

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G01N 33/574 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200780020492.9

[43] 公开日 2009 年 6 月 17 日

[11] 公开号 CN 101460850A

[22] 申请日 2007.6.4

[21] 申请号 200780020492.9

[30] 优先权

[32] 2006.6.2 [33] EP [31] 06114954.8

[32] 2006.6.27 [33] US [31] 60/816,613

[86] 国际申请 PCT/EP2007/004931 2007.6.4

[87] 国际公布 WO2007/140966 英 2007.12.13

[85] 进入国家阶段日期 2008.12.2

[71] 申请人 阿特拉斯抗体有限公司

地址 瑞典斯德哥尔摩

[72] 发明人 M·乌莱恩 F·庞蒂恩

[74] 专利代理机构 北京北翔知识产权代理有限公司

代理人 张广育 姜建成

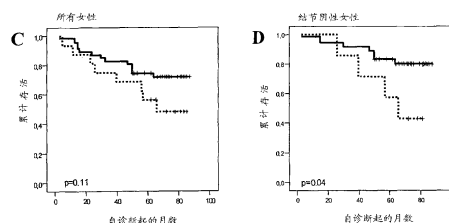
权利要求书 6 页 说明书 38 页 序列表 7 页
附图 16 页

[54] 发明名称

SATB2 蛋白作为结肠直肠癌标记的应用

[57] 摘要

本发明通过识别作为结肠直肠癌类型的标记的 SATB2 蛋白, 提供了与这种癌症的检测、鉴定和预后有关的新方法、手段和应用。



- 1 确定患有或怀疑患有结肠直肠癌的哺乳动物受试者的结肠直肠癌预后是否不良的方法，包括以下步骤：
 - a) 提供一个来自该受试者的样品；
 - b) 对存在于所述样品中的 SATB2 蛋白进行定量以得到一个样品值；
 - c) 将得自步骤 b) 的样品值与一个对照值进行比较；并且，如果所述样品值低于所述对照值，
 - d) 则判断所述受试者的结肠直肠癌预后不良。
- 2 根据权利要求 1 的方法，其中步骤 d) 中的不良预后对应五年存活的可能性为 65%或更低。
- 3 根据权利要求 2 的方法，其中步骤 d) 中的不良预后对应五年存活的可能性为 60%或更低。
- 4 对有需要的结肠直肠癌受试者的治疗方法，包括：
 - a) 提供一个来自该受试者的样品；
 - b) 对存在于所述样品中的 SATB2 蛋白进行定量，得到一个样品值；
 - c) 将得自步骤 b) 的样品值与对照值进行比较；并且，如果所述样品值低于所述对照值，
 - d) 则用一个适用于结肠直肠癌的预后不良的治疗方案治疗所述受试者。
- 5 根据权利要求 4 的方法，其中所述治疗方案选自化学疗法、新辅助疗法及它们的组合。
- 6 根据权利要求 5 的方法，其中所述治疗方案为新辅助疗法。
- 7 根据权利要求 6 的方法，其中所述新辅助疗法选自 i) 仅放射疗法和 ii) 放射疗法结合化学疗法。
- 8 根据前述权利要求任一项的方法，其中所述结肠直肠癌是结节阴性的。

9 根据前述权利要求任一项的方法，其中所述结肠直肠癌处于 Dukes A 或 B 期。

10 根据前述权利要求任一项的方法，其中所述结肠直肠癌是结肠直肠腺瘤。

11 根据权利要求 1-9 的任一项的方法，其中所述结肠直肠癌是结肠直肠上皮癌 (colo-rectal carcinoma)。

12 根据权利要求 1-7 的任一项的方法，其中所述结肠直肠癌是转移性癌。

13 根据权利要求 1-7 和 12 的任一项的方法，其中所述结肠直肠癌处于 Dukes C 期。

14 根据前述权利要求任一项的方法，其中所述受试者是人类。

15 根据权利要求 14 的方法，其中所述受试者是女性。

16 根据前述权利要求任一项的方法，其中所述对照值是一个预定值，该值与一个对照样品中的 SATB2 表达量相对应。

17 根据前述权利要求任一项的方法，其中所述对照值是 SATB2 阳性细胞的分数分值为 50% 。

18 根据权利要求 17 的方法，其中所述对照值是 SATB2 阳性细胞的分数分值为 25 %。

19 根据权利要求 1-16 任一项的方法，其中所述对照值是 SATB2 表达的自动分值为 70。

20 根据权利要求 19 的方法，其中所述对照值是 SATB2 表达的自动分值为 50。

21 根据前述权利要求任一项的方法，其中所述样品是一种体液样品。

22 根据权利要求 21 的方法，其中所述体液选自下组，其构成为血液、血浆、血清、脑液、尿液、精液和渗出液。

23 根据权利要求 1-20 任一项的方法，其中所述样品是一种粪便样品。

24 根据权利要求 1-20 任一项的方法，其中所述样品是一种组织样品。

25 根据权利要求 24 的方法，其中所述组织样品是一种结肠直肠组织样品。

26 根据权利要求 1-20 任一项的方法，其中所述样品是一种细胞学样品。

27 根据前述权利要求任一项的方法，其中所述 SATB2 蛋白的氨基酸序列包括一个选自如下的序列：

- i) SEQ ID NO: 1; 和
- ii) 一个与 SEQ ID NO: 1 至少 85%相同的序列。

28 根据前述权利要求任一项的方法，其中所述 SATB2 蛋白的氨基酸序列包括一个选自如下的序列：

- i) SEQ ID NO: 2; 和
- ii) 一个与 SEQ ID NO: 2 至少 85%相同的序列。

29 根据前述权利要求任一项的方法，其中步骤 b) 包括：

- b1) 对所述样品应用一种可定量的亲和配体，所述亲和配体能够选择性地与特定量的 SATB2 蛋白相互作用，所述应用是在使所述亲和配体能够与存在于所述样品中的任何 SATB2 蛋白结合的条件下施行的；
- b2) 除去未结合的亲和配体；并且
- b3) 对与该样品保持相关联的任何亲和配体进行定量。

30 根据权利要求 29 的方法, 其中所述可定量的亲和配体选自下组, 其构成为抗体、其片段和其衍生物。

31 根据权利要求 29 的方法, 其中所述可定量的亲和配体是源自一种支架的蛋白配体, 该支架选自下组, 其构成为葡萄球菌 A 蛋白及其结构域、脂质运载蛋白、锚蛋白重复结构域、纤维素结合结构域、 γ 晶体、绿色荧光蛋白、人类细胞毒性 T 淋巴细胞相关抗原 4、蛋白酶抑制剂、PDZ 结构域、肽适体、葡萄球菌核酸酶、淀粉酶抑肽、纤维连接蛋白 III 型结构域和锌指结构。

32 根据权利要求 29 的方法, 其中所述可定量的亲和配体是一种寡核苷酸分子。

33 根据权利要求 29 的方法, 其中所述可定量的亲和配体含有一个标签, 该标签选自下组, 其构成为荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

34 根据权利要求 29-33 的任一项的方法, 其中所述可定量的亲和配体是用一种能够识别所述可定量的亲和配体的二级亲和配体检测的。

35 根据权利要求 34 的方法, 其中所述能够识别所述可定量的亲和配体的二级亲和配体含有一个标签, 该标签选自下组, 其构成为荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

36 执行根据前述权利要求任一项的方法的试剂盒, 所述试剂盒包括:

- a) 能够选择性地与 SATB2 蛋白相互作用的可定量的亲和配体; 和
- b) 对所述亲和配体进行定量所必需的试剂。

37 根据权利要求 36 的试剂盒, 其中所述可定量的亲和配体选自下组, 其构成为抗体、其片段和其衍生物。

38 根据权利要求 36 的试剂盒,其中所述可定量的亲和配体是源自一种支架的蛋白配体,该支架选自下组,其构成为葡萄球菌 A 蛋白及其结构域、脂质运载蛋白、锚蛋白重复结构域、纤维素结合结构域、 γ 晶体、绿色荧光蛋白、人类细胞毒性 T 淋巴细胞相关抗原 4、蛋白酶抑制剂、PDZ 结构域、肽适体、葡萄球菌核酸酶、淀粉酶抑肽、纤维连接蛋白 III 型结构域和锌指结构。

39 根据权利要求 36 的试剂盒,其中可定量的亲和配体是一种寡核苷酸分子。

40 根据权利要求 36-39 的任一项的试剂盒,其中所述可检测的亲和配体含有一个标签,该标签选自荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

41 根据权利要求 36-39 的任一项的试剂盒,其中所述用于亲和配体进行定量的必需试剂包括能够识别所述可定量的亲和配体的一种二级亲和配体。

42 根据权利要求 41 的试剂盒,其中所述能够识别所述可定量的亲和配体的二级亲和配体含有一个标签,该标签选自下组,其构成为荧光染料或金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

43 根据权利要求 36-42 的任一项所述的试剂盒,所述试剂盒还包括用于提供对照值的一个对照样品。

44 SATB2 蛋白作为一种预后标记的应用。

45 SATB2 蛋白作为一种结肠直肠癌预后标记的应用。

46 SATB2 蛋白或其抗原活性片段在制造用于结肠直肠癌预后的预后试剂中的应用。

47 根据权利要求 44-46 任一项的应用，其中所述 SATB2 蛋白的氨基酸序列包含一个选自如下的序列：

- i) SEQ ID NO: 1; 和
- ii) 一个与 SEQ ID NO: 1 至少 85%相同的序列。

48 根据权利要求 44-46 任一项的应用，其中所述 SATB2 蛋白的氨基酸序列包含一个选自如下的序列：

- i) SEQ ID NO: 2; 和
- ii) 一个与 SEQ ID NO: 2 至少 85%相同的序列。

49 能够选择性地与 SATB2 蛋白相互作用的亲和配体，所述亲和配体是一种抗体或其片段或衍生物。

50 根据权利要求 49 的亲和配体，所述亲和配体可从一种包括用一种蛋白免疫动物的步骤的方法中获得，所述蛋白的氨基酸序列包含 SEQ ID NO: 1 序列。

51 根据权利要求 49 和 50 的任一项的亲和配体作为预后试剂的应用。

52 根据权利要求 49 和 50 的任一项的亲和配体用于结肠直肠癌预后的应用。

53 根据权利要求 49 和 50 的任一项的亲和配体在制造用于结肠直肠癌预后的预后试剂中的应用。

SATB2 蛋白作为结肠直肠癌标记的应用

技术领域

本发明涉及癌症诊断和预后领域。具体而言，它提供了一种通过鉴别作为结肠直肠癌的标记的 SATB2 蛋白来用于检测和鉴定结肠直肠癌的新方法。

背景技术

SATB2

编码富含 AT 的特殊序列结合蛋白 2 (special AT-rich sequence-binding protein 2, SATB2) 的基因是在 1999 年人类基因组测序的浩大工程中被鉴别出的 (Kikuno R *et al* (1999) DNA Res. 6: 197-205)。自此，SATB2 基因被认为主要在神经组织中表达。

SATB2 是一个形成核基质各部分的转录因子，并且它通过与富含 AT 的序列（也称为基质结合区 (MARs)）相互作用来调控高度有序的染色质 (high-order chromatin) 结构从而进行组织特异性的基因表达 (Dickinson LA *et al* (1992) Cell 70, 631-45; FitzPatrick DR *et al* (2003) Hum. Mol. Genet. 12, 2491-501; Yasui D, (2002) Nature 419, 641-5; Bode, J (2000) Crit. Rev. Eukaryot. Gene. Expr. 10, 73-90)。

对该基因及其蛋白产物 SATB2 蛋白的研究，指出它作为转录因子在神经组织中参与基因表达的调控 (Dobrev G *et al* (2003) Genes Dev. 17: 3048-3061; Britanova O *et al* (2005) Eur. J. Neurosci. 21: 658-668)。还有记载，SATB2 基因在腭发育和腭裂中起作用 (FitzPatrick DR *et al* (2003) Human Mol. Genet. 12: 2491-2501; van Buggenhout G *et al* (2005) Eur. J. Med. Genet. 48: 276-289)。

Salahshor 等人对一名腺瘤性结肠息肉病 (APC) 基因突变的患者进行了研究 (Salahshor *et al* (2005) BMC cancer 5: 66)。APC 患者在年轻时产生异常量的结肠腺瘤，如果不治疗，最终将发展成结肠直肠癌。总基因表达谱揭示了一组包括 SATB2 在内的 84 个基因在腺瘤内的表达相比在正常粘膜中有显著改变。SATB2 被发现显著下调但未被选用于做进一步的分

析。Int J Cancer 上关于结肠直肠癌表达图谱的新近研究同样指出 SATB2 在 mRNA 水平上的表达状态有所改变 (Groene J *et al* (2006) Int J Cancer 119, 1829-1836)。

PCT 公布文本 W003/022126 和 W02006/015742 描述了关于癌细胞表达图谱的其他相似研究。对包括 SATB2 的大量基因的表达进行分析并根据总的表达模式得出结论。

重要地是,上述研究并未提供关于 SATB2 蛋白作为特异的结肠直肠标记的应用或 SATB2 作为结肠直肠癌预后工具的应用的暗示。

癌症

癌症是西方最常见的疾病和死亡原因之一。一般而言,大多数的癌症形式的发病率是随年龄增加的。因为人类总体健康状况提高所致的寿命的继续增长,患癌症的个体会更多。尽管日益增加的知识指出环境因素(饮食、烟草、紫外线等)及遗传因素(“癌症基因”如 p53、APC、BRCA1、XP 等的种系突变)与癌症的发病风险之间有一种联系,但大多数常见类型的癌症的病因一般来说还不清楚。

尽管癌症本质上是一种细胞疾病并被定义为伴有净细胞增长和有害群体行为的转化细胞群,但从细胞生物学角度来看,关于癌症的定义都不能令人完全满意。恶性转化代表基于不可逆转的遗传改变的向恶性表型的转变。尽管还没被正式地证明,恶性转化被认为发生在一个细胞内,由此始发随后产生的肿瘤(“癌症克隆形成能力”学说)。致癌作用是指产生癌症并通常被认为包括多个最终导致恶性肿瘤生长的事件的过程。这个多步过程包括数个速率限制步骤,如突变的增加,还可能包括外遗传

(epigenetic)事件,导致癌变前增殖阶段后的癌症形成。最常见类型的癌症发生在体细胞,其中绝大多数起源于上皮组织(皮肤、前列腺、结肠和肺),然后是从血细胞生成谱系发生的癌症(白血病和淋巴瘤)和从间充质细胞发生的癌症(肉瘤)。逐步的变化包括决定细胞分裂、自我行为和细胞死亡的重要调控通路中的错误(突变)累积。每一种变化都提供相对于周围环境细胞的达尔文选择生长优势,从而导致肿瘤细胞群的净生长。需要强调的是恶性肿瘤不仅由转化的肿瘤细胞自身组成而且还包括作为支持性间质的周围正常细胞。这种补充的癌症间质构成了结缔组织、血

管和多种其他正常细胞,如炎性细胞,它们共同起到向转化的肿瘤细胞提供肿瘤继续生长所必需的信号的作用。

癌症诊断

对取自肿瘤的组织切片进行显微评价仍然是确诊癌症的金标准。对基因组 DNA、转录基因和表达蛋白的分析在显微镜检测到的组织学特征基础上又增加了重要信息。今后的诊断、预后信息和治疗选择都可能基于对形态学的整体评价和对核酸及蛋白质的分析。现在,基于人类基因组序列和生物化学通路(包括组织中细胞内和细胞间信号传导)的不断进步的认知,已经可以揭示肿瘤形成中不同阶段以及表型变化(这定义了不同的癌症类型)的某些机制。

尽管分子生物学已取得显著进步,癌症诊断仍然要依靠光学显微镜。分子工具的发展对将癌症细胞与普通细胞区分开来起到了重要的作用,而且这种重要性还在增加。除了组织切片的组织化学染色之外,最常用的方法是免疫组织化学法。免疫组织化学法使得可以使用特定的抗体来检测组织和细胞中的蛋白表达模式。免疫组织化学法在临床诊断上的应用不仅可以分析组织结构和细胞形态学,还可以检测不同细胞群中的免疫活性。这对于支持不同原发性肿瘤的精确评级和分类,以及对不明起源的转移的诊断已经变得非常重要。现在临床实践中,最常用的抗体包括抗细胞类型标记(如 PSA、MelanA、甲状腺球蛋白)的抗体和识别中间丝、分化抗原簇等和潜在恶性标记(如 Ki67、p53 和 HER-2)的抗体。除了免疫组织化学法,应用原位杂交来检测基因扩增和应用基因测序来分析突变是癌症诊断学中正在发展的技术。

结肠直肠癌

结肠直肠癌是世界上最常见的人类癌症形式之一。Parkin 等人提供的 GLOBOCAM 2002 数据库的数据显示每年有大约 100 万新的结肠直肠癌病例被发现(Parkin *et al* (2007) CA Cancer J Clin 55, 74-108)。另外,世界上结肠直肠癌的发生占全部癌症的 9.4%左右,而且结肠直肠癌是西方世界第二最常见的死亡原因。西方世界的结肠直肠癌的五年存活率约为 60%,但是在东欧和印度则低至 30%。

目前早期检测和切除肿瘤的外科手术对良好的预后是至关重要的。症状取决于肿瘤在远端胃肠道中的位置，它包括肠不良应激、腹泻、便秘、疼痛和贫血（继发于肿瘤出血至肠道中）。恶性肿瘤根据不同的分类方案可被分为几个阶段，如 TNM/UICC 分类法 I 到 IV 期或 Dukes A 到 C 期。最轻度恶性肿瘤（Dukes A 和 B 期）可获得合理有利的效果，而另一极端，某些转移性高度恶性肿瘤（Dukes C 期）的存活率则很低。目前的诊断基于患者病史、临床和内窥镜检查（直肠镜检查 and 结肠镜检查），随后任选地进行放射成像以确定肿瘤生长的范围。对可疑损伤处灶的活组织检查与内窥镜检查结合进行。

显微诊断，指从可疑肿瘤收集活检材料并置于显微镜下观察。为了获得可靠的诊断结果，再将肿瘤组织固定于福尔马林中、进行组织学处理和石蜡包埋。从所得的石蜡块获得组织切片并用组织化学方法和免疫组织化学方法进行染色。

对限定在局部区域的肿瘤，即还没有发展为转移性疾病的肿瘤，通过外科手术将肿瘤及周围的肠和组织完全切除。然后将外科手术标本送去进行粗略的病理学和显微分析。这种分析形成了肿瘤分期的基础。目前最常见形式的结肠直肠癌是代表腺源肿瘤的腺癌，它可高度、中度或低度分化。

对原发性肿瘤，苏木素-伊红染色的组织切片足够用于根据结肠直肠癌分类法作出正确的诊断和分类。但是，由于结肠直肠癌非常常见并经常在检出之前已经生长得很大，转移却不常见。肿瘤一般会转移至局部淋巴结，但是远距离转移至肝或肺的却很少见。癌症的一个常见的临床问题是患者表现出未知来源的转移。如果转移的是腺癌，可怀疑数种可能的原发性肿瘤，如乳腺癌、前列腺癌、胰腺癌、胃癌和结肠直肠癌。对于鉴别诊断，可以使用可识别来源细胞的固有特征的免疫组织化学标记。目前，细胞角蛋白 20 (CK20) 被用于作为结肠直肠癌的特征，它是一种胃肠道腺细胞中富含的中间丝标记。但是，几种其他腺癌也是对 CK20 抗体呈阳性的，尽管不是所有结肠直肠癌都为阳性。另外，现在没有可将低度恶性和低转移风险的肿瘤与威胁生命的高度恶性肿瘤区分开来的可用标记。

