

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

C12N 15/12

C07K 14/47 A61K 38/17

C12Q 1/68 G01N 33/50

C07K 16/18

# [12] 发明专利申请公开说明书

[21] 申请号 00810302. X

[43] 公开日 2002 年 11 月 6 日

[11] 公开号 CN 1378594A

[22] 申请日 2000.7.12 [21] 申请号 00810302. X

[30] 优先权

[32]1999.7.13 [33]US [31]60/143,560

[32]1999.10.1 [33]US [31]60/157,471

[86] 国际申请 PCT/US00/19039 2000.7.12

[87] 国际公布 WO01/04309 英 2001.1.18

[85] 进入国家阶段日期 2002.1.14

[71] 申请人 美国政府健康及人类服务部

地址 美国马里兰

[72] 发明人 艾拉·帕斯坦 马格努斯·埃森德

比扬库克·李 乔治·瓦斯马茨斯

库尔特·沃尔夫冈

乌尔里希·布林克曼

[74] 专利代理机构 中原信达知识产权代理有限责任公司

代理人 丁业平 王维玉

权利要求书 6 页 说明书 64 页 附图 14 页

[54] 发明名称 T - 细胞受体  $\gamma$  可变阅读框架蛋白 (TARP) 及其应用

[57] 摘要

本发明提供了一种含有来自前列腺上皮细胞和很多乳房肿瘤细胞的 TCR $\gamma$  转录本的序列的核酸以及由这些序列翻译表达的 T - 细胞受体  $\gamma$  可变阅读框架蛋白 (“TARP”)。从 TARP 制备而来的疫苗对于增加对其中该蛋白被表达的细胞(包括前列腺癌细胞和很多乳房癌变的细胞)的免疫应答是有用的。本发明提供了诊断前列腺癌和表达 TARP 的乳房癌细胞存在的方法,以及将 TARP 和编码 TARP 的核酸给药于个体的方法。

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

- 5 1. 一种分离的多肽，其含有的氨基酸序列选自：TCR $\gamma$ 可变阅读  
框架蛋白（TARP），其免疫原性片段，与 TARP 有至少 90%序列同  
一性且被特异性识别 TARP 的抗体特异性识别的多肽，以及，与 TARP  
有至少 90%序列同一性且当被加工并在主要组织相容性复合物分子体  
系中被提呈后，激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽。
- 10 2. 权利要求 1 所述的分离的多肽，其中的多肽含有 TARP 的序  
列。
3. 权利要求 1 所述的分离的多肽，其中的多肽含有 TARP 的免  
疫原性片段的序列。
- 15 4. 权利要求 1 所述的分离的多肽，其中的多肽与 TARP 至少有  
90%的序列同一性且被特异性识别 TARP 的抗体特异性识别。
- 20 5. 权利要求 1 所述的分离的多肽，其中的多肽与 TARP 至少有  
90%的序列同一性，并且，当被加工并在主要组织相容性复合物分子  
体系中提呈后，激活 T 淋巴细胞抵抗表达 TARP 的细胞。
6. 一种含有权利要求 2 的多肽和一种可药用载体的组合物。
7. 一种含有权利要求 3 的多肽和一种可药用载体的组合物。
- 25 8. 一种含有权利要求 4 的多肽和一种可药用载体的组合物。
9. 一种含有权利要求 5 的多肽和一种可药用载体的组合物。
- 30 10. 一种分离的重组核酸分子，其含有编码具有 TCR $\gamma$ 可变阅读

5 框架蛋白(“TARP”)的氨基酸序列的多肽,其免疫原性片段,与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽,以及,与 TARP 至少有 90%序列同一性且其当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽的核苷酸序列。

11. 如权利要求 10 所述的分离的重组核酸分子,含有 TARP 的序列。

10 12. 如权利要求 10 所述的分离的重组核酸分子,其中的多肽是 TARP 的免疫原性片段。

15 13. 如权利要求 10 所述的分离的重组核酸分子,其中的多肽与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别。

14. 如权利要求 10 所述的分离的重组核酸分子,其中的多肽与 TARP 至少有 90%序列同一性且其当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞。

20 15. 如权利要求 10 所述的分离的重组核酸分子是一个含有可操作地连接于核苷酸序列的启动子的表达载体。

25 16. 如权利要求 15 所述的分离的重组核酸分子,其中所述的核苷酸编码具有 TCR $\gamma$ 可变阅读框架蛋白("TARP")的氨基酸序列的多肽。

17. 如权利要求 15 所述的分离的重组核酸分子,其中所述的核苷酸序列编码具有 TARP 的免疫原性片段的氨基酸序列的多肽。

30

18. 如权利要求 12 所述的分离的重组核酸分子，其中所述的核苷酸序列编码与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽。

5           19. 如权利要求 12 所述的分离的重组核酸分子，其中所述的核苷酸序列编码与 TARP 至少有 90%序列同一性且当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽。

10           20. 一种包括将一种组合物给药于个体的方法，其中的组合物选自以下成员组成的组：分离的具有 TCR $\gamma$ 可变阅读框架蛋白（“TARP”）的氨基酸序列的多肽，其免疫原性片段，与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽，与 TARP 至少有 90%序列同一性且当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽，一种分离的编码这些多肽之一的核酸，一种用含有 TARP 的表位的多肽刺激的抗原提呈细胞，以及，体外被 TARP、其免疫原性片段、与 TARP 有至少 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽或与 TARP 有至少 90%序列同一性且当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽致敏的细胞。

25           21. 权利要求 20 的方法，包括给药于个体 TARP 或其免疫原性片段。

            22. 权利要求 20 的方法，其中的多肽与 TARP 有至少 90%序列同一性且被特异性识别 TARP 的抗体特异性识别。

30           23. 权利要求 20 的方法，其中的多肽与 TARP 至少有 90%的序列同一性且当被加工并被结合有 MHC 分子的抗原提呈细胞提呈后激

活 T 淋巴细胞抵抗表达 TARP 的细胞。

24. 权利要求 20 的方法，其中给药于患前列腺癌的个体。
- 5 25. 权利要求 20 的方法，其中给药于患乳房癌的个体。
26. 权利要求 20 的方法，其中给药于没有进行过乳房癌诊断的雌性个体。
- 10 27. 权利要求 20 的方法，其中给药包括在体外使 CD8+细胞对 TARP 蛋白的表位敏化并将致敏细胞给药于个体。
28. 权利要求 20 的方法，进一步包括共同给药于个体选自下列的免疫佐剂：非特异性的免疫佐剂，亚细胞微生物产物和成分，半抗原，免疫原性蛋白，免疫调节剂，干扰素，胸腺激素和集落刺激因子。
- 15 29. 权利要求 20 的方法，包括给药用含有 TARP 表位的多肽脉冲的抗原提呈细胞。
- 20 30. 权利要求 20 的方法，包括给药编码含有 TARP 表位的多肽的核酸序列，其中的核酸包含在重组病毒中。
31. 权利要求 20 的方法，包括给药编码含有 TARP 蛋白的表位的多肽的核酸序列。
- 25 32. 权利要求 20 的方法，包括给药表达含有 TARP 蛋白的表位的多肽的表达载体，其中的表达载体是在重组细菌细胞中。
- 30 33. 权利要求 20 的方法，包括用表达含有 TARP 蛋白的表位的多肽的表达载体免疫个体，其中的表达载体在自体重组细胞中。

34. 权利要求 27 的方法，其中的 CD8+细胞为 Tc 细胞。

35. 权利要求 34 的方法，其中的 Tc 细胞为肿瘤浸润淋巴细胞。

5

36. 一种在雄性中检测上皮来源的前列腺细胞或在雌性中检测乳房癌细胞的方法，包括检测来自所述雄性或所述雌性的细胞中的编码 TARP 的核酸转录本，或检测通过转录本的翻译产生的 TARP，其中在来自所述的男性的细胞中检测到所述的转录本或蛋白则该细胞被鉴定为前列腺上皮细胞，在来自所述的女性的细胞中检测到所述的转录本或蛋白则该细胞被鉴定为乳房癌细胞。

10

37. 如权利要求 36 所述的方法，包括检测所述的转录本。

15

38. 如权利要求 36 所述的方法，包括检测所述的蛋白。

39. 如权利要求 36 所述的方法，包括用在杂交条件下特异性杂交于转录本的核酸探针接触来自所述细胞的 RNA，并检测杂交。

20

40. 如权利要求 36 所述的方法，包括裂解所述细胞并用含有靶向元件和可检测标记的嵌合分子接触细胞内容物的部分，其中的靶向元件特异性结合于所述蛋白，并检测结合于蛋白的标记物。

41. 如权利要求 36 所述的方法，其中的细胞取自淋巴结。

25

42. 如权利要求 36 所述的方法，其中的细胞取自乳房活组织。

43. 一种特异性结合于 TCR $\gamma$ 可变阅读框架蛋白的抗体。

30

44. 一种在细胞中调节 TARP 水平的方法，包括向所述细胞导入

一种选自由以下成员组成的组中的组合物：特异性裂解编码 TARP 的核酸的核酶，特异性结合于编码 TARP 的核酸的反义寡核苷酸，特异性结合于与编码 TARP 的核酸结合的蛋白的 DNA，以及可操作地连接于启动子的编码 TARP 的核酸。

5

T-细胞受体 $\gamma$ 可变阅读框架蛋白(TARP) 及其应用

## 5 相关申请的引用

本申请要求 1999 年 7 月 13 日申请的 US 临时申请 No.60/143,560, 以及 1999 年 10 月 1 日申请的 US 临时申请 No. 60/157,471 的优先权。此两个申请的内容通过引用在此引入本文。

10 关于联邦资助研究和发展的发明的权利的声明  
是不适用的。

## 发明背景

15 本发明涉及分子生物学以及医学诊断学和治疗学领域。前列腺癌是人类最普遍的一种癌症，是导致人的因癌症死亡的第三大原因。此种疾病的早期检测和治疗的方法将降低前列腺癌患者的死亡率。由恶性肿瘤细胞表达而其它细胞很少表达的肿瘤相关蛋白，可被用作靶来用于检测和干预。几种与前列腺癌相关的蛋白已被鉴定，包括前列腺特异性抗原 (PSA)。

20 免疫疗法是一种用于抵抗癌症的有效的新武器。免疫疗法包括基于其靶抗原的产生来激发抵抗癌细胞的免疫应答。当使用人抵抗癌细胞的免疫应答时，优选用于激发细胞介导的抵抗癌细胞的免疫应答。基于细胞介导的免疫应答的免疫疗法包括对产生特定的表位的细胞产生细胞介导的应答。癌细胞产生不同的可成为免疫治疗的靶的蛋白。某些癌产生新的蛋白，例如作为突变的结果，它们是免疫原性的。然而，研究者们也发现肿瘤浸润淋巴细胞可特异性地识别癌细胞的未突变蛋白。例如，Rosenberg 等人已发现肿瘤浸润淋巴细胞靶向和识别正常黑素瘤细胞和恶性黑素瘤细胞均产生的蛋白 MART-1 的抗原决定基。此外，抗 MART-1 或其与 MHC I 类分子（具体的，HLA A2 的抗

25

30

原决定基)结合的肽的主动或被动免疫治疗导致产生 MART-1 的黑素瘤细胞以及正常细胞的破坏。Y. Kawakami 等人, *J. Immunol.* 21: 237 (1998)。

5           新的癌抗原被期望能激发免疫反应, 因为免疫系统将识别其为异己蛋白。然而, 因为免疫系统对自身蛋白具有耐受性, 其激发抵抗非突变的自身蛋白的免疫应答的能力是令人吃惊的。人们确信这种免疫应答直接抵抗通常不以足够的量暴露于免疫系统的来激发耐受性或免疫应答的表位。然而在癌中, 这些决定簇不再避开免疫系统的检测。  
10           这点可能由通过 MHC I 类分子对表位增加的提呈产生的。

          细胞介导的免疫应答涉及主要组织相容性复合物分子的活性。在人体中, 复合物被称为“HLA”(“人白细胞抗原”)复合物。在小鼠中, 其是指“H-2”复合物。主要组织相容性复合物包括 3 类蛋白,  
15           MHC I 类、MHC II 类和 MHC III 类。MHC I 类分子几乎在所有有核细胞的表面上表达。其提呈抗原肽到 T<sub>c</sub> 细胞 (CD8+)。在人体中有 3 个 MHC I 类基因位点, HLA A, HLA B 和 HLA C。每一位点为高度多态的。因此, 一个人在其细胞的表面可能有六种不同类型的 HLA 分子。MHC II 类蛋白主要在抗原提呈细胞例如巨噬细胞、树突状细胞  
20           和 B 细胞上表达, 其提呈被加工过的抗原肽到 T<sub>H</sub> 细胞。在人体中有 3 个 MHC II 类基因位点, HLA DP, HLA DQ 和 HLA DR。MHC III 类蛋白与各种的免疫过程相关, 包括可溶性血清蛋白, 以及补体系统的组分和肿瘤坏死因子。J. Kuby, Chapter 9, IMMUNOLOGY, 第三版 W. H. Freeman and Company, New York (1997)。

25           在癌细胞和正常细胞中, MHC I 类分子将自身蛋白的表位提呈到 T<sub>c</sub> 细胞。HLA A, HLA B 和 HLA C 分子结合大约长为 8 到 10 氨基酸的具有特殊锚定残基的肽。HLA I 类分子识别的锚定残基依赖于 HLA 分子的特定等位形式。CD8+ T 细胞带有能识别当被在细胞上的特定  
30           HLA 分子提呈时细胞上的特异性表位的 T 细胞受体。当能被抗原提呈

细胞刺激变为细胞毒性 T 淋巴细胞的 Tc 细胞接触带有 HLA-肽复合物的细胞时，CTL 与此细胞形成结合物并将其破坏。

5 肽通过 MHC I 类分子的提呈包括内源产生的蛋白通过蛋白酶体裂解为肽，通过 ER 和高尔基体进行加工，通过肽的锚定残基结合到 MHC I 类分子的裂口以及最终提呈到细胞的表面。除其它因素外，依据特定的锚定残基，某些肽可更紧密的结合于特定的 HLA 分子。结合较好的肽指“显性”表位，而结合较差的肽成为“亚显性”或“阴性”表位，自身蛋白或外源蛋白的显性表位激发强烈的耐受性或免疫应答。亚显性或阴性的表位产生弱的或根本不产生免疫应答。假设显性表位与 HLA 分子的紧密结合导致它们在细胞表面密集提呈，产生更多的与免疫细胞反应的机会以及更多的激发免疫应答或耐受性的可能性。

15 研究表明在自身蛋白 MART-1 蛋白存在的情况下，免疫系统产生最多的抵抗亚显性或阴性表位的 CTL 应答。Y. Kawakami 等, 1997 Immunol. Res. 16: 313。可能在癌细胞中亚显性或阴性表位比在正常细胞中提呈得密集或量更大；所以免疫系统遇见这些先前未检测到的表位时，将其识别为外源的并产生抵抗它们的免疫应答。无论什么原因，将免疫系统暴露于自身蛋白的大量的亚显性或阴性表位，作为一种激发免疫应答抵抗产生所述蛋白的细胞的方式是癌症免疫治疗的关键要素。当然，启动抗自身蛋白的免疫应答将对癌性细胞和健康细胞二者均产生破坏。因此，为使免疫治疗获得成功，健康细胞必须是维持生命不必需的或具有通过其它疗法可取代的功能。在前列腺癌和乳房癌中，通常的治疗手段是分别将前列腺或乳房手术切除。在此情况下，大多数或所有显示出前列腺或乳房抗原的组织将成为肿瘤细胞。

#### 发明概述

30 我们已发现健康和癌性的上皮来源的前列腺细胞，转录未重排的 TCR $\gamma$ 基因的一部分。而在体外，此转录产生两种蛋白，一种为 TCR $\gamma$

的截短形式，令人吃惊的是，我们的研究显示，在体内，表达的蛋白不是 TCR $\gamma$ 蛋白，而是可变阅读框架表达的蛋白。相应地，此蛋白被称为 TCR $\gamma$ 可变阅读框架蛋白或“TARP”。

5 我们已经发现 TARP 在上皮来源的前列腺细胞和前列腺癌细胞中表达。令人吃惊的是，我们发现相同的蛋白表达在许多的乳房癌细胞中。TARP 因此可作为前列腺癌和表达 TARP（表达 TARP 的乳房癌）的乳房癌细胞的标记，并作为免疫治疗的基础。本发明提供了分离的或重组形式的核酸和蛋白。虽然我们已发现此种蛋白在许多乳房癌细胞中被表达，因其被首先在前列腺细胞中鉴定，在下面通常被称为前列腺特异性（“PS”）—TCR $\gamma$ 。

10 在一方面，本发明提供了分离的含有 TCR $\gamma$ 可变阅读框架蛋白（“TARP”）的氨基酸序列的多肽。本发明进一步提供了 TARP 的免疫原性片段（此片段可产生能特异性识别和结合于全长 TARP 的抗体或其激活识别表达 TARP 的细胞的 T 细胞），与 TARP 至少有 90% 序列同一性且其被特异性识别 TARP 的抗体特异性识别的多肽，以及，与 TARP 至少有 90% 序列同一性且其当被加工并在主要组织相容性复合物分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽。本发明进一步提供了组合物，其含有 TARP，其免疫原性片段，或具有至少 90% 序列同一性且满足上述功能条件的肽，并含有可药用载体。

25 在另一个实施方案中，本发明提供了分离的重组核酸分子，其编码具有 TCR $\gamma$ 可变阅读框架蛋白（“TARP”）的氨基酸序列的多肽，编码其免疫原性片段，编码与 TARP 至少有 90% 序列同一性且其被特异性识别 TARP 的抗体特异性识别的多肽，或者，编码与 TARP 至少有 90% 序列同一性且其当被加工并在主要组织相容性复合体分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽。

30

在另一个实施方案中，本发明提供了一种方法，包括给药于个体一种组合物，该组合物选自：分离的具有 TCR $\gamma$ 可变阅读框架蛋白（“TARP”）的氨基酸序列的多肽，其免疫原性片段，与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽，以及，与 TARP 至少有 90%序列同一性且其当被加工并在主要组织相容性复合体分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽。

本发明进一步提供了给药于个体一种组合物的方法，其中的组合物选自：编码 TARP、其免疫原性片段、与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽、以及与 TARP 至少有 90%序列同一性且当被加工并在主要组织相容性复合体分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽的分离的核酸。

本发明进一步提供了将组合物给药于个体的方法，其中的组合物含有用包含 TARP 表位的多肽刺激的抗原提呈细胞。另外，本发明提供了给药于个体一种组合物的方法，其中的组合物含有在体外被 TARP、与 TARP 至少有 90%序列同一性且被特异性识别 TARP 的抗体特异性识别的多肽、或者与 TARP 至少有 90%序列同一性且当被加工并在主要组织相容性复合体分子体系中被提呈后激活 T 淋巴细胞抵抗表达 TARP 的细胞的多肽致敏的细胞。

上述方法中所述的组合物可被给药于患前列腺癌的个体，患乳房癌的个体，或未进行乳房癌诊断的雌性个体。

该方法还包括在体外致敏 CD8<sup>+</sup>细胞使其对 TARP 蛋白的表位敏感并将这些致敏细胞给药于个体。CD8<sup>+</sup> 细胞可为 Tc 细胞。在一些实施方案中，Tc 细胞可为肿瘤浸润淋巴细胞。

此外，该方法可包括共同给药于个体免疫佐剂，其选自非特异性的免疫佐剂，亚细胞微生物产物和成分，半抗原，免疫原性蛋白，免疫调节剂，干扰素，胸腺激素和集落刺激因子。

5            该方法也进一步包括给药一种用含有 TARP 表位的多肽刺激的抗原提呈细胞，或给药编码含有 TARP 的表位的多肽的核酸序列，其中的核酸存在于重组病毒中。该方法进一步包括给药编码含有 TARP 蛋白的表位的多肽的核酸序列。

10           此外，该方法包括给药一种表达载体，其表达含有 TARP 蛋白的表位的多肽，其中表达载体存在于在重组细菌细胞中。进一步地，该方法包括用表达含有 TARP 蛋白表位的多肽的表达载体免疫个体，其中的表达载体包含在自体重组细胞中。

15           在另一方面，本发明提供了一种检测方法，在雄性中为检测上皮来源的前列腺细胞，雌性中为检测乳房癌细胞，包括检测细胞中的所述雄性或所述雌性的编码 TARP 的核酸转录本，或检测转录本的翻译所产生的 TARP，因此在所述的雄性细胞中检测到所述的转录本或蛋白则该细胞为前列腺上皮细胞，而在所述雌性的细胞中检测到所述的蛋白或转录本则该细胞为乳房癌细胞。该方法可包括将来自所述细胞的 RNA 与在杂交条件下特异性杂交于转录本的核酸探针接触，并检测杂交。此外，该方法包括破坏细胞并使一部分细胞内容物与含有靶向部分和可检测标记的嵌合分子接触，其中的靶向部分特异性地结合于所述的蛋白，检测结合于蛋白的标记。靶向部分本身也可被标记（例如，抗体如 scFv 可具有放射性残基）。

20

25

被检测的细胞可取自淋巴结，或乳房活组织。

最后，本发明提供了特异性结合于 TCR $\gamma$ 可变阅读框架蛋白的表位的抗体。

30

### 附图的简要描述

附图 1. PS-TCR $\gamma$ 转录本的核苷酸序列 (SEQ ID NO: 13)。在体外翻译系统中, 此转录本产生两种多肽。第一种多肽具有推导的氨基酸序列, 开始于 MQM... ("PS-TCR $\gamma$ -1", 现在称为 "TARP", SEQ ID NO: 14)。据估计此多肽的分子量为 7.2 kD。其由与自然 TCR $\gamma$ 阅读框架不一致的阅读框架翻译。第二种多肽具有开始为 MKT...的推导氨基酸序列 ("PS-TCR $\gamma$ -2", SEQ ID NO: 15)。据预测, 其分子重为 13 kD。其由与 TCR $\gamma$ 相同的阅读框架翻译且表现为 TCR $\gamma$ 的截短形式。图中的单下划线表示转录起始位点和多聚腺苷酸位点。

附图 2. TCR $\gamma$  mRNA 表达的杂交分析。附图 2A) 多种组织的点杂交印迹显示人 TCR $\gamma$ 的不同的表达。阳性组织为前列腺 (C7), 小肠(E3), 脾 (E4), 胸腺(E5), 外周白细胞 (E6), 淋巴结 (E7), 骨髓(E8) 以及肺(F2)。附图 2B, Northern 印迹显示正常组织中 TCR $\gamma$ 转录本的大小。在前列腺中表达的 TCR $\gamma$ 转录本为 1.1 kb 和 2.8 kb, 而在脾脏、胸腺和外周白细胞中的主要转录本为 1.5 kb。底片被曝光 20 小时。

附图 3. TCR $\gamma\delta$ 表达的 Northern 印迹分析。附图 3A) TCR $\gamma$ 恒定区 (TCR C $\gamma$ ) cDNA 探针显示了 1.1 和 2.8 kb 前列腺特异性转录本 (与附图 2B 比较)。底片被曝光 20 小时。附图 3B) TCR $\delta$ 恒定区 (TCR C $\delta$ ) cDNA 探针显示 TCR $\delta$  mRNA 不在前列腺中表达, 而在脾脏、胸腺和外周白细胞中表达。底片被曝光 50 小时。附图 3C) TCR C $\gamma$  cDNA 探针表明 LNCaP 细胞系表达 TCR $\gamma$ 而 PC-3 细胞系不表达。底片被曝光 20 小时。人 $\beta$ -肌动蛋白 mRNA 的表达被分析作为对照。

附图 4. 石蜡包埋的组织通过 TCR $\gamma$ (C $\gamma$ 1-3' UTR)反义、<sup>35</sup>S-标记探针进行的原位杂交。左边照片来自暗视野显微镜检查而相应的右边的照片来自明视野显微镜检查。暗视野照片中所示的明亮颗粒是 RNA 杂交信号。4A) 来自 67 岁男性的前列腺组织显示了阳性的腺泡上皮细胞和阴性基质细胞, (5 倍放大)。4B) 4A 的明视野。4C) 高度放大 (40x), 显示在右下角的阳性区。4D) 4C 的明视野。4E) 肾组织显示没有 RNA 杂交, (5 倍放大)。4F) 4E 的明视野。

附图 5. LNCaP mRNA 的引物延伸。在 TCR $\gamma$ 恒定区（从 C $\gamma$ 1 的 5' 末端开始的 75 个核苷酸）逆转引退火。逆转录本终止于大约 128 核苷酸处，用箭头表示，其表明转录本在 C $\gamma$ 1 的上游大约 53 核苷酸处启始。含有 LNCaP 的 TCR $\gamma$ 逆转录本的泳道被曝光 72 小时而标记泳道被曝光 8 小时。

