



(12) 发明专利

(10) 授权公告号 CN 1864068 B

(45) 授权公告日 2010.10.13

(21) 申请号 200480029598.1

C07K 14/00 (2006.01)

(22) 申请日 2004.08.18

(56) 对比文件

(30) 优先权数据

60/496,154 2003.08.18 US

60/511,720 2003.10.15 US

W0 03027682 A2, 2003.04.03, 参见说明书第 37 页第 20 行至第 39 页第 23 行.

US 20030119063A1, 2003.06.26, 说明书第 [0009] — [0014] 以及 [0079] — [0080].

(85) PCT 申请进入国家阶段日

2006.04.10

W0 0186306 A2, 2001.11.15, 说明书第 21 页第 10 行至第 24 页第 5 行.

(86) PCT 申请的申请数据

PCT/US2004/027005 2004.08.18

审查员 李冰

(87) PCT 申请的公布数据

W02005/019831 EN 2005.03.03

(73) 专利权人 特提斯生物科学公司

地址 美国加利福尼亚州

(72) 发明人 M·S·乌尔代亚 G·M·兰德斯

G·T·文特

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 黄革生 隗永良

(51) Int. Cl.

G01N 33/68 (2006.01)

C07K 16/00 (2006.01)

权利要求书 2 页 说明书 43 页 附图 2 页

(54) 发明名称

用小表位抗体降低样品复杂度的方法

(57) 摘要

本发明一般涉及用于降低样品复杂度的方法。更特别地,本发明涉及蛋白质组学、生物样品中蛋白质水平的测定及用识别小表位的抗体分析样品中的蛋白质。

1. 降低包含蛋白质混合物的样品的复杂度的方法,所述方法包括从所述蛋白质的混合物分离多种不同的蛋白质,其中所述多种蛋白质在小表位抗体-蛋白质复合体中结合到一种小表位抗体,其中所述小表位抗体结合由3到5个氨基酸组成的表位,其中富集被所述小表位抗体结合的多种蛋白质,并从所述抗体-蛋白质复合体分离所述多种蛋白质。

2. 根据权利要求1的方法,其还包括将分离自所述抗体-蛋白质复合体的多种蛋白质与蛋白质切割剂接触以形成多肽片段。

3. 根据权利要求2的方法,其中所述蛋白质切割剂包含蛋白酶。

4. 根据权利要求2的方法,其中所述蛋白质切割剂包含化学剂。

5. 根据权利要求1的方法,其中所述蛋白质混合物包含通过用蛋白质切割剂切割所述样品中的蛋白质产生的多肽片段的混合物。

6. 根据权利要求5的方法,其中所述蛋白质切割剂包含蛋白酶。

7. 根据权利要求5的方法,其中所述蛋白质切割剂包含化学剂。

8. 根据权利要求1的方法制备的多种蛋白质。

9. 多肽片段,其根据权利要求2的方法制备。

10. 降低包含蛋白质混合物的样品的复杂度的方法,所述方法包括从所述蛋白质的混合物分离多种不同的蛋白质,其中所述多种蛋白质在小表位抗体-蛋白质复合体中结合到多种不同的小表位抗体,其中所述多种小表位抗体中每种小表位抗体结合由3到5个氨基酸组成的不同表位,其中富集被所述多种小表位抗体结合的所述多种蛋白质,并从所述抗体-蛋白质复合体分离所述多种蛋白质。

11. 根据权利要求10的方法,其中将小表位抗体固定在固体基质上。

12. 根据权利要求10的方法,其中可检测地标记所述小表位抗体。

13. 根据权利要求10的方法,其中所述样品与多种小表位抗体的所述接触平行地进行。

14. 根据权利要求10的方法,其中所述样品与多种小表位抗体的所述接触顺序地进行。

15. 根据权利要求10的方法,其中所述多种不同的小表位抗体包含至少约100种不同的小表位抗体。

16. 降低包含蛋白质混合物的样品的复杂度的方法,所述方法包括:

(a) 将包含蛋白质混合物的样品与至少一种小表位抗体接触以形成小表位抗体-蛋白质复合体,其中来自所述样品的多种不同的蛋白质在小表位抗体-蛋白质复合体中结合到所述小表位抗体,其中该小表位抗体结合由3到5个氨基酸组成的表位;和

(b) 分离样品中所述小表位抗体-蛋白质复合体与未结合的蛋白质。

17. 根据权利要求16的方法,其中步骤(a)和(b)顺序进行。

18. 根据权利要求16的方法,其中步骤(a)和(b)同时进行。

19. 根据权利要求16的方法,其中至少一种小表位抗体包含至少约100种小表位抗体。

20. 根据权利要求16的方法,其还包括从所述小表位抗体-蛋白质复合体分离所述多种蛋白质。

21. 确定样品中存在或不存在目的蛋白质的方法,所述方法包含如果有目的蛋白质,那么检测富集的蛋白质级分中的目的蛋白质,其中通过权利要求1或10的方法制备经富集的

蛋白质级分,并且其中目的蛋白质的检测表明样品中存在所述蛋白质。

22. 根据权利要求 21 的方法,其中所述检测包含质谱分析法。

23. 确定样品中目的蛋白质的量的方法,所述方法包括定量所富集的蛋白质级分中目的蛋白质的量,其中通过权利要求 1 或 10 的方法制备所述富集的蛋白质级分。

24. 根据权利要求 23 的方法,其中所述定量包含质谱分析法。

25. 鉴定小表位抗体-蛋白质复合体中蛋白质的方法,其中根据权利要求 1 制备所述小表位抗体-蛋白质复合体。

26. 根据权利要求 25 的方法,其中所述鉴定包含质谱分析法。

27. 用于鉴定生物标记的方法,所述方法包括比较两种或多种经富集的蛋白质级分中的蛋白质,其中根据权利要求 1 或 10 的方法从样品制备每种经富集的蛋白质级分。

28. 根据权利要求 27 的方法,其中两个或多个样品包含来自至少一名具有疾病状况的个体的样品和至少一名没有所述疾病状况的个体的样品,并且其中所述生物标记的存在或不存在指示所述疾病状况。

29. 组合物,其包含多种不同的小表位抗体,其中所述多种小表位抗体中的每种小表位抗体结合由 3 到 5 个氨基酸组成的不同表位。

30. 根据权利要求 29 的组合物,其中可检测地标记所述小表位抗体。

31. 根据权利要求 29 的组合物,其中所述多种小表位抗体包含至少约 100 种不同的小表位抗体。

32. 用于根据权利要求 1-7、10-28 的任一项的方法的试剂盒,其包含至少一种小表位抗体,其中所述至少一种小表位抗体结合由 3 到 5 个氨基酸组成的表位。

33. 根据权利要求 32 的试剂盒,其中可检测地标记所述至少一种小表位抗体。

34. 根据权利要求 32 或 33 的试剂盒,其中所述至少一种小表位抗体包含多种小表位抗体。

35. 根据权利要求 34 的试剂盒,其中所述多种小表位抗体包含至少约 100 种小表位抗体。

36. 权利要求 10 的方法,其还包括将从所述抗体-蛋白质复合体分离的多种蛋白质与蛋白质切割剂接触以形成多肽片段。

37. 权利要求 36 的方法,其中所述蛋白质切割剂包含蛋白酶。

38. 权利要求 36 的方法,其中所述蛋白质切割剂包含化学剂。

39. 权利要求 10 的方法,其中所述蛋白质混合物包含通过用蛋白质切割剂切割所述样品中的蛋白质产生的多肽片段的混合物。

40. 权利要求 39 的方法,其中所述蛋白质切割剂包含蛋白酶。

41. 权利要求 39 的方法,其中所述蛋白质切割剂包含化学剂。

42. 权利要求 1、10 和 16 任一项的方法,其中所述一种或多种小表位抗体结合的表位存在于所述多种蛋白质的每种蛋白质的氨基酸序列内。

用小表位抗体降低样品复杂度的方法

[0001] 相关申请的交互参照

[0002] 本申请要求 2003 年 8 月 18 日提交的美国临时专利申请号 60/496, 154、以及 2003 年 10 月 15 日提交的美国临时专利申请号 60/511, 720 的权利, 将它们在此全部引用作为参考。

发明领域

[0003] 本发明一般涉及用于降低样品复杂度的方法。更特别地, 本发明涉及蛋白质组学, 生物样品中蛋白质水平的测量, 和使用识别小表位的抗体对样品中蛋白质的分析。

[0004] 发明背景

[0005] 蛋白质组学与用于评估基因活性的传统方法基因组学相比, 可更直接的考察细胞或有机体的生物学功能。蛋白质组学包括通过检测及定量在蛋白质水平而不是在信使 RNA 水平上的表达定性及定量测定基因活性。蛋白质组学也包括研究非基因组编码事件, 包括蛋白质的翻译后修饰、蛋白质降解及蛋白质副产物、蛋白质之间的相互作用以及细胞中蛋白质的定位。细胞表达的蛋白质的结构、功能或活性水平也是重要的。

[0006] 蛋白质水平上基因表达的研究是重要的, 因为许多最重要的细胞过程是通过细胞的蛋白质状态调节, 而不是通过基因表达状态调节的。并且, 由于大部分药物设计为对蛋白质靶具有活性, 所以细胞的蛋白质含量与药物发现成果高度相关。

[0007] 用于分析蛋白质混合物的当前技术是基于多种蛋白质分离技术及随后经分离蛋白质的鉴定和 / 或分析, 所述蛋白质混合物为诸如细胞或细胞群的细胞内蛋白质及通过细胞或细胞群或生物学液体分泌的蛋白质。最普遍的方法是基于 2D- 凝胶电泳及随后的“胶内”(in-gel) 蛋白水解消化及质谱分析。备选地, 可用 Edman 及相关方法测序。该 2D- 凝胶技术需要大量样品量、费时且当前在可再现地分辨通过人细胞表达的蛋白质的重要级分的能力中受到限制。包括一些大型 2D- 凝胶的技术比常规 2D- 凝胶技术可产生分离较大量蛋白质的凝胶, 但重复性仍很低并且由于可利用的测序技术灵敏度的限制不能对 95% 以上的点进行测序。电泳技术也具有偏向于高丰度蛋白质的缺点。

[0008] 因此, 需要能够更完全地测定有机体中细胞或细胞群表达的蛋白质或包含蛋白质的液体(如血清、血浆、淋巴及其它生物液体)中的蛋白质, 包括直到一个或多个细胞表达的或者包含蛋白质的液体中发现的整组蛋白质。

[0009] 发明概述

[0010] 一方面, 本发明提供了用于降低样品复杂度的方法, 所述方法包括:(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触; 和(b) 分离抗体-蛋白质复合体, 由此离析、分离、富集和 / 或纯化包含被一种或多种小表位抗体结合的一种或多种表位的蛋白质。

[0011] 另一方面, 本发明提供了方法, 所述方法包括(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触;(b) 分离抗体-蛋白质复合体, 而将通过蛋白质进行离析、分离、富集和 / 或纯化, 由此离析、分离、富集和 / 或纯化包含被一种或多种小表位抗体结合的一种或多种表位的蛋白质; 和(c) 自抗体-蛋白质复合体分离蛋白质。

[0012] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:分离小表位抗体-蛋白质复合物,而将包含通过小表位抗体结合的表位的蛋白质进行富集;其中通过样品与小表位抗体接触产生所述复合物。

[0013] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:(a) 分离小表位抗体-蛋白质复合物,而将包含通过小表位抗体结合的表位的蛋白质进行富集;其中通过样品与小表位抗体接触产生所述复合物;和(b) 自抗体-蛋白质复合物分离蛋白质。

[0014] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括自小表位抗体-蛋白质复合物分离蛋白质,而将包含通过小表位抗体结合的表位的蛋白质进行富集;其中小表位抗体-蛋白质复合物通过(a) 和(b) 产生,所述(a) 为在允许结合条件下将样品与小表位抗体接触,由此产生小表位抗体-蛋白质复合物;所述(b) 为分离抗体-蛋白质复合物。

[0015] 显然地,可组合和/或顺序进行(常常以任意次序,只要能形成所需产物)一个或多个步骤,且显然地,本发明包括文中描述的步骤的多种组合。显然且在文中描述的是,本发明包括这样的方法,其中最初,或第一个步骤是文中描述的任意步骤。本发明的方法包含实施方案,在所述实施方案中较晚的“下游”步骤为最初步骤。

[0016] 在一些实施方案中,所述方法还包括用蛋白质切割剂处理样品的步骤,从而产生多肽片段。在包括自抗体-蛋白质复合物分离蛋白质的步骤的实施方案中,在样品与至少一种小表位抗体接触的步骤之前,和/或自抗体-蛋白质复合物分离蛋白质的步骤之后,可用蛋白质切割剂处理样品。用蛋白质切割剂处理的方法是本领域公知的且在文中描述。可使用一种或多种蛋白质切割剂。蛋白质切割剂可为酶(如胰凝乳蛋白酶或胰蛋白酶)或化学剂(如溴化氰)。

[0017] 因此,另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触;(b) 分离抗体-蛋白质复合物,从而将包含被一种或多种小表位抗体结合的一种或多种表位的蛋白质进行富集;(c) 自抗体-蛋白质复合物分离蛋白质;和(d) 用蛋白质切割剂处理蛋白质,从而产生多肽片段。

[0018] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触,以形成抗体-蛋白质复合物;和(b) 用蛋白质切割剂处理抗体-蛋白质复合物以产生多肽片段。

[0019] 另一方面,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括:(a) 用蛋白质切割剂处理样品,从而产生多肽片段;(b) 在允许结合的条件下将多肽片段与一种或多种小表位抗体接触,从而产生抗体-多肽复合物;和(c) 分离抗体-多肽复合物,从而将包含被一种或多种小表位抗体结合的一种或多种表位的多肽进行富集。

[0020] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:(a) 孵育反应混合物,所述反应混合物包含:(i) 小表位抗体,和(ii) 样品,其中在允许结合的条件下孵育;和(b) 分离抗体-蛋白质复合物,从而富集蛋白质。

[0021] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:分离抗体-蛋白质复合物,从而富集蛋白质;其中通过孵育反应混合物产生抗体-蛋白质复合物,所述反应混合物包含:(a) 小表位抗体;和(b) 样品,其中在允许结合的条件下孵育。

[0022] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:(a) 孵育反

应混合物,所述反应混合物包含:(i)小表位抗体,和(ii)样品,其中在允许结合的条件下孵育;(b)分离抗体-蛋白质复合物;和(c)自蛋白质-抗体复合物分离蛋白质,从而富集蛋白质。

[0023] 另一方面,本发明提供了自经分离的蛋白质-抗体复合物分离蛋白质,其中通过孵育反应混合物产生蛋白质-抗体复合物,所述反应混合物包含(a)小表位抗体,和(b)样品,其中在允许结合的条件下孵育;和分离蛋白质-抗体复合物。

[0024] 另一方面,本发明提供了用于降低包含蛋白质混合物的样品复杂度的方法,其包括分离小表位抗体-蛋白质复合物,其中将包含通过小表位抗体结合的表位的蛋白质进行富集。在一些实施方案中,所述方法还包括自抗体-蛋白质复合物分离蛋白质。在一些实施方案中,小表位抗体结合由约3至约5个氨基酸组成的表位。在一些实施方案中,将样品与多种小表位抗体接触以形成多种小表位抗体-蛋白质复合物。在一些实施方案中,将小表位抗体进行可检测地标记。在一些实施方案中,将多种小表位抗体在固体基质上固定。在一些实施方案中,样品与多种小表位抗体平行地接触。在一些实施方案中,样品与多种小表位抗体顺序接触。在一些实施方案中,样品与至少100种小表位抗体接触。在一些实施方案中,所述方法还包括将自抗体-蛋白质复合物分离的蛋白质与蛋白质切割剂接触以形成多肽片段。在一些实施方案中,方法又包括小表位抗体-蛋白质复合物与蛋白质切割剂接触以形成多肽片段。在一些实施方案中,所述方法还包括在形成小表位抗体-蛋白质复合物之前,将样品与蛋白质切割剂接触以形成多肽片段,任选还包括自小表位抗体-蛋白质复合物分离多肽片段。在一个实施方案中,蛋白质切割剂包含蛋白酶。在另一实施方案中,蛋白质切割剂包含化学剂。

[0025] 另一方面,本发明提供了用于降低包含蛋白质混合物的样品复杂度的方法,其包括(a)将样品与至少一种小表位抗体接触以形成抗体-蛋白质复合物;和(b)自样品中未结合的蛋白质分离抗体-蛋白质复合物。在一些实施方案中,步骤(a)和(b)顺序地进行。在一些实施方案中,步骤(a)和(b)同时进行。在一些实施方案中,所述方法还包括自抗体-蛋白质复合物分离蛋白质。在一些实施方案中,小表位抗体结合由约3个至约5个氨基酸组成的表位。在一些实施方案中,至少一种小表位抗体包含至少约100种小表位抗体。

[0026] 在其它方面,本发明提供了用文中描述的用于降低样品复杂度的方法制备的小表位抗体-蛋白质复合物、蛋白质和/或多肽片段。

[0027] 对本领域技术人员显然地,所提及的组合及孵育所得混合物的方面也包括方法实施方案,其包括孵育多种混合物(以多种组合和/或亚组合)以形成所希望的产物。

[0028] 本发明的方法中可使用一种或一种以上(如约两、约三、约四、约五、约十、约二十或二十种以上)小表位抗体。在一些实施方案中,样品与约20、约30、约40、约50、约75、约100、约125、约150、约200、约300、约400、约500、约1000或1000种以上小表位抗体接触。在一些实施方案中,样品与至少约20、约30、约40、约50、约75、约100、约125、约150、约200、约300、约400、约500、约1000或1000种以上小表位抗体接触。在一些实施方案中,样品与低于约100、约95、约90、约85、约80、约75、约70、约65、约60、约55、约50、约45、约40、约35、约30、约25、约20、约15、约10或10种以下小表位抗体接触。在一些实施方案中,样品与至少约10、20、30、40、50、75、100、125、150、200、300、400或500种小表位抗体接触,上限分别为约20、30、40、50、75、100、125、150、200、300、400、500或1000种小表位抗

体。

[0029] 本发明还提供了使用文中描述的任意方法制备的蛋白质的方法,例如用于表征蛋白质的方法,表达概况分析方法,鉴定蛋白质的方法;用于鉴定蛋白质降解产物的方法;用于鉴定翻译后修饰中变化的方法以及用于确定样品中蛋白质质量、量和/或同一性的方法。本发明的方法还包括基因分型(蛋白质突变检测)、鉴定剪接变体、确定目的蛋白质的存在或不存在、表达概况分析的方法;用于鉴定蛋白质降解产物的方法;用于鉴定翻译后修饰中变化以及蛋白质发现的方法。

[0030] 因此,一方面,本发明提供了用于表征蛋白质的方法,其包括:(a)用文中描述的任意方法降低样品的复杂度,从而富集和/或纯化蛋白质;和(b)分析蛋白质(可互换地称为“产物”)。

[0031] 另一方面,本发明提供了用于表征蛋白质,包括分析蛋白质的方法;其中用文中描述的任意方法制备所述蛋白质。

[0032] 在一些实施方案中,分析步骤包含确定所述蛋白质的量,从而对经制备、富集和/或分离的蛋白质的量进行定量。在一些实施方案中,分析步骤包含鉴定一种或多种所述蛋白质。在一些实施方案中,使用小表位抗体结合的表位的同一性帮助鉴定所富集的蛋白质。在一些实施方案中,用任意一种或多种以下特征鉴定蛋白质:序列、质量、 m/z 比率(在包括质谱分析的实施方案中)和/或氨基酸组成。在其它实施方案中,分析步骤包含确定一种或多种蛋白质的质量。在一些实施方案中,分析步骤检测与参考蛋白质相比,蛋白质中的改变,所述参考蛋白质与没有序列改变的蛋白质序列(至少部分)相同。序列改变包括突变(如一个或多个氨基酸的缺失、替代、插入和/或颠换)、剪接变体、降解产物及糖基化的变化。

[0033] 另一方面,本发明提供了用于用质谱分析法表征蛋白质的方法,其包括:(a)用文中描述的任意方法降低样品的复杂度,从而富集和/或纯化蛋白质;和(b)分析用文中描述的任意方法分离、纯化、制备和/或分离的蛋白质(互换地称为“产物”),其中通过质谱分析法分析。