为了让医生针对正确的癌症类型给予具体且尽可能早的治疗，提供新的只针对结肠直肠癌并可以将患者按不同风险类别进行区分的分子标记非常重要。总之，目前非常需要能够提高结肠直肠癌诊断和筛选的新方法。

发明内容

本发明的一个目的是通过提供可用于受试者结肠直肠癌的诊断和/或预后的标记来满足上述需要。

本发明的一个相关目的是提供可用于区分结肠直肠癌和其他类型癌症的标记。

本发明的另一目的是提供结肠直肠癌诊断、预后和/或治疗的新方法。

本发明的一个相关目的是提供一种可与结肠直肠癌诊断、预后和/或治疗的方法结合使用的试剂盒。

本发明的另一目的是提供可用于结肠直肠癌诊断、预后和/或治疗的新化合物。

对于这些目的以及本领域技术人员从本发明公开的内容中显而易见的其他目的,本发明在不同方面提供了确定结肠直肠癌状态和预后及其治疗的新方法。

因此,在第一方面,本发明提供了确定患有结肠直肠癌或怀疑患有结肠直肠癌的哺乳动物受试者的结肠直肠癌预后是否不良的方法,所述方法包括以下步骤:

- a) 提供取自所述受试者的一个样品;
- b) 对存在于所述样品中的 SATB2 蛋白进行定量以得到一个样品值;
- c) 将得自 b) 的样品值与对照值进行比较;并且,如果所述样品值低于所述对照值,
- d) 则判断所述受试者的结肠直肠癌预后不良。

本发明的第一方面是基于以往未认识到的事实,即取自患有结肠直肠癌或怀疑患有结肠直肠癌的受试者的样品中 SATB2 蛋白的表达可作为受试者疾病状态的指示物。更具体地,本发明第一次指出一方面的 SATB2 的低表达值与另一方面的更具攻击性或高风险的形式的结肠直肠癌之间相关性。基于以 SATB2 表达作为结肠直肠癌预后指示物的本发明有若干优点。一般来说,对于癌症,攻击性形式肿瘤的早期检测非常重要,因为早期可以治愈。对于结肠直肠癌而言尤其如此,已有几个大型研究表明患有早期即 1 期和 2 期(实际上是 Dukes A 和 B 期)癌症的受试者的预后,与患有晚期肿瘤的受试者相比要好很多。这种差别不是取决于治疗方式,因为所有类型的结肠直肠癌都会进行完全切除。这种存活上的巨大差别而是

与早期检测、正确诊断和适当的外科手术治疗明显有关。SATB2 蛋白作为标记,其某种表达水平与某种模式的疾病进程相关,它在例如用于转移性的鉴别诊断方面具有巨大的潜力。

在本发明的一个实施方案中,步骤 d) 中不良预后的结论可能包括表明,所述受试者与其不表现低 SATB2 表达值的情况相比,具有更短的预期存活时间。或者/另外,不良预后的结论可能包括表明,与不表现低 SATB2 表达值的情况相比,具有更低的 5 年存活可能性。例如,结论可能为所述受试者的 5 年存活可能性为 65%或更低,例如,60%或更低,50%或更低,40%或更低或 30%或更低。

此外,对于患有或怀疑患有淋巴结阴性肿瘤的受试者,结论可能是所述受试者的 5 年存活可能性为 73%或更低,例如,70%或更低,例如 60%或更低,50%或更低,40%或更低,或者 30%或更低。对于女性受试者,结论可能为所述受试者的 5 年存活可能性为 74%或更低,例如 70%或更低,例如 60%或更低,50%或更低,40%或更低或 30%或更低。对于患有或怀疑患有结节阴性肿瘤的女性受试者,结论可能为所述受试者的 5 年存活可能性为 80%或更低,例如 75%或更低,例如 70%或更低,例如 60%或更低,50%或更低,40%或更低或 30%或更低。

在 SATB2 的低表达与结肠直肠癌的高风险形式之间所识别的这种相关性也可构成决定对受试者相对于其不表现低 SATB2 表达值的情况施行不同治疗方案的基础。因此,在第二方面,本发明提供了治疗需要治疗的受试者的结肠直肠癌的方法,包括

- a) 提供取自该所述受试者的一个样品;
- b) 对存在于所述样品中的 SATB2 蛋白进行定量以得到一个样品值;
- c) 将得自 b) 的样品值与对照值进行比较;并且,如果所述样品值低于所述对照值,
- d) 则采用适合于预后不良的结肠直肠癌的一个治疗方案来治疗所述受试者

在本发明的一个实施方案中,所述治疗方案选自化学疗法、新辅助疗法及其组合。

因此,所述治疗方案可以是新辅助疗法。所述新辅助疗法可仅由放射治疗构成或由放射治疗与化学疗法结合构成。

在上述本发明的方法各方面中,所述受试者可以患有或怀疑患有不同形式和/或阶段的结肠直肠癌。

在这些方面的某些实施方案中,要研究的结肠直肠癌是结节阴性的结肠直肠癌,即还没有发展到结节转移阶段的结肠直肠癌。在其他类似的实施方案中,要研究的结肠直肠癌的特征是处于 Dukes A 或 B 期。在其他实施方案中,要研究的结肠直肠癌是结肠直肠腺瘤或结肠直肠癌

(carcinoma)。在这些实施方案中,确定受试者表现出低 SATB2 表达可能对疾病的未来发展预后具有重要价值,并因此形成对未来疾病处理的相关可靠决定的基础。在一组患有这种相对早期阶段疾病的受试者中,低 SATB2 表达的受试者可能有发展为更具攻击性的疾病相对较高风险。因此,患有结节阴性结肠直肠癌或 Dukes A 或 B 期结肠直肠癌的受试者中低 SATB2 表达可说明这些受试者应该得到比不表现低 SATB2 表达的受试者更密切地监控和/或被施以与不表现低 SATB2 表达的受试者不同的治疗。由于 SATB2 标记提供的补充预后信息,本发明的方法因此提供了使这种受试者能够在某段时间存活和/或存活更长时间的更大可能性。

在其他实施方案中,要研究的结肠直肠癌是转移性结肠直肠癌。在其他类似的实施方案中,要研究的结肠直肠癌的特征为处于 Dukes C 期。

在本发明的实施方案中,所述受试者是人类,如女人。如所附实施例所示,SATB2 标记的预后价值在患有结节阴性结肠直肠癌的人类女性受试者组中特别显著。

SATB2 蛋白表达的样品值低于对照值的确定,本文有时将其表达为“低 SATB2 表达”的确定。

在本发明的方法中,作为与受试者样品值相比较的对照值可通过不同的方法确定。作为一个非限制性的实例,对照值可对应进行所述预后或治疗的受试者健康组织中的 SATB2 表达量。作为另一个实例,对照值可由取自另一个对照受试者的正常组织的标准样品测得的 SATB2 表达量提供。作为另一个实例,对照值可由如 Dukes 期 A 或 B 癌症组织的肿瘤组织的标准样品测得的 SATB2 表达量提供。

对照值可通过实施本发明上述各方面的方法获得。可替代地,对照值是得自对照样品并且与所述对照样品的 SATB2 表达量一致的预定值。

对一个样品中的 SATB2 表达定量的另一种替代方法是测定 SATB2 表达高于某一水平的样品中细胞分数。所述测定例如可按下面的实施例第 4

部分“分数分值”的定义中所述进行。本发明以上各方面的方法的实施方案中，步骤 d) 的结论的标准是 SATB2 阳性细胞的核分数样品值，即“分数分值”，它低于对照值 50%，例如低于 40%，如低于 30%，如低于 25%，如低于 20%，如低于 15%，如低于 10%，如低于 5%，如低于 1%。另外，不良预后的确定可对应于在样品中基本上检测不到 SATB2 阳性细胞，即“分数分值”基本为零。

对一个样品中的 SATB2 表达定量的另一种替代方法是用自动扫描仪和图像处理软件自动测量 SATB2 表达的自动分值 (autoscore)。这种测定例如可按下面的实施例第 5 部分“自动分值”的定义中所述进行。本发明上面各方面的方法的实施方案中，步骤 d) 的结论的标准是样品细胞中 SATB2 表达的样品值，即“自动分值”，它低于对照值 70，如低于 60，如低于 50，如低于 40，如低于 30，如低于 25，如低于 20，如低于 15，如低于 10，如低于 5。

在本发明的一些实施方案中，无论是作为如上的分数分值或自动分值或作为某种其他已知的或被适当改变的变量的样品值和/或对照值的测定，都是在受试者的远端胃肠道即阑尾、结肠和/或直肠的腺细胞和/或结肠直肠癌细胞上进行的。

在本发明的另一实施方案中，不良预后的确定对应于在受试者的远端胃肠道的腺细胞中没有可检出的 SATB2 表达。

在本发明的上下文中，对术语“样品值”和“对照值”做宽泛的解释。如上所述，可通过自动方法或通过一个基于样品视觉或显微镜检查的评分系统来对 SATB2 的表达定量以获得这些值。但是，一个技术人员如组织病理学领域的技术人员，也可仅通过检查例如已进行 SATB2 表达染色的组织切片就可确定所述样品值和对照值。因此，样品值低于对照值的判断可对应于视觉或显微镜检查后的判断，即样品组织切片比对照组织切片染色密度低和/或有更少的染色细胞。在这样情况下，所述样品值和对照值被认为是技术人员在检查和比较后确定的主观值。因此，本发明不仅限于使用自动分析。

本发明的方法中用于检测 SATB2 蛋白表达的具体步骤不限于任何特定方式。在本发明方法的某些实施方案中，步骤 b) 包括：

b1) 对所述样品应用一种可定量的亲和配体,所述亲和配体可选择性地与待定量的 SATB2 蛋白相互作用,所述应用是在允许亲和配体与存在于所述样品中的任何 SATB2 蛋白结合的条件下施行的。

b2) 除去未结合的亲和配体; 并且

b3) 对于与该样品保持相关联的任何亲和配体进行定量。

在本发明的这些实施方案中,所述的取自所述受试者的样品可为体液样品,如血液、血浆、血清、脑液、尿液、精液和渗出液样品。在本发明的方法中,作为替代方案,所述样品可以是粪便样品、细胞学样品或组织样品,如结肠直肠组织的样品。

在一个优选的实施方案中,本发明的方法是在体外进行的。

技术人员会认识到本发明的应用不限于对被研究的受试者体内存在的 SATB2 蛋白任何具体变体定量,只要所述蛋白是被相关基因编码并表现相关的表达形式即可。作为一个非限制性的实例,所述 SATB2 蛋白的氨基酸序列包括一个选自如下的序列:

i) SEQ ID NO: 1; 和

ii) 一个与 SEQ ID NO: 1 至少 85%相同的序列。

在某些实施方案中,上面的序列 ii) 与 SEQ ID NO: 1 至少 90%相同,至少 91%相同,至少 92%相同,至少 93%相同,至少 94%相同,至少 95%相同,至少 96%相同,至少 97%相同,至少 98%相同,或者至少 99%相同。

作为另一个非限制性的实例,所述 SATB2 蛋白的氨基酸序列包括一个选自如下的序列:

i) SEQ ID NO: 2; 和

ii) 一个与 SEQ ID NO: 2 至少 85%相同的序列。

在某些实施方案中,上面的序列 ii) 与 SEQ ID NO: 2 至少 90%相同,至少 91%相同,至少 92%相同,至少 93%相同,至少 94%相同,至少 95%相同,至少 96%相同,至少 97%相同,至少 98%相同,或者至少 99%相同。

在根据本发明的方法的实施方案中,SATB2 蛋白是通过对样品应用一种可检测和/或可定量的亲和配体来检测和/或定量的,该亲和配体能够特异性或选择性地与 SATB2 蛋白相互作用。亲和配体的应用是在允许亲和配体与样品中的任何 SATB2 蛋白结合的条件下施行的。一旦通过本申请公开内容的教导了解了 SATB2 与结肠直肠癌之间的关系,本领域普通技术人员就能够选择或制造合适的亲和配体以及选择用于检测和/或定量的合适形

式和条件。尽管如此，为了说明，仍然在下面给出证明可用的亲和配体的实例，以及用于检测和/或定量的形式和条件的实例。

因此，在本发明的某些实施方案中，使用的亲和配体选自抗体及其片段和其衍生物，即基于免疫球蛋白支架的亲和配体。抗体包含任何来源的单克隆和多克隆抗体，包括鼠类抗体、人类抗体和其他抗体，以及含有不同物种的序列的嵌合抗体，如部分人源化的小鼠抗体。多克隆抗体是通过用选定抗原免疫动物来生产，而具有确定的特异性的单克隆抗体可通过 Köhler 和 Milstein 开发的杂交瘤技术来生产 (Köhler G and Milstein C (1976) Eur. J. Immunol. 6: 511-519)。抗体片段及衍生物包括 Fab 片段，它由完整免疫球蛋白的重链第一恒定区 (CH1)、轻链恒定区 (CL)、重链可变区 (VH) 和轻链可变区 (VL) 组成；Fv 片段，它由抗体的两个可变区 VH 和 VL 组成 (Skerra A and Plü ckthun A (1988) Science 240: 1038-1041)；单链 Fv 片段 (scFv)，它由通过可弯曲的肽接头连结在一起的 VH 和 VL 区组成 (Bird RE and Walker BW (1991) Trends Biotechnol. 9: 132-137)；Bence Jones 二聚体 (Stevens FJ *et al* (1991) Biochemistry 30: 6803-6805)；骆驼 (camelid) 重链二聚体 (Hamers-Casterman C *et al* (1993) Nature 363: 446-448) 和单链可变区 (Cai X and Garen A (1996) Proc. Natl. Acad. Sci. U.S.A. 93: 6280-6285; Masat L *et al* (1994) Proc. Natl. Acad. Sci. U.S.A. 91: 893-896)，以及单结构域链支架如铰口鲨的新抗原受体 (NAR) (Dooley H *et al* (2003) Mol. Immunol. 40: 25-33) 和基于重链可变区的微抗体 (minibody) (Skerra A and Plü ckthun A (1988) Science 240: 1038-1041)。

多克隆和单克隆抗体以及它们的片段和衍生物，代表在需要选择性生物分子识别的应用中亲和配体的常规选择，例如在根据本发明对 SATB2 蛋白的检测和/或定量中。然而，本领域的技术人员了解，因为对特异性结合配体的高通量生产和低成本生产系统的需求增加，新的生物分子多样化技术在过去十年中已经发展起来。这使免疫球蛋白源及非免疫球蛋白源的新型亲和配体得以生产，所述非免疫球蛋白源被证明同样可作为结合配体用于生物分子识别应用并可代替免疫球蛋白，或者与其一起使用。

选择亲和配体所需的生物分子多样性可通过多种可能的支架分子之一的组合工程来产生，然后通过合适的筛选平台来选择特异性和/或选择

性的亲和配体。所述支架分子可以是免疫球蛋白源 (Bradbury AR and Marks JD (2004) J. Immunol. Meths. 290:29-49)、非免疫球蛋白源 (Nygren PÅ and Skerra A (2004) J. Immunol. Meths. 290:3-28) 或寡核苷酸源 (Gold L *et al* (1995) Annu. Rev. Biochem. 64:763-797)。

许多非免疫球蛋白的蛋白支架已经被用作新结合蛋白的开发中的支持结构。可用于生产本发明中所用抗 SATB2 的亲和配体的这类结构的非限制性实例为, 葡萄球菌 A 蛋白及其结构域和这些结构域的衍生物, 如蛋白 Z (Nord K *et al* (1997) Nat. Biotechnol. 15:772-777); 脂质运载蛋白 (Beste G *et al* (1999) Proc. Natl. Acad. Sci. U.S.A. 96:1898-1903); 锚蛋白重复结构域 (Binz HK *et al* (2003) J. Mol. Biol. 332:489-503); 纤维素结合结构域 (CBD) (Smith GP *et al* (1998) J. Mol. Biol. 277:317-332; Lehtiö J *et al* (2000) Proteins 41:316-322); γ 晶体 (Fiedler U and Rudolph R, W001/04144); 绿色荧光蛋白 (GFP) (Peelle B *et al* (2001) Chem. Biol. 8:521-534); 人类细胞毒性 T 淋巴细胞相关抗原 4 (CTLA-4) (Hufton SE *et al* (2000) FEBS Lett. 475:225-231; Irving RA *et al* (2001) J. Immunol. Meth. 248:31-45); 蛋白酶抑制剂, 如 Knottin 蛋白 (Wentzel A *et al* (2001) J. Bacteriol. 183:7273-7284; Baggio R *et al* (2002) J. Mol. Recognit. 15:126-134) 和库尼兹 (Kunitz) 结构域 (Roberts BL *et al* (1992) Gene 121:9-15; Dennis MS and Lazarus RA (1994) J. Biol. Chem. 269:22137-22144); PDZ 结构域 (Schneider S *et al* (1999) Nat. Biotechnol. 17:170-175); 肽适体, 如硫氧还蛋白 (Lu Z *et al* (1995) Biotechnology 13:366-372; Klevenz B *et al* (2002) Cell. Mol. Life Sci. 59:1993-1998); 葡萄球菌核酸酶 (Norman TC *et al* (1999) Science 285:591-595); 淀粉酶抑肽 (McConnell SJ and Hoess RH (1995) J. Mol. Biol. 250:460-479; Li R *et al* (2003) Protein Eng. 16:65-72); 基于纤维连接蛋白 III 型结构域的 trinectin (Koide A *et al* (1998) J. Mol. Biol. 284:1141-1151; Xu L *et al* (2002) Chem. Biol. 9:933-942); 以及锌指结构 (Bianchi E *et al* (1995) J. Mol. Biol. 247:154-160; Klug A (1999) J. Mol. Biol. 293:215-218; Segal DJ *et al* (2003) Biochemistry 42:2137-2148)。

上面提到的非免疫球蛋白的蛋白支架的实例包括:呈现出单随机环的支架蛋白,该环用于新的结合特异性的生成;具有刚性二级结构的蛋白支架,该结构中从蛋白表面突出的侧链是随机的以生成新的结合特异性;和呈现出不连续的高变的环区的支架,该环区用于新结合特异性的生成。

除了非免疫球蛋白的蛋白,寡核苷酸也可被用作亲和配体。被称为诱饵(decoy)的单链核酸,折叠成界限清楚的三维结构并以高亲和性和特异性与它们的靶标结合(Ellington AD and Szostak JW (1990) *Nature* 346:818-822; Brody EN and Gold L (2000) *J. Biotechnol.* 74:5-13; Mayer G and Jenne A (2004) *BioDrugs* 18:351-359)。寡核苷酸配体既可以是RNA也可以是DNA,并可与很多类别的目标分子结合。

为了从上面提到的任何支架结构的变体库中选择想要的亲和配体,有许多筛选平台可用于分离特异的抗所选目标蛋白的新型配体。筛选平台包括但不限于:噬菌体展示(Smith GP (1985) *Science* 228:1315-1317),核糖体展示(Hanes J and Plü ckthun A (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94:4937-4942),酵母双杂交系统(Fields S and Song O (1989) *Nature* 340:245-246),mRNA展示(Roberts RW and Szostak JW (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94:12297-12302),SELEX(指数富集式配体系统进化)(Tuerk C and Gold L (1990) *Science* 249:505-510)和蛋白片段互补测定法(PCA)(Remy I and Michnick SW (1999) *Proc. Natl. Acad. Sci. U. S. A.* 96:5394-5399)。

