

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 15/10

C12N 15/12 C12N 15/13

C12N 15/62

//C07K19/00 C07K16/00

C07K14/435

[12] 发明专利申请公开说明书

[21] 申请号 99808906.0

[43] 公开日 2001年8月29日

[11] 公开号 CN 1310758A

[22] 申请日 1999.4.30 [21] 申请号 99808906.0

[30] 优先权

[32] 1998.5.20 [33] JP [31] 138652/1998

[32] 1998.10.1 [33] JP [31] 279876/1998

[86] 国际申请 PCT/JP99/02341 1999.4.30

[87] 国际公布 WO99/60113 日 1999.11.25

[85] 进入国家阶段日期 2001.1.19

[71] 申请人 中外制药株式会社

地址 日本东京都

[72] 发明人 土屋政幸 齐藤千良 大友俊彦

[74] 专利代理机构 柳沈知识产权律师事务所

代理人 巫肖南

权利要求书 2 页 说明书 68 页 附图页数 7 页

[54] 发明名称 基因克隆的新方法

[57] 摘要

本发明涉及用于分离膜结合蛋白的编码基因的方法,其以将一种功能性蛋白与一种融合蛋白本身相融合为特征,而且与现有的在融合蛋白中载有被抗体识别的表位的 TMT 方法不同。该方法使选择性分离膜结合蛋白的编码基因成为可能。

I S S N 1 0 0 8 - 4 2 7 4

权 利 要 求 书

1. 用于分离膜结合蛋白的编码基因的方法，所述的方法包括如下步骤：
- 5 (i)将包括编码对一种抗原具有结合亲和力的可分泌性功能蛋白的 DNA 以及连接于该功能蛋白编码 DNA 3' 侧下游的 cDNA 的载体引入细胞，
- (ii)在细胞内表达对抗原具有结合亲和力的可分泌性功能蛋白与由该 cDNA 所编码蛋白的融合蛋白，
- (iii)通过将细胞膜上表达融合蛋白的细胞与抗原相接触，筛选结合抗原的细胞，并且
- 10 (iv)从所筛选的细胞中分离插入到载体中的 cDNA。
2. 权利要求 1 所述的方法，其中通过在功能蛋白编码 DNA 3' 侧下游的限制性酶切位点处将 cDNA 引入载体来获得在步骤(i)中引入细胞的载体。
3. 权利要求 1 所述的方法，其中通过将包括编码功能蛋白的 DNA 和连接于该功能蛋白编码 DNA 3' 侧下游的 cDNA 的 DNA 引入载体来获得
- 15 在步骤(i)中引入细胞的载体，
4. 权利要求 1 至 3 中任何一个所述的方法，其中通过编码肽连接序列的 DNA 连接编码功能蛋白的 DNA 和其 3' 侧下游的 cDNA。
5. 权利要求 1 至 4 中任何一个所述的方法，其中该 cDNA 源于哺乳动物细胞的 cDNA 文库。
- 20 6. 权利要求 1 至 5 中任何一个所述的方法，其中在步骤(i)中引入细胞的载体在编码功能蛋白的 DNA 5' 侧的上游包括编码分泌信号序列的 DNA。
7. 权利要求 1 至 6 中任何一个所述的方法，其中该功能蛋白为抗体。
- 25 8. 权利要求 1 至 7 中任何一个所述的方法，其中对抗原具有结合亲和力的功能蛋白为一种单链抗体。
9. 权利要求 1 至 8 中任何一个所述的方法，其中该载体包括一种 DNA，在该 DNA 中编码抗体恒定区的 DNA 连接于编码一种单链抗体的 DNA 3' 侧的下游。
- 30 10. 权利要求 1 至 9 中任何一个所述的方法，其中将该抗原结合于一种支持物。

11. 权利要求 10 所述的方法，其中该支持物用于细胞培养。

12. 权利要求 1 至 11 中任何一个所述的方法，其包括确定从细胞获得的基因是否包含新的序列。

13. 权利要求 12 所述的方法，其包括筛选 cDNA 文库以获得由细胞所
5 获基因的全长基因，该基因包括新的序列。

14. 权利要求 13 所述的方法，其包括分离由细胞所获基因的全长基因，
该基因包括新的序列。

15. 用于分离膜结合蛋白的编码基因的试剂盒，该试剂盒包括具有用
于在编码对一种抗原具有结合亲和力的可分泌性功能蛋白的 DNA 3' 侧下
10 游插入 cDNA 的限制性酶识别位点的载体。

16. 权利要求 15 所述的试剂盒进一步还包括结合抗原的支持物和/或
待引入载体的细胞。

说明书

基因克隆的新方法

5 技术领域

本发明涉及一种选择性而且有效地分离膜结合蛋白的编码基因的新型基因克隆方法。

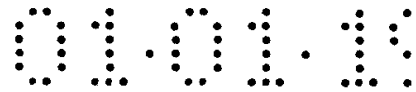
背景技术

10 细胞中合成的蛋白质根据其各自特征可以被分成定位于细胞内细胞器如细胞核、线粒体、细胞质等中的那些蛋白；通过结合至细胞膜而发挥功能的那些蛋白，例如受体和通道分子；以及通过分泌至细胞外而发挥功能的那些蛋白，例如生长因子和细胞因子等。特别是，结合至细胞膜的蛋白质分子负责重要的生物学功能，例如对生长因子和分化因子的细胞反应、
15 炎症反应、细胞-细胞相互作用、激素反应等等，因此这些分子可以成为各类疾病的诊断和治疗药物的靶分子。

近年来，如以基因组计划为代表，实施了利用随机方法的大规模基因克隆方法，而且堆积了大量的基因序列信息例如大量的 EST(表达序列标签)(Matsubara, K. 人造器官(Artificial Organs)(1996)20, 823-827)。然而，
20 通过这些 EST 鉴定具有预期功能的蛋白质决非易事，而且为了从基因序列信息预测和分析所编码蛋白的功能，需要大量的时间和努力。因此，长期以来人们期待着一种在随机 cDNA 克隆阶段筛选，或者至少在某种程度上筛选预计具有预期功能的蛋白的编码基因的方法。

利用蛋白质定位进行的克隆方法被发展成为对此类问题的一个解决方法。例如，分泌至细胞外的蛋白质具有对分泌至关重要的包括
25 左右氨基酸残基的氨基酸序列，其通常被称为分泌信号序列或引导序列。

Tashiro, K.等关注于该分泌性蛋白合成的特征，而且发展了一种特异性筛选分泌性蛋白的编码基因的克隆方法(Tashiro, K.等, 科学(Science)(1993)261, 300-603)。当将正常分泌至细胞外的蛋白质如白细胞介素-2(IL-2)受体的信号序列删除时，它们不能表达于细胞膜上。如果融合有
30 编码分泌信号序列的 cDNA，则该 IL-2 受体能够作为一种融合蛋白而被重



新表达于细胞膜上。由于用识别 IL-2 受体的抗体可以筛选表达 IL-2 受体融合蛋白的细胞，因此可以分离对引入细胞的信号序列已发挥功能的蛋白进行编码的 cDNA。该方法因其选择性地克隆信号序列的编码基因而通常被称为 SST(信号序列捕获)方法。通过基本上相同的原理还发展了一种针对酵母的克隆方法(美国专利号 5,536,637)。

然而，即使通过该方法获得了编码一种包括信号序列的蛋白的基因片段，人们还是无法知道其是否是一种分泌性蛋白，或者其是否是一种膜结合蛋白。同样，该方法需要利用包括 5' 末端的 cDNA 文库，而用于有效构建一个选择性地包括 5' 末端的 cDNA 文库的技术不一定是容易而通用的技术。

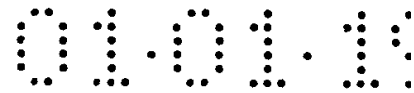
最近，Ishihara 等和 Nakauchi 等报道了 TMT(跨膜捕获)方法，其更选择性地克隆一个编码膜结合蛋白的基因(Yoshikazu Ichihara 和 Yoshikazu Kurozawa，日本分子生物学会年会摘要(1998)，No. 3-509-P-533，Nakauchi 等，WO98/03645)。Ichihara 等的方法基于同上述 SST 方法相反的原理。换句话说，将 IL-2 受体的胞外区和含有由 cDNA 所编码的细胞膜结合区的一种蛋白融合，IL-2 受体表达于细胞膜表面，而利用抗 IL-2 受体的抗体对细胞进行筛选。该方法的模式实验利用抗 IL-2 受体的抗体证实了 I 型或 II 型膜结合蛋白或者糖基磷脂酰肌醇(GPI)锚定型膜结合蛋白与 IL-2 受体的融合分子在细胞膜上的表达。

然而，当引入 cDNA 文库时，在所选的 cDNA 中还包括不含跨膜区和膜结合区的蛋白。换句话说，对该 TMT 方法获得的膜结合蛋白编码基因的克隆选择性不一定高。这表明，例如虽然理论上所有不具有跨膜区和 GPI 锚的融合蛋白可以被分泌，但根据融合蛋白的结构和氨基酸组成，并非由跨膜区和 GPI 锚所致的非特异性凝集也可以发生于细胞膜上。

此外，在该 TMT 方法的情况下，被抗体识别的表位表达于融合蛋白中。因此，即使以上述方式表达的融合蛋白被非特异性地吸收于细胞膜上，只要该表位暴露，该抗体将识别并结合该表位。同样，细胞膜表面那些将被分泌至细胞外的分子也可以被该抗体识别。因此，希望进一步提高对该 TMT 方法所获表达膜结合蛋白之细胞的选择性。

30

发明内容



本发明解决了 TMT 方法的问题并提供了一种具有更高选择性的基因克隆方法。

本发明的一个特征为通过将功能性蛋白连接至融合蛋白本身来分离编码膜结合蛋白的基因，其与携带识别抗体的表位的传统 TMT 方法不同。因此，本方法能够选择性分离编码膜结合蛋白的基因。

换言之，本发明提供：

(1)用于分离膜结合蛋白的编码基因的一种方法，该方法包括以下步骤：

(i)将包括对一种抗原具有结合亲和力的可分泌性功能蛋白的编码 DNA 以及连接于该功能性蛋白编码 DNA 3' 侧下游的 cDNA 的载体引入细胞，

(ii)对抗原具有结合亲和力的可分泌性功能蛋白与所述 cDNA 所编码蛋白的融合蛋白在细胞内表达，

(iii)通过将在细胞膜上表达该融合蛋白的细胞与抗原相接触，筛选结合该抗原的细胞，并且

(iv)从所筛选的细胞中分离插入到所述载体中的 cDNA，

(2)如(1)所述的方法，其中通过在载体上功能性蛋白编码 DNA 3' 侧下游的限制性酶切位点处引入 cDNA 来获得在步骤(i)中引入细胞的载体，

(3)如(1)所述的方法，其中通过将包括编码一功能性蛋白的 DNA 和连接于该功能性蛋白编码 DNA 3' 侧下游的 cDNA 的 DNA 引入载体来获得在步骤(i)中引入细胞的载体，

(4)如(1)至(3)中任何一个所述的方法，其中通过编码肽连接序列的 DNA 使编码所述功能性蛋白的 DNA 和其 3' 侧下游的 cDNA 连接，

(5)如(1)至(4)中任何一个所述的方法，其中所述 cDNA 源于哺乳动物细胞的 cDNA 文库。

(6)如(1)至(5)中任何一个所述的方法，其中在步骤(i)中引入细胞的载体在编码功能性蛋白的 DNA 5' 侧的上游包括编码分泌信号序列的 DNA，

(7)如(1)至(6)中任何一个所述的方法，其中该功能性蛋白为一种抗体，

(8)如(1)至(7)中任何一个所述的方法，其中对抗原具有结合亲和力的功能性蛋白为一种单链抗体，其优选为单价或二价，

(9)如(1)至(8)中任何一个所述的方法，其中该载体包括一种 DNA，在



该 DNA 中编码所述抗体恒定区的 DNA 连接于编码一种单链抗体的 DNA 3' 侧的下游，

(10)如(1)至(9)中任何一个所述的方法，其中将抗原结合于一种支持物，

(11)如(10)所述的方法，其中支持物用于细胞培养，

5 (12)如(1)至(11)中任何一个所述的方法，其包括确定从细胞获得的基因是否包含新的序列，

(13)如(12)所述的方法，其包括筛选 cDNA 文库以获得由细胞所获基因的全长基因，该基因包括一新的序列，

10 (14)如(13)所述的方法，其包括分离由细胞所获基因的全长基因，该基因包括一新的序列，

(15)用于分离膜结合蛋白的编码基因的试剂盒，该试剂盒包括具有用于在编码对一种抗原具有结合亲和力的可分泌性功能蛋白的 DNA 3' 侧下游插入 cDNA 的限制性酶识别位点的载体，以及

15 (16)如(15)所述的试剂盒进一步还包括结合抗原的支持物和/或待引入载体的细胞。

通过本发明的方法可分离的膜结合蛋白，可如 I 型或 II 型膜结合蛋白以及 GPI 锚定型膜结合蛋白等。I 型或 II 型膜结合蛋白为包括跨膜区的蛋白，而且在由所表达多肽的 N 末端侧或 C 末端侧分泌至细胞外后结合于细胞膜。跨膜区为穿透细胞膜内侧和外侧的区域，而且因为该跨膜区保持在细胞膜内，所以蛋白质以固定于细胞膜上的形式存在。跨膜区通常由蛋白质的氨基酸序列内富含疏水性氨基酸残基的区域组成。例如，一种市售的计算机程序 GCG 序列分析软件包(Genetic Computer Group, Oxford Molecular Group, Inc.)可以很容易地预测蛋白质是否具有跨膜区。GPI 锚定型膜结合蛋白为经过 GPI 的修饰并通过 GPI 锚定于细胞膜脂质层的蛋白质(GPI 锚定型膜结合蛋白)。

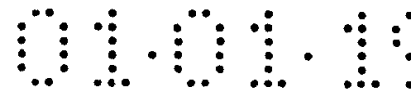
20

25

在本发明分离方法的第一步((i))中，将包括编码对一种抗原具有结合亲和力的可分泌性功能蛋白的 DNA 以及连接于功能性蛋白编码 DNA 3' 侧下游的 cDNA 的载体引入细胞。

“对一种抗原具有结合亲和力的功能性蛋白”意指功能上能够结合某种抗原的蛋白。作为功能性蛋白，优选与抗原的结合常数为 10^7M 或更高。更优选 10^8M 或更高，还更优选 10^9M 或更高。功能性蛋白特别为抗体、抗

30



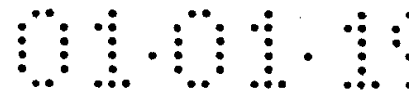
体片段、单链抗体等。抗体包括两条重链(H 链)和两条轻链(L 链), 这些 H 链和 L 链通过二硫键结合形成单个抗体分子。H 链和 L 链均由一个可变区(V 区, Fv)和一个恒定区(C 区, Fc)组成。抗体片段为抗体中对抗原具有结合亲和力的部分蛋白, 例如可以为 Fab、F(ab')₂、Fv 等。单链抗体(此后称为单链 Fv(scFv))为对抗原具有结合亲和力的蛋白, 在该蛋白中通过一个连接序列将 H 链 Fv 和 L 链 Fv 相连, 如单价单链抗体和二价单链抗体。单价单链抗体具有包括一个 H 链 Fv 和 L 链 Fv 的抗原结合位点, 而二价单链抗体具有通过一个连接序列将两个单价单链抗体分子相连的结构, 因而具有两个抗原结合位点。

10 为了各种目的, 例如改善结合常数, 可以将抗体、抗体片段或单链抗体删除、插入和/或用其它氨基酸残基置换一个或多个氨基酸残基, 或者将其与其它肽或多肽融合, 而这些都包括在本发明的功能性蛋白之中。同样, 修饰的抗体可以用作抗体、抗体片段或单链抗体。修饰抗体的实例为嵌合抗体和人源化抗体。嵌合抗体包括源自不同动物抗体的 V 区和 C 区。人源化抗体包括源自动物而并非人的抗体的互补决定区(CDR)以及源自人的抗体的框架区(FR)和 C 区。

15 对本发明的功能性蛋白具有结合亲和力的抗原可以是任何物质, 只要其具有抗原性。实例为蛋白质、肽和糖等, 优选为蛋白质。用作抗原的蛋白质如细胞或微生物表达的蛋白、血清蛋白、细胞因子、细胞内蛋白、膜蛋白等。

20 通过众所周知的方法可以获得编码抗体的 DNA。换言之, 它们可以分离自产生抗体的细胞, 例如杂交瘤、抗原致敏的永生化淋巴细胞以及在引入一种抗体基因后产生重组抗体的细胞。此外, 也可以利用已经分离并插入到载体中的 DNA。只要能够用于本发明, 编码抗体的 DNA 的来源和类型是无须讨论的。

25 通过下面通常采用的方法可以由编码抗体的 DNA 构建编码抗体片段或单链抗体的 DNA。通过连接编码抗体 H 链 V 区(H 链 Fv)、编码连接序列以及编码 L 链 V 区(L 链 Fv)的 DNA 获得编码单价单链抗体的 DNA。连接序列不限, 只要其能够在空间上重建 H 链 Fv 和 L 链 Fv 以使它们具有抗原亲和力。其优选为肽连接序列, 例如包括 12 至 19 个氨基酸残基(Huston, J.S. 等, 美国国家科学院院报(1988)85, 5879-5883)。特别为具有下述氨基酸序



列的肽连接序列：GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer((Gly₄Ser)₃)(SEQ ID NO: 1)。通过利用编码一种肽连接序列的 DNA 连接两个编码单价单链抗体的 DNA 分子的 5' 末端和 3' 末端构建编码二价单链抗体的 DNA。连接两个单链抗体的肽连接序列包括，例如

5 GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySer((Gly₄Ser)₃)(SEQ ID NO:1) 的氨基酸序列。

在本发明中为了提高克隆效率，例如当利用单链 Fv 作为功能蛋白时，优选 C 末端含有少量的疏水性氨基酸，特别是可以利用如在下面实施例中所述的转角(elbow)区已被删除的单链 Fv。同样，在本发明中还优选通过在

10 单链 Fv 的 C 末端进一步连接一个分泌性蛋白来源的结构域，例如编码下面实施例中所述抗体恒定区氨基酸的 DNA，以增加稳定性和表达效率。

分泌功能性蛋白，可以利用分泌信号序列。换言之，在编码对一种抗原具有结合亲和力的功能性蛋白的 DNA 5' 侧上游连接一个编码分泌信号序列的 DNA 便足够了。作为分泌信号序列，采用对用于 cDNA 文库表达和

15 蛋白分泌的细胞适宜的分泌信号序列。分泌信号序列可以为任何分泌性蛋白的信号序列，只要其能够分泌该功能性蛋白。优选的动物来源的分泌信号序列为那些来自哺乳动物的分泌信号序列，例如人免疫球蛋白(Kabat, E. 等, 免疫学重要蛋白的序列, US Department of Health and Human Services(1991))、细胞因子以及细胞因子受体的信号序列。

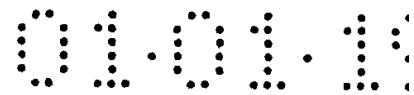
20 连接至编码功能性蛋白的 DNA 3' 侧下游的 cDNA 优选来自 cDNA 文库。cDNA 文库可使用经众所周知的方法获得的文库或者市售的文库。通过从目的样品中分离 mRNA 并由所分离的 mRNA 合成 cDNA 可以制备 cDNA 文库。

可使用能从中分离 mRNA 的来源如哺乳动物、除哺乳动物外的动物、

25 植物、酵母、细菌或蓝绿藻，优选哺乳动物。哺乳动物的实例包括人、猴、兔、大鼠、小鼠等，特别优选人。除哺乳动物外的动物为例如果蝇(Drosophila)等昆虫。