[0034] 另一方面,本发明提供了用于表征蛋白质的方法,其包括用质谱分析法分析蛋白质;其中蛋白质用文中描述的任意方法制备;其中通过质谱分析法分析。在一些实施方案中,确定了蛋白质的数量、质量和/或同一性。在一些实施方案中,所述方法还包括使用表位同一性信息。

[0035] 在一些实施方案中,质谱为基质辅助激光解吸/电离(“MALDI”)质谱分析法;表面增强激光解吸/电离(“SELDI”);和/或串连质谱分析法(例如,MS/MS、MS/MS/MS、ESI-MS/MS)。

[0036] 另一方面,本发明提供了用质谱分析法确定样品中蛋白质同一性的方法,所述方法包括:(a)用文中描述的任意方法降低样品的复杂度,从而富集和/或纯化蛋白质;(b)分析蛋白质(可互换地称为“产物”),其中通过质谱分析法分析;和(c)确定经富集的蛋白质的同一性。在一些实施方案中,所述方法还包括使用表位同一性信息。

[0037] 另一方面,本发明提供了用于用质谱分析法确定样品中蛋白质同一性的方法,所述方法包括用质谱分析法确定蛋白质的同一性,其中蛋白质用文中描述降低样品复杂度的任意方法进行制备。在一些实施方案中,所述方法还包括使用表位同一性信息。

[0038] 另一方面,本发明提供了用于蛋白质表达概况分析的方法,其中确定了一种或多种蛋白质的表达水平,其中蛋白质用文中描述用于降低样品复杂度的任意方法制备。在一些实施方案中,用质谱分析法确定表达水平。在一些实施方案中,本发明提供了用于比较两个或多个样品中蛋白质的量的方法。

[0039] 另一方面,本发明提供了用于蛋白质表达概况分析的方法,其中确定了一种或多种蛋白质的同一性,其中蛋白质用文中描述用于降低样品复杂度的任意方法制备。在一些实施方案中,用质谱分析法确定蛋白质的同一性。在一些实施方案中,所述方法还包括使用表位同一性信息。在一些实施方案中,本发明提供了用于比较两个或多个样品中蛋白质同一性的方法。

[0040] 另一方面,本发明提供了用于确定样品中存在或不存在目的蛋白质的方法,其中该方法包括在经富集的蛋白质级分中检测任意目的蛋白质,其中经富集的蛋白质级分用文中描述用于降低样品复杂度的任意方法制备,且其中目的蛋白质的检测表明样品中存在所述蛋白质。在一个实施方案中,检测包含质谱分析法。

[0041] 另一方面,本发明提供了用于确定样品中目的蛋白质的量的方法,其中方法包括定量经富集的蛋白质级分中的蛋白质的量,其中经富集的蛋白质级分用文中描述的用于降低样品复杂度的任意方法制备。在一个实施方案中,目的蛋白质的定量包含质谱分析法。

[0042] 另一方面,本发明提供了用于鉴定小表位抗体-蛋白质复合物中蛋白质的方法,其中小表位抗体-蛋白质复合物用文中描述的用于降低样品复杂度的任意方法制备。在一个实施方案中,鉴定包含质谱分析法。

[0043] 另一方面,本发明提供了用于鉴定生物标记的方法,其中该方法包括比较两个或多个经富集的蛋白质级分中的蛋白质,其中用文中描述的用于降低样品复杂度的任意方法从样品制备两个或多个经富集的蛋白质级分中的每个级分。在一些实施方案中,两个或多个样品包含来自至少一名具有疾病状况的个体和至少一名不具有疾病状况的个体的样品,生物标记的存在或不存在指示该疾病状况。在一个实施方案中,本发明提供了用于确定个体中存在或不存在疾病状况的方法,其包括确定来自该个体的样品中生物标记的水平,其中生物标记如文中描述进行鉴定,且其中生物标记的水平指示存在或不存在所述疾病状况。在一些实施方案中,两个或多个样品包含来自至少一名接受疾病状况治疗的个体和至少一名未接受疾病状况治疗的个体的样品,所述生物标记的存在或缺乏指示治疗功效。在一个实施方案中,本发明提供了用于确定个体中疾病状况治疗功效的方法,其包括确定来自该个体的样品中生物标记的水平,其中所述生物标记如文中描述进行鉴定,其中所述生物标记的水平指示治疗功效。在一些实施方案中,两个或多个样品包含来自至少一名暴露于毒素或病原体的个体及至少一名未暴露于该毒素或病原体的个体的样品,生物标记的存在或缺乏指示个体暴露于该毒素或病原体。在一个实施方案中,本发明提供了用于确定个体暴露于毒素或病原体的方法,其包括确定来自所述个体的样品中生物标记水平,其中生物标记如文中描述的进行鉴定,其中所述生物标记的水平指示暴露于所述毒素或病原体。

[0044] 另一方面,本发明提供了组合物及试剂盒,所述组合物及试剂盒包含用于本发明所述方法任一种的一种或多种小表位抗体。

[0045] 在一些实施方案中,本发明提供了包含多种小表位抗体的组合物。在一些实施方案中,多种小表位抗体结合由约3个至约5个氨基酸组成的表位。在一些实施方案中,可检

测地标记小表位抗体。在一些实施方案中,多种小表位抗体包含至少约 100 种小表位抗体。

[0046] 在一些实施方案中,本发明提供了包含多种小表位抗体的试剂盒。在一些实施方案中,多种小表位抗体结合由约 3 个至约 5 个氨基酸组成的表位。在一些实施方案中,可检测地标记小表位抗体。在一些实施方案中,多种小表位抗体包含至少约 100 种小表位抗体。

[0047] 附图简述

[0048] 图 1 分别显示对于第 2 组及第 5 组小鼠使用免疫多肽的作图多肽跨越序列的反应图式。

[0049] 图 2 显示了如在实施例 2 中描述的噬菌体 ELISA 中阳性抗体的第二次筛选结果。

[0050] 图 3 显示了如在实施例 2 中描述的,针对肽 1、6、7、8 及 9 的来源于噬菌体 L5 OP 1_15 的单链抗体的 SPR 迹线。

[0051] 发明详述

[0052] 本发明提供了方法,所述方法用一种或多种结合(一般地,特异结合)小表位的抗体,称为“小表位抗体”基于蛋白质混合物中蛋白质内小表位的存在和/或量来分级分离蛋白质混合物,从而离析、分离、制备、纯化和/或富集包含小表位的蛋白质。因此使用本发明所述方法提供了用于降低蛋白质混合物复杂度的方法,有助于样品的经富集的蛋白质组分的随后使用和/或表征。在抗体结合的小表位是已知的范围内,通过小表位抗体结合提供了与小表位抗体结合的蛋白质的氨基酸含量相关的信息。文中还描述了小表位抗体。

[0053] 总体上,所述方法包括:(a) 在允许结合的条件下,将样品与至少一种小表位抗体接触;和(b) 分离抗体-蛋白质复合体与不结合小表位抗体的蛋白质。大体上,将包含通过至少一种小表位抗体结合的一种或多种表位的蛋白质进行离析、分离、富集和/或纯化。在一些实施方案中,所述方法还包括:自抗体-蛋白质复合体分离蛋白质的步骤(c)。在一些实施方案中,所述方法还包括在将样品与至少一种小表位抗体接触的步骤(a)之前或在涉及自抗体-蛋白质复合体分离蛋白质的实施方案中,在自抗体-蛋白质复合体分离蛋白质的步骤(c)之后,用蛋白质切割剂处理样品的步骤。

[0054] 本发明所述方法对包含蛋白质的样品的分级分离是有用的,所述分级分离通过使用抗体(称为“小表位抗体”)完成,所述抗体识别在多种蛋白质中存在的表位(例如,由或基本由 3 个线性氨基酸、4 个线性氨基酸或 5 个线性氨基酸组成的表位)。适合在本发明的方法中使用的小表位抗体在文中广泛描述且在实施例中示例。由于小表位抗体的特异性,依赖通过本发明的方法中使用的小表位抗体识别的蛋白质中小表位的存在和/或量,分离、富集和/或纯化蛋白质(例如,多肽)。文中还描述了使用通过本发明的方法制备的蛋白质的方法。显然地,如文中所用的,“降低样品复杂度”包含自样品中离析、纯化、分离、富集和/或纯化蛋白质(例如多肽)。因此,本发明提供了用于纯化和/或富集蛋白质的方法,用于离析蛋白质的方法,用于分离蛋白质的方法,用于制备蛋白质用于表征的方法,用于制备蛋白质用于质谱分析法分析的方法,用于鉴定蛋白质(如一种或一组蛋白质)的方法,用于发现新蛋白质的方法及用于样品中蛋白质定量的方法。

[0055] 一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触;和(b) 分离抗体-蛋白质复合体,从而离析、分离、富集和/或纯化包含一种或多种小表位抗体结合的一种或多种表位的蛋白质。

[0056] 另一方面,本发明还提供了方法,所述方法用于纯化和/或富集蛋白质;离析蛋白

质；分离蛋白质，制备蛋白质用于表征；制备蛋白质用于质谱分析法分析；鉴定蛋白质（如一种或多种蛋白质或一组蛋白质）；发现新蛋白质；和 / 或样品中蛋白质的定量，其中所述方法包括：(a) 在允许结合的条件下将样品与一种或多种小表位抗体接触；和 (b) 分离抗体 - 蛋白质复合物。

[0057] 另一方面，本发明还包含使用通过本发明的任意方法制备的蛋白质的方法，例如，用于表征蛋白质，表达概况分析的方法，鉴定蛋白质的方法；用于鉴定蛋白质降解产物的方法；用于鉴定翻译后修饰中变化的方法，以及用于确定样品中蛋白质质量、量和 / 或同一性的方法。例如，这些方法可在诸如蛋白质发现、表达概况分析、药物发现及诊断学的领域中使用。

[0058] 在另一实施方案中，质谱分析法用于表征用本发明的任意方法制备的蛋白质。由于通过使用文中描述的小表位抗体（与起始样品相比）降低了蛋白质（包括蛋白质变体）的数目，所以用小表位抗体产生的蛋白质级分特别经得起用质谱分析法分析。在鉴定了蛋白质中存在的表位范围内，表位的氨基酸序列或含量（称为“表位序列”或“表位氨基酸含量”）还提供了用于表征及鉴定蛋白质的信息。质谱分析法方法已用于定量和 / 或鉴定蛋白质。在一些实施方案中，质谱分析法分析产生多肽质量图。用这些结果，多肽质量作图可允许鉴定对应的蛋白质。在其它实施方案中，质谱分析法分析通过串连质谱仪并产生特定序列信息。用此信息可在序列水平鉴定对应蛋白质。在一些实施方案中，用包含用本发明的任意方法制备的蛋白质的 MS 分析联合表位序列或氨基酸含量信息的方法鉴定蛋白质。

[0059] 在本发明所述方法中可使用一种或一种以上（如约 2、约 5、约 7、约 10、约 20、约 30、约 50、约 100 或 100 种以上）小表位抗体。在一些实施方案中，样品与约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体接触。在一些实施方案中，样品与至少约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体接触。在一些实施方案中，样品与低于约 100、约 95、约 90、约 85、约 80、约 75、约 70、约 65、约 60、约 55、约 50、约 45、约 40、约 35、约 30、约 25、约 20、约 15、约 10 或 10 种以下小表位抗体接触。在一些实施方案中，样品与至少约 10、20、30、40、50、75、100、125、150、200、300、400 或 500 种小表位抗体接触，上限分别为约 20、30、40、50、75、100、125、150、200、300、400、500 或 1000 种小表位抗体。

[0060] 另一方面，本发明提供了组合物及试剂盒，所述组合物及试剂盒包含在本发明的任一方法中使用的一种或多种小表位抗体。在一些实施方案中，试剂盒还包含用于文中描述的任一种方法的说明书。

[0061] 常规技术

[0062] 除了另外注明外，本发明的实践利用本领域技术范围内的分子生物学（包括重组技术）、微生物学、细胞生物学、生物化学及免疫学常规技术。此类技术在文献中充分说明，所述文献为诸如 Molecular Cloning: A Laboratory Manual, 第二版 (Sambrook 等人, 1989) Cold Spring Harbor Press; Oligonucleotide Synthesis (M. J. Gait, ed., 1984) ; Methods in Molecular Biology, Humana Press; Cell Biology: A Laboratory Notebook (J. E. Cellis, ed., 1998) Academic Press; Animal Cell Culture (R. I. Freshney, ed., 1987) ; Introduction to Cell and Tissue Culture (J. P. Mather 和

P. E. Roberts, 1998) Plenum Press; Cell and Tissue Culture: Laboratory Procedures (A. Doyle, J. B. Griffiths 和 D. G. Newell, eds., 1993-8) J. Wiley 和 Sons; Methods in Enzymology (Academic Press, Inc.); Handbook of Experimental Immunology (D. M. Weir 和 C. C. Blackwell, eds.); Gene Transfer Vectors for Mammalian Cells (J. M. Miller 和 M. P. Calos, eds., 1987); Current Protocols in Molecular Biology (F. M. Ausubel 等人, eds., 1987); PCR: The Polymerase Chain Reaction (Mullis 等人, eds., 1994); Current Protocols in Immunology (J. E. Coligan 等人, eds., 1991); Short Protocols in Molecular Biology (Wiley 和 Sons, 1999); Immunobiology (C. A. Janeway 和 P. Travers, 1997); Antibodies (P. Finch, 1997); Antibodies: a practical approach (D. Catty., ed., IRL Press, 1988-1989); Monoclonal antibodies: a practical approach (P. Shepherd 和 C. Dean, eds., Oxford University Press, 2000); Using antibodies: a laboratory manual (E. Harlow 和 D. Lane (Cold Spring Harbor Laboratory Press, 1999); 以及 The Antibodies (M. Zanetti 和 J. D. Capra, eds., Harwood Academic Publishers, 1995)。

[0063] 定义

[0064] “抗体”为免疫球蛋白分子,其能够通过位于该免疫球蛋白分子的可变区中的至少一个抗原识别位点特异结合靶标,如糖、多核苷酸、脂类、多肽等。如文中所用的,该术语不仅仅包括完整多克隆或单克隆抗体,也包括其片段(如 Fab、Fab'、F(ab')₂、Fv)、单链(ScFv)、其突变体、包含抗体部分的融合蛋白及包含具有所需要的特异性的抗原识别位点的免疫球蛋白分子的任意其它经修饰的构型。抗体包括任意类别的抗体,如 IgG、IgA 或 IgM(或其亚类),且抗体不必属于任意具体类别。取决于抗体重链的恒定结构域的氨基酸序列,免疫球蛋白可分为不同类。主要有五类免疫球蛋白: IgA、IgD、IgE、IgG 及 IgM,且这些类中的几种又可分为亚类(同种型),例如, IgG1、IgG2、IgG3、IgG4、IgA1 及 IgA2。对应于免疫球蛋白不同类的重链恒定结构域分别称为 α 、 δ 、 ϵ 、 γ 及 μ 。不同类免疫球蛋白的亚单位结构及三维构型是公知的。

[0065] “Fv”为包含完全抗原识别及抗原结合位点的抗体片段。在两链 Fv 种类中,此区域由以紧密、非共价连接的一条重链及一条轻链可变结构域的二聚体组成。在单链 Fv 种类中,一条重链及一条轻链可变结构域可通过柔性多肽接头共价连接从而轻链及重链可以在两链 Fv 种类中类似的二聚体结构连接。这种构型中,每个可变结构域的三个 CDR 相互作用以限定在 VH-VL 二聚体表面上抗原结合特异性。然而,甚至单个可变结构域(或仅仅包含对抗原特异的 3 个 CDR 的 Fv 的一半)具有识别及结合抗原的能力,即使大体上以比完整结合位点较低的亲和力识别及结合抗原。

[0066] Fab 片段也包含轻链的恒定结构域及重链的第一个恒定结构域(CH1)。Fab' 片段通过在重链 CH1 结构域的羧基末端加入几个残基,包括来自抗体铰链区的一个或多个半胱氨酸而与 Fab 片段不同。

[0067] “单克隆抗体”是指同质性抗体群,其中单克隆抗体包含涉及选择性结合抗原的氨基酸(天然存在或非天然存在的)。在它们针对单个抗原性位点方面,单克隆抗体群(相对于多克隆抗体)是高度特异的。术语“单克隆抗体”不仅仅包含完全单克隆抗体及全长单克隆抗体,也包含其片段(如 Fab、Fab'、F(ab')₂、Fv)、单链(ScFv)、其突变体、包含抗体部分的融合蛋白及包含具有所需要的特异性的抗原识别位点及结合抗原的能力的免疫球蛋白。

白分子的任意其它经修饰的构型（参见抗体定义）。抗体来源或抗体制备方式（例如，通过杂交瘤、噬菌体选择、重组表达、转基因动物，等）不欲受到限制。

[0068] 文中可互换使用的术语“多肽”、“寡肽”、“肽”及“蛋白质”是指任意长度的氨基酸聚合物。聚合物可为线性或分支的，它可包含经修饰氨基酸，它可通过非氨基酸断开。该术语也包含经天然修饰或通过干扰修饰的氨基酸聚合物；所述修饰例如二硫键形成、糖基化、脂化、乙酰化、磷酸化或任意其它操作或修饰，如用标记组分缀合。所述定义中也包括，例如包含一种或多种氨基酸类似物（包括，例如非天然氨基酸等等）以及本领域已知的其它修饰的多肽。

[0069] 与抗体“特异结合”或“优先结合”（文中互换使用）的表位是本领域充分理解的术语，且确定特异或优先结合的方法也是本领域公知的。如果分子与特定细胞或物质比与备选细胞或物质反应或结合地更频繁、更快速、持续时间更长和 / 或亲和力更高，那么称此分子显示“特异结合”或“优先结合”。如果抗体与靶结合比与其它物质结合具有更大亲和力、亲合力、更容易和 / 或持续时间更长，那么此抗体与靶“特异结合”或“优先结合”。例如与表位特异或优先结合的抗体为比与其它表位以更大亲和力、亲合力、更容易和 / 或持续时间更长结合该表位的抗体。通过阅读这个定义也应当理解，例如，与第一种靶特异或优先结合的抗体（或部分或表位）与第二种靶可以或不可以特异或优先结合。同样地，“特异结合”或“优先结合”不必需要（虽然它可包括）专一结合。一般地，但不是必需的，所谈及的结合表示优先结合。

[0070] “样品”包含多种样品类型，包括获得自个体的样品。该定义包含血液及其它生物来源的液体样品、固体组织样品，如活体解剖标本或其来源的组织培养物或细胞及其子代。样品可来自微生物（例如，细菌、酵母、病毒、类病毒、霉菌、真菌）、植物或动物，包括哺乳动物如人、啮齿动物（如小鼠和大鼠）及猴（及其它灵长类）。样品可包含单个细胞或一个以上的细胞。该定义还包括在获得样品后对其以任意方法进行操作的产品，所述方法为诸如通过用试剂处理、增溶，或富集某些组分，如蛋白质或多核苷酸。术语“样品”包含临床样品，也包括培养的细胞、细胞上清液、细胞裂解物、血清、血浆、生物液、在动物中增殖的人组织，及组织样品。样品的实例包括血液、血浆、血清、尿、粪、脑脊液、滑液、羊膜液、唾液、肺灌洗液、精液、奶、乳头抽吸液、前列腺液、粘液及泪液。

[0071] 样品的“复杂度”指不同蛋白质种类的数目，包括不同蛋白质的数目以及不同蛋白质变体（包括剪接变体、多态性及蛋白质降解产物）的数目。

[0072] “检测”是指鉴定（确定）待检测物体或物质的存在、缺乏和 / 或量，如文中描述的，检测可为定性和 / 或定量的。

[0073] 如文中所用的，除非另外注明，单数形式“一个”、“一种”及“这个”包括复数形式。例如，“一种”抗体包括一种或多种抗体，“一种蛋白质”表示一种或多种蛋白质。

[0074] 发明方法

[0075] 关于文中描述的所有方法，谈及小表位抗体时，其也包括包含一种或多种这些抗体的组合物。这些组合物可以还包含用以增强稳定性的合适的赋形剂，如可药用赋形剂以及缓冲剂和 / 或组分，它们是本领域公知的。

