



(12) 发明专利申请

(10) 申请公布号 CN 112626207 A

(43) 申请公布日 2021. 04. 09

(21) 申请号 202011083745.0

C12N 15/11 (2006.01)

(22) 申请日 2021.01.26

G01N 33/574 (2006.01)

(66) 本国优先权数据

202010772749.3 2020.08.04 CN

(71) 申请人 中国医学科学院北京协和医院

地址 100730 北京市东城区王府井帅府园1号

申请人 中国医学科学院基础医学研究所

(72) 发明人 包新杰 余佳 王任直 王耕超

王小爽 于姗

(74) 专利代理机构 北京预立生科知识产权代理

有限公司 11736

代理人 李红伟 孟祥斌

(51) Int. Cl.

C12Q 1/6886 (2018.01)

权利要求书2页 说明书16页

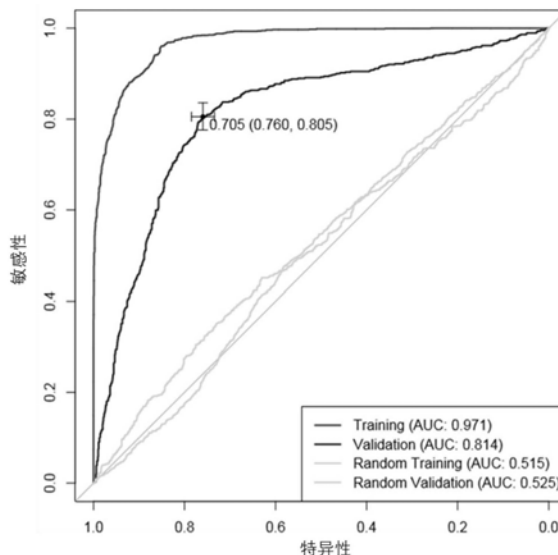
序列表11页 附图4页

(54) 发明名称

一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的基因组

(57) 摘要

本发明公开了一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的基因组,所述基因组由以下基因组成: INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A。通过检测本发明的基因组的表达水平,可以区分非侵袭性和侵袭性无功能垂体腺瘤。



1. 一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的分子标志物,所述分子标志物包括以下基因组合: INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A。

2. 检测权利要求1所述的分子标志物表达的产品在制备区分非侵袭性和侵袭性无功能垂体腺瘤的工具中的应用。

3. 根据权利要求2所述的应用,其特征在于,所述产品包括:通过反转录PCR、实时定量PCR、免疫检测、原位杂交、芯片或高通量测序平台检测所述分子标志物的表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品。

4. 根据权利要求3所述的应用,其特征在于,通过反转录PCR检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括特异扩增所述分子标志物的引物;通过实时定量PCR检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括特异扩增所述分子标志物的引物和/或探针;通过免疫检测检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:与所述分子标志物特异性结合的抗体;通过原位杂交检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:与所述分子标志物的核酸序列杂交的探针;通过芯片检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:蛋白芯片和基因芯片;其中,蛋白芯片包括与所述分子标志物特异性结合的抗体,基因芯片包括与所述分子标志物的核酸序列杂交的探针。

5. 根据权利要求4所述的应用,其特征在于,通过实时定量PCR检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品中包含的引物包括以下至少一对引物对:

- SEQ ID NO.1和SEQ ID NO.2所示的引物对;
- SEQ ID NO.4和SEQ ID NO.5所示的引物对;
- SEQ ID NO.7和SEQ ID NO.8所示的引物对;
- SEQ ID NO.10和SEQ ID NO.11所示的引物对;
- SEQ ID NO.13和SEQ ID NO.14所示的引物对;
- SEQ ID NO.16和SEQ ID NO.17所示的引物对;
- SEQ ID NO.19和SEQ ID NO.20所示的引物对;
- SEQ ID NO.22和SEQ ID NO.23所示的引物对;
- SEQ ID NO.25和SEQ ID NO.26所示的引物对;
- SEQ ID NO.28和SEQ ID NO.29所示的引物对;
- SEQ ID NO.31和SEQ ID NO.32所示的引物对;
- SEQ ID NO.34和SEQ ID NO.35所示的引物对;
- SEQ ID NO.37和SEQ ID NO.38所示的引物对;
- SEQ ID NO.40和SEQ ID NO.41所示的引物对;
- SEQ ID NO.43和SEQ ID NO.44所示的引物对;
- SEQ ID NO.46和SEQ ID NO.47所示的引物对;
- SEQ ID NO.52和SEQ ID NO.53所示的引物对;
- SEQ ID NO.55和SEQ ID NO.56所示的引物对;

SEQ ID NO.58和SEQ ID NO.59所示的引物对；

SEQ ID NO.61和SEQ ID NO.62所示的引物对。

6. 根据权利要求4所述的应用,其特征在於,通过实时定量PCR检测所述分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品中包含的探针包括以下探针序列中的至少一个:

SEQ ID NO.3所示的探针,与SEQ ID NO.1和SEQ ID NO.2所示的引物对配合使用;

SEQ ID NO.6所示的探针,与SEQ ID NO.4和SEQ ID NO.5所示的引物对配合使用;

SEQ ID NO.9所示的探针,与SEQ ID NO.7和SEQ ID NO.8所示的引物对配合使用;

SEQ ID NO.12所示的探针,与SEQ ID NO.10和SEQ ID NO.11所示的引物对配合使用;

SEQ ID NO.15所示的探针,与SEQ ID NO.13和SEQ ID NO.14所示的引物对配合使用;

SEQ ID NO.18所示的探针,与SEQ ID NO.16和SEQ ID NO.17所示的引物对配合使用;

SEQ ID NO.21所示的探针,与SEQ ID NO.19和SEQ ID NO.20所示的引物对配合使用;

SEQ ID NO.24所示的探针,与SEQ ID NO.22和SEQ ID NO.23所示的引物对配合使用;

SEQ ID NO.27所示的探针,与SEQ ID NO.25和SEQ ID NO.26所示的引物对配合使用;

SEQ ID NO.30所示的探针,与SEQ ID NO.28和SEQ ID NO.29所示的引物对配合使用;

SEQ ID NO.33所示的探针,与SEQ ID NO.31和SEQ ID NO.32所示的引物对配合使用;

SEQ ID NO.36所示的探针,与SEQ ID NO.34和SEQ ID NO.35所示的引物对配合使用;

SEQ ID NO.39所示的探针,与SEQ ID NO.37和SEQ ID NO.38所示的引物对配合使用;

SEQ ID NO.42所示的探针,与SEQ ID NO.40和SEQ ID NO.41所示的引物对配合使用;

SEQ ID NO.45所示的探针,与SEQ ID NO.43和SEQ ID NO.44所示的引物对配合使用;

SEQ ID NO.48所示的探针,与SEQ ID NO.46和SEQ ID NO.47所示的引物对配合使用;

SEQ ID NO.54所示的探针,与SEQ ID NO.52和SEQ ID NO.53所示的引物对配合使用;

SEQ ID NO.57所示的探针,与SEQ ID NO.55和SEQ ID NO.56所示的引物对配合使用;

SEQ ID NO.60所示的探针,与SEQ ID NO.58和SEQ ID NO.59所示的引物对配合使用;

SEQ ID NO.63所示的探针,与SEQ ID NO.61和SEQ ID NO.62所示的引物对配合使用。

7. 一种区分非侵袭性和侵袭性无功能垂体腺瘤的工具,其特征在於,所述工具包括检测样本中权利要求1所述的分子标志物表达水平的试剂。

8. 根据权利要求7所述的工具,其特征在於,所述试剂包括能够结合所述分子标志物的核酸或者能够结合所述分子标志物编码蛋白的物质;所述核酸能够检测所述分子标志物的mRNA表达水平;所述物质能够检测所述分子标志物编码蛋白的表达水平。

9. 根据权利要求7或8所述的工具,其特征在於,所述工具包括芯片、试剂盒、试纸或高通量测序平台。

10. 根据权利要求7-9中任一项所述的工具,其特征在於,所述样本来源于组织。

## 一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的基因组合

### 技术领域

[0001] 本发明属于生物医学领域,涉及一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的基因组合。

