

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
G01N 33/53 (2006.01)



## [12] 发明专利申请公布说明书

[21] 申请号 200680042935.X

[43] 公开日 2008年11月19日

[11] 公开号 CN 101310185A

[22] 申请日 2006.9.19

[21] 申请号 200680042935.X

[30] 优先权

[32] 2005.9.19 [33] US [31] 60/718,843

[86] 国际申请 PCT/US2006/036538 2006.9.19

[87] 国际公布 WO2007/035766 英 2007.3.29

[85] 进入国家阶段日期 2008.5.16

[71] 申请人 约翰·霍普金斯大学

地址 美国马里兰州

[72] 发明人 Z·张 D·W·陈

C·N·罗森茨魏希

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 戈 泊

权利要求书4页 说明书42页 附图25页

[54] 发明名称

前列腺癌的生物标记

[57] 摘要

本发明提供以蛋白质为基础的生物标记，也就是蛋白C抑制剂(PCI)，用于鉴别病人的前列腺癌状态。特别地，本发明的生物标记用于分类受测样本为前列腺癌或非前列腺癌。生物标记可通过SELDI质谱分析检测。

- 1、一种用于鉴别受测对象中的前列腺癌状态的方法，包括：
  - (a) 测量来自所述受测对象生物样本中的 PCI 生物标记；以及
  - (b) 将所述测量与前列腺癌的状态进行关联化。
- 2、如权利要求 1 所述的方法，其中，所述前列腺癌的状态通过 Gleason 评分确定。
- 3、如权利要求 1 所述的方法，其中，生物标记 C4a 也在从所述相同受测对象的相同生物样本中测量。
- 4、一种确定前列腺癌复发机率的方法，包括：
  - (a) 测量来自所述受测对象生物样本中的 PCI 和 C4a 生物标记；以及
  - (b) 将所述测量与前列腺癌的复发进行关联化。
- 5、如权利要求 1 所述的方法，其中，所述 PCI 生物标记通过在 SELDI 探针的吸附剂表面上捕捉所述生物标记，并且通过激光解吸-离子化质谱检测所述已捕捉的生物标记来测量。
- 6、如权利要求 1 所述的方法，其中，所述 PCI 生物标记通过免疫测定来测量。
- 7、如权利要求 1 所述的方法，其中，所述样本是血清。
- 8、如权利要求 1 所述的方法，其中，所述关联化通过软件分类算法完成。
- 9、如权利要求 1 所述的方法，其中，前列腺癌的状态选自前列腺癌和非前列腺癌。
- 10、如权利要求 1 所述的方法，其中，前列腺癌的状态选自使用 Gleason 评分为早期的 1 分和晚期的 5 分。
- 11、如权利要求 1 所述的方法，进一步包括：(c) 基于所述状态处理受测对象的治疗。

12、如权利要求 5 所述的方法，其中，所述吸附剂是 IMAC 铜吸附剂。

13、如权利要求 5 所述的方法，其中，所述吸附剂是生物特异性吸附剂。

14、如权利要求 1 所述的方法，其中，若所述测量与前列腺癌相关联，则处理受测对象的治疗包括向所述受测对象施加 PCI 拮抗剂。

15、如权利要求 1 所述的方法，进一步包括：(d) 在处理受测对象后，测量至少一种所述生物标记以及将所述测量与疾病进程进行关联化。

16、一种用于确定前列腺癌过程的方法，包括：

(a) 在第一时间，测量来自所述受测对象生物样本中的 PCI 生物标记；

(b) 在第二时间，测量来自所述受测对象生物样本中的 PCI 生物标记；以及

(c) 比较所述第一测量和第二测量；

其中，所述比较的测量决定前列腺癌的过程。

17、一种组合物，包括经纯化的 PCI。

18、一种组合物，包括特异性结合 PCI 的生物特异性捕捉剂。

19、如权利要求 18 所述的组合物，其中，所述生物特异性捕捉剂是抗体。

20、如权利要求 18 所述的组合物，其中，所述生物特异性捕捉剂结合于固相载体。

21、一种组合物，包括结合于 PCI 的生物特异性捕捉剂。

22、一种试剂盒，包括：

(a) 固相载体，包括至少一种附着其上的捕捉剂，其中，所述捕捉剂与 PCI 结合；以及

(b) 使用所述固相载体检测 PCI 的指导说明。

23、如权利要求 22 所述的试剂盒，包括使用所述固相载体检测 PCI 的指导说明。

24、如权利要求 22 所述的试剂盒，包括使用所述固相载体检测 PCI 的指导说明。

25、如权利要求 21、22 或 23 中任一项所述的试剂盒，其中，所述包括捕捉剂的固相载体是 SELDI 探针。

26、如权利要求 21、22 或 23 中任一项所述的试剂盒，其中，所述捕捉剂是 IMAC 铜吸附剂。

27、如权利要求 21、22 或 23 中任一项所述的试剂盒，还包括：  
(c) 含有所述 PCI 生物标记的容器。

28、如权利要求 21、22 或 23 中任一项所述的试剂盒，还包括：  
(c) 免疫特异性的层析吸附剂。

29、一种试剂盒，包括：

- (a) 固相载体，包括至少一种附着其上的捕捉剂，其中，所述捕捉剂与 PCI 结合；以及
- (b) 含有所述 PCI 生物标记的容器。

30、如权利要求 26 所述的试剂盒，还包括：(c) 免疫特异性的层析吸附剂。

31、一种软件产品，包括：

- a. 存取属于样本的数据的编码，所述数据包括在所述样本中 PCI 生物标记的测量；以及
- b. 执行分类算法的编码，所述分类算法依照测量的函数分类所述样本的前列腺癌状态。

32、如权利要求 31 所述的软件产品，其中，所述分类算法依照生物标记 PCI 测量的函数分类所述样本的前列腺癌状态。

33、如权利要求 31 所述的软件产品，其中，所述分类算法进一步依照分子量测量的函数分类所述样本的前列腺癌状态。

34、一种方法，包括通过质谱测定或免疫测定来检测所述生物标记 PCI。

35、一种方法，包括向受测对象传达与前列腺癌状态相关的诊断，所述前列腺癌状态通过来自所述受测对象生物样本中的生物标记关联化来确定，其中，所述生物标记选自由 PCI 和 C4a 所组成的群组。

36、如权利要求 35 所述的方法，其中，所述诊断通过计算机产生的媒体向所述受测对象传达。

37、一种用于识别与 PCI 相互作用的化合物的方法，其中，所述方法包括：

- a) 将 PCI 与测试化合物接触；以及
- b) 确定所述测试化合物是否与 PCI 相互作用。

38、一种用于调节细胞中 PCI 浓度的方法，其中，所述方法包括：

- a) 将所述细胞与抑制 PCI 表达的小分子接触。

39、一种治疗在受测对象中状态的方法，其中，所述方法包括：向受测对象施加治疗上有效剂量的小分子，其中，所述小分子抑制 PCI 的表达。

40、如权利要求 39 所述的方法，其中所述状态是前列腺癌。

## 前列腺癌的生物标记

### 相关申请案

本申请案以 2005 年 9 月 19 日申请的美国第 60/718,843 号临时申请案请求优先权，临时申请案的完整内容特别并入本案作为参考。

### 技术领域

本发明提供对于检测前列腺癌来说是重要的生物标记。此生物标记的识别是通过使用 SELDI 分析区别前列腺癌患者与健康个体的血清蛋白图谱。本发明提供与生物标记相关的系统和方法，其中，生物标记用于鉴别前列腺癌的状态。本发明也确认所述生物标记为公知的蛋白 C 抑制剂(Protein C Inhibitor, 简称 PCI)。

### 背景技术

前列腺癌是男性最常见的癌症类型。在 2002 年，预计在男性中将有 189,000 件的新癌症病例，并有 30,200 例将死于此疾病。早期检测出被限定在前列腺的前列腺癌，并通过前列腺癌根治术(手术)具有最佳的治愈可能。尽管 PSA 被认为是一种有效的肿瘤标记，并实际上具器官特异性，但是并不具癌症特异性。PSA 浓度在患有前列腺癌与患有良性前列腺疾病的男性中有相当大的重叠。PSA 不能从患有非器官限定的前列腺癌男性(不会受益于手术)中区分出患有器官限定的前列腺癌男性(会受益于手术)。因此，PSA 对于选择适合前列腺根治术的病人并不具有功效。

当前，前列腺癌的早期检测和诊断是依靠直肠指检(digital rectal examination, 简称 DRE)、前列腺特异性抗原(prostate specific antigen, 简称 PSA)测量、经直肠超声术(transrectal ultrasonography, 简称 TRUS)和经直肠针刺活检(transrectal needle biopsy, 简称 TRNB)。目前，血清 PSA 测量与 DRE 的组合为检测和诊断前列腺癌的主要方法。

商业上可得到的 PSA 分析通常以试剂盒出售，此分析在区域性或地方性的实验室中进行。然而，前列腺特异性抗原(PSA)和前列腺酸性磷酸酶(prostatic acid phosphatase, 简称 PAP)具有局限的治疗和诊断可能性。举例来说，PSA 的水平与前列腺癌的存在经常没有很好的关联，在部分包括良性前列腺增生症(benign prostatic hyperplasia, 简称 BPH)的非前列腺癌病例呈阳性反应。此外，PSA 的测量与前列腺的体积相关，并且无法指明转移的水平。

这些试剂盒在现行早期检测前列腺癌的方案上扮演部分角色。然而，当阴性直肠指检(DRE)遇到适度地非正常 PSA 值(4-10 ng/ml)时将产生问题。仅有 20-30%具有这些发现的个体在活检中证明为癌症(Kantoff 和 Talcott, 8(3) Hematol. Oncol. Clinics N Amer 555(1994))。因此，发展提高 PSA 测试阳性预测值的方案是重要的。这样的方案包括建立经年龄调整的正常范围、确定游离与全部 PSA 的比例、修正前列腺的质量(密度)，以及计算 PSA 值变化的速率(Kantoff 和 Talcott, 8(3) Hematol. Oncol. Clinics N Amer 555(1994)以及 Brawer, 45 CA-A Cancer J Clinicians 148 (1995))。尽管每种方案均有一定作用，关于如何处理 PSA 阳性和 DRE 阴性的病人仍存在相当大的不确定性。

另外，PSA 并不是一种疾病特异性标记，因为在大部分患有 BPH 和前列腺炎的病人中可检测到提高的 PSA 水平(25-86%) (Gao et al., 1997, Prostate 31: 264-281)，并且在其它非恶性疾病和一些正常男性中，它是显著局限标记的诊断特异性的因素。举例来说，在 BPH 中可观察到血清 PSA 提升为 4 到 10 ng/ml，甚至在前列腺炎中可观察到更高的数值，尤其是急性前列腺炎。BPH 是男性中极为常见的疾病。进一步混淆此情况的事实为血清 PSA 的提升可在没有任何指明为 DRE 疾病下被检测，反之亦然。另外，目前 PSA 已被认为不具前列腺特异性(Gao et al. 如前，重新检阅)。

对于晚期前列腺癌的处理更可靠和有益的阶段发展以及预防方法也是需要的。临床确定前列腺肿瘤的阶段是依靠直肠检查，确定肿瘤是否在前列腺囊的区域内(局部限定)或超出其区域(局部晚期)，并且搭配血清 PSA 测定和经直肠超声引导活检。然而，这些技术中没有一

种在预测疾病的进程上被证明为可靠的。

因此，有需要可特异性识别前列腺癌、可从良性增生症中区分出前列腺癌、可在即使低 PSA 水平下识别前列腺癌，以及识别疾病进程的阶段的方法。

## 发明内容

本发明通过测量特异性的生物标记提供用于确定前列腺癌状态的敏感且快速的方法和试剂盒。在病人样本中测量标记，提供诊断者可能的人类癌症诊断或阴性诊断(例如正常或非疾病)相关联的信息。标记的特征在于它的分子量和已知的蛋白质特性，蛋白 C 抑制剂(“PCI”)。通过使用多种分离技术可将标记从样本的其它蛋白质中分辨出，例如通过层析分离搭配质谱分析、利用固定抗体的蛋白质捕捉或是通过传统的免疫分析。在优选的具体实施例中，分离的方法包含表面增强激光解吸/离子化(Surface-Enhanced Laser Desorption/Ionization, 简称“SELDI”)质谱分析，其中，质谱分析探针的表面包括结合标记的吸附剂。

本发明提供鉴别受测对象前列腺癌状态的方法，其包括(a)测量来自受测对象样本的 PCI 生物标记，以及(b)将测量与前列腺癌的状态相关联化，特别是经由 Gleason 评分。在特定的方法中，测量的步骤包括检测样本中生物标记的存在或不存在。在其它方法中，测量的步骤包括量化样本中标记的含量。在其它的方法中，测定的步骤包括鉴别样本中生物标记的类型。

在特定的具体实施例中，本发明可在血清 PSA 水平提高的男性中确定患有前列腺癌的病人和患有良性前列腺疾病的病人的区别。绝大多数此类病患的血清游离和总 PSA 的比例范围为 10-20%(或复合物和总 PSA 比例的相当范围)，此为诊断的灰色区域。所揭示的生物标记具有在此范围内增进癌症检测的可能性。

本发明也有关于方法，其中测量的步骤包括：提供受测对象的血液或尿液或血液/尿液衍生物的样本；利用阴离子交换树脂分离样本中的蛋白质，收集含有 PCI 的馏分，PCI 则收集自包含结合此蛋白质生物标记的捕捉剂的基材表面。血液的衍生物如血清或血浆。在优选的具

体实施例中，基材是包括 IMAC 铜表面的 SELDI 探针，并且其中蛋白质标记通过 SELDI 检测。在其它的具体实施例中，基材为包括结合 PCI 的生物特异亲和性试剂的 SELDI 探针，并且其中蛋白质生物标记通过 SELDI 检测。在其它的具体实施例中，基材是包括结合 PCI 的生物特异亲和性试剂的微孔板，并且蛋白质生物标记通过免疫分析检测。

在特定的具体实施例中，此方法进一步包括基于通过此方法所确定的状态来处理受测对象的治疗。举例来说，若本发明方法的结果为不确定的或有必要进行状态的确认，医生可安排更多的测试。或者，若状态指明手术是合适的，医生可为病患进行手术的排期。同样地，若测试的结果为阳性，例如状态为晚期前列腺癌或状态为急性，则没有进一步的行动可被保证。此外，若结果显示治疗成功，则无进一步处理的必要。

本发明也提供在处理受测对象后再次测定 PCI 生物标记的方法。在这些实例中，处理治疗受测对象的方法依据所获得的结果随后重复或变更。在优选的具体实施例中，PCI 在治疗后单独或与 C4a 组合测定，以确定复发的可能性。

“前列腺癌状态”一词是指病人疾病的状态。前列腺癌状态的类型实例，包括但不限于，受测对象罹癌的风险、疾病的存在或不存在、病人疾病的阶段，以及疾病治疗的效果。其它状态与每一状态的程度为本领域所已知。

在特定的优选具体实施例中，此方法进一步包括测量来自受测对象的样本中至少一种先前已知的标记(这里以“标记 X”表示)，并将至少一种标记 X 和 PCI 的测量与前列腺癌的状态进行关联化。在特定的具体实施例中，除 PCI 标记外，仅有一种标记 X 被测量，而在其它的具体实施例中，则一种以上的标记 X 被测定。

公知的标记 X 实例包含前列腺癌的生物标记，包括但不限于 C4a、APOC1、BRCA1、CHGA、CHGB、CLU、COL1A1、COL6A1、EGF、ERBB2、ERK8、FGF1、FGF10、FGF11、FGF13、FGF14、FGF16、FGF17、FGF18、FGF2、FGF20、FGF21、FGF22、FGF23、FGF3、FGF4、FGF5、FGF6、FGF7、FGF8、FGF9、GNRH1、IGF1、IGF2、IGFBP3、IGFBP6、I112A、I11A、I11B、IL2、IL24、INHA、INSL3、INSL4、KLK10、KLK12、KLK13、KLK14、

KLK15、KLK3、KLK4、KLK5、KLK6、KLK9、MMP2、MMP9、MSMB、NTN4、ODZ1、PAP、PLAU、PRL、PSAP、SERPINA3、SHBG、TGFA、TIMP3、VEGF和标记的蛋白质变形(例如断裂形式、异构体)。

本发明的 PCI 标记以多种态样的一种或一种以上为特征。特别地在一种态样中,标记以在本文明确指示条件下的分子量为特征,特别是指经质谱分析确定。在另一种态样中,此标记以标记质谱特性的特色为特征,例如标记谱峰的尺寸(包括面积)和/或形状,其特色则包含相邻峰的邻近度、尺寸和形状等。在又一个态样中,标记以亲和性的结合特性为特征,尤其是于特定的情况下结合至 IMAC 铜吸附剂的能力,然而,其它如镍的金属也可使用。在优选的具体实施例中,本发明的标记可以这些态样中的每一个为特征,也就是分子量、质谱特性和与 IMAC-Cu 吸收剂的结合。

就本文所揭示的标记质量值来说,光谱仪的质量精确度被认为是所揭示分子量值的约正负 0.15 百分比内。另外,这些被认知为仪器的精确度变化,所确定的质谱可于约 400 至 1000 m/dm 的分辨率限制范围内变化,其中 m 为质量, dm 为 0.5 峰高处的质谱峰宽。与质谱仪及其操作相关的这些质量精确度与分辨率变化会使用「约」字在揭示标记质量时。这些质量精确度与分辨率变化是可预期的,因此与本文所揭示标记的「约」字相关的意义包括由于受测对象的性别、基因型和/或种族,以及特定癌症或起源或其阶段而存在的标记变形。

本发明进一步提供一种鉴别受测对象前列腺癌状态的方法,包括(a)测定来自受测对象样本的 PCI 生物标记,以及(b)将此测定与前列腺的状态相关联。在特定的方法中,测定的步骤包括检测样本中标记存在或不存在。在其它的方法中,测定的步骤包括量化样本中标记的含量。

诊断测试的精确度以接收者操作特征曲线(Receiver Operating Characteristic curve, 简称 ROC 曲线)为特征。ROC 是诊断测试中不同可能性切点的真阳性率相对假阳性率的图式。ROC 曲线显示敏感度与特异性间的关系,也就是,敏感度的增加将伴随特异性的减少。曲线沿着左轴且随后与 ROC 空间顶部边缘越接近,测试越精确。相反地,曲线越靠近 ROC 图 45 度对角线,测试则越不精确。ROC 下的区域为测

试精确度的方法。测试的精确度取决于测试能将待测群区分为患有及未患有所考虑疾病的程度。曲线下的区域(也就是 area under the curve, 简称 AUC)为 1 时表示为理想的测试, 而区域为 0.5 则表示为较无用的测试。因此, 本发明的生物标记与诊断方法具有大于 0.5 的 AUC, 优选的测试具有大于 0.6 的 AUC, 更优选的测试具有大于 0.7 的 AUC。

测定 PCI 生物标记的优选方法包含使用生物芯片数组。适用于本发明的生物芯片数组包含蛋白质与核酸数组。一种或多种标记被捕捉于生物芯片数组上, 并且通过激光离子化检测标记的分子量。标记的分析是通过如标记分子量相对于根据总离子电流标准化的阈值强度。优选地, 使用对数转换来降低峰强度范围, 从而限制被检测标记的数量。

在本发明优选的方法中, 将 PCI 生物标记的测定与前列腺癌状况相关联化的步骤是通过软件分类算法完成。优选地, 从生物芯片数组上固定的受测对象样本所产生的数据是将所述生物芯片数组通过激光离子化, 并检测质/荷比率的信号强度; 以及将数据转换成计算机可读取的形式; 然后, 根据使用者输入参数执行用以分类数据的算法, 作为检测表明标记存在于前列腺癌病患中却于无癌症的受测对象控制组内不存在的信号。

优选的生物芯片表面例如离子性、阴离子性、由固定的镍离子所组成、由阴、阳离子的混合物所组成、由一种或多种抗体所组成、单链或双链核酸、蛋白质、肽或其片段、氨基酸探针、或噬菌体展示库 (phage display libraries) 所组成。

在其它优选的方法中, 一种或多种标记是使用激光解吸/离子化质谱分析测定, 包括提供适用于含有吸附剂附着的质谱仪探针, 以及使受测对象样本与吸附剂接触; 然后, 从探针上解吸与离子化标记, 并用质谱仪检测经解吸/离子化的标记。

优选地, 激光解吸/离子化质谱分析: 提供包括吸附剂附着其上的基质; 将受测对象样本与吸附剂接触; 将基质放置于适用于含有吸附剂附着的质谱仪探针上, 以及从探针上解吸和离子化标记, 并用质谱仪检测经解吸/离子化的标记。

