(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119325511 A
(43)申请公布日 2025.01.17

(21)申请号 202380030039.5
(22)申请日 2023.03.24
(30)优先权数据
63/323,778 2022.03.25 US
(85)PCT国际申请进入国家阶段日
2024.09.24
(86)PCT国际申请的申请数据
PCT/US2023/016295 2023.03.24
(87)PCT国际申请的公布数据
WO2023/183611 EN 2023.09.28
(71)申请人 柏尔科学公司
地址 美国德克萨斯州
(72)发明人 K·D·艾伦 K·冈宁 A·拉巴斯 E·卡斯特罗
(74)专利代理机构 北京北翔知识产权代理有限公司 11285
专利代理师 孙占华 张广育
(51)Int.Cl.
C12Q 1/6806(2006.01)
C12Q 1/6855(2006.01)

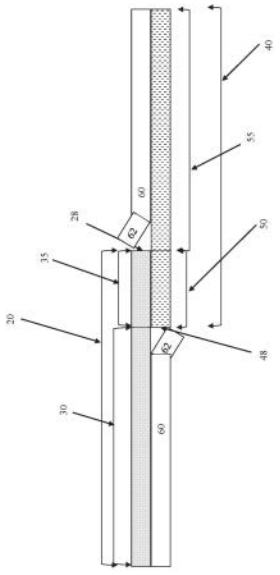
权利要求书3页 说明书26页
序列表（电子公布） 附图8页

(54)发明名称

用于抑制接头二聚体形成的方法、组合物和试剂盒

(57)摘要

根据本公开内容的方面，提供了减少测序文库中接头-二聚体的方法，其包括：使过量的3’接头与阻断剂寡核苷酸杂交以形成杂交复合体，使得过量的3’接头不能形成接头-二聚体，所述杂交复合体包括：第一阻断剂寡核苷酸3’端，其与第一未连接的3’接头的5’端处的腺苷酸化核苷酸相邻，和第二阻断剂寡核苷酸3’端，其与第二未连接的3’接头的5’端处的腺苷酸化核苷酸相邻，其中第一阻断剂寡核苷酸具有与第一未连接的3’接头互补并杂交的5’部分，第二阻断剂寡核苷酸具有与第二未连接的3’接头互补并杂交的5’部分，和第一阻断剂寡核苷酸具有与第二阻断剂寡核苷酸的3’部分互补并杂交的3’部分。第一和第二阻断剂寡核苷酸可能是相同的寡核苷酸，其包括允许自杂交的回文序列。



CN 119325511 A

1.一种在测序文库中减少接头-二聚体的方法,包括:

在连接反应条件下的混合物中,将各自具有5’末端和3’末端的样品核酸与3’接头和连接酶结合,所述3’接头各自具有长度为至少8个核苷酸的5’末端部分和一个为腺苷酸化核苷酸的5’端,产生1)第一连接产物,其各自包括连接至3’接头的5’端的样品核酸的3’末端,和2)混合物中过量的未连接的3’接头;

向所述混合物中加入阻断剂寡核苷酸,所述阻断剂寡核苷酸包括:

a)第一阻断剂寡核苷酸,其包括:

i)第一阻断剂寡核苷酸5’部分,其长度为至少8个核苷酸,和ii)第一阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第一寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个为非天然存在的核苷酸;和

b)第二阻断剂寡核苷酸,其包括:

iii)第二寡核苷酸5’部分,其长度为至少8个核苷酸,和iv)第二阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;

其中所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分长度相等且彼此互补;

其中所述第一阻断剂寡核苷酸5’部分和所述第二阻断剂寡核苷酸5’部分均与所述3’接头的5’末端部分互补;

在杂交条件下孵育混合物,从而使第一阻断剂寡核苷酸3’部分与第二阻断剂寡核苷酸3’部分杂交,使第一阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,并从而使第二阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,产生杂交复合体,其中第一阻断剂寡核苷酸3’端与第一未连接的3’接头的腺苷酸化5’末端核苷酸相邻,且第二阻断剂寡核苷酸3’端与第二未连接的3’接头的腺苷酸化5’末端核苷酸相邻,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

2.根据权利要求1的方法,其中所述第一阻断剂寡核苷酸3’端和所述第二阻断剂寡核苷酸3’端均能够连接至过量的未连接的3’接头的5’端处的腺苷酸化核苷酸;且进一步包括将所述杂交复合体的第一阻断剂寡核苷酸3’端连接至第一未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,并且将所述杂交复合体的第二阻断剂寡核苷酸3’端连接到第二未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,产生连接复合体,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

3.根据权利要求1或2的方法,其中第一阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物,且第二阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物。

4.根据权利要求1、2或3中任一项的方法,其中当所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸时,所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分各自包括回文核苷酸序列,其中所述第一阻断剂寡核苷酸3’部分的回文核苷酸序列与所述第二阻断剂寡核苷酸3’部分的回文核苷酸序列相同。

5.根据权利要求1至4中任一项的方法,进一步包括:向混合物中加入5'接头,所述5'接头具有3'端,其中所述第一连接产物各自包括与3'接头的5'端连接的样品核酸的3'末端,和与5'接头的3'端连接的样品核酸的5'末端,从而产生测序文库分子,所述测序文库分子包括位于5'接头和3'接头之间并与其结合的样品核酸插入片段。

6.根据权利要求1至5中任一项的方法,进一步包括:向纯化的混合物中加入5'接头;和在连接反应条件下孵育纯化的混合物,从而1)将所述5'接头与所述第一连接产物连接,产生测序文库分子,所述测序文库分子包括位于5'接头和3'接头之间并与其结合的样品核酸插入片段。

7.根据权利要求1至6中任一项的方法,其中减少但不清除过量的未连接的3'接头,使得所述连接包括将5'接头与过量的未连接的3'接头的至少一部分连接,产生接头-二聚体,其包括在5'接头-3'接头界面处与3'接头结合的5'接头;进一步包括抑制扩增反应中接头-二聚体的扩增,使得与接头-二聚体相比,通过扩增反应选择性富集测序文库分子。

8.根据权利要求1至7中任一项的方法,其中抑制接头-二聚体的扩增包括在扩增之前去除至少一些接头-二聚体。

9.根据权利要求7或8的方法,其中抑制接头-二聚体的扩增可包括以下中的一种或多种:接头-二聚体的特异性酶促切割;基于CRISPR/Cas9的接头-二聚体的切割;凝胶纯化;基于SPRI小珠的尺寸选择;和向混合物中加入接头-二聚体阻断剂寡核苷酸,所述接头-二聚体阻断剂寡核苷酸有效抑制接头-二聚体的扩增。

10.根据权利要求9的方法,其中所述接头-二聚体阻断剂寡核苷酸与所述接头-二聚体的一个或多个相邻部分互补,所述一个或多个相邻部分包括在5'接头-3'接头界面处的5'接头的至少一个核苷酸,和包括扩增引物结合位点的3'接头的相邻部分。

11.根据权利要求1至10中任一项的方法,其中所述第一阻断剂寡核苷酸和所述第二阻断剂寡核苷酸具有相同的核苷酸序列。

12.根据权利要求1的方法,其中所述样品核酸是或包括RNA,且进一步包括逆转录所述RNA,产生逆转录的测序文库分子。

13.根据权利要求1的方法,其中样品核酸是或包括DNA。

14.根据权利要求12或13的方法,进一步包括扩增所述测序文库分子。

15.一种用于减少测序文库中的接头-二聚体的寡核苷酸对,包括:

a)第一阻断剂寡核苷酸,其包括:

i)第一阻断剂寡核苷酸5'部分,其长度为至少8个核苷酸,和ii)第一阻断剂寡核苷酸3'部分,其长度为至少1至20个核苷酸,并包括第一寡核苷酸3'端,其中所述1至20个核苷酸中的至少1个为非天然存在的核苷酸;和

b)第二阻断剂寡核苷酸,其包括:

iii)第二寡核苷酸5'部分,其长度为至少8个核苷酸,和iv)第二阻断剂寡核苷酸3'部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3'端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;

其中所述第一阻断剂寡核苷酸3'部分和所述第二阻断剂寡核苷酸3'部分长度相等且彼此互补;和

其中所述第一阻断剂寡核苷酸5'部分和所述第二阻断剂寡核苷酸5'部分均与所述3'

接头的5’末端部分互补；

16.根据权利要求15的用于减少测序文库中接头-二聚体的寡核苷酸对，其中所述第一阻断剂寡核苷酸3’端和所述第二阻断剂寡核苷酸3’端均能够连接至过量的未连接的3’接头的5’端处的腺苷酸化核苷酸。

17.根据权利要求15或16的用于减少测序文库中接头-二聚体的寡核苷酸对，其中所述第一阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物，且所述第二阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物。

18.根据权利要求15、16或17任一项的用于减少测序文库中接头-二聚体的寡核苷酸对，其中当所述第一阻断剂寡核苷酸3’部分和第二阻断剂寡核苷酸3’部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸时，所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分各自包括回文核苷酸序列，其中所述第一阻断剂寡核苷酸3’部分的回文核苷酸序列与所述第二阻断剂寡核苷酸3’部分的回文核苷酸序列相同。

19.根据权利要求15至18任一项的用于减少测序文库中接头-二聚体的寡核苷酸对，其中所述第一阻断剂寡核苷酸和所述第二阻断剂寡核苷酸各自具有约11至约50个核苷酸范围内的长度。

20.根据权利要求15至19任一项的用于减少测序文库中接头-二聚体的寡核苷酸对，其中所述第一阻断剂寡核苷酸和所述第二阻断剂寡核苷酸具有相同的核苷酸序列。

21.一种用于减少测序文库中的接头-二聚体的试剂盒，其包括权利要求15至20中任一项的寡核苷酸对。

22.根据权利要求21的试剂盒，进一步包括有效阻断接头-二聚体扩增的接头-二聚体阻断剂寡核苷酸。

23.根据权利要求22的试剂盒，其中所述接头-二聚体阻断剂寡核苷酸与所述接头-二聚体的一个或多个相邻部分互补，所述一个或多个相邻部分包括在5’接头-3’接头界面处的5’接头的至少一个核苷酸，和包括扩增引物结合位点的3’接头的相邻部分。

24.根据权利要求1至14中任一项的方法或权利要求21至23中任一项的试剂盒，其中所述第一阻断剂和所述第二阻断剂具有相同的核酸序列。

25.根据权利要求1至14中任一项的方法或权利要求21至24中任一项的试剂盒，其中所述第一阻断剂和所述第二阻断剂具有不同的核酸序列。

26.一种基本上如本文所示或所述的方法、试剂盒或寡核苷酸对。

用于抑制接头二聚体形成的方法、组合物和试剂盒

[0001] 相关专利申请的交叉引用

[0002] 本申请要求2022年3月25日提交的序列号63/323,778的美国临时专利申请的优先权,其全部内容通过引用的方式纳入本文。

[0003] 序列表

[0004] 本申请含有序列表,该序列表已作为XML格式的文件以电子方式提交,且其全部内容通过引用的方式纳入本文。所述XML格式的文件创建于2023年3月24日,文件名为16NEN27152WO.xml,且大小为12,732字节。

技术领域

[0005] 本公开内容涉及用于减少测序文库中的接头-二聚体的方法、组合物和试剂盒。

背景技术

[0006] 遗传分析已变得越来越普遍,且其可用于多种分子生物学应用中。例如,个体的基因检测对于遗传疾病的早期检测特别有用,并且可以在特定疾病或病症的治疗选择中发挥作用。

[0007] 大规模平行核酸测序技术在靶核酸的遗传分析中起关键作用,但当非靶核酸与靶核酸一起测序时会产生不理想的结果。

[0008] 在大规模平行核酸测序技术中,靶向富集方法通常用于减少非靶序列对测序结果的污染。一种这样的富集方法依赖于靶核酸测序文库分子的杂交捕获。

[0009] 然而,用于大规模平行核酸测序的测序文库包括连接至核酸测序文库分子中的插入序列的接头序列。接头包括具有多种功能特性的核酸序列,例如延伸引物结合位点、用于样品鉴定的一个或多个索引序列、测序引物结合位点和扩增引物结合位点。在测序文库产生的方法的过程中,接头被连接至目的插入序列。接头连接至插入序列的不合需要的副产物是接头可以彼此连接,产生接头-二聚体。即使少量的接头-二聚体也会对测序质量产生不利影响,因为接头-二聚体可以通过旨在富集目的插入序列的扩增反应而有效地扩增。

[0010] 进一步地,接头-二聚体可以在富集步骤期间被杂交捕获保留,从而捕获接头-二聚体并浪费试剂和时间来测序接头-二聚体。

[0011] 大规模平行测序技术的日益扩大的应用和日益增长的在单次运行中多重处理(multiplex)大量单个文库的需求导致需要有效抑制接头-二聚体的形成,并去除那些已经形成的接头-二聚体。

[0012] 持续需要用于在用于大规模平行测序的测序文库中减少接头-二聚体的组合物和方法。

发明内容

[0013] 根据本公开内容的方面,提供了在测序文库中减少接头-二聚体的方法,其包括:在连接反应条件下的混合物中,将各自具有5’末端和3’末端的样品核酸与3’接头和连接酶

组合,所述3’接头各自具有长度为至少8个核苷酸的5’末端部分和一个为腺苷酸化核苷酸的5’端,产生1)第一连接产物,其各自包括连接至3’接头的5’端的样品核酸的3’末端,和2)混合物中过量的未连接的3’接头;向所述混合物中加入阻断剂寡核苷酸(blocker oligonucleotide),所述阻断剂寡核苷酸包括:a)第一阻断剂寡核苷酸,其包括:i)第一阻断剂寡核苷酸5’部分,其长度为至少8个核苷酸,和ii)第一阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第一寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个为非天然存在的核苷酸;和b)第二阻断剂寡核苷酸,其包括:iii)第二寡核苷酸5’部分,其长度为至少8个核苷酸,和iv)第二阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;其中所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分长度相等且彼此互补;并且其中所述第一阻断剂寡核苷酸5’部分和所述第二阻断剂寡核苷酸5’部分均与所述3’接头的5’末端部分互补;并且在杂交条件下孵育混合物,从而使第一阻断剂寡核苷酸3’部分与第二阻断剂寡核苷酸3’部分杂交,使第一阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,并从而使第二阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,产生杂交复合体,其中第一阻断剂寡核苷酸3’端与第一未连接的3’接头的腺苷酸化5’末端核苷酸相邻,且第二阻断剂寡核苷酸3’端与第二未连接的3’接头的腺苷酸化5’末端核苷酸相邻,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

