



(12) 发明专利申请

(10) 申请公布号 CN 113692449 A

(43) 申请公布日 2021. 11. 23

(21) 申请号 201980092626.0

(22) 申请日 2019.12.20

(30) 优先权数据

10201811546W 2018.12.21 SG

(85) PCT国际申请进入国家阶段日

2021.08.20

(86) PCT国际申请的申请数据

PCT/SG2019/050629 2019.12.20

(87) PCT国际申请的公布数据

W02020/130948 EN 2020.06.25

(71) 申请人 新加坡科技研究局

地址 新加坡新加坡

申请人 新加坡国立大学医院

新加坡国立大学

(72) 发明人 陈文炜 R·孙达尔 K·K·黄

A·卡玛拉 J·戈克

D·德米尔吉路

(74) 专利代理机构 北京世峰知识产权代理有限公司 11713

代理人 王思琪 王建秀

(51) Int.Cl.

*G12Q 1/6886* (2018.01)

*G01N 33/574* (2006.01)

*A61K 39/395* (2006.01)

*A61P 35/00* (2006.01)

权利要求书2页 说明书18页 附图13页

(54) 发明名称

预测免疫检查点抑制疗法益处的方法

(57) 摘要

本发明涉及鉴定患有癌症的患者是否受益于免疫检查点抑制疗法的方法,包括:测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平,基于一种或多种预选标志物的表达水平鉴定差异表达的替代启动子,计算替代启动子利用评分(APBscore),并使用APBscore确定患者是否受益于免疫检查点抑制疗法。特别是,高APBscore患者比低APBscore患者具有更低的CD8A、颗粒酶A和穿孔蛋白1表达,并具有更差的无进展生存期。与APBscore患者相比,APBscore较低的患者具有更高的客观响应率。

1. 一种鉴定患有癌症的患者是否受益于免疫检查点抑制 (ICI) 疗法的方法, 包括以下步骤:

- a) 测量从所述患者获得的癌性生物样品中一种或多种预选标志物的表达水平;
- b) 基于步骤 (a) 中测量的所述一种或多种预选标志物的表达水平, 鉴定差异表达的替代启动子;
- c) 计算替代启动子利用评分;
- d) 使用所述替代启动子利用评分鉴定所述患者是否受益于 ICI 疗法。

2. 根据权利要求1所述的方法, 其中所述预选标志物是核酸分子或肽。

3. 根据权利要求1所述的方法, 其中所述预选标志物是与癌性生物样品中体细胞启动子区相关的标志物。

4. 根据权利要求1-3中任一项所述的方法, 其中通过将所述癌性生物样品中所述一种或多种预选标志物的表达水平与一个或多个参考样品中所述一种或多种预选标志物的表达水平进行比较, 以确定所述癌性生物样品与所述一个或多个参考样品之间所述预选标志物表达水平的增加或降低, 来鉴定步骤 (b) 中的所述差异表达的替代启动子。

5. 根据权利要求4所述的方法, 其中对于获得性启动子, 与所述一个或多个参考样品中所述一种或多种预选标志物的表达水平相比, 所述癌性生物样品中所述一种或多种预选标志物表达水平的增加表示差异表达的替代启动子, 并且其中对于缺失性启动子, 与所述一个或多个参考样品中所述一种或多种预选标志物的表达水平相比, 所述癌性生物样品中所述一种或多种预选标志物表达水平的降低表示差异表达的替代启动子。

6. 根据权利要求5所述的方法, 其中所述获得性启动子是, 与非癌性生物样品相比, 在癌性生物样品中是获得性的或增加的启动子, 并且其中所述缺失性启动子是, 与非癌性生物样品相比, 在癌性生物样品中是缺失性的或降低的启动子。

7. 根据权利要求4-6中任一项所述的方法, 其中将所述癌性生物样品中所述一种或多种预选标志物的表达水平, 与所述一个或多个参考样品中所述一种或多种预选标志物的百分位数表达水平或所述一个或多个参考样品中所述一种或多种预选标志物的平均表达水平或二者进行比较。

8. 根据权利要求7所述的方法, 其中 i) 对于获得性启动子, 与中值表达水平相比, 预选标志物表达水平的增加, 或 ii) 对于缺失性启动子, 与中值表达水平相比, 预选标志物表达水平的降低, 表示差异表达的替代启动子。

9. 根据权利要求1-8中任一项所述的方法, 其中通过测定步骤 (b) 中鉴定的所述差异表达的替代启动子之和, 来计算步骤 (c) 中的所述替代启动子利用评分。

10. 根据权利要求1-9中任一项所述的方法, 其中通过测定获得性启动子的数量和缺失性启动子的数量之和, 来计算步骤 (c) 中的所述替代启动子利用评分, 与转录物中值表达水平相比, 所述获得性启动子的转录物表达水平是其至少4倍, 所述缺失性启动子的转录物表达水平不到其0.25倍。

11. 根据权利要求1-9中任一项所述的方法, 其中通过确定相对启动子活性大于中值的获得性启动子的数量和相对启动子活性小于中值的缺失性启动子的数量之和, 来计算步骤 (c) 中的所述替代启动子利用评分。

12. 根据权利要求1-11中任一项所述的方法, 还包括以下步骤: 将步骤 (c) 中计算的所

述替代启动子利用评分与参考评分进行比较,以鉴定所述患者是否受益于ICI疗法,其中与所述参考评分相比,增加的替代启动子利用评分将所述患者鉴定为未受益于ICI疗法。

13.根据权利要求12所述的方法,其中所述参考评分是一个或多个参考样品的替代启动子利用评分。

14.根据权利要求12或13所述的方法,其中所述增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。

15.根据权利要求14所述的方法,其中所述增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

16.确定正在接受或已经接受ICI疗法的患者预后的方法,所述方法包括以下步骤:

- a) 测量从所述患者获得的癌性生物样品中一种或多种预选标志物的表达水平;
- b) 基于步骤(a)中测量的所述一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;
- c) 计算替代启动子利用评分;
- d) 使用所述替代启动子利用评分确定所述患者的预后。

17.根据权利要求16所述的方法,还包括以下步骤:将步骤(c)中计算的所述替代启动子利用评分与参考评分进行比较,以确定所述患者的预后,其中与参考评分相比,增加的替代启动子利用评分表示预后较差。

18.根据权利要求17所述的方法,其中所述参考评分是一个或多个参考样品的替代启动子利用评分。

19.根据权利要求17或18所述的方法,其中所述增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。

20.根据权利要求19所述的方法,其中所述增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

21.用于将患有癌症的患者从ICI疗法中排除的生物标志物,其中所述生物标志物是高于一个或多个参考样品的替代启动子利用评分的替代启动子利用评分,其中所述替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

22.根据权利要求21所述的生物标志物,其中所述生物标志物是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的替代启动子利用评分。

23.根据权利要求21所述的生物标志物,其中所述生物标志物是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的替代启动子利用评分。

## 预测免疫检查点抑制疗法益处的方法

[0001] 对相关申请的交叉引用

[0002] 本申请要求2018年12月21日提交的第10201811546W号新加坡申请的优先权,其内容通过引用整体并入本文,用于所有目的。

### 技术领域

[0003] 本发明总体上涉及癌症领域。特别是,本发明涉及选择针对免疫检查点抑制疗法患者的方法的用途。

### 背景技术

[0004] 在癌症的各种治疗方式中,免疫检查点抑制(immune checkpoint inhibition, ICI)已经在多种肿瘤类型中取得了重大突破。在ICI疗法中,派姆单抗和纳武单抗等免疫检查点抑制剂阻断免疫检查点受体PD-1与其配体之间的相互作用,从而减少负共刺激信号并增强T效应细胞引发抗肿瘤响应的功能。

[0005] 虽然对某些肿瘤类型有益,但最近的几项ICI III期试验也证明不成功,尤其是在实体上皮肿瘤中。仅在少数患者中观察到对ICI治疗的响应。因此,需要能够预测患者对免疫疗法的响应的稳健生物标志物。目前,最成熟的ICI生物标志物是PD-L1表达、微卫星不稳定性 and 肿瘤突变负荷。这些生物标志物是阳性预测生物标志物,旨在确定对ICI疗法有响应的患者。然而,已经出现围绕这些生物标志物的争议,并且已经观察到生物标志物阴性群体中的ICI响应。这些观察结果突出强调了对用于ICI疗法的阴性预测生物标志物的补充要求,该阴性预测生物标志物可以识别可能对免疫疗法具有抗性的肿瘤。然而,对ICI的阴性预测生物标志物的描述仍然不足。

[0006] 启动子是转录起始位点(TSS)上游的基因组顺式调节元件,其功能是启动转录。启动子活性受表观遗传调节,超过一半的所有人类基因具有多重启动子,这些启动子可以由正常生物功能或疾病状态的结果而被选择性激活。使用替代启动子(alternative promoter)可以产生不同的5'非翻译区(UTR)和第一外显子,从而提高mRNA和蛋白质同种型的多样性。在癌症中,替代启动子可以生成具有致癌性的癌症特异性同种型。最近,研究表明,肿瘤可能利用替代启动子作为免疫编辑和逃避的机制。因此,表观遗传驱动的替代启动子利用是对ICI疗法抗性的潜在机制。

[0007] 因此,需要开发基于替代启动子利用的替代方法,以预测患者对免疫疗法或ICI疗法的响应。

[0008] 发明概述

[0009] 在一个方面,提供了鉴定患有癌症的患者是否受益于免疫检查点抑制(ICI)疗法的方法,包括以下步骤:

[0010] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0011] b) 基于步骤(a)中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0012] c) 计算替代启动子利用评分 (alternative promoter usage score);

[0013] d) 使用替代启动子利用评分鉴定患者是否受益于ICI疗法。

[0014] 在另一方面,提供了确定患有癌症的患者的预后的方法,包括以下步骤:

[0015] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0016] b) 基于步骤 (a) 中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0017] c) 计算替代启动子利用评分;

[0018] d) 使用替代启动子利用评分确定患者的预后。

[0019] 在另一方面,提供了用于将患有癌症的患者从ICI疗法中排除的生物标志物,其中生物标志物是高于一个或多个参考样品的替代启动子利用评分的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0020] 定义

[0021] 本文所用的术语“启动子”是指启动基因转录的DNA区。启动子可以是主要启动子、次要启动子或替代启动子。主要启动子是最常用于基因转录的启动子。

[0022] 本文所用的术语“备选启动子 (alternate promoter)”和“替代启动子”是指在不同于主要启动子或次要启动子的、于替代转录起始位点处启动基因转录的DNA区。

[0023] 本文所用的术语“获得性启动子 (gained promoter)”是指,与非癌性生物样品相比在癌性生物样品中获得性的或活性增强的启动子。获得性启动子可以是不存在于非癌性生物样品中的癌性生物样品中的启动子。获得性启动子可以是癌性生物样品中与非癌性生物样品中的启动子相比活性增强的启动子。

[0024] 本文所用的术语“缺失性启动子 (lost promoter)”是指,与非癌性生物样品相比,在癌性生物样品中缺失性的或活性降低的启动子。缺失性启动子可以是存在于非癌性生物样品中但不存在于癌性生物样品中的启动子。缺失性启动子可以是癌性生物样品中与非癌性生物样品中的启动子相比降低活性的启动子。

[0025] 在癌症的上下文中,术语“免疫疗法”是指涉及调节免疫系统以治疗癌症的治疗形式。免疫系统的调节可涉及免疫系统的激活或失活。这可能涉及使用免疫系统的成分,例如抗体、细胞因子和疫苗。

[0026] 本文所用的术语“免疫检查点抑制疗法”是指靶向免疫系统的关键调节因子的癌症免疫治疗形式。某些类型的癌细胞和免疫细胞产生的某些蛋白质有助于控制免疫应答,并可以防止T细胞杀伤癌细胞。某些癌细胞使用这些蛋白质来逃避免疫监视。术语“免疫检查点抑制疗法”是指,涉及使用阻断这些蛋白质的因子,恢复免疫系统功能并允许T细胞增强有效抗肿瘤响应的疗法。

[0027] 就本申请的目的而言,术语“预后”是指预测临床疾病状况或疾病的可能进程和结果。术语“预后”并不是指能够100%准确地预测疾病状况的进程或结果。相反,术语“预后”是指与未表现出给定疾病状况的个体相比,表现出该疾病状况的个体发生某个进程或结果的概率。可以针对一种或多种临床结果做出“预后”,例如个体疾病的进展速度、疾病的严重性、存活率、存活时间、转移几率、疾病复发几率或对治疗干预的响应。

[0028] 就本申请的目的而言,术语“表观遗传”是指在不改变个体DNA序列的情况下调节基因表达的变化。表观遗传改变的实例是DNA甲基化、组蛋白修饰和RNA相关沉默。表观遗传

改变通过控制基因表达在疾病发展中起着至关重要的作用。

[0029] 本说明书环境中的术语“生物标志物”是指生物学疾病状况或疾病的可测量指标。生物标志物包括但不限于可以在身体组织和体液中测量的物质、结构或过程。

[0030] 本文所用的术语“备选启动子利用评分(alternate promoter usage score)”和“替代启动子利用评分”是指替代启动子利用的度量。评分可以表示为“AP”评分或“APB”评分。 $AP_{high}$  ( $AP_{高}$ )和 $AP_{low}$  ( $AP_{低}$ )环境中的术语“AP”是指备选启动子利用。 $AP_{high}$ 评分表示高替代启动子利用, $AP_{low}$ 评分表示低替代启动子利用。 $APB_{high}$  ( $APB_{高}$ )和 $APB_{low}$  ( $APB_{低}$ )环境中的术语“APB”是指备选启动子利用负荷。 $APB_{high}$ 评分表示高替代启动子利用, $APB_{low}$ 评分表示低替代启动子利用。

[0031] 附图简要说明

[0032] 当结合非限制性示例和附图考虑时,参考详细描述会更好理解本发明,其中:

[0033] 图1显示了33种肿瘤类型的备选启动子利用负荷(APBscore)。APBscore是在33种肿瘤类型的10,393个样品中计算的。每个点都代表样品的APBscore(y轴)。肿瘤类型位于x轴上,按从各自肿瘤类型的最低到最高中值APBscore排序。红色水平条是肿瘤的中值APBscore,而绿色水平条代表相应正常组织的中值APBscore(仅当至少有10个正常样品可用于分析时才提供)。显示了第80个百分位数截止值,因为发现这是将样品二分为 $APB_{high}$ 和 $APB_{low}$ 组的最理想截止点。

[0034] 图2显示了胃癌中的备选启动子利用(发现队列, $n=24$ )。特别是,(A)显示了在接受纳武单抗或派姆单抗治疗的患者发现队列中备选启动子利用的热图。所有肿瘤中表达水平高于中值水平的4倍且映射到先前鉴定的获得性替代启动子位点的转录物被认为是获得性替代启动子(在热图中标记为红色)。所有肿瘤中表达水平比低于中值水平的1/4且映射到先前确定的缺失性替代启动子位点的转录物被认为是缺失性替代启动子(在热图中标记为蓝色)。(B)显示了 $AP_{high(disc)}$ 组相比 $AP_{low(disc)}$ 组与T细胞免疫相关物(correlate)之间的关联。 $AP_{high(disc)}$ 组为红色, $AP_{low(disc)}$ 组为蓝色。描绘了T细胞标志物CD8A ( $P=0.059$ )以及T细胞溶细胞性标志物GZMA ( $P=0.025$ )和PRF1 ( $P=0.011$ )的表达。 $AP_{high(disc)}$ 组显示免疫标志物的表达较低。(C)显示了无进展生存期的Kaplan Meier曲线,其比较了发现队列中的 $AP_{high}$ 组与 $AP_{low}$ 组。

[0035] 图3显示了胃癌中的备选启动子利用(派姆单抗试验队列, $n=37$ )。特别是,(A)显示了备选启动子利用的热图。所有肿瘤中表达水平高于中值水平的4倍且映射到先前确定的获得性替代启动子位点的转录物被认为是获得性替代启动子(在热图中标记为红色)。所有肿瘤中表达水平低于中值水平的0.25倍且映射到先前确定的缺失性替代启动子位点的转录物被认为是丧失性替代启动子(在热图中标记为蓝色)。(B)显示了替代启动子利用评分的图表。替代启动子利用评分计算为各样品中获得性替代启动子和缺失性替代启动子之和。高备选启动子利用被定义为>第66百分位数的那些备选启动子利用。(C)显示了 $AP_{high}$ 组相比 $AP_{low}$ 组与T细胞免疫相关物之间的关联。 $AP_{high}$ 组为红色, $AP_{low}$ 组为蓝色。描绘了T细胞标志物CD8A ( $P=0.0037$ )和T细胞溶细胞性标志物GZMA ( $P=0.0055$ )和PRF1 ( $P=0.016$ )的表达。 $AP_{high}$ 组显示免疫标志物的表达较低。(D)是 $AP_{high}$ (红色)和 $AP_{low}$ (蓝色)亚组对派姆单抗响应的瀑布图。Y轴代表了根据RECIST 1.1标准评估的最大肿瘤减少百分比。

[0036] 图4显示了基于备选启动子利用的生存曲线。特别是,(A)显示了无进展生存期的

Kaplan Meier曲线,其比较了AP<sub>high</sub>组与AP<sub>low</sub>组。(B)显示了Swimmer图,其中x轴代表每位患者的派姆单抗治疗持续时间。描绘了AP<sub>high</sub> (红色)和AP<sub>low</sub> (蓝色)子组。(C)显示了无进展生存期的Kaplan Meier曲线,其比较了AP<sub>high</sub>组与AP<sub>low</sub>组划分的TCGA亚型。

[0037] 图5显示了APBscore算法的生物信息学工作流程。针对泛癌分析中从TCGA鉴定的所有启动子(113,076个启动子),对在胃癌中鉴定的体细胞启动子区(获得和缺失)(2732个区)进行映射。总共选择了4672个启动子,并使用相对启动子活性来计算APBscore。肿瘤在APBscore的第80个百分位数处被二分为APB<sub>high</sub>和APB<sub>low</sub>。

[0038] 图6显示了APBscore组与溶细胞性T细胞活性标志物的相关性。特别是,(A)显示了APB<sub>high</sub>组相比APB<sub>low</sub>组(所有肿瘤样品,泛癌)与T细胞免疫相关物之间的关联。肿瘤在APBscore的第80个百分位数处被二分为APB<sub>high</sub>和APB<sub>low</sub>。APB<sub>high</sub>组为红色,而APB<sub>low</sub>组为蓝色。描绘了T细胞标志物CD8A和T细胞溶细胞性标志物GZMA和PRF1的表达。APB<sub>high</sub>组显示免疫标志物的表达较低(对于CD8A、GZMA和PRF1,威尔科克森检验(Wilcoxon test)  $p < 0.0001$  (\*\*\*) )。(B)显示了,对于选择的肿瘤类型,APB<sub>high</sub>组相比APB<sub>low</sub>组与CD8A、GZMA和PRF1之间的关联:膀胱癌(BLCA)、乳腺癌(BRCA)、宫颈鳞状细胞癌(CESC)、食管癌(ESCA)、头颈部鳞状细胞癌(HNSC)、肝脏肝细胞癌(LIHC)、肺鳞状细胞癌(LUSC)、卵巢癌(OV)和胃腺癌(STAD)。APB<sub>high</sub>组为红色,APB<sub>low</sub>组为蓝色。(\*\* $= p < 0.0001$ , \*\* $= p < 0.001$ , \* $= p < 0.01$ , NS $= p > 0.05$ , 威尔科克森检验)。

[0039] 图7显示了APBscore组中的差异基因表达,突出了免疫基因。在与APB<sub>high</sub>和APB<sub>low</sub>相关的该各免疫基因火山图中,x轴是APB<sub>high</sub>和APB<sub>low</sub>之间的基因表达(RSEM)  $\log_2$ 倍数变化( $\log_2FC$ )。y轴是 $-\log_{10}$ 调整后的p值结果(Bonferroni校正)。定义为 $\log_2FC > 1$ 且调整后的p值 $< 0.05$ 的显著基因以红点表示。标记了位于顶部(top)的基因( $\log_2FC > 1.3$ 且 $-\log_{10}$ 调整后的p值 $> 200$ )。

[0040] 图8显示了APBscore组和其他免疫相关物之间的关联。特别是,该图显示了与APB<sub>high</sub>和APB<sub>low</sub>相关的PanCanAtlas中约20,000个基因的火山图。x轴是APB<sub>high</sub>和APB<sub>low</sub>之间的基因表达(RSEM)的 $\log_2$ 倍数变化( $\log_2FC$ )。y轴是 $-\log_{10}$ 调整后的p值结果(Bonferroni校正)。免疫基因以红点表示。观察到两组之间整体基因表达的均等分割(equal split),伴有APB<sub>high</sub>组中免疫基因表达略低(under-expression)的倾斜。

[0041] 图9显示了淋巴细胞浸润特征评分与APBscore的相关性。淋巴细胞浸润特征由APBscore组二分。与APB<sub>high</sub>相比,APB<sub>low</sub>组的评分在统计学上较高(威尔科克森检验 $p < 0.0001$ )。

[0042] 图10显示了IFN- $\gamma$ 响应特征与APBscore的相关性。IFN- $\gamma$ 响应特征由APBscore组二分。与APB<sub>high</sub>相比,APB<sub>low</sub>组的评分在统计学上较高(威尔科克森检验 $p < 0.0001$ )。

[0043] 图11显示了,在由于与CD8A、GZMA和PRF1的相关性而选择的19种肿瘤类型中,APBscore与派姆单抗客观响应率(ORR)之间的相关性。x轴是每种肿瘤类型的中值APB评分,y轴是已公开试验中描述的使用单药剂派姆单抗的肿瘤类型的ORR。圆圈的大小对应于选择用于ORR评估的临床试验中治疗的患者数量。圆圈的颜色对应于TCGA队列中用于评估APBscore的样品的数量。COAD包括COAD和READ(结肠癌直肠癌)。

[0044] 发明详述

[0045] 在第一方面,本发明涉及鉴定患有癌症的患者是否受益于免疫检查点抑制(ICI)

疗法的方法,包括以下步骤:

[0046] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0047] b) 基于步骤(a)中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0048] c) 计算替代启动子利用评分;

[0049] d) 使用替代启动子利用评分鉴定患者是否受益于ICI疗法。

[0050] 本领域技术人员通常会理解,癌性生物样品可以是新鲜的、冷冻的、固定的或保存的样品。

[0051] 在一个实施方案中,患有癌症的患者已经接受或正在接受ICI疗法。在优选的实施方案中,患有癌症的患者没有接受ICI疗法。

[0052] 癌性生物样品可采集自被诊断为患有的一种或多种癌症的患者,癌症可包括但不限于癌、肉瘤和黑素瘤。生物样品可以是细胞、组织或流体样品。

[0053] 在一些实施方案中,癌性生物样品可采集自患有的一种或多种癌症的患者,该癌症包括但不限于肾上腺皮质癌、膀胱尿路上皮癌、乳腺浸润性癌、宫颈鳞状细胞癌和宫颈内腺癌(endocervical adenocarcinoma)、胆管癌、结肠腺癌、弥漫性大B细胞淋巴瘤、食管癌、多形性胶质母细胞瘤、头颈部鳞状细胞癌、肾嫌色细胞、肾脏肾透明细胞癌、肾脏肾乳头状细胞癌、急性髓系白血病、脑低级别胶质瘤、肝脏肝细胞癌、肺腺癌、肺鳞状细胞癌、间皮瘤、卵巢浆液性囊腺癌、胰腺癌、嗜铬细胞瘤和副神经节瘤、前列腺癌、直肠腺癌、肉瘤、皮肤黑色素瘤、胃腺癌、睾丸生殖细胞肿瘤、甲状腺癌、胸腺瘤、子宫体子宫内膜癌、子宫癌肉瘤、葡萄膜黑素瘤、肛门癌及其组合。

[0054] 在一个实施方案中,癌症是实体癌。在另一个实施方案中,实体癌是胃肠癌。在又一个实施方案中,胃肠癌是胃癌。在又一个实施方案中,胃癌属于染色体不稳定(CIN)亚型和/或基因组稳定(GS)亚型。在又一个实施方案中,胃癌是转移性胃癌。

[0055] 在一些实施方案中,预选标志物是一种或多种核酸分子。核酸分子可以是DNA或RNA。在一些实施方案中,核酸分子选自mRNA、cDNA、micorRNA和基因组DNA。预选标志物可以是一种或多种基因或一种或多种转录物。在一些实施方案中,预选标志物是一种或多种肽。在另一个实施方案中,肽是翻译后修饰的肽。

[0056] 在一个实施方案中,预选标志物是与癌性生物样品中体细胞启动子区相关的标志物。在优选的实施方案中,癌性生物样品是胃癌样品。

[0057] 可使用选自全转录组测序(whole transcriptome sequencing,WTS)、Nanostring分析、RNA测序及其组合的方法测量本文所述癌性生物样品中一种或多种预选标志物的表达水平。

[0058] 本领域技术人员应当理解,替代启动子的使用允许在不同的转录起始位点启动转录。因此,替代启动子的使用可以以各种方式影响基因表达,包括改变基因表达水平和/或导致产生改变的mRNA转录物和蛋白质亚型。

[0059] 在一个实施方案中,通过比较癌性生物样品中一种或多种预选标志物的表达水平与一个或多个参考样品中所述一种或多种预选标志物的表达水平以确定癌性生物样品与一个或多个参考样品之间所述预选标志物表达水平的增加或降低,来鉴定本文所述差异表达的替代启动子。本领域技术人员应当理解,预选标志物可以是一种或多种基因、转录物或

肽,其表达水平可以受启动子调节,并因此指示启动子的活性。

[0060] 可以通过将癌性生物样品中一种或多种预选标志物的表达水平与一个或多个参考样品中一种或多种预选标志物的表达水平进行比较,来确定差异表达的替代启动子。

[0061] 在一个实施方案中,对于获得性启动子,与一个或多个参考样品中一种或多种预选标志物的表达水平相比,癌性生物样品中一种或多种预选标志物的表达水平的增加表示差异表达的替代物启动子。

[0062] 在另一个实施方案中,对于缺失性启动子,与一个或多个参考样品中一种或多种预选标志物的表达水平相比,癌性生物样品中一种或多种预选标志物的表达水平降低表示差异表达的替代启动子。

[0063] 可以以多种方式鉴定差异表达的替代启动子。在一个实施方案中,由获得性启动子鉴定本文所述差异表达的替代启动子,该获得性启动子的转录物表达水平是其在的一组肿瘤样品中中值表达水平的至少4倍。在另一个实施方案中,由缺失性启动子鉴定本文所述差异表达的替代启动子,该缺失性启动子的转录物表达水平不到其在一组肿瘤样品中中值表达水平的0.25倍或至少4倍降低(4-fold reduction)。在一个实施方案中,每个启动子的转录物的中值表达水平是一组肿瘤样品中所有样品的启动子的中值表达水平。

[0064] 在另一个实施方案中,由相对启动子活性大于一组肿瘤样品的中值启动子活性的获得性启动子鉴定本文所述差异表达的替代启动子。在另一个实施方案中,由相对启动子活性低于一组肿瘤样品的中值启动子活性的缺失性启动子鉴定本文所述差异表达的替代启动子。在一个实施方案中,每个启动子的转录物的中值表达水平是一组肿瘤样品中所有样品的启动子的中值表达水平。

[0065] 在一个实施方案中,获得性启动子是与非癌性生物样品相比在癌性生物样品中获得性的或活性增强的启动子。获得性启动子可以是癌性生物样品中的启动子而不存在于非癌性生物样品中,或者是与非癌性生物样品中的启动子相比,在癌性生物样品中启动子活性增强的启动子。

[0066] 在另一个实施方案中,缺失性启动子是与非癌性生物样品相比在癌性生物样品中缺失性的或活性降低的启动子。缺失性启动子可以是存在于非癌性生物样品中但不存在于癌性生物样品中的启动子,或与非癌性生物样品中的启动子相比,在癌性生物样品中启动子活性降低的启动子。