吸附剂可例如为疏水性、亲水性、离子性或金属螯合吸附剂, 例

如镍或抗体、单或双链寡核苷酸、氨基酸、蛋白质、肽或其片段。

本发明的方法可在经得起所述方法试验的任何类型的病患样本中进行，如血液、血清和血浆。

本发明也提供试剂盒，包括(a)一种结合 PCI 生物标记的捕捉剂；以及(b)包括生物标记的容器。在优选的具体实施例中，捕捉剂结合生物标记。捕捉剂可为任何类型的试剂，优选的试剂为 SELDI 探针。捕捉剂也可结合其它公知的生物标记，如标记 X。在特定的优选具体实施例中，所述试剂盒进一步包括可与第一捕捉剂未结合的生物标记结合的第二捕捉剂。

在本发明特定的试剂盒中，捕捉剂包括固定的金属螯合物(“IMAC”)。

本发明的特定试剂盒进一步包括清洗溶液，在洗涤后相对于其它生物标记可选择性地保留结合到捕捉剂的生物标记。

本发明也提供试剂盒，包括(a)结合 PCI 生物标记的第一捕捉剂，及(b)使用捕捉剂测定生物标记的指导说明。在这些特定的试剂盒中，捕捉剂包括抗体。另外，一些试剂盒进一步包括捕捉剂附着或可附着的 MS 探针。在一些试剂盒中，捕捉剂包括 IMAC。试剂盒也可包含清洗溶液，在洗涤后相对于其它生物标记可选择性地保留结合捕捉剂的生物标记。优选地，所述试剂盒包括使用试剂盒确定前列腺癌状态的书面指导说明，以及提供将捕捉剂与测试样本接触和测定一种或多种由捕捉剂保留的生物标记的指导说明。

所述试剂盒也提供为抗体、单或双链寡核苷酸、氨基酸、蛋白质、肽或其片段的捕捉剂。

使用所述试剂盒测定蛋白质生物标记，是通过质谱分析或如 ELISA 的免疫分析。

本发明也提供用于检测前列腺癌和/或用于为进一步诊断检测分析而产生抗体的纯化蛋白质。纯化的蛋白质包含纯化的 PCI 肽。本发明也提供进一步包括可检测卷标的纯化蛋白质。

在另一个具体实施例中，非侵入性的医疗成像技术，如经阴道超声、正电子放射断层扫描(positron emission tomography, 简称 PET)或单光子发射计算机断层扫描(single photon emission computerized

tomography, 简称 SPECT) 成像于检测癌症、冠心病和脑疾病尤为有效。多普勒血流超声、PET 和 SPECT 成像显示器官与组织的化学功能, 而其它如 X 射线、CT 和 MRI 的成像技术主要则显示结构。血流超声、PET 和 SPECT 成像的使用对于鉴别与监控如前列腺癌的疾病发展变得日益有用。

本文所述的 PCI 生物标记或其片段可用于前文所述 PET 与 SPECT 成像的应用。在以合适的追踪剩余物修饰 PET 或 SPECT 的应用后, 与肿瘤蛋白质反应的 PCI 生物标记可用于在前列腺癌病患体内成像生物标记的沉积。

本发明的其它态样将于下文中描述。

## 附图说明

图 1 显示 PCI 生物标记的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 2 显示商业血清 PCI 生物标记的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 3 显示使用 Q 馏分 6 的 RPC 分馏的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 4 显示使用 RPC40% 馏分、YM30 过滤和 NP20 数组的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 5 显示使用 Q 分馏的商业血清和 CM10 数组的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 6 显示 Q 流出液的 RPC 分馏和 CM10 数组的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 7 显示使用重复的 Q-FT 的 RPC 分馏、RPC-40% 的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 8 显示用于 3890 Da 蛋白质 Q-TOF 光谱的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 9 显示具有 3889 m/z 的离子的 MS/MS 光谱的代表性质谱图谱。在这些图式中, 特定标记的质谱峰在描述的光谱中以分子量标明。

图 10 是蛋白质探测器搜索输入的代表性实例。

图 11 是数据库搜索结果的代表性实例。

图 12 是 MS/MS 数据的代表性解释。

图 13 是 PCI 氨基酸序列的图式。

图 14 是 PCI 的理论分子量的图式。

图 15 是使用全血清和山羊抗 PCI 抗体的以珠为基础的免疫分析结果的图式。

图 16 显示用于发现和确认的多中心样本组 (Multicenter Sample Sets for Discovery and Validation) 所得到的数据。

图 17 是通过相关网络分析发现的图式。

图 18 是与 Gleason 评分相关联化的数据以及从良性前列腺疾病分离前列腺癌的图式。

图 19 是与前列腺癌复发相关联化数据的图式。

图 20 是复发研究数据组 PCI 独立确认的图式。

图 21 是使用部分复发样本作为训练组的 Pre-RRP PSA、PCI 和 C4a 的组合的图式。

图 22 是使用部分复发样本作为训练组的测试组的 Pre-RRP PSA、PCI 和 C4a 的组合的图式。

图 23 是表示 PCI 峰的 ROC 分析的图式, 将 Gleason 评分 5-7 的前列腺癌病人与 Gleason 评分为 8 或更高的病人予以区分。

图 24 是表示 PCI 峰的 ROC 分析的图式, 将患有良性前列腺疾病的病人与前列腺癌病人予以区分。

## 具体实施方式

### 1. 引言

生物标记是一种有机分子, 其差异性地存在于具有一种表现状态 (例如患有疾病) 和另一种表现状态 (例如未患有疾病) 的受测对象样本中。当于不同的组群中, 一种生物标记的平均或中值表达水平的计算具有统计上显著的意义时, 此生物标记是差异性地存在于不同的表现状态间。于其它的测试中, 计算统计上具显著意义的通常测试包括 T-测定 (t-test)、ANOVA、Kruskal-Wallis、Wilcoxon、Mann-Whitney

及比数比。单独或组合的生物标记提供对受测对象属于一种或其它表现状态相对风险的测定。因此，生物标记对于疾病(诊断)、药物的治疗效果(治疗诊断学)及药物毒性皆为有用的标记。

## 定义

除非另行定义，于本发明所用的所有技术和科学术语通常可被技术领域的人员所理解。以下参考文献提供技术人员用于本发明许多术语的一般定义：Singleton et al., Dictionary of Microbiology and Molecular Biology (2<sup>nd</sup> ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5<sup>th</sup> Ed., R. Rieger et al. (eds.), Springer Verlag (1991); 以及 Hale & Marham, The Harper Collins Dictionary of Biology (1991)。除非另行指明，于此所用的以下术语具有其所属的含意。

于此所使用的“前列腺疾病”或“前列腺状态”是指任何前列腺疾病或状态，包括但不限于，良性前列腺增生症(BPH)、前列腺炎、前列腺上皮内瘤(prostatic intraepithelial neoplasia, 简称 PIN)和癌症。

于此所使用的“前列腺癌”是指任何前列腺恶性疾病，包括但不限于，腺癌、小细胞未分化癌和粘液腺(胶质)癌。

于此所使用的“转移性前列腺癌”和“转移性疾病”是指前列腺癌扩散至局部的淋巴结或至更远的位置，并包含在 AUA 系统下的 D 阶段疾病和在 TNM 系统下的 TxNxM+阶段。如同局部的晚期前列腺癌情况，手术通常不用于具有转移性疾病的病人，而激素(雄激素阻断)治疗是优选的治疗形式。在治疗初始的 12-18 个月内，患有转移性前列腺癌的病人最终发展为雄激素非依赖性阶段(androsten-refractory state)，并且约一半的此类病人在此后的 6 个月内死亡。最常见的前列腺癌转移位置是骨。总而言之，前列腺癌骨转移典型地为成骨性而非溶骨性(也就是导致网状骨的形成)。骨转移最常发生于脊椎，之后为大腿骨、骨盆、胸肋、头骨和肱骨。其它常见的转移位置包括淋巴结、肺、肝和脑。转移性前列腺癌典型地通过开放或腹腔镜骨盆淋巴结切除术、全身放射性核素扫描、骨骼放射线照相术和/或骨损伤活检。

于此所使用的“局部晚期前列腺癌”和“局部晚期疾病”是指前列腺癌延伸通过前列腺囊，并且包含在美国泌尿协会(American Urological Association, 简称 AUA)系统下的 C 阶段疾病、在 Whitmore-Jewett 系统下的 C1-C2 阶段疾病，以及在 TNM(肿瘤、瘤、转移瘤)系统下的 T3-T4 和 N+疾病。通常，患有局部晚期疾病的病人并不建议进行手术，并且相较于患有临床局部(器官限定)的前列腺癌病人，这些病人实质上得到较少有利的结果。局部晚期疾病通过前列腺侧部边缘外明显的硬化迹象或前列腺主要部分上的不对称或硬化在临床上识别。在前列腺根治术后，病理性地诊断局部晚期前列腺癌，是否肿瘤侵入或渗透前列腺囊、扩散至手术边缘或侵入精囊。

于此所使用的“肿瘤阶段”或“肿瘤进程”是指肿瘤的不同临床阶段。肿瘤的临床阶段通过在医学界已完善建立的不同参数定义。一些参数包含形态学、肿瘤大小、通过病人身体的转移程度及其相似者。

“Gleason 分级系统”是定前列腺癌等级的系统。Gleason 分级系统是在组织样本中两个最大癌症区域的每一个指定级别。级别的范围从 1 至 5，1 为最无侵犯性，而 5 为最具侵犯性。举例来说，3 级的肿瘤很少转移，但转移则通常为 4 级或 5 级。两个等级随后相加以产生 Gleason 评分。2-4 的评分为低级；5-7 为中级，以及 8-10 为高级。具有低 Gleason 评分的肿瘤通常生长缓慢，在病人的一生当中并不造成重大的威胁。

“蛋白 C 抑制剂”(PCI)是血浆丝氨酸蛋白酶抑制剂，抑制数种包含抗凝血酶、活化蛋白 C(APC)以及凝血酶和 Xa 因子的凝血酵素的蛋白酶，并且在精液中存在有高浓度。

“C4a”是在 C1s 分裂 C4 为 C4a 和 C4b 时所形成的较小片段。如同过敏毒素，C4a 造成立即的过敏性症状，但其活性较 C3a 或 C5a 为弱。

“气相离子谱仪”是指检测气相离子的装置。气相离子谱仪包括提供气相离子的离子源。举例来说，气相离子谱仪包括质谱仪、离子迁移谱仪和总离子电流测量装置。“气相离子谱分析”是指使用气相离子谱仪检测气相离子。

“质谱仪”是指可将测量的参数转换为气相离子的质荷比的气相离子谱仪。质谱仪一般包括离子源和质量分析仪。质谱仪的实例是飞

行时间式、扇形磁场式、四极滤质器式、离子阱式、离子回旋共振式、扇形静电分析式及这些的混合。“质谱分析”是指使用质谱仪检测气相离子。

“激光解吸质谱仪”是指使用激光能作为解吸、挥发和离子化分析物的质谱仪。

“串联质谱仪”是指任何能够进行两个连续阶段基于  $m/z$  的区分或离子的测量的质谱仪，包括离子混合物中的离子。此用语包括具有两个质量分析仪的质谱仪，其能够进行两个连续阶段基于  $m/z$  的区分或空间串联的离子测量。此用语进一步包括具有单一质量分析仪的质谱仪，其能够进行两个连续阶段基于  $m/z$  的区分或时间串联的离子测量。此用语因此明确地包括  $Qq$ -TOF 质谱仪、离子阱质谱仪、离子阱-TOF 质谱仪、TOF-TOF 质谱仪、傅立叶转换离子回旋共振质谱仪、扇形静电-扇形磁场质谱仪和其组合。

“质量分析仪”是指质谱仪的组件，包括测量可转换为气相离子质荷比的参数的工具。在飞行时间质谱仪中，质量分析仪包括光离子组件、飞行管和离子检测仪。

“离子源”是指提供气相离子的气相离子谱仪组件。在一个具体实施例中，离子源提供通过解吸/离子化过程的离子。所述的具体实施例一般包括探针接口，其以可询问的关系将探针定位于离子化能量源（例如激光解吸/离子化源），并且在大气压或次大气压下与气相离子谱仪的检测器协同传递。

用于解吸/离子化固相分析物的离子化能形式，举例来说包括：(1) 激光能；(2) 快速原子（用于快速原子轰击）；(3) 通过放射性核素  $\beta$  衰变产生的高能粒子（用于等离子体解吸）；以及(4) 初级离子产生二级离子（用于二级离子质谱分析）。优选用于固相分析物的离子化能形式是激光（用于激光解吸/离子化），特别地为氮激光、Nd-Yag 激光和其它脉冲激光源。“注量”是指询问图像每单位面积传送的能量，如激光的高注量源传送约  $1\text{mJ}/\text{mm}^2$  至  $50\text{mJ}/\text{mm}^2$ 。通常，样本放置于探针表面，探针与探针接口接合，并且探针表面被击以离子化能。能量将来自表面的分析物分子解吸形成气相，并将其离子化。

用于分析物的其它形式离子化能包括，例如：(1) 离子化气相中性

粒子的电子；(2)强电场以诱导来自气相、固相或液相中性粒子的离子化；以及(3)应用具有离子化颗粒或中性化合物电场组合的来源，以诱导固相、气相和液相中性粒子的化学离子化。

“固相载体”是指可被捕捉剂诱导或附着于捕捉剂的固态物质。代表性的固相载体包括探针、微孔板和层析树脂。

本发明内容的“探针”是指适用于接合气相离子谱仪(例如质谱仪)的探针接口，并且引起分析物离子化能以进行离子化，以及引入如质谱仪的气相离子谱仪的装置。”探针”一般包括固体基质(弹性或坚硬)，表现分析物呈现给离子化能量源的表面的样本。

“表面增强激光解吸/离子化”或“SELDI”是指如质谱仪解析/离子化气相离子谱仪的方法，其中分析物被捕捉于 SELDI 探针表面，而探针则接合于气相离子谱仪的探针接口。在“SELDI MS”中，气相离子谱仪是质谱仪。SELDI 技术描述于例如美国第 5,719,060 号专利(Hutchens 和 Yip)以及美国第 6,225,047 号专利(Hutchens 和 Yip)。

“表面增强亲和捕捉”或“SEAC”是一种涉及使用包含吸附性表面的探针(“SEAC 探针”)的 SELDI 型式。“吸附性表面”是指结合吸附剂(也称为“捕捉剂”或“亲和剂”)的表面。吸附剂是任何能结合分析物(例如目标多肽或核酸)的物质。“层析吸附剂”是指典型用于层析的物质。层析吸附剂举例来说包含离子交换物质、金属螯合剂(例如次氨基乙酸和亚氨基二乙酸)、固定化金属螯合剂、疏水作用吸附剂、亲水作用吸附剂、染料、简单生物分子(例如核苷酸、氨基酸、简单糖和脂肪酸)及混合型吸附剂(例如疏水吸引/静电排斥吸附剂)。“生物特异性吸附剂”是指包括生物分子的吸附剂，生物分子例如核酸分子(如适体)、多肽、多糖、脂质、类固醇或这些的偶联物(如糖蛋白、脂蛋白、糖脂、核酸(如 DNA)-蛋白质偶联物)。在特定的实例中，生物特异性吸附剂可为大分子结构，例如多蛋白质复合物、生物膜或病毒。生物特异性吸附剂的实例是抗体、受体蛋白质和核酸。生物特异性吸附剂典型地相对于层析吸附剂，对目标分析物具有较高的特异性。进一步用于 SELDI 的吸附剂实例可在美国第 6,225,047 号专利(Hutchens 和 Yip, “Use of retentate chromatography to generate difference maps”, May 1, 2001)中找到。

在一些具体实施例中，SEAC 探针提供作为预活化的表面，而预活化的表面可被修饰以提供吸附剂的选择。举例来说，特定探针提供有可通过共价键结合生物分子的反应基团。环氧化物和酰基咪唑是可用于共价结合如抗体或细胞受体的生物特异性吸附剂的反应基团。

“吸附”是指可检测分析物非共价结合至吸附剂或捕捉剂。

“表面增强净解吸”或“SEND”是一种涉及使用探针的 SELDI 形式，所述探针包括化学结合于探针表面的能量吸收分子(“SEND 探针”)。

“能量吸收分子”(“EAM”)是指能够从激光解吸/离子化源吸收能量，随后促成与其接触的分析物分子解吸和离子化的分子。此用语包括用于 MALDI 的分子，通常是指“基质”，并且明确地包括肉桂酸衍生物、芥子酸(SPA)、氰氧-羟基-肉桂酸(“CHCA”)和二羟基苯甲酸、阿魏酸和羟苯乙酮衍生物以及其它。也包含用于 SELDI 的 EAM。SEND 进一步描述于美国第 5,179,060 号专利和于 2002 年 9 月 4 日申请的美国第 60/408,255 号专利(Kitagawa, “Monomers And Polymers Having Energy Absorbing Moieties Of Use In Desorption/Ionization Of Analytes” )。

“表面增强光不稳定附着和释放”或“SEPAR”是一种涉及使用具有基团的探针的 SELDI 形式，所述基团附着于可共价结合分析物的表面，随后暴露于如激光光束的光下，通过打断基团内的光不稳定键而释放分析物。SEPAR 进一步描述于美国第 5,719,060 号专利。

“洗提液”或“清洗溶液”是指典型为溶液的试剂，用于影响或修饰分析物吸附至吸附剂的表面和/或从表面移除未结合的物质。洗提液的洗提特性取决于如 pH、离子强度、疏水性、混乱度、清洁剂强度和温度。

“分析物”是指任何需要被检测的样本成分。此用语可指样本中的单一成分或复数成分。

吸附于亲和捕捉探针的吸附表面的样本“复杂性”是指所吸附的不同蛋白质种类的数目。

“分子结合伙伴”和“特异性结合伙伴”是指一对分子，典型为展示特异性结合的数对生物分子。分子结合伙伴包括但不限于受体和配体、抗体和抗原、生物素和抗生物素蛋白以及生物素和链霉抗生

物素。

“监控”是指以连续变化的参数记录变化。

“生物芯片”是指一般具有吸附剂附着平面的固体基材。经常地，生物芯片的表面包括复数可寻址的位置，其各自的位置具有吸附剂结合其上。生物芯片可用于结合探针接口，并因此如同探针般作用。

“蛋白质生物芯片”是指适用于捕捉多肽的生物芯片。许多蛋白质生物芯片被描述于技术领域。这些包括如 CIPHERGEN Biosystems (Fremont, CA)、Packard BioScience Company (Meriden CT), Zyomyx (Hayward, CA) 和 Phyllos (Lexington, MA)。这些蛋白质生物芯片的实施例揭露于以下专利或专利申请中：美国第 6,255,047 号专利 (Hutchens 和 Yip, “Use of retentate chromatography to generate difference maps”, May 1, 2001)、第 W099/51773 号国际公开 (Kuimelis 和 Wagner, “Addressable protein arrays”, October 14, 1999)；美国第 6,329,209 号专利 (Wagner et al., “Arrays of protein-capture agents and methods of use thereof”, December 11, 2001) 以及第 WO 00/56934 号国际公开 (Englert et al., “Continuous porous matrix arrays”, September 28, 2000)。

CIPHERGEN Biosystems 制造的蛋白质生物芯片包含具有层析或生物特异性吸附剂在可寻址位置附着其上的表面。CIPHERGEN ProteinChip® 数组包括 NP20、H4、H50、SAX-2、WCX-2、CM-10、IMAC-3、IMAC-30、LSAX-30、LWCX-30、IMAC-40、PS-10、PS-20 和 PG-20。这些蛋白质生物芯片包含带状铝基材。带的表面涂覆二氧化硅。

在 NP-20 生物芯片的案例中，二氧化硅的作用为捕捉亲水蛋白质的亲水吸附剂。

H4、H50、SAX-2、WCX-2、CM-10、IMAC-3、IMAC-30、PS-10 和 PS-20 生物芯片进一步包括功能化、交联的聚合物，以水凝胶形式物理性地结合于生物芯片的表面或共价地通过硅烷结合于生物芯片的表面。H4 生物芯片具有疏水结合的异丙基官能基。H50 生物芯片具有疏水结合的壬基苯氧基-聚(乙二醇)甲基丙烯酸酯。SAX-2 生物芯片具有阴离子交换的四级铵官能团。WCX-2 和 CM-10 生物芯片具有阳离子交换的羧酸官能团。IMAC-3 和 IMAC-30 生物芯片具有通过螯合吸收如  $\text{Cu}^{++}$  和  $\text{Ni}^{++}$