[0014] 根据本公开内容的方面,第一阻断剂寡核苷酸3’端和第二阻断剂寡核苷酸3’端均能够连接至过量的未连接的3’接头的5’端处的腺苷酸化核苷酸;且在本公开内容的方法中进一步包括将杂交复合体的第一阻断剂寡核苷酸3’端连接至第一未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,并且将杂交复合体的第二阻断剂寡核苷酸3’端连接到第二未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,产生连接复合体,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

[0015] 根据本公开内容的方面,第一阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸,且第二阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸。

[0016] 根据本公开内容的方面,第一阻断剂寡核苷酸3’部分和第二阻断剂寡核苷酸3’部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸,且第一阻断剂寡核苷酸3’部分和第二阻断剂寡核苷酸3’部分各自包括回文核苷酸序列,其中第一阻断剂寡核苷酸3’部分的回文核苷酸序列与第二阻断剂寡核苷酸3’部分的回文核苷酸序列相同。

[0017] 根据本公开内容的方面,第一阻断剂寡核苷酸和第二阻断剂寡核苷酸各自具有11至50个核苷酸范围内的长度。

[0018] 根据本公开内容的方面,第一阻断剂寡核苷酸和第二阻断剂寡核苷酸具有相同的核苷酸序列。

[0019] 根据本公开内容的方面,减少测序文库中的接头-二聚体的方法进一步包括向混合物中加入5’接头,所述5’接头具有3’端,其中所述第一连接产物各自包括与3’接头的5’端连接的样品核酸的3’末端,和与5’接头的3’端连接的样品核酸的5’末端,从而产生测序

文库分子,所述测序文库分子包括位于5'接头和3'接头之间并与其结合的样品核酸插入片段(insert)。

[0020] 根据本公开内容的方面,减少测序文库中的接头-二聚体的方法还包括向纯化的混合物中加入5'接头;和在连接反应条件下孵育纯化的混合物,从而1)将5'接头与第一连接产物连接,产生测序文库分子,所述测序文库分子包括位于5'接头和3'接头之间并与其结合的样品核酸插入片段。

[0021] 根据本公开内容的方面,提供了减少测序文库中的接头-二聚体的方法,其中减少但不清除过量的未连接的3'接头,使得所述连接包括将5'接头与过量的未连接的3'接头的至少一部分连接,产生接头-二聚体,其包括在5'接头-3'接头界面处(interface)与3'接头结合的5'接头。因此,根据本公开内容的方面,提供了减少测序文库中的接头-二聚体的方法,其进一步包括抑制扩增反应中接头-二聚体的扩增,使得与接头-二聚体相比,通过扩增反应选择性富集测序文库分子。抑制接头-二聚体的扩增可包括以下中的一种或多种:在扩增之前去除至少一些接头-二聚体;接头-二聚体的特异性酶促切割;基于CRISPR/Cas9的接头-二聚体的切割;凝胶纯化以去除接头-二聚体;基于SPRI小珠的尺寸选择以去除接头-二聚体;以及向混合物中加入接头-二聚体阻断剂寡核苷酸,所述接头-二聚体阻断剂寡核苷酸有效抑制接头-二聚体的扩增。

[0022] 根据本公开内容的方面,提供了减少测序文库中接头-二聚体的方法,其包括向混合物中加入接头-二聚体阻断剂寡核苷酸,所述接头-二聚体阻断剂寡核苷酸有效抑制接头-二聚体的扩增,其中所述接头-二聚体阻断剂寡核苷酸与接头-二聚体的一个或多个相邻部分互补,所述一个或多个相邻部分包括在5'接头-3'接头界面处的5'接头的至少一个核苷酸,和包括扩增引物结合位点的3'接头的相邻部分。

[0023] 根据本公开内容的方面,提供了在测序文库中减少接头-二聚体的方法,其中样品核酸为RNA或包括RNA。

[0024] 根据本公开内容的方面,提供了减少测序文库中的接头-二聚体的方法,其中样品核酸为RNA或包括RNA,且所述方法包括逆转录RNA,产生逆转录的测序文库分子。

[0025] 根据本公开内容的方面,提供了减少测序文库中的接头-二聚体的方法,其进一步包括扩增测序文库分子或逆转录的测序文库分子。

[0026] 根据本公开内容的方面,提供了一种用于减少测序文库中的接头-二聚体的寡核苷酸对,其包括a)第一阻断剂寡核苷酸,其包括:i)长度为至少8个核苷酸的第一阻断剂寡核苷酸5'部分和ii)长度为至少1至20个核苷酸且包括第一寡核苷酸3'端的第一阻断剂寡核苷酸3'部分,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;和b)第二阻断剂寡核苷酸,其包括:iii)长度为至少8个核苷酸的第二寡核苷酸5'部分和iv)第二阻断剂寡核苷酸3'部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3'端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;其中第一阻断剂寡核苷酸3'部分和第二阻断剂寡核苷酸3'部分长度相等且彼此互补;其中第一阻断剂寡核苷酸5'部分和第二阻断剂寡核苷酸5'部分均与3'接头的5'末端部分互补。

[0027] 根据本公开内容的方面,第一阻断剂寡核苷酸3'端和第二阻断剂寡核苷酸3'端均能够连接至过量的未连接的3'接头的5'端处的腺苷酸化核苷酸。

[0028] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的阻断剂寡

核苷酸对,其中第一阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸,且第二阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸。

[0029] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的阻断剂寡核苷酸对,其中当第一阻断剂寡核苷酸3'部分和第二阻断剂寡核苷酸3'部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸时,第一阻断剂寡核苷酸3'部分和第二阻断剂寡核苷酸3'部分各自包括回文核苷酸序列,其中第一阻断剂寡核苷酸3'部分的回文核苷酸序列与第二阻断剂寡核苷酸3'部分的回文核苷酸序列相同。

[0030] 根据本公开内容的方面,提供了用于减少测序文库中接头-二聚体的阻断剂寡核苷酸对,其中第一阻断剂寡核苷酸和第二阻断剂寡核苷酸各自具有约11至约50个核苷酸范围内的长度。

[0031] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的阻断剂寡核苷酸对,其中第一阻断剂寡核苷酸和第二阻断剂寡核苷酸具有相同的核苷酸序列。

[0032] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的试剂盒,其包括一个或多个用于用于减少测序文库中的接头-二聚体的阻断剂寡核苷酸对。

[0033] 根据本公开内容的方面,用于减少测序文库中的接头-二聚体的试剂盒进一步包括有效阻断接头-二聚体扩增的接头-二聚体阻断剂寡核苷酸。

[0034] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的试剂盒,其中接头-二聚体阻断剂寡核苷酸与接头-二聚体的一个或多个相邻部分互补,所述一个或多个相邻部分包括在5'接头-3'接头界面处的5'接头的至少一个核苷酸,和包括扩增引物结合位点的3'接头的相邻部分。

附图说明

[0035] 图1为根据本公开内容的方面的包括阻断剂寡核苷酸和3'接头的杂交复合体的图解说明;图1包括显示杂交复合体特征的编号,所述杂交复合体包括:第一阻断剂寡核苷酸20;第一阻断剂寡核苷酸3'端28;第一阻断剂寡核苷酸5'部分30;第一阻断剂寡核苷酸3'部分35;第二阻断剂寡核苷酸5'部分55;第二阻断剂寡核苷酸40;第二阻断剂寡核苷酸3'端48;第二阻断剂寡核苷酸3'部分50;3'接头60;和3'接头的5'腺苷酸化末端62。

[0036] 图2A为示出"单接头系统"的示意图,其中阻断寡核苷酸具有相同的序列,包括阻断剂寡核苷酸5'部分(深灰色)且包括回文阻断剂寡核苷酸3'部分(浅灰色),使得回文阻断剂寡核苷酸3'部分如图所示互补并退火;例如,"单接头系统"用于测序文库生成的方法中,其中包括单一类型的接头。

[0037] 图2B为示出包括至少两种类型的阻断寡核苷酸的"多接头系统"的示意图,其中不同类型的阻断寡核苷酸具有不同的序列,包括不同的阻断剂寡核苷酸5'部分(深灰色,左上;和中灰色,左下),且包括相同的回文阻断剂寡核苷酸3'部分(浅灰色),使得回文阻断剂寡核苷酸3'部分如图所示互补并退火;例如,"多接头系统"用于测序文库生成的方法中,其中包括多种类型的接头,如每个接头包括不同的索引序列。

[0038] 图2C为示出包括至少两种类型的阻断寡核苷酸的"单接头系统"的示意图,其中不

同类型的阻断寡核苷酸具有不同的序列,包括相同的阻断剂寡核苷酸5'部分(深灰色)且包括不同的非回文阻断剂寡核苷酸3'部分(浅灰色,顶部;点阵灰色,底部),使得非回文的阻断剂寡核苷酸3'部分如图所示互补并退火;例如,"单接头系统"用于测序文库生成的方法中,其中包括单一类型的接头。

[0039] 图3为示出根据本公开内容的方面抑制接头-二聚体形成的图。图3示出了a)成功纳入测序文库分子中的所需RNA的比例(深灰色);b)形成且纳入测序文库分子中的接头-二聚体(标准二聚体(canonical dimer))的比例(中灰色);和c)在使用以下方法减少测序文库中接头-二聚体的方法中纳入测序文库分子中的其它分子的比例(浅灰色):#1)没有阻断接头二聚体形成的方法;#2)包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法;3#)包括在扩增之前去除至少一些接头-二聚体的方法;或#4)包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸且进一步包括在扩增之前去除至少一些接头-二聚体的方法。

[0040] 图4为示出结果的图,所述结果表明,与常规方法相比,由于使用包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法,纳入测序文库中的接头-二聚体的百分比显著降低,测序的二聚体的百分比显著降低。进一步示出了与常规方法相比,由于使用包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法,纳入测序文库中的接头-二聚体的百分比显著降低,测序的成熟miRNA的百分比增加。在图4中,浅灰色条1表示使用包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法获得的数据,深灰色条2表示使用常规方法获得的数据。

[0041] 图5为示出使用包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法获得的结果的图,且深灰色条2表示使用常规方法获得的数据。结果示出了对测序数据中表示的独特miRNA的数量的比较,即,与使用常规方法至少1次发现的仅509个独特单个序列相比,使用包括根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法至少1次发现的600个独特单个序列;与使用常规方法至少3次发现的仅424个独特单个序列相比,使用包括根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法至少3次发现的471个独特序列;与使用常规方法至少5次发现的仅370个独特单个序列相比,使用包括根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法至少5次发现的407个独特单个序列;以及与使用常规方法至少10次发现的仅295个独特单个序列相比,使用包括根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法至少10次发现的339个独特单个序列。

[0042] 图6为文氏(Venn)图,其示出了使用包括使用根据本公开内容的方面的一种或多种阻断剂寡核苷酸的方法的3次重复(大灰色圆圈)与使用常规方法获得的3次重复(中等大小的中灰色圆圈)之间相同的阈值为5的发现的独特miRNA的数量的比较。最小的黑色圆圈代表两种方法共有的独特miRNA的数量。

具体实施方式

[0043] 本文使用的科学和技术术语旨在具有本领域普通技术人员通常理解的含义。这些术语在多种标准参考文献的上下文中被定义和使用,其举例说明性地包括J.Sambrook and D.W.Russell,Molecular Cloning:ALaboratory Manual,Cold Spring Harbor Laboratory Press;3rd Ed.,2001;F.M.Ausubel,Ed.,Short Protocols in Molecular

Biology,Current Protocols;5th Ed.,2002;B.Alberts et al.,Molecular Biology of the Cell,4th Ed.,Garland,2002;D.L.Nelson and M.M.Cox,Lehninger Principles of Biochemistry,4th Ed.,W.H.Freeman&Company,2004;Herdewijn,P.(Ed.),Oligonucleotide Synthesis:Methods and Applications,Methods in Molecular Biology,Humana Press,2004;Remington:The Science and Practice of Pharmacy,Lippincott Williams&Wilkins,21st Ed.,2005;L.V.Allen,Jr.et al.,Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems,8th Ed.,Philadelphia,PA:Lippincott,Williams&Wilkins,2004;和L.Brunton et al.,Goodman&Gilman's The Pharmacological Basis of Therapeutics,McGraw-Hill Professional,12th Ed.,2011。

[0044] 单数术语“一”、“一个”和“所述”(“a”、“an”和“the”)并不旨在具有限制性,且包括复数指代对象,除非另有明确说明或上下文另有明确指示。

[0045] 本文涉及参考数值所使用的术语“约(about)”及其语法等同形式是指参考数值和参考数值10%以内的数值,包括参考数值的加或减(+/-)1%、+/-2%、+/-3%、+/-4%、+/-5%、+/-6%、+/-7%、+/-8%、+/-9%或+/-10%的数值。在所述参考数值范围的上下文中,本文所使用的术语“约(about)”及其语法等同形式是指包括数值范围的参考下限的+/-1%、+/-2%、+/-3%、+/-4%、+/-5%、+/-6%、+/-7%、+/-8%、+/-9%或+/-10%的数值,以及数值范围的参考上限的+/-1%、+/-2%、+/-3%、+/-4%、+/-5%、+/-7%、+/-8%、+/-9%或+/-10%的数值的范围。进一步的,除非另有说明,应理解本文的数值列表为包括所列数值的所有中间值和分数值,如,50%、60%、75%被理解为包括55%、64.5%和74%。

[0046] 术语“第一”、“第二”等可在本文中用于指元素(element)或元素的组合,且这些术语用于指这样的元素或元素的组合,而不限制第一元素和第二元素是彼此相同还是不同。

[0047] 特征的特定组合记载于权利要求中和/或公开于说明书中,且这些特征的组合不旨在限制各个方面的公开内容。这些特征的所有可能的组合没有具体记载于权利要求中和/或公开于说明书中。尽管下列每个从属权利要求可仅直接从属于一个权利要求,但是各个方面的公开内容包括每个从属权利要求与权利要求书中每个其他权利要求的组合。如本文所使用的,短语项目列表中的“至少一个”是指那些项目的任何组合,包括单个成员。例如,“a、b或c中的至少一个”旨在涵盖单独a;单独b;单独c;a和b;a、b和c;b和c;a和c,以及与多个相同元素的任何组合,例如a和a;a、a和a;a、a和b;a、a和c;a、b和b;a、c和c;及a、b和c的任何其它组合或排序。