[0067] 在鉴定本文所述差异表达的替代启动子后,然后计算替代启动子利用评分。在一个实施方案中,替代启动子利用评分可以是AP评分。在另一个实施方案中,替代启动子利用评分可以是APB评分。可以通过确定本文所述差异表达的替代启动子之和来计算替代启动子利用评分。

[0068] 替代启动子利用评分可用于鉴定患有癌症的患者是否受益于ICI疗法的方法中。鉴定患有癌症的患者是否受益于ICI疗法的方法还可以包括将本文所述的替代启动子利用评分与参考评分进行比较,以鉴定患者是否受益于ICI疗法的步骤。

[0069] 在一个实施方案中,参考评分是位于一个或多个参考样品的替代启动子利用评分的中值、三分位数或四分位数的评分。在另一个实施方案中,参考评分是位于一个或多个参考样品的替代启动子利用评分的第10、第20、第30、第40、第50、第60、第70、第80或第90个百分位数的评分。在优选实施方案中,参考评分可以是位于一个或多个参考样品的替代启动

子利用评分的第66或第80个百分位数的评分。

[0070] 在优选实施方案中,参考评分是绝对评分。与患者的替代启动子评分相比,绝对评分是一个固定的、非可变的评分。

[0071] 在一个实施方案中,与参考评分相比,增加的替代启动子利用评分将患者鉴定为未受益于ICI疗法。在一个实施方案中,增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。在另一个实施方案中,增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

[0072] 在另一个实施方案中,与参考评分相比,降低的替代启动子利用评分将患者鉴定为受益于ICI疗法。在一个实施方案中,降低的替代启动子利用评分是低于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。在一个实施方案中,降低的替代启动子利用评分是低于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

[0073] 在一个方面,本发明提供了确定正在接受或已经接受ICI疗法的患者预后的方法,该方法包括以下步骤:

[0074] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0075] b) 基于步骤(a)中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0076] c) 计算替代启动子利用评分;

[0077] d) 使用替代启动子利用评分确定患者的预后。

[0078] 在一个实施方案中,本文所述确定正在接受或已经接受ICI疗法的患者预后的方法还包括,将本文所述替代启动子利用评分与参考评分进行比较,以确定患者的预后的步骤,其中与参考评分相比,增加的替代启动子利用评分表示预后较差。

[0079] 在一个实施方案中,增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。在另一个实施方案中,增加的替代启动子利用评分是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

[0080] 在另一个实施方案中,本文所述确定正在接受或已经接受ICI疗法的患者预后的方法还包括,将本文所述替代启动子利用评分与参考评分进行比较,以确定患者的预后的步骤,其中与参考评分相比,降低的替代启动子利用评分表示预后更好。

[0081] 在一个实施方案中,降低的替代启动子利用评分是低于一个或多个参考样品的替代启动子利用评分的第66个百分位数的评分。在另一个实施方案中,降低的替代启动子利用评分是低于一个或多个参考样品的替代启动子利用评分的第80个百分位数的评分。

[0082] 在另一方面,本发明提供了用于将患有癌症的患者从ICI疗法中排除的生物标志物,其中该生物标志物是高于一个或多个参考样品的替代启动子利用评分的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0083] 在一个实施方案中,本发明提供了用于排除患有癌症的患者进行ICI疗法的生物标志物,其中该生物标志物是高于一个或多个参考样品的替代启动子利用评分的第66个百分位数的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0084] 在另一个实施方案中,本发明提供了用于选择患有癌症的患者进行ICI疗法的生物标志物,其中该生物标志物是低于一个或多个参考样品的替代启动子利用评分的第66个百分位数的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0085] 在另一方面,本发明提供了用于将患有癌症的患者从ICI疗法中排除的生物标志物,其中该生物标志物是高于一个或多个参考样品的替代启动子利用评分的第80个百分位数的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0086] 在另一个实施方案中,本发明提供了用于选择患有癌症的患者进行ICI疗法的生物标志物,其中该生物标志物是低于一个或多个参考样品的替代启动子利用评分的第80个百分位数的替代启动子利用评分,其中替代启动子利用评分是每个替代启动子位点处差异表达的替代启动子之和。

[0087] 一个或多个参考样品可以选自由以下组成的组:从一个或多个不同患者获得的一个或多个肿瘤样品、从一个或多个不同患者获得的一个或多个非癌性样品、从同一患者获得的一个或多个非癌性样品及其组合。在一个实施方案中,从同一患者获得的一个或多个非癌性样品是癌性组织附近的组织。本领域技术人员应当理解,这些参考样品的多种组合可用于本发明的方法中。

[0088] 在另一方面,本发明涉及确定患有癌症的患者是否不应接受免疫检查点抑制剂(ICI)疗法治疗的方法,该方法包括以下步骤:

[0089] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0090] b) 基于步骤(a)中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0091] c) 计算替代启动子利用评分;

[0092] d) 使用替代启动子利用评分鉴定患者未受益于ICI疗法;

[0093] e) 当所述患者被确定为未受益于ICI疗法时,用替代疗法治疗该患者。

[0094] 在另一方面,本发明涉及治疗患有癌症的患者的方法,该方法包括以下步骤:

[0095] a) 测量从患者获得的癌性生物样品中一种或多种预选标志物的表达水平;

[0096] b) 基于步骤(a)中测量的一种或多种预选标志物的表达水平,鉴定差异表达的替代启动子;

[0097] c) 计算替代启动子利用评分;

[0098] d) 使用替代启动子利用评分鉴定患者是否受益于ICI疗法;

[0099] e) 用免疫检查点抑制剂治疗根据步骤(d)鉴定为受益于ICI疗法的患者,或用替代疗法治疗根据步骤(d)鉴定为未受益于ICI疗法的患者。

[0100] 在一个实施方案中,ICI疗法可包括但不限于易普利姆玛、纳武单抗、派姆单抗、阿特珠单抗、阿维单抗(avelumab)、德瓦鲁单抗(durvalumab)、西米普利单抗(cemiplimab)、司利珠单抗(spartalizumab)、信迪利单抗(sintilimab)、卡瑞利珠单抗和替雷利珠单抗。

[0101] 在另一个实施方案中,替代疗法是除ICI疗法之外适合治疗癌症的任何疗法。例如,替代疗法可以包括化学疗法、放射疗法、干细胞移植、手术、激素疗法和靶向疗法。靶向疗法涉及使用干扰肿瘤生长和进展所需的特定分子的药物。替代疗法还可以包括在接受

ICI疗法的患者中停止ICI疗法。

[0102] 在缺少本文没有具体公开的任何一种要素或多种要素、一种限制或多种限制的情况下,可以适当地实施本文说明性描述的本发明。因此,例如,术语“包含/包括(comprising)”、“包括(including)”、“含有(containing)”等应被广泛地而非限制地解读。此外,本文中使用的术语和措辞已用作描述性而非限制性术语,并且在使用此类术语和措辞时无意排除所示和所述特征或其部分的任何等同物,而是认识到在要求保护的本发明的范围内可以进行各种修改。因此,应当理解,虽然已经通过优选实施方案和可选特征具体公开了本发明,但是本领域技术人员可以诉诸本文公开的其中体现的本发明的修改和变化,并且这样的修改和变化视为在本发明的范围内

[0103] 本文已经广泛和一般性地描述了本发明。属于一般性公开内容的每个较窄的类别和亚属归组也构成本发明的一部分。这包括本发明的一般性描述,附带条件或否定限制是从属中除去任何主题,无论文本是否具体引用了除去的材料。

[0104] 其他实施方案在以下权利要求和非限制性实施例内。此外,如果根据马库什组描述了本发明的特征或方面,则本领域技术人员应当认识到也因此根据马库什组的任何个体成员或成员子组描述了本发明。

## 实施例

[0105] 将参考具体实施例对本发明的非限制性实施例和比较例进行更详细的描述,无论如何不应将具体实施例解释为限制本发明的范围。

[0106] 材料和方法

[0107] 本研究遵循了用于验证肿瘤生物标志物的REMARK标准。

[0108] 临床队列

[0109] 发现队列

[0110] 该队列包括在韩国首尔三星医疗中心(Samsung Medical Centre)接受纳武单抗或派姆单抗治疗的连续转移性胃癌患者。在至少一种细胞毒性治疗方案无效的患者中施用ICI作为补救治疗。纳武单抗3mg/kg以每2周1小时输注施用,派姆单抗200mg以每3周30分钟静脉内输注施用,直至疾病进展或出现不可接受的毒性。获得了伦理批准,在收集原发性肿瘤的存档肿瘤组织样本并前瞻性随访生存数据之前,所有患者都提供了书面知情同意书。

[0111] 派姆单抗试验队列

[0112] 本研究招募的患者患有组织学证实的转移性和/或复发性胃腺癌,且包括铂/氟嘧啶在内的至少1线化疗对这些患者无效。该试验根据Declaration of Helsinki and the Guidelines for Good Clinical Practice(ClinicalTrials.gov identifier:NCT#02589496)(赫尔辛基宣言和良好临床实践指南(ClinicalTrials.gov标识符:NCT#02589496))进行。试验方案获得三星医疗中心(韩国首尔)机构审查委员会(Institutional Review Board of Samsung Medical Center(Seoul,Korea))批准,所有患者在招募前均提供了书面知情同意书。派姆单抗200mg以每3周30分钟静脉输注施用,直到记录到疾病进展、不可接受的毒性或长达24个月。根据RECIST1.1标准,每两个周期评估肿瘤响应。

[0113] NanoString分析

[0114] 设计NanoString nCounter Reporter CodeSets,用于80个常发生(recurrent)体

细胞备选启动子相关基因以及对应于肿瘤内溶细胞活性 (CYT)、细胞因子和免疫检查点的免疫相关基因。为每个基因设计了至少两个探针,以测量规范启动子和备选启动子驱动的转录物的表达。由未改变的H3K4me3标记的5'转录物处的规范探针,以及体细胞启动子5'转录物处的备选探针。使用供应商提供的nCounter软件 (nSolver) 进行数据分析。使用每个CodeSet中包括的内部阳性对照探针的几何平均值对原始计数进行归一化。

#### [0115] RNA测序

[0116] 在研究治疗开始前的第-42天和第1天之间获得肿瘤组织。如果在彻底病理检查后估计肿瘤含量超过40%,则使用QIAamp Mini Kit (Qiagen,Hilden,德国) 根据制造商的说明从新鲜获得的组织中提取肿瘤DNA和RNA。使用ND1000分光光度计 (Nanodrop Technologies,Thermo-Fisher Scientific,MA,美国) 测量浓度以及260/280和260/230nm比率,然后使用Qubit荧光计 (Life Technologies,CA,美国) 进一步量化DNA/RNA。

#### [0117] RNA转录组分析

[0118] 使用TopHat将RNAseq数据与GENCODE v19转录物注释比对,并使用Cufflinks生成FPKM丰度测量。然后在所有样品中合并转录物,并使用Cuffnorm进行标准化。为了分析替代启动子相关表达,将RNAseq读取 (read) 映射到先前通过表观基因组分析鉴定的基因组位置。然后量化映射到这些表观基因组定义的启动子区的RNAseq,并通过启动子长度和文库大小进行归一化。最后,计算每个肿瘤和所有肿瘤样品的中值表达水平之间每个启动子位点处表达的倍数变化。

[0119] PDL1免疫组织化学分析、MSI状态、EBV状态、TCGA亚分型和肿瘤突变负荷基于用于转移性胃癌中单药剂派姆单抗II期研究的分类。

#### [0120] 统计分析

[0121] 使用费希尔精确检验 (Fisher's Exact Test) 进行临床病理学特征与组织学亚分类的关联。根据派姆单抗首次剂量时间至疾病进展或死亡时间计算无进展生存期 (PFS), 根据派姆单抗或纳武单抗首次剂量时间至死亡时间计算总生存期 (OS)。Kaplan-Meier (KM) 曲线和对数秩检验用于生存分析。使用Cox比例风险回归模型评估每个分析的风险比 (HR) 及其95%置信区间 (CI)。所有分析均使用R (3.4.1) 完成。在验证队列中,备选启动子利用评分大于第66个百分位数的样品被定义为高备选启动子利用 ( $AP_{high}$ ), 剩余样品定义为低备选启动子利用 ( $AP_{low}$ )。

#### [0122] PDL1免疫化学

[0123] 使用Dako PD-L1 IHC 22C3 pharmDx试剂盒 (Agilent Technologies) 进行PDL1免疫组织化学。使用CPS确定PD-L1蛋白表达,CPS是PD-L1染色细胞数 (肿瘤细胞、淋巴细胞、巨噬细胞) 除以存活肿瘤细胞总数,再乘以100。

#### [0124] MSI状态

[0125] 通过MLH1和MSH2的IHC二者以及具有单核苷酸重复的5个标志物的PCR分析确定肿瘤组织MSI状态。

#### [0126] EBV状态亚型

[0127] 通过EBV编码的小RNA (EBER) 原位杂交确定EBV状态。

#### [0128] 肿瘤突变负荷

[0129] 根据全外显子组测序分析测定突变负荷。个体的突变负荷定义为通过所有过滤器

的体细胞非同义SNV的数量。体细胞突变用变异效应预测因子 (variant effect predictor) 进行注释。突变负荷计算为肿瘤外显子组数据中非同义SNV的数量。组织的ML-H阈值被设置为较高三分位数 (upper tertile)。

#### [0130] TCGA亚型定义

[0131] TCGA定义的胃癌亚型基于DNA基因组改变。这些组包括EBV (+)、MSI-H、CIN和基因组稳定肿瘤,其缺乏CIN并且在弥漫性组织学亚型中大量富集。关于CIN的代表性 (proxy), EBV (-)、MSS肿瘤根据其TP53状态分为CIN和基因组稳定的。使用R中的deconstructSigs软件包 (v1.6.0) 进行突变特征分析。

#### [0132] 数据集

[0133] 根据从TCGA的PanCanAtlas获得的RNA-Seq数据推断启动子活性,TCGA由33个肿瘤类型的10393个样品 (9668个肿瘤和725个正常样品) 组成。

