

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G01N 33/50 (2006.01)

G01N 33/574 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200680050593.6

[43] 公开日 2009 年 6 月 10 日

[11] 公开号 CN 101454668A

[22] 申请日 2006.11.14

[21] 申请号 200680050593.6

[30] 优先权

[32] 2005.11.14 [33] US [31] 60/735,854

[86] 国际申请 PCT/US2006/044117 2006.11.14

[87] 国际公布 WO2007/059094 英 2007.5.24

[85] 进入国家阶段日期 2008.7.7

[71] 申请人 拜耳医药保健有限责任公司

地址 美国纽约州

[72] 发明人 S·威尔赫姆 J·艾廷

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 韦 东

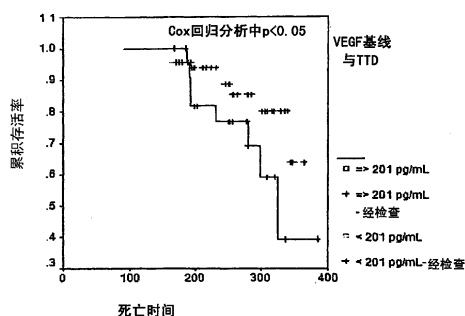
权利要求书 5 页 说明书 29 页 附图 2 页

[54] 发明名称

癌症预测与预后以及监测癌症治疗的方法

[57] 摘要

本发明也涉及用于癌症预测与预后的生物标记物及其应用以及该生物标记物在监测癌症治疗疗效中的应用。具体地说,本发明涉及使用 VEGF 和 sVEGFR 作为用索拉非尼治疗对象的生物标记物。



1. 一种监测用索拉非尼治疗的癌症患者的响应的方法，所述方法包括：检测患者样本中 VEGF 和/或 sVEGFR-2 的水平或活性以及将所述水平与标准值进行比较。

2. 如权利要求 1 所述的方法，所述方法包括检测所述索拉非尼治疗的患者样本和所述对照样本中 VEGF 和/或 sVEGFR-2 的 mRNA 水平。

3. 如权利要求 1 所述的方法，所述方法包括检测所述索拉非尼治疗的患者样本和所述对照样本中 VEGF 和/或 sVEGFR-2 的蛋白质水平。

4. 如权利要求 2 所述的方法，其特征在于，所述 mRNA 水平通过以下过程检测：使所述患者样本与能够和所述 mRNA 特异性结合的试剂相接触并测量所述特异性结合的试剂量。

5. 如权利要求 3 所述的方法，其特征在于，所述蛋白质水平通过以下过程检测：使所述患者样本与能够与所述蛋白质特异性结合的试剂相接触并测量所述特异性结合的试剂量。

6. 如权利要求 4 所述的方法，其特征在于，所述结合试剂包含至少一种多核苷酸。

7. 如权利要求 5 所述的方法，其特征在于，所述结合试剂包含至少一种抗体。

8. 如权利要求 1 所述的方法，其特征在于，所述患者样本包括体液。

9. 如权利要求 2 所述的方法，其特征在于，所述 mRNA 水平通过 RNA 印迹分析、RT-PCR 或 cDNA 微阵列来检测。

10. 如权利要求 3 所述的方法，其特征在于，所述蛋白质水平通过免疫印迹、免疫沉淀或 ELISA 分析来检测。

11. 如权利要求 8 所述的方法，其特征在于，所述体液是血液。

12. 如权利要求 2 所述的方法，其特征在于，所述癌症是原发性肿瘤或转移性肿瘤。

13. 如权利要求 2 所述的方法，其特征在于，所述癌症是癌、淋巴瘤、白血病、骨髓瘤、肉瘤、成胶质细胞瘤、星形细胞瘤、黑色素瘤或肾母细胞瘤。

14. 如权利要求 12 所述的方法，其特征在于，所述癌症是乳腺、呼吸道、脑、生殖器官、消化道、尿道、眼、肝、皮肤、头和颈、甲状腺、甲状旁腺、血液或肌

肉的癌症。

15. 如权利要求 12 所述方法，其特征在于，所述乳腺癌是浸润性导管癌、浸润性小叶癌、原位管癌和原位小叶癌。

16. 如权利要求 12 所述的方法，其特征在于，所述呼吸道癌症是小细胞肺癌、非小细胞肺癌、支气管腺瘤或胸膜肺母细胞瘤。

17. 如权利要求 12 所述的方法，其特征在于，所述脑癌是脑干和眼下胶质瘤、小脑和大脑星形细胞瘤、髓母细胞瘤、室管膜瘤、神经外胚层或松果体瘤。

18. 如权利要求 12 所述的方法，其特征在于，所述生殖器官癌症是前列腺癌、睾丸癌、子宫内膜癌、子宫颈癌、卵巢癌、阴道癌、外阴癌、或子宫肉瘤。

19. 如权利要求 12 所述的方法，其特征在于，所述消化道癌症是肛门、结肠、结肠直肠、食管、胆囊、胃、胰脏、直肠、小肠或唾液腺癌症。

20. 如权利要求 12 所述的方法，其特征在于，所述尿道癌症是膀胱、阴茎、肾、肾脏、骨盆、输尿管或尿道癌症。

21. 如权利要求 12 所述的方法，其特征在于，所述眼部癌症是眼内黑色素瘤或视网膜母细胞瘤。

22. 如权利要求 12 所述的方法，其特征在于，所述肝癌包括肝细胞癌、胆管癌、或混合型肝细胞癌。

23. 如权利要求 12 所述的方法，其特征在于，所述皮肤癌包括鳞状细胞癌、卡波济氏肉瘤、恶性黑色素瘤、Merkel 细胞皮肤癌或非黑色素瘤皮肤癌。

24. 如权利要求 12 所述的方法，其特征在于，所述头颈部癌症包括喉、下咽、鼻咽或口咽癌、唇癌或口腔癌。

25. 如权利要求 12 所述的方法，其特征在于，所述血液癌症包括 AIDS-相关淋巴瘤、非霍奇金淋巴瘤、表皮 T-细胞淋巴瘤、霍奇金病、中枢神经系统淋巴瘤、急性髓细胞性白血病、急性淋巴细胞性白血病、慢性淋巴细胞性白血病、慢性髓细胞性白血病或多毛细胞白血病。

26. 如权利要求 12 所述的方法，其特征在于，所述肉瘤包括软组织肉瘤、骨肉瘤、恶性纤维组织细胞瘤、淋巴肉瘤、横纹肌肉瘤、急性髓细胞性白血病、急性淋巴细胞性白血病、慢性淋巴细胞性白血病、慢性髓细胞性白血病或多毛细胞白血病。

27. 一种监测通过给予索拉非尼治疗肾细胞癌的患者响应的的方法，所述方法包括：

(a)测定用索拉非尼治疗之前从患者获取的第一血浆样品中生物标记物 VEGF 和/或 sVEGFR-2 的表达水平;

(b)测定开始用索拉非尼治疗后从患者获取的第二血浆样品中 VEGF 和/或 sVEGFR-2 的表达水平; 和

(c)将第二样品中 VEGF 和/或 sVEGFR-2 的表达水平与第一样品中所述生物标记物的表达水平进行比较;

其中, 与第一样品中 VEGF 和/或 sVEGFR-2 的表达水平相比, 第二样品中 VEGF 和/或 sVEGFR-2 的表达水平的改变表明用索拉非尼治疗的疗效。

28. 一种评价癌症患者的病情的方法, 所述方法包括:

a.测定从患者获取的生物样品中 VEGF 和/或 sVEGFR-2 的表达水平;

b.将 VEGF 和/或 sVEGFR-2 的表达水平与以下一种或多种进行比较:

i)从一个或多个非疑似患有癌症的对象获取的类似样品中的表达水平;

ii)从一个或多个疑似患有癌症的对象获取的类似样品中的表达水平; 或

iii)在另一时间从该患者获取的类似样品中的表达水平,

其中, 样品(a)中 VEGF 和/或 sVEGFR-2 的表达水平和一种或多种比较水平间的差异在疾病状态和/或疾病状态的预计变化方面与患者病情有关。

29. 如权利要求 28 所述的方法, 其特征在于, 所述评价患者病情包括以下一种或多种:

诊断疾病状态,

监测疾病状态的变化, 和

预测治疗或不治疗情况下疾病状态的变化。

30. 如权利要求 28 所述的方法, 其特征在于, 所述生物样品选自血液、羊水、血浆、血清、精液、骨髓、尿液或活检组织。

31. 如权利要求 28 所述的方法, 其特征在于, 所述生物样品是血浆。

32. 如权利要求 28 所述的方法, 其特征在于, 所述癌症是: (a)乳腺、呼吸道、脑、生殖器官、消化道、尿道、眼、肝、皮肤、头和颈、甲状腺、甲状旁腺的实体瘤或者它们不同的转移瘤, (b)淋巴瘤, (c)肉瘤或(d)白血病。

33. 如权利要求 32 所述方法, 其特征在于, 所述乳腺癌是浸润性导管癌、浸润性小叶癌、原位管癌和原位小叶癌。

34. 如权利要求 32 所述的方法, 其特征在于, 所述呼吸道癌症是小细胞肺癌、非小细胞肺癌、支气管腺瘤或胸膜肺母细胞瘤。

35. 如权利要求 32 所述的方法, 其特征在于, 所述脑癌是脑干和眼下胶质瘤、小脑和大脑星形细胞瘤、髓母细胞瘤、室管膜瘤、神经外胚层或松果体瘤。

36. 如权利要求 32 所述的方法, 其特征在于, 所述生殖器官癌症是前列腺癌、睾丸癌、子宫内膜癌、子宫颈癌、卵巢癌、阴道癌、外阴癌、或子宫肉瘤。

37. 如权利要求 32 所述的方法, 其特征在于, 所述消化道癌症是肛门、结肠、结肠直肠、食管、胆囊、胃、胰脏、直肠、小肠或唾液腺癌症。

38. 如权利要求 32 所述的方法, 其特征在于, 所述尿道癌症是膀胱、阴茎、肾、肾脏、骨盆、输尿管或尿道癌症。

39. 如权利要求 32 所述的方法, 其特征在于, 所述眼部癌症是眼内黑色素瘤或视网膜母细胞瘤。

40. 如权利要求 32 所述的方法, 其特征在于, 所述肝癌包括肝细胞癌、胆管癌、或混合型肝细胞癌。

41. 如权利要求 32 所述的方法, 其特征在于, 所述皮肤癌包括鳞状细胞癌、卡波济氏肉瘤、恶性黑色素瘤、Merkel 细胞皮肤癌或非黑色素瘤皮肤癌。

42. 如权利要求 32 所述的方法, 其特征在于, 所述头颈部癌症包括喉、下咽、鼻咽或口咽癌、唇癌或口腔癌。

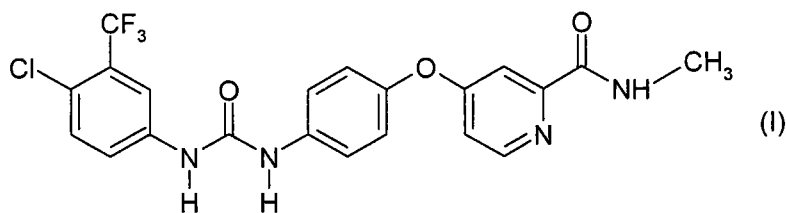
43. 如权利要求 32 所述的方法, 其特征在于, 所述血液癌症包括 AIDS-相关淋巴瘤、非霍奇金淋巴瘤、表皮 T-细胞淋巴瘤、霍奇金病、中枢神经系统淋巴瘤、急性髓细胞性白血病、急性淋巴细胞性白血病、慢性淋巴细胞性白血病、慢性髓细胞性白血病或多毛细胞白血病。