可以从其中分离 mRNA 的样品来源可以是任何来源，例如取自活体的细胞、建立的细胞系、胚胎、组织、血液或者器官。代表性的实例为成骨

30 细胞、造血干细胞、平滑肌细胞、神经元、基质细胞、ES 细胞、肝脏、肠、肺、肾脏、淋巴结等。



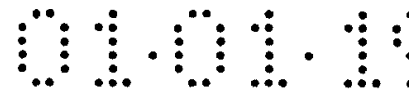
通过常用的方法在常用缓冲液存在下悬浮用于分离的样品可以进行 mRNA 的分离。为了制备完整的 mRNA 作为 mRNA 分离的第一步，可以采用例如鸟苷超速离心方法(Chirgwin, J.M.等, 生物化学(Biochemistry)(1979)18, 5294-5299)或者 AGPC 方法(Chomczynski, P.和 Sacchi, N., 分析生物化学(Anal. Biochem.)(1987)162, 156-159)等。然后, 5 为了从完整的 mRNA 中纯化 mRNA, 可以利用例如 mRNA 纯化试剂盒(Pharmacia)等。例如, QuickPrep mRNA 纯化试剂盒(Pharmacia)也可以作为用寡聚 dT 通过亲和纯化浓缩 mRNA 的市售试剂盒。

利用反转录酶由所获的 mRNA 合成 cDNA。可以使用市售的反转录酶。10 通过利用与 mRNA 的多聚 A 互补的寡聚 dT 或者利用随机序列寡核苷酸作为引物可以合成与 mRNA 互补的单链 cDNA。例如, AMV 反转录酶第一链 cDNA 合成试剂盒(Seikagaku 公司)等可以用来合成 cDNA。通过 DNA 聚合酶由所获的单链 cDNA 制备双链 cDNA。

此外, 为了特定的目的还可以利用常用的方法对 cDNA 文库进行选择15 性地富集。为了特定的目的, 例如为了获得表达量不同的基因的 cDNA, 可以利用差异克隆的方法(Lau, L.F.和 Nathans, D., 欧洲分子生物学组织杂志(EMBO J.)(1985)4, 3145-3151)、差异显示方法(Liang, P.和 Pardee, A.B., 科学(1992)257, 967-971)、扣除克隆的方法(核酸研究(Nucleic Acids Research)(1988)16, 10937)或者基因表达的系列分析方法(SAGE 方20 法)(Velculescu, V.E.等, 科学(1995)270, 484-487)。SST 方法(Tashiro, K.等, 科学(1993)261, 300-603)和在美国专利号 5,536,637 中所述的方法也可以用来富集编码分泌性蛋白的 cDNA。

载体可以为任何载体, 只要其可以转化细胞并表达其中含有的 DNA。表达载体优选在待转化细胞中可以操作的载体。表达载体的实例为质粒载25 体和病毒来源的载体。

将所获的 cDNA 连接至一个载体。在这种情况下, 可以通过将 cDNA 引入已经包含在载体中的功能性蛋白编码 DNA 的 3' 侧下游而将 cDNA 引入载体。为了这个目的, 在编码功能性蛋白的 DNA 的 3' 侧下游设计适当的限制性酶位点, 例如多克隆位点, 并将 cDNA 引入该位点。同样, 可以30 将 cDNA 首先连接到编码功能性蛋白的 DNA 的下游, 然后将所获的 DNA 引入载体。可以将 DNA 构建体引入载体 DNA 中包含的适当的限制性



酶位点。在制备载体时，可以将编码功能性蛋白的 DNA 和位于 3' 侧下游的 cDNA 直接连接，或者可以通过编码肽连接序列的 DNA 连接以便能够使功能性蛋白易于结合抗原。

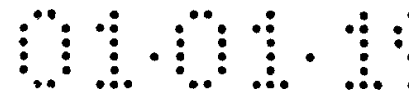
5 表达载体优选包括目的 DNA 在细胞中表达所需的表达调节区。启动子/增强子可以作为表达调节区，具体可以利用人 EF1 α 启动子、HCMV 启动子或者 SV40 启动子等。利用常用的方法可以将以这种方式制备的表达载体引入细胞。此类方法的实例为电穿孔法(欧洲分子生物学组织杂志(1982)1, 841-845)、磷酸钙法(病毒学(Virology)(1973)52, 456-467)、脂质体法、DEAE 葡聚糖法等。

10 用于转化的细胞可以为任何细胞，只要载体中包含的分泌信号序列和表达调节区在细胞内优选为动物细胞，例如 COS、CHO 或者 BAF3 等中发挥功能。

15 在本发明方法的第二步((ii))中，将对抗原具有结合亲和力的可分泌性功能蛋白与由一种 cDNA 编码的蛋白的融合蛋白在细胞内进行表达。具体为，利用含有编码上述融合蛋白的 DNA 的载体转化细胞，并在适于细胞生长的条件下进行培养。按照常用的方法进行培养。例如，DMEM、MEM、RPMI1640 以及 IMDM 可以用作培养基，而且可以与添加血清的溶液例如胎牛血清(FCS)一起使用。

20 为了在细胞内表达 DNA，可以使用诱导 DNA 表达的系统。例如，如果使用利用四环素的表达调节系统或针对如细胞因子、脂多糖(LPS)、类固醇激素等的刺激而表达的启动子/增强子，那么通过刺激细胞有可能诱导细胞内 DNA 的表达。在 DNA 表达时产生含有功能性蛋白和 cDNA 的基因产物的融合蛋白。在 cDNA 编码一种膜结合蛋白时，融合蛋白在粗面内质网(ER)上的合成过程中分泌信号序列被去除，而融合蛋白表达于细胞膜上。在将
25 编码肽连接序列的 DNA 连接在编码功能性蛋白的 DNA 和 cDNA 之间时，表达的融合蛋白在功能性蛋白和 cDNA 之间含有肽连接序列。

30 本发明方法的第三步((iii))涉及通过将细胞膜上表达融合蛋白的细胞与抗原接触来筛选结合抗原的细胞。抗原优选结合于一种支持物。支持物的实例为用于细胞培养的那些支持物，优选培养板，例如塑料板、多孔板、培养板或珠。磁珠可以用作珠。利用常用的方法可以将抗原结合于支持物。例如，通过在适当的缓冲液存在下将抗原加入培养板，放置过夜，然后洗



5 涂，可以将抗原结合于支持物。通过特异性结合抗原的抗体可以将抗原结合于支持物。例如，在加入特异性结合抗原的抗体并固定于培养板上后，可以加入抗原以使其结合于支持物。或者，首先可以结合不与支持物和细胞结合的抗原，然后利用特异性结合支持物上的固定抗原的抗体可以将细胞结合于支持物。在与未结合于支持物和细胞的抗原结合后，通过交联剂例如 DMS(dimethylsulberimidate)、BS³[双(磺基琥珀酰亚氨基)辛二酸酯(bis(sulfosuccinimididyl)suberate)]、以及 DSS(二琥珀酰亚氨基辛二酸酯, disuccinimidyl suberate)可以将抗原和细胞交联。

10 将未结合抗原的细胞移去，通过在细胞可以结合培养板上的抗原的条件下温育并在细胞结合抗原后于适当的条件下洗涤培养板，可以筛选结合于抗原的细胞。流式细胞术(FACS)也可以用来筛选结合于抗原的细胞。收集通过此类方法筛选的细胞。通过重复这些方法二至数次，可以更选择性地获得目的细胞。

15 本发明方法的第四步((iv))涉及从筛选的细胞中分离插入到载体中的cDNA。首先从已含有载体并结合于培养板上的细胞中提取载体，分离在载体中所包含的cDNA。当利用质粒载体时，提取质粒载体，引入大肠杆菌，在其中扩增，并制备来分离cDNA。随后，确定所分离基因的核苷酸序列。或者，基于载体上的核苷酸序列设计PCR引物，利用其对cDNA进行扩增，并确定核苷酸序列。当利用逆转录病毒载体时，以类似的方式通过PCR扩
20 增cDNA，并确定核苷酸序列。

本发明的方法可能包括用于确定上述分离的基因是否包含新序列的分析步骤。利用DNA数据库，例如GENBANK、EMBL等，进行序列的同源性(氨基酸残基的等同性)检索可以分析所分离DNA序列的新颖性。可以遵照在“Wilbur, W.J.和 Lipman, D.J., 美国国家科学院院报(1983)80,
25 726-730”中所述的算法来确定蛋白质的同源性。

本发明的方法还可以包括筛选cDNA文库来获得上述分离基因的全长基因的步骤。按照常用的方法，如下可以对cDNA文库进行筛选。首先，对所分离基因的一个片段进行标记，用作探针，并同cDNA文库进行杂交。然后利用该标记物检测与分离基因的片段结合的cDNA克隆。

30 本发明的方法还可以包括对上述分离基因的全长基因进行分离的步骤。这可以通过如上所述筛选cDNA文库，分离通过通常已知的方法检测

的 cDNA 克隆，并确定其核苷酸序列来进行。

此外，本发明包括一种用于分离编码上述膜结合蛋白的基因的试剂盒。本发明的试剂盒包括一种载体，其具有用于在编码对一种抗原具有结合亲和力的可分泌性功能蛋白的 DNA 3' 侧下游插入 cDNA 的限制性酶识别位点。本发明的试剂盒进一步优选包括结合有一种抗原的支持物和/或其中待引入载体的细胞。另外，还可以包括用于淘洗的溶液、用于细胞与抗原桥联的交联剂、cDNA 文库、用于通过溶解筛选的细胞收集 DNA 的溶液等。

附图简述

10 图 1 用示意图显示表达克隆载体 pTMT-SR345 的结构。图中的“SR345”表示人 IL-6 受体胞外区，“NEO^r”新霉素抗性基因，“EF1 α ”肽链延伸因子 1 α 的启动子/增强子区，“SV40E”SV40 早期启动子/增强子，而“Amp^r”为氨苄青霉素抗性基因。

15 图 2 用示意图显示表达克隆载体 pTMT-scFv 的结构。图中“scFv”表示单链抗体，而“Ig' s”为抗体分泌信号肽。其它符号同图 1。

图 3 显示利用已经引入各种类型质粒 DNA 的 COS-7 细胞通过分选回收的菌落数。

图 4 用示意图显示表达克隆载体 pTMT-BvGS3 的结构。图中“hPM1-BvGS3”表示二价单链抗体。其它符号同图 1 和图 2。

20 图 5 显示在利用抗人源化 PM-1 抗体的兔多克隆抗体通过流式细胞仪对已经引入各种类型质粒 DNA 的 COS-7 细胞进行分析时获得的直方图。

图 6 显示在利用抗小鼠抗 IL-6 受体抗体 MT-18 通过流式细胞仪对已经引入各种类型质粒 DNA 的 COS-7 细胞进行分析时获得的直方图。

25 图 7 用示意图显示表达克隆载体 pTMT-shPM1F- κ 的结构。图中“shPM1- κ ”表示一种单链抗体。其它符号同图 1 和图 2。

实施本发明的最佳模式

本发明的克隆方法可以按照下面所描述的特异地实施，但是本发明决非仅限于此。

30

实施例 1：表达克隆载体 pTMT-SR345 的构建

构建了表达克隆载体 pTMT-SR345。SR345，由表达克隆载体 pTMT-SR345 中所包含的 DNA 编码，是人 IL-6 受体的胞外区部分，包括从 N 端起

5 SR345 的核苷酸序列和氨基酸序列一起示于 SEQ IN NO: 2 中。

首先，为了从 IL-6 受体的 cDNA 扩增大约 1.1 kb 的含有编码 SR345 的 cDNA 的片段(Yamasaki, K.等, 科学(1988)241, 825-828)，设计了 PCR 引物 IL6R1(SEQ ID NO: 3)和 IL6R2(SEQ ID NO: 4)。将含有 10mM Tris-HCl(pH8.3)、50mM KCl、0.1mM dNTPs、1.5mM MgCl₂、各 100pmol 的上述引物、100ng 模板 DNA(编码 IL-6 受体的 cDNA)以及 5 单位 AmpliTaq Gold 酶的 PCR 反应混合物(100ml)在 94℃ 进行变性，按照 94℃ 1 分钟、55℃ 1 分钟以及 72℃ 1 分钟温育 30 个循环，最后于 72℃ 温育 10 分钟。收集并通过 1%的低熔点琼脂糖凝胶电泳纯化扩增的 DNA 片段，EcoRI 消化，并插入到表达载体 pCOS1 的 EcoRI 位点中。将其转染到大肠杆菌中，并制备质粒来获得 DNA 片段以正确的方向插入的那些质粒。质粒 HEF-PMh-g γ 1(见 WO92/19759)通过 EcoRI 和 SmaI 消化删除所含的基因并用 EcoRI-NotI-BamHI 接头(TaKaRa)连接构建成表达载体 pCOS1。

随后，用以下方法去除在 SR345 上游侧的 EcoRI 位点。首先，由 EcoRI 对质粒进行部分消化，收集通过在一个位点的切割所获得的线性分子。将其通过 DNA 聚合酶 I(Klenow 片段)使末端钝性化，自身连接，并转染到大肠杆菌中以获得表达克隆载体 pTMT-SR345。表达克隆载体 pTMT-SR345 的结构示于图 1 中。

实施例 2：表达载体 pTMT-scFv 的构建

25 构建了表达载体 pTMT-scFv。利用识别人 IL-6 受体的人源化单克隆抗体 PM-1 的可变区以及一个连接序列区设计由在表达载体 pTMT-scFv 中包含的 DNA 编码的单链抗体(scFv)。在表达载体 pTMT-scFv 中，由插入在编码 scFv 的 DNA 下游的 cDNA 编码的蛋白与 scFv 以融合蛋白的形式表达。scFv 基因的核苷酸序列和氨基酸序列一起示于 SEQ IN NO: 5 中。

30 1)编码抗体 V 区的 DNA 片段的扩增

通过 PCR 扩增人源化 PM1 抗体 H 链和 L 链 V 区的基因(Sato, K 等,

癌症研究(Cancer Res.)(1993)53, 851-856)。设计 H 链 V 区的反向引物 TMT1(SEQ ID NO: 6)使其可以与编码 H 链 V 区 N 末端的 DNA 杂交而且包括一个 SalI 限制性酶识别位点。设计 H 链 V 区的正向引物 LINK1(SEQ ID NO: 7)使其可以与编码 H 链 V 区 C 末端的 DNA 杂交而且包括连接序列区的 5' 末端序列。同样, 设计 L 链 V 区的反向引物 LINK3(SEQ ID NO: 8)使其可以与编码 L 链 V 区 N 末端的 DNA 杂交而且包括连接序列区的 3' 末端序列。设计 L 链 V 区的正向引物 SCP-C(SEQ ID NO: 9)使其可以与编码形成 L 链恒定区转角位点的氨基酸序列的核苷酸序列杂交而且还包括 HindIII 限制性酶识别位点、编码 FLAG 肽(SEQ ID NO: 10)的核苷酸序列以及两个重复翻译终止密码子。

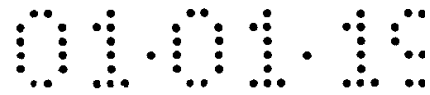
将含有 10mM Tris-HCl(pH8.3)、50mM KCl、0.1mM dNTPs、1.5mM MgCl₂、各 100pmol 的上述引物、100ng 模板 DNA 以及 5 单位 AmpliTaq Gold 酶的 PCR 反应混合物(100ml)在 94℃ 变性 9 分钟, 按照 94℃ 30 秒以及 60℃ 1 分钟温育 30 个循环, 最后于 60℃ 温育 5 分钟。利用 1.5%的低熔点琼脂糖凝胶对 PCR 产物进行纯化。

2)编码连接序列区的 DNA 片段的扩增

通过 PCR 方法利用人源化的单链抗体表达载体 pSCFVT7-hM21(见 WO95/14041)对编码包括氨基酸序列(Gly₄Ser)₃的连接序列区的 DNA 片段进行扩增。设计反向引物 LINK2(SEQ ID NO: 11)使其可以与该连接序列区的 5' 末端杂交而且还包括 H 链 V 区 3' 末端 DNA 序列。设计正向引物 LINK4(SEQ ID NO: 12)使其可以与连接序列区的 3' 末端杂交而且还包括 L 链 V 区 5' 末端 DNA 序列。利用 100ng 模板 DNA(pSCFVT7-hM21)在上述条件下进行 PCR, 并利用 1.5%的低熔点琼脂糖凝胶对 PCR 产物进行纯化。

3)人源化 PM1 抗体单链 Fv 的构建

上面制备的编码 H 链和 L 链 V 区的 DNA 片段, 以及编码连接序列区的 DNA 片段用 PCR 方法组装, 加入反向引物 TMT1 和正向引物 TMT2(SEQ ID NO: 13)来扩增编码人源化 PM1 的 scFv 的全长 DNA 片段。设计正向引物 TMT2 使其可以与编码 HindIII 限制性酶识别位点和 FLAG 肽的 DNA 序



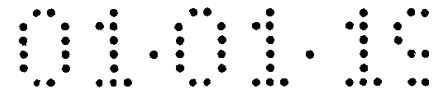
列杂交，而且还包括两个重复的翻译终止密码子以及 EcoRI 限制性酶识别位点。如下进行初级 PCR：将含有 10mM Tris-HCl(pH8.3)、50mM KCl、0.1mM dNTPs、1.5mM MgCl₂、大约各 100ng 的上述 PCR 产物以及 5 单位 AmpliTaq Gold 酶的 PCR 反应混合物 98ml 首先在 94℃ 变性，然后按照 94℃ 2 分钟、55℃ 2 分钟以及 72℃ 2 分钟进行 2 个循环来连接各 DNA 片段。以下列方式进行次级 PCR：将 100pmol 的每种引物加入到上面的 PCR 反应溶液中，按照 94℃ 30 秒和 60℃ 1 分钟进行 30 个循环，最后将混合物在 60℃ 温育 5 分钟。

PCR 产物用 1.5% 的低熔点琼脂糖凝胶纯化后，由 SalI 和 NotI 对其消化，并插入到包括人 EF1 α 启动子和抗体的前导序列(SEQ ID NO: 14)的表达载体 pSFLAG 中。进行 DNA 测序后，获得了质粒 pTMT-scFv，其含有包括正确 DNA 序列的 DNA 片段。表达载体 pTMT-scFv 的结构以示意图示于图 2 中。如下所述进行 pSFLAG 的构建。设计并合成两个分别为正义和反义方向的重叠的寡核苷酸 S-FLAG1(SEQ ID NO: 15)和 S-FLAG2(SEQ ID NO: 16)以便编码 EcoRI 限制性酶识别位点、抗体的前导序列(SEQ ID NO: 14)、FLAG 肽(SEQ ID NO: 10)以及 KpnI、NotI 和 BamHI 限制性酶识别位点。将含有合成的寡核苷酸各 100pmol 的反应混合物在 96℃ 温育 5 分钟，在 20 分钟内将温度降至 65℃，并在 65℃ 温育 5 分钟。然后，在 20 分钟内将温度降至 42℃，将混合物进一步温育 5 分钟，通过在 20 分钟内将温度降至室温使两种寡核苷酸退火。将该 DNA 片段插入到由 EcoRI 和 BamHI 消化的 pCOS1 中。

实施例 3：SR345-gp130 和 scFv-gp130 融合蛋白表达系统的构建

(A)SR345-gp130

细胞因子信号转导分子 gp130 是一种 I 型膜结合蛋白(Taga, T.等, 细胞(Cell)(1989)58, 573-581、Saito, M.等, 免疫学杂志(J. Immunol.)(1992)148, 4066-4071)。将小鼠 gp130 cDNA 的一部分连接至表达载体 pTMT-SR345 上编码可溶型 IL-6 受体(SR345)的 cDNA 的下游，以便在 COS 细胞中表达一种包括 SR345 和小鼠 gp130 的一部分序列的融合蛋白。根据两种类型的融合蛋白在 gp130 部分区域中的差别构建了它们。它们中一种为膜结合融合蛋白(SR345-mgpTMIC)，其中连接了 gp130 的跨膜区和随后的胞内区，而