### 背景技术

[0002] 垂体腺瘤(Pituitary Tumours,PA)是一组起源于垂体前叶细胞的异质性疾病,约占颅内肿瘤的10%-15%,是常见的单克隆起源的腺瘤。人群发病率为1/10万,也有的报告可高达7/10万。随着分子生物学、细胞生物学、遗传学、免疫学等学科和新技术的飞速发展,很多学者从细胞、分子水平来阐明与其发病相关等方面的因素进行研究,以期对其发生机理及侵袭、复发相关性作出预测并指导临床治疗,取得了一定的进展。目前认为,包括无功能垂体腺瘤在内的大部分垂体腺瘤都是单克隆肿瘤,由一系列基因突变导致原癌基因的激活、抑癌基因的灭活,加上激素、生长因子、转录因子的刺激、细胞受体和信号通路异常等各种因素共同作用,通过复杂机制导致单个细胞转化为肿瘤细胞,肿瘤细胞的过度增殖和分泌功能失常,最终发生垂体腺瘤。然而,垂体腺瘤确切的机理尚待阐明。

[0003] 根据肿瘤细胞有无合成和分泌有生物活性激素的功能将垂体肿瘤分为功能性垂体肿瘤和无功能肿瘤。不具备激素分泌功能的垂体腺瘤称为无功能垂体腺瘤(Non-functioning Pituitary Adenoma,NF-PA),占30%-40%。无功能垂体腺瘤的诊断要点为存在垂体腺瘤的影像学证据。由于无功能垂体腺瘤缺乏特异的血清激素标志,故确诊有时需依赖手术标本的病理检查及免疫细胞化学检查。2017年最新的WHO指南根据垂体前叶激素和转录因子的免疫组化染色结果将无功能垂体腺瘤分为零细胞腺瘤(激素和转录因子的免疫组化染色均为阴性)和静默型腺瘤(激素和转录因子的免疫组化染色结果与某一种具有分泌生物活性激素功能的垂体腺瘤一致)两大类。其中,绝大多数促性腺激素腺瘤和约20%的促肾上腺皮质激素腺瘤为静默型,是最主要的两类静默型无功能腺瘤。这种新的分型标准提示无功能垂体腺瘤患者之间存在较大的异质性。

[0004] 虽然垂体腺瘤大多数是良性的,但仍有约35%的垂体腺瘤具有侵袭性,甚至有0.1%-0.2%的垂体腺瘤能够发展为癌,其恶性程度常与预后直接相关。侵袭性腺瘤难以通过常规手术完全切除,术后为了防止肿瘤复发,需要辅助放化疗。对于部分侵袭性腺瘤,单一的辅助治疗并不足以阻止肿瘤复发,从而造成肿瘤恶化及患者死亡。另一方面,手术方式选择不当将伤及正常腺体组织,影响垂体分泌功能,造成内分泌紊乱,副作用较多。因此,准确诊断肿瘤的侵袭性将有助于指导手术方式及后续治疗方法的选择,从早期阶段预防肿瘤的复发。近期研究发现无功能垂体腺瘤的侵袭性与上述细胞谱系来源、影像学特征等均存在一定相关性;也有研究报道了与无功能垂体腺瘤侵袭性相关的分子表达变化,但是目前尚无能够准确诊断无功能垂体腺瘤侵袭性的方法。

[0005] 以高密度基因芯片和高通量测序为核心的基因组学技术迅速发展,已广泛应用到生物医学的各个领域并取得突出进展,使我们对于疾病的分子与遗传学基础的认识达到了一个新的水平。这些研究成果对于阐明疾病的病因、解析疾病发生的分子机制、寻找疾病特

异的生物标志物和药物靶点,进而提升疾病的预防、诊断和治疗水平具有重要意义。

## 发明内容

[0006] 为了弥补现有技术的不足,本发明的目的在于提供区分非侵袭性和侵袭性无功能垂体腺瘤的分子标志物。

[0007] 为了实现上述目的,本发明采用如下技术方案:

[0008] 一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的分子标志物,所述分子标志物包括以下基因组合:INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A。

[0009] 根据本发明的一个方面,本发明提供了检测前面所述的分子标志物表达的产品在制备区分非侵袭性和侵袭性无功能垂体腺瘤的工具中的应用。

[0010] 进一步,上面所提到的检测前面所述的分子标志物表达的产品包括:通过反转录PCR、实时定量PCR、免疫检测、原位杂交、芯片或高通量测序平台检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品。

[0011] 进一步,通过反转录PCR检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括特异扩增前面所述的分子标志物的引物;通过实时定量PCR检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括特异扩增前面所述的分子标志物的引物和/或探针;通过免疫检测检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:与前面所述的分子标志物特异性结合的抗体;通过原位杂交检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:与前面所述的分子标志物的核酸序列杂交的探针;通过芯片检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品包括:蛋白芯片和基因芯片;其中,蛋白芯片包括与前面所述的分子标志物特异性结合的抗体,基因芯片包括与前面所述的分子标志物的核酸序列杂交的探针。

[0012] 在本发明的具体实施方案中,通过实时定量PCR检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品中包含的引物包括以下至少一对引物对:

[0013] SEQ ID NO.1和SEQ ID NO.2所示的引物对;

[0014] SEQ ID NO.4和SEQ ID NO.5所示的引物对;

[0015] SEQ ID NO.7和SEQ ID NO.8所示的引物对;

[0016] SEQ ID NO.10和SEQ ID NO.11所示的引物对;

[0017] SEQ ID NO.13和SEQ ID NO.14所示的引物对;

[0018] SEQ ID NO.16和SEQ ID NO.17所示的引物对;

[0019] SEQ ID NO.19和SEQ ID NO.20所示的引物对;

[0020] SEQ ID NO.22和SEQ ID NO.23所示的引物对;

[0021] SEQ ID NO.25和SEQ ID NO.26所示的引物对;

[0022] SEQ ID NO.28和SEQ ID NO.29所示的引物对;

[0023] SEQ ID NO.31和SEQ ID NO.32所示的引物对;

[0024] SEQ ID NO.34和SEQ ID NO.35所示的引物对;

- [0025] SEQ ID NO.37和SEQ ID NO.38所示的引物对；
- [0026] SEQ ID NO.40和SEQ ID NO.41所示的引物对；
- [0027] SEQ ID NO.43和SEQ ID NO.44所示的引物对；
- [0028] SEQ ID NO.46和SEQ ID NO.47所示的引物对；
- [0029] SEQ ID NO.52和SEQ ID NO.53所示的引物对；
- [0030] SEQ ID NO.55和SEQ ID NO.56所示的引物对；
- [0031] SEQ ID NO.58和SEQ ID NO.59所示的引物对；
- [0032] SEQ ID NO.61和SEQ ID NO.62所示的引物对。
- [0033] 通过实时定量PCR检测前面所述的分子标志物表达水平以区分非侵袭性和侵袭性无功能垂体腺瘤的产品中包含的探针包括以下探针序列中的至少一个：
- [0034] SEQ ID NO.3所示的探针，与SEQ ID NO.1和SEQ ID NO.2所示的引物对配合使用；
- [0035] SEQ ID NO.6所示的探针，与SEQ ID NO.4和SEQ ID NO.5所示的引物对配合使用；
- [0036] SEQ ID NO.9所示的探针，与SEQ ID NO.7和SEQ ID NO.8所示的引物对配合使用；
- [0037] SEQ ID NO.12所示的探针，与SEQ ID NO.10和SEQ ID NO.11所示的引物对配合使用；
- [0038] SEQ ID NO.15所示的探针，与SEQ ID NO.13和SEQ ID NO.14所示的引物对配合使用；
- [0039] SEQ ID NO.18所示的探针，与SEQ ID NO.16和SEQ ID NO.17所示的引物对配合使用；
- [0040] SEQ ID NO.21所示的探针，与SEQ ID NO.19和SEQ ID NO.20所示的引物对配合使用；
- [0041] SEQ ID NO.24所示的探针，与SEQ ID NO.22和SEQ ID NO.23所示的引物对配合使用；
- [0042] SEQ ID NO.27所示的探针，与SEQ ID NO.25和SEQ ID NO.26所示的引物对配合使用；
- [0043] SEQ ID NO.30所示的探针，与SEQ ID NO.28和SEQ ID NO.29所示的引物对配合使用；
- [0044] SEQ ID NO.33所示的探针，与SEQ ID NO.31和SEQ ID NO.32所示的引物对配合使用；
- [0045] SEQ ID NO.36所示的探针，与SEQ ID NO.34和SEQ ID NO.35所示的引物对配合使用；
- [0046] SEQ ID NO.39所示的探针，与SEQ ID NO.37和SEQ ID NO.38所示的引物对配合使用；
- [0047] SEQ ID NO.42所示的探针，与SEQ ID NO.40和SEQ ID NO.41所示的引物对配合使用；
- [0048] SEQ ID NO.45所示的探针，与SEQ ID NO.43和SEQ ID NO.44所示的引物对配合使用；
- [0049] SEQ ID NO.48所示的探针，与SEQ ID NO.46和SEQ ID NO.47所示的引物对配合使用；