44. 如权利要求 32 所述的方法, 其特征在于, 所述肉瘤包括软组织肉瘤、骨肉瘤、恶性纤维组织细胞瘤、淋巴肉瘤和横纹肌肉瘤、急性髓细胞性白血病、急性淋巴细胞性白血病、慢性淋巴细胞性白血病、慢性髓细胞性白血病或多毛细胞白血病。

45. 如权利要求 28 所述的方法, 其特征在于, 所述癌症是肾细胞癌。

46. 如权利要求 1 所述的方法, 其特征在于, 所述患者用索拉非尼进行治疗。

47. 一种监测接受下述治疗的实体瘤患者的响应的方法, 所述治疗采用下式 I 的化合物 N-[4-氯-3-(三氟甲基)苯基]-N'-{4-[2-氨甲酰基-1-氧代-(4-吡啶基氧)]苯基}脲或其药学上可接受的盐、多晶型、水合物、溶剂合物、或其组合



所述方法包括：

a)测定从患者获取的生物样品中 VEGF 和/或 sVEGFR-2 的表达水平；

b)测定开始用式 I 的化合物或其药学上可接受的盐、多晶型、水合物、溶剂合物或其组合治疗后从患者获取的至少第二生物样品中 VEGF 和/或 sVEGFR-2 的表达水平；和

c)将第二样品中 VEGF 和/或 sVEGFR-2 的表达水平与第一样品中 VEGF 和/或 sVEGFR-2 的表达水平进行比较；其中，与第一样品中 VEGF 和/或 sVEGFR-2 的表达水平相比，第二样品中 VEGF 和/或 sVEGFR-2 的表达水平的改变表明用式 I 的化合物或其药学上可接受的盐、多晶型、水合物、溶剂合物或其组合治疗的疗效。

48. 如权利要求 47 所述的方法，其特征在于，所述生物样品是血液、羊水、血浆、血清、精液、尿液、骨髓或活检组织。

49. 如权利要求 47 所述的方法，其特征在于，所述生物样品是血浆。

50. 如权利要求 47 所述的方法，其特征在于，所述癌症是肾细胞癌。

癌症预测与预后以及监测癌症治疗的方法

相关申请的交叉引用

本申请要求 2005 年 11 月 14 日提交的美国临时专利申请第 60/735,854 号的优先权，该申请的内容以参考的方式结合在本说明书中。

发明领域

本发明涉及生物标记物及其在癌症预测与预后中的应用，以及使用生物标记物来监测癌症治疗效果的应用。具体地说，本发明涉及使用可溶性 VEGF(“VEGF”)和可溶性 VEGF 受体(sVEGFR)作为生物标记物在索拉非尼(sorafenib)治疗效果中的应用。

发明背景

许多疾病状态表现为各种基因表达水平的差异，这种差异通过改变基因组 DNA 拷贝数或者通过改变特定基因的转录水平(例如通过控制起始、RNA 前体的形成、RNA 加工等)来实现。例如，遗传物质的丧失与获取在恶性转化和进程中起着重要的作用。认为这些获取与丧失是由至少两类基因，即癌基因和肿瘤抑制基因所驱动的。癌基因是肿瘤发生的正调节子，而肿瘤抑制基因是肿瘤发生的负调节子(Marshall, Cell 64:313-326, 1991; Weinberg, Science 254:1138-1146, 1991)。因此，激活不受调节的生长的一种机制是增加编码癌基因蛋白质的基因数量或增加癌基因的表达水平(如，响应细胞内或环境的改变)，另一种机制是丧失遗传物质或降低编码肿瘤抑制子的基因的表达水平。该模型得到胶质瘤进程相关遗传物质的丧失与获取的支持 (Mikkelsen, 等, J. Cellular Biochem. 46:3-8, 1991)。因此，特定基因(如癌基因或肿瘤抑制基因)表达(转录)水平的改变适用于标示各种癌症的存在与进程。

发明概述

本发明涉及生物标记物及其在癌症预测与预后中的应用，以及使用生物标记

物来监测癌症治疗疗效的应用。具体地说,本发明涉及采用 VEGF 和 sVEGFR,更优选 sVEGFR-2 (可溶性 VEGFR-2)作为监测索拉非尼治疗疗效的生物标记物。如下面的实施例中更详细所述,发现在用索拉非尼治疗的对象中 sVEGFR-2 降低,而 VEGF 水平升高。因此,这些标记物可用于确定索拉非尼的治疗疗效。

此外,本发明的目的是提供用于癌症预测、诊断、预后及治疗的方法和试剂。

本发明的另一个实施方式是筛选药物对组织或细胞样品的作用的方法,该方法包括以下步骤:分析一种或多种基因或基因产物的表达水平,其中,分析接触药物之前和之后组织或细胞样品中所述基因和/或基因产物的表达水平,基因和/或基因产物表达水平的变化表明药物作用或提供患者诊断或预测患者对治疗的响应。在又一个实施方式中,药物是索拉非尼。在另一实施方式中,基因或基因产物是 VEGF 和 VEGFR,更优选 VEGFR-2 以及它们的可溶形式(如,检测流出的 VEGFR2)。

本发明的另一方面是发现新型药物的方法,该方法包括以下步骤:分析一种或多种基因和/或基因产物的表达水平,其中,分析接触药物之前和之后细胞中基因和/或基因产物的表达水平,基因和/或基因产物表达水平的变化表明药物疗效。又一方面,所述基因或基因产物是 VEGF 和 VEGFR,更优选 VEGFR-2,以及它们的可溶形式(如,检测流出的 VEGFR2)。

本发明还提供了鉴别适用于治疗癌症的化合物的方法,该方法包括:给予癌症患者试验化合物,并测定多肽活性,其中,所述多肽活性的变化表明试验化合物适用于治疗癌症。在又一个实施方式中,多肽是 VEGF 和 VEGFR,更优选 VEGFR-2,以及它们的可溶形式(如,检测流出的 VEGFR2),在另一实施方式中,所述化合物是索拉非尼。

因此,本发明提供了适用于鉴别可能以调节子或调节蛋白如激动剂或拮抗剂、部分激动剂、反相激动剂、激活剂、辅激活剂和抑制剂的形式起作用的化合物。因此,本发明提供了用于调节癌症相关多核苷酸或多肽表达的试剂和方法。调节多核苷酸的表达、稳定性或量以及多肽活性的试剂可以是蛋白质、肽、肽模拟物、核酸、核酸类似物(如,肽核酸,闭锁核酸)、或小分子。

本发明还提供了实施患者诊断的方法,该方法包括以下步骤:分析一种或多种基因和/或基因产物的表达水平,其中,分析正常对象和患者样品中的基因和/或基因产物表达水平,患者样品中基因和/或基因产物表达水平的变化能用于疾

病的诊断。患者样品包括但不限于：血液、羊水、血浆、精液、骨髓和组织活检液。在又一个实施方式中，基因或基因产物是 VEGF 和 VEGFR，更优选 VEGFR-2，以及它们的可溶形式(如，检测流出的 VEGFR2)。

本发明还提供了在对象中诊断癌症的方法，该方法包括：测定疑似患有癌症的对象中多肽的活性，如果该多肽活性相对于不患癌症的对象中的多肽活性存在差异，则该对象诊断为患有癌症。在又一个实施方式中，多肽是 VEGF 和 VEGFR，更优选 VEGFR-2，以及它们的可溶形式(如，检测流出的 VEGFR2)。

在另一实施方式中，本发明提供了在患者样品中检测癌症的方法，其中，使用蛋白质抗体来与患者样品中的蛋白质反应。在又一个实施方式中，抗体是特异性针对 VEGF 和 VEGFR，更优选 VEGFR-2，以及它们的可溶形式(如，检测流出的 VEGFR2)的。抗体可以是常规产生的多肽暴露区的抗体。例如，抗体可以是常规产生的 VEGFR-2(如可溶性 VEGFR-2)的胞外结构域的抗体。

本发明的另一方面是辨别正常和疾病状态的方法，该方法包括以下步骤：分析一种或多种基因和/或基因产物的表达水平，其中，分析正常和疾病组织的基因和/或基因产物的表达水平，基因和/或基因产物表达水平的变异表明疾病状态。在又一方面，基因或基因产物是 VEGF 或 VEGFR-2。

在另一实施方式中，本发明涉及鉴定细胞表型的方法，该方法包括检测至少一种基因相对于正常细胞的差异表达，其中，所述差异表达的因数至少是 2，因数至少是 5，因数至少是 20，或者因数至少是 50。在又一个实施方式中，所述基因编码 VEGF 和 VEGFR，更优选 VEGFR-2。

在又一实施方式中，本发明涉及检测细胞表型的方法，该方法包括检测至少一种多肽相对于正常细胞的差异表达，其中，所述蛋白质差异表达的因数至少是 2，因数至少是 5，因数至少是 20，一直到因数至少是 50。在又一个实施方式中，所述多肽是 VEGF 和 VEGFR，更优选 VEGFR-2，以及它们的可溶形式(如，检测流出的 VEGFR2)。

在另一实施方式中，本发明涉及检测患者细胞表型的方法，该方法包括提供由具有至少约 10 个、至少约 15 个、至少约 25 个、或至少约 40 个连续核苷酸的核苷酸序列构成的核酸探针，从患者获取细胞样品，任选地提供基本上非癌的第二细胞样品，在严谨条件下使核酸探针分别与所述第一和第二细胞样品的 mRNA 相接触，以及将(a)探针与第一细胞样品的 mRNA 杂交的量，与(b)探针与第二细

胞样品的 mRNA 杂交的量进行比较, 其中, 与第一细胞样品的 mRNA 杂交的量和与第二细胞样品的 mRNA 杂交的量之间的差异因数至少是 2, 因数至少是 5, 因数至少是 20, 或者因数至少是 50, 则表明第一细胞样品中的细胞表型。在又一个实施方式中, 核酸探针包括编码 VEGF 和/或 VEGFR, 优选 VEGFR-2 的核苷酸序列。

在另一实施方式中, 本发明提供了用于鉴别癌细胞或组织的存在的检验试剂盒, 其包括: 探针/引物, 用于测量从患者分离的细胞样品中的核酸水平。在某些实施方式中, 试剂盒还可包括: 试剂盒的使用说明, 用于混悬或固定细胞的溶液, 可检测的标签或标记, 使核酸易于杂交的溶液, 用于裂解细胞的溶液, 或用于核酸纯化的溶液。在又一个实施方式中, 所述探针/引物包括编码 VEGF 和/或 VEGFR, 优选 VEGFR-2 的片段的核苷酸序列。

在一个实施方式中, 本发明提供了用于鉴别癌细胞或组织的存在的检验试剂盒, 其包括蛋白质特异性抗体。在某些实施方式中, 试剂盒还包括试剂盒的使用说明。在某些实施方式中, 试剂盒还包括用于混悬或固定细胞的溶液, 可检测的标签或标记, 使多肽易于结合抗体的溶液, 用于裂解细胞的溶液, 或用于多肽纯化的溶液。在又一个实施方式中, 所述抗体是特异性针对 VEGF 和/或 sVEGFR, 优选针对 sVEGFR-2。

在另一实施方式中, 本发明提供了用于监测化合物或治疗剂在癌细胞或组织中的疗效的检验试剂盒, 其包括探针/引物, 用于测量从患者分离的细胞样品中的核酸水平。在某些实施方式中, 试剂盒还包括试剂盒的使用说明, 用于混悬或固定细胞的溶液, 可检测的标签或标记, 使核酸易于杂交的溶液, 用于裂解细胞的溶液, 或用于核酸纯化的溶液。在又一个实施方式中, 所述探针/引物包括编码 VEGF 和/或 VEGFR 的核苷酸序列。