附图 6. 前列腺 TCR $\gamma$ 转录本。其图示了前列腺是如何被转录和剪接的。转录本包括 J $\gamma$ 1.2 片段, C $\gamma$ 1 的 3 个外显子, 接着是非翻译序列。

附图 7. 前列腺 TCR $\gamma$ 的体外转录偶联的翻译分析。两个大约为 8 和 13 kDa 的蛋白被获得(泳道 1)。使用空载体的阴性对照反应（泳道）没有产生任何的蛋白产物。

附图 8. 用于 PS-TCR $\gamma$ 转录分析的引物(SEQ ID NOs: 1-12)。

附图 9. 前列腺特异性 TCR $\gamma$ 转录本体外翻译的分析。前列腺特异性 TCR $\gamma$ 转录本在体外编码两种蛋白。<sup>35</sup>S-Met 体外标记的翻译蛋白被置于 16.5% Tris-Tricine 凝胶上进行跑胶并通过放射自显影法进行分析。所用突变构建体的图示分析被显示在右边。空白框代表第一阅读框架且可能的启动密码子用粗体表示, 第二阅读框架用阴影框表示且可能的启动密码子用斜体表示, “X”表示一个 ATG 密码子突变为 ATA。以 kDa 表示的大小标志被显示于顶部。

附图 10. TARP 为前列腺中表达的核蛋白。获自 LNCaP 细胞、PC3 细胞或前列腺肿瘤样本(癌)的蛋白提取物的 Western 印迹。20 $\mu$ g 的每种蛋白提取物被置于 16.5% Tris-Tricine 凝胶上电泳且用抗 TARP (上图)或 TCR $\gamma$ (下图)的抗体进行探测。作为阳性对照, 1 $\mu$ g 的 His-标记的 TARP (上图)或 100 ng 的 His-标记的 TCR $\gamma$  (下图)被置于凝胶上电泳(重组)。以 kDa 表示的大小标志被显示于左边。(B) LNCaP 细胞的细胞质部分(细胞质), 膜部分(膜)以及核部分(核)的 Western 印迹。每一部分 40g 被置于 16.5% Tris-Tricine 凝胶上电泳且用抗 TARP 的抗体进行探测。作为阳性对照, 1 $\mu$ g His-标记的 TARP 被置于凝胶上跑胶(重组)。以 kDa 表示的分子大小被显示于左边。

附图 11. TARP mRNA 在乳房癌细胞中的表达。(A)用 TARP(上图)或肌动蛋白(下图)特异的引物对获自下列细胞系的 RNA 进行

RT-PCR: 前列腺(LNCaP 和 PC3), 成神经细胞瘤 (A172), 结肠 (COLO 205), 胃 (KATO III) 以及乳房 (MCF7, BT-474, Hs57Bst, SK-BR-3, CRL-1897 和 MDA-468)。没有模板的 RT-PCR 反应被显示为 dH<sub>2</sub>O。  
(B) 用 TARP (上图) 或肌动蛋白 (下图) 特异的引物对获自 12 个人  
5 乳房癌组织样本的 cDNA (1-12 泳道) 进行 PCR。没有模板的 PCR 反应被显示为 dH<sub>2</sub>O。对于两个图, 20% 的 PCR 产物被置于 1% 琼脂糖凝胶上进行电泳且用溴化乙锭染色进行显影。

附图 12. 在乳房细胞系中发现的 TARP 转录本与前列腺特异形式的相同。(A) TCR $\gamma$ 座位的图示以及 TARP 如何在前列腺细胞中被转录和剪接。用于在图 B 中 RT-PCR 分析的引物被显示了出来。(B) TARP  
10 mRNA 表达的 RT-PCR 分析。利用获自前列腺细胞系(LNCaP 和 PC3) 和乳房细胞系(MCF7, BT-474, SK-BR-3 和 Hs578Bst)的 cDNA, 并利用引物 1 和 3 (上图), TARP 引物 2 和 3 (中图) 或肌动蛋白引物 (下图) 进行 PCR 反应。进行没有模板的 RT-PCR, 并显示为 dH<sub>2</sub>O。20% 的  
15 PCR 产物被置于 1% 的琼脂糖凝胶上进行电泳并用溴化乙锭染色观察。(C) TARP 转录本的 Northern 印迹分析。获自前列腺细胞系(LNCaP 和 PC3) 和乳房细胞系(MCF7, BT-474, SK-BR-3 和 Hs578Bst)的 2 $\mu$ g poly(A)mRNA 利用恒定区片段作为探针进行分析。24 小时曝光产生放射自显影 (上图)。用人 $\beta$ -肌动蛋白 RNA 探针在同一滤膜上产生  
20 条带并进行分析以校证明相同的上样量。曝光 1 小时产生放射自显影(下图)。表示核苷酸大小的 RNA 大小标志被显示左边。

附图 13. TARP 存在于乳房癌细胞的核中。(A) 对获自 LNCaP , MCF7, BT-474, SK-BR-3 和 Hs578Bst 细胞的核提取物进行 Western 印迹分析。40 $\mu$ g 的每种核提取物被置于 16.5% 的 Tris-Tricine 凝胶上电泳并用抗 TARP (上图) 或 TCR $\gamma$ (下图)的抗体进行探测。作为阳性对照, 1 $\mu$ g 的 His-标记的 TARP (His-TARP) 和 100 ng 的 His-标记的 TCR $\gamma$  (His-TCR $\gamma$ ) 在凝胶上跑电泳。大小标志以 kDa 为单位显示在左侧。  
25

附图 14. TARP 的可能的功能区。(A) TARP 含有一个可能的亮氨酸拉链基元和磷酸化位点。可能的亮氨酸拉链基元用亮氨酸加框表示, 其后的碱性区加下划线。cAMP-和 cGMP-依赖的蛋白激酶磷酸  
30

化位点(氨基酸 46-49 和 55-58) 以及蛋白激酶 C 磷酸化位点(氨基酸 19-21 和 20-22) 用空心轮廓表示。(B) TARP 与 Tup1 的蛋白序列比较。 TARP (42-57), *Dictyostelium discoideum* Tup1 (dTup1, 521-536) 以及 *Saccharomyces cerevisiae* Tup1 (yTup1, 626-660) 的氨基酸序列被显示。保守的残基被加框。

本发明的详述

I 介绍

令人惊讶地, 人们已发现上皮来源的前列腺细胞以及很多乳房癌细胞表达 T 细胞受体 $\gamma$ 链("TCR $\gamma$ ")的 mRNA。前列腺中的主要 TCR $\gamma$ 转录本与 T 淋巴细胞中表达的大小不同。前列腺上皮细胞和很多乳房癌高水平表达被认为仅在 T 淋巴细胞中表达的基因的转录本的发现是出乎意料的。

因为 TCR $\gamma$ 阅读框架含有很好的 Kozak 序列 (Kozak, M. Cell 44 283-92 (1986)), 我们最初设想截短的 TCR $\gamma$  蛋白被编码。因此, 另一个意外的发现是在上皮前列腺细胞和乳房癌细胞中表达的 TCR $\gamma$ 座位编码一个 7Kda 的核蛋白。因为该蛋白是由不同于 TCR $\gamma$ 的阅读框架所编码的, 我们将其命名为 TCR $\gamma$ 可变阅读框架蛋白, 缩写为"TARP"。除了在 TCR $\gamma$ 座位的一个内含子中转录本来源的可变的阅读框架被翻译外, TARP 具有另外两种独特的特征。首先, 意外地在细胞中发现这样的小肽, 因为大多数是分泌型。另外, TARP 缺少很好的 Kozak 序列。

体外和体内应用

在前列腺上皮细胞、前列腺癌细胞以及很多乳房癌变的细胞中该种蛋白的存在为体内和体外的应用创造了大量的机会。首先抗该种蛋白的抗体能被用于体外测试来检测样品中表达 TARP 细胞的存在。例如下面描述的实施例中, 在正常的乳房细胞中不存在 TARP 和 TARP 转录本, 或者低水平存在, 但是, 在表达 TARP 的乳房癌细胞很容易被

检测到。在从个体取出的乳房细胞中检测到高水平的 TARP 转录本或 TARP, 将表明个体存在表达 TARP 的乳房癌。鉴于前列腺细胞中 TARP 的表达, 本领域的技术人员意识到在进行性前列腺癌中去除前列腺是一种常见的手术干预。从一个已去除前列腺的个体的细胞中检测到 TARP 则表明存在前列腺癌。(本领域的技术人员将意识到每年大约有 1,000 男人被诊断出患有乳房癌。因此, 病人单一的患有乳房癌不是不可能的。假定事实上男性的前列腺癌患者超过正常人的 200 倍, 且所关注的患者已被发现患有前列腺癌, 则病人单独患乳房癌的可能性很小。诊断可通过取样位点的一些信息被证实, 其中样品被进行组织学和细胞形态学以及其它常规诊断标准的研究。在任一情况中, 由于任一病征的存在需要进一步的鉴定和监测, 检测到 TARP 表达细胞存在于前列腺已被切除的男性中是非常有用的, 不管该个体是否已患上前列腺癌, 乳房癌, 或二者。在没有前列腺癌的男性中检测到 TARP 表达细胞是乳房癌的一个指标。)

15

TARP 本身, TARP 的免疫原性片段, 以及编码 TARP 或其免疫原性片段的核酸也可被用于在体外激活来自个体的细胞毒性 T 淋巴细胞 ("CTL"), 在灌输给所述个体后用以攻击前列腺癌和表达 TARP 的乳房癌。

20

可将 TARP 本身、TARP 的免疫原性片段、编码 TARP 或其免疫原性片段的核酸分子通常与可药用载体一起对个体给药, 以刺激产生或加强对前列腺癌或表达 TARP 的乳房癌的免疫应答。对已被确诊患有前列腺癌或表达 TARP 的乳房癌的个体可治疗性地给药这一类的组合物。

25

如下所讨论的, TARP 是一种具有调控转录的蛋白结构特征的核蛋白。人们预料细胞中 TARP 水平的调节将影响细胞的生长, 细胞的复制, 或二者。因此, TARP 水平的调节在控制癌细胞的入侵方面可能是很重要的。细胞中的 TARP 的水平可以通过不同的减少 RNA 转

30

录本被翻译为蛋白例如核酶和反义分子的形式来降低。细胞中 TARP 的水平也可被增加。例如，含有编码 TARP 的核酸的表达载体，由强组成型启动子驱动，其可被导入到细胞中。核酸的组成型转录导致蛋白在细胞中高水平的表达。

5

此部分的余下内容描述了 TARP 的不同结构特征。接下来描述的是本文所用术语的定义，然后是有关 TARP 的免疫原性片段的选择、TARP 对个体的给药、抗 TARP 抗体的形成、TARP 转录本和蛋白的检测以及药物组合物的讨论。

10

### TARP 的结构特征

TARP 在七体重复中含有五个亮氨酸，此表明 TARP 含有亮氨酸拉链二聚基元（附图 7A）。所以 TARP 必定含有一个两亲螺旋。一种表明 TARP 可含有一个两性螺旋的迹象是丝氨酸和脯氨酸残基（这些残基被认为是螺旋起始位置的残基）被发现紧挨在第一个亮氨酸重复的前面。第二，很多的带电氨基酸被发现在七体重复中，因此螺旋有两性特性且可能充当与其它螺旋的盐桥。尽管亮氨酸在七体重复中的存在是亮氨酸拉链基元的指示，有些被鉴定在七体重复中含有 5 个亮氨酸的蛋白并没被认为是亮氨酸拉链蛋白。例如，亲核素（karyopherin）的晶体结构(Chook, Y. M. 等, Nature 399: 230-237 (1999)), *B. stearothermophilus* 嘧啶核苷酸磷酸化酶(Pugmire, M. J. 等, Structure 6: 1467-1479 (1998))以及 *T. thermophilus* 苯丙氨酰 tRNA 合成酶(Mosyak, L. 等, Nat. Struct. Biol. 2: 537-547 (1995))表明这些蛋白在含有五个亮氨酸的七体重复序列中不含有 $\alpha$ -螺旋结构。需要用相互作用和结构研究来测定在 TARP 中发现的亮氨酸重复的意义。

25

TARP 氨基酸序列另一个独特的特征是碱性氨基酸的区域紧接着可能的亮氨酸拉链基元（附图 7A），此表明其可能的 DNA 结合基元。然而，其中碱性区域的方向是非常独特的，其紧接着亮氨酸重复而不是在其前面。大多数结合 DNA 的亮氨酸拉链蛋白在亮氨酸重复前面

30

具有碱性区域(参见(Chook, Y. M. 等, Nature 399: 230-237 (1999)))。TARP 的碱性区域可能仅仅充当核定位信号, 但 TARP 是一种核蛋白的事实加强了 TARP 可结合 DNA 的假说。

5           为了检测 TARP 是否与任意已知的蛋白具有同源性, 我们在 GenBank 中进行了蛋白 BLAST 检索。此检索显示 TARP 的氨基酸序列与 *Dictyostelium discoideum* Tup1(GenBank 登录号 AAC29438)以及 *Saccharomyces cerevisiae* Tup1 (Williams, F.E. 等, Mol. Cell. Biol. 10: 6500-6511 (1990)) (附图 7C)有一些同源性。酵母 Tup1 通常在  
10           Cyc8(Ssn6)复合体发现且是被葡萄糖、氧和 DNA 损害所调控的基因的转录抑制所需要的 (Tzamarias, D. 等, Genes Dev. 9: 821-831 (1995))。Cyc8 (Ssn6) 和 Tup1 都不结合 DNA, 但是它们通过特异性 DNA 结合蛋白例如 $\alpha 2$ 、Mig1、Rox1 和 al 作为共抑制剂复合物的一个部分而发挥作用(Tzamarias, D. 等, Genes Dev. 9: 821-831(1995))。Tup1  
15           的 C'-末端部分含有富含天门冬氨酸和色氨酸的 43 个氨基酸序列的 6 个重复, 已知为 WD-40 或 $\beta$ -转导素重复子 (Williams, F.E. 等, Mol. Cell. Biol. 10: 6500-6511 (1990); Fong, H. K. 等, Proc. Natl. Acad. Sci. USA 83: 2162-2166 (1986))。WD-40 重复子已在很多蛋白中被鉴定出, 在蛋白-蛋白的相互作用中起作用。值得注意的是, Tup1 已被显示通过 2  
20           个 WD-40 重复子与 $\alpha 2$  相互作用(Komachi, K. 等, Genes Dev. 8: 2857-2867(1994))。令人感兴趣的是 TARP 与 Tup1 的第 5 个 WD-40 重复子具有同源性(附图 7C)。因为 TARP 是一种核蛋白, 其与 Tup1 同源则表明 TARP 是一种涉及转录调控的功能性核蛋白复合物的成员。

25           TARP 抗体识别前列腺和乳房核提取物的双重物质(附图 6A)。最快的 7 kDa 条带与 His-TARP 重组蛋白共移动, 而最弱的条带为较大分子量。对于 9kDa 条带的一个可能解释则是存在翻译后修饰。为了确定 TARP 是否含有任意已知的翻译后修饰的位点, 我们利用 PROSITE 程序分析 TARP 氨基酸序列, 其中的 PROSITE 程序来自  
30           Swiss Institute of Bioinformatics ExpASy proteomics server (<http://www.>

expasy. ch) (Appel, R. D. 等, Trends Biochem. Sci. 19: 248-260 (1994); Hofmann, K. 等, Nucleic Acids Res. 27: 215-219 (1999))。如附图 7A 所示, 许多可能的磷酸化作用位点被发现, 包括 cAMP 和 cGMP-依赖性蛋白激酶磷酸化位点 (RRAT 和 RRGT) 以及蛋白激酶 C 磷酸化位点(SSR 和 SRR)。已在很多例子中表明在 SDS PAGE 凝胶上磷酸化作用引起蛋白显示出较大分子量。如果也是这种情况,附图 6 的结果表明未修饰的形式在 LNCaP 细胞中是普遍的, 而且仅仅磷酸化形式存在于 MCF7 和 SK-BR-3 细胞中。因此, TARP 可能在前列腺和乳房癌细胞中表达后进行了翻译后-修饰。

10

我们最初的对 TARP 转录本的研究没有揭示 TARP 在乳房中表达 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。一种可能的解释是 TARP 在正常乳房中表达的水平很低, 因此是很难被检测的。如在结果部分中所述的, 与癌样品中检测到的强烈信号相比, 在正常乳房样品的 PCR 分析中检测到非常弱的信号。因此, TARP 在乳房癌细胞中的存在可表明在乳房细胞致癌转化后 TARP 被诱导表达。此外, TARP 在乳房癌细胞中的存在表明 TARP 被雌激素所调控。在 TARP 的内含启动子中鉴定出的使雄激素应答元件 (ARE) 与雌激素应答元件(ERE)结合的元件加强了这一假说。此杂交元件含有两个半-位点, 其在 5'末端对 ARE 特异, 在 3'末端对 ERE 特异[ (Zilliacus, J. 等, Mol. Endocrinol. 9: 389-400 (1995)) 以及未公开的数据]]。另外需要进行一些实验来检测是否雌激素调控 TARP。然而有一些例子, 其中的突变体 ARE 引起某些前列腺特异性基因在乳房肿瘤中表达。例如, 前列腺特异性抗原(PSA) 已被显示在乳房肿瘤中表达 (Majumdar, S. 等, Br. J. Cancer 79: 1594-1602 (1999))。PSA 异常表达的分子分析使得发现了在 PSA 启动子中发现的 ARE 中的单一位点突变。据信, 该突变导致雄激素调控的 PSA 在乳房瘤中表达的丧失 (Majumdar, S. 等, Br. J. Cancer 79: 1594-1602 (1999))。此时是否在 3 个被检测的乳房细胞系中在 TARP 启动子中发生了类似的突变仍不得而知。

30

前列腺依据雄激素来维持其结构和功能。当前列腺细胞变为恶性细胞时，其失去了雄激素依赖性。在此研究中，我们使用两种对用于生长的雄激素具有不同依赖性的前列腺细胞系：LNCaP 和 PC3 细胞。雄激素受体存在于雄激素依赖的 LNCaP 细胞系中，但不存在于雄激素非依赖性的 PC3 细胞系中 (Tilley, W. D. 等, Cancer Res. 50: 5382-5386(1990))。如附图 3 所示, TARP 在 LNCaP 细胞而不在 PC3 细胞中表达。此结果表明 TARP 表达可通过雄激素刺激来调控。TARP 启动子中的 ARE 样元件的鉴定加强了 TARP 被雄激素诱导的观点。在 LNCaP 细胞而不在 PC3 细胞中表达表明 TARP 在调控雄激素依赖的应答中是重要的

## II. 名词定义

除非另有定义，此处所用的技术和科学术语具有本领域技术人员通常理解的含义。下面引用的文献提供了本发明所用的很多术语的基本定义: Singleton 等, 微生物学和分子生物学词典 (DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY) (第二版, 1994); 剑桥科技词典 (THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY) (Walker ed., 1988); 遗传学专业词典 (THE GLOSSARY OF GENETICS), 第 5 版., R. Rieger 等(编辑), Springer Verlag (1991); 和 Hale & Marham, HARPER COLLINS 生物学词典 (THE HARPER COLLINS DICTIONARY OF BIOLOGY) (1991)。如此处所述的，下面的术语除非另有说明，其具有上述词典中的含义。

"T 细胞受体"是指在含有一个 $\alpha$ 链和一个 $\beta$ 链或一个 $\gamma$ 链和一个 $\delta$ 链的 T 细胞表面发现的杂二聚体。T 细胞受体识别与 MHC 分子结合的抗原。

"T-细胞受体 $\gamma$ 可变阅读框架蛋白"和"TARP"是指一种多肽，其序列如例如附图 14 所描述的。此多肽由 T 细胞受体 $\gamma$ 基因翻译而来，所述基因转录本存在于上皮来源的前列腺细胞中，前列腺癌细胞中，以

及在很多乳房癌中。由于"TARP"中的最后一个缩写字母代表单词"蛋白", 所以"TARP 蛋白"是多余的。

5 此处所用的从其可翻译成 TARP 的核酸转录本是指"TARP 核酸"或"编码 TARP 的核酸", 从中转录 TARP 的基因在此处是指"TARP 基因。"

10 如此处所用的, TARP 的"免疫原性片段"是指 TARP 的一部分, 当其在主要组织相容性复合物分子体系中被细胞提呈后, 可以在 T 细胞活化测定中, 激活 T 淋巴细胞抵抗表达 TARP 的细胞。尽管较长的片段也可被使用, 典型地, 此种片段为 TARP 的 8 到 12 个连续氨基酸。

15 在上下文中将一个与另一个多肽比较时, "序列同一性"通过比较作为参考序列的 TARP 序列与一个受试序列而被测定。典型地, 两个序列进行最大和最佳比对。

"配体"是指一种特异性与靶分子结合的化合物。

20 "受体"是特异性与配体结合的化合物。

25 "细胞毒素 T 淋巴细胞" ("CTL")在对肿瘤细胞的免疫应答中是重要的。CTL 识别被在几乎所有有核细胞的细胞表面表达的 HLA I 类分子提呈的肽的表位。

30 已知肿瘤特异性辅助 T 淋巴细胞 ("HTL")对维持有效的抗肿瘤免疫时是很重要的。其在抗肿瘤免疫中的作用也已在动物模型中被证明, 其中这些细胞不仅为 CTL 的诱导和抗体应答提供帮助, 而且提供效应器功能, 该功能通过直接细胞接触和通过淋巴因子(例如 IFN $\gamma$  和 IFN- $\alpha$ )的分泌来介导。

"抗体"是指特异性识别和结合表位(例如,抗原)的至少含有轻链或重链免疫球蛋白可变区的多肽配体。包括完整的免疫球蛋白和变体和本领域已知的它们的部分例如 Fab'片段, F(ab)'<sub>2</sub> 片段, 单链 Fv 蛋白("scFv"), 以及二硫化物稳定的 Fv 蛋白("dsFv")。scFv 蛋白是一种融合蛋白, 其免疫球蛋白的轻链可变区和免疫球蛋白的重链可变区通过接头结合。天然的免疫球蛋白由免疫球蛋白基因编码。包括  $\kappa$  和  $\lambda$  轻链恒定区基因,  $\alpha$ 、 $\gamma$ 、 $\delta$ 、 $\epsilon$  和  $\mu$  重链恒定区基因, 以及无数的免疫球蛋白可变区基因。术语 "抗体"包括多克隆抗体, 单克隆抗体, 嵌合抗体和人源化抗体, 通过免疫产生的, 从杂交瘤产生的, 或重组的。

"表位"或"抗原决定基"是指 B 和/或 T 细胞对其产生应答的抗原上的位点。表位可由抗原上的连续氨基酸形成或由非相邻氨基酸通过蛋白质的三级折叠形成。由连续氨基酸形成的表位典型地在变性溶剂中可保留而由非邻接氨基酸形成的表位典型地在用变性溶剂的处理中会丢失。一个表位典型地在独特空间构象中包括至少 3 个, 和通常更多, 至少 5 个或 8-10 个氨基酸。测定表位的空间构象的方法包括, 例如, x-光衍射以及二维核磁共振。参见, 例如., 分子生物学方法, Vol. 66, Glenn E. Morris, Ed (1996) 中的表位绘图手册 (Epitope Mapping Protocols in Methods in Molecular Biology)。

当配体或受体进行结合反应时, 配体或受体 "特异性地结合于" 某一化合物, 即 "待分析物", 该结合是待分析物在异质化合物样品中存在的决定因素。因此, 配体或受体优选与特定的待分析物结合且不结合样品中的有效量的其它物质。例如, 多核苷酸特异性地结合于含有一个互补序列的待分析多核苷酸而抗体在免疫测定条件下特异地与含有抗体所针对的表位的抗原待分析物结合。

"免疫测定"是指检测样品中待分析物的方法, 其中待分析物的特异性是抗体和配体之间的特异性结合所赋予的。包括通过抗体和配体

之间的特异性结合来检测抗体待分析物。参见 Harlow 和 Lane (1988) 抗体, 实验室手册, 冷泉港出版社, New York, 描述了可被用来检测特异性免疫活性的免疫测定形式和条件。

- 5 "疫苗"是指含有有效赋予生物体仅引起较低发病率和致死率的生物免疫治疗水平的药剂或组合物。当然, 制备疫苗的方法在免疫系统的研究并预防和治疗动物和人类疾病方面是有用的。

“免疫有效量”是有效地在个体中引起免疫应答的量

10

"多肽"是指由氨基酸残基、相关的天然存在的其结构变体和合成的通过肽键连接的非天然存在的类似物、相关的天然存在的结构变体, 以及合成的非天然产生的类似物组成的聚合物。合成多肽可被例如利用自动多肽合成仪来合成。术语“蛋白质”典型地是指大的多肽。