[0050] SEQ ID NO.54所示的探针,与SEQ ID NO.52和SEQ ID NO.53所示的引物对配合使用;

[0051] SEQ ID NO.57所示的探针,与SEQ ID NO.55和SEQ ID NO.56所示的引物对配合使用;

[0052] SEQ ID NO.60所示的探针,与SEQ ID NO.58和SEQ ID NO.59所示的引物对配合使用;

[0053] SEQ ID NO.63所示的探针,与SEQ ID NO.61和SEQ ID NO.62所示的引物对配合使用。

[0054] 根据本发明的另一个方面,本发明提供了一种区分非侵袭性和侵袭性无功能垂体腺瘤的工具,所述工具包括检测样本中前面所述的分子标志物的表达水平的试剂。

[0055] 进一步,所述试剂包括能够结合前面所述的分子标志物的核酸或者能够结合前面所述的分子标志物编码蛋白的物质。所述核酸能够检测前面所述的分子标志物的mRNA表达水平;所述物质能够检测前面所述的分子标志物编码蛋白的表达水平。

[0056] 更进一步,所述核酸包括扩增前面所述的分子标志物的引物,引物可以通过化学合成来制备,通过使用本领域技术人员知道的方法参考已知信息来适当地设计,并通过化学合成来制备。

[0057] 在本发明的具体实施方案中,所述核酸为QPCR实验中使用的扩增引物,所述引物的序列如前面所述。

[0058] 上面所述的核酸还可包括探针,所述探针可以通过化学合成来制备,通过使用本领域技术人员知道的方法参考已知信息来恰当设计,并通过化学合成来制备,或者可以通过从生物材料制备含有期望核酸序列的基因,并使用设计用于扩增期望核酸序列的引物扩增它来制备。

[0059] 探针序列如前面所述的相同。

[0060] 更进一步,所述物质包括针对前面所述的分子标志物编码蛋白的抗体。

[0061] 进一步,所述诊断工具包括芯片、试剂盒、试纸、或高通量测序平台。

[0062] 其中,所述芯片包括基因芯片、蛋白质芯片;所述基因芯片包括固相载体以及固定在固相载体的寡核苷酸探针,所述寡核苷酸探针包括用于检测前面所述的分子标志物转录水平的针对前面所述的分子标志物的寡核苷酸探针;所述蛋白质芯片包括固相载体以及固定在固相载体的前面所述的分子标志物编码蛋白的特异性抗体;所述基因芯片可用于检测前面所述的分子标志物的表达水平。所述蛋白质芯片可用于检测前面所述的分子标志物编码蛋白的表达水平。

[0063] 其中,所述试剂盒包括基因检测试剂盒和蛋白免疫检测试剂盒;所述基因检测试剂盒包括用于检测前面所述的分子标志物转录水平的试剂;所述蛋白免疫检测试剂盒包括前面所述的分子标志物编码蛋白的特异性抗体。进一步,所述试剂包括使用反转录PCR、实时定量PCR、免疫检测、原位杂交或芯片方法检测前面所述的分子标志物表达水平过程中所需的试剂。优选地,所述试剂包括针对前面所述的分子标志物的引物和/或探针。根据前面所述的分子标志物的核苷酸序列信息容易设计出可以用于检测前面所述的分子标志物表达水平的引物和探针。

[0064] 与前面所述的分子标志物的核酸序列杂交的探针可以是DNA、RNA、DNA-RNA嵌合

体、PNA或其它衍生物。所述探针的长度没有限制,只要完成特异性杂交、与目的核苷酸序列特异性结合,任何长度都可以。所述探针的长度可短至25、20、15、13或10个碱基长度。同样,所述探针的长度可长至60、80、100、150、300个碱基对或更长,甚至整个基因。由于不同的探针长度对杂交效率、信号特异性有不同的影响,所述探针的长度通常至少是14个碱基对,最长一般不超过30个碱基对,与目的核苷酸序列互补的长度以15-25个碱基对最佳。所述探针自身互补序列最好少于4个碱基对,以免影响杂交效率。

[0065] 所述高通量测序平台包括检测前面所述的分子标志物表达水平的试剂。

[0066] 所述试纸包括试纸载体和固定在试纸载体上的寡核苷酸,所述寡核苷酸能够检测前面所述的分子标志物的转录水平。

[0067] 进一步,前面所述的分子标志物编码蛋白的抗体包括单克隆抗体、多克隆抗体。前面所述的分子标志物编码蛋白的抗体包括完整的抗体分子、抗体的任何片段或修饰(例如,,嵌合抗体、scFv、Fab、F(ab')<sub>2</sub>、Fv等。只要所述片段能够保留与前面所述的分子标志物编码蛋白的结合能力即可。用于蛋白质水平的抗体的制备时本领域技术人员公知的,并且本发明可以使用任何方法来制备所述抗体。

[0068] 在本发明的具体实施方案中,所述针对前面所述的分子标志物的引物序列如前面所述。

[0069] 进一步,本发明提到的样本包括组织、体液,体液包括但不限于血液、组织液、尿液、唾液、脊髓液等可以获得RNA的体液。在本发明的具体实施方案中,所述样本来源于组织。

[0070] 根据本发明的又一个方面,本发明还提供了一种区分非侵袭性和侵袭性无功能垂体腺瘤的方法,所述方法包括如下步骤:

[0071] (1) 获取待试者含有基因组DNA的样本;

[0072] (2) 检测样本中前面所述的分子标志物的表达水平;

[0073] (3) 将测得的前面所述的分子标志物的表达水平与无功能垂体腺瘤的亚型关联起来。

[0074] (4) 与非侵袭性亚型参考值相比,如果INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6表达显著上调、SFRP1、EMILIN2、CD9、SEMA7A表达显著下调,则代表该待试者属于侵袭性无功能垂体腺瘤患者。

[0075] 本发明上面所述的待试者可以是不知是否患有垂体腺瘤的人,也可以是已经确诊为垂体腺瘤但不能确定是属于垂体腺瘤哪种亚型的人,也可以是已经确诊为无功能垂体腺瘤的人。

[0076] 本发明的优点和有益效果:

[0077] 本发明首次发现了INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A表达与侵袭性无功能垂体腺瘤相关,通过检测上述基因表达,可以判断待试者是否为侵袭性无功能垂体腺瘤,从而指导临床医师给受试者提供治疗方案。

[0078] 本发明发现了一种新的分子标记物-INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A,相比传统的检测手段,基因诊断更及时、更特异、更灵敏,能够实现侵袭性无功能垂体腺瘤



亚型快速诊断。

### 附图说明

- [0079] 图1显示MRI结果图；
- [0080] 图2显示免疫组化结果图；
- [0081] 图3显示RNA-seq结果图；
- [0082] 图4显示本发明的20个差异表达基因的表达水平统计图；
- [0083] 图5显示利用qRT-PCR检测20个差异表达基因的表达水平统计图；
- [0084] 图6显示qRT-PCR和RNA-seq一致性的结果图；
- [0085] 图7显示ROC曲线图；
- [0086] 图8显示qRT-PCR中基因扩增的标准曲线在C<sub>q</sub>值和样品浓度的对数之间的线性关系图。

