



# (12) 发明专利

(10) 授权公告号 CN 113528448 B

(45) 授权公告日 2023.01.24

(21) 申请号 202010290822.3

C12R 1/91 (2006.01)

(22) 申请日 2020.04.14

(56) 对比文件

(65) 同一申请的已公布的文献号

CN 110799530 A, 2020.02.14

申请公布号 CN 113528448 A

CN 107533051 A, 2018.01.02

(43) 申请公布日 2021.10.22

WO 2019202041 A1, 2019.10.24

(73) 专利权人 同济大学

TW 201945395 A, 2019.12.01

地址 200092 上海市杨浦区四平路1239号

CN 109952318 A, 2019.06.28

CN 112004831 A, 2020.11.27

(72) 发明人 章小清 刘玲

Germán G Gornalusse. "HLA-E-expressing pluripotent stem cells escape allogeneic responses and lysis by NK cells". 《Nature Biotechnology》. 2017,

(74) 专利代理机构 上海光华专利事务所(普通合伙) 31219

Xiao Han等. Generation of hypoimmunogenic human pluripotent stem cells. 《BIOLOGICAL SCIENCES》. 2019,

专利代理师 严晨 许亦琳

审查员 修旺珊

(51) Int. Cl.

C12N 5/10 (2006.01)

C12N 15/62 (2006.01)

C12N 15/867 (2006.01)

C07K 19/00 (2006.01)

C07K 14/74 (2006.01)

权利要求书1页 说明书15页

序列表14页 附图6页

(54) 发明名称

一种人胚胎干细胞的构建方法

(57) 摘要

本发明涉及生物技术领域,特别是涉及一种人胚胎干细胞的构建方法。本发明提供一种人胚胎干细胞的构建方法,所述构建方法包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸,所述B2M-HLA-G1融合蛋白包括第一B2M片段和HLA-G1片段,整合获得的人胚胎干细胞不表达游离的B2M蛋白。本发明所提供的人胚胎干细胞的构建方法构建获得的人胚胎干细胞细胞系内源性HLA-A,B,C蛋白无法到达细胞膜表面,因此该细胞系不能被CD8<sup>+</sup> T细胞识别,可以有效阻止同种异体细胞免疫排斥,且胚胎干细胞的多能性以及增殖的能力并未受到影响,依然可以作为未来细胞移植的目的细胞来源。

1. 一种人胚胎干细胞的构建方法,所述构建方法包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸,所述B2M-HLA-G1融合蛋白包括第一B2M片段和HLA-G1片段,整合获得的人胚胎干细胞不表达游离的B2M蛋白;将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸,所述B2M-HLA-G5融合蛋白包括第二B2M片段和HLA-G5片段,所述将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸的方法具体包括:通过慢病毒载体将编码B2M-HLA-G5融合蛋白的多核苷酸整合入人胚胎干细胞的基因组;所述B2M-HLA-G1融合蛋白氨基酸序列如SEQ ID No. 4所示;所述B2M-HLA-G5融合蛋白氨基酸序列如SEQ ID No. 8所示。

2. 如权利要求1所述的构建方法,其特征在于,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸的方法具体包括:将HLA-G1片段的编码基因与人胚胎干细胞中内源性的B2M基因融合。

3. 如权利要求2所述的构建方法,其特征在于,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸的方法具体包括:将HLA-G1片段的编码基因替换位于内源B2M基因外显子3中的终止密码子。

4. 一种人胚胎干细胞,由如权利要求1-3任一权利要求所述的人胚胎干细胞的构建方法构建获得。

## 一种人胚胎干细胞的构建方法

### 技术领域

[0001] 本发明涉及生物技术领域,特别是涉及一种人胚胎干细胞的构建方法。

### 背景技术

[0002] 多能性干细胞具有能在体外培养环境中无限分裂增殖以及在特定诱导条件下定向分化为三胚层来源的所有细胞类型的能力,因此,在再生医学的未来发展中被给予了厚望。早在十年前,Thomson (Thomson et al.,1998) 与Reubinoff (Reubinoff et al.,2000) 教授等人便率先开始建立并分化人类胚胎干细胞系。在他们的研究基础上,更多的科学家开展了应用多能性干细胞作为功能性细胞来源,修复受损组织功能与治疗退行性疾病的研究。在动物模型当中,分化后的胚胎干细胞表现出了对脊髓损伤 (Soundararajan et al., 2007)、糖尿病 (Liu and Lee,2012)、帕金森氏病 (Barberi et al.,2003)、肝衰竭 (Soto-Gutierrez et al.,2006) 等疾病的治疗能力。尽管有这些令人振奋的研究结果,与多能干细胞移植植物相关的免疫排斥问题尚未有良好的解决方案,这已经成为了未来细胞替代疗法的最大临床应用障碍之一。

### 发明内容

[0003] 鉴于以上所述现有技术的缺点,本发明的目的在于提供一种人胚胎干细胞的构建方法,用于解决现有技术中的问题。

[0004] 为实现上述目的及其他相关目的,本发明一方面提供一种人胚胎干细胞的构建方法,所述构建方法包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸,所述B2M-HLA-G1融合蛋白包括第一B2M片段和HLA-G1片段,整合获得的人胚胎干细胞不表达游离的B2M蛋白。

[0005] 在本发明一些实施方式中,所述HLA-G1片段包括:

[0006] a) 氨基酸序列如SEQ ID No.1所示的多肽片段;或,

[0007] b) 氨基酸序列与SEQ ID No.1具有90%以上序列一致性且具有a)限定的多肽片段的功能的多肽片段。

[0008] 在本发明一些实施方式中,所述第一B2M片段包括:

[0009] c) 氨基酸序列如SEQ ID No.2所示的多肽片段;或,

[0010] d) 氨基酸序列与SEQ ID No.2具有90%以上序列一致性且具有a)限定的多肽片段的功能的多肽片段。

[0011] 在本发明一些实施方式中,所述B2M-HLA-G1融合蛋白还包括位于第一B2M片段和HLA-G1片段之间的第一柔性连接肽段;

[0012] 和/或,所述B2M-HLA-G1融合蛋白自N端至C端依次包括第一B2M片段和HLA-G1片段,所述B2M-HLA-G1融合蛋白包括氨基酸序列如SEQ ID No.4所示的多肽片段。

[0013] 在本发明一些实施方式中,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸的方法具体包括:将HLA-G1片段的编码基因与人胚胎干细胞中内源

性的B2M基因融合,优选为将HLA-G1片段的编码基因替换位于内源B2M基因外显子3中的终止密码子。

[0014] 在本发明一些实施方式中,所述构建方法还包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸,所述B2M-HLA-G5融合蛋白包括第二B2M片段和HLA-G5片段。

[0015] 在本发明一些实施方式中,所述第二B2M片段包括:

[0016] e) 氨基酸序列如SEQ ID No.5所示的多肽片段;或,

[0017] f) 氨基酸序列与SEQ ID No.5具有90%以上序列一致性且具有c)限定的多肽片段的功能的多肽片段。

[0018] 在本发明一些实施方式中,所述HLA-G5片段包括:

[0019] g) 氨基酸序列如SEQ ID No.6所示的多肽片段;或,

[0020] h) 氨基酸序列与SEQ ID No.6具有90%以上序列一致性且具有e)限定的多肽片段的功能的多肽片段。

[0021] 在本发明一些实施方式中,所述B2M-HLA-G5融合蛋白还包括位于第二B2M片段和HLA-G5片段之间的第二柔性连接肽段;

[0022] 和/或,所述B2M-HLA-G5融合蛋白自N端至C端依次包括第二B2M片段和HLA-G5片段,所述B2M-HLA-G5融合蛋白包括氨基酸序列如SEQ ID No.8所示的多肽片段。

[0023] 在本发明一些实施方式中,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸的方法具体包括:通过慢病毒载体将编码B2M-HLA-G5融合蛋白的多核苷酸整合入人胚胎干细胞的基因组。

[0024] 本发明另一方面提供一种人胚胎干细胞,由上述的人胚胎干细胞的构建方法构建获得。

## 附图说明

[0025] 图1显示为本发明所构建的B2M<sup>null</sup>, B2M<sup>mHLA<sub>G</sub></sup>, B2M<sup>m/sHLA<sub>G</sub></sup> hESCs细胞系及相关验证实验示意图。其中, (A) 为逐步构建B2M<sup>null</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>细胞系的基因编辑策略示意图; (B) 为敲除B2M基因的双gRNA基因编辑策略示意图,切割位点由剪刀和箭头指出, PAM序列为紫色,用于基因组PCR的设计引物的位置显示为三角形; (C) 为双gRNA基因编辑方案敲除B2M实验中所获得的7个抗药型细胞克隆的基因组PCR结果示意图,其中,双等位基因缺失的克隆标记为红色,使用的引物在(B)中指出; (D) 为在B2M基因座中敲入HLA-G1的基因编辑策略示意图,用于基因组PCR的引物显示为三角形,设计用于Southern Blot的探针显示为黑线; (E) 为HLA-G1敲入实验中所获得的5个细胞克隆的基因组DNAPCR结果示意图,纯合子敲入细胞系标记为红色,所用的引物在(D)中指示; (F) 为Southern Blot结果示意图,证明了在(E)中所示的细胞克隆4中,HLA-G1被精确的整合进了B2M基因中,且在基因组其他位点未发现整合,所使用的探针在(D)中标出。

[0026] 图2显示为本发明所构建的B2M<sup>null</sup>与B2M<sup>mHLA<sub>G</sub></sup>细胞系脱靶现象验证结果示意图。其中, (A) 为B2M<sup>null</sup> hESCs结果示意图,其中,靶向B2M的gRNA1和gRNA2各自的5个预测脱靶位点均未发现DNA突变; (B) 为B2M<sup>mHLA<sub>G</sub></sup> hESCs结果示意图,其中,靶向B2M的gRNA3的所有5个预测脱靶位点均未发现DNA突变。

[0027] 图3显示为本发明WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G与B2M<sup>m/sHLA</sup>G细胞系中B2M与HLA I蛋白质表达情况示意图。其中, (A) 为在加入(橙色)或不加(浅蓝色) IFN- $\gamma$  刺激的情况下, 使用流式细胞术检测B2M, HLA-G, HLA-A\*03 (存在于WT H9细胞中) 和HLA-BC蛋白质的表达示意图, 其中, 假阳性对照组红色, 每种hESCs细胞系用胰蛋白酶消化, 以形成单细胞悬液进行流式细胞检测; (B) 为Western Blot实验结果示意图, 证实了B2M<sup>m/sHLA</sup>G hESCs可以表达和分泌HLA-G5蛋白质。

[0028] 图4显示为本发明WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G与B2M<sup>m/sHLA</sup>G细胞系分化产生的心肌细胞中, HLA I与心肌特异性表达基因的表达情况示意图。其中, (A) 为WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs来源的心肌细胞中心肌特异性基因的表达情况示意图。(B) 为免疫染色显示了在WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs来源的心肌细胞中TNNT2的表达情况示意图。(C) 为流式细胞检验结果示意图, 显示不同hESCs来源的心肌细胞中B2M, HLA-G, HLA-A\*03和HLA-BC的表达情况。

[0029] 图5显示为本发明针对不同基因型hESCs及其分化细胞的免疫反应结果示意图。其中, (A) 为通过流式细胞术分析与WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs分化的心肌细胞共培养的同种异体PBMC中的CFSE信号结果示意图, 与PHA共培养的PBMC用作阳性对照, 加IFN- $\gamma$  组是指在MLR之前用IFN- $\gamma$  处理心肌细胞48小时; (B) 为收集与WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs分化的心肌细胞共培养的PBMC的上清液, 用于检测上清液中IFN- $\gamma$  的浓度, 加IFN- $\gamma$  组, 是指在与PBMC共培养之前, 用IFN- $\gamma$  处理心肌细胞48小时; 数据表示为三个独立实验的平均值 $\pm$ SEM。\*\*\* $p < 0.01$ , 两末端Student's t检验, ND, 未检测到; (C) 为收集与WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs共培养的上清液, 用于检测上清液中IFN- $\gamma$  的浓度, 数据表示为三个独立实验的平均值 $\pm$ SEM, \* $p < 0.05$ , \*\* $p < 0.01$ 表示指定基因型hESCs与野生型hESCs组比较, \$ $p < 0.05$ 表示B2M<sup>mHLA</sup>G hESCs与B2M<sup>m/sHLA</sup>G hESCs组比较, 两末端Student's t检验; (D) 为在B2M<sup>m/sHLA</sup>G hESCs条件培养基与RPMI-1640中共培养NK-92与WT hESCs, 检测上清液中IFN- $\gamma$  的浓度, 数据表示为三个独立实验的平均值 $\pm$ SEM, \*\* $p < 0.01$ , 两末端Student's t检验。

[0030] 图6显示为本发明由B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESC分化得到的畸胎瘤在体内不受同种异体免疫排斥的影响结果示意图。其中, (A) 为针对HLA抗原诱发同种异体PBMC免疫排斥反应体内实验方案的示意图; (B) 为将WT, B2M<sup>null1</sup>, B2M<sup>mHLA</sup>G和B2M<sup>m/sHLA</sup>G hESCs细胞注射到M-NSG小鼠皮下(每个部位大约 $2 \times 10^6$ 个细胞)。一个月后, 一半数量的小鼠移植已经激活的同种异体PBMC。两周后, 回收畸胎瘤并用相应抗体染色。通过抗CD3抗体鉴定T细胞, 并将CD3<sup>-</sup>CD56<sup>+</sup>细胞认定为NK细胞。细胞核用DAPI染色。使用同等条件对M-NSG小鼠中注射相同体积PBS的畸胎瘤进行染色, 作为阴性对照。比例尺, 50 $\mu$ m。每组n=3; (C) 对浸润畸胎瘤的T细胞和NK细胞进行计数。对三个随机选择的视野中的CD3<sup>+</sup>和CD3<sup>-</sup>CD56<sup>+</sup>细胞进行计数。数据表示为三个独立实验的平均值 $\pm$ SEM。\* $p < 0.05$ , 未配对的两尾学生t检验; (D) 为通过qPCR分析在PBMC注射后第14天收获的WT (n=3), B2M<sup>null1</sup> (n=3), B2M<sup>mHLA</sup>G (n=4) 和B2M<sup>m/sHLA</sup>G (n=4) 畸胎瘤中IL-2mRNA的表达量。数据表示为三个独立实验的平均值 $\pm$ SEM。\* $p < 0.05$ , 未配对的两尾学生t检验。

## 具体实施方式

[0031] 为了使本发明的发明目的、技术方案和有益技术效果更加清晰,以下结合实施例对本发明进行进一步详细说明,熟悉此技术的人士可由本说明书所揭露的内容容易地了解本申请发明的其他优点及功效。

[0032] 本发明发明人经过大量实践研究,发现将胚胎干细胞整合HLA-G1片段和/或HLA-G5片段以后,可以使细胞不被CD8<sup>+</sup>T细胞识别,避免CD8<sup>+</sup>T的排斥,还可以抑制NK细胞、抗原递呈细胞等淋巴细胞参与的免疫排斥作用,从而更加高效的抑制同种异体免疫排斥反应,在此基础上完成了本发明。

[0033] 本发明第一方面提供一种人胚胎干细胞的构建方法,所述构建方法包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸,所述B2M-HLA-G1融合蛋白包括B2M片段和HLA-G1片段,整合获得的人胚胎干细胞不表达游离的B2M蛋白。通过将胚胎干细胞整合HLA-G1片段以后,能够成功表达B2M-HLA-G1融合蛋白,且不表达游离的B2M蛋白(即单独地、未进行融合的B2M蛋白),B2M-HLA-G1融合蛋白质可以到达细胞膜表面,且该蛋白地表达受到内源性B2M基因启动子调控,从而可以使细胞不被CD8<sup>+</sup>T细胞识别,避免CD8<sup>+</sup>T的排斥,从而可以有效阻止同种异体细胞免疫排斥。

[0034] 本申请中,所使用的人胚胎干细胞可以是利用未经过体内发育的受精14天以内的人类胚胎分离或者获取的,例如,可以是商用的H9 hESC细胞系。

[0035] 本发明所提供的构建方法中,所述HLA-G1片段的编码基因可以包括HLA-G1的重链阅读框的编码序列,所述HLA-G1片段可以包括:

[0036] a) 氨基酸序列如SEQ ID No.1所示的多肽片段;或,

[0037] b) 氨基酸序列与SEQ ID No.1具有90%以上序列一致性且具有a)限定的多肽片段的功能的多肽片段。具体的,所述b)中的氨基酸序列具体指:如SEQ ID No.1所示的氨基酸序列经过取代、缺失或者添加一个或多个(具体可以是1-50、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,或者在N-末端和/或C-末端添加一个或多个(具体可以是1-50个、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,且具有氨基酸如SEQ ID No.1所示的多肽片段的功能的多肽片段。例如,所述HLA-G1片段通常具有完整的 $\alpha$ 重链结构,与经典HLA I分子不同,其主要具有免疫抑制的功能(例如,可以与抑制性受体相结合从而可以调控B细胞、T细胞、NK细胞和APC细胞介导的免疫反应等,这些抑制性受体主要包括ILT2/CD85j/LILRB1,ILT4/CD85d/LILRB2,和KIR2DL4/CD158d等)。所述b)中的氨基酸序列可与SEQ ID No.1具有90%、93%、95%、97%、或99%以上的一致性(Sequence identity)。所述HLA-G1片段通常来源于人。