另一种为分泌性融合蛋白(SR345-mgpIC)，其中只连接了 gp130 的胞内区。
SEQ ID NO: 17 显示全长小鼠 gp130 的氨基酸序列和核苷酸序列。

1)膜结合融合蛋白 SR345-mgpTMIC 表达载体的建立

5 将全长小鼠 gp130 cDNA 用 EcoRI 进行消化获得大约 1.1kb 的 EcoRI 片段。该 EcoRI 片段编码小鼠 gp130(C 末端)第 603 位至第 917 位的氨基酸，而且含有小鼠 gp130 胞外区的一部分(15 个氨基酸)以及随后跨膜区和胞内区的全部。将该 EcoRI 片段插入到 pTMT-SR345 表达载体的 EcoRI 位点中以建立膜结合融合蛋白 SR345-mgpTMIC 表达载体。

10

2)分泌性融合蛋白 SR345-mgpIC 表达载体的建立

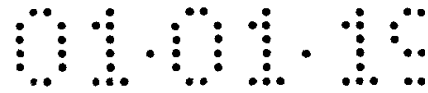
为了获得编码小鼠 gp130 胞内区的 cDNA 片段，合成已加入 HindIII 和 EcoRI 位点的 PCR 引物 mgp2(SEQ ID NO: 20; 包括编码 SEQ ID NO: 17 氨基酸序列的第 646 位氨基酸至第 651 位氨基酸的 DNA)和 mgp3(SEQ ID
15 NO: 19; 包括编码 SEQ ID NO: 17 氨基酸序列(C 末端)第 912 位氨基酸至第 917 位氨基酸的 DNA)，并利用这些引物获得了大约 1kb 的小鼠 gp130 的 cDNA 片段。该大约 1kb 的 cDNA 片段编码 SEQ ID NO: 17 的氨基酸序列中第 646 位氨基酸至第 917 位氨基酸，而且其对应于缺少从 N 末端起 6 个氨基酸的胞内区。将由此获得的 cDNA 片段用 EcoRI 进行消化，并插入到
20 pTMT-SR345 表达载体的 EcoRI 位点中以制备分泌性融合蛋白 SR345-mgpIC 表达载体。

(B)scFv-gp130

在表达载体 pTMT-scFv 中，将小鼠 gp130 cDNA 的一部分连接于 scFv
25 cDNA 的下游以便在 COS 细胞中表达一种包括 scFv 和小鼠 gp130 的部分区域的融合蛋白。根据两种类型的融合蛋白在 gp130 部分区域中的差别构建了它们。它们中一种为膜结合融合蛋白(scFv-mgpTMIC)，其中连接了 gp130 的跨膜区和随后的胞内区，而另一种为分泌性融合蛋白(scFv-mgpIC)，其中只连接了 gp130 的胞内区。

30

1)膜结合融合蛋白 scFv-mgpTMIC 表达载体的建立



5 为了获得编码小鼠 gp130 整个胞内区和跨膜区的 cDNA 片段，合成已加入 HindIII 位点的 PCR 引物 mgp1(SEQ ID NO: 18; 包括编码 SEQ ID NO: 17 氨基酸序列的第 603 位氨基酸至第 608 位氨基酸的 DNA)和已加入 NotI 位点的 mgp3(SEQ ID NO: 19; 包括编码 SEQ ID NO: 17 氨基酸序列(C 末端)第 912 位氨基酸至第 917 位氨基酸的 DNA)，并利用这些引物获得了大约 1.1kb 的小鼠 gp130 的 cDNA 片段。该大约 1.1kb 的 cDNA 片段编码 SEQ ID NO: 17 的氨基酸序列中(C 末端)第 603 位氨基酸至第 917 位氨基酸，而且该片段包括小鼠 gp130 胞外区的一部分(15 个氨基酸)和整个随后的跨膜区和胞内区。将由此获得的 cDNA 片段用 HindIII 和 NotI 进行消化，并插入到 pTMT-scFv 表达载体的 HindIII-NotI 位点中以制备膜结合融合蛋白 scFv-mgpTMIC 表达载体。

2)分泌性融合蛋白 scFv-mgpIC 表达载体的建立

15 为了获得编码小鼠 gp130 胞内区的 cDNA 片段，合成已加入 HindIII-EcoRI 位点的 PCR 引物 mgp2(SEQ ID NO: 20; 包括编码 SEQ ID NO: 17 氨基酸序列的第 646 位氨基酸至第 651 位氨基酸的 DNA)和已加入 NotI 位点的 mgp3(SEQ ID NO: 19; 包括编码 SEQ ID NO: 17 氨基酸序列(C 末端)第 912 位氨基酸至第 917 位氨基酸的 DNA)，并利用这些引物通过 PCR 获得了大约 1kb 的小鼠 gp130 的 cDNA 片段。

20 该大约 1kb 的 cDNA 片段编码 SEQ ID NO: 17 的氨基酸序列中(C 末端)第 646 位氨基酸至第 917 位氨基酸，而且其对应于缺少从 N 末端起 6 个氨基酸的胞内区。将由此获得的 cDNA 片段用 HindIII 和 NotI 进行消化，并插入到 pTMT-scFv 表达载体的 HindIII-NotI 位点中以制备分泌性融合蛋白 scFv-mgpIC 表达载体。

25

实施例 4: 由 COS 细胞的表达

30 将上述每种类型表达载体转染到 COS 细胞中，瞬时表达融合蛋白，并且证实了在细胞膜上表达融合蛋白的细胞通过淘洗选择性凝集。以不含基因的表达载体转染的 COS 细胞被用作阴性对照。阳性对照为用表达载体 P3.19 转染并用相应的抗体淘洗的 COS 细胞，载体 P3.19 是通过向载体 pCOS1 中引入编码 HM1.24 抗原蛋白 (WO 98/14580)的 DNA 制备的。

1) 转染到 COS 细胞中

利用 Lipofect AMINE PLUS™ 试剂(GIBCO-BRL)将质粒 DNA 转染到 COS-7 细胞中。即, 在转染前一天将 COS-7 细胞以 1×10^5 个细胞/孔(6 孔板) 5 接种, 培养过夜并用无血清 DMEM 培养基(GIBCO-BRL)洗涤, 然后向其中加入 0.8ml 相同的培养基。在将 $1 \mu\text{g}$ 的质粒 DNA 和 $6 \mu\text{l}$ 的 PLUS 试剂分别加至 0.1ml 无血清 DMEM 培养基中后, 在室温下将混合物温育 15 分钟, 与 0.1ml 的 LipofectAMINE 溶液($4 \mu\text{l}$ 的 LipofectAMINE/0.1ml 的无血清 DMEM 培养基)混合, 并在室温下继续温育 15 分钟。随后, 将该混合物加 10 至上述 COS-7 细胞并在 37°C 温育 3 小时。向其中加入含有 20%胎牛血清(GIBCO-BRL)的 DMEM 培养基(1ml)(终浓度 10%的血清)。培养过夜后, 将培养基更换为 3ml 含有 10%胎牛血清的 DMEM 培养基, 并在 37°C 和 5% CO_2 条件下温育 3 天。

15 2) 淘洗皿的制备

在利用表达载体 pTMT-SR345 时, 按照 Seed, B.等, 美国国家科学院院报(1987)84, 3365-3369 的方法制备用小鼠抗人 IL-6 受体抗体 MT18(参见已公开未审查的日本专利申请第 Hei 2-288898 号)包被的培养皿。即, 将小鼠抗 IL-6 受体抗体加到 50mM Tris-HCl(pH9.5)中至 $10 \mu\text{g/ml}$ 。室温下将由 20 此制备的抗体溶液(3ml)在 60mm 直径细胞培养皿中温育 2 小时。用 0.15M NaCl 溶液洗涤培养皿 3 次后, 加入含有 5%胎牛血清、1mM EDTA 和 0.02% NaN_3 的 PBS, 然后在封闭后按照下面所述进行淘洗。

在利用 pTMT-scFv 时, 制备两种类型的淘洗皿。一种用可溶型 IL-6 受体(SR344)(Yasukawa, K.等, 生物化学杂志(1990)108, 673-676)包被, 而另 25 一种用上述小鼠抗 IL-6 受体抗体包被。用 50mM Tris-HCl(pH9.5)将 SR344 的浓度调节至 $2 \mu\text{g/ml}$, 并如上所述制备淘洗皿。在利用阴性对照 pCOS-1 时, 采用由上述小鼠抗 IL-6 受体抗体包被的培养皿。在利用阳性对照 HM1.24 抗原蛋白表达载体 P3.19 时, 采用由抗 HM1.24 抗原的抗体包被的培养皿。

30 3) 淘洗

将 pCOS-1 或 pTMT-SR345 转染的 COS-7 细胞用含有 1mM EDTA 的 PBS

使其从培养板上脱离，用含有 5%胎牛血清的 PBS 洗涤一次，悬于 2ml 含有 5%胎牛血清和 0.02% NaN₃ 的 PBS 中，并加入到由小鼠抗 IL-6 受体抗体包被的淘洗板中。

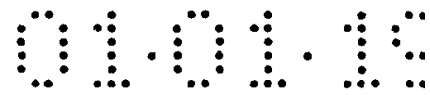
5 通过三种不同的方法对 pTMT-scFv 转染的 COS-7 细胞进行淘洗。在一种方法中，将 COS-7 细胞如上所述使其脱壁并用含有 5%胎牛血清的 PBS 洗涤一次后，悬于 500 μl 含有 2 μg/ml SR344、5%胎牛血清和 0.02% NaN₃ 的 PBS 中，并在冰上温育 1 小时。将细胞用冰冷的 PBS 洗涤三次后，重悬于含有 0.2mM 交联剂双(磺基琥珀酰亚氨)辛二酸酯(BS³; PIERCE)和 50mM Hepes(pH8.0)的 PBS 中，并在冰上进一步温育 30 分钟。然后，加入 1M
10 Tris-HCl(pH8.0)至 50mM，并在冰上继续温育 10 分钟以去除过量的交联剂。在用 PBS 洗涤细胞后，将其加至由小鼠抗 IL-6 受体抗体包被的淘洗板中。在第二种方法中，将与 SR344 预温育的 COS-7 细胞不经交联剂处理加至由小鼠抗 IL-6 受体抗体包被的淘洗板中。在第三种方法中，将 COS-7 细胞加至由 SR344 直接包被的培养板中。冰上温育、Tris-HCl 处理以及洗涤的时间在三种方法中都相似。
15

将由 HM1.24 抗原蛋白表达载体 P3.19 转染的 COS-7 细胞加至由抗 HM1.24 抗原的抗体(WO 98/14580)包被的淘洗板中。室温下将上述各种类型的 COS-7 细胞在各种淘洗板中温育 2 小时后，将淘洗板用含有 5%胎牛血清和 0.02% NaN₃ 的 PBS 轻轻洗涤三次，并利用含有 0.6% SDS 和 10mM
20 EDTA 的溶液从结合至淘洗皿上的细胞中收集质粒 DNA。通过利用电穿孔将所回收质粒 DNA 的 1/5 转染入大肠杆菌 DH5a 中，并通过出现的氨苄青霉素抗性菌落的数量对淘洗的聚集作用进行评价。结果示于图 3 中。

在利用表达载体 pTMT-SR345 时，SR345-mgpIC 产生比 SR345-mgpTMIC 多的菌落，因此未观察到针对膜结合蛋白的特异性。在另一方面，
25 在利用表达载体 pTMT-scFv 时，在所有的淘洗方法中，scFv-mgpTMIC 产生比 scFv-mgpIC 多的菌落，因此表达膜结合蛋白的细胞被特异性地富集。特别是在使用交联剂时，选择性更加明显。

因此，上述结果显示通过将功能蛋白(单链抗体)表达为细胞表面融合蛋白比通过仅仅将一种被该抗体识别的表位表达为融合蛋白更有选择性地
30 而且更有效地获得编码膜结合蛋白的 cDNA。

通常，重复淘洗数次提高克隆选择性，但是如本实施例所示，在本发



明中(其中功能蛋白表达在细胞膜上)在第一次淘洗中便观察到显著的选择性。因此,通过进一步淘洗数次可以非常有效而且选择性地完成编码一种膜结合蛋白的基因的克隆。

5 实施例 5: 利用人源化二价单链 Fv 构建融合蛋白表达系统

1. 人源化 PM1 抗体二价单链 Fv 表达载体的构建

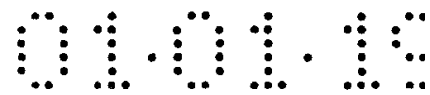
基于人源化 PM1 抗体 Fv 构建了一种二价单链 Fv 表达载体。设计具有二价可变区的人源化 PM1 抗体单链 Fv(hPM1-BvGS3)以便通过一个包括(Gly₄Ser)₃的肽连接序列(SEQ ID NO:1)连接在实施例 2 中所述的人源化 PM1 抗体单链 Fv2 的两个分子。hPM1-BvGS3 的氨基酸序列和核苷酸序列示于 SEQ ID NO: 21 中。

表达载体 pTMT-BvGS3 的构建按照下述方法进行。通过 PCR 方法对编码人源化 PM1 抗体单链 Fv 并在其 C 末端具有一个包括(Gly₄Ser)₃的连接序列的基因进行扩增。TMT-1(SEQ ID NO: 6)被用作反向引物。同样,设计正向引物 BvGS3(SEQ ID NO:22)使其可以与编码 L 链 V 区 C 末端的 DNA 杂交而且包括编码连接序列和 SalI 限制性酶识别位点的核苷酸序列。在与上述相同的条件下利用 100ng pTMT-scFv 作为模板 DNA 进行 PCR,而且利用 1.5%的低熔点琼脂糖凝胶对 PCR 产物进行纯化。

将纯化的 PCR 产物用限制性酶 SalI 进行消化,并插入到克隆载体 pBluescriptII(Stratagene)中。DNA 测序后,用限制性酶 SalI 对含有包括正确 DNA 序列的 DNA 片段的质粒进行消化,以获得编码人源化 PM1 抗体单链 Fv 并在其 C 末端具有一个包括(Gly₄Ser)₃的连接序列的基因。随后,通过将上述获得的 DNA 片段插入到 pTMT-scFv 中获得 hPM1-BvGS3 表达载体 TMT-BvGS3。hPM1-BvGS3 表达载体 TMT-BvGS3 的结构以示意图示于图 4 中。

2. 融合蛋白 hPM1-BvGS3-gp130 表达载体的构建

在表达载体 TMT-BvGS3 中,将小鼠 gp130 cDNA 的一部分连接于编码 hPM1-BvGS3 的 cDNA 的下游以构建一种包括 hPM1-BvGS3 和小鼠 gp130 的一部分区域的融合蛋白表达系统。根据所连接的 gp130 部分区域的差别构建两种类型的融合蛋白。其中一种为连接了 gp130 的跨膜区及随后的胞



内区的膜结合融合蛋白(BvGS3-mgpTMIC), 另一种为只连接了 gp130 的胞内区的分泌性融合蛋白(BvGS3-mgpIC)。BvGS3-mgpTMIC 和 BvGS3-mgpIC 分别在实施例 3(B)的 1)和 2)中制备, 而且通过将这些插入到 pTMT-BvGS3 的 HindIII-NotI 位点中, 构建了膜结合融合蛋白表达载体 pTMT-BvGS3-
5 mgpTMIC 和分泌性融合蛋白表达载体 pTMT-BvGS3-mgpIC。

实施例 6: 通过流式细胞仪对表达的分析

将上面构建的每种类型的表达载体, pTMT-BvGS3、pTMT-BvGS3-
10 mgpIC 和 pTMT-BvGS3-mgpTMIC 转染到 COS-7 细胞中, 瞬时表达该融合蛋白, 并利用流式细胞仪(FACScan, Beckton Dickinson)分析在细胞膜上的表达。通过两种类型的方法进行表达分析。一种涉及通过兔抗人源化 PM-1 抗体的多克隆抗体进行检测, 而另一种涉及在可溶性 IL-6 受体抗体的存在下通过小鼠抗 IL-6 受体抗体 MT-18 进行检测。结果证实膜结合融合蛋白 BvGS3-mgpTMIC 以可以识别可溶性 IL-6 受体的形式大量表达于细胞膜上。
15 将仅用表达载体 pCOS-1 转染的细胞用作阴性对照。

1)向 COS-7 细胞中的转染

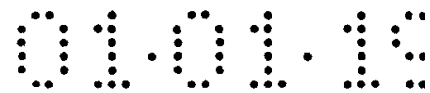
利用转染试剂盒 FuGENE™6(Boehringer-Mannheim)将质粒 DNA 转染到 COS-7 细胞中。

20 即, 在转染前一天将 COS-7 细胞以 5×10^4 个细胞/孔(6 孔板)接种, 于 37 °C 和 5% CO₂ 的条件下在 2ml 含有 10%胎牛血清的 DMEM 培养基(GIBCO-BRL)中培养过夜。在转染当天, 将 6 μl FuGENE™6 加入到 0.1ml 无血清 DMEM 培养基中, 在室温下温育 5 分钟后, 将其与 2 μg 的质粒 DNA 混合, 并在室温下继续温育 15 分钟。然后, 将该混合物加入到上述 COS-7 细胞中
25 并在 37°C 和 5% CO₂ 的条件下温育 3 天。

2)COS-7 细胞的染色

将上述 COS-7 细胞用含有 1mM EDTA 的 PBS 使其脱壁, 用含有 5%胎牛血清的 PBS 洗涤, 悬于 50 μl FACS 缓冲液(含有 2%胎牛血清和 0.05% NaN₃ 的 PBS)中, 并通过下面两种类型的方法进行染色。
30

A)用抗人源化 PM-1 抗体的兔多克隆抗体染色



将一种抗人源化 PM-1 抗体的兔多克隆抗体($2 \mu\text{g}$ /反应)加入到上述 COS-7 细胞中, 在冰上温育 30 分钟, 用 1ml FACS 缓冲液洗涤两次, 并悬于 $50 \mu\text{l}$ FACS 缓冲液中。随后加入 FITC(异硫氰酸荧光素)标记的山羊抗兔 IgG(AMERICAN QUAREX) $2 \mu\text{l}$ /反应和用于单独染色死细胞的 $2.5 \mu\text{g}$ /反应的 PI(碘化丙锭), 并在冰上避光温育 30 分钟。温育后, 将细胞用 1ml 的 FACS 缓冲液洗涤两次并重悬于 0.5ml 的 FACS 缓冲液中, 用流式细胞仪进行分析。

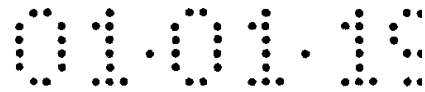
B) 用小鼠抗 IL-6 受体抗体 MT-18 染色

将可溶型 IL-6 受体($3 \mu\text{g}$ /反应)加入上述 COS-7 细胞中并于冰上温育 4 小时, 用 1 ml FACS 缓冲液洗两次, 重悬于 $50 \mu\text{l}$ FACS 缓冲液中。然后, 加入 $2 \mu\text{g}$ /反应的小鼠抗 IL-6 受体抗体 MT-18, 并于冰上温育 30 分钟。此后用 1 ml FACS 缓冲液将细胞洗两次, 重悬于 $50 \mu\text{l}$ FACS 缓冲液中。随后加入 FITC 标记的山羊抗小鼠 IgG2b(Dainippon Seiyaku) $2 \mu\text{l}$ /反应和用于单独染色死细胞的 PI(碘化丙锭) $2.5 \mu\text{g}$ /反应, 并在冰上避光温育 30 分钟。温育后, 将细胞用 1ml 的 FACS 缓冲液洗涤两次并重悬于 0.5ml 的 FACS 缓冲液中, 用流式细胞仪进行分析。