#### [0134] APBscore算法

[0135] 胃癌的最初表观遗传启动子改变研究确定了2732个体细胞启动子区 (2053个获得性启动子,679个缺失性启动子),这些同种型的转录组表达被用于创建APBscore算法。

[0136] 对于TCGA数据,使用GENCODE v19注释来确定启动子集。将每个TSS的重叠第一外显子组合以获得一组启动子。然后使用与组成性转录物的第一个内含子对齐的连接读取来量化每个启动子的活性。然后在整个数据集中对总连接读取计数进行标准化。 $\log_2$ 转化的归一化读取计数用于进一步下游分析中的启动子活性。通过总结每个基因的每个启动子的活性来获得基因表达估计。然后通过基因表达对每个启动子的活性进行归一化,以获得相对启动子活性。总共确定了113,076个启动子。

[0137] 然后从这组113,076个启动子中选择在胃癌中确定的体细胞启动子区。4672个启动子 (3263个在获得性区,1409个在缺失性区) 位于在胃癌中确定的2732个体细胞启动子区内。在这4672个启动子中,计算了所有肿瘤样品 ( $n=9668$ ) 中每个启动子的中值相对启动子活性。对于每个样品,将APBscore计算为相对启动子活性大于中值的获得性启动子和相对启动子活性小于中值的缺失性启动子的数量 (图1)。使用位于不同百分位数截止值 (第10、第20、第30……、第90) 的APBscore将肿瘤二分为 $APB_{high}$ 和 $APB_{low}$ 。第80个百分位数被确定为 $APB_{high}$ 和 $APB_{low}$ 组的理想截止值,并用于进一步的下游分析。仅将正常样品添加到涉及正常组织的分析队列中,而不用于计算中值评分和截止值。

#### [0138] 免疫相关物

[0139] CD8A、GZMA和PRF1的转录组表达水平以及700多个基因的选择,包括免疫检查点、各种免疫细胞类型的标志物、与适应性和先天免疫应答相关的基因和抗原,是从来自Broad GDAC Firehose的TCGA数据集中选择和提取的。免疫亚型、其他免疫特征和TMB是从泛癌免疫全局 (landscape) 分析中提取的。无进展生存期 (PFS) 和总生存期 (OS),包括审查数据,是从泛癌免疫全局分析中提取的。

#### [0140] 免疫检查点抑制剂临床数据

[0141] 涉及免疫检查点抑制剂的临床试验数据正在迅速扩大,其中药物作为单一药剂进行测试,并与其他疗法相组合。派姆单抗仍是已在最广泛的肿瘤类型中进行了测试并报告了临床试验数据的免疫检查点抑制剂。因此,对单药剂派姆单抗试验进行了广泛的文献回顾,并选择了针对每种肿瘤类型报道的最大研究。然后将这些试验的客观响应率 (ORR) 与每

种肿瘤类型的APBscore相关联。TMB是从泛癌免疫全局分析中提取的，PD-L1转录表达是从Broad GDAC Firehose中提取的。

#### [0142] 统计分析

[0143] 威尔科克森秩和检验 (Wilcoxon rank-sum test) 用于比较APB<sub>high</sub>组与APB<sub>low</sub>组之间CD8A、GZMA和PRF1的表达水平。“Rtsne”R软件包用于生成T-SNE图。皮尔逊检验 (Pearson test) 用于APBscore和派姆单抗ORR之间的相关性。Kaplan-Meier (KM) 曲线和对数秩检验用于生存分析。使用Cox比例风险回归模型评估每个分析的风险比 (HR) 及其95%置信区间 (CI)。使用Reactome Pathway Database完成体细胞启动子基因功能的网络映射。所有分析均使用R (3.5.2) 完成。

#### [0144] 使用RNA-Seq验证启动子活性的估计

[0145] 通过针对其他“金标准”措施 (例如可在公共数据集上获得的H3K4me3ChIP-Seq和CAGE标签数据) 进行基准对比, 来验证使用RNA-Seq估计启动子活性的准确性。H3K4me3水平与RNA-Seq启动子活性密切相关, 表明基于表观遗传和基于转录物的启动子活性估计是一致的 (Kruskal-Wallis  $p < 0.001$ )。CAGE-Tag数据证实了这一发现, 同时在RNA-Seq上确定的独特启动子具有更高的CAGE-Tag支持。

[0146] 本算法还与RNA-Seq定量和第一外显子读取计数中使用的其他生物信息学方法进行了比较。本算法产生的结果与这些算法相似, 具有高相关性水平 (皮尔逊相关系数  $> 0.85$ )。总体而言, 该分析表明, 我们的方法能够根据RNA-Seq数据对启动子活性进行定量、稳健和可重复的估计。

#### [0147] 实施例1: 转移性胃癌中的备选启动子利用

[0148] 第一个队列由24名接受纳武单抗和派姆单抗治疗的转移性胃癌患者 (最初包括29名个体, 其中24个肿瘤样品通过质量控制以获得足够的组织用于Nanostring分析) 组成。定制的Nanostring组 (Nanostring panel) 用于测量与规范启动子或备选启动子相关的转录物。差异表达的替代启动子被定义为这样一个启动子位点, 其表达水平相对于所有样品的中值显示出  $< 0.25$  倍变化 (对于缺失性体细胞启动子) 或  $> 4$  倍变化 (对于获得性体细胞启动子)。使用该算法, 发现三分之一的肿瘤 (8/24) 在超过10%的位点 ( $> 8/80$ ) 中显示出高备选启动子利用。该组被定义为AP<sub>high</sub>, 而其余的被定义为AP<sub>low</sub> (图2A)。

[0149] 先前已通过研究CD8A (CD8+肿瘤浸润淋巴细胞)、颗粒酶A (GZMA) 和穿孔蛋白1 (perforin 1, PRF1) 的表达描述了溶细胞性T细胞活性的测量。与AP<sub>high</sub>组相比, AP<sub>low</sub>组表现出GZMA ( $P = 0.025$ )、PRF1 ( $P = 0.011$ ) 和CD8A ( $P = 0.059$ ) 的表达显著增加, 表明在AP<sub>low</sub>组中细胞毒性T细胞活性增强 (图2B)。这些发现与早期胃癌中描述的早期结果一致, 表明转移性胃癌中的备选启动子利用也与抗肿瘤免疫呈负相关。值得注意的是, 尽管该队列的样品量小、治疗方案不同且具有非试验性质, 但AP<sub>high</sub>肿瘤患者与AP<sub>low</sub>肿瘤患者相比倾向于具有更差的无进展生存期 (PFS) (129天相比389天, HR 1.96 95%CI: 0.55-6.93,  $P = 0.29$ , 图2C)。基于这些发现, 在统一治疗的患者单独队列中进一步检验了这一假设。

#### [0150] 实施例2: 备选启动子利用作为派姆单抗治疗响应和生存的预测因子

[0151] 对于第二个队列, 使用了早期描述的II期研究的转录组数据。来自治疗前活检样品的转录组数据和匹配的临床数据可用于37名个体并用于分析。中值年龄为57岁, 73%为男性 ( $N = 27$ ), 4名 (11%) 为EBV阳性, 4名为MSI (11%), 其余定义为CIN或GS TCGA亚型。在11

名个体 (30%) 中观察到对治疗的完全或部分响应。使用先前在胃癌中确定的2732个体细胞备选启动子位点 [10], 差异表达的替代启动子的定义类似于第一队列 (针对已知的体细胞缺失性启动子, <0.25倍变化; 或针对已知的体细胞获得性启动子, >4倍变化)。值得注意的是, 之前已经示出用于评估备选启动子利用的RNAseq平台与Nanostring平台之间的良好一致性。计算每个样品中差异表达位点之和, 以定义备选启动子利用评分 (图3A)。评分范围从37-426 (中值136) (图3B)。使用第一队列的数据指导截止点, AP<sub>high</sub> 组被定义为>66%百分位数的样品 (n=13), 而AP<sub>low</sub> 构成剩余样品。

[0152] 与AP<sub>low</sub> 组相比, AP<sub>high</sub> 组在年龄、性别或组织学亚型的临床病理特征方面没有统计学上的显著差异。也未检测到两组在TCGA亚型、突变负荷和PDL1 CPS评分之间的差异 (表1)。与AP<sub>high</sub> 组相比, AP<sub>low</sub> 组表现出CD8A (P=0.0037)、GZMA (P=0.0055) 和PRF1 (P=0.016) 的表达显著增加, 表明AP<sub>low</sub> 组中的细胞毒性T细胞活性增强 (图3C)。客观响应率 (ORR), 定义为对治疗的部分或完全响应, 在AP<sub>low</sub> 组中与AP<sub>high</sub> 组相比较 (10/24相比1/13, P=0.03) (图3D)。值得注意的是, 在AP<sub>high</sub> 组中, 唯一的响应在MSI亚型肿瘤中。AP<sub>high</sub> 组的中值无进展生存期 (PFS) 为55天, 而AP<sub>low</sub> 组为180天 (对数秩P=0.0076) (图4A、5B)。AP<sub>low</sub> 组有17% EBV (n=4) 和12% MSI (n=3) TCGA亚型样品, 而AP<sub>high</sub> 组只有8% MSI样品 (n=1) 且没有EBV样品 (图4C)。如先前所示, 各种TCGA亚型之间的PFS不同 (P=0.0026), MSI和EBV亚型的生存期明显更长 (491天 (MSI/EBV) 相比80天 (CIN/GS))。值得注意的是, 在CIN/GS亚型中, AP<sub>low</sub> 与AP<sub>high</sub> 组之间的PFS也有统计学显著差异 (48天 (CIN/GS AP<sub>high</sub>) 相比161天 (CIN/GS AP<sub>low</sub>), P=0.0019) (图4D)。根据初步的总体生存数据, 在AP<sub>low</sub> 组中观察到改善的生存期的趋势 (340天相比292天, P=0.16)。临床病理学和备选启动子利用的多变量分析揭示, 高备选启动子利用是派姆单抗PFS的独立预测因子 (HR 0.29, (95%CI 0.099-0.85), P=0.024) (表2)。

[0153] 表1. AP<sub>high</sub> 与AP<sub>low</sub> 组之间的特征差异。

特征	各 AP 组		P-值
	AP <sub>High</sub> (n = 13) n (%)	AP <sub>Low</sub> (n = 24) n (%)	
<b>年龄, 岁</b>			
<57	6 (46)	11 (46)	1
≥57	7 (54)	13 (54)	
<b>性别</b>			
男性	11 (85)	16 (67)	0.44
女性	2 (15)	8 (33)	
<b>原发性肿瘤位置</b>			
身体	6 (46)	17 (71)	0.17
窦/其他	7 (54)	7 (29)	
<b>组织学</b>			
低分化	5 (38)	15 (62)	0.34
完全/中等分化	5 (38)	5 (21)	
印戒(Signet Ring)/其他	3 (23)	4 (17)	
<b>免疫特征</b>			
高	5 (38)	14 (58)	0.31

[0154]

	低	8 (62)	10 (42)	
	<b>间充质亚型(ACRG)</b>			
	非间充质	9 (69)	22 (92)	0.16
	间充质	4 (31)	2 (8)	
	<b>突变负荷</b>			
	低	3 (23)	13 (54)	0.13
	中等	8 (62)	7 (29)	
	高	2 (15)	4 (17)	
	<b>TCGA 亚型</b>			
	CIN	7 (54)	6 (25)	0.26
	GS	5 (38)	11 (46)	
	EBV	0 (0)	4 (17)	
	MSI	1 (8)	3 (12)	
	<b>P53 突变状态</b>			
	突变的	6 (46)	17 (71)	0.17
	野生型	7 (54)	7 (29)	
	<b>KRAS 突变状态</b>			
	突变的	3 (23)	4 (17)	0.68
	野生型	10 (77)	20 (83)	
	<b>PDL1 CPS*</b>			
	0	8 (62)	6 (29)	0.22
	1	3 (23)	5 (24)	
	2-9	0 (0)	3 (14)	
	≥10	2 (15)	7 (33)	

[0156] 费希尔检验

[0157] \*3个样品没有可用于通过免疫组织化学进行PDL1 CPS评分的组织,均来自AP<sub>low</sub>组。

[0158] 表2.单变量和多变量生存分析

变量	单变量		多变量	
	HR (95% CI)	P 值	HR (95% CI)	P 值
备选启动子组: 高 vs 低	0.36(0.17-0.78)	0.0094	0.29 (0.099 – 0.85)	0.024
TCGA 亚型: CIN/GS 相比 EBV/MSI	0.10(0.024-0.45)	0.0024	0.31 (0.052 – 1.88)	0.20
年龄	0.98 (0.95 – 1)	0.061		
性别: 女性 vs 男性	1.2 (0.52 – 2.5)	0.72		
原发性肿瘤位点: 身体 vs 窦/其他	0.61 (0.28 – 1.3)	0.22		
组织学 分化低/Signet/其他 vs 完全分化/中等分化	0.98(0.44 – 2.2)	0.97		

[0159]

[0160]	突变负荷 高 vs 低/中等	5.67(1.3 24)	0.018	3.46 (0.52 – 23)	0.20
	<i>P53</i> 突变状态 突变的 vs 野生型	0.74 (0.36 – 1.56)	0.43		
	<i>KRAS</i> 突变状态 突变的 vs 野生型	2.72 (0.92 – 7.96)	0.068		
	间充质亚型 间充质 vs 非间充质	0.73(0.29 – 1.8)	0.48		
	免疫特征 高 vs 低	2.31(1.1 – 4.8)	0.025	1.82 (0.80 – 4.12)	0.15
	PDL1 CPS ≥ 1 vs 0	7.62 (3 – 19)	<0.001	3.09 (1.15 – 8.29)	0.025

