

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/37

G01N 33/569 C07K 14/00

A61K 39/00 A61P 35/00

A61P 31/00 A61P 33/00



[12] 发明专利申请公开说明书

[21] 申请号 01812119.5

[43] 公开日 2003 年 9 月 3 日

[11] 公开号 CN 1440462A

[22] 申请日 2001.4.27 [21] 申请号 01812119.5

[30] 优先权

[32] 2000. 4. 28 [33] US [31] 09/560,465

[32] 2000. 4. 28 [33] US [31] 09/561,074

[32] 2000. 4. 28 [33] US [31] 09/561,571

[32] 2000. 4. 28 [33] US [31] 09/561,572

[86] 国际申请 PCT/US01/13806 2001.4.27

[87] 国际公布 WO01/82963 英 2001.11.8

[85] 进入国家阶段日期 2002.12.30

[71] 申请人 麦康公司

地址 美国加利福尼亚州

[72] 发明人 约翰·J·L·西马德

戴维·C·戴蒙德 雷向东

[74] 专利代理机构 中科专利商标代理有限责任公
司

代理人 程金山

权利要求书 13 页 说明书 90 页 序列表 21 页
附图 18 页

[54] 发明名称 鉴定和生产抗原肽的方法并且将其
用作疫苗

[57] 摘要

本发明公开了诱导针对癌细胞和细胞内寄生物感染细胞的免疫反应的方法和疫苗。公开了含有看家表位的疫苗。看家表位是由周围细胞中的看家蛋白酶体，而不是由专门抗原呈递细胞的看家蛋白酶体形成的。含有看家表位的疫苗可以因此指导针对靶细胞的免疫反应，所述的看家表位衍生于靶细胞相关抗原。本发明也公开了治疗方法，该方法包括含有看家表位疫苗的给予。

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

- 1、 一种含有第一看家表位的疫苗，其中的看家表位衍生于与第一靶细胞相关的第一抗原。
- 5 2、 权利要求 1 中的疫苗，该疫苗包含核酸构建体，其中的核酸构建体编码看家表位。
- 3、 权利要求 1 或者 2 的疫苗，其中的靶细胞是瘤细胞。
- 4、 权利要求 3 的疫苗，其中的瘤细胞选自白血病，癌，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm 氏瘤，
10 膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。
- 5、 权利要求 3 的疫苗，其中的抗原选自 MelanA (MART-1)， gp 100 (Pmel 17)， 酪氨酸酶， TRP-1， TRP-2， MAGE-1， MAGE-3， BAGE， GAGE-1， GAGE-2， CEA， RAGE， NY-ESO， SCP-1，
15 Hom/Mel-40， PRAME， p53， H-Ras， HER-2/neu， BCR-ABL， E2A-PRL， H4-RET， IGH-IGK， MYL-RAR， EB 病毒抗原， EBNA， 人乳头状瘤病毒(HPV)抗原 E6 和 E7， TSP-180， MAGE-4， MAGE-5， MAGE-6， p185erbB2， p180erbB-3， c-met， nm-23H1， PSA， TAG-72-4， CAM 17.1， NuMa， K-ras， β -连环蛋白，
20 CDK4， Mum-1， p16， TAGE， PSMA， PSCA， CT7， 端粒酶， 43-9F， 5T4， 79ITgp72， α -胎蛋白， β -HCG， BCA225， BTAA， CA 125， CA 15-3 (CA27.29\BCAA)， CA 195， CA 242， CA-50， CAM43， CD68\KP1， CO-029， FGF-5， G250， Ga733(EpCAM)， HTgp-175， M344， MA-50， MG7-Ag， MOV18， NB/70K， NY-
25 CO-1， RCAS1， SDCCAG 16， TA-90 (Mac-2 结合蛋白\亲环蛋白 C-相关蛋白)， TAAL6， TAG72， TLP， TPS， 和 GA733-2\KSA。
- 6、 权利要求 1 或者 2 的疫苗，其中的靶细胞被细胞内寄生物感染。
- 7、 权利要求 6 的疫苗，其中的细胞内寄生物是病毒。
- 30 8、 权利要求 7 的疫苗，其中的病毒选自腺病毒，巨细胞病毒，EB 病

- 毒，单纯疱疹病毒 I，单纯疱疹病毒 II，人类疱疹病毒 6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒 B19，多瘤病毒 BK，多瘤病毒 JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人 T 细胞白血病病毒 I 和人 T 细胞白血病病毒 II。
- 5
- 9、 权利要求 7 的疫苗，其中的抗原是病毒相关抗原。
- 10、 权利要求 6 的疫苗，其中的微生物是细菌，原生动物，真菌，或者朊病毒。
- 11、 权利要求 10 的疫苗，其中的细胞内寄生虫选自衣原体，李司忒氏菌属，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次氏体属，分支杆菌属，利什曼原虫属，锥虫属（Trypanasoma），弓形虫属和疟原虫。
- 10
- 12、 权利要求 10 的疫苗，其中的抗原是寄生虫相关抗原。
- 13、 权利要求 1 或者 2 的疫苗，其中的看家表位包含或者编码大约 6—23 个氨基酸长度的多肽。
- 15
- 14、 权利要求 13 的疫苗，其中多肽的长度为 9—10 个氨基酸。
- 15、 权利要求 1 的疫苗，其中的看家表位包含核酸。
- 16、 权利要求 1 的疫苗，其中的看家表位包含合成的多肽。
- 17、 以上权利要求中任何一项的疫苗，其中的疫苗还包含至少一种选自缓冲剂，去垢剂，表面活性剂，抗氧化剂或还原剂的组分。
- 20
- 18、 以上权利要求中任何一项的疫苗，其中的看家表位对至少一种 MHC 等位基因特异。
- 19、 权利要求 18 的疫苗，其中的等位基因编码 A2 型。
- 20、 权利要求 18 的疫苗，其中的等位基因编码的类型选自 A1， A3， A11， 和 A24。
- 25
- 21、 权利要求 18 的疫苗，其中的等位基因编码的类型选自 A26， A29， B7， B8， B14， B18， B27， B35， B44， B62， B60， 和 B51。
- 22、 以上权利要求中任何一项的疫苗，该疫苗还包含免疫表位。
- 23、 权利要求 22 的疫苗，其中的免疫表位衍生于与靶细胞相关的第二抗原。
- 30

- 24、 权利要求 21 的疫苗，其中第一抗原与第二抗原相同。
- 25、 权利要求 22, 23, 或者 24 的疫苗，其中的看家表位对 MHC 第一等位基因特异，而且免疫表位对 MHC 第二等位基因特异。
- 26、 权利要求 25 的疫苗，其中第一等位基因与第二等位基因相同。
- 5 27、 权利要求 25 的疫苗，其中第一等位基因与第二等位基因不相同。
- 28、 权利要求 22—27 中任何一项的疫苗，其中疫苗包含表位簇，而且其中的表位簇包含免疫表位。
- 29、 权利要求 28 的疫苗，其中的表位簇包含或者编码一定长度的多肽，其中多肽的长度为至少 10 个氨基酸。
- 10 30、 权利要求 29 的疫苗，其中的多肽的长度小于大约 60 个氨基酸。
- 31、 权利要求 28 的疫苗，其中的表位簇衍生于与靶细胞相关的第二抗原。
- 32、 权利要求 31 的疫苗，其中第一抗原与第二抗原相同。
- 33、 权利要求 31 或者 32 的疫苗，第二抗原具有一定长度，其中,该表位簇由具有一定长度的多肽组成或编码具有一定长度的多肽，其中
- 15 多肽的长度小于第二抗原长度的大约 80%。
- 34、 权利要求 33 的疫苗，其中多肽的长度小于第二抗原长度的大约 50%。
- 35、 权利要求 34 的疫苗，其中多肽的长度小于第二抗原长度的大约 20
- 20 %。
- 36、 以上权利要求中任何一项的疫苗，还包含第二看家表位，其中的第二看家表位衍生于与第二靶细胞相关的第二抗原。
- 37、 权利要求 36 的疫苗，其中第一抗原与第二抗原相同。
- 38、 权利要求 36 的疫苗，其中第一抗原与第二抗原不同。
- 25 39、 权利要求 36 的疫苗，其中第一靶细胞与第二靶细胞相同。
- 40、 权利要求 36 的疫苗，其中第一靶细胞与第二靶细胞不同。
- 41、 制备疫苗的方法，包含的步骤有：

通过鉴定多肽片段来选择第一看家表位，所述的多肽片段被或者可以被看家蛋白酶体从第一抗原产生，所述的第一抗原与第一靶细胞或者靶病原体相关；和

30

制备包含或者编码被选择看家表位的疫苗组合物。

- 42、一种分离的 T 细胞，该 T 细胞表达对 MHC-肽复合物特异的 T 细胞受体，所述的 MHC-肽复合物包含第一看家表位，其中该看家表位衍生于与第一靶细胞相关的第一抗原。
- 5 43、包含权利要求 42 的 T 细胞的 T 细胞克隆。
- 44、包含权利要求 42 的 T 细胞的 T 细胞多克隆群体。
- 45、由体外免疫生产的权利要求 42 的 T 细胞。
- 46、从免疫动物中分离的权利要求 42 的 T 细胞。
- 47、一种产生过继免疫治疗的方法，包括：
- 10 将权利要求 42-46 中任何一项的 T 细胞与药学上可接受的佐剂，载体，稀释剂，或赋形剂相结合。
- 48、权利要求 42-46 中任何一项的 T 细胞在用于过继免疫治疗药物生产中的用途。
- 49、一种发现表位的方法，包括下面的步骤，从与宿主中的靶相关的抗原的肽片段群中选择表位，其中的肽片段对宿主的主要组织相容性复合物受体肽结合裂隙具有已知的或者预测的亲和力，其中被选择的表位与宿主细胞抗原的蛋白酶剪切产物相对应。
- 15 50、权利要求 49 的方法，其中的靶是靶细胞，其中的主要组织相容性复合物受体是 I 类主要组织相容性复合物受体，其中蛋白酶剪切产物是蛋白酶体切割产物，而且其中的宿主细胞是靶细胞。
- 20 51、权利要求 49 的方法，其中的主要组织相容性复合物受体是 II 类主要组织相容性复合物受体。
- 52、权利要求 51 的方法，其中的靶选自瘤细胞，病原体，毒素，毒液和过敏源。
- 25 53、一种发现表位的方法，包括的步骤有：
- 从靶细胞提供一段序列，其中的序列编码或者包含该靶细胞中表达的蛋白；
- 鉴定该蛋白的肽片段群，其中肽片段群的成员对主要组织相容性复合物 I 类受体肽结合裂隙具有已知的或者预测的亲和性；
- 30 从肽片段群中选择表位，其中的表位与在靶细胞中有活性的蛋

白酶体产物相对应。

54、权利要求 50 或者 53 的方法，其中的靶细胞是瘤细胞或者抗原呈递细胞。

55、一种发现表位的方法，包括下列步骤：

5 提供瘤细胞和序列，其中的序列含有或者编码与瘤细胞相关的抗原。

鉴定抗原的肽片段群，其中的肽片段群被预测对主要组织相容性复合物 I 类受体肽结合裂隙具有亲和性；

10 从肽片段群中选择表位，其中的表位经过体外分析，被确定为是瘤细胞中有活性的蛋白酶体的蛋白酶体切割反应产物。

56、权利要求 52, 54, 或者 55 的方法，其中的瘤细胞选自白血病，癌，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm 氏瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。

15 57、权利要求 52, 54, 55 或者 56 的方法，其中的抗原选自 MelanA (MART-1), gp 100 (Pmel 17), 酪氨酸酶, TRP-1, TRP-2, MAGE-1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40, PRAME, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-20 -RAR, EB 病毒抗原, EBNA, 人乳头状瘤病毒(HPV)抗原 E6 和 E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, p185erbB2, p180erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连环蛋白, CDK4, Mum-1, 和 p16。

25 58、权利要求 50 或者 53 的方法，其中的靶细胞被细胞内寄生物感染。

59、权利要求 52 的方法，其中病原体是细胞内寄生物。

60、权利要求 58 或者 59 的方法，其中的细胞内寄生物是病毒。

61、权利要求 60 的方法，其中的病毒选自腺病毒，巨细胞病毒，EB 病毒，单纯疱疹病毒 I，单纯疱疹病毒 II，人类疱疹病毒 6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细

30

- 小病毒 B19, 多瘤病毒 BK, 多瘤病毒 JC, 丙型肝炎病毒, 麻疹病毒, 风疹病毒, 人类免疫缺陷病毒 (HIV), 人 T 细胞白血病病毒 I 和人 T 细胞白血病病毒 II, 轮状病毒, 呼肠孤病毒, 细小核糖核酸病毒、肝 DNA 病毒属 (hepadnavirus)、腺病毒, 和乳多空病毒。
- 5 62、权利要求 60 的方法, 其中的抗原选自腺病毒, 巨细胞病毒, EB 病毒, 单纯疱疹病毒 I, 单纯疱疹病毒 II, 人类疱疹病毒 6, 水痘一带状疱疹病毒, 乙型肝炎病毒, 丁型肝炎病毒, 乳头状瘤病毒, 细小病毒 B19, 多瘤病毒 BK, 多瘤病毒 JC, 丙型肝炎病毒, 麻疹病毒, 风疹病毒, 人类免疫缺陷病毒 (HIV), 人 T 细胞白血病病毒 I 和人 T 细胞白血病病毒 II。
- 10 63、权利要求 52 的方法, 其中的病原体选自细菌, 真菌, 病原动物和朊病毒。
- 64、权利要求 58 的方法, 其中的细胞内寄生物是细菌, 原生动物, 真菌, 或者朊病毒。
- 15 65、权利要求 63 或者 64 的方法, 其中的病原体或者细胞内寄生物选自衣原体, 李司忒氏菌属, 沙门氏菌属, 军团菌属, 布鲁氏菌属, 柯克斯菌属, 立克次氏体属, 分支杆菌属, 利什曼原虫属, 锥虫属 (Trypanasoma), 弓形虫属和疟原虫。
- 66、权利要求 58 的方法, 其中的抗原是寄生物相关抗原。
- 20 67、权利要求 50, 53, 或者 55 的方法, 其中的蛋白酶体是看家蛋白酶体。
- 68、权利要求 50, 53, 或者 55 的方法, 其中的蛋白酶体是免疫蛋白酶体。
- 69、权利要求 49 或者 51 的方法, 其中的蛋白酶剪切含有宿主溶酶体中的蛋白酶活性。
- 25 70、权利要求 49 或者 51 的方法, 其中的蛋白酶剪切含有宿主内体中的蛋白酶活性。
- 71、权利要求 49—70 中任何一项的方法, 其中的亲和性由体外结合分析确定。
- 30 72、权利要求 49—70 中任何一项的方法, 其中的亲和性由算法确定。

- 73、 权利要求 49—72 中任何一项的方法，其中的产物由体外结合分析确定。
- 74、 依照权利要求 49—73 的任何一项的方法发现的表位。
- 75、 权利要求 74 的表位，该表位含有序列 SEP ID NO: 78。
- 5 76、 MHC I 类限制的表位，该表位通过使用包含下列步骤的方法发现：确定靶细胞表达的蛋白酶体的类型，和鉴定该靶细胞表达的表位。
- 77、 权利要求 76 的表位，其中在确定步骤中包含的技术选自 RT-PCR，ELISA，蛋白质印迹，聚丙烯酰胺凝胶电泳，免疫组织化学，和用
10 纯化自所述靶细胞的蛋白酶体切割标准基质。
- 78、 权利要求 76 的表位，其中的确定步骤包含将未处理的靶细胞与 γ -干扰素处理的靶细胞相比较。
- 79、 权利要求 76 的表位，其中在鉴定步骤中包含的技术选自肽洗脱，体外蛋白水解，质谱分析法，HPLC，氨基酸序列测序，和免疫反
15 应性的确定。
- 80、 权利要求 79 的表位，其中免疫反应性的确定包含一种检测条件的鉴定方法，所述的条件选自细胞毒性，增殖，和细胞因子产生。
- 81、 包含权利要求 74—80 中任何一项的表位的疫苗。
- 82、 一种表位簇，该簇衍生于靶相关抗原，该簇包含或者编码至少两段
20 对 MHC 受体肽结合裂隙具有已知的或者预测的亲亲和性的序列，其中的簇是抗原的不完全片断。
- 83、 权利要求 82 的表位簇，其中的靶细胞是瘤细胞。
- 84、 权利要求 82 的表位簇，其中的靶是被细胞内寄生物感染的细胞。
- 85、 权利要求 84 的表位簇，其中的细胞内寄生物是病毒。
- 25 86、 权利要求 84 的表位簇，其中的细胞内寄生物是细菌或者原生动物。
- 87、 权利要求 82 的表位簇，其中的靶是病原体。
- 88、 权利要求 87 的表位簇，其中的病原体选自病毒，细菌，真菌，原生动物，朊病毒，毒素，毒液，和过敏源。
- 89、 权利要求 82 的表位簇，其中的 MHC 受体是 I 类 HLA 受体。
- 30 90、 权利要求 82 的表位簇，其中的 MHC 受体是 II 类 HLA 受体。

- 91、权利要求 82 的表位簇，其中的簇包含或者编码具有一定长度的多肽，该长度至少是 10 个氨基酸。
- 92、权利要求 91 的表位簇，其中多肽的长度小于大约 75 个氨基酸。
- 93、权利要求 82 的表位簇，抗原具有一定的长度，其中的簇由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度
5 小于大约 80% 抗原的长度。
- 94、权利要求 93 的表位簇，抗原具有一定的长度，其中的簇由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度小于大约 50% 抗原的长度。
- 10 95、权利要求 94 的表位簇，抗原具有一定的长度，其中的簇由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度小于大约 20% 抗原的长度。
- 96、一种发现表位簇的方法，包含的步骤有：
提供一段靶相关抗原的序列；
15 基于对 MHC 受体肽结合裂隙的已知的或者预测的亲合力，给序列中的候选肽打分，以鉴定假定的 MHC 表位；和
鉴定抗原中的表位簇，其中的簇包含至少两个假定的 MHC 表位，而且其中的簇所包含的假定 MHC 表位的密度要比整个抗原中假定 MHC 表位的密度要高。
- 20 97、权利要求 96 的方法，其中的靶细胞是瘤细胞。
- 98、权利要求 96 的方法，其中的靶是被细胞内寄生物感染的细胞。
- 99、权利要求 98 的方法，其中的细胞内寄生物是病毒。
- 100、权利要求 98 的方法，其中的细胞内寄生物是细菌或者原生动物。
- 101、权利要求 96 的方法，其中的靶是病原体。
- 25 102、权利要求 96 的方法，其中的 MHC 受体结合裂隙是 I 类 HLA 受体结合裂隙。
- 103、权利要求 96 的方法，其中的 MHC 受体结合裂隙是 II 类 HLA 受体结合裂隙。
- 104、权利要求 96 的方法，其中的簇包含或者编码具有一定长度的多肽，
30 该长度至少是 10 个氨基酸。

- 105、权利要求 104 的方法，其中多肽的长度小于大约 75 个氨基酸。
- 106、权利要求 96 的表位簇，抗原具有一定的长度，其中的簇基本上是由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度小于大约 80% 抗原的长度。
- 5 107、权利要求 106 的表位簇，抗原具有一定的长度，其中的簇基本上是由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度小于大约 50% 抗原的长度。
- 108、权利要求 107 的表位簇，抗原具有一定的长度，其中的簇基本上是由具有一定长度的多肽组成或者编码具有一定长度的多肽，其中多肽的长度小于大约 20% 抗原的长度。
- 10 109、一种核酸构建体，它包含第一编码区域，其中的第一编码区域包含至少编码第一多肽的第一序列，其中的第一多肽包含衍生于与第一靶细胞相关的第一抗原的第一看家表位。
- 110、权利要求 109 的核酸构建体，其中的第一编码区还包含编码至少一个第二多肽的第二序列，其中的第二多肽包含衍生于与第二靶细胞相关的第二抗原的第二表位。
- 15 111、权利要求 110 的核酸构建体，其中的第一多肽和第二多肽是邻接的。
- 112、权利要求 110 的核酸构建体，其中的第一多肽和第二多肽不是邻接的。
- 20 113、权利要求 110 的核酸构建体，其中的第二多肽是看家表位。
- 114、权利要求 110 的核酸构建体，其中的第二多肽是免疫表位。
- 115、权利要求 110 的核酸构建体，其中的第一抗原和第二抗原是相同的。
- 25 116、权利要求 110 的核酸构建体，其中的第一抗原和第二抗原是不相同的。
- 117、权利要求 110 的核酸构建体，其中的第一靶细胞和第二靶细胞是相同的。
- 118、权利要求 110 的核酸构建体，其中的第一靶细胞和第二靶细胞不是相同的。
- 30

- 119、权利要求 109 的核酸构建体，其中的第一靶细胞是瘤细胞。
- 120、权利要求 119 的核酸构建体，其中的瘤细胞选自白血病，癌，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm 氏瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。
- 121、权利要求 119 的核酸构建体，其中的第一抗原选自 MART-1/MelanA, gp 100 (Pmel 17), 酪氨酸酶, TRP-1, TRP-2, MAGE-1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, SCP-1, Hom/Mel-40, PRAME, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, h4-RET, IGH-IGK, MYL-RAR, EB 病毒抗原, EBNA, 人乳头状瘤病毒(HPV)抗原 E6 和 E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, p185erbB2, p180erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, b-连环蛋白, CDK4, Mum-1, p16, p15, NY-ESO, SSX 基因家组成员的产物, CT-7, 和 SCP 基因家组成员的产物。
- 122、权利要求 109 的核酸构建体，其中的靶细胞被细胞内寄生物感染。
- 123、权利要求 122 的核酸构建体，其中的细胞内寄生物是病毒。
- 124、权利要求 123 的核酸构建体，其中的病毒选自腺病毒，巨细胞病毒，EB 病毒，单纯疱疹病毒 I，单纯疱疹病毒 II，人类疱疹病毒 6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒 B19，多瘤病毒 BK，多瘤病毒 JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒 (HIV)，人 T 细胞白血病病毒 I 和人 T 细胞白血病病毒 II。
- 125、权利要求 122 的核酸构建体，其中的细胞内寄生物是细菌，原生动物，真菌，或者朊病毒。
- 126、权利要求 125 的核酸构建体，其中的细胞内寄生物选自衣原体，李司忒氏菌属，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次氏体属，分支杆菌属，利什曼原虫属，锥虫属 (Trypanasoma)，弓形虫属和疟原虫。

- 127、权利要求 109 的核酸构建体，该构建体还包括与第一编码区有效连接的第一启动子序列。
- 128、权利要求 127 的核酸构建体，其中的启动子衍生于巨细胞病毒（CMV）、SV40、或者逆转录病毒的长末端重复序列（LTR）。
- 5 129、权利要求 127 的核酸构建体，其中的启动子是双向启动子。
- 130、权利要求 127 的核酸构建体，该构建体还包括与第二编码区有效连接的第二启动子序列。
- 131、权利要求 109 的核酸构建体，该构建体还包括与第一编码区有效连接的聚合 A 序列。
- 10 132、权利要求 110 的核酸构建体，该构建体还包括内部核糖体进入位点（IRES）序列。
- 133、权利要求 132 的核酸构建体，其中的 IRES 序列位于第一和第二编码序列之间。
- 134、权利要求 110 的核酸构建体，该构建体还包含泛素序列。
- 15 135、权利要求 134 的核酸构建体，其中的泛素序列位于第一和第二编码序列之间。
- 136、权利要求 110 的核酸构建体，该构建体还包括一段编码自催化肽的序列。
- 137、权利要求 136 的核酸构建体，其中的自催化肽编码序列位于第一编码序列和第二编码序列之间。
- 20 138、权利要求 109 的核酸构建体，其还包含选自增强子，核进入序列，免疫刺激序列，和细胞因子、选择标记、或者报告分子的表达盒。
- 139、权利要求 109 的核酸构建体，其中第一多肽的长度为大约 7—15 个氨基酸。
- 25 140、权利要求 109 的核酸构建体，其中第一多肽的长度为大约 9 或者 10 个氨基酸。
- 141、权利要求 109 的核酸构建体，其中第二多肽的长度为大约 9 或者 10 个氨基酸。
- 142、权利要求 110 的核酸构建体，其中第二多肽是大约 10—大约 75 个氨基酸长度的表位簇。
- 30

- 143、权利要求 110 的核酸构建体，其中的第一表位具有对第一 MHC 等位基因的结合亲和性，并且其中的第二表位具有对第二 MHC 等位基因的结合亲和性。
- 144、权利要求 143 的核酸构建体，其中的第一等位基因和第二等位基因是相同的。
- 145、权利要求 143 的核酸构建体，其中的第一等位基因和第二等位基因是不同的。
- 146、权利要求 109 的核酸构建体，其还包含至少一个第二编码区，其中的第二编码区包含一个编码至少一个第二多肽的第二序列，其中的第二多肽包含衍生于与第二靶细胞相关的第二抗原的第二表位。
- 147、权利要求 146 的核酸构建体，其中所述的第一和第二编码区被有效地连接到单个双向启动子的任一侧。
- 148、权利要求 146 的核酸构建体，其中的至少一个编码区包含编码多个多肽的多个序列，其中的每个多肽含有至少一个衍生于与靶细胞相关的抗原的表位。
- 149、权利要求 129 的核酸构建体，其中的双向启动子还被有效地连接到第二序列，该第二序列编码一种佐剂。
- 150、权利要求 136 的核酸构建体，其中的编码的多肽具有自动蛋白水解活性。
- 151、权利要求 136 的核酸构建体，其中的编码的多肽抑制肽键形成。
- 152、含有权利要求 109—151 中任何一项的核酸构建体的疫苗。
- 153、一种制作疫苗的方法，其包含下面的步骤：制备含有权利要求 109—151 中任何一项的核酸构建体的疫苗组合物。
- 154、根据权利要求 41 或者 153 的方法制造的疫苗。
- 155、一种治疗动物的方法，其包含：将权利要求 1—40，81，152，或者 154 中任何一项的疫苗给予动物。
- 156、权利要求 155 的方法，其中的给予步骤包括选自经皮肤，结节内，结节周，经口，静脉内，真皮内，肌肉内，腹膜内，和粘膜的传送模式。
- 157、权利要求 155 的方法，其还包括一个分析步骤来确定靶细胞状态的

特征指示。

- 158、权利要求 157 的方法，其包括第一分析步骤和第二分析步骤，其中第一分析步骤先于所述的给予步骤，并且第二分析步骤在所述的给予步骤之后。
- 5 159、权利要求 158 的方法，其还包括将在第一分析步骤中确定的特征与在第二分析步骤中确定的特征相比较以获得结果。
- 160、权利要求 159 的方法，其中的结果选自：靶细胞数量的减少、包含靶细胞的肿瘤的质量和大小的小、或者感染靶细胞的细胞内寄生物的数量或者浓度的降低。
- 10 161、看家表位在疫苗生产中的用途，该疫苗用于权利要求 155—160 中任何一项的治疗方法中。

鉴定和生产抗原肽的方法并且将其用作疫苗

5

发明背景

发明领域

本发明涉及诱导抗原呈递细胞呈递特殊的靶细胞特异的表位，从而
10 促进对靶细胞有效的细胞毒性 T 细胞反应的方法和组合物。

本发明还涉及靶细胞表位和表位簇的鉴别，也涉及表位编码载体，
该载体可用于产生免疫活性药物组合物。这些组合物在被给予时，可以
刺激受试者的免疫系统对呈现靶抗原的靶细胞发动免疫反应。因此本发
明在治疗和预防肿瘤和病毒疾病中有用。

15

相关技术描述

瘤的形成与免疫系统

通常被称为癌症的肿瘤性病状，一般被认为是由于单细胞生长失控
而引起的。典型地，失控生长状态由多步骤过程导致，在该多步骤过程
20 中，发生了一系列细胞系统损坏，引起瘤细胞发生。这样产生的瘤细胞
快速自我复制，形成一个或多个肿瘤，最终可能导致宿主死亡。

由于瘤细胞的先祖具有与宿主细胞相同的遗传物质，所以瘤细胞在
很大程度上逃避于宿主的免疫系统。在宿主免疫系统检查和定位外来物
质的免疫监视过程中，瘤细胞会在宿主的免疫监视机构中呈现为自我细
25 胞。

病毒和免疫系统

与癌细胞对比，病毒感染包括明显的非我抗原的表达。结果，免疫
系统成功地处理了许多病毒感染，而只伴随着最小限度的临床后遗症。
此外，对那些引起严重疾病的感染，形成有效的疫苗是可能的。许多疫
30 苗方法已经被成功地用来与多种疾病作斗争。这些方法包括个体蛋白组

成的亚单位疫苗，该个体蛋白是通过重组DNA技术生产的。尽管有这些进展，但是，用作病毒疫苗的最小表位的选择和有效给予仍然是有困难的。

除了涉及表位选择的那些困难，还存在病毒的难题，这些病毒已经演化出了逃避宿主免疫系统的能力。许多病毒，特别是建立了持续感染的病毒，例如疱疹病毒和痘病毒家族的成员，产生了允许病毒逃避宿主免疫系统的免疫调节分子。可以通过导向作为免疫原性组合物给药的选择表位，来克服这些免疫调节分子对抗原呈递的影响。为了更好地理解肿瘤细胞和病毒感染细胞与宿主免疫系统的相互作用，对该系统的组分讨论如下。

免疫系统发挥功能，从外来的或与有机体无关的物质（“非我”分子）中，辨别有机体内源性分子（“自我”分子）。基于介导反应的组分，免疫系统对异物有两种类型的适应反应：体液反应和细胞介导的反应。体液反应由抗体介导，而细胞介导的反应涉及被分类为淋巴细胞的细胞。近来的抗癌和抗病毒策略集中在：将调动宿主的免疫系统作为抗癌和抗病毒处理或治疗的手段。

免疫系统在三个阶段发挥作用，来保护宿主免于异物侵害：认知期，活化期，和效应期。在认知期，免疫系统对体内的外来抗原或侵入者识别并且发出信号。外来抗原可以是，例如，来自于瘤细胞或病毒蛋白的细胞表面标记。一旦系统意识到有侵入体，免疫系统的抗原特异性细胞增生并且分化，以此来响应侵入者触发的信号。最后一个阶段是效应期，在效应期，免疫系统的效应细胞作出反应并消除被发现的侵入物。

一批效应细胞执行着对侵入物的免疫反应。一种类型的效应细胞——B细胞，产生针对宿主遇到的外来抗原的抗体。抗体与补体系统相结合，指导着对含有靶抗原的细胞或有机体的破坏。另一种类型的效应细胞是自然杀伤细胞（NK细胞），这是一种淋巴细胞，该淋巴细胞是具有自发识别并破坏多种病毒感染细胞和恶性细胞能力的细胞。我们对NK细胞所使用的识别靶细胞的方法知道的很少。

另一种类型的效应细胞——T细胞，含有被分为三种亚类的细胞，每一种亚类都在免疫反应中起着不同的作用。辅助性T细胞分泌刺激其它细

胞增殖的细胞因子，所述的其它细胞对发动有效的免疫反应是必需的；而抑制性T细胞对免疫反应进行负调节。第三类的T细胞——细胞毒性T细胞（CTL）能够直接将细胞表面呈现有外来抗原的靶细胞溶解。

主要组织相容性复合物和T细胞靶识别

5 T细胞是抗原特异的免疫细胞，它们响应特异抗原信号而发挥作用。B淋巴细胞和它们产生的抗体也是抗原特异的实体。然而，T细胞与B细胞不同，它不与自由或溶解形式的抗原响应。当T细胞与抗原响应的时候，它需要抗原结合到一个被称作主要组织相容性复合物(MHC)的呈递复合物上。

10 MHC复合物蛋白提供了某些方法，通过这些方法，T细胞将外来细胞与自体细胞或自我细胞鉴别开来。存在两种类型的MHC，I类MHC和II类MHC。辅助性T细胞（CD4⁺）主要与II类MHC蛋白相互作用，而细胞毒性T细胞（CD8⁺）主要与I类MHC蛋白相互作用。两种MHC复合物都是跨膜蛋白，它们的大部分结构在细胞的外表面。此外，在两种类型的MHC的外部都有一个肽结合裂隙。正是在这个裂隙中，自体或外来的小片段蛋白被结合并被呈现在细胞外环境中。

被称作抗原呈递细胞（APCs）的细胞，通过使用MHC复合物将抗原呈递给T细胞。为了使T细胞识别抗原，抗原必须呈现在MHC复合物上以供识别。这种需要称作MHC限制，而且就是通过这种机制T细胞将自我细胞从非我细胞辨别出来。如果一种抗原不能通过可识别的MHC复合物呈递，T细胞就不会识别并作用于该抗原信号。对肽（结合于可识别MHC复合物的肽）特异的T细胞，结合到这些MHC-肽复合物上并且进行下一阶段的免疫反应。

25 如上所述，瘤细胞在很大程度上被免疫系统忽视。研究者正在作出大量的努力，试图利用宿主的免疫系统来帮助与宿主体内出现的瘤细胞作斗争。这些研究的其中一个领域涉及抗癌疫苗制剂。

抗癌疫苗

对于肿瘤学家来说，受试者的免疫系统是在与癌症的斗争中可以使用的多种武器中的一种。已经做过了多种尝试来引起免疫系统与癌症或肿瘤疾病作斗争，不幸的是，迄今为止，结果很令人失望。一个特别重

要的领域涉及抗癌疫苗的生产和使用。

为了产生疫苗或其它免疫原性的组合物，必须要给受试者引入一种可以引发免疫反应的抗原或表位。尽管瘤细胞衍生于正常细胞，并且因此遗传水平上与正常细胞基本相同，但是已知许多瘤细胞呈现出肿瘤相关抗原（TuAAs）。理论上讲，受试者的免疫系统可以识别这些抗原并攻击瘤细胞。不幸的是，宿主的免疫系统看起来却忽视了瘤细胞。

研究者研究出了大量不同的策略，试图产生出针对瘤细胞的活性疫苗。这些策略包括将肿瘤相关抗原作为免疫原。例如在美国专利 No. 5,993,828中描述了一种方法，该方法通过给予受试者有效剂量的、含有失活肿瘤细胞和至少一种肿瘤相关抗原的组合物，产生针对泌尿肿瘤相关抗原特殊亚单位的免疫反应，其中所述的失活肿瘤细胞在细胞表面含有泌尿肿瘤相关抗原，所述的肿瘤相关抗原是至少一种选自GM-2，GD-2，胚胎性抗原，和黑素瘤相关抗原的肿瘤相关抗原。因而，该专利描述了使用整体的，失活的肿瘤细胞作为抗癌疫苗中的免疫原。

抗癌疫苗使用的另一策略包括：给予含有分离的肿瘤抗原的组合物。在一种方法中，MAGE-A1抗原肽被用作免疫原(见 Chaux, P., 等, “通过MAGE-A1转导的Dendritic细胞的体外刺激得到的、由细胞裂解T淋巴细胞识别的5种MAGE-A1的鉴别,” J. Immunol., 163(5):2928-2936 (1999))。尽管接种方案的效果有限，但是已经有几个治疗试验将MAGE-A1肽用于接种。这些试验中的一些试验的结果在Vose, J.M., “由T淋巴细胞识别的肿瘤抗原,” 10th European Cancer Conference, Day 2, Sept. 14, 1999.中讨论。

在将肿瘤相关抗原用作疫苗的另一个实例中，Scheinberg等对已经接受了5剂I类相关bcr-abl肽的干扰素（INF）或羟基脲的12个慢性粒性白血病（CML）患者，给予了辅助肽和助剂QS-21。Scheinberg, D.A., et al, “衍生于BCR-ABL断裂点的癌基因融合肽疫苗在慢性粒性白血病（CML）患者中产生特异性免疫反应 [Abstract 1665]”, American Society of Clinical Oncology 35th Annual Meeting, Atlanta (1999). 在新鲜血液样品中，诱导出了表明T辅助细胞活性的T细胞增生和迟发型超敏反应（DTH），但是没有观察到细胞溶解杀伤T细胞的活性。

在Cebon等和Scheibenbogen等的近期工作中，可以看到试图鉴定用作疫苗的TAAs的附加实例。Cebon等使用真皮内给予MARI-1₂₆₋₃₅肽，伴随皮下或者静脉内给予增量的IL-2的方法，对患有黑色素瘤的患者进行免疫。注意到在最初的15名患者中，1人完全缓解，1人部分缓解，一人有混合反应。T细胞产生的免疫测定包括DTH，在存在或者不存在IL-12的患者中观察到了这种反应。在伴随有临床有益迹象的患者中，可以看到阳性结果的CTL测定，但是在没有肿瘤退化的患者中却看不到。Cebon, 等, “在HLA A2+阳性的患有第III和IV期恶性黑色素瘤的患者中用Melan-A和IL-12免疫的第I期研究,” [Abstract 1671], American Society of Clinical Oncology 35th Annual Meeting, Atlanta (1999)。

Scheibenbogen等用4 HLA I类限制性酪氨酸酶肽对18名患者进行免疫，其中16名患者患有转移黑色素瘤，2名为辅助患者。Scheibenbogen, et al, “在转移黑色素瘤患者中用酪氨酸酶肽和GM-CSF接种：第II期试验,” [Abstract 1680], American Society of Clinical Oncology 35th Annual Meeting, Atlanta (1999)。在4/15的患者，2名辅助患者和2名出现了肿瘤退化迹象的患者中观察到了增加的CTL活性。就像Cebon等的试验中一样，患有累进性疾病的患者没有显示增高的免疫力。尽管迄今为止作了多种尝试来产生有效的抗癌疫苗，但是仍没有研究出这样的组合物。

保护病毒疾病的疫苗方法已经获得了许多成功。或许这些成功中最有名的就是在天花疾病中取得的进步了，天花已经灭绝。脊髓灰质炎疫苗也获得了类似程度的成功。

病毒疫苗可以被分为3类：活的衰减病毒疫苗，例如天花牛痘，萨宾脊髓灰质炎疫苗，麻疹病毒性腮腺炎和风疹；整体死亡或失活的病毒疫苗，例如Salk脊髓灰质炎疫苗，甲型肝炎病毒疫苗和典型的流感疫苗；和亚单位疫苗，例如乙型肝炎。由于亚单位疫苗缺少完全的病毒基因组，它们比基于全病毒的那些疫苗在安全性上要好多。

一个成功的亚单位疫苗的范例是基于病毒包膜蛋白的重组乙型肝炎疫苗。尽管学术界对于将亚单位的概念从单个蛋白上升到个体表位很感兴趣，但是这些努力还没有带来很多成果。病毒疫苗研究同样集中在诱导抗体反应，虽然细胞反应也会发生。但是，许多亚单位制剂在产生CTL