[0048] 根据本公开内容的方面,提供了减少测序文库中接头-二聚体的方法,其包括:使过量的3'接头与阻断剂寡核苷酸杂交以形成杂交复合体,使得过量的3'接头不能形成接头-二聚体,其中杂交复合体包括第一阻断剂寡核苷酸,其具有与第一未连接的3'接头互补并杂交的5'部分,第二阻断剂寡核苷酸,其具有与第二未连接的3'接头互补并杂交的5'部分,且其中第一阻断剂寡核苷酸具有与第二阻断剂寡核苷酸的3'部分互补并杂交的3'部分。

[0049] 任选地,3'接头通过连接与杂交复合体中的阻断剂寡核苷酸共价结合,从而形成连接复合体,其中第一阻断剂寡核苷酸3'端与第一未连接的3'接头的5'端处的相邻的腺苷酸化核苷酸连接,且第二阻断剂寡核苷酸3'端与第二未连接的3'接头的5'端处的相邻的腺苷酸化核苷酸连接,其中第一阻断剂寡核苷酸,其具有与第一未连接的3'接头互补并杂交

的5'部分,第二阻断剂寡核苷酸,其具有与第二未连接的3'接头互补并杂交的5'部分,和其中第一阻断剂寡核苷酸具有与第二阻断剂寡核苷酸的3'部分互补并杂交的3'部分。

[0050] 在本文中可互换使用的术语"阻断剂寡核苷酸(blocker oligonucleotide)"和"寡核苷酸阻断剂(oligonucleotide blocker)"是指合成的寡核苷酸,其包括i)阻断剂寡核苷酸5'部分,长度至少为8个核苷酸,与5'端相邻,和ii)阻断剂寡核苷酸3'部分,长度至少为1至20个核苷酸,与3'端相邻。根据本公开内容的方面,阻断剂寡核苷酸具有在约9个核苷酸至约50个核苷酸的范围内的总长度。根据本公开内容的方面,阻断剂寡核苷酸具有总长度为约9个核苷酸、10个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸、30个核苷酸、31个核苷酸、32个核苷酸、33个核苷酸、34个核苷酸、35个核苷酸、36个核苷酸、37个核苷酸、38个核苷酸、39个核苷酸、40个核苷酸、41个核苷酸、42个核苷酸、43个核苷酸、44个核苷酸、45个核苷酸、46个核苷酸、47个核苷酸、48个核苷酸、49个核苷酸或约50个核苷酸,且可更短或更长。

[0051] 根据本公开内容的方面,阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸。

[0052] 根据本公开内容的方面,阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是Tm增加的核苷酸和/或核苷酸类似物。根据本公开内容的方面,阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是LNA单体、PNA单体、BNA单体、其他T_m增加的核苷酸类似物,或其任何两种或更多种。根据本公开内容的方面,阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是LNA单体。

[0053] 根据本公开内容的方面,阻断剂寡核苷酸5'部分的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是非天然存在的核苷酸。根据本公开内容的方面,阻断剂寡核苷酸5'部分的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是Tm增加的核苷酸和/或核苷酸类似物。根据本公开内容的方面,阻断剂寡核苷酸5'部分的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是LNA单体、PNA单体、BNA单体、其他T_m增加的核苷酸类似物,或其任何两种或更多种。根据本公开内容的方面,阻断剂寡核苷酸5'部分的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是LNA单体。

[0054] 根据本公开内容的方面,阻断剂寡核苷酸的3'部分是或包括回文核苷酸序列。

[0055] 根据本公开内容的方面的寡核苷酸是使用化学合成和/或重组分子生物学技术以合成方式产生,例如详细记载于Herdewijn,P.(Ed.),Oligonucleotide Synthesis:Methods and Applications,Methods in Molecular Biology,Humana Press,2004。

[0056] 示例性阻断剂寡核苷酸为5'-CACCCGAGAATTCCA--C/iisodG//iMe-isodC//iisodG//iMe-isodC/G-3'(SEQ ID NO:1)。

[0057] 图1图解说明了第一阻断剂寡核苷酸20和第二阻断剂寡核苷酸40。第一阻断剂寡核苷酸20具有5'部分30和3'部分35。第一阻断剂寡核苷酸的3'端示于28。第二阻断剂寡核

苷酸具有5’部分55和5’部分50。第二阻断剂寡核苷酸3’端示于48。尽管第一和第二阻断剂寡核苷酸显示为长度与3’接头60及其末端62相等,但第一和第二阻断剂寡核苷酸的长度可短于或长于3’接头60及其末端62。

[0058] 图1示出了杂交复合体,其中第一阻断剂寡核苷酸3’部分35与第二阻断剂寡核苷酸3’部分50杂交。进一步示出的是各自具有5’腺苷酸化末端62的3’接头60。示出了3’接头60与第一阻断剂寡核苷酸30的5’部分和第二阻断剂寡核苷酸的5’部分杂交。还示出的是与第一阻断剂寡核苷酸的3’端28相邻的3’接头的5’腺苷酸化末端62和与第二阻断剂寡核苷酸的3’端48相邻的3’接头的5’腺苷酸化末端62。

[0059] 任选地,进行连接反应,使得在连接条件下和连接酶的存在下,发生3’接头的5’腺苷酸化末端62与第一阻断剂寡核苷酸的相邻3’端28的连接;且在连接条件下和连接酶的存在下,发生3’接头的5’腺苷酸化末端62与第二阻断剂寡核苷酸48的相邻3’端的连接,形成连接复合体。

[0060] 根据本公开内容的方面,第一和第二阻断剂寡核苷酸的序列相同。

[0061] 根据本公开内容的方面,第一和第二阻断剂寡核苷酸的序列不同。

[0062] 本公开内容的阻断剂寡核苷酸的5’部分与3’接头互补,使得阻断剂寡核苷酸的5’部分在杂交条件下与3’接头杂交。阻断剂寡核苷酸的长度可以长于、短于或与3’接头相同。阻断剂寡核苷酸的5’部分可以与3’接头完全互补,或者使用适于形成杂交复合体和随后连接以形成连接复合体的杂交条件,可以为50%、60%、70%、80%、90%或更高%的互补。

[0063] 本文所述的方法、阻断剂寡核苷酸和试剂盒可用于从源自多种来源中的任一种的样品核酸产生测序文库。所述样品核酸,例如DNA或RNA,通常存在于生物样品中,其可以从个体获得,例如从身体样品获得,例如血液、淋巴、口腔拭子、皮肤组织、尿液、唾液、组织、痰、腹膜液、粪便物质、胃液、精液、脑脊液、活检物材料、伤口渗出物等,及由其衍生的细胞系。产前生物样品核酸可以从羊水、受孕产物、胚泡和卵裂球、绒毛膜绒毛,母体血液中循环的胎儿细胞、胎儿DNA和/或胎儿RNA获得。所述样品核酸也可以从体外来源例如细胞系获得。

[0064] 本文所述的方法、阻断剂寡核苷酸和试剂盒可用于从多种生物体中的任一种获得的样品核酸产生测序文库,所述生物体包括但不限于人、非人灵长类动物、啮齿动物、兔、狗、猫、马、牛、猪、山羊和绵羊。也可以使用样品核酸的非哺乳动物来源,示例性地包括鱼和其他水生生物、鸟、家禽、细菌、病毒、植物、昆虫、爬行动物、两栖动物、真菌和分枝杆菌。因此,所述样品核酸,例如样品DNA或样品RNA,可从这些来源中的任一种获得。

[0065] 所述样品核酸,例如DNA或RNA,可以从任何来源获得,包括但不限于人、非人哺乳动物、脊椎动物、无脊椎动物、微生物或植物。所述样品核酸,例如DNA或RNA,可以从一种或多种离体或体外细胞获得。例如,所述样品核酸,例如DNA或RNA,可以从培养的细胞获得,包括但不限于细胞系、原代细胞或实验室操纵的细胞(laboratory manipulated cell),例如重组细胞。

[0066] 根据本公开内容的方面,样品核酸为“小RNA(small RNA)”。如本文所使用的,术语“小RNA(small RNA)”是指长度在约10个核苷酸长度至约300个核苷酸长度范围内的RNA,例如长度在约10个核苷酸长度至约250个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约200个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约150个核苷酸长度范围

内的RNA,长度在约10个核苷酸长度至约100个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约90个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约80个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约70个核苷酸长度范围内的RNA,长度在约10个核苷酸长度至约60个核苷酸长度范围内的RNA、长度在约10个核苷酸长度至约50个核苷酸长度范围内的RNA、长度在约10个核苷酸长度至约40个核苷酸长度范围内的RNA或长度在约10个核苷酸长度至约30个核苷酸长度范围内的RNA。小RNA可为以下中的任一种:RNA片段、rRNA、病毒RNA、微小RNA(miRNA)、小干扰RNA(siRNA)、转移RNA(tRNA)、短发夹RNA(shRNA)、重复序列相关RNA(rasiRNA)、核仁小RNA(snoRNA)、Piwi相互作用RNA(piRNA)、小卡哈尔体RNA(small cajal body rna,scaRNA)、异染色质siRNA(hc-siRNA)和反式作用RNA(ta-siRNA)。

[0067] 如本文所使用的,术语“测序文库(sequencing library)”是指多个线性双链DNA片段,其包括插入片段,所述插入片段为来源于样品核酸的目的测序靶标,所述样品核酸可源自或可不源自相同的来源。测序文库的DNA片段进一步包括插入片段侧翼的功能元件,其提供用于如DNA片段的鉴定(例如条形码或索引序列),扩增引物的结合,测序引物的结合,或与所需靶标的特异性结合,从而允许在测序方法(特别是大规模平行测序方法)中进行有效处理。测序文库可包括侧翼为两个接头(5’接头和3’接头)的插入片段,其中5’接头直接或间接地共价结合至插入片段的5’末端,且3’接头直接或间接地共价结合至插入片段的3’末端。接头与插入片段的间接结合是指其中另一DNA序列位于插入片段与接头之间的构型。

[0068] 术语“大规模平行测序(massively parallel sequencing)”,包括“合成测序(sequencing by synthesis)”,也称为“下一代测序(next generation sequencing)”和“高通量测序(high throughput sequencing)”,表示对许多核酸测序文库分子(例如数千个核酸测序文库分子)同时或接近同时测序的方法。根据本公开内容的方面提供的寡核苷酸阻断剂组记载于本文中,在一些情况下特别参考了Illumina Nextera和/或TruSeq的平台和配置,但不限于此。更确切地说,本公开内容的组合物和方法可用于多种测序方案中的任一种且与多种测序设备一起使用,包括但不限于半导体测序(Ion Torrent)系统。

[0069] 术语“接头(adapter)”是指线性、单链、双链或部分双链的DNA片段,其包括提供用于例如DNA片段的鉴定、扩增或与所需靶标的特异性结合的功能元件,从而允许在测序方法中进行有效处理。接头包括一个或多个功能元件,例如测序引物结合序列(也称为测序引物结合位点)、延伸引物结合序列(也称为延伸引物结合位点)和索引序列(也称为条形码序列或独特分子标识符,UMI)。术语“5’接头(5’adapter)”是指直接或间接与插入片段的5’末端共价结合或旨在与插入片段的5’末端共价结合的接头。术语“3’接头(3’adapter)”是指直接或间接与插入片段的3’末端共价结合或旨在与插入片段的3’末端共价结合的接头。接头的长度通常为约10至150个核苷酸,例如约10至25、15至30、25至40、30至50、45至75、50至100、90至125或120至150个核苷酸。应注意,术语“接头(adapter)”可选地拼写为“接头(adaptor)”。

[0070] 根据本公开内容的方面,3’接头在其5’末端处被5’腺苷酸化,以促进3’接头与样品核酸和阻断剂寡核苷酸的连接。

[0071] 根据本公开内容的方面,包括在方法、组合物和/或试剂盒中的5’腺苷酸化3’接头在3’末端被阻断,例如被2’,3’-双脱氧核苷酸或3’(C3)丙基间隔基阻断。

[0072] 如本文所使用的关于核酸分子的术语“末端(end)”是指线性核酸分子的末端,并且在本文中可与术语“端(terminus)”或语法上适当的变体如“端(termini)”互换使用。如本文所使用的关于核酸分子的术语“末端部分(end portion)”是指涵盖核酸分子的末端和相邻的连续核酸序列的核酸序列,但不包括核酸分子的核酸序列全长。根据本公开内容的方面,核酸分子的“末端部分(end portion)”是指涵盖核酸分子的末端和构成核酸分子的整个核酸序列长度的约一半或更少的相邻的连续核酸序列的核酸序列。

[0073] 术语“接头-二聚体(adapter-dimer)”是指在产生测序文库的方法中当5’接头和3’接头连接至样品核酸的5’端和3’端时形成的不合需要的副产物。具体而言,接头-二聚体是3’接头的腺苷酸化5’端和5’接头的3’端的连接产物。接头-二聚体在产生测序文库的方法中是不合需要的,因为它们在产生测序文库的方法的方面中可被扩增并与测序文库分子竞争,例如竞争流动池中的结合位点,从而显著降低测序质量。

[0074] 如本文所使用的,术语“连接(ligation)”是指由连接酶介导的两个线性核酸分子的共价键合以形成更长的线性核酸分子。

[0075] 术语“连接酶(ligase enzyme)”和“连接酶(ligase)”在本文中可互换使用,是指能够共价键合两个相邻核酸的酶,在两个核酸之间形成键,例如磷酸二酯键,或取决于末端核苷酸的特性,修饰的共价键,例如但不限于硫代磷酸酯键或磷酸酰胺酯键。连接酶包括但不限于T7连接酶、T4连接酶、大肠杆菌连接酶和Taq连接酶。连接酶包括ATP依赖性连接酶、NAD+依赖性连接酶,且包括DNA连接酶和RNA连接酶,例如由国际生物化学和分子生物学联盟(The International Union of Biochemistry and Molecular Biology)定义的酶类EC 6.5的成员,包括但不限于EC 6.5.1.1(ATP依赖性连接酶)、EC 6.5.1.2(NAD+依赖性连接酶)、EC 6.5.1.3(RNA连接酶)、EC 6.5.1.6DNA连接酶(ATP或NAD+)、EC 6.5.1.7DNA连接酶(ATP、ADP或GTP)和EC 6.5.1.8 3’-磷酸/5’-羟基核酸连接酶。