15

术语“肽”典型地是指短的多肽。

在此使用常规的符号来描述多肽序列: 多肽序列的左手端为氨基末端, 多肽序列的右手端为羧基末端

20

"融合蛋白"是指由两个或多个多肽通过肽键连接形成的多肽, 其中的肽键是由一个多肽的氨基末端和另一个多肽的羧基末端形成的。融合蛋白可典型地被表达为单一连续的融合蛋白的核酸序列所编码。然而, 融合蛋白也可由组成型多肽的化学偶联来形成。

25

"保守替代"是指多肽中的氨基酸用功能类似的氨基酸来替代。下列 6 组中均包括作为另一个氨基酸保守替代的氨基酸:

1) 丙氨酸 (A), 丝氨酸 (S), 苏氨酸 (T);

2) 天门冬氨酸 (D), 谷氨酸 (E);

3) 天冬酰胺(N), 谷氨酰胺 (Q);

30

4) 精氨酸(R), 赖氨酸 (K);

- 5) 异亮氨酸(I), 亮氨酸 (L), 甲硫氨酸 (M), 缬氨酸 (V); 和
- 6) 苯丙氨酸 (F), 酪氨酸 (Y), 色氨酸 (W)。

5 两个蛋白如存在于不同的种中且具有共同的遗传祖先和 70%的氨基酸序列同一性则彼此是"同源物"。

10 "大体上纯的"或"分离的"是指目的物质是主要存在的物质(例如, 以摩尔计算, 在组合物中比其它任意的大分子物质含量更多), 而基本上纯化的部分是一种其中含有在所有存在的大分子物质中至少 50% (以摩尔计算) 为目的物质的组分。通常, 大体上纯的组分是指组分中的大约 80%到 90%或以上的大分子物质是纯化的目的物质。如果组分主要含有一种单一的大分子物质, 则目的物质被纯化到基本上同质(通过常规的检测方法在组分中检测不到污染物)。为了定义, 溶剂、小分子 (< 500 道尔顿)、稳定剂 (例如, BSA), 以及元素离子在此不被认为是大分子。

15

20 "核酸"是指由通过磷酸二酯键连接的核苷酸 (核糖核苷酸, 脱氧核糖核苷酸, 其相关的天然存在的结构变体, 及合成的非天然存在的其类似物)单元、相关的天然存在的其结构变体及合成的非天然存在的其类似物组成的聚合物。因此, 此术语包括核苷酸聚合物, 其中的核苷酸及其之间的键, 包括非天然存在的合成类似物, 例如 (但不限制于), 硫代磷酸酯, 氨基磷酸酯, 甲基磷酸盐, 手性-甲基磷酸盐, 2-O-甲基核糖核苷酸, 肽-核酸(PNA), 等等。此种多核苷酸可以利用例如自动 DNA 合成仪来合成。术语“寡核苷酸”典型地是指短的多核苷酸, 通常不长于 50 个核苷酸。当通过 DNA 来表示核苷酸序列时(例如, A, T, G, C), 可以理解的是其也包括 RNA 序列(例如, A, U, G, C), 其中的"U"替代"T"。

25

30 此处所用的描述核苷酸序列的常规表示是: 单链核苷酸的左手边为 5'末端; 双链核苷酸序列的左手方向作为 5'方向。产生 RNA 转录

本的核苷酸 5'到 3'的增加的方向作为转录的方向。与 mRNA 具有相同序列的 DNA 链被称为“编码链”；与转录自 DNA 的 mRNA 具有相同序列的且位于 RNA 转录本的 5'末端的 5'的 DNA 链的序列为"上游序列"；与转录自 DNA 的 mRNA 具有相同序列的且位于编码 RNA 转录本的 3'末端的 3'的 DNA 链的序列为"下游序列"；

"cDNA"是指互补于或等同于 mRNA 的 DNA，其为单链或双链的形式。

10 "编码"是例如基因、cDNA 或 mRNA 充当模板来合成其它聚合物的多核苷酸和在生物学的过程中具有所定义的核苷酸序列(例如, rRNA, tRNA 和 mRNA) 或所定义的氨基酸序列大分子中的特定核苷酸序列固有的特性以及由其产生的生物学特性。因此，如果由基因产生的 mRNA 的转录和翻译在细胞或其它的生物系统中产生蛋白质则此基因

15 编码了蛋白。基因或 cDNA 的编码链（等同于 mRNA 序列且通常被显示在序列表中的核苷酸序列），以及被用作转录模板的非编码链可被称为编码蛋白或基因或 cDNA 的其它产物。除非另有说明，“编码氨基酸序列的核苷酸序列”包括所有的编码相同氨基酸序列的核苷酸序列简并形式。编码蛋白的核苷酸序列和 RNA 可包括内含子。

20 "重组核酸"是指具有非天然连接在一起的核苷酸序列的核酸。其包括含有扩增的或装配的可用于转化合适宿主细胞的核酸的核酸载体。含有重组核酸的宿主细胞称为“重组宿主细胞”。然后基因在重组宿主中表达来产生，例如，“重组多肽”。重组核酸也可行使非

25 编码的功能（例如，启动子，复制区，核糖体结合位点等）。

"表达调控序列"是指多核苷酸中的调控与其可操作地连接的核苷酸序列的表达（转录和/或翻译）的核苷酸序列。“可操作地连接”是指两个部分的功能关系，其中一个部分的活性（例如，调控转录的能力）是另一个部分作用的结果（例如，序列的转录）。表达调控序列

30

可包括,例如,用于举例而并非限制,启动子的序列(例如,诱导型的或组成型的),增强子,转录终止子,起始密码子(例如,ATG),内含子的剪接信号,以及终止密码子。

5 "表达盒"是指含有与可表达的核苷酸序列可操作相连的表达调控序列的重组核酸构建体。表达盒通常含有足够用于表达的顺式作用元件,其它用于表达的元件可由宿主细胞或体外表达系统提供。

10 "表达载体"是指含有一个表达盒的载体。表达载体包括所有本领域已知的载体,例如,粘粒,质粒(例如,裸露的或包含在脂质体中的)以及包含表达盒的病毒。

如果多核苷酸的序列是与其序列是第二序列的多核苷酸特异地杂交的第一序列,则相对于第二序列,第一序列是“反义序列”。

15

用于描述两个或多个核苷酸序列或氨基酸序列之间的关系术语包括"参考序列", "选自", "对比窗", "同一性", "序列同一性的百分比", "基本上相同", "互补"和"基本互补"。

20

为了核酸序列的序列比较,典型地,一个序列充当参考序列,来与待测序列进行比较。当使用序列比较算法时,测试序列和参考序列被输入到计算机中,指定亚序列坐标系,如果需要,还要设定序列算法程序的参数。使用默认的程序参数。序列比较的比对方法是本领域公知的。可进行序列比较的最佳比对通过如下方法进行:例如 Smith 和 Waterman, *Adv. Appl. Math.* 2: 482 (1981)的局部同源性算法, Needleman 和 Wunsch, *J. Mol. Biol.* 48: 443 (1970)的同源性比对算法, Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85: 2444 (1988)的相似性检索法,这些算法的计算机处理(GAP, BESTFIT, FASTA, 和 TFASTA, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI),或通过人工比对和肉眼检查(参

30

见,例如现代分子生物学手册 (Current Proctocols in Molecular Biology) (Ausubel 等编辑, 1995 增补本))。

5 一个有用算法的例子是 PILEUP。PILEUP 利用 Feng & Doolittle, J. Mol. Evol. 35: 351-360 (1987)的改进比对的简化方法。该方法类似于 Higgins & Sharp, CABIOS 5: 151-153 (1989)所描述的方法。利用 PILEUP, 比较参考序列与其它测试序列来测定序列同一性的百分比, 使用下面的参数: 默认缺口权值(3.00), 默认缺口长度权值(0.10), 以及加权的末端缺口。PILEUP 可从 GCG 序列分析软件包中获得, 例如, 7.0 版本 (Devereaux 等, Nuc. Acids Res. 12: 387-395 (1984))。

15 另一个适于测定百分序列同一性和相似性的算法的例子是 BLAST 和 BLAST 2.0 算法, 其被描述在 Altschul 等, J. Mol. Biol. 215: 403-410 (1990) 和 Altschul 等, Nucleic Acids Res. 25: 3389-3402 (1977)) 中。进行 BLAST 分析的软件可通过国家生物工程信息中心(<http://www.ncbi.nlm.nih.gov/>)获得。BLASTN 程序 (用于核苷酸程序)使用默认的字长(W)为 11, 校验值 (B) 为 50, 期望值 (E) 为 10, M=5, N=-4, 和两个链的比较。BLASTP 程序 (用于氨基酸序列) 使用默认的字长 (W) 长为 3, 期望值 (E) 为 10, 以及 BLOSUM62 得分矩阵(参见 Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89: 10915 (1989))。

25 "严格杂交条件"是指 50% 甲酰胺, 5 x SSC 和 1% SDS 在 42°C 下温育或 5 x SSC 和 1% SDS 65°C 下温育, 并用 0.2 x SSC 和 0.1% SDS 在 65°C 下洗涤。

30 "自然产生的"这一术语, 对研究对象而言是指研究对象可在自然中找到。例如, 可从天然源中分离得到并没有经过人在实验室中进行的有意的改造的存在于生物体 (包括病毒) 中的氨基酸或核苷酸序列。就可称其为 "天然产生的"。

“接头”是指通过共价键或离子键、范德华力或氢键将两个分子连接在一起的分子，例如，与一互补序列在 5'末端杂交并与另一互补序列在 3'末端杂交的核酸分子由此将两个非互补的序列连接起来。

5 “药物组合物”是指适于在哺乳动物中使用的组合物。药物组合物含有药学有效量的活性剂及可药用载体。

“药学有效量的”是能产生所希望的药学结果的有效药剂的量。

10 “可药用载体”是指任何标准的药物载体、缓冲剂和赋形剂如磷酸盐缓冲液、5%葡萄糖水溶液、乳液如油/水乳液或水/油乳液和各种类型的湿润剂和/或佐剂。合适的药物载体和制剂被描述于 Remington's Pharmaceutical Sciences, 第 19 版, Mack Publishing Co., Easton, 1995 中。优选的药物载体取决于所希望的活性药物的给药方式。给药的典型方式包括肠内（例如，口服）或肠胃外（例如，皮下、肌肉内、静脉内或腹膜内注射）或局部、经皮或经粘膜给药。“可药用盐”是可配制成可药用化合物的盐，例如，金属盐（钠、钾、镁、钙盐等），以及铵或有机胺的盐。

20 用于诊断或治疗的“个体”是人或非人的哺乳动物。

组合物的“给药”是指通过一个选择的途径将组合物导入到个体中，例如，如果所选的途径是静脉内施用，则组合物被通过导入组合物到个体的血管中来给药。

25

“治疗”是指预防性治疗或治疗性治疗。

“预防性”治疗是给药于没有显示疾病症状或显示早期症状的个体来减少病症发展的危险。

30

"治疗性"治疗是一种给药于表现出病症的个体进行的治疗，目的是减小或消除这些症状。

5 "诊断"是指鉴定病症状况的存在或特性。诊断的方法根据其敏感性和特异性而不同。诊断方法的“敏感性”是指进行试验呈阳性的患病个体的百分数（真阳性的百分数）。诊断方法的“特异性”是指 1 减去假阳性率，其中的假阳性率是指那些试验呈阳性但是没有患病的比例。而一种特定的诊断方法有可能不能提供出一种病况的确定性诊断，如果该方法提供了有助于诊断的阳性显示，其就足够了。

10

"预诊"是指对一种病症可能性进展（例如，严重性）的预测。

### III. TARP

15 本发明提供分离的重组的 TARP。因为我们首先发现了分离的前列腺特异性 TCR $\gamma$ 转录本，我们最初使用术语“PS-TCR $\gamma$ 蛋白”和“PS-TCR $\gamma$ 多肽”来代表能够从约 1.1kb PS-TCR $\gamma$ 转录本开始的任何阅读框架中翻译的任何多肽。具体地讲，该术语是指两种在体外翻译系统中翻译的蛋白 PS-TCR $\gamma$ -1 (SEQ ID NO: 14) 和 PS-TCR $\gamma$ -2 (SEQ ID NO: 15)。我们现在已知道只有这些阅读框架的第一个在前列腺细胞中  
20 被翻译。因为这个阅读框架不是产生 TCR $\gamma$ 链的阅读框架，该蛋白目前被称为“T 细胞受体可变阅读框架蛋白”。全长 TARP 是一个 58 氨基酸的蛋白，其序列被描述于 SEQ ID NO: 14 和附图 14 中。

25 在某些的实施方案中，本发明提供了含有包含从 TARP 的至少 5 到至少 15 个保守氨基酸的表位的多肽。这样的蛋白与抗全长 TARP 的抗体结合（在此部分，除非上下文的其它需要，“TARP”是指全长蛋白）。其它的实施方案中，本发明提供了含有第一和第二多肽部分的融合蛋白，其中一个蛋白部分含有被鉴定为 TARP 表位的至少 5 个氨基酸的氨基酸序列。在一个实施方案中，TARP 部分是指 TARP  
30 的全部或基本上全部的 TARP。另外一部分可以是一个例如免疫原性

蛋白。这样的融合子也可以被用于激发抗 TARP 的免疫反应。在本发明其它的实施方案中提供了 TARP 样肽（“TARP 类似物”），其氨基酸序列至少 90% 相同于 TARP（虽然它们可能具有相对于 TARP 91%， 92%， 93%， 94%， 95%， 或更高的序列同一性）且其被与 TARP 特异性结合的抗体特异性地结合。在其它实施方案中，本发明提供了 TARP 样肽（有时此处也称为“TARP 类似物”），其氨基酸序列至少 90% 相同于 TARP（虽然它们可能具有相对于 TARP 91%， 92%， 93%， 94%， 95%， 或更高的序列同一性）且激活 T 淋巴细胞抵抗表达 TARP 的细胞。这样的蛋白也可用作免疫原以阻断对 PS-TCR $\gamma$  蛋白的耐受。

10

在另一个实施方案中，所述多肽含有结合于 MHC 分子例如 HLA 分子或 DR 分子的表位。这些分子结合于具有由大约 8 或 9 个氨基酸所间隔的校正锚氨基酸的多肽。这些肽可通过检测 TARP 的氨基酸序列或者通过本领域技术人员所公知的有关 MHC 结合基元的知识进行鉴定。

15

TARP、其免疫原性片段和 TARP 类似物，可被重组合成。TARP 的免疫原性片段和 TARP 自身的 58 个残基也可通过常规的方法化学合成。如果需要，多肽也可通过已有技术来化学合成。一种这样的方法被描述于 W. Lu 等, *Federation of European Biochemical Societies Letters*. 429: 31- 35 (1998)。

20

#### IV. TARP 核酸

本发明的一个方面提供了含有编码 TARP 多肽的核苷酸序列的分离的重组核酸分子（参见，例如附图 14）。该核酸可被用于表达 TARP，然后所表达的 TARP 可被用于例如产生用于治疗目的的抗体。如上所述，该核酸分子具有 3 个阅读框架，每一个编码由不同可读框所定义的多肽。在此处所述的实施方案中，目的阅读框架就是编码 TARP 的阅读框架。

25  
30

如上所述，两个阅读框架在体外翻译系统中被翻译。获自 LNCaP cDNA 的约 1.1 kb PS-TCR $\gamma$  转录本 (SEQ ID NO: 13)的核苷酸序列和当转录本在核苷酸位置 74 (PS-TCR $\gamma$ -1, SEQ ID NO: 14) 以及核苷酸位置 247 (PS-TCR $\gamma$ -2, SEQ ID NO: 15)的起始密码子开始被翻译的推导的氨基酸序列被描述在附图 1 中。转录的起始点（加下划线）在 J $\gamma$ 1.2 区段的前 10 个核苷酸内。序列的数据可获自 EMBL/GenBank/DDBJ，其登录号为 AF151103。应说明的是现在已证实真正的"+1"位置是附图 1 所示的序列的第 6 个核苷酸。

研究者可以利用这个序列制备 PCR 引物，用于分离本发明的核苷酸序列。LNCaP 细胞是约 1.1 kb 转录本序列 cDNA 的有用的来源。来自尚未进行 TCR $\gamma$ 基因重排的人细胞（例如，除了 T-淋巴细胞前体之外的细胞）的基因组 DNA，可用作在转录后被加工成为约 1.1 kb 转录本的长序列。该序列可利用公知技术被修饰成为一个编码本发明相关分子的核酸。

含有本发明序列的核酸分子可用体外的方法被克隆或扩增，例如聚合酶链式反应 (PCR)，连接酶链式反应 (LCR)，以转录为基础的扩增系统 (TAS)，自我维持的序列复制系统(3SR)和 Q $\beta$ 复制酶扩增系统(QB)。例如，编码所述蛋白的多核苷酸可利用以该分子的 DNA 序列为基础的引物通过 cDNA 的聚合酶链式反应被分离。

众多的克隆和体外扩增方法是本领域技术人员所公知的。PCR 方法被描述在，例如，U. S.专利. No. 4,683,195; Mullis 等 (1987) Cold Spring Harbor Symp. Quant. Biol. 51: 263; 和 Erlich, ed., PCR TECHNOLOGY, (Stockton Press, NY, 1989)。多核苷酸也可通过在严格杂交条件下用选自目的多核苷酸序列的探针筛选基因组或 cDNA 文库进行分离。

核酸的工程化形式可通过编码所述蛋白的多核苷酸的定点诱变来

进行，或通过加入 0.1 mM MnCl<sub>2</sub> 和不平衡浓度的核苷酸增加 PCR 反应中原始多核苷酸中错误率导致的随机突变来进行。

### 1. 表达载体

5 本发明也提供了用于表达由 TARP 转录本编码的多肽的表达载体。表达载体可通过包含用于转录和 mRNA 的翻译的适宜的启动子、复制序列、标记物等来使其适用于真核和原核表达系统。表达载体的构建和在转染细胞中基因的表达涉及本领域公知的分子克隆技术的应用。Sambrook 等, MOLECULAR CLONING--A LABORATORY  
10 MANUAL, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, (1989) 和 CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, F. M. Ausubel 等编辑, (Current Protocols, a joint venture between Greene Publishing Associates, Inc. 以及 John Wiley & Sons, Inc.)。用于这样目的有用启动子包括金属硫蛋白启动子, 组成型腺病毒主要后期启动子, 地塞米松诱导型 MMTV 启动子, SV40 启动子, MRP polIII 启动子, 组成型 MPSV 启动子, 四环素诱导型 CMV 启动子(例如, 人早  
15 早期 CMV 启动子), 以及组成型 CMV 启动子。用于基因治疗的质粒可能包括其它的功能元件, 例如可选择的标记、鉴别区和其它的基因。

20 用于本发明的表达载体依赖于其目的用途。当然这样的表达载体必须含有与宿主细胞相容的表达和复制信号。用于表达生物活性结合物的表达载体包括病毒载体, 例如逆转录病毒、腺病毒和腺相关病毒, 质粒载体, 粘粒, 等等。病毒和质粒载体优选转染哺乳动物细胞。其中表达控制序列含有 CMV 启动子的表达载体 pcDNA1 (Invitrogen, San  
25 Diego, CA), 提供了较好的转染和表达率。腺相关病毒载体在本发明中适用于基因治疗的方法。

很多的方法适用于将多核苷酸传递到细胞, 包括, 例如由细胞从溶液中直接摄取分子, 通过脂质转染法(例如脂质体或免疫脂质体)  
30 协助摄取, 颗粒介导的转染, 以及从具有可操作连接于编码抑制性多

核苷酸的核苷酸序列的表达调控序列的表达盒进行胞内表达。参见 U. S. 专利 5,272,065 (Inouye 等);酶学方法, vol. 185, Academic Press, Inc., San Diego, CA (D. V. Goeddel, ed.) (1990) 或 M. Krieger, GENE TRANSFER AND EXPRESSION—A LABORATORY MANUAL, Stockton 出版社, New York, NY, (1990)。虽然通过体外化学合成法制备短的寡核苷酸更经济, 重组 DNA 表达质粒也被用于制备本发明的多核苷酸, 用于通过除基因治疗以外的方法进行传递。

构建体也可以含有一个标记物从而简化蛋白的分离程序。例如, 一个多组氨酸标记, 例如, 六个组氨酸残基可在蛋白的氨基末端被引入。多组氨酸标志使得蛋白仅通过镍-螯合层析法一步即可提取。

## 2. 重组细胞

本发明也提供了含有用于表达本发明核苷酸序列的表达载体的重组细胞。可选择高水平表达的宿主细胞来纯化蛋白。细胞可以是原核生物细胞, 例如 *E. coli*, 或真核细胞。有用的真核细胞包括酵母和哺乳动物细胞。该细胞可以是例如, 在培养物中的重组细胞或体内细胞。

表达 TARP 的细胞适用于使个体抵抗表达这些肽的细胞的主动或被动免疫。在特定的实施方案中, 这些细胞是细菌细胞。在主动免疫的情况下, 重组细胞是个体的自身固有的细胞, 其可以与 HLA 分子联合提呈多肽。例如抗原提呈细胞就适用于该目的。在此情况下, 其优选利用“自身固有的细胞”, 也即这些细胞来源于个体。这样的细胞是 MHC 相容性的。编码 TARP 的核苷酸序列在这样的细胞中应该被一个组成型启动子控制, 因为一个目的就是在细胞表面表达高密度的蛋白, 优选的, 表达比健康前列腺上皮细胞具有更大密度的多肽。

## V. 激发抗表达 TARP 细胞的细胞介导的免疫反应的方法

TARP 由上皮来源的前列腺癌细胞和很多乳房癌细胞表达。因此, TARP 可被用作干预治疗前列腺癌和表达 TARP 的乳房癌的靶, 也可

被用作分别从前列腺或乳房转移的癌细胞的标志物。本发明提供了免疫治疗学的方法用于治疗前列腺癌和表达 TARP 的乳房癌。本发明涉及对个体进行免疫，使之抗 TARP，即激发抗表达 TARP 细胞的细胞介导的免疫反应。免疫可以是主动的或被动的。在主动免疫中，免疫反应在个体体内被激发。在被动免疫中，被激活的抗多肽的 Tc 细胞在体外培养并给药于个体。预料这样的方法导致表达 TARP 的健康上皮前列腺组织被破坏。然而，前列腺并非重要的器官，其损伤必然抵消个体由于前列腺癌导致的寿命的减少，且前列腺实际上可在建立 TARP 免疫治疗前通过外科手术去除。因为正常的乳房组织尚未发现表达大量的 TARP，抗表达 TARP 细胞的免疫反应并未表现出导致女性正常细胞的损伤。因此 TARP 组合物可预防给药于女性来提供免疫，防止表达 TARP 的乳房癌发育到后期。

免疫剂可以是全长的 TARP，含有 TARP 的抗原决定基的肽，例如，TARP 的免疫原性片段，或基本上等同于 TARP 的蛋白或肽。当激发抗 TARP 的细胞介导的免疫反应时含有抗原决定基的优选肽是带有个体的 HLA 分子的结合基元的肽。这些基元是本领域公知的。例如 HLA-A2 是人类的一个普遍的等位基因。此分子的结合包括具有在第二位具有亮氨酸或甲硫氨酸且在最后位置为缬氨酸或亮氨酸的 9 或 10 个氨基酸的多肽。基于 TARP 的多肽序列，可鉴定含有任意特定的 HLA 分子的基元的氨基酸序列。可通过任一典型的方法(例如，重组法，化学法,等等)制备含有这些基元的肽。因为 TARP 是一种自身蛋白，优选的含有 HLA 结合基元的氨基酸序列为那些编码亚显性的或阴性的表位的氨基酸序列。这些表位可通过 HLA 分子的与分子中的其它表位相比的较低的结合亲和力或与其它的结合于 HLA 分子的分子比较来鉴定。

含有来自包括 HLA 结合基元的 TARP 的氨基酸序列的多肽也适用于激发免疫应答。在一定程度上，因为，这样的蛋白将被细胞加工为可结合于 HLA 分子且具有 TARP 表位的肽。