[0038] 本申请中,序列一致性(sequence identity)指参与对比的序列中相同残基的百分比。可采用本领域周知的计算软件计算两条或多条目的序列的序列一致性,这些软件可获自如NCBI。

[0039] 本发明所提供的构建方法中,所述第一B2M片段可以包括:

[0040] c) 氨基酸序列如SEQ ID No.2所示的多肽片段;或,

[0041] d) 氨基酸序列与SEQ ID No.2具有90%以上序列一致性且具有c)限定的多肽片段的功能的多肽片段。具体的,所述d)中的氨基酸序列具体指:如SEQ ID No.2所示的氨基酸序列经过取代、缺失或者添加一个或多个(具体可以是1-50、1-30个、1-20个、1-10个、1-5

个、1-3个、1个、2个、或3个)氨基酸而得到的,或者在N-末端和/或C-末端添加一个或多个(具体可以是1-50个、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,且具有氨基酸如SEQ ID No.2所示的多肽片段的功用的多肽片段。例如,所述第一B2M片段通常具有 $\beta$ 折叠的片状结构,其主要具有与主要组织相容性复合物(MHC) I类重链通过非共价键结合的功能。所述d)中的氨基酸序列可与SEQ ID No.2具有90%、93%、95%、97%、或99%以上的一致性(Sequence identity)。所述第一B2M片段通常来源于人。

[0042] 本发明所提供的构建方法中,所述B2M-HLA-G1融合蛋白还可以包括第一柔性连接肽段,所述第一柔性连接肽段通常位于HLA-G1片段和第一B2M片段之间。所述第一柔性连接肽段通常可以为一段长度合适的由甘氨酸(G)、丝氨酸(S)和/或丙氨酸(A)构成的柔性多肽,从而使相邻的蛋白质结构域可相对于彼此自由移动,例如,所述第一柔性连接肽段的氨基酸序列可以包括如(GS)<sub>n</sub>、(GGS)<sub>n</sub>、(GGSG)<sub>n</sub>、(GGGS)<sub>n</sub>A、(GGGGS)<sub>n</sub>A、(GGGGA)<sub>n</sub>A、(GGGGG)<sub>n</sub>A等序列,其中,n选自1-10之间的整数。在本发明一具体实施例中,所述第一柔性连接肽段的氨基酸序列的长度可以为5-26。在本发明一具体实施例中,所述第一柔性连接肽段可以包括氨基酸序列如SEQ ID NO.3所示的多肽片段。在本发明另一具体实施例中,所述B2M-HLA-G1融合蛋白自N端至C端依次包括第一B2M片段和HLA-G1片段,所述B2M-HLA-G1融合蛋白包括氨基酸序列如SEQ ID No.4所示的多肽片段。

[0043] GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer (SEQ ID NO.3)

[0044] 本发明所提供的构建方法中,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G1融合蛋白的多核苷酸的方法具体可以包括:将HLA-G1片段的编码基因与人胚胎干细胞中内源性的B2M基因融合。所述内源性的B2M基因的具体信息可以参照chr15:44,715,509-44,717,171GRCh38/hg38 human assembly。在本发明一具体实施例中,可以将HLA-G1片段的编码基因替换位于内源B2M基因外显子3中的终止密码子。

[0045] 本发明所提供的构建方法中,所述构建方法还可以包括:将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸,所述B2M-HLA-G5融合蛋白包括第二B2M片段和HLA-G5片段。构建获得的人胚胎干细胞可以成功表达分泌型B2M-HLA-G5融合蛋白,从而可以在有效避免CD8<sup>+</sup>T排斥的同时,还可以抑制NK细胞,抗原递呈细胞等淋巴细胞参与的免疫排斥作用,从而更加高效的抑制同种异体免疫排斥反应。

[0046] 本发明所提供的构建方法中,所述第二B2M片段可以包括:

[0047] e) 氨基酸序列如SEQ ID No.5所示的多肽片段;或,

[0048] f) 氨基酸序列与SEQ ID No.5具有90%以上序列一致性且具有e)限定的多肽片段的功用的多肽片段。具体的,所述f)中的氨基酸序列具体指:如SEQ ID No.5所示的氨基酸序列经过取代、缺失或者添加一个或多个(具体可以是1-50、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,或者在N-末端和/或C-末端添加一个或多个(具体可以是1-50个、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,且具有氨基酸如SEQ ID No.5所示的多肽片段的功用的多肽片段。例如,所述第二B2M片段通常具有 $\beta$ 折叠的片状结构,其主要具有与主要组织相容性复合物(MHC) I类重链通过非共价键结合的功能。所述f)中的氨基酸序列可与SEQ ID No.5具有90%、93%、95%、97%、或99%以上的一致性(Sequence identity)。所述第二B2M片段通常来源于人。

[0049] 本发明所提供的构建方法中,所述HLA-G5片段包括:

[0050] g) 氨基酸序列如SEQ ID No.6所示的多肽片段;或,

[0051] h) 氨基酸序列与SEQ ID No.6具有90%以上序列一致性且具有g)限定的多肽片段的功用的多肽片段。具体的,所述h)中的氨基酸序列具体指:如SEQ ID No.6所示的氨基酸序列经过取代、缺失或者添加一个或多个(具体可以是1-50、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,或者在N-末端和/或C-末端添加一个或多个(具体可以是1-50个、1-30个、1-20个、1-10个、1-5个、1-3个、1个、2个、或3个)氨基酸而得到的,且具有氨基酸如SEQ ID No.6所示的多肽片段的功用的多肽片段。例如,所述HLA-G5片段通常具有完整的 $\alpha$ 重链胞外结构,与经典型HLA I分子不同,其主要具有免疫抑制的功能(例如,可以与抑制性受体相结合从而可以调控B细胞、T细胞、NK细胞和APC细胞介导的免疫反应等,这些抑制性受体主要包括ILT2/CD85j/LILRB1,ILT4/CD85d/LILRB2,和KIR2DL4/CD158d等)。所述h)中的氨基酸序列可与SEQ ID No.6具有90%、93%、95%、97%、或99%以上的一致性(Sequence identity)。所述HLA-G5片段通常来源于人。

[0052] 本发明所提供的构建方法中,所述B2M-HLA-G5融合蛋白还可以包括第二柔性连接肽段,所述第二柔性连接肽段通常位于HLA-G5片段和第二B2M片段之间。所述第二柔性连接肽段通常可以为一段长度合适的由甘氨酸(G)、丝氨酸(S)和/或丙氨酸(A)构成的柔性多肽,从而使相邻的蛋白质结构域可相对于彼此自由移动,例如,所述第二柔性连接肽段的氨基酸序列可以包括如(GS)<sub>n</sub>、(GGS)<sub>n</sub>、(GGSG)<sub>n</sub>、(GGGS)<sub>n</sub>A、(GGGS)<sub>n</sub>A、(GGGGA)<sub>n</sub>A、(GGGGG)<sub>n</sub>A等序列,其中,n选自1-10之间的整数。在本发明一具体实施例中,所述第二柔性连接肽段的氨基酸序列的长度可以为5-26。在本发明一具体实施例中,所述第二柔性连接肽段可以包括氨基酸序列如SEQ ID NO.7所示的多肽片段。在本发明另一具体实施例中,所述B2M-HLA-G5融合蛋白自N端至C端依次包括第二B2M片段和HLA-G5片段,所述B2M-HLA-G5融合蛋白包括氨基酸序列如SEQ ID No.8所示的多肽片段。

[0053] GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer (SEQ ID NO.7)

[0054] 本发明所提供的构建方法中,将人胚胎干细胞的基因组中整合外源的编码B2M-HLA-G5融合蛋白的多核苷酸的方法具体可以包括:通过慢病毒载体将编码B2M-HLA-G5融合蛋白的多核苷酸整合入人胚胎干细胞的基因组,使其可以表达B2M-HLA-G5融合蛋白。

[0055] 本发明第二方面提供一种人胚胎干细胞,由本发明第一方面所提供的人胚胎干细胞的构建方法构建获得。

[0056] 本发明所提供的人胚胎干细胞的构建方法可以构建成功地在人类胚胎干细胞地两个内源性B2M位点整合了HLA-G1片段序列,并且通过慢病毒载体将编码B2M-HLA-G5融合蛋白的序列整合入了人胚胎干细胞的基因组中。构建获得的人胚胎干细胞细胞系内源性HLA-A,B,C蛋白无法到达细胞膜表面,因此该细胞系不能被CD8<sup>+</sup>T细胞识别,可以有效阻止同种异体细胞免疫排斥,且胚胎干细胞的多能性以及增殖的能力并未受到影响,依然可以作为未来细胞移植的目的细胞来源。此外,细胞系可以成功表达B2M-HLA-G1融合蛋白,该蛋白质可以到达细胞膜表面,且该蛋白地表达受到内源性B2M基因启动子调控。此外,能够进一步成功表达分泌型B2M-HLA-G5融合蛋白的细胞系可以在有效避免CD8<sup>+</sup>T排斥的同时,还可以抑制NK细胞,抗原递呈细胞等淋巴细胞参与的免疫排斥作用,从而更加高效的抑制同



种异体免疫排斥反应。

[0057] 下面通过实施例对本申请的发明予以进一步说明,但并不因此而限制本申请的范围。

[0058] 实施例1

[0059] 通过双gRNA基因编辑系统构建双等位B2M基因缺失的hESCs细胞系

[0060] 双gRNA基因编辑策略能够高效、特异性地敲除双等位基因,并且被切断的DNA主要通过直接末端连接进行修复而不会引入核苷酸地插入或缺失(Liu et al.,2016)。为了敲除人类胚胎干细胞(human embryonic stem cells,hESCs)中的B2M基因,设计两个相邻的gRNA,一个位于B2M基因外显子2的上游,另一个位于外显子3的下游(图1A和图1B),具体序列分别如表1中的B2M gRNA 1、B2M gRNA 2所示。绝大部分成熟的B2M蛋白由该靶向序列编码。

[0061] 表1

gRNAs	Sequences
B2M gRNA 1	TAGCCCCAAGTGAAATACCC TGG (SEQ ID No. 11)
B2M gRNA 2	GCTGTCTATAAATAGTCCTC AGG (SEQ ID No. 12)
B2M gRNA 3	ACATGTAAGCAGCATCATGG AGG (SEQ ID No. 13)

[0063] 将两个gRNA载体(Addgene#41824),CAG启动子驱动的Cas9-P2A-GFP载体(Addgene#44719)与瞬时表达嘌呤霉素的载体(SEQ ID NO.14)一起共同电转到H9 hESC(由美国威斯康星大学张素春教授实验室提供)中。经过嘌呤霉素筛选后,挑选抗性细胞克隆并扩增以进行基因组DNA PCR分析,在被挑选出来的克隆中有28.6%(2/7)是双等位基因DNA缺失(B2M<sup>null</sup>hESCs),而选择的克隆中有57.1%(4/7)是单等位基因DNA缺失(图1C)。为了检验这两条gRNA诱发的潜在脱靶作用,使用Sanger测序分析了由张锋教授实验室设计的在线程序(<http://crispr.mit.edu/>)预测的潜在脱靶位点的DNA完整性。结果显示,B2Mnull hESC中每条gRNA的五个主要可疑脱靶位点DNA均保持完整(图2A)。

[0064] 实施例2

[0065] 构建表达HLA-G1,HLA-G5 hESCs细胞系

[0066] 设计了两种表达HLA-G蛋白质的hESCs细胞系(图1A)。在B2MmHLA hESCs中,HLA-G1重链通过柔性(G<sub>4</sub>S)<sub>4</sub>连接肽段与内源性B2M蛋白质融合。为了用编码HLA-G1重链阅读框和柔性(G<sub>4</sub>S)<sub>4</sub>连接肽段的序列替换位于内源B2M基因外显子3中的终止密码子,设计跨越B2M终止密码子的B2M gRNA3(参见表1),这样成功实现敲入的基因位点将不能被B2M gRNA3结合,从而提高了建立纯合子敲入细胞系的效率。将CAG启动子驱动的Cas9载体(同实施例1),重组供体(SEQ ID No.9),B2M gRNA3载体(Addgene#41824)与瞬时表达嘌呤霉素载体(同实施例1)一起共同电转到H9 hESC中。通过这种方法,(G<sub>4</sub>S)<sub>4</sub>-HLA-G1将连接到B2M蛋白的末端,并在B2M启动子的调控下与内源B2M一起表达(图1D)。药物筛选后,挑取存活细胞克隆并扩增以进行基因组DNA PCR分析。获得了5个抗嘌呤霉素的细胞克隆,基因组PCR分析显示挑选出来的细胞克隆中,20.0%(1/5)为双拷贝敲入(B2M<sup>mHLA</sup>),40.0%(2/5)为单拷贝敲入(图1E)。通过Southern Blot实验进一步分析纯合子敲入克隆,结果显示在该细胞克隆中,(G<sub>4</sub>S)<sub>4</sub>-HLA-G1仅仅被整合进了预定的B2M基因位置(图1F)。

[0067] 为了获得另一种同时表达膜蛋白HLA-G1与分泌蛋白HLA-G5的hESCs, B2M<sup>m/sHLAG</sup>hESCs,我们将无终止密码子的B2M CDS序列, (G<sub>4</sub>S)<sub>4</sub>柔性连接肽序列和HLA-G5重链编码序列整合到具有嘌呤霉素抗性的慢病毒载体(pLVX-CAG-Puro质粒,插入位点为BamHI、MluI)中,使用该病毒感染B2M<sup>mHLAG</sup>hESCs细胞系,然后进行药物筛选和抗药性hESCs的扩增(图1A)。此外,与B2M<sup>nu11</sup>hESCs中观察到的情况类似,B2M<sup>mHLAG</sup>hESCs中的5个B2M gRNA3疑似脱靶位点DNA保持完整,说明基因编辑过程中没有脱靶现象产生(图2B)。

[0068] HLA-G1重链序列(HLA-G1片段)的氨基酸序列:

[0069] GSHSMRYFSAAVSRPGRGEPRIAMGYVDDTQFVRFSDSACPRMEPRAPWVEQEGPEYWEETRNTK  
AHAQTDRLMNLQTLRGYYNQSEASSHTLQWMIGCDLGS DGRLLRGYEQYAYDGKDY LALNEDLRSWTAADTAAQISK  
RKCEANVAEQRRAYLEGTCVEWLHRYLENGKEMLRADPPKTHVTHHPVFDYEATLRCWALGFYPAEIILTWQRD  
GEDQTQDVELVETRPAGDGTQKWA AVVVP SGEEQRYTCHVQHEGLPEPLMLRWKQSSLPTIPIMGIVAGLVVLA  
VVTGA AVAVLWRKKSSD\*(SEQ ID No.1)

[0070] 内源性B2M蛋白质(第一B2M片段)的氨基酸序列:

[0071] MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEK  
VEHSDLSFSKDW SFYLLYYTEFTPTTEKDEYACRVNHVTL SQPKIVKWDRDM (SEQ ID No.2)

[0072] 柔性(G<sub>4</sub>S)<sub>4</sub>连接肽段(第一柔性连接肽段)的氨基酸序列:

[0073] GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer (SEQ ID  
No.3)

[0074] B2M-HLA-G1融合蛋白的氨基酸序列:

[0075] MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEK  
VEHSDLSFSKDW SFYLLYYTEFTPTTEKDEYACRVNHVTL SQPKIVKWDRDMGGGGSGGGGSGGGGSGSHSM  
RYFSAAVSRPGRGEPRIAMGYVDDTQFVRFSDSACPRMEPRAPWVEQEGPEYWEETRNTKAHAQTDRLMNLQTL  
RGYYNQSEASSHTLQWMIGCDLGS DGRLLRGYEQYAYDGKDY LALNEDLRSWTAADTAAQISKRKCEANVAEQRR  
AYLEGTCVEWLHRYLENGKEMLRADPPKTHVTHHPVFDYEATLRCWALGFYPAEIILTWQRDGEDQTQDVELVET  
RPAGDGTQKWA AVVVP SGEEQRYTCHVQHEGLPEPLMLRWKQSSLPTIPIMGIVAGLVVLA AVVTGA AVAVLWR  
KKSSD\*(SEQ ID No.4)

