



(12) 发明专利

(10) 授权公告号 CN 108350484 B

(45) 授权公告日 2024. 03. 29

(21) 申请号 201680055964.3

(22) 申请日 2016.07.22

(65) 同一申请的已公布的文献号  
申请公布号 CN 108350484 A

(43) 申请公布日 2018.07.31

(30) 优先权数据  
62/196,725 2015.07.24 US

(85) PCT国际申请进入国家阶段日  
2018.03.26

(86) PCT国际申请的申请数据  
PCT/US2016/043487 2016.07.22

(87) PCT国际申请的公布数据  
W02017/019481 EN 2017.02.02

(73) 专利权人 约翰·霍普金斯大学  
地址 美国马里兰州

(72) 发明人 H·B·拉尔曼 J·克雷德尔

(74) 专利代理机构 北京戈程知识产权代理有限公司 11314  
专利代理师 程伟

(51) Int.Cl.  
C12Q 1/6806 (2018.01)

(56) 对比文件  
US 8008010 B1, 2011.08.30  
US 2010256017 A1, 2010.10.07  
US 2013157269 A1, 2013.06.20

审查员 马艳林

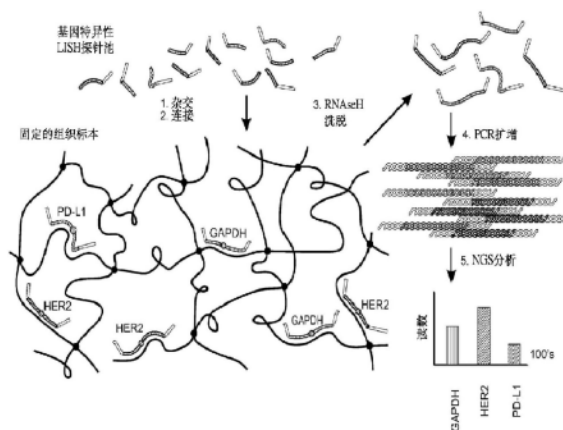
权利要求书3页 说明书22页 附图15页

(54) 发明名称

RNA分析的组合物和方法

(57) 摘要

本公开涉及RNA分析的组合物和方法。特别地,本公开提供一种RNA分析方法,该方法包括获得样品;将一种或多种多歧探针施加至该样品,其中该一个或多个多歧探针包括至少两个子探针;将所施加的一个或多个多歧探针的至少一个退火为该样品内的至少一个靶标核酸;以及将结合该至少一个退火的多歧探针的至少两个子探针连接,以创造可被检测的靶标核酸替代物。



1. 一种检测和鉴别样品中核酸序列的非疾病检测方法,包含:

获得固定的样品,其中,该固定的样品是选自由细胞、器官、组织及其任何组合所组成的群组;

将两个或多个多歧探针施加至固定的样品,其中,该两个或多个多歧探针各自包括可退火为该固定的样品内的靶标核酸中的接续靶标序列的两个子探针,其中,一个该子探针包含具有至少两个RNA碱基的3'-端的DNA寡核苷酸,且另一个该子探针包含具有5'-磷酸化的DNA寡核苷酸;

将所施加的两个或多个多歧探针的至少一个退火为该固定的样品内的至少一个靶标RNA,而不经RNA提取;

将该包含至少两个RNA碱基的至少一个子探针的至少一个退火的3'-端酶促连接至该固定的样品内的另一个该多歧探针的接续退火的5'-磷酸化端,以创造可被检测的靶标核酸替代物;

从该靶标核酸释放该靶标核酸替代物;以及

扩增该靶标核酸替代物,其中,该靶标核酸替代物可被检测。

2. 如权利要求1所述的方法,其中,该子探针包含附加引物结合位点,以促进该靶标核酸替代物的后续扩增。

3. 如权利要求1所述的方法,其中,该至少两个子探针与酶、化学反应或光反应连接。

4. 如权利要求3所述的方法,其中,该酶是连接酶。

5. 如权利要求4所述的方法,其中,该连接酶是选自由T4 RNA连接酶2(Rnl2)、T4 DNA连接酶、PBCV-1DNA连接酶、Rnl2衍生物、PBCV-1衍生物、及其任何组合组成的群组。

6. 如权利要求1所述的方法,其中,每一子探针包含寡核苷酸。

7. 如权利要求1所述的方法,其中,该固定的样品是福尔马林固定石蜡包埋(FFPE)组织样品。

8. 如权利要求1所述的方法,其中,该固定的样品是使用非福尔马林试剂固定的。

9. 如权利要求8所述的方法,其中,该非福尔马林试剂包含戊二醛、汞、氧化剂、醇、及苦味酸盐。

10. 如权利要求1所述的方法,其中,该固定的样品是由固定在悬浮体中的细胞构成。

11. 如权利要求1所述的方法,其中,该固定的样品是固定的组织培养物。

12. 如权利要求1所述的方法,其中,该固定的样品在固定之前是以高度降解的形式获得的。

13. 如权利要求1所述的方法,其中,该靶标核酸替代物是通过核酸内切酶从该靶标核酸释放的,或是通过令含有该至少两个连接的子探针的靶标核酸替代物变性而回收的。

14. 如权利要求13所述的方法,其中,该核酸内切酶包含RNase H、RNase A、RNase If、或RNase HIII。

15. 如权利要求14所述的方法,其中,该RNase H包含RNase H1、RNase H2、RNase HI及RNase HII。

16. 如权利要求1所述的方法,其中,所扩增的靶标核酸替代物是使用选自下列所组成组的技术分析:下一代测序(NGS)、深度测序、基于质谱的序列或长度分析、通过凝胶电泳或毛细管电泳进行DNA片段序列或长度的分析、以及在固定化的检测探针上杂交。

17. 如权利要求7所述的方法,其中,该福尔马林固定石蜡包埋组织样品加工为厚度为1至100微米的切片。

18. 如权利要求1所述的方法,其中,该多个多歧探针的至少两个子探针各自的长度为10至200个核苷酸。

19. 如权利要求1所述的方法,其中,该多个多歧探针的该至少两个子探针各自的长度为15至40个核苷酸。

20. 如权利要求1所述的方法,其中,该核酸替代物是使用PCR扩增的。

21. 如权利要求20所述的方法,其中,该PCR包括10至50次循环。

22. 如权利要求1所述的方法,其中,该多个多歧探针各自包括两个子探针。

23. 如权利要求1所述的方法,其中,该多个多歧探针各自包括三个子探针。

24. 如权利要求1所述的方法,其个,该方法进一步包含:

再次将两个或多个多歧探针施加至该固定的样品,其中该两个或多个多歧探针各自包括可退火为该固定的样品内的靶标核酸中的接续靶标序列的两个子探针,其中,一个该子探针包含具有至少两个RNA碱基的3'-端的DNA寡核苷酸,且另一个该子探针包含具有5'-磷酸化的DNA寡核苷酸;

将所施加的两个或多个多歧探针的至少一个退火为该固定的样品内的至少一个靶标RNA,而不经RNA提取;以及

将该包含至少两个RNA碱基的至少一个子探针的至少一个退火的3'-端酶促连接至该固定的样品内的另一个该多歧探针的接续退火的5'-磷酸化端,以创造可检测的靶标核酸替代物。

25. 一种非疾病检测方法,包含:

从受试者获得固定的组织样品;

将两个或多个多歧探针施加至该固定的组织样品,其中,该两个或多个多歧探针包括可退火为该固定的组织样品内的靶标核酸中的接续靶标序列的两个子探针,其中,一个该子探针包含具有至少两个RNA碱基的3'-端的DNA寡核苷酸,且另一个该子探针包含具有5'-磷酸化的DNA寡核苷酸;

将所施加的两个或多个多歧探针的至少一个退火为该固定的组织样品内的至少一个靶标RNA,而不经RNA提取;

将该包含至少两个RNA碱基的至少一个子探针的至少一个退火的3'-端酶促连接至该固定的组织样品内的另一个该多歧探针的接续退火的5'-磷酸化端,以创造该固定的组织样品内的靶标核酸替代物;以及

将连接的探针印在涂覆有多个固定化PCR引物的副本表面上。

26. 如权利要求25所述的方法,其特征在于,该至少两个子探针与酶、化学反应、或光反应连接。

27. 如权利要求26所述的方法,其中,该酶是连接酶。

28. 如权利要求27所述的方法,其中,该连接酶是选自T4 RNA连接酶2(Rn12)、T4 DNA连接酶、PBCV-1DNA连接酶、Rn12衍生物、PBCV-1衍生物、及其任何组合所组成的群组。

29. 如权利要求25所述的方法,其中,该固定的组织样品是福尔马林固定石蜡包埋(FFPE)组织样品。

30. 如权利要求25所述的方法,其中,该靶标核酸替代物进行固相扩增。
31. 如权利要求30所述的方法,其中,该固相扩增是桥式PCR或滚环扩增(RCA)。
32. 如权利要求25所述的方法,其中,该表面包含聚合物涂层玻璃显微镜载片,其上含有固定化的桥式PCR引物。
33. 如权利要求25所述的方法,其特征在于,该固定的组织样品包含切除的肿瘤组织或活组织检查材料。

## RNA分析的组合物和方法

[0001] 相关申请的交叉引用

[0002] 基于35U.S.C. §119(e), 本申请主张2015年7月24日提交的题为《RNA分析的组合物和方法》的第62/196,725号美国临时专利申请案的优先权, 该美国临时专利申请案通过引用而整体并入本文。

### 背景技术

[0003] 精准医学依赖于研究人员和病理学家在分子层面表征化患者标本的能力, 该标本包括切除的组织入肿瘤组织及活组织检查材料。对于与癌症相关的基因损伤的检测, 用于DNA分析的先进技术如高通量DNA测序(或下一代DNA测序(NGS))的研发功不可没。但是, 表征化固定的组织标本(例如, 福尔马林固定石蜡包埋组织切片)内基因表达谱的能力已经滞后, 这是因为固定过程降低了样品中RNA的品质, 显著抑制了从该样品中提取重要信息的能力。由于这些固定的组织标本中基因表达模式的信息非常有价值, 亟需能稳健分析固定的组织标本中RNA的平台。

### 发明内容

[0004] 本公开提供用于分析样品(如, 细胞、器官、组织等)中核酸(如, RNA)的存在和表达水平的组合物和方法。

[0005] 一方面, 本公开提供一种方法, 包括获得一样品; 将一种或多种多歧探针施加至该样品, 其中该一个或多个多歧探针包括至少两个子探针; 将所施加的一个或多个多歧探针的至少一个退火为该样品内的至少一个靶标核酸; 以及将结合该至少一个退火的多歧探针的至少两个子探针连接, 以创造可被检测的靶标核酸替代物。该方法可进一步包含从该靶标核酸释放该靶标核酸替代物, 以及扩增该靶标核酸替代物。一些情形中, 每一探针包含一寡核苷酸, 如, DNA、RNA、或DNA与RNA两者的混合物。优选地, 该靶标核酸是RNA。一些情形中, 该靶标核酸替代物经修饰。

[0006] 一方面, 该子探针包含附加引物结合位点, 其促进该靶标核酸替代物在连接后的扩增。一些方法, 每一子探针包含一寡核苷酸。其它方面, 该子探针(优选也可指代为“接纳体子探针”)具有至少两个RNA碱基的3'-端。一方面, 该子探针是具有至少两个RNA碱基的3'-端的DNA寡核苷酸。其它方面, 该子探针是具有5'-磷酸化的DNA寡核苷酸。

[0007] 一种具体例中, 该至少两个子探针可与酶、化学反应、光反应连接。该酶可以是连接酶。具体具体例中, 该酶可以是下述连接酶中的一种: T4 RNA连接酶2(Rnl2)、T4 DNA连接酶、小球藻病毒DNA连接酶(PBCV-1DNA连接酶)、Rnl2衍生物、PBCV-1衍生物、或其任何组合。

[0008] 一种具体例中, 该样品可选自由细胞、器官、组织、及其任何组合所组成的群组。一些实例中, 该样品可使用非福尔马林试剂固定, 该试剂包括, 例如, 戊二醛、汞、氧化剂、醇、及苦味酸盐。其它实例中, 该样品可包括固定在悬浮体中的细胞或固定的组织培养物。该样品可在固定之前以高度降解的形式获得。特别地, 该样品可以是福尔马林固定石蜡包埋(FFPE)组织样品。

[0009] 一种具体例中,该靶标核酸替代物可通过核酸内切酶从该靶标核酸释放,或通过将该含有该至少两个连接的子探针的靶标核酸替代物变性而回收。该核酸内切酶可以是RNase H、RNase、RNase If、或RNase HIII。

[0010] 一种具体例中,可使用下列技术分析所扩增的靶标核酸替代物,如,举例而言,下一代测序(NGS)、深度测序、基于质谱的序列或长度分析、通过凝胶电泳或毛细管电泳进行的DNA片段序列或长度分析、及/或在固定化的检测探针上杂交。

[0011] 一种具体例中,可将该固定的组织标本加工为具有1至100微米的厚度,如,1微米、5微米、10微米、15微米、20微米、25微米、30微米、35微米、40微米、45微米、50微米、60微米、65微米、70微米、75微米、80微米、85微米、90微米、95微米、或100微米。

[0012] 一种具体例中,该多个多歧探针的至少两个子探针的长度可以是约10至200个核苷酸,如长度为约10、约15、约20、约25、约30、约35、约40、约50、约75、约100、约125、约150、约175、或约200个核苷酸。其它具体例中,该多个多歧探针的至少两个子探针各自的长度可以是约15至30个核苷酸,如长度为约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、或约30个核苷酸。

[0013] 一种具体例中,该核酸替代物可使用PCR扩增。一种具体例中,该PCR包括约20至50次循环,如,约20、约25、约30、约35、约40、约45、或约50次循环。

[0014] 一种具体例中,该多个多歧探针各自包括两个子探针。相关具体例中,该多个多歧探针各自包括三个子探针。

[0015] 任选地,上述方法连续在样品上实施,即,该方法在同一样品上实施两次、三次、四次、五次、六次、七次、八次、九次、或十次。一些情形中,该标本(即,样品)在连续检验之间归档。

[0016] 本文中揭示的方法也可用于诊断感染,如病毒感染、细菌感染、或真菌感染。例如,该样品获自被怀疑患有病毒感染、细菌感染、或真菌感染或处于该感染风险下的受试者。任选地,该靶标核酸选自病毒核酸、细菌核酸、及真菌核酸所组成的组。一些情形中,该方法进一步包含从该靶标核酸释放该靶标核酸替代物;扩增该靶标核酸替代物;以及对该靶标核酸替代物测序,从而分别鉴别病毒核酸、细菌核酸或真菌核酸,并诊断该受试者体内的病毒感染、细菌感染、或真菌感染。优选在诊断后,使用抗真菌剂、抗细菌剂、或抗病毒剂治疗该受试者。

[0017] 例示性的真菌感染可包括源自下述有机体的感染:枝顶孢属(*Acremonium* sp.)、棒曲霉(*Aspergillus clavatus*)、黄曲霉(*Aspergillus flavus*)、烟曲霉(*Aspergillus fumigatus*)、灰绿曲霉(*Aspergillus glaucus*)、小巢状曲霉(*Aspergillus nidulans*)、黑曲霉(*Aspergillus niger*)、赭曲霉(*Aspergillus ochraceus*)、土曲霉(*Aspergillus terreus*)、爪甲曲霉(*Aspergillus unguis*)、焦曲霉(*Aspergillus ustus*)、球孢属(*Beauveria* sp.)、离蠕孢属(*Bipolaris* sp.)、芽裂殖菌属(*Blastoschizomyces* sp.)、皮炎芽生菌(*Blastomyces dermatitidis*)、白色念珠菌(*Candida albicans*)、光滑念珠菌(*Candida glabrata*)、吉利蒙念珠菌(*Candida guilliermondii*)、乳酒念珠菌(*Candida kefyr*)、克柔念珠菌(*Candida krusei*)、葡萄牙念珠菌(*Candida lusitanae*)、近平滑念珠菌(*Candida parapsilosis*)、热带念珠菌(*Candida tropicalis*)、金孢子菌属(*Chrysosporium* sp.)、枝孢属(*Cladosporium* sp.)、粗球孢子菌(*Coccidioides*

immitis)、新型隐球菌洛特变种(*Cryptococcus neoformans* var *gattii*)血清型B、新型隐球菌(*Cryptococcus neoformans*)血清型A、罗伦隐球菌(*Cryptococcus laurentii*)、地生隐球菌(*Cryptococcus terreus*)、弯孢属(*Curvularia* sp.)、镰孢属(*Fusarium* sp.)、蒴果线黑粉菌(*Filobasidium capsuligenum*)、新型线黑粉菌(隐球菌)杆状变种(*Filobasidiella* (*Cryptococcus*) *neoformans* var *bacillispora*)血清型C、新型线黑粉菌(隐球菌)新型变种(*Filobasidiella* (*Cryptococcus*) *neoformans* var *neoformans*)血清型D、线黑粉菌(*Filobasidium uniguttulatum*)、地丝菌属(*Geotrichum* sp.)、荚膜组织胞浆菌(*Histoplasma capsulatum*)、枝霉属(*Malbranchea* sp.)、毛霉属(*Mucor* sp.)、拟青霉属(*Paecilomyces* sp.)、巴西副球孢子菌(*Paracoccidioides brasiliensis*)、青霉属(*Penicillium* species)、卡氏肺孢子虫(*Pneumocystis carinii*)、波氏假性阿利什霉(*Pseudallescheria boydii*)、根霉属(*Rhizopus* sp.)、申克氏孢子丝菌(*Sporothrix schenkii*)、短帚霉属(*Scopulariopsis brevicaulis* sp.)、布氏帚霉(*Scopulariopsis brumptii*)、酿酒酵母(*Saccharomyces cerevisiae*)、及白吉利毛孢子菌(*Trichosporon beigeli*)。