HLA 分子和肽抗原的复合物充当被 HLA 限制性 T 细胞识别的配体(Buus, S. 等, *Cell* 47: 1071,1986; Babbitt, B. P. 等, *Nature* 317: 359,1985; Townsend, A.和 Bodmer, H., *Annu. Rev. Immunol.* 7: 601, 1989; 5 Germain, R. N., *Annu. Rev. Immunol.* 11: 403,1993)。通过单氨基酸替代的抗原类似物的研究和内源结合的自然加工的肽的测序, 与特异性结合于 HLA 抗原分子所必需的基元相对应的关键残基已被鉴定(参见, 例如., Southwood, 等, *J. Immunol.* 160: 3363, 1998; Rammensee,等, *Immunogenetics* 41: 178,1995; Rammensee 等, Sette, A.和 Sidney, J. *Curr. Opin. Immunol.* 10: 478,1998; Engelhard, V. H., *Curr. Opin. Immunol.* 6: 10 13,1994; Sette, A. 和 Grey, H. M., *Curr. Opin. Immunol.* 4: 79,1992)。

此外, HLA-肽复合物的 x-射线晶体学分析显示在 HLA 分子的肽结合性裂口中的口袋(蛋白质的立体结构)以等位基因的方式容纳通过肽配体含有的残基; 这些残基反过来决定了含有这些残基的肽与 HLA 结合的能力。(参见, 例如., Madden, D. R. *Annu. Rev. Immunol.* 13: 15 587, 1995; Smith, 等, *Immunity* 4: 203,1996; Fremont 等, *Immunity* 8: 305,1998; Stern 等, *Structure* 2: 245,1994; Jones, E. Y. *Curr. Opin. Immunol.* 9: 75,1997; Brown, J. H. 等, *Nature* 364: 33, 1993)。

20

相应地, I 类和 II 类等位基因特异性的 HLA 结合基元或 I 类和 II 类超级基元的定义, 涵盖具有结合特定的 HLA 分子的可能性的 TARP 的区域。

25

与 TARP 有高度序列同一性的分子对于激发免疫反应也是有用的。此种分子可以被免疫系统识别为“外源的”, 产生抗体或与 TARP 发生交叉反应的 CTL。与 TARP 具有高度序列同一性的分子包括非人源的 TCR $\gamma$ 同源物, 特别是来自灵长目的。氨基酸序列至少与 TARP 有 90% (虽然其可为 91%, 92%, 93%, 94%, 95%, 或甚至与 TARP 有更高的序列同一性) 相同且其特异性地结合于与 TARP 特异性结合的抗体 30

的 TARP 类似物可以被使用。此外,在此方面被认为有用的还有 TARP 类似物,也即,其氨基酸序列与 TARP 至少 90%(虽然其可为 91%, 92%, 93%, 94%, 95%, 或甚至与 TARP 有更高的序列同一性)相同且激活 T-淋巴细胞使之抵抗表达 TARP 的细胞的肽。

5

另一个与 TARP 抗原决定基基本上同源的分子可通过修饰天然 TARP 表位的序列使其以较大的亲合力结合于 HLA 分子来制备。

10 鉴定编码抗原决定基的基因的方法如下: 获自患有转移性癌症的个体的 TIL 被培养并测试其在体外识别自体固有癌的能力。将这些 TIL 给药到个体中以识别导致肿瘤消退的 TIL。这些 TIL 被用于筛选表达被 TIL 识别的表位的基因的表达文库。然后用这些基因免疫个体。可选择地,在体外致敏淋巴细胞来抵抗这些基因编码的抗原,然后致敏的淋巴细胞被转移到个体体内并测试其导致肿瘤消退的能力。  
15 Rosenberg, 等, (1997) Immunol. Today 1997 18: 175。

这些分子的应用现在已被描述。这些方法也被描述在 Rosenberg 等. (1997) Immunol. Today 18: 175 和 Restifo 等.(1999) Oncology 11: 20 50。

20

一种引起免疫应答的方法包括单独地用含有来自 TARP 的抗原决定基的多肽免疫个体,或者更优选地与佐剂结合,例如弗氏不完全佐剂,类脂或脂质体, gp96, Hsp70 或 Hsp90。多肽可以为 TARP, TARP 的抗原片段,含有抗原决定基的融合蛋白,或含有与这样的抗原决定基大体上相同的序列的肽。  
25

另一个方法包括将含有来自 TARP 的表位的肽结合到抗原提呈细胞上并将细胞给药于个体。

30

在另一个方法中,将在表达盒中含有编码具有来自 TARP 抗原决

定基的多肽的核酸序列的重组病毒给药于个体。所述病毒选择性地也可编码细胞因子(例如, IL-2)、共刺激分子或其它的增强免疫应答的基因。病毒可以是, 例如, 腺病毒、鸟痘病毒或牛痘病毒。通过感染, 感染的细胞将表达 TARP 肽并在与具有和抗原决定基相同基元的肽结合的 HLA 分子结合的情况下在细胞表面表达表位。这些细胞然后激活能识别被提呈的抗原的 CTL, 导致含有决定基的癌细胞的破坏。

在另一个方法中, 通过, 例如, 肌内、biolistic 注射或与脂类连接, 个体用编码含有来自 TARP 的抗原决定基的多肽的裸 DNA 致免疫。此方法显示产生了对细胞介导的抵抗表达所述编码多肽的细胞的应答的刺激。

在另一种方法中, 表达表位的重组细菌例如 *Bacillus Calmette-Guerin* (BCG), *Salmonella* 或 *Listera* 给药于个体, 可选择地, 重组细菌也编码细胞因子、共刺激分子或其它增强免疫应答的基因。

在另一种方法中, 表达抗原的细胞被给药于个体。包括, 例如, 用 TARP 表位刺激的树突状细胞, 用含有 TARP 抗原决定基的多肽、HLA 和 B7 基因转染的细胞。多重转染导致产生了一些对于将抗原决定基提呈到细胞表面上所必需的几个组分。在一个实施方案中, 分子为一种融合蛋白, 含有抗原决定基的多肽被融合于 HLA 分子(通常通过接头)来提高肽与 HLA 分子的结合。在一个实施方案中, 细胞为抗原提呈细胞。优选的, 细胞为真核细胞, 较优选的为哺乳动物细胞, 更优选的为人细胞, 最优选的为获自所述个体的自体固有的人细胞。

在另一种方法中, 抗原提呈细胞(APC)在体外被刺激或与含有来自 TARP 的表位的肽共温育。这些细胞被用于致敏 CD8 细胞, 如来自前列腺癌肿瘤或外周血淋巴细胞的肿瘤浸润淋巴细胞。TIL 或 PBL 优选来自所述个体。然而, 它们至少为在个体具有的 HLA 型所限制的

MHC I 类细胞。这些致敏细胞然后给药于个体。

在补充的方法中，任意的这些免疫疗法通过给药细胞因子被加强，例如 IL-2, IL-3, IL-6, IL-10, IL-12, IL-15, GM-CSF, 干扰素。

5

除了上述评价肽的免疫原性的方法外，免疫原性也可以通过评价来自正常个体的初级 T 细胞培养物来进行 (参见,例如, Wentworth, P. A. 等, Mol. Immunol.32: 603,1995; Celis, E. 等, Proc. Natl. Acad. Sci. USA 91: 2105,1994; Tsai, V. 等, J. Immunol. 158: 1796, 1997; Kawashima, I. 等, Human Immunol. 59: 1,1998); 通过 HLA 转基因鼠的免疫来进行(参见,例如, Wentworth, P. A. 等, J. Immunol.26: 97,1996; Wentworth, P. A. 等, Int. Immunol. 8: 651,1996; Alexander, J. 等, J. Immunol. 159: 4753,1997),以及通过证实已被有效接种或具有肿瘤的患者记忆 T 细胞应答来进行(参见,例如, Rehermann, B. 等, J. Exp. Med. 181: 1047,1995; Doolan, D. L. 等, Immunity 7: 97,1997; Bertoni, R. 等, J. Clin. Invest. 100: 503,1997; Threlkeld, S. C. 等, J. Immunol. 159: 1648,1997; Diepolder, H. M. 等, J. Virol. 71: 6011,1997)。

在选择用于疫苗组合物的 CTL 诱导性目的肽时，与 I 类 HLA 分子有较高结合力的肽是优选的。肽的结合被通过测试候选肽与纯化的 HLA 分子在体外结合的能力来进行评价。

为确保 TARP 类似物当用作疫苗时能在体内引起 CTL 对 TARP 的应答(或者,为 II 类表位时激发与野生型肽交叉反应的辅助 T 细胞), TARP 类似物可被用于在体外免疫来自含适当 HLA 等位基因的个体的 T 细胞。其后,评价免疫细胞诱导 TARP 致敏靶细胞的溶解的能力。

25

通常,来自 TARP 的肽或其类似物(本发明的肽)被合成,并测定其结合 HLA 蛋白和激活 HTL 或 CTL 应答或二者的能力。

常规的用于检测 T 细胞应答的试验包括增殖试验, 淋巴因子分泌试验, 直接细胞毒性试验, 以及有限稀释试验。例如, 与肽温育的抗原提呈细胞被检测其在效应器细胞群中诱导 CTL 应答的能力。

5 PBMC 可被用作 CTL 前体的效应器细胞来源。合适的抗原提呈细胞被与肽温育, 然后负载肽的抗原提呈细胞在最适培养条件下与效应器细胞温育。阳性 CTL 激活可通过在培养物中测定杀死放射性标记靶细胞的 CTL 的存在, 特异性肽刺激的靶以及表达由其获得肽序列的抗原的内源被加工形式的靶细胞来检测。

10

抗原特异性 T 细胞的直接定量检测的方法是用荧光素标记的 HLA 四聚体复合物染色法 (Altman 等, Proc. Natl. Acad. Sci. USA 90: 10330 (1993); Altman 等, Science 274: 94 (1996))。可选择的, 胞内淋巴因子的染色,  $\gamma$ 干扰素释放分析或 ELISPOT 试验, 可被用于评价 T-细胞应答。

15

HTL 活化可通过本领域公知的技术来评价, 例如 T 细胞增殖和淋巴因子例如 IL-2 的分泌(参见, 例如 Alexander 等, Immunity 1: 751-761 (1994))。

20

## VI. 抗 TARP 的抗体

本发明的一个方面提供了一种含有特异性结合 TARP 的抗体的组合物。抗体优选具有至少  $10^6 M^{-1}$ ,  $10^7 M^{-1}$ ,  $10^8 M^{-1}$ , 或  $10^9 M^{-1}$  亲和性。本发明包括多克隆和单克隆抗体组合物。

25

很多免疫原可被用来产生特异性与 TARP 结合的抗体。全长 TARP 是合适的免疫原, 典型地, 目的免疫原为至少大约 3 个氨基酸的肽, 更典型地, 肽至少为 5 个氨基酸, 优选的该片段至少为 10 个氨基酸长, 更优选的片段的长度为 15 个氨基酸。所述肽可与载体蛋白偶联 (例如, 作为融合蛋白), 或在免疫载体中进行重组表达。抗体与之

30

结合的肽上抗原决定基通长为 3-10 个氨基酸。天然存在的多肽可以纯净的或未纯化的形式使用。

在原核或真核细胞中表达重组的多肽并采用标准技术进行纯化。  
5 将该多肽或其合成的形式注射到能产生抗体的动物中，可产生单克隆或多克隆抗体，它们随后可用于免疫测定以定性和定量地测量所述的多肽。

生产多克隆抗体的方法是本领域技术人员公知的。简言之，将免疫原（优选纯净形式的多肽、与合适的载体（如 GST、镍孔血蓝蛋白等）偶联的多肽或掺入到致免疫载体的重组牛痘病毒（参见 US 专利 No.4722848）中的多肽）与佐剂混合并用混合物免疫动物。动物对免疫原制剂的免疫应答被通过取待测血液，并测定对目的多肽反应的滴度来监测。当获得对免疫原的合适的高滴度的抗体时，从动物中收集血液并制备抗血清。如需要，进一步分级分离抗血清使其富含与多肽反应的抗体，参见，例如, Coligan (1991) CURRENT PROTOCOLS IN IMMUNOLOGY Wiley/Greene, NY; 以及 Harlow 和 Lane (1989) ANTIBODIES: A LABORATORY MANUAL Cold Spring Harbor Press, NY。  
10  
15  
20

针对 TARP 预定片段的抗体，包括结合性片段及其单链重组变体，可通过用例如与上述载体蛋白结合的片段的结合物免疫动物来产生。

单克隆抗体由分泌目的抗体的细胞制备。这些抗体被筛选用于结合正常的或修饰的多肽，或用于激动或拮抗活性的筛选。在一些情况下，从不同的哺乳动物宿主例如小鼠、啮齿动物、灵长目、人等宿主中制备单克隆抗体是可行的。制备这样的单克隆抗体的技术被描述在例如, Stites 等(编辑) BASIC AND CLINICAL IMMUNOLOGY (4th ed.) Lange Medical Publications, Los Altos, CA, 此处引入作为参考; Harlow  
25  
30

和 Lane,同上; Goding (1986) *Monoclonal Antibodies : Principles and Practice* (2d ed.) Academic Press, New York, NY; 以及, Kohler 和 Milstein (1975) *Nature* 256: 495-497。

5           其它合适的技术涉及筛选噬菌体或类似载体中重组抗体文库的技术。参见, Huse 等. (1989) *Science* 246: 1275-1281; 以及 Ward, 等. (1989) *Nature* 341: 544-546。

10           同样, 重组免疫球蛋白可被生产。参见美国专利 No. 4,816,567 (Cabilly)和 Queen 等. (1989) *Proc. Nat'l Acad. Sci. USA* 86: 10029-10033。

15           通常, 多肽和抗体通过共价或非共价连接提供可检测信号的物质而被标记。大量的标记物和连接技术是已知的且被广泛公开在科学和专利文献上。因此用于检测待分析物的抗体可被可检测的部分直接标记, 或例如通过将所述抗体与被直接或间接标记的第二抗体结合而进行间接标记。

20           本发明的抗体也被用于在 TARP 分离中的亲和层析。例如, 制备带有与固体支持物例如颗粒如琼脂糖、交联葡聚糖凝胶等连接的抗体的柱子, 使细胞溶解产物通过此柱, 洗涤, 用浓度递增的温和变性剂处理, 由此纯化的 TARP 被释放出来。

25           一种可选择的方法是通过重组 DNA 技术将非人源抗体的 CDR 区连接到人保守区产生人源化的免疫球蛋白。参见美国专利 5,585,089 (Queen 等)。

30           另一个分离编码人单克隆抗体或其结合性片段的 DNA 序列的方法是通过从人 B 细胞中筛选 DNA 文库, 其相应的常规操作参见 Huse 等, *Science* 246: 1275-1281 (1989) , 然后克隆和扩增编码具有所需特

5 异性的抗体（或结合性片段）的序列。Huse 所描述的操作与噬菌体展示技术结合会更有效，参见，例如 WO 91/17271 (Dower 等)和 WO 92/01047 (McCafferty 等)。噬菌体展示技术也可被用于诱变先前显示对 TARP 有亲合力的抗体的 CDR 区。结合亲和性得以改善的抗体被选择出来。

10 在本发明的另一个实施方案中，提供了抗 TARP 或蛋白类似物的抗体的片段。典型地，这些片段表现出类似于完整免疫球蛋白的与 TARP 结合的特异性。这些抗体片段包括分离的重链，轻链 Fab, Fab'F(ab')<sub>2</sub> 和 Fv。片段通过重组 DNA 技术，或通过完整免疫球蛋白的酶解或化学分离产生。

## VII. 靶向 TARP 的嵌合分子

15 本发明提供了靶向 TARP 的嵌合分子。嵌合分子含有一个靶向部分和一个效应器部分。嵌合蛋白对于所述多肽和含有它的细胞的检测是有用的。

### A. 靶向部分

20 本发明的嵌合分子含有一个靶向部分，该靶向部分含有一个特异性结合 TARP 的配体。优选的配体为抗体，此处所用的该术语包括抗体的结合性片段。然而，其它的这些分子的天然配体也可被使用。

### B. 效应器部分

25 效应器部分可为另一个特异性结合部分例如抗体、生长因子或配体。嵌合分子充当一个高度特异的双功能接头。此接头可连接和增强融合蛋白所结合的细胞或细胞组分之间的相互作用。

30 在另一个实施方案中，效应器分子可以是药剂（例如，药物）或含有药剂的载体。因此，特异性结合于 TARP 的部分可与药物例如长春碱、阿霉素、金雀异黄素(酪氨酸激酶抑制剂)，反义分子以及其它

的本领域技术人员公知的药剂轭合，由此特异性地使药剂靶向肿瘤细胞。

5 可选择地，靶向分子可结合于含有治疗组合物的载体。这样的载体包括但不限于脂质体，胶束，各种合成的珠子等。

10 本领域的技术人员可以理解本发明的嵌合分子可包括结合一个效应器的多个靶向部分或相反结合一个靶向部分的多个效应器。在另一个实施方案中，嵌合分子也可同时包括多个靶向部分和多个效应器的分子。适于作为本发明的嵌合分子的效应器分子组分的可检测的标记包括通过光镜、光化学的、生物化学的、免疫化学的、电学的、目测的或上述所有化学方式可检测的任意组分。

15 本领域技术人员可以理解靶向分子和效应器分子可以以任意的顺序连接。因此，当其中的靶向分子为多肽，效应器分子可被连接于靶向分子的氨基端或羧基端。靶向分子也可被连接到效应器分子的中间区域，或相反，效应器分子被连接到靶向分子的中间区域，只要这种连接不影响各自分子的活性。

20 靶向分子和效应器分子可通过本领域技术人员公知的任意方式连接。典型地，效应器分子被直接或通过接头（间隔区）连接于靶向分子。然而，当效应器分子和靶向分子都是多肽时，优选重组表达嵌合分子作为单链融合蛋白。

25 在一个实施方案中，靶向分子被化学结合于效应器分子（例如细胞毒素，标记物，配体，或药物或脂质体）。化学结合分子的方法是本领域公知的。将试剂结合到抗体或其它多肽靶向分子的方法可依据试剂的化学结构而变化。多肽典型地含有不同的功能基团，例如，羧酸 (COOH) 或自由胺(-NH<sub>2</sub>) 基团，其可通过与合适的功能团在效应器分子上反应来与效应器结合。可选择地，靶向分子和/或效应器分子  
30

可被衍生化以暴露或连接其它的反应功能基团。衍生化反应可包括连接任意的接头分子，例如获自 Pierce 化学药品公司（Rockford Illinois）的那些。

5           具有一个与特定试剂的基团反应的功能性基团以及另一个与抗体反应的基团的双功能接头，可被用来形成目的免疫结合物。可选择地，衍生化反应包括靶向分子的化学处理，例如，糖蛋白抗体的糖部分用高碘酸盐进行乙二醇裂解来产生自由醛基团。抗体上的自由醛基团可与试剂上的自由胺或肼基团反应来与试剂结合。（参见美国专利  
10           4,671,958）。在多肽例如抗体或抗体片段上产生自由巯基基团的方法是已知的（参见美国专利 4,659,839）。

          不同的化合物包括放射性核素金属螯合物、毒素和药物连接到蛋白例如抗体上的多种方法和接头分子是已知的，参见，例如欧洲专利申请 No. 188,256；美国专利 4,671,958；4,659,839；4,414,148；  
15           4,699,784；4,680,338；4,569,789；以及 Borlinghaus 等，Cancer Res. 47: 4071-4075 (1987)。具体地讲，各种免疫毒素的生产在本领域是已知的且可在 "Monoclonal Antibody-Toxin Conjugates: Aiming the Magic  
20           Bullet", Thorpe 等，MONOCLONAL ANTIBODIES IN CLINICAL MEDICINE, Academic Press, 第 168-190 页 (1982), Waldmann, Science, 252: 1657 (1991), 美国专利 4,545,985 和 4,894,443 中找到。

          当靶向分子和/或效应器分子相对较短（例如不到 50 个氨基酸）时，它们可通过标准的化学肽合成技术来合成。当两个分子都相对较短时，嵌合分子作为一个单独的连续的多肽可被合成。可选择的，  
25           靶向分子和效应器分子可被分别合成然后通过一个分子的氨基末端与另一个分子的羧基末端的缩合反应形成肽键而被融合。可选择的，靶向分子和效应器分子可与肽间隔分子的一个末端缩合来形成无间隔的融合蛋白。

30

一序列的 C-末端氨基酸被连接于一个不溶性支持物然后在序列中能够依次增加氨基酸的固相合成是本发明的肽的化学合成的优选的方法。固相合成的技术被描述于 Barany 和 Merrifield, *Solid-Phase Peptide Synthesis*; 第 3-284 页。*The Peptides: Analysis, Synthesis, Biology*. Vol. 2: *Special Methods in Peptide Synthesis, Part A.*, Merrifield 等, *J. Am. Chem. Soc.*, 85: 2149-2156 (1963), 和 Stewart 等, *Solid Phase Peptide Synthesis*, 第二版, Pierce Chem. Co., Rockford, Ill. (1984)。

在一个优选的实施方案中, 嵌合融合蛋白通过重组 DNA 技术来合成。通常, 该技术包括生成编码融合蛋白的 DNA 序列, 在特定启动子的调控下将 DNA 置于表达盒中, 在宿主中表达蛋白, 分离表达的蛋白以及, 如需要的话, 让蛋白复性。

本发明的编码融合蛋白的 DNA 可通过合适的方法被制备, 包括, 例如, 克隆和适当序列的限制性酶切或通过下列的方法直接化学合成, 例如 Narang 等, *Meth. Enzymol.* 68: 90-99 (1979)中所述的磷酸三酯法; Brown 等, *Meth. Enzymol.* 68: 109-151 (1979)中所述的磷酸二酯法; Beaucage 等, *Tetra. Lett.*, 22: 1859-1862 (1981) 中所述的二乙基亚磷酰胺法; 以及美国专利 No.4,458,066 中的固相支持体法。

当两个分子优选地基本上直接相连接时, 本领域的技术人员可以理解, 分子可被含有一个或更多个氨基酸的肽间隔区隔开。通常, 间隔区除了连接蛋白或保留最小的距离或其它的空间联系外没有特异的生物活性。然而, 间隔区的组成性氨基酸可被选择来改变分子的一些特性例如折叠、净电荷或疏水性。

编码融合蛋白的核酸序列可在不同的宿主细胞中表达, 包括 *E. coli*, 其它的细菌宿主, 酵母, 以及各种高等的真核细胞例如 COS, CHO 和 HeLa 细胞系和骨髓瘤细胞系。重组蛋白的基因被可操作地连接于对每一宿主而言适当的表达调控序列。对 *E. coli* 而言包括启动

子例如 T7、trp、或  $\lambda$  启动子、核糖体结合位点，并优选地包含转录终止信号。对真核细胞而言，调控序列包括启动子和优选的来自免疫球蛋白基因的增强子，SV40，巨细胞病毒，等等，以及多聚腺苷酸化序列，也可包括剪接供体和受体的序列。本发明的质粒和载体可通过公知的方法被转移到所选择的宿主细胞中，例如氯化钙转化 E. coli 和磷酸钙处理或电穿孔处理哺乳动物细胞。

重组融合蛋白一旦被表达，其可通过本领域的常规方法进行纯化，包括硫酸铵沉淀，亲和柱层析，层析法，凝胶电泳等（参见，R. Scopes, PROTEIN PURIFICATION, Springer--Verlag, N.Y. (1982), Deutscher, METHODS IN ENZYMOLOGY Vol. 182: Guide to Protein Purification., Academic Press, Inc. N.Y.(1990))。用于药物的大体上纯化的组分优选至少含有 90%到 95%的同质物，含有 98 到 99% 或更多的同质物为最优选的。纯化后，部分均质或达到所需同质性的多肽被用于治疗。

#### VIII. 检测表达 TARP 的细胞的方法

在本发明的另一个方面，提供了检测表达 TARP 的细胞的方法。该方法包括检测 TARP 转录本或多肽。因为上皮来源的前列腺癌细胞和很多的乳房癌细胞表达 TARP，检测方法对于前列腺癌和表达 TARP 的乳房癌的检测是有用的。具体地，前列腺癌细胞和很多的乳房癌细胞可通过 TARP 的表达来区别于其它的细胞。

组织样本可分别地选自原位的或转移性癌的任一可能的位置包括前列腺或乳房，以及末梢位置例如淋巴结和其它的器官。本领域的技术人员知道男性和女性一样患乳房癌。乳房癌在男性中比较少，仅占所有乳房癌患者的 1%。因为其不常见，所以通常在晚期才被诊断，这样就影响了生存的机会。因此，改进对男性乳房癌的诊断是可取的。

在一个方法中，对个体进行活体组织检查并在体外测试收集的组

5 织。典型地，通过溶解、超声波破坏、渗透压、冷冻和融化、酶处理或其它的本领域常规的方式破坏细胞来提取没有变性的核蛋白。细胞的内容物 (或核内容物，如果内容物被分级分离了的话)然后被与例如抗-TARP 抗体接触。产生的任意免疫复合物指示 TARP 在活体组织检查样本中的存在。为了有助于检测，抗体可以被放射性标记或与被放射性标记的效应器分子偶联。在另一个方法中，通过典型地显像系统细胞可被在体内检测。例如，该方法可包括给药于个体能够到达细胞核的标记组分。然后通过任意的已知检测标记的方法来测定标记的位置。任意的显像诊断的常规方法可以被使用。例如，顺磁同位素可被用于 MRI。