[0076] 用于降低样品复杂度的方法

[0077] 本发明提供了方法，所述方法基于蛋白质混合物中的蛋白质中蛋白质内小表位的

存在或缺乏或量,用结合(一般地,特异结合)小表位的一种或多种抗体,称为“小表位抗体”以分级分离蛋白质混合物,从而产生蛋白质级分,所述蛋白质包含并富集所述小表位。如文中所用的,“经富集的”是指蛋白质或肽的浓度和/或纯度与该蛋白质或肽所来源的样品中的该蛋白质或肽的浓度和/或纯度相比增加。因此本发明所述方法的使用提供了降低蛋白质混合物的复杂度,方便样品的经富集的蛋白质组分的随后使用和/或表征的方法。在抗体结合的小表位的氨基酸序列或组成是已知的范围内,被小表位抗体结合提供了关于小表位抗体结合的蛋白质的氨基酸序列和/或含量的信息。如文中描述的,表位同一性信息(即,小表位抗体识别的氨基酸含量和/或序列)可与本发明的其它方法组合使用以例如,鉴定蛋白质。文中还描述了小表位抗体。

[0078] 本发明还提供了方法,这些方法用于纯化和/或富集蛋白质;离析蛋白质;分离蛋白质,制备蛋白质用于表征(例如,随后分析);制备蛋白质用于质谱分析法分析;鉴定蛋白质;发现新蛋白质;和/或定量样品中的蛋白质。

[0079] 总体上,所述方法包括:(a)在允许结合条件下,将样品与至少一种小表位抗体接触;和(b)分离抗体-蛋白质复合体。在一个实施方案中,步骤(a)与步骤(b)顺序发生。在另一实施方案中,步骤(a)与步骤(b)同时发生。大体上,将包含被至少一种或多种小表位抗体结合的一种或多种表位的蛋白质进行离析、分离、富集和/或纯化(即,自最初样品的环境中除去)。在一些实施方案中,所述方法还包括步骤(c):自抗体-蛋白质复合体分离蛋白质。在一些实施方案中,所述方法还包括用蛋白质切割剂处理样品。在一个实施方案中,在样品与至少一种小表位抗体接触的步骤(a)之前加入蛋白质切割剂。在另一实施方案中,在自抗体-蛋白质复合体分离蛋白质的步骤(c)之后加入蛋白质切割剂。

[0080] 本发明的方法对分级分离包含蛋白质(如多肽)的样品是有用的,所述方法通过使用识别存在于多种蛋白质中的表位(例如,由或者基本上由3个线性氨基酸、4个线性氨基酸或5个线性氨基酸组成的表位)的抗体(称为“小表位抗体”)来完成。适合在本发明的方法中使用的小表位抗体在文中详尽描述且在实施例例证。依靠小表位抗体的特异性,根据本发明方法中使用的小表位抗体识别的蛋白质中的小表位的存在和/或量,分离、富集和/或纯化蛋白质或肽(例如,多肽)。显然,如文中所用的,“降低样品复杂度”包括从样品中离析、纯化、分离、富集和/或纯化蛋白质或肽(例如多肽)(包括自样品的环境除去所述蛋白质或肽)。

[0081] 因此,一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括:(a)在允许结合的条件下将样品与一种或多种小表位抗体接触;和(b)分离抗体-蛋白质复合体,从而离析、分离、富集和/或纯化包含被一种或多种小表位抗体结合的一种或多种表位的蛋白质。在一些实施方案中,所述方法还包括:自抗体-蛋白质复合体分离蛋白质的步骤(c)。

[0082] 在一个实施方案中,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括分离小表位抗体-蛋白质复合体,从而富集包含小表位抗体结合的表位的蛋白质;其中通过将样品与小表位抗体接触产生复合体。在另一实施方案中,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括分离多种小表位抗体-蛋白质复合体,从而富集包含被多种小表位抗体结合的表位的蛋白质,其中通过将样品与多种小表位抗体接触产生复合体。

[0083] 另一方面,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括自小表位抗体-蛋白质复合体分离蛋白质,从而富集包含被小表位抗体结合的抗体的蛋白质;其中在允许结合的条件下,通过将样品与小表位抗体接触产生小表位抗体-蛋白质复合体,从而产生小表位抗体-蛋白质复合体;并分离抗体-蛋白质复合体与样品中未结合的蛋白质(如果存在)。在一个实施方案中,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括自小表位抗体-蛋白质复合体分离多种蛋白质,从而富集包含被多种小表位抗体结合的表位的蛋白质,其中在允许结合样品中蛋白质的条件下,通过将样品与多种小表位抗体接触产生小表位抗体-蛋白质复合体,从而产生小表位抗体-蛋白质复合体,并分离抗体-蛋白质复合体与样品中未结合的蛋白质(如果存在)。

[0084] 显然,一个或多个步骤可组合和/或顺序进行(常常是以任意次序,只要能形成需要的产物),并且,显然地,本发明包括文中描述的步骤的多种组合。显然且在文中描述的是,本发明包括这样的方法,其中最初,或第一个步骤是文中描述的任意步骤。本发明的方法包含实施方案,在所述实施方案中较晚的“下游”步骤为最初步骤。

[0085] 在一些实施方案中,所述方法还包括用蛋白质切割剂处理样品的步骤,从而产生多肽片段。在包括自抗体-蛋白质复合体分离蛋白质的步骤(c)的实施方案中,在样品与至少一种小表位抗体接触的步骤(a)之前,和/或自抗体-蛋白质复合体分离蛋白质的步骤(c)之后,可用蛋白质切割剂处理样品。蛋白质切割剂可为酶(如胰凝乳蛋白酶或胰蛋白酶)或化学剂(如溴化氰)。蛋白质切割剂及用蛋白质切割剂处理的方法是本领域公知的并在文中进一步描述。

[0086] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括(a)在允许结合的条件下,将样品与一种或多种小表位抗体接触;(b)分离蛋白质-抗体复合体,从而富集包含被一种或多种小表位抗体结合的一种或多种表位的蛋白质;(c)自蛋白质-抗体复合体分离蛋白质;和(d)用蛋白质切割剂处理蛋白质,从而产生多肽片段。

[0087] 另一方面,本发明提供了用于降低样品复杂度的方法,所述方法包括(a)在允许结合的条件下,将样品与一种或多种小表位抗体接触,从而形成抗体-蛋白质复合体;和(b)用蛋白质切割剂处理抗体-蛋白质复合体以产生多肽片段。

[0088] 另一方面,本发明提供了用于降低蛋白质样品复杂度的方法,所述方法包括:(a)用蛋白质切割剂处理样品,从而产生多肽片段;(b)在允许结合的条件下将多肽片段与一种或多种小表位抗体接触,从而产生抗体-多肽复合体;和(c)分离抗体-多肽复合体,从而富集包含被一种或多种小表位抗体结合的一种或多种表位的多肽。

[0089] 在本发明方法中可使用一种,或一种以上(如约两、约三、约五、约十、约二十、约一百或一百种以上)小表位抗体。在一些实施方案中,样品与约20、约30、约40、约50、约75、约100、约125、约150、约200、约300、约400、约500、约1000或1000种以上小表位抗体接触。在一些实施方案中,样品与至少约20、约30、约40、约50、约75、约100、约125、约150、约200、约300、约400、约500、约1000或1000种以上小表位抗体接触。在一些实施方案中,样品与低于约100、约95、约90、约85、约80、约75、约70、约65、约60、约55、约50、约45、约40、约35、约30、约25、约20、约15、约10或10种以下小表位抗体接触。在一些实施方案中,样品与至少约10、20、30、40、50、75、100、125、150、200、300、400或500任一种的小表位抗体接触,上限为约20、30、40、50、75、100、125、150、200、300、400、500或1000任一种

的小表位抗体。

[0090] 应当理解样品也可以与其它蛋白质结合活性剂,包括不是小表位抗体的蛋白质及其它蛋白质结合活性剂接触。此类活性剂可在用小表位抗体处理前或处理后同时、顺序使用。

[0091] 在一些实施方案中,在样品与一种或多种小表位抗体接触的步骤之前或同时,用结合一种或多种蛋白质的一种或多种抗体处理样品,优选已知在样品中丰富的蛋白质。例如,在血清样品中,预处理可包含结合白蛋白、免疫球蛋白和 / 或其它丰富蛋白质的抗体。在一个实施方案中,在与结合一种或多种已知丰富蛋白质的一种或多种抗体接触之前,将样品中的蛋白质用蛋白质切割剂切割。在另一实施方案中,在与结合一种或多种已知蛋白质,如丰富蛋白质的一种或多种抗体接触之后,将样品中的蛋白质用蛋白质切割剂切割。在一个实施方案中,在与一种或多种小表位抗体接触前,自样品中除去结合的蛋白质(如丰富蛋白质)。在一个实施方案中,所述方法包括通过用与样品中一种或多种已知蛋白质,如丰富蛋白质结合的一种或多种抗体处理(任选接着除去结合的蛋白质)来“压实”(debulking)样品,用蛋白质切割剂切割样品中的蛋白质,及将切割的蛋白质与一种或多种小表位抗体接触。在另一实施方案中,所述方法包括用蛋白质切割剂处理样品,通过用与一种或多种已知蛋白质,如丰富蛋白质和 / 或样品中经切割的多肽片段结合的一种或多种抗体处理压实样品(任选接着除去结合的蛋白质和 / 或多肽片段),及将剩余蛋白质和 / 或经切割的多肽片段与一种或多种小表位抗体接触。在另一实施方案中,所述方法包括通过用结合一种或多种已知蛋白质,如丰富蛋白质的一种或多种抗体处理(任选接着除去丰富蛋白质)压实样品,将样品与至少一种小表位抗体接触以形成抗体-蛋白质复合物,及用蛋白质切割剂处理抗体蛋白质复合物。

[0092] 还理解用小表位抗体处理后剩余的样品的蛋白质组分(即未结合的组分)也可适合在本发明的方法中使用,所述方法使用用本发明的方法产生的蛋白质。因此,在一些实施方案中,使用用本发明的方法产生的蛋白质的方法包括使用此未结合的蛋白质组分。

[0093] 用于抗体结合及分离抗体-蛋白质复合物的方法及条件是本领域公知的且在文中进一步描述。一般地,样品与小表位抗体接触时部分或全部变性,但不是在每个实施方案都需要变性。在一些实施方案中,将样品与两种或多种抗体接触的步骤(a)是顺序的(如当一种抗体与样品接触,然后除去时,另一抗体与样品接触并除去,等等)。在其它实施方案中,例如,当一组抗体与样品同时接触时,接触两种或多种抗体的步骤(a)是平行的。在一些实施方案中,两种或多种抗体的几个组与样品顺序接触,例如,组1接触并除去,组2接触并除去等等。

[0094] 如在定义中注明的,及如文中所用的,“样品”包含多种样品类型,包括获得自个体的样品类型。在一些实施方案中,样品包含血液、血浆、血清、尿、粪、脑脊液、滑液、羊水、唾液、肺灌洗液、精液、奶、乳头抽吸液、前列腺液、粘液及泪液。用于在本发明方法中使用的合适的样品在文中进一步描述。

[0095] 使用用本发明的方法分离(富集)的蛋白质的方法

[0096] 用本发明的方法分离或富集的蛋白质可用于多种目的。为了阐明,描述了用通过本发明的方法富集和 / 或纯化蛋白质表征蛋白质的方法。在一些实施方案中,用质谱分析法表征蛋白质,从而可定量和 / 或鉴定蛋白质。还描述了基因分型(蛋白质突变检测)、鉴

定剪接变体、确定目的蛋白质的存在或不存在、表达概况分析的方法；用于鉴定蛋白质降解产物的方法；用于鉴定翻译后修饰中变化的方法及蛋白质发现的方法。

[0097] 为了简便和方便，一般谈及“蛋白质”。应当理解谈及蛋白质时其包含“多肽”（可互换地称为“多肽片段”）。如从文中讨论显然的，在一些实施方案中，用蛋白质切割剂产生多肽片段。

[0098] 表征蛋白质的方法

[0099] 本发明提供了用于表征（例如，检测（存在或不存在）和 / 或定量）目的蛋白质（一般地，多肽片段）的方法。在一些实施方案中，用本发明的方法产生一种或多种样品级分，每种级分包含的蛋白质少于起始样品中的，有助于随后表征级分中包含的所述蛋白质。特别地，如文中进一步描述的，预计有待加强使用质谱分析法的表征。

[0100] 因此，本发明提供了用于表征蛋白质的方法，其包括：(a) 用文中描述的任意方法降低样品的复杂度，从而对蛋白质进行富集和 / 或纯化；及 (b) 分析通过文中描述的任意一种方法分离的蛋白质（可互换地称为“产物”）。

[0101] 另一方面，本发明提供了用于表征蛋白质的方法，其包括：分析蛋白质（可互换地称为“产物”），其中用文中描述的用于降低样品复杂度的任意方法（包括：用于纯化和 / 或富集蛋白质的方法，用于离析蛋白质的方法，用于分离蛋白质的方法，用于制备蛋白质级分用于表征的方法，用于制备蛋白质级分用于质谱分析法分析的方法，用于鉴定蛋白质（如一种或多种蛋白质，或一组蛋白质）的方法，用于发现新蛋白质的方法及用于定量样品中的蛋白质的方法）制备蛋白质。

[0102] 通过本领域已知或文中描述的任意方法进行分析步骤。用于分析蛋白质的方法是本领域公知的，且包括：十二烷基硫酸钠 - 聚丙烯酰胺凝胶电泳（“SDS-PAGE”）、等电聚焦、通过诸如高压液相层析、FPLC、薄层层析、亲和层析、凝胶过滤层析、离子交换层析的技术分离，及其它标准生物化学分析、免疫检测、蛋白质测序、用蛋白质阵列分析、质谱分析等等。因此，本发明包括如应用于文中方法的任意产物的其他分析和 / 或定量方法。在一些实施方案中，分析步骤包括确定所述蛋白质的量，从而定量所制备、富集和 / 或分离的蛋白质的量。应当理解可用定量和 / 或定性方法确定富集的蛋白质的量。确定蛋白质产物的量包括确定产物存在或不存在。在一些实施方案中，分析步骤包含鉴定一种或多种所述蛋白质。用于鉴定蛋白质的方法是本领域已知的，且包括：免疫检测、蛋白质测序等等。在一些实施方案中，鉴定基本上所有经富集的蛋白质（从样品纯化或富集）。在一些实施方案中，小表位抗体结合的表位的同一性用于帮助鉴定经富集的蛋白质。在一些实施方案中，用任意一种或多种以下特征鉴定蛋白质：序列、质量、 m/z 比率（在涉及质谱分析的实施方案中）、氨基酸组成及提供鉴定蛋白质的足够信息的任意其它方法。如文中所用的，“鉴定”包括鉴定已知（以前经表征的蛋白质）以及发现以前未知或未表征的蛋白质（包括蛋白质变体如突变蛋白质、差别修饰的蛋白质（例如，不同糖含量）及剪接变体）。在一些实施方案中，鉴定了多种、很多种或非常多种蛋白质。在其它实施方案中，鉴定了至少约 2、3、4、5、10、20、30、40、50、60、70、80、90、100、500 或 1000 或 1000 种以上的蛋白质。

[0103] 在其它实施方案中，分析步骤包括确定一种或多种蛋白质的质量。

[0104] 在一些实施方案中，分析步骤包括分析与参考蛋白质相比蛋白质中的改变，所述参考蛋白质与没有序列改变的蛋白质序列（至少部分）相同。序列改变可为基因组序列中

存在的序列改变或可为在基因组 DNA 序列中没有反映出来的序列改变,例如,由于翻译后修饰改变和 / 或 mRNA 加工的改变,包括剪接变体,和 / 或翻译后修饰如糖基化的量的改变及蛋白质降解或副产物。序列改变包括突变(如一个或多个氨基酸的缺失、替代、插入和 / 或颠换)。

[0105] 应当理解小表位抗体与表位的同一性(序列)可与文中描述的任意方法组合使用以例如鉴定蛋白质。

[0106] 用质谱分析法表征蛋白质的方法

[0107] 在一些实施方案中,质谱分析法(MS)用于表征用本发明的方法分离的蛋白质。一般地,在涉及质谱分析的实施方案中,用蛋白质切割剂处理样品(从而产生多肽片段),但在不是在每个实施方案中都需要用切割剂处理。在一些实施方案中,在样品与小表位抗体接触前用蛋白质切割剂处理样品。在其它实施方案中,在通过与小表位抗体接触、分离抗体-蛋白质复合体及从蛋白质-抗体复合体分离蛋白质富集蛋白质级分之后用蛋白质切割剂处理样品。如文中提到的,因为用本发明的方法产生的蛋白质级分复杂性低于起始样品,所以用本发明的方法产生的蛋白质(如多肽片段)特别适合用质谱分析法分析。在蛋白质中存在的表位是已知的,例如,小表位抗体识别的关连表位(cognate epitope)用于纯化和 / 或富集包含所述蛋白质的蛋白质级分的情况下,所述表位的氨基酸序列或含量(称为“表位序列”或“表位含量”)提供了用于表征及鉴定所述蛋白质的其他信息。

[0108] 用于质谱蛋白质分析的方法是本领域已知的且在文中进一步描述。质谱分析方法已经用于定量和 / 或鉴定蛋白质(参见,例如, Li 等人(2000) *Tibtech* 18:151-160; Rowley 等人(2000) *Methods* 20:383-397; 及 Kuster 和 Mann(1998) *Curr. Opin. Structural Biol.* 8:393-400.)。还研发了质谱分析技术,其允许至少部分从头测序分离的蛋白质。Chait 等人(1993) *Science* 262:89-92; Keough 等人(1999) *Proc. Natl. Acad. Sci. USA* 96:7131-6; 在 Bergman(2000) *EXS* 88:133-44 中综述。

[0109] 多肽质量作图提供了所分析的蛋白质或蛋白质级分的基于它的氨基酸组成的多肽质量指纹图。可用例如 MALDI-TOF 平台获得多肽质量作图,在所述 MALDI-TOF 平台中基质辅助激光解吸 / 电离(MALDI)用于离子化目的多肽而离子化多肽的飞行时间分布提供了每种多肽的质量与电荷比,其可用于查询蛋白质序列数据库。产生的多肽质量指纹图包含基于质量及电荷测定的氨基酸组成。用这些结果,小组多肽质量匹配可提供用于鉴定对应蛋白质的足够信息。

[0110] 在用于通过 MS 鉴定蛋白质的第二种方法中,将混合物中的个体多肽片段化以产生序列信息。通过来自液相的电喷射(ESI)将多肽离子化,且然后喷射到串联的质谱仪中,所述质谱仪通过主要断裂多肽键(碰撞诱导的解离)能溶解混合物中的多肽、分离目的多肽和将个体多肽种类解离成为组成性包含氨基末端及羧基末端的片段。产生的质谱包含亲本离子以及来自包含氨基末端及羧基末端的片段的离子的两个重叠质量梯。因为梯的每个成员的质量与电荷比(称为“ m/z ”)与该系列中它的最接近的质量邻居有一个氨基酸不同,所以可产生部分一级序列并将其用于查询蛋白质及所翻译的 DNA 序列数据库。此质谱分析平台提供了来自几种多肽的特定序列信息,所述特定序列信息对于蛋白质鉴定通常比反映多肽氨基酸组成的一组多肽质量(如通过其它平台,包括 SELDI-TOF 产生的)更有用。

[0111] 如下文进一步描述的,质谱分析方法还允许定量所分析的蛋白质。

[0112] 其它质谱分析方法是本领域公知的,且包括:基质辅助激光解吸/电离(“MALDI”)质谱分析法;表面增强激光解吸/电离(“SELDI”);串联质谱分析法(例如,MS/MS、MS/MS/MS、ESI-MS/MS等等)。在一些实施方案中,使用激光解吸/电离质谱仪进行串联质谱分析法,所述质谱仪又连接四极飞行时间质谱仪 QqTOF MS(参见例如, Krutichinsky 等人, W099/38185)。以前描述了诸如 MALDI-QqTOFMS(Krutichinsky 等人, W0 99/38185; Shevchenko 等人 (2000) Anal. Chem. 72 :2132-2141)、ESI-QqTOF MS(Figeys 等人 (1998) Rapid Comm' ns. Mass Spec. 12-1435-144) 及芯片毛细管电泳(芯片-CE)-QqTOF MS(Li 等人 (2000) Anal. Chem. 72 :599-609) 的方法。质谱分析仪及本发明方法中使用它们的技术是本领域技术人员公知的。本领域技术人员将理解质谱仪的任意组件(例如,解吸源、质量分析仪、检测等等)可与文中描述或本领域技术人员已知的其它合适的组件组合。关于质谱仪的额外信息,参见,例如, Principles of Instrumental Analysis, 3rd ed., Skoog, Saunders College Publishing, Philadelphia, 1985; 及 Kirk-Othmer Encyclopedia of Chemical Technology, 4th ed. Vol. 15 (John Wiley & Sons, New York 1995), pp. 1071-1094。