过渡金属离子的次氨基乙酸官能团。这些固定的金属离子允许通过配位键结吸附的肽和蛋白质。PS-10 生物芯片具有可与蛋白质上的基团反应用以共价结合的酰基咪唑官能基团。PS-20 生物芯片具有用于共价结合蛋白质的环氧官能基团。PS 系列的生物芯片适用于结合生物特异性吸附剂至芯片表面，用于特异性地从样本中捕捉分析物，其中生物特异性吸附剂例如抗体、受体、植物凝集素、肝磷脂、蛋白 A、生物素/链霉抗生物素和其相似者。PG-20 生物芯片是蛋白 G 结合的 PS-20 芯片。LSAX-30(阴离子交换)、LWCX-30(阳离子交换)和 IMAC-40(金属螯合)生物芯片在其表面具有功能性的乳胶珠。这样的生物芯片进一步描述于：WO 00/66265(Rich et al., “Probes for a Gas Phase Ion Spectrometer”, November 9, 2000); WO 00/67293(Beecher et al., “Sample Holder with Hydrophobic Coating for Gas Phase Mass Spectrometer”, November 9, 2000); 美国第 US20030032043A1 号专利申请(Pohl and Papanu, “Latex Based Adsorbent Chip”, July 16, 2002)和美国第 60/350,110 号专利申请(Um et al., “Hydrophobic Surface Chip”, November 8, 2001)。

一旦在生物芯片上被捕捉，分析物可通过不同的检测方法测得，检测方法选自如气相离子谱分析方法、光学方法、电化学方法、原子力显微镜和射频方法。气相离子谱分析方法在此描述。尤其感兴趣的为质谱分析的使用，特别是 SELDI。光学方法包括如荧光、冷光、化学冷光、吸光、反射、穿透和双折射或折射率(例如表面等离子共振、椭圆偏光法、共振镜法、光栅耦合波导法或干涉测量分析)。光学方法包括显微镜(共轭焦和非共轭焦)、成像方法和非成像方法。不同形式的免疫分析(如 ELISA)为普遍检测捕捉于固相的分析物的方法。电化学方法包括伏安法和电流测定法。射频方法包括多极共振光谱分析。

本发明文中的“标记”是指在得自患有癌症的病人样本相对于控制组受测对象(例如具有阴性诊断或未检测出癌症、正常或健康的受测对象)的样本中差异表现的多肽(具有特别明显的分子量)。所使用的“生物标记”可与“标记”交替使用。

“测量”一词是指包括检测样本中标记的存在或不存在、确定样本中标记的量和/或鉴别生物标记类型的方法。测量可通过技艺中所公

知的方法和进一步在此描述的方法完成，包括但不限于 SELDI 和免疫分析。任何合适的方法可用于检测和测量一种或更多在此所描述的标记。这些方法包括但不限于质谱分析(例如激光解吸/离子化质谱分析)、荧光(例如双抗体夹心免疫分析)、表面等离子体共振法、椭圆偏光法和原子力显微镜。

“检测”是指鉴别被测对象的存在、不存在或量。

“差异性地存在”一词是指在得自患有癌症的病人样本相对于控制组受测对象(例如具有阴性诊断或未检测出癌症、正常或健康的受测对象)存在标记量和/或频率的差异。例如一些在此描述的标记，相较于来自控制组受测对象的样本，在癌症病人的样本中存在有提高的水平。相对地，其它在此描述的标记，相较于来自控制组受测对象的样本，在癌症病人的样本中存在有降低的水平。此外，标记可为多肽，相对于控制组受测对象的样本，在患有癌症病人的样本中被检测存在较高或较低的频率。标记可以量、频率或两者差异性地存在。

若一个样本中的多肽量统计上显著不同于另一个样本中的多肽量，则多肽为差异性地存在于两个样本中。举例来说，若多肽存在至少约 120%、至少约 130%、至少约 150%、至少约 180%、至少约 200%、至少约 300%、至少约 500%、至少约 700%、至少约 900%或至少约 1000% 大于存在于另一个样本中，或若其在一个样本中为可检测到的，而在另一个样本中为不可检测到的，则多肽是差异性地存在于两个样本中。

可选择地或额外地，若在前列腺癌病人的样本中检测到多肽的频率统计上显著地高于或低于控制组的样本，则多肽是差异性地存在于两组样本中。举例来说，若在一组样本中被检测至少约 120%、至少约 130%、至少约 150%、至少约 180%、至少约 200%、至少约 300%、至少约 500%、至少约 700%、至少约 900%或至少约 1000% 更高频率或更低频率在另一组样本中，则多肽是差异性地存在于两组样本中。

“诊断”是指鉴别病理状态，也就是前列腺癌的存在或性质。诊断方法在其敏感度和特异性上不同。诊断分析的“敏感度”是测试为阳性(“真阳性”的百分比)的患病个体的百分比。患病的个体未被检测到是“假阴性”。在未患病的受测对象测试为阴性称为“真阴性”。诊断分析的“特异性”是 1 减去假阳性率，其中“假阳性”率被定义

为无疾病却被测试为阳性的比率。特定的诊断方法可不提供确定的状态诊断，若方法提供辅助诊断的阳性指示则为足够的。

标记的“测试量”是指存在被测试样本中标记的量。测试量可为绝对量(如  $\mu\text{g/ml}$ )或相对量(如信号的相对强度)。

标记的“诊断量”是指受测对象样本中的标记量与前列腺癌的诊断一致。诊断量可为绝对量(如  $\mu\text{g/ml}$ )或相对量(如信号的相对强度)。

标记的“控制量”可为任何量或量的范围，用于比较标记的测试量。例如，标记的控制量可为在无前列腺癌的人中的标记量。控制量可为绝对量(如  $\mu\text{g/ml}$ )或相对量(如信号的相对强度)。

于此所使用的“敏感度”一词是具有特定疾病病人的百分比。例如，在 PCa/HC 组中，本发明的生物标记具有约 98% 的敏感度。生物标记的评判为正确地从 103 个前列腺癌病人区分出 101 个，也就是  $101/103=98\%$ 。

于此所使用的“特异性”一词是正确地识别为患有特定疾病，也就是正常或健康受测对象的百分比。例如，特异性以患有特定疾病的受测对象数量和正常健康受测对象的比较计算得到。

“多肽”、“肽”和“蛋白质”在这里可替换使用，是指氨基酸残基的聚合体。此用语应用于氨基酸的聚合体，其中一个或更多氨基酸残基是对应于自然出现的氨基酸以及天然出现的氨基酸聚合体的类似物或仿真物。多肽可被修饰，例如通过糖残基的加形成糖蛋白。“多肽”、“肽”和“蛋白质”包含糖蛋白以及非糖蛋白。

“免疫分析”是使用抗体特异性结合抗原(如标记)的分析。免疫分析以使用特定抗体的特异性结合性质分离、标定和/或定量抗原为特征。

“抗体”是指由单一免疫球蛋白基因或多数免疫球蛋白基因或其片段实质上编码的多肽配体，特异性地结合并识别抗原决定簇(例如抗原)。已知的免疫球蛋白基因包括  $\kappa$  和  $\lambda$  轻链恒定区基因， $\alpha$ 、 $\gamma$ 、 $\delta$ 、 $\epsilon$  和  $\mu$  重链恒定区基因，以及众多免疫球蛋白可变区域基因。抗体例如以完整的免疫球蛋白或以一些被不同肽酶消化所产生的特征化片段存在。这包括如 Fab' 和 F(ab)'<sub>2</sub> 的片段。这里所使用的“抗体”一词也包括由完整抗体修饰或使用重组 DNA 方法重新合成所产生的抗体

片段。也包括多克隆抗体、单克隆抗体、嵌合式抗体、人类化抗体或单链抗体。抗体的“Fc”部分是指免疫球蛋白的重链部分，包括一个或更多重链恒定区、CH<sub>1</sub>、CH<sub>2</sub>和CH<sub>3</sub>，但不包括重链的可变区。

当提到蛋白质或肽时，“特异性(或选择性)结合”于抗体或“特异性(或选择性)免疫反应”是指在蛋白质或其它生物制品的异种群中确定蛋白质存在的结合反应。因此，在指定的免疫分析条件下，特定抗体结合特定蛋白质至少两倍于背景值，并且不会与存在样本中的其它蛋白质大量实质地结合。在这些条件下特异性地结合抗体需要选择对特定蛋白质具有特异性的抗体。例如，来自如大鼠、小鼠或人类的特定物种所产生标记”X”的多克隆抗体，可选择地为以仅特异性地与标记X，而不与其它除标记X的多形态变化体和等位基因外的蛋白质免疫反应的多克隆抗体。选择可通过减去与来自其它物种的标记X交叉反应的抗体来完成。不同的免疫分析形式可用于选择与特定蛋白质特异性免疫反应的抗体。例如，固相ELISA免疫分析为常规用于选择与蛋白质特异性免疫反应的抗体(参阅如Harlow & Lane, *Antibodies, A Laboratory Manual* (1988), 可用于决定特异性免疫反应的免疫分析形式和条件的说明)。通常特异性或选择性的反应产生至少两倍于背景信号或干扰，并且更典型地为大于10-100倍的背景值。

“处理受测对象的治疗”是指临床医生或医生后续为确定前列腺癌状态的行为。举例来说，若本发明方法的结果为不确定或有必要进行状态的确认时，医生可安排更多测试。另外，若状态表明手术为合适的，医生可为病患进行手术排期。同样地，若状态为阴性，例如为晚期前列腺癌，或状态为急性，则没有进一步的行动可被保证。此外，若结果显示治疗成功，则不必要进一步处理。

## 2. 前列腺癌的生物标记

### 2.1. 生物标记

本发明提供以多肽为基础的生物标记PCI，其差异性地存在于患有前列腺癌的受测对象中，尤其是前列腺癌相对于正常状态(非前列腺癌)。生物标记以分子量、通过质谱分析测得的质荷比、在飞行时间质谱分析的质谱峰形、以及对吸附剂表面的结合特性为特征。这些特征

提供确定特定被检测的生物分子是否为本发明生物标记的方法。这些特征显示生物分子的固有特性，但并非作为区分生物分子的局限。

生物标记通过使用来自 CIPHERGEN Biosystems, Inc. (Fremont, CA) (“CIPHERGEN”) 的 ProteinChip 数组以 SELDI 技术来发现。搜集的血清样本收集自被诊断为前列腺癌的受测对象及被诊断为正常的受测对象。这些样本通过阴离子交换层析进行分馏。分馏的样本涂覆于 SELDI 生物芯片上，并且样本中多肽的图谱经由 CIPHERGEN PBSII 质谱仪以飞行时间质谱分析产生。所得到的图谱以 CIPHERGEN Express™ Data Manager Software 和 Biomarker Wizard，以及取自 CIPHERGEN Biosystems, Inc. 的 Biomarker Pattern Software 分析。每个组的质谱通过散布图分析。以 Mann-Whitney 测试分析比较前列腺癌与控制组在散布图中的每一个蛋白质群集，并且选取于两组中具有显著差异 ( $p < 0.0001$ ) 的蛋白质。此方法更详细描述于下文实施例的段落。

生物标记因此被发现为 PCI。“ProteinChip®分析” 栏是指发现生物标记的层析分馏，生物标记结合的生物芯片类型以及清洗条件。

表 1

标记	P-值	前列腺癌中的上或下 调控	ProteinChip®分析
PCI (3890 DA)	<0.0001	上	CM10

本发明的生物标记以通过质谱分析决定的质荷比为特征。质荷比是通过 CIPHERGEN Biosystems, Inc. PBS II 质谱仪上所产生的质谱确定。此仪器具有约正/负 0.15 百分比的质量精确度。另外，此仪器具有约 400 至 1000 m/dm 的质量分辨率，其中 m 为质量和 dm 为于 0.5 峰高处的质谱峰宽。生物标记的质荷比使用 Biomarker Wizard™ 软件 (CIPHERGEN Biosystems, Inc.) 确定。Biomarker Wizard 通过群集所有由 PBS II 确定的分析光谱中相同峰的质荷比，取群集内极大与极小的质荷比并除以二，并指定一个质荷比给所述生物标记。因此，所提供的质量反映这些说明。

本发明的生物标记进一步以其在飞行时间质谱分析中的质谱峰的形状为特征。图 8 显示代表生物标记的质谱谱峰。

本发明的生物标记进一步以其在层析表面的结合性质为特征。大多数生物标记在以 100 mM pH 4 的醋酸钠清洗后结合于阳离子交换吸附剂(例如 CIPHERGEN® WCX ProteinChip®数组)。

由于本发明的生物标记通过质荷比、结合性质和图谱形状为特征,他们可通过质谱分析检测,而不需知道其特定特性。然而,若需要,未确定特性的生物标记可通过如确定多肽的氨基酸序列来识别。举例来说,生物标记可以一些如胰岛素或 V8 蛋白酶的酶绘制肽图,以及消化片段的分子量可用来搜索数据库的序列,所述的序列与通过不同酶所产生消化片段的分子量相匹配。可选择地,蛋白质生物标记可使用串联 MS 技术测序。在这个方法中,蛋白质通过例如凝胶电泳来分离。含有生物标记的带被切下,并使蛋白质通过蛋白酶消化。特定的蛋白质片段通过第一质谱仪分离。此片段随后经过碰撞诱导冷却,将肽片段化并产生多肽阶梯。多肽阶梯然后通过串联 MS 的第二质谱仪分析。以多肽阶梯成分的质量差异识别序列中的氨基酸。完整的蛋白质可以此方式测序,或序列片段可通过数据库挖掘寻找相同的候选序列。

检测生物标记的优选生物来源是血清。然而,在其它的具体实施例中,可于血浆或尿液中侦测生物标记。

本发明的生物标记是多肽。因此,本发明提供分离形式的多肽。生物标记可从如尿液或血清的生物体液中分离。其可通过任何技艺领域所已知基于质量和结合特性的方法分离。例如,将包括多肽的样本进行如本文所描述的层析分馏,并进一步通过如丙烯酰胺凝胶电泳进行分离。生物标记特性的知识也允许其通过免疫亲和层析进行分离。

### 3. 生物标记与蛋白质的不同形式

蛋白质通常以复数不同的形式存在于样本中。这些形式可由翻译前、后修饰的其中一者或两者产生。翻译前的修饰形式包括等位基因变体、剪接变体和 RNA 编辑形式。翻译后的修饰形式包括由蛋白质裂解(例如信号序列的裂解或母蛋白质的片段)、糖基化、磷酸化、脂化、氧化、甲基化、半胱氨酸酸化、磺化及乙酰化产生的形式。

在检测与测定样本中的蛋白质时,区分蛋白质不同形式的能力取决于差异的性质以及所使用检测或测定的方法。例如使用单株抗体的

免疫分析将检测所有含有抗原决定簇(eptiope)的蛋白质形式而无法进行区分。然而,利用二个抗体直接针对蛋白质上不同抗原决定簇的双抗体夹心免疫分析将会检测所有包含此二种抗原决定簇的蛋白质形式,而不会检测仅含有其中一种抗原决定簇的形式。

在诊断分析中,当使用特定方法所检测到的形式与任何特定的形式同样为好的生物标记时,则无法区别蛋白质不同形式所造成的影响将很小。然而,当蛋白质的特定形式(或特定形式的亚群)比通过特定方法共同检测到的不同形式的集合体为较好的生物标记时,此分析能力将受到影响。在此种情况下,采用一种能区别蛋白质不同形式,并且能特异地检测与测定所需蛋白质的单一或多种形式的分析是有用的。区别分析物的不同形式或特异地检测分析物的特定形式是指“分辨”所述分析物。

质谱分析是一种分辨蛋白质不同形式特别有效的方法,因为不同的形式通常具有可由质谱分析分辨的不同质量。因此,若蛋白质的一种形式相对其他形式对一种疾病为较好的生物标记时,质谱分析可特异地检测和测定有效的形式,而传统免疫分析却无法区分各种形式,并且无法特异地检测有效的生物标记。

一种有用的方法结合质谱分析与免疫分析。首先,使用一种生物特异性捕捉剂(例如能辨识生物标记及它的其它形式的抗体、适体或亲和体)捕捉感兴趣的生物标记。优选地,生物特异性捕捉剂结合至如珠、盘、膜或芯片的固相。在洗去未结合的物质后,通过质谱分析检测和/或测定所捕捉的分析物(此方法也会捕捉到结合至蛋白质、或可由抗体和其本身为生物标记所识别的蛋白质交互反应物)。不同形式的质谱分析可用于检测蛋白质的形式,包括如传统的 MALDI 或 SELDI 与电喷射离子化的激光解析法。

因此,当本文提到检测特定蛋白质或测定特定蛋白质的量时,其意指在有或没有分辨不同形式的蛋白质检测和测定蛋白质。举例来说,“测定 PCI”的步骤包含通过在蛋白质样本中不区分不同蛋白质形式的方法,以及通过从其它形式中区分一些形式或测定蛋白质的特定形式的方法测定 PCI。

#### 4. 前列腺癌生物标记的检测

本发明的生物标记可通过任何适当的方法检测。为达此目的，可使用的检测范例包含光学方法、电化学方法(伏安法和电流测定法技术)、原子力显微镜和如多极共振谱分析的射频方法。光学方法的例子除显微镜外，共轭焦和非共轭焦可检测荧光、冷光、化学冷光、吸光、反射、穿透和双折射或折射率(例如表面等离子共振、椭圆偏光分析、共振镜法、光栅耦合波导法或干涉测量分析)。

在一个具体实施例中，通过生物芯片分析样本。生物芯片一般包括固体基质以及具有供捕捉剂附着(也称吸附剂或亲和剂)的通常为平面的表面。通常，生物芯片的表面包括复数个分别具有捕捉剂结合的可寻址位置。

蛋白质生物芯片为适合捕捉多肽的生物芯片。许多蛋白质生物芯片已记述于本技术领域，其中包括如 CIPHERGEN Biosystems, Inc. (Fremont, CA)、Zyomyx (Hayward, CA)、Invitrogen (Carlsbad, CA)、Biacore (Uppsala, Sweden)和 Procognia (Berkshire, UK)等公司制造的蛋白质生物芯片。这些蛋白质生物芯片的实施例揭露于以下的专利或公开的专利申请中：美国第 6,255,047 号专利(Hutchens 和 Yip)、美国第 6,537,749 号专利(Kuimelis 和 Wagner)、美国第 6,329,209 号专利(Wagner et al.)、PCT 第 WO 00/56934 号国际公开案(Englert et al.)、PCT 第 WO 03/048768 号国际公开案(Boutell et al.)以及美国第 5,242,828 号专利(Bergstrom et al.)。

##### 4.1. 通过质谱分析检测

在一个优选的具体实施例中，本发明的生物标记通过质谱分析检测，为使用质谱仪检测气相离子的方法。质谱仪的实例是飞行时间式、扇形磁场式、四极滤质器式、离子阱式、离子回旋共振式、扇形静电分析式及这些的混合。

在更优选的具体实施例中，质谱仪为激光解吸/离子化质谱仪。在激光解吸/离子化质谱分析中，分析物置于质谱分析探针的表面，探针为适合接合于质谱仪的探针接口，并且把分析物呈现于离子化能量以离子化及引入质谱仪内的装置。激光解吸质谱仪利用典型为紫外线激

光但也可作为红外线激光的激光能，将分析物从表面释出，并且将其挥发和离子化，使其可为质谱仪的离子光学仪器所利用。

## SELDI

本发明所使用优选的质谱分析技术为“表面增强激光解吸和离子化”或“SELDI”，如美国第 5,719,060 号与第 6,255,047 号皆属于 Hutchens 与 Yip 的专利所记载。这是指解析/离子化气相离子谱分析（例如质谱分析）的方法，其中分析物（于此指一种或多种的生物标记）经捕捉于 SELDI 质谱分析探针的表面。有多种形式的 SELDI。

一种形式的 SELDI 被称为“亲和捕捉质谱分析”，也被称为“表面增强亲和捕捉”或“SEAC”。此方法包括使用在探针的表面具有通过物质与分析物间非共价亲和作用（吸附）捕捉分析物的探针。此物质不同地称为“吸附物”、“捕捉剂”、“亲和剂”或者“结合基团”。这些探针可为“亲和捕捉探针”以及具有“吸附表面”。捕捉剂可为任何能结合分析物的物质。捕捉剂通过物理吸附或化学吸附而黏附于探针表面。在特定的具体实施例中，探针具有已附着于表面的捕捉剂。在其它的具体实施例中，探针经预活化，并且包含能结合捕捉剂的反应基团，例如通过形成共价或配位键的反应。环氧化物和酰基咪唑为可用于共价结合如抗体或细胞受体的多肽捕捉剂的反应基团。次氨基三乙酸和亚氨基二乙酸是可用于作为与金属离子结合的螯合剂的反应基团，所述金属离子非共价地与含组氨酸的肽反应。吸附剂一般分为层析吸附剂与生物特异性吸附剂。

“层析吸附剂”是指典型使用于层析的吸附材料。层析吸附剂包含如离子交换物质、金属螯合剂（例如次氨基三乙酸和亚氨基二乙酸）、固定化金属螯合剂、疏水性作用吸附剂、亲水性作用吸附剂、染料、简单生物分子（如核苷酸、氨基酸、简单糖和脂肪酸）及混合型吸附剂（如疏水吸引/静电排斥吸附剂）。