因此,在本发明的实施方案中,可使用如下的亲和配体,它是来自任一上文所列蛋白支架的非免疫球蛋白亲和配体,或者是寡核苷酸分子。

在本发明方法的某些实施方案中,能够选择性地与SATB2蛋白相互作用的亲和配体可被检测和/或定量。这种亲和配体的检测和/或定量可以通过技术人员知道的任何在基于生物相互作用的测定中对结合试剂进行检测和/或定量的方法完成。因此,如前一部分所述的任何亲和配体,可被定量或定性地用于检测SATB2蛋白的存在。这些“一级”亲和配体自身可被各种标记物标记,或者通过可以检测、目测和/或定量的被标记的二级亲和配体间接检测。这可以使用多种标签的任何一个或多个并使用技术人员已知的许多技术中的任何一种或多种来完成,这其中并不需要任何过多的实验,所述多种标签可与能够与SATB2相互作用的亲和配体或任何二级亲和配体偶联。

可与一级和/或二级亲和配体偶联的标签的非限制性实例包括荧光染料或金属（如荧光素、罗丹明、藻红蛋白、荧光胺），发色团染料（如视紫红质），化学发光化合物（如鲁米那、咪唑）和生物发光蛋白（如萤光素、萤光素酶），半抗原（如生物素）。多种其他有用的荧光剂和发色团在 Stryer L (1968) *Science* 162: 526-533 和 Brand L 和 Gohlke JR (1972) *Annu. Rev. Biochem.* 41: 843-868 中有描述。亲和配体也可被酶（如辣根过氧化物酶、碱性磷酸酯酶和 β -内酰胺酶）、放射性同位素（如 ^3H 、 ^{14}C 、 ^{32}P 、 ^{35}S 或 ^{125}I ）和颗粒（如金）标记。不同类型的标签可通过不同的化学法与亲和配体偶联，如胺反应或巯基反应。但是，也可使用胺和巯基以外的其他反应基团，如醛、羧酸和谷氨酰胺。

本发明的方法的各方面可通过数种已知的形式和设置中的任何形式和设置进行使用，其非限制性的选择将在下面讨论。

在一种基于组织学的设置中，一种与 SATB2 靶标结合的已标记的亲配体的检测、定位和/或定量可包括视觉化技术如光学显微镜或免疫荧光显微镜。其他方法可能包括通过流式细胞术或发光法来检测。

如上所阐释，对受试者体内的 SATB2 蛋白的检测和/或定量可通过从所述受试者采集生物样品来完成，所述生物样品例如采集自例如结肠直肠组织、血样、脑液、尿液或粪便的组织样（活组织检查）。将亲和配体用于生物样品以对 SATB2 标记蛋白检测和/或定量。这种方法不仅可检测 SATB2 蛋白，还可显示其分布和相对表达水平。

亲和配体上的标签的视觉化方法包括但不限于：荧光技术、发光技术和/或酶促技术。荧光检测和/或定量是通过将荧光标签暴露在特定波长的光下，然后对发射出来的特定波长的光进行检测和/或定量而实现。发光标记的亲配体的存在可通过化学反应中发出的光来检测和/或定量。酶促反应的检测是基于化学反应过程中产生的样品颜色变化。本领域的技术人员应认识到，为了正确的检测和/或定量，可对多种不同方法作出修改。

在本发明的方法中，生物样品可被固定于固相支持物或载体上，如硝化纤维素或其他任何能够固定施加于固体支持基质本身的生物样品中存在的任何 SATB2 蛋白的固体支持基质。某些可用于本发明的熟知的固态支持材料包括玻璃、碳水化合物（如琼脂糖凝胶（Sephacrose））、尼龙、塑料、毛、聚苯乙烯、聚乙烯、聚丙烯、右旋糖苷、淀粉酶、膜、树脂、

纤维素、聚丙烯酰胺、琼脂糖、氧化铝、辉长岩和磁铁矿。如果一级亲和配体本身没有被标记,可以用本领域已知的各种缓冲液清洗支持基质,然后使其与标记的二级亲和配体接触,再用缓冲液清洗一次以除去未结合的亲和配体,然后可以用常规方法检测和/或定量选择性的亲和配体。亲和配体的结合特性会随不同的固态支持而变化,但是本领域的技术人员能够通过常规实验为每个测定确定有效的和最佳的测定条件。

本发明需要的检测和/或定量 SATB2 蛋白的方法是通过将亲和配体连接至一种随后可以在酶免疫测定(如 EIA 或 ELISA)中被检测和/或定量的酶上。这类技术已经相当成熟,而且它们的实现对于技术人员而言不存在任何过分的困难。在这类方法中,使生物样品与一种固体材料或一种偶联有抗 SATB2 蛋白亲和配体的固体材料相接触,然后它可通过酶标记的二级亲和配体进行检测和/或定量。随后,使一种合适的底物在合适的缓冲液中与酶标签反应生成一种化学成分,该化学成分可用分光光度计、荧光计、发光计或视觉化方法检测和/或定量。

如上所述,一级和任何二级亲和配体可用放射性同位素标记从而能够被检测和/或定量。本发明中合适的放射性标签的非限制性实例为 ^3H 、 ^{14}C 、 ^{32}P 、 ^{35}S 或 ^{125}I 。经标记的亲和配体的比放射性活度依赖于放射性标签的半衰期、同位素纯度和标签是如何掺入亲和配体中的。对亲和配体的标记优选地使用熟知的技术(Wensel TG and Meares CF (1983)

in: *Radioimmunoimaging and Radioimmunotherapy* (Burchiel SW and Rhodes BA eds.) Elsevier, New York, pp 185-196)。可通过在体内或体外检测放射活性,将这种经放射性标记的亲和配体用于显示 SATB2 蛋白。放射性核扫描(如用 γ 相机)、磁共振波谱或发射断层术可用于体外或体内检测,而 γ/β 计数器、闪烁计数器和射线照相术也可在体外使用。

本发明的另一方面提供了一种用于实施本发明上述各方法方面的方法的试剂盒,所述试剂盒包括:

- a) 一种能够选择性地与 SATB2 蛋白相互作用的可定量的亲和配体; 和
- b) 对所述亲和配体定量所必需的试剂。

本发明的试剂盒的各种成分按照上面本发明的方法方面的阐述进行选择 and 确定。

因此,本发明的试剂盒包含抗 SATB2 亲和配体,以及其他在特异和/或选择性的亲和配体与 SATB2 特异地和/或选择性地结合之后帮助对所述

亲和配体定量的工具。例如，本发明的试剂盒可含有一种二级亲和配体，所述二级亲和配体用于检测和/或定量任何由 SATB2 蛋白与所述能够选择性地与 SATB2 蛋白相互作用的亲和配体形成的复合物。本发明的试剂盒还可含有多种除亲和配体之外的辅助物质，以使得可容易有效地使用该试剂盒。所述辅助物质的实例包括溶解或复原该试剂盒中冻干蛋白组分的溶剂、洗涤缓冲液、当酶用作标签时用于测定酶活性的底物，以及免疫测定试剂盒中的常用物质例如反应终止剂。

本发明的试剂盒还可有利地包含用于提供对照值的对照样品，对照值用于与样品值相比较。这种对照样品可由例如一个具有预定量 SATB2 蛋白的组织样品构成，病理学领域技术人员可使用所述预定量的 SATB2 蛋白，通过目测比较或自动比较对照样品与受试者样品之间的 SATB2 表达水平来确定 SATB2 在被研究的样品中的表达状态。

本发明的另一个方面提供了 SATB2 蛋白作为一种预后标记的应用。还提供了 SATB2 蛋白作为一种结肠直肠癌预后标记的应用。

本发明的一个相关方面提供了 SATB2 蛋白或其抗原活性片段在制造用于结肠直肠癌预后的预后药剂中的应用。SATB2 蛋白的抗原活性片段是足够大的可用于产生亲和配体（如抗体）的片段，该亲和配体将与含有所述片段的 SATB2 蛋白相互作用。

在本发明的这些应用方面的实施方案中，作为一个非限制性的实例，SATB2 蛋白的氨基酸序列包括一个选自如下的序列：

- i) SEQ ID NO:1; 和
- ii) 一个与 SEQ ID NO:1 至少 85%相同的序列。

在某些实施方案中，上面的序列 ii) 与 SEQ ID NO:1 至少 90%相同，至少 91%相同，至少 92%相同，至少 93%相同，至少 94%相同，至少 95%相同，至少 96%相同，至少 97%相同，至少 98%相同，或者至少 99%相同。

作为另一个非限制性的实例，SATB2 蛋白的氨基酸序列包括一个选自如下的序列：

- i) SEQ ID NO:2; 和
- ii) 一个与 SEQ ID NO:2 至少 85%相同的序列。

在某些实施方案中，上面的序列 ii) 与 SEQ ID NO:2 至少 90%相同，至少 91%相同，至少 92%相同，至少 93%相同，至少 94%相同，至少 95%相同，至少 96%相同，至少 97%相同，至少 98%相同，或者至少 99%相同。

在另一个方面,本发明提供能够选择性地与 SATB2 蛋白相互作用的亲和配体,它是一个抗体或者抗体的片段或抗体的衍生物。这样的抗体或者其片段或衍生物,可例如通过包含用氨基酸序列含有 SEQ ID NO: 1 序列的蛋白免疫动物的步骤的方法获得。生成针对给定靶标的抗体或者其衍生物或片段的方法是本领域已知的,且可用于本发明的这一方面。任何上面讨论的 SATB2 蛋白 (SEQ ID NO: 2) 或其抗原活性的片段 (SEQ ID NO: 1) 的变体,当然可用于这种产生抗体或者其片段或衍生物的方法中。

另一个方面,本发明提供了根据本发明的亲和配体作为预后药剂的应用。该应用的一个优选的实施方案是将亲和配体用作结肠直肠癌预后的预后药剂。本发明还提供了所述亲和配体用于结肠直肠癌的预后的应用。作为其相关的方面,本发明提供了本发明的亲和配体在制造用于结肠直肠癌预后的预后药剂中的应用。

本发明还在其另一个方面提供了一种用于结肠直肠癌预后的方法,所述方法包括检测 SATB2 蛋白的步骤。本发明的这一方面是基于发现 SATB2 在一般情况下可以用作结肠直肠组织——具体而言,结肠直肠癌——的蛋白标记。如下面要详述的,针对 SATB2 蛋白片段产生的抗体在远端胃肠道即阑尾、结肠和直肠的腺细胞和结肠直肠癌中表现出选择性的强核免疫活性。最显著的发现是在 11 种结肠直肠癌中全部都为阳性。除了结肠直肠癌,只在极少数的其他肿瘤中呈弱或中度阳性。

另外, SATB2 在其他类型癌症中相对来说几乎不存在,这又使针对 SATB2 的亲和配体成为可特异地将结肠直肠癌和其他癌症区分开来的高度引人关注的工具。大多数结肠直肠癌是腺源的并因此被分类为腺癌。这是一种典型的癌症类型,并且也可以源自多种其他器官。因此,本发明的发现在用于确定肿瘤转移的类型时受到高度关注,其中所述肿瘤源自哪个器官通常是未知的。目前,可用的结肠直肠癌分子标记会与其他腺癌交叉反应,并因此对肿瘤的定位和转移源的确定造成了困难。本发明特定的结肠直肠癌标记可使医生有效定位癌症,提供更有效的治疗并最终为患者提供更可靠的预后。

本发明的另一方面包括同时用 SATB2 和 CK20 标记来检测癌症样品。如实施例的第 6 部分所详述,对 SATB2 和 CK20 表达的联合检测在分辨结肠直肠癌方面的预测价值超过对每种标记物单独检测的预测价值。因此,本发明在此方面提供了一种诊断结肠直肠癌的方法,包括检测 SATB2 蛋白

和检测 CK20 蛋白的步骤。此外，本发明还提供了一种通过检测 SATB2 蛋白和/或 CK20 蛋白的存在来检测一种转移是否源自结肠直肠癌的方法。通过综合从 CK20 和 SATB2 获得的信息，患者会更容易地获得结肠直肠疾病的精确诊断。技术人员能够对本文关于 SATB2 检测的教导作适当改变以使其适应本发明此方面的方法，并且按照本文的描述和按照例如免疫组织化学领域的知识不需要过分的负担就能进行 SATB2 和 CK20 的同时或顺序检测和/或定量。

本发明的一个引人关注的方面是预测癌症患者的 SATB2 蛋白会泄漏到血浆和粪便中。作为对比，一个熟知的前列腺癌标记 PSA（也在正常前列腺中表达）甚至在健康患者中也会从前列腺泄漏到血浆中。但是，许多前列腺癌患者的血浆的 PSA 水平有所升高，因此对于具有患前列腺癌风险的男性群体，检测 PSA 水平的升高是一种常用的早期筛选方法。SATB2 和结肠直肠癌之间被预测有同样的联系。因此，也可通过将人类血浆或其他体液或粪便作为本发明的样品从而将 SATB2 用作筛选结肠直肠癌的工具。在这一点上，本发明具有广泛的应用价值，甚至可能替代目前结肠直肠癌的筛选方法例如结肠直肠镜检查或乙状结肠镜检查，这些检查非常令人不适因此很多人放弃，尽管美国癌症研究所建议结肠直肠癌危险群体做定期检查。基于本发明的、将 SATB2 作为结肠直肠癌标记蛋白的筛选方法，对患此类癌症的患者的筛查、及早发现和治疗都有很大的好处。

在本发明的背景下，“预后”是指对一种疾病及其治疗的过程或结果的预测。预后也可指一种疾病的存活或痊愈机会的确定，以及对受试者的预期存活时间的预测。预后特别地可包括确定受试者在到将来的一段时间内存活的可能性，如三年、五年、十年或任何其他时间段。

在本发明的背景下，“诊断”是指确定一种疾病或病症的存在或对其进行鉴定。诊断也指通过这个过程得到的结论。在本文中，“诊断的”是指关于或有助于确定疾病存在或性质。在本发明的上下文中，“诊断”和“诊断的”也指监控一种疾病随时间自然发生的任何改变或者由于治疗产生的任何改变。

根据上面的定义，显然，术语“预后”和“诊断”有重叠的意义，它们不互相排斥。

在本发明的背景下，例如一个亲和配体与其靶标或抗原的“特异性的”或“选择性的”相互作用是指特异性和非特异性的相互作用之间的区别或

者选择性和非选择性的相互作用之间的区别具有意义。两种蛋白之间的相互作用有时用离解常数来衡量。离解常数描述了两个分子之间的结合（或亲和性）强度。一般，抗体及其抗原之间的离解常数是 10^{-7} 到 10^{-11} M。但高特异性不必要求高亲和性。已证明，对其对方具有低亲和性（在摩尔范围内）的分子已经被证明与具有更高亲和性的分子的特异性一样。在本发明的情况下，特异性的或选择性的相互作用是指，在天然存在的或经加工的生物流体样品中在存在其他蛋白的给定条件下，使得一种具体方法可用于确定一种特定蛋白即所述靶蛋白或其片段的存在和/或量的程度。换言之，特异性或选择性是分辨各相关蛋白的能力。本说明书中特异性和选择性有时是可互换使用的。

在本发明的背景下，“单特异性抗体”是用其自身抗原亲和纯化的一群多克隆抗体中的一种，通过所述亲和纯化可将该单特异性抗体与其他抗血清蛋白和非特异性抗体分开。这种亲和纯化可产生与自身抗原选择性结合的抗体。本发明中，用一种基于免疫亲和的两步法纯化多克隆抗血清，获得对目标蛋白有选择性的单特异性抗体。通过将固定化标签蛋白用作捕获剂，在初次消耗步骤中除去针对抗原片段通用亲和和标签的抗体。在该初次消耗步骤之后，将血清加载到以所述抗原为捕获剂的第二亲和柱，以富集特异性针对所述抗原的抗体（也可参考 Nilsson P *et al* (2005) *Proteomics* 5: 4327-4337）。

附图简要说明

图 1 说明了在含有另外 94 种不同人类蛋白（两个重复）的蛋白微芯片上，针对 SATB2 片段 (SEQ ID NO: 1) 所生成的抗体的特异性。

图 2 显示了纯化的单特异性抗体的组织蛋白质印迹分析。从人类细胞系 RT-4（泳道 1）、EF0-21（泳道 2）和 A-431（泳道 3），以及从正常人类肝脏（泳道 4）和正常人类扁桃体（泳道 5）提取的总蛋白提取物。

图 3A 和 3B 显示了 (A) 正常大脑皮层和海马体以及 (B) 正常睾丸的细胞核中的 SATB2 的免疫组织化学染色。图 3C-3E 显示了来自正常受试者的 (C) 阑尾、(D) 结肠和 (E) 直肠的粘膜的腺细胞中 SATB2 的染色。图 3F 显示了结肠粘膜的染色的高倍放大图。

图 4A 显示了所有 11 个测试的结肠直肠癌样品（两个重复）中的 SATB2 的免疫组织化学染色。图 4B 显示了图 4A 所示 6 个癌症样品的高倍放大图。

图 5 显示了中间分化型结肠直肠癌的免疫组织化学染色。图 5A(左图)显示了有 SATB2 表达的切片,而图 5B(右图)显示了无 SATB2 表达的切片。显示的两肿瘤样品都是两份。

图 6 显示了基于 122 名被诊断有结肠直肠癌的受试者的免疫组织化学染色的存活分析结果。根据 SATB2 水平的高低对组织芯(tissue core)评分。实线:核分数>25%。虚线/点线:核分数<25%。组织取自(A)所有患者、(B)仅女性、(C)所有结节阴性患者和(D)结节阴性女性。

图 7 显示了对 SATB2 表达和常规结肠癌标记 CEA、CK20、CDX2、p53、Ki67 和细胞周期蛋白 B1 (CCNB1)表达进行免疫组织化学染色的 216 个癌组织的表达数据的分级聚类结果。

图 8 显示了基于 122 个被诊断有结肠直肠癌的受试者组织的免疫组织化学染色的 CK20 和 SATB2 表达的具体比较。根据 SATB2 水平的高低和 CK20 水平的高低对组织芯评分。

图 9 显示了基于 17 个被诊断有结肠直肠癌的受试者的结节转移的免疫组织化学染色的 CK20 和 SATB2 表达的比较。根据 SATB 核分数和 CK20 染色对组织芯评分。

实例

抗 SATB2 的单特异性抗体的生成和其在正常和癌症样品中检测 SATB2 的应用

1) 抗原的生成

a. 材料和方法

以人类基因组序列为模板,用分子信息学工具设计了由 Ensembl Gene ID ENSG00000119042 编码的目标蛋白的一个合适片段(Lindskog *et al* (2005) *Biotechniques* 38:723-727, Ensembl, www.ensembl.org)。设计由对应于 SATB2 蛋白 (SEQ ID NO: 2; Ensembl 登录号为 ENSP00000260926) 第 377-499 位氨基酸(SEQ ID NO: 1)的 123 个氨基酸组成的一个片段。以人类总 RNA 库组(Human Total RNA Panel IV, BD Biosciences Clontech)为模板,用含有 Platinum® Taq 酶的