在一个实施方式中, 本发明提供了用于化合物或治疗剂在监测癌细胞或组织中的疗效的检验试剂盒, 其包括蛋白质特异性抗体。在某些实施方式中, 试剂盒还包括试剂盒的使用说明。在某些实施方式中, 试剂盒还包括: 用于混悬或固定细胞的溶液, 可检测的标签或标记, 使多肽易于结合抗体的溶液, 用于裂解细胞的溶液, 或用于多肽纯化的溶液。在又一个实施方式中, 所述抗体是特异性针对 VEGF 和/或 VEGFR-2, 如针对其可溶性胞外结构域。

发明详述

应理解，本发明并不限于所述的具体方法、方案、细胞系、动物物种或属，构建物和反应试剂，它们可以变化。还应理解，本文所述术语仅仅是为了描述具体实施方式的目的，而不是限制本发明的范围，本发明的范围仅由所附权利要求书限定。

应理解，除非另外清楚指出，本文和所附权利要求书中所用的单数形式“一”、“一个”和“这个”包括复数涵义。因此，例如，“一个基因”是指一个或多个基因，包括本领域技术人员已知的等价形式，等等。

除非另有说明，本文所用的所有科技术语具有与本发明所属领域普通技术人员通常所理解的相同的涵义。虽然本发明的实施或检测中可使用类似或等价于本文所述的任何方法、装置和材料，但现在将描述优选的方法、装置和材料。

本文所述的所有出版物和专利都以参考的方式结合于此，用于描述和阐述的目的，例如可与本发明联用的出版物中所述的构建物和方法学。提供上述及整篇内容中的出版物仅仅是因为它们公开在本发明的提交日之前。都不应解释为承认，本发明没有权利优先于这些在先发明的公开。

定义

为了方便，下面提供说明书、实施例和所附权利要求书中采用的一些术语和短语。

阵列(如微阵列)上的“地址”是指诸如寡核苷酸等元件附连于阵列固相表面的位置。

本文所用术语“激动剂”是指模拟或上调(如，赋予或补充)蛋白质生物活性的试剂。激动剂可以是野生型蛋白质或其具有至少一种野生型蛋白质的生物活性的衍生物。激动剂也可以是上调基因表达或提高至少一种蛋白质生物活性的化合物。激动剂也可以是提高多肽与另一种分子如靶肽或核酸相互作用的化合物。

本文所用术语“扩增”涉及产生额外的核酸序列拷贝。例如，扩增可采用本领域公知的聚合酶链反应 (PCR) 技术来实现(例如参见，Dieffenbach 和 Dveksler (1995) PCR Primer, 《实验室手册》(A Laboratory Manual), Cold Spring Harbor Press, Plainview, N.Y.)

本文所用“拮抗剂”是指下调(如限制或抑制)至少一种蛋白质生物活性的试

剂。例如，索拉非尼是拮抗剂的一个例子。拮抗剂可以是抑制或降低蛋白质与其它分子如靶肽或酶底物间相互作用的化合物。拮抗剂也可以是下调基因表达或减少表达的蛋白质含量的化合物。

本文所用术语“抗体”指包括任意同种型的完整抗体(IgG, IgA, IgM, IgE等), 包括其也能够与脊椎动物(如哺乳动物)蛋白质特异性反应的片段。抗体可采用常规技术片段化, 筛选的片段以上述完整抗体相同的方式使用。因此, 该术语包括抗体分子的蛋白水解裂解的片段或重组制备部分, 其能够选择性地与某些蛋白质反应。这种蛋白水解片段和/或重组片段的非限制性例子包括: Fab、F(ab')₂、Fab'、Fv 以及包含肽接头连接的 V[L]和/或 V[H]域的单链抗体(scFv)。scFv's 可共价或非共价连接形成具有两个或多个结合位点的抗体。本发明包括多克隆抗体、单克隆抗体、或其他抗体的纯化制剂和重组抗体。

术语“阵列”或“矩阵”是指装置上可寻址位置或“地址”的阵列。位置可以二维阵列、三维阵列或其它矩阵模式排列。位置的数量可以从几千到至少数十万。更重要地, 每个位置代表完全独立的反应位点。“核酸阵列”是指含有核酸探针如寡核苷酸或较大部分的基因的阵列。阵列上的核酸优选为单链。探针为寡核苷酸的阵列称为“寡核苷酸阵列”或“寡核苷酸芯片”。“微阵列”在这里也称为“生物芯片”, 是离散区域密度至少约 100/cm², 优选至少约 1000/cm²的区域阵列。微阵列区域具有典型尺寸, 例如直径约为 10-250 μm 且与阵列中其它区域隔开约相同的距离。

“生物学活性”或“生物活性”或“活性”或“生物功能”可互换使用, 在这里是指, 多肽(无论是天然构象或是变性构象)或其部分序列直接或间接执行的效应器或抗原功能。生物活性包括结合多肽的活性, 结合其它蛋白质或分子的活性, 作为 DNA 结合蛋白的活性, 作为转录调节子的活性, 结合受损 DNA 的活性等。生物活性可通过直接影响对象多肽进行调节。或者, 生物活性剂可通过调节多肽水平, 如调节相应基因的表达来改变。

本文所用术语“生物样品”是指从生物体或从生物体元件(如细胞)获取的样品。样品可以是任何生物组织或流体。样品可以是取自患者的样品。这种样品包括但不限于: 痰液、血液、血细胞(如白细胞)、组织或活检样品(如, 肿瘤活检样品)、尿液、腹膜液和胸膜液、或由它们获取的细胞。生物样品还可包括组织切片, 如为组织学目的获取的冰冻切片。

术语“生物标记物”或“标记物”包括细胞内和细胞外事件的广泛范围以及全生物体生理学改变。生物标记物本质上可以代表细胞功能的任意方面，例如但不限于：信号转导分子、转录因子、代谢物、基因转录物的产生水平或速率以及蛋白质翻译后修饰。生物标记物可包括转录物水平的全基因组分析、或蛋白质水平和/或修饰的全蛋白质组分析。

生物标记物也可指与未处理患病细胞相比，化合物处理的患者患病细胞中上调或下调了的基因或基因产物。即，基因或基因产物对处理细胞具有足够的特异性，可任选地与其它基因或基因产物联用来鉴别、预测或检测小分子的疗效。因此，生物标记物是表征化合物在患病细胞中的疗效或该患病细胞对化合物处理的反应的基因或基因产物。

如果两个序列的碱基各自匹配，即能够形成沃森-克里克(Watson-Crick)碱基对，则该核苷酸序列与另一核苷酸序列“互补”。本文所用术语“互补链”可与术语“互补序列(complement)”互换使用。核酸链的互补序列可以是编码链的互补序列或非编码链的互补序列。

“基因检测试剂”是指可用于特异性检测基因或有关的其它生物分子的试剂，所述生物分子如由该基因转录的 RNA 或该基因编码的多肽。检测试剂的例子是能与基因对应的核酸进行杂交的核酸探针、以及抗体。

术语“癌症”包括但不限于：实体瘤，例如乳房、呼吸道、脑、生殖器官、消化道、尿道、眼、肝、皮肤、头和颈、甲状腺、甲状旁腺的癌症以及它们的远端转移癌症。该术语还包括淋巴瘤、肉瘤和白血病。

乳腺癌的例子包括但不限于：浸润性导管癌、浸润性小叶癌、原位管癌和原位小叶癌。

呼吸道癌症的例子包括但不限于：小细胞和非小细胞肺癌，以及支气管腺瘤和胸膜肺母细胞瘤。

脑癌的例子包括但不限于：脑干和眼下(hypophthalmic)胶质瘤、小脑和大脑星形细胞瘤、髓母细胞瘤、室管膜瘤、以及神经外胚层和松果体瘤。

雄性生殖器官肿瘤包括但不限于：前列腺和睾丸癌。雌性生殖器官肿瘤包括但不限于：子宫内膜、子宫颈、卵巢、阴道和外阴癌症、以及子宫肉瘤。

消化道肿瘤包括但不限于：肛门、结肠、结肠直肠、食管、胆囊、胃、胰脏、直肠、小肠或唾液腺癌。

尿道肿瘤包括但不限于：膀胱、阴茎、肾、肾脏(如、肾细胞癌)、骨盆、输尿管和尿道癌。

眼部癌症包括但不限于：眼内黑色素瘤或视网膜母细胞瘤。

肝癌的例子包括但不限于：肝细胞癌(伴有或没有纤维层样变异的肝细胞癌)、胆管癌(肝内胆管癌)以及混合的肝细胞胆管癌。

皮肤癌包括但不限于：鳞状细胞癌、卡波济氏肉瘤、恶性黑色素瘤、Merkel 细胞皮肤癌和非黑色素瘤皮肤癌。

头颈部癌症包括但不限于：喉/下咽/鼻咽/口咽部癌症以及唇部和口腔癌。

淋巴瘤包括但不限于：AIDS-相关淋巴瘤、非霍奇金淋巴瘤、表皮 T-细胞淋巴瘤、霍奇金病和中枢神经系统淋巴瘤。

肉瘤包括但不限于：软组织肉瘤、骨肉瘤、恶性纤维组织细胞瘤、淋巴肉瘤和横纹肌肉瘤。

白血病包括但不限于：急性髓细胞性白血病、急性淋巴细胞性白血病、慢性淋巴细胞性白血病、慢性髓细胞性白血病或多毛细胞白血病。

“癌症患病细胞”是指癌症患者中存在的细胞。就是说，正常细胞的修饰形式且没有癌症的对象中不存在的细胞，或是癌症对象相对于非癌对象的存在数量显著较高或较低的细胞。

术语“等价”应理解为包括编码功能等价多肽的核苷酸序列。等价核苷酸序列可包括取代、增加、或缺失一个或多个核苷酸而有所不同的序列，如等位变体。

术语“表达谱”可与细胞的“基因表达谱”和“指纹图谱”互换使用，是指代表细胞中一个或多个基因的 mRNA 水平的一组值。表达谱优选包括代表至少约 10 个基因的表达水平的值，优选至少约 50 个，100 个，200 个或更多个基因。表达谱还可包括在多种细胞和条件下以类似水平表达的基因的 mRNA 水平(例如，管家基因如 GAPDH)。例如，癌症患病细胞的表达谱是指代表患病细胞中 10 个或更多个基因的 mRNA 水平的一组值。

术语“基因”是指包括产生多肽或前体所需的控制和编码序列的核酸序列。多肽可由全长编码序列或由任意部分的编码序列来编码。基因可全长或部分地来自本领域已知的任何来源，包括植物、真菌、动物、细菌基因组或附加体、真核、核或质粒 DNA、cDNA、病毒 DNA、或化学合成的 DNA。基因的编码或非翻译区中可包含一个或多个修饰，从而影响表达产物的生物学活性或化学结构、表达

速率、或表达控制的方式。这种修饰包括但不限于：突变，插入，缺失和取代一个或多个核苷酸。基因可构成不间断的编码序列或者它可包含一个或多个由合适的剪切点连接的内含子。

“杂交”是指核酸链通过碱基配对与互补链结合的任何过程。例如，两条单链核酸形成双链体时发生“杂交”。双链区域可包括两条单链核酸之一或两者的全长，或者一条单链核酸的全部和另一条单链核酸的部分序列，或者双链区域可包括两条核酸的部分序列。杂交还包括形成含有某些错配的双链体，只要两条链仍然形成双螺旋。“严谨杂交条件”是指导致基本上特异性杂交的杂交条件。

本文所用术语“分离的(isolated)”对于核酸如 DNA 或 RNA 是指，分别从大分子天然来源中存在的其它 DNA 或 RNA 分离的分子。本文所用术语“分离的”也指基本上不含细胞材料、病毒材料、培养基(通过重组 DNA 技术制备时)、或化学前体或其他化学物质(化学合成时)的核酸或肽。而且，“分离的核酸”可包括天然不以片段形式出现和自然状态下不存在的核酸片段。本文所用术语“分离的”也指从其它蛋白质分离的多肽，旨在涵盖纯化和重组的多肽。