“生物特异性吸附剂”是指包括生物分子的吸附剂，其中生物分子如核酸分子（如适体）、多肽、多糖、脂质、类固醇或其偶联物（如糖蛋白、脂蛋白、糖脂、核酸（如 DNA）-蛋白质偶联物）。在特定的实施例中，生物特异性吸附剂可为大分子结构，如多蛋白复合物、生物膜或

病毒。生物特异性吸附剂的实例是抗体、受体蛋白质和核酸。生物特异性吸附剂典型地较层析吸附剂对目标分析物具有更高的特异性。进一步用于 SELDI 的吸附剂实例可于美国第 6, 225, 047 号专利中查到。

“生物选择性吸附剂”是指与分析物具有至少  $10^{-8}$ M 亲和性的吸附剂。

由 CIPHERGEN Biosystems, Inc. 制造的蛋白质生物芯片包括具有层析或生物特异性吸附剂于可寻址的位置附着其上的蛋白质芯片。CIPHERGEN 的 ProteinChip® 数组包含 NP20 (亲水性); H4 和 H50 (疏水性); SAX-2、Q-10 和 LSAX-30 (阴离子交换); WCX-2、CM-10 和 LWCX-30 (阳离子交换); IMAC-3、IMAC-30 和 IMAC-40 (金属螯合); 以及 PS-10、PS-20 (具有醯基咪唑、环氧化物的反应表面) 和 PG-20 (通过醯基咪唑耦合的蛋白质 G)。疏水性蛋白质芯片 (ProteinChip) 数组含有异丙基或壬基苯氧基-聚(乙二醇)甲基丙烯酸酯的官能团。阴离子交换蛋白质芯片 (ProteinChip) 数组含有四级铵的官能团。阳离子交换蛋白质芯片 (ProteinChip) 数组含有羧酸的官能团。固定化金属螯合蛋白质芯片 (ProteinChip) 数组含有通过螯合吸附如铜、镍、锌和镓过渡金属的次氨基三乙酸官能团。预活化蛋白质芯片 (ProteinChip) 数组含有可与蛋白质上的基团共价结合的醯基咪唑或环氧化物的官能团。

这些生物芯片进一步记述于: 美国第 6, 579, 719 号专利 (Hutchens 和 Yip, “Retentate Chromatography,” June 17, 2003)、美国第 6, 897, 072 号专利 (Rich et al., “Probes for a Gas Phase Ion Spectrometer,” May 24, 2005)、美国第 6, 555, 813 号专利 (Beecher et al., “Sample Holder with Hydrophobic Coating for Gas Phase Mass Spectrometer,” April 29, 2003)、美国第 U.S. 2003-0032043 A1 号专利公开案 (Pohl 和 Papanu, “Latex Based Adsorbent Chip, July 16, 2002)、PCT 第 WO 03/040700 号国际公开案 (Um et al., “Hydrophobic Surface Chip,” May 15, 2003)、美国第 US 2003/-0218130 A1 号专利申请公开案 (Boschetti et al., “Biochips With Surfaces Coated With Polysaccharide-Based Hydrogels,” April 14, 2003) 以及美国第 7, 045, 366 号专利 (Huang et al., “Photocrosslinked Hydrogel Blend Surface Coatings,” May 16, 2006)。

一般来说, 具有吸附性表面的探针与样本接触一段时间, 足以允许生物标记或可能存在于样本中的生物标记结合至吸附剂。在反应一段时间后, 清洗基质以除去未结合的物质。可使用任何适合的清洗溶液, 优选地可使用水性溶液。通过调整清洗的力度可控制仍保持结合的分子范围。清洗溶液的洗提特性可取决于如 pH、离子强度、疏水性、混乱度、清洁剂强度和温度。除非探针同时具有 SEAC 与 SEND 的性质(如同本文所述), 接着施加能量吸收分子至具有结合生物标记的基质上。

通过如飞行时间质谱仪的气相离子谱仪检测结合于基质的生物标记。生物标记通过如激光的离子化源而离子化, 产生的离子由离子光学装置收集, 然后通过质量分析器分散和分析通过的离子。检测器随后将检测离子的数据转换为质荷比。生物标记的检测通常将涉及信号强度的检测。因此, 生物标记的量与质量皆可被测定。

另一种形式的 SELDI 为表面增强净解吸(“SEND”), 涉及使用包括化学结合到探针表面(“SEND 探针”)的能量吸收分子的探针。“能量吸收分子”(EAM)一词表示能从激光解吸/离子化源吸收能量, 然后提供与其接触的分析物分子解吸和离子化。EAM 的范畴包括使用于 MALDI 的分子, 通常是指“基质”, 并且以肉桂酸衍生物、芥子酸(SPA)、氰氧-羟基-肉桂酸(CHCA)和二羟基苯甲酸、阿魏酸和羟苯乙酮衍生物作为例示。在特定的具体实施例中, 能量吸收分子并入线性或交联的聚合物中, 如聚甲基丙烯酸酯。举例来说, 组合物可为  $\alpha$ -氰基-4-异丁烯酰基氧基肉桂酸和丙烯酸酯的共聚物。在另一个具体实施例中, 组合物为  $\alpha$ -氰基-4-异丁烯酰基氧基肉桂酸、丙烯酸脂和 3-(三乙氧基)硅烷基丙基甲基丙烯酸酯的共聚物。在另一个具体实施例中, 组合物为  $\alpha$ -氰基-4-异丁烯酰基氧基肉桂酸和十八烷基甲基丙烯酸酯(“C18 SEND”)的共聚物。SEND 进一步记述于美国第 6, 124, 137 号专利及 PCT 第 WO 03/64594 号国际公开案(Kitagawa, “Monomers And Polymers Having Energy Absorbing Moieties Of Use In Desorption/Ionization Of Analytes,” August 7, 2003)。

SEAC/SEND 为捕捉剂和能量吸收分子皆附着于样本表现表面的一种 SELDI 变化形式。因此, SEAC/SEND 探针允许通过亲和捕捉和离子化/解吸捕捉分析物而不需使用外界基质。C18 SEND 生物芯片是一种

SEAC/SEND 的变化形式,包括作为捕捉剂的 C18 基团和作为能量吸收基团的 CHCA 基团。

另一种变化形式的 SELDI 称为表面增强光不稳定附着和释放 (“SEPAR”),涉及使用具有基团的探针,所述基团附着于表面,并且可共价结合分析物,随后通过暴露于如激光光束的光打断基团中光不稳定键而释放分析物(参阅美国第 5,719,060 号专利)。依照本发明,SEPAR 和其它形式的 SELDI 很快地适合于检测生物标记或生物标记的形态。

### 其它质谱分析方法

在另一种质谱分析中,生物标记可先被具有结合生物标记层析性质的层析树脂捕捉。在本实施例中,可包括多种方法。例如,可于如 CM Ceramic HyperD F 树脂的阳离子交换树脂上捕捉生物标记,清洗树脂,洗提生物标记,并以 MALDI 检测。另外,此方法可在利用阳离子交换树脂前先通过阴离子交换树脂分馏样本。在另外的方法,可通过阴离子交换树脂分馏,并以 MALDI 直接检测。在又一个方法中,可于含有与生物标记结合的抗体的免疫层析树脂上捕捉生物标记,清洗树脂以除去未结合的物质,从树脂上洗提生物标记,并以 MALDI 或 SELDI 检测洗提出的生物标记。

### 数据分析

通过飞行时间质谱分析进行分析物的分析,产生飞行时间光谱。最终分析的飞行时间光谱通常并不代表样本对离子化能单一脉冲的信号,而是许多脉冲的信号的和。这将降低噪声并增加动态范围。随后对飞行时间的数据进行数据处理。在 Ciphergen 的 ProteinChip®软件中,数据处理通常包含 TOF 至 M/Z 的转换以产生质谱、基线的减运算以消除仪器偏差及高频噪声的过滤以降低高频噪声。

藉由生物标记的解吸和检测所产生的数据可使用可编程的数字计算机分析。计算机程序分析数据指明被检测生物标记的数量,以及可选择地每一个被检测生物标记的信号强度和测定的分子量。数据分析的步骤可包括确定生物标记信号强度以及移除偏离预定统计分布的数

据。例如，可通过计算相对一些参考物的每一个谱峰高度标准化所观测到的峰值。参考物可为通过装置或如能量吸收物质的化学品所产生的背景噪声，并在标度中被设为零。

计算机可转换结果的数据为不同显示的形式。可显示的标准光谱，但在有用的形式中从光谱的角度来看仅保留峰高和质量的数据，产生更清晰的画面，使分子量几乎相等的生物标记更容易被观测到。在另一个有用的形式中，比较二个或更多个光谱，方便地突显独特的生物标记以及于样本中上或下调控的生物标记。利用任何一种所述的形式，将可立即确定在样本中是否存在特定的生物标记。

分析一般包括识别表示来自分析物的信号的谱峰。谱峰的选择可通过直接观察完成，但也可使用软件，如 CIPHERGEN 的 ProteinChip® 软件套件的部分可自动检测谱峰。一般而言，软件的运作是通过识别信号-噪声比高于所选阈值的信号，以及在谱峰信号的中心标记谱峰的质量。在一个有效的应用中，比较许多光谱，以识别在选定的质谱百分比中所呈现相同的谱峰。一种变化形式的软件群集所有在指定的质量范围内出现在不同光谱的谱峰，并在临近质量 (M/Z) 群集的中点为所有谱峰指派一个质量 (M/Z)。

用于分析数据的软件可包含应用算法于信号分析的编码，以确定在信号中表示的谱峰信号是否对应于本发明的生物标记。软件也可将关于所观察生物标记谱峰的数据进行分类树或 ANN 分析处理，以确定于检查中指示特定临床参数状态的生物标记的谱峰或多个生物标记谱峰的组合是否存在。数据的分析对于不论直接或间接从样本的质谱分析所获得的多种参数可能是「关键的」。这些参数包含但不局限于一个或多个谱峰的存在或不存在、一个谱峰或一群谱峰的形状、一个或多个谱峰的高度、一个或多个谱峰的高度对数以及其它峰高数据的算法运算。

### SELDI 检测前列腺癌生物标记的一般操作步骤

优选用于检测本发明生物标记的操作步骤如下文描述。用于测试的样本如尿液，优选地于 SELDI 分析前经过预分馏。这使样本单纯化并提高敏感度。优选预分馏的方法包括使样本接触阴离子交换层析材

料, 例如 Q HyperD (BioSeptra, SA)。结合物质随后通过使用 pH 9、pH7、pH5 和 pH4 的缓冲液以分步 pH 洗提(参阅实施例 1-缓冲液列表)(生物标记被洗提的馏分也表示于表 1 中)。多种包含生物标记的分馏被收集。

被测试的样本(优选地经预分馏)随后与包括阳离子交换吸附剂(优选为 WCX ProteinChip 数组(Ciphergen Biosystems, Inc.))或 IMAC 吸附剂(优选为 IMAC3 ProteinChip 数组(Ciphergen Biosystems, Inc.))的亲合捕捉探针接触, 再一次如表 1 所示。以缓冲液清洗探针, 洗去未结合的分子而保留生物标记。适合每一个生物标记的清洗液于表 1 中所示。通过激光解吸/离子化质谱分析检测生物标记。

另外, 若可取得辨识生物标记的抗体, 例如  $\beta$  2 微球蛋白、半胱氨酸蛋白酶、转铁蛋白、转甲状腺素蛋白或血清蛋白, 可被附着于探针的表面, 如预活化的 PS10 或 PS20 ProteinChip 数组(Ciphergen Biosystems, Inc.)。这些抗体能从样本中将生物标记捕捉到探针的表面。然后, 可通过如激光解吸/离子化质谱分析检测生物标记。

#### 4.2. 通过免疫分析进行检测

在另一个具体的实施例中, 本发明的生物标记通过免疫分析测定。免疫分析需要如抗体的生物特异性捕捉剂捕捉生物标记。抗体可由本技术领域所公知的方法制造, 例如用生物标记免疫动物。生物标记可基于其结合特征从样本中分离。另外, 若已知多肽生物标记的氨基酸序列, 则多肽可经由合成, 并通过本技术领域所公知的方法产生抗体。

本发明可考虑传统的免疫分析, 包括如包含 ELISA 或基于荧光的免疫分析的双抗体夹心免疫分析, 以及其它酵素免疫分析。浊度分析是可于液相实施的分析法, 其中抗体是在溶液中。而抗体与抗原的结合使所测定的吸光度产生变化。在基于 SELDI 的免疫分析中, 生物标记的生物特异性捕捉剂附着于 MS 探针的表面, 例如预活化蛋白质芯片(ProteinChip)数组。生物标记随后通过反应剂特异性地被捕捉于生物芯片上, 并且被捕捉的生物标记通过质谱分析检测。

#### 5. 确定受测对象前列腺癌的状态

### 5.1. 单一标记

本发明的生物标记可用于评估诊断测试中受测对象的前列腺癌状态，例如诊断前列腺癌。“前列腺癌状态”一词包含任何可区别疾病的表现，包括非疾病。举例来说，疾病的状态包括但不局限于疾病的存在或不存在(例如前列腺癌相对于非前列腺癌)、患病风险、疾病的阶段、疾病的进程(例如随时间的疾病进展或疾病缓解)和疾病治疗的效果或反应。基于这个状态，可进一步表示包含额外诊断测试或治疗程序或疗法的程序。

正确预测状态的诊断测试力度通常通过分析敏感度、分析的特异性或接收工作特征(“ROC”)曲线下的面积而测定。敏感度为通过测试预测为阳性的真阳性百分比，而特异性为通过测试预测为阴性的真阴性百分比。ROC曲线提供测试的敏感度作为1-特异性的函数。ROC曲线下的区域越大，测试的预测值则越有效力。其它利用测试的有用测定为阳性的预测值和阴性的预测值。阳性的预测值为测试为阳性且确实为阳性的人的百分比。阴性的预测值为测试为阴性且确实为阴性的人的百分比。

本发明的生物标记显示至少  $p \leq 0.05$ 、 $p \leq 10^{-2}$ 、 $p \leq 10^{-3}$ 、 $p \leq 10^{-4}$  或  $p \leq 10^{-5}$  的不同前列腺癌状态的统计差异。单独或组合使用这些生物标记进行诊断测试，显示至少 75%、至少 80%、至少 85%、至少 90%、至少 95%、至少 98% 及约 100% 的敏感度和特异性。

PCI 差异性地存在于前列腺癌中，因此，每一个可个别地有助于确定前列腺癌的状态。此方法首先包括使用本文所述的方法在受测对象的样本中测定所选定的生物标记，例如在 SELDI 生物芯片上捕捉后以质谱分析检测，其次是将测定与诊断量或区分阳性前列腺癌状态与阴性前列腺癌状态的切点进行比较。诊断量代表测定的生物标记量高于或低于受测对象被分类为患有特定前列腺癌的状态。举例来说，若生物标记于前列腺癌中相较于正常为上调控，则测定量高于诊断的切点提供前列腺癌的诊断。另外，若生物标记于前列腺癌中为下调控，则测定的量低于诊断的切点提供前列腺癌的诊断。如同本技术领域所已知，通过调整分析中特定诊断的切点，可依照诊断者的偏好增强诊断分析的敏感度或特异性。特定诊断的切点可通过如在来自多个含有不

同前列腺癌状态的受测对象具统计上显著意义数量的样本中测定生物标记的含量，如同本文所实施的，可依诊断者所要求的特异性与敏感度水平取得切点。

### 5.2. 确定患病风险

在一个具体实施例中，本发明提供确定受测对象患病风险的方法。生物标记的含量或型态是不同风险状况的特征，例如高、中或低。患病的风险通过测定相关的生物标记或复数生物标记确定，随后可进行分类算法或与特定风险水平相关的生物标记的参考量和/或型态进行比较。

### 5.3. 确定疾病的阶段

在一个具体实施例中，本发明提供确定受测对象疾病阶段的方法。疾病的每一个阶段具有生物标记的特征性含量或一组生物标记(一种模式)的相对含量。疾病的阶段通过测定相关的生物标记或复数生物标记确定，随后可进行分类演算或与特定阶段相关的生物标记的参考量和/或型态进行比较。

### 5.4. 确定疾病的进程(进展/缓解)

在一个具体实施例中，本发明提供确定受测对象疾病进程的方法。疾病进程是指疾病状态随时间的变化，包含疾病的进展(恶化)和疾病的转好(改善)。生物标记的含量或相对含量(例如型态)随着时间而变化。举例来说，生物标记 PCI 随着疾病增加，而生物标记” Z” 却在疾病中减少。因此，这些生物标记的趋势，不论是随着时间增加或降低而朝着疾病或是非疾病的方向，皆表示疾病的进程。因此，此方法涉及在至少二个不同的时间点测定受测对象的一种或多种生物标记，例如第一时间和第二时间，以及若有含量的变化时，则比较含量的变化。疾病的进程基于这些比较而确定。

### 5.5. 受测对象的处理

在特定鉴别前列腺癌状态的方法具体实施例中，此方法进一步包

括基于状态而处理受测对象的治疗。这种处理包含医生或临床医生在确定前列腺癌状态后的行为。例如，若医生作出前列腺癌的诊断，随后可进行如开处方或施加治疗剂的特定治疗疗法。另外，非前列腺癌的诊断或非前列腺癌可伴随进一步的测试以确定病人所可能遭受的特定疾病。同样地，若诊断测试作出对前列腺癌的状态不确定的结果，则可进行进一步的测试。

本发明的其它具体实施例为有关于对如技术人员、医生或病患传达检测结果或诊断或二者。在特定的具体实施例中，计算机用于将检测结果或诊断或二者传达给有兴趣的一方，如医生及其病患。在一些具体实施例中，检测或检测的结果将在不同于结果或诊断被传达的国家或管辖区域中被完成或被分析。

在本发明优选的具体实施例中，基于在测试的受测对象中存在或不存在生物标记进行诊断，在取得诊断后立即传达给受测对象。此诊断可由受测对象的治疗医生传达于受测对象。另外，诊断可通过电子邮件传送或通过电话传达给测试的受测对象。计算机可用于以电子邮件或电话传达此诊断。在特定的具体实施例中，使用电信技术领域的人员所熟知的计算机硬件与软件的组合可产生并自动将含有诊断测试结果的讯息传达于受测对象。保健取向的通信系统实例已描述于美国第 6, 283, 761 号专利，然而，本发明并不局限于使用特定通信系统的方法。在本发明方法的特定具体实施例中，所有或一些方法的步骤，包含样本检测、疾病诊断，以及检测结果或诊断的传达，可在不同的(例如外国的)管辖区域中实行。

#### 5.6. 确定医药药物的疗效

在另一个具体实施例中，本发明提供确定医药药物疗效的方法。这些方法适用于实施药物的临床试验与监控病患对于药物的进程。治疗或临床试验包括在特别疗法中施加药物。此疗法包括药物的单一剂量或经过一段时间药物的多数剂量。医生或临床研究者施加的过程中监控药物对病患或受测对象的影响。若药物对于病症具有医药的影响，则本发明生物标记的含量或相对含量(例如型态或图谱)会朝非疾病图谱的方向变化。因此，可追踪在治疗过程中受测对象体内这些生

物标记含量的进程。因此，此方法包括在接受药物治疗的受测对象体内测定一种或多种生物标记，以及将生物标记的量与受测对象的疾病状态相关联化。一个本方法的具体实施例包括在药物治疗过程中的至少二个不同的时间点测定生物标记的水平，例如第一时间和第二时间，若有变化，则比较生物标记量的变化。举例来说，可在施加药物前后或药物施加期间二个不同的时间点测定生物标记。治疗的效果基于这些比较而确定。若治疗为有效的，生物标记将趋于正常，若治疗为无效的，则生物标记趋向疾病的迹象。若治疗为有效的，生物标记将趋于正常，若治疗为无效的，则生物标记趋向疾病的迹象。

## 6. 产生鉴别前列腺癌状态的分类算法

在一些具体实施例中，使用如“已知样本”的样本产生的光谱(例如质谱或飞行时间光谱)所得到的数据可用来“训练”分类模型。“已知样本”是已预分类的样本。得自光谱和用于形成分类模型的数据称作“训练数据集”。一旦被训练，分类模型可识别由使用未知样本所产生的光谱所得到数据的型态。分类模型可随后用于将未知的样本分类。这将适用于如预测特定生物样本是否与特定生物病症(例如疾病相对于非疾病)相关。

用于形成分类模型的训练数据集可包括原始数据或预处理数据。在一些具体实施例中，原始数据可直接由飞行时间光谱或质谱中得到，并且其随后视需要地如前所述进行“预处理”。

可使用任何适合的统计分类(或“学习”)方法形成分类模型，此方法试图基于存在于数据中的目标参数将数据本体分离归类。分类方法可被监督或未被监督。监督或未监督的分类过程的实例记述于Jain, "Statistical Pattern Recognition: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, January 2000, 其所教导的内容并入本文作为参考。