反应方面特别差。

发明概述

本发明涉及一些方法和组合物，这些方法和组合物用于使抗原呈递细胞呈递特殊的靶细胞特异的抗原表位，从而促进对靶细胞的延长的，
5 定向的细胞毒性 T 细胞反应。

含有看家表位的疫苗

在本发明的一个部分中，提供了一种含有看家表位的疫苗，所述的看家表位衍生于与靶细胞相关的抗原。有利地，靶细胞可以是瘤细胞。
10 瘤细胞可以是任何与实体瘤或淋巴瘤相关的转化细胞，例如白血病，淋巴瘤，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm氏瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。或者，靶细胞可以被细胞内寄生物感染。例如细胞内寄生物可以是病毒如腺病毒，巨
15 细胞病毒，EB病毒，单纯疱疹病毒，人类疱疹病毒6，水痘一带状疱疹病毒，肝炎病毒，乳头状瘤病毒，细小病毒，多瘤病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），或人T细胞白血病病毒。细胞内寄生物可以是细菌，原生动物，真菌或朊病毒。更具体而言，细胞内病毒可以是衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌属，利什曼虫属，锥虫属，弓形体和疟原虫。
20

看家表位可以衍生于与靶细胞相关的抗原。抗原可以是 MelanA (MART-I), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE- 1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40,
25 PRAME, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, EB病毒抗原, EBNA, 人乳头状瘤病毒(HPV)抗原E6和E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, pl85erbB2, pl80erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连环蛋白, CDK4, Mum-1, p16, TAGE, PSMA, PSCA, CT7, 端粒酶, 43-9F, 5T4, 79ITgp72, α -胎蛋白 ,
30 β -HCG, BCA225, BTAA, CA 125, CA 15-3 (CA27.29\BCAA), CA 195, CA

242, CA-50, CAM43, CD68\KP1, CO-029, FGF-5, G250, Ga733(EpCAM), HTgp-175, M344, MA-50, MG7-Ag, MOV18, NB/70K, NY-CO-1, RCAS1, SDCCAG 16, TA-90 (Mac-2 结合蛋白\亲环蛋白C-相关蛋白), TAAL6, TAG72, TLP, TPS, GA733-2\KSA 等等。任选地, 抗原可以是病毒相关抗原。在本发明的另一方面, 抗原可以是寄生物相关抗原。

在本发明的另一方面中, 看家表位可以包括或编码一条长度为6-23个氨基酸的多肽。优选地, 该多肽长度为9-10个氨基酸。该多肽可以是合成多肽。有利地, 该疫苗额外地包括缓冲剂, 去垢剂, 表面活性剂, 抗氧化剂或还原剂。在该疫苗的另一方面中, 看家表位包括核酸。在优选实施方案中, 看家表位对至少一条MHC等位基因特异。该等位基因可以编码类型A1, A2, A3, A11, A24, A26, A29, B7, B8, B14, B18, B27, B35, B44, B62, B60, 或B51。

在本发明的另一方面中, 疫苗可以包括免疫表位。任选地, 免疫表位衍生于与靶细胞相关的第二抗原。第一抗原和第二抗原可以相同也可以不同。有利地, 看家表位对MHC第一等位基因特异, 免疫表位对MHC的第二等位基因特异。第一等位基因与第二等位基因可以相同也可以不同。

在本发明的另一方面中, 疫苗包括含有免疫表位的表位簇(见下面)。表位簇可以衍生于与靶细胞相关的第二抗原。第一抗原和第二抗原可以相同也可以不同。有利地, 表位簇可以包括或编码长度为至少10个氨基酸, 但少于60个氨基酸的多肽。优选地, 表位簇的多肽链的长度小于大约80%, 50%, 或 20%第二抗原的长度。

在本发明的另一方面中, 疫苗还包括第二看家表位, 该第二看家表位衍生于与第二靶细胞相关的第二抗原。任选地, 第一抗原和第一抗原可以相同。或者, 第一抗原和第二抗原不同。类似地, 第一靶细胞和第二靶细胞可以相同也可以不同。

有利地, 本发明的疫苗可以包括核酸构建体, 该核酸构建体编码衍生于靶细胞相关抗原的看家表位。优选地, 靶细胞是瘤细胞。瘤细胞可以是任何与实体瘤或淋巴瘤相关的转化细胞, 例如白血病, 癌瘤, 淋巴瘤, 星形细胞瘤, 肉瘤, 胶质瘤, 视网膜母细胞瘤, 黑色素瘤, Wilm氏

瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。作为对比，靶细胞可以是被细胞内寄生物感染的细胞，细胞内寄生物可以是病毒。具体说，病毒可以是腺病毒，巨细胞病毒，EB病毒，单纯疱疹病毒I，单纯疱疹病毒II，人类疱疹病毒6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒B19，多瘤病毒BK，多瘤病毒JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人T细胞白血病毒I或人T细胞白血病毒II。任选地，细胞内寄生物可以是细菌，原生动物，真菌或朊病毒。更具体而言，细胞内病毒可以是衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌属，利什曼虫属，锥虫属，弓形体和疟原虫。

含有核酸构建体的疫苗抗原可以是MelanA (MART-I), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE- 1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40, PRAME, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, EB病毒抗原, EBNA, 人乳头状瘤病毒(HPV)抗原E6和E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, pl85erbB2, pl80erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连环蛋白, CDK4, Mum-1,和p16。任选地，抗原可以是与病毒或病毒感染相关的抗原。在另一种实施方案中，抗原是与非病毒细胞内寄生物相关的抗原。

该看家表位优选编码一条长度为6-23个氨基酸的多肽。更优选地，看家表位编码一条长度为9-10个氨基酸的多肽。有利地，看家表位对至少一个MHC等位基因特异。该等位基因可以编码类型A1, A2, A3, A11, A24, A26, A29, B7, B8, B14, B18, B27, B35, B44, B62, B60, 或B51。

在本发明的另一方面中，疫苗含有免疫表位。该免疫表位可以衍生于与靶细胞相关的第二抗原。第一抗原与第二抗原可以相同也可以不同。优选地，看家表位对MHC第一等位基因特异，免疫表位对MHC的第二等位基因特异。第一等位基因与第二等位基因可以相同也可以不同。

在本发明的另一方面中，含有核酸构建体的疫苗还额外地包括表位簇。表位簇含有免疫表位。优选地，表位簇衍生于与靶细胞相关的第二

抗原。第一抗原和第二抗原可以相同也可以不同。

有利地，表位簇包括或编码一条长度为至少10个氨基酸，但少于60个氨基酸的多肽。在优选实施方案中，表位簇包括或者编码长度小于第二抗原长度大约80%的多肽。在另一个优选实施方案中，表位簇包括或者编码长度小于第二抗原长度大约50%的多肽。在特别优选实施方案中，表位簇包括或者编码长度小于第二抗原长度大约20%的多肽。

在本发明的另一方面中，含有核酸构建体的疫苗还包括第二看家表位，其中的第二看家表位衍生于与第二靶细胞相关的第二抗原。第一抗原和第二抗原可以相同也可以不同。优选地，第一靶细胞和第二靶细胞不同。

核酸构建体

本发明提供了包括第一编码区的核酸构建体，其中的第一编码区包括编码至少第一多肽的第一序列，其中的第一多肽包括衍生于与第一靶细胞相关的第一抗原的第一看家表位。第一编码区还可以包括编码至少第二多肽的第二序列，其中的第二多肽包括衍生于与第二靶细胞相关的第二抗原的第二表位。第一多肽和第二多肽可以是邻接的或者不邻接的。第二表位可以是看家表位或者免疫表位。第一抗原和第二抗原可以相同也可以不同。同样，第一靶细胞和第二靶细胞可以相同也可以不同。

靶细胞可以是瘤细胞，例如白血病，癌瘤，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm氏瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，肾细胞癌，和脑癌。第一抗原可以是，例如，MART-I/MelanA, gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE-1, MAGE-3, BAGE, GAGE-1, GAGE-2, p15, NY-ESO, SSX基因家族成员的产物, CT-7, 和SCP基因家族成员的产物。靶细胞可以被下列病毒感染，例如，腺病毒，巨细胞病毒，EB病毒，单纯疱疹病毒1和2，人类疱疹病毒6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒B19，多瘤病毒BK，多瘤病毒JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人T细胞白血病病毒I或人T细胞白血病病毒II。

靶细胞同样可以被细菌，原生动物，真菌，朊病毒和任何其它细胞内寄生物感染，其实例包括衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌属，利什曼虫属，锥虫属，弓形体和疟原虫。

- 5 核酸构建体典型地包括与第一编码区有效连接的第一启动子序列。启动子可以是，例如巨细胞病毒（CMV），SV40和逆病毒长末端重复序列（LTR）。启动子可以是双向启动子，和/或与第二编码区有效连接的第二启动子序列。核酸构建体还可以包括与第一编码区、第二编码区或两者有效连接的聚合A序列。核酸构建体也可以包括内部核糖体进入位点（IRES）序列，泛素序列，自动催化肽序列，增强子，核进入序列，免疫刺激序列，和细胞因子、选择标记、报告分子等的表达盒，等等。第一多肽的长度可以是7-15个氨基酸，优选9或10个氨基酸。第二多肽的长度可以是9或10个氨基酸，或者可以是长度为大约10个和大约75个氨基酸之间的表位簇。第一表位和第二表位可以结合到相同或不同的MHC等位基因上。
- 10
- 15

本发明的其它实施方案包括含有任何上述核酸构建体的疫苗；通过给予此疫苗来治疗动物的方法；和制造该疫苗的方法。

表位簇的鉴定

- 20 本发明的其它实施方案涉及表位簇区域的鉴别，这些表位簇区域用来产生能够诱导受试者体内发生免疫反应的药物组合物，其中的受试者被给予了这种组合物。本发明的一种实施方案涉及衍生于与靶相关的表位簇，该表位簇含有或编码至少两个序列，这两个序列对MHC受体肽结合裂隙有已知的或者预测的亲和力，其中的表位簇为抗原片段。
- 25 在本发明的一方面中，靶是瘤细胞。或者，靶也可以是被细胞内寄生物感染的细胞。细胞内寄生物可以是病毒，细菌或者原生动物。任选地，靶是一种病原体。病原体可以包括病毒，细菌，真菌，原生动物，朊病毒，毒素或毒液。

- 在本发明的另一方面中，MHC受体可以是HLA I类受体。类似地，
- 30 MHC受体可以是HLA II类受体。

在本发明的另一方面中，这种簇包括或编码具有一定长度的多肽，其中该长度最少是10个氨基酸。有利地，多肽的长度可以少于大约75个氨基酸。

在本发明的另一方面中，提供了具有一定长度的抗原，其中的簇由一定长度的多肽组成或者编码一定长度的多肽，其中多肽的长度小于抗原长度的大约80%。优选地，多肽长度小于抗原长度的大约50%。特别优选地，多肽长度小于抗原长度的大约20%。

本发明的另一实施方案涉及表位簇的鉴别方法，方法包含以下步骤：提供与靶细胞相关的抗原序列；基于对MHC受体肽结合裂隙的已知的或者预测的亲和力，给序列中的候选肽评分，以鉴定假定的MHC表位；鉴定抗原中的一个区域，其中该区域包括至少两个假定MHC表位，而且该区域中所含有的假定MHC表位的密度要比整体抗原中假定MHC表位的密度要大。

治疗方法

本发明类似地设计了一种治疗动物的方法，方法包括给予动物含有第一看家表位的疫苗，其中的看家表位衍生于与第一靶细胞相关的第一抗原。优选地，给予步骤包括输送方法，即：经皮肤的，经结节内（intranodal），经结节周（perinodal），经口的，静脉内的，真皮内的，肌肉内的，腹膜内的，或者粘膜的。

治疗动物的方法还可以额外地包括一个测定步骤，该测定步骤用来确定表示靶细胞状态的特征。有利地，测定步骤还可以包括第一测定步骤和第二测定步骤，其中第一测定步骤先于给予步骤，第二测定步骤在给予步骤之后。优选地，将在第一测定步骤中确定的特征与在第二测定步骤中确定的特征进行比较以得出结果。结果可以是靶细胞数目减少，含有靶细胞的肿瘤的质量和大小减小，或者感染靶细胞的细胞内寄生物的数量或浓度的降低。

优选地，靶细胞是瘤细胞，瘤细胞可以是任何与实体瘤或淋巴瘤相关的转化细胞，例如白血病，癌瘤，淋巴瘤，星形细胞瘤，肉瘤，胶质瘤，视网膜母细胞瘤，黑色素瘤，Wilm氏瘤，膀胱癌，乳癌，结肠癌，肝细胞癌，胰腺癌，前列腺癌，肺癌，肝癌，胃癌，宫颈癌，睾丸癌，

肾细胞癌，和脑癌。或者，靶细胞可以被细胞内寄生物感染。例如细胞内寄生物可以是病毒如腺病毒，巨细胞病毒，EB病毒，单纯疱疹病毒1，单纯疱疹病毒2，人类疱疹病毒6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒B19，多瘤病毒BK，多瘤病毒JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人T细胞白血病毒I或人T细胞白血病毒II。细胞内寄生物可以是细菌，原生动物，真菌或朊病毒。有利地，细胞内病毒是衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌属，利什曼虫属，锥虫属，弓形体和疟原虫。

本发明的另一方面中，抗原是MelanA (MART-I), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE- 1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40, PRAME, p53, H-Ras, HER-2/neu, BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, EB病毒抗原, EBNA, 人乳头状瘤病毒(HPV)抗原E6和E7, TSP-180, MAGE-4, MAGE-5, MAGE-6, pl85erbB2, pl80erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连环蛋白, CDK4, Mum-1, p16。或者，抗原与病毒或病毒感染相关。在另一种实施方案中，抗原是与非病毒细胞内寄生物相关的抗原。

看家表位可以含有或编码长度为6-23个氨基酸的多肽。优选地，多肽的长度为9-10个氨基酸。该多肽可以是合成多肽。疫苗可以额外地包括缓冲剂，去垢剂，表面活性剂，抗氧化剂或还原剂。有利地，看家表位可以包括核酸。优选地，看家表位对至少一个MHC等位基因特异。该等位基因可以编码类型A1, A2, A3, A11, A24, A26, A29, B7, B8, B14, B18, B27, B35, B44, B62, B60, 或B51。

在本发明的另一个方面中，治疗动物的方法还包括免疫表位。该免疫表位可以衍生于与靶细胞相关的第二抗原。任选地，第一抗原和第二抗原相同。看家表位可以对MHC的第一等位基因特异，免疫表位可以对MHC的第二等位基因特异。第一等位基因与第二等位基因可以相同也可以不同。

有利地，疫苗包括含有免疫表位的表位簇。表位簇可以衍生于与靶

细胞相关的第二抗原。任选地，第一抗原和第二抗原相同。表位簇可以包含或者编码长度为至少10个氨基酸并且小于大约60个氨基酸的多肽。

优选地，表位簇包括或者编码长度少于第二抗原长度大约80%的多肽。多肽的长度可以少于第二抗原长度的大约50%。在另一方面，多肽
5 的长度可以少于第二抗原长度的大约20%。

治疗动物的方法还可以包括第二看家表位，其中的第二看家表位衍生于与第二靶细胞相关的第二抗原。第一抗原和第二抗原可以相同也可以不同。类似地，第一靶细胞与第二靶细胞可以相同也可以不同。

本发明同样设计了一种治疗动物的方法，方法包括向动物给予含有
10 核酸构建体的疫苗。有利地，该核酸构建体编码一种看家表位。其中的看家表位可以衍生于与第一靶细胞相关的第一抗原。

在本发明的另一方面中，提供了一种制造疫苗的方法，该方法包括：
通过鉴定表位（所述的表位是通过看家蛋白酶，从特殊抗原源制造或者可能制造出来的）来选择看家表位（所述的看家表位衍生于与第一靶细
15 胞相关的第一抗原）；制造含有看家表位的疫苗；和制备含有或编码所选择的看家表位的疫苗组合物。

本发明同样提供了依照前述方法制造的疫苗。该疫苗可以被给予用以治疗动物。因此，本发明类似地设计了用疫苗治疗动物的方法。

20 看家表位和其它表位的发现

本发明的其它实施方案涉及对表位的鉴别（这些表位对产生能够诱导受试者体内免疫反应的疫苗有用，其中的受试者被给予了这种组合物），尤其是对那些在发明实施方案中最有用的表位的鉴别。本发明的一种实施方案涉及发现表位的方法，该方法包括下列步骤：从大量与靶
25 细胞相关的抗原多肽片段中，选择一种表位，其中的片段对主要组织相容性复合物I类受体的肽结合裂隙有已知的或者预测的亲和力，其中被选择的表位与靶细胞的蛋白酶体切割产物相对应。

本发明的另一种实施方案涉及发现表位的方法，该方法包括：从靶细胞中提供一个序列，其中的序列编码或者包含靶细胞中表达的蛋白；
30 鉴别蛋白的肽段群，其中肽段群中的成员具有对主要组织相容性复合物I

类受体的肽结合裂隙有已知的或者预测的亲和力；从肽段群中选择表位，其中的表位与靶细胞中活性蛋白酶体的产物相对应。

该实施方案的一个方面涉及用前述的方法发现的表位。该实施方案的另一方面涉及含有被发现表位的疫苗。本发明的另一方面涉及治疗动物的方法，包括给予动物上述的疫苗。

本发明的一种实施方案涉及发现表位的方法，该方法包括的步骤有：提供瘤细胞和一段序列，其中的序列含有或者编码与瘤细胞相关的抗原；鉴别抗原的肽段群，其中的肽段群被预测对主要组织相容性复合物I类受体的肽结合裂隙有亲和力；从肽段群中选择表位，其中的表位通过体外分析被确定为：是瘤细胞中活性蛋白酶体的蛋白酶体切割反应产物。

本实施方案的一个方面涉及用前述方法发现的表位。本实施方案的另一方面涉及含有被发现表位的疫苗。本发明的另一方面涉及治疗动物的方法，包括给予动物上述的疫苗。

本发明的另一实施方案涉及发现表位的方法，该方法包括的步骤是：从与宿主中的靶相关的抗原的肽段群中，选择表位，其中的片段对宿主主要组织相容性复合物I类受体或II类受体的肽结合裂隙具有已知的或者预测的亲合性，其中被选择的表位与宿主细胞中抗原的蛋白水解切割产物相对应。

本实施方案的一个方面涉及用前述的方法发现的表位。本实施方案的另一方面涉及含有被发现表位的疫苗。本发明的另一方面涉及治疗动物的方法，包括给予动物上述的疫苗。

附图简述

图1图解描述了涉及通过蛋白酶体来进行蛋白加工和表位呈递的细胞部分。

图2是看家蛋白酶体与免疫蛋白酶体的比较。

图3图解描述了被感染的细胞与pAPCs之间的表位同步。

图4显示了pAPCs和肿瘤细胞对不同表位的呈递。

图5显示了pAPCs和被感染细胞对不同表位的呈递。

图6描述了由于IFN- γ 的诱导，肿瘤细胞对看家和免疫表位的呈递。

图7显示了诱导识别看家表位的T细胞对病毒感染细胞的攻击。

图8显示了同时针对看家和免疫表位的双重攻击。

图9描述了质粒pVAX-EP1-IRES-EP2-ISS-NIS的组分。

图10描述了质粒pVAX-EP2-UB-EP1的组分。

5 图11描述了质粒pVAX-EP2-2A-EP1的组分。

图12描述了质粒pVAX-EP1-IRES-EP2的组分。

图13描述了流式细胞仪分析的结果，证实了Melan-A表位结合了HLA。

10 图14描述了流式细胞仪分析的结果，证实了酪氨酸酶肽207-216结合了HLA。

图15描述了Melan-A (SEQ ID NO: 1)的序列，显示I类HLA表位的群集。

图16描述了SSX-2 (SEQ ID NO: 2)的序列，显示I类HLA表位的群集。

15 图17描述了NY-ESO(SEQ ID NO: 3)的序列，显示I类HLA表位的群集。

图18描述了酪氨酸酶(SEQ ID NO: 4)的序列，显示了通过序列线以上的BIMAS-NIH/Parker算法和序列线以下的SYFPEITHI/Rammensee算法进行预测的I类HLA表位的群集。

20 发明详述

本发明的实施方案提供了用来对靶细胞发动有效免疫反应的表位，疫苗和治疗方法。本发明的主要基础是这样一个新颖的和意想不到的发现：许多靶细胞呈现的表位与专门抗原呈递细胞(pAPCs)呈现的表位不同。由于这种差别，pAPCs指导T细胞对在靶细胞上不被呈现的表位反应，
25 T细胞因此无法识别靶细胞。本发明的方法和药物可以使pAPCs呈递出现在靶细胞上的相同表位，使得T细胞能够正确识别并破坏靶细胞。

本发明的实施方案还提供了鉴定靶抗原的表位的方法，这些表位可以用来产生免疫上有效的疫苗。这些疫苗可以刺激免疫系统识别并破坏呈现被选择表位的靶细胞。本发明的实施方案对于治疗和预防癌症和细胞内寄生物的细胞感染特别有用，而且对治疗和预防与其它病原体，毒
30

素和致敏原相关的病症特别有用。

某些类型的靶特别容易逃避免疫系统。包括许多种类的癌症和被细胞内寄生物感染的细胞，例如病毒，细菌和原生动物。研究者曾经作了大量的工作来鉴定对产生有效的免疫反应有用的抗原和表位，但是收效不大。本发明提供了有效发现能够有效对抗这种靶逃避现象的新一代有效表位的基础。

本发明使选择真正生物学相关的表位序列成为可能。具有生物学意义的表位（例如：在刺激免疫反应中发挥作用），必须对主要组织相容性复合物（MHC）受体肽的结合裂隙有亲和性。在本领域中，众所周知，有多种预测一个寡肽序列是否会具有MHC结合亲和性的方法。然而，预测到具有MHC结合亲和性的大多数序列没有生物学相关性，因为它们实际上没有在靶细胞或pAPC表面上呈现。

本发明的方法可以使疫苗设计者忽略那些尽管被预测具有MHC高结合亲和性，但由于不能被靶细胞呈递而根本没用的肽。因此，本发明公开的方法和教导，为疫苗设计提供了一个较大的进步，这种进步将抗原序列分析的能力与免疫学基本事实相结合。此处讲授的方法使得通过使用任何与需要的靶相应的多肽序列，简单而有效的选择对制造MHC I类和II类疫苗有意义的表位成为可能。

本发明的另一实施方案提供了用于疫苗和疫苗设计和表位的发现的表位簇区域（ECRs）。具体地，本发明的实施方案涉及对表位簇的鉴别，该表位簇用于产生指向靶细胞群的免疫活性组合物，并且用于不连续的看家表位和免疫表位的发现。在许多情况下，无数假定的I类MHC表位可以存在于单个靶相关抗原（TAA）中。这种假定的表位通常被发现在簇（ECRs）中，MHC表面抗原以相对高的密度分布在亲本TAA的氨基酸序列的某些区域里。由于这些ECRs包括多个假定的表位，这些表位对于诱导免疫反应有着潜在的有用的生物活性，因此它们代表着一种非常好的原料，该原料被用于体内或体外分析以鉴定对疫苗设计特别有用的表位。而且，由于表位簇本身可以在细胞内被加工以产生活性MHC表位，所以这些表位簇可以在疫苗中被直接使用，表位簇中的一个或多个假定表位实际上被加工成活性MHC表位。

疫苗中ECRs的使用，为重组疫苗的制造提供了重要的技术进步，还在安全性上为现有的编码整个蛋白序列的核酸疫苗提供了重要的有利条件。重组疫苗通常依靠微生物发酵罐中整个蛋白的昂贵的和工艺复杂的产物。ECRs提供了使用化学合成多肽的可供选择的方法，极大地简化了开发和生产，并且消除了多种安全上的考虑。类似地，使用编码ECRs的核酸序列（通常，这是整个序列的相对短的区域）的能力，允许在开发和操作核酸疫苗中使用合成寡核苷酸的化学方法，而不是使用涉及整个基因序列的、更昂贵的，耗时的，可能困难的分子生物学步骤。

由于与编码整个蛋白（ECR从中被发现）的核酸序列比较，ECR是由相对短的核酸序列编码的，因而可以极大地提高核酸疫苗的安全性。在核酸疫苗领域的一个重要问题是这样的事实：疫苗的序列与被给予动物的序列之间的同源性程度，决定了疫苗序列整合到动物基因组中去的概率。核酸疫苗的主要安全问题是它们整合到基因组序列的潜能，这样可能会导致基因表达异常和肿瘤转化。食品与药品管理局忠告：核酸和重组疫苗应该含有尽可能少的与人类序列同源的序列。在传送肿瘤相关抗原的疫苗的情况下，疫苗不可避免地含有与那些编码患者肿瘤细胞中蛋白的核酸具有同源性的核酸序列。然而，最好限制这样一些序列的使用，使其含有最低限度的为诱导治疗性免疫反应所必需的表位的表达的序列。因此，ECRs的使用提供了双重好处：提供了最小同源区，同时又整合了具有潜在治疗价值的多个表位。

在本发明的一些方面中，提供了编码看家表位的核酸构建体。看家表位（下面将会有非常详细的描述）包括由周围细胞的活性蛋白酶体产生的肽片段。本发明的基础是这样一个发现：任何靶细胞相关抗原可以被分化加工成两组不同的表位，用来被机体的I型主要组织相容性复合物（MHC）分子所呈递。“免疫表位”被pAPCs呈递，通常也被提呈在这样的周围细胞中，这些周围细胞被急性感染或者被干扰素（IFN）分泌细胞活性免疫攻击。相比之下，“看家表位”被所有的周围细胞（通常包括瘤（癌）细胞和慢性感染细胞）呈递。尽管宿主中存在机能免疫系统，但是这种被呈递表位的不匹配或不同步性，成为癌症和慢性感染的持续性和进展性的基础。因此，在pAPCs和靶细胞中产生表位呈递的

同步性，以引发有效的、细胞毒性T细胞介导的免疫反应，是非常重要的。

能够获得同步性的最可靠手段就是给pAPCs提供看家表位。通常，通过提供一个以上的单一表位可以引发更强的反应。此外，一旦建立了对靶细胞有效地免疫反应，IFN的分泌会引起免疫蛋白酶体的表达，因此表位呈递被转换成了免疫表位。因为这个原因，在根据上面公开内容而研制的疫苗中，除了看家表位，最好也包括免疫表位。以ECR的形式提供免疫表位更加有利。本发明的实施方案提供了以多种组合编码看家表位和/或免疫表位的表达载体。优选的表达构建体编码至少一种能够刺激对靶细胞发生免疫反应的表位。在本发明的一种实施方案中，靶细胞是瘤细胞。在另一种实施方案中，靶细胞是任何细胞内被感染的宿主细胞。细胞内感染物包括持续性病毒和任何其它具有细胞内传染期的传染性有机体。

一些实施方案的核酸构建体涉及增强受试者的免疫系统并且使之对宿主内存在的瘤细胞敏感化。另一些实施方案中，核酸构建体促进了对持续性病毒的感染和感染了细胞内寄生物的细胞的根除。

定义

除非在其它情况下，从文章的前后关系中清楚本文术语的用法，下列的术语在本说明书中通常具有下文描述的意思。

专门抗原呈递细胞（pAPC）——一种具有T细胞共刺激分子，而且能够诱导T细胞反应的细胞。被清楚表征的pAPCs是树突状细胞，B细胞，巨噬细胞。

周围细胞——不是pAPCs的细胞。

看家蛋白酶体——通常在周围细胞中有活性，而且通常在pAPCs中不出现或者没有很强活性的蛋白酶体。

免疫蛋白酶体——通常在pAPCs中有活性的蛋白酶体，免疫蛋白酶体在被感染组织中的一些周围细胞中也有活性。

表位——能够刺激免疫反应的分子或物质。在优选实施方案中，根据此处定义的表位包括但不必须限于多肽和编码多肽的核酸，其中的多肽能够刺激免疫发应。在其它的优选实施方案中，根据此处定义的表位包

括但不必须限于呈现于细胞表面、非共价结合于I类MHC袋的肽，这样它们可以与T细胞受体相互作用。

MHC表位—对哺乳动物I类主要组织相容性复合物（MHC）分子具有已知的或者预测的亲和多肽。

5 ELA表位—对人类动物I类主要组织相容性复合物（MHC）分子具有已知的或者可预测的亲和多肽。

看家表位—在优选的实施方案中，看家表位被定义为多肽片段，该片段是MHC表位而且该片段呈现在这样一种细胞上，在这种细胞中看家蛋白酶体为主要活性。在另一个优选实施方案中，看家表位被定义为含有依照上述定义的看家表位的多肽，该多肽被一个到几个附加的氨基酸侧接。在另一个优选实施方案中，看家表位被定义为编码依照上述任何一个定义的看家表位的核酸。

15 免疫表位—在优选的实施方案中，免疫表位被定义为多肽片段，该片段是MHC表位而且该片段呈现在这样一种细胞上，在这种细胞中免疫蛋白酶体为主要活性。在另一个优选实施方案中，免疫表位被定义为含有依照上述定义的免疫表位的多肽，该多肽被一个到几个附加的氨基酸侧接。在另一个优选实施方案中，免疫表位被定义为包含表位簇序列的多肽，该多肽含有至少两个对I类MHC具有已知的或者预测的亲和多肽序列。在另一个优选实施方案中，免疫表位被定义为编码依照上述任

20 何一个定义的免疫表位的核酸。

靶细胞—被本发明的疫苗和方法所靶向的细胞。依照此定义的靶细胞的实例包括但不必要限于：瘤细胞和寄居着细胞内寄生物的细胞，例如，病毒，细菌和原生动物。

靶相关抗原（TAA）—靶细胞中出现的蛋白质或多肽。

25 肿瘤相关抗原（TuAA）—一种TAA，其中的靶细胞是瘤细胞。

要注意下面的讨论阐明了发明者对发明操作的理解。但是并不能说该讨论将本专利限制于权利要求中没有阐明的任何特殊操作理论。

不同的蛋白酶体产生不同的表位

30 被I类MHC在pAPCs 或周围细胞表面所呈现的表位，是由蛋白酶体

通过消化这些细胞中的蛋白而产生的。虽然有报道称pAPCs与周围细胞中的蛋白酶体并不是相同的，但是这种不同的显著性迄今并不被人认同。本发明基于这样一个事实：当pAPCs和周围细胞加工特定的TAA时，在pAPCs中有活性的蛋白酶体所产生的表位片段，与在周围细胞中有活性的蛋白酶体所产生的表位片段不同。为了便于参考，而且依照上述的定义，在pAPCs中主要活性的蛋白酶体在此称为“免疫蛋白酶体”，而在周围细胞中有正常活性的蛋白酶体在此称为“看家蛋白酶体”。

不能过分夸大pAPCs和周围细胞对TAAs差别加工的意义。这种差别加工为“为什么某些靶细胞对免疫系统的识别和攻击有抗性”提供了统一标准的解释。尽管pAPCs能够摄取从靶细胞中流出的TAAs，并且将它们呈递到其表面，但是pAPCs将因此刺激CTLs的产生来识别“免疫表位”（免疫蛋白酶体加工TAA而得到的表位），而靶细胞呈现“看家表位”（通过看家蛋白酶体产生的明显不同的TTA片段）。结果，在生理条件下的CTL反应被误导，远离了靶细胞上的表位。

由于CTL反应是由pAPCs诱导的，按照定义，它们靶向免疫表位而不是看家表位，因此不能识别靶细胞，这样靶细胞可以在体内存留。这种基本的细胞免疫反应的“表位区域化”，就是一些肿瘤细胞能够持续形成肿瘤的原因，也是一些病毒和细胞内寄生物能够慢性感染细胞却没有被免疫系统根除的原因。关于传染物，通常它们会导致它们所感染细胞的免疫蛋白酶体表达。这会导致在细胞表面产生与那些pAPCs呈递给免疫系统相同的表位。这样，传染会导致在免疫系统和被感染细胞之间产生“表位同步性”，随后感染细胞被破坏，并且从体内清除出传染物。在一些传染物中，特别是那些能够建立慢性传染的传染物中，它们在其感染的细胞中演化出了一种防止免疫蛋白酶体表达的方法。这些细胞中的蛋白酶体保持为看家形式，从而防止表位的同步性和防止CTL的攻击。有重要的证据表明：实际上所有慢性传染物都使用一种共同的机制。

克服CTLs无法识别并根除某些靶细胞的一种方法是：提供能够“同步”表位呈递的疫苗和治疗方法。本文中的表位同步指的是使pAPCs呈递看家表位，引起CTLs识别在靶细胞上呈现的看家表位，从而攻击并消除靶细胞。

因此，本发明的实施方案对于处理包括实体瘤和淋巴瘤在内的肿瘤疾病有用。本发明其它的实施方案可用于治疗持续性病毒感染和寄生虫感染，其中的传染物有细胞内传染期。适当地施用与这种靶细胞相应的看家表位，可以激活特定的、针对靶细胞的细胞毒性T细胞反应。

5

表位差别在癌症中的作用

在一些实施方案中，本发明涉及治疗肿瘤性疾病的方法。癌症是由单个异常细胞后代的进行性的，失控的生长而导致的。此处用到的术语“癌症”包括肿瘤性疾病，瘤细胞，肿瘤，肿瘤细胞，恶性肿瘤和任何转化的细胞，同时包括实体瘤和弥散肿瘤疾病。在历史上，癌细胞曾经通常被认为逃避了免疫系统的检测和破坏，因为癌细胞含有与其它体内非肿瘤细胞相同的遗传物质。体内癌细胞与健康细胞的遗传一致性或相似性，被认为在区别癌细胞和肿瘤细胞方面引起了困难，因此免疫系统无法发动有效地免疫反应，正如体内癌细胞的持续性存留所显明的。

相反，已经有多种肿瘤相关抗原（TuAAs）被描述并且的确引发了免疫反应。无数的研究已经描述肿瘤浸润淋巴细胞(TILs)可以在体外杀死呈递衍生于多种TuAAs的肽的靶细胞。然而，就像下面的进一步详细讨论中所描述的那样，由于表位的差别，TILs无法控制癌症，所述的表位是由下面的细胞产生和呈递的：诱导CTL活性的细胞，pAPC，和期望的靶细胞，即，肿瘤细胞。为了理解这种差别，必须理解蛋白酶体的功能和动力学。

所有细胞含有降解蛋白的蛋白酶体。这些占大约1%细胞总蛋白含量的蛋白酶体，被用来调控细胞中蛋白的半衰期。在蛋白降解的过程中，蛋白酶体产生绝大部分涉及I类抗原呈递的肽段，而且蛋白酶体的切割方式影响了被呈递在I类分子上的抗原表位的有效性（图1）。这样，细胞的蛋白酶体活性产生了MHC表位。然而，蛋白水解酶的活性在pAPCs中与在周围细胞中相比，是明显不同的。作为细胞免疫反应的一部分，pAPCs含有这样的蛋白酶体，该蛋白酶体在感染时或者在接触了不同细胞因子，特别是干扰素（IFN）后，组成性地结合那些典型情况下只在周围细胞中表达的亚单位。如上所述，pAPCs和周围细胞的不同蛋白酶体活性在此

分别称为免疫和看家蛋白酶体。

免疫和看家蛋白酶体具有在类似但不同的位置切割蛋白的能力。免疫蛋白酶体结合几种使其与其看家蛋白酶体相区别亚单位。这些亚单位包括LMP2, LMP7, 和 MECL1, 这些亚单位替代了看家蛋白酶体的催化亚单位, 和PA28 α , PA28 β , 这些亚单位提供调控功能(图2)。总的来说, 这些亚单位的结合引起免疫蛋白酶体与看家蛋白酶体的活性在数量上和质量上的差别。尽管存在证据证明: 看家和免疫蛋白酶体产生的MHC表位, 在看家和免疫蛋白酶体之间存在着明显的不同, 但是迄今这些差别仅在数量上进行了阐明。其它人认为: 免疫蛋白酶体介导的最终效应是促进产生更多的肽, 而不是不同的肽。

在免疫和看家蛋白酶体之间的抗原加工的本质上的区别, 对疫苗设计有重要意义。IFN- γ 是由T淋巴细胞产生的, 它涉及促进细胞免疫反应的诱导并且诱导免疫蛋白酶体表达(如上所述)。特别是, IFN还可以在细胞被病原体感染的情况下, 由任何其它细胞产生。在自然状态下, 病毒感染会典型地导致被感染细胞产生IFN, IFN随后诱导细胞从看家蛋白酶体构型转变为免疫蛋白酶体构型。一种对此现象的解释为: 根据被呈递的抗原全部组分, 用感染和随后的IFN正调节来调整感染细胞, 而pAPC呈递的抗原全部组分参与刺激针对病毒的免疫反应。这样使被感染细胞以与pAPCs中发生的抗原加工相同的方式, 同时加工其内源性“自我”蛋白(由该细胞正常表达)和涉及感染物的那些蛋白(非我)。将被感染细胞的蛋白酶体从看家构型转变为免疫构型, 导致被感染细胞和pAPC细胞之间发生“表位同步”(图3)。

对TuAAs特异的MHC I类限制性CTLs是针对癌症免疫反应的重要组成部分。TuAAs是肿瘤特异T细胞反应的有用的靶子, 它们不在正常细胞表面表达, 或者它们被肿瘤细胞过表达, 或者它们是肿瘤细胞的重要特征。对于本领域的技术人员, 了解并且可以很容易地通过文献获得或者商购到多种TuAAs。

在一些肿瘤中, 确实已经发现它们在免疫蛋白酶体的IFN- γ 诱导中存在缺陷。在这些情况中, CTL有可能靶向那些来自于被pAPC加工的TuAAs的免疫表位。尽管这些患者中有高数量的CTL对该免疫表位特别有活性,

但是CTL却不能发现癌细胞上的表位。疾病发展，到最后积聚的CTL无法定位靶细胞，出现功能障碍(Lee, 等 *Nature Medicine* (1999) 5[6]:677-685)。通过给pAPC提供看家表位，可以使pAPCs呈递的表位和肿瘤呈递的表位同步化，并且可以活化CTL群，这些CTL群识别那些呈递在肿瘤上的看家表位。

因此，pAPCs中的免疫蛋白酶体和周围细胞中的看家蛋白酶体产生性质上不同的表位，这一发现为为什么TILs不能根除肿瘤细胞提供了解释。上述的加工机制解释了T淋巴细胞是如何找到它们进入瘤块的路径的和为什么它们对肿瘤细胞相对不起作用。pAPCs的免疫蛋白酶体和肿瘤细胞的看家蛋白酶体之间的不同抗原加工，可以解释在患有进行性癌症的患者体内，能够观察到特异针对TuAAs的高频度出现的T淋巴细胞。Lee, 等 *Nature Medicine* (1999) 5[6]:677-685. (图 4)。