[0113] 质谱的数据分析

[0114] 用质谱分析法分析获得的质谱数据可用于获得关于用本发明方法获得的经富集的蛋白质产物的量和/或同一性的信息。可用任意合适方法(例如,通过视觉,通过计算机,等等)分析通过多肽解吸及检测产生的数据。在一个实施方案中,使用可编程的数字计算机分析数据。计算机含有接收输入数据的代码,所述输入数据为在基质上特定可寻址位置接收的多种分子质量的信号强度的数据。此数据可指出所检测产物的数目,任选包括峰值信号的强度及对于每种检测产物测定的分子质量。

[0115] 数据分析可包括确定所检测的峰值(例如,特定质量-电荷值或值的范围的峰值)的信号强度(例如,峰高度)和除去“异常值”(与预定统计分布偏离的数据)的步骤。可对观察到的峰进行归一化,该方法中计算相对于某个参比的每个峰的高度。例如,参比可为通过仪器及化学品(例如,能量吸收分子)产生的本底噪声,所述本底噪声在刻度上设为零。然后对每种多肽或其它物质检测的信号强度可以以希望的刻度(例如,100)以相对强度的形式显示。备选地,可承认样品是标准的,从而来自标准的峰可用作参比以计算为所检测的每种亲和标记产物观察到的信号的相对强度。可用软件程序如 Biomarker Wizard 程序(Ciphergen Biosystems, Inc., Fremont, Calif.) 帮助分析质谱。

[0116] 在一些实施方案中,通过用可编程的数字计算机执行算法部分确定样品中存在的一种或多种蛋白质的量。所述算法鉴定第一个样品中第一个质谱中及第二个样品中的第二个质谱中的至少一个峰值。然后算法比较第一个质谱峰值的信号强度与第二个质谱峰值的信号强度。相对信号强度表明第一及第二样品中存在的蛋白质的量。含有已知量蛋白质的标准可作为第二个样品进行分析以对第一个样品中存在的蛋白质的量更好地定量。在某些实施方案中,也可确定第一及第二个样品中蛋白质的同一性(参见下文)。本发明还提供了用于确定蛋白质同一性的方法。在某些实施方案中,可编程的数字计算机用于存取包含一个或多个质谱的数据库。然后用可编程的数字计算机执行算法以对每个预计的质谱确定至少第一次测定。第一次测定表明所述蛋白质的质谱与多种预计的质谱的每一个之间的拟合接近性(closeness-of-fit)。

[0117] 通过用可编程的数字计算机执行比较 MS 数据与数据库中的记录的算法,质谱的

数据可用于鉴定蛋白质。通过 MS 方法分析时,每个分子提供了特征性的质谱 (MS) 数据 (也称作质谱“签名”或“指纹”)。可通过将此数据与包含真实的或理论 MS 数据或蛋白质序列信息等的数据库比较分析此数据。此外,可将蛋白质切割成片段用于 MS 分析。获得自片段的 MS 分析的信息也与数据库进行比较以鉴定样品中的蛋白质 (例如,蛋白质) (参见,例如, Yates (1998) *J. Mass Spec.* 33 :1-19 ;Yates 等人,美国专利号 5,538,897 ; Yates 等人,美国专利号 6,017,693 ;PCT 申请号 WO 00/11208 及 Gygi 等人 (1999) *Nat. Biotechnol.* 10 :994-999)。现在易于在因特网上获得有助于解释质谱,特别是蛋白质质谱及公共结构域序列数据库的挖掘的软件来源来促进蛋白质鉴定。这些软件来源中有 ProteinProspector (<http://prospector.ucsf.edu>)、PROWL (<http://prowl.rockefeller.edu>) 及 Mascot Search Engine (Matrix Science Ltd., London, UK, www.matrixscience.com)。

[0118] 在某些实施方案中,获得自此数据的 MS 数据及信息与由关于蛋白质的数据及信息组成的数据库进行比较。例如数据库可由核苷酸或氨基酸序列组成。数据库可由表达序列标签 (EST) 的核苷酸或氨基酸序列组成。备选地,数据库可由核苷酸或氨基酸水平上的基因序列组成。数据库可包括但不限于一组核苷酸序列、氨基酸序列或任意物种的基因组中包括的核苷酸序列的翻译。

[0119] 通常通过任选由计算机执行的计算机程序或搜索算法分析关于蛋白质信息,例如,核苷酸或氨基酸序列的数据库。对来自序列数据库的信息搜索与获得自本发明的方法的数据及信息的最佳匹配 (参见例如, Yates (1998) *J. Mass Spec.* 33 :1-19 ;Yates 等人,美国专利号 5,538,897 ;Yates 等人,美国专利号 6,017,693)。可用任意适当算法或计算机程序搜索数据库。搜索算法及数据库不断更新,并且根据本发明使用此类更新版本。程序或数据库的实例可在万维网 (WWW) 上 <http://base-peak.wiley.com/>、<http://mac-mann6.embl-heidelberg.de/MassSpec/Software.html>、<http://www.mann.emblheidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html>、<ftp://ftp.ebi.ac.uk/pub/databases/> 及 <http://donatello.ucsf.edu> 找到。美国专利号 5,632,041 ;5,964,860 ; 5,706,498 及 5,701,256 也描述了用于序列比较的算法或方法。其它数据库的实例包括 Genpept 数据库、GenBank 数据库 (在 Burks 等人 (1990) *Methods in Enzymology* 183 :3-22 中描述)、EMBL 数据库 (在 Kahn 等人 (1990) *Methods in Enzymology* 183 :23-31 中描述)、蛋白质序列数据库 (在 Barker 等人 (1990) *Methods in Enzymology* 183 :31-49 中描述)、SWISS-PROT (在 Bairoch 等人 (1993) *Nucleic Acids Res.* 21 :3093-3096 中描述) 及 PIR-International (在 (1993) *Protein Seg. Data Anal.* 5 :67-192 中描述)。

[0120] 在一些实施方案中,通过小表位抗体识别的表位的氨基酸序列 (称为“表位序列”)与数据库搜索信息及搜索算法联合使用以加强蛋白质的鉴定。例如,在 MS 数据及通过比较由涉及蛋白质的数据及信息组成的数据库获得的信息分析之前或之后,表位的氨基酸序列可用于改进数据分析。例如,通过自列表中排除不包括所述表位序列的成员可精选蛋白质同一性候选物的初步列表。在另一实例中,可编辑或理论上产生包含给定表位序列的所有蛋白质的数据库。然后用本领域已知的数据分析方法进一步分析该数据库。

[0121] 一般通过任选通过计算机执行的计算机程序或搜索算法分析涉及蛋白质信息例如,核苷酸或氨基酸序列的数据库。

[0122] 在又一实施方案中,产生新数据库用于比较质谱分析法确定的 MS 数据,例如经切割蛋白质及多肽片段的质量或质谱。例如,产生包含小表位抗体识别的表位的所有多肽片段的理论数据库。该数据库可与任意数据分析工具及文中描述的方法联合使用。

[0123] 在一些实施方案中,来自质谱的多肽的质量用于查询数据库中提供最接近拟合的来自核酸序列的蛋白质或预计蛋白质的那些质量。以这种方式,在无氨基酸序列时可快速鉴定未知蛋白质。在本发明的其它实施方案中,自其多肽片段提供的质量可与来自提供最接近拟合的核酸序列的蛋白质或预计蛋白质数据库的预计质谱比较。

[0124] 如果来自序列数据库的序列或者模拟切割序列落入与从亲本或者片段分子的 MS 数据产生的序列的相似序列同源性的所希望的范围,那么称所述来自序列数据库的序列或者模拟切割序列“匹配”或“命中”。以这种方式,可快速确定蛋白质或其片段的同一性。研究者可根据每个具体分析定制或改变可接受的序列同源性比较值的范围。

[0125] 应当理解为了方便,谈及蛋白质“同一性”。应当理解文中描述的方法同样适用于确定突变(如氨基酸替代、颠换、插入或缺失)及其它蛋白质变体,如剪接变体、降解产物和/或差别翻译后修饰(例如,在糖基化水平的变异)的存在或不存在。

[0126] 在一些实施方案中,通过检测 m/z 比相对于参照 m/z 比的改变确定突变的存在或不存在。

[0127] 在一些实施方案中,通过比较内糖基化酶(endoglycosylase)处理的样品与参考样品(例如,没有用内糖基化酶处理的样品)确定翻译后修饰的水平(或水平的改变),从而确定翻译后修饰的水平。

[0128] 表达概况分析

[0129] 本发明的方法适合用于确定样品中一种或多种蛋白质的表达水平。如以上描述地,可通过如文中和/或本领域已知的多种方法检测和/或定量经富集和/或纯化的蛋白质级分。在一些实施方案中,用质谱分析法分析(包括定量和/或鉴定)蛋白质级分。理解可用定量和/或定性方法确定蛋白质产物的量。确定产物的量包括确定产物存在或不存在。因此,表达概况分析可包括关于一种或多种目的蛋白质存在或不存在的信息。如文中所用的产物的“不存在”及“检测不到产物”包括不显著的或最低(de minimus)水平。在一些实施方案中,比较了两个或多个样品中蛋白质的量。一般地,样品具有重叠蛋白质分布图。用本发明的方法,可比较蛋白质的量以确定分布图本质的差别和存在的蛋白质的量。这些方法用于鉴定指示疾病状态(例如,疾病生物标记、PSA、BRCA1,等等)的蛋白质的性质或者量的改变,或治疗功效、或活性剂的毒性作用或病原体(例如,HIV、细菌病原体、朊病毒等等)的存在,等等。这些方法也用于发现与疾病状态相关的蛋白质,其用于药物发现目的、诊断目的等等。具体地,比较来自不同受试者或者已经受到不同条件或者处理的样品的蛋白质分布图是有用的。

[0130] 例如,在某些实施方案中,第一个样品为未处理的对照样品,第二个样品受活性剂或条件的处理。活性剂的实例包括但不限于:化学治疗剂、紫外线、医学装置(例如,支架除纤颤器)、外源基因及生长因子。本领域技术人员认识到有许多将外源基因导入到细胞中的方法(参见,例如,Ausubel 等人,eds., (1994), 同上)。在其它实施方案中,第一个样品为患病样品且第二个样品为非患病样品。此外,活性剂可为候选药物的形式。例如,第一个样品中的蛋白质用候选药物处理且可与为阴性或阳性对照的第二个样品进行比较。候选药物

对第一及第二个样品中存在的蛋白质（例如，蛋白质）的量的影响可表明候选药物的效力或毒性。本领域技术人员应当理解这些方法可适合分析任意活性剂对疾病状态或样品中存在的疾病标记物的量的影响。在一个实施方案中，所述方法用于鉴定与活性剂（如候选药物）处理相关的蛋白质。此类蛋白质可以，例如，与活性剂效力相关，且因此作为临床终点的代表。

[0131] 生物标记

[0132] 可用文中描述的概况分析及表征方法鉴定生物标记蛋白质（或蛋白质）。生物标记为目的蛋白质，其检测、监视、定量和 / 或表征是重要的。在一些实施方案中，生物标记与特定条件或处理，如疾病或病症、用药物治疗（包括药物治疗的效力和 / 或毒性）、用医学装置处理等等相关。在其它实施方案中，生物标记在目的组织或细胞（例如肿瘤、器官，等等）中表达。如文中所用的，生物标记蛋白质可为新鉴定的蛋白质或蛋白质变体（如突变蛋白质、剪接变体、具有改变的翻译后修饰的蛋白质等等）。在其它实施方案中，生物标记为组织特异标记物。

[0133] 生物标记在诊断（在一些实施方案中，包括疾病分期）、预测、评估和 / 或疗法选择、疾病发展监视、治疗效力监视和 / 或疾病治疗中可用作替代标记物。在一些实施方案中，生物标记通过本领域已知的任意方法和 / 或文中描述任意方法进行检测和 / 或定量，其中生物标记的表达（生物标记存在或不存在，或生物标记的差异表达）表明病症或疾病的存在。在一个实施方案中，生物标记的水平升高表明存在病症或疾病。在另一实施方案中，生物标记的水平降低表明存在病症或疾病。在一些实施方案中，生物标记的表达用于评估特定治疗性治疗方案在动物研究、临床试验中的效力或监测个体受试者的治疗。在一些实施方案中，生物标记作为所希望的临床终点的代表。在其它实施方案中，当生物标记的表达预示活性剂（如药物）治疗效力时，生物标记与活性剂的效力相关。在一个实施方案中，生物标记水平升高表明治疗的效力或进展。在另一实施方案中，生物标记水平降低表明治疗的效力或进展。

[0134] 生物标记可用作毒性的标记，所述毒性包括活性剂如药物、新的药物候选者、化妆品或其它化学品的毒性。在一些实施方案中，生物标记表达的检测也可用于监测对活性剂如毒素或病原体的环境暴露。在一个实施方案中，生物标记水平升高表明毒性或暴露于活性剂。在另一实施方案中，生物标记水平降低表明毒性或暴露于活性剂。

[0135] 生物标记可用于筛选具有特异结合亲和力的多种分子及化合物或分子或化合物库，包括例如 DNA 分子、RNA 分子、肽核酸、多肽、模拟物、小分子等等。在一个实施方案中，测定包括提供多种分子和 / 或化合物，在允许特异结合条件下将生物标记与多种分子和 / 或化合物组合，并检测特异结合以鉴定特异结合该生物标记的至少一种分子或化合物。

[0136] 类似地，一种或多种生物标记或其部分可在多种筛选测定的任一种中用于筛选多种分子和 / 或化合物或分子和 / 或化合物库以鉴定配体。筛选方法是本领域公知的。测定法可用于筛选，例如，与生物标记特异结合的适体、DNA 分子、RNA 分子、肽核酸、多肽、模拟物、蛋白质、抗体、激动剂、拮抗剂、免疫球蛋白、抑制剂、小分子、药学活性剂或药物化合物等等。

[0137] 在另一实施方案中，包含特异结合生物标记的抗原结合位点的一种或多种抗体可用于检测所述生物标记（包括体外及体内检测）。在另一实例中，特异结合生物标

记的抗体可连接体内成像试剂,例如³H、¹¹¹In、¹²⁵I(参见 Esteban 等人 (1987) J. Nucl. Med. 28. 861-870),并用于体内成像应用中。组合物及试剂盒

[0138] 本发明还提供了组合物,其用于在文中描述的任意方法中,所述方法为诸如降低样品复杂度的方法、纯化和 / 或富集蛋白质或多种蛋白质的方法、离析和 / 或分离蛋白质或多种蛋白质的方法,和 / 或制备用于表征的蛋白质、多种蛋白质或蛋白质级分的方法、制备用于质谱分析法分析的蛋白质、多种蛋白质或蛋白质级分的方法、鉴定蛋白质或多种蛋白质的方法、发现一种或多种新蛋白质的方法、检测和 / 或定量样品中蛋白质或多种蛋白质的方法、表征一种或多种蛋白质的方法、表达概况分析的方法、鉴定蛋白质降解产物的方法、鉴定翻译后修饰改变的方法和 / 或确定样品中蛋白质的质量、量和 / 或同一性的方法。本发明方法中使用的组合物可包含一种或多种(如约 2、约 3、约 4、约 5、约 7、约 10、约 15 或 15 种以上)小表位抗体。在一些实施方案中,组合物包含低于约 100、约 95、约 90、约 85、约 80、约 75、约 70、约 65、约 60、约 55、约 50、约 45、约 40、约 35、约 30、约 25、约 20、约 15、约 10、约 5 或 5 种以下小表位抗体。在一些实施方案中,组合物包含至少约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体。在一些实施方案中,组合物包含约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体。在一些实施方案中,组合物包含至少约 10、20、30、40、50、75、100、125、150、200、300、400 或 500 种小表位抗体,上限分别为约 20、30、40、50、75、100、125、150、200、300、400、500 或 1000 种小表位抗体。

[0139] 本发明还提供了在本方法中使用的试剂盒。本发明的试剂盒包括包含一种或多种小表位抗体的一个或多个容器。在一些实施方案中,试剂盒包含少于约 100、约 95、约 90、约 85、约 80、约 75、约 70、约 65、约 60、约 55、约 50、约 45、约 40、约 35、约 30、约 25、约 20、约 15、约 10、约 5 或 5 种以下的小表位抗体。在一些实施方案中,试剂盒包含至少约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体。在一些实施方案中,试剂盒包含约 20、约 30、约 40、约 50、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 1000 或 1000 种以上小表位抗体。在一些实施方案中,试剂盒包含至少约 10、20、30、40、50、75、100、125、150、200、300、400 或 500 种小表位抗体,上限分别为约 20、30、40、50、75、100、125、150、200、300、400、500 或 1000 种小表位抗体。在一些实施方案中,试剂盒还包含关于根据文中描述的本发明的任意方法使用的说明书,所述方法为诸如降低样品复杂度的方法、纯化和 / 或富集蛋白质或多种蛋白质的方法、离析和 / 或分离蛋白质或多种蛋白质的方法,和 / 或制备用于表征的蛋白质、多种蛋白质或蛋白质级分的方法、制备用于质谱分析法分析的蛋白质、多种蛋白质或蛋白质级分的方法、鉴定蛋白质或多种蛋白质的方法、发现一种或多种新蛋白质的方法、检测和 / 或定量样品中蛋白质或多种蛋白质的方法、表征一种或多种蛋白质的方法、表达概况分析的方法、鉴定蛋白质降解产物的方法、鉴定翻译后修饰改变的方法和 / 或确定样品中蛋白质的质量、量和 / 或同一性的方法。

[0140] 本发明还包含任意蛋白质“产物”(例如,用文中描述的任意方法富集、纯化、离析、制备、分离和 / 或分级分离的蛋白质)。本发明还提供了用文中描述的本发明的任意方法表征(例如,检测、鉴定、定量,等等)的蛋白质或蛋白质片段及包含此类产物的组合物。此类蛋白质包含小表位抗体识别的关连小表位(蛋白质结合所述小表位抗体)。本发明还提供

了通过文中描述的任意方法制备或分离的小表位抗体-蛋白质复合体或小表位抗体-蛋白质片段复合体（对于其中在与小表位抗体接触前蛋白质与蛋白质切割剂接触的方法）。本发明还提供了根据文中描述的任意方法自小表位抗体-蛋白质复合体或小表位抗体-蛋白质片段复合体分离的蛋白质或蛋白质片段，和/或从小表位抗体分离后的蛋白质制备的蛋白质片段。

[0141] 另一方面，本发明包括组合物和/或试剂盒，所述组合物和/或试剂盒包含通过本发明方法的任意方面产生的中间物（如复合体，例如小表位抗体-蛋白质复合体）。本发明还提供了孵育混合物，所述孵育混合物包含如文中描述的含有蛋白质的样品及小表位抗体和/或小表位抗体-蛋白质复合体。本发明的试剂盒在合适的包装内。合适的包装包括但不限于小瓶、瓶、罐、柔性包装（例如，密封 Mylar 或塑料袋），等等。在一些实施方案中，试剂盒包含容器和容器上或与容器结合的标签或包装说明书。标记或包装说明书可说明小表位抗体可用于文中描述的任意方法，所述方法为例如用于降低样品复杂度的方法，或用于鉴定蛋白质、表征蛋白质和/或表达概况分析的方法。可提供用于实践文中描述的任意方法的说明书。

[0142] 本发明的方法中有用的组分及反应混合物

[0143] 小表位抗体

[0144] 本发明的方法使用小表位抗体。如文中所用的，“小表位抗体”为结合（一般地特异结合）小肽表位的抗体。由于表位特异性，小表位抗体一般识别包含抗体结合的小表位的多种蛋白质。在通过抗体结合的小表位是已知的情况下，通过小表位抗体结合提供了涉及小表位抗体结合的蛋白质的氨基酸含量和/或序列的信息。小表位抗体例如在共同待决的美国专利申请号 10/687,174 中描述。小表位抗体及制备小表位抗体的方法在文中进一步讨论且在实施例中示例。