在监督的分类中，包含已知类别实例的训练数据被呈现于学习机制，其学习一组或多组定义已知类别中每一个的关系。新数据可随后用于学习机制，随后使用经学习的关系分类新数据。监督分类方法的实例包含线性回归方法(例如多元线性回归(MLR)、偏最小二乘(PLS)回

归及主成分回归(PCR))、二进制判定树(例如递归分化方法,如 CART-分类和回归树)、如反向传播网络的人工神经网络、辨别分析(例如贝叶斯分类器或费希尔分析)、逻辑分类器及支持向量分析器(支持向量机)。

优选的监督分类方法是递归分化方法。递归分化方法使用递归分化树将从未知样本取得的光谱分类。进一步关于递归分化过程的细节描述于美国第 2002 0138208 A1 号专利申请案, Paulse et al., “Methods for analyzing mass spectra.”。

在其它的具体实施例中,所建构的分类模型可使用无监督的学习方法形成。无监督的分类试图基于在训练数据组中的相似性学习分类,而不用由光谱得到的训练数据组预分类光谱。无监督的学习方法包括群集分析(cluster analysis)。群集分析尝试将数据分成「群集」或群组,其理想应为成分彼此间非常相似,并且与其它群集的成分非常不相似。随后使用一些测量数据项目间距离的距离度量测定相似性,并且群集彼此靠近的数据项目。群集技术包含 MacQueen 的 K-平均算法和 Kohonen 的自组织映图(Self-Organizing Map)算法。

声称用于分类生物信息的的学习算法已描述于如 PCT 第 WO 01/31580 号国际公开案(Barnhill et al., “Methods and devices for indentifying patterns in biological systems and methods of use thereof” )、美国第 2002 0193950 A1 号专利申请案(Gavin et al., “Method or analyzing mass spectra” )、美国第 2003 0004402 A1 号专利申请案(Hitt et al., “Process for discrimination between biological states based on hidden patterns from biological data” )、以及美国第 2003 0055615 A1 号专利申请案(Zhang 和 Zhang, “Systems and methods for processing biological expression data” )。

分类模型可在任何合适的数字计算机上形成和使用。适合的数字计算机包括微型、小型或大型利用任何标准或特殊化操作系统的计算机,例如基于 Unix、Windows™ 或 Linux™ 的操作系统。所使用的数字计算机可物理性地与产生感兴趣光谱的质谱仪分离,或可与质谱仪结合。

依照本发明具体实施例的训练数据组与分类模型可通过由数字计

计算机执行或使用的计算机编码而具体化。计算机编码能储存于任何合适计算机可读的媒体，包含光盘或磁盘、记忆棒、磁带等，并可由任何包括 C、C++、visual basic 等的合适计算机编程语言编写。

上述的学习算法适用于发展已发现生物标记的分类算法或发现前列腺癌的新生物标记。因此，分类算法通过提供单独或组合使用的生物标记诊断值(如切点)形成诊断测试的基础。

## 7. 物质的组合物

在另一个态样中，本发明提供基于本发明生物标记的物质组合物。

在另一个具体实施例中，本发明提供本发明生物标记的纯化形式。纯化的生物标记可作为抗原以制备抗体。纯化的生物标记也可作为检测过程中的标准物。在此所使用的”纯化的生物标记”是由其它蛋白质和肽，和/或发现所述生物标记的生物样本的其它物质中单离出的生物标记。可使用任何本领域所已知的技术纯化生物标记，包括但不限于机械分离(例如离心)、硫酸铵沉淀、透析(包括大小排除透析)、大小排除层析、亲和层析、阴离子交换层析、阳离子交换层析及金属螯合层析。这些方法可以任何适当的规模进行，例如于层析柱内或生物芯片上。

在另一个具体实施例中，本发明提供生物特异性的捕捉剂，可选择性为纯化的形式，特异性地与本发明的生物标记结合。在一个具体实施例中，生物特异性的捕捉剂是抗体。这些组合物适用于在如诊断的检测分析中检测生物标记。

在另一个具体实施例中，本发明提供包括结合本发明生物标记的生物特异性捕捉剂的物品，其中捕捉剂结合于固相上。举例来说，本发明可考虑为包括衍生具有生物特异性捕捉剂的珠、芯片、膜、石料或微孔板的装置。此等物品用于生物标记的检测分析。

另一个本发明的态样提供包括结合至本发明生物标记如抗体的生物特异性捕捉剂的组合物，所述组合物可选择地为纯化的形式。这些组合物适用于纯化生物标记或用于检测生物标记的分析。

在另一个具体实施例中，本发明提供包括固体基质的物品，其中如层析吸附剂或生物特异性捕捉剂的吸附剂附着于固体基质上，本发

明的生物标记进一步结合于所述吸附剂上。在一个具体实施例中，物品为用于质谱分析的生物芯片或探针，例如 SELDI 探针。这些物品用于纯化或检测生物标记。

## 8. 检测前列腺癌生物标记的试剂盒

在另一个态样中，本发明提供鉴别前列腺癌状态的试剂盒，用于检测根据本发明的生物标记。在一个具体实施例中，试剂盒包括固相载体，例如芯片、微孔板或珠或具有捕捉剂附着其上的树脂，其中捕捉剂结合本发明的生物标记。因此，举例来说，本发明的试剂盒可包括用于 SELDI 的质谱分析探针，例如 ProteinChip® 数组。在生物特异性捕捉剂的例子中，试剂盒可包括具有反应表面的固相载体，以及包括生物特异性捕捉剂的容器。

试剂盒也可包括清洗溶液或配制清洗溶液的指导说明，其中捕捉剂和清洗溶液的组合允许在固相载体上捕捉一种或多种生物标记供后续通过如质谱分析的检测。试剂盒可包含多种的吸附剂，每一种表现于不同的固相载体上。

在进一步的具体实施例中，试剂盒可包括卷标或分离夹页形式适合操作参数的指导说明。例如，指导说明可告知顾客关于如何收集样本、如何清洗探针或被检测的特定生物探针。

在又另一个具体实施例中，试剂盒可包括一个或多个具有作为校正用标准物的生物标记样本的容器。

## 9. 前列腺癌生物标记于筛选分析中的使用以及治疗前列腺癌的方法

本发明的方法也有其它的应用。例如，生物标记可用于筛选调节生物标记于活体外或活体内表达的化合物，此化合物随后可用于治疗或预防病人的前列腺癌。在另一个实例中，生物标记可用于监控对治疗前列腺癌的反应。在又另一个实例中，生物标记可用于遗传研究，以确定受测对象是否具罹患前列腺癌的风险。

因此，举例来说，本发明的试剂盒可包含具有疏水功能的固体基质，如蛋白质生物芯片(如 CIPHERGEN H50 蛋白质芯片(ProteinChip) 数组，如蛋白质芯片(ProteinChip) 数组)，和清洗基质的醋酸钠缓冲

液，以及提供于芯片上测定本发明生物标记和使用测定诊断前列腺癌的指导说明。

通过识别与一种或多种表 I 所列的生物标记反应的化合物，初步筛选出适用于治疗测试的化合物。经由实例说明，筛选可包含重组表达表 I 所列的生物标记、纯化此生物标记，以及附着此生物标记于基质上。测试化合物随后与基质通常于水性条件下接触，测试化合物与生物标记间的相互作用通过如测量洗提率作为盐浓度的函数。特定蛋白质可辨识并裂解表 I 的生物标记，在这种情况下，蛋白质可通过在标准分析中监控一种或多种生物标记的消化而检测，例如通过蛋白质凝胶电泳。

在一个相关的具体实施例中，可测定测试化合物抑制表 I 中生物标记活性的能力。熟习本领域的技术人员将知道用于测定特定生物标记活性的技术将取决于生物标记的功能和性质而有所不同。举例来说，若可取得适当的基质，并且基质的浓度或反应产物出现可容易测定时，可测试生物标记的酶活性。通过测定测试化合物存在或不存在的催化速率，可确定潜在治疗测试化合物抑制或增强所指定生物标记活性的能力。测试化合物干扰表 I 生物标记的其中之一非酶(例如结构的)功能或活性的能力也可被测定。举例来说，可于测试化合物存在或不存在的条件下通过光谱分析监控包含表 I 生物标记的其中之一多蛋白复合物的自组装。另外，若生物标记是转录的非酶增强子，则可通过在活体内或活体外测试化合物存在或不存在时测定生物标记依赖转录的水平而识别干扰生物标记增强转录的能力的测试化合物。

可调控表 I 生物标记活性的测试化合物可施予患有或有发展为前列腺癌或其它癌症风险的病人。举例来说，若于活体内特定生物标记的活性防止前列腺癌的蛋白质累积，则施予增加特定生物标记活性的测试化合物可降低病患罹患前列腺癌的风险。相反地，若生物标记增强的活性至少部分导致前列腺癌的起始，则施予降低特定生物标记活性的测试化合物可降低罹患前列腺癌的风险。

在另一种态样中，本发明提供识别适于治疗如前列腺癌病症的化合物的方法，所述前列腺癌与修饰形式 PCI 升高的水平相关。举例来说，在一个具体实施例中，催化将完整长度的 PCI 裂解形成截断形式

的 PCI 的化合物可以细胞萃取物或表达库进行筛选。在此筛选分析的具体实施例中，可通过附着荧光物于 PCI 而检测 PCI 的裂解，其中荧光物于 PCI 未裂解时保持淬灭，而于蛋白裂解时则发出荧光。另外，完整长度的 PCI 变体经修饰，使得在氨基酸 X 和 Y 间的酰胺键不可裂解，可将其用于选择性结合或「困住」可于活体内此位置裂解完整长度 PCI 的细胞蛋白酶。筛选和识别蛋白酶及其目标物的方法已于科学文献中完整地记录，例如于 Lopez-Otin et al. (Nature Reviews, 3:509-519(2002))。

在又一个具体实施例中，本发明提供治疗或降低疾病进程或可能性的方法，例如与截断的 PCI 水平增加相关的前列腺癌。举例来说，于识别一种或多种可裂解完整长度 PCI 的蛋白质后，可于组合库筛选抑制已识别蛋白质裂解活性的化合物。于化学库筛选化合物的方法为本领域所公知的技术。参阅如 Lopez-Otin et al. (2002)。另外，可基于 PCI 的结构明智地设计抑制性化合物。

N 端截断的 PCI 被认为减少 PCI 的蛋白酶抑制活性。参阅如 Abrahamson et al. (Biochem. J. 273:621-626(1991))。因此，给予具完整 PCI 功能的截断 PCI 的化合物可能可用于治疗如前列腺癌的病征，因前列腺癌与截断形式的 PCI 相关。因此，在进一步的具体实施例中，本发明提供识别增强截断 PCI 对其目标蛋白酶亲和性的化合物的方法。举例来说，化合物可经由给予截断的 PCI 完成长度 PCI 蛋白酶抑制活性的能力进行筛选。能调控 PCI 的抑制活性或与 PCI 反应的分子活性的测试化合物，可随后于活体内测试其减缓或停止受测对象中前列腺癌发展的能力。

在临床阶段，筛选测试化合物包含在受测对象接受测试化合物前与后从受测对象取得样本。样本中表 I 所列生物标记的水平可被测定与分析，以确定在接触测试化合物后生物标记的水平是否变化。样本可通过本文所述的质谱分析进行分析，或样本可通过任何技术领域所公知的适当方法进行分析。举例来说，表 I 所列生物标记的水平可直接通过使用特异性结合生物标记的放射性或荧光标记抗体的西方墨点法测定。另外，可测定编码一种或多种生物标记的 mRNA 的水平变化，并且将其与施予受测对象所给定的测试化合物相关联。在进一步的

具体实施例中，一种或多种生物标记表达水平的变化可通过活体外的方法与材料测得。举例来说，表达或能表达一种或多种表 I 生物标记的人类组织培养细胞可与测试化合物接触。经测试化合物治疗的受测对象将被常规地检查任何治疗所导致的生理影响。特别地，测试化合物将以其于受测对象中降低疾病可能性的能力进行评估。另外，若将测试化合物施予已诊断为前列腺癌的受测对象，则测试化合物将以其减缓或停止疾病进程的能力进行筛选。

## 10. 实施例

### 10.1. 实施例 1. 前列腺癌生物标记的发现

可以被理解的是，文中所述的实施例和具体实施例仅为例示的目的，本领域技术人员可经建议进行不同的修饰或变化，并且包含于本发明的精神和范围内以及权利要求的范畴。所有的公开物、专利和专利申请都在此并入整体作为参考。

#### 实施例 1

##### 蛋白 C 抑制剂的识别

3890 Da 的蛋白质通过 CM10 数组在 Q 分馏 6 中被检测。所提供的代表性图谱通过公知已识别的蛋白质内部再度校正。目标蛋白质的平均分子量(来自 7 个光谱)经计算为 3890.23 Da(图 1)。

3890 Da 的蛋白质存在于商业血清总库(图 1, 光谱“C”)。重要地，相比于临床样本，商业血清几乎没有血小板因子 4，其 2 价离子相当精确地与 3890 Da 的蛋白质重叠。1 ml 商业血清在变性的条件下通过阴离子交换层析(Q HyperD F)进行分馏。使用 NP20 数组的有机分馏图谱确认 3890 Da 蛋白质谱峰的存在(图 2)。馏分 6 进一步通过反相层析分馏。以 40% ACN 将 3890 Da 的蛋白质从 RPC 珠上洗提(图 3)。在得到的制备物中无 PF4 的存在确认 3890 Da 蛋白质谱峰是确实存在的蛋白质而非 2 价离子(尽管我们不能排除在图谱研究中 2 价的 PF4 离子可贡献 3890 Da 的蛋白质谱峰强度)。

蛋氨酸在中性碱性的 pH 下为稳定的。然而，在酸性的 pH 下，则会受到显著的氧化(CM10 和 H50 数组图谱、反相层析，以及在 ACN 和

TFA 存在下膜分馏的切点等)。来自 Q 馏分 6 的 3890 Da 蛋白质在富集的过程中可观察到逐渐氧化(图 3 和 4), 表示候选的生物标记含有蛋氨酸。

SDS-PAGE 纯化和 Q-Fr. 6 3890 Da 谱峰的直接测序对于来自 1 ml 的血清收集量都是不可行的(图 3)。我对于起始物质的最低估计为 20 ml 的参考血清或临床样本。

在低于 pH3 以下没有蛋白质可保持特异性地结合于阴离子交换树脂(最酸的理论蛋白质 EEEEEEEEE 和 DDDDDDD 分别具有 pI 3.47 和 3.08)。从过去的计画共同观测为任何在馏分 6 中的蛋白质也可于其它与蛋白质 pI 匹配的 Q 馏分中被发现(并且在大多数情况下为很大量)。因此通过 CM10 数组将 Q 馏分图谱化。的确, 3890 Da 的谱峰可清晰地 Q-流出液中被检测(图 5)。其自具有 40% ACN 的 RPC 珠上洗提, 非常相似于 Q-Fr. 6 “兄弟”(图 6)。Q-FT 3890 Da 的蛋白质与 Q-Fr. 6 3890 Da 的蛋白质相似, 在纯化的过程中逐渐被氧化, 表示其含有蛋氨酸(图 7)。

3890 Da 的蛋白质直接通过 Q-TOF MSMS 分析(图 8、9)。使用 Mascot 搜索引擎直接进行数据库搜索并不成功, 因此, 通过使用蛋白质探测器工具(Protein Prospector tool)手动解释 MSMS 数据。使用最强离子无偏见搜索导致太多的候选物。然而, 特定候选物, 天然存在的血浆蛋白 C 抑制剂的 C 端肽最符合“手动”筛选的标准, 如(1)生理相关性, (2)Met 的数量, (3)活体内产生任何有意义的肽, (4)解释个别离子 m/z 值的错误, (5)基于倾向于片段化的预测位点的离子推断, (6)b 和 y 离子的配对等(图 9-14)。

通过使用全血清和抗蛋白 C 抑制剂抗体的以珠为基础的免役分析直接确认 3890 Da 蛋白质为暂定的候选物。

## 参考文献

1. Suzuki, K. Protein C inhibitor. *Methods Enzymol*, 222: 385-399, 1993.
2. Glasscock, L. N., Rehaul, S. M., Gregory, C. W., Cooper, S. T., Jackson, T. P., Hoffman, M., and Church, F. C. Protein

C inhibitor (plasminogen activator inhibitor-3) expression in the CWR22 prostate cancer xenograft. *Exp Mol Pathol*, 79: 23-32, 2005.

3. Suzuki, K., Deyashiki, Y., Nishioka, J., and Toma, K. Protein C inhibitor: structure and function. *Thromb Haemost*, 61: 337-342, 1989.

4. Suzuki, K., Deyashiki, Y., Nishioka, J., Kurachi, K., Akira, M., Yamamoto, S., and Hashimoto, S. Characterization of a cDNA for human protein C inhibitor. A new member of the plasma serine protease inhibitor superfamily. *J Biol Chem*, 262: 611-616, 1987.

5. Suzuki, K., Nishioka, J., and Hashimoto, S. Protein C inhibitor. Purification from human plasma and characterization. *J Biol Chem*, 258: 163-168, 1983.

6. Cao, Y., Becker, C., Lundwall, A., Christensson, A., Gadaleanu, V., Lilja, H., and Bjartell, A. Expression of protein C inhibitor (PCI) in benign and malignant prostatic tissues. *Prostate*, 57: 196-204, 2003.

7. Clauss, A., Lilja, H., and Lundwall, A. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. *Biochem J*, 368: 233-242, 2002.

8. Strandberg, K., Kjellberg, M., Knebel, R., Lilja, H., and Stenflo, J. A sensitive immunochemical assay for measuring the concentration of the activated protein C-protein C inhibitor complex in plasma: use of a catcher antibody specific for the complexed/cleaved form of the inhibitor. *Thromb Haemost*, 86: 604-610, 2001.

9. Christensson, A. and Lilja, H. Complex formation between protein C inhibitor and prostate-specific antigen in vitro and in human semen. *Eur J Biochem*, 220: 45-53, 1994.

10. Laurell, M., Christensson, A., Abrahamsson, P. A., Stenflo, J., and Lilja, H. Protein C inhibitor in human body fluids. Seminal plasma is rich in inhibitor antigen deriving from cells throughout the male reproductive system. *J Clin Invest*, 89: 1094-1101, 1992.

11. Otlewski, J., Jelen, F., Zakrzewska, M., and Oleksy, A. The many faces of protease-protein inhibitor interaction. *Embo J*, 24: 1303-1310, 2005.