[0018] 例示性细菌感染可包括源自下述有机体的感染：炭疽杆菌(*Bacillus anthracis*)、百日咳杆菌(*Bordetella pertussis*)、伯氏疏螺旋体(*Borrelia burgdorferi*)、流产布鲁氏菌(*Brucella abortus*)、犬布鲁氏菌(*Brucella canis*)、羊布鲁氏菌(*Brucella melitensis*)、猪布鲁氏菌(*Brucella suis* Brucellosis)、空肠弯曲菌(*Campylobacter jejuni*)、肺炎衣原体(*Chlamydia pneumonia*)呼吸道感染、鹦鹉热衣原体(*Chlamydia psittaci*)、沙眼衣原体(*Chlamydia trachomatis*)、性病淋巴肉芽肿(*Lymphogranuloma venereum*)、肉毒梭菌(*Clostridium botulinum*)、艰难梭菌(*Clostridium difficile*)、产气荚膜梭菌(*Clostridium perfringens*)、破伤风梭菌(*Clostridium tetani*)、白喉棒状杆菌(*Corynebacterium diphtheria*)、粪肠球菌(*Enterococcus faecalis*)、屎肠球菌(*Enterococcus faecium*)、大肠杆菌(*Escherichia coli*)、土拉热杆菌(*Francisella tularensis*)、流感嗜血杆菌(*Haemophilus influenza*)、幽门螺杆菌(*Helicobacter pylori*)、嗜肺军团菌(*Legionella pneumophila*)、肾脏钩端螺旋体(*Leptospira interrogans*)、单核细胞增多性李斯特菌(*Listeria monocytogenes*)、麻风分枝杆菌(*Mycobacterium leprae*)、结合分支杆菌(*Mycobacterium tuberculosis*)、肺炎支原体(*Mycoplasma pneumonia*)、淋病奈瑟菌(*Neisseria gonorrhoeae*)、脑膜炎奈瑟菌(*Neisseria meningitidis*)、绿脓假单胞菌(*Pseudomonas aeruginosa*)、立克次体(*Rickettsia rickettsia*)、伤寒沙门菌(*Salmonella typhi*)、鼠伤寒沙门菌(*Salmonella typhimurium*)、索氏志贺菌(*Shigella sonnei*)、金黄色葡萄球菌(*Staphylococcus aureus*)、表皮葡萄球菌(*Staphylococcus epidermidis*)、腐生葡萄球菌(*Staphylococcus saprophyticus*)、无乳链球菌(*Streptococcus agalactiae*)、肺炎链球菌(*Streptococcus pneumoniae*)、酿脓链球菌(*Streptococcus pyogenes*)、梅毒螺旋体(*Treponema pallidum*)、霍乱弧菌(*Vibrio cholera*)、及鼠疫耶尔森菌(*Yersinia pestis*)。

[0019] 例示性病毒感染可包括源自下述有机体的感染：腺病毒、禽流感病毒、甲型流感病毒、乙型流感病毒、麻疹病毒、副流感病毒、呼吸道合胞体病毒(RSV)、鼻病毒、非典型肺炎病毒(SARS-CoV(严重急性呼吸系统综合征病毒))、柯萨奇病毒、肠道病毒、脊髓灰质炎病毒、

轮状病毒、乙型肝炎病毒、丙型肝炎病毒、牛病毒性腹泻病毒(代用品)、1型单纯性疱疹病毒、2型单纯性疱疹病毒、人巨细胞病毒、水痘带状疱疹病毒、1型人免疫缺陷病毒(HIV-1)、2型人免疫缺陷病毒(HIV-2)、猴免疫缺陷病毒(SIV)、猴人免疫缺陷病毒(SHIV)、登革热病毒、汉他病毒、出血热病毒、淋巴性脉络丛脑膜炎病毒、天花病毒代用品(牛痘病毒、猴痘病毒、兔痘病毒)、痘苗病毒、委内瑞拉马脑炎病毒(VEE)、西尼罗河病毒、黄热病病毒、寨卡病毒。

[0020] 一方面,本公开提供一种方法,该方法包括从受试者获得固定的组织样品;将一种或多种多歧探针施加至该固定的组织样品,其中该一个或多个多歧探针各自包括至少两个子探针;将所施加的一个或多个多歧探针的至少一个退火为该固定的组织样品内的靶标核酸;将结合该至少一个退火的多歧探针的至少两个子探针连接,以创建靶标核酸替代物;以及将连接的探针印在涂覆有多个固定化PCR引物的副本表面上。

[0021] 一种具体例中,该至少两个子探针可与酶、化学反应、或光反应连接。一种具体例中,该酶可以是连接酶。特别地,该连接酶可以是T4 RNA连接酶2(Rn12)、T4 DNA连接酶、小球藻病毒DNA连接酶(PBCV-1DNA连接酶)、Rn12衍生物、PBCV-1衍生物、或其任何组合。

[0022] 一种具体例中,该样品是福尔马林固定石蜡包埋(FFPE)组织样品。

[0023] 一种具体例中,令该靶标核酸替代物进行固相扩增。一种具体例中,该固相PCR扩增可以是桥式PCR或滚环扩增(RCA)。一种具体例中,该靶标表面包含聚合物涂层玻璃显微镜载片,且该表面含有固定化的PCR引物。

[0024] 一种具体例中,该组织样品包含切除的肿瘤或活组织检查材料。

[0025] 另一具体例中,本公开提供一种方法,该方法包括从受试者获得固定的组织样品;将一个或多个多歧探针施加至该固定的组织样品,其中,该一个或多个多歧探针各自包括至少两个子探针;将所施加的一个或多个多歧探针的至少一个退火为该固定的组织样品内的靶标核酸;将结合该至少一个退火的多歧探针的至少两个子探针连接,以创建该固定的组织样品内的靶标核酸替代物;以及,将所连接的探针印在具有多个固定化的PCR引物的副本表面上。在具体例中,这一方法可称为“印记途径”或“印记方法”。

[0026] 一方面,该子探针与酶、化学反应、或光反应连接。一些实例中,该酶可以是选自T4 RNA连接酶2(Rn12)、T4 DNA连接酶、小球藻病毒DNA连接酶(PBCV-1DNA连接酶)、Rn12衍生物、PBCV-1衍生物、及其任何组合所组成组的连接酶。

[0027] 一方面,上述印记方法可使用福尔马林固定石蜡包埋(FFPE)组织样品、切除的肿瘤组织或活组织检查材料。其它具体例中,该靶标核酸替代物进行固相扩增,包括,例如,桥式PCR或滚环扩增(RCA),该扩增使用含有固定化桥式PCR引物的聚合物涂层玻璃显微镜载片。

## 附图说明

[0028] 图1A至图1C分别描绘电泳图谱、柱状图和曲线图。图1A是显示酶促连结后的毛细管电泳的图。RNA碱基为灰色,DNA碱基为黑色,而FAM是荧光标记。T4 DNA连接酶并不产生可测量的链接产物(上图)。Rn12(T4 RNA连接酶2)在RNA模板链的存在下连接嵌合的DNA-RNA探针组(中图),但在缺少RNA模板链时不连接(下图)。图1B描绘显示在细胞RNA上评估7种以GAPDH(甘油醛3-磷酸脱氢酶)mRNA为靶向的独特的RASL(RNA退火、选择和连接测序)探针组



的柱状图。显示各探针组在Rn12与T4 DNA连接酶连接效率间的倍数差异。图1C描绘显示来自设计为测量恒定背景的细胞RNA中的添加RNA序列的基于Rn12的RASL检验的信号对SYBR qPCR检验的背景值(BKGD)的曲线图。

[0029] 图2A至图2C分别描绘图解、流程图、及图像。图2A显示一描绘原位连接杂交测序(LISH-seq)的流程图。将基因特异性探针组的池杂交至位于该固定的组织内的其靶标RNA上。洗除过量探针后,Rn12连接邻近退火的探针。随后回收连接产物,扩增并进行序列分析。图2B描绘一说明初步实验的示意图,如所揭示的,该实验通过使用冷冻组织标本和FFPE(福尔马林固定石蜡包埋)组织标本两者的RT-qPCR、RASL和LISH来比较转录本定量。图2C是说明将LISH-seq用于显微切割组织分析的图像。

[0030] 图3A至图3B分别显示模板图解和模板/副本图像。图3A显示说明将DNA印在靶标表面上的模板图解。显示了用于ssDNA微阵列的用以将导生物素化cDNA引导至合并有链霉素和素的凝胶膜涂覆的复本表面上的模板。图3B是说明被转移至副本表面的Cy5/生物素标记的cDNA的图像。(i)显示转移前的模板,(ii)显示转移后的模板。副本表面在与Cy3互补序列杂交前的图像为(iii),而杂交后的图像为(iv)。斑点为约100微米。

[0031] 图4描绘根据本公开的例示性具体例对福尔马林固定石蜡包埋(FFPE)样品的LISH印记探针分析。上图描绘该LISH印记过程。LISH探针杂交和连接将在FFPE切片上实施。中图描绘将连接产物印记在桥式PCR靶标表面(每一点表示单分子的连接产物)上。下图描绘桥式PCR,其将会原位扩增LISH信号,用于使用杂交探针板进行的后续分析。

[0032] 图5A至图5D分别描绘图解、柱状图、凝胶图谱、及柱状图。图5A揭示途径的图解,经由该途径,有效地测量FFPE标本中多种RNA序列而无需RNA纯化或逆转录。图5B是柱状图,说明DNA退火的含有二核糖核苷酸的DNA序列被RNAse H2裂解,而RNA退火的序列未被裂解。图5C是凝胶图谱,说明FFPE RNA可作为适当的模板用于Rn12介导的探针原位连接中。图5D是柱状图,该图说明,通过下一代DNA测序对RNAse H回收的连接产物的端点PCR分析显示,正确的探针对确实是原位连接的,而误配的探针组并未以显著水平连接。

[0033] 图6A至6C分别描绘活组织检查图像、凝胶图谱、及柱状图,说明LISH可用来使用保存的临床FFPE标本诊断感染。图6A描绘被喙状凸脐孢(*Exserohilum rostratum*) (左图)和烟曲霉(右图)感染的脑组织的活组织检查。图6B描绘凝胶图谱,说明喙状凸脐孢检测器探针报告来自指标患者的强上靶信号,但来自含有不相关感染如密切相关的烟曲霉感染的无关感染的组织或切片无信号。图6C是柱状图,说明喙状凸脐孢检测器探针报告来自指标患者的强上靶信号,但来自含有不相关感染如密切相关的烟曲霉感染的无关感染的组织或切片无信号。

[0034] 图7A至7C分别描绘图解、凝胶图谱、及凝胶图谱,说明FFPE RNA可作为适当的模板用于Rn12介导的探针原位连接中。图7A描绘一图解,说明将会在两个不同探针组之间出现的可能的“上靶”和“脱靶”子探针连接事件。图7B描绘一凝胶图谱,说明对强制连接反应的端点PCR分析显示,在不适当的条件下生产并轻易地检测脱靶连接产物是可能的。图7C描绘一凝胶图谱,说明为了实现两个探针之间的可检测的连接产物,对于Rn12介导的连接的绝对要求。

[0035] 图8描绘一说明用以回收LISH产物的激光捕获显微切割(LCM)的图解。评估了从LCM样品获得的大小为 $200\mu\text{m}^2$ 、 $2000\mu\text{m}^2$ 、和 $20,000\mu\text{m}^2$ 的GAPDH探针组连接产物。

## 具体实施方式

[0036] 定义

[0037] “剂”意为任何小的化合物、抗体、核酸分子,或多肽或其片段。

[0038] “改变”意为如通过该领域已知的标准方法如本文中揭示的那些方法所测,基因或多肽的表达水平或活性的变化(增加或降低)。本文中,改变包括表达水平变化10%,优选变化25%,更优选变化40%,且最优选表达水平变化50%或更高。

[0039] “桥式PCR”是局部的固相PCR扩增方法,在该方法中,通过附着至表面(而非溶液中)的引物而扩增DNA,因此形成表面固定化DNA簇。若这些簇充分分离,它们可克隆测序。

[0040] 本公开中,“包含”、“含有”和“具有”等可具有与它们在美国专利法中认为的相同意义,且可意为“包括”等;“主要由…组成”等具有与它们在美国专利法中认为的相同意义且是开放性的,允许存在超过其所引用者,只要其所设定的基本特征或新颖特征未因存在超过其所引用者而变化即可,但需排除背景技术具体例。

[0041] “检测”指的是鉴别待检测的核酸(如,RNA)存在、不存在、或量。

[0042] “可检测的标记”意为一种组合物,当链接至感兴趣的分子时,该组合物令该分子经由例如光谱学、光化学、生物化学、免疫化学、或化学手段可检测。例如,可用的标记可包括放射活性同位素、磁珠、金属珠、胶体粒子、荧光染料、高电子密度试剂、酶(例如,如常用于ELISA的酶)、生物素、异羟基洋地黄毒苷、或半抗原。

[0043] “有效量”意为相对于未处理的患者改善疾病症状所需的剂量。用以实践的用于治疗性处理疾病的活性剂的有效量,依据给药模式和受试者的年龄、体重和一般健康状况而变。基本上,主治医师或兽医将决定适宜的和计量方案。该量称为“有效”量。

[0044] “FFPE”意为福尔马林固定石蜡包埋组织。FFPE样品源自下述组织(一般为疑似肿瘤样品):使用福尔马林固定以保持结构-空间特征和生物分子特征(如,细胞支架机构和蛋白结构),随后包埋在一种类型的石蜡中,因此该组织可切片,典型为5至10微米厚。福尔马林经由氨基不可逆地交联蛋白质,因此保持细胞的结构完整,如此它们可使用染料或用以分析该组织中表明细胞状况改变如癌症的畸形的免疫染色剂染色。但是,这些交联固定剂对该样品中RNA和DNA核酸的影响,不利于当前的分子检验如使用源自FFPE样品的DNA或RNA的分子检验中可实现的敏感性和特异性。此外,可使用非福尔马林试剂,包括,例如,戊二醛、汞、氧化剂、醇、及苦味酸盐制备样品。

[0045] “片段”意为多肽或核酸分子的一部分。这一部分优选含有参考核酸分子或多肽整体长度的至少10%、20%、30%、40%、50%、60%、70%、80%、或90%。片段可含有10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、或1000个核苷酸或氨基酸。

[0046] “杂交”意为互补核酸碱基间的氢键键结,可以是Watson-Crick、Hoogsteen、或反向Hoogsteen氢键键结。例如,腺嘌呤和胸腺嘧啶是互补核酸碱基,它们通过形成氢键而配对。

[0047] “感染”意为致病剂侵入有机体内,它们的增加以及宿主与这些有机体的反应及它们产生的毒素。该感染可由任何微生物/微生物类造成,该微生物包括,例如,细菌、真菌、和病毒。微生物类可包括所有细菌物种、古生物种、和原生动物种。这一种群亦可含有真菌、藻类、及某些动物的一些物种。一些具体例中,病毒亦可归类为微生物类。

[0048] “分离的多核苷酸”意为不含基因的核酸(如,RNA、DNA、cDNA等),在本发明的核酸分子所来自的有机体的天然出现的基因型中,该核酸处于该基因侧翼。该术语因此包括,例如,合并入载体内、合并入自主复制质粒或病毒中、或合并入原核生物或真核生物的基因组DNA中的重组DNA;或作为独立于其它序列的独立分子(例如,通过PCR或限制性核酸内切酶消解产生的cDNA或基因组或cDNA片段)存在的DNA。此外,该术语包括从DNA分子转录的RNA分子,以及作为编码另外的多肽序列的杂交基因的一部分的重组DNA。

[0049] “激光捕获显微切割”或“LCM”是用于从组织、细胞或有机体的显微切割区域分离特定细胞的方法。LCM是用以在直接显微镜观察下取得组织细胞的亚种群的方法。LCM技术可直接收获感兴趣的细胞,或可通过切除不想要的细胞以给出组织学纯富集的细胞群落,从而分离特定的细胞。

[0050] 原位连接杂交(Ligation in situ Hybridization)或“LISH”是本文中公开的用于基因表达的多重测量的方法,适用于固定的组织标本的分析。一种LISH具体例使用T4 RNA连接酶2(Rn12)和嵌合DNA-RNA杂交探针,该探针当退火时则变为连接,其中,该探针作为可靠的替代品用于各靶标的表达水平。

[0051] “LISH-印记”是将LISH产物印记并扩增在靶标表面上,以保存空间信息的方法。

[0052] “标记物”意为表达水平和活性具有与疾病或病变相关的改变的任何蛋白质或多核苷酸。

[0053] “调节”意为改变(增加或降低)。通过该领域已知的标准方法如本文中揭示的那些方法检测这些改变。

[0054] “瘤形成”意为以过度增殖或细胞凋亡减少为特征的疾病或病变。本发明可应用的例示性肿瘤可包括,但不限于,胰腺癌、白血病(如,急性白血病、急性淋巴细胞性白血病、急性髓细胞性白血病、急性成髓细胞性白血病、急性前髓细胞性白血病、急性骨髓单核细胞性白血病、急性单核细胞性白血病、急性红白血病、慢性白血病、慢性粒细胞性白血病、慢性淋巴细胞性白血病)、真性红细胞增多症、淋巴瘤(霍奇金氏病、非霍奇金氏病)、巨球蛋白血症、重链病、以及实体瘤如肉瘤和癌(如,纤维肉瘤、粘液肉瘤、脂肪肉瘤、软骨肉瘤、骨源性肉瘤、脊索癌、血管肉瘤、内皮肉瘤、淋巴管肉瘤、淋巴管内皮肉瘤、滑膜瘤、间皮瘤、尤文氏肿瘤、平滑肌肉瘤、横纹肌肉瘤、结肠癌、乳腺癌、卵巢癌、前列腺癌、鳞状细胞癌、基底细胞癌、腺癌、汗腺癌、皮脂腺癌、乳头状癌、乳头状腺癌、囊腺癌、髓样癌、支气管癌、肾细胞癌、肝癌、胆管癌、绒毛膜癌、精原细胞瘤、胚胎性癌、威尔姆斯瘤、宫颈癌、子宫癌、睾丸癌、肺癌、小细胞肺癌、膀胱癌、上皮细胞癌、胶质瘤、多形性成胶质细胞瘤、星形细胞瘤、成神经管细胞瘤、颅咽管瘤、室管膜瘤、松果体瘤、成血管细胞瘤、听神经瘤、少突神经胶质瘤、神经鞘瘤、脑膜瘤、黑素瘤、成神经细胞瘤、及成视网膜细胞瘤)。

[0055] “NGS”意为下一代测序。NGS平台大量地实施并行测序,在该过程中,来自单一样品的数百万DNA片段得以一致地测序。大规模并行测序技术促进高通量测序,令整个基因组在不足一天内测序完成。NGS平台的创建已经令测序为更多实验室所接受,并使得使用核酸测序实施的研究和临床诊断的量快速增长。