由于蛋白酶体活性的差别，包括肿瘤细胞和一些被病毒或其它细胞内寄生物感染细胞在内的周围靶细胞（所有这些细胞表达看家蛋白酶体），必须呈现与通过pAPCs调节T细胞而识别的表位信号不同的表位信号。鉴于本发现，出现了令人感兴趣的蛋白酶体免疫调节作用。此发现为免疫系统，特别是为pAPCs的操作提供了方法，该方法的目的是：诱导由细胞介导的、对靶细胞发动有效地和致死地攻击。

抗原加工的差异，解释了为什么对TuAAs特异的CTLs常常会在TILs中被发现却没有根除疾病。T淋巴细胞反应主要是针对已经被pAPCs加工的TuAAs。在TILs中发现的CTLs拼命地靶向那些呈现在pAPCs上的I型TuAAs，但却不能靶向肿瘤细胞（图4）。

肿瘤细胞在体内的行为，也就是迁移、抗原流出、和炎症反应诱导等，会引起强烈的免疫反应。不幸的是，肿瘤细胞和pAPCs之间不同抗原加工的自然机制，导致肿瘤表位隔离——即：肿瘤具有与pAPCs不同的加工TuAA的表位标记，因此肿瘤抗原从这样的表位中被“隔离”，所述的表位被pAPCs诱导的CTLs所识别。预测和对抗此表位隔离效应的能力，对于形成新一代的治疗性癌疫苗非常重要。克服表位隔离就会使表位同步。

表位差异在病毒感染和其它细胞内寄生物感染中的作用

为了在宿主有机体体内建立慢性感染，持续性病原体使用了很多机制。一个常见的特征就是降低或改变抗原表达。在一些实施方案中，本发明涉及治疗和预防多种病原体引起的细胞内感染。这些病原体的实例
5 包括，但不仅仅限于：任何病毒，细菌，原生动物，朊病毒或其它在宿主体内具有细胞内感染期的有机体。

pAPCs呈递的病毒抗原起始于使用蛋白酶体将病毒抗原消化成肽。在蛋白酶体将蛋白消化成肽以后，一些肽被负载到内质网中的I型复合物上并且被传送到细胞表面。在细胞表面上，CTLs表面的T细胞受体识别I
10 型一肽复合物，被感染的细胞被杀死。

疱疹病毒和反转录病毒通过限制病毒基因表达来逃避宿主免疫系统的监测和随后对其的消除。某些病毒还可以通过其它机制来逃避免疫系统，包括在重要病毒肽中的病毒感染和抗原变异的“免疫特权”位点。虽然这些模型可以解释某些病毒的存留，但是表位同步的概念或者与之
15 相反的表位区域化的概念，提供了一种解决办法。也就是，这个概念为疫苗提供了基础，该疫苗用来指导针对任何病毒或其它细胞内寄生物的有效细胞免疫反应，所述的病毒或其它细胞内寄生物通过阻断宿主细胞的免疫蛋白酶体表达或通过防止感染细胞和pAPCs之间的有效的表位同步来逃避免疫系统（图5）。

20 由于任何被病原体感染的细胞通常会导致感染细胞产生IFN，所以典型地，被感染组织内的蛋白酶体从看家构型转变为免疫构型。因而，根据呈现于被感染细胞表面的抗原全部组分，感染具有对感染细胞进行调整的作用，并且pAPC表面的抗原的全部组分参与刺激针对病毒或其它细胞内病原体的免疫反应。当病毒感染的细胞或寄生物感染的细胞被诱导
25 表达免疫蛋白酶体，而不是看家蛋白酶体时，结果就是被感染细胞和pAPCs之间“表位同步”，进而由CTL根除感染的细胞。

但是，某些病毒和其它细胞内寄生物可以下列方法来逃避T细胞识别：负调节那些对T细胞有效识别感染细胞所必需的宿主分子的表达。存在证据显示：许多慢性病毒感染干扰了IFN级联（见表1）。因此，由于
30 看家蛋白酶体、免疫蛋白酶体和表位区域化的作用，在许多慢性感染中，

本发明的一些实施方案同样适用于对任何相关细胞内寄生物的疫苗的设计，这些寄生物包括但不限于病毒，细菌和原生动物。所有细胞内寄生物都是这种疫苗设计的靶。这些细胞内寄生物包括但不限于：
5 病毒例如腺病毒，巨细胞病毒，EB病毒，单纯疱疹病毒1，单纯疱疹病毒2，人类疱疹病毒6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒B19，多瘤病毒BK，多瘤病毒JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人T细胞白血病病毒I或人T细胞白血病病毒II；细菌例如衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌
10 属；原生动物例如利什曼虫属，锥虫属，弓形体和疟原虫。

表1

病毒	病毒基因产物	IFN（免疫蛋白酶体）抑制机制
腺病毒	E1A 蛋白	阻断信号传导
	VA RNA	阻断PKR活化
牛痘病毒	CrmA 蛋白	丝氨酸蛋白酶抑制剂，蛋白抑制剂阻断I1-1 β 的活化
EB病毒	EBNA-2蛋白	阻断信号传导
	EBER RNA	阻断PKR活化
	BCRFL(VIRAL IL-10)	IL-10 同系物（抑制Ifn- γ ， I1-1， I1-2， Tnf合成）
乙型肝炎病毒	末端蛋白	阻断信号传导
人类免疫缺陷病毒	Tat蛋白	降低PKR活性
	TAR RNA	阻断PKR活性
单纯疱疹病毒	未知	阻断RnaseL活化
流感病毒	NSI	结合dsRNA， 阻断PKR活化
	未知	活化叫做p58的PKR细胞抑制剂
粘液瘤病毒	M-T7蛋白	可溶Ifn- γ 受体（德夸(decoy)）
	T2蛋白	可溶Tnf受体（德夸(decoy)）
脊髓灰质炎病毒	未知	活化PKR细胞抑制剂
呼肠孤病毒	Sigma3蛋白	结合dsRNA和阻断PKR活化
	K31蛋白	PKR假底物
	B15R蛋白	可溶I1-1受体（德夸(decoy)）
	A18R蛋白	调节dsRNA产物

获得表位同步的疫苗和方法

- 5 如上讨论，有效地细胞免疫基于pAPCs和被感染周围细胞之间的同步化表位呈递。在缺少表位同步的时候，即使T细胞被导向TAAs，靶细胞也不能被T细胞所识别。由于可以避免表位同步，癌细胞和寄居有持续

性细胞内寄生物的细胞逃避细胞内免疫反应。“自然的”表位同步包括在被感染细胞中的免疫蛋白酶体活化，这样被感染的细胞呈递免疫表位并且因此可以被经pAPCs诱导的T细胞识别。但是癌症和感染了持续性细胞内寄生物的细胞没有活化的免疫蛋白酶体，因此不能被正常的诱导的T细胞所识别。

因此，本发明优选实施方案的疫苗和方法提供了一种“倒转(reverse)”表位同步，造成pAPCs呈递看家表位来面对靶细胞不呈现免疫表位的情况(图6和7)。某些实施方案还通过诱导pAPCs同时呈现对应于被选择靶细胞的看家表位和免疫表位，提供了的第二波表位同步。这样，在这些双重表位实施方案中，一旦靶细胞被识别看家表位的T细胞有效攻击，靶细胞中的免疫蛋白酶体加工的转换，不会导致免疫识别的损失。这是因为疫苗中存在免疫表位，该免疫表位的作用是诱导识别免疫表位的T细胞群。

本发明的优选实施方案涉及引起pAPC和pAPCs群呈递看家表位的疫苗和方法，所述的看家表位与呈现在特殊靶细胞上的表位相符合。在一种实施方案中，表位是由特殊肿瘤类型的看家蛋白酶体加工的TuAA表位。在另一种实施方案中，看家表位是由病毒感染细胞的看家蛋白酶体加工的病毒相关表位。这促进了特异T细胞对靶细胞的反应。pAPCs的多个表位同时表达，对应于不同的诱导状态(攻击前和攻击后)，可以在靶细胞呈现看家或免疫表位的时候，引发CTL对靶细胞的有效反应。

通过在pAPC上同时呈现看家表位和免疫表位，该实施方案可以使细胞毒性T细胞对靶细胞的反应最优化。伴随着双重表位表达，当肿瘤细胞在IFN的诱导下，其看家蛋白酶体转换为免疫蛋白酶体的时候，pAPCs仍然可以连续维持对免疫型表位的CTL反应，IFN可以由例如肿瘤浸润CTLs产生。

在优选实施方案中，用包含看家表位的疫苗对患者进行免疫。许多优选的TAAs只与靶细胞相关，特别是被感染的细胞。在另一种实施方案中，许多优选的TAAs是在转化细胞中失控的基因表达的结果，但是也可以在睾丸、卵巢和胎儿组织中发现。在另一种实施方案中，有用的TAAs在靶细胞中比在其它细胞中以更高的水平表达。在另一种实施方案中，

TAA与其它细胞比较，在靶细胞中并不差异表达，但是由于它们涉及细胞的特殊功能而且可以将靶细胞从其它大多数周围细胞中区别出来，所以仍然是有用的；在这些实施方案中，同样呈现TAA的健康细胞可能也会附带地被诱导的T细胞反应所攻击，但是这样的附带伤害比靶细胞所引起伤害要小得多。

当瘤细胞是靶的时候，优选的抗原包括TuAAs。适合使用的蛋白抗原包括分化抗原例如MelanA (MART-I), gp 100 (Pmel 17), 酪氨酸酶, TRP-1, TRP-2, , 和肿瘤特异的多系抗原 (multilineage antigens) 例如MAGE-1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40, 和PRAME。类似地, TuAAs包括过表达的癌基因, 和突变的肿瘤抑制基因例如 p53, H-Ras, HER-2/neu。此外, 染色体易位产生的独特的TuAAs例如 BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, 和病毒抗原例如EB病毒抗原EBNA, 并且包括人乳头状瘤病毒(HPV) 抗原E6和E7。其它有用的蛋白抗原包括但是不仅仅限于 TSP-180, MAGE-4, MAGE-5, MAGE-6, pl85erbB2, pl80erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连环蛋白, CDK4, Mum-1, p16。对于本领域的技术人员, 这些和其它的TuAAs和病原相关抗原可以通过文献获得或者商购。

在另一的实施方案中, TAA是对病毒特异的抗原。见表2。在本发明的另一种实施方案中, TAA是对非病毒细胞内寄生物特异的抗原。寄生物特异的抗原的实例包括核酸, 蛋白, 或其它与细胞内寄生物相关的基因产物。适用的核酸和蛋白可以在NCBI分类数据库中找到, 网址是 <http://www.ncbi.nlm.nih.gov/Taxonomy/tax.html/>。此网址提供了对寄生物和其它病原体基因产物的更为详细的描述。

25

30

表 2

病毒	候选基因产物
单纯疱疹病毒	ICP4, VP16, ICPO, $\gamma 1^{34.5}$, g13
EBV	ZTA, EBNA-2, EBNA-1, LMP-1, LMP-2, LMP-2a, LMP-2b EBNA-3, EBNA-4, EBNA-LP, EBNA-3A, 3C, BZLF-1
痘病毒	VeTF, K3L, p37, A14L, A13L, A17L, A18R
SV40	大T抗原, 小T抗原, VPZ
腺病毒	E1A, E3L, E1B, E4 (OEF6), E4 (ORF1), gp19K, ADP, RID α , RID β
乙肝病毒	pX, L(pre-Si)
Htlv-1	Tax
HIV	TAT, GAG, MA, ENV, TM, NEF, VIF, VPR, REV, VPX
Hepatis B	NS5A
呼肠孤病毒	δ -3
劳氏肉瘤病毒	pp60 ^{src}
哈维肉瘤病毒	p21 ^{ras}
HPV	E6, E7, E5
多瘤病毒	LT, mT, sT

- 此处描述的化合物和方法在靶细胞呈递看家表位的任何范围内都有效。此处公开了发现有效表位的方法, 所述的有效表位被应用于本发明的疫苗和治疗实施方案。

抗原的蛋白水解加工

- 被MHC呈递在靶细胞上或pAPCs上的表位是大蛋白抗原前体的切割产物。对于MHC I 表位, 蛋白抗原被细胞内的蛋白酶体消化。见图1。细胞内蛋白酶消化典型地产生3-23个氨基酸长度的肽段。细胞中或细胞

外环境中的其它的蛋白水解活性，可以进一步修饰和加工这些片段。MHC II 表位的加工是通过来自溶酶体区室（compartment）/内体区室的细胞内蛋白酶发生的。

推测起来，多数由蛋白酶体或其它蛋白酶活性处理的蛋白加工产物，
5 对特殊MHC受体肽结合裂隙只有很少的或没有亲和性。可是，通过细胞表面的MHC，可能会出现一定丰度的、具有这种亲和性的加工产物。相反，如果特定的寡肽序列不能完整地来自细胞的抗原加工活性中出现，它就不能被呈递到细胞表面，这与此序列对MHC的预测亲和性无关。

完全集中在MHC亲和性上的疫苗设计是根本错误的。仅仅因为肽具有MHC结合亲和性这样一个事实，并不能保证这样的肽将会成为功能免疫原。为了提供能够诱导对TAA进行有效免疫反应的表位，肽必须具有MHC结合亲和性而且肽必须是细胞肽产生系统的产物。本发明的方法同时利用了MHC结合亲和性分析和抗原加工分析方法来鉴别感兴趣的新表位。

15

使预测的或已知的MHC结合与抗原蛋白水解加工相关联

为了鉴定作为免疫化合物的、潜在有效的表位，只是对MHC结合进行预测通常是不利的，因为许多预测结合的片段在细胞中实际上从来没有形成。本发明的实施方案将MHC结合分析与蛋白水解加工分析结合了起来，来鉴定同时具有MHC亲和性和正确蛋白水解加工两个重要性质的有用表位。同时具有这些特性的肽段是疫苗和免疫治疗的有力的候选肽段。缺乏其中任何一种性质的肽段都不太可能成为有效表位。

本发明的实施方案能够鉴定用于疫苗的、衍生于TAAs的表位。靶抗原可能衍生于瘤细胞，感染了病毒或其它细胞内寄生物的细胞，或者感染了其它致病物质如细菌，真菌，原生动物，病毒，朊病毒，毒素，毒液，过敏源等等的细胞。简而言之，本方法的实施方案可以被应用于实际上所有的蛋白序列，以鉴定其中的能够通过蛋白水解作用产生的并且能够结合于MHC的表位。因此，本发明不限制于任何特殊的靶或疾病，相反，本发明包括从任何有用的来源来发现生物学相关的MHC表位。

30 在优选实施方案中，TAA是瘤细胞的特征，而且因此被定义为肿瘤

相关抗原 (TuAA)。优选的TuAAs包括：分化抗原例如MelanA (MART-1), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE- 1, MAGE-3, BAGE, GAGE-1, GAGE-2, p15；通常过表达的胚胎抗原；过表达的癌基因和突变的肿瘤抑制基因例如p53, Ras, HER-2/neu；染色体易位产生的独特
5 TuAAs例如 BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR1, 和病毒抗原例如EB病毒抗原 (EBVA) 和人乳头状瘤病毒(HPV)抗原E6和E7。其它感兴趣的抗原包括前列腺特异抗原 (PSA), 前列腺干细胞抗原 (PSCA), MAAT- 1, GP- 100, TSP-180, MAGE-4, MAGE-5, MAGE-6, RAGE, pl85erbB2, pl80erbB-3, c-met, nm-23H1, TAG-72, CA19-9, CA72-
10 4, CAM 17.1, NuMa, K-ras, -连环蛋白, CDK4, Mum-1, p15, p16。也可以考虑其它靶抗原。

本领域中有许多众所周知的方法可以用来鉴定TuAAs, 这些技术的实例包括差异杂交, 包括使用微排列技术；减数杂交克隆；或者在mRNA水平或者在蛋白表达水平表达的差异显示；EST测序；SAGE (基因表达
15 序列分析)。这些核酸技术已经被Carulli, J.P.等在J. Cellular Biochem Suppl. 30/31:286-296, 1998.中综述。蛋白的差异显示包括, 例如, 对从正常组织和肿瘤组织中细胞溶解物的二维聚丙烯酰胺胶电泳的比较, 对肿瘤中独特的或过表达的蛋白斑点定位, 从胶中回收蛋白, 和使用传统的生化技术或者质谱分析测序技术鉴定蛋白。鉴定TuAAs的另外一种技术是SEREX
20 技术, 在Türeci, Ö., Sahin, U., 和 Pfreundschuh, M.,的“人类肿瘤抗原的血清学分析：分子定义及其意义”, Molecular Medicine Today, 3:342, 1997.中讨论。对这些方法和其他方法的使用, 给本领域的技术人员提供了这样的技术, 这些技术对于鉴定产生看家与免疫I类表位的有用抗原和用于疫苗的II类表位的有用抗原是必需的。但是, 在实施本发明时, 并不是必
25 须要鉴定新的TuAA或TAA。更确切地说, 本发明的实施方案使得从任何相关蛋白序列中鉴定出有用的表位成为可能, 不管这个序列是新的还是已知的。

对MHC结合的TAA片段的分析

30 为了鉴定生物学上相关的表位, TAA中对MHC具有已知的或预测的

亲和性的片段被鉴别出来。可以用许多不同的技术来分析TAA的氨基酸序列，这些技术用来鉴定对MHC肽结合裂隙具有已知的或预测的亲和性的肽段。在本发明的一种实施方案中，使用计算机算法来分析TAA片段，预测它们结合到MHC肽结合裂隙的能力。MHC的每一个等位基因都对一个特殊的表位结合结构域特异。因此，对于任何特定的MHC等位基因，
5 可以筛查候选肽段与它的预测的亲和性。用于此目的的合适的计算机算法的实例包括：Hans-Georg Rammensee, Jutta Bachmann, Niels Emmerich, Stefan Stevanovic的万维网页：SYFPEITHI: MHC配体和肽基元的因特网数据库 (http://access via: syfpeithi.bmi-heidelberg.com/scripts/MHCServer.dll/home.htm)。从此方法得到的结果在
10 Rammensee等的“MHC Ligands and Peptide Motifs,” Landes Bioscience Austin, TX, 224-227, 1997. 中讨论。另一个感兴趣的站点是“http://bimas.dcrt.nih.gov/molbio/hla_bind,”，此站点也包含合适的算法。该站点中的方法在Parker等的“Scheme for ranking potential HLA-A2
15 binding peptides based on independent binding of individual peptide side-chains,” J. Immunol. 152:163-175.中讨论。

使用NIH (Parker) 算法和本发明的方法来选择肽段将使用各种可能的驻留时间来显示结合序列。在一种实施方案中，将选择具有无限驻留时间的肽。在另一种实施方案中，将选择驻留时间为25分钟或更多时间的
20 肽段来显示结合序列。在另一种实施方案中，将选择驻留时间为15分钟或更多时间的肽段来显示结合序列。在另一种实施方案中，将选择驻留时间为10分钟或更多时间的肽段来显示结合序列。也可以考虑9, 8, 7, 6, 5, 4, 3, 2, 和1分钟的驻留时间。

作为预测算法的一种选择，一些标准的体外受体结合亲和性分析可以
25 用来鉴定对MHC特殊等位基因有亲和性的肽。因此，通过本发明的这种方法，可以将起初的肽片段群范围缩小到只包括那些对所选择的MHC等位基因有实际的或预测的亲和性的肽。

起初，用于分析的候选肽可以包括来自TAA全部蛋白序列的大约6—24个邻接氨基酸的每一种可能序列。在优选实施方案中，序列的长度可
30 以是7—20个氨基酸。在更优选方案种，序列的长度可以是8—15个氨基

酸。对于通过预测对MHC I的亲性和来鉴定片段的序列分析方法，最优选的实施方案是分析TAA的9或10个邻接氨基酸片段的所有可能序列。片段的MHC亲和性分析，可以通过体外进行或通过片段计算机分析来进行。

- 5 在下面的表3—5中，报告了所选择的常见MHC I等位基因，及其近似出现频度。

表3.HLA—A抗原的估计基因出现频度

10

15

20

25

抗原	CAU		AFR		ASI		LAT		NAT	
	Gf ^a	SE ^b	Gf	SE	Gf	SE	Gf	SE	Gf	SE
A1	15.1843	0.0489	5.7256	0.0771	4.4818	0.0846	7.4007	0.0978	12.0316	0.2533
A2	28.6535	0.0619	18.8849	0.1317	24.6352	0.1794	28.1198	0.1700	29.3408	0.3585
A3	13.3890	0.0463	8.4406	0.0925	2.6454	0.0655	8.0789	0.1019	11.0293	0.2437
A28	4.4652	0.0280	9.9269	0.0997	1.7657	0.0537	8.9446	0.1067	5.3856	0.1750
A36	0.0221	0.0020	1.8836	0.0448	0.0148	0.0049	0.1584	0.0148	0.1545	0.0303
A23	1.8287	0.0181	10.2086	0.1010	0.3256	0.0231	2.9269	0.0628	1.9903	0.1080
A24	9.3251	0.0395	2.9668	0.0560	22.0391	0.1722	13.2610	0.1271	12.6613	0.2590
A9 未分裂	0.0809	0.0038	0.0367	0.0063	0.0858	0.0119	0.0537	0.0086	0.0356	0.0145
A9 总共	11.2347	0.0429	13.2121	0.1128	22.4505	0.1733	16.2416	0.1382	14.6872	0.2756
A25	2.1157	0.0195	0.4329	0.0216	0.0990	0.0128	1.1937	0.0404	1.4520	0.0924
A26	3.8795	0.0262	2.8284	0.0547	4.6628	0.0862	3.2612	0.0662	2.4292	0.1191
A34	0.1508	0.0052	3.5228	0.0610	1.3529	0.0470	0.4928	0.0260	0.3150	0.0432
A43	0.0018	0.0006	0.0334	0.0060	0.0231	0.0062	0.0055	0.0028	0.0059	0.0059
A66	0.0173	0.0018	0.2233	0.0155	0.0478	0.0089	0.0399	0.0074	0.0534	0.0178
A10 未分裂	0.0790	0.0038	0.0939	0.0101	0.1255	0.0144	0.0647	0.0094	0.0298	0.0133
A10 总共	6.2441	0.0328	7.1348	0.0850	6.3111	0.0993	5.0578	0.0816	4.2853	0.1565
A29	3.5796	0.0252	3.2071	0.0582	1.1233	0.0429	4.5156	0.0774	3.4345	0.1410
A30	2.5067	0.0212	13.0969	0.1129	2.2025	0.0598	4.4873	0.0772	2.5314	0.1215
A31	2.7386	0.0221	1.6556	0.0420	3.6005	0.0761	4.8328	0.0800	6.0881	0.1855
A32	3.6956	0.0256	1.5384	0.0405	1.0331	0.0411	2.7064	0.0604	2.5521	0.1220
A33	1.2080	0.0148	6.5607	0.0822	9.2701	0.1191	2.6593	0.0599	1.0754	0.0796
A74	0.0277	0.0022	1.9949	0.0461	0.0561	0.0096	0.2027	0.0167	0.1068	0.0252
A19 未分裂	0.0567	0.0032	0.2057	0.0149	0.0990	0.0128	0.1211	0.0129	0.0475	0.0168
A19 总共	13.8129	0.0468	28.2593	0.1504	17.3846	0.1555	19.5252	0.1481	15.8358	0.2832
AX	0.8204	0.0297	4.9506	0.0963	2.9916	0.1177	1.6332	0.0878	1.8454	0.1925

^a基因频率

30 ^b标准误差

5

表4.HLA-B抗原的估计基因出现频度

抗原	CAU		AFR		ASI		LAT		NAT	
	Gf ^a	SE ^b	Gf	SE	Gf	SE	Gf	SE	Gf	SE
B7	12.1782	0.0445	10.5960	0.1024	4.2691	0.0827	6.4477	0.0918	10.9845	0.2432
B8	9.4077	0.0397	3.8315	0.0634	1.3322	0.0467	3.8225	0.0715	8.5789	0.2176
B13	2.3061	0.0203	0.8103	0.0295	4.9222	0.0886	1.2699	0.0416	1.7495	0.1013
B14	4.3481	0.0277	3.0331	0.0566	0.5004	0.0287	5.4166	0.0846	2.9823	0.1316
B18	4.7980	0.0290	3.2057	0.0582	1.1246	0.0429	4.2349	0.0752	3.3422	0.1391
B27	4.3831	0.0278	1.2918	0.0372	2.2355	0.0603	2.3724	0.0567	5.1970	0.1721
B35	9.6614	0.0402	8.5172	0.0927	8.1203	0.1122	14.6516	0.1329	10.1198	0.2345
B37	1.4032	0.0159	0.5916	0.0252	1.2327	0.0449	0.7807	0.0327	0.9755	0.0759
B41	0.9211	0.0129	0.8183	0.0296	0.1303	0.0147	1.2818	0.0418	0.4766	0.0531
B42	0.0608	0.0033	5.6991	0.0768	0.0841	0.0118	0.5866	0.0284	0.2856	0.0411
B46	0.0099	0.0013	0.0151	0.0040	4.9292	0.0886	0.0234	0.0057	0.0238	0.0119
B47	0.2069	0.0061	0.1305	0.0119	0.0956	0.0126	0.1832	0.0159	0.2139	0.0356

10

15

	抗原	CAU		AFR		ASI		LAT		NAT	
		Gf	SE ^b	Gf	SE	Gf	SE	Gf	SE	Gf	SE
5	B48	0.0865	0.0040	0.1316	0.0119	2.0276	0.0575	1.5915	0.0466	1.0267	0.0778
	B53	0.4620	0.0092	10.9529	0.1039	0.4315	0.0266	1.6982	0.0481	1.0804	0.0798
	B59	0.0020	0.0006	0.0032	0.0019	0.4277	0.0265	0.0055	0.0028	0 ^c	
	B67	0.0040	0.0009	0.0086	0.0030	0.2276	0.0194	0.0055	0.0028	0.0059	0.0059
	B70	0.3270	0.0077	7.3571	0.0866	0.8901	0.0382	1.9266	0.0512	0.6901	0.0639
	B73	0.0108	0.0014	0.0032	0.0019	0.0132	0.0047	0.0261	0.0060	0 ^c	
10	B51	5.4215	0.0307	2.5980	0.0525	7.4751	0.1080	6.8147	0.0943	6.9077	0.1968
	B52	0.9658	0.0132	1.3712	0.0383	3.5121	0.0752	2.2447	0.0552	0.6960	0.0641
	B5 未分裂	0.1565	0.0053	0.1522	0.0128	0.1288	0.0146	0.1546	0.0146	0.1307	0.0278
	B5 总共	6.5438	0.0435	4.1214	0.0747	11.1160	0.1504	9.2141	0.1324	7.7344	0.2784
	B44	13.4838	0.0465	7.0137	0.0847	5.6807	0.0948	9.9253	0.1121	11.8024	0.2511
	B45	0.5771	0.0102	4.8069	0.0708	0.1816	0.0173	1.8812	0.0506	0.7603	0.0670
15	B12 未分裂	0.0788	0.0038	0.0280	0.0055	0.0049	0.0029	0.0193	0.0051	0.0654	0.0197
	B12 总共	14.1440	0.0474	11.8486	0.1072	5.8673	0.0963	11.8258	0.1210	12.6281	0.2584
	B62	5.9117	0.0320	1.5267	0.0404	9.2249	0.1190	4.1825	0.0747	6.9421	0.1973
	B63	0.4302	0.0088	1.8865	0.0448	0.4438	0.0270	0.8083	0.0333	0.3738	0.0471
	B75	0.0104	0.0014	0.0226	0.0049	1.9673	0.0566	0.1101	0.0123	0.0356	0.0145
	B76	0.0026	0.0007	0.0065	0.0026	0.0874	0.0120	0.0055	0.0028	0	
20	B77	0.0057	0.0010	0.0119	0.0036	0.0577	0.0098	0.0083	0.0034	0 ^c	0.0059
	B15 未分裂	0.1305	0.0049	0.0691	0.0086	0.4301	0.0266	0.1820	0.0158	0.0059	0.0206
	B15 总共	6.4910	0.0334	3.5232	0.0608	12.2112	0.1344	5.2967	0.0835	0.0715	0.2035
	B38	2.4413	0.0209	0.3323	0.0189	3.2818	0.0728	1.9652	0.0517	1.1017	0.0806
	B39	1.9614	0.0188	1.2893	0.0371	2.0352	0.0576	6.3040	0.0909	4.5527	0.1615
	B16 未分裂	0.0638	0.0034	0.0237	0.0051	0.0644	0.0103	0.1226	0.0130	0.0593	0.0188
25	B16 总共	4.4667	0.0280	1.6453	0.0419	5.3814	0.0921	8.3917	0.1036	5.7137	0.1797
	B57	3.5955	0.0252	5.6746	0.0766	2.5782	0.0647	2.1800	0.0544	2.7265	0.1260
	B58	0.7152	0.0114	5.9546	0.0784	4.0189	0.0803	1.2481	0.0413	0.9398	0.0745
	B17 未分裂	0.2845	0.0072	0.3248	0.0187	0.3751	0.0248	0.1446	0.0141	0.2674	0.0398
	B17 总共	4.5952	0.0284	11.9540	0.1076	6.9722	0.1041	3.5727	0.0691	3.9338	0.1503
	B49	1.6452	0.0172	2.6286	0.0528	0.2440	0.0200	2.3353	0.0562	1.5462	0.0953
30	B50	1.0580	0.0138	0.8636	0.0304	0.4421	0.0270	1.8883	0.0507	0.7862	0.0681
	B21 未分裂	0.0702	0.0036	0.0270	0.0054	0.0132	0.0047	0.0771	0.0103	0.0356	0.0145
	B21 总共	2.7733	0.0222	3.5192	0.0608	0.6993	0.0339	4.3007	0.0755	2.3680	0.1174
	B54	0.0124	0.0015	0.0183	0.0044	2.6873	0.0660	0.0289	0.0063	0.0534	0.0178
	B55	1.9046	0.0185	0.4895	0.0229	2.2444	0.0604	0.9515	0.0361	1.4054	0.0909
	B56	0.5527	0.0100	0.2686	0.0170	0.8260	0.0368	0.3596	0.0222	0.3387	0.0448
35	B22 未分裂	0.1682	0.0055	0.0496	0.0073	0.2730	0.0212	0.0372	0.0071	0.1246	0.0272
	B22 总共	2.0852	0.0217	0.8261	0.0297	6.0307	0.0971	1.3771	0.0433	1.9221	0.1060
	B60	5.2222	0.0302	1.5299	0.0404	8.3254	0.1135	2.2538	0.0553	5.7218	0.1801
	B61	1.1916	0.0147	0.4709	0.0225	6.2072	0.0989	4.6691	0.0788	2.6023	0.1231
	B40 未分裂	0.2696	0.0070	0.0388	0.0065	0.3205	0.0230	0.2473	0.0184	0.2271	0.0367
	B40 总共	6.6834	0.0338	2.0396	0.0465	14.8531	0.1462	7.1702	0.0963	8.5512	0.2168
40	BX	1.0922	0.0252	3.5258	0.0802	3.8749	0.0988	2.5266	0.0807	1.9867	0.1634

^a基因频率 ^b标准误差 ^c观察到的基因数是零

5

表5. HLA—DR抗原的估计基因出现频度

10

抗原	CAU		AFR		ASI		LAT		NAT	
	Gf ^a	SE ^b	Gf	SE	Gf	SE	Gf	SE	Gf	SE
DR1	10.2279	0.0413	6.8200	0.0832	3.4628	0.0747	7.9859	0.1013	8.2512	0.2139
DR2	15.2408	0.0491	16.2373	0.1222	18.6162	0.1608	11.2389	0.1182	15.3932	0.2818
DR3	10.8708	0.0424	13.3080	0.1124	4.7223	0.0867	7.8998	0.1008	10.2549	0.2361
DR4	16.7589	0.0511	5.7084	0.0765	15.4623	0.1490	20.5373	0.1520	19.8264	0.3123
DR6	14.3937	0.0479	18.6117	0.1291	13.4471	0.1404	17.0265	0.1411	14.8021	0.2772
DR7	13.2807	0.0463	10.1317	0.0997	6.9270	0.1040	10.6726	0.1155	10.4219	0.2378
DR8	2.8820	0.0227	6.2673	0.0800	6.5413	0.1013	9.7731	0.1110	6.0059	0.1844
DR9	1.0616	0.0139	2.9646	0.0559	9.7527	0.1218	1.0712	0.0383	2.8662	0.1291
DR10	1.4790	0.0163	2.0397	0.0465	2.2304	0.0602	1.8044	0.0495	1.0896	0.0801
DR11	9.3180	0.0396	10.6151	0.1018	4.7375	0.0869	7.0411	0.0955	5.3152	0.1740
DR12	1.9070	0.0185	4.1152	0.0655	10.1365	0.1239	1.7244	0.0484	2.0132	0.1086
DR5 未分裂	1.2199	0.0149	2.2957	0.0493	1.4118	0.0480	1.8225	0.0498	1.6769	0.0992
DR5 总共	12.4449	0.0045	17.0260	0.1243	16.2858	0.1516	10.5880	0.1148	9.0052	0.2218
DRX	1.3598	0.0342	0.8853	0.0760	2.5521	0.1089	1.4023	0.0930	2.0834	0.2037

15

a基因频率

b标准误差

20

表3， 4和5来源于北美群体HLA 基因和单倍体频率：： The National Marrow Donor Program Donor Registry, Mori, M. 等

确定具有MHC亲和性的片段是否是有用表位

如上所述，本方法的最初步骤是从最初的肽片段群中选择出带有实际的或预测的MHC亲和性的肽亚群。进一步对这些选择的片段进行分析以

25 确定哪些片段可以由细胞在体内条件下产生，所述的体内条件可以导致肽与被选择的MHC等位基因结合。所有同时符合MHC亲和性和正确蛋白水解加工这两个标准的肽被叫做“被发现的表位”。许多方法可以被用来确定哪些肽段可以在体内被蛋白水解加工所产生。这些方法包括从被溶解的MHC和完整细胞中洗脱，蛋白水解切割基元的计算机测序分析，

30 和对由细胞蛋白水解机制所产生的实际肽段进行体外分析。

在优选实施方案中，可以产生一系列的在中间含有单个的或者簇集的候选肽序列的合成肽。这样的肽的典型长度范围为大约10个到75个氨基酸。在优选实施方案中，合成肽的长度为大约20—60个氨基酸。在更优选实施方案中，簇的长度为大约30—40个氨基酸。通过使用标准肽合成化学，包括t-Boc保护化学（t-Boc protection chemistry），Fmoc保护化学（Fmoc protection chemistry）等，本领域的普通技术人员可以生产出用于以后筛选的候选肽群。

作为选择，可以在体外通过蛋白酶消化或化学切割TTA或其片段来产生含有候选肽的肽片段。用来制备这些TAAs片段的蛋白酶消化可以使用很多已知的蛋白酶，包括但不限于蛋白酶体蛋白酶，胰蛋白酶，一糜蛋白酶，菠萝蛋白酶，梭菌蛋白酶，弹性蛋白酶，内切蛋白酶，外切蛋白酶，蛋白酶K，无花果蛋白酶，木瓜蛋白酶，胃蛋白酶，血纤维蛋白溶酶，嗜热菌蛋白酶，凝血酶，胰蛋白酶，组织蛋白酶和其它酶。也可以使用化学方法来产生候选肽。用于切割肽键的适合的化学制剂或化学反应包括：弱酸切割，溴化氰，羟胺，亚碘酰基苯甲酸，2-硝基-5-氰硫基苯甲酸盐等等。在一种实施方案中，可以使用未被片段化的TAA，尽管使用特别大的起始序列会使分析复杂化。

不管含有候选肽的片段是如何产生的，确定哪些肽是由细胞机制制造的是重要的。在本发明的一种实施方案中，蛋白酶体消化被用来估计细胞表位产生。在这个实施方案中，纯化免疫蛋白酶体和看家蛋白酶体，用于体外使用以评价由这两种蛋白酶体天然产生的抗原的所有成分。

通常，通过使用亲和纯化法从细胞抽提物中制备蛋白酶体。在优选实施方案中，使用标准技术来制备细胞溶解产物。如果红血球不是最初的原料，就用超速离心法来清除细胞溶解物。而后，使用任何一种纯化技术（包括多种形式的层析法），将制备的细胞溶解产物从其它细胞组分中纯化出来。

在一种实施方案中，用亲和层析法来纯化蛋白酶体。将细胞溶解产物加到含有单克隆抗体（mAb）的亲和柱，所述的单克隆抗体对应一种蛋白酶体的亚单位。而后冲洗柱子以将结合的蛋白酶体从其它细胞材料中纯化出来。冲洗后，将蛋白酶体从柱子上洗脱下来。根据标准底物上

的蛋白含量和蛋白水解活性来表征洗脱液。

同时使用看家和免疫蛋白酶体的切割分析，从多种TAA中产生I类表位。被pAPCs呈递的表位与免疫蛋白酶体切割产物相对应，而被肿瘤细胞和被细胞内寄生物慢性感染的细胞呈递的表位则与看家蛋白酶体的切割产物相对应。一旦进行了消化反应，产生的特殊分子种类就被鉴定出来。在优选实施方案中，这是由质谱分析来完成的。这样允许快速鉴定由两种蛋白酶体中的任何一种所产生的天然肽片段。在另一种实施方案中，通过使用计算机模型（此计算机模型基于相关蛋白溶解活性的切割基元），来预测对靶抗原或其片段所进行的切割，这些切割是通过使用免疫和看家蛋白酶体或者通过使用内体蛋白酶/溶酶体蛋白酶（见下面）进行的。

由于I类MHC在内质网中起始折叠的时候，主要装载蛋白酶体来源的肽，II类MHC的结合裂隙被此区域中所谓的恒定链（Ii）所阻断。II类MHC的肽的装载主要在内体区域中，利用内体蛋白酶和溶酶体蛋白酶所产生的肽进行。因此，如果需要体外鉴定MHC II类表位时，可以用来自内体和/或溶酶体片段的蛋白酶制剂替代蛋白酶体。多种用来完成这一替代的方法在文献中有描述。例如在Kido & Ohshita, *Anal. Biochem.*, 230:41-7 (1995); Yamada, 等, *J. Biochem. (Tokyo)*, 95:1155-60 (1984); Kawashima, 等, *Kidney mt.*, 54:275-8 (1998); Nakabayshi & Ikezawa, *Biochem. mt.* 16:1119-25 (1988); Kanaseki & Ohkuma, *J. Biochem. (Tokyo)*, 110:541-7 (1991); Wattiaux, 等, *J. Cell Biol.*, 78:349-68 (1978); Lisman, 等, *Biochem. J.* 178:79-87 (1979); Dean, B., *Arch. Biochem. Biophys.*, 227:154-63 (1983); Overdijk, 等, *Adv. Exp. Med. Biol.*, 101:601-10 (1978); Stromhaug, 等, *Biochem. J.*, *Biochem. J.*, 335:217-24 (1998); Escola, 等, *J. Biol. Chem.* 271:27360-5 (1996); Hammond, 等, *Am. J. Physiol.*, 267:F516-27 (1994); Williams & Smith, *Arch. Biochem. Biophys.* 305:298-306 (1993); Marsh, M., *Methods Cell Biol.*, 31:319-34 (1989); 和Schmid & and Mellman, *Prog. Clin. Biol. Res.*, 270:35-49 (1988)中均公开了制备合适的蛋白水解制剂的方法。