[0161] 实施例3:派姆单抗治疗后的备选启动子利用演变

[0162] 配对活检样品可用于第二队列的8名个体,这提供了监测ICI治疗压力导致的肿瘤演变的机会。治疗后的活检采集自于派姆单抗进展点的原发性胃肿瘤。在这8名个体中,两名有部分响应(PR),响应持续时间为211和491天(均为AP<sub>low</sub>),一名病情稳定(SD),响应持续时间为167天(AP<sub>low</sub>),5名有进展性疾病(PD)(AP<sub>high</sub> N=3;AP<sub>low</sub> N=2)作为最佳响应。有趣的是,基于临床响应,在替代启动子利用的方向性中观察到非常一致的变化。具体而言,与治疗前活检样品相比,具有PR和SD的肿瘤在治疗后的活检样品中表现出的备选启动子利用评分增加了1.5倍或更高,而所有5个具有PD的肿瘤在治疗后的活检样品中表现出备选启动子利用评分减少(费希尔精确检验,P=0.018)。这些结果进一步支持了替代启动子全局与ICI治疗压力之间的关系。

[0163] 实施例4:多种肿瘤类型的备选启动子利用

[0164] 为了研究除胃癌以外这些发现是否还适用于其他肿瘤类型,使用最近描述的算法推断PanCanAtlas RNA-seq数据库中33种肿瘤类型的10,393个样品(9668个肿瘤和725个正常样品)的启动子活性(表3)。简而言之,使用Gencode(第19版)注释编译了一组113,076个可能的启动子。假设具有相同或非常接近的TSS的同种型受相同启动子调节,通过使用独特的连接读取量化在每个启动子处开始的表达来推断启动子活性。

[0165] 表3. TCGA队列描述和样品大小

肿瘤类型	TCGA代码	肿瘤样品 (n)	正常样品 (n)	肿瘤+正常 (n)
肾上腺皮质癌	ACC	79	0	79
膀胱尿路上皮癌	BLCA	405	19	424
乳腺浸润性癌	BRCA	1095	112	1207
宫颈鳞状细胞癌和宫颈内腺癌	CESC	306	3	309
胆管癌	CHOL	36	9	45
结肠腺癌	COAD	281	41	322
弥漫性大B细胞淋巴瘤	DLBC	48	0	48
食管癌	ESCA	185	11	196
多形性胶质母细胞瘤	GBM	165	5	170
头颈部鳞状细胞癌	HNSC	504	44	548

肾嫌色细胞	KICH	66	25	91
肾脏肾透明细胞癌	KIRC	530	72	602
肾脏肾乳头状细胞癌	KIRP	291	32	323
急性髓系白血病	LAML	171	0	171
脑低级别胶质瘤	LGG	528	0	528
肝脏肝细胞癌	LIHC	373	50	423
肺腺癌	LUAD	506	59	565
肺鳞状细胞癌	LUSC	500	51	551
间皮瘤	MESO	87	0	87
卵巢浆液性囊腺癌	OV	307	0	307
胰腺癌	PAAD	179	4	183
嗜铬细胞瘤和副神经节瘤	PCPG	184	3	187
前列腺癌	PRAD	498	52	550
直肠腺癌	READ	95	10	105
肉瘤	SARC	263	2	265
皮肤黑色素瘤	SKCM	472	1	473
胃腺癌	STAD	415	35	450
睾丸生殖细胞肿瘤	TGCT	156	0	156
甲状腺癌	THCA	509	59	568
胸腺瘤	THYM	120	2	122
子宫体子宫内膜癌	UCEC	177	24	201
子宫癌肉瘤	UCS	57	0	57
葡萄膜黑色素瘤	UVM	80	0	80
组合		9668	725	10393

[0167] 在113,076个启动子中,4672个启动子映射到先前在胃癌中定义的2732个肿瘤相关启动子区。假设这组启动子的改变可能代表对宿主免疫识别的普遍泛肿瘤响应,不仅限于胃癌。为了验证这个假设,用4672个启动子计算每个肿瘤的APB水平(APBscore)(图5)。如图1所示,在肿瘤类型内和肿瘤类型之间都观察到了宽范围的APBscore。整个队列的中值APBscore为178(范围:46-241)。中值APBscore最低的肿瘤类型包括甲状腺癌(THCA)(中值165,范围:54-207)、弥漫性大B细胞淋巴瘤(DLBC)(中值:166,范围:108-196)和肺腺癌(LUAD)(中值:167,范围:141-215)。相比之下,胶质母细胞瘤(GBM)(中值:207,范围:172-241)、低级别胶质瘤(LGG)(中值:204,范围:130-230)和睾丸生殖细胞肿瘤(TGCT)(中值:189,范围:83-218)的中值APBscore最高。

[0168] 肿瘤APBscore与T细胞溶细胞活性的标志物相关:CD8A、GZMA和PRF1。与其他泛癌TCGA研究类似,对肿瘤类型不可知的整个TCGA样品队列进行主要分析,但也提供了肿瘤类型特异性结果。以前,观察到在中值或最高三分位数处对肿瘤进行二分产生了与CD8A、GZMA和PRF1的显著相关性。然而,由于这些研究是用较小的样品量完成的,因此对于与这三个基因的相关性,选择其来重新建立二分肿瘤的理想截止值。因此,在不同的百分位数(第10、第20、第30……、第90)处测试APBscore阈值,以二分 $APB_{high}$ 组和 $APB_{low}$ 组。对于整个队列(所有肿瘤类型),在二分法的所有九个截止值处,与 $APB_{low}$ 肿瘤相比, $APB_{high}$ 肿瘤的CD8A、GZMA和

PRF1水平显著较低 ( $p < 0.0001$ ) (图6A)。发现第80个百分位数与这三个基因的相关性最强, 33种肿瘤类型中有19种与CD8A、GZMA和PRF1显著相关。对于随后的分析, 在整个队列的第80个百分位数 (APBscore截止值:190) 处将肿瘤二分为APB<sub>high</sub>组和APB<sub>low</sub>组。重要的是, 即使使用了不同的技术来询问备选启动子, 先前在胃癌 (STAD) 中的发现在该分析中也得到了有力的复制 (图6B)。为了解决肿瘤类型的潜在混淆, 还进行了肿瘤类型特异性分析, 其中在该特定肿瘤类型的中值APBscore处将每种肿瘤类型二分为APB<sub>high</sub>组和APB<sub>low</sub>组。对于33种肿瘤类型中的20种, 与CD8A、GZMA和PRF1存在显著相关性。8种肿瘤类型在任何截止值处都没有显著相关性: 肾嫌色细胞 (KICH)、DLBC、胸腺瘤和胸腺癌 (THYM)、肾脏肾细胞癌 (KIRC)、急性髓系白血病 (LAML)、胆管癌 (CHOL)、子宫癌肉瘤 (UCS)、嗜铬细胞瘤和副神经节瘤 (PCPG)。在这8种中, CHOL、KICH和UCS的样品量相对较小 ( $n < 100$ ), 而DLBC、THYM、LAML是血液学/免疫起源的。

[0169] 值得注意的是, 与最初在用于限定APBscore的胃癌中定义的4672个启动子的组相比, 如果使用所有113,076个已确定的启动子或根据置换检验 (经验 $p < 0.001$ ) 获得的其他类似大小的启动子子集推断APBscore, 则未观察到与CD8A/GZMA/PRF1表达的相似相关性强度。这一观察结果支持这样一个假设, 即虽然APBscore最初源自胃癌, 但APBscore算法可以适用于多种肿瘤类型, 这可能反映了对宿主免疫的保守泛癌响应。为了研究肿瘤相关启动子同种型的具体功能, 分析了在至少100个肿瘤 (9668个肿瘤队列的约1%) 中上调的570个启动子 (4672个启动子的12%)。不同的肿瘤类型利用特定的启动子同种型簇, 基因功能的网络分析揭示了, 替代启动子影响具有不同作用的基因。这可能表明, 在全基因组范围内, 肿瘤中的备选启动子的选择可能较少受内在基因功能的驱动, 而更多受外在选择压力 (例如宿主抗肿瘤免疫) 的驱动。还将肿瘤APBscore与非恶性组织进行了比较。在PanCanAtlas分析的725个匹配的正常组织样品中, 只有5% ( $n = 36$ 个正常样本) 的样品被归类为APB<sub>high</sub> (在第80个百分位数截止值处)。当正常样品和肿瘤样品通过交替启动子利用聚类时, 正常样品聚类成接近但不同于其相应的肿瘤类型。在16种具有至少10个正常样品的肿瘤类型中的15种中, 与正常样品相比, 肿瘤样品更有可能被归类为APB<sub>high</sub> (费希尔精确,  $p < 0.001$ )。

[0170] 实施例5: 肿瘤和正常样品之间的备选启动子利用

[0171] 为了进一步探索备选启动子利用和肿瘤免疫之间的相互作用, 分析扩展到研究APBscore与广泛的涵盖多种免疫细胞类型、免疫检查点和抗原的约700种免疫相关基因之间的关系。与APB<sub>high</sub>肿瘤相比, 这些基因中的大多数 (78%) 都在APB<sub>low</sub>肿瘤中表现出显著更高的表达 (图7), 表明APB<sub>high</sub>肿瘤具有免疫耗竭表型。PanCanAtlas中可用的所有约20,000个基因的差异表达分析证实, APB<sub>high</sub>亚组中基因的下调仅限于免疫基因 (图8A)。在使用TCGA样品的泛癌免疫全局研究中, 肿瘤被分类为六种免疫亚型, 其特征是不同的免疫特征——伤口愈合、IFN- $\gamma$ 显性、炎性、淋巴细胞耗竭、免疫安静和TGF- $\beta$ 显性。当APBscore与这些免疫亚型相关时, 发现淋巴细胞耗竭或免疫安静亚型更有可能是APB<sub>high</sub>肿瘤 (费希尔精确 $p < 0.0001$ ) (图8B)。值得注意的是, 淋巴细胞耗竭的APB<sub>high</sub>肿瘤包含多种肿瘤亚型, 而免疫安静的APB<sub>high</sub>肿瘤主要由LGG主导。相比之下, IFN- $\gamma$ 亚型和炎性亚型更可能是APB<sub>low</sub>肿瘤 (费希尔精确 $p < 0.0001$ ) (图8B), 并且APB<sub>low</sub>肿瘤也更有可能是具有更高的淋巴细胞浸润特征评分和IFN- $\gamma$ 响应特征 ( $p < 0.0001$ , 威尔科克森检验) (图9、图10)。重要的是,

APBscore与TMB水平无关( $r=0.02, p=0.02$ ),表明不同的过程驱动备选启动子利用和DNA体细胞突变的获得。

[0172] 实施例6:备选启动子利用和生存结果

[0173] 研究了APBscore与根据TCGA Pan-Cancer分析获得的无进展生存期(PFS)和总生存期(OS)数据之间的关联。为了避免在泛癌水平上出现混淆,在这种水平上,生存结果可能受到内在组织特性或位点特异特性的影响,进行了个体肿瘤类型特异性分析。当在肿瘤类型特异性水平上进行分析时,大多数肿瘤类型在APBscore与PFS或OS之间没有表现出任何相关性。只有少数肿瘤类型在APB组(KIRC、LGG、LUAD、THYM)之间表现出生存差异。APBscore和疾病亚型的多变量分析也表明,APBscore不是未接受ICI疗法的患者生存的独立预后预测因子。

[0174] 实施例7:备选启动子利用和派姆单抗抗性

[0175] 为了测试APB<sub>high</sub>肿瘤是否如先前在转移性胃癌中示出的那样对ICI更具抵抗性,对独立研究中报道的APBscore与派姆单抗客观响应率(ORR)的关联进行了量化。在所有肿瘤类型中,观察到ORR与APBscore之间的显著负相关性(皮尔逊R(Pearson's R) = -0.46,  $p=0.025$ )。对于CD8A( $r=0.18, p=0.4$ )、GZMA( $r=0.17, p=0.4$ )或PRF1( $r=0.3, p=0.1$ )未观察到类似的相关性。与之前的研究一致,对于TMB( $r=0.53, p=0.0078$ )和PD-L1表达( $r=0.42, p=0.042$ )观察到正关联。因此,这些结果表明APBscore和TMB在确定不同的肿瘤子集方面可能是互补的,TMB选择具有更好ICI响应的类型,而APBscore确定那些缺乏响应的类型,因为APBscore和TMB不相关。有趣的是,当仅选择表现出与CD8A、GZMA和PRF1的显著APBscore相关性的肿瘤类型进行分析时(19种肿瘤类型),APBscore与ORR之间的相关性在强度和显著性上都有所提高:皮尔逊R = -0.55,  $p=0.019$ (图11)。总的来说,这些结果扩展了先前在胃癌中显示的发现,并建立了APBscore与派姆单抗抗性之间的泛癌关系。

[0176] 等同物

[0177] 上述实施例是出于说明本发明的目的而提出,不应被解释为对本发明的范围施加任何限制。很明显,在不脱离本发明的基本原理的情况下,可以对上文所述和实施例中说明的本发明的具体实施方案进行各种修改和改变。所有这些修改和改变都旨在被本申请所涵盖。