Superscript™ One-Step RT-PCR 扩增试剂盒 (Invitrogen) 分离编码目标蛋白的多核苷酸, 该多核苷酸含有长 SATB2 基因转录物 (SEQ ID NO: 3; EnsEMBL 登录号为 ENST00000260926) 的第 1542-1910 位核苷酸。通过 PCR 扩增引物将侧翼限制酶切位点 NotI 和 AscI 引入到该片段中, 使其可框内 (in frame) 克隆到表达载体中 (正向引物: GTGTCCCAAGCTGTCTTTG, 反向引物: CTTGGCCCTTTTCATCTCC)。然后, 下游引物被生物素化以允许如前所述的固相克隆, 并且产生的生物素化 PCR 产物被固定在 Dynabeads M280 Streptavidin (Dynal Biotech) 上 (Larsson M *et al* (2000) J. Biotechnol. 80: 143-157)。通过 NotI-AscI (New England Biolabs) 消化将该片段从固体支持物上释放出来, 该片段被框内的连入到具有双亲和标签的 pAff8c 载体 (Larsson M *et al*, 见上文) 中, 所述双亲和标签包括一个用于固定化金属离子色谱 (IMAC) 纯化的六组氨酰标签和一个来自链球菌蛋白 G 的免疫增强白蛋白结合蛋白 (ABP) (Sjölander A *et al* (1997) J. Immunol. Methods 201: 115-123; Ståhl S *et al* (1999) Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis and Bioseparation (Fleckinger MC and Drew SW, eds) John Wiley and Sons Inc., New York, pp 49-63), 并转化到大肠杆菌 (*E. coli*) BL21 (DE3) 细胞 (Novagen) 中。按照厂商建议使用 TempliPhi DNA 测序扩增试剂盒 (GE Healthcare, Uppsala, Sweden), 对扩增的质粒 DNA 进行染料终止循环测序以确认克隆的序列。

将 1ml 含有该表达载体的 BL21 (DE3) 细胞的过夜培养物加入到添加有 5g/l 酵母提取物 (Merck KGaA) 和 50 mg/l 卡那霉素 (Sigma-Aldrich) 的 100 ml 30 g/l 的胰蛋白酶大豆肉汤 (Merck KGaA) 中进行接种, 所述过夜培养物在相同的培养基中。将该细胞培养物在 1 升摇瓶中于 37 ° C, 150 rpm 条件下培养, 直到 600nm 处的光密度达到 0.5-1.5。然后加入异丙基-β-D 硫乳吡喃糖苷 (isopropyl-β-D-thiogalactopyranoside) (Apollo Scientific) 至终浓度为 1mM 以诱导蛋白表达, 并且在 25 ° C 和 150 rpm 下继续培养过夜。以 2400 g 离心来收集细胞, 再用 5 ml 裂解缓冲液 (7 M 盐酸胍, 47 mM Na₂HPO₄, 2.65 mM NaH₂PO₄, 10 mM Tris-HCl, 100 mM NaCl, 20 mM β-巯基乙醇; pH=8.0) 重新悬浮沉淀物, 然后在 37 ° C, 150 rpm 下培养 2 小时。在以 35300 g 离心之后, 收集含有变性和溶解的基因产物的上清液。

通过 ASPEC XL4™ 上 (Gilson) 的一个自动蛋白纯化程序 (Steen J *et al* (2006) *Protein Expr. Purif.* 46:173-178), 使用加有 1 ml Talon® 金属 (Co²⁺) 亲和树脂 (BD Biosciences Clontech) 的色谱柱对具有 His₆ 标签的融合蛋白进行固定化金属离子亲和色谱 (IMAC) 纯化。用 20 ml 变性洗涤缓冲液 (6M 盐酸胍, 46.6 mM Na₂HPO₄, 3.4 mM NaH₂PO₄, 300 mM NaCl, pH 8.0-8.2) 平衡树脂。然后用最小量为 31.5 ml 的洗涤缓冲液冲洗树脂, 再用 2.5 ml 洗脱缓冲液 (6 M 尿素, 50 mM NaH₂PO₄, 100 mM NaCl, 30 mM 乙酸, 70 mM 乙酸钠, pH 5.0) 洗脱。洗脱下来的物质被分为 500、700 和 1300 μl 三个级分。700 μl 的级分含有抗原, 该级分以及收集的 500 和 1300 μl 的级分都储存备用。

用磷酸缓冲盐溶液 (PBS; 1.9 mM NaH₂PO₄, 8.1 mM Na₂HPO₄, 154 mM NaCl) 将该抗原级分稀释到尿素的终浓度为 1 M, 然后用截留分子量为 7500 Da 的 Vivapore 10/20 ml 浓缩器 (Vivascience AG) 进行一个浓缩步骤以提高蛋白浓度。根据厂商的建议, 用以一种牛血清白蛋白为标准品的二喹啉甲酸 (bicinchoninic acid, BCA) 微测定方法 (Pierce), 测定蛋白浓度。用一台 Bioanalyzer 仪器通过用蛋白 50 或 200 测定 (Agilent Technologies) 来分析该蛋白质量。

b) 结果

用特异性针对所述蛋白片段的引物通过 RT-PCR 将对应于 STAB2 基因长转录物 (SEQ ID NO: 3) 的第 1542-1910 位核苷酸的基因片段从人类 RNA 库中成功地分离出来, 所述基因片段编码由目标蛋白 SATB2 (SEQ ID NO: 2) 的第 377-499 位氨基酸组成的一个肽 (SEQ ID NO: 1)。但是, 与 EnSEMBL 中 ENSG00000119042 的序列相比, 该序列中有一个单核苷酸沉默突变。目标蛋白 (SEQ ID NO: 2) 的 123 个氨基酸片段 (SEQ ID NO: 1) 被设计为不含跨膜区域, 以保证其在大肠杆菌中有效表达; 并且被设计为不含任何信号肽, 因为它们在成熟蛋白中会被切除。此外, 所述蛋白片段被设计为由一段与其他人类蛋白低同源性的独特序列组成, 以尽可能减少所生成的亲和试剂的交叉反应; 并且被设计为具有合适的大小, 以允许构象表位的形成以及细菌系统内的有效克隆和表达。

鉴定出一个编码该正确氨基酸序列的克隆, 并且, 在大肠杆菌中表达后产生了正确大小的一个单一蛋白, 随后用固定化金属离子色谱纯化。将

洗脱出的样品稀释到尿素终浓度为 1M 之后并且浓缩样品到 1 ml, 测定该蛋白片段浓度为 7.4 mg/ml, 纯度分析结果为 98% 的纯度。

2) 抗体的生成

a) 材料和方法

按照国家标准 (Swedish permit no. A 84-02) 用上面得到的纯化的 SATB2 片段作为抗原来免疫家兔。用含 200 μg 抗原的弗氏完全佐剂对家兔进行肌肉免疫作为初次免疫, 然后用含 100 μg 抗原的弗氏不完全佐剂加强免疫三次, 每次之间间隔为四周。

用基于免疫亲和的三步法纯化来自该被免疫的动物的抗血清 (Agaton C *et al* (2004) J. Chromatogr. A 1043: 33-40; Nilsson P *et al* (2005) Proteomics 5: 4327-4337)。第一步, 用 10x PBS 将 10 ml 总抗血清缓冲至 1x PBS 的终浓度 (1.9 mM NaH_2PO_4 , 8.1 mM Na_2HPO_4 , 154 mM NaCl), 用 0.46 μm 孔径的过滤器 (Acrodisc®, Life Science) 过滤, 并将其加入到一个亲和柱中, 该柱中含有 5 ml 偶合有双亲和标签蛋白 His₆-ABP (一个六组氨酰基标签和一个白蛋白结合蛋白标签) 的 N-羟基琥珀酰亚胺活化的 Sepharose™ 4 Fast Flow 填料 (GE Healthcare), 该标签蛋白 His₆-ABP 是用 pAff8c 载体表达并用与上述纯化抗原蛋白片段同样的方法纯化。第二步, 将流过全柱的去除了抗双亲和标签 His₆-ABP 的抗体以 0.5 ml/分钟的流速载入 1 ml Hi-Trap NHS 活化的 HP 柱 (GE Healthcare) 中, 该柱偶合有用作免疫抗原的 SATB2 蛋白片段 (SEQ ID NO: 1)。按照厂商的建议, 将所述 His₆-ABP 蛋白和所述蛋白片段抗原偶合到 NHS 活化的基质上。用 1x PBST (1x PBS, 0.1 % 吐温 20, pH 7.25) 洗去没有结合的材料, 并且用低 pH 的甘氨酸缓冲液 (0.2 M 甘氨酸, 1 mM EGTA, pH 2.5) 洗脱捕获到的抗体。洗脱的抗体级分被自动收集, 并载入到两个串联的 5 ml HiTrap™ 脱盐柱 (GE Healthcare) 中, 这样在第三步可进行有效的缓冲液置换。第二和第三纯化步骤在 ÄKTExpress™ 平台 (GE Healthcare) 上进行。用添加有终浓度分别为 50 % 和 0.02 % 的甘油和 NaN_3 的 PBS 缓冲液洗脱该抗原选择性 (单特异性) 抗体 (msAb), 在 -20°C 长期储存 (Nilsson P *et al* (2005) Proteomics 5: 4327-4337)。

通过对自身抗原及蛋白质阵列设置中的其他 94 种人类蛋白片段的结合分析, 对亲和纯化的抗体级分的特异性和选择性进行分析 (Nilsson P *et*

a1 (2005) *Proteomics* 5:4327-4337)。将这些蛋白片段在含 0.1 M 尿素的 1x PBS (pH 7.4) 中稀释至 40 $\mu\text{g}/\text{ml}$ ，并且每种取 50 μl 转移到一个 96 孔点样板的孔中。用销环阵列点样仪 (pin-and-ring arrayer) (Affymetrix 427) 将这些蛋白片段点加并固定化在环氧树脂载片 (SuperEpoxy, TeleChem) 上。用 1x PBS 漂洗该载片 5 分钟并且封闭 (SuperBlock®, Pierce) 其表面 30 分钟。将一个粘性 16 孔硅酮掩罩 (Schleicher & Schuell) 加在玻璃上，然后加入该单特异性抗体 (用 1x PBST 以 1:2000 稀释至大约 50 ng/ml) 并置于摇床上孵育 60 分钟。使亲和标签特异的 IgY 抗体与单特异性抗体共同孵育以确定每个点中的蛋白量。用 1x PBST 和 1x PBS 漂洗该载片两次，每次 10 分钟。二抗 (偶联有 Alexa 647 的山羊抗-兔抗体和偶联有 Alexa 555 的山羊抗-鸡抗体, Molecular Probes) 用 1x PBST 以 1:60000 稀释至 30 ng/ml ，并孵育 60 分钟。在进行与首次孵育时的相同洗涤步骤之后，该载片被甩干并扫描 (G2565BA 阵列扫描仪, Agilent)，然后用图像分析软件 (GenePix 5.1, Axon Instruments) 对图像进行定量。结果在下面讨论并示于图 1 中。

此外，用蛋白质印迹法分析该亲和纯化的抗体的特异性和选择性。根据厂商的建议，通过于还原条件下在预制的 10-20% Criterion™ SDS-PAGE 梯度胶 (Bio-Rad Laboratories) 上分离来自选定人类细胞系和组织的总蛋白提取物，然后电转导至 PVDF 膜 (Bio-Rad Laboratories) 上，来进行蛋白质印迹法。此膜在室温下封闭 (5 % 奶粉, 1x TBST; 0.1 M Tris-HCl, 0.5 M NaCl, 0.5 % 吐温 20) 1 小时，与用封闭缓冲液以 1:500 稀释的一级亲和纯化抗体一起孵育，再用 TBST 漂洗。将偶联 HRP 的二抗 (猪抗兔免疫球蛋白/HRP, DakoCytomation) 用封闭缓冲液以 1:3000 稀释，并按照厂商的方法，用 Chemidoc™ CCD 相机 (Bio-Rad Laboratories) 和 SuperSignal® West Dura Extended Duration 底物 (Pierce) 进行化学发光检测。结果在下面讨论并被示于图 2。

b) 结果

已证明该多克隆抗体制剂的质量依赖于该抗体纯化的严格程度，并且先前已证明，针对非目标蛋白源的表位的抗体的去除对于避免与其他蛋白的交叉反应和背景结合而言是必需的 (Agaton C *et al* (2004) *J. Chromatogr. A* 1043: 33-40)。

因此, 进行蛋白微阵列分析, 确保通过去除针对该 His₆ 标签的抗体和针对该 ABP 标签的抗体, 已生成具有高特异性的单特异性多克隆抗体。随后在带有固定化的抗原的一个亲和柱上进行抗原选择性抗体的亲和捕获。

为对蛋白阵列的每个点的蛋白定量, 使用了一套联合使用一抗和二抗的双色染料标记系统。用 Alexa 555 荧光染料标记的山羊抗鸡的二抗来检测在鸡体内生成的标签特异的 IgY 抗体。用 Alexa 647 荧光标记的山羊抗兔抗体检测兔 msAb 与该阵列上其抗原的特异性结合。在图 1 中, 阵列结果以条柱显示, 对应从阵列上每点检测到的 Alexa 647 的荧光强度 (y 轴) 的量。每种蛋白片段都重复点样两份, 并且图中 x 轴上的每个条柱都代表一个蛋白点。蛋白阵列分析显示该亲和纯化的抗 SATB2 的单特异性抗体对正确的蛋白片段具有高度选择性, 并对该阵列上分析的所有其他蛋白片段呈现一个很低的背景。

蛋白质印迹法分析的结果 (图 2) 显示该抗体可在一个膀胱癌细胞系 (RT-4)、一个卵巢囊腺癌细胞系 (EO-21) 以及一个表皮样细胞系 (A-431) 中特异性地检测出一条大约 100 kDa 的条带 (1-3 泳道)。另外, 在肝脏和扁桃腺组织样品中观察到一个较弱的特异性条带 (4-5 泳道)。SATB2 的理论分子量为 82 kDa (按 SATB2 的氨基酸序列 SEQ ID NO: 2 计算), 考虑到该被分析的蛋白在分析条件下可能被糖基化或作其他修饰, 获得的结果与理论充分符合。

3) 免疫组织化学的组织分析

a) 材料和方法

总计用该单特异性抗体样品分析了 576 个含人类组织的石蜡块。作为供体块 (donor block) 用于制作组织微阵列 (TMA) 的所有组织都选自乌普萨拉大学医院病理科的档案, 并征得当地伦理委员会的同意。检查相应的组织切片以确诊并在供体块中选择代表区域。正常组织被定义为显微镜下正常 (无肿瘤), 其中绝大部分选自从手术切除肿瘤的附近区域采集的标本。检查癌症组织以诊断和分类。所有组织用福尔马林固定、石蜡包埋并进行诊断切片。

TMA 的制作基本上如前人所述 (Kononen J *et al* (1998) *Nature Med.* 4: 844-847; Kallioniemi OP *et al* (2001) *Hum. Mol. Genet.*

10: 657-662)。简言之, 在 TMA 受体块上制造一个洞。需要来自供体块的一个核心组织样品并且将该样品置于 TMA 受体块中。在一个 Beecher Instrument (ATA-27, Beecher Instruments, Sun Prairie, CA, USA) 的自动组织阵列点样仪中重复进行此步骤, 直到产生一个完整的 TMA 设计。TMA 受体块在 42℃ 温育 2 小时, 然后切片。

TMA 设计的重点在于从大范围的代表性正常组织中获得样品, 以及在于包括代表性癌组织。这在以前的文献 Kampf C *et al* (2004) Clin Proteomics 1: 285-300 中有详述。简言之, 样品选自 48 种正常组织和 20 种人类最常患的癌症类型。总计制造了 8 种不同设计的 TMA 块, 每种都包含 72 个 1mm 直径的组织芯。2 个 TMA 代表正常组织, 对应于 48 种从不同个体获得的不同的正常组织 (设三个重复)。剩下的 6 个 TMA 代表选自 20 种不同癌症类型的癌组织。对于 20 种癌症类型中的 17 种, 采集了各不相同的 12 个肿瘤, 并且对于剩下的 3 种癌症类型, 采集了各不相同的 4 个肿瘤, 所有样品对于同一肿瘤均设两个重复。用悬瀑切片机

(waterfall microtome) (Leica) 将所述 TMA 块切成 4 μm 的厚度并置于 SuperFrost® (Roche Applied Science) 玻片上用于免疫组织化学分析。

自动免疫组织化学分析的实现如前人所述 (Kampf C *et al* (2004) Clin. Proteomics 1: 285-300)。简言之, 玻片在 60℃ 孵育 45 分钟, 在二甲苯中脱蜡 (2 x 15 min), 然后在分级乙醇中水化。为回收抗原, 将玻片浸于 TRS (目标回收溶液, pH 6.0, DakoCytomation) 中, 并在 Decloaking chamber® (Biocare Medical) 中于 125℃ 下煮沸 4 分钟。将玻片置于 Autostainer® (DakoCytomation) 中, 并用 H_2O_2

(DakoCytomation) 初步阻断内源过氧化物酶。该一抗和山羊抗兔过氧化物酶偶联的 Envision® 都在室温下孵育 30 分钟。每步之间, 都用洗涤缓冲液 (DakoCytomation) 漂洗玻片。最后, 用二氨基联苯胺 (DakoCytomation) 作为发色团, 并用哈里斯 (Harris) 苏木精 (Sigma-Aldrich) 进行复染色。用 Pertex® (Histolab) 封固玻片。

用 ScanScope T2 自动切片扫描系统 (Aperio Technologies) 扫描所有免疫组织化学染色的 8 种不同 TMA 的切片。为了表示 8 种 TMA 的总含量, 生成了 576 张数字图像。扫描是 20 倍放大的。数字图像被分离并提取成独立标记的图像文件格式 (TIFF) 以储存原始数据。为了能够在一个基于网络的注释系统中处理图像, 将各图像从 TIFF 格式压缩为 JPEG 格式。免疫

组织化学染色的组织的所有图像均在显微镜下被人工评估并被经委员会认证的病理学专家注释或被专业人员注释（然后由病理专家验证）。用一种对免疫组织化学结果分类的简化方案对每种不同的正常组织以及癌组织进行注释。检查每种组织的代表性和免疫反应性。注释包括在每种正常组织类型中的不同组织特异性的细胞类型。对每种癌症，注释肿瘤细胞和间质。基本注解参数包括评估：i) 染色强度，ii) 染色细胞分数和 iii) 亚细胞定位（细胞核和/或细胞质/细胞膜）。按照在临床组织病理学诊断中使用的标准主观评估染色强度，并且结果作如下分类：阴性=没有免疫反应性，弱=微弱免疫反应性，中度=中等免疫反应性或者强=明显或强烈的免疫反应性。染色细胞分数被分为免疫反应活性细胞占代表性细胞群体的比例为<2%，2-25%，26-75%或>75%。基于强度和免疫反应活性细胞分数，对每个组织样品给出“染色分值”：0=阴性，1=弱，2=中度以及 3=强。

b) 结果

用针对该人类目标蛋白 SATB2 的一个重组蛋白片段产生的单特异性抗体进行组织分析的结果显示在几种正常组织和结肠直肠癌中（表 1-4 及图 3-4）有一种特殊的免疫反应性（深灰）。