在另一种实施方案中，用来决定细胞机制产生哪些表位的消化作用发生在表达TAA或其片断的细胞中。对于I型表位，优选，例如蛋白印

迹，来确定由细胞表达的蛋白酶体类型。随后，产生的MHC表位可以从溶解的和纯化的MHC中洗脱下来，在Falk, K. 等 Nature 351:290, 1991中有描述，或者直接从完整细胞中洗脱下来，在美国专利5,989,565中有描述。随后用质谱法鉴定洗脱的片段。

5

靶蛋白片段分析

将用质谱法测定的分子种类与上面预测的候选肽相比较。在I型表位的情况中，期望的分子种类与候选肽等长度或者比候选肽长，而且有共同的C末端；至少25个氨基酸的N末端修饰可以不依赖于蛋白酶体而发生 (Craiu, A. 等 Proc. Natl. Acad. Sci. USA 94:10850-55, 1997)。II类MHC对其所结合的肽的长度非常耐受，所以，基本标准并不是正确末端的产生，而是表位中间没有被切割。

随后，合成被选择的消化产物并在分析方法中（如HPLC对消化等分）将其用作标准。这提供了对消化产物同一性的进一步检查，而且使得其产量得到确定。在很少的情况下，多于一种的潜在的产物可能具有十分类似的质量和化学性质，使它们不能被这些方法可靠地区分开。在这些情况下，可以收集HPLC峰并且进行直接测序。

对MHC结合肽的分析

合成表位并测试其与MHC 受体结合的能力。例如，在一种优选分析中，呈现MHC I受体的细胞可以用于测定与被标记了放射性核素的候选肽的结合亲和性。另一种优选方法使用基于细胞培养的分析来测量肽与MHC I 结合的能力。在该分析中，使用抗原加工相关转运蛋白（TAP）缺陷的细胞来确定候选肽是否有结合到MHC I 受体的能力。TAP⁻细胞具有这样的表型：I类MHC蛋白不总是能够正确折叠，因此降低或破坏了MHC I的表面表达。当细胞充满了能够结合到MHC I裂隙的外源肽的时候，受体的表达被恢复。这可以通过例如RIA, FACS等几种方法来监控。通过使用TAP⁻细胞，本领域中的技术人员可以在不需要进行详细结合亲和性分析的情况下，筛查大量的用于受体结合的可能的候选肽。

本发明中各种实施方案中的分析方法，对于检查由各种方法产生的

候选肽非常有用。例如，所述的分析可以用来评价通过体外方法或者计算机分析产生的复合候选肽，以鉴定那些具有MHC受体结合特性的候选序列。本发明中此实施方案的优选候选肽是这样的肽：早已明确它们是看家和/或免疫蛋白酶体的蛋白水解的产物。显示的或预测的与MHC结合5 的切割产物，无论是体内切割产物还是体外切割产物，都被适当地称为“发现的表位”。此处公开了与本发明相关的表位簇。

ECRs被加工到pAPCs中的MHC结合表位

免疫系统不断地观察体内是否出现外来抗原，部分是通过pAPCs的10 活性。pAPCs吞噬在细胞外环境中发现的物质，将所发现的物质从多肽的形式加工成为短一些的、长度大约是3—23个氨基酸的寡肽，并且将一些所得到的肽通过pAPCs的MHC复合物，呈递给T细胞。例如，溶解的肿瘤细胞将其细胞物质（包括各种蛋白），释放到细胞外环境中。这些释放的蛋白可以被pAPCs吞噬并加工为分散的肽，而后通过MHC呈现到15 pAPCs的表面。通过这种机制，不是整个靶蛋白呈现在pAPCs的表面，而是只有靶蛋白的一个或更多分散片段呈现为MHC结合表位。如果T细胞识别了被呈现的表位，该T细胞被活化而且引起免疫反应。

类似地，pAPCs上的清道夫受体（scavenger receptor）可以摄取裸露的核酸序列或含有靶核酸序列的重组有机体。核酸序列被摄取到pAPC20 中，随后引起编码产物的表达。如上，当ECR可以被加工到一种或多种有用的表位时，这些产物可以被作为MHC表位呈递，供T细胞识别。

结合MHC的表位常常以簇的形式不均一地分布在蛋白序列中。本发明的实施方案涉及对靶蛋白特殊区域中的表位簇区域（ECRs）的识别。候选ECRs可能是不同蛋白水解酶的天然底物，而且可能被加工到一种或25 多种呈现在pAPC表面的MHC表位中。与传递整个蛋白或者生物制剂传统疫苗相比，ECRs可以作为疫苗给予，以很高的概率引起至少一种表位呈现在MHC上，而不需要使用全长序列。

ECRs在鉴定分散MHC结合表位中的应用

30 鉴定用于疫苗的假定MHC表位通常包括：使用分析蛋白或基因序列

的可用的预测性算法来预测肽段对MHC的亲合性。这些算法根据预测的亲合性或其它与MHC结合相关的特性，对假定表位进行排序。此种分析的典型算法包括Rammensee和NIH (Parker)算法。可是已经证明，将自然呈现在细胞表面的表位从那些通过使用这些算法预测的假定表位中分离出来，是一个困难且繁重的过程。ECRs在表位鉴定过程中的应用，可以大大简化鉴别分散的MHC结合表位的任务。

在优选实施方案中，在自动肽合成器上合成ECR多肽，随后，使用蛋白水解酶体外消化这些合成的ECRs，所用的这些蛋白水解酶参与加工用于表位呈递的蛋白。而后，使用质谱和/或分析HPLC来鉴别消化产物，并使用体外MHC结合研究来评价这些产物实际结合MHC的能力。一旦ECRs中含有的表位显示出结合了MHC，就可以将它们整合到疫苗中或者用于诊断，或者以分散表位的形式或者以ECRs结合的形式。

在这种优选实施方案中对ECR的使用（因为序列相对短，它可以通过化学合成生产），比起在其他情况下所需要使用整个蛋白的情况来说，是一个重要的进步。这是因为必须使用重组表达载体体系和/或复合物纯化步骤才能够生产整个蛋白。使用化学合成的ECRs如此简单，使得分析和鉴定大量的表位成为可能，而与当前使用的其它方法相比，大大降低了操作过程的时间和费用。使用定义的ECR同时也大大简化了对消化产物的质谱分析，因为ECR消化产物仅是整个蛋白消化产物的一小部分。

在另一种实施方案中，编码ECRs的核酸序列被用来表达细胞或细胞系中的多肽，以评价哪些表位呈现在表面。可以使用多种方法来检测表面的表位。优选实施方案包括细胞溶解和MHC亲和纯化，随后对肽从MHC洗脱和分析；或从完整细胞洗脱表位；（分别地，FaLk, K. 等 Nature 351:290, 1991, 和美国专利 5,989,565）。在一种对肽（以此种方法从MHC中洗脱的肽）进行分析的灵敏方法中，使用毛细或纳毛细HPLC ESI 质谱分析和在线测序。

含有ECRs的靶相关抗原

可以限定ECRs 的TAAs，包括那些来自TuAAs的TAAs，包括胎性癌，睾丸癌，失控基因（deregulated genes），来自错误易位的融和基因，

分化抗原, 胚胎抗原, 细胞周期蛋白, 突变的肿瘤抑制基因和过表达的基因产物, 包括癌基因。此外, ECRs可以衍生于病毒基因产物, 特别是那些与导致慢性疾病或者致瘤相关的病毒, 例如疱疹病毒, 人乳头状瘤病毒, 人类免疫缺陷病毒, 和人T细胞白血病病毒。ECRs还可以衍生于寄生生物的基因产物, 例如锥虫, 利什曼虫, 和其它细胞内的或寄生的生物。

一些TuAA包括 α -胎蛋白, 癌胚抗原(CEA), 衍生于食道癌的NY-ESO-1, 和SSX基因, SCP-1, PRAME, MART-1/MelanA (MART-1), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2, MAGE- 1, MAGE-2, MAGE-3, BAGE, GAGE-1, GAGE-2, p15; 过表达的癌基因, 和突变的肿瘤抑制基因例如 p53, Ras, HER-2/neu; 染色体易位产生的独特肿瘤抗原例如 BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR1, 和病毒抗原, EBNA1, EBNA2, HPV-E6,-E7; 前列腺特异抗原(PSA), 前列腺干细胞抗原(PSCA), MAAT-1, GP- 100, TSP- 180, MAGE-4, MAGE-5, MAGE-6, RAGE, pl85erbB2, pl80erbB-3, c-met, nm-23H1, TAG-72, CA 19-9, CA72-4, CAM 17.1, NuMa, K-ras, -连环蛋白, CDK4, Mum-1, p15, 和p16。

对于病原体和肿瘤, 还可以考虑许多其它TAAs。就TuAAs来说, 本领域中许多熟知的和可用的方法可以被用来鉴别那些与正常细胞相比, 在瘤细胞中分化表达的基因与基因产物。这些技术的实例包括使用微排列技术的差异杂交, 减数杂交克隆; 或者在mRNA水平, 或者蛋白表达水平的差异显示; EST测序; 和SAGE(基因表达的序列分析)。这些核酸技术在Carulli, J.P. 等, J. Cellular Biochem Suppl. 30/31:286-296, 1998中被综述。蛋白的差异显示包括, 例如, 肿瘤和正常组织细胞溶解物的二维聚丙烯酰胺凝胶电泳比较, 定位肿瘤中独特的和过表达的蛋白质渍点, 从胶中回收蛋白, 和使用传统的生化技术或者质谱分析测序技术鉴定蛋白。额外的鉴定TuAAs的技术是SEREX技术, 在Türeci,., Sahin, U., and Pfreundschuh, M., “人类肿瘤抗原的血清学分析: 分子限定及其意义”, Molecular Medicine Today, 3:342, 1997中讨论。

这些方法和其他方法的使用, 给本领域的技术人员提供了这样的技术, 这些技术对于鉴定包含在靶细胞内的基因和基因产物是必需的, 所

述的这些基因和基因产物可以被用作产生本发明表位的可能候选蛋白。但是，在本发明的实施中，并不是必须要鉴定新的TuAA或TAA。更确切地说，本发明的实施方案使得从任何相关蛋白序列中鉴定出有用的表位成为可能，不管这个序列是新的还是已知的。

5

蛋白序列分析来鉴定表位簇

在本发明的优选实施方案中，ECRs的鉴定包括两个主要步骤：1.鉴定好的假定表位；和2.限定任何簇的范围，其中表位定位在这些簇中。这两个步骤中的每一个步骤都有多种优选实施方案，并且第一个步骤中选择的实施方案可以与第二个步骤中选择的实施方案任意组合。此处公开的每一个步骤中的方法和实施方案仅仅是例证性的，并不以任何方式限制本发明的范围。本领域的技术人员能理解这些可以用来分析特异TAA的特殊方法，而且可以依照本发明用很多方法来实现这种分析。

用来鉴别好的假定表位的优选实施方案包括使用任何可用的预测性算法，这些算法用来分析蛋白或基因的序列以预测肽段对MHC的亲15 和性，或者根据预测的亲和性或其它与MHC结合相关的特征来给假定表位排序。如上所述，可以用于这种类型分析的典型算法包括Rammensee和NIH (Parker)算法。同样，可以通过直接或间接分析MHC结合来鉴定假定表位。为了选择“好的”假定表位，有必要根据预测软件报告的得分或者根据分析得到的结合亲和性而设定截断（cut off）点。在一些实施方案中，这种截断是绝对的。例如，该截断可以基于测定的或者预测的表位与所选择的MHC等位基因之间的分离半衰期。在这些情况下，截断的实施方案可以是任何大于例如0.5分钟的分离半衰期；在优选实施方案中，大于2.5分钟；在更优选实施方案中大于5分钟；在非常严格的实施方案中，可以大于10或20或25分钟。在这些实施方案中，好的假定表位是那些被预测或者被鉴别为具有良好MHC结合特征的表位，限定在定义截断点的期望的一边。同样，截断可以基于表位与所选择的MHC等位基因之间的测定的或预测的结合活性。此外，绝对截断可以仅仅是假定表位的一个选择的数字。

30 在其它实施方案中，截断是相对的。例如，可以用选择的假定表位

总数的百分比来建立截断，该截断用来将候选序列定义为好的假定表位。表位排序的性质还是来自测定的或预测的MHC结合；用于这种决定的性质可以是任何与MHC结合相关的或指示的性质。在优选实施方案中，可以将好的假定表位的鉴别结合排序候选序列的多种方法。在这样的实施方案中，典型的好的表位是这样的表位，它们或者基于不同的方法和参数，一致被认为是好的表位；或者至少在其中的一种方法中得到特别高的排序。

当几种好的表位被鉴定出来时，可以通过它们之间的相对位置分析来确定用于疫苗或疫苗设计的最理想的簇。此分析基于TAA序列中所选择表位的特征的密度。将这样的区域叫做ECRs：该区域具有特征的最高密度，或者具有高于某种选择截断的密度。本发明不同的实施方案使用不同的特征来进行密度分析。例如，一种优选的特征只是出现任何好的表位（用任何合适的方法所定义的那样）。在这个实施方案中，在密度分析中平等地处理所有高于截断的假定表位，最好的簇是这样的簇：具有最高密度的好的假定表位/氨基酸残基的簇。在另一种实施方案中，优选的特征基于以前被用于计分的参数或用于排序假定表位的参数。在这种实施方案中，具有比另一种假定表位分数高一倍的假定表位，相对于另外的假定表位，在密度分析中被双倍加权。其它的实施方案还考虑到得分或排序，只是进行了减权处理，例如，在密度分析中，使用得分的log或者平方根来给予一些假定表位比其它更多的权数。

依照所要分析的TAA的长度，可能的候选表位的数量，好的假定表位的数量，好的假定表位得分的可变性，和在任何特定分析中变得明显的其它因素，本发明的各种实施方案可以单独使用或者结合使用来鉴定在特定应用中最有用的ECRs。使用复合方法的迭代分析或平行分析在许多情况下有益。ECRs是鉴定真正MHC表位提高效率的工具，并且是将MHC表位有效“包装”到疫苗中的工具。因此，此处描述的任何实施方案，或者基于本发明的、本领域内的技术人员非常清楚的其它实施方案，对于提高通过在疫苗和疫苗设计中使用ECRs而不是使用整个TAAs所作的这些努力的效率，非常有用。

由于许多或大多数TAAs具有预测的MHC表位的低密度区域，所以

使用ECRs提供了一种有价值的方法学，该方法学避免了在疫苗和表位鉴定方法中包含低表位密度区域。这样有用的ECRs也可以被定义为不是整个TAA的TAA的任何部分，这些部分具有比整个TAA更高的假定表位密度，或者比任何具有特别低密度的假定表位的TAA区域更高的假定表位密度。因此，在本发明的此部分中，ECR可以是具有升高的表位密度的任何TAA片段。在一些实施方案中，ECR可以包括大约80%TAA长度的区域。在优选实施方案中，ECR可以包括大约50%TAA长度的区域。在更优选实施方案中，ECR可以包括大约30%TAA长度的区域。在最优选实施方案中，ECR可以包括5%—15%TAA长度的区域。

在本发明的另一方面中，ECR可以根据绝对长度定义。因此，依照这种定义，9聚物（9-mer）表位的最小簇包括10个氨基酸残基并且是具有8个共有氨基酸的两个重叠9聚物。在优选实施方案中，簇的长度在大约15-75个氨基酸之间。在更优选实施方案中，簇的长度在大约20-60个氨基酸之间。在最优选实施方案中，簇的长度在大约30-40个氨基酸之间。

实际上，如上所述，ECR鉴定可以使用简单密度函数例如表位的数目除以那些表位跨度的氨基酸的数目。并不是必须要求表位重叠，但是单个表位的值是没有意义的。如果只使用百分比截断的单个值，而不使用表位预测的绝对截断，在此步骤设定单一界限来限定簇是可能的。可是，同时使用绝对截断和使用不同的百分比截断来执行第一步骤，可以在候选肽的总密度中产生变量。这些变量可能需要进一步计算或处理。例如，对于任何特定长度的蛋白，只考虑三个候选表位，与考虑30个候选表位相比，两个表位的重叠对于前者更有意义。考虑到这个特点，为了提高计算的意义，给予特殊簇的权数还可以除以实际被考虑的可能肽段部分。这将结果度量为亲本蛋白中预测的表位的平均密度。

类似地，一些实施方案将好的假定表位的得分基于蛋白中每个氨基酸所考虑的肽的平均数。所得到的比率代表了这样一种因素，通过这种因素，假定簇中预测表位的密度与蛋白中的平均密度不同。因此，ECR在一种实施方案中被定义为任何含有两个或更多比率超过2的预测表位，也就是说，任何具有两倍表位平均密度的区域。在其它实施方案中，如果比率超过1.5，3，4或5，或更多，此区域就被定义为ECR。

用靶蛋白中每个氨基酸的肽平均数来计算ECR的出现，可以使高度聚集的ECRs更加突出，而不用考虑单个组分的得分/亲和性。这对于使用基于得分的截断最为合适。可是，只具有少量数目高排序候选者的ECR可以比具有几个密集但排序较低的候选者，更有生物学意义，特别是在只有小百分比的候选肽总数被指定为好的假定表位的时候。这样，在一些实施方案中，考虑这些单个肽的得分是合适的。在上述的计算中，这可以非常容易地通过用假定簇中肽的得分之和，替代假定簇中肽的数目来完成。

这种得分和的方法对于含有高分表位的分散填充的簇更加灵敏。因为由BIMAS-NIH/Parker算法产生的分数（即：分离半衰期）的宽范围可以导致单个高分的肽，使其它可能的表位的贡献相形见绌，在这一过程中优选使用分数的log值，而不是分数本身。

多种其它运算在一种或另一种条件下可以被修改。一般而言，建立表位密度函数以使其与假定簇中预测的表位的数目，得分，排序成比例，并且与氨基酸的数目或者包含在假定簇中的蛋白片段成反比。作为选择，被选择数目的邻接氨基酸的一个窗口可以用来评价函数。无论哪种情况，函数都要被整个蛋白中的所有预测的表位评价。如果假定簇（或窗口）与整个蛋白的值的比率大于例如1.5，3，4或5，或更多，ECR被定义。

20 对MHC结合的靶基因产物的分析

一旦TAA被鉴定出来，就可以使用蛋白序列来鉴定对MHC肽结合裂隙具有已知的或者预测的亲和性的假定表位。肽片段的试验可以体外进行，或者使用可以由计算机分析的序列来确定的肽片段与MHC受体结合。在本发明的一种实施方案中，对基于靶蛋白氨基酸序列的肽段进行分析，分析它们结合MHC肽结合裂隙的预测的能力。用于此目的合适的计算机算法的实例包括Rammensee/SYPFPEITHI 和NIH (Parker)站点，这些实例在以前的“表位发现部分”中作过讨论。

作为预测性算法的选择，可以使用许多标准的体外受体结合亲和性分析来鉴定对特定MHC等位基因有亲和性的肽。因此，通过使用本发明此部分的方法，可以将起始的肽群范围缩小到这样的范围：只包括含有

对选择的MHC等位基因具有实际或预测亲和性的假定表位的肽段。所选择的MHC的常见的等位基因，及其大约出现频度，示于上面的表3—5中。

已经发现，预测的表位通常聚集在TAA氨基酸序列的一个或多个特殊区域。对这些ECRs的鉴别，为设计刺激细胞免疫的有效疫苗，提供了一个简单并且实用解决方法。对于需要免疫表位的疫苗，ECR被作为疫苗是直接有用的。这是因为pAPCs的免疫蛋白酶体可以对该簇进行正确的加工，释放一个或多个含有的MHC结合肽，以同样方法，含有免疫蛋白酶体的细胞加工和呈递衍生于完全TAA的肽。该簇也是一种有用的原材料，用来鉴定由活跃在周围细胞中的看家蛋白酶体产生的看家表位。

10

对疫苗设计的额外考虑

在人群中存在着无数的MHC I等位基因。因此，在优选实施方案中，疫苗设计可以考虑患者的MHC I 基因型，以传递含有对特殊患者MHC 等位基因有合适结合亲和性的表位。由于患者的相关基因座可能是纯和的或者是杂合的，所以在本发明的一些实施方案中，优选对单个MHC I等位基因最理想的表位，而在其它实施方案中，可以优选对应不同MHC等位基因的表位。主要的I类MHC类型的部分列表（每种通常由复合等位基因编码）和它们大约的出现频度，在表6中报告。

20

表 6

HLA型	在混和人种群中的出现频度
A1	30%
A2	47%
A3	23%
A11	15%
A24	14%
A29	7%
A26	6%
B7	22%

B8	21%
B14	9%
B18	10%
B27	8%
B35	20%
B44	26%
B62	14%
B60	10%
B51	10%

5

在本发明的另一种实施方案中，给pAPCs提供看家表位和表位簇。表位簇是含有或者编码至少两条具有对MHC I有已知的或者预测的亲

10 性的序列。虽然优选将看家表位以完全被加工的状态提供给pAPCs，或者作为一种前体提供给pAPCs（以某种方式设计该前体，以使其可以在pAPCs中被加工为有效的看家表位），免疫表位可以被pAPCs从更大的前体加工而成。这是因为免疫蛋白酶体在pAPC中有组成性活性，而且完全有能力将大概任何长度的合适前体加工成“正确”的免疫表位。

15 为了提供免疫表位，可能的表位通常但不总是在含有多个表位的TAA的分散片段的簇中被发现。仅仅给pAPC提供含有一个可能表位簇的多肽，或者编码簇的核酸，或者表达簇的重组重组的有机体，就能够使pAPCs产生至少一种合适的免疫表位。由于表位簇通常对超过一种的I类MHC等位基因含有可能的表位，所以在许多实施方案中，可以使用单个簇来产生对超过一种I类MHC等位基因有用的免疫表位。

20

在优选实施方案中，给一位患者接种疫苗，该疫苗含有来源于所选

择TAA的看家表位。该看家表位可以是多肽，或者编码多肽的核酸，或者被设计来表达分散表位的重组的有机体。除了这一含有看家表位（或者是多肽，核酸，或是重组重组的有机体）的“最小”的疫苗之外，本发明的实施方案包括额外含有一种或多种其它看家表位或者一种或多种免疫表位或者其任意组合的疫苗。这些表位可以衍生于相同TAA，或者衍生于不同的TAAs。

本发明的优选实施方案包括疫苗的给予方法，该疫苗包含看家表位以诱导治疗性免疫反应。疫苗给予患者的方法可按照本领域公知的方法进行。对给予TAAs表位的方法没有限制，包括经皮肤的，结节内，结节周，经口的，静脉内的，真皮内的，肌肉内的，腹膜内的，和粘膜给予。一种诱导CTL反应的特别有用的疫苗传送方法，在PCT出版物No. WO 99/01283中公开，题目为“一种诱导CTL反应的方法”，申请日为1998年7月10日。

因为表位同步系统在诱导细胞介导的免疫反应中 useful，所以诱导特异的T细胞对靶细胞进行反应的疫苗，同样包含在本发明的优选实施方案中。该疫苗含有能够有效引起pAPC或pAPCs群呈现看家表位的看家表位浓度。有利地，疫苗可以包含多个看家表位，或者一种或多种任选地与一种或多种免疫表位结合的看家表位。疫苗制剂所含有的肽和/或核酸的浓度，要足以引起pAPCs呈现表位。该制剂优选含有表位的总浓度大约为1 μ g-1mg/100 μ l疫苗制剂。本发明可使用常规的肽疫苗和/或核酸疫苗的剂量和配量，而且剂量服法在本领域被人熟知。在一种实施方案中，有利地，成人的单次剂量可以为该组合物的1—5000 μ l，一次给予或多次给予，例如2，3，4或更多剂量给予，间隔为1周，2周，1个月或更多。在特别优选实施方案中，该组合物被连续给予，通过使用胰岛素注射器，直接给予到淋巴结，给予率为至少1 μ l每小时，持续超过几天。这种给予可以阶段性反复进行以保持CTL反应，在PCT出版物No. WO 99/0 1283.中有更全面的描述。

本发明所公开的组合物和方法还可以考虑将助剂整合到制剂中以提高疫苗的性能。具体地说，设计在制剂中添加助剂，以提高pAPCs对表位的传送或摄取。本发明考虑的助剂在本领域中被技术人员熟知，例如 GMCSF, GCSF, IL-2, IL-12, BCG, 破伤风类毒素, 和骨桥接素/ETA-1。

在另一实施方案中，可以将与看家表位反应T细胞作为采用的免疫治疗 (adoptive immunotherapy) 给予患者。这种T细胞最容易通过使用来自原初供体细胞的体外免疫得到，通过将患者用作供体也是可行的。体外免疫的技术在本领域中被人熟知，例如Stauss 等, Proc. Natl. Acad. Sci. USA 89:7871-7875, 1992; Salgaller 等 Cancer Res. 55:4972-4979, 1995; Tsai 等, J. Immunol. 158:1796-1802, 1997; 和 Chung 等, J. Immunother. 22:279-287, 1999。也可以考虑将免疫的供体或者患者本身用作T细胞原材料。(实例见: Oelke, M. 等 Clin cancer Res. 6:1997-2005, 2000; Gervois, N. 等 Clin cancer Res. 6:1459-1467, 2000; Valniori, D. 等 cancer Res. 59:2167-3173, 1999; Tsai, V. 等 Crit. Rev. Immunol. 18:65-75, 1998; Matsunaga, K. 等 Jpn. JJ Cancer Res. 90:1007-1015, 1999; van Elsas, A. 等 Eur J. Immunol. 26:1683-1689, 1996; 和 Alters, S.E. 等 Adv. Exp. Med. Biol. 417:519-524, 1997)。一旦产生，可以通过使用本发明的疫苗和/或细胞因子来进行刺激，利用体外扩充以得到足够数量的这种T细胞(实例见: Kurokawa, T. 等, Int. J. Cancer 91:749-746, 2001)。这些T细胞可以组成识别一个或多个表位的克隆或多克隆群。典型地，大约 10^5 - 10^8 数量级的细胞传递到小鼠中， 10^8 - 10^{11} 数量级的细胞被传递到人体中(实例见: Drobyski, W.R. 等 Blood 97:2506-2513, 2001; Seeley B.M. 等 Otolaryngol. Head Neck Surg. 124:436-441, 2001; 20 Kanwar, J.R. 等 Cancer Res. 61:1948-1956, 2001; Plautz, G.E. 等 Clin. Cancer Res. 6:2209-2218, 2000; Plautz, G.E. 等, J. Neurosurg. 89:42-51, 1998; 和 Plautz, G.E. 等, Urology 54:617-623, 1999)。克隆和其它更富集的群体通常需要更少细胞的传递。识别的表位可以是看家表位或者是看家和免疫表位的结合。同样可以预想，可以用遗传工程在细胞系中表达克隆的TCRs，所述的细胞系适合用在采用的免疫治疗当中。可以克隆有用TCRs的来源的实施例包括上述的T细胞，和用本发明的疫苗免疫了的HLA-转基因小鼠。本领域的技术人员很清楚额外的变异。

在本发明的一些实施方案中，疫苗可以包括通过遗传改造在宿主内表达表位的重组有机体，例如病毒，细菌或者寄生虫。例如，单核细胞基因李斯特杆菌 (*Listeria monocytogene*) (革兰阳性,兼性细胞内细菌)

是一种强有力的将TuAAs靶向免疫系统的载体。在优选实施方案中，该载体可以被改造来表达看家表位以诱导治疗反应。这种有机体的正常感染途径是通过内脏，而且可以经口传递。在另一种实施方案中，编码TuAA看家表位的腺病毒（Ad）载体可以用来诱导抗病毒和抗肿瘤反应。衍生于骨髓的树突状细胞可以被病毒构建体转导并被随后注入，或者病毒可以经由皮下注射直接被传送到动物体内而引起可能的T细胞反应。另一种实施方案使用重组痘苗病毒，该病毒被设计成为编码对应TAA看家表位的氨基酸序列。以微基因结构携带具有合适核苷酸替代构建体的痘苗病毒，可以指导看家表位的表达，引起治疗性T细胞针对表位的反应。

10 依照本发明，在此公开了被用作疫苗的特别有用的核酸构建体。

编码表位的载体构建体

本发明提供了用作治疗性疫苗的核酸构建体。该构建体包括一个编码区，该编码区含有编码多肽的序列。此多肽是TAA的表位。在一种优选实施方案中，靶细胞是瘤细胞而且多肽是TuAA的表位或者表位的前体。在另一个实施方案中，靶细胞是任何感染了细胞内寄生物的细胞。此处用到的术语“寄生物”包括任何有机体或者感染物例如：在宿主体内有细胞内感染期的病毒。这些寄生物包括但不限于：病毒例如腺病毒，巨细胞病毒，EB病毒，单纯疱疹病毒1，单纯疱疹病毒2，人类疱疹病毒6，水痘一带状疱疹病毒，乙型肝炎病毒，丁型肝炎病毒，乳头状瘤病毒，细小病毒B19，多瘤病毒BK，多瘤病毒JC，丙型肝炎病毒，麻疹病毒，风疹病毒，人类免疫缺陷病毒（HIV），人T细胞白血病病毒I或人T细胞白血病病毒II；细菌例如衣原体，李司忒氏菌，沙门氏菌属，军团菌属，布鲁氏菌属，柯克斯菌属，立克次体，分支杆菌属，和原生动物例如利什曼虫属，锥虫属，弓形体和疟原虫。

25 该核酸构建体编码的多肽可以包括TAA的看家表位。在优选实施方案中，该核酸构建体编码复合看家表位。当该构建体编码这样的复合表位时，所有的复合表位都可以与单个TAA的不同片断对应，或者他们可以与不同TAAs相对应。在优选实施方案中，该核酸构建体含有看家表位和免疫表位。在另一种优选实施方案中，该核酸构建体含有看家表位和

表位簇区域。

在疫苗构建体同时编码看家和免疫表位的优选实施方案中，疫苗可以刺激针对呈递任何一种表位的靶细胞的免疫反应，一即，免疫反应可以识别最初呈现在靶细胞上的看家表位，并且随后还可以在IFN诱导后，
5 识别靶细胞上呈现的免疫表位。

有利地，该核酸构建体还可以包括第三或者第四或者更多序列，这些序列分别编码第三或者第四或者附加表位。这些表位可以衍生于单个TAA或者衍生于两个或者更多不同TAAs，而且可以是看家或者免疫表位的任何组合。可以通过设计这些构建体来使之编码相应于任何蛋白酶体活性的表位，所述的蛋白酶体活性可以在任何靶细胞或pAPC的抗原加工
10 中起作用。

编码的MHC表位的长度优选7—15个氨基酸，更优选的长度为大约9或10个氨基酸。虽然通常与MHC I结合的肽的大小优选9个氨基酸，短一些或者长一些的肽也可以在一些情况下结合MHC I。同样，许多比9个氨基酸长得多的肽可以被细胞中固有的外肽酶或者其它蛋白酶所修饰，以
15 产生能够非常有效地结合MHC I的片段。含有免疫表位的肽的大小并不重要，只要序列包含此表位。这是因为pAPCs固有的免疫蛋白酶体，与外肽酶和其它蛋白酶协同修饰，以其正常的功能正确地加工全长TAAs以产生免疫表位。因此，编码免疫表位的核酸序列实际上可以编码大得多的前体，包括整个TAA。这种构建体优选也可以编码看家表位。
20

适用于本发明的TuAAs和其它TAAs的实例包括但不限于：分化抗原例如MelanA (MART-I), gp 100 (Pmel 17), 酪氨酸酶, TRP- 1, TRP-2 , 和肿瘤特异的多系抗原例如MAGE- 1, MAGE-3, BAGE, GAGE-1, GAGE-2, CEA, RAGE, NY-ESO, SCP-1, Hom/Mel-40, 和PRAME。类似地，TuAAs
25 包括过表达的癌基因，和突变的肿瘤抑制基因例如 p53, H-Ras, HER-2/neu。此外，染色体易位产生的独特TuAAs例如 BCR-ABL, E2A-PRL, H4-RET, IGH-IGK, MYL-RAR, 和病毒抗原例如EB病毒抗原EBNA和人乳头状瘤病毒(HPV)抗原E6和E7。其它有用的蛋白抗原包括但是不仅仅限于 TSP-180, MAGE-4, MAGE-5, MAGE-6, RAGE, NY-ESO, pl85erbB2,
30 pl80erbB-3, c-met, nm-23H1, PSA, TAG-72-4, CAM 17.1, NuMa, K-ras, β -连

环蛋白, CDK4, Mum-1, p16。对于本领域的技术人员, 这些和其它的TuAAs和病原相关抗原可以通过文献获得或者商购。

在另一实施方案种, TAA是病毒特异的抗原。见上面的的表2。在本发明的另一种实施方案中, TAA是对非病毒细胞内寄生物特异的抗原。
5 寄生物特异的抗原的实例包括核苷酸, 蛋白, 和其它与细胞内寄生物相关的基因产物。合适的核苷酸或者蛋白可以在NCBI分类数据库中找到, 网址是<http://www.ncbi.nlm.nih.gov/Taxonomy/tax.html/>。此网址提供了对寄生物和其它病原体基因产物的更为详细的描述。

特别优选的肽的长度大约是7-15个氨基酸。在Han-Georg Rammensee,
10 Jutta Bachmann, 和 Stefan Stevanovic, 的“MHC Ligands and Peptide Motifs,” Springer-Verlag, Germany, (1997) Landes Bioscience, Austin, Texas. 中, 提供了含有MHC结合基元的肽的广泛列表。

被该构建体编码的表位对一种或多种MHC I等位基因具有亲和性。在一些实施方案中(这些实施方案中的患者是对MHC I杂合的), 该构建体可以编码对应于不同MHC I等位基因的表位。
15

优选的核酸构建体包括至少一个有效连接到该构建体编码区5'末端的启动子序列。本领域的技术人员能够理解到可使用在哺乳动物细胞中有活性的任何启动子。优选的启动子包括但不限于, CMV启动子, SV40启动子和逆转录病毒LTR启动子序列, 并且还可以包括EF-1A, UbC, β -肌动蛋白启动子。在一些实施方案中, 该构建体可以包括两个或者更多的有效连接到不同多肽编码序列5'末端的启动子。同样, 该构建体还可以使用增强子, 核进入序列(nuclear import sequences), 免疫刺激序列(immunostimulatory sequences), 和细胞因子、选择标记、报告分子等的表达盒, 等等。此外, 可以通过稳定杂交的PNA肽核酸, 将免疫刺激
20 序列和其它调控序列加到载体上。在优选实施方案中, 本发明的核酸构建体还包括有效连接到编码区3'末端的聚合A序列。图9中描述了包括核进入序列和免疫刺激序列的核酸构建体。

在某些实施方案中, 该核酸构建体编码mRNA, 该mRNA被翻译成单个多肽, 而后被切割。在这样的一个实施方案中, 多肽由表位线性排列组成, 其中第一(N末端)序列为一个或多个免疫表位或者表位簇,
30

而且第二（C末端）序列为看家表位，这样，正确的看家表位的C末端被终止密码确定，而且所有其它HLA表位末端被蛋白酶体加工和外肽酶修饰所确定。

在另一种优选实施方案中，该核酸构建体编码氨基酸序列，其中免疫表位或者表位簇被连接到泛素序列。泛素序列被类似地连接到看家表位。表位之间存在的泛素，促进了将免疫表位有效传递给蛋白酶体以进行表位加工。泛素序列（带有或者不带有N末端间隔物以保证前述肽的完整）位于第一和第二序列之间的结构中，或者位于任何其它表位编码序列之间的结构中。这样产生的序列1—泛素—序列2多肽在泛素—序列2连接处被泛素特异的加工蛋白酶快速（共翻译）切割，产生序列1—泛素和序列2。（见图10）

生理学上，泛素主要作为一种信号，指导蛋白酶降解蛋白。它是真核细胞中最为保守的蛋白之一，在酵母和人类之间只有三个保守的氨基酸替代。尽管泛素的精确序列可能有些变化，但是优选实施方案的序列用SEQ ID NO: 5代表。76个氨基酸长度的泛素具有两个重要特征：1、C末端的甘氨酸残基，涉及蛋白底物的赖氨酸支链与泛素接合，和 2、赖氨酸残基，在48位，用来形成多—泛素链。

所有的泛素基因都是作为其它多肽的（包括其它泛素）融和体被合成的，从这种意义上说，泛素基因是独一无二的。在酵母*S. Cerevisiae*中，鉴别出了4中泛素基因：前三种（UBI1-3）融和到核糖体蛋白，第四种基因（UBI4）作为泛素序列的5个相同重复的融和体被合成。这样，通过到处表达的泛素特异性水解酶，在共翻译的蛋白水解加工后，天然地产生了无功能的泛素。这样的天然有机体已经通过在单个泛素部分和任何需要的多肽之间产生C末端融和体被利用。

泛素可以以两种构象存在：上述的第一种构象由单个泛素与任何所需多肽的线性融和组成，其中泛素的C末端甘氨酸，通过肽键，被连接到选择的多肽的N末端氨基酸。第二种构象涉及通过形成甘氨酸—赖氨酸键，泛素部分与蛋白底物的接合。在这种情况下，泛素中赖氨酸的羧基被连接到暴露于溶剂的底物（或者另一个泛素部分）赖氨酸的 ϵ (epsilon)支链。降解底物的泛素信号与第二种构象相关。这样，在上述的序列1—

泛素一序列2构建体中，典型地，蛋白酶体不靶向序列2。因此，序列2的位置被优选用于完全加工的表位，或者只需要N末端修饰的表位，典型为看家表位。仍然连接在上述构建体中序列1上的泛素部分，可以在赖氨酸48位被聚遍在蛋白化（polyubiquitinated），因此片段被靶向蛋白酶体得以加工，并且导致序列1中含有的表位被释放。应该注意到，如果超过两个以上的序列被泛素部分线性连接到一起，那么通常只有最后的序列以序列2的方式工作；所有上游序列的加工类似于序列1的加工。在这个意义上，此处描述的构建体在pAPCs中表达，其中免疫蛋白酶体是主要活跃的，通过这些构建体对看家表位的正确表达，从序列2位置中的看家表位中获益，或者从相对应位置中的看家表位中获益，其中所述的相对应位置中的看家表位不需要pAPC中的蛋白酶体加工。

