 (import)”残基,其通常取自于“输入”可变区。人源化基本上可以按 Winter 和其同事的方法进行,通过用将相应的非人类免疫球蛋白序列取代为高变区序列而实现。因此,这种“人源化”免疫球蛋白是嵌合免疫球蛋白,其中,完整的人源可变区仅有很小部分被源自于非人类物种的相应序列所取代。实际上,人源化的免疫球蛋白通常是其中的某些高变区残基和可能某些框架区残基,被源自啮齿动物和非人类灵长类动物免疫球蛋白的相似位点的残基取代的免疫球蛋白。

[0035] 依据其重链恒定区的氨基酸序列,免疫球蛋白被分类为 :IgA、IgD、IgE、IgG、和 IgM。其中一些可以被进一步分为子类 (同种型) 例如,IgG1、IgG2、IgG3、和 IgG4,IgA1 和 IgA2。根据重链恒定区,不同类的免疫球蛋白的重链分别被标示为 α -、δ -、ε -、γ -、和 μ 链。IgA、IgD 和 IgG 类全长人免疫球蛋白重链恒定区,由恒定区 1(本文标示为 C_H1)、铰链区、恒定区 2(C_H2)、以及恒定区 3(C_H3) 组成。IgE 和 IgM 类人免疫球蛋白包括另外的第四恒定区 (C_H4)。此外,IgM 类免疫球蛋白聚合体包括多种免疫球蛋白,如 5 或 6 种,由二硫键共价相连。这些不同的人免疫球蛋白类的 C- 末端恒定区氨基酸序列被列于表 1 中。SEQ ID NO :01 至 08 的第一个氨基酸可能存在或不存在,因为该氨基酸由两个外显子编码,编码 C- 末端恒定区的外显子和编码前面区域的外显子。

[0036] 表 1 :不同种类免疫球蛋白的 C- 末端氨基酸序列。

[0037] 下划线是甘氨酸 - 赖氨酸二肽。免疫球蛋白类氨基酸序列 C- 末端部分

[0038]

免疫球蛋白种类 (C-末端恒定区)	C-末端部分的氨基酸序 (从 N-末端至 C-末端方向)	SEQ ID NO:
IgA1 (C_H3), IgA2 (C_H3)	GNTFRPEVHL LPPPSEELAL NELVTLTCLA RGFSPKDVLV RWLQGSQELP REKYL TWASR QEPSQGTTTF AVTSILRVAA EDWKKGDTFS CMVGHEALPL AFTQK TIDRL AGKPTHVNVS VVMAEVDGTC Y	01
IgD (C_H3)	AAQAPVKLSLN LLASSDPPEA ASWLLCEVSG FSPPNILLMW LEDQREVNTS GFAPARPPPQ PGSTTFWAWS VLRVPAPPSP QPATYTCVVS HEDSRTLLNA SRSLEVSY	02
IgE (C_H4)	GPRAAPEVYA FATPEWPGSR DKRTLACLIQ NFMPEDISVQ WLHNEVQLPD ARHSTTQPRK TKGSGFFVFS RLEVTRAEWE QKDEFICRAV HEAASPSQTV QRAVSVNPGK	03
IgM (C_H4)	GVALHRPDVY LLPPAREQLN LRESATITCL VTGFSPADVF VQWMQRGQPL SPEKYVTSAP MPEPQAPGRY FAHSILTVSE EEWNTGETYT CVVAHEALPN RVT ERTVDK STGKPTLYNV SLVMSDTAGT CY	04
IgG1 (C_H3)	GQPREPQVYT LPPSRDELTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	05

[0039]

免疫球蛋白种类 (C-末端恒定区)	C-末端部分的氨基酸序 (从 N-末端至 C-末端方向)	SEQ ID NO:
IgG2 (C_H3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPMLDS DGSFFLYSKL TVDKSRWQQG NVFSCSVMHE ALHNHYTQKS LSLSPGK	06
IgG3 (C_H3)	GQPREPQVYT LPPSREEMTK NQVSLTCLVK GFYPSDIAVE WESSGQPENN YNTTPPMLDS DGSFFLYSKL TVDKSRWQQG NIFSCSVMHE ALHNRFTQKS LSLSPGK	07
IgG4 (C_H3)	GQPREPQVYT LPPSQEEMTK NQVSLTCLVK GFYPSDIAVE WESNGQPENN YKTTPPVLDS DGSFFLYSRL TVDKSRWQEG NVFSCSVMHE ALHNHYTQKS LSLSLGK	08

[0040] 术语“IgA 或 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分，或 IgE 或 IgM 类免疫球蛋白 C_H4- 结构域的 C- 末端部分”指免疫球蛋白重链的氨基酸序列，其位于全长或天然存在的免疫球蛋白重链的 C- 末端末尾，即在 C- 末端部分的 C- 末端与免疫球蛋白重链初级氨基酸序列的 C- 末端是一致的。术语“初级氨基酸序列”表示，相应 mRNA 翻译后的免疫球蛋白重链的氨基酸序列。该初级氨基酸序列在 mRNA 翻译后可能会在表达细胞中被进一步修饰，例如由肽酶从初级氨基酸序列裂解一个或多个 C- 末端氨基酸。因此，初级氨基酸序列和分泌的氨基酸序列可能不相同，但可能相差在某些 C- 末端的氨基酸。在一个实施方案中，该 C- 末端部分包括免疫球蛋白重链初级氨基酸序列的至少 100 个 C- 末端氨基酸，或者优选免疫球蛋白重链初级氨基酸序列的至少 50 个 C- 末端氨基酸，或优选免疫球蛋白重链初级氨基酸序列的至少 20 个 C- 末端氨基酸。在一个实施方案中，根据本发明的核酸编码 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分，或 IgE 类免疫球蛋白 C_H4- 结构域的 C- 末端部分的氨基酸序列。在另一实施方案中根据本发明的核酸编码 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分的氨基酸序列。

[0041] 不同人类免疫球蛋白的 C- 末端恒定区氨基酸序列由相应的 DNA 序列进行编码。在基因组中，这些 DNA 序列包含编码（外显子）和非编码（内含子）序列。DNA 转录为前体 mRNA 后，前体 mRNA 也包含这些内含子和外显子序列。在翻译之前，非编码内含子序列在 mRNA 加工期间被移除，其通过将它们剪接出初级 mRNA 转录物以产生成熟的 mRNA。初级 mRNA 的剪接，是通过剪接供体位点与以适当间隔分隔开的剪接受体位点的结合控制的。剪接供体位点位于内含子序列的 5′ 端，剪接受体位点位于内含子序列的 3′ 端，并且两者在前体 mRNA 剪接过程中都被部分地移除。

[0042] 术语“以适当间隔分隔开”表示，核酸中的剪接供体位点与剪接受体位点以这样一种方式排列，以便剪接过程所有必需的元件是可用的并在适当位置上，使剪接过程得以发生。

[0043] 本发明包括编码 IgA 或 IgG 类免疫球蛋白 C_H3- 结构域的 C- 末端部分，或编码 IgE 或 IgM 类免疫球蛋白 C_H4- 结构域的 C- 末端部分的氨基酸序列的核酸，其中包含在所述 C_H3- 或 C_H4- 结构域的 C- 末端部分的氨基酸序列中的甘氨酸 - 赖氨酸二肽，由核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa 编码。

[0044] 令人惊讶地发现，用根据本发明的核酸，可以减少通过不需要的剪接事件形成的不需要的副产物。

[0045] 在本申请中使用的术语“甘氨酸 - 赖氨酸二肽”是指在 N- 末端到 C- 末端的方向上包括由肽键连接的两种氨基酸甘氨酸和赖氨酸的肽。在本申请中使用的术语“甘氨酸 - 赖氨酸二肽”是指一个较大多肽或蛋白质的二肽部分，起可在该较大多肽的起点、内部、或末端发现。氨基酸赖氨酸可以由核酸 aaa 和 aag 编码。因此，本发明的另一个实施方案是，包含在免疫球蛋白重链 CH3- 或 CH4- 结构域的 C- 末端部分的氨基酸序列中的甘氨酸 - 赖氨酸二肽，由核酸 ggaaag、或核酸 ggcaag、或核酸 gggaag 编码。在另一实施方案中，编码甘氨酸 - 赖氨酸二肽的核酸的前面是核苷酸 a 或 g。

[0046] 编码不同人免疫球蛋白类和亚类的 C- 末端恒定区的核酸序列在表 2 中列出。对于人类 IgG1 和 IgG2 的 C_H3 区域，已知两种的变体形式。在一个实施方案中核酸编码免疫球蛋白重链 C- 末端恒定区的一部分。

[0047] 表 2 :编码不同人免疫球蛋白类重链 C- 末端部分的核酸序列。

[0048]

免疫球蛋白种类 (C 末端恒定区)	编码免疫球蛋白重链的氨基酸序列 C-末端部分的核酸序列	SEQ ID NO:
IgA1 (C_H3), IgA2 (C_H3)	ggcttcagcc ccaaggacgt gctggttcgc tggctgcagg ggtcacagga gctgccccgc gagaagtacc tgacttgggc atccccggcag gagccccagcc agggcaccac caccttcgct gtgaccagca tactgcgcgt ggcagccgag gactggaaga agggggacac cttctcctgc atggtgggcc acgaggccct gccgctggcc ttcacacaga agaccatcga ccgcttggcg ggtaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac	09
IgD (C_H3)	taccaccccaa cgtccgtgac tgtcacctgg tacatgggga cacagagcca gccccagaga accttcccctg agatacaaag acggggacagc tactacatga caagcagcca gctctccacc ccctccagc agtggcgcca aggcgagtac aaatgcgtgg tccagcacac cgccagcaag agtaagaagg agatcttccg ctggccaggt aggtcgcacc ggagatcacc cagaagggcc ccccaggacc cccagcacct tccactcagg gcctgaccac aaagacagaa gcaagggctg	10
IgE (C_H4)	tttgcgacgc cggagtggcc ggggagccgg gacaagcgca ccctcgcctg cctgatccag aacttcatgc ctgaggacat ctcggtgcag tggctgcaca acgaggtgca gctcccggac gccccggcaca gcacgacgca gccccgcaag accaaggget ccggcttctt cgtcttcagc cgcctggagg tgaccagggc cgaatgggag cagaaagatg agttcatctg ccgtgcagtc catgaggcag cgagccccctc acagaccgtc cagcgagcgg tgtctgtaaa tccccggtaaa	11
IgM (C_H4)	acgggcttct ctcccgcgga cgtcttcgtg cagtggatgc agaggggggca gccccttgtcc ccggagaagt atgtgaccag cgccccaatg cctgagcccc aggccccagg ccggtacttc gcccacagca tcctgaccgt gtccgaagag gaatggaaca cgggggagac ctacacctgc gtggtggccc atgaggccct gcccaacagg gtcaccgaga ggaccgtgga caagtccacc ggtaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac	12

[0049]