[0056] 可用于本发明方法中的核酸分子包括编码本发明的多肽或其片段的任何核酸分子。这些核酸分子无需与内源性核酸序列100%一致,但典型显现基本一致性。具有与内源性序列“基本一致性”的多核苷酸典型地能与单链或双链核酸分子的至少一条链杂交。“杂

交”意为在互补的多核苷酸序列(如,本文中揭示的基因)或其部分之间在各种严格条件下配对,以形成双链分子(见,如,Wahl,G.M.and S.L.Berger (1987) *Methods Enzymol.* 152: 399;Kimmel,A.R. (1987) *Methods Enzymol.* 152:507)。

[0057] 例如,严格的盐浓度通常低于约750mM NaCl和75mM柠檬酸三钠,优选低于约500mM NaCl和50mM柠檬酸三钠,更优选低于约250mM NaCl和25mM柠檬酸三钠。低严格度杂交可在有机溶剂如甲酰胺的不存在下获得;而高严格度杂交可在至少约35%甲酰胺,更优选至少约50%甲酰胺的存在下获得。严格的温度条件通常包括至少约30℃的温度,更优选至少约37℃,且最优选至少约42℃。改变其它参数,如杂交时间、洗涤剂如十二烷基硫酸钠(SDS)的浓度、以及包含或不饱和载体DNA,是该领域技术人员所周知的。通过如所需者合并这些多种条件而完成各种水平的严格度。一种优选具体例中,杂交将在30℃出现于750mM NaCl、75mM柠檬酸三钠、及1% SDS中。一种更优选的具体例中,杂交将在37℃出现于500mM NaCl、50mM柠檬酸三钠、1% SDS、35%甲酰胺、及100μg/ml变性的大马哈鱼精子DNA(ssDNA)中。一种最优选的具体例中,杂交将在42℃出现于250mM NaCl、25mM柠檬酸三钠、1% SDS、50%甲酰胺、及200μg/ml ssDNA中。该领域技术人员可轻易获知对这些条件的可用改变。

[0058] 对于大多数应用,杂交后的洗涤步骤的迫切性亦可变。洗涤的迫切性条件可由盐浓度和温度而定。如上,洗涤的迫切性可由于盐浓度增加或温度升高而增加。例如,对于洗涤步骤的严格的盐浓度优选为低于约30mM NaCl和3mM柠檬酸三钠,且最优选低于约15mM NaCl和1.5mM柠檬酸三钠。对于洗涤步骤的严格温度通常包括至少约25℃的温度,更优选至少约42℃,且甚至更优选至少约68℃。一种优选具体例中,洗涤步骤将在25℃出现于30mM NaCl、3mM柠檬酸三钠、及0.1% SDS中。一种更优选的具体例中,洗涤步骤将在42℃出现于15mM NaCl、1.5mM柠檬酸三钠、及0.1% SDS中。一种更优选的具体例中,洗涤步骤将在68℃出现于15mM NaCl、1.5mM柠檬酸三钠、及0.1% SDS中。该领域技术人员可轻易获知对这些条件的其它改变。杂交技术是该领域技术人员周知的,且揭示于例如Benton and Davis (*Science*196:180,1977)、Grunstein and Hogness (*Proc.Natl.Acad.Sci.,USA* 72:3961, 1975)、Ausubel et al. (*Current Protocols in Molecular Biology*,Wiley Interscience,New York,2001)、Berger and Kimmel (*Guide to Molecular Cloning Techniques*,1987,Academic Press,New York)、和Sambrook et al.,*Molecular Cloning: A Laboratory Manual*,Cold Spring Harbor Laboratory Press,New York)中。

[0059] 本文中使用的“测序”或任何语法等效物可指代用以将所扩增的靶标核酸替代物测序的方法。测序技术可包括,例如,下一代测序(NGS)、深度测序、基于质谱的序列或长度分析、或通过凝胶电泳或毛细管电泳进行的DNA片段序列或长度分析。可使用的兼容测序技术包括,单分子实时测序(太平洋生命科学(Pacific Biosciences))、离子半导体(离子激流测序(Ion Torrent sequencing))、焦磷酸测序(454)、通过合成测序(伊诺米纳公司(Illumina))、通过连接测序(SOLiD测序)、链终止(Sanger测序)、纳米孔DNA测序(牛津纳米科技(Oxford Nanosciences Technologies))、Heliscope单分子测序(Helioscope inc)、使用质谱测序、DNA纳米球测序、通过杂交测序、以及隧道电流DNA测序。

[0060] “基本一致”意为多肽或核酸分子展现与参考氨基酸序列(例如,本文中揭示的任一种氨基酸序列)或核酸序列(例如,本文中揭示的任一种核酸序列)的至少50%一致性。优选地,该序列在氨基酸水平或核酸水平上具有与用于比较的序列的至少60%,更优选80%

或85%，且更优选90%、95%或甚至99%一致性。

[0061] 典型地，使用序列分析软件（例如，威斯康星大学生物技术中心（University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705）基因计算机小组（Genetics Computer Group）的序列分析软件包BLAST、BESTFIT、GAP、或PILEUP/PRETTYBOX程序）测量序列一致性。该软件通过设定各种替换、删除、及/或其它修饰的同源性程度来匹配一致的序列或相似的序列。保守替换典型包括在下列组别中的替换：甘氨酸、丙氨酸；缬氨酸、异亮氨酸、亮氨酸；天冬氨酸、谷氨酸、天冬酰胺、谷酰胺；丝氨酸、苏氨酸；赖氨酸、精氨酸；以及苯丙氨酸、酪氨酸。在用以确定一致性程度的例示性途径中，可使用BLAST程序，评分为 $e^{-3}$ 与 $e^{-100}$ 之间的可能性说明密切相关的序列。

[0062] “引物组”意为一组可用于例如PCR的寡核苷酸。引物组可由至少2、4、6、8、10、12、14、16、18、20、30、40、50、60、80、100、200、250、300、400、500、600、或更多个引物组成。

[0063] 本文中提供的范围理解为该范围内所有值的略写。例如，1至50的范围理解为来自1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、或50所组成组的任何数字、数字的组合、或子范围，以及前述整数之间的所有具有小数的中间值如，举例而言，1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、和1.9。对于子范围，从该范围一个端点延伸的“嵌套子范围”明确处于预期之内。例如，例示性的1至50范围的嵌套子范围可包含一个方向的1至10、1至20、1至30、和1至40，或另一方向的50至40、50至30、50至20、和50至10。

[0064] “降低”意为负向变更至少10%、25%、50%、75%、或100%。

[0065] “参考”意为标准或对照条件。

[0066] “参考序列”为所定义的用作序列比对或基因表达比对基准的序列。参考序列可以是特定序列的子集或整体；例如，全长度cDNA或基因序列的节段，或完整的cDNA或基因序列。对于多肽，参考多肽序列的长度通常为至少约16个氨基酸，优选至少约20个氨基酸，更优选至少约25个氨基酸，且甚至更优选约35个氨基酸、约50个氨基酸、或约100个氨基酸。对于核酸，参考核酸的长度一般为至少约40个核苷酸，优选至少约60个核苷酸，更优选至少约75个核苷酸，且甚至更优选约100个核苷酸、或约300、或约500个核苷酸、或接近该数字或该数字间的任何整数个核苷酸。

[0067] 如本文中所使用，“获得剂”中的“获得”包括合成、购买、或以另外途径撷取该剂。

[0068] “受试者”意为哺乳动物，包括但不限于，人或非人哺乳动物，如牛、马、犬、羊或猫。

[0069] 如本文中所使用，术语“子探针”可指结合该接续靶标序列而不留下任何未键结的介入核苷酸的该两个或多个探针的任何探针。一些具体例中，本文中揭示的多歧探针可包括至少量“子探针”。另一具体例中，该多个多歧探针的至少两个子探针各自的长度可以是约10至50个核苷酸。一旦该探针被连接，所连接的多歧探针（或者，“所连接的子探针”）可从该RNA释放。一些具体例中，该子探针可含有附加引物结合位点（如，适配因子），以促进该靶标核酸替代物的后续扩增。其它具体例中，该两个或多个子探针的至少一个可称为“接纳子探针”，其具有至少两个RNA碱基的3' -端。

[0070] 如本文中所使用，“附加引物结合位点”可指本文中揭示的多歧探针或子探针中促进该靶标核酸替代物的扩增的结合位点。“附加引物结合位点”亦可称为“适配因子”。

[0071] 如本文中所使用，术语“处理”等指的是减轻或缓解病变及/或其相关症状。应理

解,尽管未排除,但处理病变或病症并不需要完全消除该病变、病症或其相关症状。

[0072] 如本文中所使用,术语“防止(prevent)”、“防止(preventing)”、“防止(prevention)”、“预防性处理”等指的是,降低未罹患病变或病症但处于发展病变或病症的风险下或易患疾病或病症的受试者体内发展出该病变或病症的可能性。

[0073] 除非从语境中明确指出或显而易见,本文中,术语“或”理解为内含的。除非从语境中明确指出或显而易见,本文中,术语“一”和“该”理解为单数或复数。

[0074] 除非从语境中明确指出或显而易见,本文中,术语“约”理解为处于该领域正常公差的范围,例如,处于均值的2标准偏差内。“约”可理解为处于所修饰值的10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%、或0.01%以内。除非从语境中明确排除,本文中所提供的所有数值可以术语“约”修饰。

[0075] 本文中对变量的任何定义中所列举化学基团的叙述,包括该变量作为任何单一基团或所列基团组合的定义。对本文中变量或方面的具体例的叙述,包括该具体例作为任何单一具体例或与其它具体例或其一部分组合。

[0076] 本文提供的任何组合物或方法可与一种或多种本文提供的任何其它组合物和方法组合。

[0077] 从下述对本发明优选具体例的说明以及权利要求书明显可见本发明的其它特征和优点。除非另做定义,本文中使用的所有科技术语均具有与本发明所属领域技术人员一般理解相同的意义。尽管可使用与本文中揭示的那些相似或等效的方法和材料来实践和测试本发明,但适当的方法和材料揭示如下。本文中所引用的所有已公开的外国专利和专利申请通过引用并入本文。本文中引用的以登录号表示的基因银行和NCBI文件通过引用并入本文。本文中引用的所有其它已出版的参考文献、文档、手稿和科学文献通过引用并入本文。若有冲突,以包括定义在内的本说明书为准。此外,该材料、方法、和实施例仅用于例示性说明而非意图限制。

[0078] 本发明提供可用于实施固定组织标本中多重基因表达分析的组合物和方法,该标本为例如,举例而言,福尔马林固定石蜡包埋(FFPE)组织标本。特别地,本发明提供一种RNA分析的方法,该方法包括获得一样品;将一种或多种多歧探针施加至该样品,其中该一个或多个多歧探针包括至少两个子探针;将所施加的一个或多个多歧探针的至少一个退火为该样品内的至少一个靶标核酸;以及将结合该至少一个退火的多歧探针的至少两个子探针连接,以创造可被检测的靶标核酸替代物。例如,本文的技术提供,原位连接杂交(LISH)可利用新颖的基于Rn12的连接化学的发展,其能高效地将交联的RNA序列直接转化为基因特异性的、PCR可扩增的DNA序列板,而不依赖于RNA提取或逆转录。另一方面,本文中的技术提供,将LISH与下一代DNA测序(NGS)分析(“LISH-seq”)组合,可能令以空前的敏感性和精确性对FFPE组织标本中的mRNA板进行定量测量成为可能,潜在地将笃定的大量患者标本档案解锁用于新分析。LISH-seq可克服先前的与显微切割组织的多重基因表达分析相关的限制以及对固定的细胞悬浮体的限定。将LISH与独特的DNA印记和桥式PCR扩增方案组合(“LISH-印记”),可以令以单细胞解析度对FFPE组织学切片或固定的细胞学样品进行空间解析度的多重基因表达分析成为可能。本公开至少部分地基于下述发现:多歧探针(如,包括两个或更多个子探针的探针)可退火为FFPE样品中的具有高度特异性的核酸靶标(如,RNA),随后连接(如,通过T4RNA连接酶2连接),以创建用于扩增的模板,而无需RNA提取或逆

转录。本文的技术将固定的组织内的RNA的共价交联用作待开发的有益特征。重要的是,本文的技术(如,LISH)对于降解和加合物形成是稳健的,且完全依赖于可轻易获得的试剂和设备。

#### [0079] 精准肿瘤学

[0080] 精准肿瘤学取决于研究人员和病理学家在分子层面表征化切除的肿瘤组织和活组织检查材料的能力。用于DNA分析的先进技术如高通量DNA测序或下一代DNA测序(NGS)的发展,已经对检测与癌症相关的基因损伤产生了巨大贡献。但是,表征化患者的固定标本中基因表达谱的能力滞后,这主要是由于这些样品的RNA品质差。

#### [0081] 临床病原体诊断

[0082] 可快速鉴别人类病原体的分子检验的发展仍是诊断行业的重要目标。组织样品通常从被怀疑罹患尚未被鉴别的真菌感染、细菌感染、或病毒感染的个体。然而,在常规固定过程中导致的这些样品的生物分子降解和加工的程度使得它们不适合用于当前可用性最高的分子检验中。能耐受有害固定过程同时还维持敏感性且是高度多重性的替代性检验的发展,将会极大地改善与病原性感染的诊断和及时处理相关的结局。

[0083] 临床组织防腐最常采用福尔马林固定石蜡包埋(FFPE)过程,且FFPE样品是病理学家检查的最常见患者标本。它们的相对稳定性提供了用于染色的理想底物,并令患者组织的低成本归档成为可能。但不幸的是,这一用于令组织架构防腐的方法降解并共价地修饰组织RNA(如,令其成为高度降解的形式),妨碍使用基于RNA纯化及逆转录的传统途径对其进行分析。对来自培养细胞和未固定组织的RNA的测量典型始于RNA纯化,之后为逆转录(RT)和聚合酶链反应(PCR)。但是,当对来自化学固定标本的RNA实施RT-PCR时,RT-PCR是不充分且不可靠的(Kashofer, K., et al. *PLoS One*. 8(7) (2013); Abrahamsen, HN., et al. *J Mol Diagn*. 5(1):34-41. (2013); Stephen. A., et al. *JBiomol Tech*. 15(3):155-166. (2004))。由于FFPE对组织架构、生物分子稳定性、剂与包括免疫染色剂的多种染色剂的相容性,FFPE是最广泛应用的保存临床组织标本的方法。福尔马林固化经由加成加合物如羟甲基、以及经由形成亚甲基桥而将RNA交联至其自身及其它分子,从而损害RNA。此外,RNA片段化典型出现在组织保存之前或保存过程中,且随后在存储过程中继续以低速率出现(Evers, D.L., et al. *J Mol Diagn*. 13(3):282-8. (2011); Masuda, N., *Nucleic acids Res*. 27(22):4436-43. (1999))。由于RT-PCR需要最小修饰RNA的相对长的伸张,传统技术苦于降低的敏感性和不可预知的测量偏差。

[0084] 尽管FFPE标本如此重要,但主要由于与这些样品的RNA提取和逆转录相关的挑战,极少有技术能实现对患者的固定标本的稳健且高度多重的mRNA表达分析。现有途径或采用无需靶标扩增的探针的杂交(如,NanoString的Prosigna检验),或采用靶向逆转录(RT)和PCR产物分析(如,Genomic Health的Oncotype DX qPCR检验; Assuragen的Signature LTx 12重融合转录流式细胞小球微阵列检验)。鉴于该无扩增的途径受限于其对于昂贵且私有的仪器和试剂的需要,该RT依赖技术受限于RNA的低品质(如,由于片段化、交联、共价加成物的形成等造成的低品质),因此降低敏感性且压抑检验计划。使用NGS的无偏差的转录组学分析是新兴的替代方法,但亦受限于对RT的需要、无偏差测序的高成本、与计算转录定量相关的挑战、和数据解释的难度。此外,这些检验各自取决于RNA模板的提取,这与额外的降解、材料的损失、及空间解析度的损失相关。

**[0085]** 分子病理学

**[0086]** 分子病理学聚焦于通过检查请、组织、或体液中的分子对疾病进行研究和诊断。一些实践方面中,分子病理学具有与解剖病理学和临床病理学两者、分子生物学、生物化学、蛋白质组学、及遗传性有共通之处。关键原因是,当基于组织的形态学变化和该组织细胞中细胞变化两者而诊断时,更正确的诊断时可能的。分子病理学常用于癌症和感染疾病的诊断。例示性技术包括定量聚合酶链反应(qPCR)、多重PCR、DNA微阵列、原位杂交、DNA测序、基于抗体的免疫荧光组织检验、病原体的分子性能分析、及用于耐药抗性的细菌基因分析。

**[0087]** 分子病理学的出现已经引出了精准医疗的新领域。应用于基因型及/或表型肿瘤组织的技术,举例而言,现在常规用以提供准确诊断和优化的治疗性干预。多重基因表达测量具有提供丰富的表型信息的潜力,但由于与这些标本中RNA相关的限制,迄今为止,该测量仍未在固定的组织标本上广泛实施。

**[0088]** 原位连接杂交(LISH)

**[0089]** 本文的技术提供了用于基因表达的多重测量的强有力的平台,该平台理想地适用于固定组织标本的分析,且该平台克服了与FFPE RNA分析相关的关键技术限制。一种具体例中,LISH使用T4 RNA连接酶(Rn12)以有效地将序列特异性嵌合RNA-DNA探针组原位接合在固定的RNA靶标序列上。使用RNase H对该组织进行后续的处理,释放RNA模板化的连接产物,同时同步地摧毁背景DNA模板化的连接产物。如图5A中可见,步骤I,探针对于退火为mRNA靶标或DNA靶标并无偏好,而该RNA靶标与DNA靶标具有互补序列。因此,在RNase H处理(步骤III)过程中,存在两个截然不同的池,连接的探针:RNA模板和连接的探针:DNA模板。RNase H特异性地消解RNA:DNA双链的RNA。因此,RNase H将识别并消解连接的探针:RNA模板双链中的RNA模板,并反向消解和摧毁连接的探针:DNA模板双链中连接的探针的RNA成分(内有二核糖)。