表 1 显示了正常人类组织中的该 SATB2 蛋白表达模式。应用免疫组织化学和 TMA 技术，筛选了代表 48 种不同正常组织类型的 144 个点（直径 1 mm）的 SATB2 表达。表 1 示出了在不同组织中的表达水平。在远端胃肠道组织和脑中两个区域组织中发现强表达（染色分值为 3）。在睾丸和附睾中检测到中度（染色分值为 2）表达水平。局灶淋巴细胞显示出中度或弱（染色分值为 1）表达。所有其他细胞和组织为阴性（染色分值为 0）。N.R. 指无代表性组织。SATB2 也在某些神经组织和睾丸中表达。

表 1: 正常组织中 SATB2 的表达模式

组织类型	细胞类型	染色分值
肾上腺	皮质细胞	0
	髓质细胞	N. R.
阑尾	腺细胞	3
	淋巴组织	0
骨髓	骨髓 poetic 细胞	0
乳腺	腺细胞	0
支气管	表面上皮细胞	0
小脑	颗粒层细胞	0
	分子层细胞	0
	浦肯野细胞	0
大脑皮质	神经元细胞	3
	非神经元细胞	0
子宫颈、子宫	腺细胞	0
	表面上皮细胞（鳞状）	N. R.
结肠	腺细胞	3
十二指肠	腺细胞	0
子宫内膜 1	子宫内膜基质/细胞外基质（ECM）细胞	0
	子宫肌层/细胞外基质细胞	0
	腺细胞	0
子宫内膜 2	子宫内膜基质/细胞外基质细胞	0
	子宫肌层/细胞外基质细胞	0
	腺细胞	0
附睾	腺细胞	2
食道	表面上皮细胞	0
输卵管	腺细胞	0
胆囊	腺细胞	0

心肌	肌细胞	0
海马	神经元细胞	3
	非神经元细胞	0
肾	肾小球细胞	0
	小管细胞	0
侧脑室	神经元细胞	0
	非神经元细胞	0
肝	胆管细胞	0
	肝细胞	0
肺	肺泡细胞	0
	巨噬细胞	0
淋巴结	滤泡细胞（皮质）	0
	非滤泡细胞（副皮质）	2
鼻咽	表面上皮细胞	0
口腔粘膜	表面上皮细胞	0
卵巢	滤泡细胞	0
	卵巢基质细胞	0
胰	胰腺外分泌细胞	0
	胰岛细胞	0
甲状旁腺	腺细胞	0
胎盘	蜕膜细胞	0
	滋养层细胞	0
前列腺	腺细胞	0
直肠	腺细胞	3
唾液腺	腺细胞	0
精囊	腺细胞	0
骨骼肌	肌细胞	0
皮肤	附属细胞（adnexal cell）	0
	表皮细胞	0
小肠	腺细胞	0
平滑肌	平滑肌细胞	0
软组织 1	间充质细胞	0

软组织 2	间充质细胞	0
脾	红髓细胞	0
	白髓细胞	0
胃 1	腺细胞	0
胃 2	腺细胞	0
睾丸	输精管 (ductus seminiferous) 细胞	2
	睾丸间质细胞	0
甲状腺	腺细胞	0
扁桃体	滤泡细胞 (皮质)	0
	非滤泡细胞 (副皮质)	1
	表面上皮细胞	0
膀胱	表面上皮细胞	0
阴道	表面上皮细胞	0
阴门/肛周皮肤	表面上皮细胞	0

图 3A 显示了一个显微镜放大图，其表示大脑皮层和海马体的神经元细胞核为阳性（深灰）。周围组织和神经胶质细胞为阴性（浅灰）。睾丸组织切片显示输精管中为中度和大部分核阳性（深灰）（图 3B）。

在关于本发明的组织学阵列中具体发现了阑尾（图 3C）、结肠（图 3D）和直肠（图 3E）中的粘膜腺细胞为明显、强烈的核阳性（深灰）。注意其他细胞类型如炎性细胞、内皮细胞的阴性染色（浅灰），也存在于粘膜中。图 3F 示出了结肠粘膜的两个高倍放大图，显示出所有腺细胞具有该 SATB2 蛋白的强核表达（深灰）。

表 2 显示了 216 个不同癌症组织中 SATB2 的表达水平。作为代表的所有 11 个结肠直肠癌都显示阳性，并且其中 8 个为强表达。图 4A 示出了表示分析的 11 个结肠直肠癌病例的免疫组织化学染色组织切片的低倍显微镜放大图，同时 4B 示出了 6 个结肠直肠癌的代表性区域的高倍放大图。大部分癌细胞显示强核染色（深灰），表明 SATB2 相对周围含有正常细胞的阴性（浅灰）组织呈高水平表达。

表 2: SATB2 在 20 种癌症中的表达模式													
癌症类型	受试者编号												
	1	2	3	4	5	6	7	8	9	10	11	12	
乳腺癌	1	0	0	0	0	0	0	0	0	0	0	0	N. R.
子宫颈癌	1	0	0	0	0	0	0	0	0	0	0	0	0
结肠直肠癌	3	3	3	3	3	3	3	3	2	1	1	1	N. R.
子宫内膜癌	0	0	0	0	0	0	0	0	0	0	0	0	0
头颈癌	0	0	0	0									
肾癌	3	1	1	0	0	0	0	0	0	0	0	0	N. R.
肝癌	0	0	0	0	0	0	0	0	0	0	0	0	0
肺癌	2	0	0	0	0	0	0	0	0	0	0	0	0
恶性类癌	0	0	0	0									
恶性胶质瘤	2	1	0	0	0	0	0	0	0	0	0	0	0
恶性淋巴瘤	0	0	0	0	0	0	0	0	0	0	0	0	0
恶性黑素瘤	0	0	0	0	0	0	0	0	0	N. R.	N. R.	N. R.	N. R.
卵巢癌	1	0	0	0	0	0	0	0	0	0	0	0	0
胰癌	0	0	0	0	0	0	0	0	0	0	0	0	N. R.
前列腺癌	1	0	0	0	0	0	0	0	0	0	0	0	0
皮肤癌	0	0	0	0	0	0	0	0	0	0	0	0	0
胃癌	2	2	0	0	0	0	0	0	0	0	0	0	0
睾丸癌	2	1	0	0	0	0	0	0	0	0	0	0	N. R.
甲状腺癌	0	0	0	0									
尿道上皮细胞癌	2	2	0	0	0	0	0	0	0	0	0	0	N. R.

4) 结肠直肠癌 TMA

a) 材料和方法

取自 122 名在 1999 到 2002 年被诊断有结肠直肠癌的患者（63 名女性和 59 名男性）的福尔马林固定石蜡包埋的存档组织，是从瑞典马尔默

大学医院病理科收集的。患者的年龄中位数为 75 (32-88) 岁。39 个肿瘤处于 Dukes A 期, 42 个处于 Dukes B 期, 41 个处于 Dukes C 期。死亡日期的信息得自地区所有患者的死因登记。伦理许可得自地方伦理委员会。

所有 122 个结肠直肠癌病例都基于用苏木精和曙红染色的切片重新进行组织病理学评价。从每个病例侵袭性癌症的代表区域取 2 个 1.0 mm 的芯样以构建 TMA。TMA 的制备和自动免疫组织化学分析的进行如上面第 3 部分所述, 所用 SATB2 抗体的制备如上面第 2 部分所述。

组织注释的进行基本上如上面第 3 部分所述, 除了染色强度只划分为阴性 (无或弱免疫反应性) 或阳性 (中度或强反应性)。然后计算在细胞核中表现阳性染色强度的细胞分数, 产生被称为“分数分值”的样品值。因此, “分数分值”对应样品中根据本部分中的定义表现为阳性染色强度的细胞百分比。

基于所有不同层级的存活趋势, 建立一个二元变量用于进一步的统计分析, 将高/阳性的 SATB2 表达定义为阳性细胞核 $> 25\%$, 将低/阴性的 SATB2 表达定义为阳性细胞核 $< 25\%$ 。然后基于该分数分值将样品分为两组, 划分标准为分数分值为 25% 。因此, 若组织样品 (芯) 中的根本没有信号或细胞的阳性染色强度 $< 25\%$, 则该样品属于“ $< 25\%$ ”组; 而若细胞阳性染色强度 $> 25\%$, 则该样品属于“ $> 25\%$ ”组。

上述样品分类被用于基于 Kaplan-Meier 估计量的总存活分析, 且时序检验 (the log-rank test) 也被用于比较不同层级的存活情况。所有统计检验都是双侧检验, 且 p 值 $< 0.05\%$ 时被认为具有显著性。所有计算都是用统计软件包 SPSS 12.0 (SPSS Inc. Illinois, USA) 进行。

b) 结果

对 122 个结肠直肠癌的组织微阵列分析显示, 99 个肿瘤 (81 %) 为 SATB2 阳性。令人惊讶地, 不能通过常规切片或组化染色来预测 SATB2 的低表达或无表达, 如图 5 所示。两个肿瘤样品 (一式两份的切片) 都被诊断为中间分化型直肠腺癌。图 5A 显示出一个 SATB2 强表达的切片; 而在图 5B 中, 样品无 SATB2 表达。

存活分析的结果示于图 6, 该图表显示了不同患者组随时间的累计存活。基于全部患者的存活分析显示了一个趋势 ($p = 0.14$), 即患有 SATB2 低表达肿瘤的患者的总存活时间 (OS) 更短 (图 6A)。还考察了 SATB2 表

达与临床病理学变量性别及 Dukes 期之间的关系。患有 SATB2 低表达肿瘤的女性患者 ($n = 63$) 相对全部受试者其 OS 更短的趋势增加 ($p = 0.11$) (图 6B)。图 6C 显示了在结节阴性 (Dukes A 和 B 期) 的患者 ($n = 80$) 中观察到相似的趋势 ($p = 0.10$)。在结节阴性的女性患者 ($n = 44$) 的亚群中, 此趋势很显著 ($p = 0.04$) (图 6D)。

提供存活分析数据的其他方法可以是应用上面第 3 部分所述的“染色分值”。这种情况下, 分值为 0 和 1 的样品会被定义为 SATB2 低表达, 分值为 2 和 3 的样品会被定义为 SATB2 高表达。预期其结果类似于图 6 所显示。

5) TMA 数据的定量图像分析

为了进行定量表达测量, 使用 Aperio ScanScope CS Slide Scanner (Aperio Technologies, Vista, CA, USA) 系统来捕获按上面第 4 部分所述制备的杂交的 TMA 玻片的数字图像。进行 20 倍放大扫描且图像被储存为多层 TIFF 格式。使用 ImageScope (Aperio) 来浏览这些适合分析的数字图像。使用 TMA Lab (Aperio) 对这些图像去排列以观察单个组织芯。首先, 使用 Color Deconvolution 算法 (Aperio) 将每个图像分离到三个通道, 即红、绿和蓝 (RGB)。这使得每种染色可被分别测量, 并因此可从二氨基联苯胺铬精 (chromogene) 染色中减去苏木精复染色。

然后, 将许多不同的算法用于对细胞核、细胞质或细胞膜的染色定量。IHC Nuclear 算法 (Aperio) 被用于定量 SATB2 的核染色。基于强度来识别核。使用一种边界阈值法来识别核的边界, 这种方法可根据边界像素的平均值来自动调整阈值。从 Aperio Technologies 可得到所有算法的完整介绍。TMA 上每个芯的伪色补偿图像被生成和评估以确认每种算法的精确性。

Nuclear 算法的输出值是阳性核的一个百分比和 TMA 玻片上每个芯的一个核 RGB 强度值。通过用每个芯的核 RGB 强度乘以阳性核的百分比来计算每个组织芯上 SATB2 表达水平的一个自动分值 (AS)。对于例如上面第 4 部分所述的癌症 TMA 样品和根据 Kaplan-Meier 方法估计的总存活进行 AS 分析。时序检验用于比较不同层级的存活。统计计算是用统计软件包 SPSS 12.0 (SPSS Inc. Illinois, USA) 进行的。

6) 聚类分析

a) 材料和方法

为了研究已知的结肠癌标记和 SATB2 之间在蛋白表达上的一致性,进行分级聚类分析。聚类是一种评估初始挖掘步骤的数据的趋势和结构的合适方法。只浏览数据集时不明显的分组和分类可以容易地通过非监督式方法例如分级聚类检测出来。在生命科学中,聚类已经十分广泛地用于 RNA 转录分析,如微阵列数据。

第 3 部分所述的 6 种癌症 TMA 被再次使用,即共计 216 个癌组织样品。除了按上面第 2 部分制备的识别 SATB2 的抗体,针对已知标记 CEA (DAKO, Glostrup, Denmark)、CK20 (DAKO, Glostrup, Denmark)、CDX2 (Novocastra, Newcastle upon Tyne, UK)、p53 (DAKO, Glostrup, Denmark)、Ki67 (DAKO, Glostrup, Denmark) 和 Cyclin B1 (Transduction laboratories, Lexington, USA) 的抗体也通过一种自动免疫组织化学方法被检测,该方法类似于第 3 部分中所述。

如第 3 部分所述由病理学家注释 TMA,并对每个芯给出 0-3 范围的一个染色分值,其中 3 是强(黑)染色而 0 是无(白)染色。用统计计算语言 R 进行聚类分析。将聚类算法用于数据矩阵的两个维度,即组织和抗体。聚类过程总计使用了 7 个抗体和 216 个组织。8 个组织由于没有可定量表达水平的图像而被去掉。使用自上而下的分级方法进行群聚分析,该方法使用了基于欧几里得距离公制的平均集聚,其中群之间的距离在每个阶段都使用 Lance-Williams 差异更新公式根据平均连接重新计算。聚类所用算法决定了子树(sub-tree)从而使更紧的群示于每个节点的左侧。

为了进一步研究 SATB2 和 CK20 的异同,使用具有第 4 部分所述的 122 个癌组织芯的癌组织 TMA 进行免疫组织化学分析。用按第 2 部分所述制备的 SATB2 抗体和获自 DAKO (Glostrup, Denmark) 的 CK20 抗体对 TMA 染色。根据第 4 部分中定义的“分数分值”注释两个 TMA,然后比较。

b) 结果

在 216 个不同肿瘤中检查 SATB2 与已知标记 CEA、CK20、CDX2、p53、Ki67 和 Cyclin B1 相比较作为结肠直肠癌标记的特异性。对这 7 种不同蛋白的表达谱的数据的分级聚类得出了热图和伴随的树状图,如图 7 所示。从肿瘤的热图和树状图清楚地看出,大部分结肠直肠癌形成一个群集,该群集基于 SATB2、CK20、CDX2 和 CEA 高表达水平以最高水平被分离。另

外,分析显示 SATB2 群集以及 CK20 和 CDX2 群集都具有比其他显示出更一般的表达模式的被测标记更特异的表达。在 8 个结肠直肠癌的群集中,也有强 SATB2 阳性的一个宫颈腺癌和一个胆管细胞型肝癌病例。除该群集外,有三个结肠直肠癌为 SATB2 阴性。引人注意的是, SATB2 表达模式与 CK20 的表达没有明显的关联,因此能够在结肠直肠癌鉴定中作为 CK20 的补充。

在第 4 部分所描述的 122 个癌组织的 TMA 上 对 SATB2 和 CK20 进行了更详细的分析。其中,按照分值分值 $> 25\%$ 的标准,单独的 CK20 确认为 122 个结肠直肠癌的 86% (105/122),且单独的 SATB2 确认为 81% (99/122) (图 8)。引人注意的是,结合两种标记的染色数据, 93% (113/122) 的结肠直肠癌明显地对两种标记中的一种或两种呈阳性。只有 5 名患者完全缺乏 CK20 或 SATB2 的表达。这个信息在诊断癌症时——更具体来说在试图鉴定转移时——是有价值的,因为癌症的一个临床常见问题是患者出现来源未知的转移。因此,通过结合得自 CK20 和 SATB2 的信息,患者可更容易地获得结肠直肠腺癌的精确诊断。

另外,还分析了 SATB2 和 CK20 在 17 个结肠直肠癌患者的淋巴结转移中的表达。按照染色分值为 2 或 3,单独的 CK20 确认为 88% (15/17) 的转移的来源。按照分数分值 $> 25\%$,单独的 SATB2 确认为 82% (14/17) 的转移的来源 (图 9)。结合两种标记的染色数据,确认了 94% (16/17) 的转移的来源。这进一步支持了在确定转移是否源自结肠直肠癌时,关于 SATB2 和 CK20 表达的信息是人们希望得到的。

本发明实施方案的逐项列表

下面是本发明实施方案的非限制性逐项列表,目的是提供关于本发明某些方面的各个特征和组合的进一步的信息。

1. 诊断结肠直肠癌的方法,包含检测 SATB2 蛋白的一个步骤。
2. 根据第 1 项的方法,其中所述 SATB2 蛋白的氨基酸序列包括选自如下的序列:

- i) SEQ ID NO: 1; 和
- ii) 一个与 SEQ ID NO: 1 至少 85%相同的序列。

3. 根据前述项中任一项的方法, 其中所述 SATB2 蛋白的氨基酸序列包括选自如下的序列:

- i) SEQ ID NO: 2; 和
- ii) 一个与 SEQ ID NO: 2 至少 85%相同的序列。

4. 根据前述项的任一项的方法, 包括下列步骤:

- a) 提供来自怀疑患有结肠直肠癌的患者的样品;
- b) 对所述样品应用可检测的亲配体, 所述亲配体可选择性地与待检测的 SATB2 蛋白相互作用, 所述应用是在允许亲配体与存在于样品中的任何 SATB2 蛋白结合的条件下施行的。
- c) 除去未结合的亲配体; 并且
- d) 检测与该样品保持关联的任何亲配体。

5. 根据第 4 项的方法, 其中所述样品是一种体液样品。

6. 根据第 5 项的方法, 其中所述体液选自下组, 其构成为血液、血浆、血清、脑液、尿液、精液和渗出液。

7. 根据第 4 项的方法, 其中所述样品是一种粪便样品。

8. 根据第 4 项的方法, 其中所述样品是一种组织样品。

9. 根据第 4 项的方法, 其中所述样品是一种细胞学样品。

10. 根据第 4-9 项的任一项的方法, 其中所述可检测的亲配体选自抗体、其片段和其衍生物。

11. 根据第 4-9 项的任一项的方法, 其中所述可检测的亲配体是源自一种支架的蛋白配体, 该支架选自下组, 其构成为葡萄球菌 A 蛋白及其结构

域、脂质运载蛋白、锚蛋白重复结构域、纤维素结合结构域、 γ 晶体、绿色荧光蛋白、人类细胞毒性T淋巴细胞相关抗原 4、蛋白酶抑制剂、PDZ结构域、肽适体、葡萄球菌核酸酶、淀粉酶抑肽、纤维连接蛋白 III 型结构域和锌指结构。

12. 根据第 4-9 项的任一项的方法,其中所述可检测的亲配体是一种寡核苷酸分子。

13. 根据 4-12 项的任一项的方法,其中所述可检测的亲配体含有一个标签,该标签选自下组,其构成为荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

14. 根据第 4-12 项的任一项的方法,其中所述可检测的亲配体是用一种能够识别所述可检测的亲配体的二级亲配体检测的。

15. 根据第 14 项的方法,其中所述能够识别所述可检测的亲配体的二级亲配体含有一个标签,该标签选自下组,其构成为荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

16. 执行第 1-15 项的任一项的方法的试剂盒,所述试剂盒包括:

- a) 能够选择性地与 SATB2 蛋白相互作用的可检测的亲配体; 和
- b) 用于检测亲配体存在的必需的试剂。

17. 根据第 16 项的试剂盒,其中所述可检测的亲配体选自抗体、其片段和其衍生物。

18. 根据第 16 项的试剂盒,其中所述可检测的亲配体是源自一种支架的蛋白配体,该支架选自下组,其构成为葡萄球菌 A 蛋白及其结构域、脂质运载蛋白、锚蛋白重复结构域、纤维素结合结构域、 γ 晶体、绿色荧光蛋白、人类细胞毒性T淋巴细胞相关抗原 4、蛋白酶抑制剂、PDZ 结构域、肽适体、葡萄球菌核酸酶、淀粉酶抑肽、纤维连接蛋白 III 型结构域和锌指结构。

19. 根据第 16 项的试剂盒, 其中所述可检测的亲配体是一种寡核苷酸分子。

20. 根据第 16-19 项的任一项的试剂盒, 其中所述可检测的亲配体含有一个标签, 该标签选自下组, 其构成为荧光染料和金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

21. 根据第 16-19 项的任一项的试剂盒, 其中所述用于检测所述亲配体存在所必需的试剂包括能够识别所述可检测的亲配体的二级亲配体。

22. 根据第 21 项的试剂盒, 其中所述能够识别所述可检测的亲配体的二级亲配体含有一个标签, 该标签选自下组, 其构成为荧光染料或金属、发色团染料、化学发光化合物和生物发光蛋白、酶、放射性同位素和颗粒。

23. SATB2 蛋白作为一种结肠直肠癌诊断标记的应用。

24. SATB2 蛋白或其抗原活性的片段在制造用于诊断结肠直肠癌的诊断试剂中的应用。

25. 根据第 23 和 24 的任一项的应用, 其中所述 SATB2 蛋白的氨基酸序列包括选自如下的序列:

- i) SEQ ID NO: 1; 和
- ii) 一个与 SEQ ID NO: 1 至少 85% 相同的序列。

26. 根据第 23 和 24 项的任一项的应用, 其中所述 SATB2 蛋白的氨基酸序列包括选自如下序列:

- i) SEQ ID NO: 2; 和
- ii) 一个与 SEQ ID NO: 2 至少 85% 相同的序列。

27. 能够选择性地与 SATB2 蛋白相互作用的亲和配体,所述亲和配体是一种抗体或其片段或其衍生物。

28. 根据第 27 项的亲和配体,所述亲和配体可从一种包括用一种蛋白免疫动物的步骤的方法中获得,所述蛋白的氨基酸序列包含 SEQ ID NO: 1 序列。

29. 根据第 27-28 项的任一项的亲和配体作为诊断试剂的应用。

30. 根据第 27-28 项的任一项的亲和配体在制造用于诊断结肠直肠癌的诊断试剂中的应用。

<110> 阿特拉斯抗体有限公司

<120> SATB2蛋白作为结肠直肠癌标记的应用

<130> 21031658

<150> EP06114954.8

<151> 2006-06-02

<150> US 60/816 613

<151> 2006-06-27

<160> 3

<170> PatentIn version 3.3

<210> 1

<211> 123

<212> PRT

<213> 人类 (Homo sapiens)

<400> 1

Val Ser Gln Ala Val Phe Ala Arg Val Ala Phe Asn Arg Thr Gln Gly
1 5 10 15

Leu Leu Ser Glu Ile Leu Arg Lys Glu Glu Asp Pro Arg Thr Ala Ser
20 25 30

Gln Ser Leu Leu Val Asn Leu Arg Ala Met Gln Asn Phe Leu Asn Leu
35 40 45

Pro Glu Val Glu Arg Asp Arg Ile Tyr Gln Asp Glu Arg Glu Arg Ser
50 55 60

Met Asn Pro Asn Val Ser Met Val Ser Ser Ala Ser Ser Ser Pro Ser
65 70 75 80

Ser Ser Arg Thr Pro Gln Ala Lys Thr Ser Thr Pro Thr Thr Asp Leu
85 90 95

Pro Ile Lys Val Asp Gly Ala Asn Ile Asn Ile Thr Ala Ala Ile Tyr
100 105 110

Asp Glu Ile Gln Gln Glu Met Lys Arg Ala Lys
115 120

<210> 2

<211> 733

<212> PRT

<213> 人类 (Homo sapiens)

<400> 2

Met Glu Arg Arg Ser Glu Ser Pro Cys Leu Arg Asp Ser Pro Asp Arg
1 5 10 15

Arg Ser Gly Ser Pro Asp Val Lys Gly Pro Pro Pro Val Lys Val Ala

20	25	30
Arg Leu Glu Gln Asn Gly Ser Pro Met Gly Ala Arg Gly Arg Pro Asn 35 40 45		
Gly Ala Val Ala Lys Ala Val Gly Gly Leu Met Ile Pro Val Phe Cys 50 55 60		
Val Val Glu Gln Leu Asp Gly Ser Leu Glu Tyr Asp Asn Arg Glu Glu 65 70 75 80		
His Ala Glu Phe Val Leu Val Arg Lys Asp Val Leu Phe Ser Gln Leu 85 90 95		
Val Glu Thr Ala Leu Leu Ala Leu Gly Tyr Ser His Ser Ser Ala Ala 100 105 110		
Gln Ala Gln Gly Ile Ile Lys Leu Gly Arg Trp Asn Pro Leu Pro Leu 115 120 125		
Ser Tyr Val Thr Asp Ala Pro Asp Ala Thr Val Ala Asp Met Leu Gln 130 135 140		
Asp Val Tyr His Val Val Thr Leu Lys Ile Gln Leu Gln Ser Cys Ser 145 150 155 160		
Lys Leu Glu Asp Leu Pro Ala Glu Gln Trp Asn His Ala Thr Val Arg 165 170 175		
Asn Ala Leu Lys Glu Leu Leu Lys Glu Met Asn Gln Ser Thr Leu Ala 180 185 190		
Lys Glu Cys Pro Leu Ser Gln Ser Met Ile Ser Ser Ile Val Asn Ser 195 200 205		
Thr Tyr Tyr Ala Asn Val Ser Ala Thr Lys Cys Gln Glu Phe Gly Arg 210 215 220		
Trp Tyr Lys Lys Tyr Lys Lys Ile Lys Val Glu Arg Val Glu Arg Glu 225 230 235 240		
Asn Leu Ser Asp Tyr Cys Val Leu Gly Gln Arg Pro Met His Leu Pro 245 250 255		
Asn Met Asn Gln Leu Ala Ser Leu Gly Lys Thr Asn Glu Gln Ser Pro 260 265 270		
His Ser Gln Ile His His Ser Thr Pro Ile Arg Asn Gln Val Pro Ala 275 280 285		
Leu Gln Pro Ile Met Ser Pro Gly Leu Leu Ser Pro Gln Leu Ser Pro 290 295 300		

Gln Leu Val Arg Gln Gln Ile Ala Met Ala His Leu Ile Asn Gln Gln
 305 310 315 320

Ile Ala Val Ser Arg Leu Leu Ala His Gln His Pro Gln Ala Ile Asn
 325 330 335

Gln Gln Phe Leu Asn His Pro Pro Ile Pro Arg Ala Val Lys Pro Glu
 340 345 350

Pro Thr Asn Ser Ser Val Glu Val Ser Pro Asp Ile Tyr Gln Gln Val
 355 360 365

Arg Asp Glu Leu Lys Arg Ala Ser Val Ser Gln Ala Val Phe Ala Arg
 370 375 380

Val Ala Phe Asn Arg Thr Gln Gly Leu Leu Ser Glu Ile Leu Arg Lys
 385 390 395 400

Glu Glu Asp Pro Arg Thr Ala Ser Gln Ser Leu Leu Val Asn Leu Arg
 405 410 415

Ala Met Gln Asn Phe Leu Asn Leu Pro Glu Val Glu Arg Asp Arg Ile
 420 425 430

Tyr Gln Asp Glu Arg Glu Arg Ser Met Asn Pro Asn Val Ser Met Val
 435 440 445

Ser Ser Ala Ser Ser Ser Pro Ser Ser Ser Arg Thr Pro Gln Ala Lys
 450 455 460

Thr Ser Thr Pro Thr Thr Asp Leu Pro Ile Lys Val Asp Gly Ala Asn
 465 470 475 480

Ile Asn Ile Thr Ala Ala Ile Tyr Asp Glu Ile Gln Gln Glu Met Lys
 485 490 495

Arg Ala Lys Val Ser Gln Ala Leu Phe Ala Lys Val Ala Ala Asn Lys
 500 505 510

Ser Gln Gly Trp Leu Cys Glu Leu Leu Arg Trp Lys Glu Asn Pro Ser
 515 520 525

Pro Glu Asn Arg Thr Leu Trp Glu Asn Leu Cys Thr Ile Arg Arg Phe
 530 535 540

Leu Asn Leu Pro Gln His Glu Arg Asp Val Ile Tyr Glu Glu Glu Ser
 545 550 555 560

Arg His His His Ser Glu Arg Met Gln His Val Val Gln Leu Pro Pro
 565 570 575

Glu Pro Val Gln Val Leu His Arg Gln Gln Ser Gln Pro Ala Lys Glu
580 585 590

Ser Ser Pro Pro Arg Glu Glu Ala Pro Pro Pro Pro Pro Pro Thr Glu
595 600 605

Asp Ser Cys Ala Lys Lys Pro Arg Ser Arg Thr Lys Ile Ser Leu Glu
610 615 620

Ala Leu Gly Ile Leu Gln Ser Phe Ile His Asp Val Gly Leu Tyr Pro
625 630 635 640

Asp Gln Glu Ala Ile His Thr Leu Ser Ala Gln Leu Asp Leu Pro Lys
645 650 655

His Thr Ile Ile Lys Phe Phe Gln Asn Gln Arg Tyr His Val Lys His
660 665 670

His Gly Lys Leu Lys Glu His Leu Gly Ser Ala Val Asp Val Ala Glu
675 680 685

Tyr Lys Asp Glu Glu Leu Leu Thr Glu Ser Glu Glu Asn Asp Ser Glu
690 695 700

Glu Gly Ser Glu Glu Met Tyr Lys Val Glu Ala Glu Glu Glu Asn Ala
705 710 715 720

Asp Lys Ser Lys Ala Ala Pro Ala Glu Ile Asp Gln Arg
725 730

<210> 3
<211> 5318
<212> DNA
<213> 人类 (Homo sapiens)

<400> 3
gagagagaga gagagagact gagaaagtcg aactcgagct taggtaacca acaaacccaa 60
ttactaggct ttattgatac taaaaggaaa agttcccctc gtaaaatttc tctgagctta 120
ttagttatatt aaatcagccc cagggttgga cgcgtgacgc cgaaggcacc ttctgtcggc 180
tcctggaggt gagcagccag cgtccctgcg ccgaccaca gccctcgcct ccgtcctggg 240
cacttggggc gaaggcgggc gcagcgagcc gagaagtact gaactcctaa taccaccaat 300
tcttctaagt tcctggacat tgatccggag gaggattcgc agttcaacat caaggtccct 360
gtgcgtttta ttgcgacctg ccggtgggaa ctttgtctcc gagtcggagc agcatggagc 420
ggcggagcga gagcccgtgt ctgcgggaca gccccgaccg gcggagcggc agcccgagc 480
tcaaggggcc tccccagtg aaggtggccc ggctggagca gaacggcagc cccatgggag 540
cccgcgggag gcccacggc gccgtggcca aggccgtggg aggtttgatg attcctgtct 600
tttgtgtcgt ggagcagttg gacggctctc ttgaatatga caacagagaa gaacacgccg 660

agtttgcctt ggtgcggaaa gatgtgcttt ttagccagct ggtggagact gcgctcctgg	720
ccctggggta ttctcacagc tctgcggccc aggccaagg aataatcaag ctgggaaggt	780
ggaaccctct cccctcagc tatgtgacag atgcaccca cgcgacagtg gccgacatgc	840
tacaagatgt ctatcatgtt gtgacgttga aaatccaatt acaaagttgt tcaaagttgg	900
aagacttgcc tgcggagcag tggaacatg ccacagtcg caatgcctta aaggaactgc	960
tcaaagagat gaaccagagc acattagcca aagaatgccc tctctcccag agtatgattt	1020
catccattgt aaatagcaca tattatgcca atgtgtcagc aaccaagtgc caggagtgtg	1080
ggagatggta taaaaagtac aagaagatta aagtggaaag agtggaacga gaaaaccttt	1140
cagactattg tgttctgggc cagcgtccaa tgcatctacc aaatatgaac cagctggcat	1200
ccctggggaa aaccaacgaa cagtctctc acagccaaat tcaccacagt actccaatcc	1260
gaaaccaagt gcccgcatca cagcccatca tgagccctgg tcttctttct cccagctta	1320
gtccacaact tgtaaggcaa caaatagcca tggcccatct gataaaccaa cagattgccg	1380
ttagccggct cctggctcac cagcatctc aagccatcaa ccagcagttc ctgaaccatc	1440
caccatccc cagagcagtt aagccagagc caaccaactc ttccgtggaa gtctctccag	1500
atatctacca gcaagtcaga gatgagctga agaggccag tgtgtccaa gctgtctttg	1560
caagagtggc attcaaccgc acacagggat tgtgtctga gattctgcgt aaggaagaag	1620
accctcgac agcctctcag tctcttctag taaacctgag ggccatgcag aatttctca	1680
atctgccaga agtggagcga gatgcacatc accaggatga gagggagcgg agcatgaatc	1740
ccaatgtgag catggtctcc tcggcctcca gcagtcaccag ctctcccga acccctcagg	1800
ccaaaacctc gacaccgaca acagacctcc ctattaaggt ggacggcgcc aacatcaaca	1860
tcacagctgc catttatgac gagatccaac aggagatgaa aagggccaaag gtgtctcaag	1920
ccctgtttgc caaagtggct gcaaataaaa gtcagggtg gctgtgtgaa ctgctccgct	1980
ggaaggagaa cccaagccca gaaaaccgca ccctctggga aaacctctgt accatccgtc	2040
gcttctgaa ctttcccag catgagaggg atgtcatcta tgaggaggag tcaaggcatc	2100
accacagcga acgcatgcaa cagtggtcc agcttcccc tgagccgggtg caggctctc	2160
atagacagca gtctcagcca gccaaaggaga gttccccctc cagagaagaa gcgcctcccc	2220
cacctctcc gactgaagac agttgtgcca aaaagccccg gtctcgaca aagatctcct	2280
tagaagccct ggggatctc caaagcttta ttcattgatgt aggcctgtac ccagaccagg	2340
aagccatcca cactctttcg gctcagctgg atctcccaa acacaccatc atcaagttct	2400
tccagaacca gcggtaccac gtgaagcacc acgggaagct gaaagagcac ctgggctccg	2460
cggtggacgt ggctgaatat aaggacgagg agctgctgac cgagtcagag gagaacgaca	2520
gcgaggaagg ctccgaggag atgtacaaag tggaggctga ggaggaaaat gctgacaaaa	2580
gcaaggcagc acctgccgaa attgaccaga gataatgtga acttctacta ggcaaagcaa	2640
tacatcggtc caaggatttt ctgctttcat ttctttaaaa gttttttgtt agtttgtttt	2700
ttgtttttgt ttttgggttt ttttggcttt atttttgtct ttttatgtct gttttgtttt	2760

tcttaccctt ttggacattt ctttgttgca caggatacac ctatagactg aataagttca	2820
gtatttccga atcagacatc gccttggcaa agacactaaa gcgttacact ttatcccgtc	2880
tctatgactg gatcatagtc attataatca caggagactc tgccttcatt atccttgcac	2940
ttaacggaag ttacatcagg caagtaccag gatgaaaaga actatgaaat aaatgaagga	3000
agctacaagt gtgtgtgtat atgtatatgt atatatctct atatttacat atatatatta	3060
aaattgcatg ggacagagac tttgcaatcc gaaagaatag actgtgaaat gagttcttaa	3120
agaaaagact tgtttatgta ttaaaaaaac cacttcacag tgagtcgctt tggctttttg	3180
ataaactgcg gcctgctctc aggtgtgggt gactattttt gaattcctat ttattttttg	3240
tgtttgtccc tgattttttt ttttaattc tatggcttcc tatctggcag cttaatgggt	3300
aatttttgag gtatgtattt aacaaaataa acgacactgc cgaaaaaaa aaaagtgaag	3360
tgaaaacaat cagggcacat taaaatgata caagtcaa ataatcttaa gacacaatgc	3420
acacttaaaa tgactcaata aaatgacttg ctacgttccg ttattcaatt tgtcattact	3480
gtagtgaaca gatgcatttc tgtggaattc caaataagta aaactgaaat tcagtgcaga	3540
gaaaactttg tccactagtg caagtcttga tcaaatgaca ttttgacatt ggacatatgg	3600
aattcatagt atgagccaca ttttgttgtg aaatttatat acctgcttgt ggcttcaaat	3660
ctgaaaatta ataagcctgc tcgtttaaaa gttgtttgtt gttgctgttt ttttgcctt	3720
ttgtttttta ctagaaaata gttcagtgtg atattaagtt agaaaagaag ttgctgcca	3780
gttaaagggg ctccctctca aataaatctc catccttccc tctcccaaaa gacatttctg	3840
atttctgctt cactttgggc ttcctcttct tcgtacacat tccatctacc taatcaaaca	3900
ttttcagtcc ctgactctc ctgtcccttt tcttgggatg acagccctaa caagaactgt	3960
ttttgaatcg ttgtgcagct ccaggcaata gagtatgtga agcgatttca gtagaatcac	4020
ttactcatcc taaaagaaaa cattatccca gttacctaca tcgcaattac cttatgtaaa	4080
gcagaactaa tgctgactgg atgtttaatg ggatgagcat taaagctgca atctactata	4140
gtactccaga tctctttcgg cttcctatga gaaacaccag aagcattact ttccacttct	4200
acttacagta attgcaagag gagacctcac attcaggact ggcctagtga acgtaatcca	4260
tgctttaaac tggccattaa acagtccac atggttggat ttttttttt tttttgagtt	4320
gtgctttcac aaaaccttgt caaagacctc atgcaatata actttgaaag ttattttctg	4380
tttactacac aaacattgta atataactgt taatactatt tatatatattg aaaggtataa	4440
aaggtaggag ttaaaaaaaa aacctctatg ttagatatt aactcagaac ttacaatata	4500
cagggagaag acatgttgca atacaagcta attctagctg ctacgtaacc tctggagttt	4560
ttaaagggac attttctgt actttttcaa ataagtatgt ttaaaaatta tcttgacata	4620
agcgtcatat acctttgcaa aaggatgggt gtttgagtt agccctggcc ccatccttcc	4680
tatttctgta gtatgctgca gctttaatca gaaagtccat gggtgctgct tctgactctc	4740
cgagttactc ttcccaaatt gtcttcttac actgttgctg aaggtcactc tgtacacgta	4800

atggaaactg attttgccaa gctcttaca ggtggttcat ctatcgatgg catccgcatt	4860
tggtatcttt tacacttcaa ccaaaaattt attaggtatt tttcaatgct aagtcttgcc	4920
ttttattttt taatttcact gccaagtttg cagtggttct aagtgaatct gtgggcattt	4980
tagcctgtgg tcttgccaga tctttgcgaa ttacaatgca tatatgtcta tttattcaat	5040
atctgtcata taatatctat ttggaagaag aaactttctc ttgtagtgcc tcttgacaaa	5100
gcacaatttc cgccttttt tttttttgt gaaatgaaaa aaacaaattg tgttttattg	5160
cggtatcaac aatgtgaata aggattaaca tattgtaaat gttctttttt ccatgtaaat	5220
caactatctt tgttatcact aagtgataat taatttttaa cttatgtgca ttgttaggct	5280
gttagaattt ttggttggtt aaaataaacg cattcaat	5318