[0076] B2M-HLA-G1融合蛋白的编码序列如下:

atgctcgcctccgtggccttagctgtgctcgcgctactctctcttctggcctggaggctatccagcgtgagctctctctaccctcccgtct  
ggctccttctctcccgtctgcaccctctgtggccctcgtgtgctctctcgtcctgacttccttctccaagttctccttgggtggcccgcc  
gtggggctagtcagggtggatctcggggaagcgggggggtggcctgggagtggggaaggggtgcgcacccgggacgcgcgcta  
cttgccttctggcggggagcaggggagaccttggcctacggcgacgggagggctgggacaaaagtttagggcgtcgataagcgtca  
gagcgcgaggttgggggaggggttctcttccgctcttctcggggcctctggtcctcccagcgcagctggagtgggggacgggtagc  
tcgtcccaaagcgcggcgtgaggttgtgaacgcgtggagggcgttgggtctggggagcgtgcgccggtaagcctgtctg

ctgggctctgcttcccttagactggagagctgtggacttctgtctagggcccgctaagttcgcatgtcttagcactctgggtctatgtgg  
ggccacaccgtggggaggaaacagcagcgcgacgtttgtagaatgcttggctgtgatacaaaagcggttcgaataaactattttgtc  
ccatcacatgtcacttttaaaaaattataagaactaccggtattgacatcttctgtgtccaaggactttatgtgctttgctcatttaatt  
ttgaaaacagttatcttccgccatagataaactactatggttatcttctgcctctcacagatgaagaaactaaggcaccgagatttaagaa  
acttaattacacaggggataaatggcagcaatcgagattgaagcaagcctaaccagggttttggggagcgcgatgcttttggctgt  
aattcgtgcatttttttaagaaaaacgcttctctgctgtgagattctccagagcaaaactggggcgcatgggcccctgtggtcttttctg  
acagagggcttctcttttggctctttgctgggtgttccaagatgtaactgtgctcttactttcgggtttgaaaacatgagggggtggggc  
tgtagcttaacgctgtaatccagcacttagggagggcgagggggaggatggcttgaggtccgtagttgagaccagcctggccaac  
atggtgaagcctggctctcaaaaaataataaaaaaattagccgggtgtggtggctcgtgctgtggteccagctgctccgggtgct  
gagggggaggatctcttgagcttaggctttgagctatcatggcgccagtgcaactccagcgtgggcaacagagcgagaccctgtctct  
caaaaaagaaaaaaaaaaaaaaaaagaaagagaaaaagaaagaaagaaagtgagggttgcagtcagggggagctgtaaaac  
cattaataaagataatccaagatggttaccaagactgttgaggacgccagagatcttgagcactttctaagtaacctggcaatacactaa  
gctgctcactcttctctgcaaaacatgatcgaagcagaatgtttgatcatgagaaaatgcatttaatttgaatacaatttatta  
caacataaaggataatgtatataaccaccacttctggtatttctggttatgttagatgtcatttataaaaaatacaactctgatattta  
aaaaaaaaatcttattttgaaaatttcacaagtaatacatgccatgcatagaccatttctggaagataccacaagaaacatgtaattgatg  
tgctctgaaggctatttctctctgacctgtgtggtgtttgttttcttactgtgggcataaattaatttctcagtttaagtttggaa  
gcttaataaactctcaaaagtcataaagccagtaactgggtgagccaaattcaaaaccagcctgtctgatacttctctctcttagaa  
aagattacagtgatgctctcacaanaatctccgctctcccaaacagagagttccaggcaggatgaatctgtgctctgatccctgagg  
catttaatatgttctattattagaagctcagatgcaagagctctcttagcttttaattgatgaaaaaatcaggtcttcattagatcc  
caatccacctcttgatggggctagtagccttcttaatgatagggtgttctagagagatatctggcaaggtggcctggtactctctc  
ttctccccacagcctcccagacaaggaggagtagctgcttttagtgatcatgtacctgaataaagtgtatttaaaagaattttataca  
catatatttagtgcaatctgtatatttagtagcactaacctctctcttcaatgaaaaatagagttataatatttcttcccact  
ccccatggatggctagtcactctcattttgaaagtaactgtttctgaaacattaggcaatataattccaacctggctagttacagca  
atcactgtggatgctaattaaacgcaaatcccactgtcacatgcacttccattgatcataatggaaagtatgttctgctccatttgc  
catagctctacatccctgtgtattttatcgggtccaactcaaccatttaaggtatttgcagctctgtatgcatttaggtttgtttctt  
gttttttagctcatgaaatagggtacaaagtcagagaggggtctggcataaaaacctcagcagaaataaaggggtttgtgttggta  
agaacataccttgggttgggtgggcacggtggctgctgtaatcccaacactttgggaggccaaggcaggctgatcactgaagtt  
gggagttcaagaccagcctggccaacatgggtgaaatccgctctctactgaaaatacaaaaataaccaggcatggtggtgtgtgctgt  
agtcccaggaaatcactgaaccaggaggcggaggttcagtgagctgagatctcaccactgcacactgcactccagcctgggcaatg  
gaatgagattccatccaaaaataaaaaataaaaaataaagaacataccttgggtgatccacttaggaacctcagataataaca  
tetgcccagtatagagcaattgctatgtcccaggcactctactagacacttcatacagtttagaaaaatcagatgggtgtagatcaaggca  
ggagcaggaacccaaaaagaaaggcataaacataagaaaaaaatggaaggggtgaaacagagtacaataacatgagtaatttg  
atgggggtattatgaaactgagaaatgaaactttgaaaagtatcttggggccaaatcatgtagactcttgagtgatgtttaaggaatgct  
atgagtgctgagagggcatcagaagctctgagagcctccagagaaaggctcttaaaaatgcagcgcgaatctccagtgacagaagat

[0077]

actgctagaatctgctagaaaaaaacaaaaaggcatgtatagaggaattatgagggaaagataccaagtcacggttattcttca  
 aatggagggtgctgttgggaaggtggaagctcatttggccagagtggaatggaattgggagaaatcgatgaccaaatgtaaaca  
 cttggtgctgatatagcttgacaccaagttagcccaagtgaataccctggcaatattaatgtgtcttttccgatattctcaggtactc  
 caaagattcaggttactcacgtcatccagagagaatggaaagtcaaatctgaattgctatgtctgtgggttcaatccacattgaagttg  
 acttactgaagaatggagagagaatgaaaaagtggagcattcagactgtcttcaagcaaggactggcttctatctctgtactacactgaatca  
 cccccactgaaaaagatgagatgccctgccgtggaaccatgtgactttgtcacagcccaagatagttaagtgaggtaagttacattctttgt  
 aagetgctgaaagttgtatgagtagcatatcataaagctgctttgatataaaaaaggtctatggccatactaccctgaatgagtcce  
 atcccatctgatataaacaatctcatattgggattgtcagggaaatgttctaaagatcagattagtgccacctgctgagatactgatgca  
 cagcatgggttctgaaccagtagttccctgcagttgagcagggagcagcagcagcacttgcacaaatacatatacactcttaaacctt  
 ttactactggcttctctagcttttggcagctcaggtatatttagcactgaacgaacatctcaagaaggtataggccttggtttgaag  
 tctctgtctctagatcctataatctggacttctcagacttcttggctggattggtatctgaggetagtaggaagggctgttctctgt  
 gggtagctctaaacaatgtattcatgggtaggaacagcagcctattctgccagcctatttctaccattttagacattttagtacatgg  
 tattttaaagtaaaacttaatgtcttcttttttctccactgtcttttcatagatcgagacatgGGTGGAGGTGGAAGTGG  
 TGAGGTGGAAGTGGTGGAGGTGGAagtGGtgagggtggaagtggCTCCCACTCCATGAGGTATtcc  
 agcgcgccctgtcccggccggccgcggggagccccgctcatcgccatgggctacgtggacgacacgcagttcgtgcggttcgacagc  
 actcggcgtgtccgaggatggagccgcggcgccgtgggtggagcaggagggggccggagtattgggaaggagacacggaacaccaa  
 ggccccgcacagactgacagaatgaacctgcagaccctgcgcggtactacaaccagagcggagccagttctcacacccctcagtgatga  
 ttggctgcgacctgggtccgacggacgctcctccgcgggtatgaacagatgcctacgatggcaaggattacctgcctgaacgaggacc  
 tgcgctcctggaccgcagcggactgcggctcagatcccaagcgaagtgtgagggcggcaatgtgctgaacaaggagagcctacct  
 ggagggcacgtgctggagtggctccacagatactggagaacgggaaggagatgctgcagcgcgcggaccccccaagacacactgga  
 cccaccacctgtctttagctatgagccacctgaggtgctggccctgggcttctaccctgcggagatcatactgacctggcagcgggatgg  
 ggagggaccagaccagcagcgtggagctgctggagaccagcctgcaggggatggaacctccagaagtggcagcgtgtggtggtccttct  
 ggagaggagcagagatacagctgcatgtgcagcatgaggggctgccggagccccctcatgctgagatggaagcagcttccctgccaccat  
 ccccatcatgggtatcgttctggcctgttctcctgcagctgtagtcactggagctgcggtcgtctgtgtGGAGAAAGAAG  
 AGCTCAGATTGA (SEQ ID No. 9)

[0078] 其中,内源性B2M基因外显子DNA序列加粗,内含子DNA序列加下划线

[0080] HLA-G5重链序列(第二B2M片段)的氨基酸序列:

[0081] GSHSMRYFSAAVSRPGRGEPRFIAMGYVDDTQFVRFDSDSACPRMEPRAPWVEQEGPEYWEEETRNTK  
 AHAQTDRMNLQTLRGYYNQSEASSHTLQWMIGCDLGSGRLLRGYEQYAYDQKDYLLNEDLRSWTAADTAAQISK  
 RKCEANVAEQRRAYLEGTVEWLHRYLENGKEMLRADPPKTHVTHHPVFDYEATLRCWALGFYPAEIILTWQRD  
 GEDQTQDVELVETRPAGDGTQKWAAVVPSGEEQRYTCHVQHEGLPEPLMLRWSKEGDGIMSRESRSLSEDL\*  
 (SEQ ID No.6)

[0082] 第二B2M片段的氨基酸序列:

[0083] MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEK  
 VEHSDLSFSKDWSFYLLYYTEFTPTTEKDEYACRVNHVTLSPKIVKWDRDM (SEQ ID No.5)

[0084] 柔性(G<sub>4</sub>S)<sub>4</sub>连接肽段(第二柔性连接肽段)的氨基酸序列:

[0085] GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer (SEQ ID  
 No.7)

[0086] B2M-HLA-G5融合蛋白的氨基酸序列:

[0087] MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLLKNGERIEK  
VEHSDLSFSKDWSFYLLYYTEFTPTTEKDEYACRVNHVTLSPKIVKWDRDMGGGGSGGGGSGGGGSGGGGSGSHSM  
RYFSAAVSRPGRGEPRFIAMGYVDDTQFVRFDSDSACPRMEPRAPWVEQEGPEYWEEETRNTKAHAQTDRMNLQTL  
RGYYNQSEASSHTLQWMIGCDLGS DGRLLRGRYEQYAYDGKDYALNEDLRSWTAADTAAQISKRKCEANVAEQRR  
AYLEGTCVEWLHRYLENGKEMLRADPPKTHVTHHPVFDYEA TLRCWALGFYP AEI ILTWQRDGEDQTQDVELVET  
RPAGDGTQKWA AVVPSGEEQRYTCHVQHEGLPEPLMLRWSKEGDG GIMSVRESRSLSEDL\* (SEQ ID No.8)

[0088] B2M-HLA-G5融合蛋白的编码序列:

[0089] ATGTCTCGCtccgtggccttagctgtgctcgcgctactctctctttctggcctggaggctatccagcg  
tactccaaagattcaggtttactcacgtcatccagcagagaatggaaagtaaatctctgaattgctatgtgtct  
gggtttcatccatccgacattgaagttgacttactgaagaatggagagagaatgaaaaagtgagcattcagact  
tgcttttcagcaaggactggcttttctatctcttgtactacactgaattcacccccactgaaaagatgagtatgc  
ctgccgtgtgaacctgtgactttgtcacagcccaagATAGTTAAGTGGGATCGAGACATGGGTGGAGGTGGAAGT  
GGTGGAGGTGGAAGTGGTGGAGGTGGAAGTGGTGGAGGTGagtgGCTCCCACTCCATGAGGTATttcagcgccg  
ccgtgtcccggcccggccgcccggggagccccgcttcatcgccatgggctacgtggacgacacgcagttcgtgcggtt  
cgacagcgactcggcgtgtccgaggatggagccgcccggcggcgggtgggagcaggagggccggagtattgggaa  
gaggagacacggaacaccaaggcccacgcacagactgacagaatgaacctgcagacctgcgcggtactacaacc  
agagcgaggccagttctcacacctccagtgatgattggctgcgacctgggggtccgacggacgcctcctccgcg  
gtatgaacagtatgcctacgatggcaaggattacctcgccctgaacgaggacctgcgctcctggaccgcagcggac  
actgcggtcagatctccaagcgcaagtgtgaggcggccaatgtggctgaacaaaggagagcctacctggagggca  
cgtgctggagtggtccacagatacctggagaacgggaaggagatgctgcagcgcgagccccccaagacaca  
cgtgacccaccacctgtctttgactatgaggccacctgaggtgctgggcccctgggcttctacctgcgagatc  
atactgacctggcagcgggatggggaggaccagaccaggacgtggagctcgtggagaccaggcctgcaggggatg  
gaaccttcagaagtgggcagctgtggtggtgccttctggagaggagcagagatacacgtgccatgtgcagcatga  
gggctgcccggagccccctcatgctgagatggagtaaggaggagatggaggcatcatgtctgttagggaaagcagg  
agcctctctgaagacCTTTAA (SEQ ID No.10)

[0090] 实施例3

[0091] 基因修饰细胞系中HLA I蛋白表达情况检测

[0092] 为了研究 $\beta 2m$ 和HLA I类蛋白质在野生型(WT),  $B2M^{nu11}$ ,  $B2M^{mHLA G}$ 和 $B2M^{m/sHLA G}$  hESCs细胞系中的表达情况,进行Western Blot和流式细胞分析(FACS)实验。进行流式细胞实验之前,向hESCs细胞培养液中加入IFN- $\gamma$ 至终浓度25ng/ml,培养48小时后,进行流式细胞实验。IFN- $\gamma$ 刺激情况下,野生型,  $B2M^{mHLA G}$ 和 $B2M^{m/sHLA G}$  hESCs在加入IFN- $\gamma$ 处理后,细胞中 $\beta 2m$ 蛋白质在细胞膜表面的表达显著增加(图3A),证明HLA-G1在B2M基因中的整合不会影响内源性B2M的表达。然而,即使在IFN- $\gamma$ 刺激的情况下,在 $B2M^{nu11}$  hESCs中也无法检测到 $\beta 2m$ 蛋白的细胞膜表面表达(图3A)。因此表明,已经成功地通过双gRNA基因编辑策略敲除了 $B2M^{nu11}$  hESCs中的B2M基因。另外,由于将HLA-G1重链阅读框与内源性B2M基因融合在一起,因此 $B2M^{mHLA G}$ 和 $B2M^{m/sHLA G}$ 细胞系中的HLA-G1表达水平也应该受到IFN- $\gamma$ 的影响。在IFN- $\gamma$ 存在的情况下,  $B2M^{mHLA G}$ 和 $B2M^{m/sHLA G}$ 细胞系中HLA-G1的细胞膜表面表达量增加,而在无IFN- $\gamma$ 刺激的条件下,  $B2M^{mHLA G}$ 和 $B2M^{m/sHLA G}$ 细胞系中HLA-G的表达水平均比较低(图3A)。然而,在WT和 $B2M^{nu11}$  hESCs中均未检测到HLA-G的细胞膜表面表达(图3A)。至于经典型HLA I分子的表

达情况,即便在IFN- $\gamma$ 刺激的情况下,在B2M<sup>nu11</sup>,B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs的细胞膜表面未检测到HLA-A,-B或-C(图3A)。表明与野生型hESCs相比,这些细胞中没有可供经典型HLA I重链连接的游离 $\beta$ 2m蛋白,而在B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs中,内源B2M基因位点仅表达B2M-HLA-G1融合蛋白。

[0093] 对于cell lysate组,WT,B2M<sup>mHLA<sub>G</sub></sup>B2M<sup>m/sHLA<sub>G</sub></sup>hESC在RIPA<sup>+</sup>缓冲液中裂解,收集裂解产物用于Western Blot。对于Medium组,将WT,B2M<sup>mHLA<sub>G</sub></sup>B2M<sup>m/sHLA<sub>G</sub></sup>hESC培养至长满孔板,吸弃培养基,加入新鲜的DMEM/F12。24小时后,收集每组的上清液,加入1 $\times$ loading buffer,加入 $\beta$ -巯基乙醇使蛋白质变性,用于Western Blot,检测HLA-G5的抗体为5A6G7(Abcam, ab76869)。Western blot进一步证实B2M<sup>m/sHLA<sub>G</sub></sup>hESCs表达分泌型蛋白质HLA-G5,而WT或B2M<sup>mHLA<sub>G</sub></sup>hESCs中没有HLA-G5蛋白表达或分泌(图3B)。

[0094] 实施例4

[0095] 表达HLA-G1,HLA-G5 hESCs细胞系具有低免疫原性

[0096] 进一步研究B2M<sup>nu11</sup>,B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs细胞系的同种免疫原性。已有研究结果表明,未分化的hESCs与PBMC共培养时不能诱发PBMC增殖(Li et al.,2004;Riolobos et al.,2013)。因此,本实施例中参照Lian et al.,2013中的方法将WT,B2M<sup>nu11</sup>,B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs定向分化为心肌细胞,具体方法如下:

[0097] 1.在消化前一个小时向hESCs培养液中加入Y27632至终浓度10 $\mu$ M。取提前准备好预备接种hESCs细胞的滋养层MEF,吸弃6孔板中的MEF培养液,然后用2ml的DMEM/F12清洗一次,吸弃DMEM/F12,然后加入2ml含有终浓度为10 $\mu$ M的Y27632和bFGF(4ng/ml)的hESCs培养液,放入培养箱中待用。

[0098] 2.吸弃含Y27632的hESCs培养液,加入1ml 0.05%的胰酶,吸弃胰酶后,再加入1ml37 $^{\circ}$ C水浴锅预热过的0.05%胰酶,放入37 $^{\circ}$ C培养箱消化3min,吸弃胰酶,放入37 $^{\circ}$ C培养箱继续消化3min,显微镜下观察hESCs克隆边缘开始出现游离的单个细胞时,停止消化。