本文所用术语“标记”和“可检测的标记”指能够检测的分子，包括但不限于：放射性同位素、荧光团、化学发光部分、酶、酶底物、酶辅因子、酶抑制剂、染料、金属离子、配体(如，生物素或半抗原)等。术语“荧光剂”是指在可检测的范围内能显示荧光的物质或其一部分。适用于本发明的标记的具体例子包括：荧光素、罗丹明、丹酰、伞形酮、德克萨斯红、发光氨、NADPH、 α - β -半乳糖苷酶和辣根过氧化物酶。

本文所用术语“表达水平”是指给定核酸可测量的表达水平。核酸的表达水平可通过本领域公知的方法来确定。术语“差异性表达的”或“差异表达”是指给定核酸可测量的表达水平的升高或降低。如本文所用“差异性表达的”或“差异表达”表示相对于同一正常标准样品而言，用于比较的两个样品中核酸表达水平的差异至少是 1.4-倍或更高。本发明“差异性表达的”或“差异表达”也指用于比较的两个样品中核酸表达水平的差异为 1.4-倍或更高，一直到并包括 2-倍、5-倍、10-倍、20-倍、50-倍或更高。如果两个样品之一不包含给定核酸可检测的表达，则认为该核酸在两种样品中“差异性表达”，只要可检测的表达的核酸以 $\pm$ 至少 1.4 倍的形式表达。用于比较两个或多个样品中核酸表达水平的差异发生改变而导致差异不再是至少 1.4 倍，则核酸序列的差异表达受到“抑制”。核酸

表达水平的绝对值可通过以下过程得到：包含已知浓度的一种或多种对照核酸物质，基于对照核酸的量产生标准曲线，以及从未知核酸的杂交强度相对于标准曲线来推断未知核酸物质的表达水平。

本文所用术语“核酸”是指多核苷酸如脱氧核糖核酸(DNA)，适当时也指核糖核酸(RNA)。该术语还应理解为包括从核苷酸类似物制备的 RNA 或 DNA 的等价物、类似物，正如可应用于本文所述的实施方式，包括单链(有义(sense)或反义)以及双链多核苷酸。染色体、cDNAs、mRNAs、rRNAs 和 ESTs 是可称为核酸的分子的代表性例子。

本文所用术语“寡核苷酸”是指包含例如约 10-1000 个核苷酸的核酸分子。适用于本发明的寡核苷酸优选为约 15-150 个核苷酸，更优选约 150-1000 个核苷酸长度。寡核苷酸可以是天然产生的寡核苷酸或合成寡核苷酸。寡核苷酸可通过磷酸胺酸法制备(Beaucage 和 Carruthers, Tetrahedron Lett. 22:1859-62, 1981)，或通过三酯法制备(Matteucci, 等, J. Am. Chem. Soc. 103:3185, 1981)，或通过本领域已知的其它化学方法制备。

本文所用术语“患者”或“对象”包括哺乳动物(如，人和动物)。

如本文所用，附连于阵列的核酸或其它分子称为“探针”或“捕获探针”。当阵列包含多个对应于一个基因的探针时，这些探针称为“基因-探针系列”。基因-探针系列可包括例如约 2-20 个探针，优选约 2-10 个探针，最优选约 5 个探针。

细胞生物学状态的“谱”是指已知响应药物治疗或其它细胞生物学状态干扰发生改变的各种细胞成分的水平。细胞成分包括例如：RNA 水平、蛋白质丰度水平或蛋白质活性水平。

术语“蛋白质”在这里可与“肽”和“多肽”互换使用。

如果两种谱中的基因表达水平足够类似，以至于相似性表明共同特征，例如相同类型的细胞，则一个细胞中的表达谱“类似于”另一细胞中的表达谱。因此，如果第一个细胞中表达的基因至少 75%，在第二个细胞中以相对于第一个细胞的因数小于等于 2 的水平表达，则第一个细胞和第二个细胞的表达谱是类似的。

本文所用“小分子”是指分子量小于约 5 kD，最优选小于约 4 kD 的组分。小分子可以是核酸、肽、多肽、肽模拟物、糖、脂质或其他有机或无机分子。许多制药公司具拥有大量化学和/或生物混合物的库，常常是真菌、细菌或藻类

提取物,可用任何本发明的的分析方法进行筛选来鉴别能够调节生物活性的化合物。

术语探针与模板核酸靶位点的“特异性杂交”是指,大部分靶向该位点的探针的杂交,从而清楚解释杂交信号。如本文进一步所述,引起特异性杂交的条件根据同源性区域的长度、区域 GC 含量以及杂交体的解链温度(“ T_m ”)而变化。因此,杂交条件中杂交液和洗涤液的盐含量、酸度和温度可变化。

多肽“变体”是指氨基酸序列中一个或多个氨基酸残基发生改变的多肽。变体可具有“保守性”变化,其中,取代的氨基酸具有类似的结构或化学性质(如,亮氨酸与异亮氨酸的置换)。变体还具有“非保守性”变化(如,甘氨酸与色氨酸的置换)。类似的微小变异可包括氨基酸缺失或插入,或同时缺失和插入。确定取代、插入或缺失哪个氨基酸残基而不破坏生物学活性或免疫活性的指导可采用本领域公知的计算机程序进行鉴别,例如 LASER 基因软件(DNASTAR)。

当术语“变体”用于多核苷酸序列的内容中时,可包括与特定基因或其编码序列有关的多核苷酸序列。该定义还可包括例如“等位变体”、“剪接变体”、“物种变体”和“多态性变体”。剪接变体与参比分子具有显著的同一性,但通常由于 mRNA 加工期间外显子的选择性剪接而具有较多或较少数量的多核苷酸。对应的多肽可具有额外的功能区域或缺少功能区。物种变体是种与种之间不尽相同的多核苷酸序列。所得多肽彼此间通常具有显著的氨基酸同一性。多态性变体是指给定物种的个体之间特定基因的多核苷酸序列的变体。多态性变体还可包括多核苷酸序列中一个碱基发生改变的“单核苷酸多态性”(SNP)。SNP 的存在表明某一群体、疾病状态或疾病状态的倾向性。

本发明的一方面涉及鉴别能够调节具有异常增殖特征的细胞的分化与增殖的试剂。更具体说,本发明涉及筛选能够调节核酸序列差异表达的候选化合物的方法。就是说,如果一核酸序列在癌细胞中过度表达,则筛选能够降低表达的候选化合物,如果一核酸序列在癌细胞中表达不足,则筛选能够增加表达的试验化合物。此外,本发明涉及鉴别能够调节由本文所述的差异性表达序列所编码的一种或多种多肽的活性的试验化合物或物质的筛选分析方法。关于这一点,本发明提供了用于测定能够调节标记物核酸的表达和/或改变所编码多肽的生物活性的化合物的分析方法。

筛选差异表达的调节

通过将试验化合物(例如索拉非尼及其二芳基脲衍生物)加入到细胞样品中并监测其作用来进行药物筛选。同时监测不接受试验化合物的平行样品作为对照。然后通过任何合适的表型标准对处理和未处理的细胞进行比较,这些标准包括但不限于:显微镜分析、存活率试验、复制能力、组织学检查、细胞相关特定RNA或多肽的水平、细胞或细胞裂解物表达的酶活性水平、以及细胞与其他细胞或化合物相互作用的能力。处理和未处理细胞间的差异表明该试验化合物的作用。

所需试验化合物的作用包括对癌症相关标记物核酸序列的任意表型的作用。例子包括限制 mRNA 过量、限制编码蛋白质的产生、或限制蛋白质功能作用的试验化合物。对处理和未处理细胞间的结果进行比较之后,试验化合物的作用将显而易见。

因此,本发明还包括筛选能够抑制或提高体外核酸标记物的表达的试剂(如,索拉非尼和及其二芳基脲衍生物)的方法,该方法包括使细胞或组织与试剂相接触以确定试剂是否能够抑制或提高 mRNA 的产生,所述细胞中标记物核酸 mRNA(如 VEGF 或 VEGFR-2)在培养的细胞中是可检测的;以及测定接触后的细胞或组织中 mRNA 的水平,其中,细胞接触试剂后 mRNA 水平降低表明标记物核酸 mRNA 的产生受到抑制,而 mRNA 水平升高则表明标记 mRNA 的产生增加。

或者,筛选方法可包括体外筛选疑似能够抑制或提高标记物蛋白质的产生的试剂的方法,该方法包括使细胞或组织与疑似能够抑制或提高标记物蛋白质的产生的试剂相接触,所述细胞或组织中标记物蛋白质在培养的细胞中是可检测的;以及测定细胞或组织中标记物蛋白质的水平,其中,细胞或组织接触试剂后标记物蛋白质的水平降低表明标记物蛋白质的产生受到抑制,而蛋白质水平升高则表明标记物蛋白质的产生增加。

本发明还包括能够抑制或提高标记物核酸的表达的试剂的体内筛选方法,该方法包括:使存在肿瘤细胞的对象与疑似能够抑制或提高标记物 mRNA 或蛋白质的产生的试剂相接触,所述对象的肿瘤细胞中标记物 mRNA 或蛋白质是可检测的;以及测定接触后哺乳动物肿瘤细胞中标记物 mRNA 或蛋白质的水平。对象接触试剂后标记物 mRNA 或蛋白质的水平降低表明标记物核酸的表达受到抑制,而标记物 mRNA 或蛋白质水平升高则表明标记物核酸的表达增加。

因此,本发明提供了一种方法,该方法包括:将表达标记物核酸的细胞与试验化合物一起培养,并测定 mRNA 或蛋白质水平。本发明还提供了定量测定细胞群中标记物核酸的表达水平的方法,以及测定试剂是否能够增加或降低细胞群中标记物核酸的表达水平的方法。测定试剂是否能够增加或降低细胞群中标记物核酸的表达水平的方法包括以下步骤:(a)从对照和试剂处理的细胞群获取细胞提取物,(b)将标记物多肽从细胞提取物中分离出来,和(c)测定标记物多肽与所述多肽的特异性抗体间形成的免疫复合物的量(如平行测定)。本发明标记物多肽也可通过测定其生物活性来定量。也可通过与对照细胞中形成的免疫复合物的量相比,处理细胞中形成的免疫复合物的量增加来鉴别能够诱导标记物核酸表达增加的试剂。以类似的方式,通过与对照细胞中相比,处理细胞提取物中形成的免疫复合物的量降低,来鉴别能够降低标记物核酸表达的试剂。

本发明提供了在癌症中差异性调节的分离核酸序列以及鉴别该序列的方法。本发明提供了鉴别在癌症患者中差异性调节的核苷酸序列的方法,该方法包括:使对应于从患者获取的 RNA 的核酸样品与包含一种或多种已知身份核酸分子杂交;以及测量核酸样品与一种或多种已知身份核酸分子的杂交,其中,核酸样品与一种或多种已知身份核酸分子杂交相对于从没有癌症的对象获取的核酸样品存在两倍差异表明,在癌症患者中该核苷酸序列差异表达。

通常,本发明提供了鉴别在癌症患者中差异调节的核酸序列的方法,该方法包括从对象分离信使 RNA,从该 mRNA 样品产生 cRNA,使 cRNA 与包含许多稳定地结合在阵列离散位置上的核酸分子的微阵列杂交,以及鉴别 cRNA 与阵列的杂交模式。根据本发明,如果与没有患癌症的对象获取的核酸样品相比,同一阵列相同位置上杂交信号至少高两倍或低两倍,则认为与阵列上给定位置杂交的核酸分子为差异性调节。