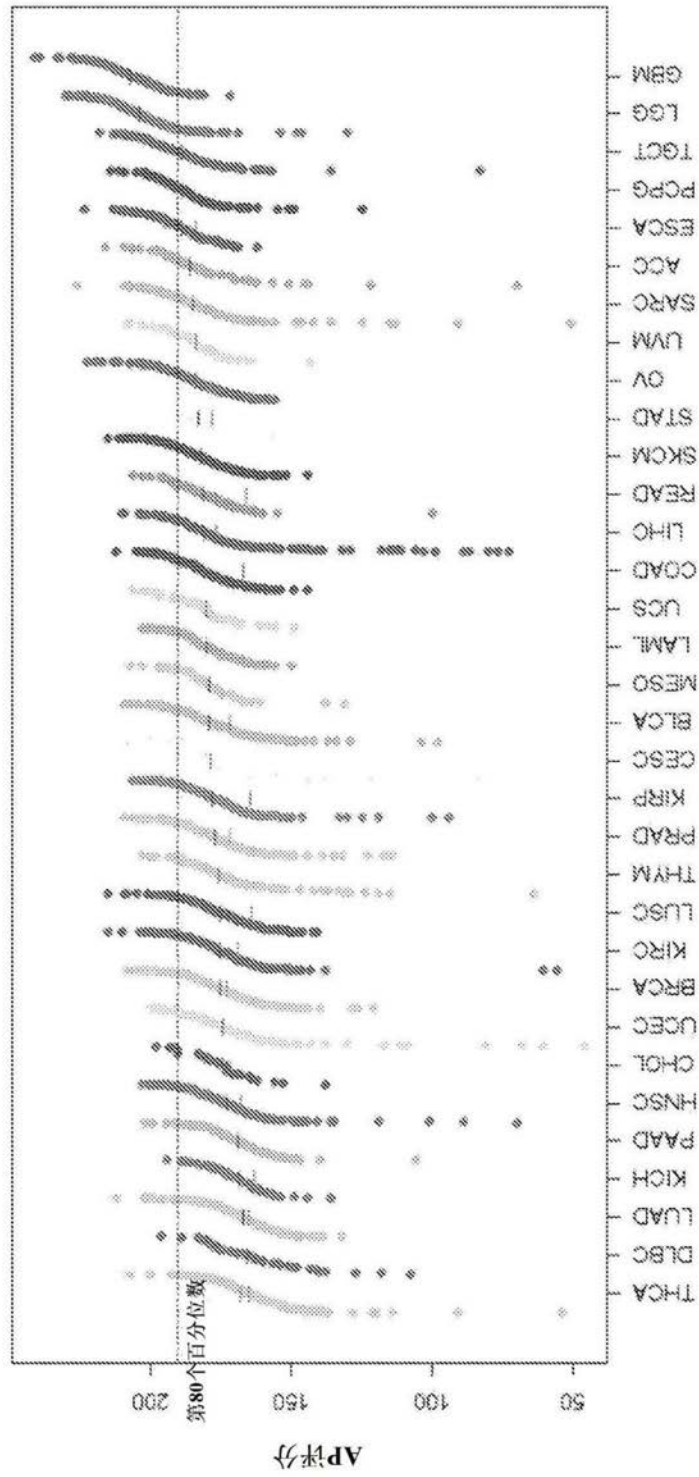


图1

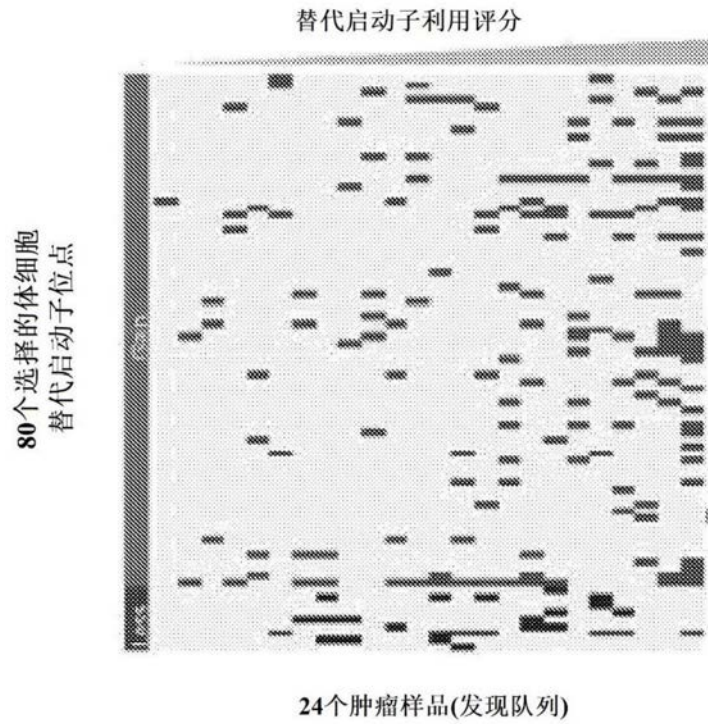


图2A

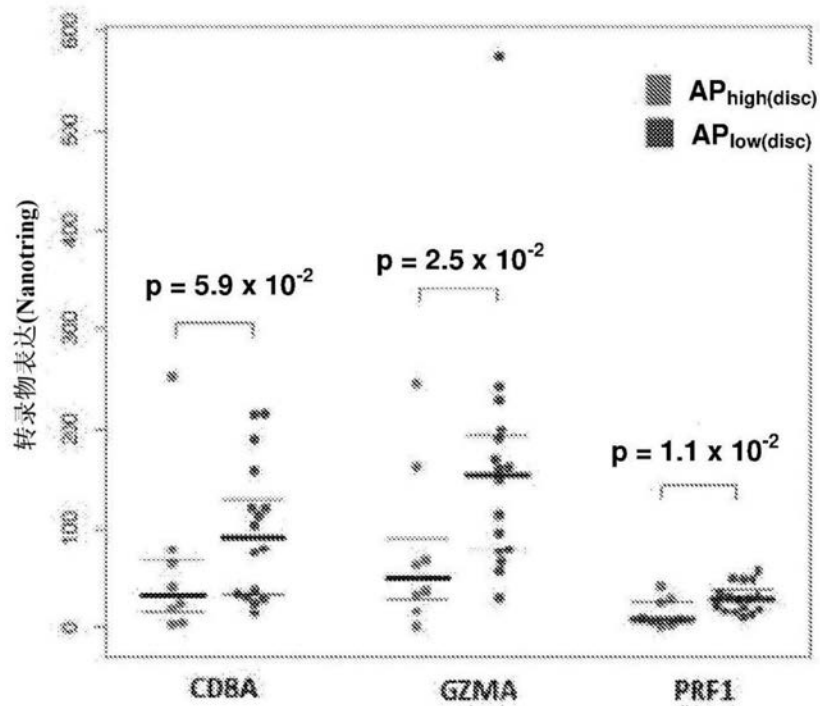


图2B

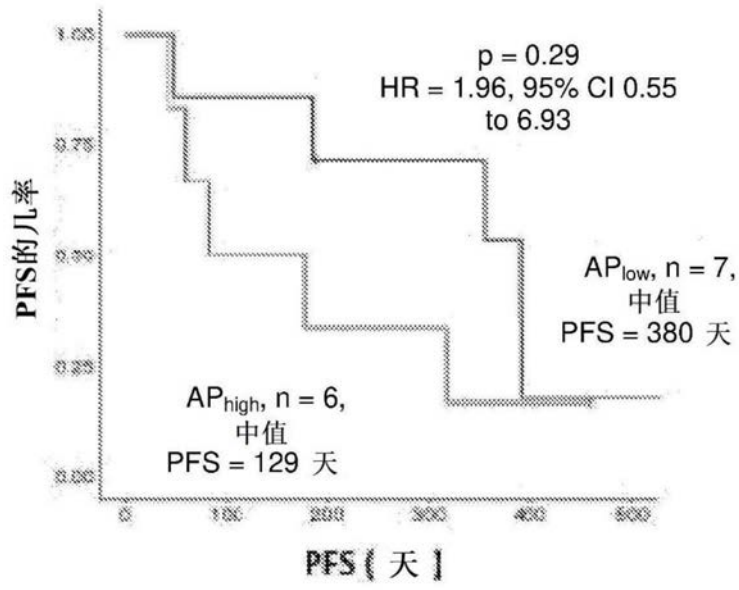


图2C

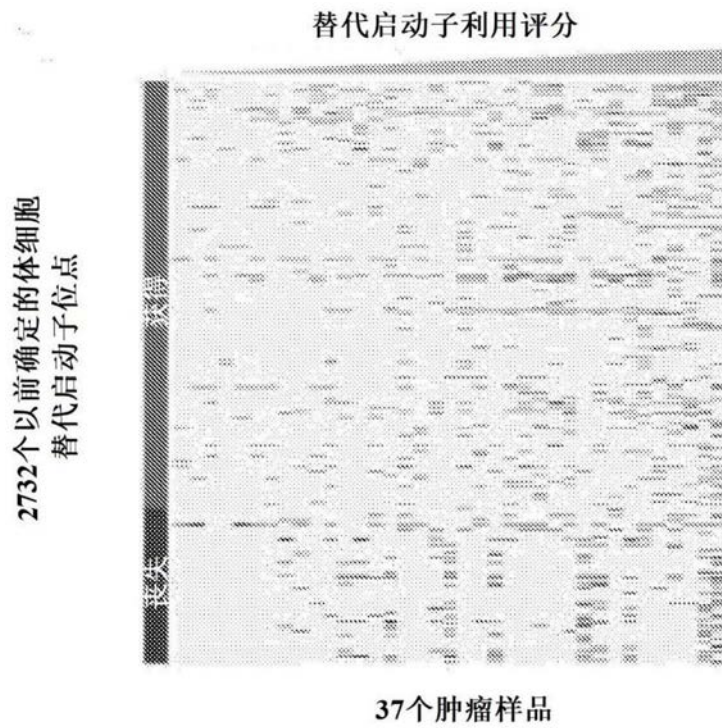


图3A

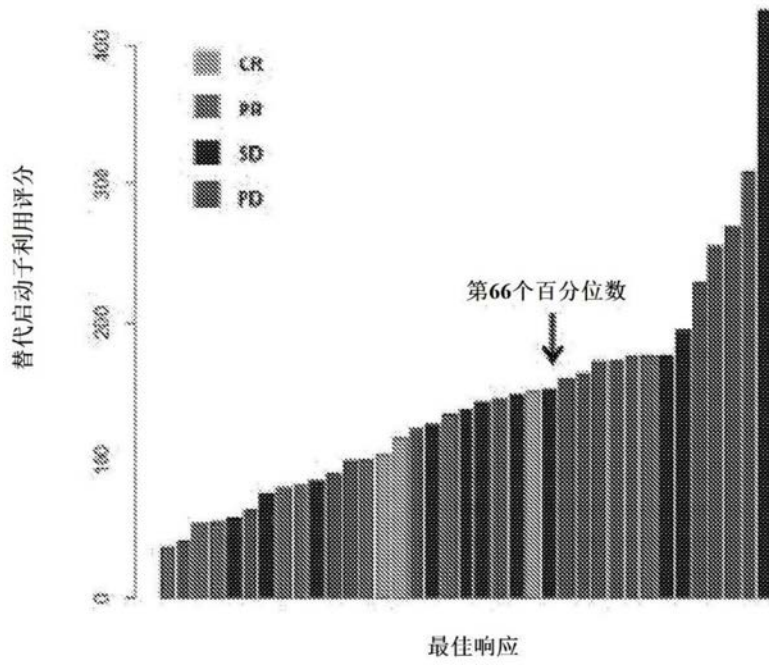


图3B

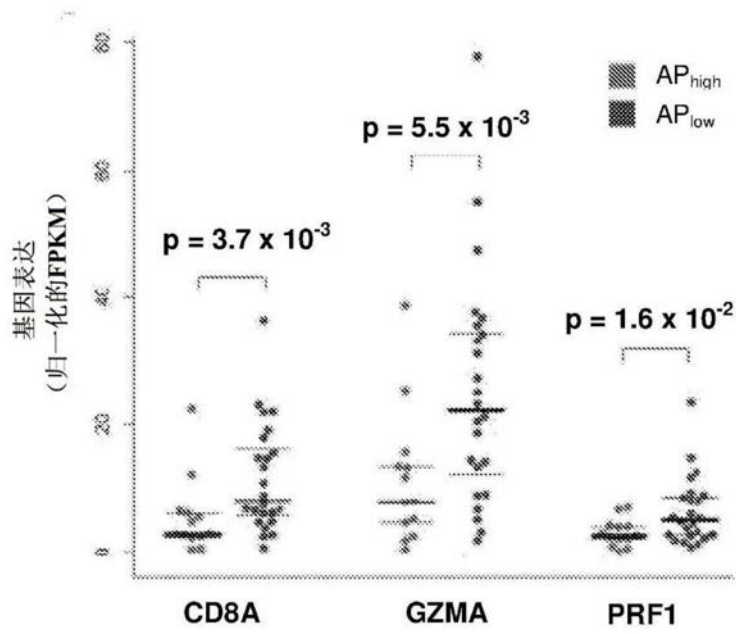