**[0090]** 一方面,本文的技术提供可施加至固定的组织标本(如,FFPE)的多歧核酸探针。该多歧探针可各自包括两个或更多个核酸探针(如,子探针),该核酸探针配置为退火为靶标核酸(如,RNA)中的接续靶标序列,因此,该两个或更多个探针各自结合该接续靶标序列,而不在该两个或更多个探针中任何相邻探针之间的接续靶标序列中留下任何未结合的核苷酸。例如,如果多歧探针包括两个探针(如,双歧探针),则一旦该两个探针结合至接续靶标序列,一个该探针的3'端将紧邻另一探针的5'端,故该两个探针以端至端的配置方式结合在靶标序列中。

**[0091]** 一旦该多歧探针结合至靶标序列,连接酶(例如,T4 RNA连接酶2(Rn12))可用以将相邻的被结合多歧探针连接为可扩增的cDNA副本,而无需RNA提取或依赖于逆转录。本文的技术(如,LISH)可遵从高度多重的mRNA分析方法。例如,可使用10至10,000个多歧探针同步分析典型的固化标本,这预期处于本发明范畴内。本文的技术(如,LISH)拓展了先前的RNA模板化DNA探针连接化学(见,如,Larman,H.B.,etal.Nucleic acids research 42,9146-9157(2014))的工作,以将交联的RNA分子转化为PCR可扩增的DNA序列。较佳地,本文中揭示的技术对于降解(如,高度降解的样品)和加合物形成是稳健的,且完全依赖于可轻易获得的试剂和设备。

**[0092]** 探针设计和校验

**[0093]** 可使用与引物-BLAST相似的探针组设计途径设计多歧探针,且使用引物3



(BLASTN, Melting, pandas和Python标准库)执行该设计。将探针设计为退火为尽可能接近靶标mRNA多聚(A)尾部。执行用户引物3设定,以设计相对于该靶标转录本为反义的高达20个独立的探针,每一探针具有36个核苷酸的。将引物3设计的探针在该探针的5'方向(朝向该多聚(A)尾部)延长4个碱基对。随后,将每一个40个核苷酸的序列对半分裂,并附加常用的适配因子(用于接纳体探针的AD1和用于供体探针的RCAD2)。引物3随后被用来计算每一探针寡核苷酸加上适配因子的性质。用于引物3计算的阈值被用来过滤备选探针。移除其脱靶 $T_m$ 与预定上靶解链温度相差 $10^{\circ}\text{C}$ 内的剩余探针。随后,采用至多聚(A)尾部和引物3惩罚性代价(基于原始的36个核苷酸引物3设计探针)的距离的非参数性排序方案,来选择两组预测最佳的探针组,对于每一靶标转录本,该探针组退火的彼此距离为至少10个核苷酸。接纳体寡核苷酸3'端碱基和3'端第二个碱基被变为其RNA副本。

[0094] 探针设计的关键参数包括探针长度和该序列的解链温度。对于给定的靶标序列,增加探针长度则增加特异性结合作用的强度,可亦可增加通过非特异性结合至脱靶序列的不适宜的连接和/或降低该反应中探针的有效浓度。

[0095] 创建供体和接纳体探针库,以探索从12至22个核苷酸范围内的退火序列长度的影响。为了消除潜在的混杂变量如核苷酸序列偏差,将接合位点保持恒定。对于供体探针和接纳体探针两者,上靶连接良率取决于该探针的长度。因为探针的成本随着长度而增加,又因为在大多数探针的约20个核苷酸后的相对定量中观察到逐渐减少的改善效果,研发一种探针设计算法,以鉴别靶标转录本中相邻的20核苷酸的序列。然而,多歧探针的探针长度可为10至200个核苷酸的范围内也预期处于本公开的范畴内。

[0096] 探针诱饵策略被研发出来,以降低在非常高水平表达的转录本的采样,从而优化序列分析的效率。每一探针组的侧翼存在常用的引物结合序列,因此对于梅一孔,连接产物可通过含有特异性短DNA条码的引物独立进行扩增。诱饵探针缺乏该引物结合序列,因此它们以所希望的水平形成不可扩增的连接产物。因此,来自多个样品的PCR产物可合并在一起进行测序,且个体读数随后被它们响应的条码简化。有限数的DNA测序读数获得高丰度转录本过度采样的结果。该探针诱饵策略克服了这一取样偏差。

[0097] 多歧探针

[0098] 多歧探针呈现一种将两个或更多个核酸序列带至一起并本体化的途径。多歧探针可包括能与靶标基因序列杂交的靶标核酸结合序列,以及与该靶标核酸结合序列相邻的另一靶标核酸探针结合序列。例如,多歧探针可各自包括配置为退火为靶标核酸(如, RNA)中接续靶标序列的两个或更多个核酸探针,因此,该两个或更多个探针各自结合该接续靶标序列,而不在该接续靶标序列中的该两个或更多个探针的相邻探针(如,子探针)之间留下未结合的核苷酸。多歧探针可被杂交至包括该靶标基因序列的基因样品并与之连接,以包括可检测的扩增分子。该可检测的扩增分子可包括捕获探针序列,该序列是能与该用于双歧探针的核酸结合探针序列杂交的核酸序列。多歧探针中的每一探针的长度可以是10至200个碱基。例如,10至200理解为包括任何数字、数字的组合或数字的子范围,以及从该范围任一端点延伸的“嵌套子范围”。例如,10至200的范围的嵌套子范围可包含10至20、10至30、10至40、10至50、10至60、10至70、10至80、10至90、10至100、10至110、10至120、10至130、10至140、10至150、10至160、10至170、10至180、10至190、及10至200。

[0099] 多重技术

[0100] 多重基因表达分析提供使用检测系统对多个核酸序列同步进行的直接和定量测量。多重检验使用下述策略：从单一样品等量小样扩增超过一种靶标并定量。在多重PCR中，使用含有荧光染料的多探针在单一PCR反应中探究样品。这增加了可从该样品提取的信息量。使用本文中揭示的LSIH技术，可在相同时间分析数百至数千个探针。

[0101] 连接酶

[0102] Rn12(dsRNA连接酶)是ATP依赖性dsRNA连接酶，该酶有效地密封双链RNA中的3'-OH/5' PO<sub>4</sub>缺口。这一过程经由该连接酶的腺苷酰化作用而出现，AMP转移至供体链的5' PO<sub>4</sub>，接纳体链3'-OH附着在5'-腺苷酰化的供体链上，导致共价磷酸二酯链接的形成。Rn12允许以脱氧核糖核苷酸完全替代其双链RNA底物，限制条件为该接纳体链的3'端以二核糖核苷酸结束。该技术采用具有5'-PO<sub>4</sub>末端的全脱氧核糖核苷酸供体探针，该探针进行高效的模板依赖连接成为杂交的脱氧核糖核苷酸-3'-二核糖核苷酸接纳体探针。总之，Rn12可连接RNA探针或RNA-DNA杂交探针，其中，当在RNA或DNA上退火时，一个探针具有3'两个碱基。Rn12不能有效地将在DNA模板上退火全DNA探针或在RNA模板上退火全DNA探针。

[0103] 可引导连接酶加入邻近退火的接纳体和供体探针组。酶促连接共价地加入该探针，该探针随后用作基于PCR的信号扩增模板。在典型条件下，该连接反应的所有成分均相对于该靶标mRNA过量，确保模板分子与连接事件之间成正比。随后，可在PCR过程中扩增该连接产物并以条码表示。

[0104] 可根据本文的技术使用其它类型的连接酶，包括，但不限于，T4 DNA连接酶和PBCV-1DNA连接酶(小球藻病毒DNA连接酶)，这在本公开的范畴预期之内。

[0105] RNase H(核糖核酸酶H)

[0106] 一旦探针被连接，则所连接的多歧探针(如，子探针)可从RNA模板释放出来，如此，该探针可被回收并用作扩增源。根据本文的技术，RNase H可用来释放所连接的多歧探针(如，子探针)。RNase H属于非序列特异性核酸内切酶家族，其催化RNA的经由水解机制的裂解。RNase H家族的成员可在从细菌到古生菌到真核生物的几乎所有有机体内发现。因为RNase H仅特异性降解RNA:DNA杂交体中的RNA，其通常用于分子生物学中以及在通过逆转录合成第一链cDNA后以及在诸如核酸酶保护检验的过程中摧毁该RNA模板。当该cDNA寡核苷酸被杂交时，RNase H亦用来降解特定的RNA链，如从杂交至寡核苷酸(dT)的mRNA移除多腺苷碱(polyadenine)尾部，或摧毁活体细胞内外的所选择的非编码RNA。RNase H特异性水解杂交至DNA的RNA的磷酸二酯键。这种酶并不消解单链或双链DNA。为了终止该反应，一般加入螯合剂如EDTA以隔绝该反应混合物中的金属离子，或可热摧毁该酶。

[0107] 在连接邻近退火的探针之后，可通过以RNase H孵化该样品而回收该连接产物(LISH产物)，其中，RNase H摧毁RNA/DNA杂交螺旋的RNA成分且将连接的子探针释放入溶液中。随后，可扩增该产物，并最终进行序列分析。用来回收所连接的探针组的备选方法包括，但不限于，其它RNase酶(例如，RNase A、RNase If、RNase HII)、在PCR扩增前进行热处理以令该杂交螺旋解链或进行机械组织提取(如，通过激光捕获显微切割或刮擦)。

[0108] 本文的技术具有极大拓展分析固定组织标本的研究人员和病理学家可用的分子工具箱的潜力。当前用于FFPE组织的高度多重基因表达分析的途径需要提取受损的、相对低品质的RNA。相比之下，基于LISH的分析使用交联的RNA分子，以获得PCR可扩增的序列，并因此不需要RNA提取或逆转录，该分析提供数个与检验灵敏度和流程简洁性相关的明显优

点。重要的是,该系统使用可广泛获得的设备和试剂,令其可即时执行。随着高通量DNA定序成本的下降和实用性的增长,本申请提出的LISH-seq检验的接受度将日益提高。

[0109] 对于其中基因表达分析的空间解析度很重要的应用,在敏感性和可同步分析的转录本数目两方面,LISH-印记具有显著胜过供选择的检验方法的潜力。通过显著地拓展测量固定组织中基因表达的能力,LISH平台将极大地改善表征化肿瘤表型和理解与免疫系统细胞的复杂的相互作用的能力。最后,因为LISH平台理想地适合固定的组织标本,该平台将解锁存到的巨量患者样品用于以其它方式行不通的新的多维RNA分析。

[0110] [实施例]

[0111] 实施例1:原位连接杂交测序(LISH-seq)的研发

[0112] 确定了有效地在FFPE切片中产生探针连接产物的最佳方案。采用看家基因探针组来探究各种方案差十年后对于该检验的信号(上靶探针连接)和干扰(脱靶探针连接)的影响。通过与RT-qPCR的直接比较来评估LISH的效能。混合的LISH产物(“LISH-seq”)的高通量DNA测序表明了该系统提供的多重化程度(如,每一组织切片中有数百至数千个转录本)。通过LISH-seq分析一批长期存储的标本。这一数据集显示了LISH-seq解锁存档的巨量患者肿瘤标本用于高度多重mRNA分析的程度。存在一种策略,其使用LISH-seq来克服先前与分析从显微切割肿瘤和小群体的固定悬浮体中提取的RNA相关的限制。

[0113] RNA退火、选择和连接测序(RASL-seq)检验

[0114] 2012年,UCSD的Fu小组介绍了“RNA退火、选择和连接测序”(RASL-seq),并阐明其以更低成本和比当前使用供选择的途径所得者更高的通量来监控数千个样品中几百个转录本的能力(见,如,Li,H.,et al.Proc Natl Acad Sci U S A 109,4609-4614(2012)和Zhou,Z.,et al.Molecular cell 47,422-433(2012))。

[0115] 通过使用基因特异性探针组的DNA池来实施这一技术壮举,其中,该探针组的配对被邻近地退火在靶标转录本上。进行转录本模板号的酶促探针组连接后,实施连接产物的PCR扩增。典型地,使用样品特异性DNA序列来条码化该PCR引物,令来自大量样品的PCR产物可进行后续的池化,用于同步的NGS分析,从而显著降低样品的预处理成本。RASL检验的一个重要特征是,通过RNA的短序列直接令探针连接模板化,因此避免对于逆转录(RT)的需求。

[0116] 先前公布的RASL-seq方案依赖于T4 DNA连接酶以实施临床连接步骤。但是,已经观察到,DNA链的这一RNA模板化连接导致非常低的产率(见,如,Bullard,D.R.&Bowater,R.P.The Biochemical journal 398,135-144(2006)中的表1和Larman,H.B.,et al.Nucleic acid s research 42,9146-9157(2014)中的图1A),并因此导致检验的灵敏度极差。

[0117] 为了解决这一问题,如本文中所述,研发了一种使用T4 RNA连接酶2(Rn12)的连接化学,T4 RNA连接酶2(Rn12)是自然修复双链RNA中断裂的酶。重要的是,Rn12容许以切口的dsDNA替代其被切口的dsRNA底物,条件是接纳体链的3'端具有两个RNA碱基(图1A)(见,如,Nandakumar,J.&Shuman,S.Molecular cell 16,211-221(2004))。如下文所详述,Rn12以定量的产率将嵌合DNA-RNA探针组连接在RNA模板上,该产率平均比T4 DNA连接酶高400倍(图1B)。LISH亦可与额外的步骤组合,以映射固定标本的RNA表达水平。这将允许将涉及靶标分布的空间信息与靶标表达水平组合,而这是使用RASL或基于RT的分子检验不可能实现的特

征。尽管含有该二核糖核苷酸元素,但连接产物是通过所测试的多种DNA聚合酶得以PCR扩增的。重要的是,RASL检验能跨越大于6个数量级的线性动态范围而测量所添加的示踪转录本(图1C)。

[0118] LISH的优化和表征化

[0119] 将源自人细胞系的肿瘤异种移植物加工为4至40厚的FFPE连续切片,并将其用于通过原位连接杂交(LISH,图2A)手段进行单一看家基因转录本的分析。当前,使用JHH(Johns Hopkins Hospital)临床病理学实验室所采用的用于杂交RNAscope DNA探针以原位检测HPV RNA的相同方案(见,如,Wang,F.,et al.The Journal of molecular diagnostics:JMD 14,22-29(2012))。RNAscope探针组与LISH探针组高度相似,支持这一途径。Rn12介导的连接后跟随修饰抗体染色方案。提取连接产物后,使用SYBR绿色qPCR进行连接产物的定量。使用手头的基本功能性方案,所评估的效果包括:

[0120] 切片厚度、

[0121] RNA二级结构的预杂交解链、

[0122] LISH探针组浓度、

[0123] 探针组杂交的温度和持续时间、洗涤的严格性、Rn12连接酶的浓度、

[0124] 连接步骤的温度和持续时间、以及

[0125] 回收已连接的LISH探针组的方法。

[0126] 通过对正确配对的探针组的SYBR绿色qPCR分析来评估检验“信号”,而通过对不正确配对的探针组的分析来评估检验“干扰”。随着该方案参数如上述改变,一系列看家基因探针组的检验信号干扰比得以最大化。研发与传统下游分析如免疫组化相容的LISH方案。

[0127] 可使用RNase H有效地回收LISH连接产物(见,如,图5A中的步骤3),该酶摧毁RNA/DNA杂交螺旋中的RNA成分,以将连接的探针组释放入溶液中。通过机械刮削或使用疏水性障碍笔分离载片上感兴趣的具体区域。使用RNase H溶液进行剪短孵化后,直接将PCR预混液加入洗脱液中,并启动PCR热循环。对于使用NGS进行的LISH探针组分析,使用DNA条码化引物组进行PCR扩增,以令样品多重化。

[0128] 建立LISH的灵敏度

[0129] 实施下述两个实验:I.使用LISH-seq解锁存档的固定组织标本;以及II.实施对显微切割肿瘤的LISH-seq分析,以建立LISH的灵敏度。将存档的来自固定脾组织的连续切片用于LISH检验中,或将相同大小的连续切片用于总RNA的提取。加工单组织切片,以提取RNA作为参考;加工第二连续切片用于LISH检验中,以比较两种方法的灵敏度。源自FFPE切片的RNA进行集中看家基因转录本的标准单基因RT-qPCR分析和单基因RASL检验分析(图2B)。当用于RT-qPCR或RASL反应中时,使用从冷冻未固定人类脾组织获得的RNA来测定可能来自等量的起始mRNA的最大信号。在该LISH检验中使用相同的RASL探针组,提供了对固定组织中RNA分子的分数定量评估,其中,该固定组织被使用LISH方案转化为可扩增的信号。此外,对固定RNA的RT-qPCR分析与LISH分析间的直接比对可证明基于LISH的检验的接触的敏感性。

[0130] I.使用LISH-seq来解锁固定的组织标本档案

[0131] 因为LISH连接产物通过使用单一引物对的PCR得以扩增,其轻易地与NGS测序兼容(LISH-seq)。此外,即使固定RNA模板的总量由于降解而随着时间而损失,但相对基因表达

谱得以良好地保存。因此,将LISH-seq分析的效能评估为FFPE切片在室温下存储时间的函数。

[0132] 对于第一系列实验,将先前以使用RT-qPCR的基因组表征化的FFPE人肿瘤切片用在LISH实验中,以比较不同分析之间的基因表达谱。对看家基因mRNA和先前表征化的基因组的分析将帮助建立用于解释来自历史组织标本的LISH-seq数据的重要框架,显示的程度为LISH-seq解锁存档的固定组织标本用于新的、高度多重的转录本分析。

[0133] II. 显微切割组织和固定的悬浮体细胞的LISH-seq分析。

[0134] 激光捕获显微切割(LCM)固定组织的基因表达分析是有价值的,但由于可提取的RNA的量极小且品质低,该分析仍特别具有挑战性(见,如,Pagedar,N.A.,et al.Brain research 1091,289-299(2006)及Kaneko,T.et al.Microscopy research and technique 72,908-912(2009))。在LCM之前对完整的FFPE切片上实施LISH,克服了这一限制。发展了直接从LCM材料PCR扩增连接产物的方案,并使用上述的相同FFPE切片评估该方案。