[0076] 使用连接酶的条件是众所周知的。

[0077] 核苷酸、核苷酸类似物和非天然存在的核苷酸

[0078] 术语“核苷酸(nucleotide)”是指包括核碱基、糖和磷酸的分子。核苷酸是核酸序列(如DNA或RNA序列)的单体单元。术语核苷酸包括核糖核苷三磷酸,例如ATP、TTP、UTP、CTP和GTP,以及脱氧核糖核苷三磷酸,例如dATP、dCTP、dUTP、dGTP和dTTP。关于核碱基,核苷酸通常被称为A、T、G、C或U作为缩写,参见表I。

[0079] 本文所使用的标准IUPAC核苷酸代码,如表I所示。

[0080]

表I

符号	描述	代表的碱基					互补物
A	腺嘌呤	A					T
C	胞嘧啶		C				G
G	鸟嘌呤			G			C

[0081]

表I

符号	描述	代表的碱基				互补物
T	胸腺嘧啶				T	A
U	尿嘧啶				U	A
W	弱	A			T	W
S	强		C	G		S
M	氨基	A	C			K
K	酮			G	T	M
R	嘌呤	A		G		Y
Y	嘧啶		C		T	R
B	非A		C	G	T	V
D	非C	A		G	T	H
H	非G	A	C		T	D
V	非T	A	C	G		B
N	任何核苷酸（非间隔）	A	C	G	T/U	N

[0082] 术语“核酸(nucleic acid)”、“核苷酸序列(nucleotide sequence)”和“核酸序列(nucleic acid sequence)”是指具有多于一个核苷酸和/或任何形式的核苷酸类似物的RNA、DNA或RNA/DNA分子,包括线性、单链、双链、寡核苷酸或多核苷酸。

[0083] 术语“核苷酸类似物(nucleotide analog)”是指修饰的或非天然存在的核苷酸,特别地为核苷酸类似物,所述核苷酸类似物可以通过由核酸聚合酶催化的模板指导的核酸聚合或非模板指导的核酸聚合,用天然存在的核苷酸和/或非天然存在的核苷酸来聚合。术语“核苷酸类似物(nucleotide analog)”是指含有对碱基部分和/或糖部分和/或磷酸部分的一种或多种修饰的核苷酸,与参考标准核苷酸相比,所述修饰修饰了核苷酸类似物的化学性质的至少一个方面,同时保留了允许核苷酸类似物执行其预期功能的其他性质。核苷酸类似物及其掺入核酸在本领域中是公知的,且可以根据标准程序合成和/或商购获得。

[0084] 特定的核苷酸类似物能够通过氢键与互补核苷酸进行Watson-Crick配对,且举例说明性地包括但不限于天然存在的和非天然存在的核苷酸及其类似物,包括那些含有核苷酸碱基的类似物,例如取代的嘌呤或嘧啶、脱氮嘌呤、甲基嘌呤、甲基嘧啶、氨基嘌呤、氨基嘧啶、硫代嘌呤、硫代嘧啶、吲哚、吡咯、7-脱氮鸟嘌呤、7-脱氮腺嘌呤、7-甲基鸟嘌呤、次黄嘌呤、假胞嘧啶、假异胞嘧啶、异胞嘧啶、异鸟嘌呤、2-硫代嘧啶、4-硫代胸腺嘧啶、6-硫代鸟嘌呤、硝基吡咯、硝基吲哚和4-甲基吲哚。核苷酸类似物包括含有脱氧核糖的那些,例如取代的脱氧核糖、取代或未取代的阿拉伯糖、取代或未取代的木糖及取代或未取代的吡喃糖。核苷酸类似物包括含有磷酸酯的类似物的那些,例如硫代磷酸酯、二硫代磷酸酯、磷酸酰胺酯(phosphoroamidate)、硒代磷酸酯(phosphoroselenoate)、苯胺硫代磷酸酯(phosphoroanilothioate)、苯胺磷酰胺酯(phosphoroanilidate)、磷酸酰胺酯、硼代磷酸酯(boronophosphate)、磷酸三酯和烷基膦酸酯,例如甲基膦酸酯。

[0085] 非天然存在的核苷酸(例如iso-dG和iso-dC)提供彼此独特的碱基配对,且在非天然存在的核苷酸与天然存在的核苷酸碱基之间没有碱基配对(或杂交),如iso-dG不会与dC配对,且iso-dC不会与dG配对。术语“iso-dG”是指异鸟苷。值得注意的是,术语“iisodG”在本文中也用于指异鸟苷,如纳入序列中的异鸟苷。术语“iso-dC”是指异胞嘧啶。值得注意的是,术语“iMe-isodC”在本文中也用于指异胞嘧啶,如纳入序列中的异胞嘧啶。

[0086] 术语“T_m”和“解链温度”是指双链核酸分子群体的50%(一半)被分离(即单链)的温度。计算T_m的方法是本领域中众所周知的。

[0087] 术语“T_m增加的核苷酸类似物”是指与不含核苷酸类似物的相同双链寡核苷酸相比增加双链寡核苷酸的解链温度(T_m)的核苷酸类似物。T_m增加的核苷酸类似物可包括修饰的核碱基、修饰的糖、修饰的磷酸或任何两种或更多种此类修饰的组合。

[0088] T_m增加的核苷酸类似物包括但不限于锁核酸(LNA)单体、肽核酸(PNA)单体和桥接核酸(BNA)单体。

[0089] 根据本公开内容的方面,可以使用多种T_m增加的核苷酸类似物中的任何一种,包括但不限于LNA单体、PNA单体、BNA单体或其任何两种或更多种。

[0090] 术语“LNA单体(LNA monomer)”在本文中与术语“LNA核苷酸(LNA nucleotide)”和“LNA碱基(LNAbase)”可互换使用,且是指核苷酸类似物,所述核苷酸类似物由于核糖环的2’氧和4’碳之间存在共价键而包括锁核糖,即2’-O,4’-C-亚甲基-β-D-呋喃核糖基单体,其可以纳入寡核苷酸中,产生“锁核酸(locked nucleic acid)”。与常规核苷酸相比,LNA单体具有相似或改进的Watson-Crick碱基配对选择性。LNA单体的实例包括但不限于腺嘌呤双环核苷单体、胞嘧啶双环核苷单体、鸟嘌呤双环核苷单体、5-甲基胞嘧啶双环核苷单体、胸腺嘧啶双环核苷单体和尿嘧啶双环核苷单体。包括一个或多个LNA单体的寡核苷酸的Tm由于增强的性质(例如增强的碱基堆叠)而增加。通常,纳入LNA寡核苷酸中的每个LNA单体使a的Tm增加2至8℃。LNA单体可使用众所周知的方法合成或商购获得。术语“锁核酸(locked nucleic acid)”和LNA在本文中可互换使用,是指包括一个或多个LNA单体的寡核苷酸。可以使用多种T_m增加的核苷酸类似物中的任何一种,包括但不限于,5-甲基dC、2,6-二氨基嘌呤、丙炔基-脱氧尿苷、5-羟基丁炔基-2’-脱氧尿苷以及其任意两种或更多种的组合。可以包括任何Tm增加的核苷酸类似物和/或其衍生物,只要与不含核苷酸类似物的相同双链寡核苷酸相比衍生物保留增加包括核苷酸类似物的双链寡核苷酸的解链温度(T_m)的功能。

[0091] 根据本公开内容的方面,阻断剂寡核苷酸具有55℃至100℃范围内的T_m,例如60℃至100℃,例如65℃至100℃,例如70℃至100℃,例如75℃至99℃,例如75℃至90℃,例如75℃至85℃,例如55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99或100℃。

[0092] 根据本公开内容的方面,一组阻断剂寡核苷酸具有55℃至100℃范围内的T_m,例如60℃至100℃,例如65℃至100℃,例如70℃至100℃,例如75℃至99℃,例如75℃至90℃,例如75℃至85℃,例如55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99或100℃。

[0093] 根据本公开内容的方面,例如,在杂交后捕获PCR步骤中,任选地修饰一种或多种

阻断剂寡核苷酸的3'末端以防止从3'末端的聚合酶延伸。可以使用用于修饰一种或多种阻断剂寡核苷酸的3'末端以防止从3'末端的聚合酶延伸的多种技术中的任一种,包括但不限于去除或修饰寡核苷酸的3'末端的羟基,例如包括3'双脱氧-C(ddC)、磷酸基团或3'间隔基。根据本公开内容的方面,间隔基存在于阻断剂寡核苷酸的3'末端。一个非限制性实施例是3'间隔基C3(缩写为3SpC3),其为连接至寡核苷酸阻断剂的末端3'羟基的3碳链(C3)。或者,阻断剂寡核苷酸的3'末端未被修饰以防止从3'末端的聚合酶延伸,例如如果不使用杂交后捕获PCR步骤或杂交后捕获的其他聚合酶介导的延伸步骤(other polymerase-mediated extension step post-hybridization capture)。

[0094] 术语"杂交(hybridization)"和"杂交的(hybridized)"是指互补核酸的配对和结合。在两种核酸之间发生不同程度的杂交,这取决于本领域中众所周知的因素,例如核酸的互补程度、核酸的解链温度T_m和杂交条件的严格程度和杂交后洗涤条件。术语"杂交条件的严格性(stringency ofhybridization conditions)"是指相对于特定常见添加剂(例如甲酰胺、甜菜碱和聚乙二醇),温度、离子强度和杂交介质的组成的条件。术语"杂交后洗涤条件的严格性(stringency of post-hybridization wash conditions)"是指温度、离子强度和杂交介质的组成的条件,尽管添加剂(例如表面活性剂、甲酰胺、甜菜碱和聚乙二醇)通常不用于洗涤条件。可以使用的特定表面活性剂包括,例如SDS(十二烷基硫酸钠(sodium dodecyl sulfate)),月桂酰肌氨酸钠(sodium lauroyl sarcosinate)(肌氨酰(Sarkosyl))和CTAB(十六烷基三甲基溴化铵(cetyltrimethylammonium bromide))。与特定核酸相关的特定杂交条件和杂交后洗涤条件的确定是常规的,且是本领域中众所周知的,例如,记载于J.Sambrook and D.W.Russell,Molecular Cloning:ALaboratory Manual,Cold Spring Harbor Laboratory Press;3rd Ed.,2001;和F.M.Ausubel,Ed.,Short Protocols in Molecular Biology,Current Protocols;5th Ed.,2002。高严格性杂交条件为仅允许实质性互补的核酸杂交的那些条件。高严格性杂交后洗涤条件为仅维持实质性互补的核酸杂交的那些条件。如本文所使用的,术语"退火(annealing)"是指寡核苷酸与靶核酸的配对和结合。术语"退火(annealing)"和"杂交(hybridizing)"之间没有刻意的区别。

[0095] 如本文所使用的,术语"互补(complementary)"涵盖但不限于核苷酸之间、核苷酸类似物之间或核苷酸和核苷酸类似物之间的Watson-Crick碱基配对,其中核苷酸和/或核苷酸类似物彼此氢键键合,例如通过两个氢键连接至腺嘌呤残基的胸腺嘧啶或尿嘧啶残基,通过三个氢键连接的胞嘧啶和鸟嘌呤残基。通常,核酸包括核苷酸序列,其被描述为与指定的第二核苷酸序列具有"百分比互补性(percent complementarity)"。例如,核苷酸序列可与特定的第二核苷酸序列具有80%、90%或100%互补性,表明序列的10个核苷酸中的8个、10个核苷酸中的9个或10核苷酸个中的10个与指定的第二核苷酸序列互补。例如,核苷酸序列3'-TCGA-5'与核苷酸序列5'-AGCT-3'100%互补。此外,核苷酸序列3'-TCGA-5'与核苷酸序列5'-TTAGCTGG-3'的区域100%或完全互补。如本文所使用的术语"互补序列(complement)"是指与给定核苷酸序列互补的核苷酸序列。

[0096] 通常,认为具有约85-100%互补性的核酸是高度互补的且在高严格条件下杂交。中等严格条件的实例是具有中等互补性(约50-84%互补性)的核酸以及具有高度互补性的核酸杂交的条件。相反,低严格性杂交条件为具有低度互补性的核酸杂交的条件。术语"特

异性杂交(specific hybridization)”和“特异性杂交(specifically hybridizes)”是指特定核酸与靶核酸的杂交,而没有与样品中非靶核酸的核酸的实质性杂交。

[0097] 术语“回文核苷酸序列(palindromic nucleotide sequence)”是指为反向重复的核苷酸序列,使得核苷酸序列无论在5’至3’方向上读取还是在互补链上从5’至3’读取时都是相同的。例如,从5’至3’读取的回文核苷酸序列是5’-GGATCC-3’,其在互补链上从5’至3’读取时是相同的,即3’-CCTAGG-5’。

[0098] 阻断剂寡核苷酸与过量接头的退火是在退火温度下进行。退火温度取决于多种因素,包括阻断剂寡核苷酸的核酸序列和靶核酸,以及反应介质的组成,包括多种因素,如盐浓度和添加剂的浓度,所述添加剂例如但不限于甲酰胺、甜菜碱、聚乙二醇、表面活性剂(例如SDS)和DMSO。通常,退火温度是在30℃至65℃的范围内,但可以更高或更低。退火温度可以高于或低于阻断剂寡核苷酸解链温度(T_m),然而,通常根据本公开内容的方面的阻断剂寡核苷酸具有高于退火温度的T_m以确保阻断剂寡核苷酸在退火期间与其靶标结合,即杂交。类似地,一次或多次杂交后洗涤可以在低于阻断剂寡核苷酸的T_m的温度下进行。

[0099] 在阻断剂寡核苷酸与过量接头退火后,产生杂交复合体,其中第一阻断剂寡核苷酸3’端与第一未连接的3’接头的腺苷酸化5’末端核苷酸相邻,第二阻断剂寡核苷酸3’端与第二未连接的3’接头的腺苷酸化5’末端核苷酸相邻,从而减少过量的未连接的3’接头。

[0100] 任选地,杂交复合体的第一阻断剂寡核苷酸3’端与第一未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸连接,且杂交复合体的第二阻断剂寡核苷酸3’端与第二未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸连接,产生连接复合体,从而减少过量的未连接的3’接头。

[0101] 可以将杂交复合体或连接复合体与所需的测序文库分子分离。杂交复合体或连接复合体与所需的测序文库分子的分离可为基于尺寸的分离,例如通过基于柱的分离、基于过滤器的分离、基于凝胶的分离、连接复合体或所需的测序文库分子与固体或半固体载体的结合,然后未结合的材料与结合的材料的分离。根据本公开内容的方面,所述固体载体包括磁珠。