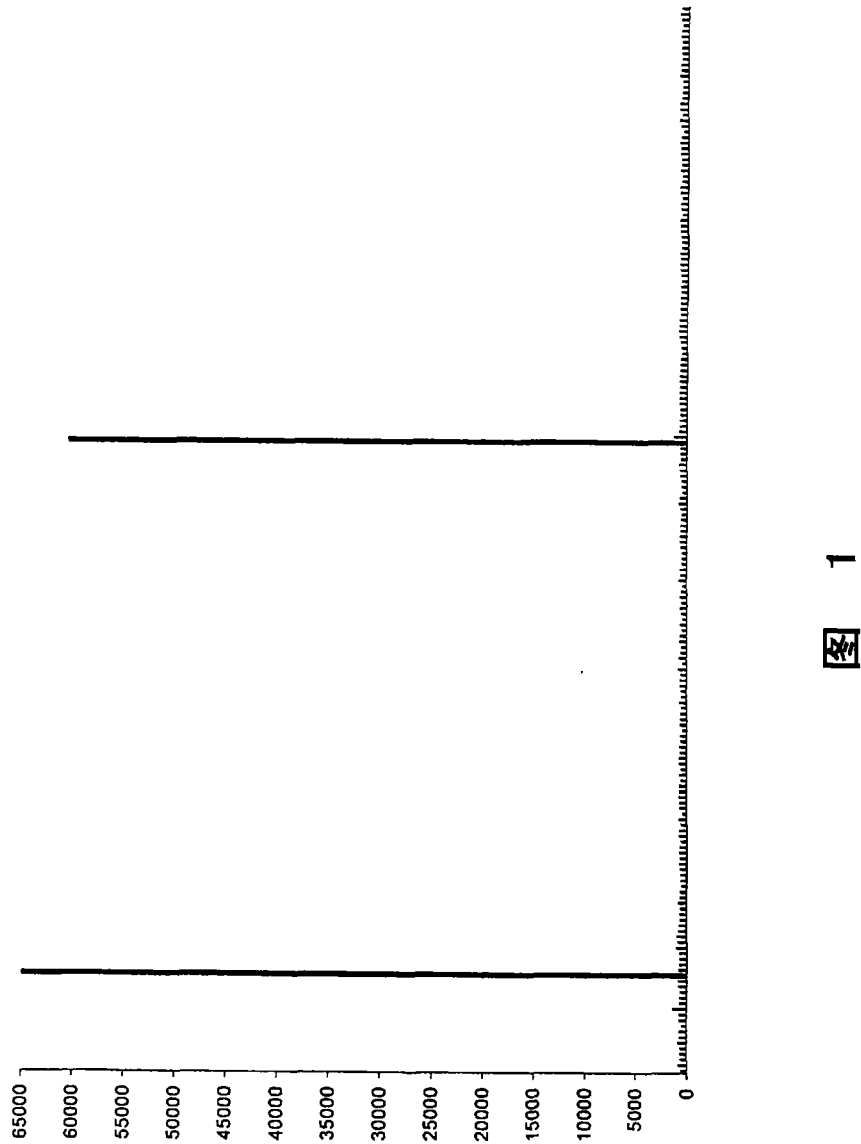


图 1

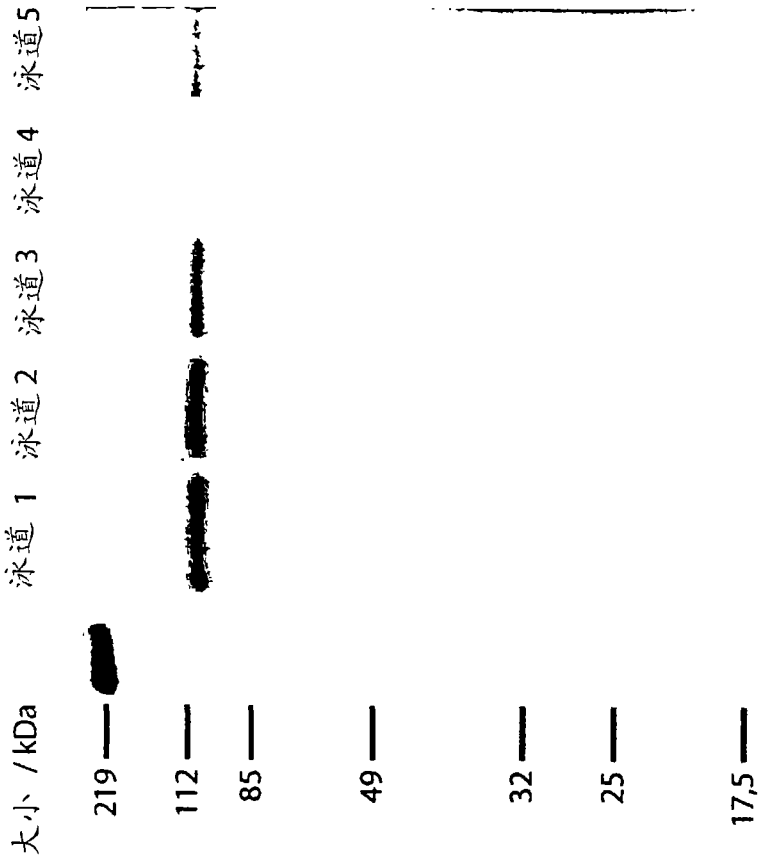
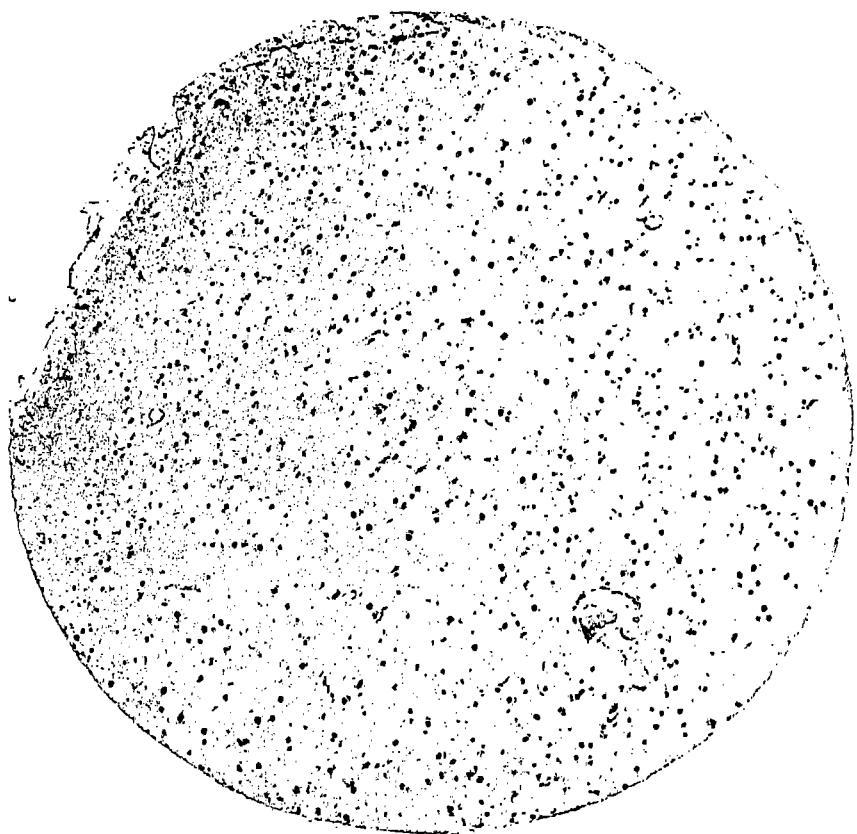
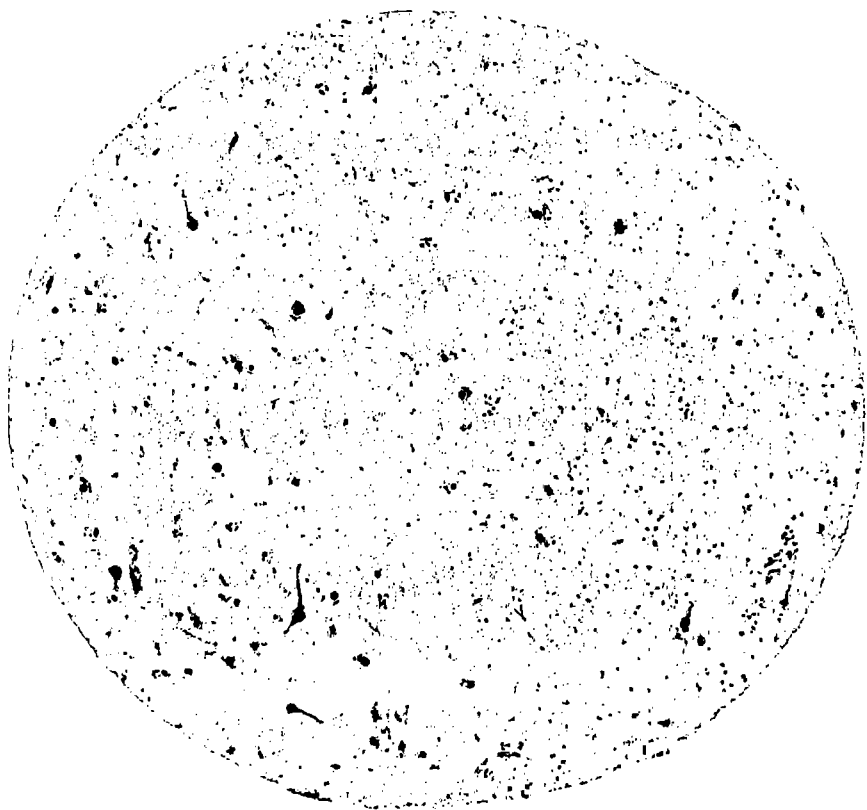