测定基因表达水平的微阵列

测定基因表达水平可采用微阵列来实现。通常,包括以下步骤:(a)从对象获取 mRNA 样品并由其制备标记的核酸(“靶核酸”或“靶点”);(b)在足以使靶核酸与阵列上的对应探针发生结合(例如杂交或特异性结合)的条件下,使靶核酸与阵列相接触;(c)任选地去除阵列上未结合的靶核酸;(d)检测结合靶点;以及(e)采用基于计算机的分析方法分析结果。如本文所用,“核酸探针”或“探针”指

附连在阵列上的核酸，而“靶核酸”指与阵列杂交的核酸。

可采用“侵入性”或“非侵入性”取样方式从待测对象获取核酸样本。如果取样方法涉及从动物(包括鼠、人、羊、马、牛、猪、犬、或猫)皮肤或器官内收集核酸，则是侵入性的。侵入性方法的例子包括例如：收集血液、收集精液、穿刺活检、胸膜抽吸、脐带活检。这种方法的例子在 Kim 等(J. Virol. 66:3879-3882, 1992); Biswas 等(Ann. NY Acad. Sci. 590:582-583, 1990); 和 Biswas 等(J. Clin. Microbiol. 29:2228-2233, 1991)中描述。

相反，“非侵入性”取样方法指从动物内表面或外表面回收核酸分子的方法。这种“非侵入性”取样方法的例子包括例如，“擦抹”、收集泪液、唾液、尿液、粪便、汗或排汗、头发。

在本发明的一个实施方式中，从待测对象获取一种或多种细胞并分离细胞的 RNA。在优选实施方式中，从对象获取外周血白细胞(PBL)样品。还可从对象获取细胞样品然后富集所需的细胞类型。例如，可采用许多技术从其他细胞分离细胞，例如采用能够与所需细胞类型的细胞表面上的表位相结合的抗体。如果所需细胞位于实体组织中，则可通过显微切割或通过激光捕获显微切割(LCM)获取特定细胞(参见例如，Bonner 等，Science 278:1481, 1997; Emmert-Buck 等，Science 274:998, 1996; Fend 等，Am. J. Path. 154:61, 1999; 和 Murakami 等，Kidney Int. 58:1346, 2000)。

RNA 可通过许多方法从组织或细胞样品提取，例如，硫氰酸胍(guanidium thiocyanate)裂解细胞然后 CsCl 离心(Chirgwin 等，Biochemistry 18:5294-5299, 1979)。单细胞 RNA 根据从单细胞制备 cDNA 库的方法中所述来获取(参见例如，Dulac, Curr. Top. Dev. Biol. 36:245, 1998; Jena 等，J. Immunol. Methods 190:199, 1996)。

可进一步富集 RNA 样品中的特定物种。在一个实施方式中，例如，可从 RNA 样品中分离聚(A)+RNA。在另一实施方式中，可通过引物特异性 cDNA 合成，或基于 cDNA 合成的多轮线性扩增以及模板介导的体外转录来富集 RNA 群中的感兴趣序列(参见例如，Wang 等，Proc. Natl. Acad. Sci. USA 86:9717, 1989; Dulac 等，同上; Jena 等，同上)。此外，富集或没有富集特定物种或序列的 RNA 可通过许多扩增方法进一步扩增，这些方法包括例如，PCR; 连接酶链反应(LCR)(参见例如，Wu 和 Wallace, Genomics 4:560, 1989; Landegren 等，Science 241:1077,

1988); 自主序列复制(SSR) (参见例如, Guatelli 等, Proc. Natl. Acad. Sci. USA 87:1874, 1990); 基于核酸的序列扩增(NASBA)和转录扩增(参见例如, Kwok 等, Proc. Natl. Acad. Sci. USA 86:1173, 1989)。PCR 技术方法是本领域众所周知的(参见例如, 《PCR 技术: DNA 扩增的原理和应用》(PCR Technology: Principles and Applications for DNA Amplification) (H. A. Erlich 编, Freeman Press, N.Y., N.Y., 1992); 《PCR 方法: 方法和应用指南》(PCR Protocols: A Guide to Methods and Applications) (Innis 等编, Academic Press, San Diego, Calif., 1990); Mattila 等, Nucleic Acids Res. 19:4967, 1991; Eckert 等, PCR Methods and Applications 1:17, 1991; PCR (McPherson 等编, IRL Press, Oxford); 和美国专利第 4,683,202 号)。扩增方法参见例如, Ohyama 等(BioTechniques 29:530, 2000); Luo 等(Nat. Med. 5:117, 1999); Hegde 等(BioTechniques 29:548, 2000); Kacharina 等(Meth. Enzymol. 303:3, 1999); Livesey 等, Curr. Biol. 10:301, 2000); Spirin 等(Invest. Ophthalmol. Vis. Sci. 40:3108, 1999); 和 Sakai 等(Anal. Biochem. 287:32, 2000)。RNA 扩增和 cDNA 合成也可在细胞中原位进行(参见例如, Eberwine 等, Proc. Natl. Acad. Sci. USA 89:3010, 1992)。

可对核酸分子进行标记从而检测核酸分子与微阵列的杂交。就是说, 探针可包含一信号产生系统的成员, 因而可直接或通过与信号产生系统中的一种或多种其他成员联合作用来进行检测。例如, 可用荧光标记的 dNTP (参见例如, Kricka, 1992, 《非同位素 DNA 探针技术》Nonisotopic DNA Probe Techniques, Academic Press San Diego, Calif.)、生物素化 dNTPs 或 rNTP 再加入标记的亲合素、化学发光标记、或同位素来标记核酸。标记的另一个例子包括“分子信标”, 如 Tyagi 和 Kramer(Nature Biotech. 14:303, 1996)所述。杂交也可通过例如等离子共振来测定(参见例如, Thiel 等, Anal. Chem. 69:4948, 1997)。

在一个实施方式中, 标记多组(如 2、3、4、5 或更多)靶核酸并在一次杂交反应中使用(“多重”分析)。例如, 一组核酸可对应于一种细胞中的 RNA, 而另一组核酸则对应于另一种细胞中的 RNA。用不同的标记例如具有独特发射光谱的不同的荧光标记(如, 荧光素和罗丹明)对多组核酸进行标记, 从而实现区别。然后混合各组并与一个微阵列同时杂交(参见例如, Shena 等, Science 270:467-470, 1995)。

本发明中使用的微阵列包括一种或多种表征小分子疗效的基因的探针。在优

选实施方式中，微阵列包含对应于一种或多种选自下组的基因的探针：癌症中上调的基因和癌症中下调的基因。微阵列可包括对应于至少 10 种，优选至少 20 种，至少 50 种，至少 100 或至少 1000 种表征小分子疗效的基因的探针。

微阵列上对应于每种基因可存在一种或一种以上的探针。例如，微阵列可包含 2-20 种对应于一种基因的探针，优选约 5-10 种。探针可对应于表征小分子疗效的基因的全长 RNA 序列或其互补序列，或者探针可对应于其一部分，该部分的长度足以实现特异性杂交。这种探针可包含约 50 个核苷酸到约 100、200、500 或 1000 个核苷酸或 1000 个核苷酸以上。如本文进一步所述，微阵列可包含由约 10-50 个核苷酸，优选约 15-30 个核苷酸，更优选约 20-25 个核苷酸组成的寡核苷酸探针。探针优选是单链的，对靶点具有足够的互补序列数量以提供所需的序列特异性杂交水平。

典型地，本发明中使用的阵列的位点密度大于每平方厘米 100 种不同的探针。优选地，阵列的位点密度大于 $500/\text{cm}^2$ ，更优选大于约 $1000/\text{cm}^2$ ，最优选大于约 $10,000/\text{cm}^2$ 。优选地，阵列对单一底物具有 100 种以上不同的探针，更优选大于约 1000 种不同的探针，甚至更优选大于约 10,000 种不同的探针，最优选对单一底物具有大于 100,000 种不同的探针。

许多不同的微阵列构型及其制备方法是本领域技术人员已知的，参见美国专利：5,242,974；5,384,261；5,405,783；5,412,087；5,424,186；5,429,807；5,436,327；5,445,934；5,556,752；5,405,783；5,412,087；5,424,186；5,429,807；5,436,327；5,472,672；5,527,681；5,529,756；5,545,531；5,554,501；5,561,071；5,571,639；5,593,839；5,624,711；5,700,637；5,744,305；5,770,456；5,770,722；5,837,832；5,856,101；5,874,219；5,885,837；5,919,523；6,022,963；6,077,674；和 6,156,501；Shena 等，Tibtech 16:301，1998；Duggan 等，Nat. Genet. 21:10，1999；Bowtell 等，Nat. Genet. 21:25，1999；Lipshutz 等，21 Nature Genet. 20-24，1999；Blanchard 等，11 Biosensors 和 Bioelectronics，687-90，1996；Maskos 等，21 核酸 s Res. 4663-69，1993；Hughes 等，Nat. Biotechnol. (2001) 19:342；这些文献的内容以参考的方式结合在本说明书中。描述阵列在各种应用中的使用方法的专利包括：美国专利 5,143,854；5,288,644；5,324,633；5,432,049；5,470,710；5,492,806；5,503,980；5,510,270；5,525,464；5,547,839；5,580,732；5,661,028；5,848,659；和 5,874,219；它们的内容以参考的方式结合在本说明书中。

阵列优选包括对照和参比核酸。对照核酸包括例如，原核基因如 bioB, bioC 和 bioD, 来自 P1 噬菌体的 cre 或 polyA 对照, 如 dap、lys、phe、thr 或 trp。参比核酸能够将多次实验的结果进行标准化并在定量水平上比较多次实验结果。示例性的参比核酸包括已知表达水平的管家基因, 例如, GAPDH、己糖激酶和肌动蛋白。

在一个实施方式中, 可在固相支持物上合成寡核苷酸阵列。示例性的固相支持物包括玻璃、塑料、聚合物、金属、类金属、陶瓷、有机物等。采用芯片掩蔽技术和光防护化学, 可产生核酸探针的有序阵列。这些阵列是已知的, 例如“DNA 芯片”或非常大规模的固化聚合物阵列(“VLSIPSTM”阵列), 约 1 平方厘米到几个平方厘米面积的底物上可包含数以百万计的指定探针区域, 从而包含几个到几百万个探针(参见例如, 美国专利 5,631,734)。

为了比较表达水平, 在足以使靶化学与阵列上的探针相结合的条件下使标记的核酸与阵列接触。在优选的实施方式中, 可选择杂交条件以实现所需的杂交特异性水平; 就是说, 足以使标记的核酸与微阵列上的探针之间发生杂交的条件。

杂交可在允许基本上特异性杂交的条件下进行。核酸的长度和 GC 含量将决定热熔点, 以及实现探针与靶核酸特异性杂交所必需的杂交条件。这些因素是本领域技术人员熟知的, 也在通过实验测定。核酸杂交的通用法则可参见 Tijssen 等(《生物化学与分子生物学的实验技术》(Laboratory Techniques in Biochemistry and Molecular Biology), 第 24 卷: 《与核酸探针杂交》Hybridization With Nucleic Acid Probes, P. Tijssen 编, Elsevier, N.Y., (1993))。

上述方法导致阵列表面形成标记靶核酸的杂交模式。所得标记的核酸的杂交模式肉眼可见或者可以多种方法检测, 基于靶核酸特定的标记来选择特定的检测方法。代表性的检测方法包括: 闪烁计数法、放射自显影、荧光测量、比色测定、光发射测量、光散射等。

一种检测方法采用市售的阵列扫描仪(Affymetrix, Santa Clara, CA), 例如, 417TM 阵列仪、418TM 阵列仪或安捷伦基因阵列TM 扫描仪。该扫描仪由装有界面及易用软件工具的计算机系统所控制。输出结果可直接输入多种应用软件或直接由其读取。优选的扫描装置参见, 例如美国专利 5,143,854 和 5,424,186。