免疫球蛋白种类 (C 末端恒定区)	编码免疫球蛋白重链的氨基酸序列 C-末端部分的核酸序列	SEQ ID NO:
IgG1 (C_H3) 变体 1	gtgtacaccc tgccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	13
IgG1 (C_H3) 变体 2	gggcagcccc gagaaccaca ggtgtacacc ctgccccat cccgggatga gctgaccaag aaccaggtca gcctgacctg cctggtcaaa ggcttctatc ccagcgacat cgccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cgcctcccgt gctggactcc gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac acagaagagc ctctccctgt ctccgggtaa a	28
IgG2 (C_H3) 变体 1	gtgtacaccc tgccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac acctcccatg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa	14
IgG2 (C_H3) 变体 2	gggcagcccc gagaaccaca ggtgtacacc ctgccccat cccgggagga gatgaccaag aaccaggtca gcctgacctg cctggtcaaa ggcttctacc ccagcgacat ctccgtggag tgggagagca atgggcagcc ggagaacaac tacaagacca cacctcccat gctggactcc gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc ctctccctgt ctccgggtaa a	29
IgG3 (C_H3)	gtgtacaccc tgccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcag cgggcagccg gagaacaact acaacaccac gcctcccatg ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg cagcagggga acatcttctc atgctccgtg atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggtaaa	15

[0050]

免疫球蛋白种类 (C 末端恒定区)	编码免疫球蛋白重链的氨基酸序列 C-末端部分的核酸序列	SEQ ID NO:
IgG4 (C_H3)	gtgtacaccc tgccccatc ccaggaggag atgaccaaga accaggtcag cctgacctgc ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac agcaggctaa ccgtggacaa gagcaggtgg caggaggggа atgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggtaaa	16

[0051] 在一个实施方案中根据发明的核酸编码的氨基酸序列选自于SEQ IDNO:1、3、4、5、6、7 或 8 的氨基酸序列。

[0052] 在一个实施方案中，核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白重链 C- 末端恒定区的一部分并选自于 SEQ ID NO:17、18、19、20、21、22、或 23、或 30、或 31 的核酸。在另一个实施方案中核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白重链 C- 末端恒定区的一部分并选自于 SEQ IDNO:17、18、19、22、23、30、或 31 的核酸。在另外的实施方案中核酸编码核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白重链 C- 末端恒定区的一部分并选自于 SEQ ID NO:17、18、19、22 或 23 的核酸。

[0053] 具有 SEQ ID NO:17 至 23 和 30 至 31 的核苷酸序列的核酸也是本发明的一个方面。在一个实施方案中，SEQ ID NO:17、18、19、22、23、30、或 31 的核苷酸序列是本发明的一个方面。在在另一个实施方案中 SEQID NO:17、18、19、22、或 23 的核苷酸序列是本发明的一个方面。

[0054] 在一个实施方案中核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白 C- 末端恒定区，并且包括的核酸选自于 SEQ ID NO:17、18、19、20、21、22、或 23、或 30、或 31 的核酸，其编码该免疫球蛋白重链 C- 末端区域的一部分。在另一个实施方案中核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白 C- 末端恒定区，并且包括的核酸选自于 SEQ ID NO:17、18、19、22、23、30、或 31 的核酸，其编码该免疫球蛋白重链 C- 末端区域的一部分。在另外的实施方案中核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白 C- 末端恒定区，并且包括的核酸选自于 SEQ ID NO:17、18、19、22、或 23 的核酸，其编码该免疫球蛋白重链 C- 末端区域的一部分。

[0055] 表 3 :编码不同类免疫球蛋白重链 C- 末端恒定区氨基酸序列一部分的根据本发明的核酸序列。

[0056]

免疫球蛋白种类	编码 C-末端氨基酸序列的核酸序列	SEQ ID NO:
IgA	ggcaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac	17
IgE	catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggcaaa	18
IgM	ggcaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac	19
IgG1 变体 1	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	20
IgG1 变体 2	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tccgggcaaa	30
IgG2 变体 1	atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa	21
IgG2 变体 2	atctccgtgg agtgggagag caatgggcag ccggagaaca actacaagac cacacctccc atgctggact ccgacggctc cttcttcctc tacagcaagc tcaccgtgga caagagcagg tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg aggctctgca caaccactac acgcagaaga gcctctccct gtctccgggc aaa	31
IgG3	atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggcaaa	22
IgG4	atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggcaaa	23

[0057] 根据发明的核酸编码 IgA、IgE、IgM、或 IgG 类免疫球蛋白 C- 末端恒定区的至少一部分。术语“C- 末端恒定区”是指 IgA 或 IgG 类免疫球蛋白重链的 C_H3- 结构域，或 IgE 或 IgM 类免疫球蛋白重链的 C_H4- 结构域。术语组“一部分”表示 IgA、IgE、IgM、或 IgG 类免疫球蛋白重链 C- 末端恒定区的 C 末端部分，是初级氨基酸序列中从免疫球蛋白重链的 C- 末端至 N- 末端方向计数的至少 20 个连续氨基酸、或至少 50 个连续氨基酸、或至少 100 个连续氨基酸。

[0058] 免疫球蛋白的重组生产是本领域现有技术中众所周知的，并被描述在，例如，综述性文章 Makrides, S. C, Protein Expr. Purif. 17(1999)183-202 ;Geisse, S. ,等, Protein Expr. Purif. 8(1996)271-282 ;Kaufman, RJ. , Mol. Biotechnol. 16(2000)151-160 Werner, R. G. , Arzneimittelforschung-DrugResearch 48(1998)870-880 中。

[0059] 为了生产包括由根据发明的核酸编码的氨基酸序列的免疫球蛋白，本发明进一步包括了用于在哺乳动物细胞中生产免疫球蛋白的方法，其包括下列步骤：

[0060] a) 用编码免疫球蛋白重链的核酸转染哺乳动物细胞，其中选自于 SEQID NO :17、

18、19、20、21、22、或 23、或 30、或 31 的核酸编码免疫球蛋白重链 C- 末端区域的一部分，

[0061] b) 在适合免疫球蛋白表达的条件下培养转染的哺乳动物细胞，

[0062] c) 从培养物或细胞回收免疫球蛋白。

[0063] 术语“在适合免疫球蛋白表达的条件下”表示这样的条件，将其用于培养表达免疫球蛋白的哺乳动物细胞，且其是本领域技术人员已知的或可以很容易地被本领域技术人员确定。本领域技术人员还熟知，这些条件可以依据所培养的哺乳动物细胞类型和所表达的免疫球蛋白类型而改变。一般地，哺乳动物细胞被培养在一温度，例如：在 20℃和 40℃之间，并持续一段足以得到免疫球蛋白的有效蛋白质产生的时间，例如持续 4 至 28 天。

[0064] 在一个实施方案中，是用编码免疫球蛋白重链的核酸转染哺乳动物细胞，其中选自于 SEQ ID NO：17、18、19、22、23、30、或 31 的核酸编码免疫球蛋白重链 C- 末端区域的一部分。在另一个实施方案中，是用编码免疫球蛋白重链的核酸转染哺乳动物细胞，其中选自于 SEQ ID NO：17、18、19、22、或 23 的核酸编码免疫球蛋白重链 C- 末端区域的一部分。在一个实施方案中，用一种核酸转染哺乳动物细胞，其中该核酸包含编码免疫球蛋白重链的表达盒和编码免疫球蛋白轻链的表达盒。在另一个实施方案中，用两种核酸转染哺乳动物细胞，一种包含编码免疫球蛋白轻链的表达盒，一种包括编码免疫球蛋白重链的表达盒。

[0065] 使用根据本发明的方法生产的免疫球蛋白，优选是异源免疫球蛋白。术语“异源免疫球蛋白”是指不是由所述的哺乳动物细胞天然产生的免疫球蛋白。根据本发明的方法产生的免疫球蛋白是通过重组方法生产的。这种方法是在本领域现有技术中众所周知的，并包括在原核和真核细胞中的蛋白质表达及随后对异源免疫球蛋白的回收和分离，通常地净化到可药用的纯度。为了生产，即表达免疫球蛋白，通过标准方法将每一个包含根据本发明的核酸的编码轻链的核酸和编码重链的核酸，插入到表达盒中。使用常规的方法很容易分离和测序编码免疫球蛋白的核酸。杂交瘤细胞可以作为这种核酸的来源。表达盒可以插入到表达质粒中，然后将该质粒转染到宿主细胞中，否则该宿主细胞不产生免疫球蛋白。在适当的原核或真核宿主细胞中进行表达，并从裂解后的细胞中或从培养上清液中回收免疫球蛋白。

[0066] 不同的方法，被充分地建立和广泛地应用于蛋白质的回收和净化，如使用微生物蛋白质的亲和层析（如蛋白质 A 或蛋白 G 亲和层析）、离子交换层析（例如：阳离子交换（羧甲基树脂）、阴离子交换（氨乙基树脂）和混合模式交换）、嗜硫吸附（如用 β- 巯基乙醇和其他巯基配体）、疏水作用或芳香吸附色谱法（例如用苯基琼脂糖、aza-arenophilic 树脂、或 m- 氨基苯硼酸）、金属螯合亲和层析（例如，用镍 (II)- 和铜 (II)- 亲和材料）、体积排阻色谱法、和电泳方法（如，凝胶电泳、毛细管电泳）(Vijayalakshmi, M.A. Appl. Biochem. Biotech. 75(1998)93-102)。

[0067] 本发明进一步包括核酸，其编码包含根据本发明的核酸的免疫球蛋白重链。此外本发明包括质粒和包含该质粒的细胞，其中该质粒包含根据本发明的核酸。

[0068] 本发明的另一个方面是用于改善免疫球蛋白在哺乳动物细胞中的表达的方法，其中编码该免疫球蛋白重链的核酸，包括核酸 ggaaaa、或核酸 ggcaaa、或核酸 gggaaa、或核酸 ggaaag、或核酸 ggcaag、或核酸 gggaag，以编码包含于 C_H3- 或 C_H4- 结构域中的甘氨酸 - 赖氨酸 - 二肽。应用该核酸可以减少或抑制不需要的剪接事件。

[0069] 在本发明的一个实施方案中，该免疫球蛋白重链是 IgG4 亚类人类抗体的免疫球

蛋白重链或者是 IgG1、IgG2、或 IgG3 亚类人类抗体的免疫球蛋白重链。在一个实施方案中，该免疫球蛋白重链是人类免疫球蛋白重链并优选源自于人类 IgG4 亚类或源自人类 IgG1 亚类的突变的免疫球蛋白重链。在另一个实施方案中，免疫球蛋白重链源于人类 IgG1 亚类，具有突变 L234A 和 L235A。在另一实施方案中，免疫球蛋白重链是人类 IgG4 免疫球蛋白重链具有突变 S228P。在一实施方案中，免疫球蛋白重链属于 IgG4 亚类或属于 IgG1 或 IgG2 亚类，在 L234、L235、和 / 或 D265 具有突变，和 / 或包含 PVA236 突变。在其它实施方案中，突变是 S228P、L234A、L235A、L235E、和 / 或 PVA236(PVA236 意思是氨基酸序列 ELLG(用单字母的氨基酸代码给定) 从 IgG1 氨基酸的位点 233 至 236 或 IgG4 的 EFLG 被 PVA 取代)。优选为 IgG4 的突变 S228P，和 IgG1 的 L234A 和 L235A(根据 Kabat 的 EU 指数编号)。