[0135] 与LCM类似,它对在FACS分类之后的悬浮细胞小群体(如,MHC-肽四聚物标记的肿瘤浸润性CD4+/IFN $\gamma$ +T细胞)内测量成组转录本极具价值。细胞固定化典型用于细胞内细胞因子和转录因子的染色,从所分类的细胞中提取高品质的RNA非常困难。发展了用于固定且渗透化的悬浮细胞的LISH方案,该方案亦与下游分析如细胞内细胞因子染色兼容。随后,对所分类的细胞(或可能单一的细胞)直接进行PCR反应(无需核酸纯化),用于高维LISH-seq分选。

[0136] 替代策略

[0137] 用于在FFPE组织切片中退火LISH探针组的方案使用优化的条件(温度、时间)和缓冲剂,以确保最大的严密性和敏感性。如果没有使用Rn12实现被退火的LISH探针的有效原位连接,则评估近年来作为“SplintR”连接酶而商业化的小球藻病毒DNA连接酶,也称为PBCV-1DNA连接酶的活性(见,如,Lohman,G.J.,et al.Nucleic acids Res42,1831-1844(2014))。过去已经直接比较了PBCV-1连接酶与Rn12,且发现Rn12的执行力更佳(见,如,Larman,H.B.,et al.Nucleic acids research 42,9146-9157(2014)),但预计PBCV-1连接酶在LISH应用中均足够良好地发挥功能。

[0138] 此外,非特异性探针连接是不希望的反应,且可通过包括与用于非特异性链结的功能探针竞争的“诱饵”探针而得以强烈抑制。诱饵探针被设计为以不相干的序列为靶标,且它们缺乏该功能LISH探针上存在的PCR引物结合适配因子。当将过量的诱饵探针加至该反应时,可显著降低所扩增的连接产物的脱靶分数(见,如,Larman,H.B.,etal.Nucleic acids research 42,9146-9157(2014))。

[0139] 实施例2:维持所扩增的LISH产物的空间分辨率:LISH-印记

[0140] 如果该转录水平的空间分布得以保存,则极大地提升多重基因表达分析的值。该技术尤其可用于表征化肿瘤组织与免疫系统细胞之间的复杂的相互作用。能对FFPE组织切片中的基因表达进行高度多重的、空间解析分析的平台,对于癌症研究人员和病理学家是可变形的。

[0141] LISH的使用性被拓展至用以将LISH产物印记在靶标表面上并扩增以保存空间信息的方法(“LISH-印记”)。这一尝试建立在将DNA分子印记在副本高密度DNA微阵列上的基础上。为了局部地扩增所印记的LISH产物,使用桥式PCR,而桥式PCR已经被揭示在文献中,

并且是Illumina高通量DNA测序技术的基础。在定义LISH印记方案后,研发了用以分析桥式PCR扩增产物的技术。这些研究使用源自肿瘤异种移植物切片的人细胞系,并从固定悬浮细胞系扩散。通过将桥式PCR产物与多色荧光寡核苷酸杂交,同步测量5至10种基因的表达水平。使用杂交、成像、和寡核苷酸移除的相继循环来分析来自单一FFPE载片的转录本的空前的数目。

[0142] 对于理解癌症病理学,尤其是在复杂的肿瘤-免疫细胞相互作用情况下,用来表征组织微环境空间架构的工具极具价值。

[0143] 超分子纳米印记(SuNS):

[0144] 超分子纳米印记(SuNS)是Francesco Stellacci等人研发的方法,该方法将来自一个表面的图案化的化学信息(如,DNA)可逆地转移至另一表面(图3A)(见,如,Yu,A.A.,et al.Nano letters 5,1061-1064(2005);Akbulut,O.,et al.Nano letters 7,3493-3498(2007);和Kim,J.&Crooks,R.M.Analytical chemistry 79,7267-7274(2007))。SuNS方法被拓展至酶促“复制”显微镜载片大小的DNA微阵列(Cuppoletti,A.&Larman,H.B.(见,如2009年的专利PCT/US2008/076717))。该实验采用下述策略:

[0145] (i) 在ssDNA模板微阵列上进行生物素标记的cDNA链的酶促合成;

[0146] (ii) 令该dsDNA微阵列与涂覆有包含链霉亲和素的凝胶膜的捕获表面接触;

[0147] (iii) 将生物素化的cDNA热释放至该凝胶中,在该处,cDNA结合链霉亲和素;以及

[0148] (iv) 分离载片,并制备用于下一个复制循环的模板。

[0149] 初步研究表明了对该副本表面解析度及其功能性两者的忠实保持(图3B)。

[0150] 本文中,调整DNA印记以令其适用于FFPE切片的空间解析基因组表达分析。

[0151] 实验设计

[0152] 初步研究采用将来自模板微阵列的DNA印记在涂覆有凝胶膜的副本表面上的经验(见,如,Cuppoletti,A.&Larman,H.B.2009年的专利PCT/US2008/076717)和Larman,H.B.&Stellacci,F.(见,如,2010年的专利PCT/US2008/076723))。但在本文的技术中,将来自“凝胶”(如,FFPE组织切片)的DNA(LISH连接产物)印记在涂覆有共价固定化PCR引物的副本表面上。

[0153] “桥式”PCR

[0154] 桥式PCR是局部化的固相PCR扩增方法,该方法中,DNA片段通过附着在表面上(而非在溶液中)的引物扩增,从而形成表面固定化的DNA簇(见,如,Adams,C.P.&Kron,S.J.(1997年的专利US 08/285,385,))。如果这些簇足够分离,则它们可克隆测序(见,如,Mitra,R.D.,et al.Analytical biochemistry 320,55-65(2003)和Mitra,R.D.,et al.Proc Natl Acad Sci U S A 100,5926-5931(2003)),这是Illumina高通量DNA测序技术的基础(见,如,Shendure,J.,et al.Science309,1728-1732(2005)和Porreca,G.J.,et al.《分子生物学实验指南》(Current protocols in molecular biology),Frederick M.Ausubel等人编着,第7章第7、8单元(2006))。但如果这些簇完全或部分地重叠,则通过使用荧光标记的cDNA序列的杂交测定每一序列的相对贡献,这是LISH-印记技术的基础(图4)。

[0155] 第一系列实验建立了桥式PCR靶标表面。使用涂覆有环氧硅烷的玻璃显微镜载片(Schott),这种载片对于5'-氨基封端的寡核苷酸引物(IDT)的共价固定化是理想的载片。

为了评估桥式PCR靶标表面的品质(表面密度和均匀度),使用荧光引物结合的DNA探针。

[0156] 使用靶标表面,该用于LISH产物的桥式PCR方案得以从已出版的方法优化。预先连接的LISH探针的稀溶液被直接手动置于载片表面上斑点化,以模拟印记。在桥式PCR后,使用互补杂交探针分析载片,以测量该扩增。

[0157] 在使用被斑点化的连接产物的桥式PCR成功后,启动使用固定组织标本作为LISH探针连接产物来源的实验。首先构建实现组织载片与靶标载片之间的可靠的、大面积保形接触的最简单的方法。研发了用于印记DNA的温控按压系统(见,如,Larman,H.B.& Stellacci,F.(2010年的专利PCT/US2008/076723)),并探索其对于LISH-印记的使用性。亦可实现与使用简单食品真空封口机相似的结果。一旦接触,加热载片以允许该LISH连接产物从FFPE切片的热释放,随后冷却至45°C,保持足够时间,以允许被靶标表面上的桥式PCR引物捕获。分离该载片,并对靶标载片实施桥式PCR扩增。实施产率对热循环数的分析。使用源自肿瘤异种移植物的人细胞系的FFPE切片来发展该方法。

[0158] LISH-印记研究

[0159] LISH-印记实验采用多个5至10个探针组,该探针组使用以荧光染料标记的cDNA寡核苷酸同步分析,其中,该荧光染料可在约翰霍普金斯医院保有的Zeiss Axiovert 200倒置显微镜(具有LSM510-Meta共焦模块,7条激光线,和4组滤波器)上分辨。研究主要集中在两种类型的组织上:(i)人细胞系的细胞离心涂片,和(ii)源自肿瘤异种移植物的人细胞系。细胞离心涂片制备允许LISH-印记的空间分辨率分析。重要的是,通过手工传递分离独立的桥式PCR载片用作多重化LISH-印记分析的荧光补偿控制,个体探针连接产物被斑点化。采用标准DNA微阵列“剥离”方案以令探针组相继杂交成为可能,从而可以令待分析的LISH探针数比另外同步成像的探针数大得多。

[0160] 替代策略

[0161] 在LISH-印记载片上进行的差异极大的基因表达水平的分析可能有问题。在某种程度上,这是由于桥式PCR引物在来自高表达基因的LISH产物与来自低表达基因的LISH产物之间的竞争而造成的。通过使用基因特异性诱饵RASL探针,先前在RASL-seq的发展中进行了类似的挑战(见,如,Larman,H.B.,et al.Nucleic acids research 42,9146-9157(2014))。因此,部分诱饵看家基因探针组被预期增加桥式PCR引物对于较低丰度转录本的LISH产物的应用性。在极度不平衡的荧光信号中,通过将标记的杂交探针稀释于未标记的杂交探针背景中,可衰减来自高丰度转录本的信号。

[0162] 实施例3:FFPE中RNA序列的测量

[0163] 近年来的报道已经表明,RNA模板化探针连接化学对于大规模多重分析是敏感的、特异性的、和适当的(Larman,H.B.,et al.Nucleic acids Research 42,9146-9157(2014),通过引用而以其整体并入本文)。该途径采用T4 RNA连接酶2(Rn12),该酶将邻近退火的序列特异性寡核苷酸探针组连接在RNA“夹板”(splint)上。在这一背景下,Rn12需要该接纳体寡核苷酸是以两个核糖核苷酸进行3'-端的(本文中称为“3'-二核糖探针”)。一些实例中,该接纳体寡核苷酸可具有一至少两个RNA碱基的3'封端。磷酸化的供体探针(本文中称为“5'-磷酸探针”)可以完全是DNA(图5A)。因此,连接的3'-二核糖探针组和5'-磷酸探针组是含有内部二核糖核苷酸序列的DNA分子。

[0164] 尚无已知的DNA连接酶展现对RNA夹板相对于DNA夹板的优先性,甚至Rn12以大致

相等的效率将嵌合探针组连接在RNA夹板和DNA夹板上。因此,对RNA序列的精确定量可被来自DNA夹板化连接的背景信号混淆,除非将DNA夹板化连接的产物特异性地消除。RNase H是特异性消解RNA-DNA杂交螺旋中RNA成分的核酸酶。因此,通过同步地(i)摧毁DNA夹板化的含二核糖核苷酸的连接产物(图5A(i))和(ii)将RNA夹板化的含二核糖核苷酸的连接产物从固定RNA释放入溶液中进行分析(图5A(ii)),RNase H用作双重目的。事实上,DNA退火的含二核糖核苷酸的DNA序列被RNase H2裂解,而RNA退火的序列不裂解(图5B)。RNase H1需要RNA的至少4个接续碱基用于消解,所以如预期,这一酶并不摧毁DNA退火的含二核糖核苷酸的DNA序列。因此,RNase H2和RNase H1可用来回收连接的探针组,分别造成或不造成DNA夹板化连接产物的同步摧毁。这一测量固定组织中RNA序列的新系统称为“原位连接杂交”(“LISH”)。

[0165] FFPE RNA用作Rn12介导的探针原位连接的模板

[0166] 使用分别以GAPDH和RPL19看家基因为靶标的两个探针组来检验人类脾脏的10微米切片。对RNase H回收连接产物的端点PCR分析显示,正确的探针对确实被原位连接,而误配的探针组未以可检测的水平连接(图5C和5D)。

[0167] 为了量化自多重反应过程中形成的特异性连接的量,对四种可能的连接产物(两个“上靶”和两个“脱靶”)各自实施定量PCR检验。随后,使用这些检验来测定信号(上靶探针连接)干扰(脱靶探针连接),同时优化LISH方案参数。

[0168] FFPE标本的临床归档评估

[0169] 全世界存在FFPE标本的大量临床档案;因此,能恢复来自存档组织的转录信息的技术将会极其有用,尤其是用来研究限制性或高价值的标本。而预计所恢复的连接产物的量级将随着时间而谨慎下滑,应使用LISH分析忠实地保持相对的基因表达谱。

[0170] 在同一组织切片上连续实施的独立LISH检验

[0171] 使用RNase H回收LISH连接产物应分出并不用作探针连接模板的RNA分子。结果,先前经LISH分析的切片的非靶标RNA序列应保留在第二次的非重叠性LISH分析中的可用性。这是LISH用于限制性或高价值标本的研究的有用特征。可对相同的组织切片依次实施独立的LISH分析,因此,来自一次LISH检验的结果,可用来指导对将在同一FFPE组织切片的第二次LISH检验中使用的信息探针的选择。

[0172] LISH检验与多种分析染色剂的相容性

[0173] 使用丰富多样性的分析染色剂来探究FFPE标本。这些染色剂中的多数,包括苏木紫和曙红(H&E)以及免疫染色剂,将可用来与LISH分析协同作用。连接的LISH探针相对稳定,并因此被预期在后续染色方案进行的过程中保持原位完整。染色之后,可使用RNase H、激光捕获显微切割(LCM)、或其它手段恢复LISH产物。因此,LISH分析可与使用多种常用染色剂对相同组织切片的分析组合。

[0174] LISH检验中使用的显微切割组织的大小

[0175] 多重LISH的产物大小均匀,且应因此可以最小PCR偏差而扩增。LISH的这一品质将会促进极小固定组织,如通过激光捕获显微切割(LCM)获得的那些组织的多重分析。确定仍会提供基因表达的稳定测量的最小量的显微切除组织。获得了大小为20,000平方微米下至200平方微米(分别约等于约300个至3个脾细胞的面积)范围内的10微米脾切片组织片段。对看家探针连接产物的定量显示了可靠的信号,该信号与LCM片段大小成正比(图8)。因此,



就获得对基因表达的稳健的多重测量所需的组织量而言,实施LISH是有利的。另一优点是,在显微切割为所希望的组织面积之前,整个切片被均匀地杂交并连接。这一途径最小化了技术差异,且规避了与从个体碎片中提取RNA相关的成本和材料损失。

[0176] 实施例4:用于诊断感染的LISH检验

[0177] 对于感染的诊断,用于检测FFPE标本中大数目的不同RNA序列的高敏感性技术将会极其有用。在临床真菌性中,例如,化学染色剂往往显示真菌有机体的存在,但该物种的后续定义可能非常困难。这是2012年的真实案例,当时,真菌性脑膜炎的离奇爆发后来被追踪至受污染的小瓶甲基强的松龙。

[0178] 从约翰霍普金斯医院(Johns Hopkins Hospital)中见到的患者获得极小的脑活组织检查切片。使用设计为区分三种密切相关真菌物种的多个探针组实施多重LISH检验。嘴图脐孢检测器探针记录了来自通过注射被嘴图脐孢污染的甲基强的松龙而被感染的索引或者的强上靶信号,但来自未感染组织或含有不相关感染如密切相关的烟曲霉的切片的活组织检查则无信号(图6A至6C)。因此,LISH被用来使用所保存的临床标本检测真菌感染。用这种方法,高度复杂的探针池可用于综合性诊断检验。在这一背景下,举例而言,可使用实时纳米孔测序或质谱实施对连接产物的快速分析。

[0179] 实施例5:用于免疫肿瘤学中诊断的LISH检验

[0180] 免疫肿瘤学是快速发展的领域,且近年的临床试验表明了引人注目的患者受益(Couzin-Frankel, J. Science (2013) 342:1432-1433)。举例而言,免疫肿瘤学可使用作为以身体免疫系统为靶向以帮助抗癌的免疫疗法而已知的药物。罹患貌似相似肿瘤的个体之间的应答可能变化各异,提示肿瘤微环境的特征奠定或破坏免疫治疗方案生效的基础。肿瘤免疫接口通常具有高度复杂的组织架构(如,结构),在该处,大量细胞类型、表面受体、和分泌产物来到一起以管理患者的抗肿瘤免疫应答。因此,显微切割标本中基因表达的高度多重测量可用以满足患者群体并指引治疗性干预。重要的是,本实施例中揭示的简单的LISH方法不需要专门的仪器或专门知识,并可立即整合入当前的临床病理学工作流程中。随着高通量DNA测序变得日益普遍和廉价,LISH平台和衍生物技术将被广泛用于常规临床应用和研究应用中。

[0181] 含有嵌合二核糖核苷酸的探针与固定RNA的基于Rn12的连接,将具有超出其本文中提出的应用。例如,在免疫染色之前,淋巴细胞被频繁固定并进行渗透化处理。一种具体实施例中,在染色和分拣之前实施LISH,以令细胞亚群或甚至可能在感兴趣的单一细胞中的基因表达成为可能。此外,对固定细胞培养物实施高通量LISH,可以用于基于多重基因表达的筛选中。将可扩增的DNA有效连接在固定RNA靶标序列上的能力,也将令多重新兴技术成为可能或促进该技术发展。例如,用于进行靶标转录本测序(Lee JH et al., Science (2014); 343(6177):1360-3)、实施邻近连接(Frei AP et al., Nature Methods (2016); 13(3):269-75)或将分子转位探针分发到固定RNA模板上(O'Roak BJ et al., Science (2012); 338(6114):1619-22)的原位方法,全部可能从本文中揭示的途径获益。

[0182] 其它等效物

[0183] 尽管已经协同其详细说明书而揭示本发明,前述说明书意图例示性说明而非限制本发明的范畴,本发明范畴为所附权利要求书的范畴所定义。其它方面、优点、及修饰处于所附权利要求书范畴内。

[0184] 本文中提及的专利和科学文献建立了该领域技术人员可获得的知识。本文中引用的所有美国专利和已公开或未公开的美国专利申请通过引用并入本文。本文中引用的所有已公开的外国专利和专利申请通过引用并入本文。本文中引用的以登录号表达的基因银行和NCBI文件通过引用并入本文。本文中引用的所有其它已出版的参考文献、文档、手稿和科学文献通过引用并入本文。

[0185] 尽管已经参考其优选具体例具体显示并说明了本发明,但该领域技术人员应理解,可对本发明做出各种形式和细节的变更而不悖离由所附权利要求书所涵盖的本发明范畴。