[0102] 使用体外扩增方法实现模板DNA的扩增。术语“扩增(amplify)”、“扩增(amplification)”和“扩增(amplifying)”通常用于指用于复制模板核酸从而产生包括模板核酸的全部或一部分的拷贝的核酸的方法或技术,所产生的核酸也称为扩增子。本发明的实施方案中包括的扩增方法为包括由核酸聚合酶催化的模板指导的引物延伸,使用一对位于靶核酸侧翼的引物,举例说明性地包括但不限于聚合酶链式反应(PCR)、逆转录PCR(RT-PCR)。连接介导的PCR(LM-PCR)、phi-29PCR、实时定量PCR(qPCR)、全基因组扩增和其他核酸扩增方法,例如,如C.W.Dieffenbach et al.,PCR Primer:ALaboratory Manual,Cold Spring Harbor Laboratory Press,2003;V.Demidov et al.,DNA Amplification:Current Technologies and Applications,Taylor&Francis,2004;和Kroneis,T.(Ed.),Whole Genome Amplification:Methods and Protocols(Methods in Molecular Biology),2015,Humana Press ISBN-10:1493929895中所记载的。

[0103] 术语“等温扩增(isothermal amplification)”是指核酸扩增,其包括单一温度扩增,并因此避免了对热循环的需求(如在PCR中)。等温扩增的实例包括但不限于环状解旋酶依赖性扩增(cHDA)、基因组指数扩增反应(GEAR)、解旋酶依赖性扩增(HDA)、等温多重置换

扩增(IMDA)、环介导等温扩增(LAMP)、多重置换扩增(MDA)、切口酶扩增反应(NEAR)、基于核酸序列的扩增(NASBA)、分支扩增(ramification amplification,RAM)、重组酶聚合酶扩增(RPA)、滚环扩增(RCA)、自主序列复制(3SR)、信号介导的RNA技术扩增(SMART)、链置换扩增(SDA)、单引物等温扩增(SPIA)和转录介导的扩增(TMA)。

[0104] 术语“扩增的核酸(amplified nucleic acid)”、“扩增的DNA(amplified DNA)”和“扩增子(amplicon)”及其复数是指复制靶核酸模板的方法的产物。

[0105] 扩增的核酸可选地含有额外的材料,例如但不限于核酸序列、用于化学反应的官能团和可检测的标记,其存在于引物中而不存在于原始核酸模板中。可以纳入扩增子中且从而纳入所产生的文库中的引物衍生的核酸序列的非限制性实施例包括通用序列、接头、索引序列、识别序列、检测序列、分选序列、捕获序列、限制性酶切位点、测序引物结合位点序列和扩增引物结合位点序列。

[0106] 术语“通用序列(universal sequence)”是指存在于多个核酸分子中的核酸序列,所述核酸分子还含有不是所述多个核酸分子共有的核酸序列。通用序列允许多个核酸分子共享共同的功能方面,例如与特定引物或捕获部分的结合。通用延伸引物结合位点的非限制性实例包括与P5和P7引物相同或互补的序列。P5和P7引物、它们的互补物和用途是本领域已知的(例如在用于下一代测序(NGS)的流动池中用于捕获到流动池底物上),例如记载于WO2015106941中。

[0107] 根据本公开内容的方面,在文库制备期间将索引序列纳入扩增子中。索引序列是一组扩增子共有的独特核酸序列,如用于鉴定该组扩增子为源于特定来源。索引序列允许多重处理(multiplexing),因为可以汇合来自不同来源的多个核酸用于测序,且如果需要的话,可以稍后“解复用(demultiplexed)”用于数据分析。

[0108] 术语“引物(primer)”是指寡核苷酸,其能够在诱导与模板核酸互补的寡核苷酸引物延伸产物的合成的条件下充当模板指导的引物延伸产物的合成起始位点。这种条件包括在合适的温度和pH下存在核苷酸和/或核苷酸类似物及合适的聚合酶。

[0109] 引物的长度通常为约10-30个连续的核苷酸,且可更长或更短。寡核苷酸引物与模板核酸的区域完全或实质性互补,使得在杂交条件下,寡核苷酸引物与模板核酸的互补区域退火。用于合成根据本公开内容的方面的引物延伸产物的合适的反应条件包括存在合适的反应组分,包括但不限于聚合酶;和核苷酸和/或核苷酸类似物。

[0110] 多种引物可用于根据本公开内容的方面的方法中,例如扩增引物、测序引物和用于逆转录的引物。

[0111] 引物设计是本领域技术人员所熟知的,且根据众所周知的方法和标准设计。例如,当使用一对引物时,例如在扩增反应中,引物的退火温度应该大致相同(在几度内)。通常,引物不应与彼此形成二聚体,且引物不应形成二级结构,例如发夹。用于引物设计和扩增程序的方法和考虑因素详细记载于Yuryev,A.,PCR Primer Design,Methods in Molecular Biology,vol.42,Human Press,2007;C.W.Dieffenbachet la.,PCR Primer:ALaboratory Manual,Cold Spring Harbor Laboratory Press,2003;和V.Demidov et al.,DNAAmplification:Current Technologies and Applications,Taylor&Francis,2004。

[0112] 产生第一链互补DNA的引物延伸是使用聚合酶在与产生互补链的聚合酶活性相容的延伸反应条件下完成。引物延伸的具体条件和方案详细记载于J.Sambrook and

D.W.Russell,Molecular Cloning:ALaboratory Manual,Cold Spring Harbor Laboratory Press;3rd Ed.,2001;和F.M.Ausubel,Ed.,Short Protocols in Molecular Biology,Current Protocols;5th Ed.,2002。

[0113] 根据本公开内容的方面,样品核酸是RNA,即样品RNA,且样品RNA被逆转录,产生cDNA。样品RNA可在加入接头之前或之后被逆转录。

[0114] 根据本公开内容的方法的方面,所述方法包括使逆转录引物与靶核酸退火,所述靶核酸包括RNA;和逆转录RNA,产生与RNA互补的第一链互补DNA(cDNA)。

[0115] 在退火温度下进行逆转录引物或逆转录引物组与靶核酸的退火。退火温度取决于多种因素,包括逆转录引物或引物的核酸序列和靶核酸,以及反应介质的组成,包括多种因素,例如盐浓度和添加剂的浓度,所述添加剂例如但不限于甲酰胺、甜菜碱、聚乙二醇、SDS和DMSO。通常,退火温度是在30℃至65℃的范围内,但可以更高或更低。退火温度可高于或低于逆转录引物解链温度(Tm)。通常退火温度高于逆转录引物的Tm以提高所述反应的严格性。

[0116] 根据本公开内容提供了产生测序文库的方法,其包括逆转录靶RNA,产生第一链互补DNA(cDNA),所述方法是使用逆转录酶在与转录靶RNA的逆转录酶活性相容的反应条件下完成。逆转录的具体条件和方案详细记载于J.Sambrook and D.W.Russell,Molecular Cloning:ALaboratory Manual,Cold Spring Harbor Laboratory Press;3rd Ed.,2001;和F.M.Ausubel Ed.,Short Protocols in Molecular Biology,Current Protocols;5th Ed.,2002。

[0117] 根据本公开内容提供了产生测序文库的方法,其进一步包括聚合与第一链cDNA互补的第二DNA链,产生双链cDNA。聚合与第一链cDNA互补的的第二DNA链可包括提供合适的DNA聚合酶并在与产生双链cDNA的DNA聚合酶活性相容的反应条件下聚合以产生双链cDNA。

[0118] 合适的聚合酶可包括细菌DNA聚合酶、真核30DNA聚合酶、古细菌DNA聚合酶、病毒DNA聚合酶、Taq聚合酶、DNA聚合酶I、T4 DNA聚合酶、Pfu聚合酶和噬菌体DNA聚合酶,包括DNA聚合酶I的Klenow片段、SEQUENASE 1.0和SEQUENASE 2.0(U.S.Biochemical)、T5 DNA聚合酶和Phi29 DNA聚合酶等。

[0119] 扩增子的测序是使用多种测序方法中的任一种完成,包括传统的Sanger测序和大规模平行测序方法(“下一代测序”)。

[0120] 有利地,进行配对末端读长。术语“配对末端读取(paired end reads)”是指包括来自待测序扩增子的每个末端的一个测序“读段”的测序技术。这允许更大的测序数据产量和增加测序结果的置信度。

[0121] 根据本公开内容的方面,通过形成杂交复合体或连接复合体减少但不消除过量的未连接的3’接头,使得将5’接头连接至文库插入片段的步骤还包括将5’接头连接至一些过量的未连接的3’接头,产生接头-二聚体,其中5’接头是通过在5’接头-3’接头界面处连接至3’接头而共价结合。因此,根据本公开内容的方面,减少测序文库中的接头-二聚体的方法进一步包括抑制扩增反应中接头-二聚体的扩增,使得与接头-二聚体相比,通过扩增反应选择性富集测序文库分子。

[0122] 根据本公开内容的方面,抑制接头-二聚体的扩增包括在扩增之前去除至少一些接头-二聚体。

[0123] 根据本公开内容的方面,在扩增之前去除至少一些接头-二聚体包括以下中的一种或多种:1)接头-二聚体的特异性酶促切割;2)接头-二聚体的基于CRISPR/Cas9的切割;3)凝胶纯化;4)基于SPRI-小珠的尺寸选择;和5)向混合物中加入接头-二聚体阻断剂寡核苷酸,所述接头-二聚体阻断剂寡核苷酸有效抑制接头-二聚体的扩增。

[0124] 接头-二聚体的特异性酶促切割是指使用限制性核酸内切酶将接头-二聚体切割成两个或更多个片段,从而减少接头-二聚体的数量。可以分析接头-二聚体的核苷酸序列以确定哪种限制性核酸内切酶将切割接头-二聚体,或者可以测定多种限制性核酸内切酶以确定哪种限制性核酸内切酶有效切割接头-二聚体,防止接头-二聚体的扩增。

[0125] 基于CRISPR/Cas的切割是指使用CRISPR/Cas系统将接头-二聚体切割成两个或更多个片段,从而减少接头-二聚体的数量。根据本公开内容的方面,接头-二聚体含有CRISPR茎环和Cas蛋白识别茎环,切割接头-二聚体,或以高亲和力结合茎环而不切割接头-二聚体,从而防止接头-二聚体的扩增。

[0126] 基于固相可逆固定化(Solid Phase Reversible Immobilization,SPRI)-小珠的尺寸选择可用于将接头-二聚体与混合物中的测序文库分子分离,或者通过将接头-二聚体固定在固体载体上并从混合物中去除具有固定的接头-二聚体的固体载体,或者通过将测序文库分子固定在固体载体上并从混合物中去除具有固定的测序文库分子的固体载体,随后从固体载体上洗脱测序文库分子。

[0127] 术语“接头-二聚体阻断剂寡核苷酸(adapter-dimer blocker oligonucleotide)”是指与含有扩增引物结合位点的接头-二聚体的至少一部分特异性杂交从而防止接头-二聚体扩增的寡核苷酸。接头-二聚体阻断剂寡核苷酸的长度在约10至约300个核苷酸的范围内,这取决于多种因素,例如接头-二聚体的大小和待使用的杂交条件。根据本公开内容的方面,接头-二聚体阻断剂寡核苷酸的长度在约10至约300个核苷酸长度、约10至约200个核苷酸长度、约10至约100个核苷酸长度、约10至约75个核苷酸长度、约10至约50个核苷酸长度、约10至约40个核苷酸长度、约10至约30个核苷酸长度或约10至约20个核苷酸长度范围内。根据本公开内容的方面,接头-二聚体阻断剂寡核苷酸的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是非天然存在的核苷酸。根据本公开内容的方面,接头-二聚体阻断剂寡核苷酸的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是T_m增加的核苷酸和/或核苷酸类似物。根据本公开内容的方面,接头-二聚体阻断剂寡核苷酸的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是LNA单体、PNA单体、BNA单体、其他T_m增加的核苷酸类似物,或其任何两种或更多种。根据本公开内容的方面,接头-二聚体阻断剂寡核苷酸的约10%、20%、30%、40%、50%、60%、70%、80%、90%或更多的核苷酸是LNA单体。

[0128] 接头-二聚体阻断剂寡核苷酸的实例为5'-C+G+AG+AA+T+T+CC+A+GA+T+CG+T+C+G+GA+C+T+G-3'(SEQ ID NO:2),其中+表示LNA碱基(即纳入的LNA单体)。在优选的实例中,3SpC3(3'(C3)丙基间隔基)存在于3'末端上。

[0129] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的试剂盒,其包括用于减少测序文库中的接头-二聚体的一个或多个阻断剂寡核苷酸和/或阻断剂寡核苷酸对。

[0130] 根据本公开内容的方面,用于减少测序文库中的接头-二聚体的试剂盒进一步包

括有效阻断接头-二聚体扩增的一个或多个接头-二聚体阻断剂寡核苷酸。

[0131] 根据本公开内容的方面,提供了用于减少测序文库中的接头-二聚体的试剂盒,其包括与接头-二聚体的一个或多个相邻部分互补的接头-二聚体阻断剂寡核苷酸,所述一个或多个相邻部分包括在5'接头-3'接头界面处的5'接头的至少一个核苷酸,和包括扩增引物结合位点的3'接头的相邻部分。

[0132] 根据本公开内容的方面的试剂盒包括水性的、缓冲的或冻干形式的一个或多个阻断剂寡核苷酸和/或一个或多个接头-二聚体阻断剂寡核苷酸。所述一个或多个阻断剂寡核苷酸和/或一个或多个接头-二聚体阻断剂寡核苷酸可一起或分开包装。可将一个或多个阻断剂寡核苷酸和/或一个或多个接头-二聚体阻断剂寡核苷酸等分成适合用于测序文库制备的量,或可大量包装(packaged in bulk)以供用户进行等分。额外的试剂盒组分可包括但不限于缓冲液、稀释剂、稳定剂、防腐剂、试剂例如核苷酸、核苷酸类似物和酶。

[0133] 根据本公开内容的方面的组合物和方法在许多应用中具有效用,包括但不限于减少在测序方案中测序的接头-二聚体,特别地在包括小插入片段的测序文库中。

[0134] 根据本公开内容的方面的组合物和方法提供了高灵敏度,使得可以更频繁地检测样品中存在的稀有核酸。稀有核酸的可靠检测允许进行分析患者样品中的核酸的诊断和治疗方法。