[0070] 本文所使用的“调控元件”，指顺式存在的核苷酸序列，是转录和 / 或翻译编码目的多肽的核酸序列所必需的。转录调控元件通常包括待表达核酸序列上游的启动子、转录起始位点和终止位点、和多聚腺苷酸信号序列。术语“转录起始位点”是指核酸中的核酸碱基，其对应于并入初级转录物，即 mRNA 前体的第一个核酸，转录起始位点可能与启动子序列重叠。术语“转录终止位点”是指通常存在于待转录的目的基因 3′ 端的核苷酸序列，其使得 RNA 聚合酶终止转录。多聚腺苷酸信号序列，或加多聚 A 信号为真核 mRNA 在 3′ 端的特定位点提供了切割信号，在切割的 3′ 端核心提供了约 100-200 腺嘌呤核苷酸序列的转录后附加物 (多聚 A 尾巴)。该多聚腺苷酸信号序列可以在距离切割位点上游约 10-30 个核苷酸的位置包括共有序列 AATAAA。

[0071] 为了产生分泌型多肽，目的核酸包括编码信号序列 / 前导肽的 DNA 片段。信号序列指导新合成的多肽到达并通过可以多肽分泌途经的 ER 膜。在蛋白质穿过 ER 膜期间，信号序列被信号肽酶切割。就信号序列的功能来说，被宿主细胞的分泌机制识别是非常重要的。因此，使用的信号序列，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) 和终止密码子 (TAA、TAG 或 TGA)。内部核糖体进入位点 (IRES) 可以被包含在某些构建体中。

[0073] “启动子”是指控制基因 / 结构基因或核酸序列转录的多核苷酸序列，其中该多苷酸序列被有效连接于该基因 / 结构基因或核酸序列。启动子包括用于 RNA 聚合酶结合和转录起始的信号。使用的启动子将在宿主细胞的细胞类型中起作用，在其中预期表达选定的序列。包括源于各种不同来源的组成型、诱导型和阻抑型启动子的大量启动子，是本领域内众所周知的 (并在数据库如 GenBank 中鉴定)，并可以用作或处于克隆的多核苷酸内 (从，例如保藏处如 ATCC 以及其他商业或个人来源) 获得。“启动子”包括指导结构基因转录的核苷酸序列。通常情况下，启动位于基因的 5′ 非编码或非翻译区，邻近结构基因转录起始位点。启动子内在转录起始中起作用的序列元件往往以共有的核苷酸序列为特征。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, Mol.Endocrinol.7(1993)551-560)、环 AMP 反应元件 (CRE)、血清反应元件 (SRE ;Treisman,R., Seminars in Cancer Biol.1(1990)47-58)、糖皮质激素反应元件 (GRE)、和其他转录因子的结合位点，如 CRE/ATF(Treisman, R., Seminars in Cancer Biol.1(1990)47-58 ;O′ Reilly, M.A., 等, J.Biol.Chem.267(1992)19938)、AP2(Ye, J., 等, J.Biol.Chem.269(1994)25728)、SP1、cAMP 反应元件结合蛋白 (CREB ;Loeken, M.R., Gene Expr.3(1993)253) 和八聚体因子 (一般而言见 Watson 等，编，Molecular Biology

of the Gene，第四版.(The Benjamin/CummingsPublishing 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，那么转录率响应诱导因子而增加。相比之下，如果启动子是组成型启动子，转录率不受诱导因子的调控。阻抑型启动子也是熟知的。例如，当生长激素在细胞表面与其受体结合后c-fos启动子特异地活化。四环素(tet)控制的表达可以通过人工杂交启动子而获得，其中人工杂交启动子由例如CMV启动子接着两个Tet操纵基因位点组成。Tet阻抑物结合到两个Tet操纵基因位点并阻断转录。当添加诱导物四环素后，Tet阻抑物从Tet操纵基因位点被释放，转录继续(Gossen，M. 和Bujard，H.PNAS 89(1992)5547-5551)。对于其他诱导型启动子，包括金属硫蛋白和热激启动子见例如Sambrook等（同上）和Gossen，M.，Curr.Opin.Biotech.5(1994)516-520。在真核启动子中已被鉴定为用于高水平表达的强启动子是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1α(CHEF-I，见例如美国5,888,809)、人EF-Iα、泛素、和人类巨细胞病毒立即早期启动子(CMV IE)。

[0074] “启动子”可以是组成型或诱导型。增强子（即顺式作用DNA元件，作用于启动子以增强转录）可能有必要与启动子共同起作用，以增强了单独用启动子获得的表达水平，并可能作为转录调控元件而被包括。通常，含有启动子的多核苷酸片段也会包含增强子序列（如CMV或SV40)。

[0075] 本文所使用的“增强子”是指增强基因或编码序列转录的多核苷酸序列，其中该多苷酸序列被有效连接于该基因或编码序列。与启动子不同，增强子相对而言不依赖于方向和位置，并已在对于转录单元而言的5′或3′(Lusky，M.，等，MoI.Cell Bio.，3(1983)1108-1122)、内含子(Banerji，J.，等，Cell,33(1983)729-740)中、以及编码序列本身(Osborne,T.F.，等,MoI.Cell Bio.,4(1984)1293-1305)中发现。因此，增强子可能被置于转录起始位点的上游或下游或距启动子相当远，但是实际上增强子可能物理上和功能上与启动子重叠。源于各种不同来源的大量增强子，是本领域内众所周知的（并在数据库如GenBank中鉴定），并可用作或处于克隆的多核苷酸序列内（从，例如，保藏处如ATCC以及其他商业或个人来源）获得。许多包含启动子序列（如常用的CMV启动子）的多核苷酸也包含增强子序列。例如，所有上面列出的强启动子也可包含强增强子（见，例如Bendig，M.M.,Genetic Engineering,7(Academic Press,1988)91-127)。

[0076] “内部核糖体进入位点”或“IRES”描述这样的序列，其功能上促进独立于IRES 5′基因的翻译起始，并且允许动物细胞内从单个转录物翻译两个顺反子（开放阅读框）。该IRES提供了独立的核糖体进入位，以翻译在其紧邻下游（本文下游与3′是互换使用的）的开放式阅读框。动物细胞mRNA不同于细菌mRNA，细菌mRNA可以是多顺反子，即编码几种不同的从mRNA顺序翻译的多肽，大多数动物细胞mRNA是单顺反子，只编码一种蛋白质的合成。用真核细胞中的单顺反子转录物，翻译将从最5′端的翻译起始位点开始，在第一个终止密码子结束，而且转录物将从核糖体被释放，导致只有mRNA中第一个编码的多肽得到翻译。在真核细胞中，多顺反子转录物具有IRES有效地连接于转录物中第二个或其后的开放阅读框，使得开放阅读框下游的相继翻译，以产生两个或更多的由同一转录物编码的多肽。先前已经描述了IRES元件在载体构建中的使用，见，例如，Pelletier，J.，等，Nature 334(1988)320-325；Jang，S.K.，等，J.Virol.63(1989)1651-1660；Davies，

M.V.， 等，J.Virol.66(1992)1924-1932；Adam，M.A.， 等 J.Virol.65(1991)4985-4990；Morgan，R.A.， 等 Nucl.AcidsRes.20(1992)1293-1299；Sugimoto，Y， 等 Biotechnology 12(1994)694-698；Ramesh，N.， 等 Nucl.Acids Res.24(1996)2697-2700；andMosser，D.D. 等，BioTechniques 22(1997)150-161)。

[0077] “有效连接”是指两个或多个组件的并置，其中这样描述的组件之间的关系，允许它们以其预期的方式起作用。例如，如果启动子和/或增强子顺式作用以控制或调节连接的序列转录，则它被有效连接于编码序列。通常，但非必需地，“有效连接”的DNA序列是邻接的，并在有必要时连接两个蛋白质编码区（如分泌型前导区和多肽），其为邻接的并符合读框。然而，尽管有效连接的启动子通常位于编码序列的上游，但其并不一定与之相连。增强子不必是相连的。如果增强子增强了编码序列的转录，则该增强子是有效连接到编码序列的。有效连接的增强子可以位于编码序列的上游、其中或下游，并距离启动子相当远处。多腺苷酸化位点被有效连接到编码序列，如果其位于编码序列下游端，这样转录经过编码序列进行到多腺苷酸化序列中。翻译终止密码子被有效连接到外显子核酸序列如果其位于编码序列的下游端（3′末端），这样翻译通过编码序列进行至终止密码子并在那里终止。连接是通过行业内熟知的重组的方法实现的，例如，使用PCR方法和/或在方便的限制性位点的连接。如果不存在方便的限制性位点，那么使用与传统方法一致的合成的寡核苷酸连接物或接头。

[0078] “异源DNA”或“异源多肽”是指DNA分子或多肽，或一群DNA分子或一群多肽，其不是自然地存在于给定的宿主细胞中。特定宿主细胞的异源DNA分子可能包含源于宿主细胞物种的DNA(即内源性DNA)，只要宿主DNA结合于非宿主DNA(即外源DNA)。例如，含有编码多肽的非宿主DNA片段的DNA分子，其中该DNA片段有效连接于包含启动子的宿主DNA片段，被认为是异源DNA分子。相反地，异源DNA分子可以包含与外源启动子有效连接的内源结构基因。

[0079] 由非宿主DNA分子编码的肽或多肽是“异源”肽或多肽。

[0080] “表达质粒”是核酸分子，其编码将在宿主细胞中被表达的蛋白质。通常，表达质粒包含原核质粒繁殖单元，如对于大肠杆菌，其中包含复制起点和选择标记、真核选择标记、以及一个或多个用于目的结构基因表达的表达盒，每个包括启动子、结构基因、和包括多腺苷酸化信号的转录终止子。基因表达通常处于启动子的控制之下，这样的结构基因被认为是“有效连接于”该启动子。同样，如果调控元件调节核心启动子的活性，则调控元件和核心启动子是有效连接的。

[0081] “分离的多肽”是一种基本上没有细胞成分污染的多肽，污染的细胞成分如碳水化合物、脂、或其他与多肽天然有关的蛋白质的杂质。通常，分离的多肽制剂包含高度纯化形式的多肽，即至少约80％纯的、至少约90％纯的、至少约95％纯的、大于95％纯的、或高于99％纯的。表明一种特定的蛋白质制剂包含分离的多肽的一种方法是，通过该蛋白质制剂的十二烷基硫酸钠(SDS)-聚丙烯酰胺凝胶电泳和凝胶的考马斯亮蓝染色后，出现单一的条带。但是，术语“分离的”不排除存在相同多肽的其他物理形式，如二聚体或可选择的糖基化或衍生形式的存在。