#### TARP 的检测

15 TARP 可通过本领域公知的任意方法来鉴定。在一个实施方案中，该方法包括检测带有能特异性识别多肽的配体的多肽 (例如，免疫测定)。本发明的抗体特别适用于 TARP 的特异性检测。各种基于抗体的检测方法是本领域已知。它们包括，例如，放射免疫测定，夹心免疫检测 (包括 ELISA)，免疫荧光检测，Western 印迹，亲和层析法 (亲和性配体与固相结合)，以及带有标记的抗体的原位检测。另一个检测 TARP 的方法包括通过其质量鉴定多肽，例如，通过凝胶电泳，质谱分析或 HPLC。个体的样本可取自各种适宜的来源，例如唾液、腹液、血或血产物(例如，血清)、尿、活组织检查组织(例如，淋巴结组织)，等等。

25 TARP 可在细胞中、在体外、在取自活组织检查的样本中以及在体内使用上述的显像系统来检测。

#### 编码 TARP 的转录本的检测

30 表达 TARP 转录本的细胞的检测可被通过使之接触带有与转录本特异性杂交的核酸探针的样本并检测杂交来进行。包括，例如，原位杂交法，其中的被标记的探针与样本接触并通过检测结合了的标记

来检测杂交。然而，转录本在样本存在的量可能很少，因此，其它的方法被用来扩增，例如 RT-PCR。在这些方法中，选择探针，其起到从 mRNA 特异性扩增 TARP 序列的扩增引物的作用。然后，扩增的序列被通过典型的方法来检测。

5

探针被选择以特异性地与 TARP 转录本杂交。通常，使用互补探针。然而，探针不需要完全的互补，只要其具有足够的序列同源性和长度来在严格条件下杂交即可。

## 10 IX. 药物组合物

在另一个方面，本发明提供了一种药物组合物，其含有可药用的载体以及本发明的组分。

15 在一个实施方案中，药物组合物含有有效量的能在个体中激发细胞介导的免疫反应或体液应答的 TARP，其免疫原性片段例如含有 TARP 表位的多肽，或 TARP 类似物，例如含有 MHC 结合基元的多肽。这样的药物组合物可用作本发明治疗方法中的疫苗和制备抗体。

20 在另一个实施方案中，药物组合物含有包含编码以有效量来在个体中引发抗表达 TARP 的细胞的免疫反应的 TARP 多肽的核苷酸的核酸分子。这样的组合物在本发明的治疗方法中也是有用的。

25 在另一个实施方案中，所述药物组合物含有可以特异性裂解编码 TARP 的核苷酸序列的核酶，一种可与此核酸结合的反义分子，或含有编码 TARP 的核酸的表达盒，来调整 TARP 在目的细胞中的表达。

30 在另外一个实施方案中，所述药物组合物可含有包含靶向分子和指示分子的嵌合分子以检测表达 TARP 的细胞。如指示分子是能够特异性与编码 TARP 的核酸结合分子（例如能特异性地与编码 TARP 的 DNA 结合的 DNA 结合蛋白），则组合物可被用于检测表达此核酸

的细胞。

本发明的药物组合物可被制备成单位剂量的形式来给药于个体。给药的量和时间取决于医师欲达到的治疗目的。

5

## 实施例

### 实施例 1. 前列腺细胞中 T-细胞受体 $\gamma$ -链的检测

我们发现 T-细胞受体 $\gamma$ 链 (TCR $\gamma$ ) mRNA 在人前列腺中表达, 并显示其最初来自前列腺上皮细胞而不是浸润的 T-淋巴细胞。相反, T-细胞受体 $\delta$ 链(TCR $\gamma\delta$ )基因在人前列腺中是沉默的。前列腺中的大部分 TCR $\gamma$ 转录本与在胸腺、脾脏和白细胞中表达的转录本大小不同。其在正常的前列腺上皮中、前列腺的腺癌以及前列腺腺癌细胞系 LNCaP 中表达。RNA 来源于未重排的 TCR $\gamma$ 座位且在紧挨 J $\gamma$ 1.2 基因区段上游的内含子序列中起始。含有 J $\gamma$ 1.2 和 C $\gamma$ 1 基因区段的前列腺特异性 TCR $\gamma$ 转录本, 其具有在 3'末端包括聚腺苷酸化信号和 poly(A)序列的非翻译序列。前列腺上皮细胞表达了被认为仅由 T 淋巴细胞表达的高水平的来自基因的转录本, 此发现是新颖的并大大地出乎意料。

10

15

#### 1. 材料和方法

##### 20 RNA 点印迹和 Northern 印迹杂交

对各种人的组织进行 RNA 点印迹 (RNA master blot, Clontech, Palo Alto, CA)和 Northern 印迹 (MTN, Clontech, Palo Alto, CA)。

对来自前列腺癌细胞系, LNCaP 和 PC-3 (ATCC, Rockville, MD) 的 mRNA 进行 Northern 印迹。poly (A) RNA 的分离使用 FastTrack 试剂盒 (InVitrogen, Carlsbad, CA)来进行。按照现行的做法, 将 RNA 在 1%的琼脂糖凝胶上跑电泳并被转到尼龙基膜 (GeneScreen Plus, DuPont, Wilmington, DE) 上。Ausubel, 同上。由 EST 质粒 ng79d11(Genome Systems, St. Louis, MO)制备对 TCR $\gamma$ 转录本的非翻译 3'末端 (3'UTR) 特异的 cDNA 探针。对 TCR $\gamma$ 转录本 (TCR C $\gamma$ ) 恒

25

30

定区特异的探针由 LNCaP cDNA 制备而成,对 TCR $\delta$ 转录本 (TCR C $\delta$ ) 的恒定区特异的探针由 TCR $\delta$ 质粒制备而成。人 $\beta$ -肌动蛋白探针被用作 mRNA 制备物的定量对照。探针被用  $^{32}\text{P}$  标记(Lofstrand Labs Limited, Gaithersburg, MD) 通过随机引物延伸到  $1\mu\text{Ci}/\text{ng}$  的比活性。RNA 膜被在含有 50%甲酰胺的杂交溶液中  $45^\circ\text{C}$  封闭 2 个小时(Hybrisol I,Oncor, Gaithersburg, MD)然后在 20ml 杂交溶液中用  $20\mu\text{Ci}$  cDNA 探测 15 小时。膜在室温下用  $2\times\text{SSC}/0.1\%$  SDS 洗涤两次,每次 15 分钟,以及  $55-65^\circ\text{C}$  在  $0.1\%$  SSC/ $0.1\%$  SDS 中洗涤两次,每次 20 分钟,膜在显影前被在  $-80^\circ\text{C}$  下给底片成像显影 (X-OMAT, Kodak, Rochester, NY)。

10

### RNA 原位杂交

TCR $\gamma$ 恒定区和 TCR $\gamma$ 的非翻译 3'末端核苷酸序列通过逆转录酶 PCR (RT-PCR) 从 LNCaP mRNA 来扩增,克隆到 pBluescript II SK (Stratagene, La Jolla, CA)中并通过 DNA 测序来证实。反义和有义 TCR $\gamma$  $^{35}\text{S}$ -核苷酸探针分别通过 T7 和 T3 RNA 聚合酶来制备。来自 NCI 的 8 个存档的前列腺经尿道切除的样本的石蜡块被恢复。所述的病例包括恶性和良性前列腺导管的病例两种。这些病例的平均年龄为 69 岁,且肿瘤的 Gleason 评分为  $3+3=6/10$  到  $4+5=9/10$ 。在玻片上进行石蜡块的处理且使用核苷酸探针杂交 (Molecular Histology, Gaithersburg, MD)。然后,杂交的玻片用苏木紫和酸性曙红复染色并利用配备能提供明视野和暗视野的可变聚光器的 Zeiss Axiophot 显微镜进行镜检。

20

### RT-PCR 分析

从 150-250ng 的 LNCaP 和 PC-3 poly(A) mRNA 利用 oligo-dT 引发分别制备单链 cDNA (Pharmacia-Biotech, Piscataway, NJ)。PCR 引物被设计来扩增 TCR $\gamma$ 转录本的不同部分。为了仅扩增 cDNA 而不扩增可能存在于 mRNA 制备物中的痕量的基因组 DNA,引物对总是结合使用来产生跨两个或更多外显子的 PCR 产物。一个完整的 PCR 用来扩增两个 TCR $\gamma$ 恒定区基因中的一个 (C $\gamma$ 1 或 C $\gamma$ 2), 在外显子 C1

30

(TCRC $\gamma$ .F)中使用正向的引物和在外显子 CIII (TCRC $\gamma$ .R4)中使用反向的引物, 附图 8。利用正向的特异于 TCR $\gamma$ 可变基因区段的四个亚型 (TCRV $\gamma$ I.F, TCRV $\gamma$ II.F, TCRV $\gamma$ III.F, TCRV $\gamma$ IV.F)的引物与 TCR $\gamma$ 恒定基因区段(TCRC $\gamma$ .RI)中的反向引物结合, 进行跨恒定区的 PCR 的变动, 附图 8。利用高保真的 PCR 组分, 蜡-介导的热启动 PCR 被进行 30 个循环 (Expand, Boehringer-Mannheim, Indianapolis, IN)。在含有 0.5 $\mu$ g/ml EtBr 的 1.2% 琼脂糖凝胶上分析 PCR 产物, 特定的 PCR 产物被凝胶纯化(Qiagen, Valencia, CA), 进行 T/A 克隆 (InVitrogen, Carlsbad, CA) 以及在自动毛细管测序仪 (Perkin Elmer Applied Systems, Foster City, CA) 上测序, 使用 Perkin-Elmer's dRhodamine 终止子循环测序试剂盒。

#### TCR $\gamma$ VJ 基因重排的分析

由  $5 \times 10^7$  LNCaP 细胞通过既定操作步骤制备基因组 DNA。Ausubel, 同上。一系列 12 个 PCR 被进行, 每个使用来自 V $\gamma$ 基因区段的四个亚型(TCRV $\gamma$ I.F, TCRV $\gamma$ II.F, TCRV $\gamma$ III.F, TCRV $\gamma$ IV.F)之一的正向引物, 并结合来自三个 J $\gamma$ 1 基因区段(TCRJ $\gamma$ 1.1.R, TCRJ $\gamma$ 1.2.R, TCRJ $\gamma$ 1.3.R)之一的反向引物, 附图 8。热启动的 PCR 被进行 30 个循环, 使用 500ng 的基因组 DNA 并在 1.2%琼脂糖凝胶上用 0.5 $\mu$ g/ml 的 EtBr 进行检查 PCR 产物。人胎盘 DNA (Clontech, Palo Alto, CA) 被用作引物的阴性对照 J $\gamma$ 1.1 至 J $\gamma$ 1.2 的基因组 DNA 进行 PCR 扩增以用作模板的阳性对照。

#### RNA 的引物延伸分析

通过 LNCaP mRNA 的引物延伸分析来测定前列腺 TCR $\gamma$ 转录本的起始位点。将 5 $\mu$ g 的 mRNA 与 0.08pmol  $^{32}$ P-末端标记的 TCRC $\gamma$ .R $_2$  引物混合, 与 C $\gamma$ 1 的 5'末端的 48-75 个核苷酸退火。使用 20 U 的 MMLV-逆转录酶(Superscript,Gibco-BRL, Gaithersburg, MD)根据既定的步骤进行分析。C. P. George 等/ (1996)"Primer-extension analysis of RNA" In A LABORATORY GUIDE To RNA, ISOLATION, ANALYSIS AND

SYNTHESIS., P. A. Krieg 编辑(Wiley-Liss, Inc., New York, NY),第 133-139 页。样本与  $^{32}\text{P}$ -末端标记的分子量标记物(MspI digested pBR322, Lofstand Labs Limited, Gaithersburg, MD)一起在 6% 聚丙烯酰胺-脲 DNA 测序凝胶上电泳。电泳结束后将凝胶印迹到 Whatman 纸上,干燥并放射自显影。

### 5'-RACE PCR 分析

通过利用 Marathon cDNA 扩增试剂盒(Clontech, Palo Alto, CA)和 25pmole 的 TCR $\gamma$ 基因特异性引物(TCRC $\gamma$ .R3)由 500 ng 的 LNCaP poly(A) mRNA 制备双链 cDNA, 附图 8。Marathon-连接物然后被结合到合成的 cDNA 的末端。使用基因特异性引物(TCRC $\gamma$ .R2)进行 5'-cDNA 末端 (5'-RACE) PCR 的快速扩增, 附图 8。用于逆转录的引物的上游与连接物特异性引物退火。热启动条件被应用(Advantage, Clontech, Palo Alto, CA) 并且 PCR 产物如 RT-PCR 所述被分析和克隆。来自 5'-RACE PCR 分析凝胶的 DNA 被转到尼龙膜上并且与上游序列杂交的  $^{32}\text{P}$ -末端标记的引物(TCRC $\gamma$ .R1), 附图 8 所示, 被应用来检测 EtBr /UV 没有检测到的可能的条带。

### 体外转录与翻译

通过 RT-PCR 和 5'RACE PCR 获得的完整的前列腺 TCR $\gamma$ 转录本, 被使用 T7 RNA 聚合酶和小麦胚芽提取物(TNT, Promega, Madison, WI)通过 RT-PCR 扩增, 克隆到 pBluescript II SK (Stratagene, La Jolla, CA) 中, 进行测序并在体外转录翻译偶联的系统中检测。  $^{35}\text{S}$ -Met (ICN, Costa Mesa, CA) 被加到反应中用于翻译产物的显影。反应在还原条件下在聚丙烯酰胺凝胶 (16.5% Tris/Tricine, BioRad, Hercules, CA)上与预染色的标记物 (Gibco-BRL, Gaithersburg, MD)一起被分析。凝胶被干燥并用于放射自显影。

### 结果

A. 通过数据库分析鉴定存在的 TCR $\gamma$ 的前列腺 EST

我们从 20 个来自 6 个肿瘤和 2 个正常前列腺 cDNA 文库的 cDNA 克隆中鉴定了 23 个 TCR $\gamma$  ESTs。由前列腺 EST 装配成的 TCR $\gamma$ 复合序列由具有 76 个核苷酸的 TCR $\gamma$ 恒定区序列、448 个核苷酸的非翻译 3' 区序列和 poly (A)序列。通过前列腺 EST 与外周血 T-淋巴细胞细胞系的成熟 TCR $\gamma$ 转录本(GenBank Acc. No. M16768, M16804 和 M30894)的  
5 的比对, 我们发现前列腺 EST 复合序列等同于来自外周血 T-淋巴细胞的 TCR $\gamma$ 转录本。dbEST 数据库分析显示 TCR $\gamma$ 基因在人前列腺中是高度转录的。

#### 10 B. 通过 RNA 点印迹验证 TCR $\gamma$ (3'UTR)在人前列腺中的表达

为了分析人前列腺中有转录活性的 TCR $\gamma$ , 来自 TCR $\gamma$ 转录本的非翻译 3'末端 (3'UTR) 的 cDNA 探针被用于检测来自 50 个不同人组织的 mRNA, 附图 2A。我们确证正常的前列腺 (C7 位置) 表达 TCR $\gamma$ mRNA 并且我们进一步观察到在所有在点印迹上显示的组织中前列腺具有最强的表达。TCR $\gamma$ 基因表达也被发现在小肠(E3), 脾脏 (E4),  
15 胸腺 (E5), 外周白细胞 (E6), 淋巴结(E7), 骨髓 (E8), 以及肺(F2) 中。

#### 20 C. Northern 显示人前列腺中的两种大小特异性的 TCR $\gamma$ 转录本

利用 3' UTR 探针的 Northern 印迹杂交显示前列腺具有两种大约 1.1 和 2.8kb 的 TCR $\gamma$ 转录本, 附图 2B (泳道 3), 而脾脏, 胸腺, 小肠和白细胞中的主要转录本为 1.5 kb。1.5 kb 大小的转录本与来自 $\gamma\delta$  T-淋巴细胞的 TCR $\gamma$ mRNA (GenBank Acc. No. M16768, M16804, (Krangel 等, Science 237,64-67 (1987));M30894, (Littman 等. Nature 326,85-88  
25 (1987))一致。由于数据库分析表明了 TCR $\gamma$ 的恒定区是前列腺转录本的一部分, 所以我们也使用 TCR $\gamma$ 恒定区探针(TCR C $\gamma$ )。我们也在前列腺中发现了同样的 1.1 kb 和 2.8 kb 条带, 附图 3A (泳道 3)。

#### 30 D. 表达 TCR $\gamma$ 的前列腺细胞不表达 TCR $\delta$ 或 CD3 转录本

TCR $\gamma$ -链蛋白通常与 TCR $\delta$ 链蛋白共表达。由于 TCR $\gamma$ 基因在人前

列腺中是转录活性的，我们继续分析 TCR $\delta$ 基因的转录活性。使用 TCR $\delta$ 转录本核苷酸序列进行 dbEST 分析 (<http://www.ncbi.nlm.nih.gov/BLAST>)。来自前列腺 cDNA 文库的 EST 与 TCR $\delta$ -链转录本的任一部分都不匹配。此外，Northern 印迹分析没有检测到任何 TCR $\delta$  mRNA 的前列腺表达，附图 3B (泳道 3)。我们得出结论：TCR $\delta$ 基因在前列腺中是沉默的。如所预料的，TCR $\delta$ 在脾脏，胸腺和血液白细胞中表达，附图 3B。

E. LNCaP 细胞，而不是 PC-3 细胞，表达前列腺-特异性 TCR $\gamma$ 转录本  
假定 TCR $\gamma$ mRNA 在正常的前列腺中表达，我们接下来分析是否其也在前列腺癌中表达。前列腺-特异性 1.1 kb 转录本被发现在来自 LNCaP 的 mRNA 制备物中，但在来自 PC-3 的 mRNA 制备物中，附图 3C。在正常的前列腺中表达的前列腺-特异性 2.8 kb 转录本，也存在于 LNCaP 中，尽管其浓度较低。

F. RNA 原位杂交显 TCR $\gamma$ 在前列腺上皮细胞中表达

前列腺由腺泡组织和可变混合的初级导管衬上皮细胞群组成，所述上皮细胞在腺隔室内排成复杂的增生导管。这些隔室被紧紧地连接于光滑的肌肉细胞、成纤维细胞和前列腺基质的其它细胞类型。为了测定人前列腺 TCR $\gamma$ 表达的细胞内位置，用 TCR (C $\gamma$ -3' UTR)有义和反义核苷酸探针进行 RNA 原位杂交。我们发现 TCR $\gamma$  mRNA 在前列腺腺泡导管的上皮细胞中高度表达而在前列腺基质细胞和其它类型的细胞中为阴性，附图 4A, 4C。前列腺的增生和瘤形成的区域也被检测到 TCR $\gamma$ 的表达。在良性的和肿瘤的腺泡上皮中的表达是相当的。在人肾组织中或在人大脑中并没有观察到 TCR $\gamma$ 的表达，附图 4E。

G. 前列腺 TCR $\gamma$ 转录本含有 C $\gamma$ 1 但不含有任何 VJ $\gamma$ 基因

在已经确定前列腺的 TCR $\gamma$  $\delta$ 表达概况后，我们继续鉴定主要的 1.1kb 前列腺特异的 TCR $\gamma$ 转录本的特性。LNCaP 细胞系被用于特性分析，因为人们不能排除来自大量前列腺组织获得的 mRNA 制备物

中有浸润 T-细胞 mRNA 污染的可能性。我们从数据库分析可知前列腺 TCR $\gamma$ 转录本的 3'末端序列等同于来自外周血液白细胞的 TCR $\gamma$ 转录本的 3'末端序列，且聚腺苷酸化信号的位置是相同的。因此，前列腺和白细胞之间的转录本大小的不同取决于通过前列腺 EST 鉴定的上游序列的不同。用 RT-PCR 扩增 TCR $\gamma$ 转录本的恒定区部分来鉴定 TCRC $\gamma$ 1 基因。较大的 TCRC $\gamma$ 2 不在 LNCaP 中表达。跨越可变区 (V $\gamma$ ) 到恒定区(C $\gamma$ )的 RT-PCR 没有产生任何的产物，表明 V $\gamma$  不是前列腺特异的 TCR $\gamma$ 转录本的一部分。

#### 10 H. LNCaP 在 TCR $\gamma$ 座位没有发生 VJ 基因重排。

由于 RT-PCR 扩增 TCR $\gamma$ 的可变区没有产生任何产物，我们接下来分析 TCR $\gamma$ 座位。在 $\gamma\delta$ T-细胞发育的过程中，TCR 座位发生 V (D) J 基因重排来使得基因区段连接在一起，形成受体的可变区。为了说明 LNCaP 细胞是否发生 TCR $\gamma$  VJ 基因重排，联合使用 TCRV $\gamma$ 和 TCRJ $\gamma$ 引物对基因组 DNA 进行 PCR，来涵盖每一种可能的重排(参见材料和方法部分)。引物联合没有产生任何的 PCR 产物，此说明了 LNCaP 细胞没有发生 TCR $\gamma$ 座位的 VJ 基因的重排。TCR $\gamma$  VJ 的重排没有在前列腺上皮细胞中发生说明了前列腺表达不同于成熟的 $\gamma\delta$ T-淋巴细胞的表达。

20

#### I. 前列腺上皮细胞表达 TCR(JC) $\gamma$ 转录本

既然确定了前列腺 TCR $\gamma$ 转录本含有 C $\gamma$ 但不含有任何 V $\gamma$ 基因区段，我们下面分析 C $\gamma$ 1 的上游序列是什么。进行 RNA 引物延伸和 5' RACE PCR 来获得转录的起始点。LNCaP mRNA 的引物延伸试验，显示大约 128 个核苷酸的大量的带和 130-135 个核苷酸的少量的带，附图 5。从 C $\gamma$ 1 的 5'末端碱基 75 处开始逆转录 (参见：材料和方法)，转录本具有大约 C $\gamma$ 1 上游序列的 53 个核苷酸。LNCaP cDNA 的 5' RACE PCR 发现了一个特异性 PCR 产物。扩增的产物被发现含有恰当地与 C $\gamma$ 1 基因区段剪接的 J $\gamma$ 1.2 基因区段。大量通过 RACE PCR 分离的克隆被测序。它们从引物延伸试验所定义的起始位点附近启动。转

30

录本的可变的起始点与在引物延伸试验中鉴定的少量的带比大量的带稍长的结果一致。前列腺 TCR $\gamma$ 如何转录和剪接的图示见附图 6。获自 LNCaP 的 TCR $\gamma$ 转录本的核苷酸序列被显示于表 1 中。该复合序列是 1020 $\pm$ 3 个核苷酸长。其含有来自 J $\gamma$ 1.2 基因区段的 53 个碱基，C $\gamma$ 1 的 519 个碱基，接着是含有聚核苷酸化信号的非翻译序列以及在 3' 末端 poly (A) 序列的 448 个碱基。

#### J. 前列腺特异性 TCR $\gamma$ 转录本在体外的翻译

前列腺转录本在原始 TCR $\gamma$ 阅读框架中具有 4 个翻译起始密码子 (ATG)，表 1 中加双下划线表示。由 4 种不同起始点计算的蛋白大小分别为 12.8, 12.0, 7.2 和 3.2 kDa。为了分析前列腺转录本的翻译活性，利用全长前列腺 TCR $\gamma$ cDNA 进行体外转录偶联的翻译。结果获得两个大约 8 和 13 kDa 的蛋白，附图 7 (泳道 1)。阴性对照没有产生任何蛋白产物。

15

#### 讨论

##### 前列腺上皮细胞中 TCR $\gamma$ 转录本的特异性表达细胞

我们在人前列腺中鉴定到 T-细胞受体 $\gamma$ 链 (TCR $\gamma$ ) mRNA 的表达，表明其来源于前列腺上皮细胞而不是来源于浸润 $\gamma\delta$  T-淋巴细胞。我们也阐明了 T-细胞受体 $\delta$ 链 (TCR $\delta$ ) 基因在前列腺中是沉默的。TCR $\gamma$  mRNA 在前列腺腺泡管的上皮细胞中表达，也在前列腺癌的上皮细胞中表达。2 个存在于人前列腺中的 TCR $\gamma$ 转录本的大小分别为 1.1 kb 和 2.8 kb。与在脾脏，胸腺和外周白细胞中发现的 1.5kb 的 TCR $\gamma$ 转录本相比有着大小的不同。通常 TCR $\gamma\delta$  mRNA 的表达图谱表明其在前列腺中的转录与 $\gamma\delta$  T-淋巴细胞的途径不同。前列腺 TCR $\gamma$ 表达最初的发现是通过分析公开 EST 数据库来进行的。我们的结果表明 EST 库是鉴定新的和未知基因表达的强大工具。代表 TCR $\gamma$ 转录本的前列腺 EST 均来自自由激光捕捉显微解剖所获得的细胞建立的 cDNA 文库 (Emmert-Buck 等, Science 274, 998-1001 (1996))。TCR $\gamma$ 转录本被证明来源于前列腺上皮细胞而不是来源于浸润 $\gamma\delta$  T-淋巴细胞的事实证实显