对于荧光标记的探针, 每个转录物阵列位点的荧光发射优选地由共聚焦激光扫描显微镜来检测。或者, 可采用允许在两种荧光团的特异性波长处同时样本发

光的激光,并且同时测定两种荧光团的发射光(参见例如,Shalon等,Genome Res. 6:639-645, 1996)。在优选实施方式中,可使用装有计算机控制的 X-Y 轴和显微镜目镜的激光荧光扫描仪来扫描阵列。荧光激光扫描装置在 Shalon 等(同上)中描述。

基因表达数据的分析存在许多算法,例如,进行比较的类型。在某些实施方式中,需要对协同调节的基因进行分组。这就需要比较大量的谱。鉴别这种基因组的一个优选的实施方式涉及群集算法(群集算法的综述参见例如, Fukunaga, 1990, Statistical Pattern Recognition, 第 2 版, Academic Press, San Diego; Everitt, 1974, Cluster Analysis, 伦敦: Heinemann Educ. Books; Hartigan, 1975, Clustering Algorithms, 纽约: Wiley; Sneath 和 Sokal, 1973, Numerical Taxonomy, Freeman; Anderberg, 1973, Cluster Analysis for Applications, Academic Press: 纽约)。

生物标记物的发现

可利用表达模式来导出一系列可用于预测患者中药物治疗疗效的生物标记物。生物标记物包括对从生物样品中分离的 RNA、从冻存的肿瘤活检样品分离的 RNA 的微阵列实验得到的基因表达水平,或血清中质谱-获得的蛋白质质量。

虽然数据分析的精确机制取决于数据的精确性质,一种典型的获取一系列生物标记物的方法如下所述。治疗之前收集每位患者的数据(基因表达水平或质谱)。随着研究的深入,根据对药物治疗的响应对患者进行分两类;有效或无效。数据模型中可存在多种疗效水平,但认为二元比较是最佳的,尤其是当患者群体数量小于几百时。假定每一类存在足够数量的患者,则可通过本领域公知的各种技术来分析蛋白质和/或基因表达数据。许多技术来源于传统统计学以及机器学习领域。这些技术用于实现以下两个目的:

1. 减少数据维数—在质谱或基因表达微阵列的情况下,将数据从几千个数据点降低至一次 3-10 个。这种降低基于视作一组时各数据点的预测能力。
2. 训练—然后用上述 3-10 个数据点来训练多种机器学习算法,在这种情况下使它们“学会”识别区别有效药物治疗与无效治疗的蛋白质质量或基因表达模式。所有患者样品都可用于训练算法。

然后测试所得经训练的算法以测定其预测能力。典型地,如果训练样品小于几百个,则进行一些形式的交叉验证。例如,考虑 10 倍交叉验证。在这种情况下

下, 将患者样品随机分配至 10 个接受器(bin)之一。在第一轮验证中, 将九个接受器中的样品用于训练, 剩余的第十接受器中的样品用于测试算法。该过程重复九次, 每次剔除在不同接受器中的样品用于测试。组合所有十轮结果(正确预测和误差), 然后评估预测能力。该研究以这种方式可比较不同的算法以及不同的小组。然后选择“最佳”算法/小组比较的组合。然后这种“聪明”算法可用于未来的研究, 以选择最有可能响应治疗的患者。

许多算法受益于取自患者的其他信息。例如, 性别或年龄可用于改善预测能力。并且, 数据转化如标准化和校平也可用于降低噪音。因此, 可采用许多不同的参数来训练各种算法以优化结果。如果数据中存在可预测的模式, 则可发展为最佳、或接近最佳的“聪明”算法。如果存在更多患者样品, 则可以运用新数据重新训练算法。

作为使用质谱的一个例子, 将血浆(1 μ l)应用于疏水 SELDI 靶, 水中彻底洗涤并通过质谱 SELDI-T 进行分析。可对 100 或更多个患者样品重复该过程。可对每个样品中一些 16,000 m/z 值的强度所产生的蛋白质谱进行统计学分析, 以鉴别能够预测药物疗效的特异性 m/z 值的组。采用其他 SELDI-靶也可进行相同的实验, 如离子交换或 IMAC 表面。它们能够捕获血浆中存在的不同的蛋白质亚组。而且, 应用于到 SELDI 靶上之前可使血浆变性和预先分馏血浆。

诊断与预后分析

本发明提供了确定对象是否存在发展成表现为不希望的细胞增殖的疾病或症状的风险的方法, 该方法包括检测生物标记物(如, VEGF 或 VEGFR, 如 VEGFR-2), 即, 癌症核酸和/或多肽标记物。

在临床应用中, 可筛选人体组织样品以确定是否存在本文所鉴别的生物标记物。所述样品可包括穿刺活检核、外科切除样品、淋巴结组织或血清。例如, 所述方法包括穿刺活检组织, 任选地通过冷冻切片分割以富集肿瘤细胞至约总细胞群的 80%。在某些实施方式中, 可采用本领域公知的技术扩增从这些样品提取的核酸。将检测的选定标记物的水平与转移性、非转移性的恶性、良性或正常组织样品的统计学有效组进行比较。

在一个实施方式中, 诊断方法包括: 例如通过 RNA 印迹分析、反转录-聚合酶链反应(RT-PCR)、原位杂交、免疫沉淀反应、蛋白质印迹杂交或免疫组织化学

方法来测定对象是否具有异常的生物标记物(如, VEGF 或 VEGFR, 如 VEGFR-2, 包括它们的可溶性形式)的 mRNA 和/或蛋白质水平。根据该方法, 从对象获取细胞, 测定生物标记物、蛋白质水平或 mRNA 水平并将这些水平与健康对象中的标记物水平进行比较。异常的生物标记物多肽或 mRNA 水平可能指示癌症。

因此, 在一个方面, 本发明提供了对本文所述独特的核酸标记物特异的探针和引物。因此, 核酸探针的核苷酸序列包含长度至少 10 个核苷酸, 优选至少 15 个核苷酸, 更优选 25 个核苷酸, 最优选至少 40 个核苷酸, 以及一直到与标记物核酸序列的编码序列的一部分互补的序列全长或接近全长。

在一个实施方式中, 该方法包括采用核酸探针来测定患者组织癌细胞的存在。具体地说, 该方法包括:

1. 提供核酸探针, 所述核酸探针的核苷酸序列包含长度至少 10 个核苷酸, 优选至少 15 个核苷酸, 更优选 25 个核苷酸, 最优选至少 40 个核苷酸, 以及一直到与核酸序列的编码序列的一部分互补且在肿瘤细胞中差异性表达的序列全长或接近全长;

2. 从潜在地包含癌细胞的患者获取组织;

3. 提供基本上所有细胞为非癌细胞的第二组织样品;

4. 在严谨条件下使核酸探针分别与所述第一和第二组织样品的 RNA 相接触(如, RNA 印迹或原位杂交分析); 和

5. 将(a)探针与第一组织样品的 RNA 杂交的量, 和(b)探针与第二组织样品杂交的量进行比较; 其中, 与第一组织样品的 RNA 杂交的量和第二组织样品的 RNA 杂交的量相比存在统计学显著性差异表明, 第一组织样品中存在癌细胞。

一方面, 所述方法包括: 与来源于给定标记物核酸序列(如, VEGF 或 VEGFR-2)的探针原位杂交。该方法包括: 使标记的杂交探针与潜在地含有癌细胞或癌前细胞以及正常细胞的给定类型的组织相接触, 并确定探针标记给定组织类型的一些细胞的程度与标记相同组织类型的其他细胞的程度是否存在统计学差异(例如, 因数至少是 2, 或因数至少是 5, 或因数至少是 20, 或因数至少是 50)。

本发明还提供了确定来自人体组织的测试细胞表型的方法, 例如无论所述细胞是(a)正常细胞, 或是(b)癌细胞或癌前细胞, 该方法包括: 使试验细胞的 mRNA 与核酸探针相接触, 所述核酸探针的长度至少是 12 个核苷酸, 优选至少 15 个核

苷酸,更优选至少 25 个核苷酸,最优选至少 40 个核苷酸的,以及一直到与核酸序列的编码序列的一部分互补且与给定组织类型的正常细胞相比、在肿瘤细胞中差异性表达的序列全长或接近全长;和测定探针与 mRNA 杂交的大致量,杂交的量比与该组织类型正常细胞的 mRNA 中看到的更多或更少表明,测试细胞是癌细胞或癌前细胞。

或者,可使用抗体进行上述诊断分析以检测由标记物核酸序列编码的蛋白质产物(如, VEGF 或 sVEGFR-2)。因此,在一个实施方式中,分析过程包括:使试验细胞的蛋白质与核酸基因产物特异性的抗体相接触,标记物核酸是与试验细胞相同组织类型的正常细胞中以给定控制水平表达的核酸,以及测定抗体与试验细胞的蛋白质形成的免疫复合物的大致量,其中,与试验细胞的蛋白质形成的免疫复合物的量和与相同组织类型的正常细胞相比存在统计学显著性差异表明,试验细胞是癌细胞或癌前细胞。优选地,抗体是对 VEGF 或 sVEGFR-2,尤其是其细胞外结构域特异的抗体。

产生与本发明中使用的多肽特异性结合的多克隆和/或单克隆抗体的方法是本领域技术人员已知的,例如可参见 Dymecki 等(J. Biol. Chem. 267:4815, 1992); Boersma 和 Van Leeuwen, (J. Neurosci. Methods 51:317, 1994); Green 等(Cell 28:477, 1982); 和 Arnheiter 等(Nature 294:278, 1981)。

另一种方法包括以下步骤:提供对标记物核酸序列的基因产物特异的抗体,所述基因产物在给定组织类型的癌组织中存在的水平大于或小于在相同组织类型的非癌组织中基因产物的水平;从患者获取给定组织类型的第一组织样品,该样品潜在地含有癌细胞;提供相同组织类型的第二组织样品(可来自同一患者或来自正常对照,例如另一个体或培养的细胞),该第二样品含有正常细胞且基本上是非癌细胞;在允许抗体与样品中存在的标记物核酸序列产物之间形成免疫复合物的条件下,使抗体与第一和第二样品的蛋白质(可部分地纯化,在裂解但未分馏细胞中,或原位)相接触;和比较(a)第一样品中形成的免疫复合物的量,与(b)第二样品中形成的免疫复合物的量,其中,第一样品中形成的免疫复合物的量与第二样品中形成的免疫复合物的量之间存在统计学显著性差异表明,第一组织样品中存在癌细胞。

本发明还提供了测定从对象获取的细胞样品是否具有异常量的标记物多肽的方法,该方法包括:(a)从对象获取细胞样品,(b)定量测定获取的样品中标记

物多肽的量, 和(c)将测定的标记物多肽的量与已知标准进行比较, 从而确定从对象获取的细胞样品是否具有异常量的标记物多肽。所述标记物多肽可通过免疫组织化学分析、斑点印迹分析、ELISA 等方法检测。

免疫分析常用于定量细胞样品中的蛋白质水平, 许多其他免疫分析技术也是本领域已知的。本发明并不限于特定的分析方法, 因而旨在包括均质和非均质方法。适用于本发明的示例性的免疫分析方法包括: 荧光极化免疫分析(FPIA)、荧光免疫分析(FIA)、酶免疫分析(EIA)、抑制免疫散射浊度分析(NIA)、酶联免疫吸附分析(ELISA)和放射免疫测定(RIA)。指示剂部分, 或标记基团可附连于抗体, 经选择以符合常常由分析设备的可获得性以及兼容的免疫分析方法所决定的各种使用方法的需要。进行上述各种免疫分析所用的一般技术是本领域普通技术人员已知的。