图3C

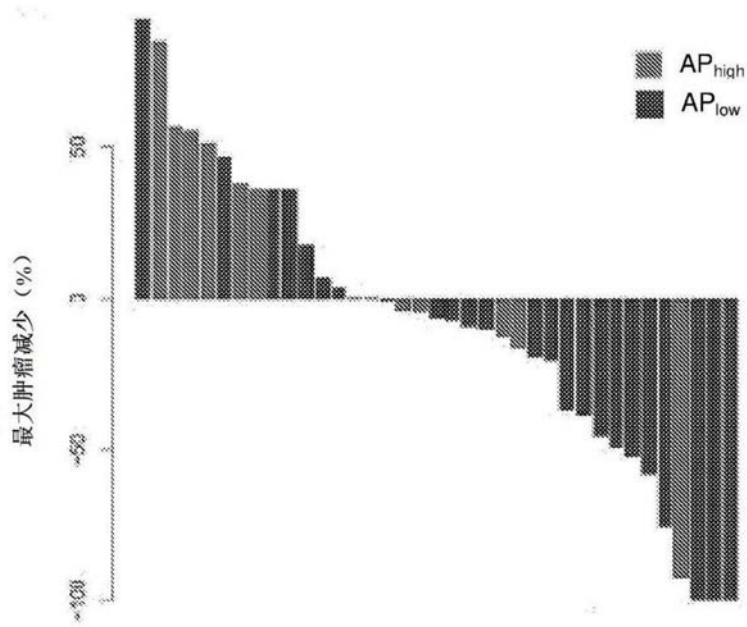


图3D

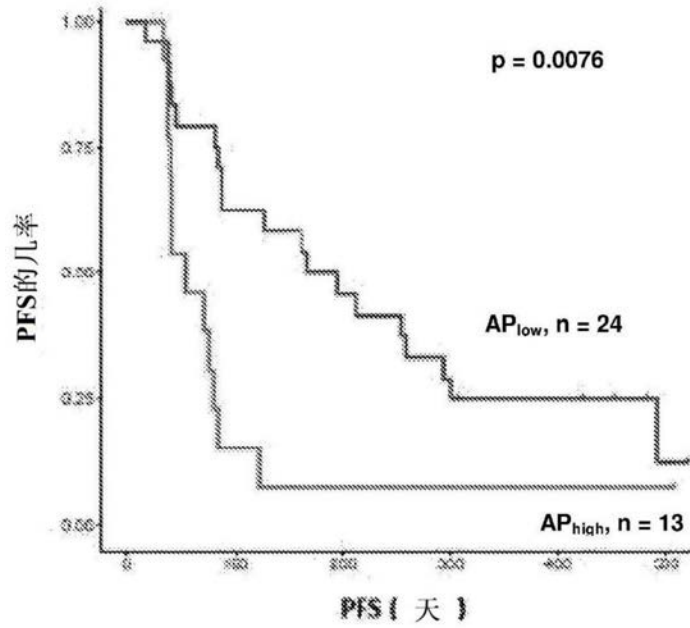


图4A

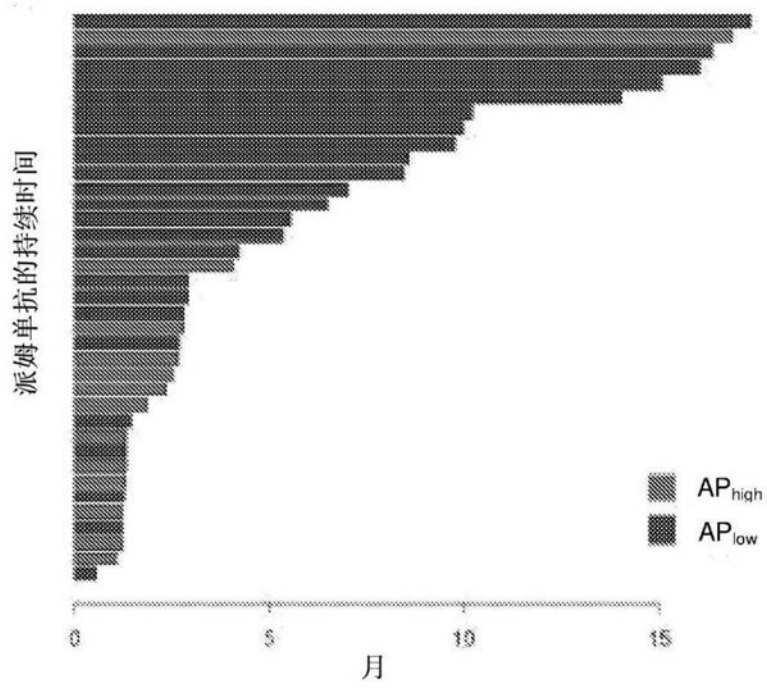


图4B

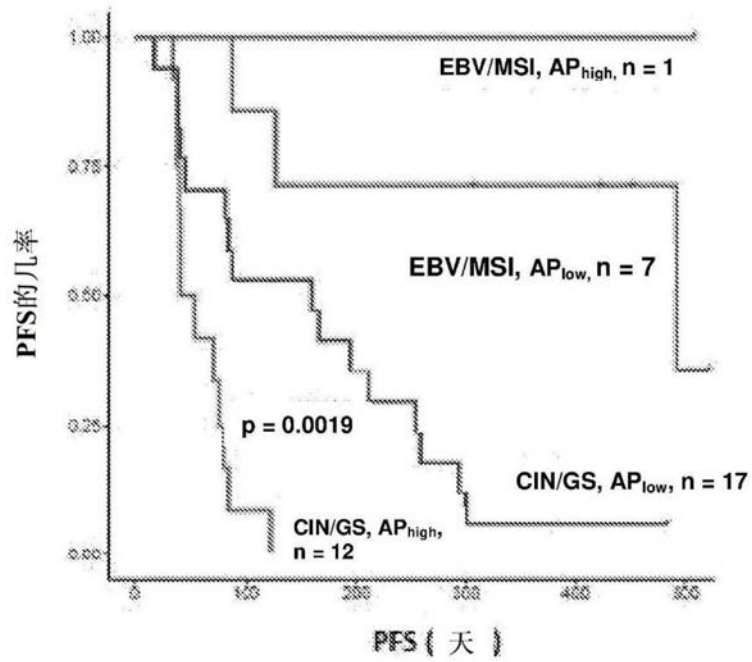


图4C

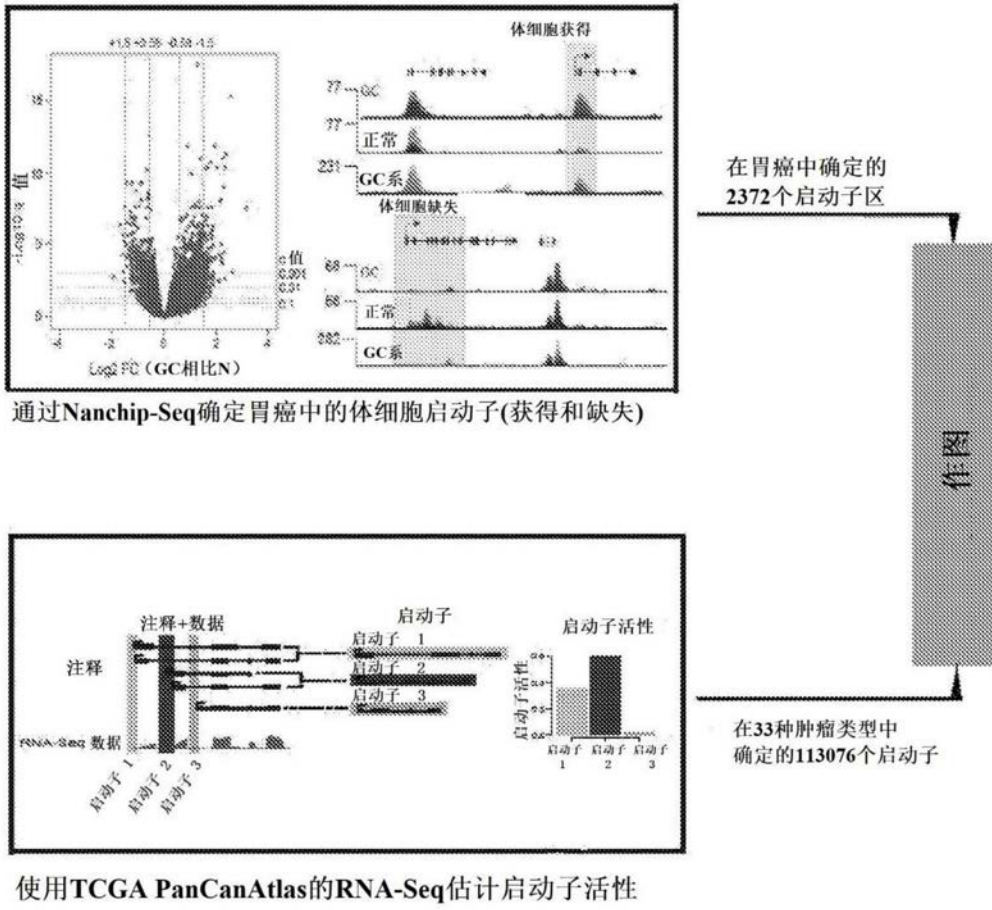


图5

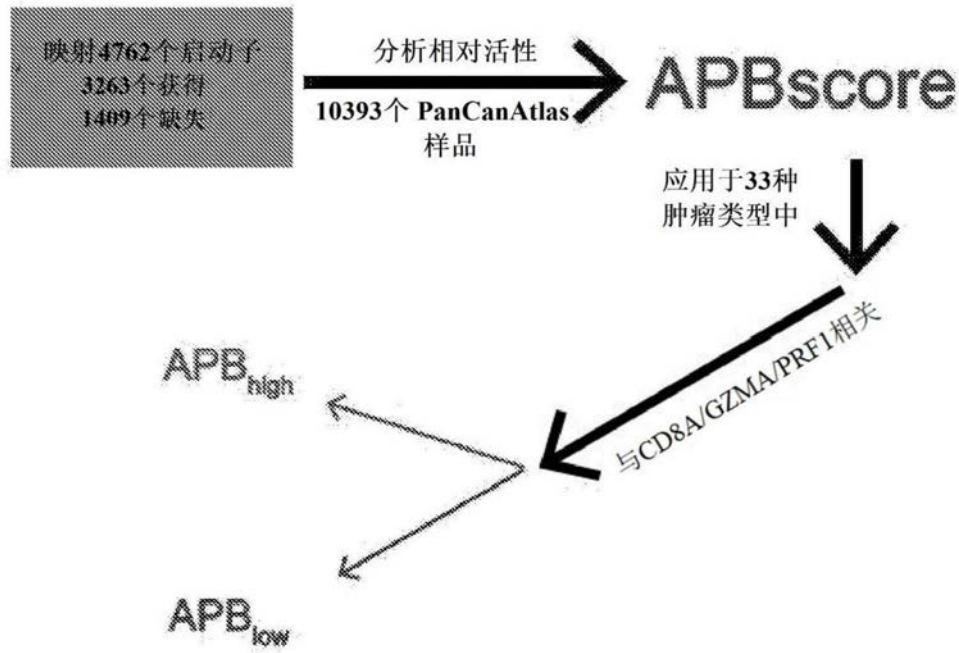


图5(续)