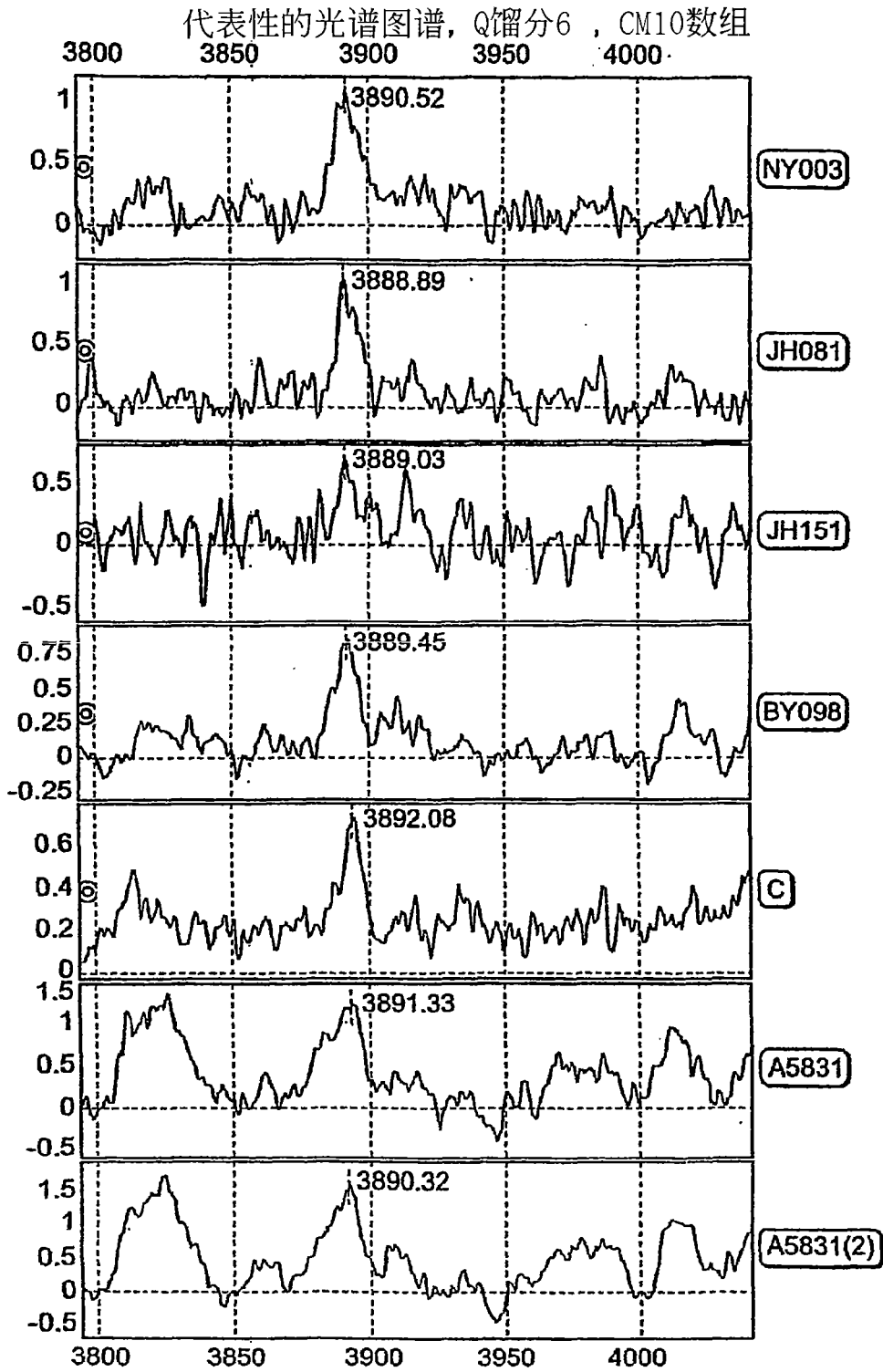


图 1

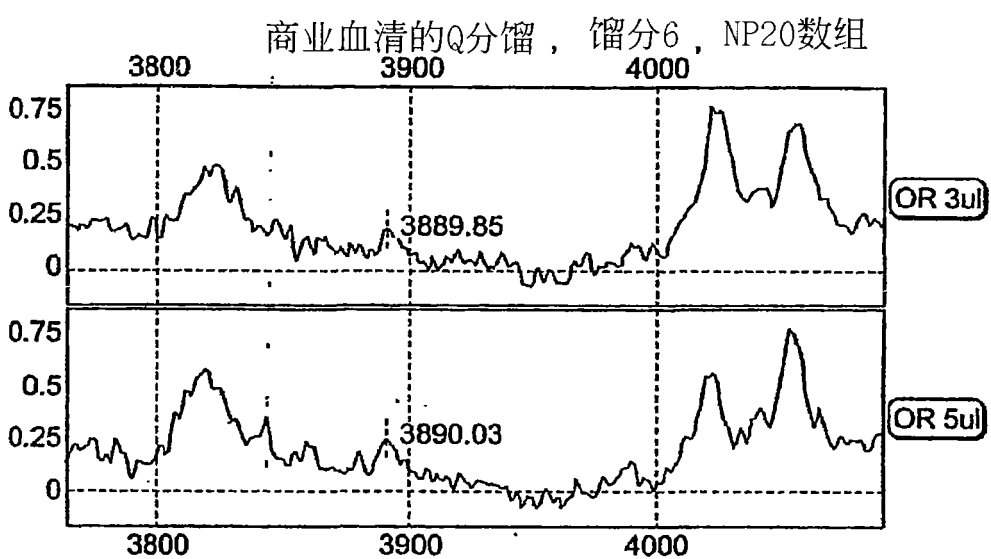


图 2

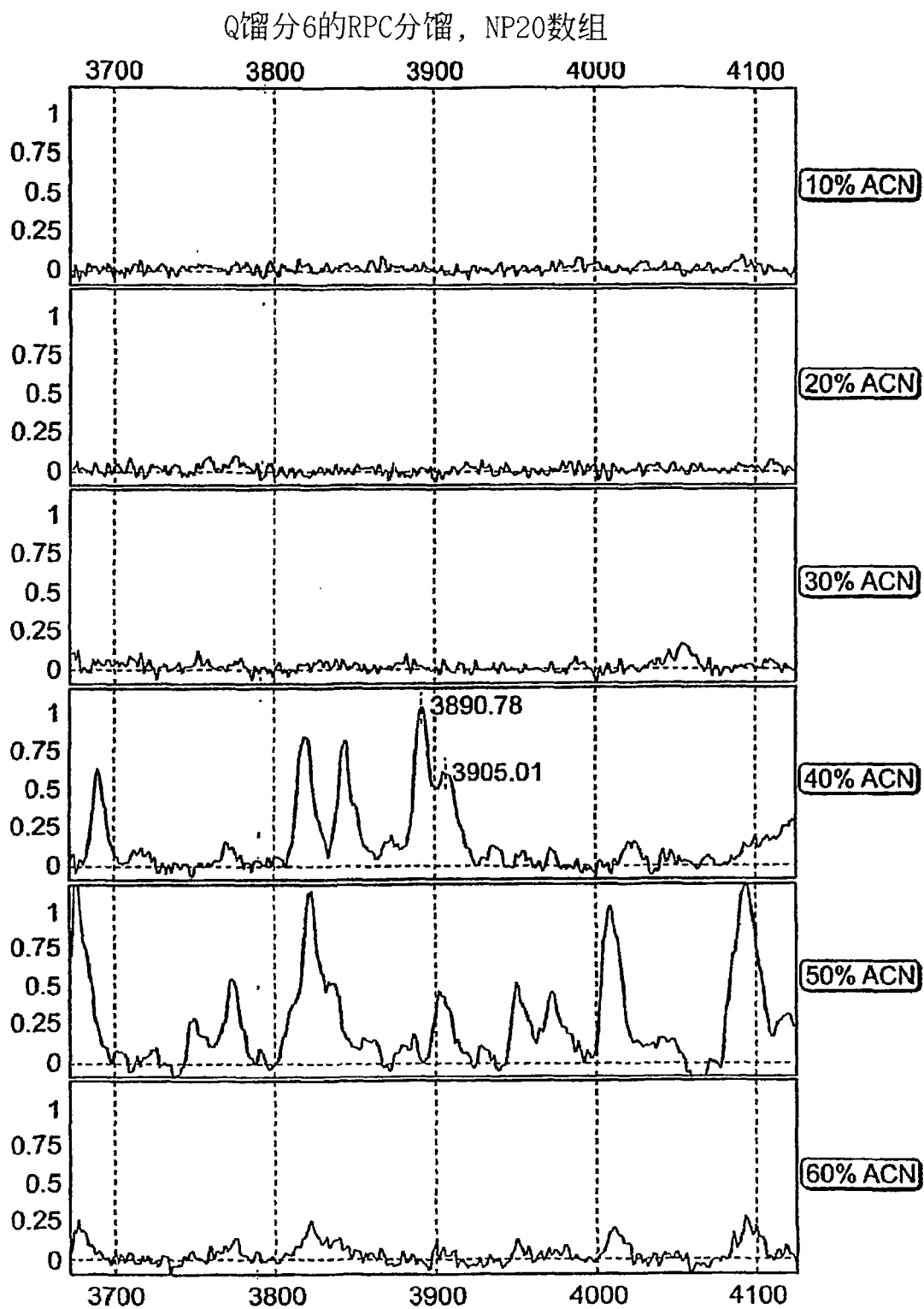


图 3

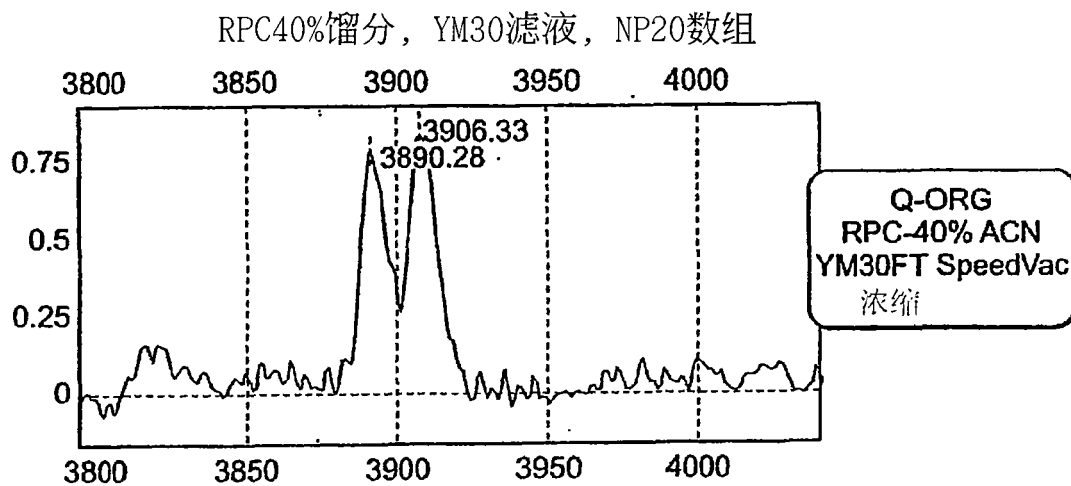


图 4

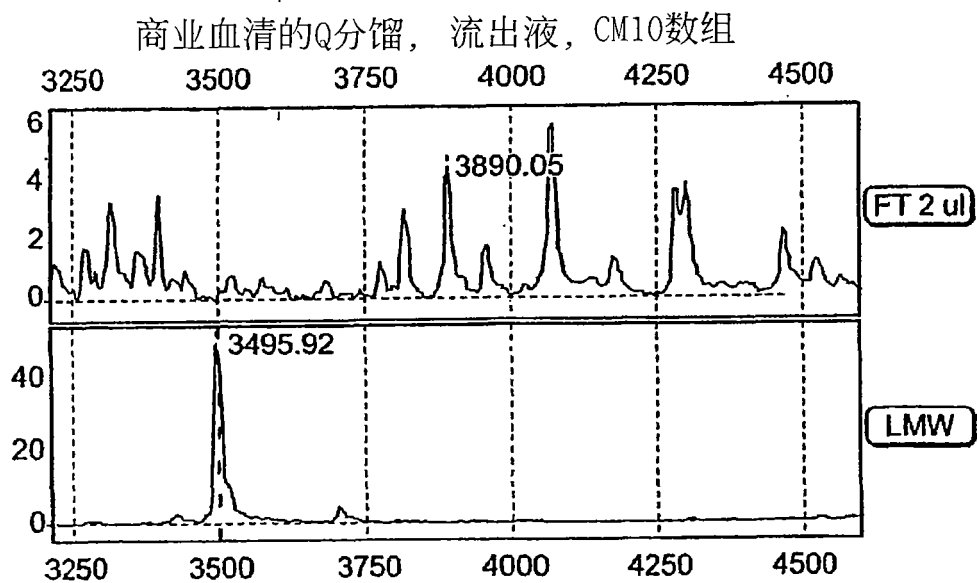


图 5



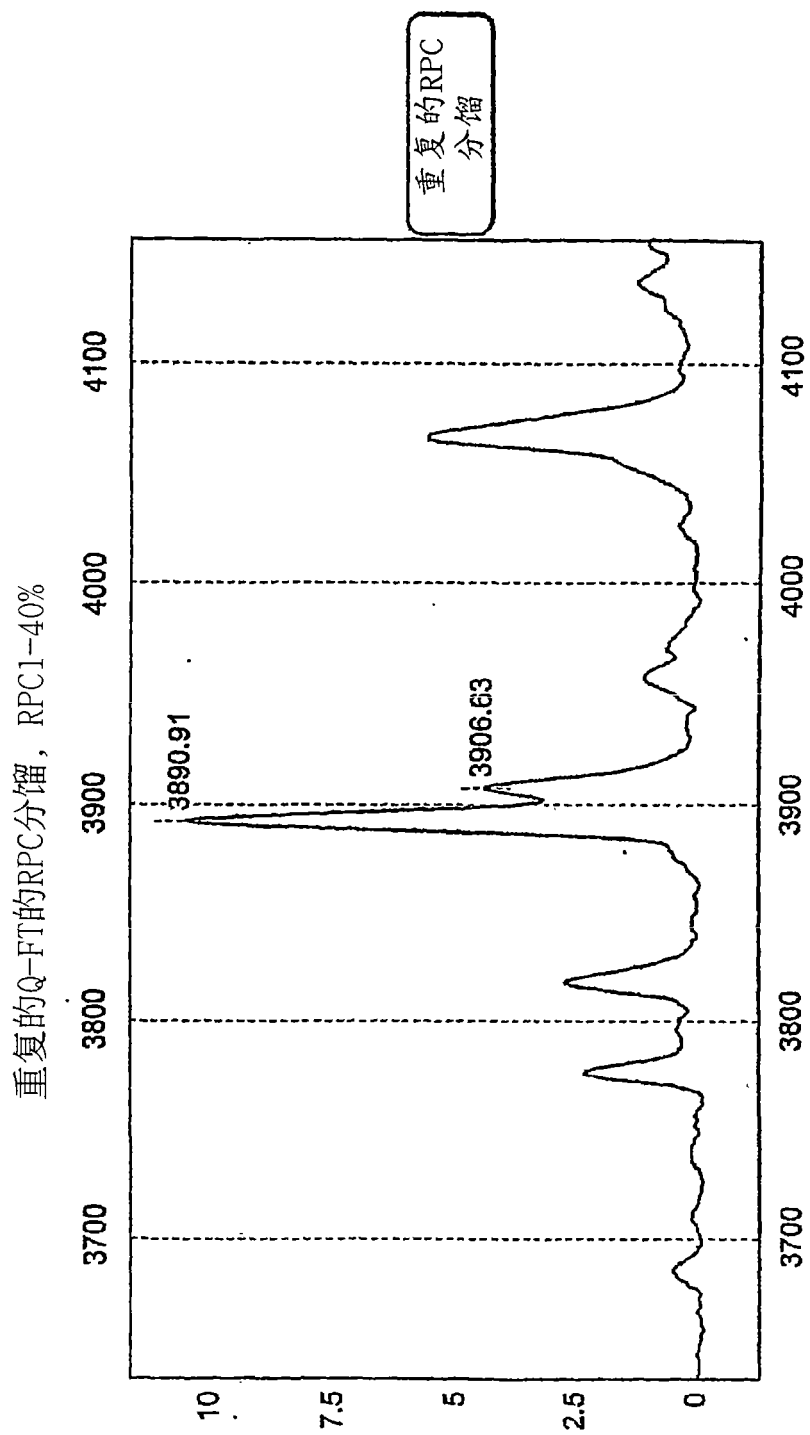


图 7

3890Da蛋白质的Q-TOF光谱

033005 3890 MS2 56(1.056) Sb(10.40.00); Sm(SG, 4x4.00); Sm(SG, 4x4.00); Sm(SG, 4x4.00); Sm(SG, 4x4.00); S

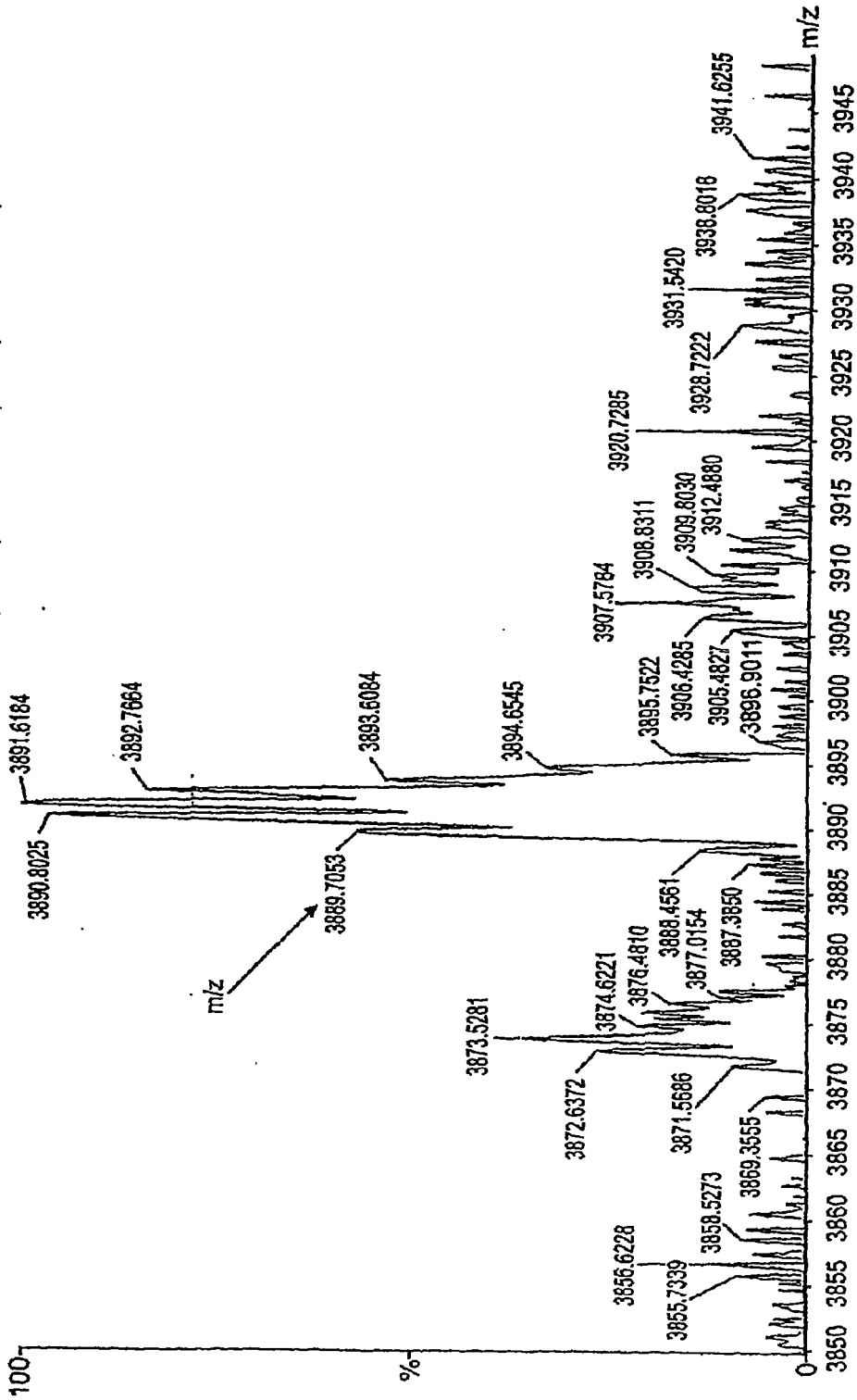


图 8

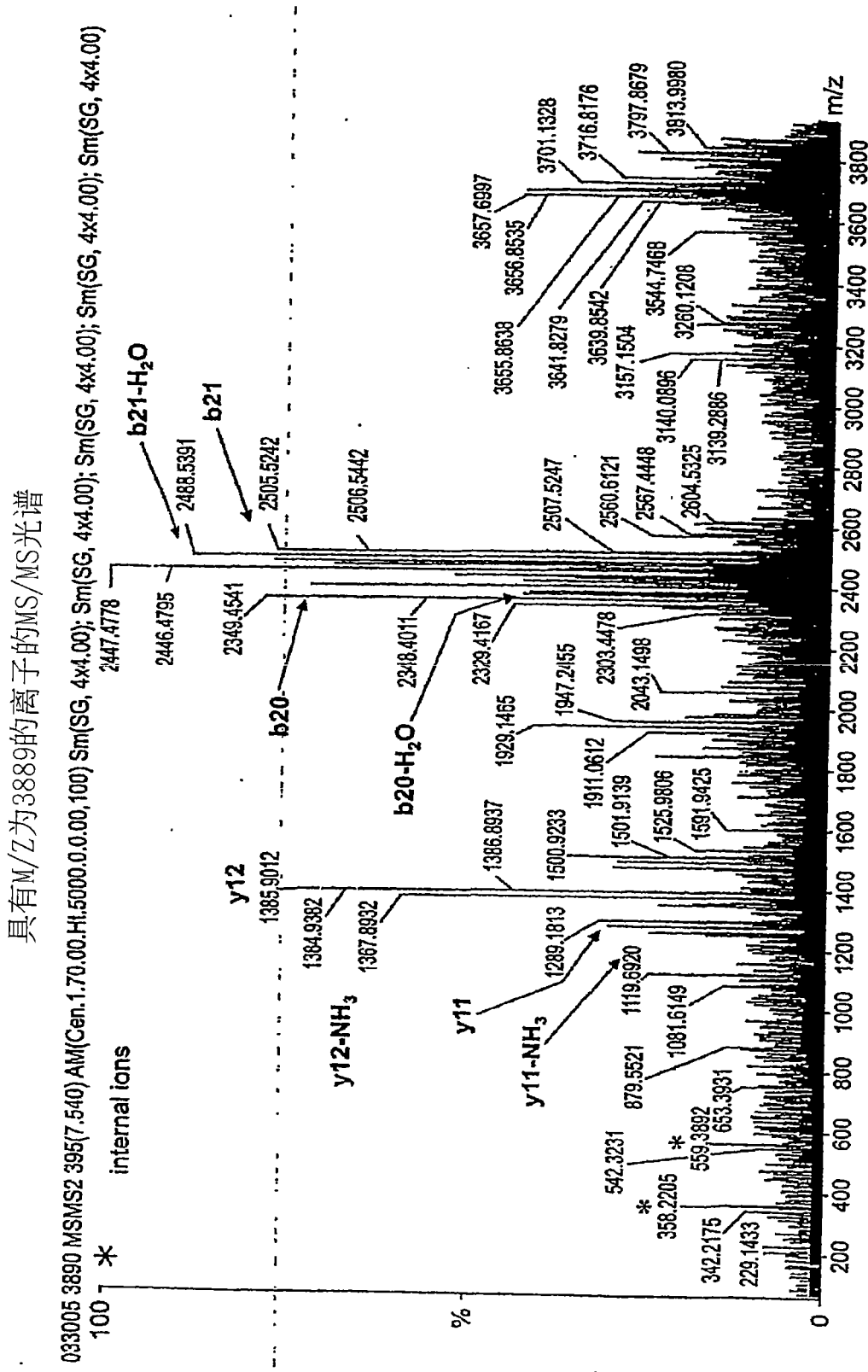


图 9

蛋白质探测器搜索输入

数据库	SwissProt2005.01.06	DNA框架翻译	3	消化	无酶
搜索点击	<input type="checkbox"/> 自 msfit	文件名	lastres1	最大的未裂解数	1
种类				Cys修饰是由	丙烯酰胺
蛋白质 MW (自	1000	至	100000 Da) 所有 <input checked="" type="checkbox"/>	最大未配对离子%	30
			开始搜索	最大报导的点击数	25
母种偏差度	150	ppm			
片段偏差度	100				
显示图表 <input type="checkbox"/>					
片段离子标签	仪器 MALDI-TOF				
质量为	自动计算AA组成 (基于亚氨基酸和相关离子)				
质量 (m, z)	样本ID (评论)				
电荷 (z)					
3889.70					
358.22					
559.39					
1253.84					
1270.83					
1272.83					
1289.78					
1367.89					
1384.9					
2348.40					
2372.52					