30

微解剖是一种有用的技术来从组织特异性显微区获得纯化的细胞亚群。

#### 前列腺 TCR(JC) $\gamma$ 转录本

5 前列腺腺癌细胞系, LNCaP, 其分离自表达易于检测水平的 1.1kb  
前列腺特异性 TCR $\gamma$ 转录本的淋巴结转移(Horoszewicz 等, *Cancer Res.*  
43,1809-1818 (1983))。在 LNCaP 细胞中表达显示转录本来源于上皮  
细胞并且其可在前列腺恶性肿瘤的发育中表达。 LNCaP 转录本含有  
10 约 53 个碱基的 J $\gamma$ 1.2 基因区段, 3 个 C $\gamma$ 1 外显子,非翻译序列, 然后是  
poly(A)序列。前列腺转录本不同于成熟的 T-淋巴细胞转录本, 因为其  
缺少一个 V $\gamma$ 基因区段并且其起始于紧挨 J $\gamma$ 1.2 上游区的内含子序列中  
(数据没有显示)。启动子启动前列腺 TCR $\gamma$ 转录且其在前列腺上皮  
细胞中的活化机制正在研究之中。2.8kb 的前列腺特异性 TCR $\gamma$ 转录本  
15 在 LNCaP 非常微弱且在 5' RACE PCR 实验中并没有获得任何含有  
2.8kb 转录本的产物, 因此 2.8kb 转录本需要进一步的研究。

#### T 淋巴细胞中 TCR(JC) $\gamma$ 转录本的比较

20 很多的研究表明可能在造血细胞中发生 V (D) J 重排之前或重  
排时能检索到 TCR 基因的转录 (Wang 等, *Mol. Immunol.* 33, 957-964  
(1996); Shimamura, M.,和 Ohta, S., *Eur. J. Immunol.* 25,1541-1546 (1995);  
Villey 等, *Eur. J. Immunol.* 27,1619-1625 (1997); Sikes 等, *J. Immunol.*  
161,1399-1405 (1998))。已有报道 TCR $\gamma$ 基因在带有未重排的 $\gamma$ 座位的鼠  
骨髓 T 淋巴细胞前体细胞中具有转录活性, 产生纯粹的 TCR C $\gamma$ 转录  
25 本(Wang 等, *Mol. Immunol.* 33,957-964 (1996))。此外, 未重排 TCR V $\gamma$   
转录本的表达也在个体发育中有报道(Goldman 等. *J. Exp. Med.* 177,  
729-739 (1993))。TCR 的非翻译转录本和免疫球蛋白基因区段在很大  
程度上被限制于来自淋巴细胞系的细胞中(Lauzurica, P.,和 Krangel, M.  
S., *J.Exp. Med.* 179, 1913-1921 (1994))。此外, 几乎所有 TCR 和免疫  
30 球蛋白的座位的种系转录的激活均暂行性地与座位重组的激活相关  
(Sikes 等, *J. Immunol.* 161,1399-1405 (1998); Goldman 等. *J. Exp. Med.*

177,729-739 (1993); Lauzurica, P., and Krangel, M. S., J. Exp. Med. 179,1913-1921 (1994); Sleckman 等, Annu. Rev. Immunol. 14, 459-481 (1996)。我们已经阐明, 通过基因组 DNA 和 cDNA 的独立性试验, 前列腺上皮细胞并未在 TCR $\gamma$ 座位发生基因重排。因此在前列腺上皮细胞中 TCR(JC) $\gamma$ 转录本的表达并不与重排相关, 并且与 T-淋巴前体细胞中观察到的非翻译转录本相比可能具有不同的功能。

#### TCR $\gamma$ 座位的新的前列腺特异性蛋白的可能性的最初假说

前列腺 TCR $\gamma$ 转录本是高度表达的, 我们设想其必定具有生物学上的重要原因。VJ 基因重排并未发生在前列腺上皮细胞的 TCR $\gamma$ 座位的事实排除了成熟 TCR $\gamma$ 链蛋白产生的可能性。因为在 C $\gamma$ 上游并没有发现翻译起始密码子(ATG), 我们还排除了没有 TCR $\gamma$ 可变区的 TCR $\gamma$ 恒定区蛋白产生的可能性。在 TCR $\gamma$ 链蛋白中, J $\gamma$ 区段编码 16-20 个氨基酸的可变区, 而可变区的绝大部分由 V $\gamma$ 区段编码。除非由 J $\gamma$ 区段编码的氨基酸与由 V $\gamma$ 基因区段编码的氨基酸结合, 它们便不能在 MHC 识别中作为 TCR 发挥作用。这增加了由 C $\gamma$ 编码的新的前列腺特异性蛋白的可能性。我们的最初假设就是, 从原始 TCR $\gamma$ 阅读框架的一个 ATG 密码子处开始翻译, 虽然也可使用其它不同的阅读框架或较少使用的起始密码子。

20

利用前列腺 TCR $\gamma$ cDNA 进行的体外转录偶联翻译试验显示该转录本具有全部的功能。共得到了 2 个蛋白。13KDa 的蛋白最可能起始于附图 1 所示第一个加双下划线 ATG, 其产生一个 12.8KDa 大小的蛋白(PS-TCR $\gamma$ -1)。8KDa 的蛋白最可能来自第二个双下划线 ATG, 计算的蛋白大小约 7.2 kDa (PS-TCR $\gamma$ -2)。在下一个实施例中进一步讨论对这两个蛋白的研究报道。总之, 前列腺上皮细胞, 或非淋巴来源的细胞类型表达高水平的被认为仅由淋巴细胞系所表达的基因的转录本的事实是一个非常出乎意料的发现。

25

#### 30 实施例 2. TCR $\gamma$ 可变阅读框架蛋白的发现

上一个实施例阐述了意外发现了在前列腺和前列腺癌细胞中的 TCR $\gamma$ 转录本，这些转录本的体外翻译，并且最初假这一转录本导致细胞中存在 TCR $\gamma$ 链的截短形式。本实施例阐述另一个意外的发现，即该转录本产生以前未知的由可变阅读框架表达的蛋白，现命名为“TARP”。更出乎意料的是，下面的研究显示 TARP 是一个核蛋白，存在于很多乳房癌细胞中。

#### 材料和方法：

引物 TCR $\gamma$ -upATGmut#1 (5'-TTACAGATAAACAACCTTGATAC  
 10 AGATGTTTCCCCCAAGCCC-3'); TCR $\gamma$ -upATGmut#2 (5'-GGGCTTGGGGGAAAC  
 ATCTGTATCAAGTTGTTTATCTGTAA-3'); TCR $\gamma$ -upATGmut#3 (5'-  
 GATAAACAACCTTGATGCAGATATTTCCCCCAAGCCC-3'); TCR $\gamma$ -upATGmut#4  
 (5'-GGGCTTGGGGGAAATATCTGCATCAAGTTGTTTATC-3'); TCR $\gamma$ -  
 upATGmut#5 (5'-GATAAACAACCTTGATACAGATATTTCCCCCAAGCCC-3');  
 15 TCR $\gamma$ -upATGmut#6 (5'-GGGCTTGGGGGAAATATCTGTATCAAGTTGTTTATC-  
 3'); TCR $\gamma$ -downATGmut#1 (5'-CCCAGGAGGGGAACACCATAAAGACTAAC  
 GACACATAC-3'); TCR $\gamma$ -downATGmut#2 (5'-GTATGTGTCGTTAGTCTTT  
 ATGGTGTTCCCCTCCTGGG-3'); TCR5.1 (5'-GATAAACAACCTTGATGCA  
 GATGTTTCC-3'); TCR3.1 (5'-TTATGATTTCTCTCCATTGCAGCAG-3');  
 TCR $\gamma$ 1.2R (5'-AAGCTTTGTTCCGGGACCAAATAC); B-肌动蛋白正向引物 (5'-  
 20 ATCTGGCACCACACCTTCTACAATGAGCTGCG-3'); B-肌动蛋白反向引物 (5'-  
 CTCATACTCCTGCTTGCTGATCCACATCTGC-3'). 通过 Sigma-  
 Genosys (The Woodlands, TX)和 Lofstrand Labs Limited (Gaithersbur  
 g, MD)合成引物。

#### 25 构建体

如前面所述进行构建被克隆到 pBluescript II SK (+) (Stratagene, La Jolla, CA)中的 TARP 转录本(Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。在该过程中，该质粒被称为 pBSSK-TCR $\gamma$ 。核苷酸位置 69 处的 ATG 被变异为 ATA 的 pBSSK-TCR $\gamma$ mutATGup1，  
 30 通过 Quickchange Site-Directed Mutagenesis 试剂盒(Stratagene)构建。

利用 TCR $\gamma$ -upATGmut#1 和 TCR $\gamma$ -upATGmut#2 作为引物并以 pBSSK-TCR $\gamma$  为模板进行 PCR 反应。在核苷酸位置 73 处的 ATG 突变为 ATA 的 pBSSK-TCR $\gamma$ mutATGup2, 如上所述采用 TCR $\gamma$ -upATGmut#3 和 TCR $\gamma$ -upATGmut#4 为引物并以 pBSSK-TCR $\gamma$ 为模板来构建。在核苷酸位置 69 和 73 处的两个 ATG 都突变为 ATA 的 pBSSK-TCR $\gamma$ mutATGup-both, 如上所述通过用 TCR $\gamma$ upATGmut#5 和 TCR $\gamma$ -upATGmut#6 作为引物并以 pBSSK-TCR $\gamma$ mutATGup1 为模板来构建。在位置 242 处的 ATG 突变为 ATA 的 pBSSK-TCR $\gamma$ mutATGdown, 如上所述通过用 TCR $\gamma$ -downATGmut#1 和 TCR $\gamma$ -downATGmut#2 作为引物并以 pBSSK-TCR $\gamma$  为模板来构建。pET-TCR $\gamma$ 含有被亚克隆到 pET23a 载体(Novagen, Madison, WI)中的 TARP 转录本 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999)) 的 242-469 位核苷酸。pET-TARP 含有被亚克隆到 pET23a 载体中的 TARP 转录本 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999)) 的 56-242 位核苷酸。pVC4D-TARP 含有被亚克隆到 pVC4D 载体(Bruggemann, E. P. 等, BioTechniques 10: 202-209 (1991))的 TARP 转录本 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))中的 69-242 位核苷酸。

## 20 逆转录聚合酶链式反应 (RT-PCR)

根据生产商的操作说明利用 Micro-FastTrack™ 2.0 试剂盒 (Invitrogen, Carlsbad, CA)进行分离 poly (A) RNA。500 ng 的 poly (A) RNA 或 5 $\mu$ g 的总 RNA 在 50 pmol oligo-dT 引物(Invitrogen) 存在的条件下 70 $^{\circ}$ C 变性 2 分钟。在 10 $\mu$ l 含有 250 uM dNTPs, 2 mM DTT, 8 U RNasin (Roche Molecular Biochemicals, Indianapolis, IN), 50 U Superscript II™ RT (Life Technologies, Rockville, MD)的反应混合物中制备单链 cDNA 并 42 $^{\circ}$ C 温育 90 分钟。样本然后用 75 $\mu$ l 10mM 的 Tris-HCl [pH 7.5]稀释并 72 $^{\circ}$ C 温育 10 分钟。3 $\mu$ l 的 cDNA 用于含有 250 $\mu$ M dNTPs、25 pmol 的各种引物、1 单位的 AmpliTaq® DNA 聚合酶(Roche)的 PCR 并扩增 35 个循环。相似的 PCR 条件用于人乳房 RAPID-SCAN™

基因表达部分(OriGene Technologies, Rockville, MD)。引物 TCR $\gamma$ J1.2R, TCR5.1 和 TCR3.1 被用于检测 TARP 转录本, 而正向引物 B-肌动蛋白和反向 B-肌动蛋白引物被用于肌动蛋白转录本的检测。

## 5 Northern 印迹杂交

如前所述利用 2 $\mu$ g 的 poly(A)RNA 进行 Northern 印迹杂交 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。

## 体外转录偶联的翻译

10 在体外偶联的翻译反应已被在前面描述 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。pBSSK-TCR $\gamma$ , pBSSK-TCR $\gamma$ mutATGdown, pBSSK-TCR $\gamma$ mutATGup1, pBSSK-TCR $\gamma$ mutATGup2 和 pBSSK-TCR $\gamma$ mutATGup-both 被用作模板。

## 15 细胞培养

LNCaP, PC3, MCF7, BT-474 和 SK-BR-3 细胞于 37 $^{\circ}$ C 在 5%CO<sub>2</sub> 条件下培养在 RPMI-1640 培养基(Quality Biological, Inc., Gaithersburg, MD)中。培养基含 10%的胎牛血清 (FBS, Quality Biological, Inc.), 2 mM L-谷氨酰胺, 1 mM 丙酮酸钠和青霉素/链霉素。Hs57Bst 细胞于 37 $^{\circ}$ C  
20 用 5% CO<sub>2</sub> 温孵培养在 RPMI-1640 培养基中。该培养基含有 10% FBS, 30 ng/ml 表皮生长因子(EGF, Harlan, Cincinnati, OH), 2 mM L-谷氨酰胺, 1 mM 丙酮酸钠和青霉素/链霉素。

## 抗体产物.

25 多克隆 $\Delta$ PE-TARP 抗体的制备如下。含有与假单胞菌外毒素( $\Delta$ PE) (Bruggemann, E. P. 等, BioTechniques 10: 202-209 (1991))的催化性失活形式的 C'末端融合的完整 TARP 可读框的 pVC4D-TARP, 在 Epicurian Coli $\text{\textcircled{R}}$  BL21-CodonPlus<sup>TM</sup> (DE3)-RIL 细胞(Stratagene)中表达。包涵体的制备以及兔免疫如前面所述(Brinkmann, U. 等, Proc. Natl. Acad. Sci.  
30 USA 88: 8616-8620 (1991))。依据制造商(Pierce, Rockford, IL)的说明用

ImmunoPure® IgG (Protein A)纯化试剂盒纯化抗血清。

5 利用 pET-TCR $\gamma$ 如前面所述的进行制备 TCR $\gamma$  抗体, pET-TCR $\gamma$ 是一个含有与 C'末端 6 个组氨酸标记融合的 TCR $\gamma$ 细胞外结构域的表达质粒。在免疫前,利用 Ni-NTA 琼脂糖柱按照制造商 (QIAGEN, Valencin, CA) 的说明纯化该组氨酸标记的 TCR $\gamma$ 蛋白。

#### 细胞提取物的制备

10 如下制备全细胞蛋白提取物。收获来自各细胞系的  $5 \times 10^6$  个培养的细胞并将其悬浮在含有蛋白酶抑制剂的  $1 \times$  RIPA 缓冲液 (50mM Tris-HCl[pH7.5], 150mM NaCl, 1mM EDTA, 0.1% Triton X-100, 1mM PMSF, 1 $\mu$ g/ml aprotinin, 1 $\mu$ g/ml 亮抑酶) 中。提取物直接用超声破碎, 并进行离心分离。根据制造商 (Pierce) 的说明利用 Coomassie® Plus 蛋白试验试剂来测定蛋白浓度。通过用冷的研钵和研棒将-80 $^{\circ}$ C 冰冻的  
15 0.5 g 前列腺癌组织研磨成粉末来制备前列腺组织的蛋白提取物。收集粉末状组织, 悬浮于  $1 \times$  RIPA 中如上所述进行处理。

前列腺和乳房细胞中的核、膜以及细胞质提取物的制备如目前已公开的方法进行 (Dignam, J. D. 等, Nucleic Acids Res. 11: 1475-1489  
20 (1983); Sladek, F. M. 等, Genes Dev. 4: 2353-2365 (1990)).

#### Western 印迹分析.

25 20 或 40 $\mu$ g 的蛋白提取物, 1 $\mu$ g 的重组 His-TARP 或 100 ng 的重组 His-TCR $\gamma$  在 16.5% Tris Tricine 凝胶上电泳(BIO-RAD, Hercules, CA)然后在转移缓冲液 (25 mM Tris, 192 mM 甘氨酸, 20% (v/v) 甲醇, pH 8.3) 4 $^{\circ}$ C 和 30 V 下 4 小时转移到 0.2 $\mu$ m Immun-Blot™ PVDF 膜 (BIO-RAD)上。用 10 $\mu$ g/ml  $\Delta$ PE-TARP 抗血清或 1 $\mu$ g/ml TCR $\gamma$  抗血清来对滤膜进行探测, 依据制造商 (Roche) 的说明, 用化学荧光 western 印迹试剂盒检测各自的信号。

30

TARP 是在前列腺癌细胞中表达的核蛋白。

5 为了证实 TARP 或 TCR $\gamma$ 是否存在于前列腺癌细胞中，我们制备抗两种蛋白的抗体并在不同的前列腺癌细胞提取物中进行 western 印迹反应。如附图 3A（上图所示），在前列腺癌 LNCaP 细胞系和一个前列腺癌肿瘤提取物中检测到 TARP。与重组物 His-TARP 共迁移的 7kDa 大小的片段表明在 LNCaP 和癌提取物中检测到的产物是 TARP。以前，我们证实了前列腺特异性 TCR $\gamma$ 转录本在前列腺癌细胞系 PC3 中并不表达(Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。因此我们使用 PC3 细胞提取物作为阴性对照从而证实  
10 7kDa 片段在这些提取物中并不存在（附图 3A, 上图）。重要的是，当利用取血前抗血清和抗假单胞菌外毒素 (PE, 参见材料和方法)的抗血清时，没有检测到 7kDa 片段（数据未显示）。在这些提取物中没有检测到 TCR $\gamma$ ，虽然重组蛋白与所用抗体表现很强的信号（附图 3A, 下图）。这些结果表明前列腺特异性 TCR 转录本编码 TARP。

15

为了鉴定 TARP 在细胞中的位置，我们从 LNCaP 细胞中制备核、细胞质和膜部分。如附图 3B 所示，在核部分而非细胞质或膜中检测到 TARP。通过蔗糖垫进行的细胞提取物分级分离纯化的核中我们得到了相似的结果 (Sladek, F. M. 等, Genes Dev. 4: 2353-2365 (1990))  
20 (数据没有显示)。

TARP 转录本在乳房细胞中表达。

我们在以前已有报道 TCR $\gamma$ EST 库也包括一些来自脑文库的 EST (Vasmatzis, G. 等, Proc. Natl. Acad. Sci. USA 95: 300-304 (1998))。在这些最初的报道之后，其它的 EST 亦被存放到这个数据库中，目前这个库包括来自乳房，结肠，肾以及胃的文库的 EST。为了鉴定这些 EST  
25 的出现是否说明 TARP 转录本在这些细胞中就表达或者当这些文库建成时是否就归因于浸润 $\gamma$ BT-淋巴细胞的存在，我们对不同的细胞系进行 RT-PCR 来检测 TARP 转录本的存在。如附图 4A 所示，在乳房细胞系 MCF7, BT-474, SK-BR-3 和 CRL-1897 中检测到 TARP 的表达。  
30

在成纤维细胞瘤细胞系 A172, 成胶质细胞瘤细胞系 IMR32, 结肠细胞系 COLO 205, 胃细胞系 KATO III 或肾细胞系 COS7 和 293 中没有检测到信号 (附图 4A, 数据未显示)。为了鉴定 TARP 转录本是否在除上述细胞系以外的人乳房组织中表达, 我们利用 RAPID-SCAN™ (OriGeneTechnologies, Rockville, MD) 检测了 12 个不同的乳房癌细胞的 cDNA。显示出 TARP mRNA 在一些乳房样品中是大量存在的 (附图 4B, 上图), 而正常乳房中的 TARP mRNA 在 35 个 PCR 循环后只达到可检测的水平 (结果未显示)。值得注意的是, 在缺少 cDNA 的反应中未检测到信号。用于显示相同 cDNA 量的肌动蛋白存在于每一个泳道 (下图)。在正常乳房样品中的微弱信号与附图 4A 和 5 的 Hs57Bst 细胞系所显示的 TARP 信号的缺乏相关, Hs57Bst 细胞系来自正常的乳房组织。这些结果表明乳房发生致癌性转化后 TARP 转录本的表达提高。但是在做出任何最后的结论前必须进行更多的研究。

为了确定在乳房细胞系中观察到的 TARP 转录本与前列腺细胞系中发现的转录本是否相同, 我们利用抗不同 TARP 转录本区域的引物进行 PT-PCR。如附图 5A 所示, 前列腺中的 TARP 转录本含有 Jy1.2 基因区段部分, 3 个 Cyl 外显子以及跟随一个 poly(A)尾(7)的一些非翻译序列。引物组 1 和 3 扩增了全部 TARP 转录本 (附图 5B, 上图) 而引物组 2 和 3 仅扩增了 Cyl 区域 (附图 5B, 中间部分)。如附图 5B 所示, 与利用任一个引物组的前列腺细胞系(LNCaP) 相比在三个乳房细胞系(MCR7, BT-474 和 SK-BR-3)中检测到相似大小的条带。重要地, 在缺少 cDNA (dH<sub>2</sub>O) 的反应中没有检测到信号, 且使用了肌动蛋白对照所示的相同量的 cDNA (附图 5B, 下图)。这些数据表明在乳房细胞系中发现的 TARP 转录本与在前列腺细胞系中发现的转录本相同。为了进一步证明这个结论, 我们用 Northern 印迹分析了来自每个细胞系的 TARP 转录本的大小。如上所述, 1100 和 2800 个核苷酸转录本存在于 LNCaP 细胞系中, 其中 1100 个核苷酸的转录本主要的形式 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。如附图 5C 所示, 与前列腺细胞系 LNCaP 相比, 在 3 个乳房

细胞系 (MCF7, BT-474 和 SK-BR-3)中都发现了相似大小的 TARP 转录本, 尽管强度较弱。因此我们得出结论 TARP 在前列腺和乳房癌细胞中表达。

5            为了确定 TARP 蛋白是否存在于乳房癌细胞系中, 我们用抗 TARP 的抗体和乳房癌细胞核提取物进行 Western 印迹。如附图 6 (上图) 所示, 在 MCF7, BT-474 和 SK-BR-3 细胞中检测到 TARP 反应条带。在这些乳房癌细胞学中的膜或细胞质部分没有检测到 TARP (数据未列出)。重要地, TARP 是乳房细胞系的 TARP 转录本所编码的蛋白产物的一部分, 因为在任一这些核提取物中都未检测到 TCR $\gamma$ , 即使重组蛋白与所用抗体表现出很强的信号 (附图 6, 下图)。这些数据表明 TARP 也存在于乳房癌细胞中。

10

我们报道了鉴定到一个来自 TCR $\gamma$ 座位由特异性转录本编码的在前列腺和乳房癌细胞中表达的 7KDa 的核蛋白。因为该蛋白由不同于 TCR $\gamma$ 的读框编码, 我们把 TCR $\gamma$ 可变阅读框架蛋白称之为 TARP。除了从起始于 TCR $\gamma$ 一个内含子内的位点可变阅读框架被翻译外, TARP 还有其它的两个特征。首先, 令人意外的发现这样的小肽存在于细胞中, 因为大多数是分泌型的。其次, TARP 缺少一个较好的 Kozak 序列(Kozak, M. Cell 44: 283-92 (1986))。由于 TCR $\gamma$ 阅读框架含有好的 Kozak 序列, 我们最初设想的截短的 TCR $\gamma$ 蛋白被编码。但是, 如附图 3 所示, 我们最初的设想是错误的, 有趣的是体外翻译结果表明 TARP 蛋白的偏好和 TARP 阅读框架中的两个 ATG 之一可被用于起始蛋白的合成。蛋白应被测序来鉴定哪个 ATG 被用于启动 TARP 蛋白的合成。

15

20

25

TARP 蛋白序列的有趣之处在于七体重复序列中有五个亮氨酸, 这表明 TARP 可能含有一个二聚化亮氨酸的拉链基元 (附图 7A。) 如果是这样, TARP 必须含有一个两亲螺旋。其一暗示 TARP 可能含有一个两亲螺旋, 据信其中丝氨酸和脯氨酸残基作为螺旋的起点, 这