3)分析利用流式细胞仪进行的表达

通过 PI 和 FSC(前向散射)的分析表明存在由 PI 染色的细胞群(死细胞)。死细胞因其被 FITC 非特异性地染色而干扰分析。因此, 圈定不被 PI 染色的细胞群(活细胞), 并仅对该群进行分析。

用抗人源化 PM-1 抗体的兔多克隆抗体染色, 在细胞膜上未观察到分泌性蛋白 hPM1-BvGS3 的表达, 然而, 膜结合融合蛋白 BvGS3-mgpTMIC 的表达显示为最强。根据该事实推测具有一个跨膜区的 BvGS3-mgpTMIC 并未被分泌而是插入细胞膜中。然而, 另一方面, 虽然 BvGS3-mgpIC 为一种分泌性融合蛋白, 但是在细胞膜上仍检测到 BvGS3-mgpIC 的表达。这被认为是由于 BvGS3-mgpIC 在分子大小和结构、连接于其下游的 mgpIC 中所含的疏水区等特征上与 BvGS3 的特征不同。换言之, 推测由于在特征上的差别, BvGS3-mgpIC 不会向 BvGS3 那样迅速通过细胞膜, 而需要更长的时间来通过。结果, 定位于细胞膜上的数量增加, 通过抗人源化 PM-1 的兔多克隆抗体检测到所有表位向细胞外突出的分子。



另一方面，在用小鼠抗 IL-6 受体的抗体 MT-18 染色的情况下，与由抗人源化 PM-1 抗体的兔多克隆抗体染色的结果相似，虽然未检测到分泌性蛋白 hPM1-BvGS3 的表达，但膜结合融合蛋白 BvGS3-mgpTMIC 的表达最强。然而，对于分泌性融合蛋白 BvGS3-mgpIC，结果与通过抗人源化 PM-1 的兔多克隆抗体所得的结果不同，在细胞膜上几乎未检测到任何表达。这提示，虽然膜结合融合蛋白 BvGS3-mgpTMIC 表达于细胞膜上形成一种可以识别可溶型 IL-6 受体的功能性构象，但是大多数分泌性融合蛋白 BvGS3-mgpIC 尽管定位于细胞膜上，仍不具有可以识别可溶型 IL-6 受体的功能性构象。结果示于图 6 中。

10 因此，通过流式细胞仪获得的结果表明当将单纯一个由抗体识别的表位表达为融合蛋白时，即使分泌性融合蛋白如果其定位于细胞膜上也将被筛选为假阳性。另一方面，当将本发明的功能性蛋白(例如，单链抗体)以融合蛋白的形式表达于细胞表面时，揭示了更有选择性且更有效地克隆编码膜结合蛋白的 cDNA 的可能性。

15

实施例 7

1. 人源化 PM1 抗体单链 Fv 的设计

为了提高克隆效率，设计了三种其它类型的单链 Fv 及其二价单链 Fv。由于在构建上述人源化 PM1 抗体单链 Fv 时加入的转角区(SEQ ID NO: 5, 20 从第 242 位到第 256 位的氨基酸序列)含有具有高疏水性的氨基酸残基，所以设计下面三种类型的人源化 PM1 抗体单链 Fv 以便更稳定的细胞外表达。换言之，为了去除 C 末端的疏水区，设计了转角区缺失的单链 Fv 并指定为 shPM1(Δ EL)(SEQ ID NO: 23)。同样，由于发现通过在单链 Fv 的 C 末端增加某种分泌性蛋白衍生的结构域可以增加稳定性和表达效率，将由人 κ 链恒定区或人膜型 μ 链恒定区的第 4 外显子(Dorai, H 和 Gillies, S.D., 核 25 酸研究, 17, 6412, 1989)编码的氨基酸序列加到单链 Fv 的 C 末端(SEQ ID NO: 23)。虽然人 κ 链恒定区的第 107 位氨基酸原本为半胱氨酸，但此次由丝氨酸残基取代(SEQ ID NO: 24)。同样，将已删除跨膜区和胞内区的序列 (SEQ ID NO: 25)用作由人膜型 μ 链恒定区的第 4 外显子编码的氨基酸序 30 列。将在单链 Fv C 末端增加了上述各序列的单链 Fv 称为 shPM1-Kappa(SEQ ID NO: 26)和 shPM1-MCH4(SEQ ID NO: 27)。

2. shPM1(Δ EL)表达载体的构建

通过 PCR 方法扩增编码 shPM1(Δ EL)的基因。利用了反向引物 EF-1(SEQ ID NO: 28)和正向引物 SCP-C2(SEQ ID NO: 29)。在上述条件下利用 100ng 的 pTMT-scFv 作为模板 DNA 进行 PCR，并利用 1.5%低熔点琼脂糖凝胶对 PCR 产物进行纯化。正向引物 SCP-C2 与编码 L 链 V 区 C 末端的 DNA 杂交，并向其中加入 HindIII-NotI 识别位点和 FLAG 肽(SEQ ID NO: 10)。

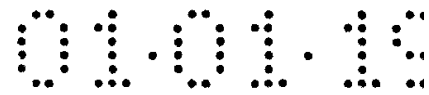
在用 EcoRI 和 NotI 对纯化的 PCR 产物进行消化后，将消化的产物插入到 pSFLAG 载体中以获得 shPM1(Δ EL)表达载体 pTMT-shPM1F。此外，通过与实施例 5 中的相似的方法获得了具有一个二价可变区并且缺失转角区的 shPM1(Δ EL)-BvGS3(SEQ ID NO: 30)的表达载体 pTMT-shPM1F-BvGS3。

3. shPM1- κ 表达载体的构建

通过 PCR 组装构建了编码人源化 PM1 抗体单链 Fv(SEQ ID NO: 23)与人 κ 链恒定区(SEQ ID NO: 24)的融合蛋白的基因。换言之，在通过 PCR 分别扩增编码人源化 PM1 抗体单链 Fv 和人 κ 链恒定区的基因后，通过它们的互补性进行组装并通过外部引物对全长基因进行扩增。

首先，通过 PCR 扩增编码人 κ 链恒定区的基因。设计反向引物 κ 1(SEQ ID NO: 31)使其可以与编码人 κ 链恒定区的转角区和下面氨基酸序列(SEQ ID NO: 24)的第 12 位 Pro 至第 21 位 Gly 的核苷酸序列杂交。设计正向引物 κ 2(SEQ ID NO: 32)使其可以与编码人 κ 链恒定区 C 末端(SEQ ID NO: 24)的第 101 位 Ser 至第 111 位 Ser 的核苷酸序列杂交并包括编码限制性酶 HindIII 和 NotI 识别位点和 FLAG 肽(SEQ ID NO: 10)的核苷酸序列以及两个终止密码子。通过利用这些引物，将原本为一个半胱氨酸残基的 SEQ ID NO: 24 的第 107 位氨基酸被一个丝氨酸残基取代。在与上述方式相似的条件，以人源化 PM1 抗体 L 链表达载体 RV1-PM1a(见 WO92/19759)作为模板利用上述两类引物进行 PCR。利用 1.5%低熔点琼脂糖凝胶对 PCR 产物进行纯化。

然后，以相同的方式扩增编码人源化 PM1 抗体单链 Fv 的基因。在与



上述相同的条件下，以 pTMT-scFv 为模板 DNA 利用 EF1(SEQ ID NO: 28) 作为反向引物，SCP-K(SEQ ID NO: 33)作为正向引物进行 PCR。设计正向引物 SCP-K 使其可以与在 SEQ ID NO: 5 中所示的编码单链 Fv C 末端的核苷酸序列杂交并包括与 PCR 扩增的 κ 链基因 5' 末端互补的核苷酸序列。

5 以同样的方式对 PCR 产物进行纯化。

利用在实施例 2-3)中所示的方法对编码 shPM1- κ 的全长 cDNA 片段进行扩增。即，通过初次 PCR 对各 100ng 的上述 DNA 片段进行组装，然后加入反向引物 EF-1(SEQ ID NO: 28)和正向引物 κ 2(SEQ ID NO: 32)各 100ng 来扩增全长 cDNA 片段。

10 将 PCR 产物用 1.5%低熔点琼脂糖凝胶纯化后，利用限制性酶 EcoRI 和 NotI 进行消化，并将其插入到 pSFLAG 载体中以获得 shPM1- κ 表达载体 pTMT-shPM1F-K(图 7)。此外，利用与在实施例 5 中所述相似的方法，获得了具有一个二价可变区的单链 Fv 和 shPM1-Kappa-BvGS3(SEQ ID NO: 34)表达载体 pTMT-shPM1FK-BvGS3。

15

4. shPM1-MCH4 表达载体的构建

通过 PCR 组装构建编码人源化 PM1 抗体单链 Fv(SEQ ID NO: 23)和人 μ 链恒定区部分序列(SEQ ID NO: 25)的融合蛋白的基因。即，通过 PCR 方法分别扩增编码人源化 PM1 抗体单链 Fv 和人 μ 链恒定区部分序列的基因，并通过它们的互补性进行组装。然后通过外部引物扩增全长基因。

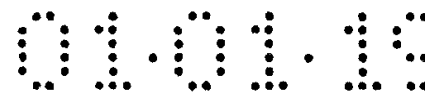
20

首先，通过 PCR 扩增编码人 μ 链恒定区的基因。设计反向引物 MCH4-1(SEQ ID NO: 35)使其可以与编码人 μ 链恒定区第 4 外显子的 5' 末端核苷酸序列杂交而且包括人源化 PM1 抗体单链 Fv(SEQ ID NO: 23)3' 末端的核苷酸序列。设计正向引物 MCH4-2.1(SEQ ID NO: 36)使其可以与编码人膜型 μ 链恒定区的胞外区的核苷酸序列杂交，而且包括限制性酶 HindIII 识别位点。在与上述方式相似的条件，以通过常用方法从人骨髓瘤细胞系 CL-4 细胞中获得的 cDNA 作为模板利用上述两类引物进行 PCR。利用 1.5%低熔点琼脂糖凝胶对 PCR 产物进行纯化。

25

然后，以相同的方式扩增编码人源化 PM1 抗体单链 Fv 的基因。在与上述相同的条件下，以 pTMT-scFv 为模板 DNA 利用 EF1(SEQ ID NO: 28)作为反向引物，SCP-Mu(SEQ ID NO: 37)作为正向引物进行 PCR。设计正

30



向引物 SCP-Mu 使其可以与在 SEQ ID NO: A 中所示的编码单链 Fv C 末端的核苷酸序列杂交，而且包括与通过 PCR 扩增的 μ 链部分序列基因 5' 末端互补的核苷酸序列。以同样的方式对 PCR 产物进行纯化。

5 利用在实施例 2-3) 中所示的方法对编码 shPM1-MCH4 的全长 cDNA 片段进行扩增。即，通过初次 PCR 对各 100ng 的上述 DNA 片段进行组装，然后加入反向引物 EF-1(SEQ ID NO: 28)和正向引物 MCH4-2.2(SEQ ID NO: 38)各 100pmol 来扩增全长 cDNA 片段。设计正向引物 MCH4-2.2 使其可以与上面扩增的编码人膜型 μ 链部分序列的核苷酸序列的 3' 末端杂交，而且包括编码 FLAG 肽的核苷酸序列、两个终止密码子以及限制性酶 NotI
10 识别位点。

将 PCR 产物用 1.5%低熔点琼脂糖凝胶纯化后，利用限制性酶 EcoRI 和 NotI 进行消化，并将其插入到 pSFLAG 载体中以获得 shPM1-MCH4 表达载体 pTMT-shPM1F-MCH4(图 7)。此外，利用与在实施例 5 中所述相似的方法，获得了具有一个二价可变区的单链 Fv 和 shPM1-MCH4-BvGS3(SEQ
15 ID NO: 39)表达载体 pTMT-shPM1FM-BvGS3。

实施例 8:通过 TMT 方法利用 shPM1- κ 表达载体(图 7)对 STX561 cDNA 文库的筛选

1. STX561 cDNA 文库的制备

20 利用常用的方法制备来自小鼠造血基质细胞系 STX561 的 mRNA，并将由其合成的 cDNA 插入到 TMT 表达载体 shPM1- κ 中来制备 STX561 cDNA 文库。利用 cDNA 合成试剂盒(STRATAGENE, cDNA 合成试剂盒)制备 cDNA 文库。基本上按照 STRATAGENE cDNA 合成试剂盒的操作，并进行了下述的调整。换言之，将 GIBCO-BRL 的 Superscript II 用作反转录酶，将 NotI-dT
25 引物(Pharmacia Biotech, 第一链 cDNA 合成试剂盒所附引物)作为第一链合成的引物，将 HindIII-SmaI 位点接头用作加至 cDNA 5' 末端的接头，而且将 Pharmacia Biotech 的 Size sep 400 离心层析柱用作大小分级的柱子。

具体如下所述制备 cDNA 文库：起始材料为 5 μ g 的 mRNA，首先利用 NotI-dT 引物(Pharmacia Biotech, 第一链 cDNA 合成试剂盒所附引物)由
30 反转录酶(Superscript II, GIBCO-BRL)从 3' 多聚 A 尾合成第一链。然后，在用 DNA 聚合酶合成第二链后，将 cDNA 的两个末端钝性化，并加入

HindIII-SmaI 位点接头(Takara)。用 HindIII 和 NotI 对两个末端消化后, 进行大小分级分离(Pharmacia Biotech, Size sep 400 离心层析柱)以去除大小为 0.5kb 或更小的 cDNA 片段。将收集的 cDNA 插入到 TMT 表达载体 shPM1- κ 的 HindIII-NotI 位点中, 并通过电穿孔方法将载体引入大肠杆菌 DH10B(electroMAX DH10B, GIBCO-BRL)中来制备 STX561 cDNA 文库。

通过分成 1000 个克隆/库将 STX561 cDNA 文库分库, 其中两个库(库号: #kappa-1, #kappa-6)一共 2000 克隆用于通过 TMT 方法进行筛选。

2. 通过淘洗对 STX cDNA 文库的筛选

10 1)向 COS-7 细胞中的转染

利用 FuGENETM6(Boehringer-Mannheim)将从#kappa-1 和#kappa-6 制备的质粒 DNA 各 2 μ g 转染到 COS-7 细胞中。

即, 在转染前一天将 COS-7 细胞以 1×10^5 个细胞/孔(6 孔板)接种, 于 37 $^{\circ}$ C 和 5% CO₂ 的条件下在 2ml 含有 10%胎牛血清的 DMEM 培养基(GIBCO-BRL)中培养过夜。在转染当天, 将 6 μ l FuGENETM6 加入到 0.1ml 无血清 DMEM 培养基中, 在室温下温育 5 分钟后, 将其与 2 μ g 的质粒 DNA 混合, 并在室温下继续温育 15 分钟。然后, 将该混合物加入到上述 COS-7 细胞上并在 37 $^{\circ}$ C 和 5% CO₂ 的条件下温育 3 天。

20 2)淘洗板的制备

按照 Seed, B.等, 美国国家科学院院报(1987)84, 3365-3369 的方法制备由山羊抗小鼠 IgG 抗体(Dainippon Seiyaku, 山羊抗小鼠 IgG(H+L 链))包被的淘洗板。即, 将山羊抗小鼠 IgG 抗体加到 50mM Tris-HCl(pH9.5)中至 10 μ g/ml。将由此制备的抗体溶液(3ml)加至 60mm 直径细胞培养皿中, 并在 25 室温下温育 3 小时。用 0.15M NaCl 溶液洗涤 3 次后, 加入含有 5%胎牛血清、1mM EDTA 和 0.02% NaN₃ 的 PBS, 然后在封闭后按照下面所述进行淘洗。

3)淘洗

30 将如上所述转染的 COS-7 细胞用含有 1mM EDTA 的 PBS 使其从培养板上脱离, 用含有 5%胎牛血清的 PBS 洗涤一次, 悬于 50 μ l FACS 缓冲液(含

有 2%胎牛血清和 0.05% NaN_3 的 PBS) 中。

将可溶型 IL-6R($2\ \mu\text{g}$)加入到细胞悬液中，并在冰上放置 90 分钟。随后，在用 FACS 缓冲液洗涤两次后，将细胞悬于 $50\ \mu\text{l}$ FACS 缓冲液中。然后，向细胞悬液中加入 $1.5\ \mu\text{g}$ 小鼠抗 IL-6 受体抗体 MT-18，并将细胞悬液
5 在冰上放置 30 分钟。将细胞用 FACS 缓冲液洗涤两次，悬于 2ml 含有 5% 胎牛血清和 0.02% NaN_3 的 PBS 中，并加入到用山羊抗小鼠 IgG 抗体包被的淘洗板中。

室温下将上述各种 COS-7 细胞在各种淘洗板中温育大约 2 小时后，将淘洗板用含有 5%胎牛血清和 0.02% NaN_3 的 PBS 轻轻洗涤三次，并利用 Hirts
10 溶液(含有 0.6% SDS 和 10mM EDTA 的溶液)从结合至淘洗板上的细胞中收集质粒 DNA。通过电穿孔方法将所回收质粒 DNA 的一半转染到 $40\ \mu\text{l}$ 大肠杆菌 DH10B(electroMAX DH10B, GIBCO-BRL)中，在 1ml SOC 培养基中温育 1 小时后，取 $50\ \mu\text{l}$ 样品用于滴度检测，并接种于 LB-氨苄青霉素($100\ \mu\text{g}/\text{ml}$)培养板上。另一方面，将其余的培养物转移至 LB-氨苄青霉素($100\ \mu\text{g}/\text{ml}$)液体培养基 500ml 中并培养。培养过夜后，利用质粒 DNA 纯化试剂盒(Plasmid-Maxi, QIAGEN)制备质粒 DNA，并将其在 -20°C 冻存。
15

利用 $3\ \mu\text{l}$ FuGENETM6(Boehringer-Mannheim)将每个库获得的质粒 DNA 各 $1\ \mu\text{g}$ 转染到 COS-7 细胞中，并如上所述进行第二次淘洗以及质粒 DNA 的回收和制备。

20

3. 所获 cDNA 克隆的核苷酸序列以及推导的氨基酸序列的分析

在第一次和第二次淘洗后，从用于滴度检测的培养板中随机收集菌落，分别于 2ml LB-氨苄青霉素($100\ \mu\text{g}/\text{ml}$)液体培养基中培养后，制备质粒 DNA。然后，通过限制性酶分析利用 SmaI 和 NotI 筛选另一个 cDNA 插入
25 片段，并从 5' 侧进行测序，对它们的核苷酸序列和推导的氨基酸序列分析的结果显示，通过利用 TMT 方法可以选择性地筛选膜结合蛋白的基因。结果示于表 1 中。

表 1

第一次淘洗

库名称(1000个克隆/库)	分析的克隆数	具有跨膜区的克隆数	详细资料(插入片段大小、氨基酸残基数、跨膜区数)
kappa-1	11	1	细胞色素氧化酶 (0.75kb, 44aa, 1TM)
kappa-6	7	3	NADH 脱氢酶 (1.7kb, 88aa, 2TM) NADH 脱氢酶 (1.7kb, 88aa, 2TM) ATP 合成酶 (0.85kb, 42aa, 1TM)

第二次淘洗

库名称(1000个克隆/库)	分析的克隆数	具有跨膜区的克隆数	详细资料(插入片段大小、氨基酸残基数、跨膜区数)
kappa-1	11	4	ATP 合成酶 (0.85kb, 42aa, 1TM) ATP 合成酶 (0.85kb, 42aa, 1TM) ATP 合成酶 (0.85kb, 42aa, 1TM) 细胞色素氧化酶 (0.75kb, 44aa, 1TM)
kappa-6	11	3	NADH 脱氢酶 (1.2kb, 81aa, 2TM) NADH 脱氢酶 (3.5kb, 58aa, 2TM) 多聚 T (0.9kb, 35aa, 1TM)