[0082] 本申请中指出的“转录终止子”是一个长50-750个碱基对的DNA序列，其向RNA聚合酶发出mRNA合成的终止信号。在表达盒3′末端的非常高效的（强）终止子适于防

止 RNA 聚合酶通读，特别是在使用强启动子的时候。低效转录终止子可导致形成操纵子样的 mRNA，其可以成为不希望的（如质粒编码的）基因表达的原因。

[0083] 提供下面的实施例、序列表和图以辅助理解本发明，本发明真正的范围在所附的权利要求中阐述。应当理解可以对所述的程序做出修改而不背离本发明的精神。

[0084] 序列描述

[0085] SEQ ID NO:01 人 IgA 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0086] SEQ ID NO:02 人 IgD 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0087] SEQ ID NO:03 人 IgE 类免疫球蛋白重链恒定区 (C_H4) 氨基酸序列的 C- 末端部分

[0088] SEQ ID NO:04 人 IgM 类免疫球蛋白重链恒定区 (C_H4) 氨基酸序列的 C- 末端部分

[0089] SEQ ID NO:05 人 IgG1 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0090] SEQ ID NO:06 人 IgG2 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0091] SEQ ID NO:07 人 IgG3 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0092] SEQ ID NO:08 人 IgG4 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列的 C- 末端部分

[0093] SEQ ID NO:09 编码人 IgA 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列

[0094] SEQ ID NO:10 编码人 IgD 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列

[0095] SEQ ID NO:11 编码人 IgE 类免疫球蛋白重链恒定区 (C_H4) 氨基酸序列 C- 末端部分的核酸序列

[0096] SEQ ID NO:12 编码人 IgM 类免疫球蛋白重链恒定区 (C_H4) 氨基酸序列 C- 末端部分的核酸序列

[0097] SEQ ID NO:13、28 编码人 IgG1 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列（变体 1 和 2）

[0098] SEQ ID NO:14、29 编码 IgG2 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列（变体 1 和 2）

[0099] SEQ ID NO:15 编码 IgG3 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列

[0100] SEQ ID NO:16 编码人 IgG4 类免疫球蛋白重链恒定区 (C_H3) 氨基酸序列 C- 末端部分的核酸序列

[0101] SEQ ID NO:17 根据本发明的核酸序列，其编码 IgA 类免疫球蛋白 C- 末端恒定区 (C_H3) 的氨基酸序列的一部分

[0102] SEQ ID NO:18 根据本发明的核酸序列，其编码 IgE 类免疫球蛋白 C- 末端恒定区 (C_H4) 的氨基酸序列的一部分

[0103] SEQ ID NO:19 根据本发明的核酸序列，其编码 IgM 类免疫球蛋白 C- 末端恒定区 (C_H4) 的氨基酸序列的一部分

[0104] SEQ ID NO:20、30 根据本发明的核酸序列，其编码 IgG1 类免疫球蛋白 C- 末端恒定区 (C_H3) 的氨基酸序列的一部分（变体 1 和 2）

[0105] SEQ ID NO:21、31 根据本发明的核酸序列，其编码 IgG2 类免疫球蛋白 C- 末端恒定区 (C_H3) 的氨基酸序列的一部分（变体 1 和 2）

[0106] SEQ ID NO:22 根据本发明的核酸序列,其编码 IgG3 类免疫球蛋白 C- 末端恒定区 (C_H3) 的氨基酸序列的一部分

[0107] SEQ ID NO:23 根据本发明的核酸序列,其编码 IgG4 类免疫球蛋白 C- 末端恒定区 (C_H3) 的氨基酸序列的一部分

[0108] SEQ ID NO:24 至 27 实施例中使用的核酸引物。

附图说明

[0109] 图 1 注释了的 p4831 的质粒图谱。

[0110] 图 2 克隆 #23 分泌抗体的 SDS 聚丙烯酰胺凝胶电泳 (SDS-PAGE) 和蛋白质印迹分析 ;a) 考马斯染色, b) 与过氧化物酶偶联的抗人免疫球蛋白 γ 链抗体蛋白质印迹分析, c) 与过氧化物酶偶联的抗人免疫球蛋白 κ 轻链抗体蛋白质印迹分析。泳道 1 :人抗 IGF-IR- 参照抗体 ;泳道 2 :包含克隆 #23 抗体的 CHO-DG44 克隆 #23 的培养上清液。

[0111] 图 3 注释了的 p4817 的质粒图谱。

[0112] 图 4 用 p4831 或 4855 转染的 CHO-DXB11 分泌的免疫球蛋白的蛋白质印迹分析。泳道 1 :用 p4831 转染的 CHO 细胞,泳道 2 :用 p4855 转染的 CHO 细胞。

[0113] 图 5 注释的 p5031 的质粒图谱。

[0114] 图 6HEK-293-EBNA 细胞转染质粒 p5031 或 p5032 后分泌的抗体的 SDS-PAGE 分析 ;泳道 1+2 :p5031,泳道 3+4 :p5032,泳道 5-8 :人抗 IGF-IR- 抗体作为参照抗体 ;5 :0.2 微克, 6 :0.7 微克,7 :2 微克,8 :6 微克。

[0115] 图 7 用 p5032 转染和用甲氨蝶呤 (MTX) 选择稳定整合的质粒的 CHO-DG44 细胞。对源自 6 个克隆的抗体进行纯化和用 SDS-PAGE 和考马斯染色分析。

[0116] 材料与方法

[0117] 重组 DNA 技术

[0118] 用如 Sambrook 等,1989(同上) 中所叙述的标准方法来操作 DNA。所有分子生物学试剂为商业购得 (如果不是则另外指出),并按照制造商的说明使用。

[0119] 核酸序列测定

[0120] 利用 MediGenomix GmbH(Martinsried, Germany) 进行的双链测序测定 DNA 序列。

[0121] DNA 和蛋白质序列分析和序列数据管理

[0122] 将 GCG(Genetics Computer Group,Madison,Wisconsin,USA) 的软件包版本 10.2 和 Infomax' s Vector NTI Advancesuite 版本 8.0 用于序列生成、绘图、分析、注释和说明。

[0123] 细胞培养技术

[0124] 将 CHO-DXBII 细胞生长在具有 10% FCS(胎牛血清获得自 Hyclone,Thermo Fisher Scientific Inc.,目录号 :SH3007.03) 的 MEM α 培养基 (Invitrogen Corp.,**Gibco®**,目录号 :22571 中)。

[0125] HEK-293-EBNA 细胞 (ATCC # CRL-10852) 培养在 DMEM 中,添加 2 毫摩谷氨酰胺 (**Gibco®**,目录号 :25030)、1% (v/v)MEM 非必需氨基酸 (**Gibco®**,目录号 :11140)、10% (v/v) 超低 IgG FCS(**Gibco®**,目录号 :16250) 和 250 微克 / 毫升 G418(Roche Applied

Sciences, RocheDiagnostics GmbH, Germany, 目录号 :1464981)。

[0126] 用于 CHO-DG44 细胞培养的培养基是 MEM α 培养基 (**Gibco®**, 目录号 :22561) 添加 10 % (v/v) 透析的 FCS(**Gibco®**, 目录号 :26400) 和 2 % (v/v)HT 补剂 (**Gibco®**, 目录号 :41065)。为了选择稳定转染的 CHO DG44 细胞系, 省略 HT 补剂, 并单独加入或与 400 μ g/ml 潮霉素 B(RocheDiagnostics GmbH, Roche Applied Sciences, Germany, 目录号 : 10843555001) 组合加入 20 至 500nM 甲氨喋呤 (MTX)。

[0127] 所有的细胞系均保持在 37℃和 5 % CO_2 的加湿孵化器中。利用 FuGENE 6(Roche Diagnostics GmbH, Roche Applied Sciences, Germany, 目录号 :1815075) 通过核转染 (Amaxa GmbH, Germany) 或者通过脂质转染法进行细胞转染。

[0128] 此外应用标准细胞培养技术, 如描述于例如 Bonifacino, J.S., 等 (编)(2000) Current Protocols in Cell Biology, John Wiley and Sons, Inc. 中。

[0129] 蛋白 A- 沉淀, SDS-PAGE 和蛋白质印迹

[0130] 用蛋白质 A- 琼脂糖珠沉淀源于细胞培养上清液的免疫球蛋白, 然后通过 SDS/ 聚丙烯酰胺凝胶电泳 (十二烷基硫酸钠, SDS-PAGE) 和蛋白质印迹分析。

[0131] 为了沉淀免疫球蛋白, 含有达 7 微克免疫球蛋白的细胞培养上清用 TBS 缓冲液 (50mM TRIS/HCl, pH 7.5, 补加 150mM NaCl),1 % (v/v) 乙基苯基聚乙二醇 (Nonidet-P40) (Roche Diagnostics GmbH, RocheApplied Sciences, 目录号 :1754599) 稀释至终体积 1 毫升。然后与 15 μ l 湿体积蛋白质 A- 琼脂糖珠孵育 1 小时。通过离心和用具有 1 % (v/v) NonidetP-40 的 TBS 洗涤珠, 此后用 2 倍浓度的 PBS(磷酸盐缓冲盐溶液), 最后以 100mM 柠檬酸钠缓冲液, pH 值 5 洗涤回收珠。最后一步清洗后, 上清液被完全从珠移除。用 20 μ l 含有 50mM DTT(二硫苏糖醇) 的 2 倍浓度的 LDS(十二烷基硫酸锂) 样本缓冲液 (Invitrogen 公司) 洗脱结合的免疫球蛋白。在 95℃孵育 5 分钟后, 离心悬液, 回收上清液作进一步分析。

[0132] SDS-PAGE :根据制造商的建议使用 **NuPAGE®** 凝胶系统 (Invitrogen 公司) 进行 SDS-PAGE。将样本上样到 10 % **NuPAGE®** Novex Bis/TRIS 凝胶 (Invitrogen 公司, 目录号 :NP0301) 上, 在还原 **NuPAGE®** MES SDS(4- 吗啉乙磺酸 / 十二烷基硫酸钠) 电泳缓冲液中分离蛋白质。通常每泳道上样 2 至 3 微克免疫球蛋白用 Simply Blue Safe **Stain®** (Invitrogen 公司) 用于考马斯亮蓝染色和 0.4 至 0.6 微克用于蛋白质印迹。