30

些残基被发现恰好在第一个亮氨酸重复之前。第二，很多带电的氨基酸被发现存在于七体重复中，即为螺旋提供了两亲特征并可能充当与其它螺旋盐桥。即使亮氨酸存在于七体重复中是亮氨酸拉链基元的好指示，但是也存在被鉴定的在七体重复中含有 5 个亮氨酸的蛋白其并非亮氨酸拉链蛋白的情况。例如 karyopherin 的晶体结构 (Chook, Y. M. 等, Nature 399: 230-237 (1999)), 嗜热脂肪芽孢杆菌 (*B. stearothermophilus*) 嘧啶核苷磷酸化酶 (Pugmire, M. J. 等, Structure 6: 1467-1479 (1998)) 以及嗜热柄热菌 (*T. Thermophilus*) 苯丙氨酰-tRNA 合成酶 (Mosyak, L. 等, Nat. Struct. Biol. 2: 537-547 (1995)) 表明这些蛋白在七体重复中含有 5 个亮氨酸的区域并不含有螺旋结构。需要研究相互作用和结构来探索 TARP 中发现的亮氨酸重复的意义。

TARP 氨基酸序列的另一个独特的特征是碱性氨基酸区连接了一个可能的亮氨酸拉链基元 (图 7A)，此表明可能是一个 DNA 结合元。但是碱性区的定位是独特的，其在亮氨酸重复的后面而不是在前面。大多数结合 DNA 的亮氨酸拉链蛋白在亮氨酸重复的前面具有碱性区域 (参见(Chook, Y. M. 等, Nature 399: 230-237 (1999)))。TARP 的碱性区域可能仅仅行使核定位信号的功能，但是 TARP 是一种核蛋白的事实加强了 TARP 可能结合 DNA 的假说。在作出最后的结论之前有必要进行功能的研究。

为了测定 TARP 与任何已知的蛋白是否具有同源性，我们在 GenBank 中进行了蛋白质的 BLAST 检索。检索的结果显示 TARP 的氨基酸序列与 *Dictyostelium discoideum* Tup1 (GenBank 入藏号 AAC29438) 和 *Saccharomyces cerevisiae* Tup1 (Williams, F. E. 等, Mol. Cell. Biol. 10: 6500-6511 (1990)) 具有一些同源性 (附图 7C)。酵母 Tup1 通常被发现与 Cyc8 (Ssn6) 结合且由葡萄糖、氧和 DNA 损伤调控基因的转录表达所需要的 (Tzamarias, D. 等, Genes Dev. 9: 821-831 (1995))。既不是 Cyc8 (Ssn6) 也不是 Tup1 结合 DNA，而是通过与特异性结合 DNA 的蛋白例如 a2, Mig1, Rox1 等蛋白相互作用，各自充当

共表达复合物的一部分 (Tzamarias, D. 等, *Genes Dev.* 9: 821-831 (1995))。Tup1 的 C'末端部分含有六个富含门冬氨酸盐和色氨酸的 43 氨基酸序列的重复, 已知为 WD-40 或 $\beta$ -转导素重复 (Williams, F.E. 等, *Mol. Cell. Biol.* 10: 6500-6511 (1990); Fong, H. K. 等, *Proc. Natl Acad. Sci. USA* 83: 2162-2166 (1986))。WD-40 重复已在很多的蛋白中被鉴定出, 且在蛋白与蛋白之间的相互作用中起作用。重要的是, Tup1 已被显示通过其两个 WD-40 重复与 a2 相互作用(Komachi, K. 等, *Genes Dev.* 8: 2857-2867 (1994))。有趣的是, TARP 与 Tup1 的第 5 个 WD-40 重复有同源性 (附图 7C)。因为 TARP 是一种核蛋白, 其与 Tup1 的同源性表明 TARP 可能是一个与转录调控有关的功能性核蛋白复合物。因此, 鉴定 TARP 相互作用蛋白来检测其功能是必要的。

TARP 抗体识别前列腺和乳房细胞核提取物中的双联体 (附图 6A)。最快的 7kDa 条带与 His-TARP 重组蛋白共迁移, 而最弱的条带为分子量较大的分子。一种可能的解释是 9kDa 的条带是翻译后修饰的。为了检测 TARP 是否含有任何已知的翻译后的修饰位点, 我们利用瑞士生物信息学研究所 ExPASy 蛋白组学服务器 (Swiss Institute of Bioinformatics ExPASy proteomics server) (<http://www.expasy.ch>) 的 PROSITE 分析程序来分析 TARP 氨基酸序列 (Appel, R. D. 等, *Trends Biochem. Sci.* 19: 248-260 (1994); Hofmann, K. 等, *Nucleic Acids Res.* 27: 215-219 (1999))。如附图 7A 所示, 发现很多可能的磷酸化作用位点包括 cAMP 和 cGMP-依赖的蛋白激酶磷酸化位点 (RRAT 和 RRGTT) 以及蛋白激酶 C 磷酸化位点 (SSR 和 SRR)。磷酸化作用被显示在很多的情况下引起蛋白质在 SDS PAGE 凝胶上以较大分子量运行。如果这种情况属实, 附图 6 的结果表明未修饰的形式在 LNCaP 细胞中是普遍的且磷酸化的形式仅存在于 MCF7 和 SK-BR-3 细胞中。需要进行另一个试验来检测 9kDa 条带的特性以及当在前列腺和乳房癌细胞中表达时 TARP 是否是翻译后修饰的。

我们在此报道了 TARP mRNA 和蛋白在乳房癌细胞中的表达。我

们最初的对 TARP 转录本的研究并没有揭示 TARP 在乳房中的表达 (Essand, M. 等, Proc. Natl. Acad. Sci. USA 96: 9287-9292 (1999))。一种可能的解释是 TARP 在正常的乳房中低水平表达且很难去检测。如结果部分所述的, 与癌样本中检测到的强烈信号相比正常乳房样本的 PCR 分析检测到非常弱的信号。因此, TARP 在乳房癌细胞中的存在说明 TARP 的表达是在乳房细胞被致癌转化后被诱导的。此外, TARP 在乳房癌细胞中的存在可以表明 TARP 由雌激素调控。这种假说通过鉴定结合雄激素应答元件 (ARE) 和雌激素应答元件 (ERE) 的 TARP 的内含启动子的元件得以加强。此种杂合的元件由两个半位点组成, 它们对 ARE 在 5' 末端特异且对 ERE 在 3' 末端特异 [ (Zilliacus, J. 等, Mol.Endocrinol. 9: 389-400 (1995)) 以及未公开的数据]]。需要进行另一个实验, 来检测是否雌激素调控 TARP。然则, 有一些例子, 其中的突变体 ARE 引起某些前列腺特异的基因在乳房肿瘤中的表达。例如, 前列腺特异性抗原(PSA) 已显示在乳房肿瘤中表达 (Majumdar, S. 等, Br. J. Cancer 79: 1594-1602 (1999))。PSA 异常表达的分子分析导致发现了在 PSA 启动子中发现的 ARE 中的单一位点突变。突变导致雌激素调控的 PSA 在乳房瘤中的表达的丢失是可信的(Majumdar, S. 等, Br. J. Cancer 79: 1594-1602 (1999))。此时 TARP 启动子中的类似突变是否发生在 3 个被检测的乳房细胞系中仍不得而知。

20

前列腺依赖雄激素来维持其结构和功能。当前列腺细胞变为恶性细胞时, 其经常失去了雄激素依赖。在此研究中, 我们使用两种对雄激素具有不同依赖性的前列腺细胞系: LNCaP 和 PC3 细胞。雄激素受体存在于雄激素非依赖的 LNCaP 细胞系中, 但不存在于雄激素依赖性的 PC3 细胞系中 (Tilley, W. D. 等, Cancer Res. 50: 5382-5386 (1990))。如附图 3 所示, TARP 在 LNCaP 细胞而不在 PC3 细胞中表达。此结果表明 TARP 表达可通过雄激素刺激来调控。TARP 启动子中的 ARE 类元件的鉴定加强了 TARP 被雄激素诱导的观点。确定雄激素是否诱导 TARP mRNA 表达的实验正在进行中。在 LNCaP 细胞而不在 PC3 细胞中表达表明 TARP 在调控雄激素依赖的应答中是重

30

要的。

5 本发明提供了新的涉及前列腺细胞，前列腺癌，以及乳房细胞和乳房癌的材料和方法。特异性的样本已被提供，上面所述的是例证且非限制的。在本发明说明书基础上的许多变化对本领域的技术人员来说是显而易见的。本发明的保护范围，因此，不应局限于上面的具体描述而应参照后附的权利要求书及其等同物的全部含义来确定。

10 本申请中所引证的所有出版物和专利文献以其全文引入本文做参考，就各种目的而言，其引用程度就如同每一出版物或专利文献被单个地整个引入一样。在本专利申请中所引入的任何文献不应看成是本发明的现有技术。

## 来自TCR Y座位的前列腺特异性转录本

```

GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
<----- JY 1.2 ----->
      M Q M F P P S P L F F F L Q L L K Q S S R R      20
GATAACAACCTTGATGCAGATGTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAAACAAAGCTCCAGAAGGC 140
<----- CY 1(外显子CI) ----->
      L E H T F V F L R N F S L M L L R Y I G K K R R A T R      40
TGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGA 220
-----
      F W D P R R G T P      58
      M K T N D T Y M K F S W L T V P E K
TTCTGGGATCCCAGGAGGGGAACACCAATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
-----
      S L D K E H R C I V R H E N N K N G V D Q E I I F P P      40
TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAACGGAGTTGATCAAGAAATTATCTTCTCTCC 380
-----
      I K T D V I T M D P K D N C S K D A N D T L L L Q L      60
AATAAAGACGGATGTCATCACAATGGATCCCAAAGACAATTGTTCAAAGATGCAAATGATACACTACTGCTGCAGCTCA 460
-----><----- CY 1(外显子CII) ----->
      T N T S A Y Y M Y L L L L L K S V V Y F A I I T C C L      80
CAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTGCCATCATCACCTGTGTCTG 540
----- CY 1(外显子CIII) -----
      L R R T A F C C N G E K S      111
CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCATCGG 620
----->
TTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGCTTCTTTCTGGGTTTGGGCCATTTTCAGTTCTCATGTGTGT 700
ACTATTCTATCATTATTGTATAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780
AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCCAAATAGCGCTGTATAGTG 860
TAGACATCCTGCGGCTTCTAGCCTTGCCCTCTCTTAGTGTCTTTAATCAGATAACTGCTGGAAGCCTTTTCATTTTAC 940
ACGCCCTGAAGCAGTCTTCTTTGCTAGTTGAATTATGTGGTGTGTTTTCCGTAATAAGCAAAATAAATTTAAAAAATG 1020
AAAAGTT 1027

```

加下划线的序列为:

- 转录起始位点(在GCAAGAG序列内)
- 多聚腺苷酸化信号(AATAAA)

双下划线的序列为:

- 可能的转录起始密码子(ATG)