图 10

## 数据库搜索结果

## 参数

数据库开始 SwissProt2005.01.06

使用的消化：无酶

最大的未裂解数：1

肽链N末端：氢

肽链C末端：游离酸

半胱氨酸修饰：丙烯酰胺

仪器名：MALDI-TOF

最大未配对离子%：30

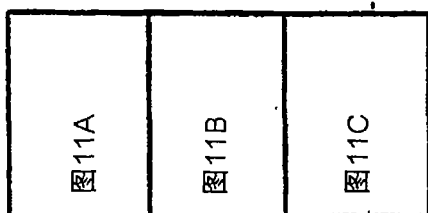
考虑的离子类型

搜索模式：相等性

肽链质量为单同位素

a-NH3, a, b, b-NH3, b-H2O, y, y-NH3 y-H2O, l, i, m, P, S

图11



## 预搜索结果

数据库内的登录数：161074

完整的分子量范围：161074次登录

完整的pI范围：161074次登录

预搜索结果：161074次登录

## 数据组1结果

通过母质量过滤的数目或序列：555317

MS标签搜索选择1登录

母质量：3889.7000(+/-0.58 Da)

用于搜索的15离子

358.22, 559.39, 1253.84, 1270.83, 1289.78, 1367.89, 1384.90, 2348.40, 2372.52, 2390.44,

2446.18, 2471.47 2487.52, 2505.52(+/- 100.00 ppm)

最大未配对数=4

图 11A

### 结果总结

等级	未配对离子数	测得序列	计算的MH+ (Da)	错误的MH+ (ppm)	MS消化 指数	蛋白质分子量 (Da)/pI	登录号	物种	蛋白名
1	4	(R)SARLNSORLVFN RPFLMFIVDNDNFL GKVNRF(-)	1889.1539	140	122.634	457829.3	P05154	人类	血浆丝氨酸蛋白酶抑制前体 (PCT) (蛋白C抑制剂(血浆酶源激活剂抑制剂-3) (PAD3)(顶体丝氨酸蛋白酶抑制剂)

### 详细结果

等级	未配对离子数	测得序列	计算的MH+ (Da)	错误的MH+ (ppm)	MS消化 指数	蛋白质分子量 (Da)/pI	登录号	物种	蛋白名
1	4	(R)SARLNSQRLVFN RPFLMFIVDNDNFL GKVNRF(-)	1889.1539	140	122.634	457829.3	P05154	HUMAN	血浆丝氨酸蛋白酶抑制前体 (PCT) (蛋白C抑制剂(血浆酶源激活剂抑制剂-3) (PAD3)(顶体丝氨酸蛋白酶抑制剂)

图 11B

片段离子	(m/z)	350.22	559.39	1253.84	1270.03	1272.83	1289.78	1367.89	1384.90	2348.40	2372.52
离子类型		PFL 19	LFLGK 52	Y <sub>11</sub> NH <sub>3</sub> 82	Y <sub>11</sub> 52	b <sub>11</sub> 88	RLNSQRLVFN- NH <sub>3</sub> 91 Y <sub>12</sub> NH <sub>3</sub> 80	DNNILFLGKVN- 94 RLNSQRLVENR 78 Y <sub>12</sub> 68			b <sub>20</sub> H <sub>2</sub> O 85
Δppm											

	2390.44	2446.48
SQRLVFNRPFLMFTVDNNIL 28 51 b <sub>20</sub> 47		RLVFNRPFLMFIVDNNILFL NH <sub>3</sub> 54
ARLNSQRLVFNRPFLMFIVD 28 47		

片段离子	(m/z)	2471.47	2487.52	2505.52
离子类型			b <sub>21</sub> <sup>-</sup> H <sub>2</sub> O 70	b <sub>21</sub> 66
Δppm				

图 11C



蛋白C抑制剂内的反应键

### Swiss-Prot: IPSP\_HUMAN

你所选择的IPSP HUMAN序列的片段相应于:

位点      373      374      反应键

以一字母编码:

1	11	21	31	41	51	
1	MQLELLCLV	LLSPQGASLH	RHHPREMKKR	VEDLHVGVTV	APSSRRDFTF	DLYRAKASAA      60
61	PSQNIFFSPV	SISMSLAMS	LGAGSSTKMQ	I LEGLGLNLQ	KSSEKELHRG	FQQLLQELNQ      120
121	PRDGFQLSLG	NALFTDLVVD	LQDTFVSAMK	TLYLADTFPT	NFRDSAGAMK	QINDYVAKOT      180
181	KGKIVDLLKN	LDSNAVVTMV	NYIFFKAKWF	TSFNHKGTOE	QDFYVTSETV	VRVPMMSRED      240
241	QYHYLLDRNL	SCRVVGVPYQ	GNATALFILP	SEKMQQVEN	GLSEKTLRKK	LKMFKKRQLE      300
301	LYLPKFSIEG	SYQLEKVLPS	LGISNVFTSH	ADLSGISNHS	NIQVSEMVHK	AVVEVDESST      360
361	RAAAATGTIF	TFERSARLNSQ	RLVFNRPFLM	FIVDNNILFL	GKVNRF	

图 13

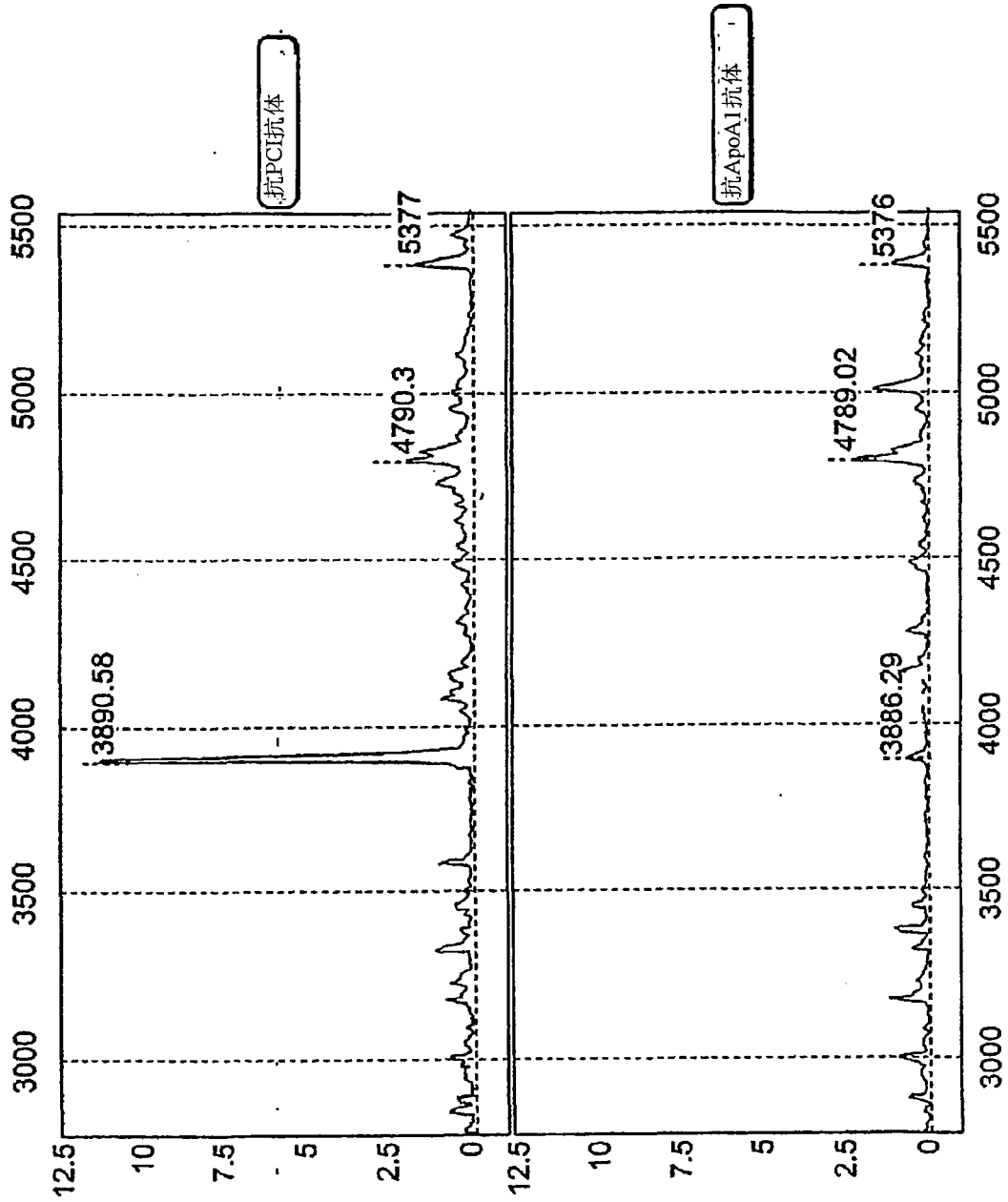
蛋白C抑制剂的C末端片段的理论MW

蛋白质序列的理论pI/Mw

**SARLNSQRLY ... ILEFLGKVNRP:**

理论pI/Mw      12.00 / 3890.61

图 14



使用全血清和山羊  
抗PCI抗体以珠为基础  
的免疫分析

图 15

用于发现和确认的多中心样本组

位点	N	基于切片的诊断			
		良性	前列腺癌/Gleason评分		
			5-6	7	8-9
1	106	56	40	7	3
2	111	63	26	17	5
3	68	35	16	9	3
4	93	52	17	20	4
总共	378	206	104	53	15

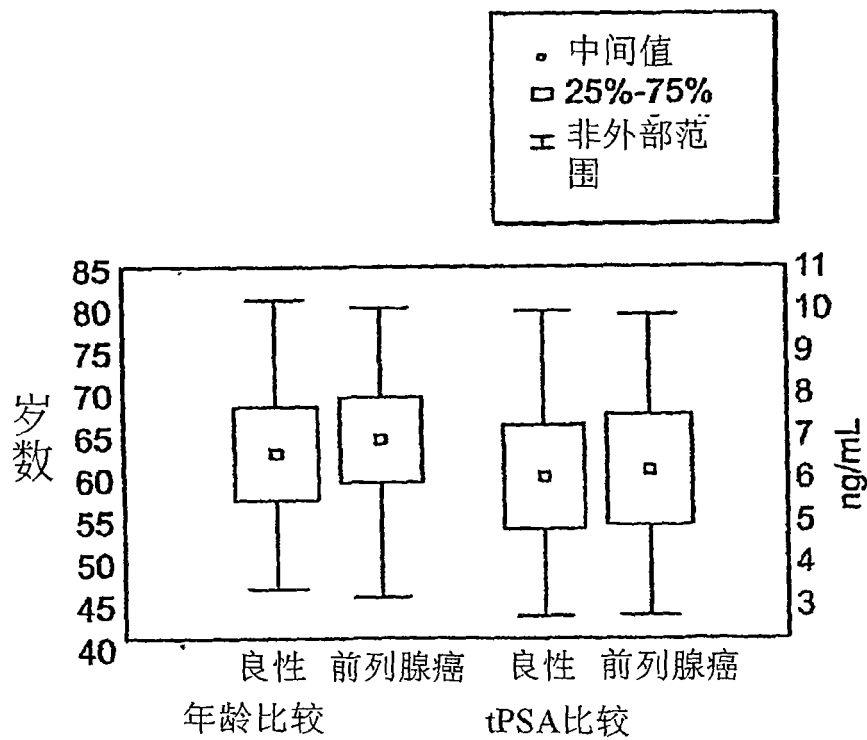


图 16

### 通过相关网络分析的发现 (仅癌症样本)

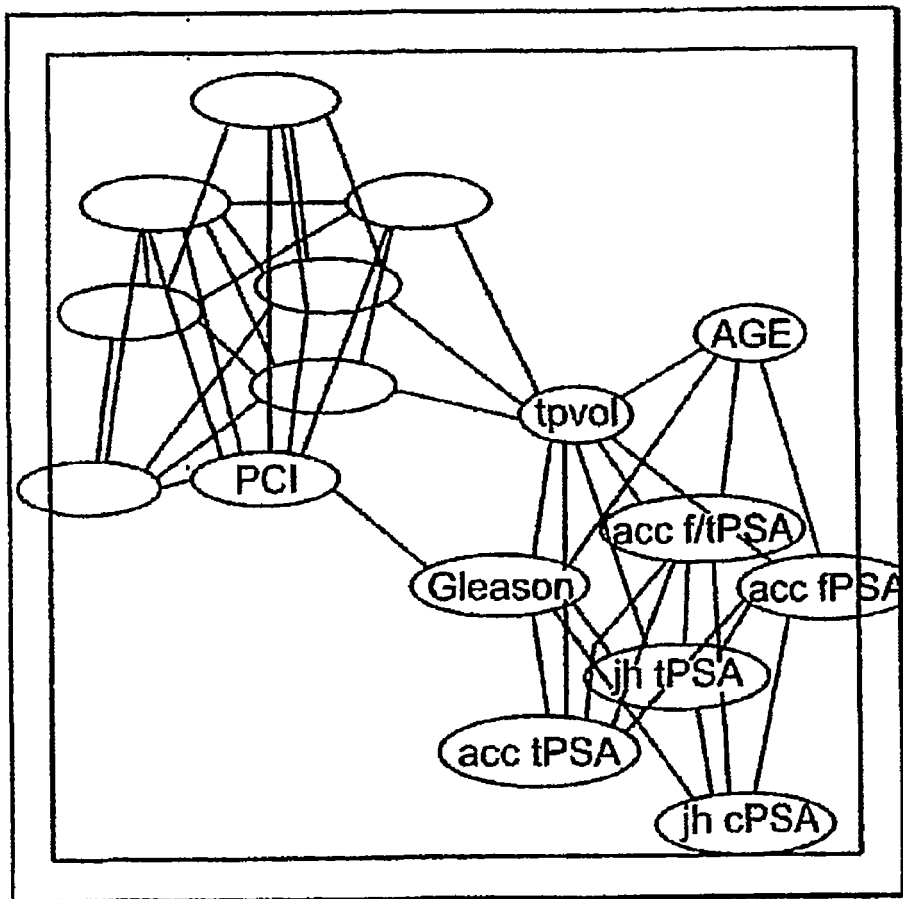
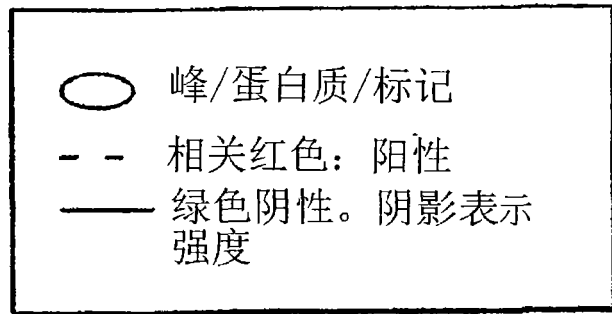
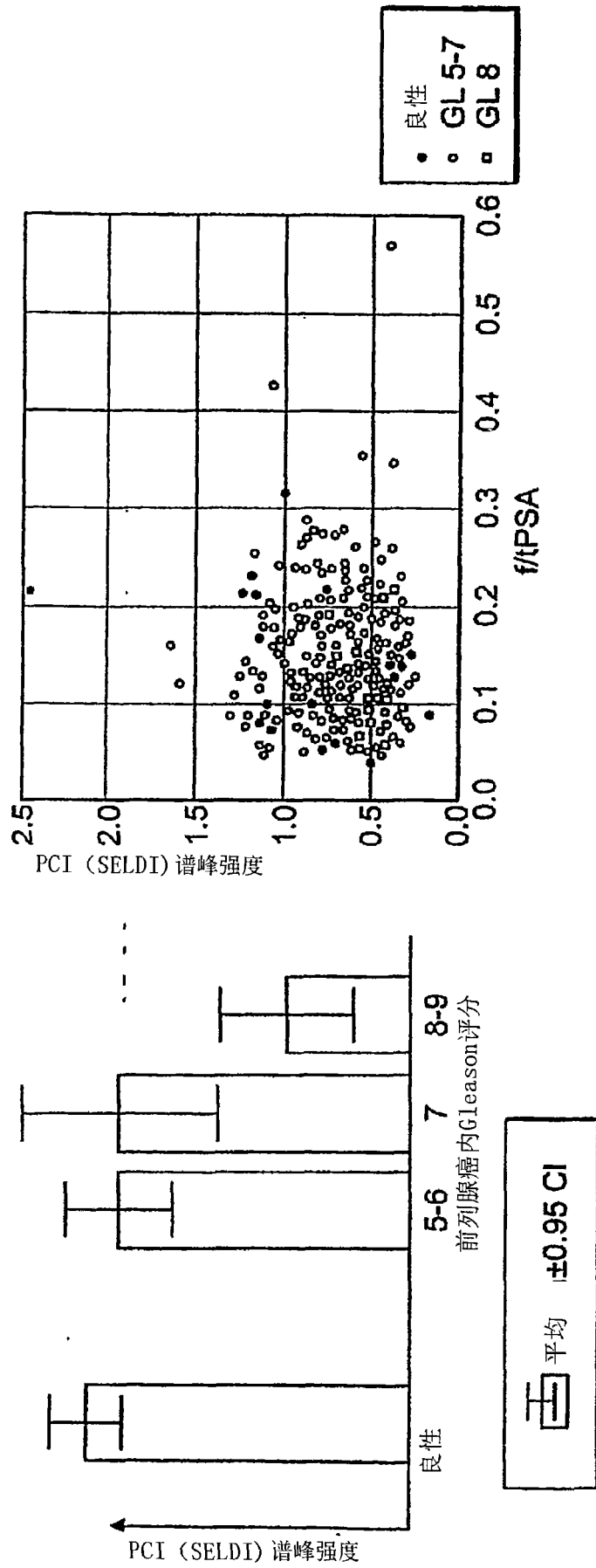


图 17

原始数据组的结果

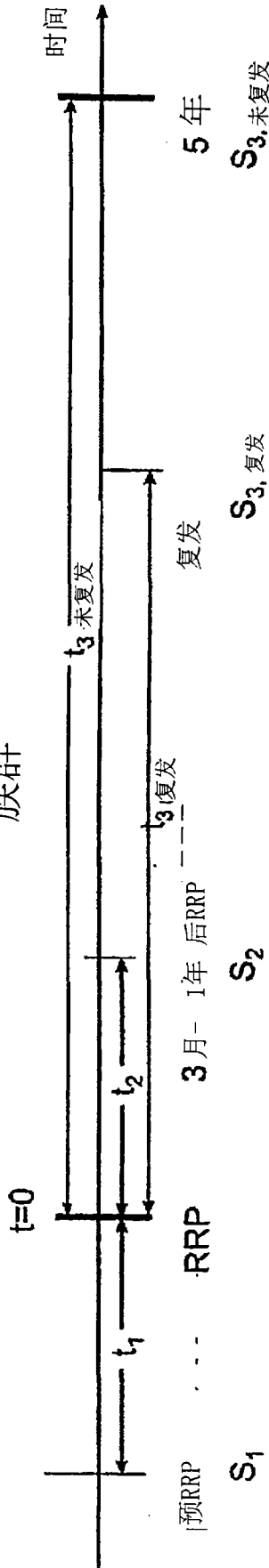


在灰色区 (10-20%) 中从具有游离/总 PSA 的男性良性前列腺癌疾病中分离前列腺癌

在前列腺癌病人中以Gleason 评分校正

图 18

### 前列腺癌复发研究 族群



	N	平均年龄	Gleason评分			临床阶段*			SI**中PSA (ng/ml) 中 间值, 25/75 百分点	平均时间 (天)				
			5-6	7	8-9	2	2a	2b		3a	3b	t <sub>1</sub>	t <sub>2</sub>	t <sub>3</sub>
复发	52	57.0	15	34	3	5	3	5	35	4	8.7 (5.8, 14.2)	57	182	1264
未复发	52	59.4	15	34	3	5	3	5	35	4	7.2 (4.8, 10.9)	59	215	1975

\* 96% NO. 100% MO.