[0133] 蛋白质印迹 :为了从 SDS/ 聚丙烯酰胺凝胶电转移蛋白质, 使用标准 PVDF(聚偏二氟乙烯) 或硝酸纤维素膜。电转移后在 TBS(tris 缓冲盐溶液, 50mM TRIS/HCl, pH 7.5, 150mM NaCl) 中洗涤膜。通过在具有 1 % (w/v)Western 阻断剂 (Roche Diagnostics GmbH, Roche AppliedSciences, 目录号 :11921673001) 的 TBS 中孵育封闭非特异性结合位点。用在具有 0.5 % (w/v)Western 阻断剂的 TBS 中稀释的过氧化物酶偶联多克隆检测抗体 (见下面的段落给出进一步详细资料) 检测人免疫球蛋白 γ 重链和 κ 轻链。经过三次使用具有 0.05 % (v/v) 吐**温®** 20 的 TBS 的洗涤步骤, 和一次使用 TBS 的洗涤步骤后, 使用 LumiLightPlus 底物溶液 (RocheDiagnostics GmbH, Roche Applied Sciences, 目录号 :12015196001) 和 LUMI-Imager F1 分析仪 (Roche Diagnostics GmbH, Roche AppliedSciences) 检测结合的过氧化物酶偶联的检测抗体。

[0134] 同时或单独地检测人免疫球蛋白 γ 重链 (H) 和 κ 轻链 (L)。对于同时检测，使用以 1 ： 2500(v/v) 稀释的过氧化物酶偶联的山羊抗人 IgG(H+L)- 抗体 (Jackson ImmunoResearch Laboratories 公司，Cat.No. ：109-035-088)。对于连续检测，首先使用过氧化物酶偶联的兔抗人 Igγ - 抗体 (DAKO GmbH，Germany，code no.P0214) 在 1 ： 1000(v/v) 稀释下或使用过氧化物酶偶联的 $F(ab')_2$ 山羊抗人 Fcγ - 抗体 (JacksonImmunoResearch Laboratories，目录号 ：109-036-008) 在 1 ： 7500(v/v) 稀释下用探测膜。在检测免疫球蛋白 γ 重链后，在 62.5mM TRIS/HCl，pH 6.7，辅以 2% (w/v)SDS 和 100mM β - 巯基乙醇中在 50℃剥离膜 30 分钟。对于第二个检测，用过氧化物酶偶联的兔抗人 Igκ - 抗体 (DAKO GmbH，codeno.P0129) 在 1 ： 1000(v/v) 稀释下再探测膜。

[0135] 实施例 1

[0136] IgG1 类免疫球蛋白表达质粒的制备

[0137] 质粒 4831(以下记为 p4831) 是表达质粒，用于在真核细胞中表达抗 -IGF-1R- 抗体 (具有保留的外显子 - 内含子组织的基因组水平上有组织的表达盒)(对于序列见例如 US 2005/0008642 或者 EP 1646720)。它包括以下功能元件 ：

[0138] - 源于载体 pUC18 的复制起点 (pUC 起点)，

[0139] - β (beta)- 内酰胺酶基因，其在大肠杆菌中提供氨苄青霉素抗性 (Amp)，

[0140] - 用于表达 γ 1 重链的表达盒，包括下列元件

[0141] - 源于人巨细胞病毒的主要即刻早期启动子和增强子 (hCMV IE1 启动子)，

[0142] - 包括 Kozak 序列的合成的 5′ UTR，

[0143] - 包括信号序列内含子的鼠免疫球蛋白重链信号序列 (L1_ 内含子 _L2)，

[0144] - 重链可变区 (VH) 的 cDNA，在 3′ 末端安排有剪接供体位点

[0145] - 鼠免疫球蛋白 μ - 增强子区，

[0146] - 人免疫球蛋白重链 γ 1 基因 (IGHG1)，包含外显子 CH1、铰链区、CH2 和 CH3、间插内含子、以及 3′ UTR，其携带多聚腺苷酸化信号序列并任选地含有突变，

[0147] - 用于表达 κ 轻链的表达盒，包括下列元件

[0148] - 源于人巨细胞病毒的主要即刻早期启动子和增强子 (hCMV IE1 启动子)，

[0149] - 包括 Kozak 序列的合成的 5′ UTR，

[0150] - 包括信号序列内含子的鼠免疫球蛋白重链信号序列 (L1_ 内含子 _L2)，

[0151] - 轻链可变区的 cDNA，在 3′ 末端安排有剪接供体位点 (VL)

[0152] - 内含子鼠 Ig-κ 增强子区，

[0153] - 人免疫球蛋白 κ 基因 (IGK) 包含 IGKC 外显子和携带多聚腺苷酸化信号序列的 IGK 3′ UTR。

[0154] - 适合在真核细胞中筛选的潮霉素 B 磷酸转移酶转录单元，其包括

[0155] -SV40 早期启动子和起点，

[0156] - 潮霉素 B 磷酸转移酶的编码序列 (HygB)，

[0157] -SV40 早期多聚腺苷酸化信号

[0158] - 适合在真核细胞中营养缺陷型筛选的用于表达鼠二氢叶酸还原酶 (DHFR) 的表达盒包括

[0159] -SV40 早期启动子和起点的缩短形式，

[0160] - 鼠 DHFR 的编码序列，

[0161] -SV40 早期多聚腺苷酸化信号。

[0162] 图 1 中显示了附加注释的 p4831 质粒图谱。

[0163] 将 p4831 转染至 CHO-DG44 细胞，用潮霉素 B 和甲氨喋呤 (MTX) 筛选后分离稳定的细胞系。用蛋白质 A- 琼脂糖珠沉淀并用 SDS-PAGE 和考马斯染色 (图 2a)) 分析筛选的克隆 #23 分泌的抗体。除了预期的 50kDa 的免疫球蛋白 γ-1 重链和 25kDa 的免疫球蛋白 κ 轻链，还检测到相当数量的 80kDa 副产品蛋白质。这种蛋白质由抗人免疫球蛋白 γ 链抗体 (图 2b)) 以及抗人免疫球蛋白 κ 链抗体 (图 2c)) 识别。

[0164] 实施例 2

[0165] 制备具有修饰的 CH3- 结构域的 IgG1 类免疫球蛋白的表达质粒

[0166] 为了防止产生源自异常剪接的 γ1 前体 mRNA 的副产品，p4831 的 CH3 外显子内部的剪接位点通过 4573 位置的 T 突变成 C 而破坏。同时 T4567 被 C 取代，以除去 BsmA I 限制位点。

[0167] P4855 被构建如下。p4817 是 p4831 的祖先质粒，具有相同的 γ1 重链转录单元，p4817 由以下元件组成 :

[0168] - 源自载体 pUC18 的复制起点，其让这个质粒得以在大肠杆菌中复制 (pUC Ori)

[0169] -β- 内酰胺酶基因，其赋予在大肠杆菌中的氨苄青霉素抗性 (Amp)

[0170] - 用于表达 γ1 重链的表达盒，其包括下列元件

[0171] - 源自人巨细胞病毒的主要即刻早期启动子和增强子 (hCMV IE1 启动子)，

[0172] - 包括 Kozak 序列的合成的 5′ UTR，

[0173] - 包括信号序列内含子的鼠免疫球蛋白重链信号序列 (L1_ 内含子 _L2)，

[0174] - 重链可变区 (VH) 的 cDNA，在 3′ 末端安排有剪接供体位点

[0175] - 鼠免疫球蛋白 μ- 增强子区，

[0176] - 人免疫球蛋白重链 γ1 基因 (IGHG1)，包含外显子 CH1、铰链区、CH2 和 CH3、间插内含子、以及 3′ UTR，其携带多聚腺苷酸化信号序列并任选地含有突变，

[0177] p4817 质粒图谱如图 3 所示。

[0178] 操纵 P4817 是通过定点突变使用 QuickChange 定点诱变试剂盒 (Stratagene，目录号 :200518) 和序列特异性寡核苷酸 1

[0179] agcctctccc tgtccccggg caaatgagtg cgacggccg SEQ ID NO :24

[0180] 和 2

[0181] cggccgtcgc actcatttgc ccggggacag ggagaggct SEQ ID NO :25。

[0182] 切除突变的 p4817 的 839bp 的 SfiI/SgrAI 片段，并与 p4831 的 13133bp 的 SgrAI/SfiI 片段的连接形成 p4855。

[0183] 实施例 3

[0184] 根据实施例 1 和 2 的核酸的表达，分离所生产的免疫球蛋白，以及分析产生的免疫球蛋白

[0185] 质粒 p4831 及 p4855 被瞬时转染至 CHO-DXB11 细胞。将细胞培养在非选择性的条件下。培养 3 天后收获细胞培养上清液，使用蛋白质 A- 琼脂糖珠纯化分泌的免疫球蛋白。用抗人 IgG 抗体 (H+L) 对免疫球蛋白的蛋白质印迹分析表明，经 p4831 转染的细胞表达了

80kDa 的副产物，但经过 p4855 转染的细胞没有表达（图 4）。

[0186] 实施例 4

[0187] IgG4 类免疫球蛋白的表达质粒的制备

[0188] 设计质粒 5031 用于在真核的组织培养细胞中瞬时和稳定表达人抗 -P- 选择蛋白 - 抗体。对于示例性抗 -P- 选择蛋白 - 抗体见例如 EP 1737891 或 US2005/0226876。P5031 是由下列元件组成

[0189] - 源自载体 pUC18 的复制起点，其让该质粒得以在大肠杆菌中复制 (pUC 起点)

[0190] - β - 内酰胺酶基因，其赋予在大肠杆菌中的氨苄青霉素抗性 (Amp)，

[0191] - 用于表达人 γ4 重链的表达盒，其包括下列元件 ：

[0192] - 源自人巨细胞病毒的主要即刻早期启动子和增强子 (hCMV IE1 启动子)，

[0193] - 包括 Kozak 序列的合成的 5′ UTR，

[0194] - 包括信号序列内含子的鼠免疫球蛋白重链信号序列 (L1_ 内含子 _L2)，

[0195] - 重链可变区 (VH) 的 cDNA，在 3′ 末端安排有剪接供体位点

[0196] - 鼠 Igμ - 增强子区，

[0197] - 人免疫球蛋白重链 γ4 基因 (IGHG4)，包含外显子 CH1、铰链区、CH2 和 CH3、间插内含子、以及 3′ UTR，其携带多聚腺苷酸化信号序列并任选地含有突变，

[0198] - 用于表达人 κ 轻链的表达盒，包含下列元件

[0199] - 源自人巨细胞病毒的主要即刻早期启动子和增强子 (hCMV IE1 启动子)，

[0200] - 包括 Kozak 序列的合成的 5′ UTR，

[0201] - 包括信号序列内含子的鼠免疫球蛋白重链信号序列 (L1_ 内含子 _L2)，

[0202] - 轻链可变区的 cDNA，在 3′ 末端安排有剪接供体位点 (VL)