在另一种实施方案中，本发明的核酸构建体可以含有编码自动蛋白水解肽的序列。这些序列处在第一序列和第二序列之间或者在任何其它表位编码序列之间。这些自动蛋白水解序列的实例包括inteins；还包括细小核糖核酸病毒，包括脊髓灰质炎病毒和其它肠道病毒，鼻病毒，心肌病毒和aphoviruse的3C^{Pro}和2A^{Pro}蛋白酶，和equivalent cornoviridae 蛋白酶。这些蛋白酶催化对该病毒家族制造的大前体多蛋白进行的翻译后切割。

在一种实施方案中，自身催化蛋白序列被插入到两个或多个表位之间。在另一实施方案中，该序列被插入到两个或多个表位后，但是切割信号在表位之间被发现，这样，它们被切割成两个或者多个完全功能的表位。蛋白酶的类型并不重要，唯一重要的是：合适的切割信号对于正确的表位加工应该是可用的。

因为切割位点和自身催化蛋白的序列是已知的(近来由Seipelt, J.等综述：Virus Research 62:159-168, 1999)，所以它们可以很容易地用来构建产生多蛋白和双蛋白的载体。简要的说，尽管其它非常临近的位置也可能是重要的，但是3C^{Pro}主要将Q-G位点识别为切割信号。而且，一些这些病毒的3C^{Pro}并不完全遵守这种常用模式，这就给设计上提供了较大的灵活性。被这些需要强加的这种限制比实际更正式，特别是如果蛋白酶被置于表位之间表达的时候。在这种排列中，即使病毒蛋白酶没能够

切割其N-末端，上游免疫表位仍可以被蛋白酶体加工释放。用于C末端切割的主要残基位于3C^{Pro}自身的内侧，如果有的话，通常留下正好1—4个残基，被外肽酶修饰从看家表位下游的N末端去除。2A^{Pro}可以以同样方式使用，只是需要理解的是，虽然切割位点倾向于G-P，但是在这些病毒中有些更加易变。必须还要考虑：2A^{Pro}的表达可以导致宿主细胞的蛋白合成迅速，完全地关闭，这取决于它所衍生的病毒株。

严格的讲，来自心肌病毒和aphoviruses (即：口蹄疫病毒(FMDV))的2A蛋白不是蛋白酶，更确切的说，它在不导致翻译终止的情况下，防止肽键在其C末端形成 (Ryan, M.D., 等, Bioorganic Chemistry 27:55-79, 1999)。因此，通过将这些2A蛋白放置在表位之间，可以导致在单一读码框内发生切断。来自FMDV的2A蛋白是很小的，只有18个氨基酸，这使得它特别适合于复合表位表达。在图11中描述了使用2A蛋白的质粒。

在某些其它实施方案中，核酸构建体编码被翻译成2个或多个肽段的mRNA。在一种这样的实施方案中，转录产物可以含有一个或多个内部核糖体进入位点 (IRES) 序列，这些序列位于第一和第二序列之间或者任何其它表位编码序列之间。IRES序列被细小核糖核酸病毒天然地使用，用来指导mRNA的内部帽一独立翻译。这些IRES序列还可以允许独立翻译两个或者更多来自相同mRNA的连续开放阅读框。尽管不同构建体的IRES序列可能不同，在SEQ ID NO: 6中提供了一种优选实施方案的IRES序列。每个表达表位的C末端由终止密码决定。因此，编码看家表位和编码免疫表位的序列顺序并不重要，这为质粒构建提供了灵活性。任选地，编码看家表位的序列可以在IRES序列之前，而且编码免疫表位的序列可以连接于另一个IRES序列的末端。这样的载体还可以有用地编码两个或更多看家表位。它们还可以允许各种上述的单个多肽构建体相互结合，以有效地表达复合表位。见图12。

在某些其它实施方案中，核酸构建体编码两种或更多mRNA转录产物。每种转录产物可以编码单个表位或者上述实施方案中描述的任何双或者复合表位转录产物。两种或者多种转录产物可能是使用复合启动子的结果。本领域的技术人员很清楚，使用超过一个拷贝的单个启动子会导致质粒在繁殖中不稳定。因此通常优选使用两种（或更多）不同的启

动子。

两种或者多种转录产物也可能是使用双向启动子的结果。双向启动子可以在很多种类的有机物中发现。这些启动子的实例包括来自人的PDGF-A，来自产黄青霉的pcbAB和pcbC，嗜神经性JC病毒，来自小鼠，
5 狗和人的BRCA1。尽管对双向启动子的深入研究最近才开展，但是有关它们的序列、调控方面的信息，和其它有关它们是如何工作方面的错综复杂的信息越来越多。例如，具有全部转录活性的人转钴胺素 II启动子需要含有GC盒和E盒的69个碱基对（bp）的片段。二肽酰基酶IV启动子可以从两边以同样的效率激活转录。大鼠线粒体RNA伴侣蛋白60和10头
10 与头相连并且公用一个双向启动子。因而，已经以这样一种方式鉴别、测序并且克隆出了多种起作用的双向启动子，该方式可以使它们被用于核酸构建体来表达两个基因。

因此，在优选实施方案中，该核酸构建体含有双向启动子（例如，上面列出的那些），这些双向启动子被连接到编码看家表位或者其前体的核酸序列上。在特别优选的实施方案中，核酸构建体含有双向启动子，
15 这些双向启动子被连接到编码多个看家表位的核酸序列上。在另一种实施方案中，核酸构建体含有双向启动子，这些双向启动子被连接到编码看家表位和免疫表位的核酸序列上，或者被连接到表位簇区域的核酸序列上。此外，双向启动子可以被正调控或者负调控。

20 当核酸构建体含有超过一个表位的时候，双向启动子可以以相近的数量表达多个表位，或者一些表位可以比另一些表位的表达水平高。作为选择，一种表位是可诱导的而另一种是组成型的。这样，可以实现表位的暂时调控，其中一种表位在治疗中表达的早而另一种表达的晚。

25 表位表达分析

文献中描述的几种方法可以用来确定表位是否呈现在pAPCs中。一种不直接但是功效大的方法是使用I型四聚物分析来确定T细胞在给予看家表位前和给予看家表位后，在动物中的出现频度。响应一种表位的T细胞克隆性扩增，只有在表位被pAPCs呈递给T细胞的时候，才优先地发生。
30 因此，在对动物给予看家表位前和给予看家表位后，对特异性T细胞相对

于看家表位的出现频度进行测量，是一种确定表位是否呈现于pAPC的方法。给予表位后，对表位特异的T细胞的出现频度增加，这说明表位被呈现到pAPCs上。其它确定T细胞出现频率的方法例如限制稀释分析或ELISPOT可以以相同的方式评价pAPCs呈递的看家表位。类似地，可以使用任何确定动物中T细胞出现频度的方法。

确定呈递在pAPCs上的看家表位的直接方法涉及将pAPCs从给予了表位的动物中纯化出来。在用看家表位给动物接种以后，通过使用指向呈现在pAPCs上的特异标记的单克隆抗体和亲和性纯化，例如使用固定到磁珠上的单克隆抗体，可以从周围血单核细胞，脾细胞，淋巴细胞中收获pAPCs。这种收获的最佳时间是不定的，可以取决于被接种的动物，疫苗的本性，和其它包括剂量，给予位点，药代动力学等因素。可以使用这种技术来富集天然血和脾细胞制剂中的pAPCs。随后可以将富集的pAPCs用于指向被产生的T细胞克隆的增殖分析，其中的T细胞克隆对感兴趣的看家表位有特异性。将pAPCs与T细胞克隆共同温浴并且监测T细胞的增殖活性，例如通过测量T细胞整合的放射性胸腺嘧啶。增殖现象说明pAPCs上的看家表位刺激了对看家表位特异的T细胞。

下面实施例的意图只是为了阐明目的，而不能理解为以任何方式限制了本发明的范围。

20 实施例

实施例1.作为看家或者免疫表位的HLA表位的蛋白水解特性

使用下面描述的步骤，制备了13个氨基酸或者更多的、在中心含有候选HLA表位的合成肽。蛋白酶体从表达各自类型蛋白酶体的细胞中制备，例如红细胞和Raji细胞分别用来制备看家蛋白酶体和免疫蛋白酶体。用蛋白酶体制剂消化肽，并且用质谱分析鉴定所得到的片段。如果在那些片段中有一个片段与HLA表位共C末端，并且由含有看家蛋白酶体的制剂以有意义的产量产生，那么该HLA表位就是看家表位。类似地，如果那些片段中有一个片段与HLA表位共C末端，并且由含有免疫蛋白酶体的制剂以有意义的产量产生，那么该HLA表位就是免疫表位。

30 A 肽的合成

构建合成多肽或者重组多肽，在这些多肽中包含有HLA表位和至少两个临近其末端的残基。这些被加到特定HLA表位末端的残基，为的是要保证蛋白酶体复合物处在与细胞内类似的加工环境中，从而提高了其正常执行其蛋白水解功能的可能性。如果必要，可以增加通常在临近HLA表位末端发现的额外的残基，以增加肽的溶解性。

由于其高疏水性，一些HLA表位呈现难溶性。一些肽特别难以纯化，因为它们不溶于正常的层析洗脱液，或者由于它们不溶于消化缓冲液，因而一旦被纯化，很难使用。可以通过仔细选择HLA表位周围序列的哪一部分被包含在特定肽构建体中来避免这一问题，或者通过上一自然段提到的延长该序列来避免这一问题。如果HLA表位的临近末端没有可以帮助增加溶解性的残基存在，可以用一个短的亲水性序列来替代（例如—EAEAE）。这段序列要添加在HLA表位末端后的至少2—5个残基，以保持蛋白酶体的天然末端切割位点。

在优选实施方案中，使用标准Fmoc固相合成方法，在应用生物系统433A肽合成器上合成肽。合成器装备有传导回馈监测系统（conductivity feedback monitoring system），该合成器可以增加序列的残基延伸的反应时间，所述的残基含有难于去保护和/或难于连接。合成了肽以后，使用三氟乙酸，在适当的清除剂存在的情况下，将肽从其支撑物上分离下来，用醚沉淀，而后冻干。

在首先用类似分析联苯HPLC系统（similar analytical diphenyl HPLC system）形成一个梯度后，将粗肽在制备的联苯HPLC柱（preparative diphenyl HPLC column）上纯化。用电喷射质谱测量法分析来自肽第一制备注射液的主要HPLC部分，以鉴定靶化合物。收集，集中并冻干来自后来注射液的相应的峰，用分析HPLC分析样品的驻留时间和层析纯度。这些被纯化的肽为蛋白酶消化做好了准备。

B、蛋白酶测定

在下面的实施例2中，详细描述了免疫或者看家蛋白酶体复合物的分离。

将纯化的肽溶于适当的缓冲液中，使浓度大约为1mM，而后加到大约2体积的蛋白酶体制剂中。制备复制消化液：一份用于光谱测定分析，

一份用于HPLC分析，使用阳性对照肽来制备额外的消化液，来证实所使用蛋白酶制剂的正常功能。下列肽适于被用作免疫蛋白酶体鉴定的对照肽：

MLLAVLYCLLSFQTS (SEQ ID NO: 7);
 HSYTTAEAAAGITILTVILGVL (SEQ ID NO: 8); EAASSSTLVEVTLGEVPAAESPD (SEQ ID
 NO: 9); EFLWGPRALVETSYVKVLHHMVKI (SEQ ID NO: 10); APEEKIWEELSVLEVFEGR
 (SEQ ID NO: 11); 和 ELMEVDPIGHLYIFAT (SEQ ID NO: 12).

带有下列

线的残基表示蛋白水解切割位点。肽FLWGPRALVETSYVK (SEQ ID NO: 13) 适于用作看家蛋白酶体鉴定的对照肽。将消化液在37°C下平行温育一段时间，而后加入三氟乙酸稀释液停止消化，在干冰上冻结样品。用
 Lasermat 2000(Finnigan Mat, LID, U.K.)基质辅助的激光解吸—电离飞行时间(MALDI-TOF)质谱仪分析一种复制消化液和阳性对照，其它的消化液留给HPLC。

C. 消化液的MALDI-TOF质谱分析

或者使用“肽”软件(Lighthouse Data)，或者可以从ThermoBioanalysis Ltd., U.K.得到的软件来对消化液进行分析。此软件可以产生所有可能片段的序列和分子量，这些可能片段能够同时满足：具有任何预测表位的正确C末端；而且含有该表位的全长序列或者更长的序列。

例如，如果HLA表位包含的肽是下面的序列：

AAMLLAVLYCLLSEIIAAAEEE,

其中带下划线的序列是HLA表位，那么该程序将所有下面的序列鉴定为可能有用的序列，而且会给每个序列指定分子量。

AAMLLAVLYCLLSEI

AMLLAVLYCLLSEI

MLLAVLYCLLSEI

LLAVLYCLLSEI

LAVLYCLLSEI

AVLYCLLSEI

如果MALDI-TOF 的结果显示一种或多种这些分子量呈现在消化混合物中，那么就将相应的肽合成，纯化，用质谱仪鉴定，而后将其用于分析HPLC以同时建立标准驻留时间和质量与峰面积的近似比率。随后，使用相同的分析HPLC方法将保留的消化液稀释在适当的溶剂中并且注入。如果消化液产生了驻留时间与标准驻留时间相同的好的产量峰，那么基本上可以确定这是由于消化液中的存在的那个序列导致的。如果出现了由其他可能产生片段所引起的模糊（这些片段会产生相同或者类似质谱测量结果），那么可以收集所怀疑的组分并且将它们用于C末端测序以确定同一性。

10 实施例2.蛋白酶体复合物的纯化。

A.来自血细胞的蛋白酶体

浓缩的红血球袋来自地方血库(HemaCare, Van Nuys, CA)。将每个袋子里的物质倒入200ml离心管中，用PBS冲洗3遍，通过下列方式冲洗：在Megafuge2.0水平转头 (Heraeus, Southplainfield, NJ) 中，以2000RPM在室温下离心10分钟。在最后一遍冲洗后，样品被集中到一个容器中，以使试管中的可变形降低到最小，而后将其重新分到几个离心管中。细胞被以2000RPM的转速再次离心10分钟。吸出剩余的PBS。将片状沉淀物储存在-70°C直到使用。

B. 来自肿瘤细胞的蛋白酶体复合物

20 Raji细胞（一种Burkitt淋巴瘤细胞系），是从ATCC（美国典型培养物保藏中心，Manassas, VA）得到的。该细胞使用标准的细胞培养方法培养，而且用干扰素- γ (100-500U/ml)刺激(Pharmingen, San Diego, CA)。通过在培养物上使用免疫组织化学方法和通过对细胞裂解物样品的SDS-PAGE，分别确证免疫蛋白酶体亚单位的表达。离心收集细胞，用PBS清洗，储存在-70°C直到使用。

C. 蛋白酶体复合物的进一步加工

将血细胞或者淋巴瘤细胞的片状沉淀（冰冻的）在37°C水浴融化，在每个试管中加入双蒸水。将细胞悬液在40ml邓氏匀浆器中匀浆。此外，对于肿瘤细胞，匀浆的细胞在2000rpm下离心以去除细胞碎片。将上清液在40°C，10,000 rpm的转速下离心10分钟，而后进一步在T-1270转子

(Sorval, Newtown, CT)中, 40°C条件下以50,000 rpm的转速离心30分钟。

将匀浆通过滤膜以去除碎片, 而后集中在一起。在集中的匀浆样品中加入68%的蔗糖溶液。室温下, 将抗体-琼脂糖凝胶制剂与匀浆在转子中培育3个小时。将该悬液离心并用TBS清洗3遍, 进一步在真空漏斗中清洗超过6-8遍。在PBS(pH 7.6)中洗脱蛋白酶体, 而且测定洗脱液的光密度。在40°C, 使用纤维素膜MWCO 1000, 在20mM的Tris溶液(pH7.6)中将蛋白酶体制剂透析过夜。第二天, 通过微孔ULTRAFREE-I离心装置(Millipore, Danbury, CT)的超滤作用, 浓缩蛋白酶体制剂。而后将浓度为4mg/ml的蛋白酶体等分储存在-20°C直到使用。通过对荧光底物或者对照肽(生产已知片段的对照肽)的消化作用, 来测试蛋白酶体的活性和特异性。下面的肽适合作为用于免疫蛋白酶体测定的对照肽:

MLLAVLYCLLWSFQTS (SEQ ID NO: 14); HSYTTAEEAAGITILTVLGVL (SEQ ID NO: 15);
EAASSSTLVEVTLGEVPAAESPD (SEQ ID NO: 16); EFLWGPRALVETSYVKVLHHMVKI
(SEQ ID NO: 17); APEEKIWEELSVLEVFEGR (SEQ ID NO: 18); 和 ELMEVDPIGHLYIFAT
(SEQ ID NO: 19).

带有下划线的残基表示蛋白切割位点。肽FLWGPRALVETSYVK (SEQ ID NO: 20) 适于用作看家蛋白酶体鉴定的对照肽。

D.蛋白酶制剂的定量和活性分析

酶联免疫反应(ELISA)被用来给上述的蛋白酶制剂定量。ELISA技术在本领域中被人熟知, 而且在Ausubel, 等的“分子生物学简短方案(Short Protocols in Molecular Biology)” 3rd Ed., Unit 11.2 (1997)中被概论。产生单克隆抗人蛋白酶体抗体的杂交瘤细胞(MCP-21), 是从欧洲细胞培养收藏中心((ECACC), UK)获得的, 并且该细胞通过使用标准细胞培养技术和设备保存。杂交瘤添加物(Gibco BRL, Rockville, MD)被加入到抗体产生细胞中。在大约2-3公升体积中, 当细胞密度达到500,000细胞/ml的时候, 通过离心作用去除细胞并收集上清液。通过使用Lambda 20 分光光度计(Perkin Elmer, Norwalk, CT)来测定光密度(O.D.), 阶段性地监测培养基中mAb的分泌。

将上清液通过蛋白G琼脂糖凝胶柱(Amersham/Pharincia Biotech Piscataway, NJ)。用PBS清洗柱子, 抗体被洗脱在pH 2.2的0.1M甘氨酸缓

冲液中。在280nm处测定洗脱部分的光密度，收集阳性部分。在4°C，将抗体在2升的PBS中透析2天并且储存直到使用。

将抗体结合到CNBr-活化的琼脂糖凝胶4B (Amersham Pharmacia biotech, Piscataway, NY)上。用pH=4的0.1M醋酸钠盐溶液, 和 pH=8的
5 0.1M硼酸钠盐溶液交替清洗5—7遍, 最后悬浮于pH=8的Tris缓冲盐溶液 (TBS) 中。该制剂在4°C储存直到使用。

实施例3. 使用算法模型产生预测的MHC I肽裂隙结合肽

通过使用运算法则, 产生了大量的候选MHC I结合肽, 这些结合肽产生自人癌胚细胞抗原前体 (CEA) 的氨基酸序列 (GENBANK
10 ACCESSION P06731)。如上所述, 这种特殊的算法可以在
<<<http://134.2.96.221/scripts/hlaserver.dll/EpPredict.htm>>>中得到。一旦得到了算法, 就可以提供CEA的氨基酸序列。接下来, 选择感兴趣的表位长度 (十体) 和特定的MHC等位基因 (H2-Db) 的参数。随后, 提交数据来进行算法分析。结果见表7。

15

表7
对H2—Db有预测亲和性的CEA片段

	POS	1 2 3 4 5 6 7 8 9 0	得分	序列ID号		POS	1 2 3 4 5 6 7 8 9 0	得分	序列ID号
5	547	L Q L S N G N R T L	26	21		496	V S A E L P K P S I	16	51
	369	L Q L S N D N R T L	26	22		465	S N I T E K N S G L	16	52
	191	L Q L S N G N R T L	26	23		453	G N I Q Q H T Q E L	16	53
	53	L L V H N - L P Q H L	26	24		441	S N P P A Q Y S W L	16	54
	371	L S N D N R T L T L	25	25		393	C G I Q N E L S V D	16	55
10	549	L S N G N R T L T L	24	26		242	I S P L N T S Y R S	16	56
	193	L S N G N R T L T L	24	27		91	P G P A Y S G R E I	16	57
	299	C Q A H N S D T G L	23	28		43	F N V A E G K E V L	16	58
	100	I I Y P N A S L L I	21	29		693	I G V L V G V A L I	15	59
	578	S A N R S D P V T L	19	30		684	S A G A T V G I M I	15	60
15	576	S V S A N R S D P V	19	31		510	S K P V E D K D A V	15	61
	504	S I S S N N S K P V	18	32		482	S A S G H S R T T V	15	62
	356	W W V N N Q S L P V	18	33		428	R P G V N L S L S C	15	63
	178	W W V N N Q S L P V	18	34		399	L S V D H S D P V I	15	64
	148	S I S S N N S K P V	18	35		372	S N D N R T L T L L	15	65
20	127	S D L V N E E A T G	18	36		332	S N P V E D E D A V	15	66
	645	I T P N N N G T Y A	17	37		329	S N N S N P V E D E	15	67
	540	S L P V S P R L Q L	17	38		307	G L N R T T V T T I	15	68
	362	S L P V S P R L Q L	17	39		289	I T V N N S G S Y T	15	69
	326	F I T S N N S N P V	17	40		280	S T Q E L F I P N I	15	70
25	250	R S G E N L N L S C	17	41		277	F Q Q S T Q E L F I	15	71
	184	S L P V S P R L Q L	17	42		222	S A R R S D S V I L	15	72
	140	V Y P E L P K P S I	17	43		221	V S A R R S D S V I	15	73
	40	S T P F N V A E G K	17	44		154	S K P V E D K D A V	15	74
	29	N P P T T A K L T I	17	45		135	T G Q F R V Y P E L	15	75
30	655	C F V S N L A T G R	16	46		49	K E V L L L V H N L	15	76
	608	G A N L N L S C H S	16	47		34	A K L T I E S T P F	15	77
	606	L S G A N L N L S C	16	48					
	604	S Y L S G A N L N L	16	49					
	571	C G I Q N S V S A N	16	50					

上面的表任意去除小于15的得分。该算法可以产生小于15的得分。

实施例4. 使用免疫和看家蛋白酶体来消化肽前体以确定由蛋白水解消化产生的片段

25 使用433A ABI合成器合成肽。使用Fastmoc化学，以0.25m摩尔的量产生肽。测定这些肽的溶解性，一旦溶解，制备2mM的溶液并且等分成25—30μl，—20°C储存用于以后使用。进行定时的消化反应（典型地，由2μl的肽和4μl的蛋白酶体组成），用t=0作为对照，而且用水代替蛋白酶体与肽培育作为进一步对照。反应在37°C进行，通过在干冰上加入10
30 %的TFA（三氟乙酸）来终止反应。而后用下面实施例5中描述的MALDI

—TOF质谱法(MS)分析冰冻的样品。

在MS分析前,通过使用ZIP—TIP方法(Millipore, Boston, MA),可以任选地对消化液执行除盐步骤。ZIP—TIP是一种特殊设计的、含有硅球树脂床的吸管尖。样品结合到用0.1%TFA预平衡过的尖上,而后,用50
5 %乙腈1%TEA洗脱缓冲液洗脱。

实施例5.用HPLC和质谱测量法对相关蛋白水解片段进行鉴定和定量

A 治疗学上感兴趣序列的鉴定

将感兴趣的蛋白的氨基酸序列输入计算机,使用Rammensee等的算法来产生9或10个氨基酸长度的、预测结合特殊HLA受体的序列。该算法
10 还根据这些预测表位与结合基元的匹配程度将其排序。

而后构建合成肽(该合成肽含有被鉴定的可能表位的序列)使之包含表位候选序列和临近其末端的至少3—5个残基。这些加到特定HLA表位末端的残基是要保证蛋白酶体复合物处在与细胞内类似的加工环境中,从而提高了其正常执行其蛋白水解功能的可能性。如果必要,可以
15 增加通常在临近HLA表位末端发现的额外的残基,以增加肽的溶解性。

使用标准Fmoc固相合成方法,在应用生物系统433A肽合成器(应用生物系统, Norwalk, CT)上合成肽。合成器装备有传导回馈监测系统(conductivity feedback monitoring system),该合成器可以增加序列的残基延伸的反应时间,所述的残基含有难于去保护和/或难于连接。合成了
20 肽以后,使用三氟乙酸,在适当的清除剂存在的情况下,将肽从其支撑物上分离下来,用醚沉淀,而后冻干。

而后将粗肽以0.5mg/ml的浓度溶解到适合的溶剂中。而后,使用0.1%TFA水—乙腈梯度在Shimadzu分析反相HPLC系统(Shimadzu 科学仪器哥伦比亚, MD)上对5微升(5 μ l)该溶液进行分析。典型地, C—18硅柱
25 (Machery-Nagel # 720051.40, (Machery-Nagel GmbH, Germany))用于亲水性肽,苯基硅柱(Vydac # 219TP5415 (The Separations Group, Inc., Hesperia, CA))用于疏水性肽。所使用的梯度变化范围为从亲水性肽的0—40%乙腈,到疏水性肽的30—70%乙腈。随后,使用类似的梯度和上述柱子(Machery Nagel #715802,和 Vydac 219TP5 10)的半制备形式(semi-preparative versions),将肽在Varian Prostar HPLC系统(Varian, Inc., Palo
30

Alto, CA)中纯化。使用MALDI-TOF质谱仪来分析来自肽第一制备注射液的主要HPLC部分,以鉴定需要的组分。收集,集中并冻干来自后来注射液的相应的峰,用使用上述系统的分析HPLC来分析样品的驻留时间和层析纯度。这些被纯化的肽为蛋白酶消化做好了准备。

5 B.蛋白酶体测定

通过上述的Levy方法(Morel, S., 等,免疫12:107-117 (2000), 和此处引用的文献)分离免疫或看家蛋白酶体复合物。将纯化的肽溶于适当的缓冲液中,使浓度大约为1-2mM,而后加到大约2体积的蛋白酶体制剂中。所选择的缓冲液必须使肽溶剂化而并不干扰消化过程。如前所述,制备
10 额外的消化液用来作阳性对照肽,以证实所使用蛋白酶制剂的正常功能。将它们在37°C下平行温育至120分钟,而后加入三氟乙酸稀释液停止消化;立刻用质谱仪分析样品,或者将它们在干冰上冻结直至用于分析。也可以通过将样品置于冰上来中断消化反应,用质谱仪立即分析。

C. 消化液的MALDI-TOF质谱分析

15 在样品载玻片上,将大约0.5 μ l的每种消化液与等体积的基质溶液(70% EtOH中有10mg/ml的二羟基苯酸, pH=2-3)直接混合,在大约40°C使其风干。而后将样品在被适合分子量标准校准了的Lasermat™ MALDI-TOF质谱仪(Thermo Bioanalysis, Santa Fe, NM)上分析。

用于蛋白酶体分析的计算机程序(或者是“肽”软件(Lighthouse
20 Data),或者是“Dynamo”(ThermoBioanalysis Ltd., U.K.))产生了所有可能片段的序列和分子量,这些可能片段能够同时满足:具有任何预测表位的正确C末端;含有该表位的全长序列或者更长的序列。

如果MALDI-TOF 的结果显示一种或多种这些分子量呈现在消化混合物中,那么就将相应的肽合成,纯化,用MALDI-TOF鉴定,而后将其
25 用于分析HPLC以同时建立标准驻留时间和质量与峰面积的近似比率。这些步骤与上述那些步骤很类似。而后,将复制的蛋白酶体消化液稀释在适当的溶剂中并且使用相同的分析HPLC方法进行分析。如果消化液产生了驻留时间与标准驻留时间相同的好的产量峰,那么基本上可以确定这是由于消化液中的存在的那个序列导致的。如果出现了由其他可能产生
30 片段所引起的模糊(这些片段会产生相同或者类似质谱测量结果),那

么可以收集所怀疑的组分并且将它们用于C末端测序以确定同一性。分析HPLC同样重要地提供了消化液中肽产物的相对准确的定量，这就可以确定某一特定肽究竟是消化的主要产物还是次要产物，这显示了表位是否可以由蛋白酶体有效地生产。使用上述的方法，看家表位被鉴别出来。

5 图13显示了流式细胞分析的结果，来证实这些表位与HLA结合。该测定的结果在实施例6中讨论。

实施例6.被选择肽的MHC结合能力的确定

依照Stauss等的方法(Proc Natl Acad Sci USA 89(17):7871-5 (1992))分析候选肽对HLA-A2.1的结合。将T2细胞（该细胞在其表面表达空的或者不稳定的MHC分子）清洗2遍并以 5×10^6 细胞/ml的浓度悬浮于无血清的完全Iscove's modified Dulbecco's 培养基 (IMDM)中。加入 β_2 微球蛋白 (Sigma, St. Louis, MO)使浓度为 $5 \mu\text{g/ml}$ ，将细胞分布到96孔U型底板中使细胞浓度为 5×10^5 细胞/孔。将肽以100，10，1和 $0.1 \mu\text{g/ml}$ 的浓度加入。将板轻柔摇动2分钟，而后在 37°C ，5% CO_2 的培养箱内培养4个小时。用

15 IMDM清洗两遍以去除未结合的肽，而后加入饱和数量的单克隆抗体W6/32 (Sigma)。在 4°C 培养30分钟后，用添加了2%热失活FCS，0.1%(w:v)叠氮化钠，pH7.4-7.6（着色缓冲液）的PBS清洗细胞，在 4°C 下与结合了异硫氰酸荧光素 (FITC) 的羊F(ab')抗鼠免疫球蛋白G(Sigma)一起培养30分钟，而后象以前那样清洗4遍。将细胞重悬于着色的缓冲液并且

20 通过加入四分之一体积的2%仲甲醛来使之固定。使用FACScan (Becton Dickinson, San Jose, CA)，对被肽结合所稳定的表面HLA-A2.1分子进行流式细胞分析。

实验结果见图14。使用上述讨论的方法，发现：由蛋白酶体消化所鉴定的候选酪氨酸酶看家表位(酪氨酸酶 207-216, FLPWHRLFLL SEQ ID NO: 78)与HLA-A2.1结合的程度，与已知A2.1结合物FLPSDYFPSV (SEQ ID NO: 79)（阳性对照）的结合程度类似。将HLA-B44 结合肽AEMGKYSFY (SEQ ID NO: 80)作为阴性对照。从阴性对照获得的荧光，与在分析中没有使用肽时所获得的信号类似。阳性对照肽与阴性对照肽选自通用免疫学方案(Current Protocols in Immunology) p.18.3.2, John Wiley

30 和Sons, 纽约, 1998. 中的表18.3.1。

实施例7.肿瘤，组织样品，永生化细胞系，或肿瘤细胞系中的HLA表位的洗脱

并不是通过体外蛋白水解来产生HLA表位，而是可以通过质谱分析方法，在这些表位从肿瘤，组织样品，肿瘤细胞系或其它永生化细胞系的HLA中被洗脱后，对其进行鉴定。尽管可以使用许多这样的方法，但是从细胞表面鉴定表位的最有力的一种方法包括毛细或者非毛细HPLC ESI质谱分析和在线测序，该方法在出版文献中有描述。溶解的HLA和完整细胞的洗脱步骤分别描述在Falk, K.等 Nature 351:290, 1991和美国专利5,989,565中。但是，在文献中没有描述如何鉴别那些表达在细胞（经历肽洗脱和分析的细胞）中的蛋白酶体的类型，进而确定所鉴别的表位是否是看家表位，这是制造有效疫苗所需要的。为了明确地鉴别HLA表位是看家表位或者是免疫表位，通常必须要知道源细胞所表达的是哪种蛋白酶体。优选使用蛋白印迹法来评价蛋白酶体的表达，该方法在下面详述，还可以用RT-PCR，免疫组织化学，或者原位杂交方法来评价蛋白酶体的表达。

实施例8. 干扰素诱导测试

区别看家表位和免疫表位的另一种分析方法，是测定抗肽CTL杀死表达所述TAA的细胞的能力。IFN可以被用于诱导免疫蛋白酶体的表达（假设其并不是组成性表达的），而且被诱导的细胞和未被诱导的细胞的CTL识别能够进行比较。如上，蛋白酶体的类型应该由例如蛋白印迹法确定。如果干扰素诱导的细胞被优先杀死，那么肽组成免疫表位。如果干扰素未诱导的细胞被优先杀死，那么肽组成看家表位。一些表位可以同时被两种蛋白酶体以不同的效率产生，在这种情况下要同时观察两种类群的细胞溶解活性。因为它们呈现在周围靶细胞中，所以这些表位被分类为看家表位。

实施例9 使用人周围血单核细胞(PBMCs)或肿瘤浸润淋巴细胞(TILs)鉴定看家表位

通过使用出版文献中描述的常用方法，可以使用从患者组织活检中分离的TILs，或者使用来自供体或患者血液的PBMCs来鉴定表位。为了鉴定看家表位，需要证实：用于测试被PBMCs或TILs活性杀死的靶细胞

只表达看家蛋白酶体而且不以有意义的水平表达免疫蛋白酶体。使用一组这样的肽抗原来体外刺激来自供体血的PBMCs：该抗原对所使用的血细胞上表达的I类HLA等位基因具有预测的亲和性。每一个PBMC样品都用特异I型肽抗原刺激一周，优选结合使用细胞因子例如IL-2或者IL-12来提高T细胞活性。该刺激至少要重复3次以诱导T细胞特异指向该肽的无性扩增。使用这样的靶细胞执行标准的铬释放分析：已知该靶细胞表达含有表位的蛋白而且只表达看家蛋白酶体。使用铬释放来测定杀死靶细胞的证据，证据显示：用于刺激PBMCs的肽，被作为看家表位呈现在靶细胞表面。因此，表达这种蛋白的肿瘤成为含有这种表位的疫苗的候选靶。

实施例10.用蛋白印迹法鉴定看家蛋白酶体和/或免疫蛋白酶体

下面的方案都是起始于一张薄膜，感兴趣细胞中抽提出来的蛋白在进行了电泳分离后，被转移到这张薄膜上。

A. 显色方案:

1. 室温下，在轨道振动器(RT/shaker)上，在20 ml PBS-T (磷酸缓冲盐溶液, pH 7.4 + 0.1% 吐温-20)中洗膜5分钟。
- PBS (Sigma, Cat. No. P-3813)
- (自始至终，体积可以随着容器类型的变化而变化).
2. 在RT/shaker上，在3%H₂O₂的20mlPBS-T中，将膜培育5分钟：
- 2 ml 30%H₂O₂ + 18 ml PBS-T
3. 在RT/shaker上，用PBS-T洗膜5分钟，洗3遍。
4. 在4⁰C，在RT/shaker上，在含有5%脱脂奶粉的20mlPBS-T中封闭过夜：
- 20mlPBS-T+1g奶
5. 在PBS-T中冲洗膜。
6. 在RT/shaker上，将膜在5ml的、含有一抗(亲和性研究产物公司，英国联合王国(Affinity Research Products Ltd, United Kingdom))的封闭缓冲液里培育2小时:
- | | |
|------------------------------------|---------|
| α-LMP 2 抗血清 (小鼠) (Cat. No.PW8205) | 1:5000 |
| α-LMP 2 抗血清 (人) (Cat. No. PW 8345) | 1:10000 |

α -LMP 7 抗血清 (Cat. No. PW 8200) 1:20000

α -20S 蛋白酶体 α 2 亚单位单克隆抗体 (Cat. No. PW 8105) 1:1000

这些条件只是对于前面的抗体。每种抗体的条件必须以经验确定。

5 7. 按照步骤3清洗膜。

8. 在RT/shaker上, 将膜在5ml的、含有二抗(载体实验室, 有限公司.(Vector Laboratories, Inc.), Burlingame, CA)的封闭缓冲液里培育30分钟:

10 GARB (羊抗兔) (对于抗血清) (载体实验室(Vector Laboratories) Cat. No. BA- 1000) 1:2000

马抗鼠 (对于单克隆抗体) (载体实验室(Vector Laboratories) Cat. No. BA-2000) 1:1000

9. 按照步骤3清洗膜。

15 10. 将膜在5ml的、含有ABC(载体实验室(Vector Laboratories), Cat. No. PK-6100)的PBS-T里培育30分钟:

至少在使用前30分钟, 如下制作ABC:

A=5ul/1ml=25ul/5ml

B=5ul/1ml=25ul/5ml

5ul A+5ul B >混和>4°C静置>加入990ul PBS-T

20 使用前在PBS-T中稀释ABC。

使用前, 在PBS-T中稀释ABC

11. 按照步骤3清洗膜。

12. 检测:

25 1) 将5ml 0.2M的PB转移到第一个15ml试管中

0.4M磷酸缓冲液:

90.4 ml磷酸二氢钠(1M)

619.2 ml磷酸一氢钠(0.5M)

pH到7.4

30 QS到1L

- 2) 将2.8ml 0.2M的PB转移到第二个15ml试管中
3) 将2ml 1%的葡萄糖转移到第三个15ml试管中
4) 称量6mg的ANS (铵基硫酸镍)并且将其转移到第一个
15ml试管中; 搅拌
5) 将110 l的葡萄糖氧化酶(Sigma, Cat. No.G-6891)加入到
eppendorf管中
6) 将110 l的DAB基质(Diaminobenzidine HCl, KPL, Maryland
Cat. No.71-00-46)加入到另一个eppendorf管中。
7) 在通风橱中混和: 5ml PB+ 2ml葡萄糖
+110 l GO
+110 l DAB
+2.8 ml 0.2M PB
13. 将检测混合物加在膜上, 设置定时器。记录色原中的培育时
间长度。
14. 在条带变得清晰可见时, 用0.2M PB将膜清洗3遍。
15. 在RT上, 将膜在PBS中摇动过夜。

B. 化学发光方案:

1. 在TBS-T (Tris-缓冲盐溶液pH7.6 + 0.1% 吐温-20)。Tri缓冲盐
溶液:
2.42g Tris 碱 (20mM)
8g 氯化钠(137mM)
3.8ml 1M 氢氯酸
2. 20ml的封闭缓冲液中封闭过夜 (TBS-T/5% 脱脂奶粉)
4 °C/振荡器:
20ml TBS-T + 1g奶
体积取决于容器的类型
3. 用TBS-T冲洗膜2遍