由第一次淘洗，获得了已知的膜结合蛋白，细胞色素氧化酶(1个克隆)、NADH 脱氢酶(2个克隆)以及 ATP 合成酶(1个克隆)。在另一方面，由第二次淘洗获得了已知的膜结合蛋白，细胞色素氧化酶(1个克隆)、NADH 脱氢酶(2个克隆)以及 ATP 合成酶(3个克隆)。上述所有均为定位于线粒体内膜上的膜结合蛋白。例如，已知 ATP 合成酶为一个跨膜区型蛋白，细胞色素氧化酶为两个跨膜区型蛋白，而 NADH 脱氢酶为 15 个跨膜区型蛋白。这些结果揭示 TMT 方法不仅能够分离 I 型膜结合型蛋白，而且可以分离具有多个跨膜区的蛋白。

在第二次淘洗中获得的多聚 T 序列的克隆可能是由于包括多聚 A 的 cDNA 以相反的方向插入。多聚 T 被翻译成疏水性苯丙氨酸串联排列的氨基酸序列，而且由于其极其富于疏水性而被认为已经分离。

此外，在所收集克隆中含有的膜结合蛋白的百分率在第二次淘洗中比在第一次淘洗中高。这表明通过反复淘洗选择性地富集了膜结合蛋白。

因此，在一个真正的 cDNA 文库筛选系统中，TMT 方法被表明是用于选择性地克隆 I 型膜结合蛋白和包括多个跨膜区的膜结合蛋白的一种有效方法。

工业实用性

由于结构问题，人们认为抗体分子在分泌过程中累积于细胞膜上时以及由于融合蛋白的非天然结构和由于氨基酸的组成而凝集时不容易发挥抗原结合活性。因此，如在本发明中，利用由识别一种抗原的抗体制备的淘洗板可以选择性地筛选在细胞表面功能性地表达抗体融合蛋白的细胞。即，本发明提供了一种通过有效去除在细胞表面具有低或无抗原结合活性的融合蛋白的细胞而非常具有选择性地克隆编码细胞膜结合蛋白的基因的方法。

序列表

<110> CHUGAI SEIYAKU KABUSHIKI KAISHA

<120> 基因克隆的新方法

<130> C1-001DP1PCT

<150> JP 1998-138652

<151> 1998-05-20

<150> JP 1998-279876

<151> 1998-10-01

<160> 39

<210> 1

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> 对人工序列的描述：肽连接序列

<400> 1

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15

<210> 2

<211> 1035

<212> DNA

<213> 人

<220>

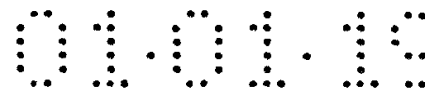
<221> CDS

<222> (1)..(1035)

<400> 2

atg ctg gcc gtc ggc tgc gcg ctg ctg gct gcc ctg ctg gcc gcg ccg 48
Met Leu Ala Val Gly Cys Ala Leu Leu Ala Ala Leu Leu Ala Ala Pro
1 5 10 15

gga gcg gcg ctg gcc cca agg cgc tgc cct gcg cag gag gtg gca aga 96
Gly Ala Ala Leu Ala Pro Arg Arg Cys Pro Ala Gln Glu Val Ala Arg
20 25 30



ggc gtg ctg acc agt ctg cca gga gac agc gtg act ctg acc tgc ccg Gly Val Leu Thr Ser Leu Pro Gly Asp Ser Val Thr Leu Thr Cys Pro 35 40 45	144
ggg gta gag ccg gaa gac aat gcc act gtt cac tgg gtg ctc agg aag Gly Val Glu Pro Glu Asp Asn Ala Thr Val His Trp Val Leu Arg Lys 50 55 60	192
ccg gct gca ggc tcc cac ccc agc aga tgg gct ggc atg gga agg agg Pro Ala Ala Gly Ser His Pro Ser Arg Trp Ala Gly Met Gly Arg Arg 65 70 75 80	240
ctg ctg ctg agg tcg gtg cag ctc cac gac tct gga aac tat tca tgc Leu Leu Leu Arg Ser Val Gln Leu His Asp Ser Gly Asn Tyr Ser Cys 85 90 95	288
tac cgg gcc ggc cgc cca gct ggg act gtg cac ttg ctg gtg gat gtt Tyr Arg Ala Gly Arg Pro Ala Gly Thr Val His Leu Leu Val Asp Val 100 105 110	336
ccc ccc gag gag ccc cag ctc tcc tgc ttc cgg aag agc ccc ctc agc Pro Pro Glu Glu Pro Gln Leu Ser Cys Phe Arg Lys Ser Pro Leu Ser 115 120 125	384
aat gtt gtt tgt gag tgg ggt cct cgg agc acc cca tcc ctg acg aca Asn Val Val Cys Glu Trp Gly Pro Arg Ser Thr Pro Ser Leu Thr Thr 130 135 140	432
aag gct gtg ctc ttg gtg agg aag ttt cag aac agt ccg gcc gaa gac Lys Ala Val Leu Leu Val Arg Lys Phe Gln Asn Ser Pro Ala Glu Asp 145 150 155 160	480
ttc cag gag ccg tgc cag tat tcc cag gag tcc cag aag ttc tcc tgc Phe Gln Glu Pro Cys Gln Tyr Ser Gln Glu Ser Gln Lys Phe Ser Cys 165 170 175	528
cag tta gca gtc ccg gag gga gac agc tct ttc tac ata gtg tcc atg Gln Leu Ala Val Pro Glu Gly Asp Ser Ser Phe Tyr Ile Val Ser Met 180 185 190	576
tgc gtc gcc agt agt gtc ggg agc aag ttc agc aaa act caa acc ttt Cys Val Ala Ser Ser Val Gly Ser Lys Phe Ser Lys Thr Gln Thr Phe 195 200 205	624
cag ggt tgt gga atc ttg cag cct gat ccg cct gcc aac atc aca gtc Gln Gly Cys Gly Ile Leu Gln Pro Asp Pro Pro Ala Asn Ile Thr Val	672

210	215	220	
act gcc gtg gcc aga aac ccc cgc tgg ctc agt gtc acc tgg caa gac			720
Thr Ala Val Ala Arg Asn Pro Arg Trp Leu Ser Val Thr Trp Gln Asp			
225	230	235	240
ccc cac tcc tgg aac tca tct ttc tac aga cta cgg ttt gag ctc aga			768
Pro His Ser Trp Asn Ser Ser Phe Tyr Arg Leu Arg Phe Glu Leu Arg			
	245	250	255
tat cgg gct gaa cgg tca aag aca ttc aca aca tgg atg gtc aag gac			816
Tyr Arg Ala Glu Arg Ser Lys Thr Phe Thr Thr Trp Met Val Lys Asp			
	260	265	270
ctc cag cat cac tgt gtc atc cac gac gcc tgg agc ggc ctg agg cac			864
Leu Gln His His Cys Val Ile His Asp Ala Trp Ser Gly Leu Arg His			
	275	280	285
gtg gtg cag ctt cgt gcc cag gag gag ttc ggg caa ggc gag tgg agc			912
Val Val Gln Leu Arg Ala Gln Glu Glu Phe Gly Gln Gly Glu Trp Ser			
	290	295	300
gag tgg agc ccg gag gcc atg ggc acg cct tgg aca gaa tcc agg agt			960
Glu Trp Ser Pro Glu Ala Met Gly Thr Pro Trp Thr Glu Ser Arg Ser			
305	310	315	320
cct cca gct gag aac gag gtg tcc acc ccc atg cag gca ctt act act			1008
Pro Pro Ala Glu Asn Glu Val Ser Thr Pro Met Gln Ala Leu Thr Thr			
	325	330	335
aat aaa gac gat gat aat att ctc ttc			1035
Asn Lys Asp Asp Asp Asn Ile Leu Phe			
	340	345	
<210> 3			
<211> 40			
<212> DNA			
<213> 人工序列			
<220>			
<223> 对人工序列的描述：“IL6R1”，一种人工合成的引物序列			
<400> 3			
ttcgaattcc caccatgctg gccgtcggct gccgcgtgct			40

<210> 4
 <211> 36
 <212> DNA
 <213> 人工序列

<220>

<223> 对人工序列的描述：“IL6R2”一种人工合成的引物序列

<400> 4
 ttcgaattcg aagagaatat tatcatcgtc tttatt 36

<210> 5
 <211> 768
 <212> DNA
 <213> 人工序列

<220>

<221> CDS

<222> (1)..(768)

<220>

<223> 对人工序列的描述：一种特别设计的单链Fv基因序列

<400> 5
 cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga cct agc cag 48
 Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro Ser Gln
 1 5 10 15
 acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att acc agc gat 96
 Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile Thr Ser Asp
 20 25 30
 cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt ctt gag tgg 144
 His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu Trp
 35 40 45
 att gga tac att agt tat agt gga atc aca acc tat aat cca tct ctc 192
 Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn Pro Ser Leu
 50 55 60
 aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac cag ttc agc 240
 Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn Gln Phe Ser
 65 70 75 80
 ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt tat tat tgt 288

Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys
85 90 95
gca aga tcc cta gct cgg act acg gct atg gac tac tgg ggt caa ggc 336
Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp Gly Gln Gly
100 105 110
agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt ggt ggt ggt 384
Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
115 120 125
tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc cca agc agc 432
Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser
130 135 140
ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt aga gcc agc 480
Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser
145 150 155 160
cag gac atc agc agt tac ctg aat tgg tac cag cag aag cca gga aag 528
Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys
165 170 175
gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac tct ggt gtg 576
Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val
180 185 190
cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc acc ttc acc 624
Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr
195 200 205
atc agc agc ctc cag cca gag gac atc gct acc tac tac tgc caa cag 672
Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln
210 215 220
ggt aac acg ctt cca tac acg ttc ggc caa ggg acc aag gtg gaa atc 720
Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
225 230 235 240
aaa cga act gtg gct gca cca tct gtc ttc atc ttc ccg cca tct gat 768
Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
245 250 255

<210> 6
<211> 32
<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“TMT1”一种人工合成的引物序列

<400> 6

ggtgtcgact cccaggtcca actgcaggag ag

32

<210> 7

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“LINK1”一种人工合成的引物序列

<400> 7

ctcgtcacag tctcctcagg tgggtgtgt tc

32

<210> 8

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“LINK3”一种人工合成的引物序列

<400> 8

gacatccaga tgaccagag cccaagcagc ctgagcgc

38

<210> 9

<211> 63

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“SCP-C”一种人工合成的引物序列

<400> 9

gctgaattct tattatttat cgtcatcgtc ttgtagtca agcttatcag atggcgggaa

60

gat

63

<210> 10
 <211> 9
 <212> PRT

<400> 10
 Met Asp Tyr Lys Asp Asp Asp Asp Lys
 1 5

<210> 11
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 对人工序列的描述: "LINK2", 一种人工合成的引物序列

<400> 11
 aaccaccacc acctgaggag actgtgacga ggct 34

<210> 12
 <211> 35
 <212> DNA
 <213> 人工序列

<220>
 <223> 对人工序列的描述: "LINK4", 一种人工合成的引物序列

<400> 12
 aggctgcttg ggctctgggt catctggatg tccga 35

<210> 13
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 对人工序列的描述: "TMT2", 一种人工合成的引物序列

<400> 13
 atccggggcc gcttattatt tatcgtcatc gtcttt 36

<210> 14

<211> 19
<212> PRT

<400> 14
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
1 5 10 15

Val Asp Ser

<210> 15
<211> 106
<212> DNA
<213> 人工序列

<220>
<223> 对人工序列的描述: "S-FLAG1", 一种人工合成的寡核苷酸序列

<400> 15
aattccacc atgggatgga gctgtatcat cctcttcttg gtagcaacag ctacaggtg 60
cgactccgac tacaagagc atgacgataa aggtaccgcg gccgcg 106

<210> 16
<211> 106
<212> DNA
<213> 人工序列

<220>
<223> 对人工序列的描述: "S-FLAG2", 一种人工合成的寡核苷酸序列

<400> 16
gatccgggc cgggtacct ttatcgtcat cgtcttttga gtcggagtcg acacctgtag 60
ctgttgctac caagaagagg atgatacagc tccatcccat ggtggg 106

<210> 17
<211> 2995
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (29)..(2839)

<400> 17
 gaattccgga catctagagg cagcgaactt gttccgatt catgctttat catttcttaa 60

tttcgtatgt tgggaacatc cctgcaag atg tca gca cca agg att tgg cta 112
 Met Ser Ala Pro Arg Ile Trp Leu
 1 5

gcg caa gct ttg ctt ttt ttc ctc acc act gaa tet ata ggt caa ctt 160
 Ala Gln Ala Leu Leu Phe Phe Leu Thr Thr Glu Ser Ile Gly Gln Leu
 10 15 20

ttg gaa ccg tgt ggt tac atc tac cct gaa ttt cca gtt gtc cag cgc 208
 Leu Glu Pro Cys Gly Tyr Ile Tyr Pro Glu Phe Pro Val Val Gln Arg
 25 30 35 40

ggc tcg aac ttc act gcc att tgt gtg ctg aag gag gcg tgt ctg cag 256
 Gly Ser Asn Phe Thr Ala Ile Cys Val Leu Lys Glu Ala Cys Leu Gln
 45 50 55

cat tac tac gtg aat gcc agc tac atc gtg tgg aag acc aac cat gct 304
 His Tyr Tyr Val Asn Ala Ser Tyr Ile Val Trp Lys Thr Asn His Ala
 60 65 70

gct gtt ccc agg gag cag gtc act gtc atc aac aga acc acg tcc agt 352
 Ala Val Pro Arg Glu Gln Val Thr Val Ile Asn Arg Thr Thr Ser Ser
 75 80 85

gtc acg ttc aca gac gtg gtc ctc ccg agc gtg cag ctc acc tgc aac 400
 Val Thr Phe Thr Asp Val Val Leu Pro Ser Val Gln Leu Thr Cys Asn
 90 95 100

atc ctg tcc ttt ggg cag atc gag cag aat gtg tat gga gtc acc atg 448
 Ile Leu Ser Phe Gly Gln Ile Glu Gln Asn Val Tyr Gly Val Thr Met
 105 110 115 120

ctt tca ggc ttt cct cca gat aaa cct aca aat ttg act tgc att gtg 496
 Leu Ser Gly Phe Pro Pro Asp Lys Pro Thr Asn Leu Thr Cys Ile Val
 125 130 135

aat gag ggg aag aat atg ctg tgc cag tgg gac ccc gga agg gag act 544
 Asn Glu Gly Lys Asn Met Leu Cys Gln Trp Asp Pro Gly Arg Glu Thr
 140 145 150

tac ctt gaa aca aac tac act ttg aaa tca gag tgg gca aca gag aag 592
 Tyr Leu Glu Thr Asn Tyr Thr Leu Lys Ser Glu Trp Ala Thr Glu Lys

155	160	165	
ttt cct gat tgc cag tca aag cat ggc act tca tgt atg gtc agc tac			640
Phe Pro Asp Cys Gln Ser Lys His Gly Thr Ser Cys Met Val Ser Tyr			
170	175	180	
atg ccc acc tat tat gtc aac att gaa gtc tgg gtg gaa gca gag aat			688
Met Pro Thr Tyr Tyr Val Asn Ile Glu Val Trp Val Glu Ala Glu Asn			
185	190	195	200
gcc ctt ggg aag gtc tcc tca gag tct atc aat ttt gac ccc gtg gat			736
Ala Leu Gly Lys Val Ser Ser Glu Ser Ile Asn Phe Asp Pro Val Asp			
205	210	215	
aaa gtg aaa ccc acc cca cca tat aat tta tca gtg acc aac tca gaa			784
Lys Val Lys Pro Thr Pro Pro Tyr Asn Leu Ser Val Thr Asn Ser Glu			
220	225	230	
gaa tta tcc agt ata tta aag cta tca tgg gtc agt tca ggg ctg ggc			832
Glu Leu Ser Ser Ile Leu Lys Leu Ser Trp Val Ser Ser Gly Leu Gly			
235	240	245	
ggt ctt tta gat cta aag tct gac atc caa tat agg acc aaa gat gcc			880
Gly Leu Leu Asp Leu Lys Ser Asp Ile Gln Tyr Arg Thr Lys Asp Ala			
250	255	260	
tca act tgg atc cag gtc cct ctt gaa gat aca atg tct cct cga act			928
Ser Thr Trp Ile Gln Val Pro Leu Glu Asp Thr Met Ser Pro Arg Thr			
265	270	275	280
tcc ttc act gtg cag gac ctc aag cct ttt aca gaa tat gtg ttt agg			976
Ser Phe Thr Val Gln Asp Leu Lys Pro Phe Thr Glu Tyr Val Phe Arg			
285	290	295	
atc cgg tcc att aag gac agt ggg aag ggc tac tgg agt gac tgg agt			1024
Ile Arg Ser Ile Lys Asp Ser Gly Lys Gly Tyr Trp Ser Asp Trp Ser			
300	305	310	
gag gag gct agt ggg acc aca tac gaa gac aga cca tcc aga cca cca			1072
Glu Glu Ala Ser Gly Thr Thr Tyr Glu Asp Arg Pro Ser Arg Pro Pro			
315	320	325	
agt ttc tgg tat aag aca aat cca tcc cat ggg cag gaa tat aga tct			1120
Ser Phe Trp Tyr Lys Thr Asn Pro Ser His Gly Gln Glu Tyr Arg Ser			
330	335	340	

gta cgg ctc ata tgg aag gca ctg cct ctt tct gaa gcc aat ggg aaa 1168
 Val Arg Leu Ile Trp Lys Ala Leu Pro Leu Ser Glu Ala Asn Gly Lys
 345 350 355 360

atc ttg gat tat gaa gtg att ctt acg cag tca aag tcc gtc tca caa 1216
 Ile Leu Asp Tyr Glu Val Ile Leu Thr Gln Ser Lys Ser Val Ser Gln
 365 370 375

acg tac aca gtc act ggc aca gag ctg acc gtg aat ctc acc aat gac 1264
 Thr Tyr Thr Val Thr Gly Thr Glu Leu Thr Val Asn Leu Thr Asn Asp
 380 385 390

cgc tat gtc gcg tct cta gca gca aga aac aag gtg ggc aaa tca gct 1312
 Arg Tyr Val Ala Ser Leu Ala Ala Arg Asn Lys Val Gly Lys Ser Ala
 395 400 405

gca gct gtc ctc acc atc ccc agc ccc cac gtc aca gct gct tat tct 1360
 Ala Ala Val Leu Thr Ile Pro Ser Pro His Val Thr Ala Ala Tyr Ser
 410 415 420

gta gtg aat ctt aaa gca ttt cca aaa gat aac ctg ctc tgg gtg gaa 1408
 Val Val Asn Leu Lys Ala Phe Pro Lys Asp Asn Leu Leu Trp Val Glu
 425 430 435 440

tgg aca cct cca cct aaa ccc gtg agc aag tac atc tta gag tgg tgt 1456
 Trp Thr Pro Pro Pro Lys Pro Val Ser Lys Tyr Ile Leu Glu Trp Cys
 445 450 455

gtg ttg tca gag aac gca ccc tgt gtt gaa gac tgg cag cag gaa gac 1504
 Val Leu Ser Glu Asn Ala Pro Cys Val Glu Asp Trp Gln Gln Glu Asp
 460 465 470

gct acc gtg aat cgg acc cac ttg aga gga cgc ctc ctg gag agc aag 1552
 Ala Thr Val Asn Arg Thr His Leu Arg Gly Arg Leu Leu Glu Ser Lys
 475 480 485

tgc tat caa atc aca gta act ccc gta ttc gcc acg ggg ccc gga ggc 1600
 Cys Tyr Gln Ile Thr Val Thr Pro Val Phe Ala Thr Gly Pro Gly Gly
 490 495 500

tct gag tcc ttg aag gcg tac ctc aaa caa gcc gct cct gcc aga gga 1648
 Ser Glu Ser Leu Lys Ala Tyr Leu Lys Gln Ala Ala Pro Ala Arg Gly
 505 510 515 520

ccg act gtt cgg aca aag aaa gtg ggg aaa aat gaa gct gtc tta gcg 1696
 Pro Thr Val Arg Thr Lys Lys Val Gly Lys Asn Glu Ala Val Leu Ala