### 具体实施方式

[0087] 下面结合附图和实施例对本发明作进一步详细的说明。以下实施例仅用于说明本发明而不适用于限制本发明的范围。实施例中未注明具体条件的实验方法,通常按照常规条件,例如Sambrook等人,分子克隆:实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件,或按照制造厂商所建议的条件。

[0088] 实施例1侵袭性无功能垂体腺瘤相关基因筛选与验证

[0089] 1、样品采集(临床诊断)和研究监督

[0090] 从该研究所有个体参与者那里获取知情同意书。研究招募过程和方案经北京协和医院伦理审查委员会批准(编号S-551)。收集2012年5月至2017年7月间在北京协和医院进行经蝶窦手术的患者,取其垂体腺瘤。61例NF-PA进行RNA-seq, 143例NF-PA进行TaqMan定量RT-PCR和组织微阵列。根据临床表现、激素测试和磁共振成像(MRI)检查确诊NF-PAs。另外,对所有垂体前叶激素和转录因子进行免疫组织化学染色以区分NF-PA亚型(表1)。在本研究中,所有NF-PA的最大直径均大于20毫米。根据Knosp分类诊断非侵袭性PA(NNF-PA)和侵袭性PA(INF-PA),并在术中予以确认。非侵袭性PA为Knosp 0、1或2级,未侵入海绵窦。所有侵袭性PA均选取Knosp 4级,且侵入了海绵窦的病例。记录每个患者的临床特征,手术结果,术后并发症,病理结果和随访数据。

[0091] 表1用于RNA-seq实验的NF-PAs亚型区分

垂体细胞谱系	垂体前叶激素	转录因子	No. NNF	No. INF	总
促肾上腺皮质激素细胞	ACTH	TPIT	1	2	3
	无	TPIT	7	15	22
[0092] 促性腺激素细胞	FSH $\beta$ 和/或 LH $\beta$	SF1	7	0	7
	FSH $\beta$ 和/或 LH $\beta$	无	20	3	23
裸细胞腺瘤	无	无	4	2	6
总			39	22	61

[0093] 2、组织微阵列(TMA) (包括免疫组织化学和评分)

[0094] 使用以下抗体对人垂体腺瘤组织微阵列进行免疫组织化学研究:T-PIT (ZM-0318, 稀释度1:100;ZSJQ Corp.);SF-1 (ZM-0089, 稀释度1:100;ZSJQ Corp.);和PIT-1 (ZM-0208, 稀释度1:100;ZSJQ Corp.),根据文献报道的方法在连续切割的组织阵列切片中染色 [Wallace TJ, Qian J, Avital I, Bay C, Man Y-G, Wellman LL, et al. Technical Feasibility of Tissue Microarray (TMA) Analysis of Tumor-Associated Immune Response in Prostate Cancer. J Cancer. 2018;9:2191-202]. TMA在二甲苯中脱蜡,并在乙醇中水合,然后水中冲洗。用3% H<sub>2</sub>O<sub>2</sub>阻断内源过氧化物酶活性30分钟。用Envision™ FLEX过氧化物酶封闭试剂 (Dako#SM801, 安捷伦, 加利福尼亚州圣克拉拉) 封闭3小时后,一抗过夜。使用Envision™ FLEX/HRP (Dako#SM802, Agilent) 作为二抗,并使用Envision™ FLEX Substrate Buffer+Envision™ FLEX DAB+色原 (Dako#SM803, Agilent) 进行显色检测。图像使用Axio Scan.Z.1 (德国蔡司) 拍摄。使用染色H评分法评估Asporin染色 [Hirsch FR, Varella-Garcia M, Bunn PA, Di Maria MV, Veve R, Bremmes RM, et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol. 2003;21:3798-807]。计算每个细胞的染色强度 (0、1、2、3) 和每种强度染色的垂体细胞比例 (0-100) 使用以下公式获得分数: 3x强染色细胞百分比+2x中度染色细胞百分比+1x弱染色细胞百分比。

[0095] 3、RNA提取和测序

[0096] 从PA组织中提取总RNA,并使用NanoDrop 2000/2000c (Thermo Fisher Scientific Inc., Waltham, MA, USA) 检查RNA完整性。根据制造商说明,使用TruSeq® RNA LT样品制备试剂盒v2 (Illumina Inc., San Diego, CA, USA) 构建RNA库。简言之,基于多聚腺苷酸化的mRNA富集和RNA片段化之后,使用First Strand Master Mix和SuperScript II合成第一链互补DNA (cDNA),然后使用Second Strand Master Mix合成第二链cDNA。对双链cDNA进行末端修复,然后对3'末端进行腺苷酸化,并连接illumina衔接子。用15个聚合酶链

反应循环扩增连接产物。测定产量和片段长度分布后,使用TrusSeq SBS试剂盒v3-HS对文库进行测序,并在HiSeq 2500或3000测序系统(均来自Illumina Inc.)上生成。用Trim Galore修剪原始读段的衔接子序列。使用TopHat2将测序读段映射到具有Gencode v28注释的人类基因组GRCh38上。基因表达水平通过Cufflinks和HTSeq进行定量。

[0097] 4、差异基因表达

[0098] 使用DESeq2分析差异基因表达(DGE)。DGE分析之前检查技术协变量和批次效应。FDR<0.05,|fold change|>2代表差异显著。

[0099] 5、实时定量PCR(qRT-PCR)

[0100] 使用基于TaqMan探针的基因表达分析(Thermo Fisher,Foster City,CA)评估mRNA表达。引物和探针序列列于表2中,使用UBC和GAPDH进行目的基因表达量的标准化。每个样品的相对表达水平记录为基因表达与两个参考基因(GAPDH和UBC)表达的几何平均值之比。每个样品进行三次重复。GAPDH,UBC和其他mRNA的标准曲线在C<sub>q</sub>值和样品浓度的对数之间显示出良好的线性关系(图8)。

[0101] 表2引物和探针序列

基因名称	基因 ID	正向引物 (5'-3')	反向引物 (5'-3')	探针 (5'-3')
SFRP1	6422	GAGGAGGTTTGCCT GTGATTT (SEQ ID	TCCCAGGACTCTATC CAGAATG (SEQ ID	AATTATCGCTAGGGCCA AGGTGGG (SEQ ID NO.3)

[0102]

		NO.1)	NO.2)	
EMILIN2	84034	CAGCAGCAGTGTGA TGACTAT (SEQ ID NO.4)	CGGGCTCGGAGGTT ATTTATT (SEQ ID NO.5)	TCTATCACTCCCAGGTAG CTGCTCC (SEQ ID NO.6)
CD9	928	CGGCTTCCTCTGGT GATATT (SEQ ID NO.7)	GCTTGTTGTAGGTGT CCTTGTA (SEQ ID NO.8)	ATGGCCGCAGCTATTTCA ATGGC (SEQ ID NO.9)
SEMA7A	8482	TTCCGTACCTCCTCA CTCAA (SEQ ID NO.10)	TGGAAGGTCTCTGT GGGTAT (SEQ ID NO.11)	CTACCACTCAAGCCTTCC CAACCC (SEQ ID NO.12)
JUN	3725	CACAGAGAGACAGA CTTGAGAAC (SEQ ID NO.13)	ACTTGGATACCCTTG GCTTTAG (SEQ ID NO.14)	CTTGACAAGTTGCGACG GAGAGAAA (SEQ ID NO.15)
JUNB	3726	CCATCAACATGGAA GACCAAGA (SEQ ID NO.16)	CTTGAGCGTCTTCAC CTTGT (SEQ ID NO.17)	CATCAAAGTGGAGCGCA AGCGGC (SEQ ID NO.18)
RHOU	58480	GCAGGGAGGTGAAT ACTCTTG (SEQ ID NO.19)	GAAGTACATCTTGG CCGACTTA (SEQ ID NO.20)	AAGGCCAACAGCAAGTG TTTGTGGGA (SEQ ID NO.21)
INSM1	3642	GTATGTAGCGTACG GGTTGTT (SEQ ID NO.22)	CAGCAGTTCACAAG CCATAATAAA (SEQ ID NO.23)	AGGCTGTAAATTGAACT TCCCACACGAT (SEQ ID NO.24)
RHOB	388	TTCGGAGCTAAGAT GGTGTTAT (SEQ ID NO.25)	GGAAAGATGGTCAA GTCCTGT (SEQ ID NO.26)	TAAGGGTGGTGATGGGT GAGCGC (SEQ ID NO.27)