图 2



3A

图



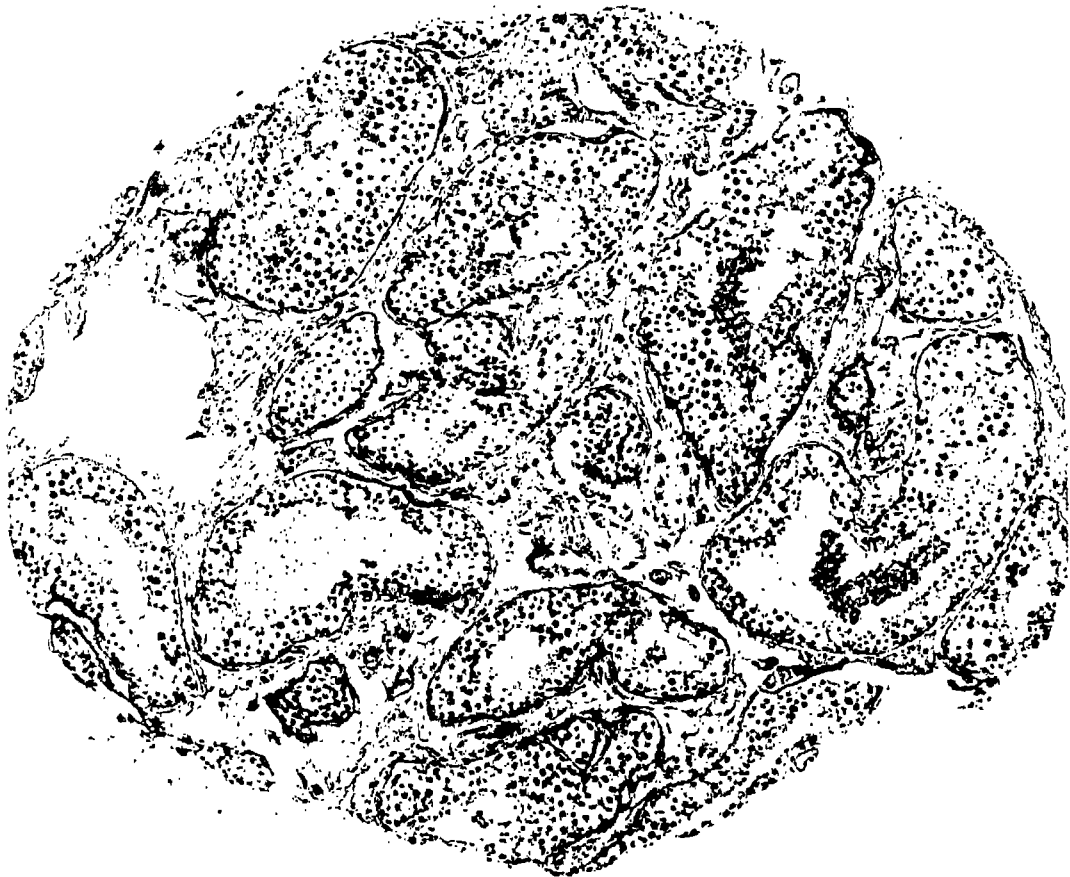


图 3B



图 3C

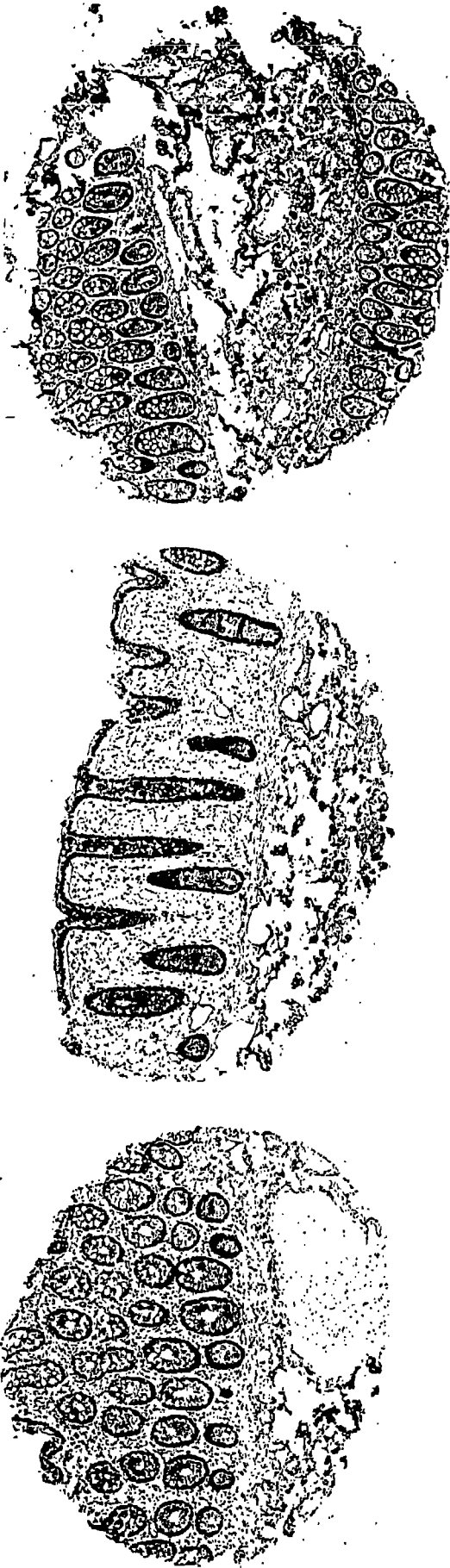


图 3D



图 3E



图 3F

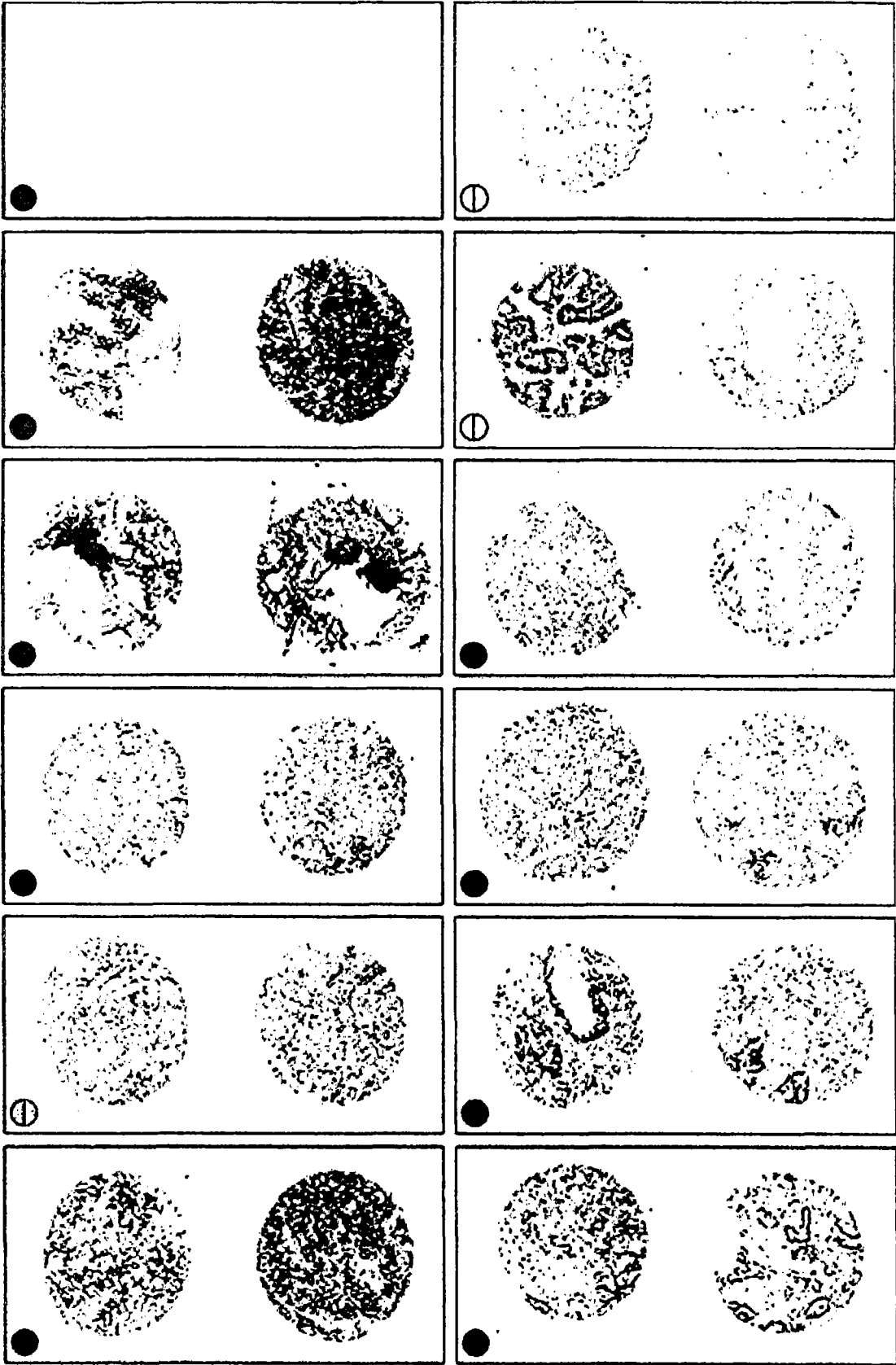


图 4A

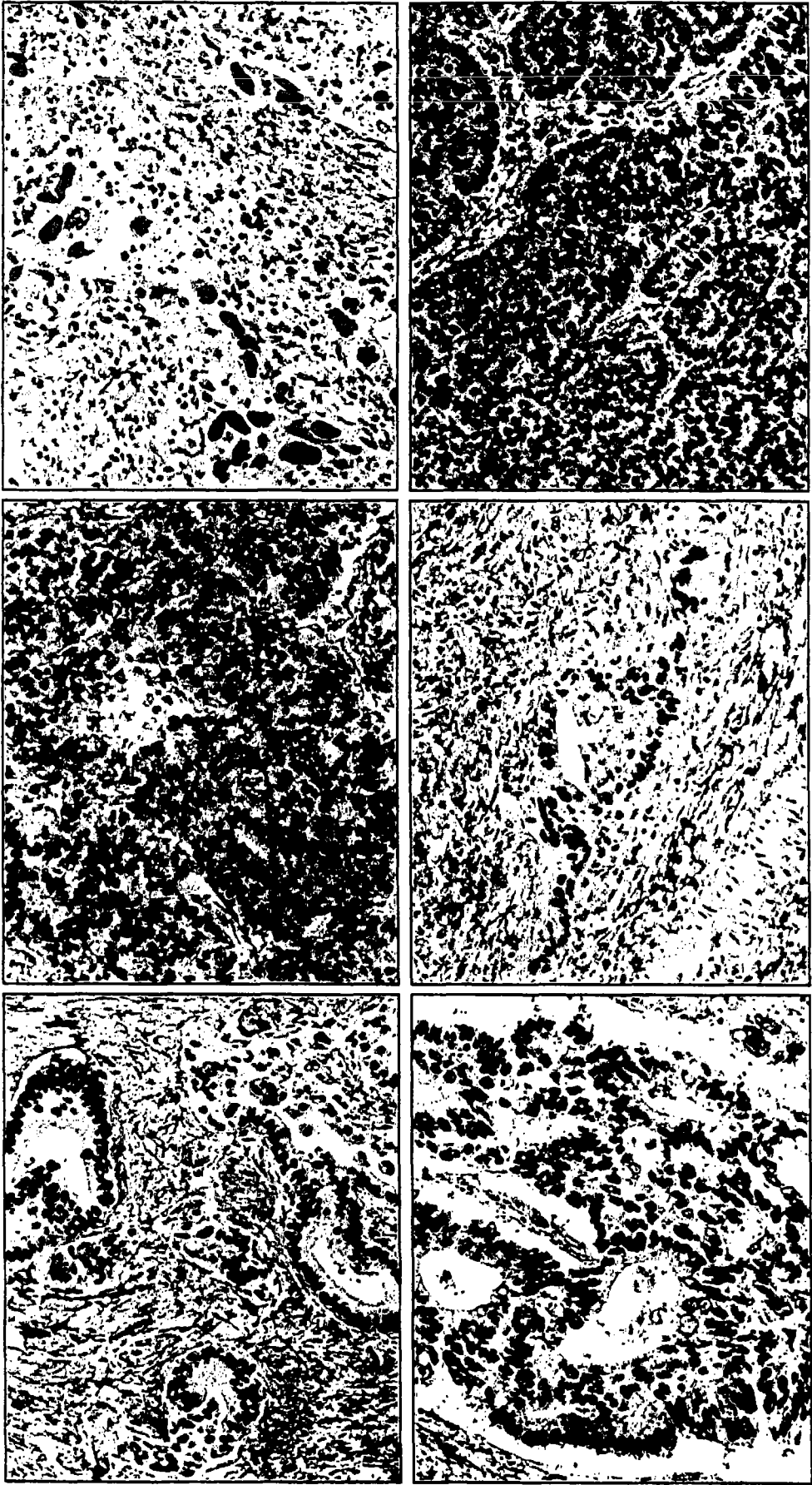


图 4B

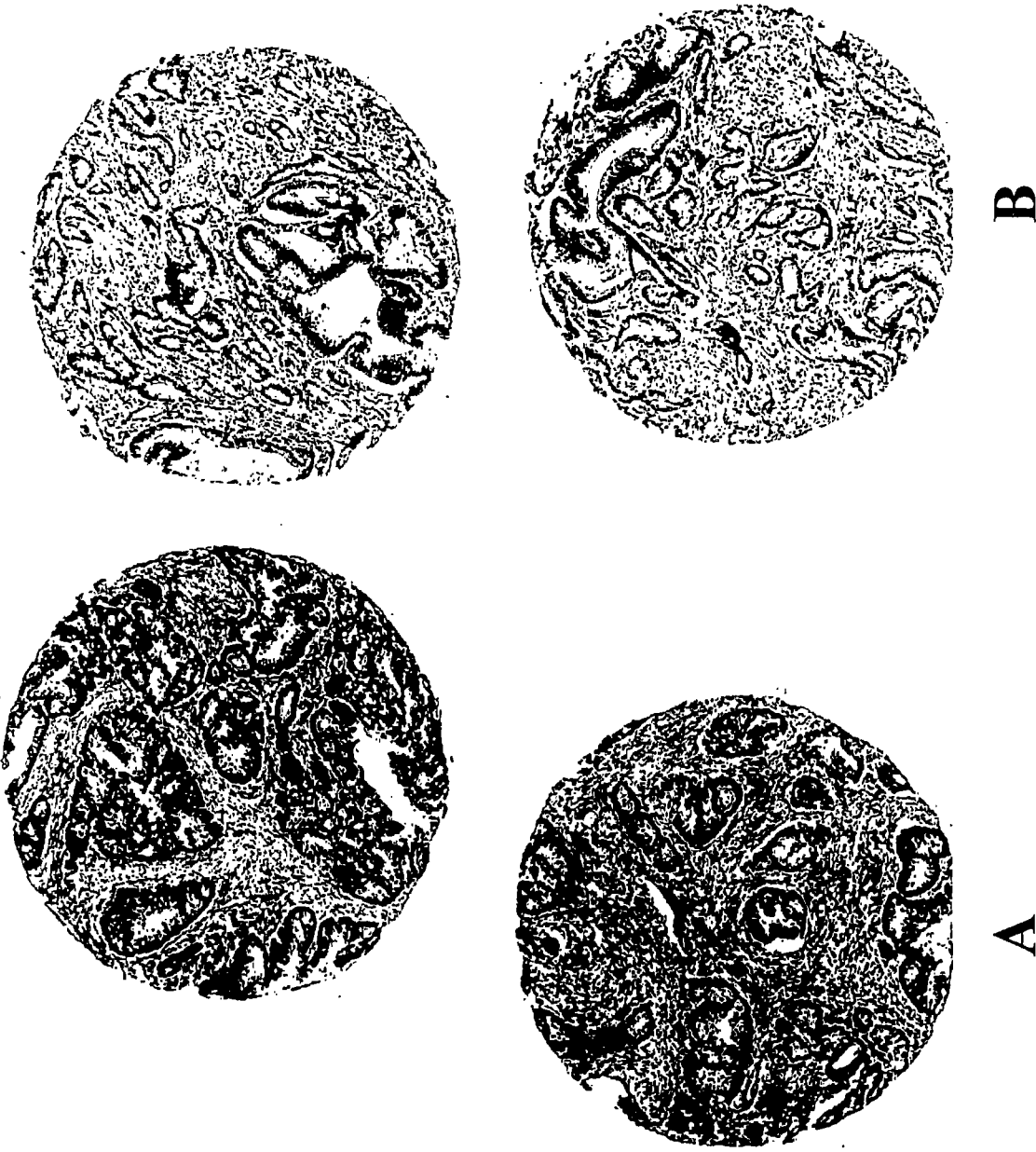


图 5

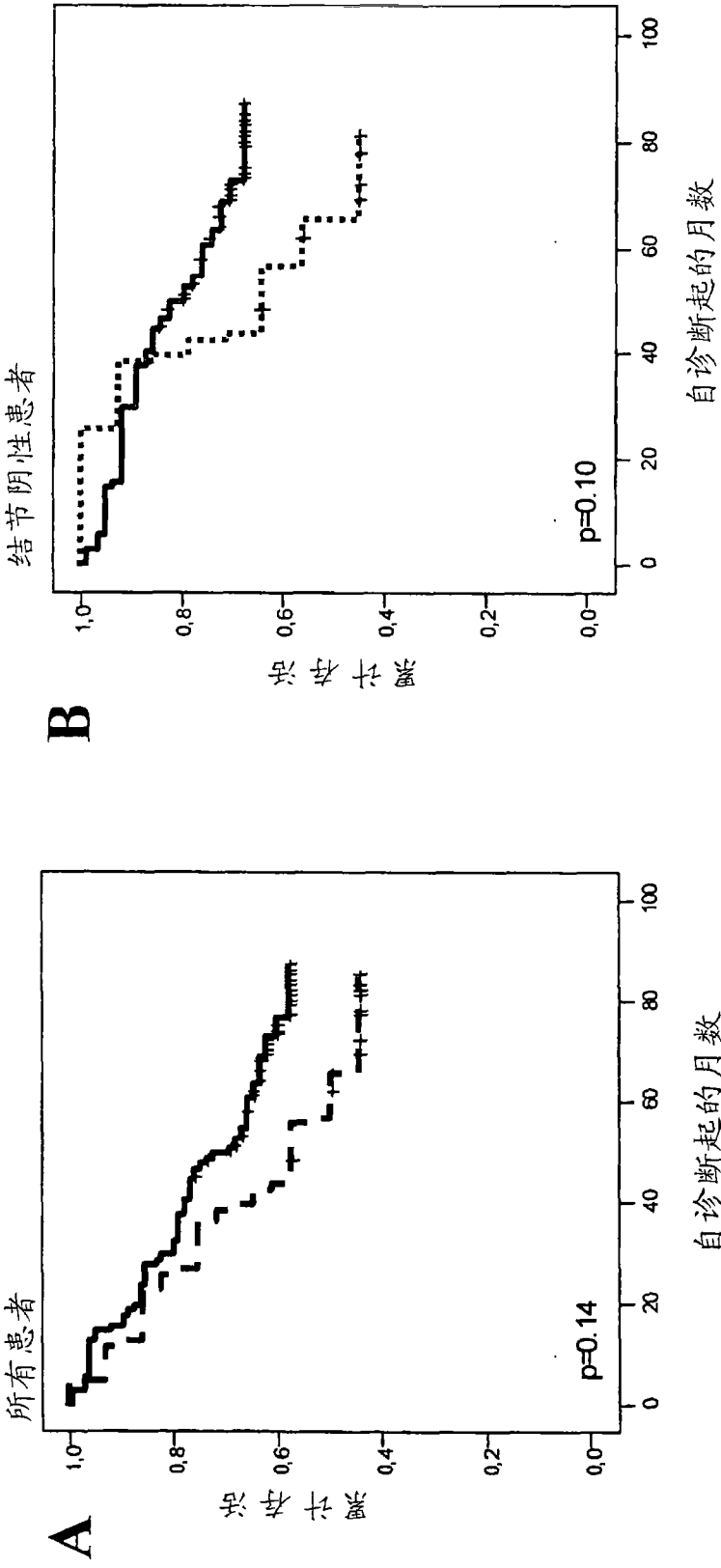


图 6A和6B

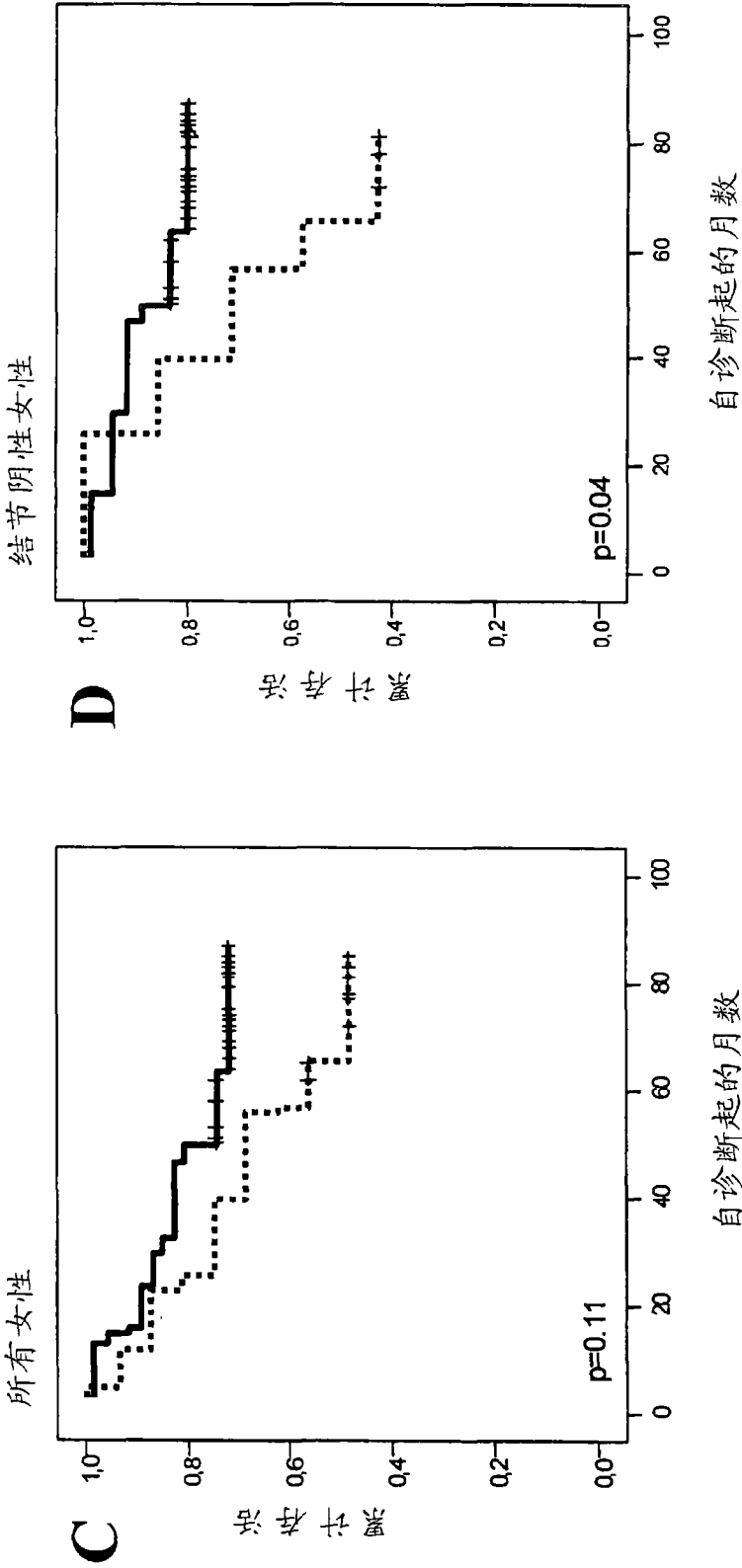


图 6C和6D

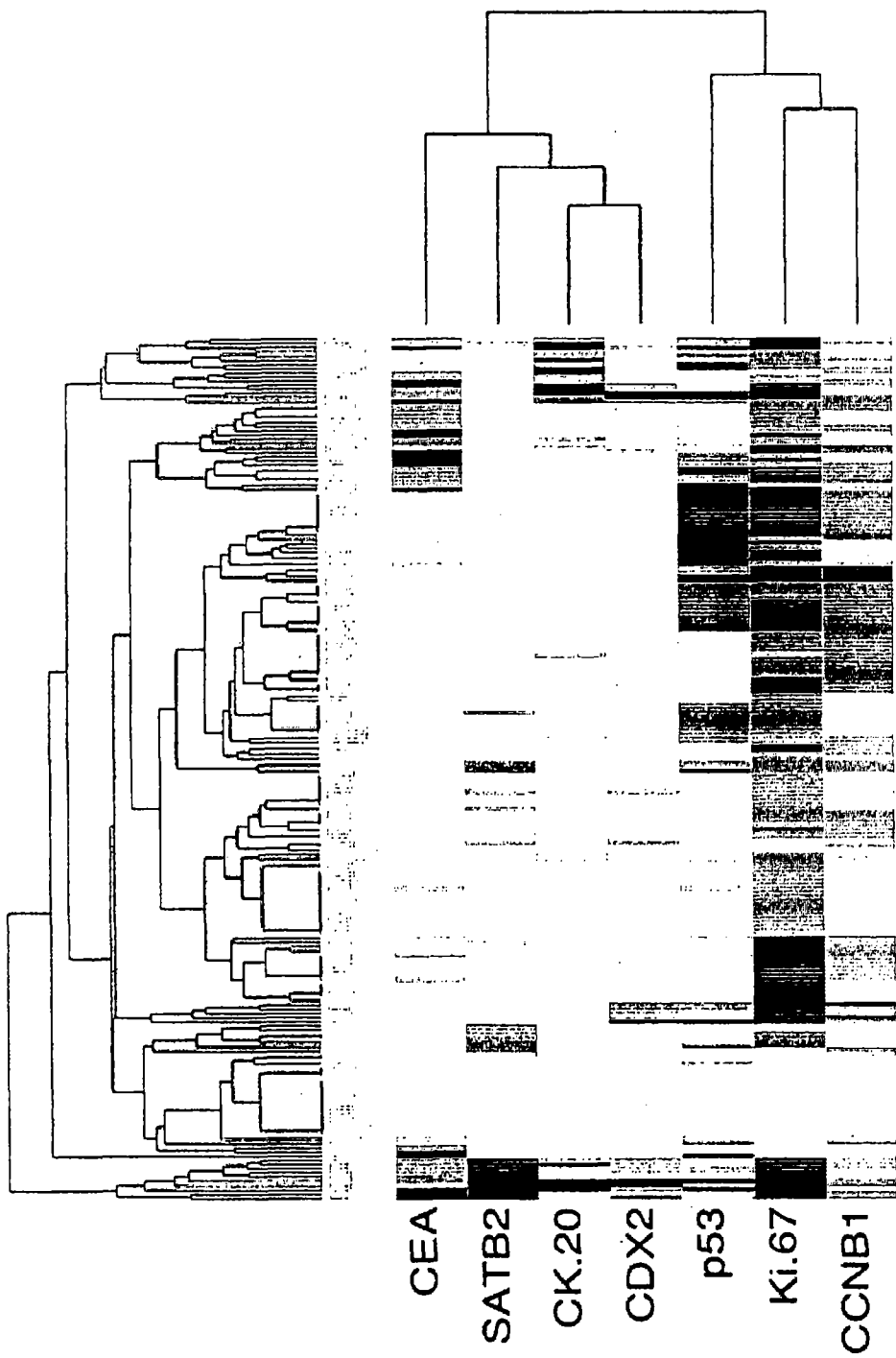


图 7

		SATB2 核分数				
		0-1%	2-25%	26-75%	>75%	总计
CK 20 核分数	0-1%	5	1	4	7	17
	2-25%	3		5	12	20
	26-75%	4	1	6	21	32
	>75%	11	3	16	23	53
	总计	23	5	31	63	122

图 8

淋巴结中Satb2和CK20的对比

计数		Satb2 核分数			总计
		0-1%	26-75%	>75%	
CK20	0	1			1
核分数	1			1	1
	2	1	2	3	6
	3	1	3	5	9
总计		3	5	9	17

图 9