525	530	535	
tgg gac cag att cct gtg gac gac cag aat ggc ttc att aga aac tac			1744
Trp Asp Gln Ile Pro Val Asp Asp Gln Asn Gly Phe Ile Arg Asn Tyr			
540	545	550	
tcc ata tct tac aga acc agc gtg gga aag gag atg gtt gtg cat gtg			1792
Ser Ile Ser Tyr Arg Thr Ser Val Gly Lys Glu Met Val Val His Val			
555	560	565	
gat tct tct cac acg gag tac acg ctg tcc tct ctg agt agt gat acg			1840
Asp Ser Ser His Thr Glu Tyr Thr Leu Ser Ser Leu Ser Ser Asp Thr			
570	575	580	
ttg tac atg gtc cga atg gcc gcg tac aca gat gaa ggt ggg aaa gat			1888
Leu Tyr Met Val Arg Met Ala Ala Tyr Thr Asp Glu Gly Gly Lys Asp			
585	590	595	600
ggg ccg gaa ttc act ttt aca aca cca aag ttc gct caa gga gaa ata			1936
Gly Pro Glu Phe Thr Phe Thr Thr Pro Lys Phe Ala Gln Gly Glu Ile			
605	610	615	
gaa gcc ata gtc gtg cct gtg tgc tta gcc ttc ctc ctg aca acc ctg			1984
Glu Ala Ile Val Val Pro Val Cys Leu Ala Phe Leu Leu Thr Thr Leu			
620	625	630	
ctg ggc gtc ttg ttc tgc ttt aac aaa cga gac cta att aaa aaa cac			2032
Leu Gly Val Leu Phe Cys Phe Asn Lys Arg Asp Leu Ile Lys Lys His			
635	640	645	
atc tgg cct aat gtt cct gat cct tcc aag agt cat att gcc cag tgg			2080
Ile Trp Pro Asn Val Pro Asp Pro Ser Lys Ser His Ile Ala Gln Trp			
650	655	660	
tca cct cac acc ccc cca agg cac aat ttt aac tcc aaa gat caa atg			2128
Ser Pro His Thr Pro Pro Arg His Asn Phe Asn Ser Lys Asp Gln Met			
665	670	675	680
tac tcg gac ggc aat ttc act gat gta agc gtt gtg gaa ata gaa gca			2176
Tyr Ser Asp Gly Asn Phe Thr Asp Val Ser Val Val Glu Ile Glu Ala			
685	690	695	
aac aac aag aag cct tgt cca gat gac ctg aag tcc gtg gac ctg ttc			2224
Asn Asn Lys Lys Pro Cys Pro Asp Asp Leu Lys Ser Val Asp Leu Phe			
700	705	710	

aag aag gag aaa gtg agt aca gaa ggg cac agc agt ggc atc ggg ggc Lys Lys Glu Lys Val Ser Thr Glu Gly His Ser Ser Gly Ile Gly Gly 715 720 725	2272
tct tca tgc atg tcc tcc tcc agg ccc agc atc tcc agc aac gag gag Ser Ser Cys Met Ser Ser Ser Arg Pro Ser Ile Ser Ser Asn Glu Glu 730 735 740	2320
aat gag tct gct cag agc acc gcc agc acg gtc gag tac tcc act gtg Asn Glu Ser Ala Gln Ser Thr Ala Ser Thr Val Glu Tyr Ser Thr Val 745 750 755 760	2368
gtg cac agc ggc tac agg cac cag gtc ccg tcc gtg caa gtg ttc tca Val His Ser Gly Tyr Arg His Gln Val Pro Ser Val Gln Val Phe Ser 765 770 775	2416
agg tcc gag tcc acc cag ccc ctg cta gac tcg gag gag cgg cca gaa Arg Ser Glu Ser Thr Gln Pro Leu Leu Asp Ser Glu Glu Arg Pro Glu 780 785 790	2464
gac ctg cag ctg gtg gat agt gta gac ggt ggg gat gag atc ttg ccc Asp Leu Gln Leu Val Asp Ser Val Asp Gly Gly Asp Glu Ile Leu Pro 795 800 805	2512
agg caa ccg tat ttc aag cag aac tgc agt cag cct gaa gcc tgt cca Arg Gln Pro Tyr Phe Lys Gln Asn Cys Ser Gln Pro Glu Ala Cys Pro 810 815 820	2560
gag att tca cat ttt gaa agg tca aac cag gtt ttg tcc ggc aat gag Glu Ile Ser His Phe Glu Arg Ser Asn Gln Val Leu Ser Gly Asn Glu 825 830 835 840	2608
gag gat ttt gtc aga ctg aag cag cag cag gtt tca gat cac att tct Glu Asp Phe Val Arg Leu Lys Gln Gln Gln Val Ser Asp His Ile Ser 845 850 855	2656
cag ccc tat gga tcc gag caa cgg agg ctg ttt cag gaa ggc tct aca Gln Pro Tyr Gly Ser Glu Gln Arg Arg Leu Phe Gln Glu Gly Ser Thr 860 865 870	2704
gcg gat gct ctt ggc acg ggg gct gat gga cag atg gag aga ttt gaa Ala Asp Ala Leu Gly Thr Gly Ala Asp Gly Gln Met Glu Arg Phe Glu 875 880 885	2752
tct gtt gga atg gag acc aca att gat gaa gaa att ccc aaa agt tac Ser Val Gly Met Glu Thr Thr Ile Asp Glu Glu Ile Pro Lys Ser Tyr	2800

890	895	900	
ttg cca cag act gta aga caa ggt ggc tac atg ccg cag tgaaggactg			2849
Leu Pro Gln Thr Val Arg Gln Gly Gly Tyr Met Pro Gln			
905	910	915	

gctcctgaac ttcagcagga actgcaaat aaagctaaag acgagtggct tcagatgaga	2909
--	------

aacagtcctc actcctgaa gataggcatt gcctctaagg acaaagtcac acctgggccg	2969
--	------

tctccattcc agagtagctg gaattc	2995
------------------------------	------

<210> 18

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: "mgp1", 一种人工合成的引物序列

<400> 18

cccaagcttg aattcacttt tacaaca	27
-------------------------------	----

<210> 19

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: "mgp3", 一种人工合成的引物序列

<400> 19

tttgcggccg cgaattccag ctactctgg	29
---------------------------------	----

<210> 20

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: "mgp2", 一种人工合成的引物序列

<400> 20

cccaagcttg aattcaaaaa acacatctgg ctt

33

<210> 21
 <211> 1662
 <212> DNA
 <213> 人工序列

<220>
 <221> CDS
 <222> (11)..(1648)

<220>
 <223> 对人工序列的描述: "gpM1-BvGS3", 一种特别设计的单链Fv基因序列

<400> 21

gaattccacc atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca 49
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr
 1 5 10

gct aca ggt gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt 97
 Ala Thr Gly Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly
 15 20 25

ctt gtg aga cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc 145
 Leu Val Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly
 30 35 40 45

tac tca att acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct 193
 Tyr Ser Ile Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro
 50 55 60

gga cga ggt ctt gag tgg att gga tac att agt tat agt gga atc aca 241
 Gly Arg Gly Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr
 65 70 75

acc tat aat cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc 289
 Thr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr
 80 85 90

agc aag aac cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac 337
 Ser Lys Asn Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp
 95 100 105

acc gcg gtt tat tat tgt gca aga tcc cta gct cgg act acg gct atg 385
 Thr Ala Val Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met

110	115	120	125	
gac tac tgg ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt				433
Asp Tyr Trp Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly				
	130	135	140	
ggt tcg ggt ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg				481
Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met				
	145	150	155	
acc cag agc cca agc agc ctg agc gcc agc gtc ggt gac aga gtc acc				529
Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr				
	160	165	170	
atc acc tgt aga gcc agc cag gac atc agc agt tac ctg aat tgg tac				577
Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr				
	175	180	185	
cag cag aag cca gga aag gct cca aag ctg ctg atc tac tac acc tcc				625
Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser				
	190	195	200	205
aga ctg cac tct ggt gtc cca agc aga ttc agc ggt agc ggt agc ggt				673
Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly				
	210	215	220	
acc gac ttc acc ttc acc atc agc agc ctc cag cca gag gac atc gct				721
Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala				
	225	230	235	
acc tac tac tgc caa cag ggt aac acg ctt cca tac acg ttc ggc caa				769
Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln				
	240	245	250	
ggg acc aag gtc gaa atc aaa tct aga ggt ggt ggt ggt tcg ggt ggt				817
Gly Thr Lys Val Glu Ile Lys Ser Arg Gly Gly Gly Gly Ser Gly Gly				
	255	260	265	
ggt ggt tcg ggt ggt ggc gga tcg gtc gac tcc cag gtc caa ctg cag				865
Gly Gly Ser Gly Gly Gly Gly Ser Val Asp Ser Gln Val Gln Leu Gln				
	270	275	280	285
gag agc ggt cca ggt ctt gtc aga cct agc cag acc ctg agc ctg acc				913
Glu Ser Gly Pro Gly Leu Val Arg Pro Ser Gln Thr Leu Ser Leu Thr				
	290	295	300	

tgc acc gtg tct ggc tac tca att acc agc gat cat gcc tgg agc tgg 961
 Cys Thr Val Ser Gly Tyr Ser Ile Thr Ser Asp His Ala Trp Ser Trp
 305 310 315

gtt cgc cag cca cct gga cga ggt ctt gag tgg att gga tac att agt 1009
 Val Arg Gln Pro Pro Gly Arg Gly Leu Glu Trp Ile Gly Tyr Ile Ser
 320 325 330

tat agt gga atc aca acc tat aat cca tct ctc aaa tcc aga gtg aca 1057
 Tyr Ser Gly Ile Thr Thr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr
 335 340 345

atg ctg aga gac acc agc aag aac cag ttc agc ctg aga ctc agc agc 1105
 Met Leu Arg Asp Thr Ser Lys Asn Gln Phe Ser Leu Arg Leu Ser Ser
 350 355 360 365

gtg aca gcc gcc gac acc gcg gtt tat tat tgt gca aga tcc cta gct 1153
 Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ser Leu Ala
 370 375 380

cgg act acg get atg gac tac tgg ggt caa gcc agc ctc gtc aca gtc 1201
 Arg Thr Thr Ala Met Asp Tyr Trp Gly Gln Gly Ser Leu Val Thr Val
 385 390 395

tcc tca ggt ggt ggt ggt tcc ggt ggt ggt ggt tcc ggt ggt gcc gga 1249
 Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
 400 405 410

tcc gac atc cag atg acc cag agc cca agc agc ctg agc gcc agc gtg 1297
 Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
 415 420 425

ggt gac aga gtg acc atc acc tgt aga gcc agc cag gac atc agc agt 1345
 Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Ser
 430 435 440 445

tac ctg aat tgg tac cag cag aag cca gga aag gct cca aag ctg ctg 1393
 Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu
 450 455 460

atc tac tac acc tcc aga ctg cac tct ggt gtg cca agc aga ttc agc 1441
 Ile Tyr Tyr Thr Ser Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser
 465 470 475

ggt agc ggt agc ggt acc gac ttc acc ttc acc atc agc agc ctc cag 1489
 Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln

480	485	490	
cca gag gac atc gct acc tac tac tgc caa cag ggt aac acg ctt cca			1537
Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro			
495	500	505	
tac acg ttc ggc caa ggg acc aag gtg gaa atc aaa cga act gtg gct			1585
Tyr Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala			
510	515	520	525
gca cca tct gtc ttc atc ttc ccg cca tct gat aag ctt gac tac aaa			1633
Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Lys Leu Asp Tyr Lys			
530	535	540	
gac gat gac gat aaa taataagcgg ccgc			1662
Asp Asp Asp Asp Lys			
545			

- <210> 22
- <211> 72
- <212> DNA
- <213> 人工序列

<220>

<223> 对人工序列的描述: "BvGS3", 一种人工合成的引物序列

<400> 22

ggagtcgacc gatccgccac caccggaacc accaccaccc gaaccaccac cacctttgat 60

ttccaccttg gt 72

- <210> 23
- <211> 780
- <212> DNA
- <213> 人工序列

<220>

<221> CDS

<222> (1)..(780)

<220>

<223> 对人工序列的描述: "shPM1(Δ EL)", 一种特别设计的单链Fv基因序列

<400> 23

atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt 48
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1 5 10 15

gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga 96
 Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg
 20 25 30

cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att 144
 Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile
 35 40 45

acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt 192
 Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly
 50 55 60

ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat 240
 Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn
 65 70 75 80

cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac 288
 Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn
 85 90 95

cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt 336
 Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val
 100 105 110

tat tat tgt gca aga tcc cta gct cgg act acg gct atg gac tac tgg 384
 Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp
 115 120 125

ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt 432
 Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
 130 135 140

ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc 480
 Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser
 145 150 155 160

cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt 528
 Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
 165 170 175

aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag 576
 Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys

180	185	190	
cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac			624
Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His			
195	200	205	
tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc			672
Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe			
210	215	220	
acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac			720
Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr			
225	230	235	240
tgc caa cag gga aat act tta cca tac acg ttc ggc caa ggg acc aag			768
Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys			
245	250	255	
gtg gaa atc aaa			780
Val Glu Ile Lys			
260			
<210> 24			
<211> 321			
<212> DNA			
<213> 人			
<220>			
<221> CDS			
<222> (1)..(321)			
<400> 24			
cga act gtg gct gca cca tct gtc ttc atc ttc ccg cca tct gat gag			48
Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu			
1	5	10	15
cag ttg aaa tct gga act gcc tct gtt gtg tgc ctg ctg aat aac ttc			96
Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe			
20	25	30	
tat ccc aga gag gcc aaa gta cag tgg aag gtg gat aac gcc ctc caa			144
Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln			
35	40	45	
tcg ggt aac tcc cag gag agt gtc aca gag cag gac agc aag gac agc			192

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
50 55 60

acc tac agc ctc agc agc acc ctg acg ctg agc aaa gca gac tac gag 240
Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
65 70 75 80

aaa cac aaa gtc tac gcc tgc gaa gtc acc cat cag ggc ctg agc tcg 288
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
85 90 95

ccc gtc aca aag agc ttc aac agg gga gag tct 321
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Ser
100 105

<210> 25
<211> 363
<212> DNA
<213> 人

<220>
<221> CDS
<222> (1)..(363)

<400> 25

gtg gcc ctg cac agg ccc gat gtc tac ttg ctg cca cca gcc cgg gag: 48
Val Ala Leu His Arg Pro Asp Val Tyr Leu Leu Pro Pro Ala Arg Glu
1 5 10 15

cag ctg aac ctg cgg gag tgc gcc acc atc acg tgc ctg gtg acg ggc 96
Gln Leu Asn Leu Arg Glu Ser Ala Thr Ile Thr Cys Leu Val Thr Gly
20 25 30

ttc tct ccc gcg gac gtc ttc gtg cag tgg atg cag agg ggg cag ccc 144
Phe Ser Pro Ala Asp Val Phe Val Gln Trp Met Gln Arg Gly Gln Pro
35 40 45

ttg tcc ccg gag aag tat gtg acc agc gcc cca atg cct gag ccc cag 192
Leu Ser Pro Glu Lys Tyr Val Thr Ser Ala Pro Met Pro Glu Pro Gln
50 55 60

gcc cca ggc cgg tac ttc gcc cac agc atc ctg acc gtg tcc gaa gag 240
Ala Pro Gly Arg Tyr Phe Ala His Ser Ile Leu Thr Val Ser Glu Glu
65 70 75 80

gaa tgg aac acg ggg gag acc tac acc tgc gtg gcc cat gag gcc ctg 288
 Glu Trp Asn Thr Gly Glu Thr Tyr Thr Cys Val Ala His Glu Ala Leu
 85 90 95

ccc aac agg gtc acc gag agg acc gtg gac aag tcc acc gag ggg gag 336
 Pro Asn Arg Val Thr Glu Arg Thr Val Asp Lys Ser Thr Glu Gly Glu
 100 105 110

gtg agc gcc gac gag gag ggc ttt gag 363
 Val Ser Ala Asp Glu Glu Gly Phe Glu
 115 120

<210> 26

<211> 1101

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)..(1101)

<220>

<223> 对人工序列的描述：“shPM1-Kappa”，一种特别设计的单链Fv基因序列

<400> 26

atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt 48
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1 5 10 15

gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga 96
 Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg
 20 25 30

cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att 144
 Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile
 35 40 45

acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt 192
 Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly
 50 55 60

ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat 240
 Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn

65	70	75	80	
cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac				288
Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn				
	85	90	95	
cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt				336
Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val				
	100	105	110	
tat tat tgt gca aga tcc cta gct cgg act acg gct atg gac tac tgg				384
Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp				
	115	120	125	
ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcc ggt				432
Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Ser Gly				
	130	135	140	
ggt ggt ggt tcc ggt ggt ggc gga tcc gac atc cag atg acc cag agc				480
Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser				
	145	150	155	160
cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt				528
Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys				
	165	170	175	
aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag				576
Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys				
	180	185	190	
cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac				624
Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His				
	195	200	205	
tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc				672
Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe				
	210	215	220	
acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac				720
Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr				
	225	230	235	240
tgc caa cag gga aat act tta cca tac acg ttc ggc caa ggc acc aag				768
Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys				
	245	250	255	

gtg gaa atc aaa cga act gtg gct gca cca tct gtc ttc atc ttc ccg 816
 Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro
 260 265 270

cca tct gat gag cag ttg aaa tct gga act gcc tct gtt gtg tgc ctg 864
 Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
 275 280 285

ctg aat aac ttc tat ccc aga gag gcc aaa gta cag tgg aag gtg gat 912
 Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp
 290 295 300

aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc aca gag cag gac 960
 Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp
 305 310 315 320

agc aag gac agc acc tac agc ctc agc agc acc ctg acg ctg agc aaa 1008
 Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys
 325 330 335

gca gac tac gag aaa cac aaa gtc tac gcc tgc gaa gtc acc cat cag 1056
 Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln
 340 345 350

ggc ctg agc tcg ccc gtc aca aag agc ttc aac agg gga gag tct 1101
 Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Ser
 355 360 365

<210> 27

<211> 1143

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)..(1143)

<220>

<223> 对人工序列的描述：“shPM1-MCH4”，一种特别设计的单链Fv基因序列

<400> 27

atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt 48
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly

1	5	10	15	
gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga				96
Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg				
	20	25	30	
cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att				144
Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile				
	35	40	45	
acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt				192
Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly				
	50	55	60	
ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat				240
Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn				
	65	70	75	80
cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac				288
Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn				
	85	90	95	
cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt				336
Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val				
	100	105	110	
tat tat tgt gca aga tcc cta gct cgg act acg gct atg gac tac tgg				384
Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp				
	115	120	125	
ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt				432
Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly				
	130	135	140	
ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc				480
Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser				
	145	150	155	160
cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt				528
Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys				
	165	170	175	
aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag				576
Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys				
	180	185	190	

cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac 624
 Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His
 195 200 205

tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc 672
 Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
 210 215 220

acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac 720
 Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr
 225 230 235 240

tgc caa cag gga aat act tta cca tac acg ttc ggc caa ggg acc aag 768
 Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys
 245 250 255

gtg gaa atc aaa gtg gcc ctg cac agg ccc gat gtc tac ttg ctg cca 816
 Val Glu Ile Lys Val Ala Leu His Arg Pro Asp Val Tyr Leu Leu Pro
 260 265 270

cca gcc cgg gag cag ctg aac ctg cgg gag tcg gcc acc atc acg tgc 864
 Pro Ala Arg Glu Gln Leu Asn Leu Arg Glu Ser Ala Thr Ile Thr Cys
 275 280 285

ctg gtg acg ggc ttc tct ccc gcg gac gtc ttc gtg cag tgg atg cag 912
 Leu Val Thr Gly Phe Ser Pro Ala Asp Val Phe Val Gln Trp Met Gln
 290 295 300

agg ggg cag ccc ttg tcc ccg gag aag tat gtg acc agc gcc cca atg 960
 Arg Gly Gln Pro Leu Ser Pro Glu Lys Tyr Val Thr Ser Ala Pro Met
 305 310 315 320

cct gag ccc cag gcc cca ggc cgg tac ttc gcc cac agc atc ctg acc 1008
 Pro Glu Pro Gln Ala Pro Gly Arg Tyr Phe Ala His Ser Ile Leu Thr
 325 330 335

gtg tcc gaa gag gaa tgg aac acg ggg gag acc tac acc tgc gtg gcc 1056
 Val Ser Glu Glu Glu Trp Asn Thr Gly Glu Thr Tyr Thr Cys Val Ala
 340 345 350

cat gag gcc ctg ccc aac agg gtc acc gag agg acc gtg gac aag tcc 1104
 His Glu Ala Leu Pro Asn Arg Val Thr Glu Arg Thr Val Asp Lys Ser
 355 360 365

acc gag ggg gag gtg agc gcc gac gag gag ggc ttt gag 1143
 Thr Glu Gly Glu Val Ser Ala Asp Glu Glu Gly Phe Glu

370

375

380

<210> 28

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: "EF-1", 一种人工合成的引物序列

<400> 28

cagacagtgg ttcaaagt

18

<210> 29

<211> 107

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述: "SCP-C2", 一种人工合成的引物序列

<400> 29

aaagcggccg cttattattt atcgatcgcg tctttgtagt ctgaagcttt gatttccacc 60

ttggtccctt gcccgaacgt gtaggtaaa gtatttcctt gttggca

107

<210> 30

<211> 1557

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)..(1557)

<220>

<223> 对人工序列的描述: "shPM1(Δ EL)-BvGS3", 一种特别设计的单链Fv基因序列

<400> 30

atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt 48
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1 5 10 15

gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga 96
 Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg
 20 25 30

cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att 144
 Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile
 35 40 45

acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt 192
 Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly
 50 55 60

ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat 240
 Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn
 65 70 75 80

cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac 288
 Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn
 85 90 95

cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt 336
 Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val
 100 105 110

tat tat tgt gca aga tcc cta gct cgg act acg gct atg gac tac tgg 384
 Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp
 115 120 125

ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt 432
 Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
 130 135 140

ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc 480
 Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser
 145 150 155 160

cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt 528
 Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
 165 170 175

aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag 576
 Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys

180	185	190	
cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac			624
Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His			
195	200	205	
tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc			672
Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe			
210	215	220	
acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac			720
Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr			
225	230	235	240
tgc caa cag ggt aac acg ctt cca tac acg ttc ggc caa ggg acc aag			768
Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys			
245	250	255	
gtg gaa atc aaa ggt ggt ggt ggt tgc ggt ggt ggt ggt tgc ggt ggt			816
Val Glu Ile Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly			
260	265	270	
ggc gga tgc gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt			864
Gly Gly Ser Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly			
275	280	285	
ctt gtg aga cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc			912
Leu Val Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly			
290	295	300	
tac tca att acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct			960
Tyr Ser Ile Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro			
305	310	315	320
gga cga ggt ctt gag tgg att gga tac att agt tat agt gga atc aca			1008
Gly Arg Gly Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr			
325	330	335	
acc tat aat cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc			1056
Thr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr			
340	345	350	
agc aag aac cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac			1104
Ser Lys Asn Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp			
355	360	365	

acc gcg gtt tat tat tgt gca aga tcc cta gct cgg act acg gct atg 1152
 Thr Ala Val Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met
 370 375 380

gac tac tgg ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt 1200
 Asp Tyr Trp Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly
 385 390 395 400

ggt tcg ggt ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg 1248
 Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met
 405 410 415

acc cag agc cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc 1296
 Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr
 420 425 430

atc acc tgt aga gcc agc cag gac atc agc agt tac ctg aat tgg tac 1344
 Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr
 435 440 445

cag cag aag cca gga aag gct cca aag ctg ctg atc tac tac acc tcc 1392
 Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser
 450 455 460

aga ctg cac tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt 1440
 Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
 465 470 475 480

acc gac ttc acc ttc acc atc agc agc ctc cag cca gag gac atc gct 1488
 Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala
 485 490 495

acc tac tac tgc caa cag gga aat act tta cca tac acg ttc ggc caa 1536
 Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln
 500 505 510

ggg acc aag gtg gaa atc aaa 1557
 Gly Thr Lys Val Glu Ile Lys
 515

<210> 31
 <211> 29
 <212> DNA
 <213> 人工序列

<220>

<223> 对人工序列的描述：“Kappa1”，一种人工合成的引物序列

<400> 31

ccgccatctg atgagcagtt gaaatctgg

29

<210> 32

<211> 54

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“Kappa2”，一种人工合成的引物序列

<400> 32

ttatttatcg tcatcgtctt tgtagtcaag cttagactct cccctgttga agct

54

<210> 33

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“SCP-K”，一种人工合成的引物序列

<400> 33

ttcaactgct catcagatgg cgggaagat

29

<210> 34

<211> 1878

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)..(1878)

<220>

<223> 对人工序列的描述：“shPM1-Kappa-BvGS3”，一种特别设计的单链Fv基因序列

<400> 34
 atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt 48
 Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
 1 5 10 15

 gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga 96
 Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg
 20 25 30

 cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att 144
 Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile
 35 40 45

 acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt 192
 Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly
 50 55 60

 ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat 240
 Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn
 65 70 75 80

 cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac 288
 Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn
 85 90 95

 cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt 336
 Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val
 100 105 110

 tat tat tgt gca aga tcc cta gct cgg act acg gct atg gac tac tgg 384
 Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp
 115 120 125

 ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt 432
 Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly
 130 135 140

 ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc 480
 Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser
 145 150 155 160

 cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt 528
 Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys
 165 170 175

aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag 576
 Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys
 180 185 190

cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac 624
 Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His
 195 200 205

tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc 672
 Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
 210 215 220

acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac 720
 Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr
 225 230 235 240

tgc caa cag ggt aac acg ctt cca tac acg ttc ggc caa ggc acc aag 768
 Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys
 245 250 255

gtg gaa atc aaa ggt ggt ggt ggt tgc ggt ggt ggt ggt tgc ggt ggt 816
 Val Glu Ile Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
 260 265 270

ggc gga tgc gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt 864
 Gly Gly Ser Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly
 275 280 285

ctt gtg aga cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc 912
 Leu Val Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly
 290 295 300

tac tca att acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct 960
 Tyr Ser Ile Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro
 305 310 315 320

gga cga ggt ctt gag tgg att gga tac att agt tat agt gga atc aca 1008
 Gly Arg Gly Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr
 325 330 335

acc tat aat cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc 1056
 Thr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr
 340 345 350

agc aag aac cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac 1104

Ser Lys Asn Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp
 355 360 365
 acc gcg gtt tat tat tgt gca aga tcc cta gct cgg act acg gct atg 1152
 Thr Ala Val Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met
 370 375 380
 gac tac tgg ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt 1200
 Asp Tyr Trp Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly
 385 390 395 400
 ggt tcg ggt ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg 1248
 Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met
 405 410 415
 acc cag agc cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc 1296
 Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr
 420 425 430
 atc acc tgt aga gcc agc cag gac atc agc agt tac ctg aat tgg tac 1344
 Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr
 435 440 445
 cag cag aag cca gga aag gct cca aag ctg ctg atc tac tac acc tcc 1392
 Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser
 450 455 460
 aga ctg cac tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt 1440
 Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly
 465 470 475 480
 acc gac ttc acc ttc acc atc agc agc ctc cag cca gag gac atc gct 1488
 Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala
 485 490 495
 acc tac tac tgc caa cag gga aat act tta cca tac acg ttc ggc caa 1536
 Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln
 500 505 510
 ggg acc aag gtg gaa atc aaa cga act gtg gct gca cca tct gtc ttc 1584
 Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
 515 520 525
 atc ttc ccg cca tct gat gag cag ttg aaa tct gga act gcc tct gtt 1632
 Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val
 530 535 540

gtg tgc ctg ctg aat aac ttc tat ccc aga gag gcc aaa gta cag tgg 1680
 Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp
 545 550 555 560

aag gtg gat aac gcc ctc caa tcg ggt aac tcc cag gag agt gtc aca 1728
 Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr
 565 570 575

gag cag gac agc aag gac agc acc tac agc ctc agc agc acc ctg acg 1776
 Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr
 580 585 590

ctg agc aaa gca gac tac gag aaa cac aaa gtc tac gcc tgc gaa gtc 1824
 Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val
 595 600 605

acc cat cag ggc ctg agc tcg ccc gtc aca aag agc ttc aac agg gga 1872
 Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly
 610 615 620

gag tct 1878
 Glu Ser
 625

<210> 35

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“MCH4-1”，一种人工合成的引物序列

<400> 35

gtggaaatca aagtggcct gcacagccc

29

<210> 36

<211> 68

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“MCH4-2.1”，一种人工合成的引物序列

<400> 36

tagtcaagct tctcaaatcc ctcttcgctc gcgctaacct ctcttcggt ggacttgcc 60

acggctct

68

<210> 37

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“SCP-Mu”，一种人工合成的引物序列

<400> 37

tgcaggcca cttgatttc caccttgg

29

<210> 38

<211> 53

<212> DNA

<213> 人工序列

<220>

<223> 对人工序列的描述：“MCH4-2.2”，一种人工合成的引物序列

<400> 38

aaagcggccg cttattattt atcgatcgcg tctttgtagt caagcttctc aaa

53

<210> 39

<211> 1920

<212> DNA

<213> 人工序列

<220>

<221> CDS

<222> (1)..(1920)

<220>

<223> 对人工序列的描述：“shPM1-MCH4-BvGS3”，一种特别设计的单链Fv基因序列

<400> 39

atg gga tgg agc tgt atc atc ctc ttc ttg gta gca aca gct aca ggt	48
Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly	
1 5 10 15	
gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt ctt gtg aga	96
Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg	
20 25 30	
cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc tac tca att	144
Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Tyr Ser Ile	
35 40 45	
acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct gga cga ggt	192
Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro Gly Arg Gly	
50 55 60	
ctt gag tgg att gga tac att agt tat agt gga atc aca acc tat aat	240
Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr Thr Tyr Asn	
65 70 75 80	
cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc agc aag aac	288
Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr Ser Lys Asn	
85 90 95	
cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac acc gcg gtt	336
Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val.	
100 105 110	
tat tat tgt gca aga tcc cta get cgg act acg get atg gac tac tgg	384
Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met Asp Tyr Trp	
115 120 125	
ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt ggt tcg ggt	432
Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly	
130 135 140	
ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg acc cag agc	480
Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met Thr Gln Ser	
145 150 155 160	
cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc atc acc tgt	528
Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys	
165 170 175	
aga gcc agc cag gac atc agc agt tac ctg aat tgg tac cag cag aag	576

Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr Gln Gln Lys	
180	185
190	
cca gga aag gct cca aag ctg ctg atc tac tac acc tcc aga ctg cac	624
Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser Arg Leu His	
195	200
205	
tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt acc gac ttc	672
Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe	
210	215
220	
acc ttc acc atc agc agc ctc cag cca gag gac atc gct acc tac tac	720
Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr	
225	230
235	240
tgc caa cag ggt aac acg ctt cca tac acg ttc ggc caa ggg acc aag	768
Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln Gly Thr Lys	
245	250
255	
gtg gaa atc aaa ggt ggt ggt ggt tcg ggt ggt ggt tcg ggt ggt	816
Val Glu Ile Lys Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly	
260	265
270	
ggc gga tcg gtc gac tcc cag gtc caa ctg cag gag agc ggt cca ggt	864
Gly Gly Ser Val Asp Ser Gln Val Gln Leu Gln Glu Ser Gly Pro Gly	
275	280
285	
ctt gtg aga cct agc cag acc ctg agc ctg acc tgc acc gtg tct ggc	912
Leu Val Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly	
290	295
300	
tac tca att acc agc gat cat gcc tgg agc tgg gtt cgc cag cca cct	960
Tyr Ser Ile Thr Ser Asp His Ala Trp Ser Trp Val Arg Gln Pro Pro	
305	310
315	320
gga cga ggt ctt gag tgg att gga tac att agt tat agt gga atc aca	1008
Gly Arg Gly Leu Glu Trp Ile Gly Tyr Ile Ser Tyr Ser Gly Ile Thr	
325	330
335	
acc tat aat cca tct ctc aaa tcc aga gtg aca atg ctg aga gac acc	1056
Thr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Met Leu Arg Asp Thr	
340	345
350	
agc aag aac cag ttc agc ctg aga ctc agc agc gtg aca gcc gcc gac	1104
Ser Lys Asn Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp	
355	360
365	

acc gcg gtt tat tat tgt gca aga tcc cta gct cgg act acg gct atg	1152
Thr Ala Val Tyr Tyr Cys Ala Arg Ser Leu Ala Arg Thr Thr Ala Met	
370 375 380	
gac tac tgg ggt caa ggc agc ctc gtc aca gtc tcc tca ggt ggt ggt	1200
Asp Tyr Trp Gly Gln Gly Ser Leu Val Thr Val Ser Ser Gly Gly Gly	
385 390 395 400	
ggt tcg ggt ggt ggt ggt tcg ggt ggt ggc gga tcg gac atc cag atg	1248
Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Asp Ile Gln Met	
405 410 415	
acc cag agc cca agc agc ctg agc gcc agc gtg ggt gac aga gtg acc	1296
Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr	
420 425 430	
atc acc tgt aga gcc agc cag gac atc agc agt tac ctg aat tgg tac	1344
Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Ser Tyr Leu Asn Trp Tyr	
435 440 445	
cag cag aag cca gga aag gct cca aag ctg ctg atc tac tac acc tcc	1392
Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser	
450 455 460	
aga ctg cac tct ggt gtg cca agc aga ttc agc ggt agc ggt agc ggt	1440
Arg Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly	
465 470 475 480	
acc gac ttc acc ttc acc atc agc agc ctc cag cca gag gac atc gct	1488
Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala	
485 490 495	
acc tac tac tgc caa cag gga aat act tta cca tac acg ttc ggc caa	1536
Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro Tyr Thr Phe Gly Gln	
500 505 510	
ggg acc aag gtg gaa atc aaa gtg gcc ctg cac agg ccc gat gtc tac	1584
Gly Thr Lys Val Glu Ile Lys Val Ala Leu His Arg Pro Asp Val Tyr	
515 520 525	
ttg ctg cca cca gcc cgg gag cag ctg aac ctg cgc gag tcg gcc acc	1632
Leu Leu Pro Pro Ala Arg Glu Gln Leu Asn Leu Arg Glu Ser Ala Thr	
530 535 540	
atc acg tgc ctg gtg acg ggc ttc tct ccc gcg gac gtc ttc gtg cag	1680

Ile	Thr	Cys	Leu	Val	Thr	Gly	Phe	Ser	Pro	Ala	Asp	Val	Phe	Val	Gln	
545					550					555					560	
tgg	atg	cag	agg	ggg	cag	ccc	ttg	tcc	ccg	gag	aag	tat	gtg	acc	agc	1728
Trp	Met	Gln	Arg	Gly	Gln	Pro	Leu	Ser	Pro	Glu	Lys	Tyr	Val	Thr	Ser	
			565					570					575			
gcc	cca	atg	cct	gag	ccc	cag	gcc	cca	ggc	egg	tac	ttc	gcc	cac	agc	1776
Ala	Pro	Met	Pro	Glu	Pro	Gln	Ala	Pro	Gly	Arg	Tyr	Phe	Ala	His	Ser	
			580				585					590				
atc	ctg	acc	gtg	tcc	gaa	gag	gaa	tgg	aac	acg	ggg	gag	acc	tac	acc	1824
Ile	Leu	Thr	Val	Ser	Glu	Glu	Glu	Trp	Asn	Thr	Gly	Glu	Thr	Tyr	Thr	
		595				600					605					
tgc	gtg	gcc	cat	gag	gcc	ctg	ccc	aac	agg	gtc	acc	gag	agg	acc	gtg	1872
Cys	Val	Ala	His	Glu	Ala	Leu	Pro	Asn	Arg	Val	Thr	Glu	Arg	Thr	Val	
	610				615					620						
gac	aag	tcc	acc	gag	ggg	gag	gtg	agc	gcc	gac	gag	gag	ggc	ttt	gag	1920
Asp	Lys	Ser	Thr	Glu	Gly	Glu	Val	Ser	Ala	Asp	Glu	Glu	Gly	Phe	Glu	
625				630					635				640			