[0103]

EZR	7430	GATCGAGGATACTC ACCAGAAA (SEQ ID NO.28)	CCAGTTGTATTTGGC TGGATTG (SEQ ID NO.29)	CACCATGGATGCAGAGC TGGAGTTT (SEQ ID NO.30)
LLGL2	3993	TGATCTCACCCCTCGG AGTTT (SEQ ID NO.31)	GCGGTGGTTCTTGGT TTCT (SEQ ID NO.32)	CACTTGGTGGAGAGAGA GAAGCGC (SEQ ID NO.33)
SNX33	257364	CCGTAACCTCAACC GTTTCT (SEQ ID NO.34)	GGAGTATGTCTCAG CGATCTTG (SEQ ID NO.35)	ATGCTTTGTGCGTTCTGG AGTGGA (SEQ ID NO.36)
SPIRE2	84501	GTCTGCACTTCCTGT AGCATAA (SEQ ID NO.37)	GATACCCTCTGAGG ACTCTCAA (SEQ ID NO.38)	CACATCCCTGTCTACACA CTGGGC (SEQ ID NO.39)
MLF1	4291	CTGGCTCCCGAGAA CTTAAA (SEQ ID NO.40)	GCCTTTGATGTGGA GTTTGTC (SEQ ID NO.41)	AGGGAGAAACCTCAACA AAGTCCAGC (SEQ ID NO.42)
HSPA2	3306	CCCTGGAGTCCTATA CCTACAA (SEQ ID NO.43)	GACACTTGTCGAGG ATCTTGTT (SEQ ID NO.44)	AAGCAGACGGTGGGAAGA CGAGAAA (SEQ ID NO.45)
CDKN1A	1026	CGGAACAAGGAGTC AGACATT (SEQ ID NO.46)	AGTGCCAGGAAAGA CAACTAC (SEQ ID NO.47)	AGATGGTGGCAGTAGAG GCTATGGA (SEQ ID NO.48)
ID2	3398	CAAGAAGGTGAGCA AGATGGA (SEQ ID NO.49)	GGTGATGCAGGCTG ACAATA (SEQ ID NO.50)	TGCAGCACGTCATCGAC TACATCTTG (SEQ ID NO.51)
ERN1	2081	CCCATCAACCTCTCT TCTGTATC (SEQ ID	AGGCCGCATAGTCA AAGTAG (SEQ ID	TGGTGATGGTGTATTCTG

[0104]

		NO.52)	NO.53)	TTCGCCC (SEQ ID NO.54)
NR2F2	7026	CGTCTCCTCCTCAGT CATAGA (SEQ ID NO.55)	TGCCATATACGGCC AGTTAAA (SEQ ID NO.56)	TGCTGCCGGACAGTAAC ATATCCC (SEQ ID NO.57)
GAS2	2620	GATTGCAGCCAGGT ATGGT (SEQ ID NO.58)	GAAGGAGAAGGTGA AGGAGAAG (SEQ ID NO.59)	TGGAGCCTCCTGGTTTGA TAAAGCTG (SEQ ID NO.60)
[0105] CDK6	1021	CTAGCAACCATCCCT CCATTAC (SEQ ID NO.61)	CTCAGAGCATTCTG AAGACAGTAG (SEQ ID NO.62)	ACACAGAAAGCCCTCTT GAAGCAA (SEQ ID NO.63)
GAPDH	2597	GGTGTGAACCATGA GAAGTATGA (SEQ ID NO.64)	GAGTCCTTCCACGA TACCAAAG (SEQ ID NO.65)	AGATCATCAGCAATGCC TCCTGCA (SEQ ID NO.66)
UBC	7316	GTACCCTGTCTGACT ACAACATC (SEQ ID NO.67)	GTCTTGCCAGTGAG TGTCTT (SEQ ID NO.68)	AGAAAGAGTCCACCTTG CACCTGG (SEQ ID NO.69)

## [0106] 6、支持向量机分类器

[0107] 本研究使用由R软件包e1071运行的支持向量机(SVM)算法对INF-PA和NNF-PA进行分类。实验验证基因表达量均一化后用于构建SVM算法。在训练集上对SVM算法进行训练以构建模型,并在验证集中预测样本(训练:验证=2:1)。在训练集上进行留一法交叉验证(LOOCV)以评估模型的质量。通过R包pROC进行受试者工作曲线(ROC)分析评估预测强度。

## [0108] 7、统计分析

[0109] 数据显示为平均值±标准偏差。根据理论频率,使用卡方检验或费舍尔精确检验对垂体腺瘤患者的分类临床特征进行统计学检验。通过F检验确认方差相等之后,将Mann-Whitney U检验或t检验用于连续变量。

## [0110] 8、结果

### [0111] 8.1NF-PA患者的临床特征

[0112] 收集61例无任何内分泌过度分泌综合征的临床NF-PA患者,并检查肿瘤的病理特征。为了确定浸润海绵窦的状态,采用Knosp等[Knosp E,Steiner E,Kitz K,Matula C.Pituitary adenomas with invasion of the cavernous sinus space:a magnetic

resonance imaging classification compared with surgical findings. Neurosurgery. 1993;33] 描述的分类标准通过MRI确定放射学特征,并在手术过程中得到确认(图1)。其中,39例是0、1或2级(非侵袭性NF, NNF),而其他22例是4级(侵袭性NF, INF)。两组之间的苏木精和曙红染色, p53阳性率和Ki-67检测的特征无显著性差异(表3和图2)。

[0113] NNF-PA患者和INF-PA患者的手术结果,术后并发症和预后明显不同。所有NNF-PA患者进行了全切除(74.4%)或次全切除(25.6%),而超过一半(54.5%)的INF-PA患者仅进行了部分切除。结果,更多INF-PA患者接受了后续治疗,包括放射治疗,药物治疗和手术治疗。在随访期间,22例INF-PA患者中有14例残留稳定,而8例肿瘤生长,其中2例在平均60.4个月后死亡。相反,在平均58.2个月的随访中,NNF组没有肿瘤生长或死亡。此外,除8.6%的NNF-PA患者外,所有INF-PA患者均出现术后症状,如头痛,视力障碍,单眼失明,垂体功能减退,脑神经麻痹等(表3)。

[0114] 表3患者的临床特征

[0115]

变量	NNF	INF	P 值
性别			
女 (%)	21 (53.8)	18 (81.8)	0.029
男(%)	18 (46.2)	4 (18.2)	
年龄 (年)	55.2±10.6	47.2±10.2	0.006
身高 (cm)	164.0±6.6	162.3±8.2	0.380

[0116]

体重 (kg)	66.6±10.8	67.5±13.8	0.779
最大肿瘤直径(mm)	29.2±6.5	39.7±12.0	0.0007
侵袭深度			
Knosp 0 级 (%)	16 (41.0)	0 (0)	<0.001
Knosp 1-2 级 (%)	23 (59.0)	0 (0)	<0.001
Knosp 3 级(%)	0 (0)	0 (0)	
Knosp 4 级(%)	0 (0)	22 (100.0)	<0.001
肿瘤切除范围			
全切除(%)	29 (74.4)	0 (0)	<0.001
次全切除 (%)	10 (25.6)	10 (45.5)	0.270
部分切除(%)	0 (0)	12 (54.5)	<0.001
Ki-67	1.7±0.7	1.7±1.0	1
≤3% (%)	39 (100.0)	21 (95.5)	0.120
>3% (%)	0 (0)	1 (4.5)	
P53			
阳性 (%)	1 (2.6)	2 (9.1)	0.262
阴性 (%)	38 (97.4)	20 (90.9)	
随访平均月数(范围)	58.2 (45-72)	60.4 (46-75)	0.162
术后症状			