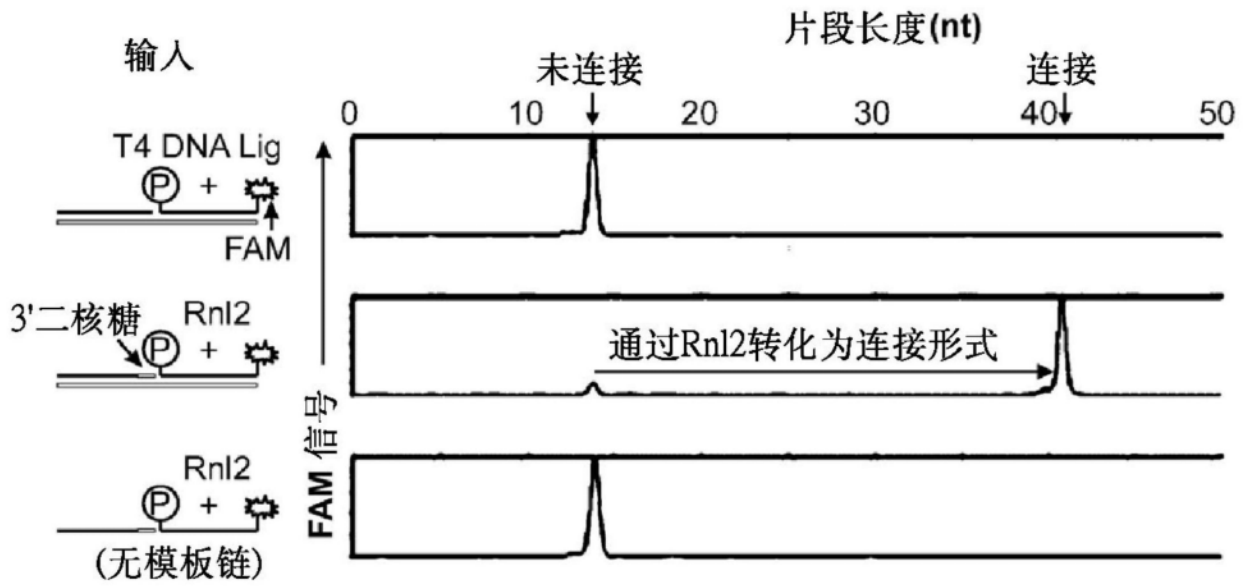


图1A

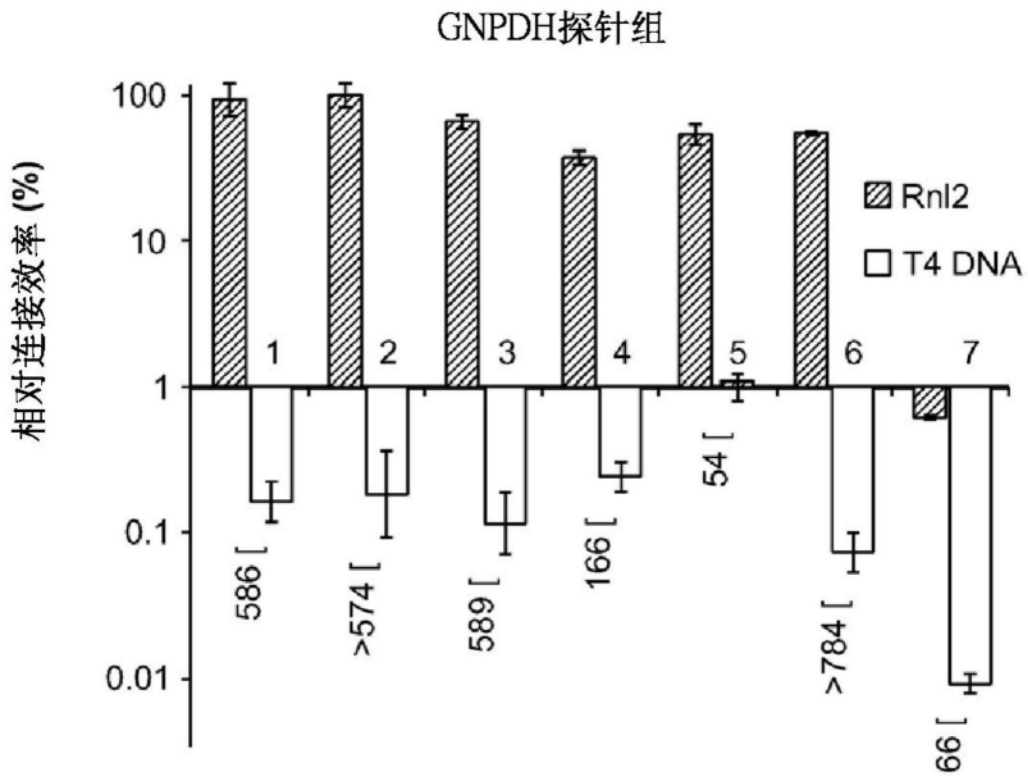


图1B

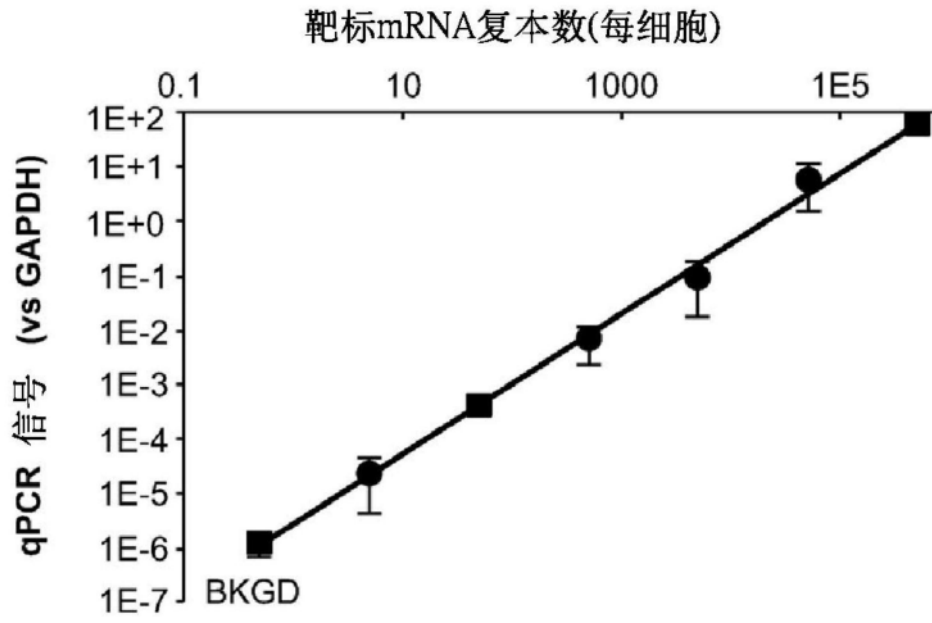


图1C

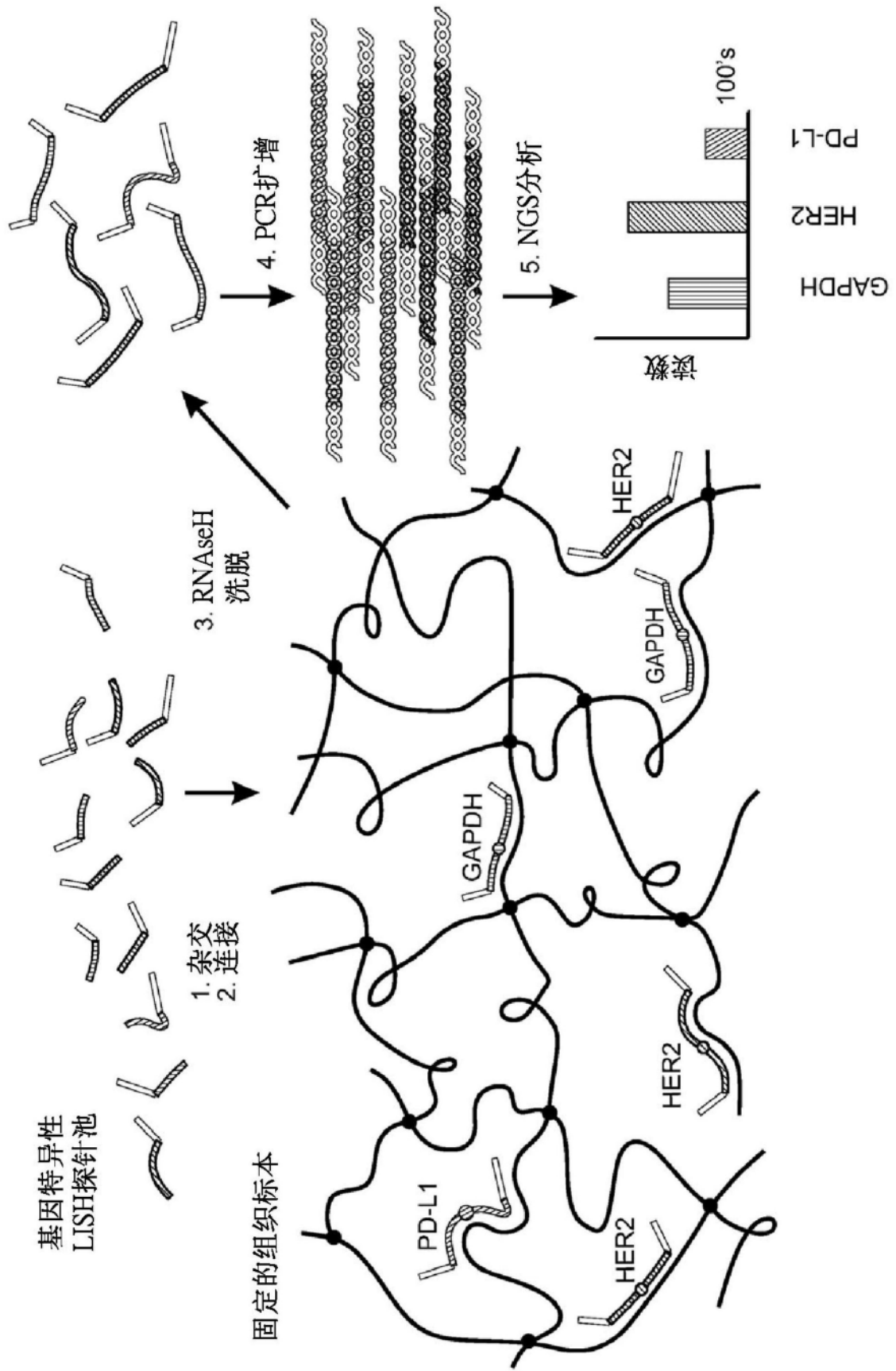


图2A

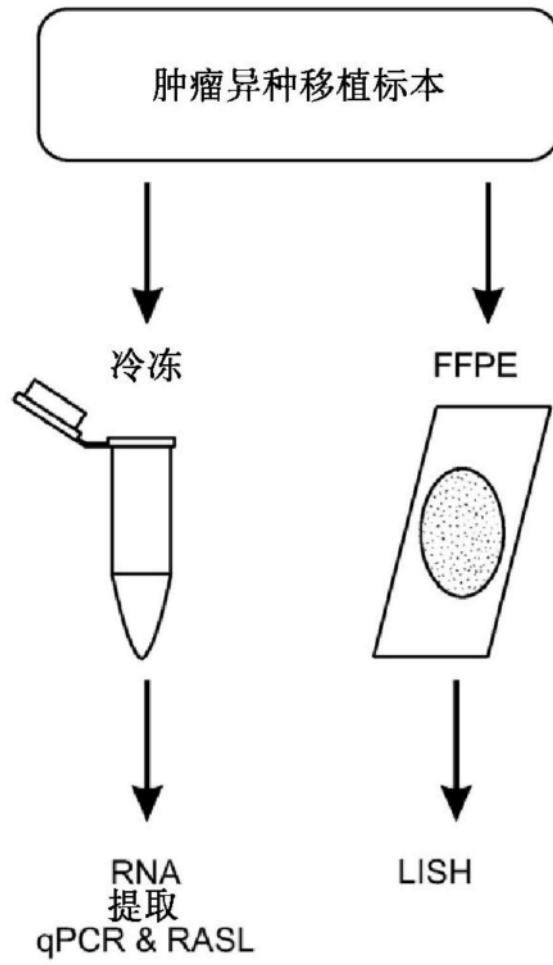


图2B

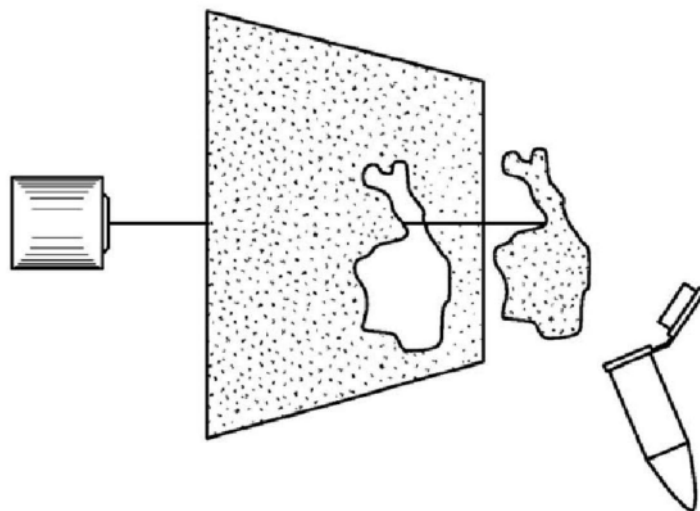


图2C

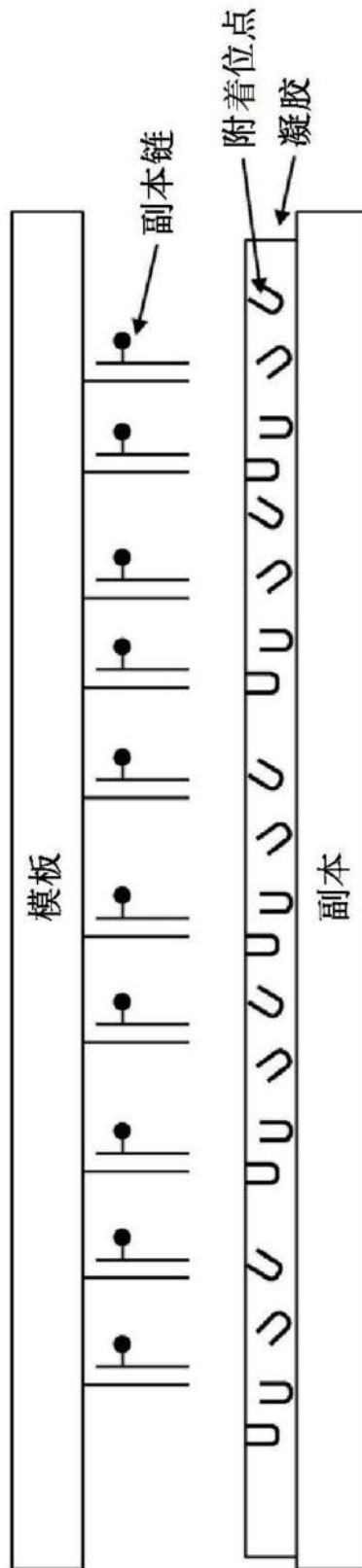