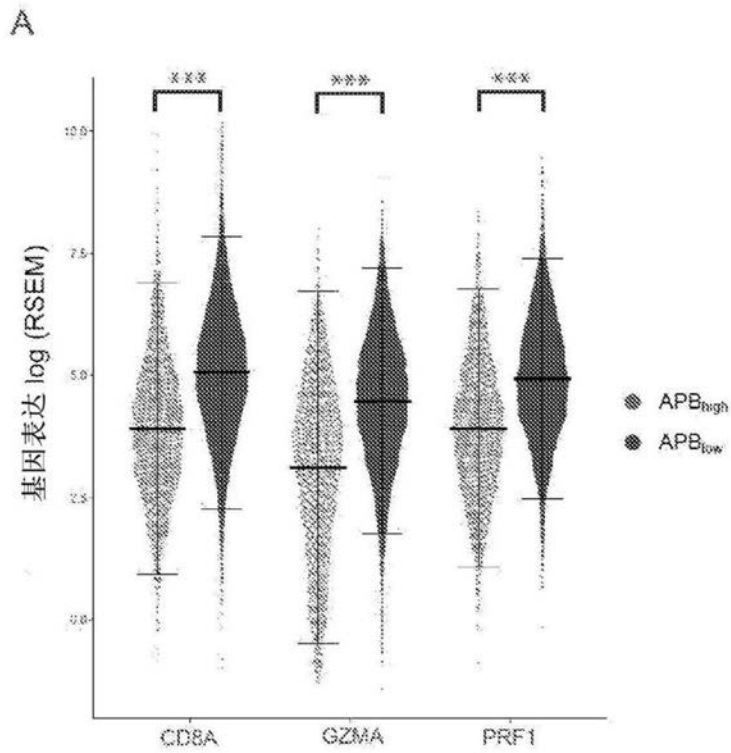


图6A

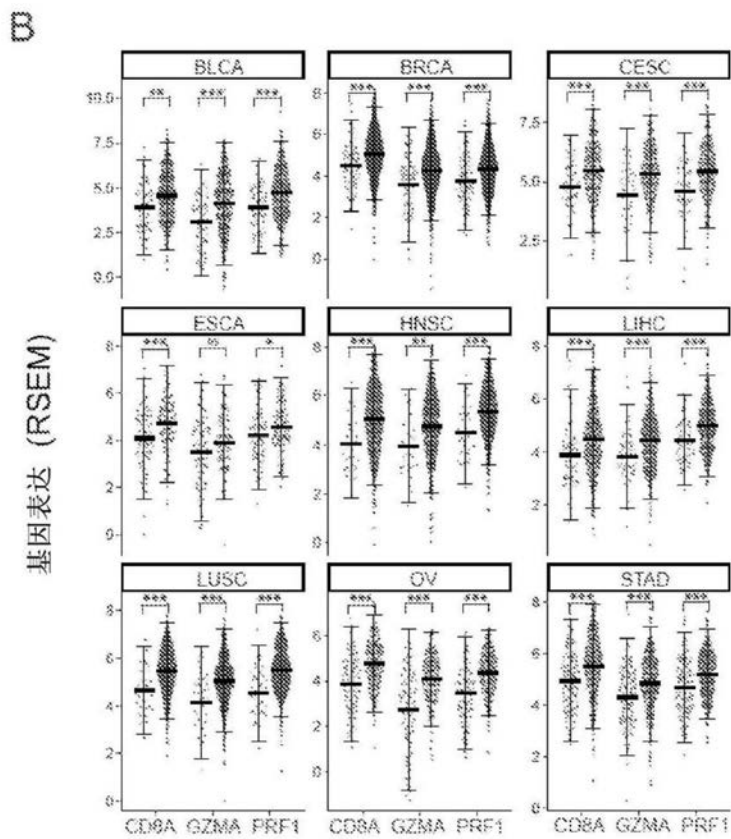


图6B

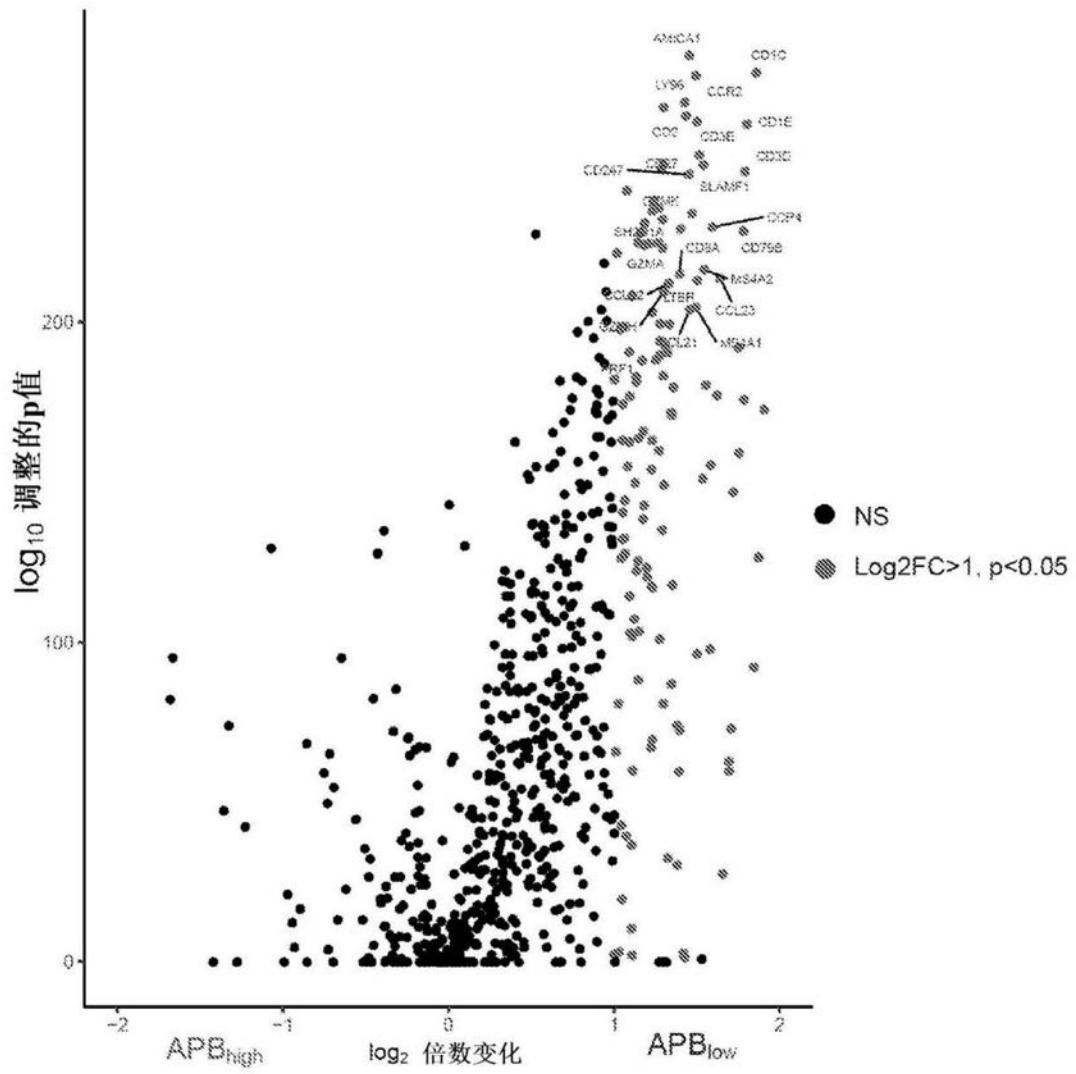


图7

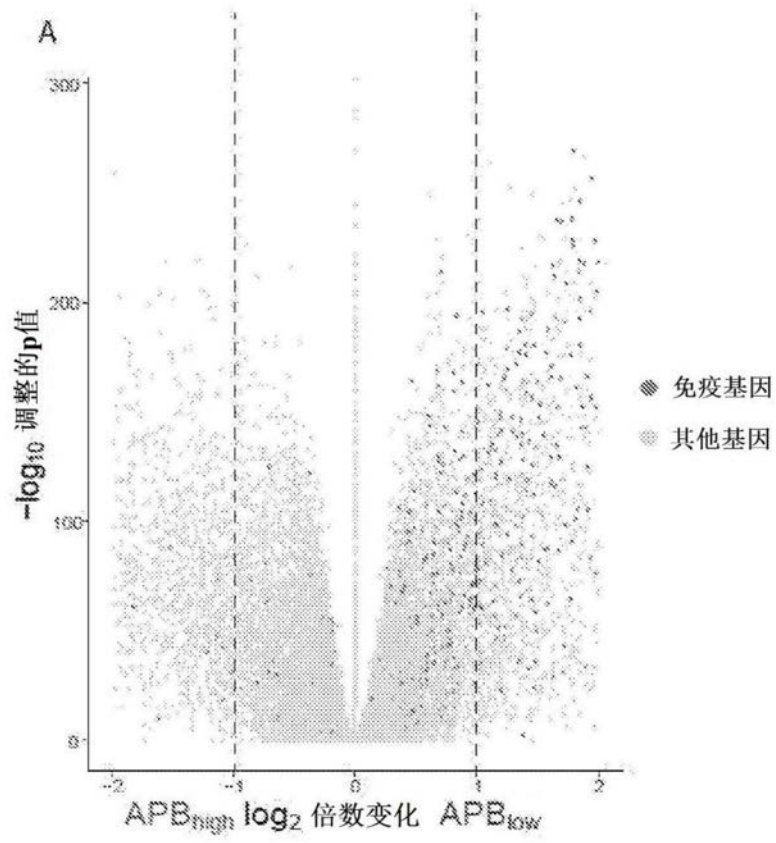


图8

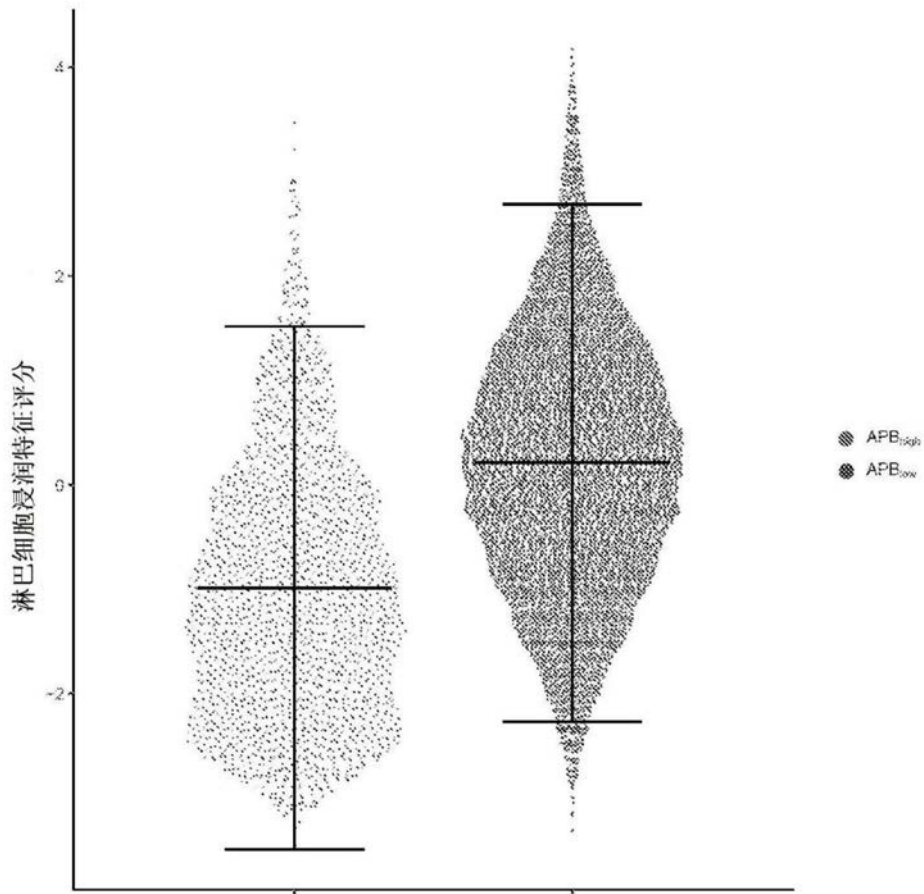


图9

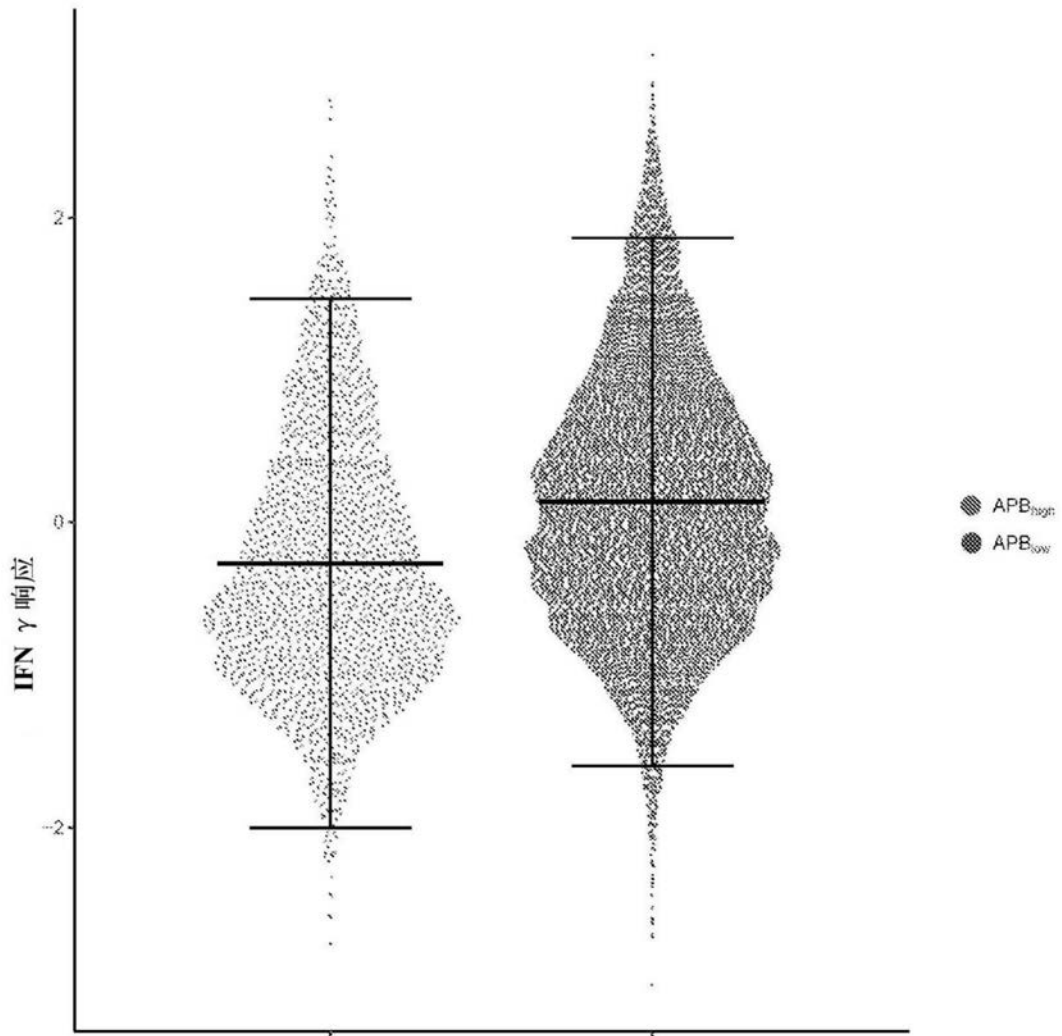


图10

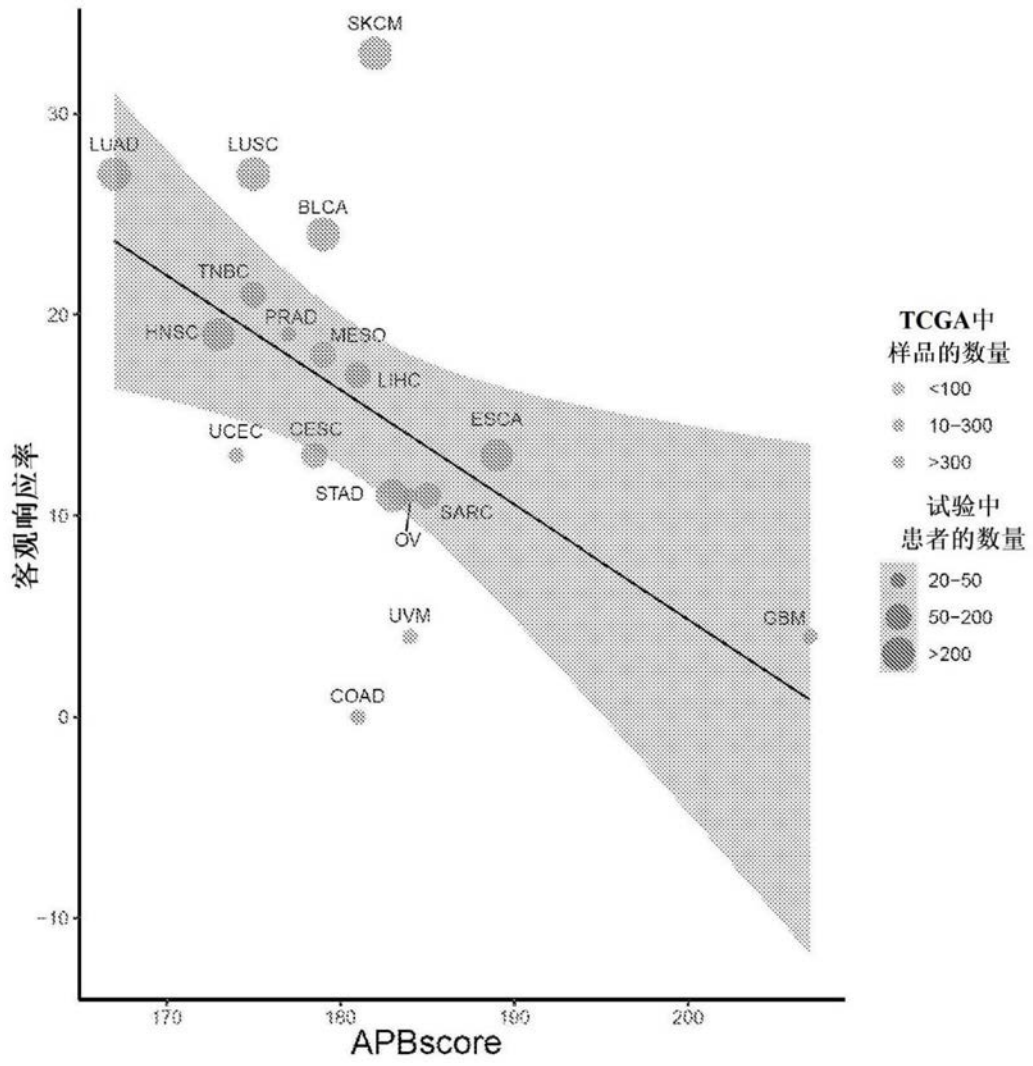


图11