[0145] 抗体可包括单克隆抗体、多克隆抗体、抗体片段（例如，Fab、Fab'、F(ab')₂、Fv、Fc 等等）、嵌合抗体、单链（ScFv）、其突变体、包含抗体部分的融合蛋白及包含具有所需要的特异性的抗原识别位点的免疫球蛋白分子的任意其它修饰的构型。抗体可为鼠、大鼠、人或任意其它来源（包括人源化抗体）。可通过本领域已知的许多方法产生，包括例如，通过杂交瘤产生、重组产生或化学合成产生小表位抗体。

[0146] 一般地，小表位抗体结合 3、4 或 5 个顺序（连续）氨基酸的短的、线性肽表位。备选地，在一些实施方案中，小表位抗体结合多肽中的非连续氨基酸序列。在一些实施方案中，小表位抗体结合由或基本由约 2、3、4、5、6、7、8、9 或 10 个氨基酸组成。在一些实施方案中，小表位抗体结合由或基本由 2 至 10、3 至 8，或 3 至 5 个氨基酸组成的表位。在一些实施方案中，小表位抗体结合由或基本由少于约 10、9、8、7、6、5、4 或 3 个氨基酸组成的表位。在一些实施方案中，小表位抗体群结合由或基本由约 3 至约 5 个氨基酸组成的表位。在一些实施方案中，小表位抗体群结合由或基本由 2 至 10、3 至 8 或 3 至 5 个氨基酸组成的表位。在一些实施方案中，小表位抗体群结合由或基本由约 2、3、4、5、6、7、8、9 或 10 个氨基酸组成的表位。在一些实施方案中，小表位抗体群结合由或基本由少于约 10、9、8、7、6、5、4 或 3 个氨基酸组成的表位。小表位抗体群包含多种小表位抗体。在一个实施方案中，多种小表位抗体结合相同数目氨基酸的表位。在其它实施方案中，多种小表位抗体结合不同数目氨基酸混合物的表位。在文中描述的任意实施方案中，表位可为如以下描述的多肽中的连续或

不连续序列。在一些实施方案中,在抗体混合物中可包含一种或多种小表位抗体,所述抗体混合物包含的抗体结合的表位多于所述一种或多种小表位抗体识别的表位。

[0147] 在一些实施方案中,小表位抗体结合由或基本由 3 个连续氨基酸(称为 3mer)、4 个连续氨基酸(称为 4mer)或 5 个连续氨基酸(称为 5mer)组成的表位。在其它实施方案中,小表位抗体结合小的“不连续”或“简并”线性肽序列,如线性肽序列 YCxC,其中 x 代表 20 种天然氨基酸任一种(简并线性序列)。在其它实施方案中,小表位抗体结合多肽中非连续(不连续)序列,所述结合是基于该多肽中氨基酸的构象接近性而形成表位(例如,由于折叠多肽中的二级结构,通过氨基酸残基的接近形成的构象表位)。在其它实施方案中,小表位抗体可结合由氨基酸序列组成的表位,所述氨基酸序列用本领域公知的用于预测抗原性的方法预测为具有抗原性。如以下表 2 中显示的,以前描述了结合小线性肽表位的抗体。在一些实施方案中,相同抗体可结合一种或多种蛋白质上的连续序列及一种或多种蛋白质上的不连续序列。

[0148] 小表位抗体一般识别多种蛋白质,所述蛋白质包含所述抗体结合的小表位。在一些实施方案中,小表位抗体结合样品中约 0.1%、0.5%、1%、2%、3%、4%、5%、10% 或 10% 以上的蛋白质中一次或多次出现的表位。在其它实施方案中,小表位抗体结合样品中约 0.1% 至 1% 的蛋白质中一次或多次出现的表位。在其它实施方案中,小表位抗体结合在样品中约 1-5% 的蛋白质中一次或多次出现的表位。在其它实施方案中,小表位抗体结合样品中约 0.1% 至 1% 的蛋白质中一次或多次出现的表位,其中小抗体表位结合由 3 个氨基酸、4 个氨基酸或 5 个氨基酸组成或一般由 3 个氨基酸、4 个氨基酸或 5 个氨基酸组成的线性肽表位。在其它实施方案中,小表位抗体结合样品中约 1-5% 的蛋白质中一次或多次出现的表位,其中小抗体表位结合一般由或基本由 3 个氨基酸、4 个氨基酸或 5 个氨基酸组成的线性肽表位。在其它实施方案中,小表位抗体结合样品中约 5-7% 或约 5-10% 的蛋白质中一次或多次出现的表位,其中小表位抗体结合由或基本由 3 个氨基酸、4 个氨基酸或 5 个氨基酸组成的线性肽表位。在一些实施方案中,多种小表位抗体一起结合样品中至少约 0.1%、0.5%、1%、2%、3%、4%、5%、10% 或 10% 以上的蛋白质中的任意一种中一次或多次出现的一种或多种表位。在一些实施方案中,多种小表位抗体结合样品中约 0.1 至 1%、1 至 5%、5 至 7% 或 5 至 10% 的蛋白质中一次或多次出现的表位。用于经验评估样品中表位频率的方法包括:用生物化学方法评估,如结合抗体后用例如 2D 凝胶或质谱分析法分析,及用例如氨基酸或核酸序列数据库如 GenBank 及 SwissProt 的基于序列的分析。文中还描述了合适的数据库。

[0149] 在一些实施方案中,通过小表位抗体识别的表位还包含作为通过内肽酶识别为切割位点的 C 末端氨基酸。例如,表位可包含 C 末端精氨酸和 / 或赖氨酸,其每个都被胰蛋白酶识别为切割位点。内肽酶消化蛋白质混合物后,通过内肽酶识别的氨基酸一般存在于靶肽的 C 末端;因此,包含此氨基酸的肽也存在于靶多肽的 C 末端,其可增加免疫原性,和增加与抗体-靶肽结合相关的结合能。

[0150] 在一些实施方案中,小表位抗体结合它的关连表位,结合反应的亲和力为至少约 10^{-7} M、至少约 10^{-8} M 或至少约 10^{-9} M 或 10^{-9} M 以下。通过本领域公知的方法,包括例如通过表面等离子体共振 (Malmberg 和 Borrebaeck(1995) J Immunol. Methods 183(1):7-13; Lofas 和 Johnson(1990) J Chem. Soc. Chem. Commun. 1526) 测定结合亲和力。在一些实施方案中,

反应中结合相互作用与偶发结合相互作用相差至少 2 倍,至少 5 倍、至少 10 倍到至少 100 倍或 100 倍以上。

[0151] 本发明的方法中可使用一种小表位抗体、或一种以上(如约二、约三、约四、约五、约十、约二十、约一百或一百种以上)的小表位抗体群。因此,在一些实施方案中,所述方法包括使用至少一种小表位抗体。在其它实施方案中,所述方法包括使用至少两种小表位抗体。在其它实施方案中,本发明方法中使用至少约 5、约 10、约 20、约 30、约 40、约 50、约 60、约 75、约 100、约 125、约 150、约 200、约 300、约 400、约 500、约 750、约 1000 或 1000 种以上小表位抗体。在一些实施方案中,样品与少于约 100、约 95、约 90、约 85、约 80、约 75、约 70、约 65、约 60、约 55、约 50、约 45、约 40、约 35、约 30、约 25、约 20、约 15、约 10、约 5 或 5 种以下小表位抗体接触。在一些实施方案中,样品与至少约 20、约 30、约 40、约 50、约 75、约 100、约 500、约 1000 或 1000 种以上小表位抗体接触。在一些实施方案中,样品与至少约 5、10、20、30、40、50、60、75、100、125、150、200、300、400、500 或 750 种小表位抗体接触,上限为约 10、20、30、40、50、75、100、125、150、200、300、400、500、750 或 1000 种小表位抗体。理解可使用小表位抗体与其它蛋白质结合剂(如不是小表位抗体的抗体)混合物。

[0152] 理解小表位抗体结合的表位的同一性(序列)可与文中描述的任意方法组合使用以例如鉴定蛋白质。在一些实施方案中,小表位同一性是已知的。在其它实施方案中,表位的同一性可用本领域已知的方法预测。

[0153] 如文中讨论的,抗体可与样品一次一种接触或与两种或多种抗体构成的组接触。在一些实施方案中,接触为顺序的(相继的或反复的),例如一种抗体或一组抗体与样品接触并分离,第二种抗体或第二组抗体与样品接触并分离。在其它实施方案中,接触是平行的,例如,一组抗体与样品接触并分离。不同组抗体与样品顺序接触时,可理解接触可为平行的和顺序的。抗体组在组成中可以是重叠的(例如,第 1 组=抗体 A、B、C、D;第 2 组=抗体 B、C、D、E 等等)。

[0154] 显然在降低样品复杂度的方法中有用的的小表位抗体数目取决于用一种或多种小表位抗体制备的蛋白质预期的用途、应用和/或随后的分析。在一些应用中,如包含通过小抗体识别的关连表位的蛋白质的检测中,一种小表位抗体(或在一些实施方案中,少数几种小表位抗体)可用于制备、纯化和/或富集蛋白质级分,所述蛋白质级分包含随后检测(或其它分析)所希望的蛋白质。然后,可进一步分析分离的蛋白质。在其它实施方案中,用一组两种或多种小表位抗体是有用的。例如,在如蛋白质发现和,在一些实施方案中,表达概况分析的应用中,希望使用多种小表位抗体,从而富集和/或纯化多种蛋白质(如起始样品中基本上所有蛋白质)。在希望纯化和/或富集新蛋白质或蛋白质形式(例如,由于关于靶蛋白质序列的信息是未知的)的应用中也可以使用多种小表位抗体。如涉及与样品中多种蛋白质分级分离有关的实施方案的示例性实例中所示的,小表位抗体识别的关连氨基酸表位的序列和/或长度的知识允许估计蛋白质样品中小表位抗体识别的表位的期望频率。如表 1 中显示的,对 3mer、4mer 及 5mer 线性肽序列总共分别有 8,000 (20^3)、160,000 (20^4) 及 3,200,000 (20^5) 种随机组合。考虑到蛋白质的平均长度为 500 个氨基酸,它通过一种抗 -3mer 抗体检测到的概率为 0.0625,使用 15 种抗 -3mer 抗体时,概率增加到约 1,用 100 种抗 -3mer 抗体时,概率增加到 6.25。此类计算是常规的。小表位抗体也可识别简并线性表位,例如短肽,如 YCxC,其中 x 代表 20 种标准氨基酸的两种或多种。

[0155] 表 1. 短线性氨基酸肽的分布性质

[0156]

	表位氨基酸长度(n)			
	2	3	4	5
随机组合 # (20 ⁿ)	400	8, 00	160,000	3,200,000
500mer 蛋白质出现率 (500/20 ⁿ)	1.25	0.0625	0.003125	0.00015625
通过 100 种抗-nmer 抗体的 检测率(100 × 500/20 ⁿ)	125	6.25	0.3125	0.015625
通过 1000 种抗-nmer 抗 体的检测率 (1000 × 500/20 ⁿ)	1, 250	62.5	3.125	0.15625

[0157] 因此,理解本发明的方法中有用的表位抗体数目取决于多种因素,包括,例如通过小表位抗体结合的蛋白质级分预期的用途、应用和 / 或随后的分析、样品的复杂度 (根据预计或估计或以前确定的蛋白质,包括蛋白质变体如剪接变体的数目)、样品中蛋白质的平均大小、样品中出现或预计出现的关连表位的频率、小表位抗体的结合亲和力和 / 或特异性;靶蛋白的知识及小表位抗体的稳定性。此类因素是本领域公知的并在文中进一步讨论。

[0158] 如表 2 中显示的,以前已经描述了结合小线性肽表位的抗体。

[0159] 表 2. 公开的短抗体肽序列

[0160]

表位序列	源蛋白质	抗体	参考文献
NKS	脑膜炎奈瑟氏菌 (N.	U623,	Malorney, B., 等人

[0161]

	meningitides) 的 Opa	U506	(1998) J Bacteriol 180 (5) :1323-30.
NRQD	脑膜炎奈瑟氏菌的 Opa	0512	同上
TTFL	脑膜炎奈瑟氏菌的 Opa	AB419	同上
NIP	脑膜炎奈瑟氏菌的 Opa	W320/15, W124	同上

	meningitides) 的 Opa	U506	(1998) J Bacteriol 180(5) :1323-30.
GAT	脑膜炎奈瑟氏菌的 Opa	P515	同上
EQP	解脲脲原体 (U. urealyticum) 的 MB	3B1.5	Zhang, X., 等人, (1996) Clin Diagn Lab Immunol 3(6) : 774-8.
WQDE	猪 ZP3 β	mAb-30	Afzalpurkar, A. 等人 (1997) Am J Reprod Immunol 38(1) :26-32.
GPGR	HIV-1 的 Gp120	9x mAbs	Akerblom, L., 等人 (1990) Aids 4(10) : 953-60.
D(A/S)F [*]	果糖磷酸激酶 -1	α -F3	Hollborn, M., 等人 (1999) J Mol Recognit 12(1) :33-7.
(D/S)GY(A/G) ^{**}	响尾蛇毒素	A-56.36	Demangel, C., 等人 (2000) Eur J Biochem 267(8) :2345-53

[0162] ^{*}:DAF 及 DSF

[0163] ^{**}:指 DGYA、DGYG、SGYA 及 SGYG。

[0164] 制备小表位抗体的方法是本领域已知的。另一方面,如在实施例中的示例的,可通过用免疫原制备小表位抗体(例如人抗体、人源化抗体、小鼠抗体、嵌合抗体),所述免疫原表达一种或多种小肽表位,如由或基本由 3、4 或 5 个氨基酸组成的小线性肽表位。

[0165] 可例如通过化学合成产生免疫原。用于合成多肽的方法是本领域公知的。在一些实施方案中,如本领域已知的,合成末端为半胱氨酸的多肽免疫原,以促进与 KLH 或 BSA 的偶联。可在多肽的氨基末端掺入(在免疫及筛选期间降低位阻效应)或羧基末端掺入末端半胱氨酸。在其它实施方案中,多肽免疫原作为多种抗原多肽或 MAP 合成。

[0166] 如文中进一步描述的,免疫宿主动物的途径及方案一般按照用于抗体刺激及产生

的成熟和常规的技术。用于产生人及小鼠抗体的一般技术是本领域已知的并在文中描述。通常,宿主动物用一定量的免疫原,包括如文中描述的免疫原腹膜内接种。

[0167] 用 Kohler, B. 和 Milstein, C. (1975) *Nature* 256 :495-497 或如通过 Buck, D. W. 等人 (1982) *In Vitro*, 18 :377-381 改进的一般体细胞杂交技术自淋巴细胞及永生化骨髓瘤细胞制备杂交瘤。在杂交中可使用可利用的骨髓瘤细胞系,包括但不限于 X63-Ag8.653 及来自 Salk Institute, Cell Distribution Center, San Diego, Calif., USA 的骨髓瘤细胞系。一般地,所述技术包括用融合剂如聚乙二醇,或通过本领域技术人员公知的电学方法融合骨髓瘤细胞及淋巴细胞。融合后,将细胞自融合培养基分离并在选择性生长培养基如次黄嘌呤-氨基蝶呤-胸苷 (HAT) 培养基上生长以除去未杂交的亲代细胞。添加或没有添加血清的文中描述的任意培养基可用于培养分泌单克隆抗体的杂交瘤。作为细胞融合技术的另一备选方案,EBV 永生化的 B 细胞可用于产生主题发明的小表位抗体。将杂交瘤扩大及亚克隆,如果希望,通过常规免疫测定方法(例如,放射免疫测定法、酶免疫测定法或荧光免疫测定法)分析上清液的抗免疫原活性。

[0168] 产生小表位抗体(如单克隆抗体)的杂交瘤或亲代杂交瘤的子代细胞可用作抗体或其衍生物或其部分的来源。

[0169] 产生此类抗体的杂交瘤可用已知方法在体外或体内生长。单克隆抗体可根据需要用常规免疫球蛋白纯化方法如硫酸铵沉淀、凝胶电泳、透析、层析及超滤从培养基或体液分离。如果存在不希望的活性,那么例如通过将制备物在由连接到固相制备的免疫原制备的吸附剂上流动并从免疫原洗脱或释放所希望的抗体除去所述活性。用人或其它物种的小表位受体,或人或其它物种小表位受体的片段或用双官能剂或衍生剂缀合到在待免疫的物种中为免疫原性的蛋白质的人或其它物种的小表位受体或含有靶氨基酸序列的片段免疫宿主动物可产生抗体群(例如单克隆抗体),所述在待免疫的物种中为免疫原性的蛋白质为例如钥孔血蓝蛋白、血清白蛋白、牛甲状腺球蛋白或大豆胰蛋白酶抑制剂,所述双官能剂或衍生剂为例如马来酰亚胺苯甲酰基磺基琥珀酰亚胺酯(通过半胱氨酸残基缀合)、N-羧基琥珀酰亚胺(通过赖氨酸残基缀合)、戊二醛、琥珀酐、 SOCl_2 或 $\text{R}_1\text{N}=\text{C}=\text{NR}$, 其中 R 及 R_1 为不同烷基。

[0170] 如果希望,可对目的小表位抗体(单克隆或多克隆的)测序,然后将多核苷酸序列克隆到载体用于表达或增殖。编码目的抗体的序列可在宿主细胞中的载体中维持并然后将宿主细胞扩大及冷冻用于以后使用。在备选方案中,多核苷酸序列可用于遗传操作以“人源化”抗体或改善抗体的亲和力或其它特征。例如,可对恒定区进行改造以更像人恒定区以避免抗体用于临床试验及人体治疗时产生免疫反应。希望遗传操作抗体序列以获得对小表位的更大亲和力和/或对小表位的更大和/或改变的特异性。对本领域技术人员显而易见的是,可以对小表位抗体进行一个或多个多核苷酸改变并仍维持它对小表位的结合能力。

[0171] 已经描述了包含来自非人免疫球蛋白的抗原结合位点的许多“人源化”抗体分子,包括嵌合抗体,其具有与人恒定结构域融合的啮齿动物或经修饰的啮齿动物 V 区及它们的相关互补性决定区(CDR)。参见,例如 Winter 等人 *Nature* 349 :293-299(1991), Lobuglio 等人 (1989) *Proc. Nat. Acad. Sci. USA* 86 :4220-4224, Shaw 等人 (1987) *J Immunol.* 138 :4534-4538 及 Brown 等人 (1987) *Cancer Res.* 47 :3577-3583。其它参考文献描述了在与适当人抗体恒定区融合前移植到人支持构架区(FR)的啮齿动物 CDR。参见,例如, Riechmann

等人 (1988) *Nature* 332 :323-327, Verhoeyen 等人 *Science* (1988) 239 :1534-1536 以及 Jones 等人 *Nature* (1986) 321 :522-525。另一参考文献描述了通过重组镶嵌的啮齿动物构架区支持的啮齿动物 CDR。参见,例如,欧洲专利公开号 519,596。设计这些“人源化”分子以将啮齿动物抗人抗体分子的不需要的免疫反应降到最小,所述不需要的免疫反应限制了这些部分在人受者中治疗性应用的持续时间及有效性。例如,可对抗体恒定区进行改造使它称为免疫惰性的(例如,不触发补体裂解)。参见,例如 PCT/GB99/01441;英国专利申请号 9809951.8。人源化单克隆抗体有四个一般步骤。它们是 (1) 确定起始抗体轻链及重链可变结构域的核苷酸序列及预测的氨基酸序列 (2) 设计人源化抗体,即决定人源化过程期间使用哪种抗体构架区 (3) 实际的人源化方法/技术及 (4) 人源化抗体的转染及表达。参见,例如,美国专利号 4,816,567;5,807,715;5,866,692;6,331,415;5,530,101;5,693,761;5,693,762;5,585,089 及 6,180,370。也可使用的其它人源化抗体的方法如通过 Daugherty 等人, *Nucl. Acids Res.* (1991) 19 :2471-2476 及在美国专利号 6,180,377;6,054,297;5,997,867;5,866,692;6,210,671;6,350,861 及 PCT 公开号 WO 01/27160 中公开。

[0172] 另一备选方案中,通过用经过改造以表达特异人免疫球蛋白蛋白质的通过商业途径可获得的小鼠可获得完全人抗体。设计产生更希望的(例如完全人抗体)或更强免疫反应的转基因动物也可用于产生人源化或人抗体。此类技术的实例为来自 Abgenix, Inc. (Fremont, CA) 的 Xenomouse™ 和来自 Medarex, Inc. (Princeton, NJ) 的 HuMAb-Mouse® 及 TC Mouse™。