\*\* 在S1收集物中的p=0.091 (Mann-Whitney U测试) ; 在S2收集物中所  
有病人的PSA是不可探测的 (<0.1ng/ml)

图 19

### 复发研究数据组的比率独立确认

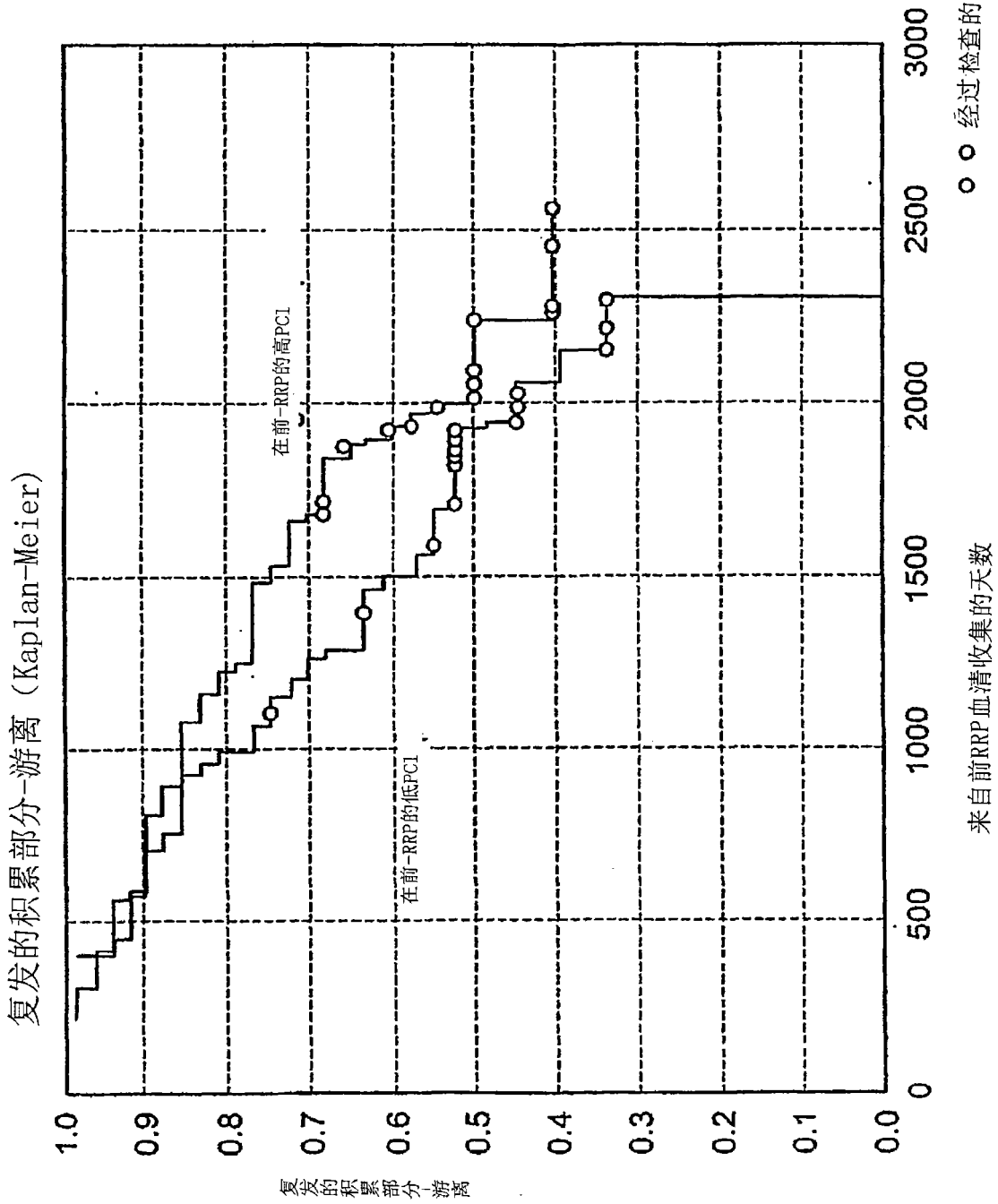


图 20

使用部分复发样本作为训练组的Pre-RRP PSA、PCI和C4a组合

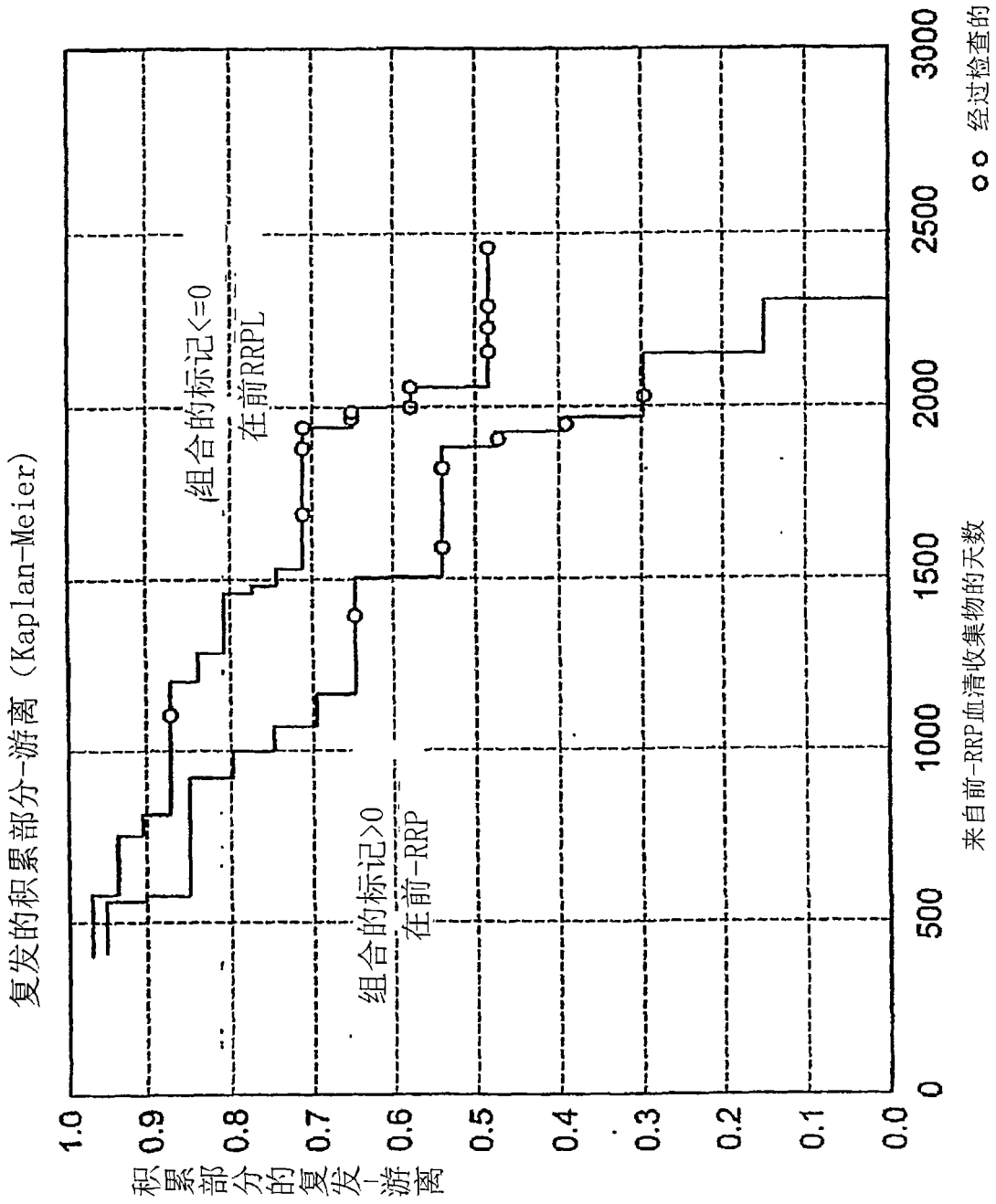


图 21

使用部分复发样本作为训练组的测试组的Pre-RRP PSA、PCI和C4a组合

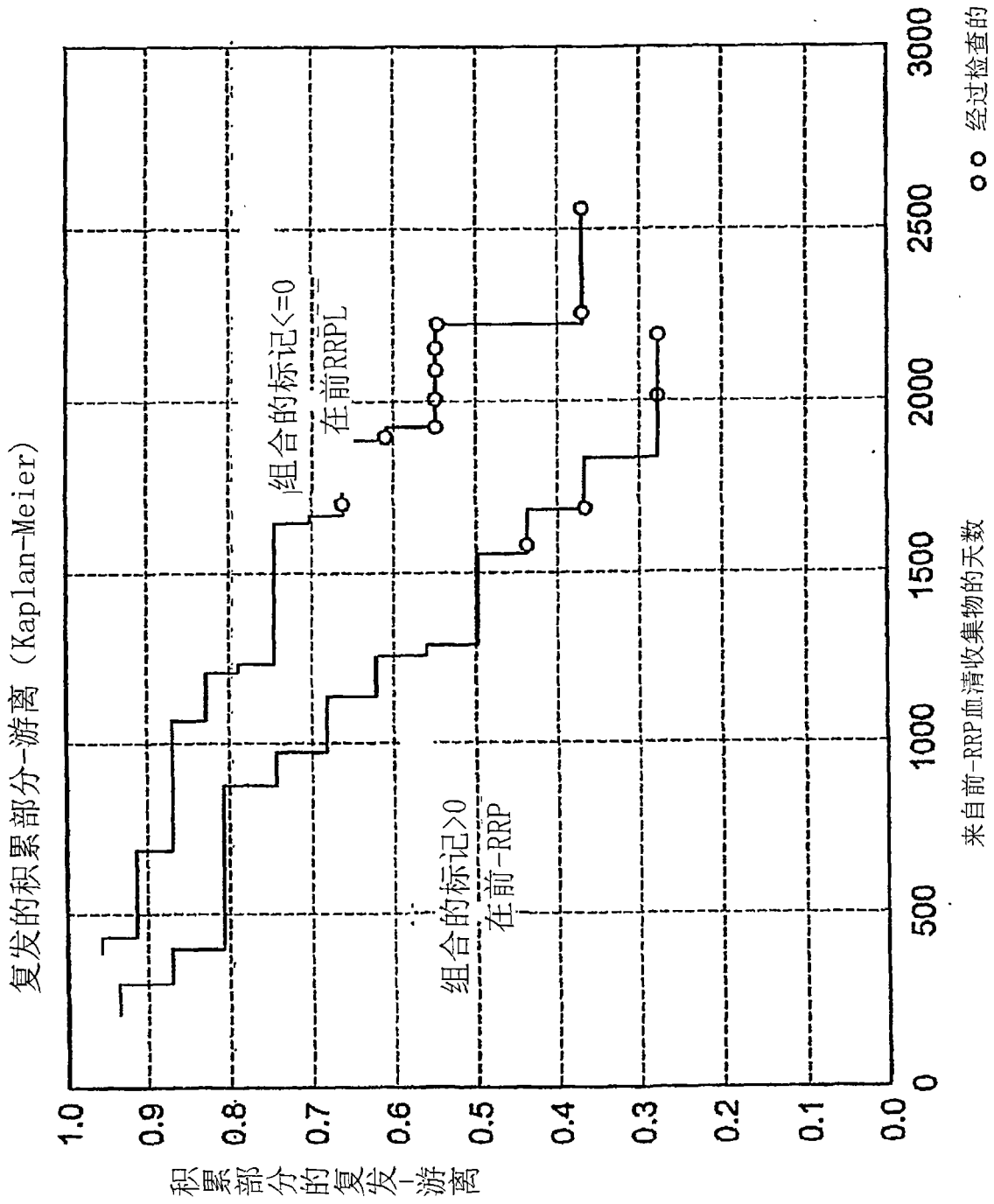
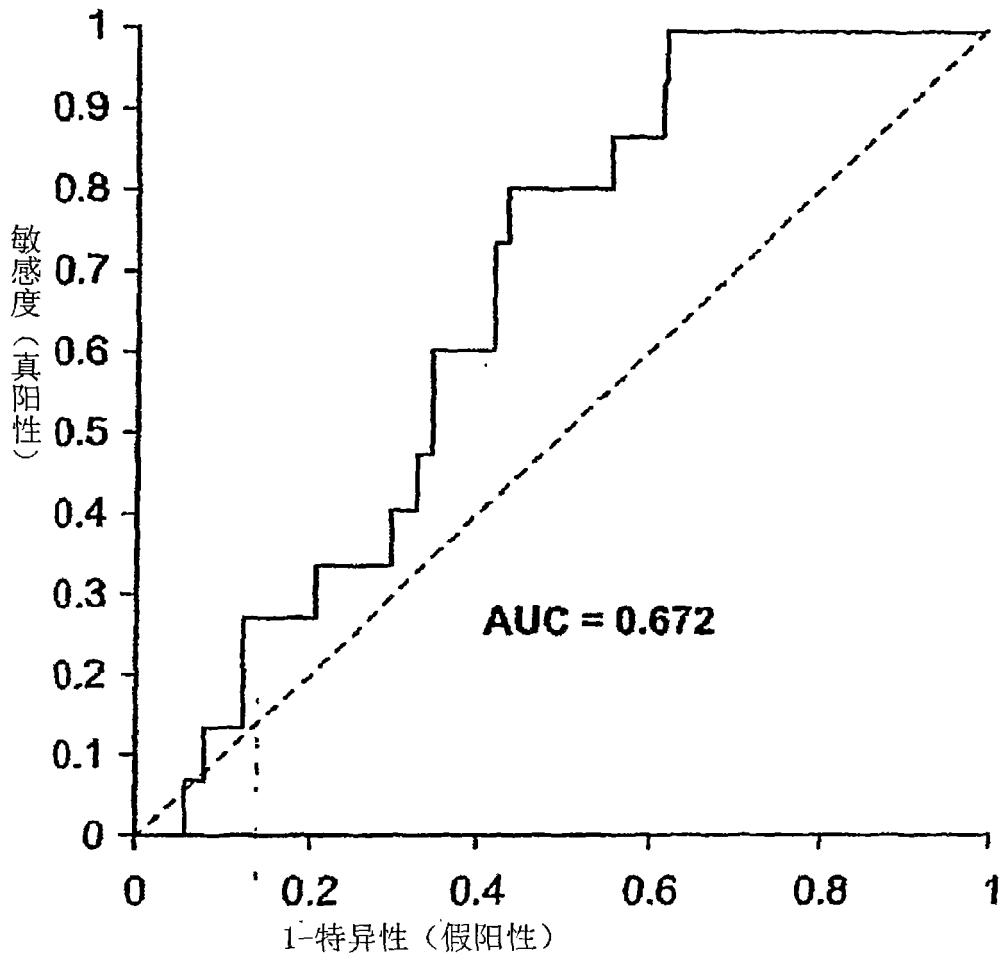


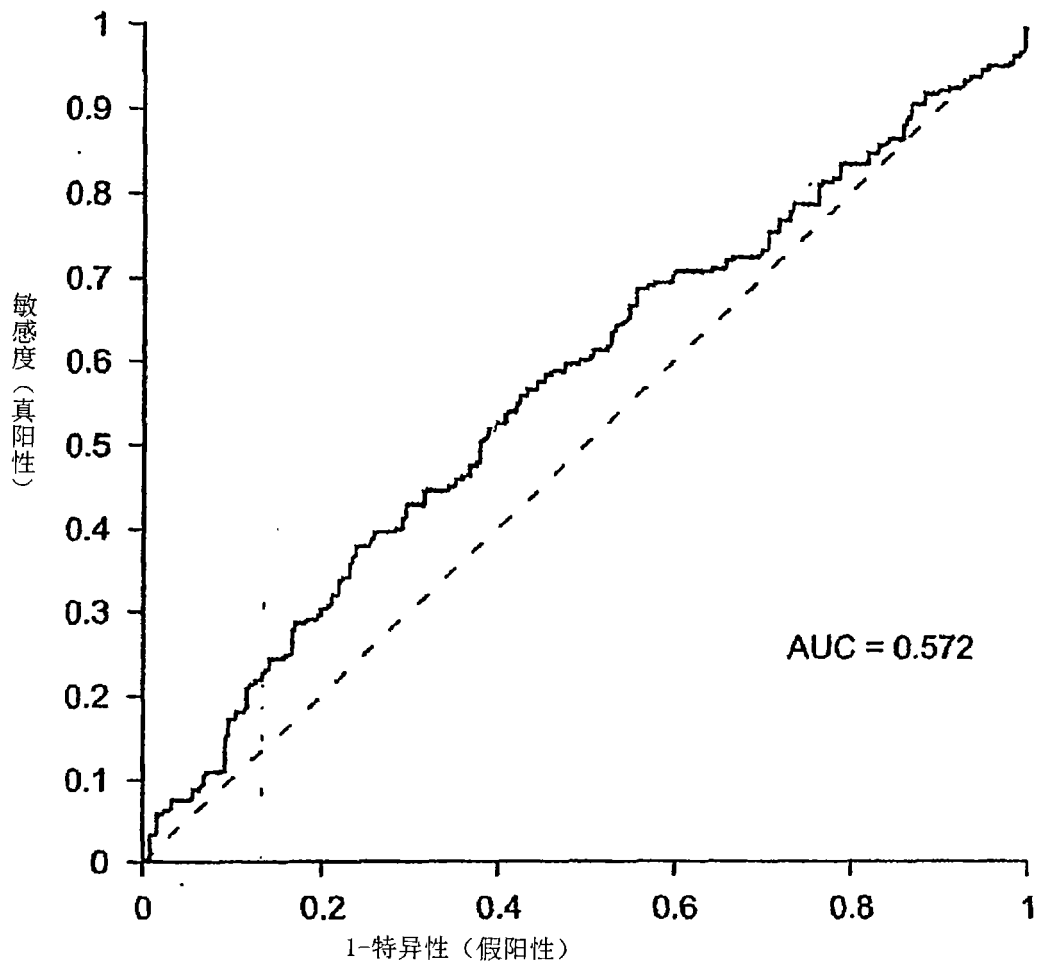
图 22



**All Ca: GL=5-7 vs GL=8+**

PCI峰的ROC分析，将Gleason评分5-7的前列腺癌病人与Gleason评分为8或更高的病人予以区分

图 23



PCI峰的ROC分析，将患有良性前列腺疾病的病人与前列腺癌病人予以区分

图 24

专利名称(译)	前列腺癌的生物标记		
公开(公告)号	<a href="#">CN101310185A</a>	公开(公告)日	2008-11-19
申请号	CN200680042935.X	申请日	2006-09-19
[标]申请(专利权)人(译)	约翰霍普金斯大学		
申请(专利权)人(译)	约翰·霍普金斯大学		
当前申请(专利权)人(译)	约翰·霍普金斯大学		
[标]发明人	Z张 DW陈 CN罗森茨魏希		
发明人	Z·张 D·W·陈 C·N·罗森茨魏希		
IPC分类号	G01N33/53		
CPC分类号	G01N33/57434 G01N2333/8121 A61P13/08 A61P35/00		
优先权	60/718843 2005-09-19 US		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明提供以蛋白质为基础的生物标记，也就是蛋白C抑制剂(PCI)，用于鉴别病人的前列腺癌状态。特别地，本发明的生物标记用于分类受测样本为前列腺癌或非前列腺癌。生物标记可通过SELDI质谱分析检测。

标记	P-值	前列腺癌中的上或下 调控	ProteinChip®分析
PCI (3890 DA)	<0.0001	上	CM10