4. 在RT/shaker上, 将膜在5ml的、含有一抗(亲和性研究产物公司, 英国联合王国(Affinity Research Products Ltd, United Kingdom))的封闭缓冲液里培育2小时:
- | | | |
|---|---|---------|
| | α -LMP 2 抗血清 (小鼠) (Cat. No. PW8205) | 1:5000 |
| 5 | α -LMP 2 抗血清 (人) (Cat. No. PW 8345) | 1:10000 |
| | α -LMP 7 抗血清 (Cat. No. PW 8200) | 1:20000 |
| | α -20S 蛋白酶体 α 2 亚单位单克隆抗体 (Cat. No. PW 8105) | 1:1000 |
5. 在RT/shaker上, 在20ml TBS-T中清洗膜。
- 10 在室温下, 使用两次TBS-T简单地清洗膜两次, 而后使用新鲜的缓冲液将膜用15分钟洗一遍, 5分钟洗两遍。
6. 在RT/shaker上, 将膜在5ml的、被封闭液以1: 1000稀释的HRP标记(辣根过氧化酶标记)的二抗(Amersham; Cat# N11E 824 or N1F 825)里培育1小时。
- 15 7. 按照步骤5清洗膜。
8. 混和等体积的测试溶液1 (Amersham, Cat # RPN2109)和测试溶液2(Amersham, Cat#RPN2109)混和(1ml+1ml)。
9. 将过量的缓冲液从被清洗的膜中排出, 将膜放置在一片Saran Wrap上, 蛋白面向上。加入测试试剂, 覆盖膜。
- 20 10. 室温培养1分钟, 不搅拌。
11. 将过量的测试试剂排出, 将膜转移到Kodak Digital Science Image Station 440CF上, 蛋白面向下。根据产品说明书对信号进行显影并定量。
- 用下面的抗体, 对看家一特异的亚单位(每个方案中的)的存在进行直接测定:
- 25 α - β 1(Y)亚单位单克隆抗体(Cat. No. PW 8140) 1: 1000
- α - β 2 (Z)亚单位单克隆抗体(Cat. No. PW 8145) 1: 1000
- (亲和性研究产物公司, 英国联合王国(Affinity Research Products Ltd, United Kingdom)).
- 30 实施例11.看家表位疫苗的制备

使用商用的肽合成器来合成被鉴定为看家表位的序列。用不同的方法制备感兴趣的肽，肽的给予可以是单独给予，或者与佐剂或细胞因子结合给予以得到刺激T细胞指向动物体内表位的效果。佐剂如CFA, TEA, 或者三聚氨胺；细胞因子如IL-2, IL-12, 或者GM-CSF。肽也可以用控释物质，例如PLGA微球体或其它生物可降解物质来制备，这些物质改变了肽的药代动力学，还提高了其免疫原性。通过使用这些物质，还可以制备经口传送的肽，以促进在GALT(肠相关淋巴细胞组织)吸收的时候，启动免疫反应。肽还可以粘附于微小金颗粒上，这样可以使“基因枪”来运送它们。

10 A. GMP级肽的合成

使用FMOC或者tBOC固相合成方法来合成肽。合成以后，在适当保护清除剂存在的情况下，分别使用三氟乙酸或者氟化氢将肽从其支撑物上切割下来。用蒸发作用去除酸，用醚抽提肽以去除清除剂和原材料，而后冻干沉淀的肽。用HPLC，序列分析，氨基酸分析，反离子含量分析（counterion content analysis）和其它合适的方法来确定粗肽的纯度。如果粗肽足够纯（纯度大于或等于大约90%），就可以使用它们。如果纯化需要达到药物物质规范，那么使用下列的一种方法或几种方法组合来纯化肽：再沉淀；反相，离子交换，大小排阻（size exclusion）或疏水作用色谱；或反流分布（counter-current distribution）。

20 B. 药物产品制剂

将GMP级肽制作在肠胃外可接受的水的，有机的，水—有机的缓冲系统或者溶剂系统中，在这些系统中，它们同时保持着物理上、化学上的稳定和生物潜能。通常，缓冲液或者缓冲液组合或者缓冲液和有机溶剂组合都是合适的。典型的pH范围为6—9。可以加入有机效应物或其它赋形剂来帮助溶解和稳定肽。这些物质包括去垢剂，脂类，共溶剂，抗氧化剂，螯合剂和还原剂。对于冻干制品，可以加入蔗糖或甘露醇或其它冻干酸。通过膜过滤将肽溶液灭菌，将肽溶液滤到其最终的封闭容器系统中，或者冻干用于临床溶解，或者储存直到使用。

实施例12.看家表位肽疫苗的运输

30 A. 结节内传送（intranodal delivery）

使用用于胰岛素传送的小型泵系统(MiniMed; Northridge, CA)将这样的一种制剂连续注射到腹股沟淋巴结, 该制剂含有水性缓冲液中的肽, 缓冲液中有抗微生物制剂, 抗氧化剂和免疫调节细胞因子。注射时间超过数天, 选择这样的注入周期为的是模拟自然注射过程中抗原呈递的动力学。

B.控释

使用受控的PLGA微球体传送肽制剂, 这些微球体改变了肽的药代动力学并且提高了免疫原性。该制剂被注射或者口服。

C.基因枪传送

10 制备肽制剂, 其中的肽被粘到金微粒上。该颗粒被传送到基因枪中, 被加以很高的速度以穿透皮肤, 基因枪携带粒子进入含有pAPCs的皮肤组织。

D.喷雾传送

15 肽制剂被作为喷雾剂吸入, 这样可以被肺中适当的血管或者淋巴组织吸收。

实施例13. 核酸疫苗的制备

携带型质粒载体pVAX1 (Invitrogen, Carlsbad, CA)含有卡那霉素抗性基因和CMV启动子, 将该质粒修饰以加入两个含有需要表位的序列。此外, 该质粒含有一个位于两个表位之间的IRES序列, 以使其在使用一个启动子的情况下, 两个序列能够同时表达。而后, 用该质粒转化合适的大肠杆菌菌株, 并将其铺在选择性培养基上。悬液培养基中长出几个集落, 用限制性图谱鉴定阳性克隆。而后培养阳性克隆, 将其等分到储存瓶中, 在-70°C储存。

25 从这些细胞的一个样品中, 进行质粒的小量制备(QIAprep Spin Miniprep: Qiagen, Valencia, CA), 并且用自动荧光双脱氧序列分析来确定该结构含有所需要的序列。其它的核酸疫苗载体和制剂在本说明书的前面部分和下面的实施例18-20中描述。

实施例14.核酸疫苗的运送

30 使用小型泵系统例如MiniMed胰岛素泵将核酸疫苗注射到淋巴结中。将核酸构建体制备在这样的水缓冲液溶液中, 该缓冲液含有抗微生

物制剂，抗氧化剂和免疫调节细胞因子。该核酸构建体的传送时间超过数天，选择这样的注射周期是为了模拟自然注射过程中抗原呈递的动力学。

- 任选地，使用控释物质例如PLGA微球体或者其它生物可降解物质来
5 传送核酸构建体。这些物质被注射或者口服。核酸疫苗的给予使用经口传输，在吸收到GALT组织过程中引发免疫反应。作为选择，使用基因枪来传送疫苗，其中核酸疫苗粘附于微小金颗粒上。核酸构建体还可以以喷雾剂形式吸入，被肺内的适当的血管或淋巴组织吸收。

实施例15.pAPCs上的看家表位呈递的分析

10 A.四聚物分析

- I类四聚物分析用来确定：在给予看家表位前后，动物体内T细胞的出现频度。对应于表位的T细胞的无性扩增，说明表位被pAPCs呈递给了T细胞。在给予动物表位之前和之后，测量指向看家表位的特异性T细胞的出现频度，可以确定表位是否呈现在pAPCs上。在给予表位后，如果
15 表位特异的T细胞出现频度增加，说明表位被呈递在pAPCs。

B.增殖测定

- 在用看家表位对动物进行免疫约24小时后，使用指向呈现在pAPCs上的特异性标记的单克隆抗体，从PBMCs，脾细胞或淋巴结细胞中收获pAPCs，将其固定在磁珠上进行亲和纯化。使用此技术富集天然血或脾
20 细胞制剂来得到pAPCs。而后将富集的pAPCs用来进行T细胞克隆的增殖测定，该T细胞克隆已经产生并且对感兴趣的看家表位特异。将pAPCs与T细胞克隆共培育，通过测定T细胞掺入放射性胸腺嘧啶的量来检测T细胞的增殖活性。如果增殖，则说明：pAPCs上的表位刺激了对看家表位特异的T细胞。

25 实施例16.看家表位用于抗癌治疗的有效性分析

用替代末端（surrogate endpoint）或存活来确定癌症治疗中表位同步疫苗的效果。

A.T细胞出现频度分析

- 有用的替代末端是T细胞出现频度的确定，所述的T细胞出现频度是指向免疫中使用的看家表位的T细胞出现频度。T细胞（该T细胞是指向
30

使用于肿瘤免疫治疗中的特异TuAA表位的T细胞)出现频度升高的患者,与T细胞出现频度没有升高的患者(该患者被免疫了相同的表位)相比,有明显好的存活性。使用四聚物分析,ELISPOT分析或者有限稀释分析来评价在用表位免疫前后,T细胞对看家表位的出现频度。此出现频度显示了疫苗中看家表位的抗癌作用。

B.肿瘤负担/存活分析

在用看家表位免疫前后,对存在肿瘤的动物进行肿瘤负担评价。部分的或者完全的肿瘤退化表明了有效的治疗干预,并且与增高的存活性相关。在实验室环境中,将几只动物平行地接种肿瘤。其中一些动物随后用看家表位疫苗免疫。将用看家表位免疫的动物的存活率与那些接受了对照表位或者安慰剂的动物的存活率相比较,以确定疫苗的作用。

C.铬释放测定

使用看家表位,对表达人I类MHC的遗传工程动物进行免疫。使用人肿瘤靶或者被设计表达相同I类MHC的靶,将来自这些动物中的T细胞用于标准铬测定。T细胞对靶的杀死,说明刺激患者体内的T细胞可以有效地杀死表达类似TuAA的肿瘤。

实施例17:独特看家表位的鉴定

如此处所公开的那样,在本发明的疫苗和方法中,有用的表位可以被很容易地鉴别出来。例如,三种不是由pAPCs产生的、独特的看家表位被鉴定如下:

A.分离和纯化

分离免疫或看家蛋白酶体复合物。将纯化的肽溶于适当的缓冲液中,使浓度大约为1—2mM,而后加到大约2体积的蛋白酶体制剂中。所选择的缓冲液必须使肽溶剂化而并不干扰消化过程。如前所述,制备额外的消化液用来作阳性对照肽,以证实所使用蛋白酶制剂的正常功能。将它们在37°C下平行温育至120分钟,而后加入三氟乙酸稀释液停止消化;立刻用质谱仪分析样品,或者将它们在干冰上冻结直至用于分析。也可以通过将样品置于冰上来中断消化反应,用质谱仪立即分析。

B.消化液的MALDI-TOF质谱分析

在样品载玻片上,将大约0.5 μ l的每种消化液与等体积的基质(70%

EtOH中有10mg/ml的二羟基苯酸, pH=2-3) 溶液直接混合, 在大约40 ° C使其风干。而后将样品在被合适的分子量标准校准了的Lasermat™ MALDI-TOF质谱仪(Thermo Bioanalysis, Santa Fe, NM)上分析。

用于蛋白酶体分析的计算机程序产生了所有可能片段的序列和分子量, 这些可能片段能够同时满足: 具有任何预测表位的正确C末端; 含有该表位的全长序列或者更长的序列。

如果MALDI-TOF 的结果显示一种或多种这些分子量呈现在消化混合物中, 那么就将相应的肽合成, 纯化, 用MALDI-TOF鉴定, 而后将其用于分析HPLC以同时建立标准驻留时间和质量与峰面积的近似比率。这些步骤与上述那些步骤很类似。而后, 将复制的蛋白酶体消化液稀释在适当的溶剂中并且使用相同的分析HPLC方法进行分析。如果消化液产生了驻留时间与标准驻留时间相同的好的产量峰, 那么基本上可以确定这是由于消化液中的存在的那个序列导致的。如果出现了由其他可能产生片段所引起的模糊(这些片段会产生相同或者类似质谱测量结果), 那么可以收集所怀疑的组分并且将它们用于C末端测序以确定同一性。使用上述方法, 看家表位被鉴定出来。

C. 额外的独特看家表位

依照Stauss等的方法(Proc Natl Acad Sci USA 89(17):7871-5 (1992))分析候选肽对HLA-A2.1的结合。将T2细胞(该细胞在其表面表达表达空的或者不稳定的MHC分子)清洗2遍并以 5×10^6 细胞/ml的浓度悬浮于无血清的完全Iscove's modified Dulbecco's 培养基 (IMDM)中。加入 β_2 微球蛋白使浓度为5 μ g/ml, 将细胞分布到96孔U-底板中使细胞浓度为 5×10^5 细胞/孔。将肽以100, 10, 1和0.1 μ g/ml的浓度加入。将板轻柔摇动2分钟, 而后在37°C, 5% CO₂ 的培养箱内培养4个小时。用IMDM清洗两遍以去除未结合的肽, 而后加入饱和数量的单克隆抗体W6/32 (Sigma)。在4°C培养30分钟后, 用补充了2%热失活FCS, 0.1%(w:v)叠氮化钠, pH7.4-7.6 (着色缓冲液)的PBS清洗细胞, 在4°C下与结合了异硫氰酸荧光素(FITC)的羊F(ab')抗鼠免疫球蛋白G(Sigma)一起培养30分钟, 而后象以前那样清洗4遍。将细胞重悬于着色的缓冲液并且通过加入四分之一体积的2%仲甲醛来使之固定。使用FACScan (Becton Dickinson, San Jose,

CA), 对被肽结合所稳定的表面HLA-A2.1分子进行流式细胞分析。

使用上述讨论的方法, 发现: 由蛋白酶体消化所鉴定的候选酪氨酸酶看家表位(酪氨酸酶 207-216, FLPWHRLFLL SEQ ID NO: 78)与HLA-A2.1结合的程度, 与已知A2.1结合物FLPSDYFPSV (SEQ ID NO: 79) (阳性对照)的结合程度类似。将HLA-B44 结合肽 AEMGKYSFY (SEQ ID NO: 80)作为阴性对照。从阴性对照获得的荧光, 与在分析中没有使用肽时所获得的信号类似。阳性对照肽与阴性对照肽选自通用免疫学方案(Current Protocols in Immunology) p.18.3.2, John Wiley and Sons, 纽约, 1998.中的表18.3.1。

10 实施例18. pVAX-EP1-IRES-EP2的构建

概述:

此构建体的起始质粒是从Invitrogen.(Carlsbad, CA)购买的pVAX1。表位EP1和EP2由GIBCO BRL (Rockville, MD)合成。IRES从由Clontech (Palo Alto, CA)购买的pIRES中切下。

15 步骤:

1. 用EcoRI和NotI消化pIRES。用琼脂糖凝胶分离消化的片段, 通过凝胶纯化来纯化IRES片段;
2. 用EcoRI和NotI消化pVAX1。凝胶纯化pVAX1片段;
3. 建立含有纯化的pVAX1和IRES片段的连接反应;
- 20 4. 用连接产物转化感受态DH5 α ;
5. 挑选4个集落进行小量制备;
6. 对小量制备的DNA进行限制酶消化分析。一个被IRES插入的重组集落被用来进一步插入EP1和EP2。该中间结构称为pVAX-IRES;
- 25 7. 合成EP1和EP2;
8. 将EP1亚克隆到pVAX-IRES的AflII和EcoRI位点之间, 制备出pVAX-EP1-IRES;
9. 将EP2亚克隆到pVAX-EP1-IRES的SalI和NotI位点之间, 制备出pVAX-EPI-IRES-EP2;
- 30 10. 对EP1-IRES-EP2插入物测序以确证其序列。

实施例19. pVAX-EP1-IRES-EP2-ISS-NIS的构建

概述:

此构建体的起始质粒是pVAX-EP1-IRES-EP2（实施例18）。被引入该构建体的ISS（免疫刺激序列）(SEQ ID NO: 82)是AACGTT，使用的NIS
5 （代表核进入序列）(SEQ ID NO: 81)是SV40 72bp重复序列。ISS-NIS由GIBCO BRL合成。见图9。

步骤:

1. 用NruI消化pVAX-EP1-IRES-EP2。凝胶纯化线性质粒;
2. 合成ISS-NIS;
- 10 3. 建立含有纯化的线性pVAX-EP1 -IRES-EP2和合成的ISS-NIS片段的连接反应;
4. 用连接产物转化感受态DH5 α ;
5. 挑选集落进行小量制备;
6. 对小量制备的质粒进行限制酶消化;
- 15 7. 对含有插入物的质粒进行测序。

实施例20 pVAX-EP2-UB-EP1的构建

概论:

此构建体的起始质粒是从pVAX1（Invitrogen）。EP1和EP2由GIBCO BRL合成。在该构建体中编码76个氨基酸的野生型泛素cDNA是从酵母克
20 隆而来的。

步骤:

1. 使用酵母mRNA进行RT-PCR。设计引物来扩增酵母泛素的整个编码序列。
2. 用琼脂糖凝胶分析RT-PCR产物。凝胶纯化预测大小的条带。
- 25 3. 将纯化的DNA条带亚克隆到pZERO1的EcoRV位点，将得到的克隆命名为pZERO-UB;
4. 测序pZERO-UB的几个克隆。在进一步操作前确定泛素的序列。
5. 合成EP1和EP2;
6. 连接EP2，泛素和EP1，并且将该插入物克隆到pVAXI的BamHI
30 和EcoRI位点之间，使之处于CMV启动子的控制下。

7. 通过测序，确认插入物EP2-UB-EP1的序列。

IRES(SEQ ID NO: 6)

AATTCGCCCCCTCTCCCTCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAA
GGCCGGTGTGCGTTTGTCTATATGTGATTTTCCACCATATTGCCGTCTTTTGGCAATGTG
5 AGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCTAGGGGTCTTTCCCCTCT
CGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCT
TCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTG
GCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGC
ACAACCCAGTGCCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCC
TCAAGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTTGTATGGGAT
10 CTGATCTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAACGT
CTAGGCCCCCGAACCACGGGGACGTGGTTTTCTTTGAAAAACACGATGATAA

参考文献：Clontech PT3266-5

UBIQUITIN (SEQ ID NO: 5)

15 ATGCAGATTTTCGTCAAGACTTTGACCGGTAAAACCATAACATTGGAAGTTGAATCTTC
CGATACCATCGACAACGTAAAGTCGAAAATTCAAGACAAGGAAGGTATCCCTCCAGAT
CAACAAAGATTGATCTTTGCCGGTAAGCAGCTAGAAGACGGTAGAACGCTGTCTGATT
ACAACATTCAGAAGGAGTCCACCTTACATCTTGTGCTAAGGCTAAGAGGTGGC

参考文献：Ozkaynak, E., Finley, D., Solomon, M.J. and Varshavsky, A., The yeast ubiquitin
genes: a family of natural gene fusions. EMBO J. 6 (5), 1429-1439 (1987).

NIS (SEQ ID NO: 81)

20 TGGTTGCTGACTAATTGAGATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGG
GACTTTCCACACC

参考文献：Dean DA, Dean BS, Muller S, Smith LC, Sequence requirements for plasmid nuclear
import. Exp. Cell Res. 253 (2): 713-22 (1999).

ISS (SEQ ID NO: 82)

25 AACGTT

参考文献：Sato Y, Roman M, Tighe H, Lee D, Corr M, Nguyen M, Silverman GJ, Lotz M,
Carson DA and Raz E, Immunostimulatory DNA sequences necessary for effective intradermal
gene immunization. Science, 273: 352-354 (1996).

30

- 5 实施例21—24都与由HLA-A2.1呈递的9聚表位的预测有关，尽管该步骤同样适用于任何HLA型，或者任何表位长度，在这个步骤中，可以使用预测性算法或者MHC结合分析。

实施例21. Melan-A/MART- I (SEQ ID NO: 1)

- 10 这种黑色素瘤肿瘤相关抗原（TAA）的长度为118个氨基酸。在100个可能的9聚中，通过SYFPEITHI/Rammensee算法计算，其中的16个9聚为16分。（见表8）。它们占可能肽的14.5%而且代表了蛋白的平均表位密度为0.136每氨基酸。它们当中有12个重叠，覆盖了SEQ ID NO: 1的22—49的氨基酸，在这里，簇的表位密度为0.428，如上所述，两者之间的比率为3.15。另外两个预测的表位重叠氨基酸SEQ ID NO: 1的56—69的
- 15 氨基酸，所得到簇的表位密度为0.143，这与平均值的差别不大，比率仅为1.05。见图15。

表8

Melan-A/MART-1 (SEQ ID NO: 1)的
SYFPEITHI (Rammensee 算法) 结果

顺序	起始	得分
1	31	27
2	56	26
3	35	26
4	32	25
5	27	25
6	29	24
7	34	23
8	61	20
9	33	19
10	22	19
11	99	18
12	36	18
13	28	18
14	87	17
15	41	17
16	40	16

15

用BIMAS-NIH/Parker算法, 将9聚体(这些9聚体被预测的半分离时间是5分钟)限制分析为只剩下了5个。(见表9)。蛋白中的表位的平均密度现在只有0.042每氨基酸。3个重复肽覆盖了SEQ ID NO: 1的31—48氨基酸, 另外两个覆盖了SEQ ID NO: 1的56—69个氨基酸, 如上, 分别得到的比率为3.93和3.40。(见表10)

20

表 9

Melan-A/MART-1 (SEQ ID NO: 1)的BIMAS-NIH/Parker 算法结果

顺序	起始	分数	Log (分数)
1	40	1289.01	3.11
2	56	1055.104	3.02
3	31	81.385	1.91
4	35	20.753	1.32
5	61	4.968	0.70

25

30

表10
Melan-A/MART-1 (SEQ ID NO: 1)的
预测的表位簇

5	计算 (表位/氨基酸)				
	簇	AA	肽	簇	整个蛋白 比率
	1	31-48	3, 4, 1	0.17	0.042 3.93
	2	56-69	2, 5	0.14	0.042 3.40

实施例22. SSX-2/HOM-MEL-40 (SEQ ID NO: 2)

10 这种黑色素瘤肿瘤相关抗原 (TAA) 的长度为188个氨基酸。在180个可能的9聚体中, 通过SYFPEITHI/Rammensee算法计算, 其中的11个9聚体为16分。(见表11)。它们占可能肽的6.1%而且代表了蛋白的平均表位密度为0.059每氨基酸。它们中有3个重叠, 覆盖了SEQ ID NO: 2的99—114的氨基酸, 在这里, 簇的表位密度为0.188, 如上, 两者之间的比率为3.18。在SEQ ID NO: 2的16—28, 57—67, 和167—183氨基酸, 同样存在预测表位重叠对, 得到的比率分别是2.63, 3.11和2.01。还有一个额外的预测表位簇覆盖了SEQ ID NO: 2的5—28氨基酸。对含有3个表位的SEQ ID NO: 25—28区域进行评价, 得到表位的密度为0.125, 比率为2.14。(见表12)。

20 用BIMAS-NIH/Parker算法, 将9聚体(这些9聚体被预测的半分离时间是5分钟)限制分析为只剩下了6个。(见表13)。蛋白中的表位的平均密度现在只有0.032每氨基酸。只有一个单对重叠, 在SEQ ID NO: 2的167—180, 比率为4.48。但是, 如果区域 (SEQ ID NO: 2的41-65氨基酸) 被评价得到的比率为2.51 (这个值代表了与平均值有实质不同), 那么
25 排序名次最高的肽与另一个预测的表位接近。(见图16和表14)。

表11

SSX-2/HOM-MEL-40 (SEQ ID NO: 2)的

SYFPEITHI (Rammensee 算法) 结果

5

10

顺序	起始	得分
1	103	23
2	167	22
3	41	22
4	16	21
5	99	20
6	59	19
7	20	17
8	5	17
9	175	16
10	106	16
11	57	16

表12

15

计算 (表位) (SEQ ID NO: 2)

20

计算 (表位/氨基酸)					
簇	AA	肽	簇	整个蛋白	比率
1	5 to 28	8, 4, 7	0.125	0.059	2.14
2	16-28	4, 7	0.15	0.059	2.63
3	57-67	11, 6	0.18	0.059	3.11
4	99-114	5, 1, 10	0.19	0.059	3.20
5	167-183	2, 9	0.12	0.059	2.01

表13

BIMAS-NIH/Parker算法(SEQ ID NO: 2)

25

顺序	起始	分数	Log (分数)
1	41	1017.062	3.01
2	167	21.672	1.34
3	57	20.81	1.32
4	103	10.433	1.02
5	172	10.068	1.00
6	16	6.442	0.81

30

表14
计算（表位）（SEQ ID NO: 2）

	簇	AA	肽	簇	整个蛋白	比率
5	1	41-65	1, 3	0.08	0.032	2.51
	2	167-180	2, 5	0.14	0.032	4.48

实施例23.NY-ESO（SEQ ID NO: 3）

该肿瘤相关抗原（TAA）的长度为180个氨基酸。在172个可能的9聚体中，通过SYFPEITHI/Rammensee算法计算，其中的25个9聚体为16分。（见表15）。它们占可能肽的14.5%而且代表了蛋白的平均表位密度为0.136每氨基酸。但是它们的分布非常不同。近一半蛋白是空的，在起始的78个氨基酸中，只有一个预测的表位。与Melan-A不同，在Melan-A中，存在高度重叠肽的非常紧凑的簇，而在NY-ESO中重叠很少而且延伸到绝大部分蛋白剩下的部分。一组19个重叠肽覆盖了SEQ ID NO: 3的108—174的氨基酸，所得到的比率为2.04。另外5个预测的表位覆盖了SEQ ID NO: 3的79—104的氨基酸，比率仅为1.38。（见表16）。

如果相反，研究者只考虑预测表位排名最靠前的5%，在这种情况下是9个肽，研究者可以检查好的簇是否被所预测的、不大可能结合MHC的肽所隐藏。当只考虑这些预测的表位的时候，我们可以看到SEQ ID NO: 3的区域108—140包含6个重叠态，比率为3.64。在SEQ ID NO: 3的148—167区域，还有2个附近的肽，比率为2.00。这样可以把大簇SEQ ID NO: 3的108—174断开成为两个覆盖了很多相同序列的小簇。（见表16）。

用BIMAS-NIH/Parker算法，将9聚体（这些9聚体被预测的半分离时间是5分钟）限制分析为只剩下了14个。（见表17）。蛋白中的表位的平均密度现在只有0.078每氨基酸。观察到一个单组的10个重叠肽，覆盖了SEQ ID NO: 3的144—171氨基酸，比率为4.59。所有14个肽处在SEQ ID NO: 3的86—171区域之间，该区域是总蛋白中平均表位密度的2.09倍。虽然这个大簇比我们期望的大，但是相对于研究整个蛋白，它仍然提供了明显的有利条件。（见图17和表18）。

表 15

NY-ESO (SEQ ID NO: 3) 的SYFPEITHI (Rammensee)算法结果

顺序	起始	得分
1	108	25
2	148	24
3	159	21
4	127	21
5	86	21
6	132	20
7	122	20
8	120	20
9	115	20
10	96	20
11	113	19
12	91	19
13	166	18
14	161	18
15	157	18
16	151	18
17	137	18
18	79	18
19	139	17
20	131	17
21	87	17
22	152	16
23	144	16
24	129	16
25	15	16

表16

计算 (表位/氨基酸) (SEQ ID NO: 3)

簇	AA	肽	簇	整个蛋白	比率
1	108-140	1, 9, 8, 7, 4, 6	0.18	0.05	3.64
2	148-167	2, 3	0.10	0.05	2.00
3	79-104	5 12, 10, 18, 21	0.19	0.14	1.38
4	108-174	1, 11, 9, 8, 7, 4, 6, 17, 2, 16, 15, 3, 14, 13, 24, 20, 19, 23, 22	0.28	0.14	2.04

表17

NY-ESO (SEQ ID NO: 3) 的BIMAS-NIH/Parker算法结果

5

10

顺序	起始	分数	Log (分数)
1	159	1197.321	3.08
2	86	429.578	2.63
3	120	130.601	2.12
4	161	83.584	1.92
5	155	52.704	1.72
6	154	49.509	1.69
7	157	42.278	1.63
8	108	21.362	1.33
9	132	19.425	1.29
10	145	13.624	1.13
11	163	11.913	1.08
12	144	11.426	1.06
13	148	6.756	0.83
14	152	4.968	0.70

表18

15

计算 (表位/氨基酸) (SEQ ID NO: 3)

簇	AA	肽	簇	整个蛋白	比率
1	86-171	2, 8, 3, 9, 10, 12, 13, 14, 6, 5, 7, 1, 4, 11	0.163	0.078	2.09
2	144-171	10, 12, 13, 14, 6, 5, 7, 1, 4, 11	0.36	0.078	4.59

20 实施例24.酪氨酸酶(SEQ ID NO: 4)

这种黑色素瘤肿瘤相关抗原 (TAA) 的长度为529个氨基酸。在521个可能的9聚体中, 通过SYFPEITHI/Rammensee算法计算, 其中的52个9聚体为16分。(见表19)。它们占可能肽的10%而且代表了蛋白的平均表位密度为0.098每氨基酸。存在有5组重叠肽, 每组肽都含有2—13个预测表位, 比率分别从2.03到4.41。还有额外的7组重叠肽, 每组含有2—4个预测的表位, 比率分别从1.20到1.85。在SEQ ID NO: 4的区域444—506中的17个肽 (包括上面的13个重复肽) 组成了一个比率为2.20的簇。(见表20)。

用BIMAS-NIH/Parker算法, 将9聚体 (这些9聚体被预测的半分离时间是5分钟) 限制分析为只剩下了28个。(见表21)。蛋白中的表位的平

均密度现在只有0.053每氨基酸。在此密度下，任何重叠都代表超过两倍的表位平均密度。存在有5组重叠肽，每组肽都含有2—7个预测表位，比率分别从2.22到4.9。（见表22）。在这些簇中，只有3个簇能被两个算法共同得到。这些簇中有几个，但不是全部，可以通过评价含有其的区域和邻近被预测的表位将其扩大。见图18。

表19
酪氨酸 (SEQ ID NO: 4)的SYFPEITHI/Rammensee 算法结果

10	<table><tr><th>顺序</th><th>起始</th><th>得分</th></tr><tr><td>1</td><td>490</td><td>34</td></tr><tr><td>2</td><td>491</td><td>31</td></tr><tr><td>3</td><td>487</td><td>28</td></tr><tr><td>4</td><td>1</td><td>27</td></tr><tr><td>5</td><td>2</td><td>25</td></tr><tr><td>6</td><td>482</td><td>23</td></tr><tr><td>7</td><td>380</td><td>23</td></tr><tr><td>8</td><td>369</td><td>23</td></tr><tr><td>9</td><td>214</td><td>23</td></tr><tr><td>10</td><td>506</td><td>22</td></tr><tr><td>11</td><td>343</td><td>22</td></tr><tr><td>12</td><td>207</td><td>22</td></tr><tr><td>13</td><td>137</td><td>22</td></tr><tr><td>14</td><td>57</td><td>22</td></tr><tr><td>15</td><td>169</td><td>20</td></tr><tr><td>16</td><td>118</td><td>20</td></tr><tr><td>17</td><td>9</td><td>20</td></tr><tr><td>18</td><td>488</td><td>19</td></tr><tr><td>19</td><td>483</td><td>19</td></tr><tr><td>20</td><td>480</td><td>19</td></tr><tr><td>21</td><td>479</td><td>19</td></tr><tr><td>22</td><td>478</td><td>19</td></tr><tr><td>23</td><td>473</td><td>19</td></tr><tr><td>24</td><td>365</td><td>19</td></tr><tr><td>25</td><td>287</td><td>19</td></tr><tr><td>26</td><td>200</td><td>19</td></tr></table>	顺序	起始	得分	1	490	34	2	491	31	3	487	28	4	1	27	5	2	25	6	482	23	7	380	23	8	369	23	9	214	23	10	506	22	11	343	22	12	207	22	13	137	22	14	57	22	15	169	20	16	118	20	17	9	20	18	488	19	19	483	19	20	480	19	21	479	19	22	478	19	23	473	19	24	365	19	25	287	19	26	200	19	<table><tr><th>顺序</th><th>起始</th><th>得分</th></tr><tr><td>27</td><td>5</td><td>19</td></tr><tr><td>28</td><td>484</td><td>18</td></tr><tr><td>29</td><td>476</td><td>18</td></tr><tr><td>30</td><td>463</td><td>18</td></tr><tr><td>31</td><td>444</td><td>18</td></tr><tr><td>32</td><td>425</td><td>18</td></tr><tr><td>33</td><td>316</td><td>18</td></tr><tr><td>34</td><td>187</td><td>18</td></tr><tr><td>35</td><td>402</td><td>17</td></tr><tr><td>36</td><td>388</td><td>17</td></tr><tr><td>37</td><td>346</td><td>17</td></tr><tr><td>38</td><td>336</td><td>17</td></tr><tr><td>39</td><td>225</td><td>17</td></tr><tr><td>40</td><td>224</td><td>17</td></tr><tr><td>41</td><td>208</td><td>17</td></tr><tr><td>42</td><td>186</td><td>17</td></tr><tr><td>43</td><td>171</td><td>17</td></tr><tr><td>44</td><td>514</td><td>16</td></tr><tr><td>45</td><td>494</td><td>16</td></tr><tr><td>46</td><td>406</td><td>16</td></tr><tr><td>47</td><td>385</td><td>16</td></tr><tr><td>48</td><td>349</td><td>16</td></tr><tr><td>49</td><td>184</td><td>16</td></tr><tr><td>50</td><td>167</td><td>16</td></tr><tr><td>51</td><td>145</td><td>16</td></tr><tr><td>52</td><td>139</td><td>16</td></tr></table>	顺序	起始	得分	27	5	19	28	484	18	29	476	18	30	463	18	31	444	18	32	425	18	33	316	18	34	187	18	35	402	17	36	388	17	37	346	17	38	336	17	39	225	17	40	224	17	41	208	17	42	186	17	43	171	17	44	514	16	45	494	16	46	406	16	47	385	16	48	349	16	49	184	16	50	167	16	51	145	16	52	139	16
顺序	起始	得分																																																																																																																																																																		
1	490	34																																																																																																																																																																		
2	491	31																																																																																																																																																																		
3	487	28																																																																																																																																																																		
4	1	27																																																																																																																																																																		
5	2	25																																																																																																																																																																		
6	482	23																																																																																																																																																																		
7	380	23																																																																																																																																																																		
8	369	23																																																																																																																																																																		
9	214	23																																																																																																																																																																		
10	506	22																																																																																																																																																																		
11	343	22																																																																																																																																																																		
12	207	22																																																																																																																																																																		
13	137	22																																																																																																																																																																		
14	57	22																																																																																																																																																																		
15	169	20																																																																																																																																																																		
16	118	20																																																																																																																																																																		
17	9	20																																																																																																																																																																		
18	488	19																																																																																																																																																																		
19	483	19																																																																																																																																																																		
20	480	19																																																																																																																																																																		
21	479	19																																																																																																																																																																		
22	478	19																																																																																																																																																																		
23	473	19																																																																																																																																																																		
24	365	19																																																																																																																																																																		
25	287	19																																																																																																																																																																		
26	200	19																																																																																																																																																																		
顺序	起始	得分																																																																																																																																																																		
27	5	19																																																																																																																																																																		
28	484	18																																																																																																																																																																		
29	476	18																																																																																																																																																																		
30	463	18																																																																																																																																																																		
31	444	18																																																																																																																																																																		
32	425	18																																																																																																																																																																		
33	316	18																																																																																																																																																																		
34	187	18																																																																																																																																																																		
35	402	17																																																																																																																																																																		
36	388	17																																																																																																																																																																		
37	346	17																																																																																																																																																																		
38	336	17																																																																																																																																																																		
39	225	17																																																																																																																																																																		
40	224	17																																																																																																																																																																		
41	208	17																																																																																																																																																																		
42	186	17																																																																																																																																																																		
43	171	17																																																																																																																																																																		
44	514	16																																																																																																																																																																		
45	494	16																																																																																																																																																																		
46	406	16																																																																																																																																																																		
47	385	16																																																																																																																																																																		
48	349	16																																																																																																																																																																		
49	184	16																																																																																																																																																																		
50	167	16																																																																																																																																																																		
51	145	16																																																																																																																																																																		
52	139	16																																																																																																																																																																		

表20

计算(表位/氨基酸)(SEQ ID NO: 4)

族	AA	肽	簇	整个蛋白	比率
1	1 to 17	4, 5, 27, 17	0.24	0.098	2.39
2	137-153	13, 52, 51	0.18	0.098	1.80
3	167-179	15, 43, 50	0.23	0.098	2.35
4	184-195	34, 42, 49	0.25	0.098	2.54
5	200-222	26, 41, 9, 12	0.17	0.098	1.77
6	224-233	39, 40	0.20	0.098	2.03
7	336-357	38, 11, 37, 48	0.18	0.098	1.85
8	365-377	24, 8	0.15	0.098	1.57
9	380-396	7, 47, 36	0.18	0.098	1.80
10	402-414	35, 46	0.15	0.098	1.57
11	473-502	29, 28, 23, 22, 21, 20, 6, 19, 3, 18, 1, 2, 45	0.43	0.098	4.41
12	506-522	10, 44	0.12	0.098	1.20
	444-522	31, 30, 23, 29, 22, 21, 20, 6, 19, 28, 3, 18, 1, 2, 45, 10, 44	0.22	0.098	2.20

表21

BIMAS-NIH/Parker算法(SEQ ID NO: 4)

顺序	起始	分数	Log(分数)
1	207	540.469	2.73
2	369	531.455	2.73
3	1	309.05	2.49
4	9	266.374	2.43
5	490	181.794	2.26
6	214	177.566	2.25
7	224	143.451	2.16
8	171	93.656	1.97
9	506	87.586	1.94
10	487	83.527	1.92
11	491	83.527	1.92
12	2	54.474	1.74
13	137	47.991	1.68
14	200	30.777	1.49
15	208	26.248	1.42
16	460	21.919	1.34
17	478	19.425	1.29
18	365	17.14	1.23
19	380	16.228	1.21
20	444	13.218	1.12
21	473	13.04	1.12
22	57	10.868	1.04
23	482	8.252	0.92
24	483	7.309	0.86
25	5	6.993	0.84
26	225	5.858	0.77
27	343	5.195	0.72
28	514	5.179	0.71

表22

计算（表位/氨基酸）（SEQ ID NO: 4）

簇	AA	肽	簇	整个蛋白	比率
1	1 to 17	3, 12, 25, 4	0.24	0.053	4.45
2	200-222	14, 1, 15, 6	0.17	0.053	3.29
3	224-233	7, 26	0.20	0.053	3.78
4	365-377	18, 2	0.15	0.053	2.91
5	473-499	21, 17, 23, 24, 10, 5, 11	0.26	0.053	4.90
6	506-522	9, 28	0.12	0.053	2.22
7	365-388	18, 2, 19	0.13	0.053	2.36
8	444-499	20, 16, 21, 17, 23, 24, 10, 5, 11	0.16	0.053	3.03
9	444-522	20, 16, 21, 17, 23, 24, 10, 5, 11, 9, 28	0.14	0.053	2.63
10	200-233	14, 1, 15, 6, 7, 26	0.18	0.053	3.33