[0173] 在备选方案中,可用本领域已知的任意方法重组制备及表达抗体。在另一备选方案中,可通过噬菌体展示技术重组制备抗体。参见,例如,美国专利号 5,565,332;5,580,717;5,733,743 及 6,265,150;以及 Winter 等人 (1994) *Annu. Rev. Immunol.* 12 :433-455。

[0174] 备选地,可用噬菌体展示技术 (McCafferty 等人 (1990) *Nature* 348 :552-553) 从来自非免疫供者免疫球蛋白可变 (V) 结构域基因所有组成成分产生人抗体。例如可针对多种合成多肽平行淘洗现有的抗体噬菌体展示库。根据此项技术,将抗体 V 结构域基因框内克隆到丝状噬菌体,如 M13 或 fd 的主要或次要外壳蛋白基因中并在噬菌体颗粒表面上作为功能性抗体片段展示。由于丝状颗粒含有噬菌体基因组的单链 DNA 拷贝,所以基于抗体功能性质的筛选也导致选择编码显示此类特性的抗体的基因。因此,噬菌体模拟了 B 细胞的一些性质。可以以多种形式进行噬菌体展示;对于综述,参见,例如 Johnson, Kevin S. 和 Chiswell, David J., *Current Opinion in Structural Biology* (1993) 3,564-571。一些 V 基因片段来源可用于噬菌体展示。Clackson 等人 *Nature* (1991) 352 :624-628 从来自经免疫小鼠脾的 V 基因的小随机组合文库中分离了一组不同的抗噁唑酮抗体。基本根据 Mark 等人 (1991) *J. Mol. Biol.* 222 :581-597 或 Griffith 等人 (1993) *EMBO J.* 12 :725-734 描述的技术,可构建来自未免疫的人供者的 V 基因库和分离针对多种抗原(包括自身抗原)的多种抗体。在天然免疫反应中,抗体基因以高速率积累突变(体细胞高变)。一些导入的改变将赋予高亲和力,并且展示高亲和力表面免疫球蛋白的 B 细胞在随后抗原攻击期间优先复制及分化。此自然过程可通过利用称为“链改组”(Marks, 等人 (1992) *Bio/Technol.* 10 :779-783) 的技术模拟。在此方法中,通过噬菌体展示获得的“初次”人抗体的亲和力可通过

用获得自未免疫供者的 V 结构域基因的天然存在的变体库（所有组成成分）顺序置换重链及轻链 V 区基因得到提高。此技术允许产生具有 pM-nM 范围内亲和力的抗体及抗体片段。用于制备非常大噬菌体抗体库（也称为“所有文库之母”）的方法已通过 Waterhouse 等人 (1993) Nucl. Acids Res. 21 :2265-2266 描述。基因改组也用于自噬齿动物抗体衍生人抗体，而人抗体与起始噬齿动物抗体具有相似亲和性及特异性。根据也称为“表位印记”的此方法，通过噬菌体展示技术获得的噬齿动物抗体的重链或轻链 V 结构域基因用人 V 结构域基因库置换，产生噬齿动物-人嵌合体。抗原的选择导致能恢复功能性抗原结合位点的人可变区的分离，即表位支配（印记）配偶体的选择。为了置换剩余噬齿动物 V 结构域而重复该过程时，获得了人抗体（参见 1993 年 4 月 1 日公开的 PCT 公开号 WO 9306213）。与通过 CDR 移植对噬齿动物抗体的常规人源化不同，此技术提供了完全的人抗体，其没有噬齿动物来源的构架或 CDR 残基。很显然，虽然以上讨论是关于人抗体的，但是所讨论的一般原理适用于定制抗体用于例如狗、猫、灵长动物、马及牛中。

[0175] 通过首先从宿主动物分离抗体，获得基因序列，并用基因序列在宿主细胞（例如，CHO 细胞）中重组表达抗体可重组制备抗体。可利用的另一方法为在植物（例如，烟草）、转基因乳或在其它生物中表达抗体序列。用于在植物或乳中重组表达抗体的方法已经公开。参见，例如，Peeters 等人 (2001) Vaccine 19 :2756 ;Lonberg, N. 和 D. Huszar (1995) Int. Rev. Immunol 13 :65 ;以及 Pollock 等人 (1999) J Immunol Methods 231 :147。用于制备抗体衍生物，例如人源化、单链等等抗体的方法是本领域已知的。

[0176] 也可利用免疫测定法及流式细胞术分选技术如荧光激活细胞分类术 (FACS) 分离对所希望的小表位特异的抗体。

[0177] 抗体可与许多不同载体结合。载体可为活性和 / 或惰性的。公知载体的实例包括聚丙烯、聚苯乙烯、聚乙烯、葡聚糖、尼龙、淀粉酶、玻璃、天然的及改性纤维素、聚丙烯酰胺、琼脂糖及磁铁矿。对于本发明，载体的性质可为可溶的或不溶的。本领域技术人员知道用于结合抗体的其它合适载体或用常规实验能够确定此类载体。

[0178] 如本领域已知的，可对编码小表位抗体的 DNA 进行分离及测序。一般地，用常规方法（例如，通过使用能够特异结合编码单克隆抗体的重链及轻链的基因的寡核苷酸探针）易于对单克隆抗体进行分离及测序。杂交瘤细胞作为此类 cDNA 的优选来源。一旦分离，可将 DNA 置于表达载体中，然后将所述载体转染到宿主细胞如大肠杆菌 (E. coli) 细胞、猴 COS 细胞、中国仓鼠卵巢 (CHO) 细胞或不产生免疫球蛋白蛋白质的骨髓瘤细胞中，以在重组宿主细胞中获得单克隆抗体的合成。也可对 DNA 进行修饰，例如，通过用人重链及轻链恒定结构域的编码序列替代同源鼠序列，Morrison 等人 (1984) Proc. Nat. Acad. Sci. 81 :6851 或通过共价连接免疫球蛋白编码序列与非免疫球蛋白多肽的所有或部分编码序列。以那种方式，制备了具有文中的小表位抗体（如单克隆抗体）的结合特异性的“嵌合”或“杂合”抗体。

[0179] 可用本领域公知的方法表征小表位抗体，一些所述方法在实施例中描述。例如，在 Harlow 和 Lane, Using Antibodies, a Laboratory Manual, ColdSpring Harbor Laboratory Press, Cold Spring Harbor, New York, 1999 第 11 章描述的一种方法是鉴定它结合的表位，包括解决抗体-抗原复合体的晶体结构、竞争测定法、基因片段表达测定法及基于合成多肽的测定法。在额外实例中，表位作图可用于确定小表位抗体结合的序列。表位作图可通

过商业途径从多种来源获得,例如 Pepscan Systems (Edelhertweg 15, 8219PH Lelystad, 荷兰)。可分离或合成(例如,重组)不同长度(例如,至少 4-6 个氨基酸长)的多肽并用于使用抗小表位抗体的结合测定中。在另一实例中,通过使用来自小表位细胞外序列的重叠多肽和测定小表位抗体的结合可以在系统筛选中确定确定小表位抗体结合的表位。如本领域公知的,可通过用在噬菌体颗粒表面展示的随机多肽序列的大文库(噬菌体文库)也可鉴定某些表位。

[0180] 可用于表征抗小表位抗体的另一方法是使用竞争测定法,其使用已知结合相同抗原的其它抗体,即确定抗小表位抗体是否结合与其它抗体结合的表位相同的表位。竞争测定法是本领域技术人员公知的。

[0181] 本发明中有用的的小表位抗体可连接标记活性剂(备选称为“标记”),如荧光分子(如半抗原或荧光珠)、结合配偶体、固相支持体或本领域已知的有助于分离的其它活性剂。此类活性剂在文中进一步描述。

[0182] 在一些实施方案中,使用一种或多种以下考虑设计小表位抗体(设计为单独使用或在群体中使用),导致具有足够冗余性的表位频率以产生样品中存在的蛋白质的最佳覆盖度。在一个实施方案中,根据一种或多种以下考虑设计的一组小表位抗体能结合样品中至少约 10、20、30、40、50、60、70、80、90 或 95% 的蛋白质上的关连表位。

[0183] • 表位大小:小表位抗体,其足够小以在蛋白质组中经常发生但足够大以在通过它们的关连表位识别时给予充足结合能。在一些实施方案中,通过每种抗体识别的表位大小为 3、4 或 5 个氨基酸。

[0184] • 表位多度:在一个实施方案中,最佳表位多度能使每个小表位抗体结合长度约 20 至约 100 个氨基酸的约 100 至约 150 种来自血清的多肽。此多度水平与多数质谱仪的分辨力匹配,不需要 MS-MS 及碰撞诱导解离(CID)。适当多度的表位优选用于实现所希望的 MS 分辨率及灵敏性。

[0185] • 样品丰余性:在一些实施方案中,使用足够多种的小表位抗体组以允许结合每个目的蛋白质组的每个蛋白质约 3 至约 5 个表位。此设计特征提供了取样丰余性以适应组中不同蛋白质的表达水平和每种抗体的结合效率所预计的可变性。

[0186] • 亲和力:在一些实施方案中,小表位抗体及它们的表位之间的结合紧密性影响蛋白质分布图的灵敏性。在一些实施方案中,组中每种抗体以足够高的亲和力结合以确保捕获了用于 MS 分析的足够分析物。

[0187] • 结合频率:在一些实施方案中,小表位抗体的结合频率是高的从而每种结合的肽级分中存在的肽含有共同表位。这提供了取样丰余性并有助于肽同一性的生物信息学确定。

[0188] 样品与小表位抗体接触和从蛋白质-抗体复合体分离蛋白质

[0189] 用于抗体与样品中蛋白质接触的方法及条件是本领域公知的。抗体可与样品一次一种或两种或多种抗体构成的组接触。在一些实施方案中,接触为顺序的(相继的或反复的),例如一种抗体或一组抗体与样品接触并分离,第二种抗体或第二组抗体与样品接触并分离等等。在其它实施方案中,接触是平行的,例如,一组抗体与样品接触并分离。不同组抗体与样品顺序接触时,可理解接触可为平行的和顺序的。抗体组在组成中可以是重叠的(例如,第 1 组=抗体 A、B、C、D;第 2 组=抗体 B、C、D、E 等等)或者在组成中不同。抗体与

蛋白质的接触可在抗体及蛋白质均在液体基质的情况下发生或一种组分（抗体或蛋白质）与固相支持体结合另一种组分在液体基质中的情况下发生。在一个实施方案中，含有液体（例如，含水）蛋白质的样品与结合固相支持体的小表位抗体接触。

[0190] 在涉及平行接触的一些实施方案中，希望可单独分离小表位抗体，例如，通过将抗体与可检测的不同小珠连接、使用单独可分离的结合配偶体、将抗体固定在例如在多孔板的不同孔中、使用抗体阵列等等来分离。在通过抗体结合的小表位是已知的情况下，通过小表位抗体结合提供了关于通过小表位抗体结合的蛋白质的氨基酸含量和 / 或序列的信息。在其中关连小表位的知识是所希望的实施方案中，方便的是单独分离小表位（从而通过每个小表位抗体结合的蛋白质保持分离）。然而，个别的分离或可分离性不是每个实施方案中都需要的。例如，小表位抗体可与有重叠抗体组成的两种或多种抗体的小库组合，所述抗体为诸如 (1) 抗体 ABC；(2) 抗体 CDE；(3) 抗体 FGH，和 (4) 抗体 HIJ。抗体 - 蛋白质复合体分离及自抗体 - 蛋白质复合体分离抗体后，可基于特定组中的成员推断关于特定小表位的存在或不存在的信息。

[0191] 为有助于抗体 - 蛋白质复合体与样品中未结合蛋白质的分离，抗体可与有助于分离的媒介物如结合配偶体（例如，生物素、寡核苷酸、适体）、固相支持体（如小珠或基质，包括微阵列或多孔板）；或本领域已知的任意其它媒介物连接。结合可为共价的或非共价的，并且可为直接的或间接的。用于连接抗体到此类媒介物的方法是本领域公知的。参见，例如 Kennedy 等人 (1976) Clin. Chim. Acta 70 :1-31，及 Schurs 等人 (1977) Clin. Chim Acta 81 :1-40（描述了偶联技术，包括戊二醛方法、高碘酸盐方法、二马来酰亚胺方法、间马来酰亚胺苯甲基 -N- 羟基 - 琥珀酰亚胺酯方法，所有所述方法在此引用作为参考）。

[0192] 用于自样品分离抗体 - 蛋白质复合体的方法是本领域已知的并包括用结合结合配偶体的捕获剂（例如，捕获连接生物素的抗体的亲和素；捕获连接抗体的寡核苷酸的寡核苷酸）；也可用物理分离，如沉淀、过滤、FACS（例如，用光谱标记物标记的小珠）及磁性分离（抗体与具有磁性的基质，如磁珠连接时）。

[0193] 可在本发明中使用的许多结合配偶体是本领域已知的（例如，二硝基苯基、洋地黄毒苷、荧光团、俄勒冈绿染料 (Oregon Green dyes)、Alexa Fluor 488 (Molecular Probes)、荧光素、丹酰基、Marina Blue (Molecular Probes)、四甲基罗丹明、德克萨斯红 (Molecular Probes)、BODIPY (4,4- 二氟 -4- 硼 -3a,4a- 二氮杂 -s-indacene；美国专利号 4,774,339) 染料等等)。可用作特异结合结合剂的捕获试剂的抗体可自销售商如 Molecular Probes, Eugene, Oreg 通过商业途径获得。这些抗体包括可特异结合二硝基苯基、洋地黄毒苷、荧光团、俄勒冈绿染料、Alexa Fluor 488 (Molecular Probes)、荧光素、丹酰基、Marina Blue (Molecular Probes)、四甲基罗丹明、德克萨斯红 (Molecular Probes)、及 BODIPY 染料 (Molecular Probes) 的抗体。也可用任意合适的配体及抗配体。

[0194] 寡核苷酸可用作结合配偶体及捕获试剂。寡核苷酸包括核酸如 DNA、RNA 及混合的 RNA/DNA 分子。用作亲和力标记的寡核苷酸应能与捕获试剂上存在的寡核苷酸序列杂交。本领域技术人员认识到可设计互相杂交的许多不同的寡核苷酸序列。对于设计此类寡核苷酸对的重要考虑包括实际的核苷酸序列、寡核苷酸长度、杂交条件（例如，温度、盐浓度、有机化学品的存在等等）及寡核苷酸的解链温度。

[0195] 合适用于固定（连接）来自样品的抗体或蛋白质的固相支持体（及使固相支

持体适于固定抗体的修饰)是本领域公知的。固相支持体的实例包括:小珠(包括磁性小珠)、微孔板及蛋白质微阵列(例如,Zyomyx, Inc. 拥有的技术,参见,例如美国专利号 6,365,418)。因此,例如,二氧化硅外壳中包裹的 CdSe-CdS 核心壳纳米晶体易于衍生用于偶联生物分子。Bruchez 等人(1998)Science 281:2013-2016。类似地,将高度荧光量子点(硫化锌封端的硒化镉)与生物分子共价偶联用于超灵敏生物学检测。Warren 和 Nie(1998)Science 281:2016-2018。荧光标记小珠可以从 Luminex and Quantum Dot 通过商业途径获得。

[0196] 用常规免疫亲和洗脱条件如酸性 pH、离子强度、洗涤剂或以上组合可自抗体-蛋白质复合物释放结合的蛋白质(或在一些实施方案中,多肽片段)。一般地,对肽或蛋白质脱盐用于随后的分级分离、表征或其它分析。

[0197] a. 蛋白质切割剂

[0198] 在一些实施方案中,本发明的方法还包括用蛋白质切割剂处理样品,从而产生多肽片段。在一个实施方案中,样品与至少一种小表位抗体接触前,将样品与蛋白质切割剂接触。在另一实施方案中,蛋白质从抗体-蛋白质复合物分离后,将所述蛋白质与蛋白质切割剂接触。

[0199] 蛋白质切割剂处理产生蛋白质切割片段(如多肽),其可以有助于随后的样品中蛋白质的量及蛋白质同一性的质谱分析。特别地,用蛋白质切割剂处理的处理可有助于分析分子量超过 25kDa 的蛋白质。蛋白质切割剂处理也可有助于小表位抗体与关联表位的接近性和/或接近。蛋白质切割剂是本领域公知的且在文中进一步讨论。在一些实施方案中,使用一种蛋白质切割剂。在其它实施方案中,使用一种以上的蛋白质切割剂。在一些实施方案中,对于一种样品使用一种以上类型的蛋白质切割剂(例如两种或多种类型的蛋白酶、两种或多种化学切割剂或一种或多种蛋白酶和一种或多种化学切割剂的组合)。用蛋白质切割剂处理的条件是本领域公知的。在一个实施方案中,蛋白质切割剂为蛋白酶。可用作蛋白质切割剂的蛋白酶的实例包括但不限于:胰凝乳蛋白酶、胰蛋白酶(精氨酸、赖氨酸切割序列)、嗜热菌蛋白酶(苯丙氨酸、亮氨酸、异亮氨酸、缬氨酸切割序列)、V8 蛋白酶、内蛋白酶 Glu-C、内蛋白酶 Asp-N、内蛋白酶 Lys-C、内蛋白酶 Arg-C、内蛋白酶 Arg-N、因子 Xa 蛋白酶、凝血酶、肠激酶、V5 蛋白酶及烟草蚀纹病毒蛋白酶。可对在本发明方法中有用的蛋白酶进行遗传改造和/或化学修饰以防止自溶。理解可对酶蛋白质切割剂(如蛋白酶)进行修饰以有助于多肽切割后从多肽切割产物中除去蛋白酶。此类修饰是本领域已知的并包括:(1) 结合小珠(例如,胶乳、二氧化硅或磁珠)的蛋白酶、(2) 半抗原化蛋白酶、(3) 蛋白酶的亲和力耗竭(例如,结合小珠的抗蛋白酶或结合小珠的不可切割的底物)和/或(4) 大小排阻层析。例如可通过用热、蛋白酶抑制剂、金属螯合剂(例如,EGTA、EDTA)等等处理抑制蛋白酶的活性。

[0200] 在另一实施方案中,蛋白质切割剂为化学切割剂,如切割多肽及肽键的化学底物及化合物。化学切割剂的非限制性实例包括溴化氰(在甲硫氨酸残基切割)、羟胺(在天冬酰胺及甘氨酸残基之间切割)及酸性 pH(可在精氨酸-脯氨酸键切割)(参见,例如,Ausubel 等人,同上)。

[0201] 在其它实施方案中,磷酸酶(例如,碱性磷酸酶、酸性磷酸酶、蛋白丝氨酸磷酸酶、蛋白酪氨酸磷酸酶、蛋白苏氨酸磷酸酶、等等)、脂酶及其它酶可用作蛋白质切割剂。

[0202] 样品

[0203] 如在定义中提及及在文中使用的,“样品”包含多种样品类型和 / 或来源,如血液及其它生物来源的液体样品、固体组织样品如活检样品或来自其的组织培养物或细胞及其子代。所述定义还包括在获得它们后以任意方式操作,如对某些组分如蛋白质或多核苷酸用试剂处理、增溶或富集的样品。术语“样品”包括临床样品,也包括培养的细胞、细胞上清液、细胞裂解物、血清、血浆、生物液、和组织样品。样品可来自微生物,例如,细菌、酵母、病毒、类病毒、霉菌、真菌、植物、动物,包括哺乳动物如人。样品可包含单个细胞或一个以上的细胞。样品的实例包括血液、血浆、血清、尿、粪、脑脊液、滑液、羊水、唾液、肺灌洗液、精液、乳、乳头抽吸物、前列腺液、粘液、颊拭子和 / 或泪液。

[0204] 可通过本领域已知的方法如裂解、分级分离、纯化,包括亲和纯化、FACS、激光捕获显微解剖 (LCM) 或等密度离心制备这些样品。在一些实施方案中,亚细胞分级分离方法用于产生富集的细胞或亚细胞级分,如亚细胞细胞器,包括核、线粒体、重膜及轻膜和细胞质。

[0205] 在样品与一种或多种小表位抗体接触前,可用能变性和 / 或溶解蛋白质的试剂,如洗涤剂 (离子和非离子的)、离液剂和 / 或还原剂处理样品。此类试剂是本领域已知的。