[0099] 3.用5ml的DMEM/F12将细胞吹起,使用移液器反复吹打细胞,将消化后的克隆尽可能地都吹成单细胞,将细胞悬液转移到15ml离心管中。

[0100] 4.将离心管800rpm离心30s,细胞团块留在离心管底部,将单细胞悬液转移到一个新的15ml离心管中,2500rpm离心1.5min。

[0101] 5.吸弃上清液,加入1ml hESCs培养液重悬细胞,细胞计数,此时大约每个6孔板孔可以收到1-1.5百万细胞,补加hESCs培养液至10ml,2500rpm离心1.5min。

[0102] 6.将细胞加入到准备好的6孔板中,依次向各个孔中加入1,1.5,2.0,2.5,3million细胞。放入37 $^{\circ}$ C培养箱中继续培养。

[0103] 7.按照培养hESCs的方案正常换液3天。

[0104] 8.吸弃旧培养基,加入含有CHIR99021(终浓度12 $\mu$ M)的RPMI/B-27without insulin溶液,每孔4ml。

[0105] 9.24小时后,吸弃旧培养基,加入RPMI/B-27without insulin溶液,每孔4ml。

[0106] 10.48小时后,从每个孔中吸出2ml培养液,补上等体积新RPMI/B-27without insulin培养液,加入IWP2至终浓度5 $\mu$ M。吸弃孔中剩余的培养基,加入配好的培养基。

[0107] 11.2天后,吸弃旧培养基,加入新的RPMI/B-27without insulin培养液,每孔4ml。

[0108] 12. 2天后,吸弃旧培养基,加入新的RPMI/B-27培养液,每孔4ml。以后每3天换一次液。大概开始分化后5天后,可以看到跳动的心肌细胞。

[0109] 针对心肌特异性基因的qPCR检验和针对TNNT2的染色结果都表明,以上所有细胞系均具有高效分化为心肌细胞的能力(图4B和6C)。

[0110] 同时,使用IFN- $\gamma$ 对心肌细胞进行刺激的具体方法如下:混合培养开始之前,向心肌细胞培养液中加入IFN- $\gamma$ 至终浓度25ng/ml,培养48小时后,收集心肌细胞进行混合培养。仅B2M<sup>mHLA</sup>和B2M<sup>m/sHLA</sup>hESCs分化的心肌细胞表现出在IFN- $\gamma$ 刺激下的HLA-G1表达量升高,同时缺失位于细胞膜表面的经典HLA I类分子表达,与未分化的hESCs相似(图4A)。

[0111] 从健康志愿者处收集到的外周血单个核细胞(PBMC)进行了HLA-A,-B,-C配型检测(n=4)。然后,将供体#01,#02,#03来源的PBMC与分化的心肌细胞共培养,以进行基于羧基荧光素琥珀酰亚胺酯(CFSE)的细胞增殖检测。在混合淋巴细胞反应(MLR)实验中,PBMC在与HLA I不匹配的WT心肌细胞以及PHA (Phytohemagglutinin-L, invitrogen, 00-4977) (阳性对照)共培养时,出现增殖现象,而PBMC与B2M<sup>nu11</sup>, B2M<sup>mHLA</sup>和B2M<sup>m/sHLA</sup>分化的心肌细胞共培养时,其增殖现象明显减弱(图5A),共培养实验方案具体如下:

[0112] 1.取 $3.5 \times 10^6$ 分化15天以上的心肌细胞(部分心肌细胞进行此步骤前,体外培养过程中加入IFN- $\gamma$ 处理48小时,最终MLR实验中作为+IFN- $\gamma$ 组的靶细胞,MLR实验即图5A所示实验中,+IFN- $\gamma$ 组的心肌细胞进行了IFN- $\gamma$ 处理,no IFN- $\gamma$ 组的心肌细胞,未进行相关处理),加入胰蛋白酶消化3min,加入等体积RPMI-5中和胰蛋白酶,移液管吹打成单细胞,2500rpm离心1.5min。

[0113] 2.吸弃上清后加入5ml RPMI-5培养液重悬细胞,离心管缠上封口膜后放入X射线辐照仪中进行辐照,辐照剂量为20Gy。

[0114] 3.辐照后的细胞用5ml RPMI清洗两次,每次洗完后2500rpm离心1.5min。

[0115] 4.使用1ml RPMI-5重悬心肌细胞,计数,加入RPMI-5将细胞密度调整为 $1 \times 10^6$ cells/ml备用。

[0116] 5.300g,10min离心收集好的PBMC,使用PBS/0.1%HSA重悬PBMC,将细胞密度调整为 $1-10 \times 10^6$ cells/ml。

[0117] 6.在避光条件下,将CFSE溶液用PBS稀释成6 $\mu$ M终浓度。

[0118] 7.向PBMC溶液中加入等体积CFSE溶液,混匀后避光放入37 $^{\circ}$ C培养箱,孵育10min。

[0119] 8.孵育结束后,向PBMC溶液中加入6倍体积的4 $^{\circ}$ C预冷的RPMI-1640/10%HSA溶液,混匀后室温静置5min。

[0120] 9.21 $^{\circ}$ C800g离心15min。

[0121] 10.弃上清,使用RPMI-5重悬PBMC,将细胞密度调整为 $1 \times 10^6$ cells/ml。

[0122] 11.将PBMC混匀后,加入到圆底96孔板孔中,每孔加100 $\mu$ l。

[0123] 12.将辐照过的心肌细胞加入到已加入PBMC的96孔板孔内,每孔加100 $\mu$ l,此时靶细胞与效应细胞的比例为1:1。混匀孔内细胞。阳性对照组的孔内加入1 $\mu$ g/ml终浓度的PHA。

[0124] 13.将96孔板放入37 $^{\circ}$ C培养箱内,培养7天。

[0125] 14.培养结束后,收集96孔板孔内的细胞,每孔细胞加入1ml PBS混匀,300g离心10min。

[0126] 15.吸弃上清,加入100 $\mu$ l流式细胞检测固定液重悬细胞,加入对应浓度的Human

Fc Block抗体冰上孵育10min,加入对应浓度的anti-human CD45抗体,冰上孵育30min。

[0127] 16.加入2ml PBS重悬细胞,300g离心10min。

[0128] 17.吸弃上清液,加入500 $\mu$ l PBS重悬细胞,流式细胞仪检测PBMC增殖情况。

[0129] 进一步进行ELISA实验,以检测与WT, B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>心肌细胞体外共培养48小时后PBMC(供体#01, #02, #04)的IFN- $\gamma$ 分泌情况。如图5B所示, WT心肌细胞刺激PBMC大量分泌IFN- $\gamma$ , 尤其是在测定IFN- $\gamma$ 之前, 用IFN- $\gamma$ 刺激WT心肌48h时的情况下, 当PBMC与B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>心肌细胞共培养时, 几乎无法检测到分泌的IFN- $\gamma$ 。

[0130] 由于HLA-G是表达在NK细胞上的LILRB1和KIR2DL4抑制型受体的配体, 所以HLA-G表达应该可以抑制NK的活化。将NK-92细胞与WT, B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs孵育, 以评估IFN- $\gamma$ 的释放。结果显示, B2M<sup>nu11</sup>hESCs可以诱发大量的IFN- $\gamma$ 释放(图5C), 表明细胞表面缺失HLA I类分子会引起NK细胞攻击。与B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs共培养的NK-92细胞明显减少了IFN- $\gamma$ 的分泌, 而同时表达HLA-G1和HLA-G5的B2M<sup>m/sHLA<sub>G</sub></sup>hESCs诱发的IFN- $\gamma$ 释放量甚至低于B2M<sup>mHLA<sub>G</sub></sup>hESCs的结果(图5C), 这些结果证明HLA-G表达可抑制NK细胞活化。

[0131] 由于HLA-G5是可分泌的蛋白质, 所以进一步研究B2M<sup>m/sHLA<sub>G</sub></sup>hESCs来源的条件培养基的免疫抑制特性。将B2M<sup>m/sHLA<sub>G</sub></sup>细胞传代后收集, 使用RPMI-1640重悬后调整细胞密度为10<sup>7</sup> cells/ml, 37 $^{\circ}$ C培养24小时, 然后收集来自B2M<sup>m/sHLA<sub>G</sub></sup>hESCs的上清液作为条件培养基, 并将NK-92细胞, WT hESCs分别在RPMI-1640或B2M<sup>m/sHLA<sub>G</sub></sup>hESCs条件培养基中共培养, 以评估在各自环境中IFN- $\gamma$ 的释放。实验结果显示, B2M<sup>m/sHLA<sub>G</sub></sup>hESCs条件培养基中的IFN- $\gamma$ 浓度要低得多(图5D), 这表明B2M<sup>m/sHLA<sub>G</sub></sup>hESCs分泌的HLA-G5可溶性蛋白质可抑制NK细胞介导的IFN- $\gamma$ 分泌。

[0132] 综上所述, HLA I类分子缺失细胞诱发的同种异体MLR反应减弱, 这些结果还表明, 表达HLA-G蛋白质的细胞在体外实验中表现出低免疫原性。

[0133] 使用体内模型检验PBMCs介导的针对B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>, B2M<sup>m/sHLA<sub>G</sub></sup>细胞的同种异体免疫反应。将WT和基因编辑过的hESCs(WT和B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs)注射到M-NSG小鼠皮下以形成畸胎瘤(M-NSG小鼠购买自上海南方模式生物科技股份有限公司)。将人PBMC与由野生型H9 ESC分化得来的拟胚体共培养14天。然后, 通过尾静脉注射将被H9 ESC抗原激活过的同种异体PBMC转移进小鼠体内, 具体方案如下: 预激活的3.4 $\times$ 10<sup>6</sup> PBMCs使用100 $\mu$ l PBS重悬, 使用胰岛素针将细胞悬液通过尾静脉注射到每只小鼠体内, 给对照组小鼠注射等体积PBS。20天后, 处死小鼠并收集畸胎瘤用于免疫染色(图6A)。免疫荧光染色结果表明, 浸润到由B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESC分化产生畸胎瘤的人类T细胞数量相比起野生型组大大减少(图6B, 图6C)。针对畸胎瘤内人IL-2的qPCR分析反映了相同的结论(图6D), 表明B2M<sup>nu11</sup>, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESCs分化产生的畸胎瘤受到保护, 免受同种异体T细胞介导的免疫排斥。然而, 尽管在野生型, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>hESC来源的畸胎瘤中检测到很少的NK细胞浸润, 但在源自B2M<sup>nu11</sup>hESC的畸胎瘤中却检测到显著增强的NK细胞浸润(图6B, 图6C)。这些结果表明, B2M<sup>nu11</sup>hESC来源的畸胎瘤细胞膜表面HLAI分子表达的丧失引发了NK细胞的攻击。但是, B2M<sup>mHLA<sub>G</sub></sup>和B2M<sup>m/sHLA<sub>G</sub></sup>来源的细胞畸胎瘤表达的HLA-G蛋白有效地阻断了同种异体NK细胞介导的排斥。

[0134] 综上所述, 本发明有效克服了现有技术中的种种缺点而具高度产业利用价值。

[0135] 上述实施例仅例示性说明本发明的原理及其功效, 而非用于限制本发明。任何熟



悉此技术的人士皆可在不违背本发明的精神及范畴下,对上述实施例进行修饰或改变。因此,举凡所属技术领域中具有通常知识者在未脱离本发明所揭示的精神与技术思想下所完成的一切等效修饰或改变,仍应由本发明的权利要求所涵盖。

[0001] 序列表  
 [0002] <110> 同济大学  
 [0003] <120> 一种人胚胎干细胞的构建方法  
 [0004] <160> 14  
 [0005] <170> SIPOSequenceListing 1.0  
 [0006] <210> 1  
 [0007] <211> 314  
 [0008] <212> PRT  
 [0009] <213> 人工序列(Artificial Sequence)  
 [0010] <400> 1  
 [0011] Gly Ser His Ser Met Arg Tyr Phe Ser Ala Ala Val Ser Arg Pro Gly  
 [0012] 1 5 10 15  
 [0013] Arg Gly Glu Pro Arg Phe Ile Ala Met Gly Tyr Val Asp Asp Thr Gln  
 [0014] 20 25 30  
 [0015] Phe Val Arg Phe Asp Ser Asp Ser Ala Cys Pro Arg Met Glu Pro Arg  
 [0016] 35 40 45  
 [0017] Ala Pro Trp Val Glu Gln Glu Gly Pro Glu Tyr Trp Glu Glu Glu Thr  
 [0018] 50 55 60  
 [0019] Arg Asn Thr Lys Ala His Ala Gln Thr Asp Arg Met Asn Leu Gln Thr  
 [0020] 65 70 75 80  
 [0021] Leu Arg Gly Tyr Tyr Asn Gln Ser Glu Ala Ser Ser His Thr Leu Gln  
 [0022] 85 90 95  
 [0023] Trp Met Ile Gly Cys Asp Leu Gly Ser Asp Gly Arg Leu Leu Arg Gly  
 [0024] 100 105 110  
 [0025] Tyr Glu Gln Tyr Ala Tyr Asp Gly Lys Asp Tyr Leu Ala Leu Asn Glu  
 [0026] 115 120 125  
 [0027] Asp Leu Arg Ser Trp Thr Ala Ala Asp Thr Ala Ala Gln Ile Ser Lys  
 [0028] 130 135 140  
 [0029] Arg Lys Cys Glu Ala Ala Asn Val Ala Glu Gln Arg Arg Ala Tyr Leu  
 [0030] 145 150 155 160  
 [0031] Glu Gly Thr Cys Val Glu Trp Leu His Arg Tyr Leu Glu Asn Gly Lys  
 [0032] 165 170 175  
 [0033] Glu Met Leu Gln Arg Ala Asp Pro Pro Lys Thr His Val Thr His His  
 [0034] 180 185 190  
 [0035] Pro Val Phe Asp Tyr Glu Ala Thr Leu Arg Cys Trp Ala Leu Gly Phe  
 [0036] 195 200 205  
 [0037] Tyr Pro Ala Glu Ile Ile Leu Thr Trp Gln Arg Asp Gly Glu Asp Gln  
 [0038] 210 215 220  
 [0039] Thr Gln Asp Val Glu Leu Val Glu Thr Arg Pro Ala Gly Asp Gly Thr  
 [0040] 225 230 235 240  
 [0041] Phe Gln Lys Trp Ala Ala Val Val Val Pro Ser Gly Glu Glu Gln Arg

[0042]		245		250		255
[0043]	Tyr Thr Cys His Val Gln His Glu Gly Leu Pro Glu Pro Leu Met Leu					
[0044]		260		265		270
[0045]	Arg Trp Lys Gln Ser Ser Leu Pro Thr Ile Pro Ile Met Gly Ile Val					
[0046]		275		280		285
[0047]	Ala Gly Leu Val Val Leu Ala Ala Val Val Thr Gly Ala Ala Val Ala					
[0048]		290		295		300
[0049]	Ala Val Leu Trp Arg Lys Lys Ser Ser Asp					
[0050]	305		310			
[0051]	<210> 2					
[0052]	<211> 119					
[0053]	<212> PRT					
[0054]	<213> 人工序列(Artificial Sequence)					
[0055]	<400> 2					
[0056]	Met Ser Arg Ser Val Ala Leu Ala Val Leu Ala Leu Leu Ser Leu Ser					
[0057]	1	5		10		15
[0058]	Gly Leu Glu Ala Ile Gln Arg Thr Pro Lys Ile Gln Val Tyr Ser Arg					
[0059]		20		25		30
[0060]	His Pro Ala Glu Asn Gly Lys Ser Asn Phe Leu Asn Cys Tyr Val Ser					
[0061]		35		40		45
[0062]	Gly Phe His Pro Ser Asp Ile Glu Val Asp Leu Leu Lys Asn Gly Glu					
[0063]		50		55		60
[0064]	Arg Ile Glu Lys Val Glu His Ser Asp Leu Ser Phe Ser Lys Asp Trp					
[0065]	65		70		75	80
[0066]	Ser Phe Tyr Leu Leu Tyr Tyr Thr Glu Phe Thr Pro Thr Glu Lys Asp					
[0067]		85		90		95
[0068]	Glu Tyr Ala Cys Arg Val Asn His Val Thr Leu Ser Gln Pro Lys Ile					
[0069]		100		105		110
[0070]	Val Lys Trp Asp Arg Asp Met					
[0071]		115				
[0072]	<210> 3					
[0073]	<211> 20					
[0074]	<212> PRT					
[0075]	<213> 人工序列(Artificial Sequence)					
[0076]	<400> 3					
[0077]	Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly					
[0078]	1	5		10		15
[0079]	Gly Gly Gly Ser					
[0080]		20				
[0081]	<210> 4					
[0082]	<211> 453					
[0083]	<212> PRT					