在另一实施方式中, 可测定患者生物流体(如血液或尿液)中编码产物的水平, 或可选地多肽的水平, 作为监测患者细胞中标记物核酸序列表达水平的一种方式。该方法可包括以下步骤: 从患者获取生物流体的样品, 使该样品(或样品的蛋白质)与对编码的标记物多肽特异的抗体相接触, 和测定抗体形成免疫复合物的量, 免疫复合物形成的量指示了样品中标记物编码产物的水平。将其与从正常个体获取的对照样品或者事先或随后从同一对象获取的一种或多种样品中相同抗体形成的免疫复合物的量进行比较时, 上述测定尤其具有指导意义。

在另一实施方式中, 该方法可用于测定细胞中存在的标记物多肽的量, 而这种量又与过度增殖疾病的进程相关。标记物多肽的水平可用于预测性地评估细胞样品是否含有变异细胞或倾向于形成变异细胞的细胞。而且, 该方法还可用于评价已知发生转化的细胞表型, 表型结果可用于规划具体的治疗方案。例如, 样品细胞中非常高的标记物多肽水平是有力的癌症诊断和预后标记物。标记物多肽水平的观察结果可用于作出是否采取更加强烈的治疗等方面的决定。

如上所述, 本发明的一方面涉及在从患者分离的细胞内容物中测定样品细胞中标记物多肽的水平是否显著降低的诊断分析。术语“显著性降低”是指细胞表型中相对于类似组织来源的正常细胞, 细胞标记物多肽的量降低。例如, 与正常对照细胞相比, 细胞标记物多肽的量减少约 50%、25%、10%或 5%。具体地说, 该分析方法可评估试验细胞中标记物多肽的水平, 优选地与至少一种对照细胞如正常细胞和/或已知表型的转化细胞中检测到的标记物多肽的测量水平相比较。

对本发明尤其重要的是能够通过正常或异常标记物多肽水平相关的细胞数量的测定来定量标记物多肽的水平。然后将具有特定标记物多肽表型的细胞数量与患者预后建立相关性。在本发明的一个实施方式中，测定损伤的标记物多肽表型作为活检中发现具有异常高/低的标记物多肽水平的细胞的百分比。这种表达可通过免疫组织化学分析、斑点印迹分析、ELISA 等检测。

当采用组织样品时，用免疫组织化学染色法来测定具有标记物多肽表型的细胞数量。对于这种染色，从活检或其他组织样品获取多块组织，并采用诸如蛋白酶 K 或胃蛋白酶等试剂进行蛋白水解。在某些实施方式中，可能需要从样品细胞分离核部分，并检测核部分中标记物多肽的水平。

用诸如福尔马林、戊二醛、甲醇等试剂处理以固定组织样品。然后将样品与能够和标记物多肽特异性结合的抗体一起孵育，优选单克隆抗体。该抗体可偶联有用于后续结合检测的标记。将样品孵育足以形成免疫复合物的时间。然后，利用偶联于抗体的标记来检测抗体的结合。如果抗体未标记，则采用标记的第二抗体，例如能够与抗标记物多肽抗体同种型特异性结合的第二抗体。可采用的标记的例子包括放射性核素、荧光剂、化学发光剂、酶等。

如果采用酶，则将酶底物加入到样品中以形成有色或荧光产物。适用于偶联物的合适的酶的例子包括辣根过氧化物酶、碱性磷酸酶、苹果酸脱氢酶等。如果不可购得，这种抗体-酶偶联物可容易地通过本领域技术人员已知技术来制备。

在一个实施方式中，分析方法是斑点印迹分析。斑点印迹分析尤其适用于测定采用组织样品的应用，因为它能够通过从预定量的细胞产生的无细胞提取物中的标记物多肽的量建立关联，来测定与单细胞有关的标记物多肽的平均量。

癌症文献中已有论述：相同类型的肿瘤细胞(例如乳房和/或结肠肿瘤细胞)中各个癌基因的表达可能不是统一地增加或各个肿瘤抑制基因的表达不是统一地降低。甚至是给定类型的癌症细胞之间也可能存在给定标记物基因的表达水平差异，进一步强调需要依赖于成套测试而不是单次测试。因此，在一个方面，本发明提供了采用多种本发明探针的成套测试，以改善诊断试验的可靠性和/或准确性。

在一个实施方式中，本发明还提供了一种方法，该方法包括：将核酸探针以有序阵列固定到 DNA 芯片上。核苷酸可通过包括平板印刷术等的许多方法结合到固体支持物上。例如，芯片可容纳有高达 250,000 个寡核苷酸。这些核酸探针

包含的核苷酸序列长度至少约 12 个核苷酸, 优选至少约 15 个核苷酸, 更优选至少约 25 个核苷酸, 最优选至少约 40 个核苷酸, 以及一直到与标记物核酸序列的编码序列的一部分互补且在肿瘤细胞中差异性表达的序列全长或接近全长。本发明提供了超出现有的各种癌症试验的显著性优点, 即通过在单一芯片上提供核酸标记物阵列而增加了试验的可靠性。

该方法包括: 获取活检组织, 任选地通过冷冻切片分割以富集肿瘤细胞至整个细胞群的约 80%。然后提取 DNA 或 RNA, 扩增并用 DNA 芯片分析以确定是否存在标记物核酸序列。

在一个实施方式中, 将核酸探针以二维矩阵或阵列打点到基片上。可将核酸样品进行标记, 然后与探针杂交。洗去未结合的样品部分之后, 可检测标记的样品核酸与探针核酸结合的双链核酸。

探针核酸可打点到包括玻璃、硝基纤维素等的基片上。探针可通过共价键或通过非特异性相互作用如疏水相互作用与基片结合。然后用放射活性标记、荧光团、生色团等标记样品核酸。

构建阵列的技术和使用这些阵列的方法参见例如, 欧洲专利 0 799 897; PCT 专利 WO 97/292 12; PCT 专利 WO 97127317; 欧洲专利 0 785 280; PCT 专利 WO 97/02357; 美国专利 5,593,839; 美国专利 5,578,832; 欧洲专利 0 728 520; 美国专利 5,599,695; 欧洲专利 0 721 016; 美国专利 5,556,752; PCT 专利 WO 95/22058; 和美国专利 5,631,734。

并且, 阵列可用于检查基因的差异表达, 用于确定基因功能。例如, 核酸序列阵列可用于确定任何核酸序列是否在正常细胞与癌细胞之间差异性表达。如果癌细胞中特定信息表达增加, 而在对应的正常细胞中未观察到, 则表明该信息是癌症特异性蛋白质。

在一个实施方式中, 使用核酸分子在固体表面(如膜)形成微阵列, 使得阵列化的核酸分子可用于确定任何核酸是否在正常细胞或组织与癌细胞或组织之间差异性表达。在一个实施方式中, 本发明核酸分子可以是 cDNA 或者可用于产生 cDNA 分子, 随后通过 PCR 扩增并打点到尼龙膜上。然后, 使膜与从癌细胞和正常组织或细胞的等价样品获取的放射标记的靶核酸分子反应。产生 cDNA 和微阵列制备的方法是本领域技术人员已知的, 参见例如 Bertucci 等 (Hum. Mol. Genet. 8:2129, 1999); Nguyen 等 (Genomics 29:207, 1995); Zhao

等(基因 156:207); Gress 等(Mammalian 基因组 3:609, 1992); Zhumabayeva 等(Biotechniques 30:158, 2001); 和 Lennon 等(Trends Genet. 7:314, 1991)。

在又一实施方式中, 本发明考虑使用一组针对本发明标记物多肽的抗体。优选地, 抗体针对 VEGF 或 sVEGFR-2。该组抗体可用作可靠的癌症诊断探针。本发明分析方法包括: 使含细胞(例如肺细胞)活检样品与一组针对一种或多种编码产物的抗体相接触, 以确定是否存在标记物多肽。

本发明诊断方法也可在治疗后进行, 例如定量标记物多肽的水平, 表明当前或先前采用的癌症疗法的有效性以及这些疗法对患者预后的效果。

此外, 标记物核酸或标记物多肽可用作初始检测、追踪筛选、复发检测、化疗或外科手术治疗的治疗后监测的诊断系列的一部分。

因此, 本发明提供了用于监测标记物多肽的获取和/或丧失的诊断分析和反应试剂, 以助于如细胞致癌转化等导致的增殖疾病的诊断和表型。

上述诊断分析也可适用作预后分析。这种应用利用了本发明分析方法对肿瘤进程中特征性阶段发生的事件的灵敏度。例如, 在非常早期阶段, 可能在细胞不可逆地发展成恶性之前, 给定标记物基因可上调或下调, 而另一标记物基因可能只在较晚期阶段特征性地上调或下调。所述方法可包括以下步骤: 使试验细胞的 mRNA 与在肿瘤进程的不同阶段在癌细胞或癌前细胞中差异性表达的来源于给定标记物核酸的核酸探针相接触, 和测定探针与细胞 mRNA 杂交的大致量, 所述量表明细胞中基因表达水平, 从而表明细胞肿瘤进程阶段; 或者所述分析可用对给定标记物核酸的基因产物特异的抗体来进行, 该方法包括使抗体与测试细胞的蛋白质相接触。一系列试验不仅能指明了肿瘤的存在和位置, 而且使临床医生能够选择最合适的肿瘤治疗模式和预测该治疗成功的可能性。

本发明方法也可用于追踪肿瘤临床病程。例如, 本发明分析方法可应用于来自患者的组织样品; 癌症患者经治疗后, 获取另一组织样品并再次进行试验。成功的治疗将导致所有表现为癌细胞或癌前细胞特征性差异表达的细胞消失, 或使那些细胞中显著增加的基因表达接近或者甚至是超过正常水平。

预测性分析

基于能够预测给定抗癌药临床疗效的实验室分析试验将大大改善癌症患者的临床控制。为了评价该作用, 在治疗之前、治疗期间和治疗之后分析生物样品

(例如肿瘤样品、血浆)中与抗癌药有关的生物标记物。

例如,可检测血浆中 VEGF 或 sVEGFR,优选 VEGFR-2 多肽的表达。因此,可监测癌症患者中这些多肽基线血浆浓度的变化。例如, VEGF 水平增加、sVEGFR-2 水平降低可能与索拉非尼疗效有关。

此外,也可采用石蜡包埋的肿瘤活检组织,通过定量免疫组织化学试验来监测多肽水平。

监测治疗的另一种方法是评价血清蛋白质组谱图(serum proteomic spectra)。具体地说,将血浆样品进行质谱(如表面增强的激光解析和电离)分析,得到每位患者的蛋白质组谱图。通过迭代搜索算法分析治疗之前和治疗期间来自患者血浆分析结果的一系列图谱,所述算法可鉴别能够完全区别治疗样品和未治疗样品的蛋白质组特征。然后,用所得特征来预测治疗后的临床效果。

可采用生物样品(例如肿瘤活检样品、血液样品)的全基因表达谱和生物信息学来源的模式鉴别来预测临床效果和灵敏度,以及抗癌药耐受性的发展。例如,采用 Affymetrix 基因芯片技术和算法,使用治疗之前和治疗期间从患者全血细胞分离的 RNA 来产生全血细胞基因表达谱。这些基因表达谱可预测用特定抗癌药治疗后的临床效果。

也可通过 1D $^1\text{H-NMR}$ (核磁共振),采用尿液生物化学组成分析进行预测分析。模式识别技术可用于评价用抗癌药治疗后的代谢效应,并将该效应与临床终点相关联。尿液中分泌的生物化学或内源性代谢物已通过正常对象的质子 NMR 很好地表征(Zuppi 等, Clin Chim Acta 265:85-97, 1997)。这些代谢物(约 30-40 种)代表了主要代谢途径的副产物,例如柠檬酸和尿素循环。已表明药物-、疾病-和遗传-刺激能够在基线尿谱中导致代谢特异性改变,指示刺激代谢响应的时间线和幅度。这些分析方法是多变量的,因而采用模式识别技术以改善数据处理过程。尿中代谢谱可与临床终点相关,用于确定临床效果。