<110> CTL 免疫治疗公司
 约翰, J., L. 西马德
 戴维, C. 戴蒙德
 雷向东

<120> 抗原呈递细胞中的表位同步

<130> CTLIMM. 004VPC

<150> US 09/561, 572

<151> 2000-04-28

<150> US 09/561, 571

<151> 2000-04-28

<150> US 09/561, 074

<151> 2000-04-28

<150> US 09/560, 465

<151> 2000-04-28

<160> 82

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 118

<212> PRT

<213> 智人

<400> 1

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

Pro Pro Pro Tyr Ser Pro
115

<210> 2
<211> 188
<212> PRT
<213> 智人

<400> 2
Met Asn Gly Asp Asp Ala Phe Ala Arg Arg Pro Thr Val Gly Ala Gln
1 5 10 15
Ile Pro Glu Lys Ile Gln Lys Ala Phe Asp Asp Ile Ala Lys Tyr Phe
20 25 30
Ser Lys Glu Glu Trp Glu Lys Met Lys Ala Ser Glu Lys Ile Phe Tyr
35 40 45
Val Tyr Met Lys Arg Lys Tyr Glu Ala Met Thr Lys Leu Gly Phe Lys
50 55 60
Ala Thr Leu Pro Pro Phe Met Cys Asn Lys Arg Ala Glu Asp Phe Gln
65 70 75 80
Gly Asn Asp Leu Asp Asn Asp Pro Asn Arg Gly Asn Gln Val Glu Arg
85 90 95
Pro Gln Met Thr Phe Gly Arg Leu Gln Gly Ile Ser Pro Lys Ile Met
100 105 110
Pro Lys Lys Pro Ala Glu Glu Gly Asn Asp Ser Glu Glu Val Pro Glu
115 120 125
Ala Ser Gly Pro Gln Asn Asp Gly Lys Glu Leu Cys Pro Pro Gly Lys
130 135 140
Pro Thr Thr Ser Glu Lys Ile His Glu Arg Ser Gly Pro Lys Arg Gly
145 150 155 160
Glu His Ala Trp Thr His Arg Leu Arg Glu Arg Lys Gln Leu Val Ile
165 170 175
Tyr Glu Glu Ile Ser Asp Pro Glu Glu Asp Asp Glu
180 185

<210> 3
<211> 180
<212> PRT
<213> 智人

<400> 3
Met Gln Ala Glu Gly Arg Gly Thr Gly Gly Ser Thr Gly Asp Ala Asp
1 5 10 15
Gly Pro Gly Gly Pro Gly Ile Pro Asp Gly Pro Gly Gly Asn Ala Gly
20 25 30
Gly Pro Gly Glu Ala Gly Ala Thr Gly Gly Arg Gly Pro Arg Gly Ala
35 40 45
Gly Ala Ala Arg Ala Ser Gly Pro Gly Gly Gly Ala Pro Arg Gly Pro

```

      50              55              60
His Gly Gly Ala Ala Ser Gly Leu Asn Gly Cys Cys Arg Cys Gly Ala
65              70              75              80
Arg Gly Pro Glu Ser Arg Leu Leu Glu Phe Tyr Leu Ala Met Pro Phe
      85              90              95
Ala Thr Pro Met Glu Ala Glu Leu Ala Arg Arg Ser Leu Ala Gln Asp
      100             105             110
Ala Pro Pro Leu Pro Val Pro Gly Val Leu Leu Lys Glu Phe Thr Val
      115             120             125
Ser Gly Asn Ile Leu Thr Ile Arg Leu Thr Ala Ala Asp His Arg Gln
      130             135             140
Leu Gln Leu Ser Ile Ser Ser Cys Leu Gln Gln Leu Ser Leu Leu Met
145             150             155             160
Trp Ile Thr Gln Cys Phe Leu Pro Val Phe Leu Ala Gln Pro Pro Ser
      165             170             175
Gly Gln Arg Arg
      180

```

<210> 4

<211> 529

<212> PRT

<213> 智人

<400> 4

```

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

```

Glu Ile Trp Arg Asp Ile Asp Phe Ala His Glu Ala Pro Ala Phe Leu
 195 200 205
 Pro Trp His Arg Leu Phe Leu Leu Arg Trp Glu Gln Glu Ile Gln Lys
 210 215 220
 Leu Thr Gly Asp Glu Asn Phe Thr Ile Pro Tyr Trp Asp Trp Arg Asp
 225 230 235 240
 Ala Glu Lys Cys Asp Ile Cys Thr Asp Glu Tyr Met Gly Gly Gln His
 245 250 255
 Pro Thr Asn Pro Asn Leu Leu Ser Pro Ala Ser Phe Phe Ser Ser Trp
 260 265 270
 Gln Ile Val Cys Ser Arg Leu Glu Glu Tyr Asn Ser His Gln Ser Leu
 275 280 285
 Cys Asn Gly Thr Pro Glu Gly Pro Leu Arg Arg Asn Pro Gly Asn His
 290 295 300
 Asp Lys Ser Arg Thr Pro Arg Leu Pro Ser Ser Ala Asp Val Glu Phe
 305 310 315 320
 Cys Leu Ser Leu Thr Gln Tyr Glu Ser Gly Ser Met Asp Lys Ala Ala
 325 330 335
 Asn Phe Ser Phe Arg Asn Thr Leu Glu Gly Phe Ala Ser Pro Leu Thr
 340 345 350
 Gly Ile Ala Asp Ala Ser Gln Ser Ser Met His Asn Ala Leu His Ile
 355 360 365
 Tyr Met Asn Gly Thr Met Ser Gln Val Gln Gly Ser Ala Asn Asp Pro
 370 375 380
 Ile Phe Leu Leu His His Ala Phe Val Asp Ser Ile Phe Glu Gln Trp
 385 390 395 400
 Leu Arg Arg His Arg Pro Leu Gln Glu Val Tyr Pro Glu Ala Asn Ala
 405 410 415
 Pro Ile Gly His Asn Arg Glu Ser Tyr Met Val Pro Phe Ile Pro Leu
 420 425 430
 Tyr Arg Asn Gly Asp Phe Phe Ile Ser Ser Lys Asp Leu Gly Tyr Asp
 435 440 445
 Tyr Ser Tyr Leu Gln Asp Ser Asp Pro Asp Ser Phe Gln Asp Tyr Ile
 450 455 460
 Lys Ser Tyr Leu Glu Gln Ala Ser Arg Ile Trp Ser Trp Leu Leu Gly
 465 470 475 480
 Ala Ala Met Val Gly Ala Val Leu Thr Ala Leu Leu Ala Gly Leu Val
 485 490 495
 Ser Leu Leu Cys Arg His Lys Arg Lys Gln Leu Pro Glu Glu Lys Gln
 500 505 510
 Pro Leu Leu Met Glu Lys Glu Asp Tyr His Ser Leu Tyr Gln Ser His
 515 520 525
 Leu

<210> 5

<211> 228

<212> DNA

<213> 酿酒酵母

<400> 5

```

atgcagattt tcgtcaagac tttagaccgtt aaaaccataa cattggaagt tgaatcttcc 60
gataccatcg acaacgttaa gtcgaaaatt caagacaagg aaggtatccc tccagatcaa 120
caaagattga tctttgccgg taagcagcta gaagacggta gaacgctgtc tgattacaac 180
attcagaagg agtccacctt acatcttgtg ctaaggctaa gaggtggc 228

```

<210> 6

<211> 581

<212> DNA

<213> 脑心肌炎病毒

<400> 6

```

aattccgccc ctctccctcc ccccccccta acgttactgg ccgaagccgc ttggaataag 60
gccggtgtgc gtttgtctat atgtgatit ccacatatt gccgtctttt ggcaatgtga 120
gggcccggaa acctggccct gtcttcttga cgagcattcc taggggtctt tcccctctcg 180
ccaaagggaat gcaaggctcg ttgaatgtcg tgaaggaagc agttcctctg gaagcttctt 240
gaagacaaac aacgtctgta gcgaccttt gcaggcagcg gaacccccca cctggcgaca 300
ggtgcctctg cggccaaaag ccacgtgtat aagatacacc tgcaaaggcg gcacaacccc 360
agtgccacgt tgtgagttgg atagtgtgg aaagagtcaa atggctctcc tcaagcgtat 420
tcaacaaggg gctgaaggat gcccagaagg taccatttg tatgggatct gatctggggc 480
ctcgggtcac atgctttaca tgtgttttagt cgaggttaaa aaaacgtcta ggccccccga 540
accacgggga cgtggttttc ctttgaaaaa cacgatgata a 581

```

<210> 7

<211> 16

<212> PRT

<213> 智人

<400> 7

```

Met Leu Leu Ala Val Leu Tyr Cys Leu Leu Trp Ser Phe Gln Thr Ser
1           5           10           15

```

<210> 8

<211> 22

<212> PRT

<213> 智人

<400> 8

```

His Ser Tyr Thr Thr Ala Glu Glu Ala Ala Gly Ile Thr Ile Leu Thr
1           5           10           15
Val Ile Leu Gly Val Leu
20

```

<210> 9

<211> 24

<212> PRT

<213> 智人

<400> 9

Glu	Ala	Ala	Ser	Ser	Ser	Ser	Thr	Leu	Val	Glu	Val	Thr	Leu	Gly	Glu
1				5					10					15	
Val	Pro	Ala	Ala	Glu	Ser	Pro	Asp								
				20											

<210> 10

<211> 24

<212> PRT

<213> 智人

<400> 10

Glu	Phe	Leu	Trp	Gly	Pro	Arg	Ala	Leu	Val	Glu	Thr	Ser	Tyr	Val	Lys
1				5					10					15	
Val	Leu	His	His	Met	Val	Lys	Ile								
				20											

<210> 11

<211> 19

<212> PRT

<213> 智人

<400> 11

Ala	Pro	Glu	Glu	Lys	Ile	Trp	Glu	Glu	Leu	Ser	Val	Leu	Glu	Val	Phe
1				5					10					15	
Glu	Gly	Arg													

<210> 12

<211> 16

<212> PRT

<213> 智人

<400> 12

Glu	Leu	Met	Glu	Val	Asp	Pro	Ile	Gly	His	Leu	Tyr	Ile	Phe	Ala	Thr
1				5					10					15	

<210> 13

<211> 15

<212> PRT

<213> 智人

<400> 13

Phe	Leu	Trp	Gly	Pro	Arg	Ala	Leu	Val	Glu	Thr	Ser	Tyr	Val	Lys
1				5					10					15

<210> 14

<211> 16

<212> PRT

<213> 智人

<400> 14

Met	Leu	Leu	Ala	Val	Leu	Tyr	Cys	Leu	Leu	Trp	Ser	Phe	Gln	Thr	Ser
1				5					10					15	

<210> 15

<211> 22

<212> PRT

<213> 智人

<400> 15

His	Ser	Tyr	Thr	Thr	Ala	Glu	Glu	Ala	Ala	Gly	Ile	Thr	Ile	Leu	Thr
1				5					10					15	
Val	Ile	Leu	Gly	Val	Leu										
					20										

<210> 16

<211> 24

<212> PRT

<213> 智人

<400> 16

Glu	Ala	Ala	Ser	Ser	Ser	Ser	Thr	Leu	Val	Glu	Val	Thr	Leu	Gly	Glu
1				5					10					15	
Val	Pro	Ala	Ala	Glu	Ser	Pro	Asp								
				20											

<210> 17

<211> 24

<212> PRT

<213> 智人

<400> 17

Glu	Phe	Leu	Trp	Gly	Pro	Arg	Ala	Leu	Val	Glu	Thr	Ser	Tyr	Val	Lys
1				5					10					15	
Val	Leu	His	His	Met	Val	Lys	Ile								
				20											

<210> 18
 <211> 19
 <212> PRT
 <213> 智人

<400> 18
 Ala Pro Glu Glu Lys Ile Trp Glu Glu Leu Ser Val Leu Glu Val Phe
 1 5 10 15
 Glu Gly Arg

<210> 19
 <211> 16
 <212> PRT
 <213> 智人

<400> 19
 Glu Leu Met Glu Val Asp Pro Ile Gly His Leu Tyr Ile Phe Ala Thr
 1 5 10 15

<210> 20
 <211> 15
 <212> PRT
 <213> 智人

<400> 20
 Phe Leu Trp Gly Pro Arg Ala Leu Val Glu Thr Ser Tyr Val Lys
 1 5 10 15

<210> 21
 <211> 10
 <212> PRT
 <213> 智人

<400> 21
 Leu Gln Leu Ser Asn Gly Asn Arg Thr Leu
 1 5 10

<210> 22
 <211> 10
 <212> PRT
 <213> 智人

<400> 22

Leu	Gln	Leu	Ser	Asn	Asp	Asn	Arg	Thr	Leu
1				5					10

<210> 23

<211> 10

<212> PRT

<213> 智人

<400> 23

Leu	Gln	Leu	Ser	Asn	Gly	Asn	Arg	Thr	Leu
1				5					10

<210> 24

<211> 10

<212> PRT

<213> 智人

<400> 24

Leu	Leu	Val	His	Asn	Leu	Pro	Gln	His	Leu
1				5					10

<210> 25

<211> 10

<212> PRT

<213> 智人

<400> 25

Leu	Ser	Asn	Asp	Asn	Arg	Thr	Leu	Thr	Leu
1				5					10

<210> 26

<211> 10

<212> PRT

<213> 智人

<400> 26

Leu	Ser	Asn	Gly	Asn	Arg	Thr	Leu	Thr	Leu
1				5					10

<210> 27

<211> 10

<212> PRT

<213> 智人

<400> 27

Leu Ser Asn Gly Asn Arg Thr Leu Thr Leu
1 5 10

<210> 28

<211> 10

<212> PRT

<213> 智人

<400> 28

Cys Gln Ala His Asn Ser Asp Thr Gly Leu
1 5 10

<210> 29

<211> 10

<212> PRT

<213> 智人

<400> 29

Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile
1 5 10

<210> 30

<211> 10

<212> PRT

<213> 智人

<400> 30

Ser Ala Asn Arg Ser Asp Pro Val Thr Leu
1 5 10

<210> 31

<211> 10

<212> PRT

<213> 智人

<400> 31

Ser Val Ser Ala Asn Arg Ser Asp Pro Val
1 5 10

<210> 32

<211> 10

<212> PRT

<213> 智人

<400> 32

Ser Ile Ser Ser Asn Asn Ser Lys Pro Val
1 5 10

<210> 33

<211> 10

<212> PRT

<213> 智人

<400> 33

Trp Trp Val Asn Asn Gln Ser Leu Pro Val
1 5 10

<210> 34

<211> 10

<212> PRT

<213> 智人

<400> 34

Trp Trp Val Asn Asn Gln Ser Leu Pro Val
1 5 10

<210> 35

<211> 10

<212> PRT

<213> 智人

<400> 35

Ser Ile Ser Ser Asn Asn Ser Lys Pro Val
1 5 10

<210> 36

<211> 10

<212> PRT

<213> 智人

<400> 36

Ser Asp Leu Val Asn Glu Glu Ala Thr Gly
1 5 10

<210> 37

<211> 10

<212> PRT
<213> 智人

<400> 37
Ile Thr Pro Asn Asn Asn Gly Thr Tyr Ala
1 5 10

<210> 38
<211> 10
<212> PRT
<213> 智人

<400> 38
Ser Leu Pro Val Ser Pro Arg Leu Gln Leu
1 5 10

<210> 39
<211> 10
<212> PRT
<213> 智人

<400> 39
Ser Leu Pro Val Ser Pro Arg Leu Gln Leu
1 5 10

<210> 40
<211> 10
<212> PRT
<213> 智人

<400> 40
Phe Ile Thr Ser Asn Asn Ser Asn Pro Val
1 5 10

<210> 41
<211> 10
<212> PRT
<213> 智人

<400> 41
Arg Ser Gly Glu Asn Leu Asn Leu Ser Cys
1 5 10

<210> 42

<211> 10
<212> PRT
<213> 智人

<400> 42
Ser Leu Pro Val Ser Pro Arg Leu Gln Leu
1 5 10

<210> 43
<211> 10
<212> PRT
<213> 智人

<400> 43
Val Tyr Pro Glu Leu Pro Lys Pro Ser Ile
1 5 10

<210> 44
<211> 10
<212> PRT
<213> 智人

<400> 44
Ser Thr Pro Phe Asn Val Ala Glu Gly Lys
1 5 10

<210> 45
<211> 10
<212> PRT
<213> 智人

<400> 45
Asn Pro Pro Thr Thr Ala Lys Leu Thr Ile
1 5 10

<210> 46
<211> 10
<212> PRT
<213> 智人

<400> 46
Cys Phe Val Ser Asn Leu Ala Thr Gly Arg
1 5 10

<210> 47
<211> 10
<212> PRT
<213> 智人

<400> 47
Gly Ala Asn Leu Asn Leu Ser Cys His Ser
1 5 10

<210> 48
<211> 10
<212> PRT
<213> 智人

<400> 48
Leu Ser Gly Ala Asn Leu Asn Leu Ser Cys
1 5 10

<210> 49
<211> 10
<212> PRT
<213> 智人

<400> 49
Ser Tyr Leu Ser Gly Ala Asn Leu Asn Leu
1 5 10

<210> 50
<211> 10
<212> PRT
<213> 智人

<400> 50
Cys Gly Ile Gln Asn Ser Val Ser Ala Asn
1 5 10

<210> 51
<211> 10
<212> PRT
<213> 智人

<400> 51
Val Ser Ala Glu Leu Pro Lys Pro Ser Ile
1 5 10

<210> 52
<211> 10
<212> PRT
<213> 智人

<400> 52
Ser Asn Ile Thr Glu Lys Asn Ser Gly Leu
1 5 10

<210> 53
<211> 10
<212> PRT
<213> 智人

<400> 53
Gly Asn Ile Gln Gln His Thr Gln Glu Leu
1 5 10

<210> 54
<211> 10
<212> PRT
<213> 智人

<400> 54
Ser Asn Pro Pro Ala Gln Tyr Ser Trp Leu
1 5 10

<210> 55
<211> 10
<212> PRT
<213> 智人

<400> 55
Cys Gly Ile Gln Asn Glu Leu Ser Val Asp
1 5 10

<210> 56
<211> 10
<212> PRT
<213> 智人

<400> 56
Ile Ser Pro Leu Asn Thr Ser Tyr Arg Ser
1 5 10

<210> 57
<211> 10
<212> PRT
<213> 智人

<400> 57
Pro Gly Pro Ala Tyr Ser Gly Arg Glu Ile
1 5 10

<210> 58
<211> 10
<212> PRT
<213> 智人

<400> 58
Phe Asn Val Ala Glu Gly Lys Glu Val Leu
1 5 10

<210> 59
<211> 10
<212> PRT
<213> 智人

<400> 59
Ile Gly Val Leu Val Gly Val Ala Leu Ile
1 5 10

<210> 60
<211> 10
<212> PRT
<213> 智人

<400> 60
Ser Ala Gly Ala Thr Val Gly Ile Met Ile
1 5 10

<210> 61
<211> 10
<212> PRT
<213> 智人

<400> 61
Ser Lys Pro Val Glu Asp Lys Asp Ala Val

1	5	10
---	---	----

<210> 62
 <211> 10
 <212> PRT
 <213> 智人

<400> 62
 Ser Ala Ser Gly His Ser Arg Thr Thr Val
 1 5 10

<210> 63
 <211> 10
 <212> PRT
 <213> 智人

<400> 63
 Arg Pro Gly Val Asn Leu Ser Leu Ser Cys
 1 5 10

<210> 64
 <211> 10
 <212> PRT
 <213> 智人

<400> 64
 Leu Ser Val Asp His Ser Asp Pro Val Ile
 1 5 10

<210> 65
 <211> 10
 <212> PRT
 <213> 智人

<400> 65
 Ser Asn Asp Asn Arg Thr Leu Thr Leu Leu
 1 5 10

<210> 66
 <211> 10
 <212> PRT
 <213> 智人

<400> 66

Ser	Asn	Pro	Val	Glu	Asp	Glu	Asp	Ala	Val
1				5					10

<210> 67
<211> 10
<212> PRT
<213> 智人

<400> 67
Ser Asn Asn Ser Asn Pro Val Glu Asp Glu
1 5 10

,
<210> 68
<211> 10
<212> PRT
<213> 智人

<400> 68
Gly Leu Asn Arg Thr Thr Val Thr Thr Ile
1 5 10

<210> 69
<211> 10
<212> PRT
<213> 智人

<400> 69
Ile Thr Val Asn Asn Ser Gly Ser Tyr Thr
1 5 10

<210> 70
<211> 10
<212> PRT
<213> 智人

<400> 70
Ser Thr Gln Glu Leu Phe Ile Pro Asn Ile
1 5 10

<210> 71
<211> 10
<212> PRT
<213> 智人

<400> 71

Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile
1 5 10

<210> 72

<211> 10

<212> PRT

<213> 智人

<400> 72

Ser Ala Arg Arg Ser Asp Ser Val Ile Leu
1 5 10

<210> 73

<211> 10

<212> PRT

<213> 智人

<400> 73

Val Ser Ala Arg Arg Ser Asp Ser Val Ile
1 5 10

<210> 74

<211> 10

<212> PRT

<213> 智人

<400> 74

Ser Lys Pro Val Glu Asp Lys Asp Ala Val
1 5 10

<210> 75

<211> 10

<212> PRT

<213> 智人

<400> 75

Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
1 5 10

<210> 76

<211> 10

<212> PRT

<213> 智人

<400> 76

Lys Glu Val Leu Leu Leu Val His Asn Leu
1 5 10

<210> 77

<211> 10

<212> PRT

<213> 智人

<400> 77

Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe
1 5 10

<210> 78

<211> 10

<212> PRT

<213> 智人

<400> 78

Phe Leu Pro Trp His Arg Leu Phe Leu Leu
1 5 10

<210> 79

<211> 10

<212> PRT

<213> 智人

<400> 79

Phe Leu Pro Ser Asp Tyr Phe Pro Ser Val
1 5 10

<210> 80

<211> 9

<212> PRT

<213> 智人

<400> 80

Ala Glu Met Gly Lys Tyr Ser Phe Tyr
1 5

<210> 81

<211> 72

<212> DNA

<213> SV40

<220>

<223> 合成的核苷酸序列

<400> 81

tggttgctga ctaattgaga tgcatgcttt gcatacttct gcctgctggg gagcctgggg 60
actttccaca cc 72