[0084]	<213> 人工序列(Artificial Sequence)															
[0085]	<400> 4															
[0086]	Met	Ser	Arg	Ser	Val	Ala	Leu	Ala	Val	Leu	Ala	Leu	Leu	Ser	Leu	Ser
[0087]	1			5					10					15		
[0088]	Gly	Leu	Glu	Ala	Ile	Gln	Arg	Thr	Pro	Lys	Ile	Gln	Val	Tyr	Ser	Arg
[0089]				20					25					30		
[0090]	His	Pro	Ala	Glu	Asn	Gly	Lys	Ser	Asn	Phe	Leu	Asn	Cys	Tyr	Val	Ser
[0091]				35					40					45		
[0092]	Gly	Phe	His	Pro	Ser	Asp	Ile	Glu	Val	Asp	Leu	Leu	Lys	Asn	Gly	Glu
[0093]			50					55					60			
[0094]	Arg	Ile	Glu	Lys	Val	Glu	His	Ser	Asp	Leu	Ser	Phe	Ser	Lys	Asp	Trp
[0095]	65					70					75				80	
[0096]	Ser	Phe	Tyr	Leu	Leu	Tyr	Tyr	Thr	Glu	Phe	Thr	Pro	Thr	Glu	Lys	Asp
[0097]						85					90				95	
[0098]	Glu	Tyr	Ala	Cys	Arg	Val	Asn	His	Val	Thr	Leu	Ser	Gln	Pro	Lys	Ile
[0099]				100						105				110		
[0100]	Val	Lys	Trp	Asp	Arg	Asp	Met	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly
[0101]				115						120				125		
[0102]	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Ser	His	Ser	Met
[0103]			130										140			
[0104]	Arg	Tyr	Phe	Ser	Ala	Ala	Val	Ser	Arg	Pro	Gly	Arg	Gly	Glu	Pro	Arg
[0105]	145					150						155			160	
[0106]	Phe	Ile	Ala	Met	Gly	Tyr	Val	Asp	Asp	Thr	Gln	Phe	Val	Arg	Phe	Asp
[0107]						165						170			175	
[0108]	Ser	Asp	Ser	Ala	Cys	Pro	Arg	Met	Glu	Pro	Arg	Ala	Pro	Trp	Val	Glu
[0109]				180							185				190	
[0110]	Gln	Glu	Gly	Pro	Glu	Tyr	Trp	Glu	Glu	Glu	Thr	Arg	Asn	Thr	Lys	Ala
[0111]				195							200				205	
[0112]	His	Ala	Gln	Thr	Asp	Arg	Met	Asn	Leu	Gln	Thr	Leu	Arg	Gly	Tyr	Tyr
[0113]				210										220		
[0114]	Asn	Gln	Ser	Glu	Ala	Ser	Ser	His	Thr	Leu	Gln	Trp	Met	Ile	Gly	Cys
[0115]	225					230							235		240	
[0116]	Asp	Leu	Gly	Ser	Asp	Gly	Arg	Leu	Leu	Arg	Gly	Tyr	Glu	Gln	Tyr	Ala
[0117]						245						250			255	
[0118]	Tyr	Asp	Gly	Lys	Asp	Tyr	Leu	Ala	Leu	Asn	Glu	Asp	Leu	Arg	Ser	Trp
[0119]						260								270		
[0120]	Thr	Ala	Ala	Asp	Thr	Ala	Ala	Gln	Ile	Ser	Lys	Arg	Lys	Cys	Glu	Ala
[0121]				275										285		
[0122]	Ala	Asn	Val	Ala	Glu	Gln	Arg	Arg	Ala	Tyr	Leu	Glu	Gly	Thr	Cys	Val
[0123]				290										300		
[0124]	Glu	Trp	Leu	His	Arg	Tyr	Leu	Glu	Asn	Gly	Lys	Glu	Met	Leu	Gln	Arg
[0125]	305					310								315		320

[0126]	Ala Asp Pro Pro Lys Thr His Val Thr His His Pro Val Phe Asp Tyr
[0127]	325 330 335
[0128]	Glu Ala Thr Leu Arg Cys Trp Ala Leu Gly Phe Tyr Pro Ala Glu Ile
[0129]	340 345 350
[0130]	Ile Leu Thr Trp Gln Arg Asp Gly Glu Asp Gln Thr Gln Asp Val Glu
[0131]	355 360 365
[0132]	Leu Val Glu Thr Arg Pro Ala Gly Asp Gly Thr Phe Gln Lys Trp Ala
[0133]	370 375 380
[0134]	Ala Val Val Val Pro Ser Gly Glu Glu Gln Arg Tyr Thr Cys His Val
[0135]	385 390 395 400
[0136]	Gln His Glu Gly Leu Pro Glu Pro Leu Met Leu Arg Trp Lys Gln Ser
[0137]	405 410 415
[0138]	Ser Leu Pro Thr Ile Pro Ile Met Gly Ile Val Ala Gly Leu Val Val
[0139]	420 425 430
[0140]	Leu Ala Ala Val Val Thr Gly Ala Ala Val Ala Ala Val Leu Trp Arg
[0141]	435 440 445
[0142]	Lys Lys Ser Ser Asp
[0143]	450
[0144]	<210> 5
[0145]	<211> 119
[0146]	<212> PRT
[0147]	<213> 人工序列(Artificial Sequence)
[0148]	<400> 5
[0149]	Met Ser Arg Ser Val Ala Leu Ala Val Leu Ala Leu Leu Ser Leu Ser
[0150]	1 5 10 15
[0151]	Gly Leu Glu Ala Ile Gln Arg Thr Pro Lys Ile Gln Val Tyr Ser Arg
[0152]	20 25 30
[0153]	His Pro Ala Glu Asn Gly Lys Ser Asn Phe Leu Asn Cys Tyr Val Ser
[0154]	35 40 45
[0155]	Gly Phe His Pro Ser Asp Ile Glu Val Asp Leu Leu Lys Asn Gly Glu
[0156]	50 55 60
[0157]	Arg Ile Glu Lys Val Glu His Ser Asp Leu Ser Phe Ser Lys Asp Trp
[0158]	65 70 75 80
[0159]	Ser Phe Tyr Leu Leu Tyr Tyr Thr Glu Phe Thr Pro Thr Glu Lys Asp
[0160]	85 90 95
[0161]	Glu Tyr Ala Cys Arg Val Asn His Val Thr Leu Ser Gln Pro Lys Ile
[0162]	100 105 110
[0163]	Val Lys Trp Asp Arg Asp Met
[0164]	115
[0165]	<210> 6
[0166]	<211> 295
[0167]	<212> PRT

[0168] <213> 人工序列(Artificial Sequence)  
 [0169] <400> 6  
 [0170] Gly Ser His Ser Met Arg Tyr Phe Ser Ala Ala Val Ser Arg Pro Gly  
 [0171] 1 5 10 15  
 [0172] Arg Gly Glu Pro Arg Phe Ile Ala Met Gly Tyr Val Asp Asp Thr Gln  
 [0173] 20 25 30  
 [0174] Phe Val Arg Phe Asp Ser Asp Ser Ala Cys Pro Arg Met Glu Pro Arg  
 [0175] 35 40 45  
 [0176] Ala Pro Trp Val Glu Gln Glu Gly Pro Glu Tyr Trp Glu Glu Glu Thr  
 [0177] 50 55 60  
 [0178] Arg Asn Thr Lys Ala His Ala Gln Thr Asp Arg Met Asn Leu Gln Thr  
 [0179] 65 70 75 80  
 [0180] Leu Arg Gly Tyr Tyr Asn Gln Ser Glu Ala Ser Ser His Thr Leu Gln  
 [0181] 85 90 95  
 [0182] Trp Met Ile Gly Cys Asp Leu Gly Ser Asp Gly Arg Leu Leu Arg Gly  
 [0183] 100 105 110  
 [0184] Tyr Glu Gln Tyr Ala Tyr Asp Gly Lys Asp Tyr Leu Ala Leu Asn Glu  
 [0185] 115 120 125  
 [0186] Asp Leu Arg Ser Trp Thr Ala Ala Asp Thr Ala Ala Gln Ile Ser Lys  
 [0187] 130 135 140  
 [0188] Arg Lys Cys Glu Ala Ala Asn Val Ala Glu Gln Arg Arg Ala Tyr Leu  
 [0189] 145 150 155 160  
 [0190] Glu Gly Thr Cys Val Glu Trp Leu His Arg Tyr Leu Glu Asn Gly Lys  
 [0191] 165 170 175  
 [0192] Glu Met Leu Gln Arg Ala Asp Pro Pro Lys Thr His Val Thr His His  
 [0193] 180 185 190  
 [0194] Pro Val Phe Asp Tyr Glu Ala Thr Leu Arg Cys Trp Ala Leu Gly Phe  
 [0195] 195 200 205  
 [0196] Tyr Pro Ala Glu Ile Ile Leu Thr Trp Gln Arg Asp Gly Glu Asp Gln  
 [0197] 210 215 220  
 [0198] Thr Gln Asp Val Glu Leu Val Glu Thr Arg Pro Ala Gly Asp Gly Thr  
 [0199] 225 230 235 240  
 [0200] Phe Gln Lys Trp Ala Ala Val Val Val Pro Ser Gly Glu Glu Gln Arg  
 [0201] 245 250 255  
 [0202] Tyr Thr Cys His Val Gln His Glu Gly Leu Pro Glu Pro Leu Met Leu  
 [0203] 260 265 270  
 [0204] Arg Trp Ser Lys Glu Gly Asp Gly Gly Ile Met Ser Val Arg Glu Ser  
 [0205] 275 280 285  
 [0206] Arg Ser Leu Ser Glu Asp Leu  
 [0207] 290 295  
 [0208] <210> 7  
 [0209] <211> 20

[0210]	<212>	PRT
[0211]	<213>	人工序列(Artificial Sequence)
[0212]	<400>	7
[0213]	Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly	
[0214]	1	5 10 15
[0215]	Gly Gly Gly Ser	
[0216]		20
[0217]	<210>	8
[0218]	<211>	434
[0219]	<212>	PRT
[0220]	<213>	人工序列(Artificial Sequence)
[0221]	<400>	8
[0222]	Met Ser Arg Ser Val Ala Leu Ala Val Leu Ala Leu Leu Ser Leu Ser	
[0223]	1	5 10 15
[0224]	Gly Leu Glu Ala Ile Gln Arg Thr Pro Lys Ile Gln Val Tyr Ser Arg	
[0225]		20 25 30
[0226]	His Pro Ala Glu Asn Gly Lys Ser Asn Phe Leu Asn Cys Tyr Val Ser	
[0227]		35 40 45
[0228]	Gly Phe His Pro Ser Asp Ile Glu Val Asp Leu Leu Lys Asn Gly Glu	
[0229]		50 55 60
[0230]	Arg Ile Glu Lys Val Glu His Ser Asp Leu Ser Phe Ser Lys Asp Trp	
[0231]		65 70 75 80
[0232]	Ser Phe Tyr Leu Leu Tyr Tyr Thr Glu Phe Thr Pro Thr Glu Lys Asp	
[0233]		85 90 95
[0234]	Glu Tyr Ala Cys Arg Val Asn His Val Thr Leu Ser Gln Pro Lys Ile	
[0235]		100 105 110
[0236]	Val Lys Trp Asp Arg Asp Met Gly Gly Gly Gly Ser Gly Gly Gly Gly	
[0237]		115 120 125
[0238]	Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Ser His Ser Met	
[0239]		130 135 140
[0240]	Arg Tyr Phe Ser Ala Ala Val Ser Arg Pro Gly Arg Gly Glu Pro Arg	
[0241]		145 150 155 160
[0242]	Phe Ile Ala Met Gly Tyr Val Asp Asp Thr Gln Phe Val Arg Phe Asp	
[0243]		165 170 175
[0244]	Ser Asp Ser Ala Cys Pro Arg Met Glu Pro Arg Ala Pro Trp Val Glu	
[0245]		180 185 190
[0246]	Gln Glu Gly Pro Glu Tyr Trp Glu Glu Glu Thr Arg Asn Thr Lys Ala	
[0247]		195 200 205
[0248]	His Ala Gln Thr Asp Arg Met Asn Leu Gln Thr Leu Arg Gly Tyr Tyr	
[0249]		210 215 220
[0250]	Asn Gln Ser Glu Ala Ser Ser His Thr Leu Gln Trp Met Ile Gly Cys	
[0251]		225 230 235 240

[0252]	Asp Leu Gly Ser Asp Gly Arg Leu Leu Arg Gly Tyr Glu Gln Tyr Ala
[0253]	245 250 255
[0254]	Tyr Asp Gly Lys Asp Tyr Leu Ala Leu Asn Glu Asp Leu Arg Ser Trp
[0255]	260 265 270
[0256]	Thr Ala Ala Asp Thr Ala Ala Gln Ile Ser Lys Arg Lys Cys Glu Ala
[0257]	275 280 285
[0258]	Ala Asn Val Ala Glu Gln Arg Arg Ala Tyr Leu Glu Gly Thr Cys Val
[0259]	290 295 300
[0260]	Glu Trp Leu His Arg Tyr Leu Glu Asn Gly Lys Glu Met Leu Gln Arg
[0261]	305 310 315 320
[0262]	Ala Asp Pro Pro Lys Thr His Val Thr His His Pro Val Phe Asp Tyr
[0263]	325 330 335
[0264]	Glu Ala Thr Leu Arg Cys Trp Ala Leu Gly Phe Tyr Pro Ala Glu Ile
[0265]	340 345 350
[0266]	Ile Leu Thr Trp Gln Arg Asp Gly Glu Asp Gln Thr Gln Asp Val Glu
[0267]	355 360 365
[0268]	Leu Val Glu Thr Arg Pro Ala Gly Asp Gly Thr Phe Gln Lys Trp Ala
[0269]	370 375 380
[0270]	Ala Val Val Val Pro Ser Gly Glu Glu Gln Arg Tyr Thr Cys His Val
[0271]	385 390 395 400
[0272]	Gln His Glu Gly Leu Pro Glu Pro Leu Met Leu Arg Trp Ser Lys Glu
[0273]	405 410 415
[0274]	Gly Asp Gly Gly Ile Met Ser Val Arg Glu Ser Arg Ser Leu Ser Glu
[0275]	420 425 430
[0276]	Asp Leu
[0277]	<210> 9
[0278]	<211> 5798
[0279]	<212> DNA
[0280]	<213> 人工序列 (Artificial Sequence)
[0281]	<400> 9
[0282]	atgtctcgct ccgtggcctt agctgtgctc gcgctactct ctctttctgg cctggaggct 60
[0283]	atccagcgtg agtctctcct accctcccgc tctggtcctt cctctcccgc tctgcacct 120
[0284]	ctgtggccct cgctgtgctc tctcgctccg tgacttcctt tctccaagtt ctcttggtg 180
[0285]	gcccgcctg gggctagtcc agggctggat ctcggggaag cggcggggtg gcctgggagt 240
[0286]	ggggaagggg gtgcgcaccc gggacgcgcg ctacttgccc ctttcggcgg ggagcagggg 300
[0287]	agaccttggt cctacggcga cgggagggtc gggacaaagt ttagggcgtc gataagcgtc 360
[0288]	agagcggcga ggttggggga gggtttctct tccgctcttt cgcggggcct ctggctcccc 420
[0289]	cagcgcagct ggagtggggg acgggtaggc tcgtcccaaa ggcgcggcgc tgaggtttgt 480
[0290]	gaacgcgtgg aggggcgctt ggggtctggg ggaggcgtcg cccgggtaag cctgtctgct 540
[0291]	gcggctctgc ttcccttaga ctggagagct gtggacttcg tctaggcgcc cgctaagttc 600
[0292]	gcatgtccta gcacctctgg gtctatgtgg ggccacaccg tggggaggaa acagcagcgc 660
[0293]	acgtttgtag aatgcttggc tgtgatacaa agcggtttcg aataattaac ttatttgttc 720