[0135] 本发明的组合物和方法的实施方案在以下实施例中举例说明。提供这些实施例是为了举例说明的目的,且不被认为是对本发明组合物和方法的范围的限制。

[0136] 实施例

[0137] 实施例1

[0138] 3'接头连接

[0139] a.在冰上组合以下成分:

[0140] i.1μL输入RNA(1ng/μL)

[0141] ii.9.5μL无核酸酶水

[0142] iii.1μL 5uM 3'接头C3/5rApp/TGGAATTCTCGGGTGCCAAGG/3SpC3/(SEQ ID NO:3)

[0143] iv.7μL 3'连接缓冲液

[0144] v.1.5μL 3'连接酶混合物

[0145] vi.20μL总反应体积

[0146] b.孵育如下:

[0147] i.25℃下1小时

[0148] ii.保持在4℃下

[0149] 未连接的3'接头的失活

[0150] a.在冰上,向反应物中加入具有回文3'末端的1μL 10μM寡核苷酸阻断剂,并混合至均匀5'-CACCCGAGAATTCCAC/iisodG//iMe-isodC//iisodG//iMe-isodC/G-3'(SEQ ID NO:1)。注意,在本实施例中,寡核苷酸阻断剂的回文3'末端允许两个寡核苷酸阻断剂分子的3'部分彼此杂交,且阻断剂寡核苷酸的5'部分在杂交条件下与过量的3'接头杂交,形成杂交复合体。

[0151] b.孵育如下:

[0152] i.70℃下10分钟

[0153] ii.保持在4℃下

[0154] 5’接头连接

[0155] a.在冰上组合以下成分：

[0156] i.21μL源自先前步骤的反应物

[0157] ii.1.5μL 2.5uM 5’接头5’-rGrUrUrCrArGrArGrUrUrCrUrArCrArGrUrCrCrGrArCrGrArUrC-3’(SEQ ID NO:4)-注意,这里使用“r”来强调所示出的核苷酸为核糖核苷酸。

[0158] iii.1.5μL 10mM ATP(NEB cat#P0756L)

[0159] iv.2μL 5’连接酶混合物

[0160] v.26μL总反应体积

[0161] b.孵育如下：

[0162] i.20℃下1小时

[0163] ii.保持在4℃下

[0164] 逆转录

[0165] a.在热循环仪上使用来自先前步骤的26μL 5’连接反应物开始。

[0166] b.编程热循环仪如下：

[0167] i.70℃下2分钟

[0168] ii.50℃下30分钟

[0169] iii.90℃下5分钟

[0170] iv.保持在4℃下

[0171] c.制备RT主混合物(master mix)如下：

[0172] i.13μL RT缓冲液

[0173] ii.1μL 25uM RT引物5’-CCTTGGCACCCGAGAATTCCA-3’(SEQ ID NO:5)

[0174] iii.2μL EnzScript逆转录酶(Enzymatics)

[0175] iv.16μL总体积。

[0176] d.在热循环仪上使用26μL来自上述5’接头连接部分中a的混合物,启动程序。

[0177] e.在热循环仪逐渐下降到50℃后、加入16μL RT主混合物并在热循环仪上混合。不要从热循环仪中移出反应物。

[0178] f.使孵育完成并逐渐下降至4℃。

[0179] 小珠纯化(Bead cleanup)

[0180] a.加入并混合40μL接头耗尽溶液

[0181] b.加入并混合40μLNEXTFLEX Cleanup小珠

[0182] c.加入并混合90μL异丙醇

[0183] d.在室温下孵育5分钟

[0184] e.置于磁性支架(magnetic stand)上5分钟或直至溶液澄清

[0185] f.抽吸去除上清液

[0186] g.通过加入200μL 80％乙醇洗涤,孵育30秒,并抽吸去除上清液

[0187] h.重复上述步骤,共洗涤2次

[0188] i.在磁性支架上干燥共计3分钟。1分钟后,抽吸并去除收集在孔底部处的任何液体

[0189] j.从磁性支架中取出并重悬于40μL无核酸酶水中

[0190] k.孵育2分钟

[0191] l.放置在磁性支架上

[0192] m.一旦小珠形成沉淀,取18μL并加入试管(strip tube)中。这个纯净的cDNA保存作为备份。

[0193] n.再取18μL并移至新孔中。该cDNA将在以下步骤中扩增。

[0194] PCR

[0195] a.在冰上组合以下成分:

[0196] i.18μL源自先前步骤5n的cDNA

[0197] ii.2μL NEXTFLEX UDI Barcoded引物混合物

[0198] iii.5μL NEXTFLEX PCR主混合物

[0199] iv.1μL 10μM长dblock02C3(或水)5'-C+G+AG+AA+T+T+CC+A+GA+T+CG+T+C+G+GA+C+T+G/3SpC3/(SEQ ID NO:2)

[0200] v.26μL总反应体积

[0201] b.孵育反应物如下:

[0202] i.98℃下30秒

[0203] ii.98℃下10秒

[0204] iii.65℃下20秒℃重复共22个循环

[0205] iv.72℃下15秒

[0206] v.72℃下2分钟

[0207] vi.保持在4℃下

[0208] 小珠纯化

[0209] a.取25μL源自先前步骤的PCR反应物

[0210] b.加入并混合48.6μLNEXTFLEX Cleanup小珠

[0211] c.在室温下孵育5分钟

[0212] d.置于磁性支架上5分钟或直至溶液澄清

[0213] e.抽吸去除上清液

[0214] f.通过加入200μL 80%乙醇洗涤,孵育30秒,并抽吸去除上清液

[0215] g.重复上述步骤,共洗涤2次

[0216] h.在磁性支架上干燥共计3分钟。1分钟后,抽吸并去除收集在孔底部处的任何液体

[0217] i.从磁性支架中取出并重悬于20μL无核酸酶水中

[0218] j.孵育2分钟

[0219] k.放置在磁性支架上

[0220] l.一旦小珠形成沉淀,移出18μL。这是你的测序文库。

[0221] 测序参数(MiSeq)

[0222] a.2%phiX

[0223] b.1x 76

[0224] 结果如图3所示。

[0225] 实施例2

[0226] 1.3’接头连接

[0227] a.在冰上组合以下成分：

[0228] i.“X”μL输入RNA(1ng总RNA)

[0229] ii.1μL tRNA/YRNA阻断剂

[0230] iii.“Y”μL无核酸酶水

[0231] iv.1μL 3’腺苷酸化接头5’-rApp/TGGAATTCTCGGGTGCCAAGG/3SpC3/-3’(SEQ ID NO:3)

[0232] v.12.5μL 3’连接缓冲液

[0233] vi.1.5μL 3’连接酶混合物

[0234] vii.20μL反应总体积(水和RNA应等于4μL总体积)

[0235] viii.通过涡旋至少3秒混合粘性反应物直至均匀。

[0236] b.孵育如下：

[0237] i.25℃下1小时

[0238] ii.保持在4℃下

[0239] 2.未连接的3’接头的失活

[0240] a.在冰上,向反应物中加入4μL接头失活混合物,包括具有回文3’末端的寡核苷酸阻断剂:5’-CACCCGAGAATTCCAC/iisodG//iMe-isodC//iisodG//iMe-isodC/G-3’(SEQ ID NO:1),并混合直至均匀。

[0241] b.孵育如下：

[0242] i.70℃下2分钟

[0243] ii.保持在4℃下

[0244] (反应物应转变为黄色)

[0245] 3.5’接头连接

[0246] a.在冰上组合以下成分：

[0247] i.24μL源自先前步骤的反应物

[0248] ii.1μL 5’接头5’-UCUUUCCCUACACGACGCUCUUCCGAUCU-3’(SEQ ID NO:6)

[0249] iii.3μL 5’连接缓冲液

[0250] iv.2μL 5’连接酶混合物

[0251] v.30μL总反应体积

[0252] b.孵育如下：

[0253] i.20℃下1小时

[0254] ii.保持在4℃下

[0255] (反应物应转变为橙色)

[0256] 4.逆转录

[0257] a.在冰上组合以下成分：

[0258] i.7μL RT缓冲液

[0259] iii.2μL RT酶混合物

[0260] iv.1μL 25uM RT引物5’-CCTTGGCACCCGAGAATTCCA-3’(SEQ ID NO:5)

[0261] v.40μL总反应体积

[0262] b.孵育如下：

[0263] ii.50℃下60分钟

[0264] iii.90℃下5分钟

[0265] iv.保持在4℃下

[0266] c.使孵育完成并逐渐下降至4℃。

[0267] 5.小珠纯化

[0268] a.加入并混合40μL接头耗尽溶液

[0269] b.加入并混合40μLNEXTFLEX Cleanup小珠

[0270] c.加入并混合90μL异丙醇

[0271] d.在室温下孵育5分钟

[0272] e.置于磁性支架上5分钟或直至溶液澄清

[0273] f.抽吸去除上清液

[0274] g.通过加入200μL 80％乙醇洗涤,孵育30秒,并抽吸去除上清液

[0275] h.重复上述步骤,共洗涤2次

[0276] i.在磁性支架上干燥共计3分钟。1分钟后,抽吸并去除收集在孔底部处的任何液体

[0277] j.从磁性支架中取出并重悬于18μL无核酸酶水中

[0278] k.孵育2分钟

[0279] l.放置在磁性支架上

[0280] m.取16μL并移至新孔中。该cDNA将在以下步骤中扩增。

[0281] 6.PCR

[0282] a.在冰上组合以下成分：

[0283] i.16μL源自上述小珠纯化部分的先前步骤m的cDNA

[0284] ii.4μL RNAUDI Barcoded引物混合物12.5μM

[0285] iii.6μL NEXTFLEX小RNA PCR主混合物

[0286] iv.26μL总反应体积

[0287] b.孵育反应物如下：

[0288] i.98℃下30秒

[0289] ii.98℃下10秒

[0290] iii.65℃下20秒

[0291] iv.72℃下15秒

[0292] 重复ii、iii和iv,共22个循环

[0293] v.72℃下2分钟

[0294] vi.保持在4℃下

[0295] 7.双面小珠纯化

[0296] a.取26μL源自先前步骤的PCR反应物

[0297] b.加入并混合34μL NEXTFLEX Cleanup小珠

[0298] c.在室温下孵育5分钟

[0299] d.将56μL上清液转移至新孔

[0300] e.加入并混合13μL NEXTFLEX Cleanup小珠

[0301] d.置于磁性支架上5分钟或直至溶液澄清

[0302] e.抽吸去除上清液

[0303] f.通过加入200μL 80％乙醇洗涤,孵育30秒,并抽吸去除上清液

[0304] g.重复上述步骤,共洗涤2次

[0305] h.在磁性支架上干燥共计3分钟。1分钟后,抽吸并去除收集在孔底部处的任何液体

[0306] i.从磁性支架中取出并重悬于17μL无核酸酶水中

[0307] j.孵育2分钟

[0308] k.放置在磁性支架上

[0309] l.一旦小珠形成沉淀,移出15μL。这是你的测序文库。

[0310] 8.测序参数(MiSeq)

[0311] a.2％phiX

[0312] b.1x 75

[0313] 结果如图4所示。

[0314] 项目

[0315] 项目1.一种在测序文库中减少接头-二聚体的方法,包括:

[0316] 在连接反应条件下的混合物中,将各自具有5’末端和3’末端的样品核酸与3’接头和连接酶结合,所述3’接头各自具有长度为至少8个核苷酸的5’末端部分和一个为腺苷酸化核苷酸的5’端,产生1)第一连接产物,其各自包括连接至3’接头的5’端的样品核酸的3’末端,和2)混合物中过量的未连接的3’接头;

[0317] 向所述混合物中加入阻断剂寡核苷酸,所述阻断剂寡核苷酸包括:

[0318] a)第一阻断剂寡核苷酸,其包括:

[0319] i)第一阻断剂寡核苷酸5’部分,其长度为至少8个核苷酸,和ii)第一阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第一寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个为非天然存在的核苷酸;和

[0320] b)第二阻断剂寡核苷酸,其包括:

[0321] iii)第二寡核苷酸5’部分,其长度为至少8个核苷酸,和iv)第二阻断剂寡核苷酸3’部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3’端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;

[0322] 其中所述第一阻断剂寡核苷酸3’部分和所述第二阻断剂寡核苷酸3’部分长度相等且彼此互补;

[0323] 其中所述第一阻断剂寡核苷酸5’部分和所述第二阻断剂寡核苷酸5’部分均与所述3’接头的5’末端部分互补;

[0324] 在杂交条件下孵育混合物,从而使第一阻断剂寡核苷酸3’部分与第二阻断剂寡核苷酸3’部分杂交,使第一阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,并从而使第二阻断剂寡核苷酸5’部分与过量的未连接的3’接头杂交,产生杂交复合体,其中第一阻断剂寡核苷酸3’端与第一未连接的3’接头的腺苷酸化5’末端核苷酸相邻,且第二阻断剂寡

核苷酸3’端与第二未连接的3’接头的腺苷酸化5’末端核苷酸相邻,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

[0325] 项目2.根据项目1的方法,其中第一阻断剂寡核苷酸3’端和第二阻断剂寡核苷酸3’端均能够连接至过量的未连接的3’接头的5’端处的腺苷酸化核苷酸;且进一步包括将杂交复合体的第一阻断剂寡核苷酸3’端连接至第一未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,并且将杂交复合体的第二阻断剂寡核苷酸3’端连接到第二未连接的3’接头的5’端处的相邻的腺苷酸化核苷酸,产生连接复合体,从而减少混合物中过量的未连接的3’接头,产生纯化的混合物。

[0326] 项目3.根据项目1或2的方法,其中第一阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物,且第二阻断剂寡核苷酸3’部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物。

[0327] 项目4.根据项目1、2或3中任一项的方法,其中当第一阻断剂寡核苷酸3’部分和第二阻断剂寡核苷酸3’部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸时,第一阻断剂寡核苷酸3’部分和第二阻断剂寡核苷酸3’部分各自包括回文核苷酸序列,其中第一阻断剂寡核苷酸3’部分的回文核苷酸序列与第二阻断剂寡核苷酸3’部分的回文核苷酸序列相同。