	头痛 (%)	2 (5.1)	7 (31.8)	0.005
	功能障碍(%)	1 (2.6)	5 (22.7)	0.011
	单眼失明 (%)	0 (0)	2 (9.1)	0.027
	垂体功能减退(%)	1 (2.6)	5 (22.7)	0.011
	永久性尿崩症(%)	1 (2.6)	1 (4.5)	0.683
	脑神经麻痹(%)	0 (0)	6 (27.3)	<0.001
	后续治疗			
[0117]	放射治疗(%)	4 (10.3)	20 (90.9)	<0.001
	药物治疗 (%)	0 (0)	6 (27.3)	<0.001
	手术治疗 (%)	3 (7.7)	8 (36.4)	0.002
	肿瘤控制			
	没有肿瘤(%)	29 (74.4)	0 (0)	<0.001
	肿瘤稳定性(%)	10 (25.6)	14 (63.6)	0.004
	肿瘤生长 (%)	0 (0)	8 (36.4)	<0.001
	死亡 (%)	0 (0)	2 (9.1)	0.027

[0118] 8.2RNA-seq区分INF-PA和NNF-PA

[0119] RNA-seq结果显示,与NNF-PA相比,INF-PA中843个基因表达上调,1435个基因表达下调(FDR<0.05,且|fold change|>2,图3)。

[0120] 8.3候选基因验证

[0121] 经过信号通路分析,在差异表达基因中选择17个上调的和4个下调的候选基因(图4),在另一组NF-PA样本(临床特征见表4)中进行qRT-PCR分析以验证其表达,这些候选基因包括INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ID2、ERN1、NR2F2、GAS2、CDK6、SFRP1、EMILIN2、CD9、SEMA7A。结果显示,与NNF-PA相比,INSM1、Jun、JunB、RHOU、RHOB、EZR、LLGL2、SNX33、SPIRE2、MLF1、HSPA2、CDKN1A、ERN1、NR2F2、GAS2、

CDK6、SFRP1、EMILIN2、CD9、SEMA7A基因这20个基因在INF-PA中差异表达,结果同RNA-seq(图5)。本研究证实独立队列中的RNA-seq和qRT-PCR存在一致性(图6)。

[0122] 表4用于验证的NF-PA患者临床特征

变量	RT-PCR 分组		P 值
	NNF	INF	
	92	51	
年龄(年)			0.154 <sup>a</sup>
平均值	52.8	46.4	
SD	14.2	14.0	
性别			0.007 <sup>b</sup>
男性	59	20	
女性	33	31	
P53			1.000 <sup>c</sup>
阴性	89	49	
阳性	3	2	
Ki-67			0.0869 <sup>c</sup>
≤3%	86	43	

[0123]

>3%	6	8	
最大直径			0.001 <sup>a</sup>
平均值	28.6	36.4	
SD	7.4	12.6	

[0124]

## [0125] 8.4诊断效能检测

[0126] 利用基于留一交叉验证 (LOOCV) 的支持向量机 (SVM) 算法, 结合肿瘤亚型, 检测 NNF-PAs 和 INF-PAs 之间经验证的 20 个差异表达基因的预测潜力。简言之, 本研究从 n-1 训练分类器, 其中 N 表示受试者的数量, 并且预测了第 N 个受试者的侵袭性。整个方法重复 N 次, 以预测队列中每个样本的侵袭性。RNA-seq 分析中的 61 例样本选择其中 N=41 例样本作为训练集, 训练集中 20 个差异表达基因的预测准确率为 90.24% (表 5)。随后在其他 20 例样本 (验证集) 中进行验证, 20 个差异表达基因的预测准确率为 80% (表 6)。ROC 曲线分析显示: 训练集中, 利用 20 个差异表达基因区分 INFs 和 NNFs 的 AUC 为 0.977, 验证集中; 利用 20 个差异表达基因区分 INFs 和 NNFs 的 AUC 为 0.843, 当 cutoff 值为 0.705 时, 敏感性为 0.805, 特异性为 0.760 (图 7)。上述结果表明, 本发明的 20 个差异表达基因可作为临床诊断侵袭性无功能垂体腺瘤的分子标志物。

## [0127] 表 5 训练集诊断效能

训练集	实际患病情况		
	NNF	INF	总 (n)
预测患病情况	25 (96%)	3 (20%)	28
	1 (4%)	12 (80%)	13

## [0129] 表 6 验证集诊断效能

验证集	实际患病情况		
	NNF	INF	总 (n)
预测患病情况	11 (85%)	2 (29%)	13
	2 (15%)	5 (71%)	7

[0132] 上述实施例的说明只是用于理解本发明的方法及其核心思想。应当指出, 对于本领域的普通技术人员来说, 在不脱离本发明原理的前提下, 还可以对本发明进行若干改进和修饰, 这些改进和修饰也将落入本发明权利要求的保护范围内。

## 序列表

<110> 中国医学科学院北京协和医院

中国医学科学院基础医学研究所

<120> 一种用于区分非侵袭性和侵袭性无功能垂体腺瘤的基因组合

<141> 2020-10-12

<150> 2020107727493

<151> 2020-08-04

<160> 69

<170> SIPOSequenceListing 1.0

<210> 1

<211> 21

<212> DNA

<213> 人工序列 (Artificial Sequence)

<400> 1

gaggaggttt gcctgtgatt t 21

<210> 2

<211> 22

<212> DNA

<213> 人工序列 (Artificial Sequence)

<400> 2

tcccaggact ctatccagaa tg 22

<210> 3

<211> 24

<212> DNA

<213> 人工序列 (Artificial Sequence)

<400> 3

aattatcgct agggccaagg tggg 24

<210> 4

<211> 21

<212> DNA

<213> 人工序列 (Artificial Sequence)

<400> 4

cagcagcagt gtgatgacta t 21

<210> 5

<211> 21

<212> DNA

<213> 人工序列 (Artificial Sequence)

<400> 5

cgggctcgga ggttatttat t 21  
<210> 6  
<211> 25  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 6  
tctatcactc ccaggtagct gctcc 25  
<210> 7  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 7  
cggcttcctc ttggtgatat t 21  
<210> 8  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 8  
gcttgttgta ggtgtccttg ta 22  
<210> 9  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 9  
atggccgcag ctatttcaat ggc 23  
<210> 10  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 10  
ttccgtacct cctcactcaa 20  
<210> 11  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 11  
tggaaggctc ctgtgggtat 20  
<210> 12  
<211> 24

<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 12  
ctaccactca agccttccca accc 24  
<210> 13  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 13  
cacagagaga cagacttgag aac 23  
<210> 14  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 14  
acttgatac ccttgcttt ag 22  
<210> 15  
<211> 25  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 15  
cttgacaagt tgcgacggag agaaa 25  
<210> 16  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 16  
ccatcaacat ggaagaccaa ga 22  
<210> 17  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 17  
cttgagcgtc ttcaccttgt 20  
<210> 18  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 18

catcaaagtg gagcgcaagc ggc 23  
<210> 19  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 19  
gcagggaggt gaatactctt g 21  
<210> 20  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 20  
gaagtacatc ttggccgact ta 22  
<210> 21  
<211> 26  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 21  
aaggccaaca gcaagtgttt gtggga 26  
<210> 22  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 22  
gtatgtagcg tacgggttgt t 21  
<210> 23  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 23  
cagcagttca caagccataa taaa 24  
<210> 24  
<211> 28  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 24  
aggctgtaaa ttgaactcc cacacgat 28  
<210> 25  
<211> 22

<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 25  
ttcggagcta agatggtgtt at 22  
<210> 26  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 26  
ggaaagatgg tcaagtctg t 21  
<210> 27  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 27  
taagggtggt gatgggtgag cgc 23  
<210> 28  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 28  
gatcgaggat actcaccaga aa 22  
<210> 29  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 29  
ccagttgtat ttggctggat tg 22  
<210> 30  
<211> 25  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 30  
caccatggat gcagagctgg agttt 25  
<210> 31  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 31



tgatctcacc ctcggagttt 20  
<210> 32  
<211> 19  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 32  
gcggtggttc ttggtttct 19  
<210> 33  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 33  
cacttgggtg agagagagaa gcgc 24  
<210> 34  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 34  
ccgtaacctc aaccgtttct 20  
<210> 35  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 35  
ggagtatgtc tcagcgatct tg 22  
<210> 36  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 36  
atgctttgtg cgttctggag tgga 24  
<210> 37  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 37  
gtctgcactt cctgtagcat aa 22  
<210> 38  
<211> 22