[0294]	ccatcacatg	tcacttttaa	aaaattataa	gaactaccg	ttattgacat	ctttctgtgt	780
[0295]	gccaaggact	ttatgtgctt	tgcgtcattt	aattttgaaa	acagttatct	tccgccatag	840
[0296]	ataactacta	tggttatctt	ctgcctctca	cagatgaaga	aactaaggca	cagagatfff	900
[0297]	aagaaactta	attacacagg	ggataaatgg	cagcaatcga	gattgaagtc	aagcctaacc	960
[0298]	agggtctttg	cgggagcgca	tgccttttgg	ctgtaattcg	tgcatttttt	tttaagaaaa	1020
[0299]	acgcctgcct	tctgcgtgag	attctccaga	gcaaaactgg	cggcatgggc	cctgtggtct	1080
[0300]	tttcttacag	agggtctcct	ctttggetct	ttgcctgggt	gtttccaaga	tgtactgtgc	1140
[0301]	ctcttacttt	cggttttgaa	aacatgaggg	ggttgggctg	ggtagcttac	gcctgtaate	1200
[0302]	ccagcactta	gggaggccga	ggcgggagga	tggcttgagg	tccgtagttg	agaccagcct	1260
[0303]	ggccaacatg	gtgaagcctg	gtctctacaa	aaaataataa	caaaaattag	cgggtgtggt	1320
[0304]	tggctcgtgc	ctgtggctcc	agctgctccg	gtggctgagg	cgggaggatc	tcttgagctt	1380
[0305]	aggcttttga	gctatcatgg	cgccagtgca	ctccagcgtg	ggcaacagag	cgagaccctg	1440
[0306]	tctctcaaaa	aagaaaaaaa	aaaaaaaaaga	aagagaaaag	aaaagaaaaga	aagaagttaa	1500
[0307]	ggtttgtcag	tcaggggagc	tgtaaaacca	ttaataaaga	taatccaaga	tggttaccaa	1560
[0308]	gactgttgag	gacgccagag	atcttgagca	ctttctaagt	acctggcaat	acactaagcg	1620
[0309]	cgctcacctt	ttctctggtc	aaaacatgat	cgaaagcaga	atgttttgat	catgagaaaa	1680
[0310]	ttgcatftaa	tttgaataca	atftatftac	aacataaagg	ataatgtata	tatcaccacc	1740
[0311]	attactggta	tttgcctggt	atgttagatg	tcattfttaa	aaataacaat	ctgatattta	1800
[0312]	aaaaaaaaatc	ttatfttgaa	aatfttcaaaa	gtaatacatg	ccatgcatag	accatftctg	1860
[0313]	gaagatacca	caagaaacat	gtaatgatga	ttgcctctga	aggctctatft	tctctctctg	1920
[0314]	acctgtgtgt	gggtfttgtt	tttgtftftac	tgtgggcata	aatftaattft	tcagfttaagt	1980
[0315]	tttggaaact	taaataactc	tccaaaagtc	ataaagccag	taactggttg	agcccaaatt	2040
[0316]	caaaccagc	ctgtctgata	cttgcctctc	tcttagaaaa	gattacagtg	atgctctcac	2100
[0317]	aaaatcttgc	cgcttccct	caaacagaga	gttccaggca	ggatgaatct	gtgctctgat	2160
[0318]	ccctgaggca	tttaatatgt	tcttattatt	agaagctcag	atgcaaagag	ctctcttagc	2220
[0319]	ttftaatgft	atgaaaaaaa	tcaggtcttc	attagattcc	ccaatccacc	tcttgatggg	2280
[0320]	gctagtagcc	tttcttaat	gatagggtgt	ttctagagag	atatactctg	tcaaggtggc	2340
[0321]	ctggtactcc	tcttctccc	cacagcctcc	cagacaagga	ggagtagctg	cctftfttagtg	2400
[0322]	atcatgtacc	ctgaatataa	gtgtatftta	aagaatftta	tacacatata	tttagtgtca	2460
[0323]	atctgtatat	ttagtagcac	taaaccttct	cttcatfttc	aatgaaaaat	atagagftta	2520
[0324]	taatatfttc	ttcccacttc	cccatggatg	gtctagtcac	gcctctcatt	ttggaagta	2580
[0325]	ctgtftctga	aacattagcc	aatatattcc	caacctggct	agftttacagc	aatcacctgt	2640
[0326]	ggatgctaat	taaaacgcaa	atcccactgt	cacatgcatt	actccatttg	atcataatgg	2700
[0327]	aaagtatgft	ctgtcccatt	tgccatagtc	ctcacctatc	cctgttgtat	tttatcgggt	2760
[0328]	ccaactcaac	cattftaaggt	atfttgcagc	tcttgtatgc	atfttaggtft	tgtfttctftg	2820
[0329]	ttftfttagct	catgaaatta	ggtacaaaagt	cagagagggg	tctggcatat	aaaacctcag	2880
[0330]	cagaaataaa	gaggtfttgt	tgtfttgtaa	gaacatacct	tgggttggft	gggcacgggt	2940
[0331]	gctcgtgcct	gtaatcccaa	cactfttggga	ggccaaggca	ggctgatcac	ttgaagfttg	3000
[0332]	gagtcaaga	ccagcctggc	caacatgggt	aaatcccgtc	tctactgaaa	atacaaaaaat	3060
[0333]	taaccaggca	tgggtggtgt	tgcctgtagt	cccaggaatc	acttgaacc	aggaggcgga	3120
[0334]	ggttgcagtg	agctgagatc	tcaccactgc	acactgcact	ccagcctggg	caatggaatg	3180
[0335]	agattccatc	ccaaaaaata	aaaaaataaa	aaaataaaga	acataccttg	ggttgatcca	3240

[0336]	cttaggaacc	tcagataata	acatctgcc	cgtatagagc	aattgctatg	tcccaggcac	3300
[0337]	tctactagac	acttcataca	gtttagaaaa	tcagatgggt	gtagatcaag	gcaggagcag	3360
[0338]	gaacccaaaa	gaaaggcata	aacataagaa	aaaaaatgga	aggggtggaa	acagagtaca	3420
[0339]	ataacatgag	taatttgatg	ggggctatta	tgaactgaga	aatgaacttt	gaaaagtatc	3480
[0340]	ttggggccaa	atcatgtaga	ctcttgagtg	atgtgttaag	gaatgctatg	agtgtctgaga	3540
[0341]	gggcatcaga	agtccttgag	agcctccaga	gaaaggctct	taaaaatgca	gcgcaatctc	3600
[0342]	cagtgcacaga	agatactgct	agaaatctgc	tagaaaaaaa	acaaaaaagg	catgtataga	3660
[0343]	ggaattatga	gggaaagata	ccaagtcacg	gtttattctt	caaaatggag	gtggcttggt	3720
[0344]	gggaagggtg	aagctcattt	ggccagagtg	gaaatggaat	tgggagaaat	cgatgaccaa	3780
[0345]	atgtaaacac	ttggtgcctg	atatagcttg	acaccaagtt	agccccaagt	gaaatacctc	3840
[0346]	ggcaatatta	atgtgtcttt	tcccgatatt	cctcaggtac	tccaaagatt	caggtttact	3900
[0347]	cacgtcatcc	agcagagaat	ggaaagtcaa	atctctgaa	ttgctatgtg	tctgggtttc	3960
[0348]	atccatccga	cattgaagtt	gacttactga	agaatggaga	gagaattgaa	aaagtggagc	4020
[0349]	attcagactt	gtctttcagc	aaggactggg	ctttctatct	cttgtactac	actgaattca	4080
[0350]	ccccactga	aaaagatgag	tatgctgcc	gtgtgaacca	tgtgactttg	tcacagccca	4140
[0351]	agatagttaa	gtggggtaag	tcttacattc	ttttgtaagc	tgctgaaagt	tgtgtatgag	4200
[0352]	tagtcatatc	ataaagctgc	tttgatataa	aaaaggctca	tggccatact	accctgaatg	4260
[0353]	agtcccatcc	catctgatat	aaacaatctg	catattggga	ttgtcagga	atgttcttaa	4320
[0354]	agatcagatt	agtggcacct	gctgagatac	tgatgcacag	catggtttct	gaaccagtag	4380
[0355]	tttccctgca	gttgagcagg	gagcagcagc	agcacttca	caaatacata	tacactctta	4440
[0356]	acacttctta	cctactggct	tcctctagct	tttgtggcag	cttcaggtat	atntagcact	4500
[0357]	gaacgaacat	ctcaagaagg	tataggcctt	tgtttgtaag	tcctgctgtc	ctagcatcct	4560
[0358]	ataatcctgg	acttctccag	tactttctgg	ctggattggg	atctgaggct	agtaggaagg	4620
[0359]	gcttgttctc	gctgggtagc	tctaaacaat	gtattcatgg	gtaggaacag	cagcctattc	4680
[0360]	tgccagcctt	atcttaacc	atcttagaca	tttgttagta	catggtattt	taaaagtaaa	4740
[0361]	acttaatgtc	ttctttttt	ttctccactg	tctttttcat	agatcgagac	atgggtggag	4800
[0362]	gtggaagtgg	tggaggtgga	agtgggtggg	gtggaagtgg	tggaggtgga	agtggctccc	4860
[0363]	actccatgag	gtatttcagc	gccgccgtgt	ccggccccgg	ccgccccggg	ccccgcttca	4920
[0364]	tcgccatggg	ctacgtggac	gacacgcagt	tcgtgcggtt	cgacagcgac	tcggcgtgtc	4980
[0365]	cgaggatgga	gccgcgggcg	ccgtgggtgg	agcaggaggg	gccggagtat	tgggaagagg	5040
[0366]	agacacggaa	caccaaggcc	cacgcacaga	ctgacagaat	gaacctgcag	accctgcgcg	5100
[0367]	gctactacaa	ccagagcgag	gccagtcttc	acaccctcca	gtggatgatt	ggctgcgacc	5160
[0368]	tgggttccga	cggacgcctc	ctccgcgggt	atgaacagta	tcctacgat	ggcaaggatt	5220
[0369]	acctgcctc	gaacgaggac	ctgcgtcctc	ggaccgcagc	ggacactgcg	gctcagatct	5280
[0370]	ccaagcgcaa	gtgtgaggcg	gccaatgtgg	ctgaacaaag	gagagcctac	ctggagggca	5340
[0371]	cgtgcgtgga	gtggctccac	agatacctgg	agaacgggaa	ggagatgctg	cagcgcgcg	5400
[0372]	accccccaa	gacacacgtg	accaccacc	ctgtctttga	ctatgaggcc	accctgaggt	5460
[0373]	gctgggcct	ggccttctac	cctgcggaga	tcatactgac	ctggcagcgg	gatggggagg	5520
[0374]	accagacca	ggacgtggag	ctcgtggaga	ccaggcctgc	aggggatgga	accttccaga	5580
[0375]	agtgggcagc	tgtggtggg	ccttctggag	aggagcagag	atacacgtgc	catgtgcagc	5640
[0376]	atgaggggct	gccggagccc	ctcatgctga	gatggaagca	gtcttccctg	cccaccatcc	5700
[0377]	ccatcatggg	tatcgttgct	ggcctggttg	tccttgcagc	tgtagtcaet	ggagctgcgg	5760

- [0378] tcgctgctgt gctgtggaga aagaagagct cagattga 5798
- [0379] <210> 10
- [0380] <211> 1305
- [0381] <212> DNA
- [0382] <213> 人工序列(Artificial Sequence)
- [0383] <400> 10
- [0384] atgtctcgct ccgtggcctt agctgtgctc gcgctactct ctctttctgg cctggagget 60
- [0385] atccagcgta ctccaaagat tcaggtttac tcacgtcatc cagcagagaa tggaaagtca 120
- [0386] aatttctga attgctatgt gtctgggttt catccatccg acattgaagt tgacttactg 180
- [0387] aagaatggag agagaattga aaaagtggag cattcagact tgtctttcag caaggactgg 240
- [0388] tctttctatc tcttgacta cactgaattc acccccactg aaaaagatga gtatgcctgc 300
- [0389] cgtgtgaacc atgtgacttt gtcacagccc aagatagtta agtgggatcg agacatgggt 360
- [0390] ggaggtggaa gtggtggagg tggaagtgtt ggaggtggaa gtggtggagg tggaagtggc 420
- [0391] tcccactcca tgaggtatth cagcgcgcc gtgtcccggc ccggcccgcg ggagccccgc 480
- [0392] ttcatcgcca tgggctacgt ggacgacacg cagttcgtgc gtttcgacag cgactcggcg 540
- [0393] tgtccgagga tggagccgag ggcgccgtgg gtggagcagg aggggcccga gtattgggaa 600
- [0394] gaggagacac ggaacaccaa ggcccacgca cagactgaca gaatgaacct gcagaccctg 660
- [0395] cgcggtact acaaccagag cgaggccagt tctcacacc tccagtggat gattggctgc 720
- [0396] gacctggggt ccgacggacg cctcctccgc gggatgaac agtatgccta cgatggcaag 780
- [0397] gattacctcg cctgaacga ggacctgcgc tctggaccg cageggacac tgcggctcag 840
- [0398] atctccaagc gcaagtgtga ggcggccaat gtggctgaac aaaggagagc ctacctggag 900
- [0399] ggcacgtgcg tggagtggct ccacagatac ctggagaacg ggaaggagat gctgcagcgc 960
- [0400] gcggaccccc ccaagacaca cgtgaccac caccctgtct ttgactatga ggccaccctg 1020
- [0401] aggtgctggg ccttgggctt ctacctcgc gagatcatac tgacctggca gcgggatggg 1080
- [0402] gaggaccaga cccaggacgt ggagctcgtg gagaccagc ctgcagggga tggaaacctc 1140
- [0403] cagaagtggg cagctgtggt ggtgccttct ggagaggagc agagatacac gtgccatgtg 1200
- [0404] cagcatgagg ggctccgga gccctcatg ctgagatgga gtaaggagg agatggaggc 1260
- [0405] atcatgtctg ttagggaaag caggagctc tctgaagacc tttaa 1305
- [0406] <210> 11
- [0407] <211> 23
- [0408] <212> DNA
- [0409] <213> 人工序列(Artificial Sequence)
- [0410] <400> 11
- [0411] tagccccaag tgaataacc tgg 23
- [0412] <210> 12
- [0413] <211> 23
- [0414] <212> DNA
- [0415] <213> 人工序列(Artificial Sequence)
- [0416] <400> 12
- [0417] gctgtctata aatagtcctc agg 23
- [0418] <210> 13
- [0419] <211> 23

[0420]	<212> DNA	
[0421]	<213> 人工序列 (Artificial Sequence)	
[0422]	<400> 13	
[0423]	acatgtaagc agcatcatgg agg 23	
[0424]	<210> 14	
[0425]	<211> 8395	
[0426]	<212> DNA	
[0427]	<213> 人工序列 (Artificial Sequence)	
[0428]	<400> 14	
[0429]	caggtggcac ttttcgggga aatgtgcgcg gaaccctat ttgtttattt ttctaaatac	60
[0430]	attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa	120
[0431]	aaaggaagag tatgagtatt caacatttcc gtgtcgcctt tattcccttt tttgcggcat	180
[0432]	tttgcttcc tgtttttgct caccagaaa cgctggtgaa agtaaaagat gctgaagatc	240
[0433]	agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga	300
[0434]	gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg	360
[0435]	cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgata cactattctc	420
[0436]	agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag	480
[0437]	taagagaatt atgcagtgtc gccataacca tgagtataa cactgcggcc aacttacttc	540
[0438]	tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg ggggatcatg	600
[0439]	taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg	660
[0440]	acaccacgat gcctgtagca atggcaaca cgttgcgcaa actattaact ggcgaaactac	720
[0441]	ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac	780
[0442]	cacttctgcg ctcgccctt ccggctggct ggtttattgc tgataaatct ggagccgggtg	840
[0443]	agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg	900
[0444]	tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg	960
[0445]	agataggtgc ctactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac	1020
[0446]	tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atcctttttg	1080
[0447]	ataatctcat gacaaaatc cttaacgtg agttttcgtt cactgagcg tcagaccccg	1140
[0448]	tagaaaagat caaaggatct tcttgagatc cttttttct gcgcgtaatc tgctgcttgc	1200
[0449]	aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag ctaccaactc	1260
[0450]	ttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt	1320
[0451]	agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctgctctgc	1380
[0452]	taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact	1440
[0453]	caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac	1500
[0454]	agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag	1560
[0455]	aaagcggcac gcttccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg	1620
[0456]	gaacaggaga gcgcacgagg gagcttcag ggggaaacgc ctggtatctt tatagtctctg	1680
[0457]	tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga	1740
[0458]	gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt	1800
[0459]	ttgctacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgctt	1860
[0460]	ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcagc tcagtgagcg	1920
[0461]	aggaagcggg agagcggcca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt	1980

[0462]	aatgcagctg	gcacgacagg	tttcccgact	ggaaagcggg	cagtgagcgc	aacgcaatta	2040
[0463]	atgtgagtta	gctcactcat	taggcacccc	aggctttaca	ctttatgctt	ccggctcgtg	2100
[0464]	tgtttgtgtg	aattgtgagc	ggataacaat	ttcacacagg	aaacagctat	gaccatgatt	2160
[0465]	acgccaagcg	cgcaattaac	cctcactaaa	gggaacaaaa	gctggagctg	caagcttaat	2220
[0466]	gtagtcttat	gcaatactct	tgtagtcttg	caacatggta	acgatgagtt	agcaacatgc	2280
[0467]	cttacaagga	gagaaaaagc	accgtgcatg	ccgattggtg	gaagtaaggt	ggtacgatcg	2340
[0468]	tgcttatta	ggaaggcaac	agacgggtct	gacatggatt	ggacgaacca	ctgaattgga	2400
[0469]	ggcgtggcct	ggcggggact	ggggagtggc	gagccctcag	atcctgcata	taagcagctg	2460
[0470]	ctttttgcct	gtactgggtc	tctctggtta	gaccagatct	gagcctggga	gctctctggc	2520
[0471]	taactaggga	accactgct	taagcctcaa	taaagcttgc	cttgagtgct	tcaagtagtg	2580
[0472]	tgtgccgctc	tgtttgtgga	ctctggtaac	tagagatccc	tcagaccctt	ttagtcagtg	2640
[0473]	tggaaaatct	ctagcagtgg	cgcccgaaca	gggacctgaa	agcgaaaggg	aaaccagagc	2700
[0474]	tctctcgacg	caggactcgg	cttgctgaag	cgcgcacggc	aagaggcgag	ggcgggcgac	2760
[0475]	tggtgagtac	gccaaaaatt	ttgactagcg	gaggctagaa	ggagagagat	gggtgcgaga	2820
[0476]	gcgtcagtat	taagcggggg	agaattagat	cgcgatggga	aaaaattcgg	ttaaggccag	2880
[0477]	ggggaaagaa	aaaatataaa	ttaaaacata	tagtatgggc	aagcagggag	ctagaacgat	2940
[0478]	tcgcagttaa	tcttgccctg	ttagaacat	cagaaggctg	tagacaaata	ctgggacagc	3000
[0479]	tacaaccatc	ccttcagaca	ggatcagaag	aacttagatc	attatataat	acagtagcaa	3060
[0480]	ccctctattg	tgtgcatcaa	aggatagaga	taaaagacac	caaggaagct	ttagacaaga	3120
[0481]	tagaggaaga	gcaaaacaaa	agtaagacca	ccgcacagca	agcggccgct	gatcttcaga	3180
[0482]	cctggaggag	gagatatgag	ggacaattgg	agaagtgaat	tatataaata	taaagtagta	3240
[0483]	aaaattgaac	cattaggagt	agcaccacc	aaggcaaaga	gaagagtgg	gcagagagaa	3300
[0484]	aaaagagcag	tgggaatagg	agctttgttc	cttgggttct	tgggagcagc	aggaagcact	3360
[0485]	atggcgcgag	cctcaatgac	gctgacggta	caggccagac	aattattgtc	tggtatagtg	3420
[0486]	cagcagcaga	acaatttgc	gagggtatt	gaggcgaac	agcatctgtt	gcaactcaca	3480
[0487]	gtctggggca	tcaagcagct	ccaggcaaga	atcctggctg	tggaaagata	cctaaaggat	3540
[0488]	caacagctcc	tggggatttg	gggttgctct	ggaaaactca	tttgcaccac	tgctgtgcct	3600
[0489]	tggaatgcta	gttggagtaa	taaattctctg	gaacagattg	gaatcacacg	acctggatgg	3660
[0490]	agtgggacag	agaaattaac	aattacacaa	gcttaataca	ctccttaatt	gaagaatcgc	3720
[0491]	aaaaccagca	agaaaagaat	gaacaagaat	tattggaatt	agataaatgg	gcaagtttgt	3780
[0492]	ggaattgggt	taacataaca	aattggctgt	ggtatataaa	attattcata	atgatagtag	3840
[0493]	gaggcttggt	aggtttaaga	atagtttttg	ctgtactttc	tatagtgaat	agagttaggc	3900
[0494]	agggatattc	accattatcg	tttcagaccc	acctccaac	cccgagggga	cccgacaggc	3960
[0495]	ccgaaggaat	agaagaagaa	ggtggagaga	gagacagaga	cagatccatt	cgattagtga	4020
[0496]	acggatctcg	acggtatcga	tcacgagact	agcctcgaca	caaatggcag	tattcatcca	4080
[0497]	caattttaa	agaaaagggg	ggattggggg	gtacagtgca	ggggaaagaa	tagtagacat	4140
[0498]	aatagcaaca	gacatacaaa	ctaaagaatt	acaaaaacaa	attacaaaaa	ttcaaaattt	4200
[0499]	tcgggtttat	tacagggaca	gcagaaatcc	actttggctc	gagaagcttg	atatcggtc	4260
[0500]	cggtgcccg	cagtgggcag	agcgcacatc	gccacagtc	cccgagaagt	tggggggagg	4320
[0501]	ggtcggcaat	tgaaccggtg	cctagagaag	gtggcgcggg	gtaaactggg	aaagtgatgt	4380
[0502]	cgtgtactgg	ctccgccttt	ttcccagggg	tgggggagaa	ccgtatataa	gtgcagtagt	4440
[0503]	cgccgtgaac	gttctttttc	gcaacgggtt	tgccgccaga	acacaggatc	cgccaccatg	4500