图3A

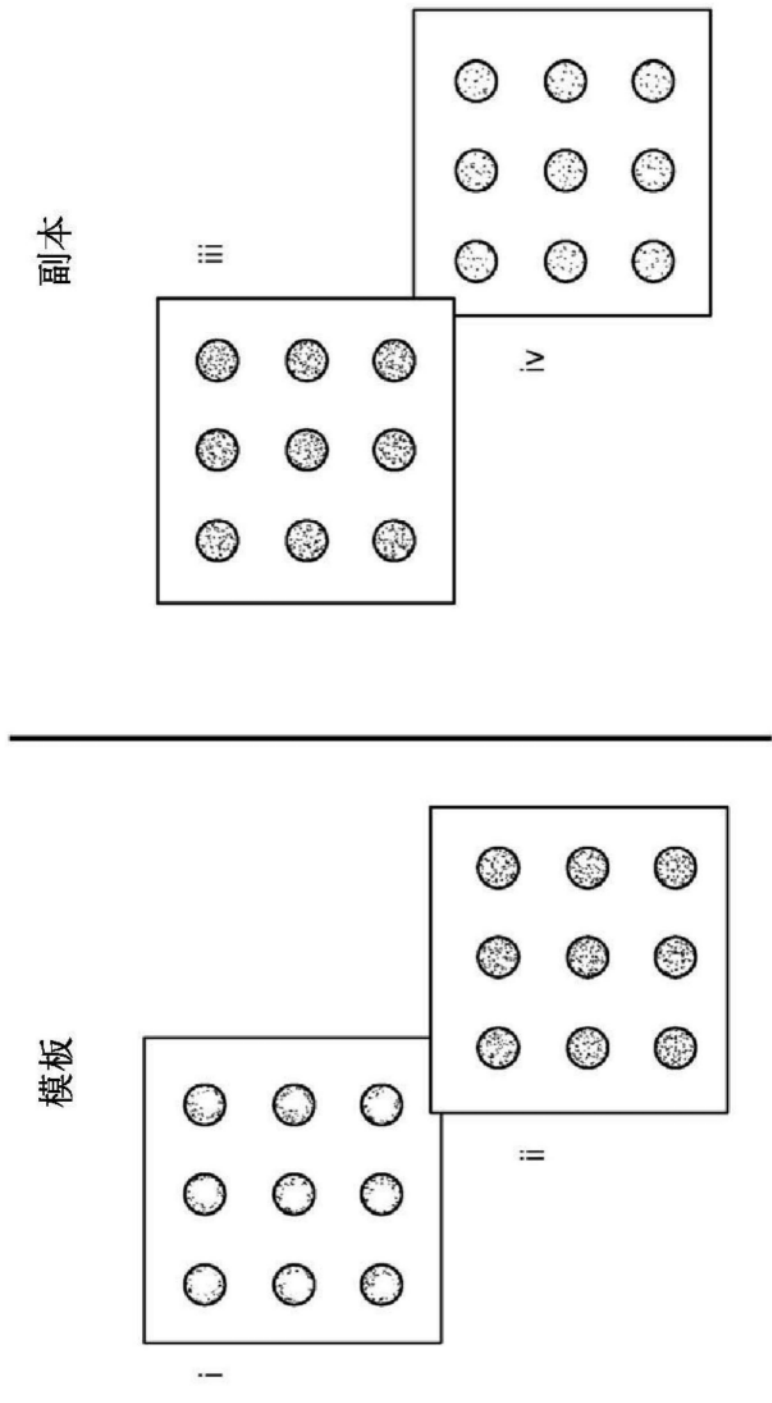


图3B



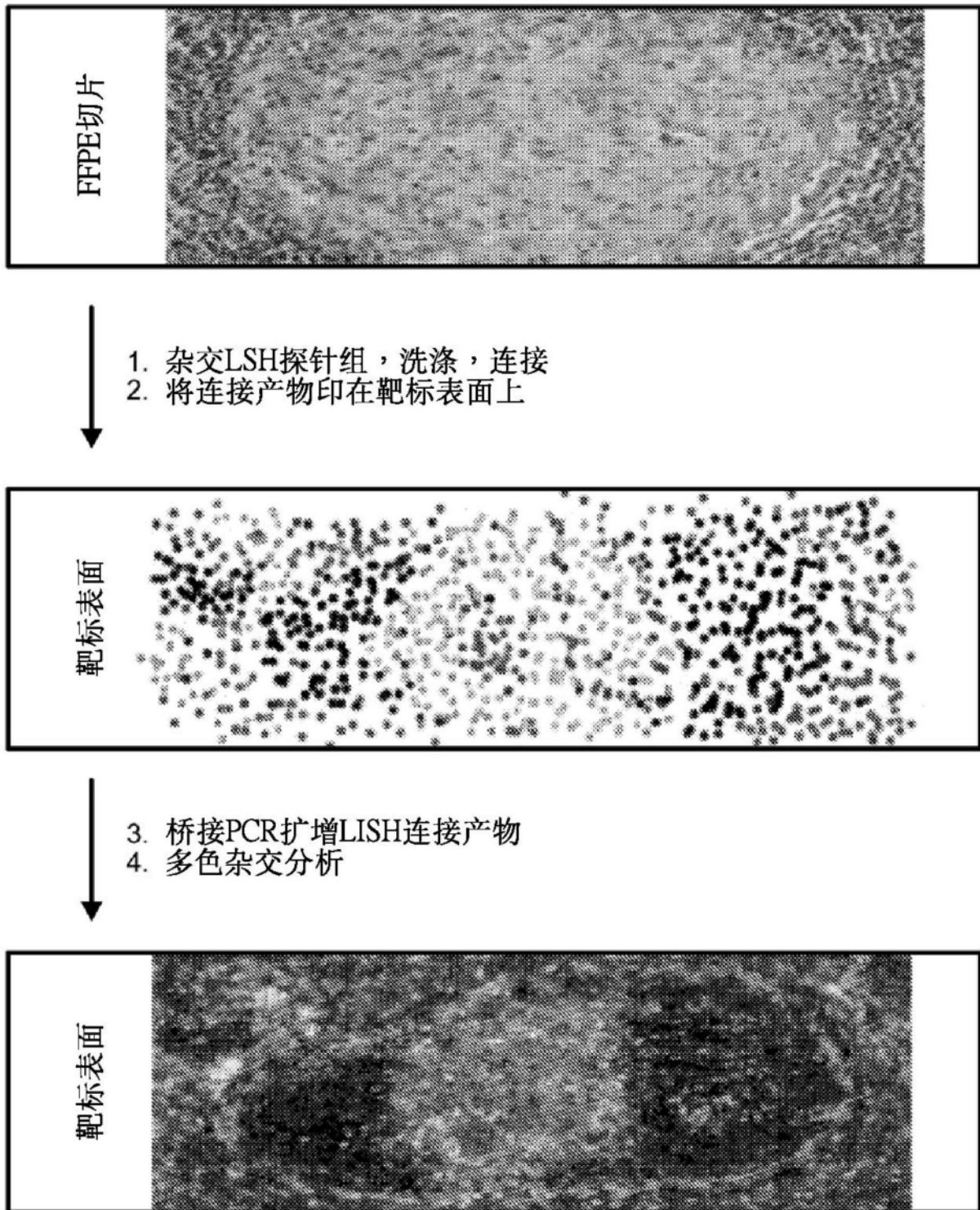


图4

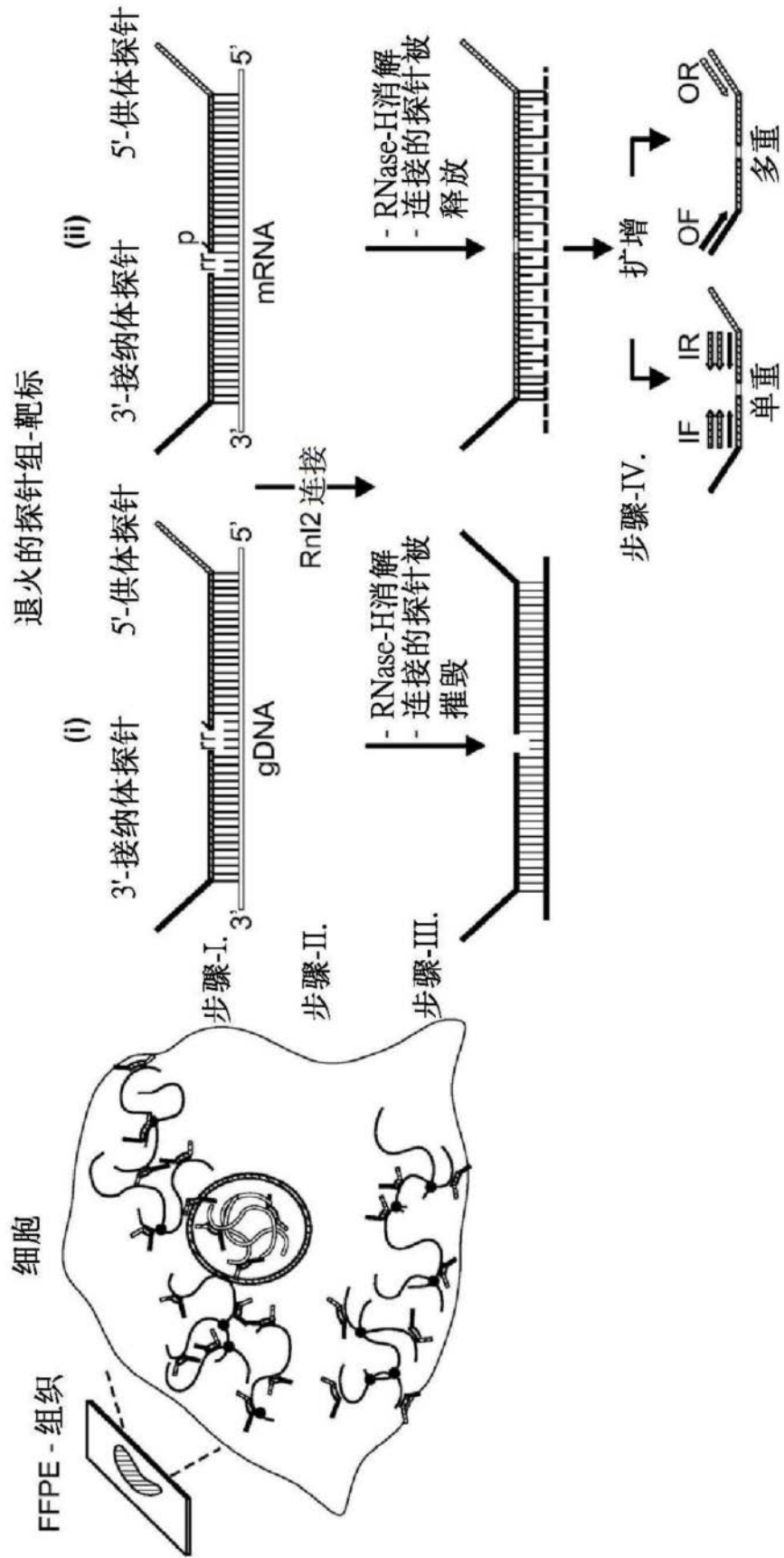


图5A

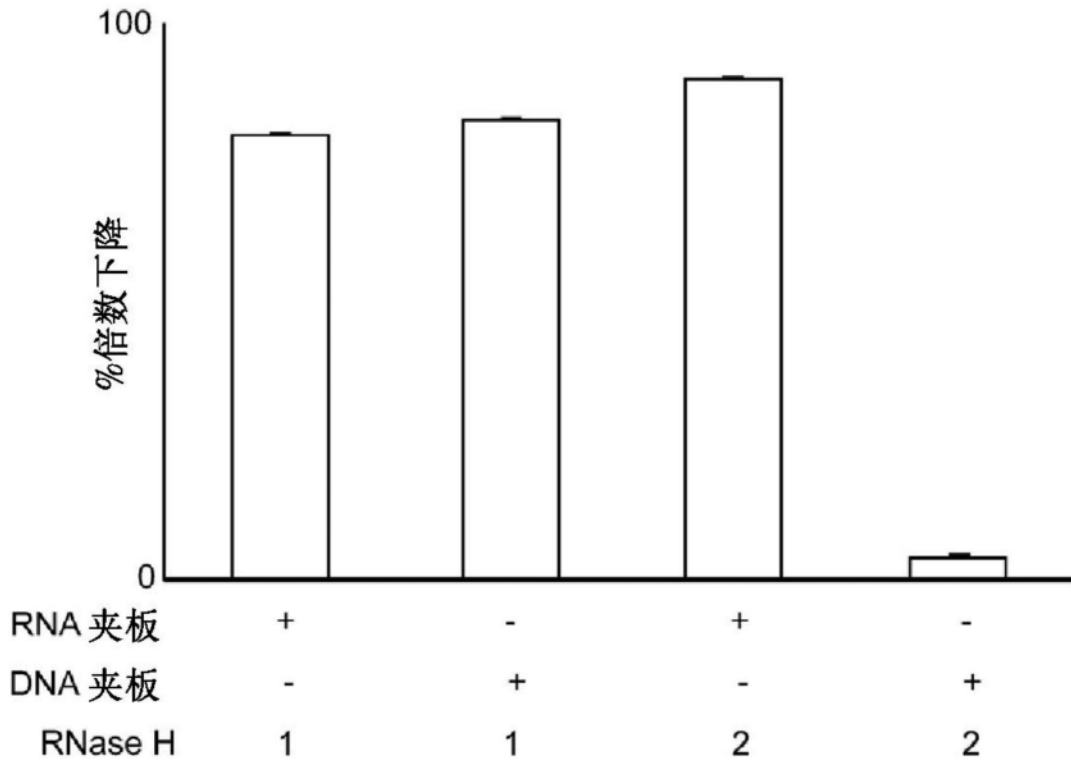


图5B

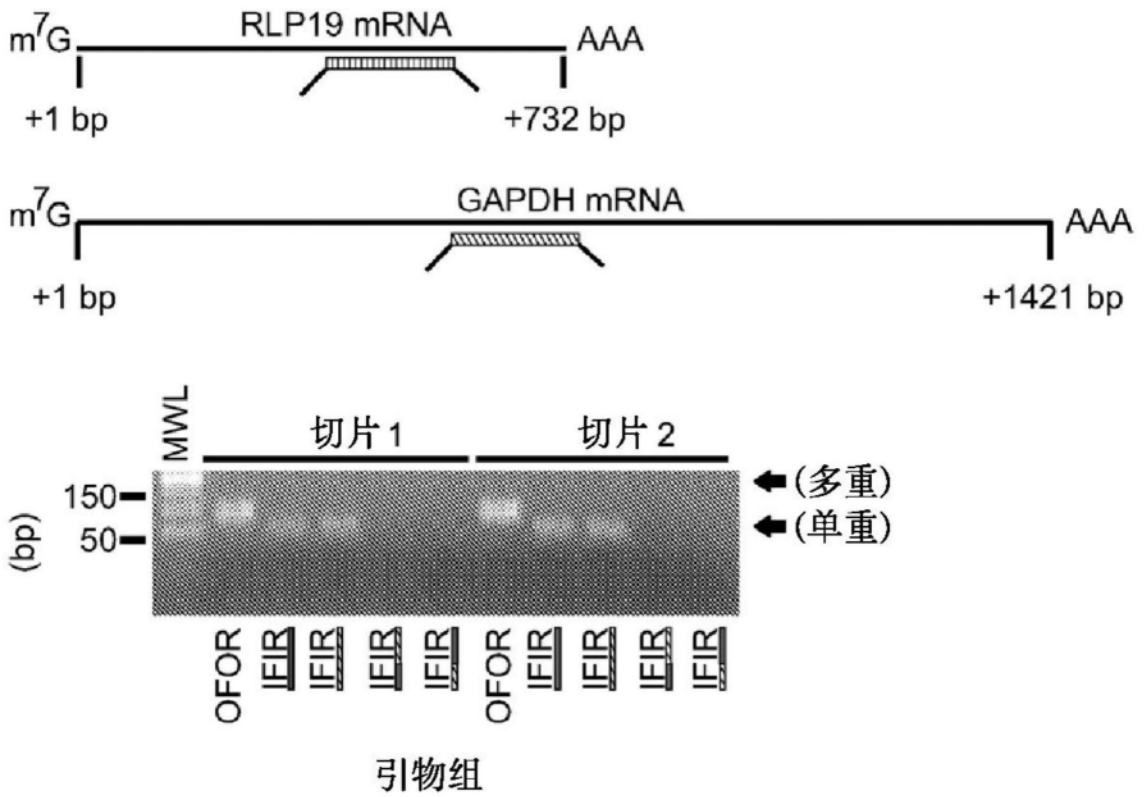


图5C

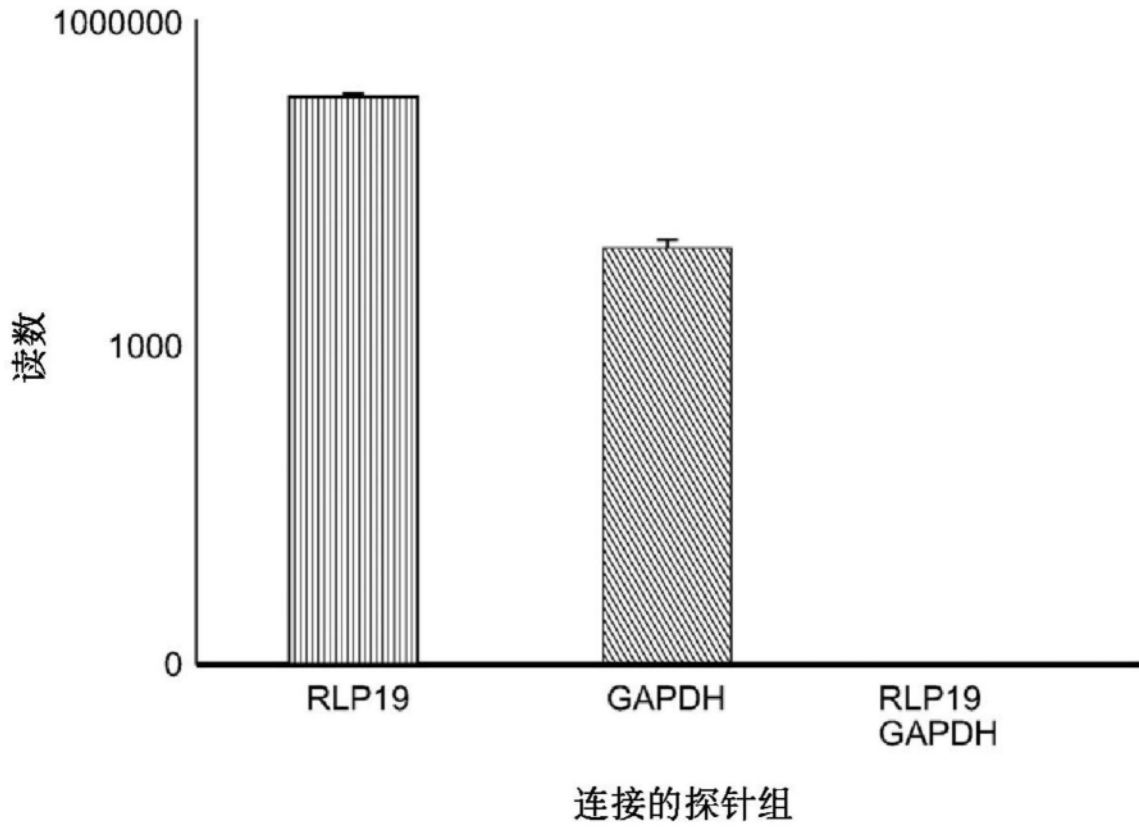


图5D