[0203] - 内含子鼠 Ig-κ 增强子区，

[0204] - 人免疫球蛋白 κ 基因 (IGK)，包含 IGKC 外显子和携带多聚腺苷酸化信号序列的 IGK 3′ UTR，

[0205] - 适合在真核细胞中筛选的潮霉素 B 磷酸转移酶转录单元，其包括

[0206] -SV40 早期启动子和起点，

[0207] - 潮霉素 B 磷酸转移酶的编码序列 (HygB)，

[0208] -SV40 早期多聚腺苷酸化信号

[0209] - 适合在真核细胞中进行营养缺陷型筛选的用于表达鼠二氢叶酸还原酶 (DHFR) 的表达盒，包括

[0210] -SV40 早期启动子和起点的缩短形式，

[0211] - 鼠 DHFR 的编码序列，

[0212] -SV40 早期多聚腺苷酸化信号。

[0213] 图 5 中显示了 p5031 的质粒图谱。

[0214] 当用质粒 p5031 转染 HEK-293-EBNA 细胞，该细胞产生的免疫球蛋白，包含 80kDa 的副产品（图 6，泳道 1 和 2）。该蛋白质在蛋白质印迹分析中与抗人 Igγ - 抗体和抗人 Igκ - 抗体结合。

[0215] 实施例 5

[0216] 制备具有修饰的 CH3- 结构域的 IgG4 类免疫球蛋白的表达质粒

[0217] 根据实施例 3 引入修饰。简言之，T4565 被突变为 C，连同第二个核苷酸 T4559，也被换为 C 以便移除 BsmAI 限制位点。寡核苷酸 3

[0218] gcctctccct gtccctgggc aaatgagtgc cagg SEQ ID NO :26

[0219] 和寡核苷酸 4

[0220] cctggcactc atttgcccag ggacagggag aggc SEQ ID NO :27

[0221] 用于定点诱变。所获得的质粒命名为 p5032。

[0222] 实施例 6

[0223] 根据实施例 5 的核酸的表达，分离所生产的免疫球蛋白，以及分析产生的免疫球蛋白

[0224] 质粒 p5031 及 p5032 被瞬时转染至 HEK-293-EBNA 细胞。将细胞培养在非选择性的条件下。培养 3 天后收获细胞培养上清液，使用蛋白质 A- 琼脂糖珠纯化分泌的免疫球蛋白。用抗人 IgG 抗体 (H+L) 对免疫球蛋白的蛋白质印迹分析表明，用 p5031 转染的细胞表达了 80kDa 的副产物 (图 6，泳道 1+2)，但是用 p5032 转染的细胞没有表达副产物 (图 6，泳道 3+4)。

[0225] 质粒 p5032 被转染至 HEK-293-EBNA 细胞，没有表达 80kDa 的蛋白质 (图 6，泳道 3 和 4)。这清楚地表明，80kDa 的融合蛋白质是异常的前体 mRNA 剪接的结果，且这种瞬时表达期间的不期望的蛋白质的产生，可被免疫球蛋白重链 γ 4 基因 (IGHG4) 内部的 CH3 剪接位点的突变有效地抑制。

[0226] IGHG4 内部的 C_H3 剪接位点的突变，也阻止了稳定细胞系中的 80kDa 的表达。用 p5032 转染 CHO-DG44 细胞，用甲氨蝶呤 (MTX) 筛选质粒的稳定整合。纯化并通过 SDS-PAGE 和考马斯染色 (图 7) 分析源自 6 个克隆的抗体。没有抗体含有 80kDa 的亚单位。

序列表

<110> 弗·哈夫曼－拉罗切有限公司

<120> 导致免疫球蛋白生产改善的重链突变体

<130>24355FT

<150>EP 07012774.1
<151>2007-06-29

<160>31

<170>PatentIn version 3.2

<210>1
<211>131
<212>PRT
<213> 智人

<400>1

Gly Asn Thr Phe Arg Pro Glu Val His Leu Leu Pro Pro Pro Ser Glu
1 5 10 15
Glu Leu Ala Leu Asn Glu Leu Val Thr Leu Thr Cys Leu Ala Arg Gly
20 25 30
Phe Ser Pro Lys Asp Val Leu Val Arg Trp Leu Gln Gly Ser Gln Glu
35 40 45
Leu Pro Arg Glu Lys Tyr Leu Thr Trp Ala Ser Arg Gln Glu Pro Ser
50 55 60
Gln Gly Thr Thr Thr Phe Ala Val Thr Ser Ile Leu Arg Val Ala Ala
65 70 75 80
Glu Asp Trp Lys Lys Gly Asp Thr Phe Ser Cys Met Val Gly His Glu
85 90 95
Ala Leu Pro Leu Ala Phe Thr Gln Lys Thr Ile Asp Arg Leu Ala Gly
100 105 110
Lys Pro Thr His Val Asn Val Ser Val Val Met Ala Glu Val Asp Gly
115 120 125
Thr Cys Tyr

130

<210>2
<211>109
<212>PRT
<213> 智人

<400>2

Ala Ala Gln Ala Pro Val Lys Leu Ser Leu Asn Leu Leu Ala Ser Ser
1 5 10 15
Asp Pro Pro Glu Ala Ala Ser Trp Leu Leu Cys Glu Val Ser Gly Phe
20 25 30
Ser Pro Pro Asn Ile Leu Leu Met Trp Leu Glu Asp Gln Arg Glu Val
35 40 45
Asn Thr Ser Gly Phe Ala Pro Ala Arg Pro Pro Pro Gln Pro Gly Ser
50 55 60
Thr Thr Phe Trp Ala Trp Ser Val Leu Arg Val Pro Ala Pro Pro Ser
65 70 75 80
Pro Gln Pro Ala Thr Tyr Thr Cys Val Val Ser His Glu Asp Ser Arg
85 90 95
Thr Leu Leu Asn Ala Ser Arg Ser Leu Glu Val Ser Tyr
100 105

<210>3
<211>110
<212>PRT
<213> 智人

<400>3

Gly Pro Arg Ala Ala Pro Glu Val Tyr Ala Phe Ala Thr Pro Glu Trp
1 5 10 15
Pro Gly Ser Arg Asp Lys Arg Thr Leu Ala Cys Leu Ile Gln Asn Phe
20 25 30
Met Pro Glu Asp Ile Ser Val Gln Trp Leu His Asn Glu Val Gln Leu
35 40 45
Pro Asp Ala Arg His Ser Thr Thr Gln Pro Arg Lys Thr Lys Gly Ser
50 55 60

Gly Phe Phe Val Phe Ser Arg Leu Glu Val Thr Arg Ala Glu Trp Glu
65 70 75 80
Gln Lys Asp Glu Phe Ile Cys Arg Ala Val His Glu Ala Ala Ser Pro
85 90 95
Ser Gln Thr Val Gln Arg Ala Val Ser Val Asn Pro Gly Lys
100 105 110

<210>4
<211>131
<212>PRT
<213> 智人

<400>4

Gly Val Ala Leu His Arg Pro Asp Val Tyr Leu Leu Pro Pro Ala Arg
1 5 10 15
Glu Gln Leu Asn Leu Arg Glu Ser Ala Thr Ile Thr Cys Leu Val Thr
20 25 30
Gly Phe Ser Pro Ala Asp Val Phe Val Gln Trp Met Gln Arg Gly Gln
35 40 45
Pro Leu Ser Pro Glu Lys Tyr Val Thr Ser Ala Pro Met Pro Glu Pro
50 55 60
Gln Ala Pro Gly Arg Tyr Phe Ala His Ser Ile Leu Thr Val Ser Glu
65 70 75 80
Glu Glu Trp Asn Thr Gly Glu Thr Tyr Thr Cys Val Val Ala His Glu
85 90 95
Ala Leu Pro Asn Arg Val Thr Glu Arg Thr Val Asp Lys Ser Thr Gly
100 105 110
Lys Pro Thr Leu Tyr Asn Val Ser Leu Val Met Ser Asp Thr Ala Gly
115 120 125
Thr Cys Tyr
130

<210>5
<211>107
<212>PRT
<213> 智人

<400>5

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp
1 5 10 15
Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45
Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe
50 55 60
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210>6
<211>107
<212>PRT
<213> 智人
<400>6

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
1 5 10 15
Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45
Asn Asn Tyr Lys Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe
50 55 60
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210>7
<211>107
<212>PRT

<213> 智人

<400>7

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu
1 5 10 15
Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Ser Gly Gln Pro Glu
35 40 45
Asn Asn Tyr Asn Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe
50 55 60
Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly
65 70 75 80
Asn Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Phe
85 90 95
Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
100 105

<210>8
<211>107
<212>PRT
<213> 智人

<400>8

Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu
1 5 10 15
Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe
20 25 30
Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu
35 40 45
Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe
50 55 60
Phe Leu Tyr Ser Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly
65 70 75 80
Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr
85 90 95
Thr Gln Lys Ser Leu Ser Leu Ser Leu Gly Lys

100 105

<210>9
<211>300
<212>DNA
<213> 智人
<400>9
ggcttcagcc ccaaggacgt gctggttcgc tggctgcagg ggtcacagga gctgccccgc 60
gagaagtacc tgacttgggc atcccggcag gagcccagcc agggcaccac caccttcgct 120
gtgaccagca tactgcgcgt ggcagccgag gactggaaga agggggacac cttctcctgc 180
atggtgggcc acgaggccct gccgctggcc ttcacacaga agaccatcga ccgcttggcg 240
ggtaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac 300

<210>10
<211>300
<212>DNA
<213> 智人

<400>10
taccacccaa cgtccgtgac tgtcacctgg tacatgggga cacagagcca gccccagaga 60
accttccctg agatacaaag acgggacagc tactacatga caagcagcca gctctccacc 120
cccctccagc agtggcgcca aggcgagtac aaatgcgtgg tccagcacac cgccagcaag 180
agtaagaagg agatcttccg ctggccaggt aggtcgcacc ggagatcacc cagaagggcc 240
ccccaggacc cccagcacct tccactcagg gcctgaccac aaagacagaa gcaagggctg 300

<210>11
<211>300
<212>DNA
<213> 智人

<400>11
tttgcgacgc cggagtggcc ggggagccgg gacaagcgca ccctcgcctg cctgatccag 60
aacttcatgc ctgaggacat ctcggtgcag tggctgcaca acgaggtgca gctcccggac 120
gcccggcaca gcacgacgca gccccgcaag accaagggct ccggcttctt cgtcttcagc 180
cgcctggagg tgaccagggc cgaatgggag cagaaagatg agttcatctg ccgtgcagtc 240
catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggtaaa 300
<210>12
<211>300
<212>DNA

<213> 智人

<400>12

acgggcttct ctcccgcgga cgtcttcgtg cagtggatgc agaggggca gcccttgtcc 60
ccggagaagt atgtgaccag cgccccaatg cctgagcccc aggccccagg ccggtacttc 120
gcccacagca tcctgaccgt gtccgaagag gaatggaaca cgggggagac ctacacctgc 180
gtggtggccc atgaggccct gcccaacagg gtcaccgaga ggaccgtgga caagtccacc 240
ggtaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac 300

<210>13
<211>300
<212>DNA
<213> 智人

<400>13

gtgtacaccc tgcccccatc ccgggatgag ctgaccaaga accaggtcag cctgacctgc 60
ctggtcaaag gcttctatcc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 120
gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 180
agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 240
atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa 300