<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 38  
gataccctct gaggactctc aa 22  
<210> 39  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 39  
cacatccctg tctacacact gggc 24  
<210> 40  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 40  
ctggctcccg agaacttaa 20  
<210> 41  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 41  
gcctttgatg tggagttgt c 21  
<210> 42  
<211> 26  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 42  
agggagaaac ctcaacaaag tccagc 26  
<210> 43  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 43  
ccctggagtc ctatacctac aa 22  
<210> 44  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 44

gacacttgtc gaggatcttg tt 22  
<210> 45  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 45  
aagcagacgg tggaagacga gaaa 24  
<210> 46  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 46  
cggaacaagg agtcagacat t 21  
<210> 47  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 47  
agtgccagga aagacaacta c 21  
<210> 48  
<211> 25  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 48  
agatggtggc agtagaggct atgga 25  
<210> 49  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 49  
caagaaggtg agcaagatgg a 21  
<210> 50  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 50  
ggtgatgcag gctgacaata 20  
<210> 51  
<211> 26

<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 51  
tgcagcacgt catcgactac atcttg 26  
<210> 52  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 52  
cccatcaacc tctcttctgt atc 23  
<210> 53  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 53  
aggccgcata gtcaaagtag 20  
<210> 54  
<211> 25  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 54  
tggatgatgg gtattctgtt cgccc 25  
<210> 55  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 55  
cgtctcctcc tcagtcata a 21  
<210> 56  
<211> 21  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 56  
tgccatatac ggccagttaa a 21  
<210> 57  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 57

tgctgccgga cagtaacata tccc 24  
<210> 58  
<211> 19  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 58  
gattgcagcc aggtatggt 19  
<210> 59  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 59  
gaaggagaag gtgaaggaga ag 22  
<210> 60  
<211> 26  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 60  
tggagcctcc tggtttgata aagctg 26  
<210> 61  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 61  
ctagcaacca tccctccatt ac 22  
<210> 62  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 62  
ctcagagcat tctgaagaca gtag 24  
<210> 63  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 63  
acacagaaag ccctcttgaa gcaa 24  
<210> 64  
<211> 23

<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 64  
ggtgtgaacc atgagaagta tga 23  
<210> 65  
<211> 22  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 65  
gagtccttcc acgataccaa ag 22  
<210> 66  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 66  
agatcatcag caatgcctcc tgca 24  
<210> 67  
<211> 23  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 67  
gtaccctgtc tgactacaac atc 23  
<210> 68  
<211> 20  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 68  
gtcttgccag tgagtgtctt 20  
<210> 69  
<211> 24  
<212> DNA  
<213> 人工序列 (Artificial Sequence)  
<400> 69  
agaaaagagtc caccttgcac ctgg 24