无需进一步的详细描述,认为本领域技术人员能够通过上述说明,最大程度地利用本发明。因此,下面的具体优选的实施方式应解释为仅仅是示例性的而非以任何方式限制本发明的其余部分。

在上述说明和下面的实施例中,除非另有说明,所有温度是未校正的摄氏度,所有份数和百分比是重量份数和百分比。

实施例

下面将参考以下非限制性实施例来解释本发明。

介绍：III 期靶向研究，在晚期肾细胞癌患者中进行随机的安慰剂对照的研究，探究索拉非尼对总存活率的作用。本研究的第二个目的是评价对特定生物标记物的治疗作用和评价这些生物标记物与患者表现间的相关性。**方法：**收集样品用于鉴别候选生物标记物，具体地包括肿瘤活检样本、全血、血浆和尿液。以 IRB 批准的方案对所有分析物进行分析。通过 DNA 测序和免疫组织化学分析(IHC)的磷酸-ERK (pERK)水平，分析组织样本中的 *VHL* 基因突变状态(只对同意进行遗传分析的患者进行)。从血中分离 mRNA 并通过微阵列分析与患者表现相关的基因表达谱。此外，采用基于质谱的方法来评价血浆中的蛋白质标签，通过 $^1\text{H-NMR}$ 测定尿中与患者表现有关的小分子的谱。ELISA 法测定治疗前和治疗后血浆样品中的 VEGF 和可溶性 VEGFR-2 (sVEGFR-2)，以及这些分子与索拉非尼治疗有关的变化。

结果：在用索拉非尼治疗的患者中，治疗 3 周后血浆中的 sVEGFR-2 显著降低($p < 2\text{E}^{-16}$)，这种降低一直持续到治疗第 8 周。药物处理患者中 sVEGFR-2 的降低与目标损伤减少弱相关 ($p=0.028$)。相同时间段内安慰剂处理患者血浆中 sVEGFR-2 水平无变化。也测定了上述两组患者的 VEGF 水平。在接受索拉非尼的患者中，治疗 3 周后 VEGF 的水平从基线显著升高($p=3.2 \text{E}^{-5}$)，但安慰剂处理患者中没有观察到升高。安慰剂组中，高基线 VEGF 患者(>250 pg/mL)的无进展存活率时间比低基线 VEGF 患者(<250 pg/mL) ($p=0.030$)显著更短，而接受索拉非尼的患者中，高或低基线 VEGF 水平的患者之间未观察到 PFS 的显著性差异。IHC 染色 pERK 显示，大多数患者样品具有高的最大染色强度(0 到 4+的量表上为 4+)。类似地，大多数样品中 pERK 染色肿瘤细胞的百分比低(染色的肿瘤细胞<25%)。

进行研究检查患者中的 VEGF 水平，以鉴别与临床表现的相关性。

发现有用的患者表现测量是采用 Cox 回归分析的死亡时间(TTD)和无进展存活率(PFS)。测定 VEGF 的基线水平，并在治疗第 18 周测定基线水平的变化。观察结果经年龄、性别、ECOG 状态或 Motzer 评分调整。这些调整是最小程度的。

VEGF 基线水平与 TTD 的关系如图 1 所示(Kaplan-Meyer 分析)，第 18 周 VEGF 的变化与 TTD 的关系如图 2 所示(Kaplan-Meyer 分析)。如图所示，较高的基线 VEGF 水平与较短的死亡时间相关(如果测定 VEGF 作为连续变量)。第 18

周 VEGF 水平的显著增加也与较短的死亡时间有关(如果测定 VEGF 作为连续变量)。

图 1 是 VEGF 基线与 TTD 的图;

图 2 是第 18 周 VEGF 从基线的变化与 TTD 的图。

无需进一步的详细描述,认为本领域技术人员能够通过上述说明,最大程度地利用本发明。因此,前述的具体优选的实施方式应解释为仅仅是示例性的而非以任何方式限制本发明的其余部分。

替换上述实施例中使用的本发明的一般或具体描述的反应物和/或操作条件,上述实施例可取得类似的成功。

通过上述说明,本领域技术人员能够在不背离本发明的主旨和范围的情况下,容易地确定本发明的本质特征,能够进行各种改变和改进以适应各种用途和条件。

相信根据上述信息和本领域已知的信息,本领域技术人员能够最大程度地利用本发明。本领域普通技术人员明白,可对本发明进行各种改变和改进而不背离本文所述的本发明的主旨或范围。上文和下文提及的标题用作本申请中出现的某些信息的导引,而不是本申请中这些标题中出现的信息的唯一来源。上文引用的所有出版物和专利以参考的方式结合在本文中。

生物标记物			表现测量	比较 p 值				
				未调整	年龄调整	性别调整	ECOG 状态调整	Motzer 评分调整
VEGF	基线	连续	TTD	0.015	0.017	0.014	0.016	0.038
VEGF	基线	连续	PFS	0.030	0.033	0.020	0.025	0.052
VEGF	第 18 周的 ΔBL	连续	TTD	0.017	0.038	0.023	0.034	0.035

生物标记物变量			比较 p 值			解释
			TTD	PFS	BR	
VEGF	基线水平	连续	0.015	0.030	0.431	较高的基线 VEGF 水平与较短的 TTD 和较短的 PFS 有关
		固定 (binned)	0.081	0.171	ND	趋势：较高的基线 VEGF 水平与较短的 TTD 有关
	从基线到第 18 周的变化	连续	0.017	0.637	0.575	第 18 周 VEGF 水平的较大增加与较短的 TTD 有关
		固定 (binned)	0.020	0.190	ND	第 18 周 VEGF 水平的较大增加与较短的 TTD 有关

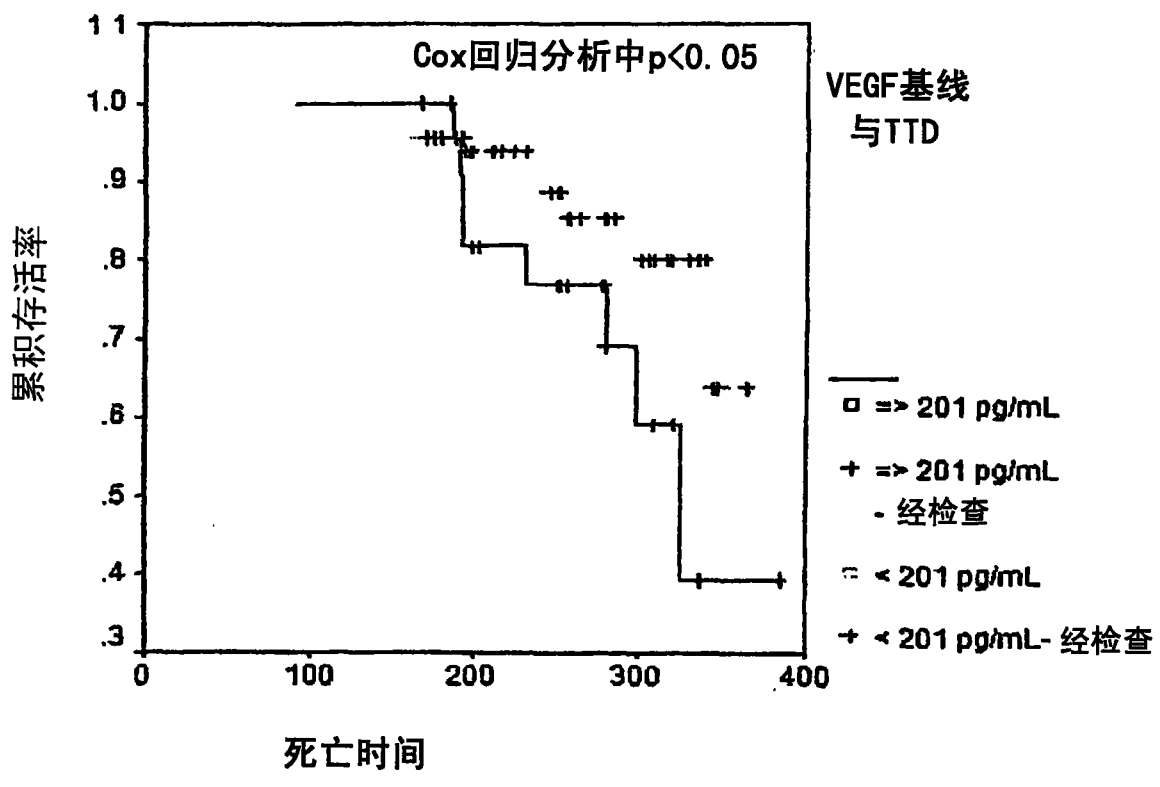


图 1

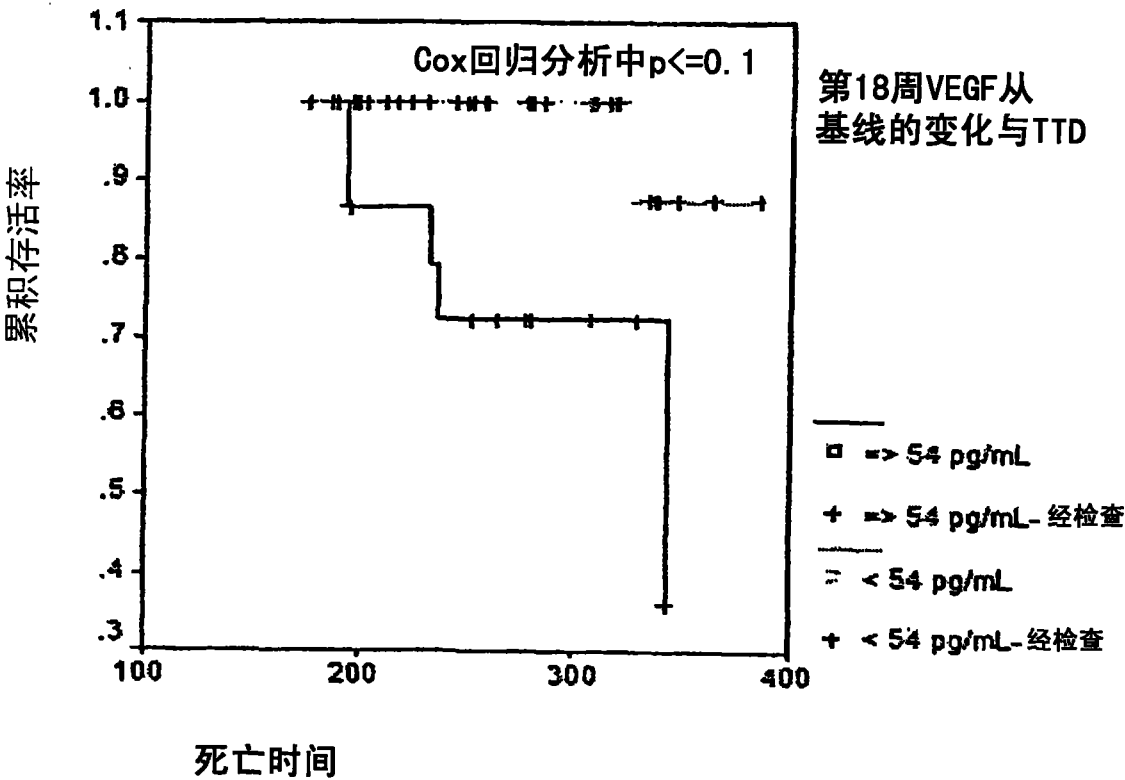


图 2