<210>14
<211>300
<212>DNA
<213> 智人

<400>14

gtgtacaccc tgcccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc 60
ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 120
gagaacaact acaagaccac acctcccatg ctggactccg acggctcctt cttcctctac 180
agcaagctca ccgtggacaa gagcaggtgg cagcagggga acgtcttctc atgctccgtg 240
atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggtaaa 300

<210>15
<211>300
<212>DNA
<213> 智人

<400>15

gtgtacaccc tgcccccatc ccgggaggag atgaccaaga accaggtcag cctgacctgc 60
ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcag cgggcagccg 120
gagaacaact acaacaccac gcctcccatg ctggactccg acggctcctt cttcctctac 180
agcaagctca ccgtggacaa gagcaggtgg cagcagggga acatcttctc atgctccgtg 240
atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggtaaa 300

<210>16
<211>300
<212>DNA
<213> 智人

<400>16
gtgtacaccc tgcccccatc ccaggaggag atgaccaaga accaggtcag cctgacctgc 60
ctggtcaaag gcttctaccc cagcgacatc gccgtggagt gggagagcaa tgggcagccg 120
gagaacaact acaagaccac gcctcccgtg ctggactccg acggctcctt cttcctctac 180
agcaggctaa ccgtggacaa gagcaggtgg caggagggga atgtcttctc atgctccgtg 240
atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggtaaa 300

<210>17
<211>60
<212>DNA
<213> 人工

<220>
<223>IgA

<400>17
ggcaaaccca cccatgtcaa tgtgtctgtt gtcatggcgg aggtggacgg cacctgctac 60

<210>18
<211>60
<212>DNA
<213> 人工

<220>
<223>IgE

<400>18
catgaggcag cgagcccctc acagaccgtc cagcgagcgg tgtctgtaaa tcccggcaaa 60

<210>19
<211>60
<212>DNA
<213> 人工

<220>
<223>IgM

<400>19
ggcaaaccca ccctgtacaa cgtgtccctg gtcatgtccg acacagctgg cacctgctac 60

<210>20
<211>60
<212>DNA
<213> 人工

<220>
<223>IgG1(变体 1)

<400>20
atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa 60

<210>21
<211>60
<212>DNA
<213> 人工

<220>
<223>IgG2(变体 1)

<400>21
atgcatgagg ctctgcacaa ccactacacg cagaagagcc tctccctgtc tccgggcaaa 60

<210>22
<211>60
<212>DNA
<213> 人工

<220>
<223>IgG3

<400>22
atgcatgagg ctctgcacaa ccgcttcacg cagaagagcc tctccctgtc tccgggcaaa 60

<210>23
<211>60
<212>DNA
<213> 人工

<220>
<223>IgG4

<400>23
atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tctgggcaaa 60

<210>24
<211>39
<212>DNA
<213> 人工
<220>
<223> 诱变引物 1

<400>24
agcctctccc tgtccccggg caaatgagtg cgacggccg 39

<210>25
<211>39
<212>DNA
<213> 人工

<220>
<223> 诱变引物 2

<400>25
cggccgtcgc actcatttgc ccggggacag ggagaggct 39

<210>26

<211>34
<212>DNA
<213> 人工

<220>
<223> 诱变引物 3

<400>26
gcctctccct gtccctgggc aaatgagtgc cagg 34

<210>27
<211>34
<212>DNA
<213> 人工

<220>
<223> 诱变引物 4

<400>27
cctggcactc atttgcccag ggacagggag aggc 34

<210>28
<211>321
<212>DNA
<213> 智人

<400>28
gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggatga gctgaccaag 60
aaccaggtca gcctgacctg cctggtcaaa ggcttctatc ccagcgacat cgccgtggag 120
tgggagagca atgggcagcc ggagaacaac tacaagacca cgcctcccgt gctggactcc 180
gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg 240
aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac acagaagagc 300
ctctccctgt ctccgggtaa a 321

<210>29
<211>321
<212>DNA
<213> 智人

<400>29

gggcagcccc gagaaccaca ggtgtacacc ctgcccccat cccgggagga gatgaccaag 60
aaccaggtca gcctgacctg cctggtcaaa ggcttctacc ccagcgacat ctccgtggag 120
tgggagagca atgggcagcc ggagaacaac tacaagacca cacctcccat gctggactcc 180
gacggctcct tcttcctcta cagcaagctc accgtggaca agagcaggtg gcagcagggg 240
aacgtcttct catgctccgt gatgcatgag gctctgcaca accactacac gcagaagagc 300
ctctccctgt ctccgggtaa a 321

<210>30
<211>60
<212>DNA
<213> 人工

<220>
<223>IgG1(变体 2)
<400>30
atgcatgagg ctctgcacaa ccactacaca cagaagagcc tctccctgtc tccgggcaaa 60

<210>31
<211>213
<212>DNA
<213> 人工

<220>
<223>IgG2(变体 2)

<400>31
atctccgtgg agtgggagag caatgggcag ccggagaaca actacaagac cacacctccc 60
atgctggact ccgacggctc cttcttcctc tacagcaagc tcaccgtgga caagagcagg 120
tggcagcagg ggaacgtctt ctcatgctcc gtgatgcatg aggctctgca caaccactac 180
acgcagaaga gcctctccct gtctccgggc aaa 213

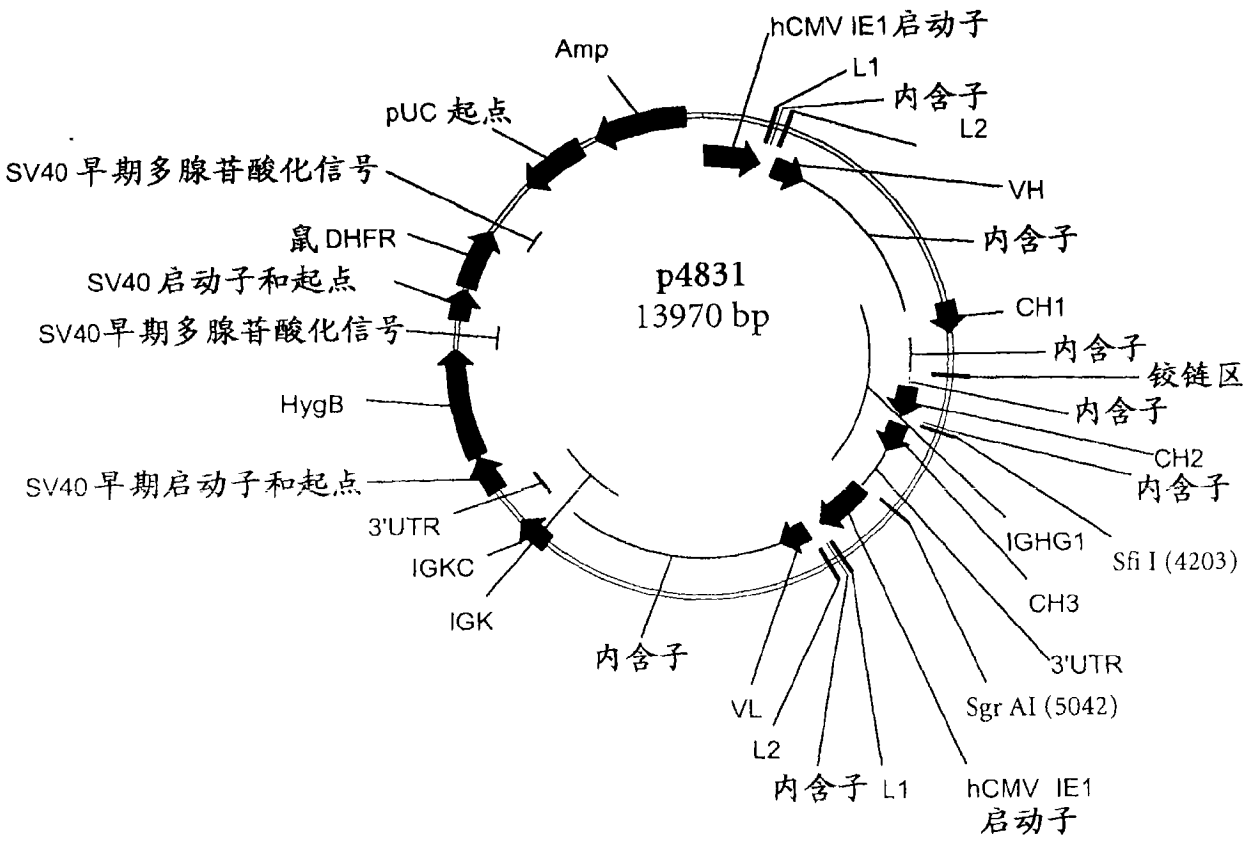

hCMV IE1启动子
Amp
L1
内含子
L2
pUC 起点
SV40 早期多腺苷酸化信号
VH
鼠 DHFR
内含子
SV40 启动子和起点
p4831
13970 bp
CH1
SV40早期多腺苷酸化信号
内含子
铰链区
HygB
内含子
CH2
SV40 早期启动子和起点
内含子
IGHG1
3'UTR
Sfi I (4203)
IGKC
CH3
IGK
3'UTR
内含子
Sgr AI (5042)
VL
L2
内含子 L1
hCMV IE1
启动子

图 1

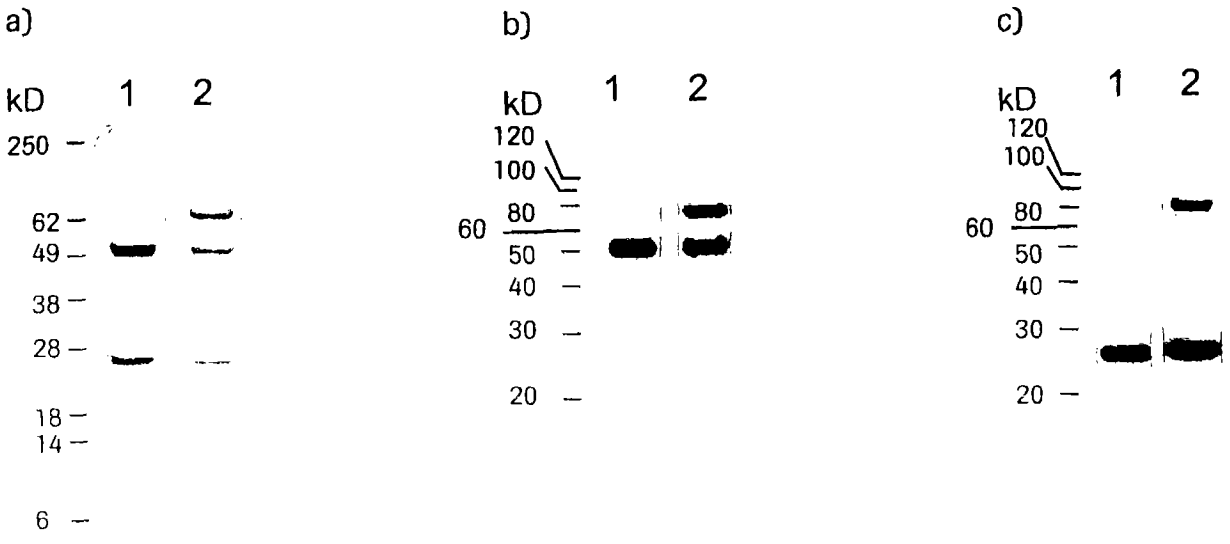

a)
kD
1
2
250
62
49
38
28
18
14
6
b)
kD
1
2
120
100
80
60
50
40
30
20
c)
kD
1
2
120
100
80
60
50
40
30
20