图1

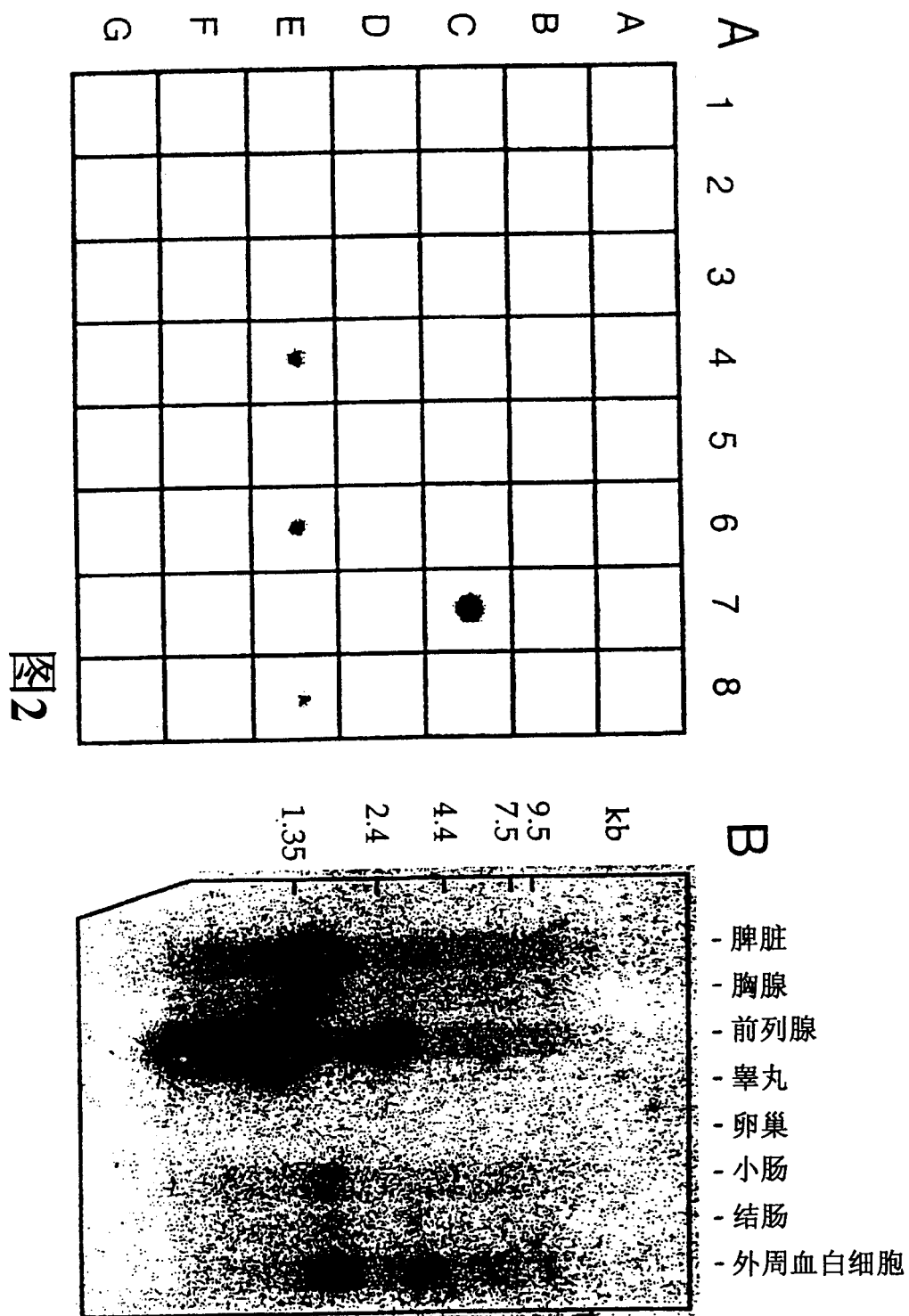


图2

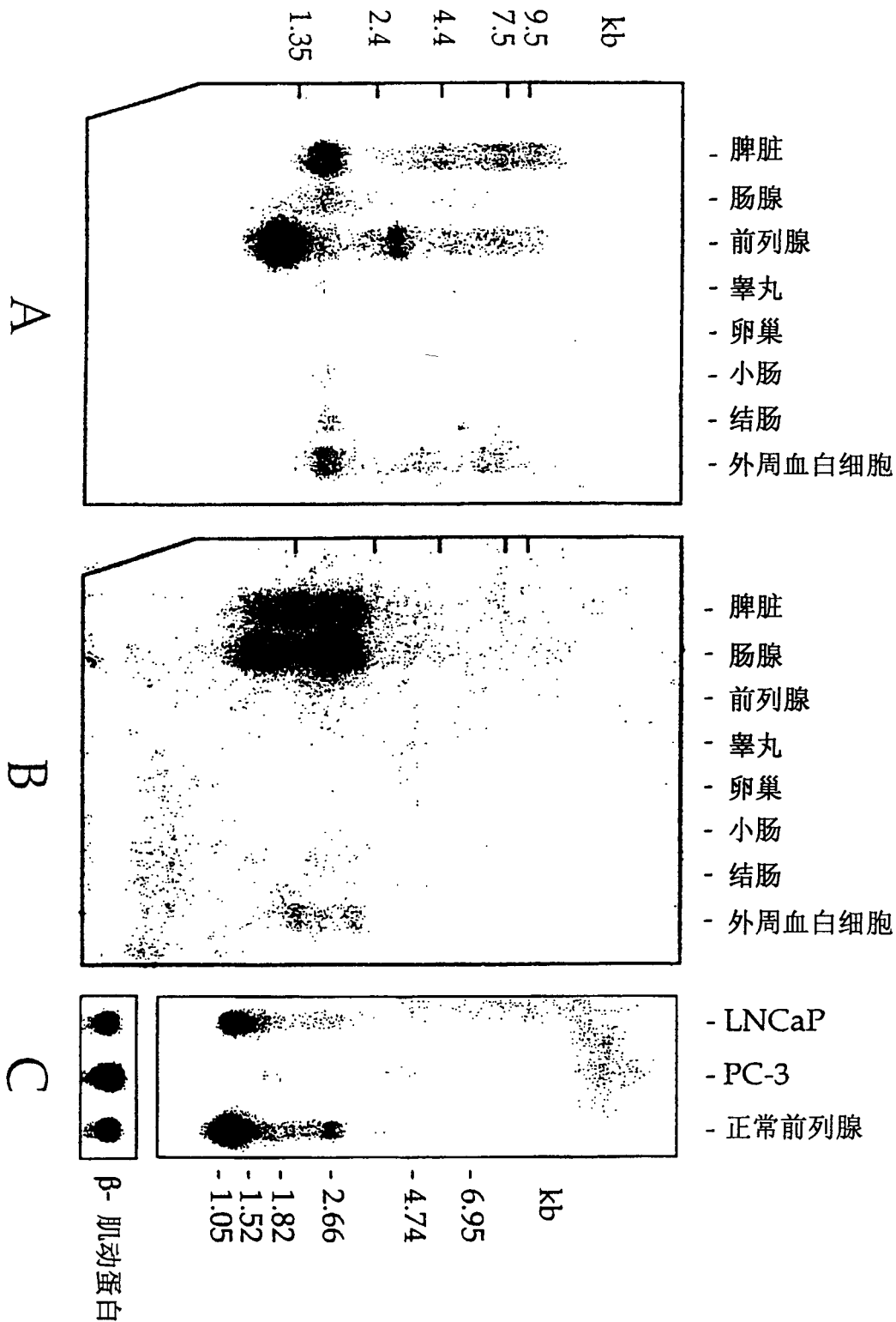


图3

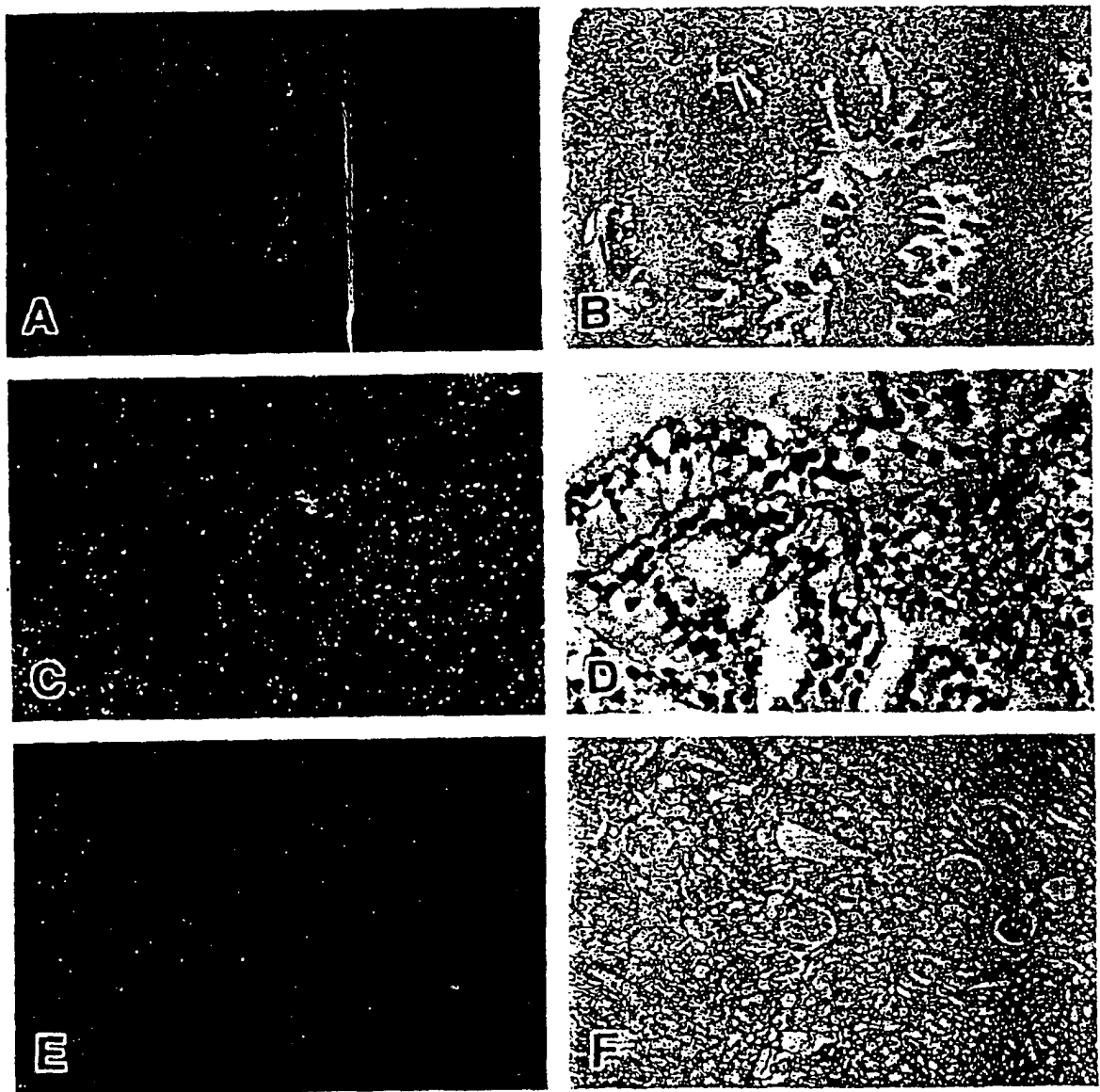


图4

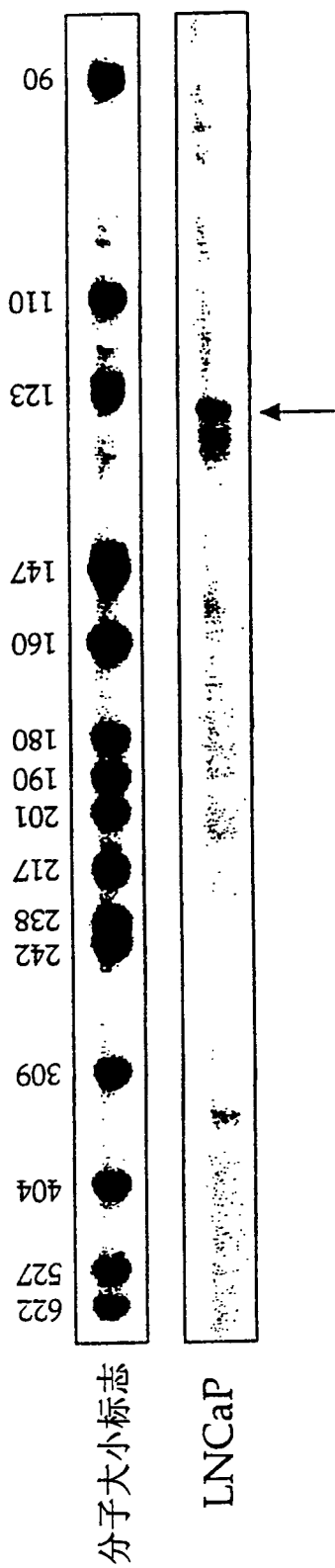


图5

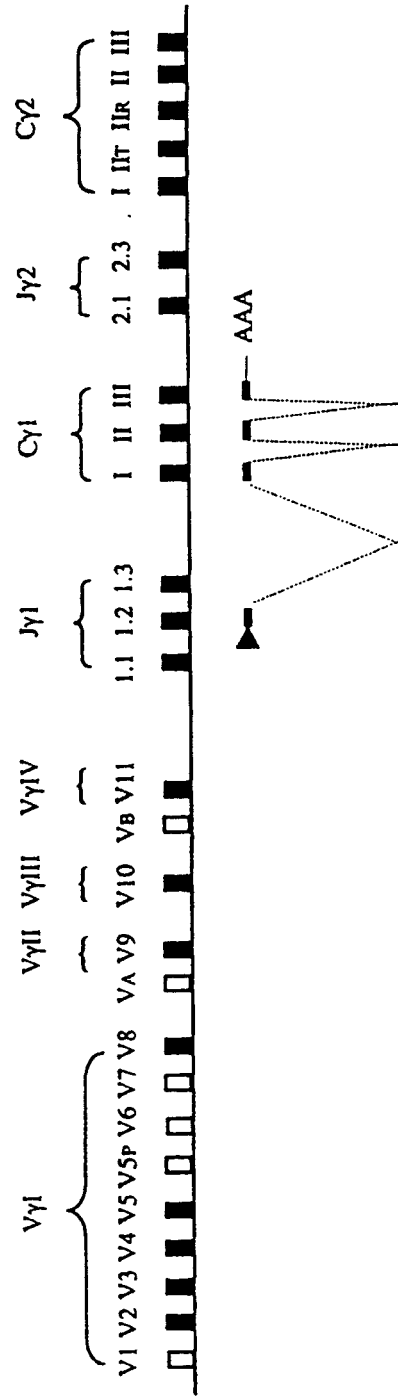


图6

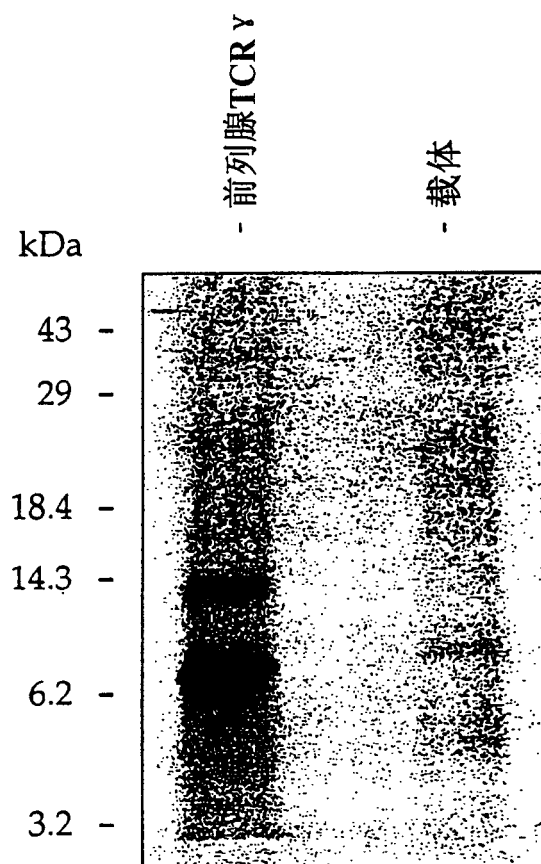
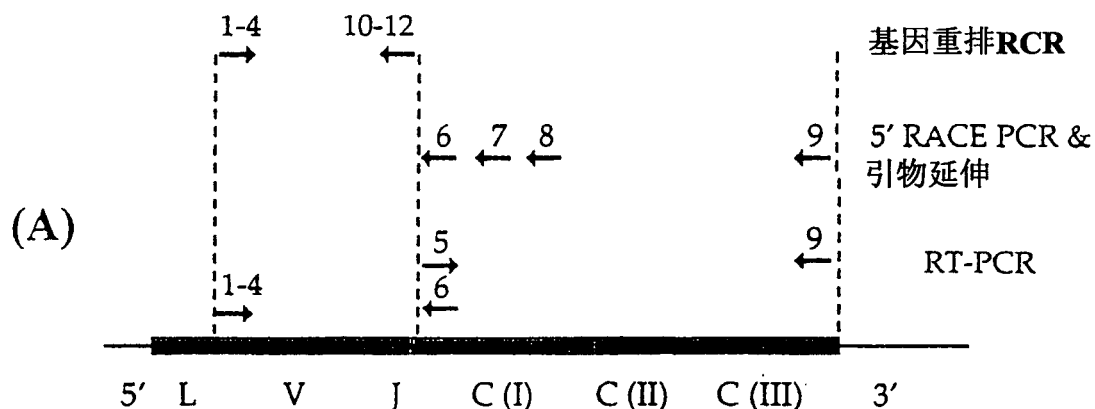


图7

表1: 用于分析前列腺TCR  $\gamma$  转录本的引物(→)

(B)

名称	退火	引物序列 5' -> 3'
1. TCRV $\gamma$ I.F	V $\gamma$ , 亚组I	AACTTGGAAGGGRGAACRAAGTCAGTC
2. TCRV $\gamma$ II.F	V $\gamma$ , 亚组II	AGTACTAAAACGCTGTCAAAAACAGCC
3. TCRV $\gamma$ III.F	V $\gamma$ , 亚组III	TTGGACTTGGATTATCAAAAGTGG
4. TCRV $\gamma$ IV.F	V $\gamma$ , 亚组IV	TTGGGCAGTTGGAACAACCTGAAA
5. TCRC $\gamma$ .F	C $\gamma$ , 外显子CI	GATAACAACCTTGATGCAGATGTTTCCC
6. TCRC $\gamma$ .R1	C $\gamma$ , 外显子CI	GGGAAACATCTGCATCAAGTTGTTTATC
7. TCRC $\gamma$ .R2	C $\gamma$ , 外显子CI	CTGGAGCTTTGTTTCAGCAATTGAAGG
8. TCRC $\gamma$ .R3	C $\gamma$ , 外显子CI	CTCAAGAAGACAAAGGTATGTTCCAGC
9. TCRC $\gamma$ .R4	C $\gamma$ , 外显子CIII	TTATGATTTCTCTCCATTGCAGCAG
10. TCRJ $\gamma$ 1.1.R	J $\gamma$ 1.1	GAAGTTACTATGAGCTTAGTCCCCTT
11. TCRJ $\gamma$ 1.2.R	J $\gamma$ 1.2	AAGCTTTGTTCCGGGACCAAATAC
12. TCRJ $\gamma$ 1.3.R	J $\gamma$ 1.3	TACCTGTGACAACAAGTGTGTTC

R=A+G

图8

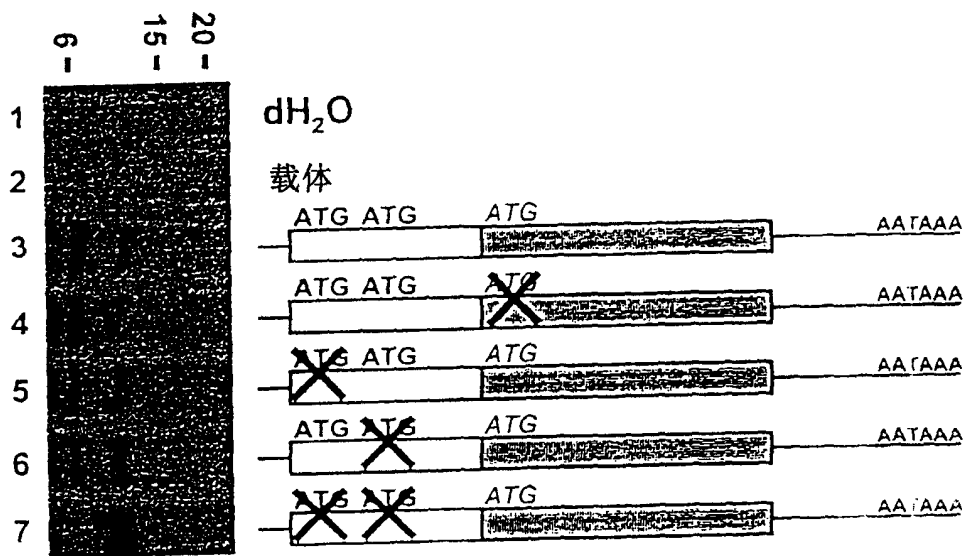


图9

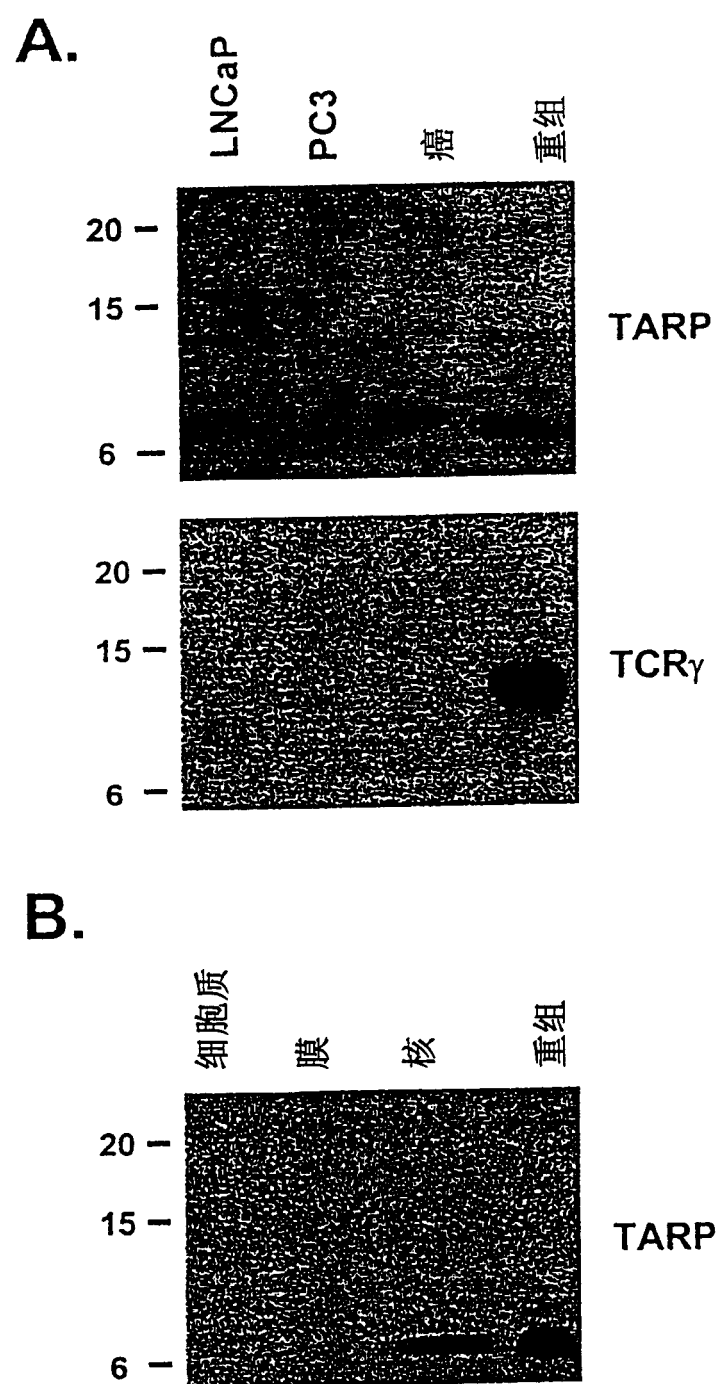


图10

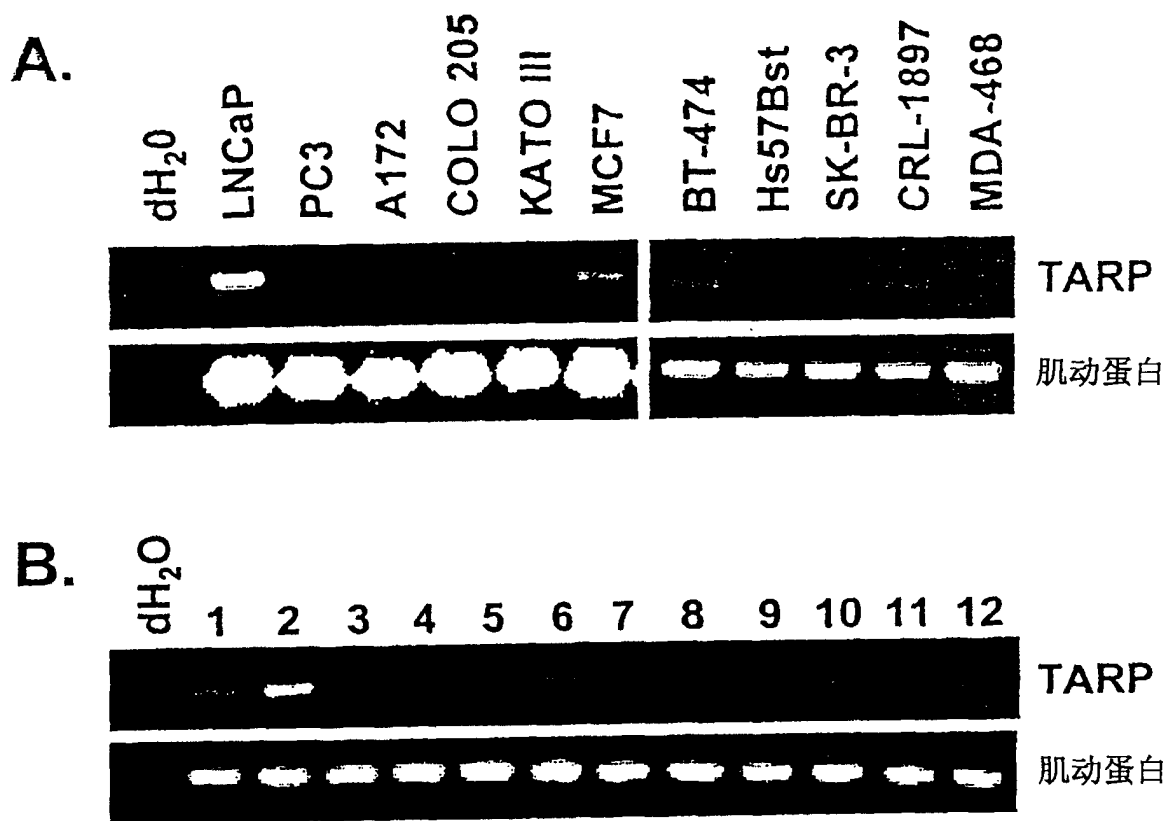


图11

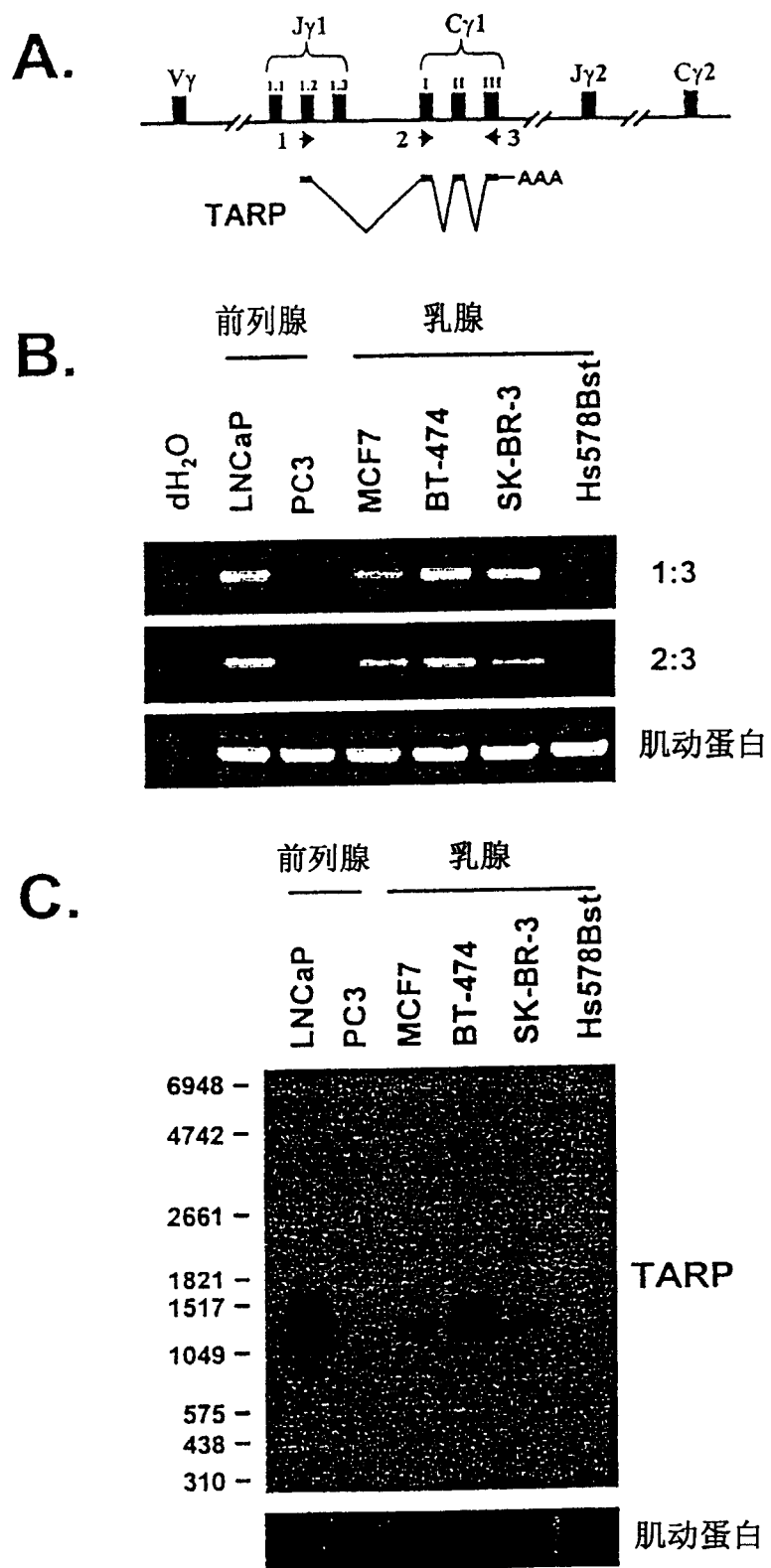


图12

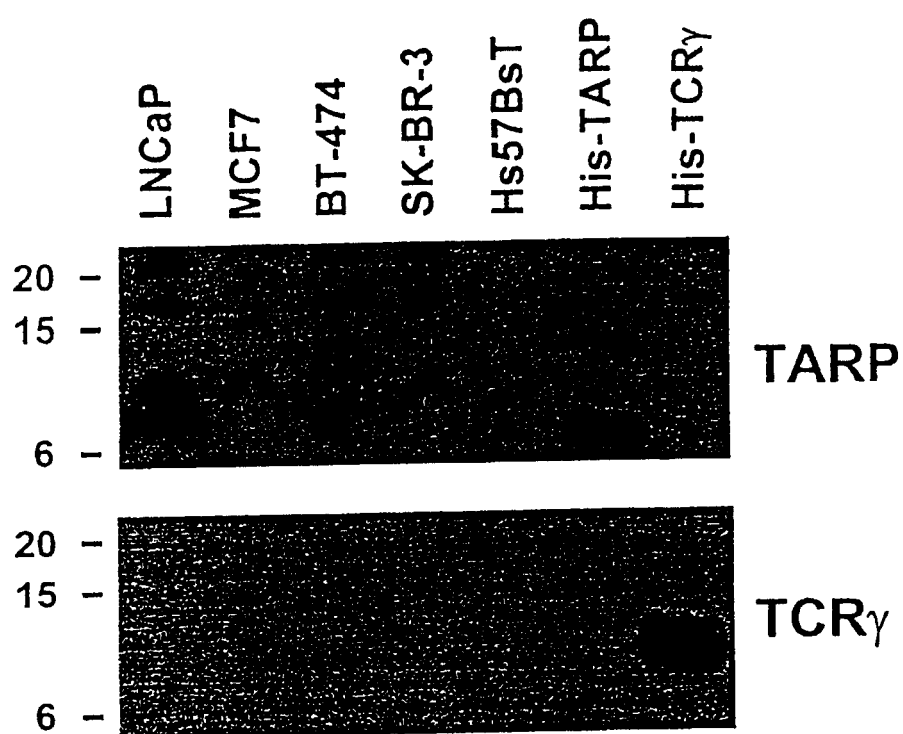


图13

A.

MQMFPSP□FFFLQL□KQSSRR  
 □EHTFVF□RNFSLM□LRYIGKKR  
RATRFWDPPRRGTP

B.

TARP	G	K	R	R	A	T	R	F	W	D	P	R	R	G	T
DTUP1	G	S	K	R	S	V	Q	F	W	D	P	R	N	G	T
YTUP1	G	S	K	R	G	V	L	F	W	D	K	K	S	G	N

图14

专利名称(译)	T - 细胞受体γ可变阅读框架蛋白(TARP)及其应用		
公开(公告)号	<a href="#">CN1378594A</a>	公开(公告)日	2002-11-06
申请号	CN00810302.X	申请日	2000-07-12
[标]申请(专利权)人(译)	美国卫生及公共服务部		
申请(专利权)人(译)	美国政府健康及人类服务部		
当前申请(专利权)人(译)	美国政府健康及人类服务部		
[标]发明人	艾拉帕斯坦 马格努斯埃森德 比扬库克李 乔治瓦斯马茨斯 库尔特沃尔夫冈 乌尔里希布林克曼		
发明人	艾拉·帕斯坦 马格努斯·埃森德 比扬库克·李 乔治·瓦斯马茨斯 库尔特·沃尔夫冈 乌尔里希·布林克曼		
IPC分类号	G01N33/53 A61K31/7088 A61K38/00 A61K38/17 A61K39/00 A61K39/39 A61K48/00 A61P35/00 C07K14/47 C07K14/82 C07K16/18 C07K16/30 C12N15/09 C12N15/12 C12Q1/02 C12Q1/68 G01N33 /50 G01N33/566 G01N33/574		
CPC分类号	C07K14/47 A61K38/00 A61K39/00		
代理人(译)	丁业平 王维玉		
优先权	60/157471 1999-10-01 US 60/143560 1999-07-13 US		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明提供了一种含有来自前列腺上皮细胞和很多乳房肿瘤细胞的TCRγ转录本的序列的核酸以及由这些序列翻译表达的T - 细胞受体γ可变阅读框架蛋白(“TARP”)。从TARP制备而来的疫苗对于增加对其中该蛋白被表达的细胞(包括前列腺癌细胞和很多乳房癌细胞的)的免疫应答是有用的。本发明提供了诊断前列腺癌和表达TARP的乳房癌细胞存在的方法,以及将TARP和编码TARP的核酸给药于个体的方法。

```

GGGCAAGAGTTGGGCAAAAAATCAAGGTATTGGTCCCGAACAAGCTTATCATACA 60
<----- JY1.2 ----->
M Q M F P P S P L F F F L Q L L K Q S S R R
GATAAAACAATGAGCAGATGTTTCCCAAGCCACTADTTTCTTCCTCAATTGCTGAACAAGCTCCAGAAGCC 140
<----- C Y I(外显子CI) ----->
L E H T F V F L R N F S L M L L R Y I G K K R R A T R
TGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTAATAAGATACATTGGCAAGAAAAGAAGCAACACGA 220
F W D P R R G T P M K T N D T Y M K F S W L T V P E K
TTCGGGATCCCGAGGAGGGGAACACCATAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTCCGAGAAAG 300
S L D K E H R C I V R H E N N K N G V D D Q E I I F P P
TCACCTGGACAAGAACAAGACAGATGATCTCCAGACTGGAAATATAAACAACGGAGTTGATCAAGAAATATCTTCTCTCC 380
I K T D V I T H D P K D N C S K D A N D T L L L L Q L
AATAAAGCGGATGTCATCACAAATGGATCCCAAGACAATGTCCAAAGATGCAAAATGATACACTGCTGCGAGCTCA 460
<----- C Y I(外显子CII) ----->
T N T S A Y Y M Y L L L L L L K S V V Y F A I I T C C L
CAAACACCTCGCATATACATGACCTCCTCCTCAAGAGTGTGGTATATTTGGCATCATCACCTGCTGCTG 540
<----- C Y I(外显子CIII) ----->
L R R T A F C C N G E K S
C T T A G A A G A C G G C T T T C T G C A A T G G A G A A A T C A T A C A G A C G G T G C A C A A G G A G C C A T C T T T C C T C A T C G G 620
T T A T G T C C C T A G A A G C G T C T C T G A G G A T C T A G T T G G G C T T T C T T C T G G G T T G G G C A T T T C A G T T C T C A T G T G T 700
A C T A T T C T A T C A T T A T G T A T A C G G T T T C A A A C A G T G G C A C A G A G A A C C T C A C T C T G T A A T A C A A T G A G G A A T 780
A G C C A G C G G A T C C A G C A C C A A T C T C C A T G T T T T C C A G A G C T C C C A G C A A C C C A A A T A G C G C C T G C T A T A G T G 860
T A G A C A T C C T G C G G C T T C T A G C C T T G T C C T C T T A G T G T C T T T A A T C A G A T A A C T G C C T G G A A G C C T T C A T P T T T A C 940
A C G C C T G A A G C A G T C T C T T T G C T A G T G A A T A T G T G T G T T T T C C G T A A T A A G C A A A A T A A A T T T T A A A A A A A T G 1020
A A A A G T T 1027

```