<210> 82

<211> 6

<212> DNA

<213> 人工序列

<220>

<223> 合成的核苷酸序列

<400> 82

aacgtt

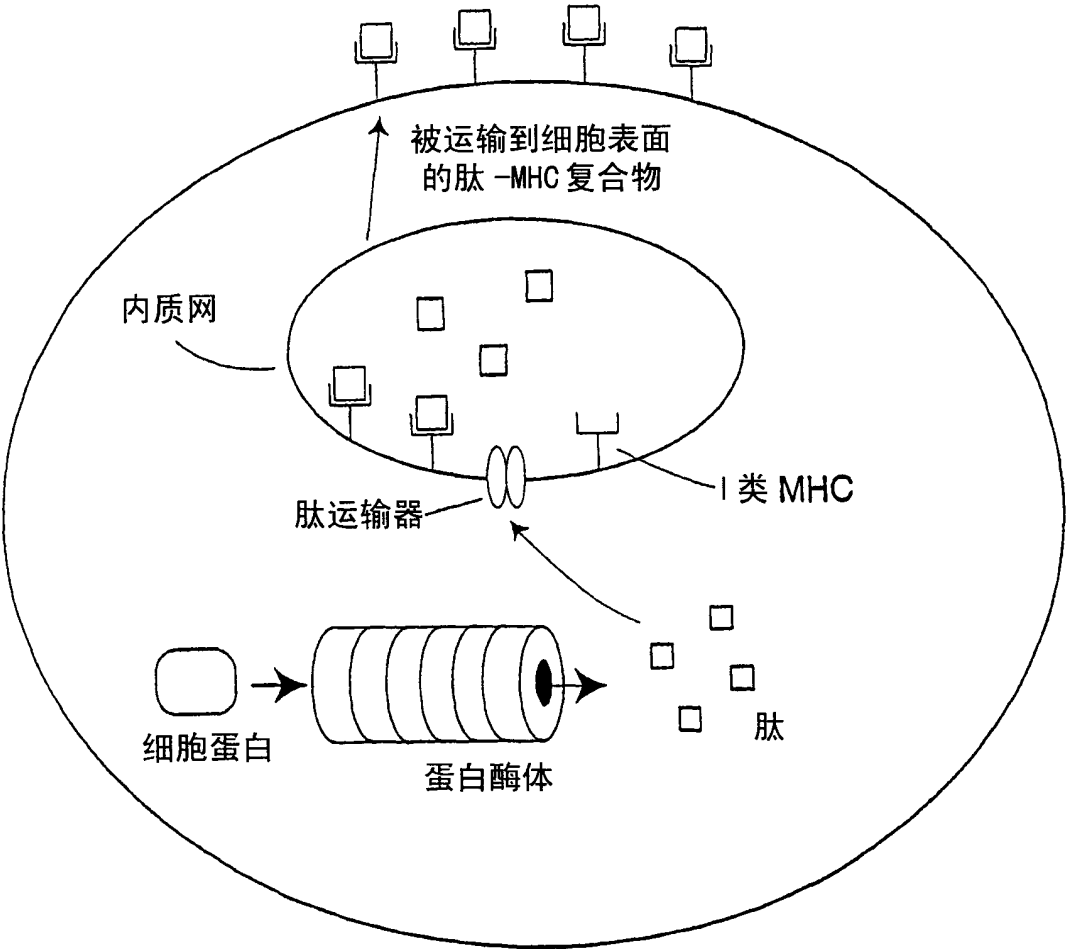


图 1

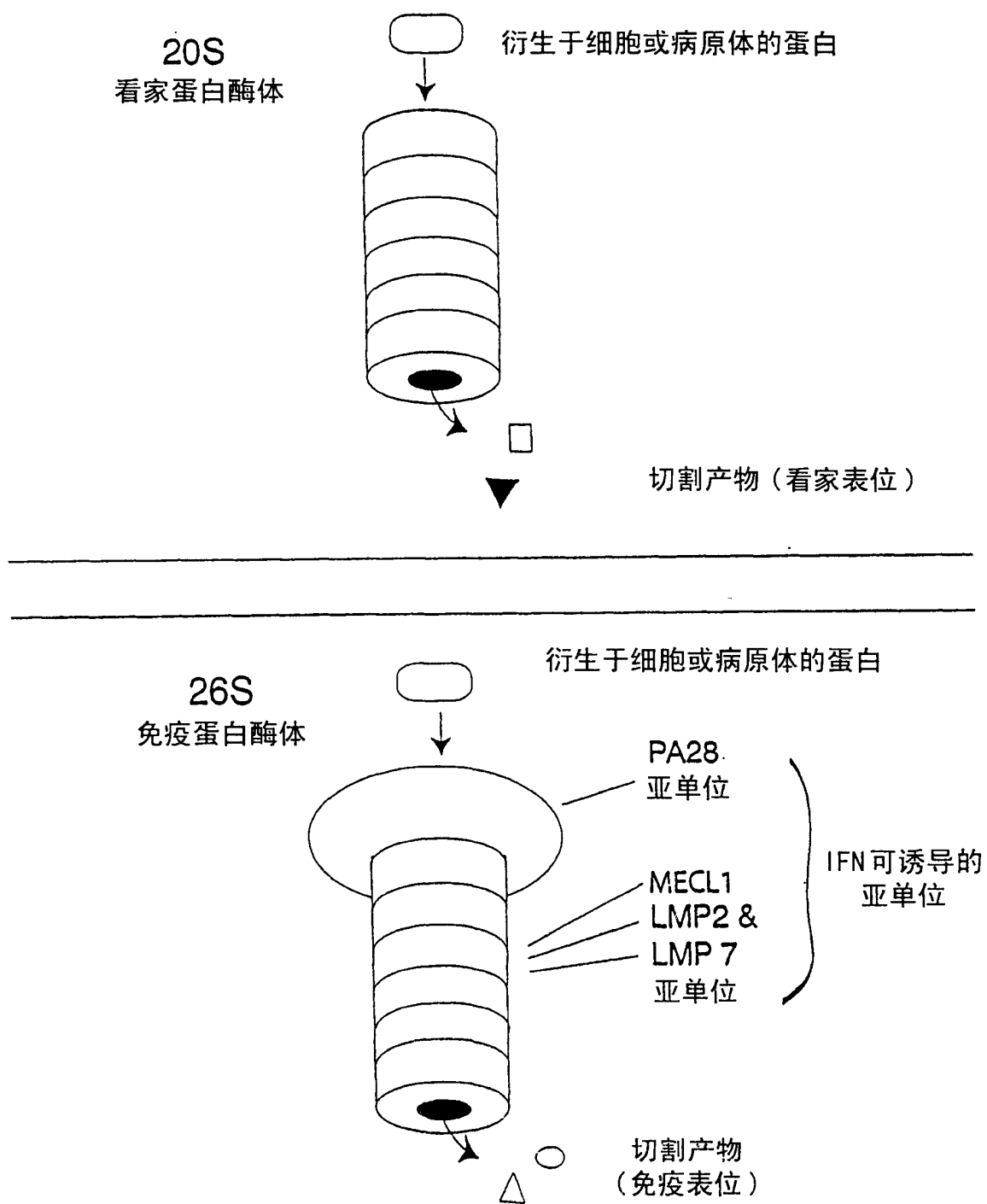


图 2

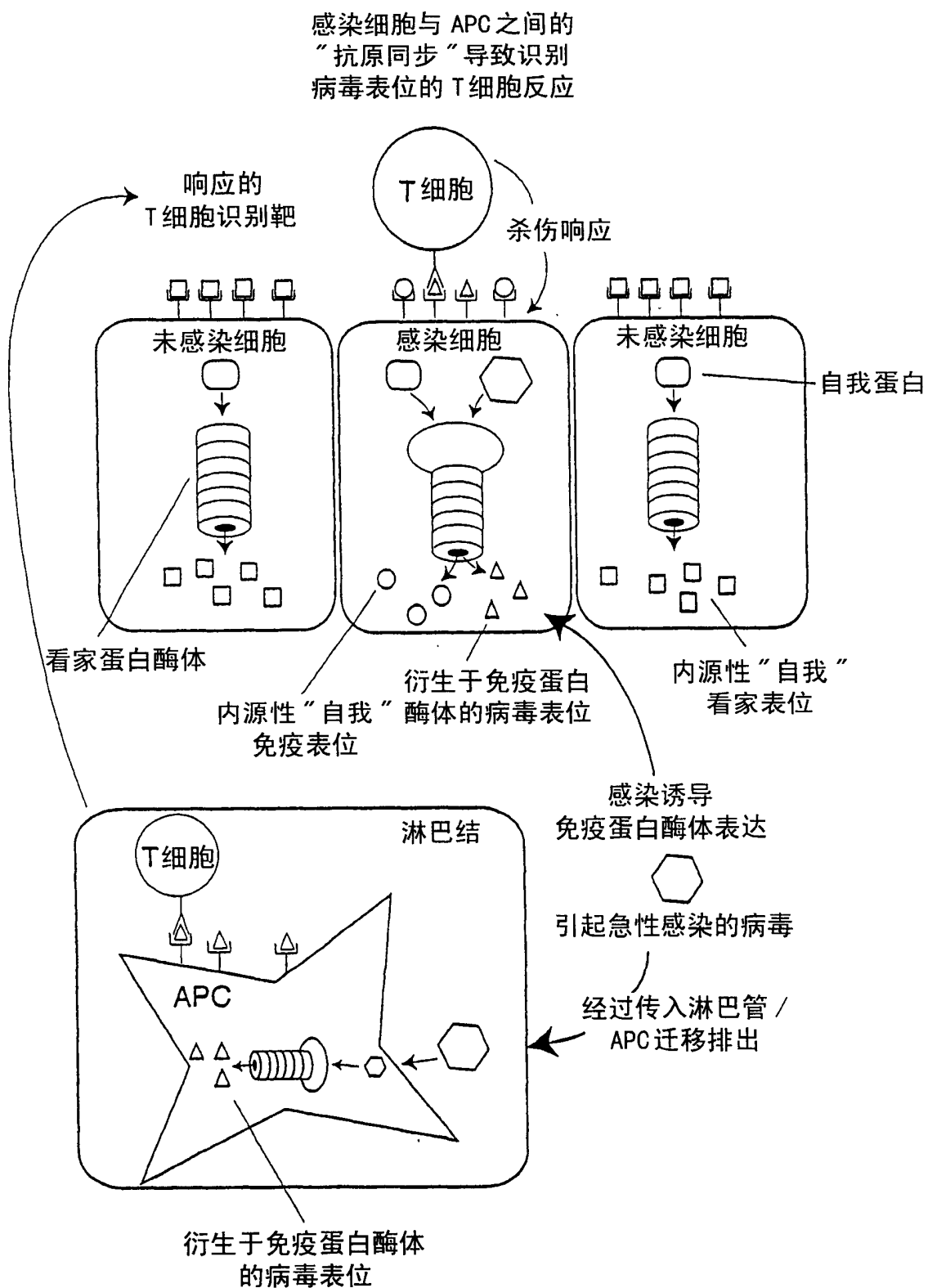


图 3

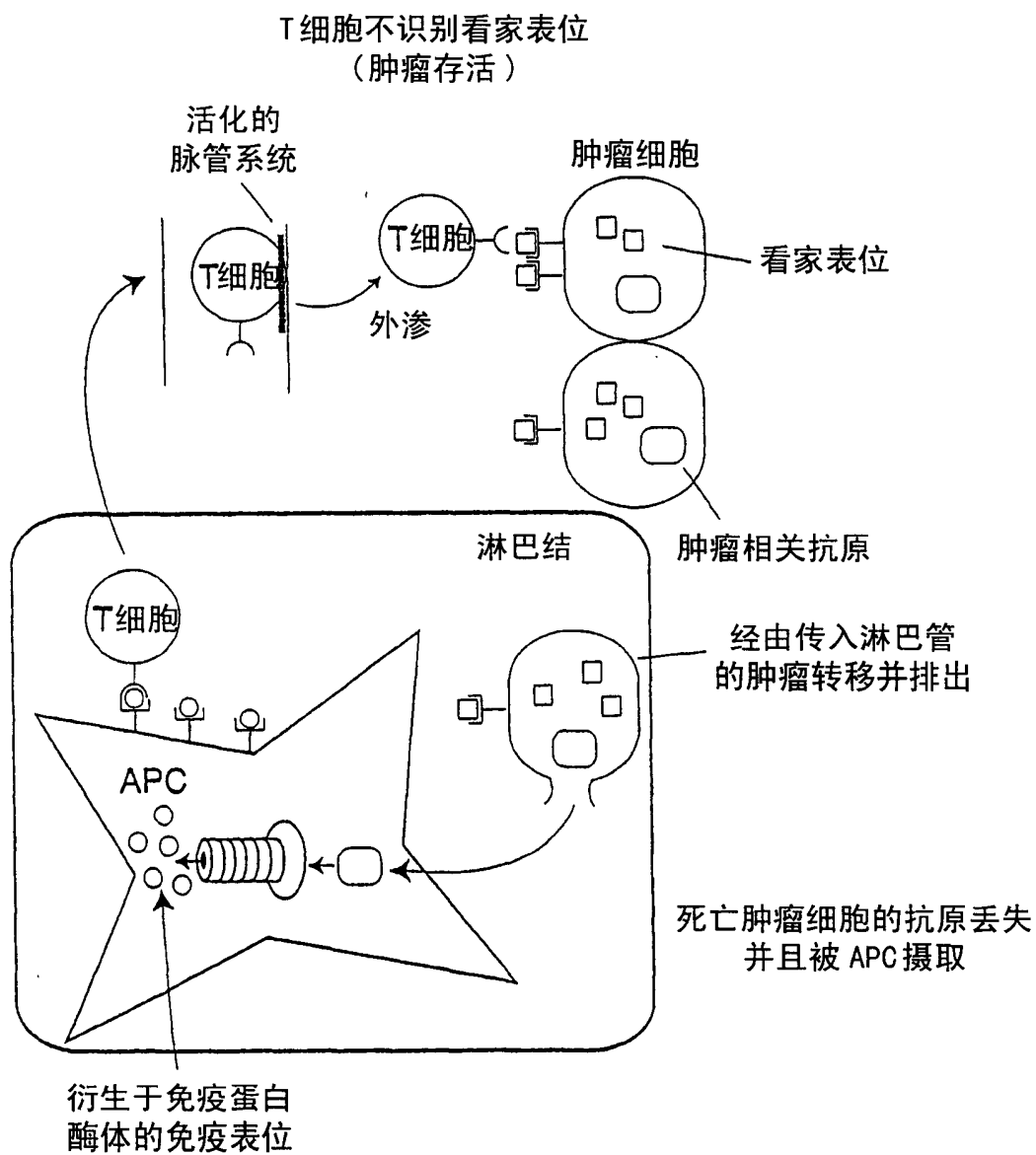


图 4

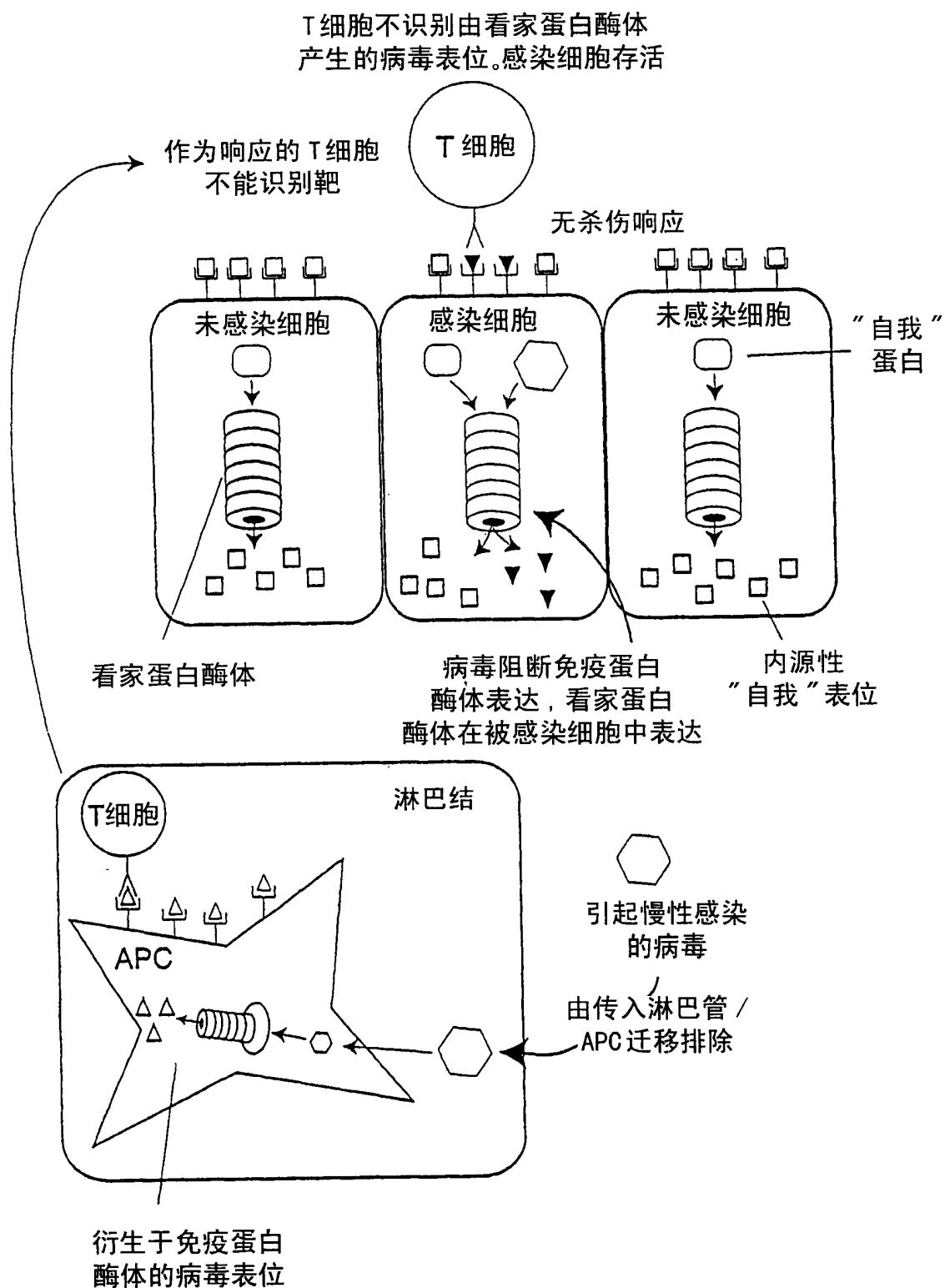


图 5

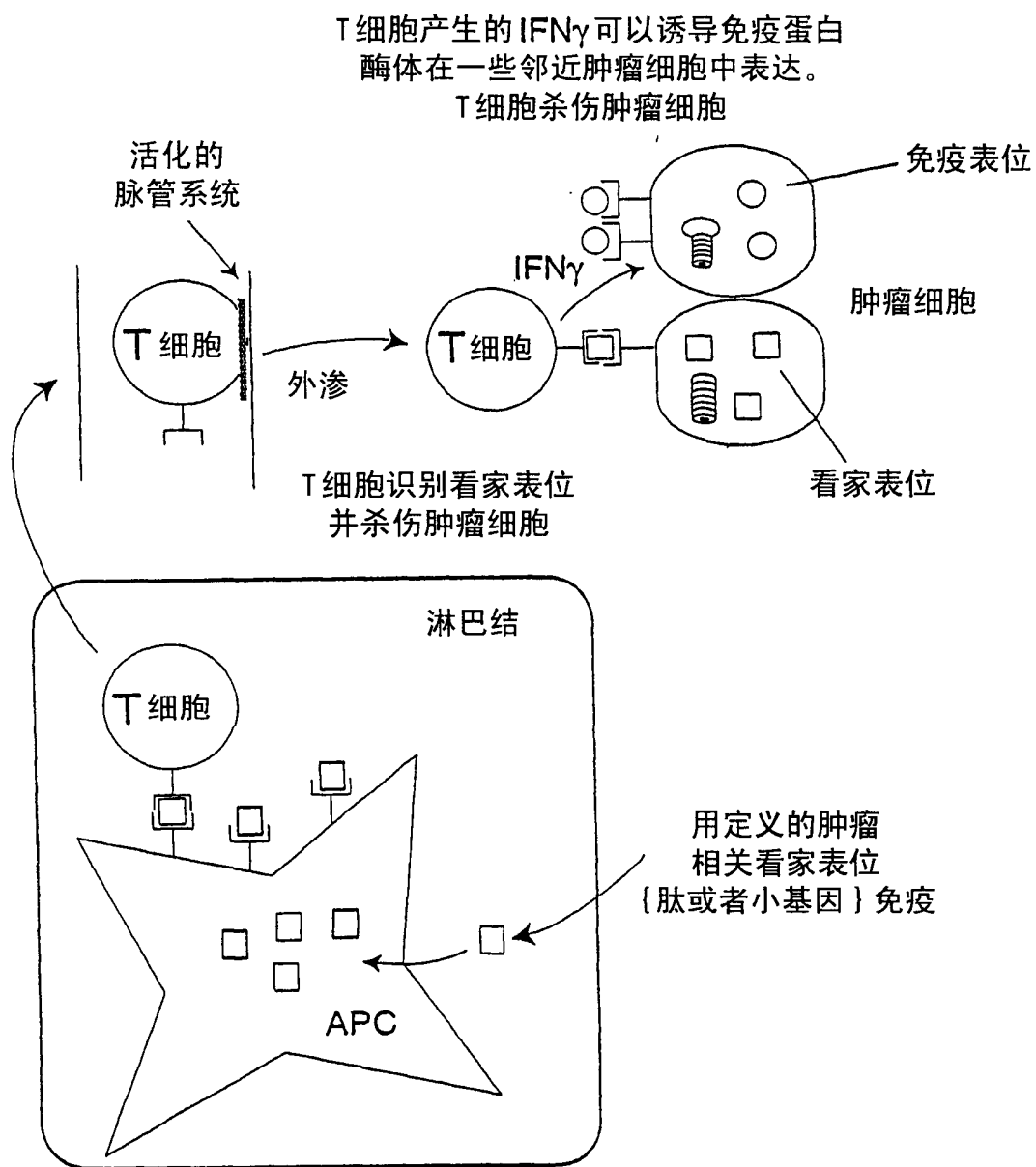


图 6

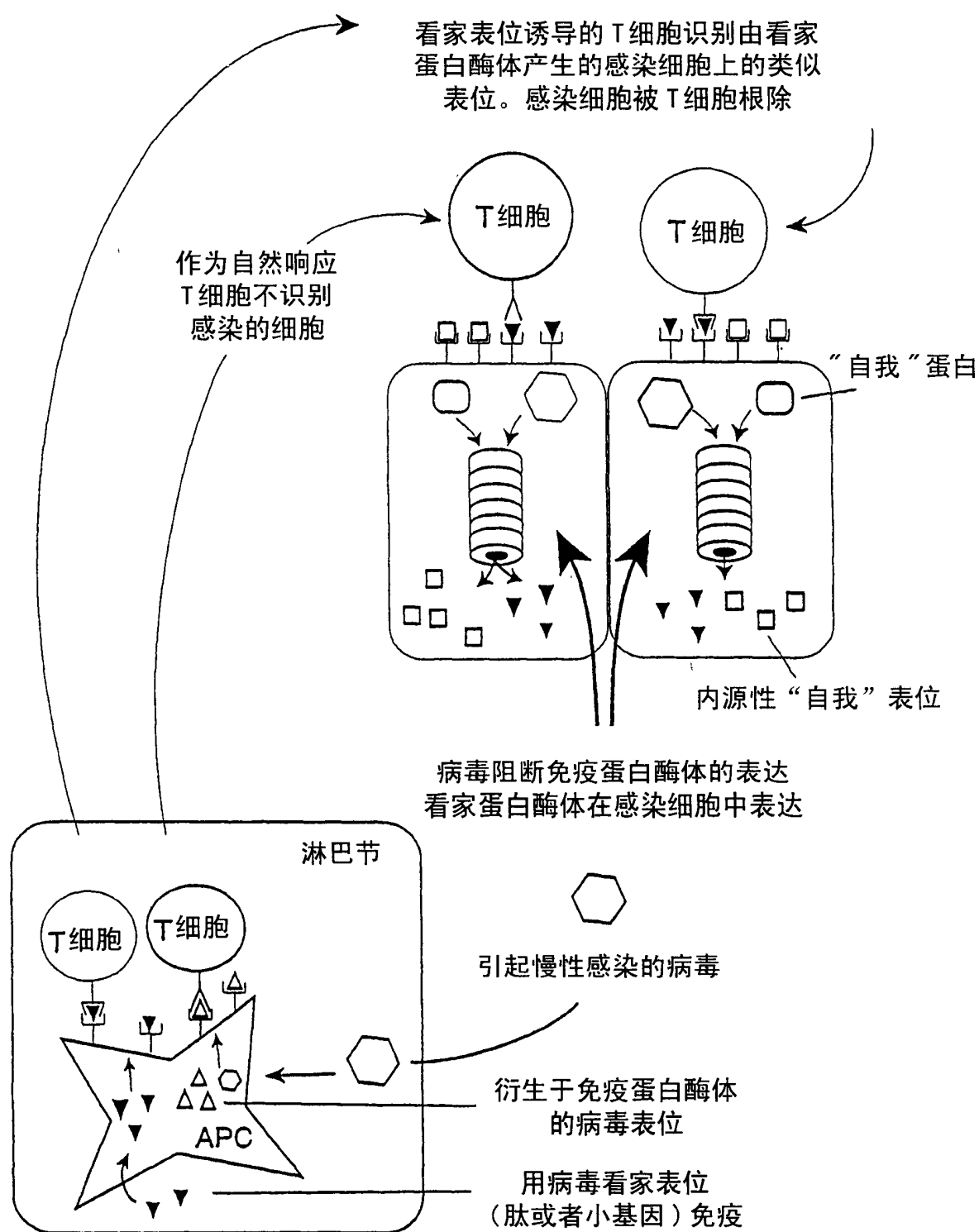


图 7

针对看家表位和免疫表位的双重攻击，T细胞杀伤所有肿瘤细胞

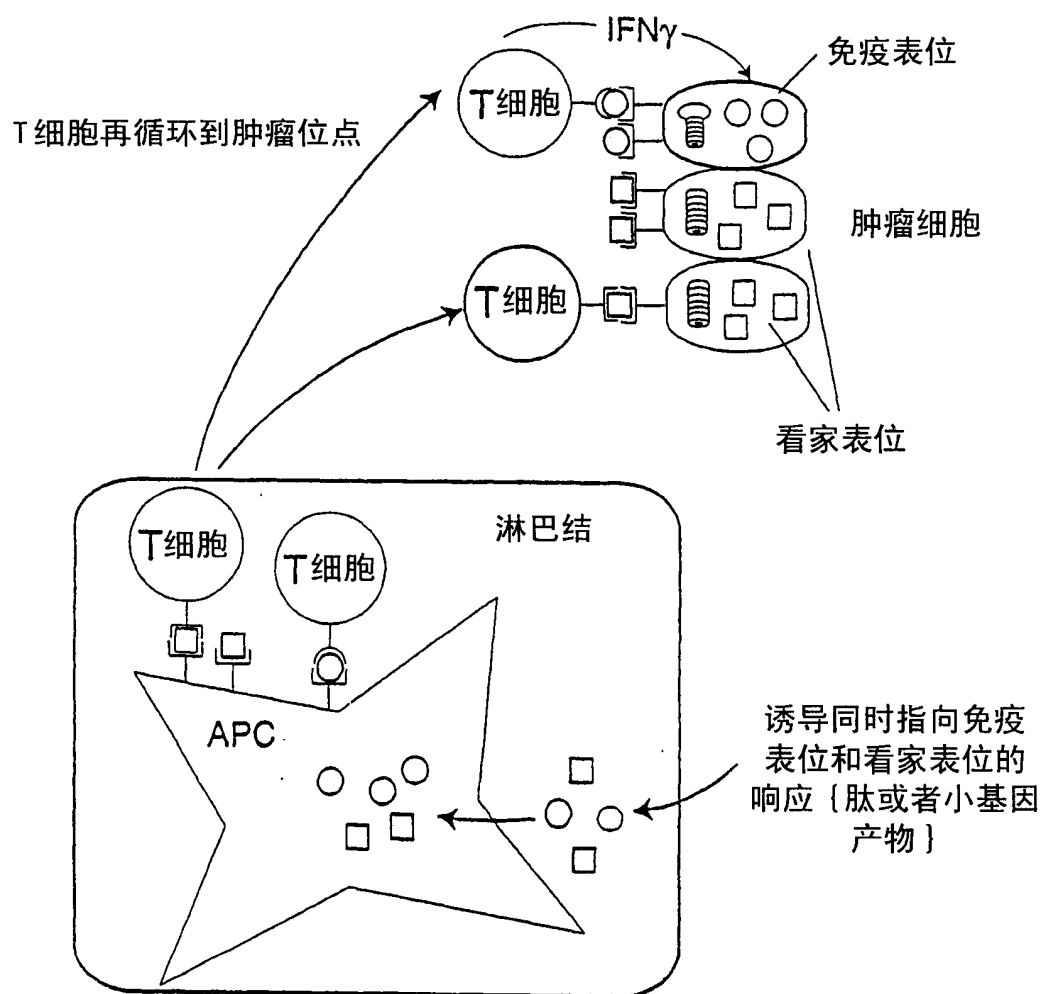


图 8

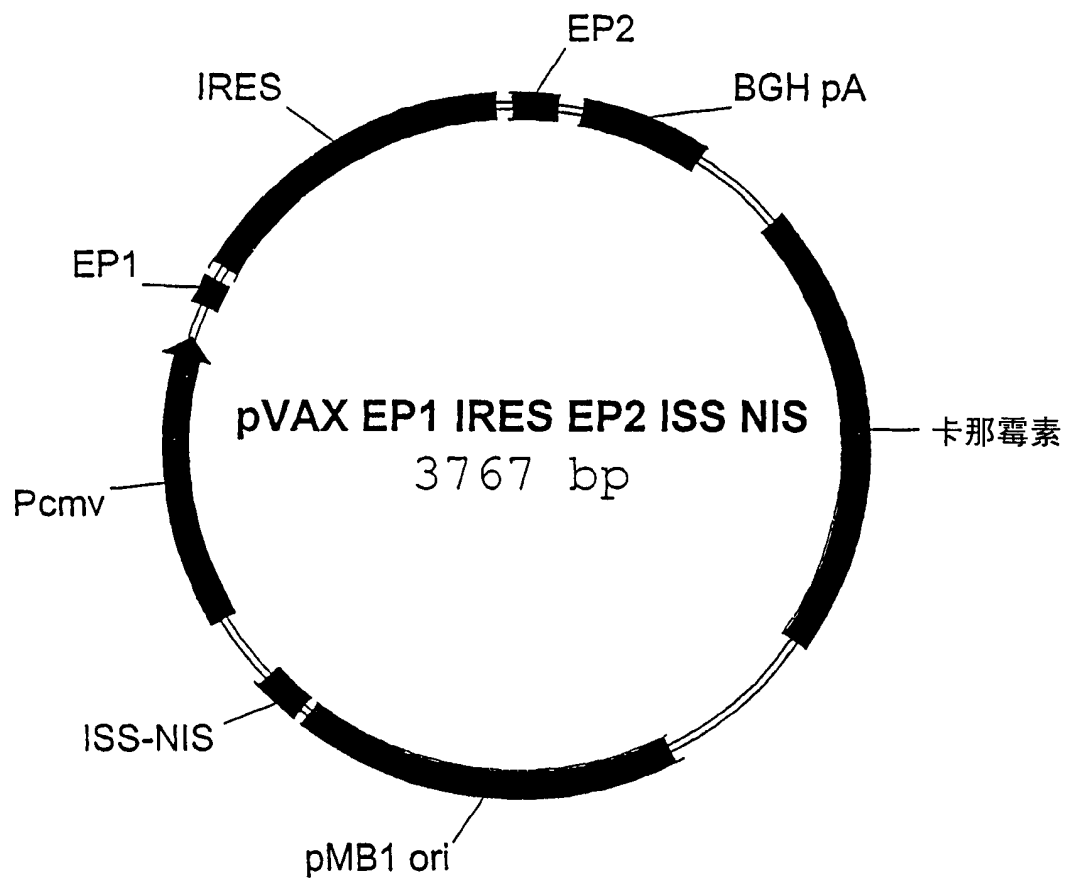


图 9

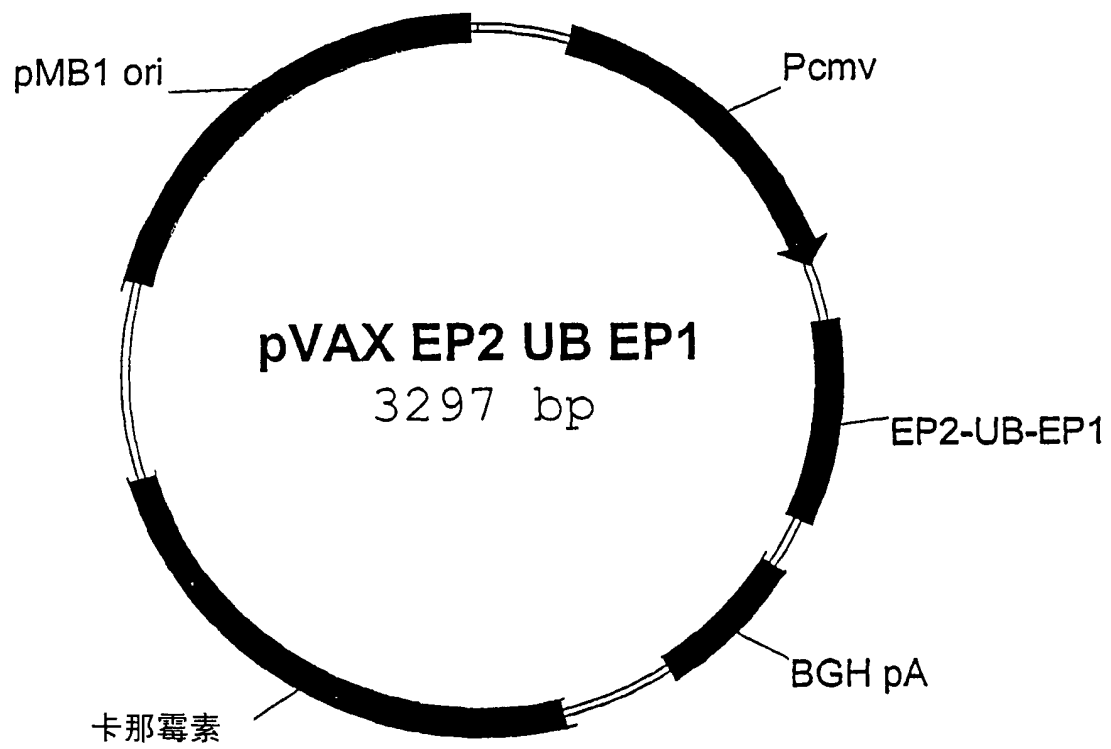


图 10

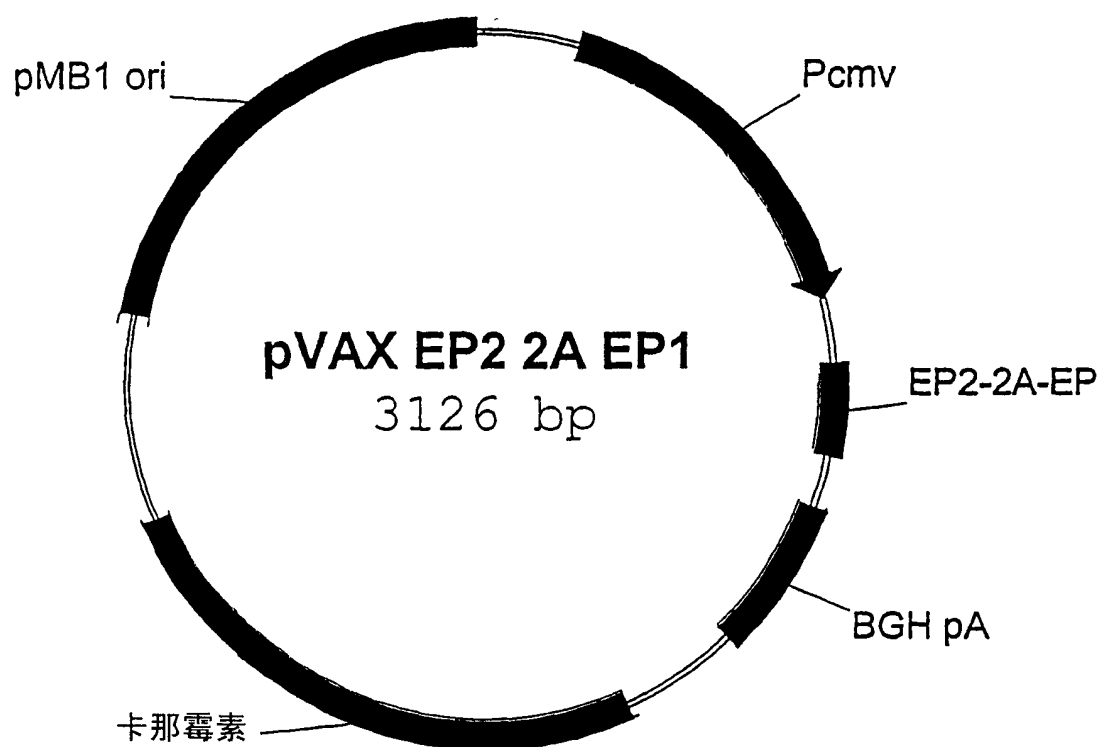


图 11

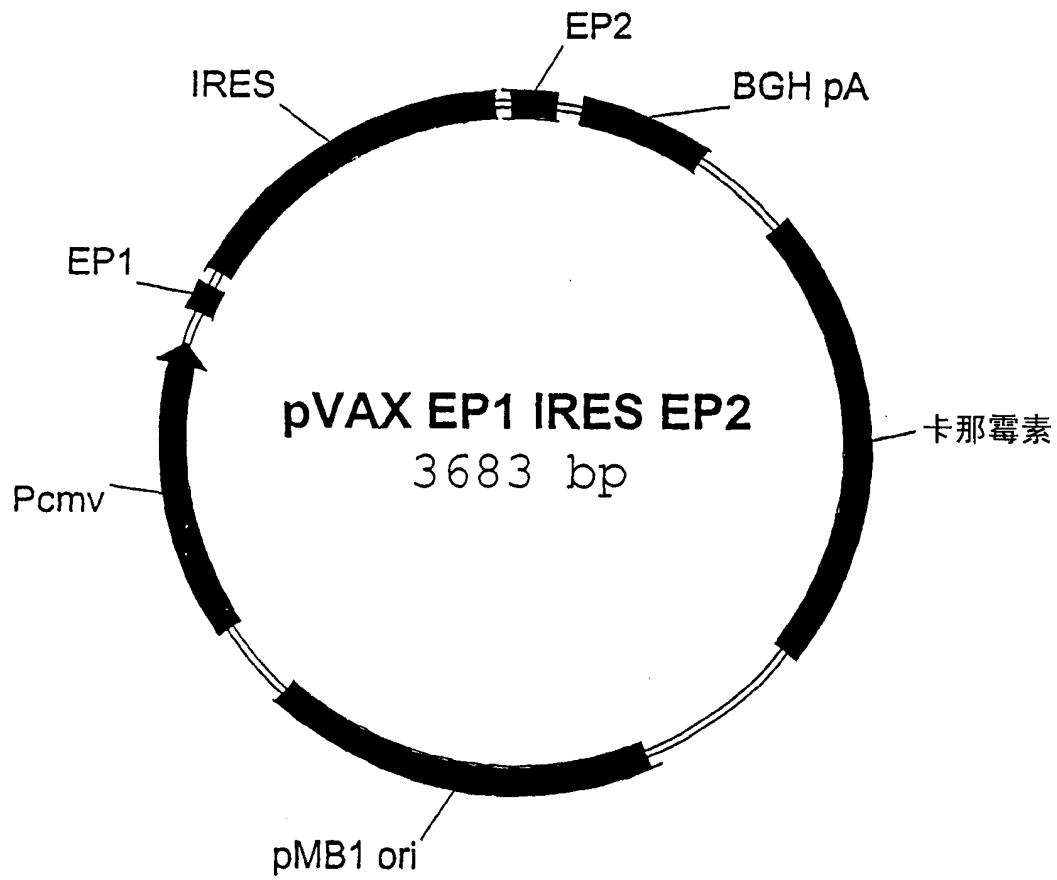


图 12

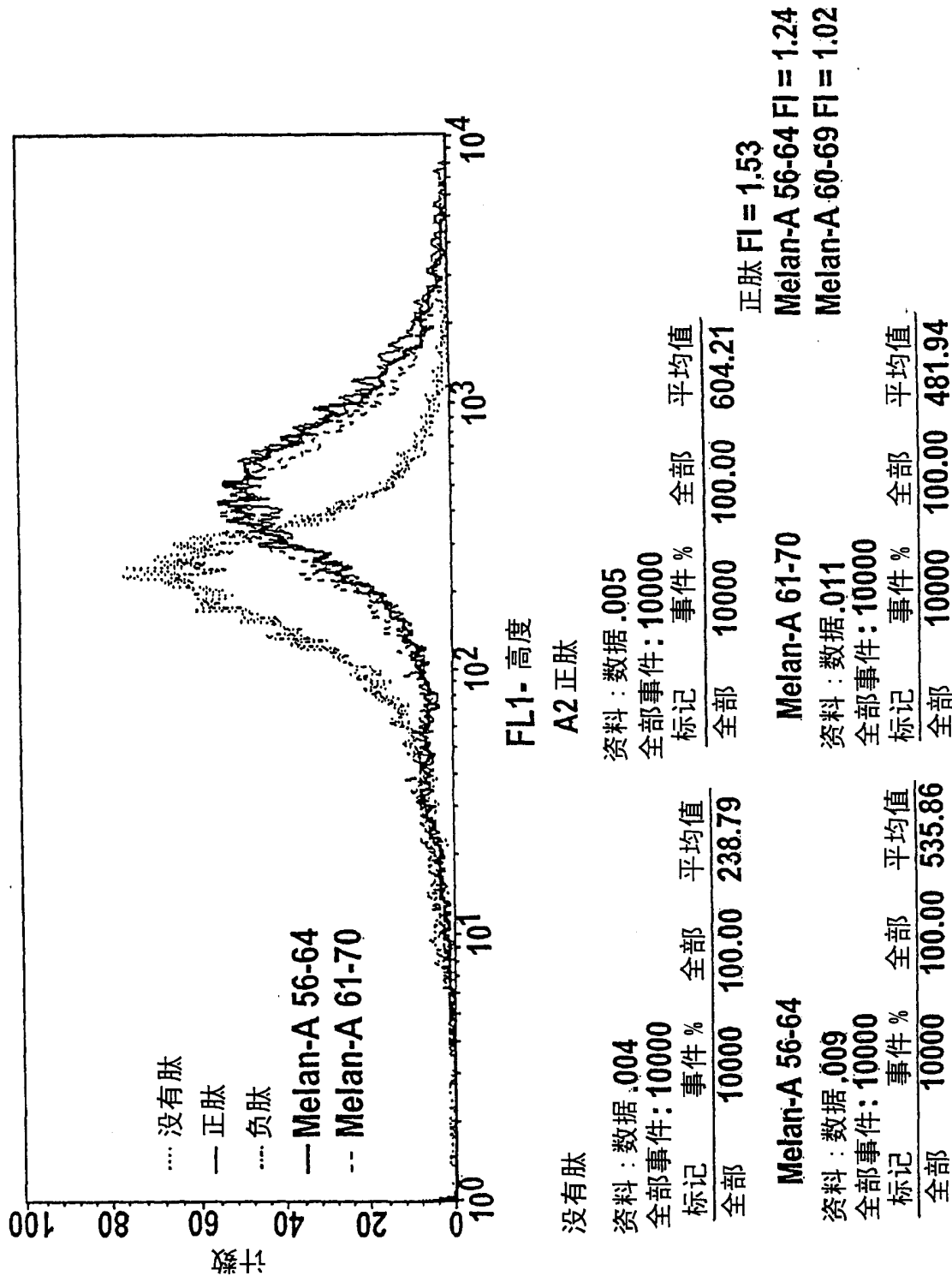


图 13

酪氨酸酶肽 207-216 结合 T₂ 细胞上的人类 MHC I
并且提高了它的表达

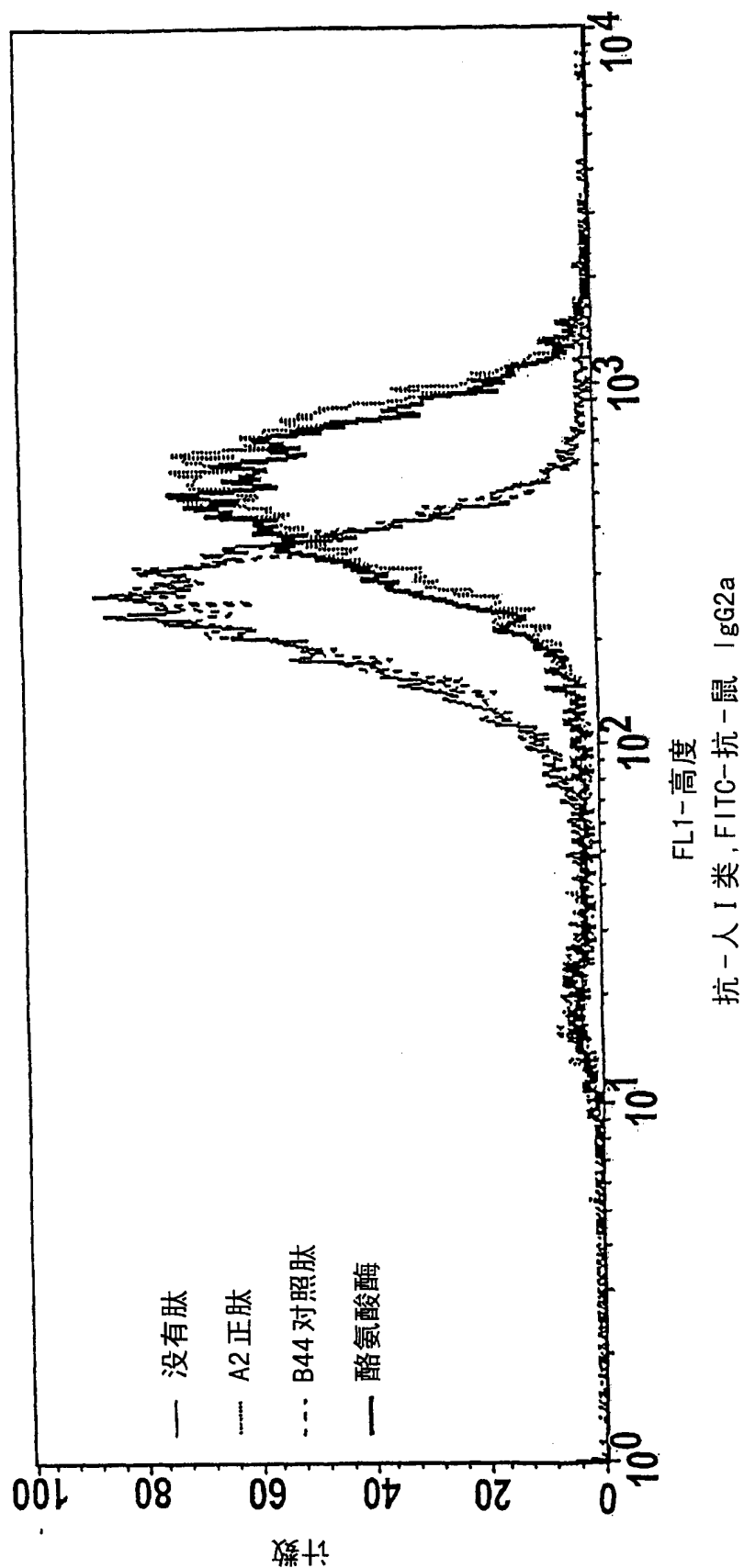


图 14

Melan-A 类 I HLA-A2 表位

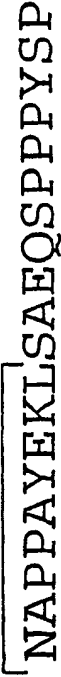
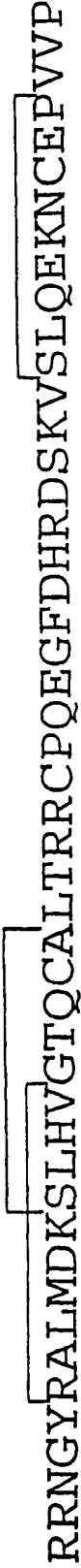
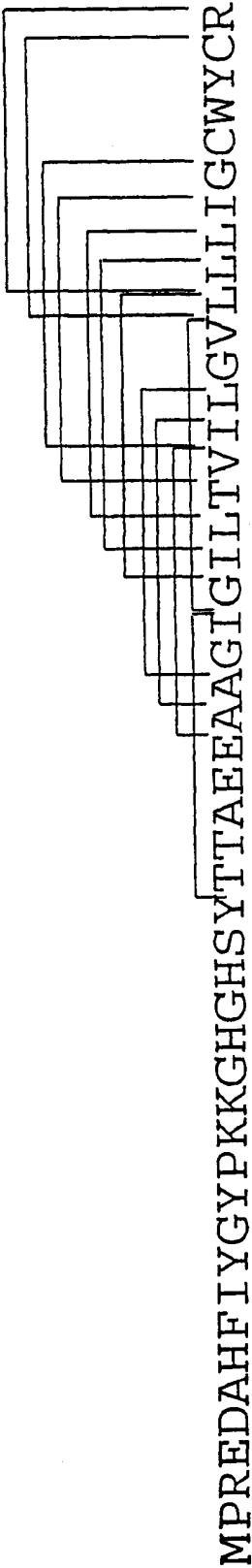


图 15

SSX-2

MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEEWEKM

KASEKIFYVMKRKYEAMTKLGFKATLPPFMCNKRAEDFQ

GNDLDNDPNRGNOVERPQMTFGRLQGISPKIMPKKPAEEG

NDSEEVPEASGPQNDGKELCPPGKPTTSEKIHERSGPKRG

EHAWTHRLRERKQLVYEEISDPPEEDDE

图 16

MQAEGRGTTG STGDADGPGG PGIPDGPGGN AGGPGEAGAT

GGRGPRGAGA ARASGPGGA PRGPHGGAAS GLNGCCRCCA

RGPEs RLLEFYLAM PFATPMEAE^{LARR}SLAQDAPPLPVP

GVLLKEFTVSGN ILTIRLTA^A DHR
QLQLSISSCLQ QLSLLMWIT QCFLPV FLAQ PPSGQRR

图 17

酪氨酸酶

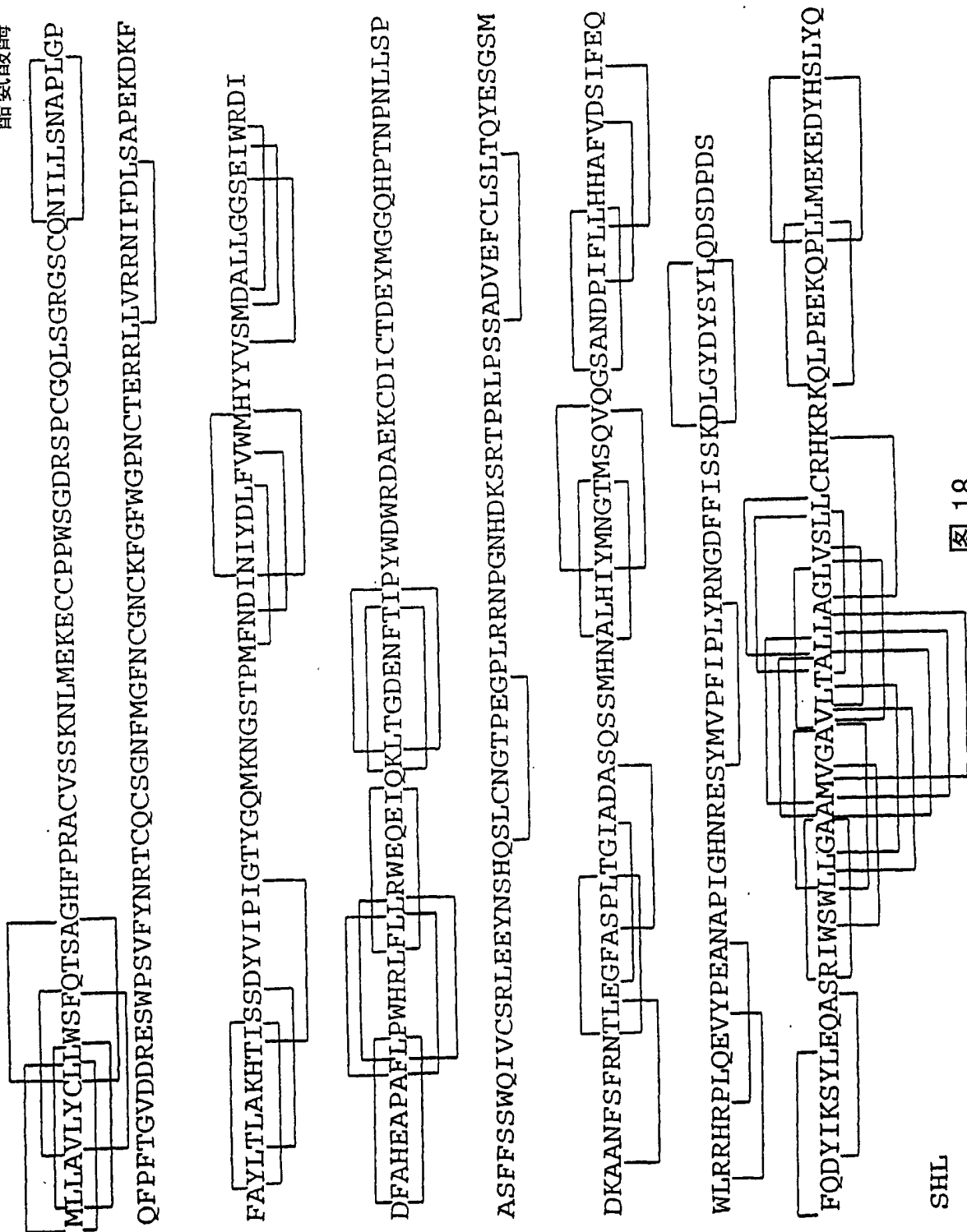


图 18