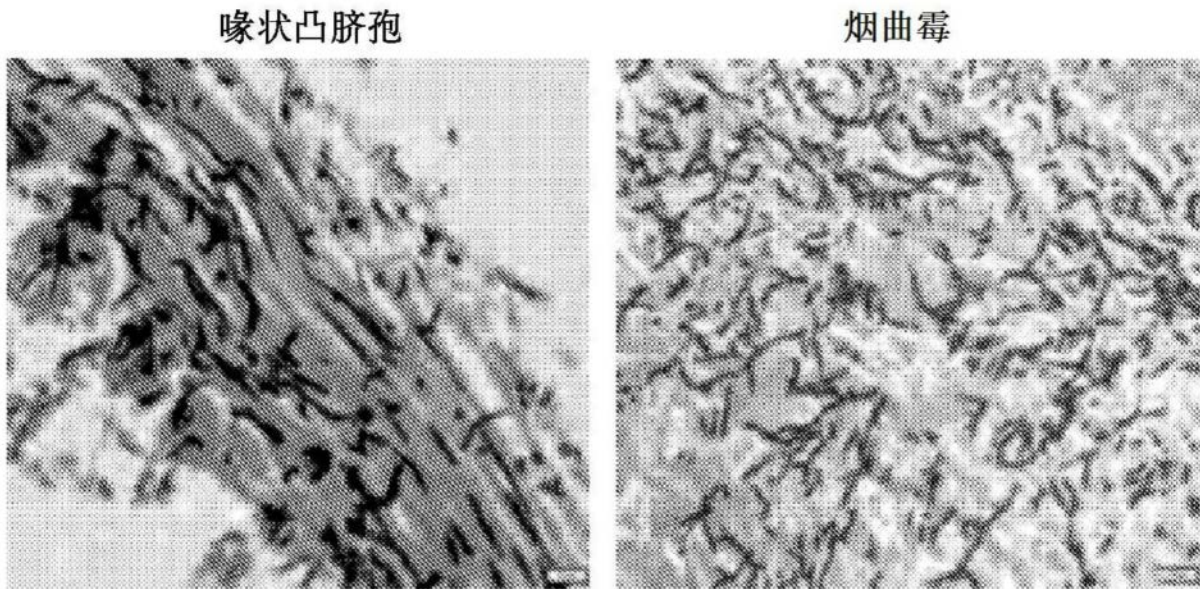


图6A

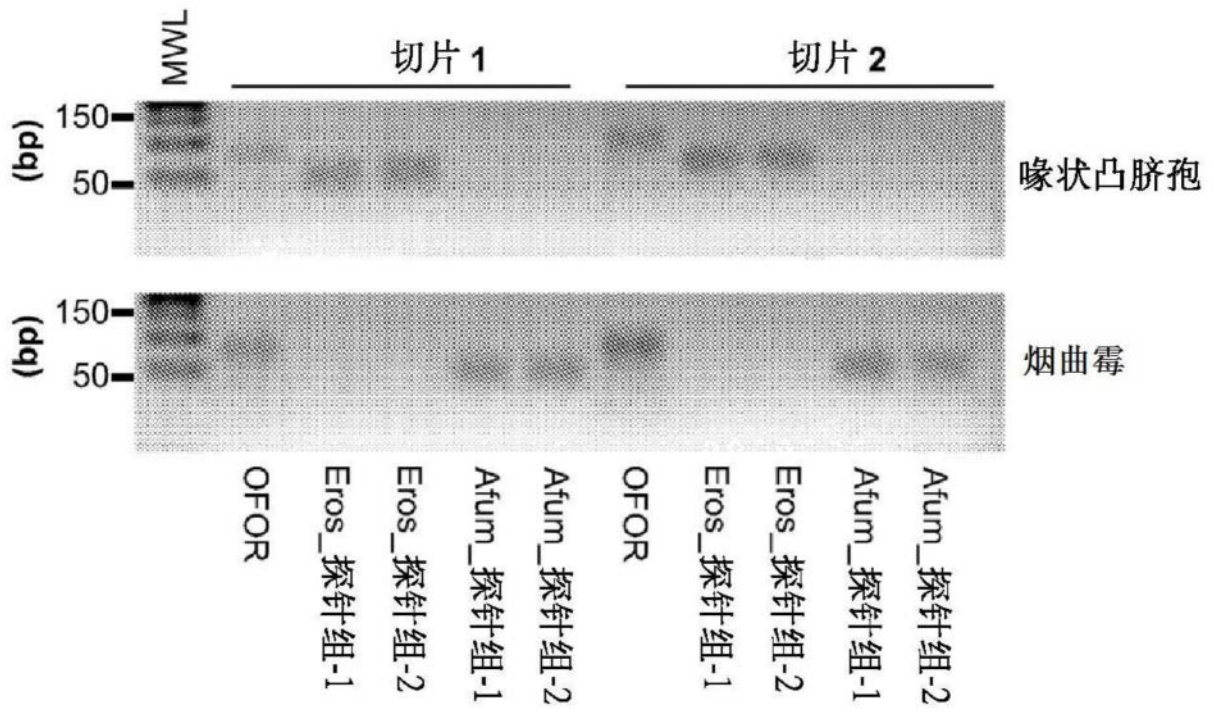


图6B

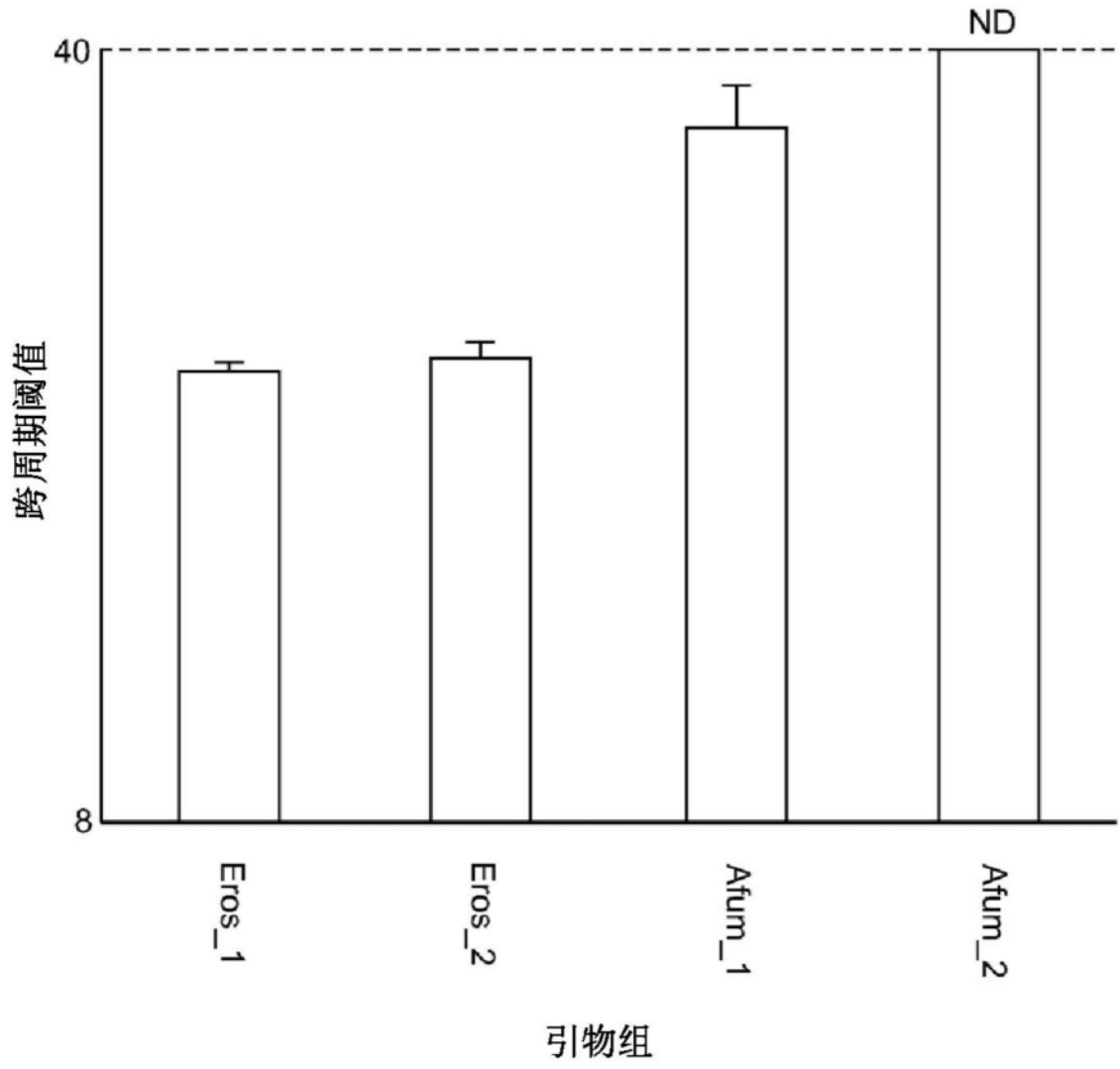


图6C



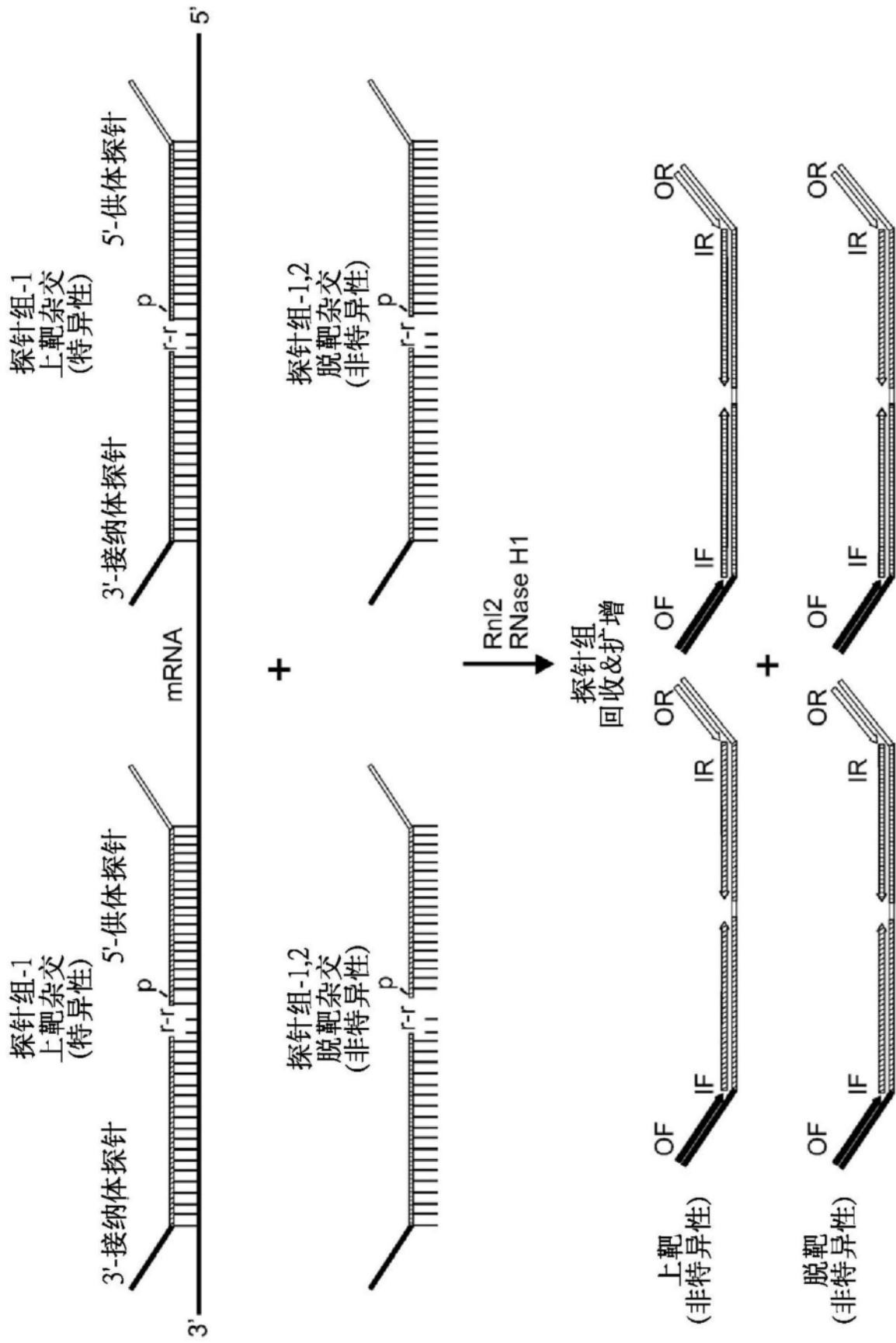


图7A

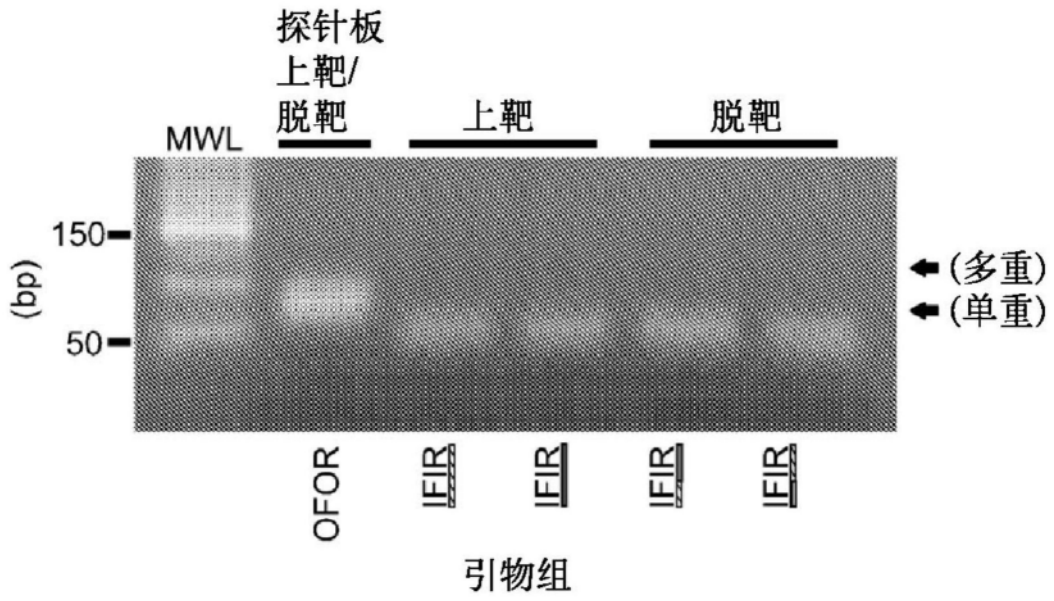


图7B

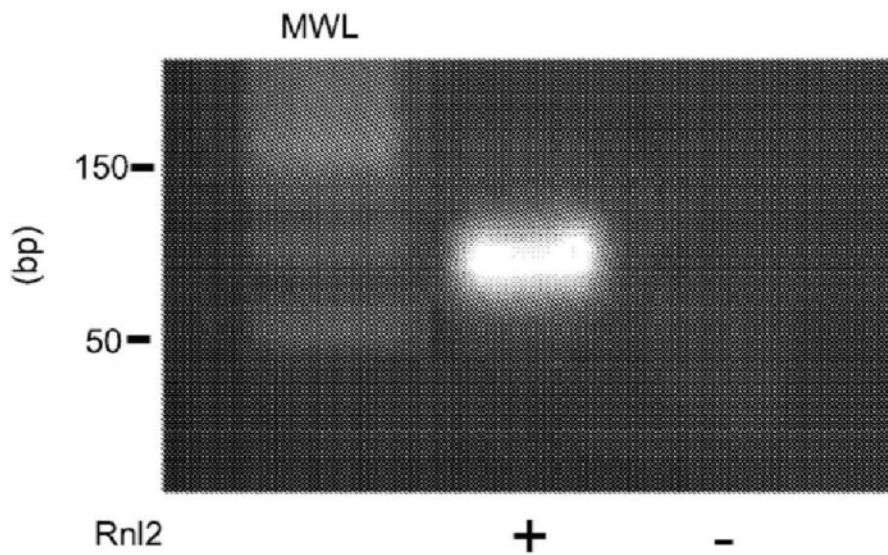


图7C



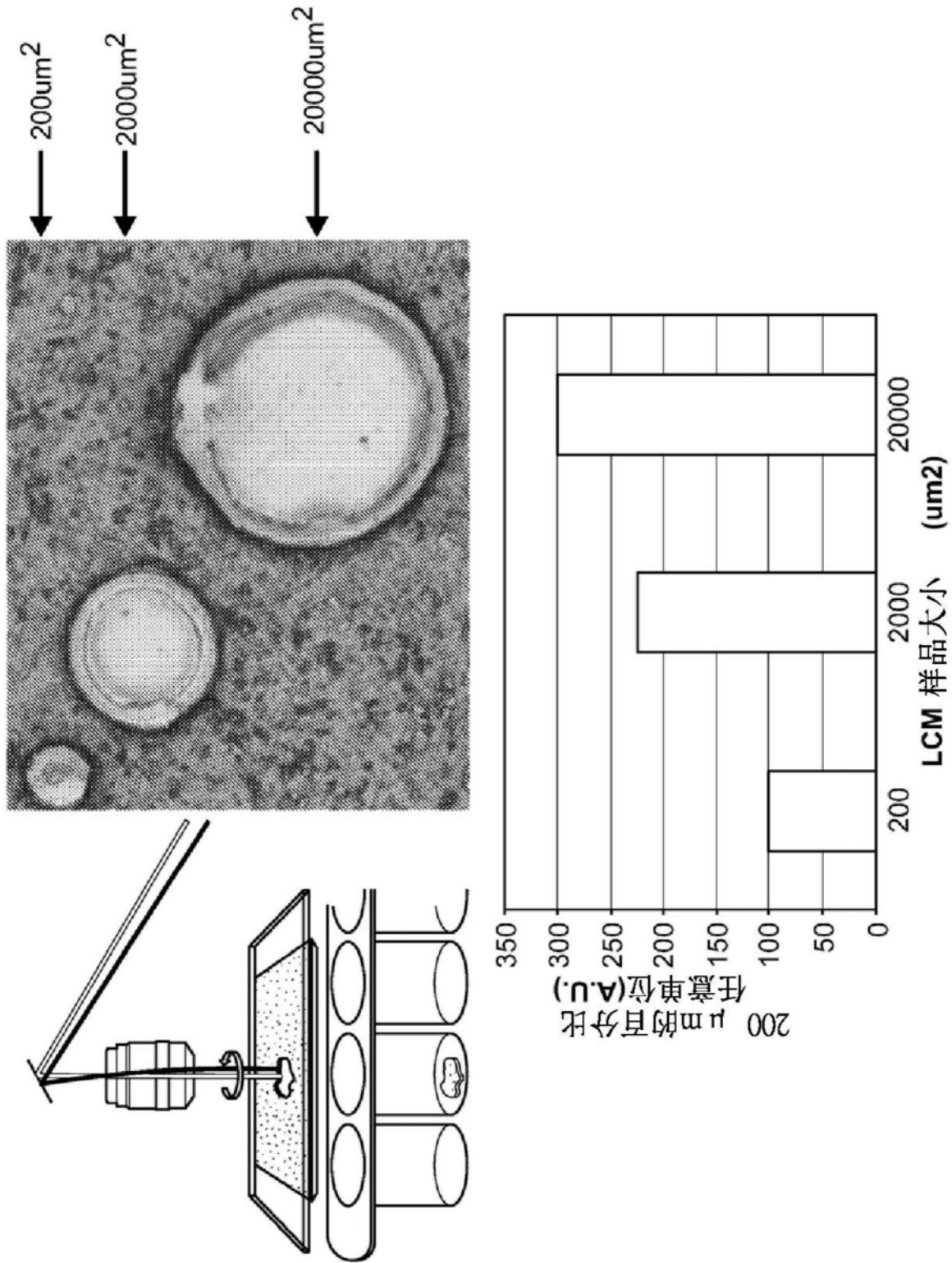


图8