[0206] 在某些情况下,例如,通过定向免疫耗竭或本领域已知的其它方法将样品中存在的蛋白质除去或降到最低是所希望的。一般地,在样品与一种或多种小表位抗体接触前发生此类除去 (或降低) (然而,在用小表位抗体处理期间或处理之后可发生此类降低或除去)。可使用任意合适试剂,包括一种或多种小表位抗体。在一个实施方案中,通过用一种或多种小表位抗体处理样品实现除去和 / 或减少一种或多种样品组分。

[0207] 在一些实施方案中,用多糖切割剂处理样品例如,以降低、最小化和 / 或去除样品蛋白质的糖基化是所希望的。可用化学或酶的方法完成任意糖部分的除去。多糖切割剂的实例包括糖苷酶、内切糖苷酶、外切糖苷酶及化学品如三氟甲磺酸。内切糖苷酶如内切糖苷酶 H (New England Biolabs, Beverly, Mass.) 及 Endo Hf (New England Biolabs) 可通过商业途径获得。这些内切糖苷酶 N- 连接的糖蛋白切割高甘露糖及一些杂合寡糖的壳二糖核心。外切糖苷酶也可从销售商如 New England Biolabs 通过商业途径获得并且包括, β -N-乙酰基氨基己糖苷酶、 α -1-2-岩藻糖苷酶、 α -1-3,4 岩藻糖苷酶、 α -1-2,3 甘露糖苷酶、 α -1-6 甘露糖苷酶、神经氨酸酶、 α -2-3 神经氨酸酶、 β 1-3 半乳糖苷酶及 α -N-乙酰基-氨基半乳糖苷酶。

[0208] 提供了以下实施例用以阐明,但不限制本发明。

[0209] 实施例

[0210] 实施例 1: 小表位抗体的制备及表征

[0211] 如表 3 中显示的,设计了多抗原肽 (MAP) 形式的五种免疫多肽。由于不同 MAP 中不同位置包括相同序列,所以这些序列组合也用于评估所诱导的抗体的交叉反应性。使用标准方法,每种免疫多肽用于免疫 4 只 Balb/C 小鼠。

[0212] 表 3. 免疫多肽的设计

[0213]

肽	组	序列	SEQ ID NO
MAP1	1	乙酰化 -HSLFH-PEDTGQV-KKTTNV-MAP	
MAP2	2	乙酰化 -PEDTGQV-KKTTNV-HSLFH-MAP	
MAP3	3	乙酰化 -LTPKKTTNVLTVPTNIPG-MAP	
MAP4	4	乙酰化 -LTPKKLTQENQNRGTHIYNQ-MAP	
MAP5	5	乙酰化 -TIYNLTQENQNRGTHIYNQ-MAP	

[0214] 表 3 注释：

[0215] 多肽 MAP1 :HSLFHPEDTGQV :来自 PSA, 氨基酸 #79-89。KKTTNV :来自脑膜炎球菌 (Meningococcal) Opa 蛋白, 含有 KTT- 公开的 3mer 抗体表位 (Malorny, Morelli 等人 1998)。

[0216] 多肽 MAP2 :MAP1 的交替序列。

[0217] 多肽 MAP3 :LTPKK :PSA 的基序 1 (Nagasaki, Watanabe 等人 1999)。KKTTNVLTVPTNIPG :来自脑膜炎球菌 Opa 蛋白, 含有两个公开的 3mer 抗体表位 :KTT 及 NIP 和一个 4mer 表位 :TNIP (Morelli, 等人 (1997) MolMicrobiol 25 (6) :1047-64。

[0218] 多肽 MAP4 :LTPKK :来自 PSA, 与肽 MAP3 中相同。LTQENQNRGTH :通过 DNASTar 计算机程序选择的 α -1-ACT 的免疫原性序列。IYNQ :来自脑膜炎球菌 Opa 蛋白, 含有一个 2mer 表位 IY 及 5mer 表位 TIYNQ 及 7mer 表位 TPTIYNQ 的四个氨基酸 (Morelli, 等人, 同上)。

[0219] 多肽 MAP5 :TIYNLTQENQNRGTH :来自脑膜炎球菌 Opa 蛋白 (Morelli, 等人, 同上)。LTQENQNRGTH :与肽 MAP4 中相同。

[0220] 设计了两组筛选多肽 : (1) 与免疫多肽相同序列的 5 种 C-末端生物素化序列 (表 4 中显示) ; 和 (2) 43 种 10mer 生物素化多肽, 其序列淘汰所有五种免疫多肽 (表 5 中显示)。

[0221] 表 4. 生物素化的筛选多肽 (大约 90% 纯度)

[0222]

肽	Mers	序列
Pep1-0	18	乙酰化 -HSLFHPEDTGQVKKTTNV- 生物素
Pep2-0	18	乙酰化 -PEDTGQVKKTTNVHSLFH- 生物素
Pep3-0	18	乙酰化 -LTPKKTTNVLTVPTNIPG- 生物素
Pep4-0	20	乙酰化 -LTPKKLTQENQNRGTHIYNQ- 生物素

肽	Mers	序列
Pep5-0	20	乙酰化 -TIYNTNIPGLTQENQNRGTH- 生物素

[0223] 表 5. 43 种 10mer 生物素化作图多肽 (大约 70%纯度)

[0224]

序号	肽名称	序列	免疫肽中位置
1	Pep1-1	乙酰化 -HSLFHPEDTG- 生物素	MAP1 1-10
2	Pep1-2	乙酰化 -SLFHPEDTGQ- 生物素	MAP1 2-11
3	Pep1-3	乙酰化 -LFHPEDTGQV- 生物素	MAP1 3-12
4	Pep1-4	乙酰化 -FHPEDTGQVK- 生物素	MAP1 4-13
5	Pep1-5	乙酰化 -HPEDTGQVKK- 生物素	MAP1 5-14
6	Pep2-1	乙酰化 -PEDTGQVKKT- 生物素	MAP1 6-15, MAP2 1-10
7	Pep2-2	乙酰化 -EDTGQVKKTT- 生物素	MAP1 7-16, MAP2 2-11
8	Pep2-3	乙酰化 -DTGQVKKTTN- 生物素	MAP1 8-17, MAP2 3-12
9	Pep2-4	乙酰化 -TGQVKKTTNV- 生物素	MAP1 9-18, MAP2 4-13
10	Pep2-5	乙酰化 -GQVKKTTNVH- 生物素	MAP2 5-14
11	Pep2-6	乙酰化 -QVKKTTNVHS- 生物素	MAP2 6-15
12	Pep2-7	乙酰化 -VKKTTNVHSL- 生物素	MAP2 7-16
13	Pep2-8	乙酰化 -KKTTNVHSLF- 生物素	MAP2 8-17
14	Pep2-9	乙酰化 -KTTNVHSLFH- 生物素	MAP2 9-18
15	Pep3-1	乙酰化 -LTPKKTTNVL- 生物素	MAP3 1-10
16	Pep3-2	乙酰化 -TPKKTTNVL- 生物素	MAP3 2-11
17	Pep3-3	乙酰化 -PKKTTNVLTV- 生物素	MAP3 3-12

序号	肽名称	序列	免疫肽中位置
18	Pep3-4	乙酰化 -KKTTNVLTVPT- 生物素	MAP3 4-13
19	Pep3-5	乙酰化 -KTTNVLTVPT- 生物素	MAP3 5-14
20	Pep3-6	乙酰化 -TTNVLTVPTN- 生物素	MAP3 6-15
21	Pep3-7	乙酰化 -TNVLTVPTNI- 生物素	MAP3 7-16
22	Pep3-8	乙酰化 -NVLTVPNTIP- 生物素	MAP3 8-17
23	Pep3-9	乙酰化 -VLTVPNTIPG- 生物素	MAP3 9-18
24	Pep4-1	乙酰化 -LTPKKLTQEN- 生物素	MAP4 1-10
25	Pep4-2	乙酰化 -TPKKLTQENQ- 生物素	MAP4 2-11
26	Pep4-3	乙酰化 -PKKLTQENQN- 生物素	MAP4 3-12
27	Pep4-4	乙酰化 -KKLTQENQNR- 生物素	MAP4 4-13
28	Pep4-5	乙酰化 -KLTQENQNRG- 生物素	MAP4 5-14
29	Pep4-6	乙酰化 -LTQENQNRGT- 生物素	MAP4 6-15, MAP5 10-19
30	Pep4-7	乙酰化 -TQENQNRGTH- 生物素	MAP4 7-16, MAP5 11-20
31	Pep4-8	乙酰化 -QENQNRGTHI- 生物素	MAP4 8-17
32	Pep4-9	乙酰化 -ENQNRGTHIY- 生物素	MAP4 9-18
33	Pep4-10	乙酰化 -QENQNRGTHI- 生物素	MAP4 10-19
34	Pep4-11	乙酰化 -ENQNRGTHIY- 生物素	MAP4 11-20
35	Pep5-1	乙酰化 -TIYNTNIPGL- 生物素	MAP5 1-10
36	Pep5-2	乙酰化 -IYNTNIPGLT- 生物素	MAP5 2-11

28	Pep4-5	乙酰化 -KLTQENQNRG- 生物素	MAP4 5-14
37	Pep5-3	乙酰化 -YNTNIPGLTQ- 生物素	MAP5 3-12
38	Pep5-4	乙酰化 -NTNIPGLTQE- 生物素	MAP5 4-13
39	Pep5-5	乙酰化 -TNIPGLTQEN- 生物素	MAP5 5-14
40	Pep5-6	乙酰化 -NIPGLTQENQ- 生物素	MAP5 6-15
41	Pep5-7	乙酰化 -IPGLTQENQN- 生物素	MAP5 7-16
42	Pep5-8	乙酰化 -PGLTQENQNR- 生物素	MAP5 8-17
43	Pep5-9	乙酰化 -GLTQENQNRG- 生物素	MAP5 9-18

[0226] 免疫的标准时期之后,用标准方法从每只小鼠收集免疫血清,如以下用 ELISA 进行检测:

[0227] 用 100 μ l/ 孔或 50 μ l/ 孔链亲和素 (Sigma Catalog No. S4762 或相似产品, 50mM 的碳酸盐缓冲液, pH 9.6 中 5 μ g/ml) 包被 ELISA 板 (Corning 3369 或相似产品)。将平板于 4°C 孵育过夜或室温孵育 2 小时。孵育后,将平板用 PBS+0.05% Tween-20 (PBST 缓冲液) 漂洗 3 次。漂洗后,用 250 μ l/ 孔 PBST 对平板进行封闭,于室温孵育 1 小时或 4°C 孵育过夜。除去 PBST,加入 100 μ l/ 孔或 50 μ l/ 孔浓度为 5 μ g/ml (在 PBS 中稀释) 的选自表 4 的检测生物素化多肽。将平板于室温孵育约 30 至 60 分钟。孵育后,平板用 PBST 漂洗 3 次。然后,加入 100 μ l/ 孔或 50 μ l/ 孔检测血清 (即来自检测血液的血清),将平板于室温孵育 1 小时或于 4°C 孵育过夜。为滴定免疫反应性,在检测前一般将血清稀释到 1 : 500、1 : 2000、1 : 8000 或 1 : 32000。孵育后,平板用 PBST 漂洗 3 次。为检测抗体结合,向每孔加入 1 : 10,000 稀释的山羊抗 - 小鼠 IgG (和 IgM) -HRP 缀合物 (Jackson Immuno orderNo. 115-036-071, 或相似产品)。将平板于室温再孵育 1 小时,然后用 PBST 漂洗 5 次。加入 HRP 底物 (Sigma Fast OPD) 并在黑暗中于室温孵育 30-60 分钟。如果 HRP 反应没有终止,用 96- 孔比色检测仪在 OD450 对平板进行读数。备选地,用 1.25M 硫酸终止 HRP 反应,并在 OD492 对平板进行读数。

[0228] 检测了来自第 1、2 及 3 组小鼠的 12 份检测血液。没有观察到来自第 1 及第 3 组小鼠的免疫反应,对这些小鼠不再进一步研究。第 2 组的所有 4 只小鼠对筛选多肽 Pep2-0 显示强免疫反应 ($> 1 : 32,000$)。此外,由于 MAP2 及 MAP1/MAP3 之间的序列同源性,来自第 2 组 4 只小鼠中的 2 只小鼠 (小鼠 #2-1 及 #2-4) 的免疫血清与对第 1 组及第 3 组设计的筛选多肽显示交叉反应性。这些结果与小鼠 #2-1 及 #2-4 表达抗体一致,所述抗体识别在 ELISA 测定法中使用的一种以上的筛选抗原中存在的不同而简明的表位。#2-1 及 #2-4 血清对跨越第 1、第 2 及第 3 组小鼠的所有 3 种免疫多肽序列的 23 种 10mer 生物素化多肽的检测也显示宽的交叉反应性。

[0229] 通过 ELISA 检测了来自第 4-5 组的 8 份检测血液。第 4 组小鼠对它们相关的筛选多肽 Pep4-0 显示中等反应,而对第 3 组设计的筛选多肽 Pep3-0 显示强交叉反应性。虽然 Pep4-0 与 Pep5-0 之间有显著的序列同一性,但是第 4 组小鼠对 Pep5-0 一般不显示实质的交叉反应性。相比,第 5 组 4 只小鼠中的 3 只(小鼠 5-2、5-3、5-4)对它们的筛选多肽 Pep5-0 及相关筛选多肽 Pep4-0 均显示强免疫反应性。虽然有 5 个氨基酸区的序列同一性,但是来自第 5 组中应答小鼠的血清对 Pep3-0 一般不显示实质的交叉反应。#5-2 及 #5-3 血清对跨越第 4 组及第 5 组小鼠所有三种免疫多肽序列的 23 种 10mer 生物素化多肽的检测显示出宽的但有区别的反应模式,作图多肽跨越第 4 及 5 组小鼠免疫多肽的序列。

[0230] 如在表 6 及图 1 中概述的,第 2 组小鼠 #1 及 #4,和第 5 组,小鼠 #2 及 #3 显示最佳的免疫反应。选择这些小鼠用于杂交瘤融合。

[0231] 表 6. 第 2 及第 5 组中选择的小鼠对筛选多肽 1-5 的免疫反应性及交叉反应性

小鼠	肽 1	肽 2	肽 3	肽 4	肽 5
2-1	0.726	0.850	0.323	未测试	未测试
2-2	0.250	1.167	0.213	未测试	未测试
2-3	0.222	0.685	0.141	未测试	未测试
2-4	0.776	0.970	0.353	未测试	未测试
5-1	未测试	未测试	0.178	0.28	0.979
5-2	未测试	未测试	0.146	1.714	1.548
5-3	未测试	未测试	0.13	1.479	1.773
5-4	未测试	未测试	0.128	1.915	1.464

[0233] 处死动物,收获淋巴结及脾,然后用标准方法,用 P3 小鼠骨髓瘤细胞系作为融合配偶体产生 B 细胞杂交瘤融合物,将融合物平板接种并孵育 11-14 天后进行筛选。

[0234] 在第一轮筛选中,一般如以上描述地,用对应的筛选多肽 2-0 及 5-0,对来自第 2 组及第 5 组小鼠的杂交瘤在 96 孔板中通过 ELISA 分析。几轮筛选后,鉴定了 48 种阳性杂交瘤细胞系,并转移到 24 孔板用于扩大及额外表征,包括表位作图。在 48 种阳性细胞系中,33 种来自接受 MAP2 免疫原的第 2 组动物而剩余 15 种源自第 5 组动物。大部分杂交瘤细胞系(~94%)为自脾收获的 B 细胞的融合产物。48 种杂交瘤细胞系的中 13 种表达 IgG,25 种表达 IgM,剩余 10 种杂交瘤细胞系表达 IgG 和 IgM 或不表达 IgG 或 IgM 并因此表达 IgA 或 IgE。

[0235] 在第二轮筛选中,选择用于扩大的杂交瘤对相关筛选多肽(多肽 2-0 或多肽 5-0)进行了重新检测。在 24 孔扩大期后 48 种杂交瘤中的 13 种显示序列特异地结合筛选多肽 2-0。其它杂交瘤非特异结合(即结合多种寡肽序列)、不结合(反映了假阳性或在 24 孔板转移及随后增殖期间克隆不稳定性及丢失)或结合含有 BSA 的对照孔。

[0236] 如以上描述,用 3 组不同的 10mer C-末端生物素化作图多肽:多肽 1-1 至 1-5;2-1 至 2-9;及 3-1 至 3-9(参见表 5),用 ELISA 对特异结合筛选多肽 2-0 的 13 种杂交瘤进行表位作图。12 种杂交瘤细胞系中 10 种与一种作图多肽 2-1 显示出最大反应性,杂交瘤 2.03 及 2.11 与不同重叠组的作图多肽:多肽 2-1 至 2-3 及 2-7 至 2-9 显示强结合。因为这些数

据显示与用于大部分杂交瘤细胞系的单一作图多肽的强反应性,我们考虑到可能与作图多肽固定化(特别地,生物素-亲和素固定化)相关的位阻防止了抗体结合到 10mer 关连系列中存在的表位,因此导致可能偏离的 ELISA 表位作图结果。因此,我们用竞争结合测定法评估了表位特异性。

[0237] 评估了个别作图多肽抑制抗体结合附着在链亲和素包被的 96 孔板上的 2-0 筛选多肽的能力。在此形式下,10mer 作图多肽没有束缚在链亲和素的结合袋中,因此与 13 种杂交瘤组中存在的反应抗体的相互作用应该没有位阻。用标准方法进行竞争实验,所述标准方法用附着在链亲和素包被的 96 孔板上的 2-0 筛选多肽并向每孔加入 10mer 作图多肽。

[0238] 用竞争结合测定法确定了通过 13 种杂交瘤中的 10 种杂交瘤识别的表位。8 种杂交瘤对表位 PEDTG 特异,杂交瘤 2.03 对表位 DTG 特异并且杂交瘤 2.11 识别表位 KKTTN。杂交瘤 2.31 显示复杂抑制模式,表明此细胞系为 2 种或多种特异性的混合物并且应进行亚克隆以分开个别反应性。最后,在竞争抑制测定中,杂交瘤 1.02 及 2.12 显示弱鉴别力。此分析结果在表 7 中概述。

[0239] 表 7. 通过竞争抑制预测的表位

[0240]

1.01, 2.01, 2.04, 2.06, 2.07, 2.08, 2.10 和 2.23 的模式: PEDTG	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
2.03 模式: DTG	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-3	LFHPEDTGQV
P1-4	FHPEDTGQVK
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
P2-2	EDTGQVKKTT
P2-3	DTGQVKKTTN
2.11 模式: KKTTN	
P1-4	FHPEDTGQVK ???
P2-3	DTGQVKKTTN
P2-4	TGQVKKTTNV
P2-5	GQVKKTTNVH
P2-6	QVKKTTNVHS
P2-7	VKKTTNVHSL
P2-8	
P2-9	
P3-1	LTPKKTTNVL
P3-2	TPKKTTNVLT
P3-3	PKKTTNVLTV
P3-4	KKTTNVLTVP
2.31 模式: 两种克隆的混合物?	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
P2-7	VKKTTNVHSL
P2-8	KKTTNVHSLF
P2-9	KTINVHSLFH
P3-2	TPKKTTNVLT
P3-3	PKKTTNVLTV
1.02 和 2.12 模式: 模式不清楚	

[0241] 竞争结合测定重复两次,证实杂交瘤 2.11 识别表位 KTTN,而不是如最初实验中提示的KKTTN。表位竞争结合测定证实了对其它杂交瘤如上描述的表位特征。此最新分析结果在表 8 中概述。正在制备杂交瘤 2.03(也称为 DA001-2.03)、2.04(DA001-2.04)及 2.11(也称为 DA001-2.11)用于在 ATCC 保藏。

[0242] 表 8. 通过竞争抑制预测的表位的最新的且经证实的表

[0243]

1.01, 2.01, 2.04, 2.06, 2.07, 2.08, 2.10 和 2.23 的模式: PEDTG	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
2.03 模式: DTG	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-3	LFHPEDTGQV
P1-4	FHPEDTGQVK
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
P2-2	EDTGQVKKTT
P2-3	DTGQVKKTTN
2.11 模式: KKTTN	
P1-4	FHPEDTGQVK ???
P2-3	DTGQVKKTTN
P2-4	TGQVKKTTNV
P2-5	GQVKKTTNVH
P2-6	QVKKTTNVHS
P2-7	VKKTTNVHSL
P2-8	
P2-9	
P3-1	LTPKKTTNVL
P3-2	TPKKTTNVLT
P3-3	PKKTTNVLTV
P3-4	KKTTNVLTVP
2.31 模式: 两种克隆的混合物?	
P1-1	HSLFHPEDTG
P1-2	SLFHPEDTGQ
P1-5	HPEDTGQVKK
P2-1	PEDTGQVKKT
P2-7	VKKTTNVHSL
P2-8	KKTTNVHSLF
P2-9	KTINVHSLFH
P3-2	TPKKTTNVLT
P3-3	PKKTTNVLTV
1.02 和 2.12 模式: 模式不清楚	