说明书附图

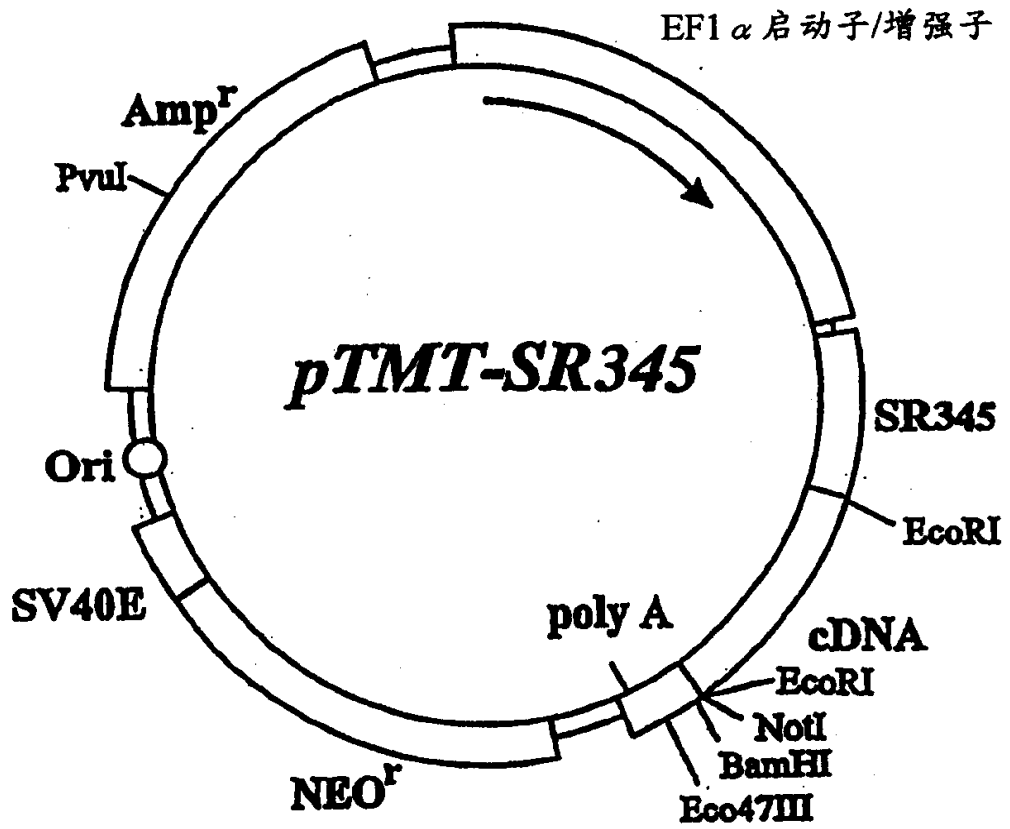


图 1

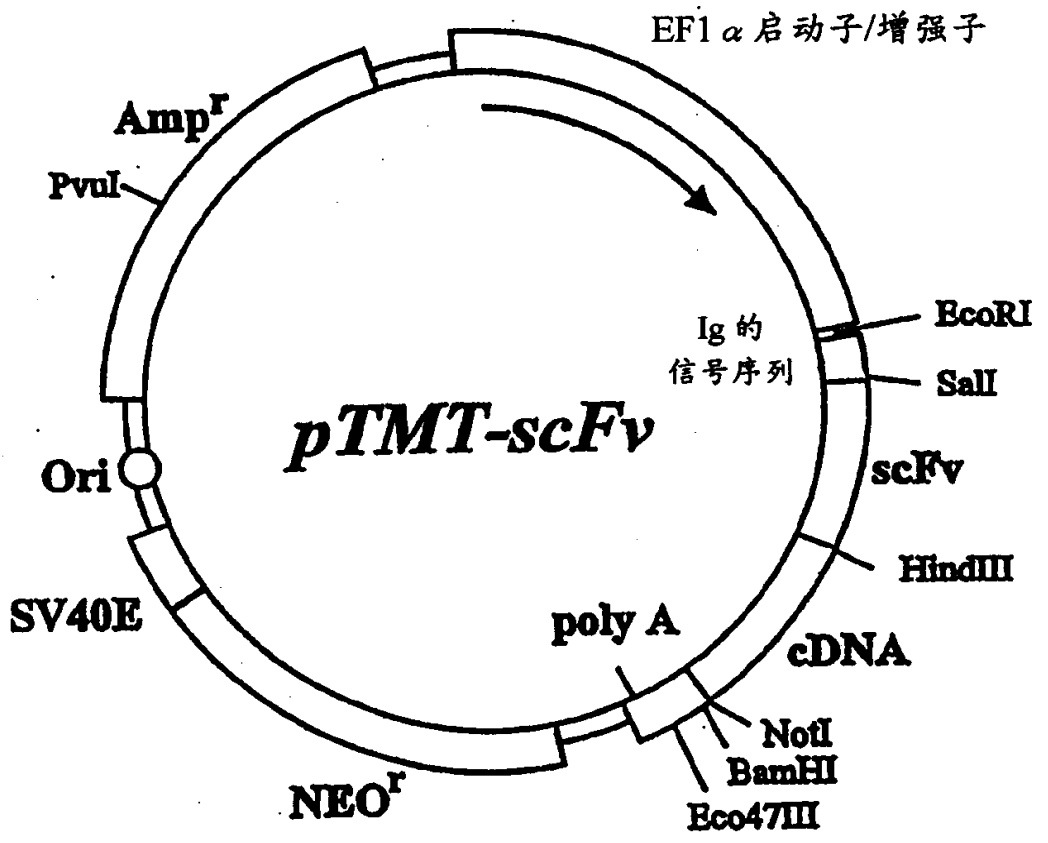


图 2

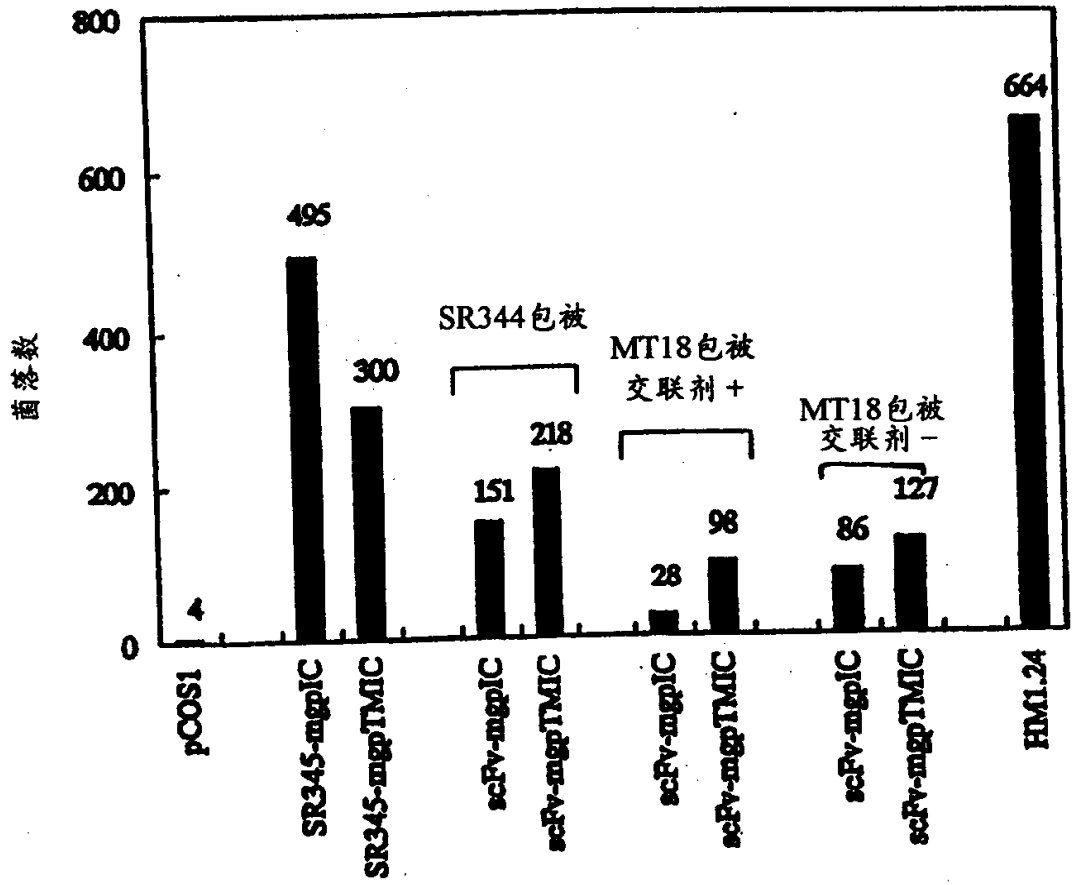


图 3

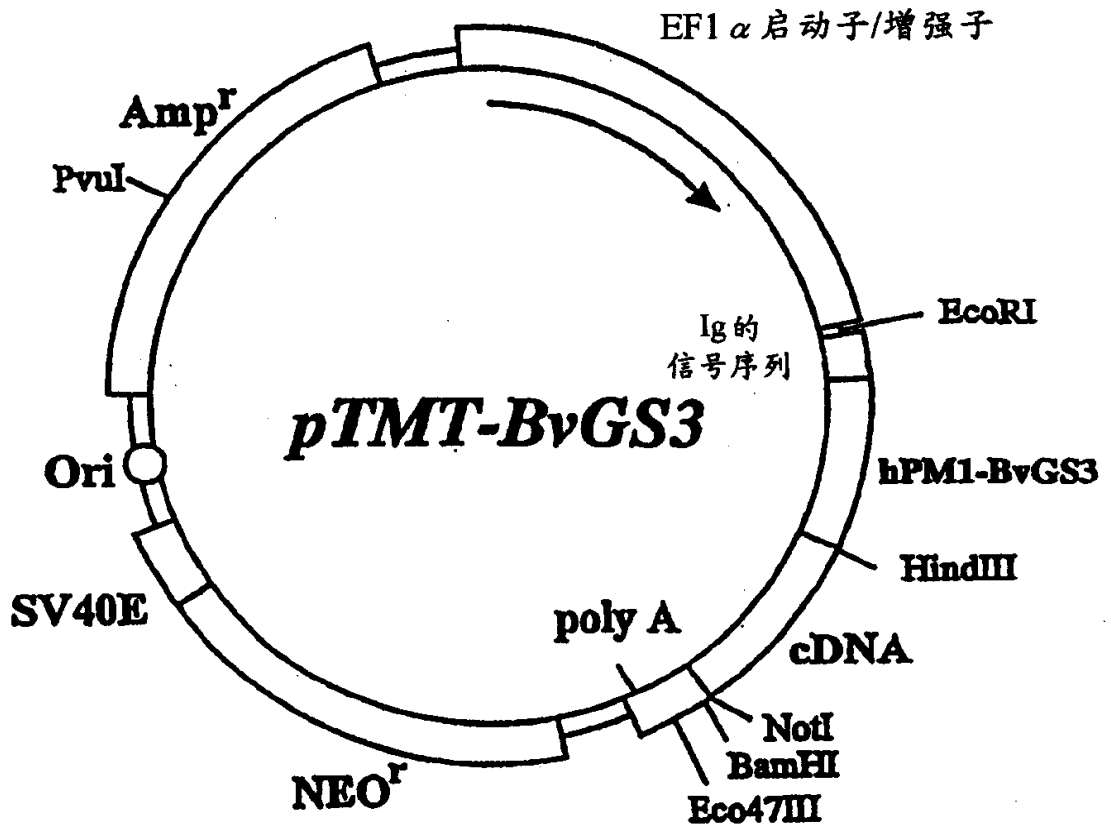


图 4

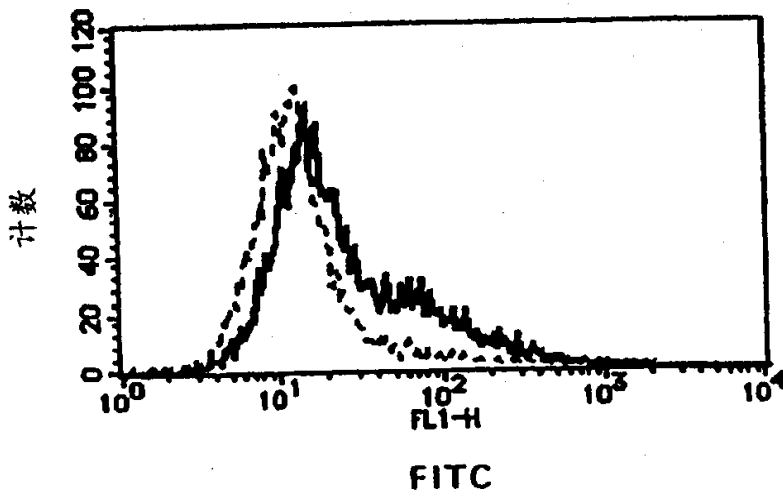
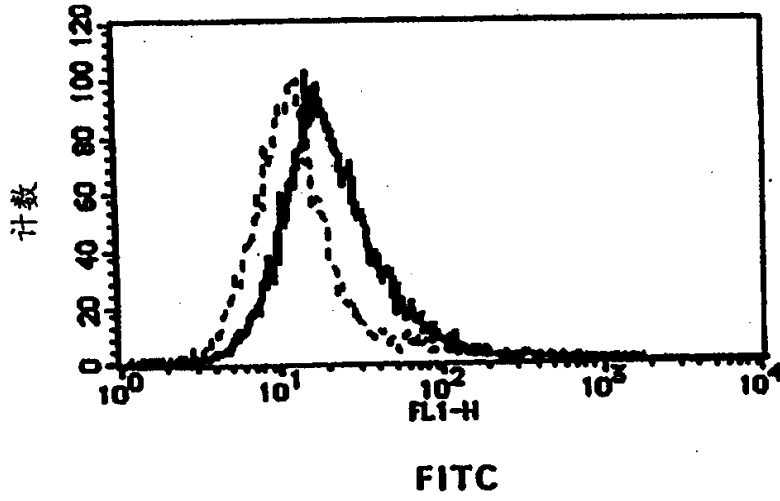
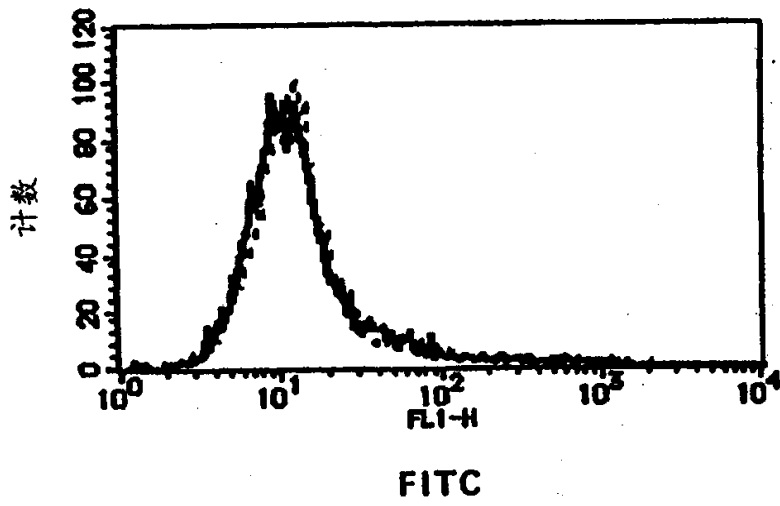


图 5

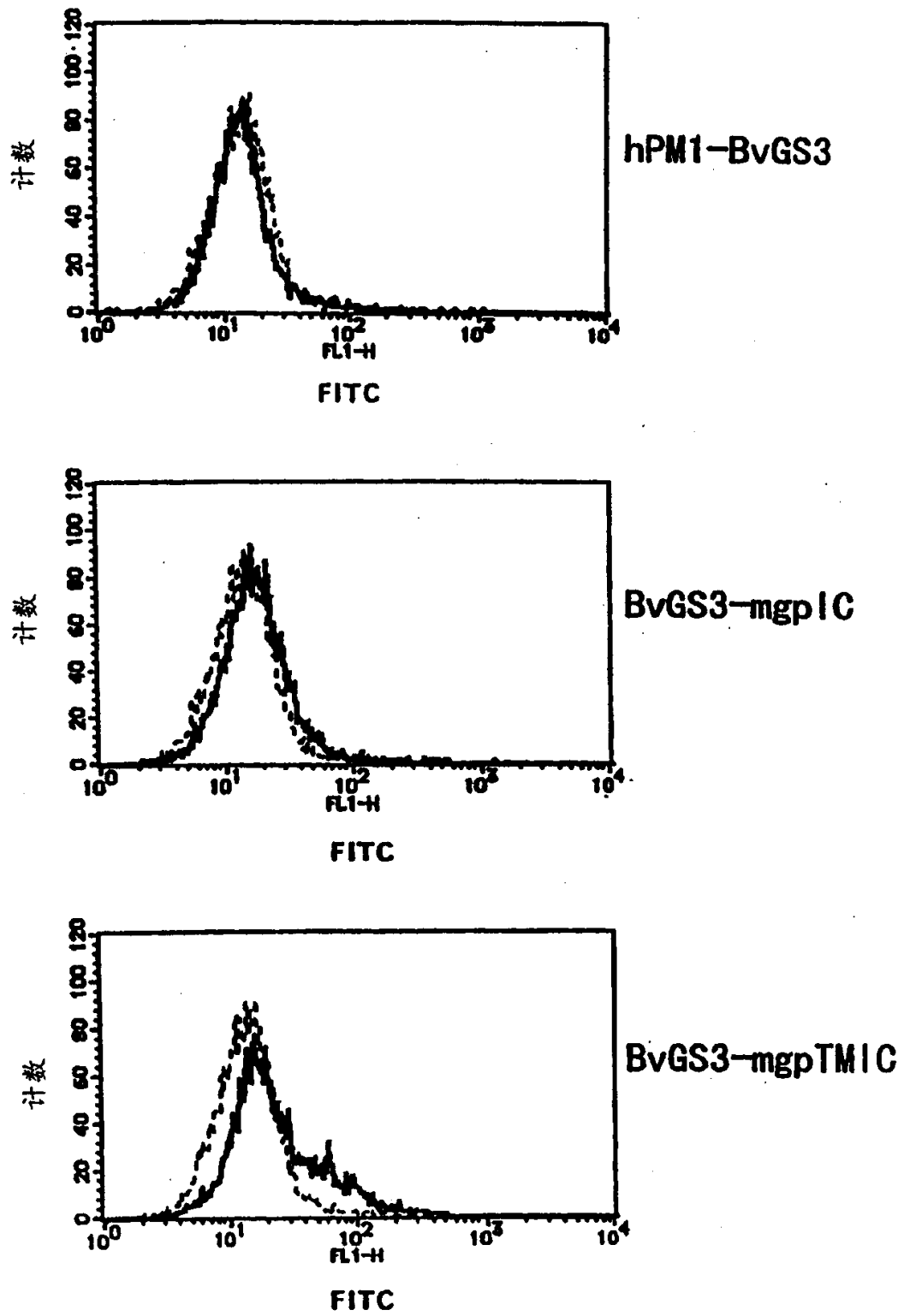


图 6

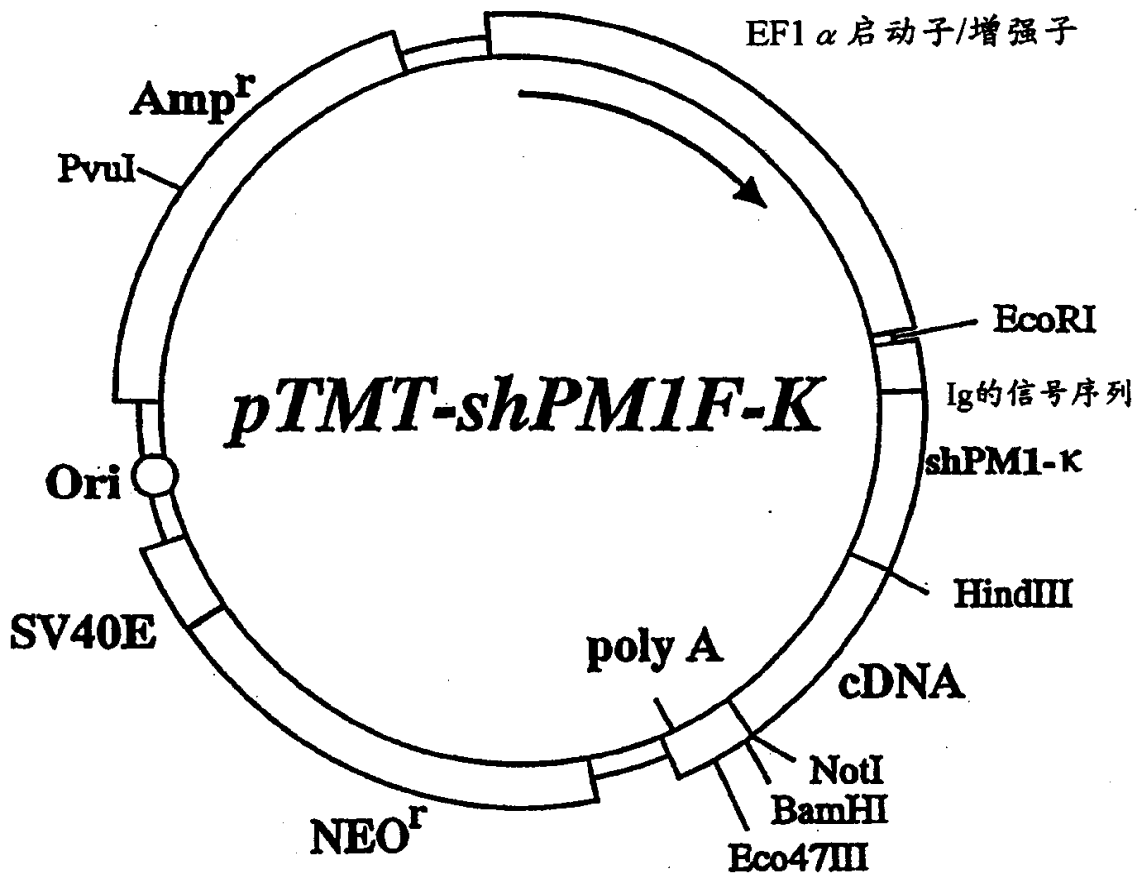


图 7