[0504]	gtgagcaagg gcgaggagct gttcaccggg gtggtgccca tcctggtcga gctggacggc	4560
[0505]	gacgtaaacc gccacaagtt cagcgtgtcc ggcgagggcg agggcgatgc cacctacggc	4620
[0506]	aagctgacct tgaagttcat ctgcaccacc ggcaagctgc ccgtgccctg gccaccctc	4680
[0507]	gtgaccaccc tgacctacgg cgtgcagtgc ttcagccgct accccgacca catgaagcag	4740
[0508]	cacgacttct tcaagtccgc catgcccgaa ggctacgtcc aggagcgcac catcttcttc	4800
[0509]	aaggacgacg gcaactacaa gacccgcgcc gaggtgaagt tcgagggcga caccctgggtg	4860
[0510]	aaccgcatcg agctgaaggc catcgacttc aaggaggacg gcaacatcct ggggcacaag	4920
[0511]	ctggagtaca actacaacag ccacaacgtc tatatcatgg ccgacaagca gaagaacggc	4980
[0512]	atcaaggtga acttcaagat ccgccacaac atcgaggacg gcagcgtgca gctcgccgac	5040
[0513]	cactaccagc agaacacccc catcgcgac ggccccgtgc tgctgcccgca caaccactac	5100
[0514]	ctgagcacc agtccgccct gagcaaagac cccaacgaga agcgcgatca catggtcctg	5160
[0515]	ctggagtctg tgaccgccgc cgggatcact ctcgcatgg acgagctgta caagtaagct	5220
[0516]	agcacgcgtg gcgcgccgtt aacggcgcgc ccctctccct cccccccc taacgttact	5280
[0517]	ggccgaagcc gcttgaata aggccggtgt gcgtttgtct atatgttatt ttccaccata	5340
[0518]	ttgccgtctt ttggcaatgt gagggcccgg aaacctggcc ctgtcttctt gacgagcatt	5400
[0519]	cctaggggtc tttcccctct cgccaaagga atgcaagtc tgttgaatgt cgtgaaggaa	5460
[0520]	gcagttctc tggaagcttc ttgaagacaa acaacgtctg tagcgacct ttgcaggcag	5520
[0521]	cggaacccc cacctggcga caggtgcctc tgcggccaaa agccacgtgt ataagataca	5580
[0522]	cctgcaaagg cggcacaacc ccagtccac gttgtgagtt ggatagttgt ggaaagagtc	5640
[0523]	aatggctct cctcaagcgt attcaacaag gggctgaagg atgcccagaa ggtacccat	5700
[0524]	tgtatgggat ctgatctgg gcctcggtag acatgcttta catgtgttta gtcgaggtta	5760
[0525]	aaaaaacgtc taggcccc gaaccacggg gacgtggtt tcctttgaaa aacacgatga	5820
[0526]	taatatggcc acaacctct agaatgacc agtacaagcc cacggtgcgc ctcgccacc	5880
[0527]	gcgacgacgt cccagggc gtacgaccc tcgccgccgc gttcgccgac taccgccca	5940
[0528]	cgcccacac cgtcgatccg gaccgccaca tcgagcgggt caccgagctg caagaactct	6000
[0529]	tcctcacgcg cgtcgggtc gacatcgga aggtgtgggt cgcggacgac ggcgccgcg	6060
[0530]	tggcgtctg gaccacccg gagagcgtc aagcggggc ggtgttcgcc gagatcggc	6120
[0531]	cgcgatggc cgagttgagc gttcccggc tggccgcga gcaacagatg gaaggcctcc	6180
[0532]	tggcggcga ccggccaag gagcccgt gtttctggc caccgtcggc gtctcggcc	6240
[0533]	accaccagg caaggctc ggcagcgcg tcgtgctccc cggagtggag gcggccgagc	6300
[0534]	gcgccgggt gccgcctc ctggagacct ccgcgcccg caacctccc ttctacgagc	6360
[0535]	ggctcggctt caccgtacc gccgacgtc aggtgcccg aggaccgc acctggtgca	6420
[0536]	tgaccgcaa gcccggtgc tgagtcgaca ctagtaatca acctttggat taaaaattt	6480
[0537]	gtgaaagatt gactggtatt cttactatg ttgctcttt tacgctatgt ggatacgtg	6540
[0538]	ctttaatgcc tttgtatcat gctattgct cccgtatggc tttcatttc tcctcctgt	6600
[0539]	ataaatctg gttgctgtct ctttatgagg agttgtggc cgttgtcagg caacgtggc	6660
[0540]	tgggtgcac tgtgtttgct gacgcaacc ccaactggtg gggcattgc accacctgc	6720
[0541]	agctccttc cgggacttc gtttcccc tcctattgc cacggcgga ctcategccg	6780
[0542]	cctgcctgc ccgtgctgg acaggggtc ggctgttgg cactgacaat tccgtggtg	6840
[0543]	tgctgggaa gctgacgtc tttccatggc tgctgcctg tgttgccacc tggattctg	6900
[0544]	gcgggacgtc cttctgtac gtccctcgg ccctcaatcc agcggacct cttcccgcg	6960
[0545]	gctgctgcc ggctctcgg cctcttcgc gtcttcgct tcgccctcag acgagtcgga	7020

[0546] tctccctttg ggccgcctcc cgcctgccc cggaattcga gctcgggtacc ttttaagacca 7080  
[0547] atgacttaca aggcagctgt agatcttagc cactttttta aagaaaagg gggactggaa 7140  
[0548] gggetaattc actcccaacg aagacaagat ctgctttttg cttgtactgg gtctctctgg 7200  
[0549] ttagaccaga tctgagcctg ggagctctct ggctaactag ggaaccact gcttaagcct 7260  
[0550] caataaagct tgccttgagt gcttcaagta gtgtgtgcc gtctgtttgt tgactctggt 7320  
[0551] aactagagat cctcagacc cttttagca gtgtggaaaa tctctagcag tagtagttca 7380  
[0552] tgtcatctta ttattcagta tttataactt gcaaagaaat gaatatcaga gagtgagagg 7440  
[0553] aactgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 7500  
[0554] aataaagcat tttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct 7560  
[0555] tatcatgtct ggctctagct atcccgcctc taactccgcc cagttccgcc cattctccgc 7620  
[0556] cccatggctg actaattttt tttatttatg cagaggccga ggccgcctcg gcctctgagc 7680  
[0557] tattccagaa gtagtgagga ggctttttg gaggcctagg cttttgcgct gagacgtacc 7740  
[0558] caattcggcc tatagtgagt cgtattacgc gcgctcactg gccgtcgttt tacaacgtcg 7800  
[0559] tgactgggaa aaccctggcg ttaccaact taatcgctt gcagcacatc cccctttcgc 7860  
[0560] cagctggcgt aatagcgaag agggccgcac cgatcgccct tccaacagt tgcgcagcct 7920  
[0561] gaatggcgaa tggcgcgac gcctcttag cgccgcatta agcgcggcgg gtgtggtggt 7980  
[0562] tacgcgcagc gtgaccgcta cacttgccag gcctctagcg cccgtcctt tcgctttctt 8040  
[0563] cccttccttt ctgccacgt tcgcccgtt tcccgtcaa gctctaaatc gggggctccc 8100  
[0564] tttagggttc cgatttagtg ctttacgca cctcgacccc aaaaaacttg attagggtga 8160  
[0565] tggttcacgt agtgggcat gccttgata gacggtttt cgcctttga cgttgagtc 8220  
[0566] cacgttcttt aatagtgac tcttgtcca aactggaaca acactcaacc ctatctcggt 8280  
[0567] ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct 8340  
[0568] gattaacaa aaatttaacg cgaatttta caaaatatta acgtttacaa tttcc 8395