[0244] 实施例 2. 小表位抗体的制备

[0245] 进行了基于噬菌体展示抗体筛选鉴定抗体的方法。用于筛选阳性抗体的五种肽序列在表 9 中显示。这些序列组合也用于评估选择的抗体的交叉反应性。

[0246] 表 9. 筛选多肽的设计

[0247]

肽	序列
P1	CXXXXXDTGXXXXX

肽	序列
P6	CXXXXXDTGXXXXXX
P7	CXXXXXAQVXXXXXX
P8	CXXXXXIARXXXXXX
P9	CXXXXXLSHXXXXXX

[0248] 表 9 说明：字母“X”表示除半胱氨酸、甲硫氨酸及色氨酸外的天然存在的 L-氨基酸的混合物。

[0249] 在六轮富集后选择阳性噬菌体。对五种筛选肽的噬菌体 ELISA 筛选的结果在表 10 中显示。对 P1 筛选了总共 96 种噬菌体；48 种噬菌体用于筛选多肽 P6-P9。在所用情况下，都鉴定了背景上的阳性噬菌体。

[0250] 表 10. 经富集噬菌体对筛选多肽的反应性

[0251]

多肽 1		多肽 6		多肽 7		多肽 8		多肽 9	
噬菌体	OD	噬菌体	OD	噬菌体	OD	噬菌体	OD	噬菌体	OD
L50P1_1	0.0781	L50P6_1	1.6477	I50P7_1	0.0791	L50P8_1	0.5249	L50P9_1	0.0813
L50P1_2	0.0737	L50P6_2	1.6612	I50P7_2	0.3119	L50P8_2	0.4247	L50P9_2	0.4743
L50P1_3	0.0684	L50P6_3	1.5365	I50P7_3	0.2111	L50P8_3	0.8174	L50P9_3	0.6882

[0252]

L50P1_4	0.3906	L50P6_4	1.4133	I50P7_4	1.6251	L50P8_4	0.6231	L50P9_4	0.5747
L50P1_5	0.3333	L50P6_5	0.9797	I50P7_5	1.3357	L50P8_5	0.5497	L50P9_5	0.4527
L50P1_6	0.0667	L50P6_6	0.1036	I50P7_6	0.2128	L50P8_6	0.7834	L50P9_6	0.6045
L50P1_7	0.0668	L50P6_7	0.5592	I50P7_7	1.4445	L50P8_7	0.4143	L50P9_7	0.0944
L50P1_8	0.0689	L50P6_8	1.5017	I50P7_8	0.0694	L50P8_8	0.8192	L50P9_8	0.0762
L50P1_9	0.0714	L50P6_9	1.1022	I50P7_9	0.7113	L50P8_9	0.5725	L50P9_9	0.3449
L50P1_10	0.0683	L50P6_10	1.1577	I50P7_10	0.1787	L50P8_10	0.6108	L50P9_10	0.0721
L50P1_11	0.0813	L50P6_11	0.4477	I50P7_11	0.1912	L50P8_11	0.2095	L50P9_11	0.6566

L50P1_4	0.3906	L50P6_4	1.4133	I50P7_4	1.6251	L50P8_4	0.6231	L50P9_4	0.5747
L50P1_12	0.1168	L50P6_12	1.2041	I50P7_12	0.1158	L50P8_12	0.6757	L50P9_12	0.0831
L50P1_13	0.0717	L50P6_13	1.6751	I50P7_13	0.0729	L50P8_13	0.5143	L50P9_13	0.4898
L50P1_14	0.4481	L50P6_14	1.1052	I50P7_14	0.1238	L50P8_14	0.659	L50P9_14	0.5458
L50P1_15	0.8381	L50P8_15	0.218	I50P7_15	0.0679	L50P8_15	1.0582	L50P9_15	0.0702
L50P1_16	0.2818	L50P6_16	0.0787	I50P7_16	0.0688	L50P8_16	0.8478	L50P9_16	0.4297
L50P1_17	0.4623	L50P6_17	0.066	I50P7_17	0.0847	L50P8_17	0.7276	L50P9_17	0.3535
L50P1_18	0.0614	L50P6_18	0.1961	I50P7_18	1.0256	L50P8_18	0.7266	L50P9_18	0.0757
L50P1_19	0.0595	L50P6_19	1.1042	I50P7_19	1.5344	L50P8_19	0.6607	L50P9_19	0.07
L50P1_20	0.0821	L50P6_20	0.0618	I50P7_20	0.4507	L50P8_20	0.8016	L50P9_20	0.547
L50P1_21	0.08	L50P6_21	1.155	I50P7_21	0.2637	L50P8_21	0.754	L50P9_21	0.5593
L50P1_22	0.0632	L50P6_22	1.4566	I50P7_22	0.1088	L50P8_22	0.4702	L50P9_22	0.6068
L50P1_23	0.0643	L50P6_23	0.129	I50P7_23	1.0236	L50P8_23	0.3573	L50P9_23	0.5225
L50P1_24	0.0817	L50P6_24	1.2605	I50P7_24	0.1236	L50P8_24	0.7595	L50P9_24	0.8072
L50P1_25	0.0917	L50P6_25	0.0583	I50P7_25	0.0965	L50P8_25	0.7424	L50P9_25	0.5658
L50P1_26	0.0791	L50P6_26	0.0848	I50P7_26	0.898	L50P8_26	0.7334	L50P9_26	0.0758
L50P1_27	0.0619	L50P6_27	0.0805	I50P7_27	0.1256	L50P8_27	0.7748	L50P9_27	0.3991
L50P1_28	0.4974	L50P6_28	1.5586	I50P7_28	0.7453	L50P8_28	0.6577	L50P9_28	0.5235
L50P1_29	0.0596	L50P6_29	0.0778	I50P7_29	0.1149	L50P8_29	0.5632	L50P9_29	0.0699

L50P1_4	0.3906	L50P6_4	1.4133	I50P7_4	1.6251	L50P8_4	0.6231	L50P9_4	0.5747
L50P1_30	0.0582	L50P6_30	1.5647	I50P7_30	0.076	L50P8_30	0.5071	L50P9_30	0.516
L50P1_31	0.4591	L50P6_31	0.0962	I50P7_31	1.4382	L50P8_31	0.5892	L50P9_31	0.2835
L50P1_32	0.0566	L50P6_32	0.0603	I50P7_32	1.5916	L50P8_32	0.6455	L50P9_32	0.0733
L50P1_33	0.0622	L50P6_33	0.0815	I50P7_33	0.8539	L50P8_33	0.4008	L50P9_33	0.5253
L50P1_34	0.0584	L50P6_34	0.1512	I50P7_34	1.0193	L50P8_34	0.4515	L50P9_34	0.5407
L50P1_35	0.7212	L50P6_35	0.1344	I50P7_35	0.1178	L50P8_35	0.4302	L50P9_35	0.0744
L50P1_36	0.0843	L50P6_36	0.1644	I50P7_36	1.2705	L50P8_36	0.3179	L50P9_36	0.613
L50P1_37	0.4181	L50P6_37	1.2164	I50P7_37	0.4899	L50P8_37	0.4526	L50P9_37	0.5239
L50P1_38	0.4914	L50P6_38	1.3835	I50P7_38	0.142	L50P8_38	0.7307	L50P9_38	0.1844
L50P1_39	0.0607	L50P6_39	0.1062	I50P7_39	0.5033	L50P8_39	0.7737	L50P9_39	0.0804
L50P1_40	0.5813	L50P6_40	0.0615	I50P7_40	0.7136	L50P8_40	0.6617	L50P9_40	0.3825
L50P1_41	0.3373	L50P6_41	1.3978	I50P7_41	0.2031	L50P8_41	0.6766	L50P9_41	0.0748
L50P1_42	0.0561	L50P6_42	0.0758	I50P7_42	0.0669	L50P8_42	0.6741	L50P9_42	0.2942
L50P1_43	0.3979	L50P6_43	0.0831	I50P7_43	0.1266	L50P8_43	0.6942	L50P9_43	0.0707
L50P1_44	0.0587	L50P6_44	1.5906	I50P7_44	0.0693	L50P8_44	0.6275	L50P9_44	0.0722
L50P1_45	0.0576	L50P6_45	0.081	I50P7_45	0.1209	L50P8_45	0.3312	L50P9_45	0.5045
L50P1_46	0.0699	L50P6_46	1.4628	I50P7_46	0.4689	L50P8_46	0.3838	L50P9_46	0.2859
L50P1_47	0.4785	L50P6_47	0.1462	I50P7_47	0.0686	L50P8_47	0.3922	L50P9_47	0.4253
L50P1_48	0.6597	L50P6_48	0.0738	阴性对照	0.0634	L50P8_48	0.5962	阴性对照	0.1297

L50P1_4	0.3906	L50P6_4	1.4133	L50P7_4	1.6251	L50P8_4	0.6231	L50P9_4	0.5747
阴性对照	0.0738	阴性对照	0.1297			阴性对照	0.1297		

[0253] 在初步筛选鉴定为阳性的第二轮筛选中,对所有五种多肽进行噬菌体 ELISA 测定。选择了多达五种阳性噬菌体用于第二轮筛选。图 2 显示了使用此测定法的大多数选择克隆的结果。所有五种阳性噬菌体对多肽产生高于 BSA 的显著信号,选自 P1 (L50P1_15)、P8 (L50P8_5) 及 P9 (L50P9_5) 的噬菌体在此半定量测定中显示特异性。

[0254] 将用于 L50P1_15 的反应性抗体亚克隆到载体中用于单链抗体的细菌表达。用表面等离子体共振 (SPR) 生物传感器测定法对粗周质制备物进行分析以监测复合体形成和蛋白质从固定化肽的解离 (Malmberg 等人, 1995)。图 3 显示了针对五种多肽及 BSA 的单链抗体的 SPR 分布图。抗体对肽 1 具有最高亲和力,估计 K_d 为 2×10^{-8} 。

[0255] 实施例 3. 蛋白质分布图及生物标记的研发

[0256] 在用于蛋白质分布图的一个示例性方法中,来自健康个体及来自受具临床重要性的疾病侵袭个体的血清进行 (a) 压实最丰富的蛋白质成分;(b) 剩余的丰度较低的蛋白质的去糖基化;(c) 压实的蛋白质组中存在的半胱氨酸残基的还原及烷基化;(d) 完全消化压实的蛋白质组;(e) 如以上描述用小表位抗体对产生的多肽片段分级分离;和 (f) 比较肽成分组成及相对多度以鉴定与特定疾病相关的候选生物标记,所述肽成分来自健康及受侵袭患者的表位富集的级分。

[0257] 用小表位抗体的分级分离用一组大约 100 种不同特异性的小表位抗体平行进行。基于一组标准选择每种抗体,所述标准包括表位大小、血清蛋白质组中的表位多度、特异性、亲和性及采样丰余性。虽然一些 4mer 或 5mer 的表位满足多度标准,但是通过抗体识别的表位主要为 3mer,每种表位在血清蛋白质组的 0.5-3% 成分中存在。每种抗体以不依赖背景的方式及高亲和力识别它的关连表位。用于分级分离的全组小表位抗体提供了 3-5 倍采样丰余性以适应对于不同蛋白质表达水平及对于组中每种抗体的捕获效率预期的变异性。

[0258] 质谱分析法用于分析每个小表位抗体级分的肽组成及肽组分的表达水平。鉴定了在健康及患病个体中差别表达的生物标记。研发了能区分健康个体及受侵袭个体的 ELISA 测定法,所述 ELISA 测定法基于血浆或血清中存在的经鉴定的生物标记的特定水平。

[0259] 虽然为了清楚理解,上述发明通过阐明及实施例方式进行了详细描述,但是本领域技术人员显而易见的是可进行某些改变及修饰而不背离本发明精神及范围。因此,不应将本说明书理解为限制本发明的范围,本发明范围通过所附权利要求书描绘。

[0260] 此处引用的所有出版物、专利和专利申请为了所有目的和相同程度地引入本文作为参考,就像每种单独的出版物、专利和专利申请被专门和分别指出引入作为参考一样。

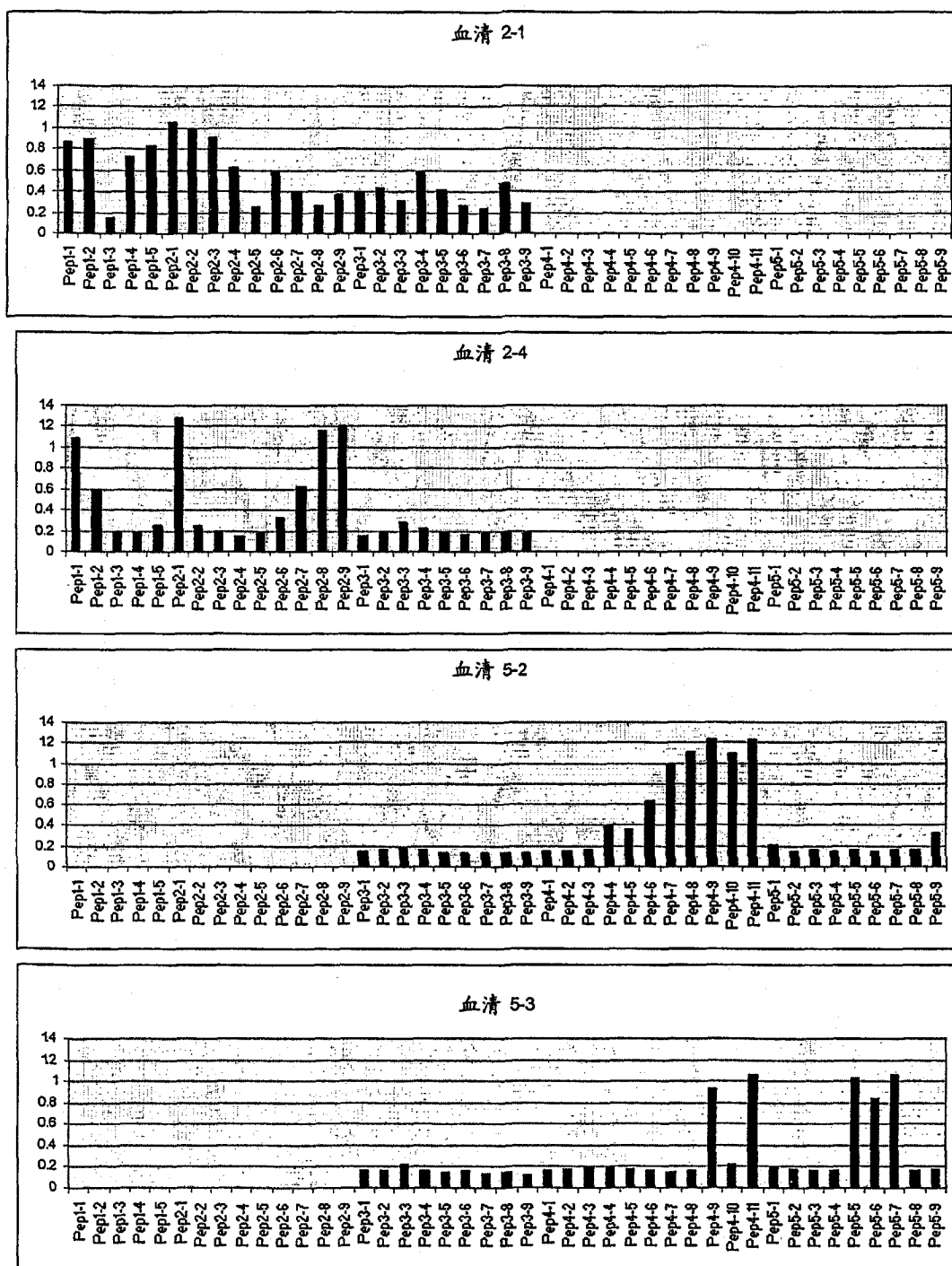


图 1

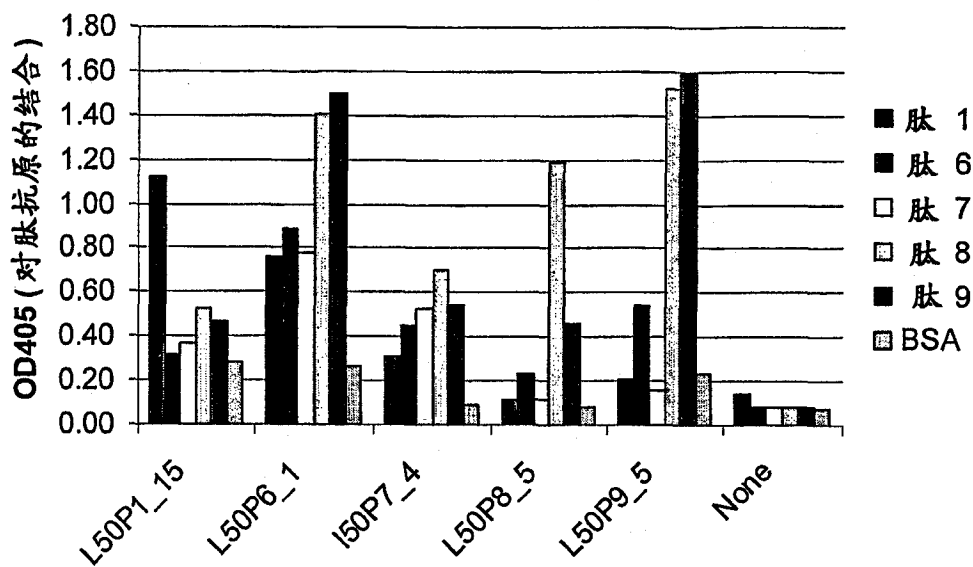


图 2

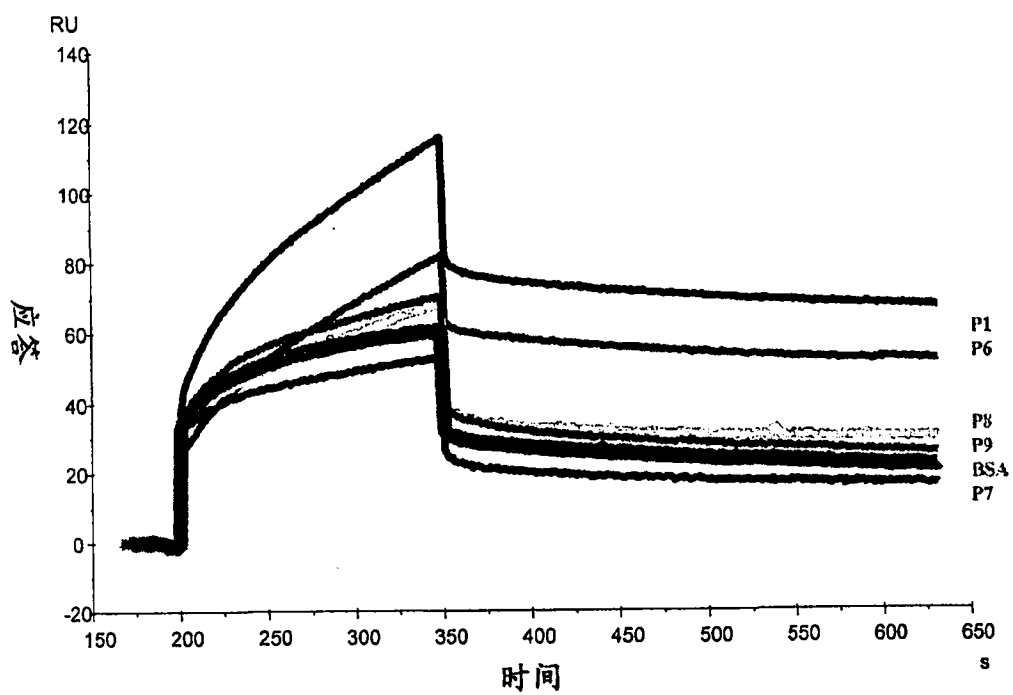


图 3