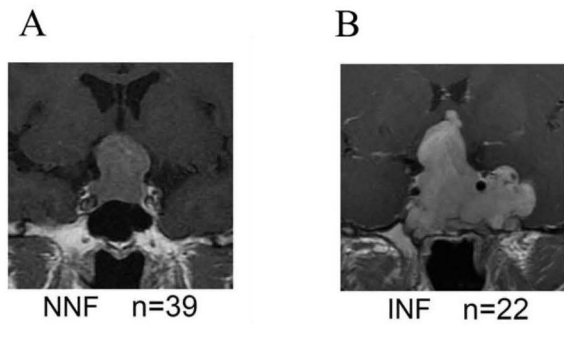


图1

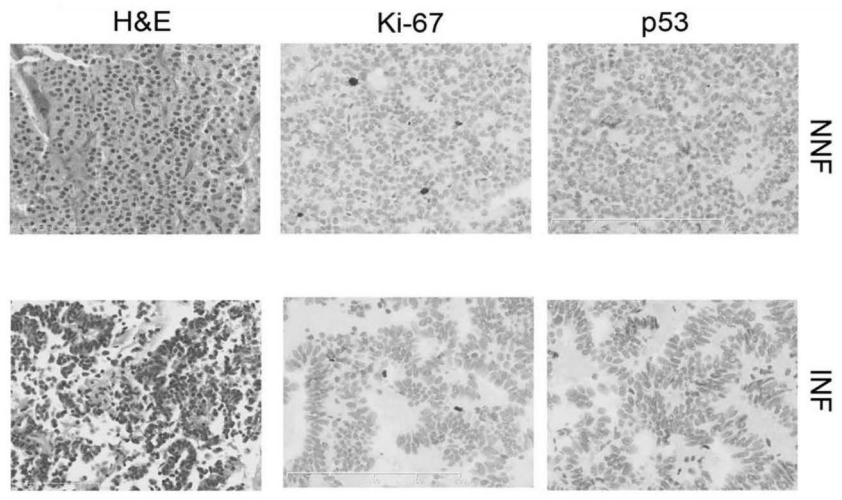


图2

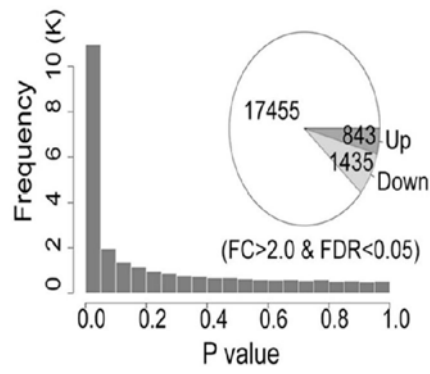


图3

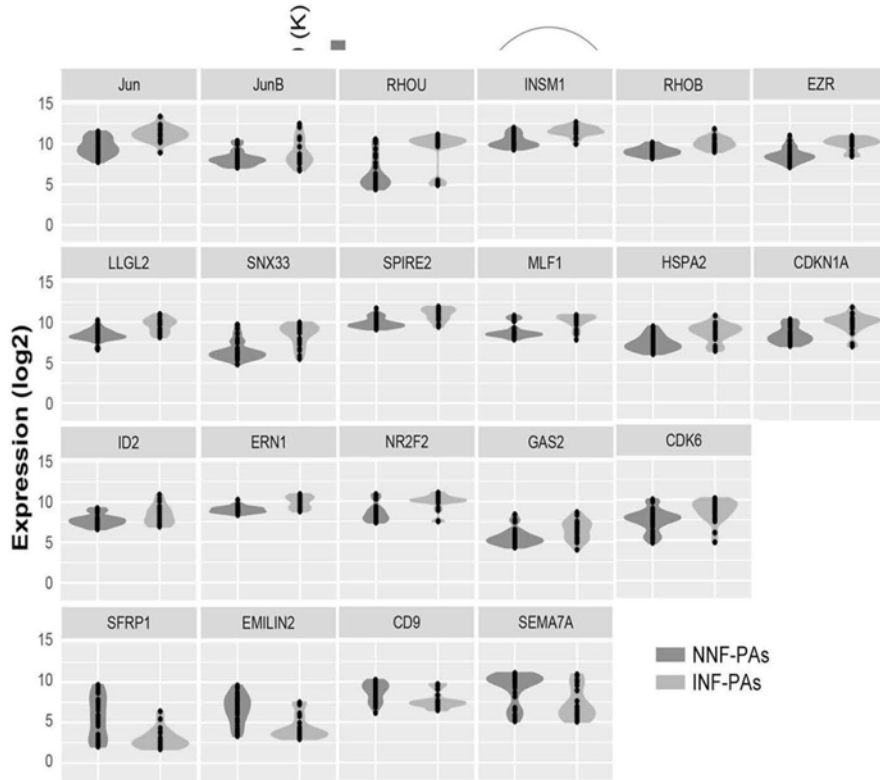


图4

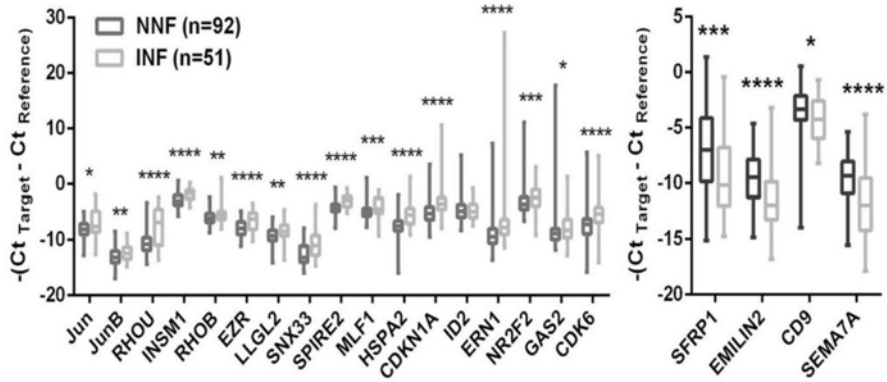


图5



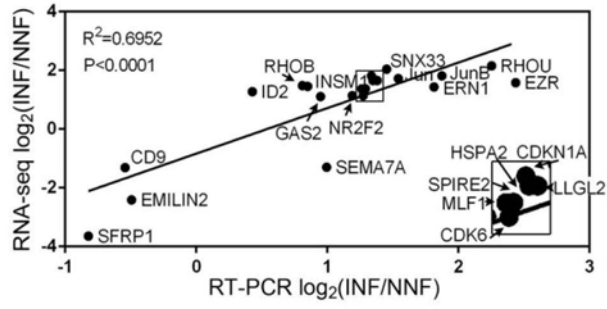


图6

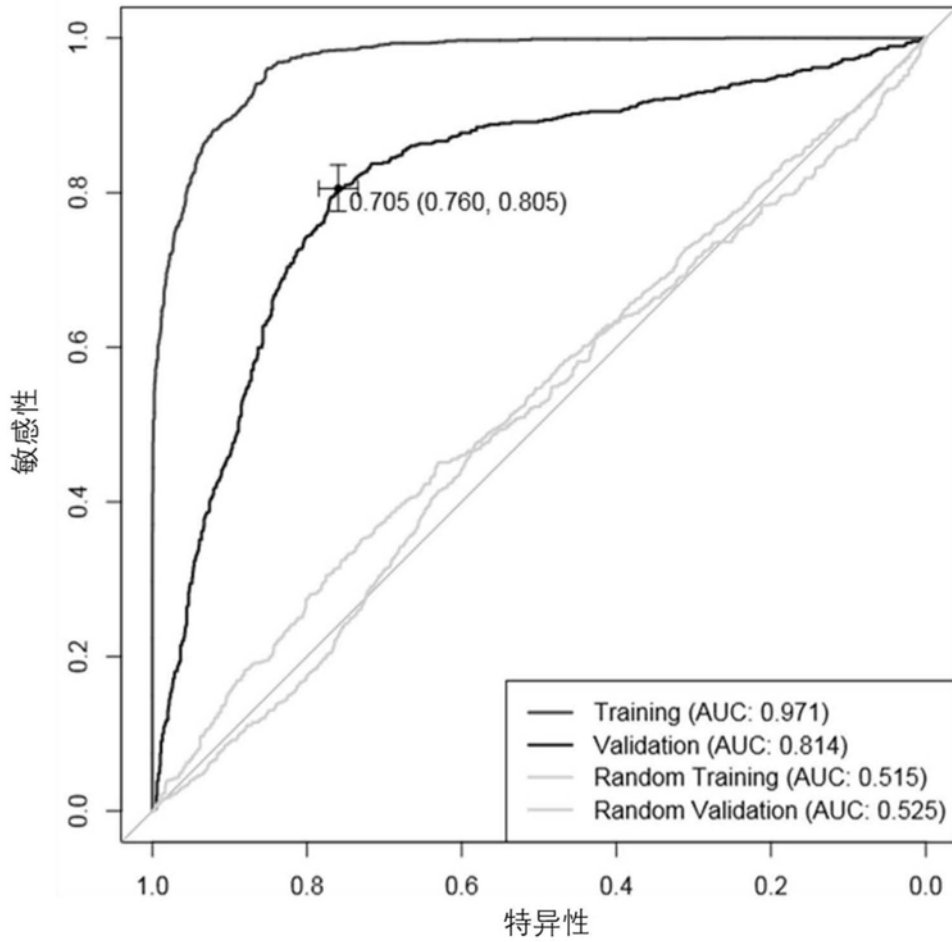


图7

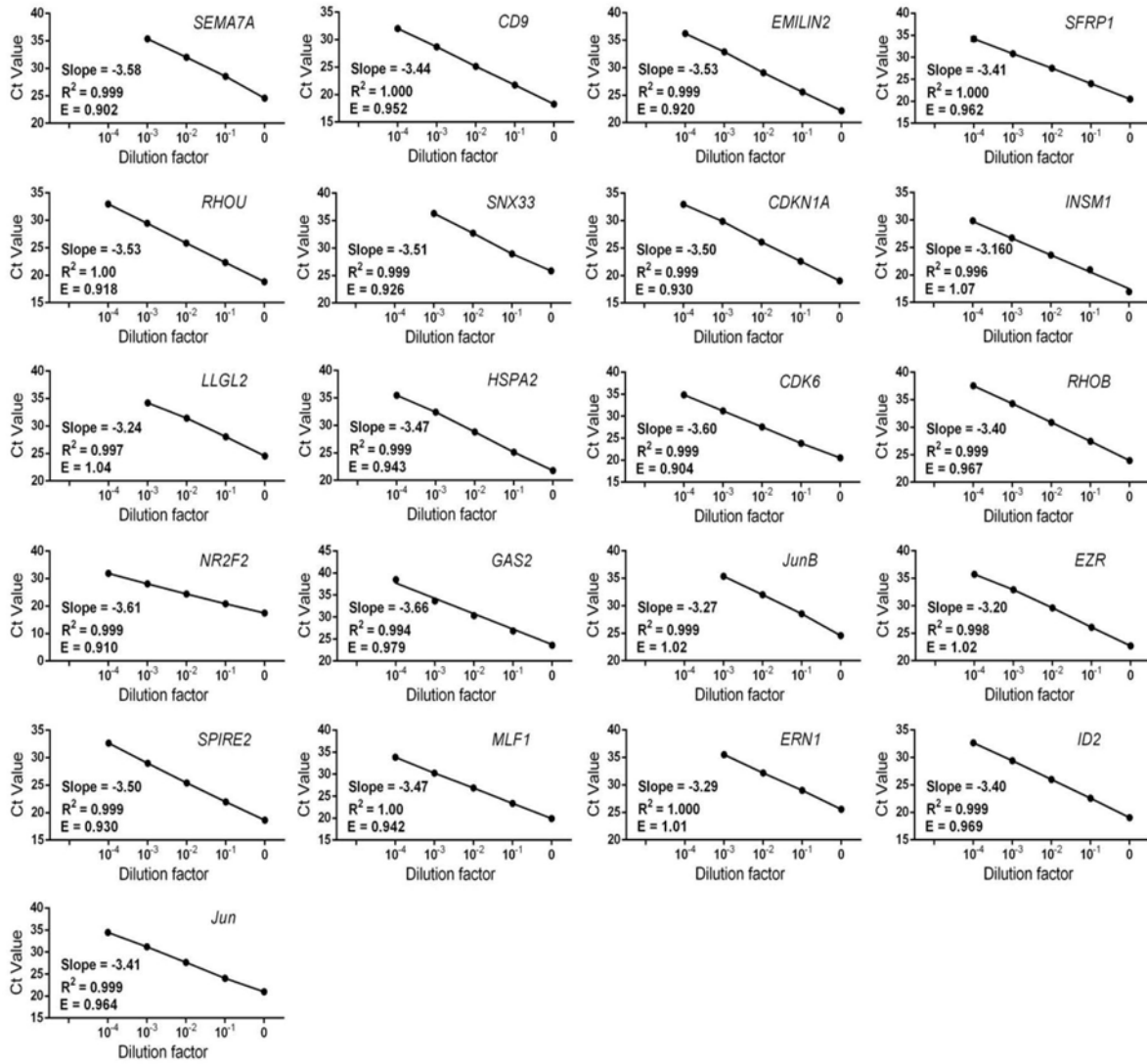


图8