



(12) 发明专利

(10) 授权公告号 CN 114277111 B

(45) 授权公告日 2024. 08. 16

(21) 申请号 202111675724.2

(22) 申请日 2021.12.31

(65) 同一申请的已公布的文献号
申请公布号 CN 114277111 A

(43) 申请公布日 2022.04.05

(73) 专利权人 深圳市核子基因科技有限公司
地址 518055 广东省深圳市南山区桃源街
道南山智园A7栋12楼

(72) 发明人 程云阳 巴颖 操利超 张核子
卢晓萍

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205
专利代理师 刘燚

(51) Int. Cl.
C12Q 1/686 (2018.01)

(56) 对比文件

CN 111379033 A, 2020.07.07

Assaf Rotem等. High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. PloS One. 2015, 第10卷(第5期), 摘要、第2页第3段、第8页第3段至第9页第4段.

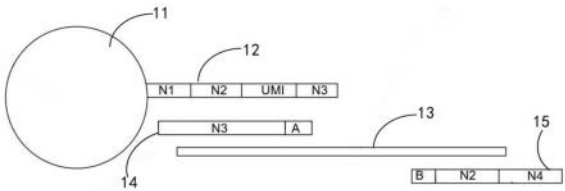
审查员 白晓岩

权利要求书1页 说明书7页
序列表2页 附图3页

(54) 发明名称
一种引入标签序列的方法

(57) 摘要

本发明属于分子生物技术领域,公开了一种引入标签序列的方法。本发明通过乳液PCR将不同的UMI标签序列配置在不同的乳液液滴中,并保证每个乳液液滴中有一个DNA分子做模板,使每次特异性扩增过程中,向同一模板的扩增产物引入的是同一种UMI标签序列,且通过第二次扩增把样本标签序列Index引入到扩增产物中,使最终的扩增产物同时带有UMI标签序列和样本标签序列Index,避免出现多个扩增产物对应大量UMI标签序列的情况,解决了相关技术中在测序时的数据浪费问题。



1. 一种引入标签序列的方法,其特征在于,包括以下步骤:

S1:提供第一油相和第一水相,所述第一水相为制备包括UMI标签序列的核酸片段的乳液PCR反应体系,混合所述第一油相和所述第一水相形成第一乳液液滴,将所述包括UMI标签序列的核酸片段配置于不同的所述第一乳液液滴中,使每个所述第一乳液液滴中配置的所述核酸片段包括不同的所述UMI标签序列;

S2:提供样本和第一引物对,将所述样本中的每个DNA分子、所述第一引物对和所述包括UMI标签序列的核酸片段配置于不同的第二乳液液滴中,通过乳液PCR进行特异性扩增,得第一扩增产物,使位于一个所述第二乳液液滴中的所述第一扩增产物携带有相同的所述UMI标签序列,经所述乳液PCR扩增得到所述第一扩增产物后,需采用磁珠分离纯化法对所述第一扩增产物进行纯化;

S3:提供包括样本标签序列Index的第二引物对,以步骤S2所述第一扩增产物为模板进行PCR扩增,得第二扩增产物,使所述第二扩增产物携带所述样本标签序列Index和所述UMI标签序列,所述PCR结束后,需采用磁珠分离纯化法对扩增产物进行纯化,得到干净的所述第二扩增产物;

其中,所述包括UMI标签序列的核酸片段被配置在磁珠上,所述磁珠连接在所述核酸片段的5'端;

所述第一乳液液滴是油包水乳液液滴,按第一油相和第一水相的体积比为(3~3.5):1配制而成;

所述第一油相按稳定剂1:稳定剂2:乳化剂=75:4:(800~1000)的体积比配制而成;所述稳定剂1包括司盘80、吐温80、吐温60或吐温20中的至少一种;所述稳定剂2包括TritonX-100或TritonX-114中的至少一种;所述乳化剂包括三聚甘油单硬脂酸酯或甘油硬脂酸酯中的至少一种;

所述第二乳液液滴按第二油相和第二水相的体积比为(3~3.5):1配制而成;

所述第二油相按稳定剂1:稳定剂2:乳化剂=75:4:(800~1000)的体积比配制而成;所述稳定剂1包括司盘80、吐温80、吐温60或吐温20中的至少一种;所述稳定剂2包括TritonX-100或TritonX-114中的至少一种;所述乳化剂包括三聚甘油单硬脂酸酯或甘油硬脂酸酯中的至少一种;

所述第二水相为步骤S2中所述乳液PCR的反应液;

所述第一引物对包括上游引物和下游引物;

所述上游引物的5'端与所述包括UMI标签序列的核酸片段的3'端序列相同,所述上游引物的3'端为用于扩增所述样本中的所述DNA分子所需的上游引物序列;

所述下游引物的3'端为用于扩增所述样本中的所述DNA分子所需的下游引物序列,连接所述下游引物的3'端的一段序列与所述包括UMI标签序列的核酸片段的一段序列相同。

2. 根据权利要求1所述的方法,其特征在于,所述UMI标签序列被配置为包括8~12个碱基的随机序列。

一种引入标签序列的方法

技术领域

[0001] 本发明属于分子生物技术领域,具体涉及一种引入标签序列的方法。

背景技术

[0002] 下一代测序(NGS)可以同时为数百万个DNA片段进行测序,在分析测序结果时,常需要区分多个目的片段是否来源于同一模板,常用的扩增法需要在目的片段上引入用于区分文库的样本标签序列Index外,还需引入用于区分模板的UMI标签序列。