[0328] 项目5.根据项目1至4中任一项的方法,进一步包括:向混合物中加入5’接头,所述5’接头具有3’端,其中所述第一连接产物各自包括与3’接头的5’端连接的样品核酸的3’末端,和与5’接头的3’端连接的样品核酸的5’末端,从而产生测序文库分子,所述测序文库分子包括位于5’接头和3’接头之间并与其结合的样品核酸插入片段。

[0329] 项目6.根据项目1至5中任一项的方法,进一步包括:向纯化的混合物中加入5’接头;和在连接反应条件下孵育纯化的混合物,从而1)将5’接头与第一连接产物连接,产生测序文库分子,所述测序文库分子包括位于5’接头和3’接头之间并与其结合的样品核酸插入片段。

[0330] 项目7.根据项目1至6中任一项的方法,其中减少但不清除过量的未连接的3’接头,使得所述连接包括将5’接头与过量的未连接的3’接头的至少一部分连接,产生接头-二聚体,其包括在5’接头-3’接头界面处与3’接头结合的5’接头;进一步包括抑制扩增反应中接头-二聚体的扩增,使得与接头-二聚体相比,通过扩增反应选择性富集测序文库分子。

[0331] 项目8.根据项目1至7中任一项的方法,其中抑制接头-二聚体的扩增包括在扩增之前去除至少一些接头-二聚体。

[0332] 项目9.根据项目7或8的方法,其中抑制接头-二聚体的扩增可包括以下中的一种或多种:接头-二聚体的特异性酶促切割;基于CRISPR/Cas9的接头-二聚体的切割;凝胶纯化;基于SPRI小珠的尺寸选择;和向混合物中加入接头-二聚体阻断剂寡核苷酸,所述接头-二聚体阻断剂寡核苷酸有效抑制接头-二聚体的扩增。

[0333] 项目10.根据项目9的方法,其中接头-二聚体阻断剂寡核苷酸与接头-二聚体的一个或多个相邻部分互补,所述一个或多个相邻部分包括在5’接头-3’接头界面处的5’接头的至少一个核苷酸,和包括扩增引物结合位点的3’接头的相邻部分。

[0334] 项目11.根据项目1至10中任一项的方法,其中第一阻断剂寡核苷酸和第二阻断剂

寡核苷酸具有相同的核苷酸序列。

[0335] 项目12.根据项目1的方法,其中样品核酸是或包括RNA,且进一步包括逆转录所述RNA,产生逆转录的测序文库分子。

[0336] 项目13.根据项目1的方法,其中样品核酸是或包括DNA。

[0337] 项目14.根据项目12或13的方法,进一步包括扩增测序文库分子。

[0338] 项目15.一种用于减少测序文库中的接头-二聚体的寡核苷酸对,包括:

[0339] a)第一阻断剂寡核苷酸,其包括:

[0340] i)第一阻断剂寡核苷酸5'部分,其长度为至少8个核苷酸,和ii)第一阻断剂寡核苷酸3'部分,其长度为至少1至20个核苷酸,并包括第一寡核苷酸3'端,其中所述1至20个核苷酸中的至少1个为非天然存在的核苷酸;和

[0341] b)第二阻断剂寡核苷酸,其包括:

[0342] iii)第二寡核苷酸5'部分,其长度为至少8个核苷酸,和iv)第二阻断剂寡核苷酸3'部分,其长度为至少1至20个核苷酸,并包括第二阻断剂寡核苷酸3'端,其中所述1至20个核苷酸中的至少1个是非天然存在的核苷酸;

[0343] 其中所述第一阻断剂寡核苷酸3'部分和所述第二阻断剂寡核苷酸3'部分长度相等且彼此互补;和

[0344] 其中所述第一阻断剂寡核苷酸5'部分和所述第二阻断剂寡核苷酸5'部分均与所述3'接头的5'末端部分互补;

[0345] 项目16.根据项目15的用于减少测序文库中接头-二聚体的寡核苷酸对,其中第一阻断剂寡核苷酸3'端和第二阻断剂寡核苷酸3'端均能够连接至过量的未连接的3'接头的5'端处的腺苷酸化核苷酸。

[0346] 项目17.根据项目15或16的用于减少测序文库中接头-二聚体的寡核苷酸对,其中第一阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物,且第二阻断剂寡核苷酸3'部分的至少1至20个核苷酸中的2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个是非天然存在的核苷酸类似物。

[0347] 项目18.根据项目15、16或17任一项的用于减少测序文库中接头-二聚体的寡核苷酸对,其中当第一阻断剂寡核苷酸3'部分和第二阻断剂寡核苷酸3'部分的长度各自为2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个核苷酸时,第一阻断剂寡核苷酸3'部分和第二阻断剂寡核苷酸3'部分各自包括回文核苷酸序列,其中第一阻断剂寡核苷酸3'部分的回文核苷酸序列与第二阻断剂寡核苷酸3'部分的回文核苷酸序列相同。

[0348] 项目19.根据项目15至18任一项的用于减少测序文库中接头-二聚体的寡核苷酸对,其中第一阻断剂寡核苷酸和第二阻断剂寡核苷酸各自具有约11至约50个核苷酸范围内的长度。

[0349] 项目20.根据项目15至19任一项的用于减少测序文库中接头-二聚体的寡核苷酸对,其中第一阻断剂寡核苷酸和第二阻断剂寡核苷酸具有相同的核苷酸序列。

[0350] 项目21.一种用于减少测序文库中的接头-二聚体的试剂盒,其包括项目15至20中任一项的寡核苷酸对。

[0351] 项目22.根据项目21的试剂盒,进一步包括有效阻断接头-二聚体扩增的接头-二

聚体阻断剂寡核苷酸。

[0352] 项目23.根据项目22的试剂盒,其中接头-二聚体阻断剂寡核苷酸与接头-二聚体的一个或多个相邻部分互补,所述一个或多个相邻部分包括在5'接头-3'接头界面处的5'接头的至少一个核苷酸,和包括扩增引物结合位点的3'接头的相邻部分。

[0353] 项目24.根据项目1至14中任一项的方法或项目21至23中任一项的试剂盒,其中第一阻断剂和第二阻断剂具有相同的核酸序列。

[0354] 项目25.根据项目1至14中任一项的方法或项目21至24中任一项的试剂盒,其中第一阻断剂和第二阻断剂具有不同的核酸序列。

[0355] 项目26.一种基本上如本文所示或所述的方法、试剂盒或寡核苷酸对。

[0356] 项目27.根据项目1至14任一项的方法,其中第一阻断剂寡核苷酸和第二阻断剂寡核苷酸各自具有约11至约50个核苷酸范围内的长度。

[0357] 本说明书中提及的任何专利或出版物均通过引用的方式纳入本文,其与每个单独的出版物被具体地且单独地被指出通过引用的方式纳入的程度相同。

[0358] 本文记载的组合物和方法目前代表优选的实施方案,是示例性的,且不旨在作为对本发明范围的限制。本领域技术人员将意识到其中的改变和其他用途。可以在不脱离权利要求中阐述的发明范围的情况下进行这种改变和其他用途。