图 2

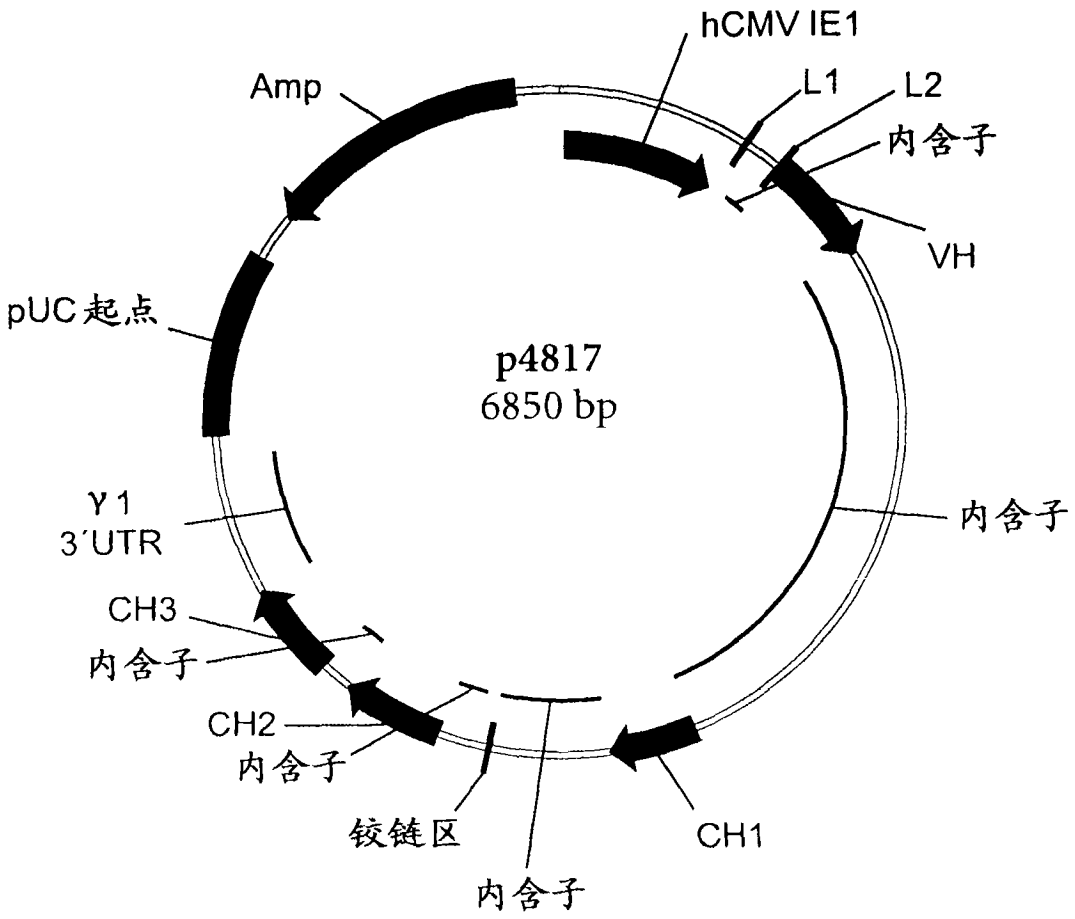

hCMV IE1
Amp
L1
L2
内含子
VH
pUC起点
p4817
6850 bp
γ1
3′UTR
内含子
CH3
内含子
CH2
内含子
铰链区
CH1
内含子

图 3

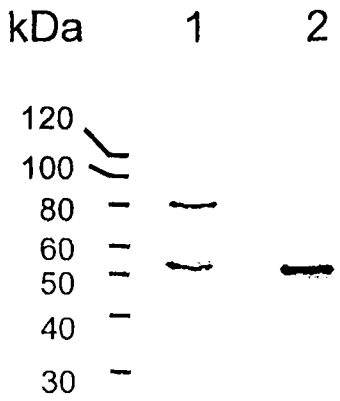

kDa
1
2
120
100
80
60
50
40
30

图 4

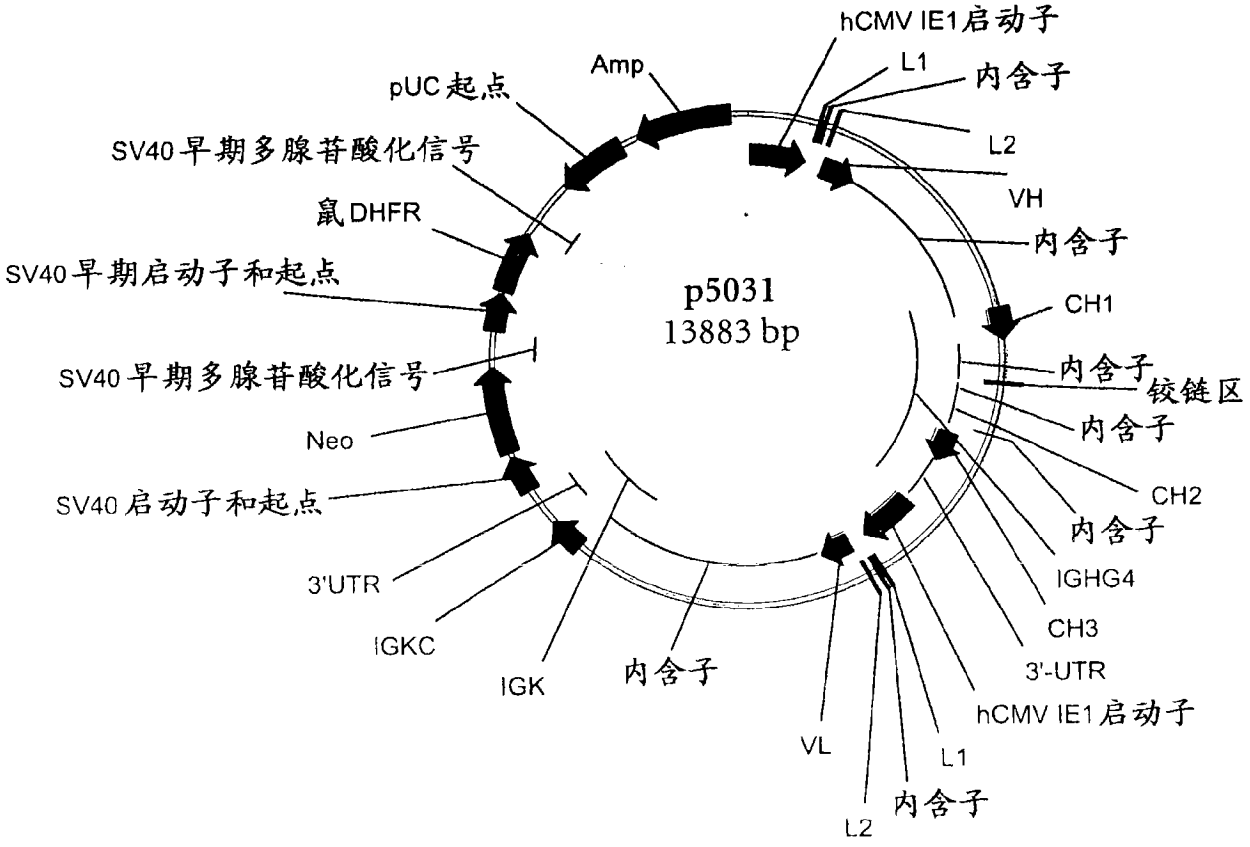

hCMV IE1启动子
L1
内含子
L2
VH
内含子
CH1
内含子
铰链区
内含子
CH2
内含子
IGHG4
CH3
3'-UTR
hCMV IE1启动子
L1
内含子
L2
VL
内含子
IGK
IGKC
3'UTR
SV40 启动子和起点
Neo
SV40早期多腺苷酸化信号
SV40早期启动子和起点
鼠DHFR
SV40早期多腺苷酸化信号
pUC起点
Amp
p5031
13883 bp

图 5

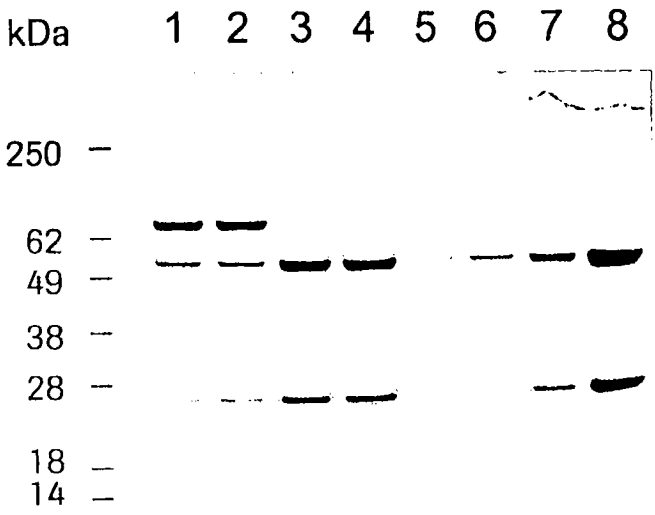

kDa
1 2 3 4 5 6 7 8
250
62
49
38
28
18
14

图 6

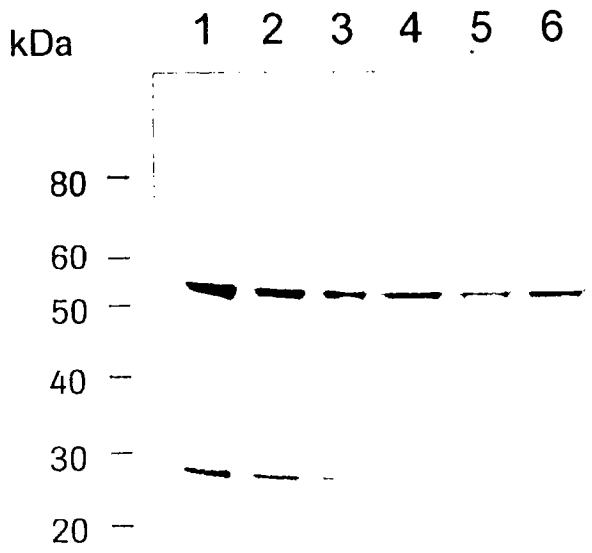


图 7