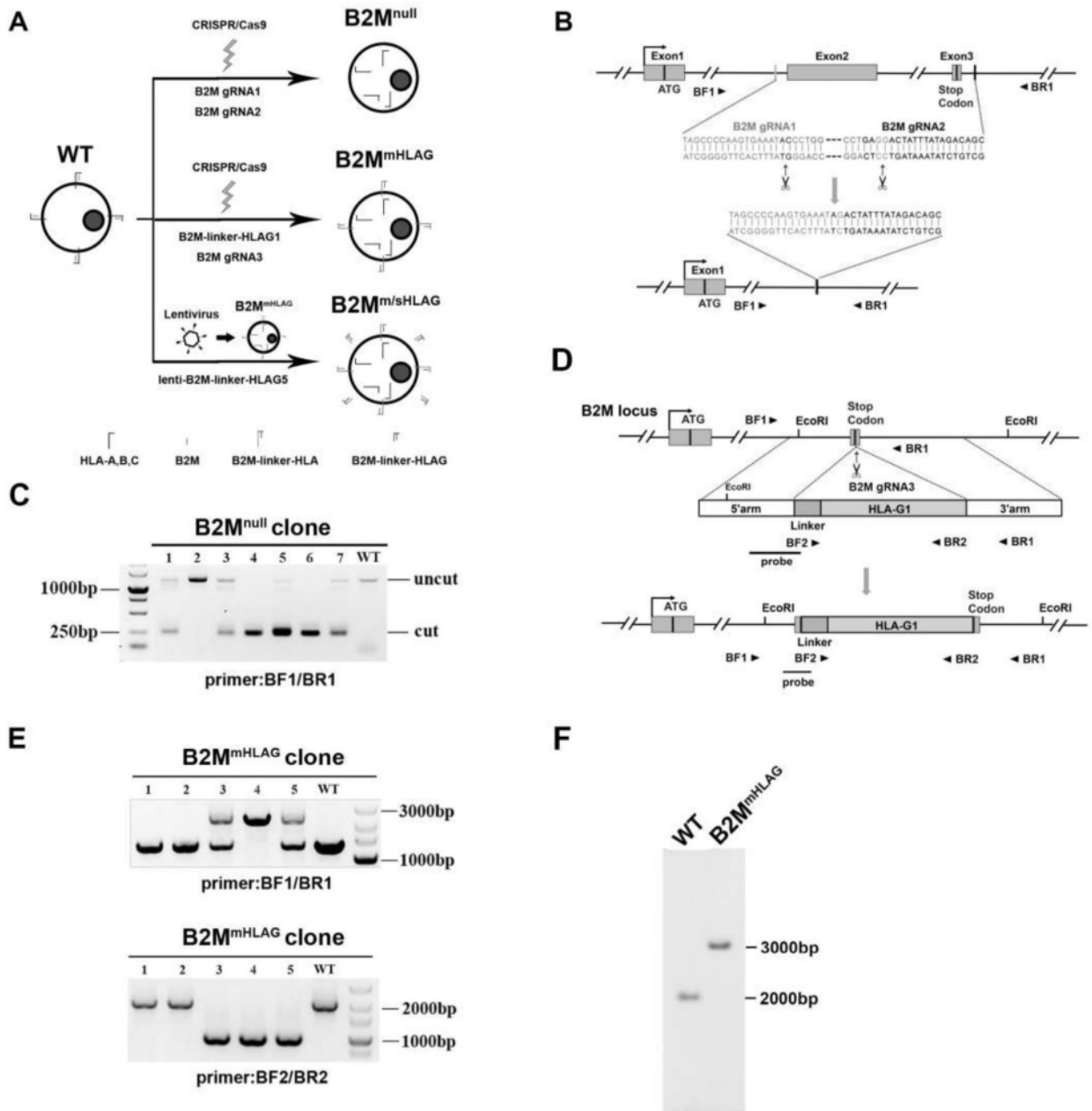


图1



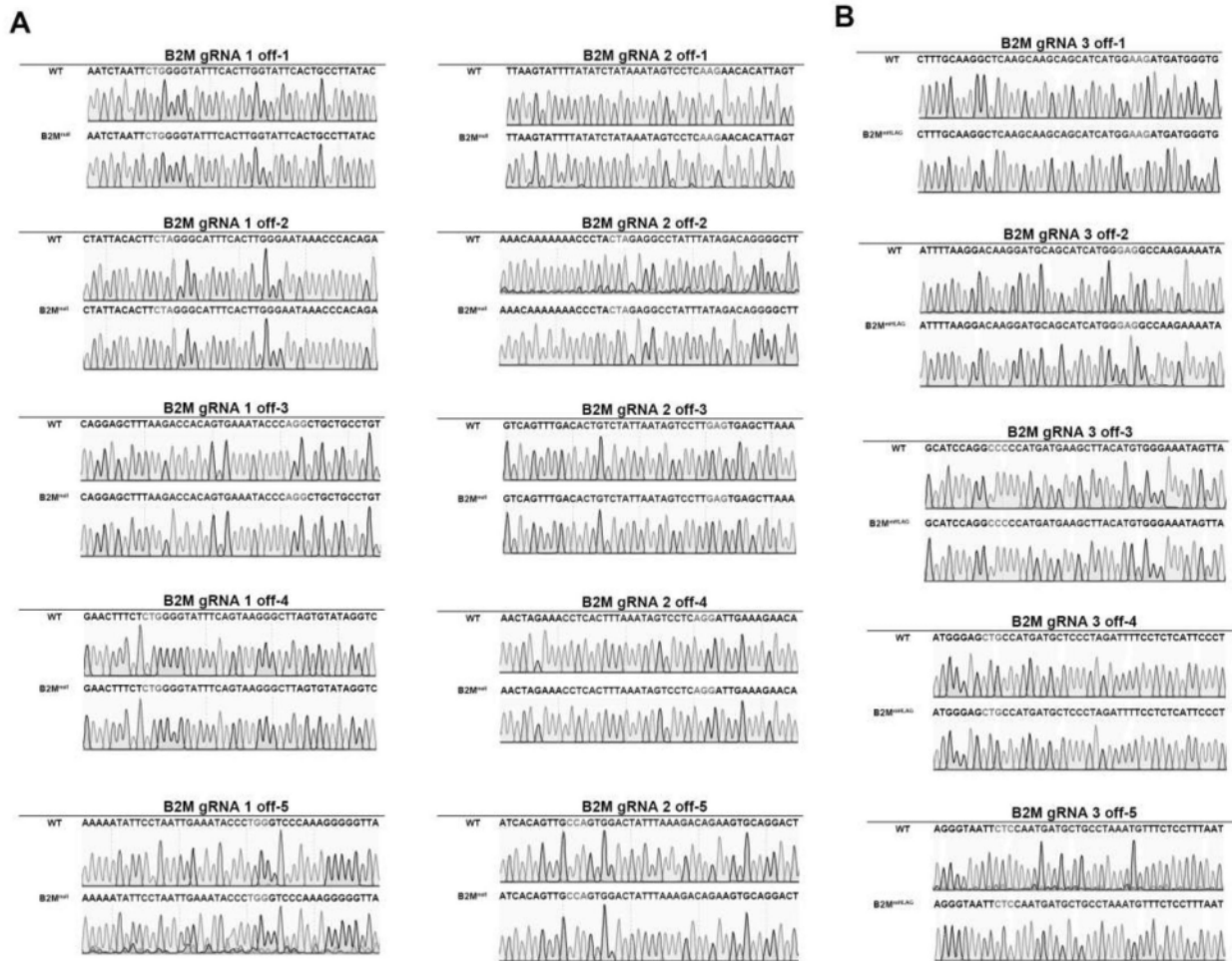


图2

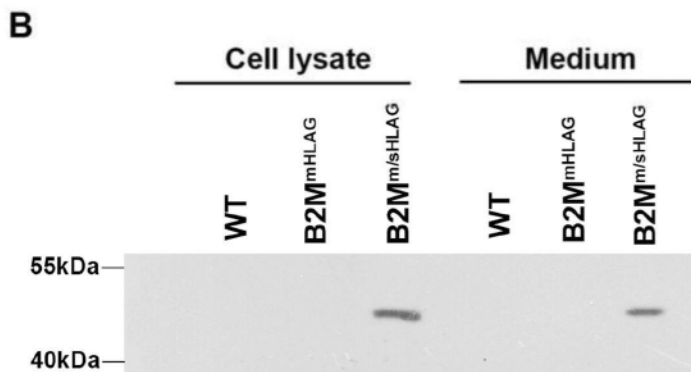
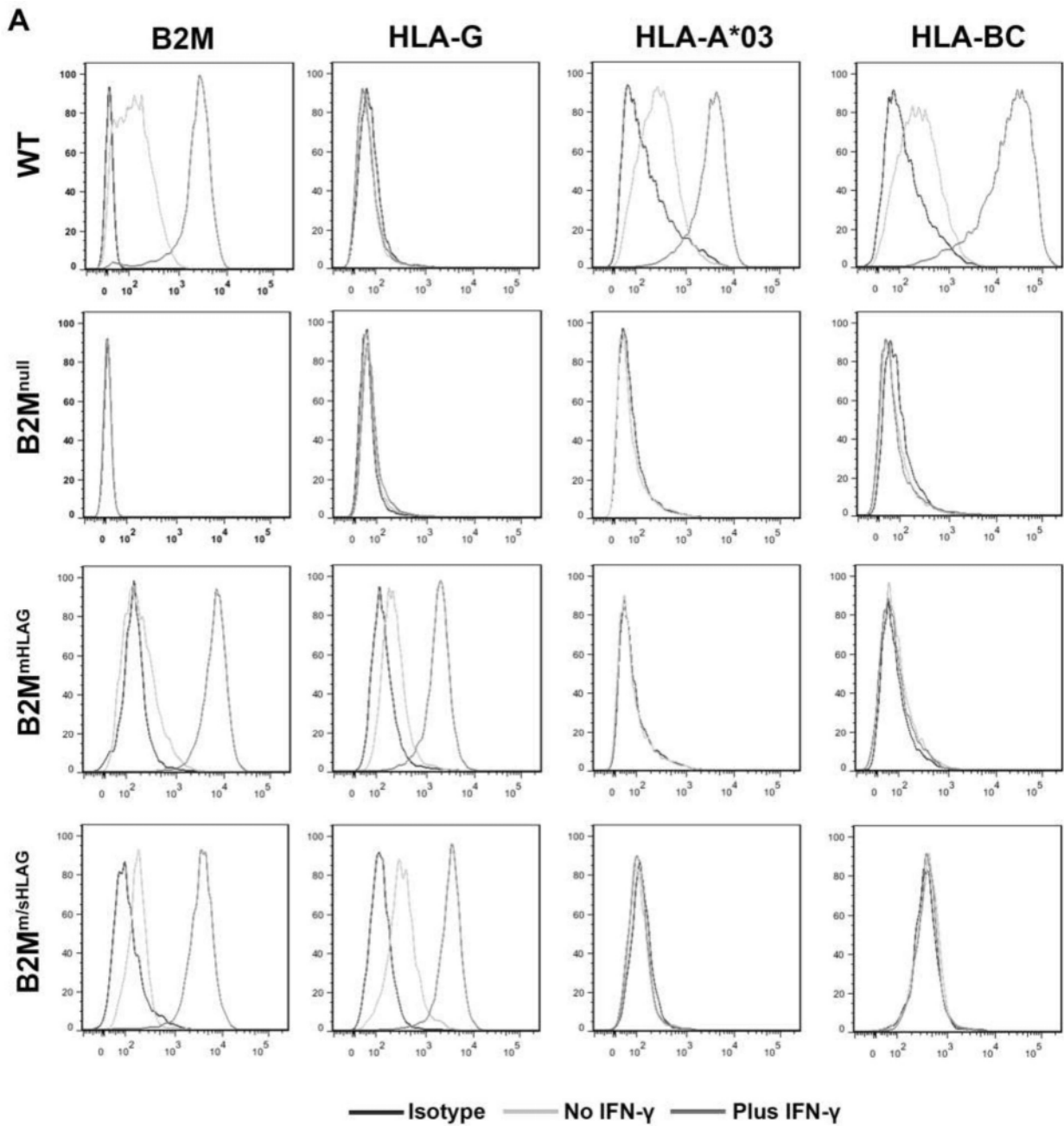


图3

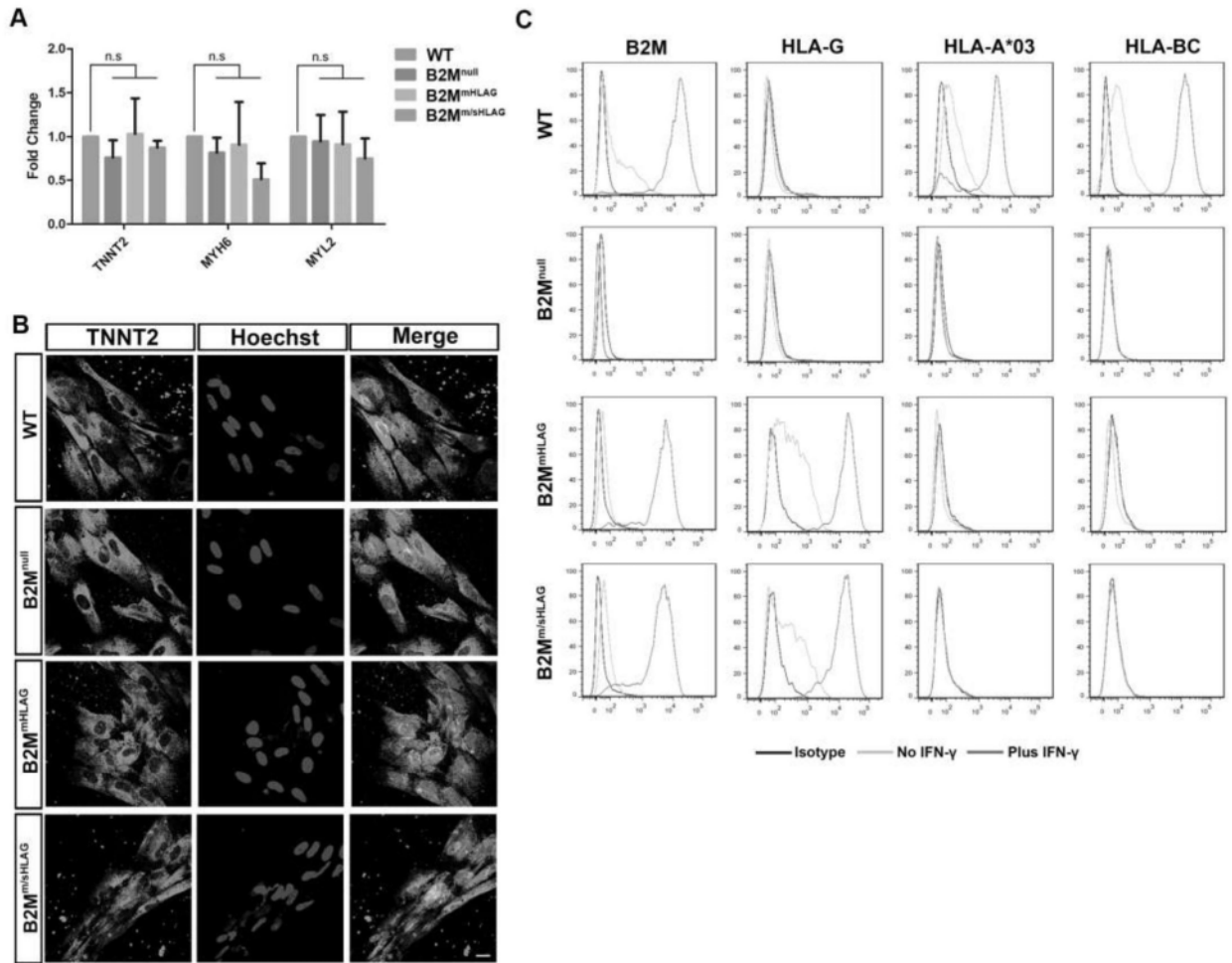
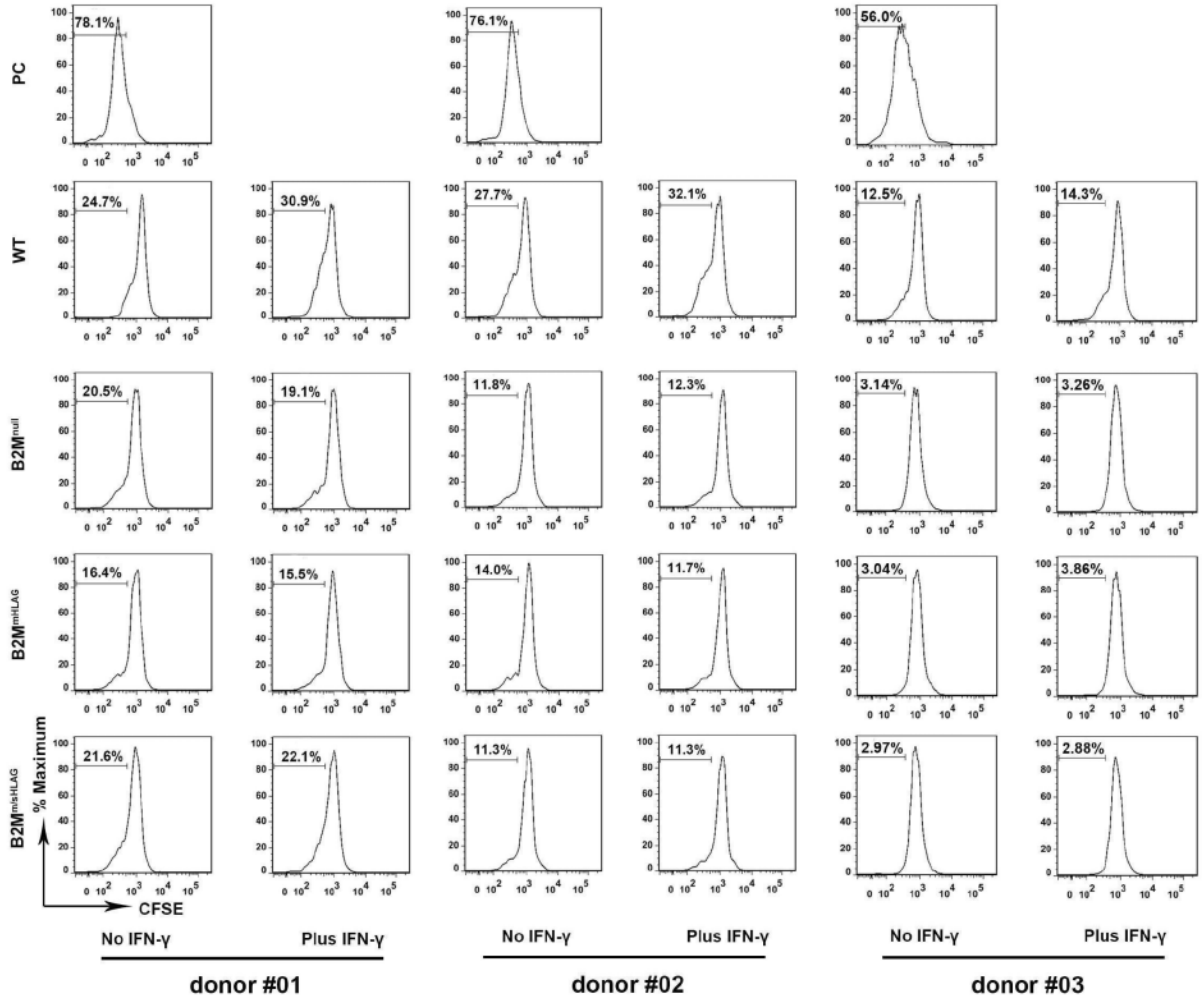
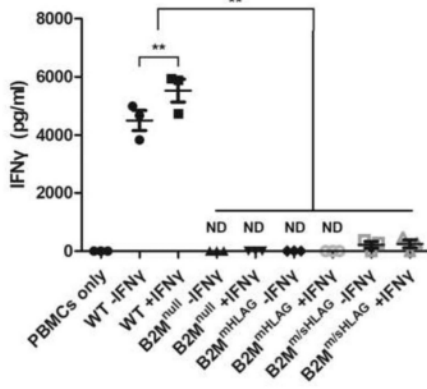


图4

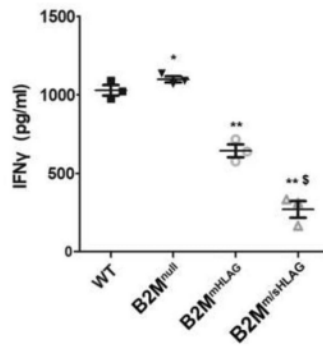
A



B



C



D

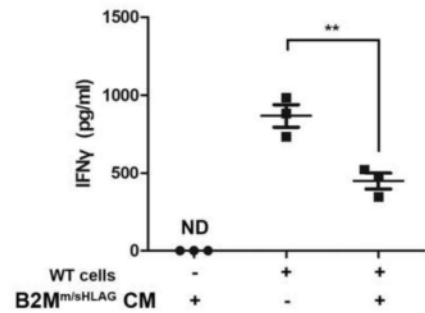


图5

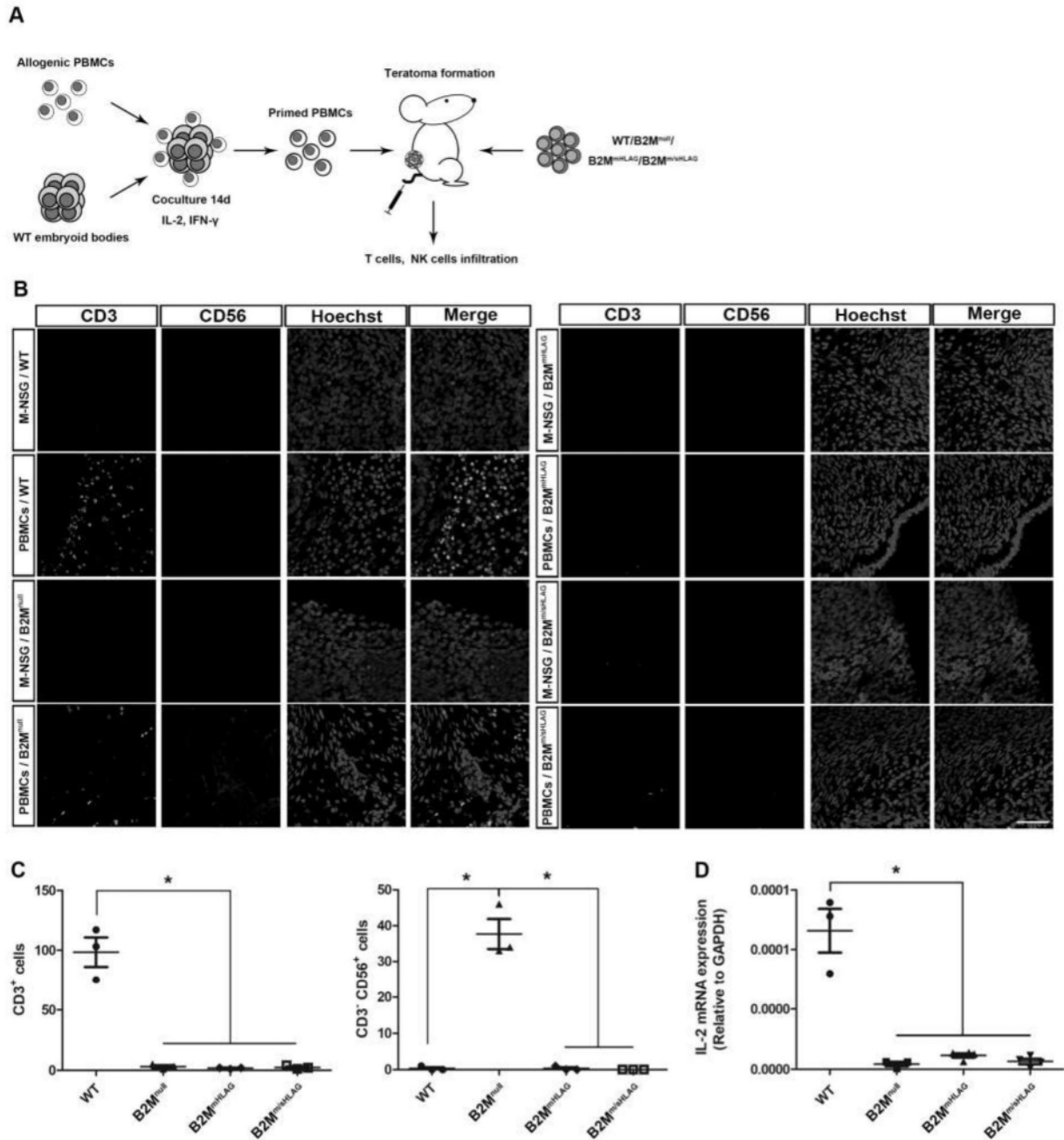


图6