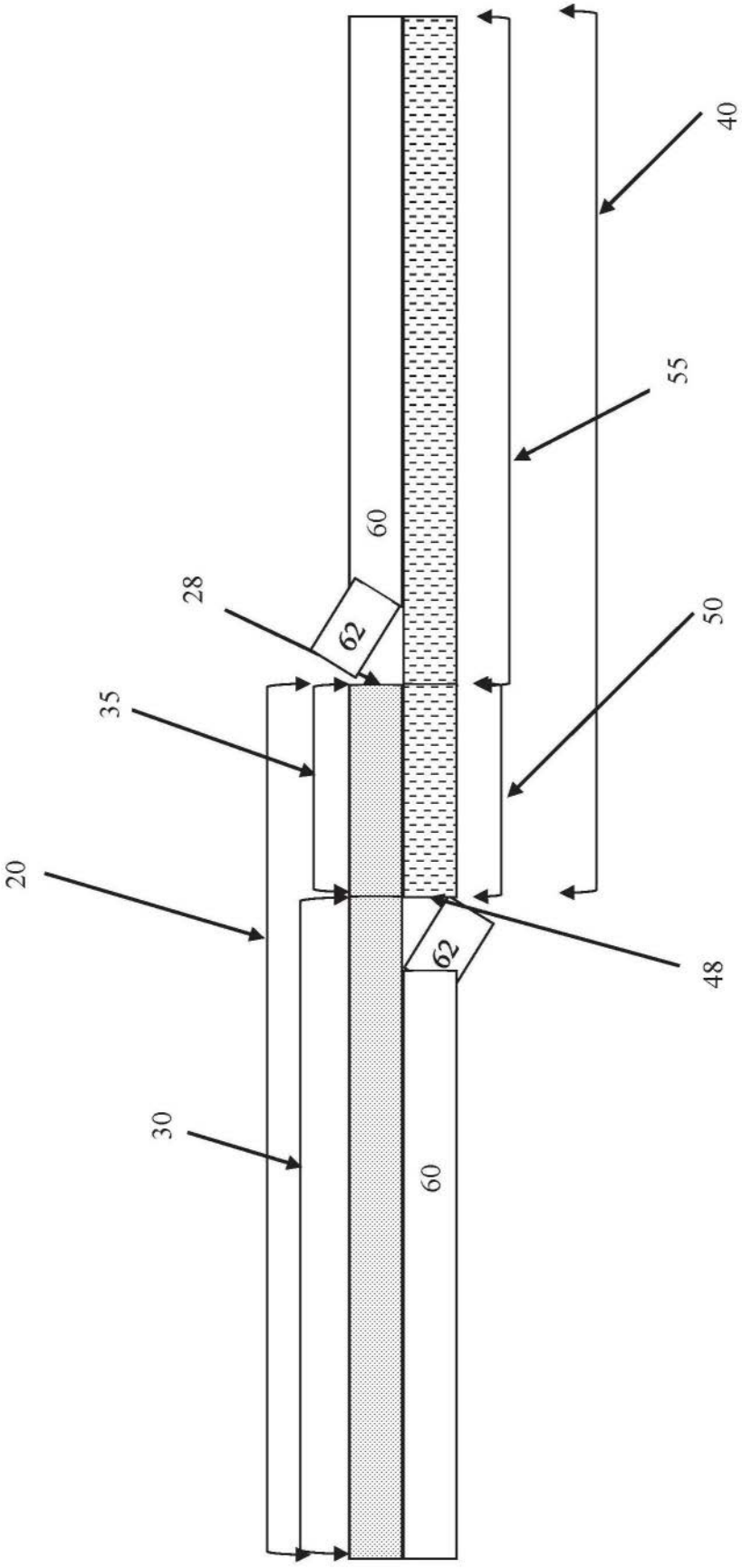


图1

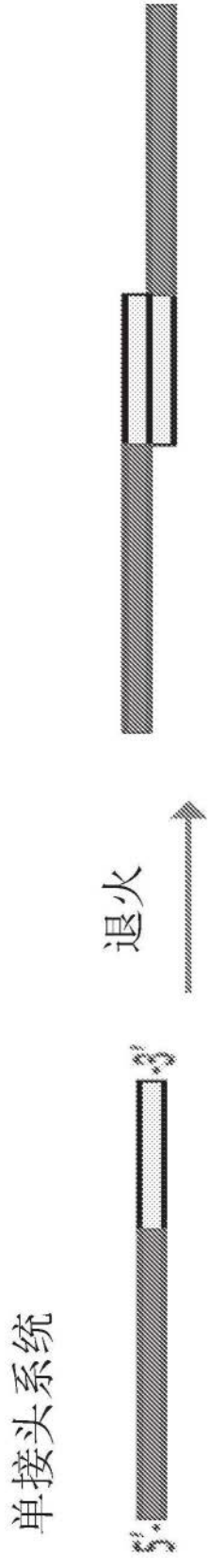


图2A

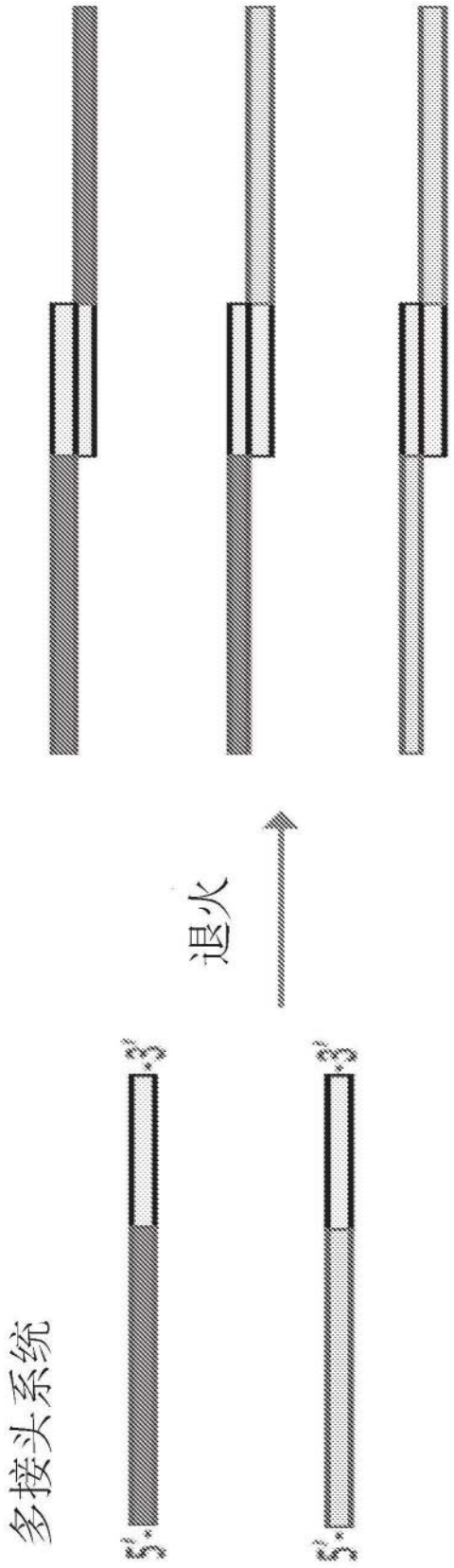


图2B

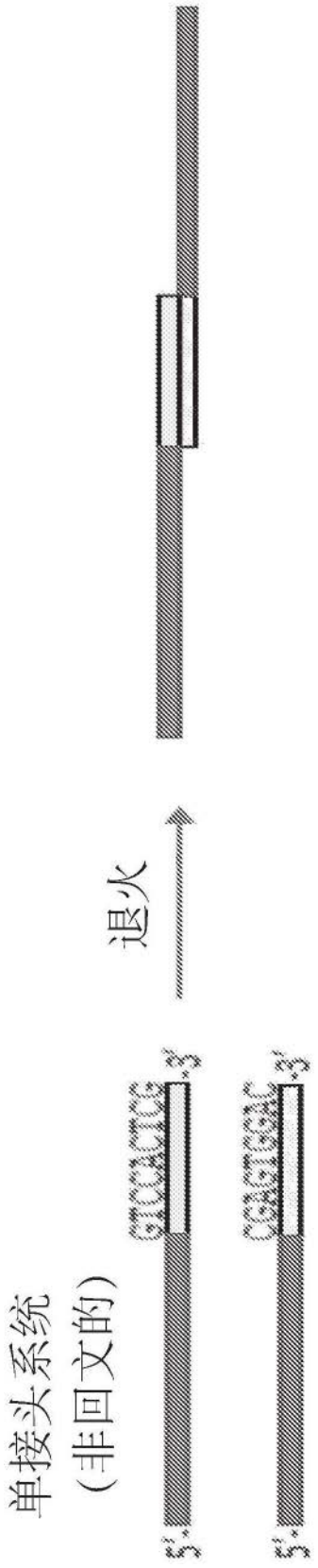


图2C

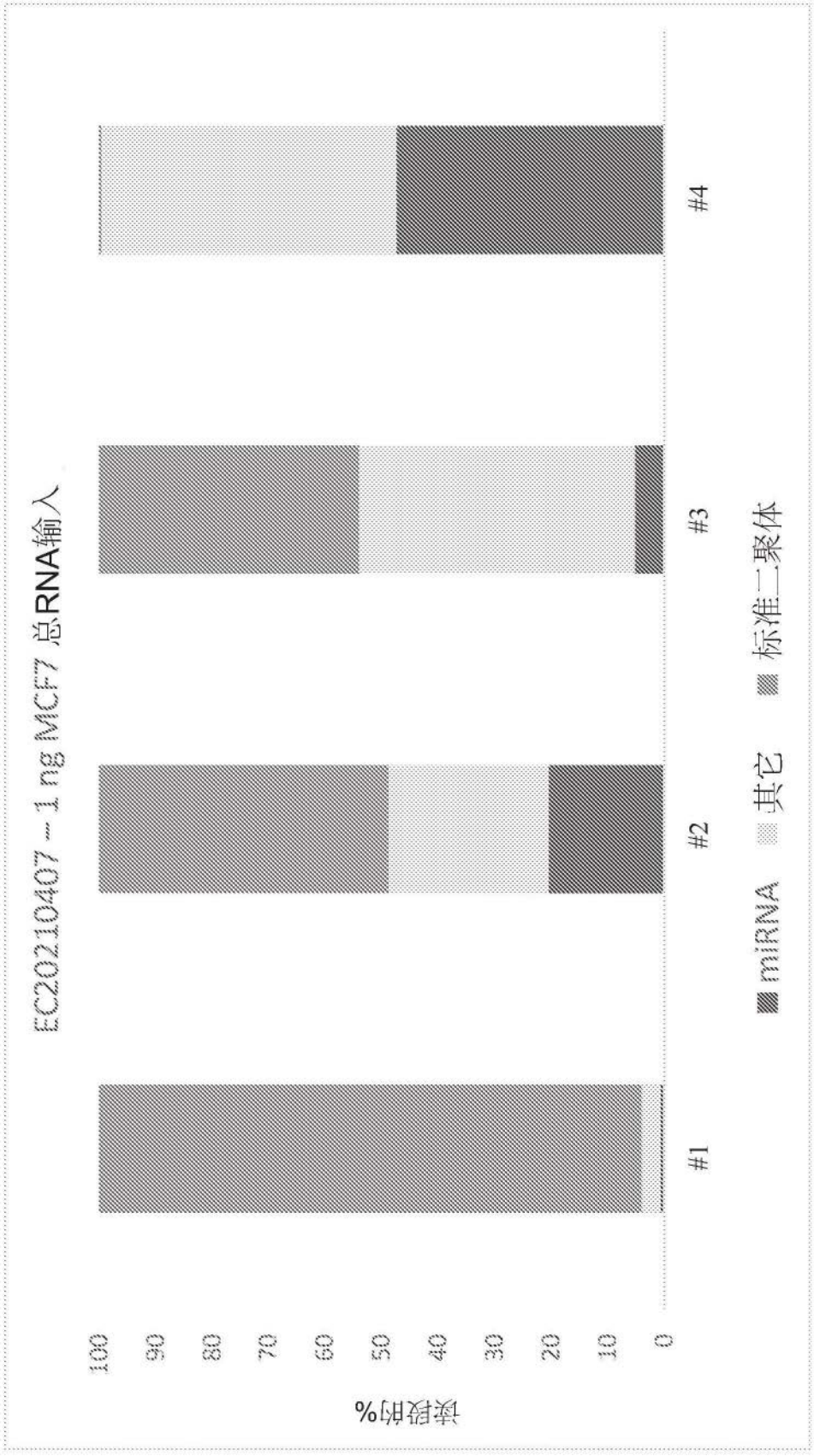

EC20210407 – 1 ng MCF7 总RNA输入
读段的%
100
90
80
70
60
50
40
30
20
10
0
#1
#2
#3
#4
miRNA
其它
标准二聚体

图3

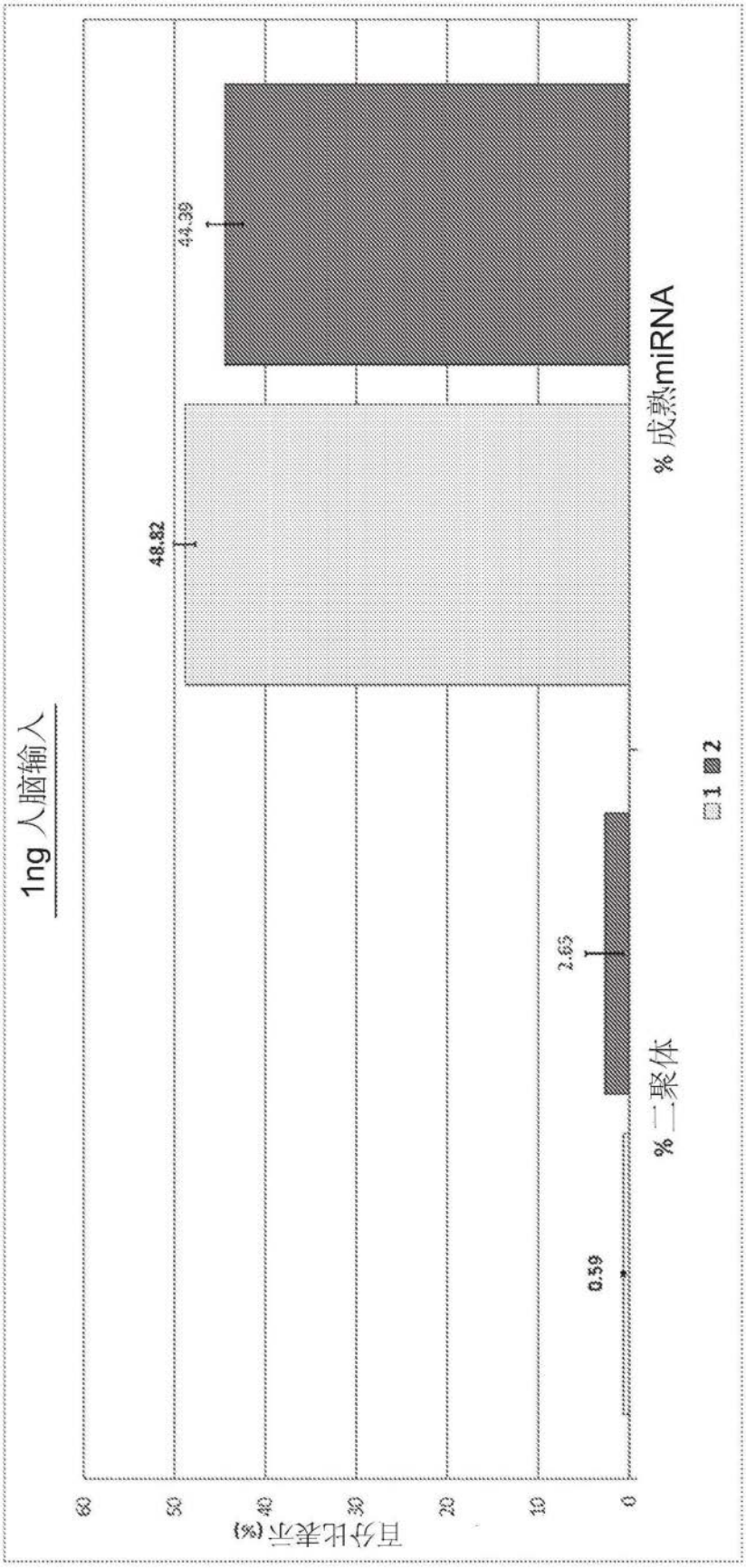

1ng 人脑输入
百分比表示(%)
60
50
40
30
20
10
0
44.39
48.82
2.65
0.59
% 成熟miRNA
% 二聚体
1
2

图4

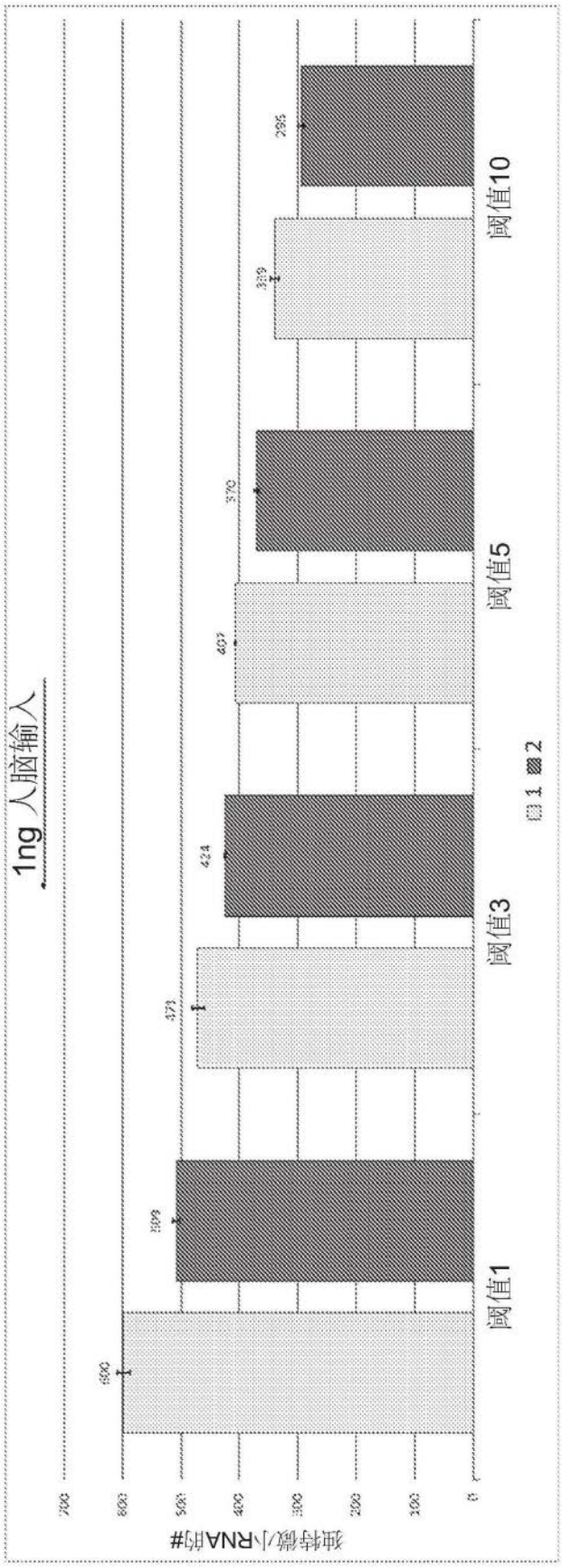

1ng 人脑输入
独特微小RNA的#
阈值1
阈值3
阈值5
阈值10
1
2
889
473
424
402
370
329
285

图5

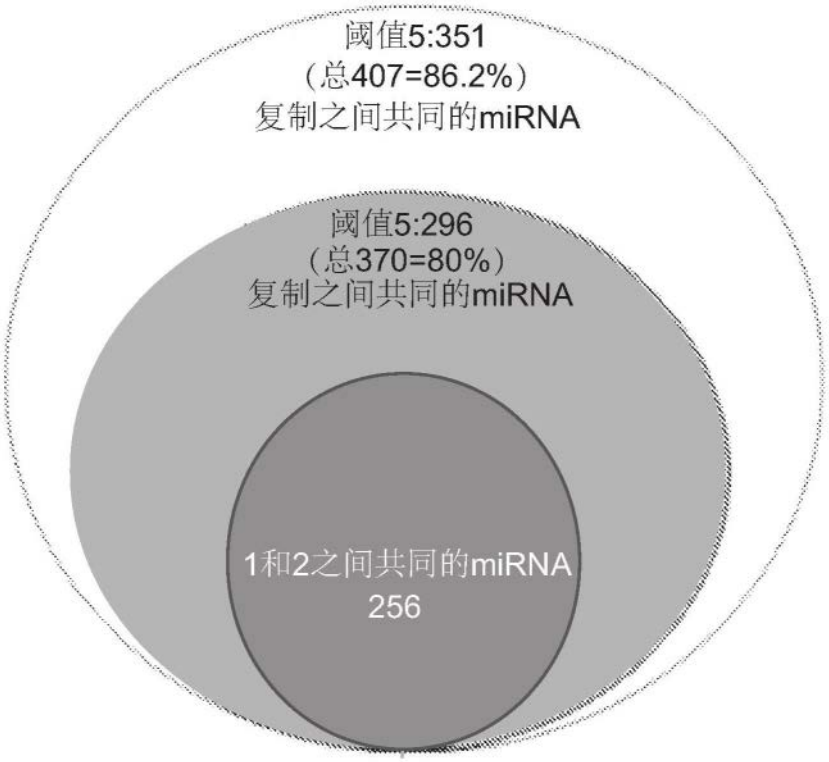

阈值5:351
（总407=86.2%）
复制之间共同的miRNA
阈值5:296
（总370=80%）
复制之间共同的miRNA
1和2之间共同的miRNA
256

图6