[0003] 在具体操作时需要进行两轮扩增,首先在第一轮特异性扩增中引入UMI标签序列,然后在第二轮通用引物扩增中引入样本标签序列Index,从而完成对样本标签的引入,以区分来自不同样本的片段。但上述方法存在一定的缺陷:由于第一轮扩增的目的是特异性扩增,扩增循环数不能太少(一般为15轮左右),但是每轮特异性扩增都会在扩增产物中随机引入UMI标签序列,如果扩增效率为2,扩增15个循环就会出现来源于同一模板的扩增产物与约 2^{15} 个UMI标签序列对应的情况,使得在后续的聚类分析中,同一种扩增产物对应大量聚类,造成数据的大量浪费。因此,有必要提供一种每次特异性扩增向同一模板的扩增产物中引入同一种UMI标签序列的方法,以减少数据的浪费。

发明内容

[0004] 本发明旨在至少解决上述现有技术中存在的技术问题之一。为此,本发明提出一种引入标签序列的方法,能够保证在每次特异性扩增过程中,向同一模板的扩增产物引入的是同一种UMI标签序列。

[0005] 在说明书中提到的核酸序列都为5'-3'的方向。

[0006] 根据本发明的一个方面,提出了一种引入标签序列的方法,包括以下步骤:

[0007] S1:提供第一油相和第一水相,所述第一水相为制备包括UMI标签序列的核酸片段的反应体系,混合所述第一油相和所述第一水相形成第一乳液液滴,将所述包括UMI标签序列的核酸片段配置于不同的所述第一乳液液滴中,使每个所述第一乳液液滴中配置的所述核酸片段包括不同的所述UMI标签序列;

[0008] S2:提供样本和第一引物对,将所述样本中的每个DNA分子、所述第一引物对和所述包括UMI标签序列的核酸片段配置于不同的第二乳液液滴中,通过乳液PCR进行特异性扩增,得第一扩增产物,使位于一个所述第二乳液液滴中的所述第一扩增产物携带有相同的所述UMI标签序列;

[0009] S3:提供包括样本标签序列Index的第二引物对,以步骤S2所述第一扩增产物为模板进行扩增,得第二扩增产物,使所述第二扩增产物携带所述样本标签序列Index和所述UMI标签序列。

[0010] 根据本发明的一种优选的实施方式,至少具有以下有益效果:

[0011] 本发明通过乳液PCR将不同的UMI标签序列配置在不同的乳液液滴中,并保证每个乳液液滴中有一个DNA分子做模板,使每次特异性扩增过程中,向同一模板的扩增产物引入

的是同一种UMI标签序列,且通过第二次扩增把样本标签序列Index引入到扩增产物中,使最终的扩增产物同时带有UMI标签序列和样本标签序列Index,避免出现多个扩增产物对应大量UMI标签序列的情况,解决了相关技术中在测序时的数据浪费问题。

[0012] 在本发明的一些实施方式中,所述第一乳液液滴是油包水乳液液滴,按第一油相和第一水相的体积比为(3~3.5):1配制而成。

[0013] 在本发明的一些实施方式中,所述第一油相按稳定剂1:稳定剂2:乳化剂=75:4:(800~1000)的体积比配制而成;所述稳定剂1包括司盘80、吐温80、吐温60或吐温20中的至少一种;所述稳定剂2包括TritonX-100或TritonX-114中的至少一种;所述乳化剂包括三聚甘油单硬脂酸酯或甘油硬脂酸酯中的至少一种。

[0014] 在本发明的一些实施方式中,所述UMI标签序列被配置为包括8~12个碱基的随机序列。

[0015] 在本发明的一些优选的实施方式中,所述UMI标签序列被配置为包括8个碱基的随机序列。

[0016] 在本发明的一些实施方式中,所述包括UMI标签序列的核酸片段被配置在载体上,所述载体连接在所述核酸片段的5'端。

[0017] 在本发明的一些实施方式中,所述载体被配置为磁珠。

[0018] 在本发明的一些实施方式中,经所述乳液PCR扩增得到所述第一扩增产物后,需采用磁珠分离纯化法对所述第一扩增产物进行纯化。

[0019] 在本发明的一些实施方式中,所述第一引物对包括上游引物和下游引物。

[0020] 在本发明的一些实施方式中,所述上游引物的5'端与所述包括UMI标签序列的核酸片段的3'端序列相同,所述上游引物的3'端为用于扩增所述样本中的所述DNA分子所需的上游引物序列。

[0021] 在本发明的一些实施方式中,所述下游引物的3'端为用于扩增所述样本中的所述DNA分子所需的下游引物序列,连接所述下游引物的3'端的一段序列与所述包括UMI标签序列的核酸片段的一段序列相同。

[0022] 在本发明的一些实施方式中,所述第二乳液液滴按第二油相和第二水相的体积比为(3~3.5):1配制而成。

[0023] 在本发明的一些实施方式中,所述第二油相按稳定剂1:稳定剂2:乳化剂=75:4:(800~1000)的体积比配制而成;所述稳定剂1包括司盘80、吐温80、吐温60或吐温20中的至少一种;所述稳定剂2包括TritonX-100或TritonX-114中的至少一种;所述乳化剂包括三聚甘油单硬脂酸酯或甘油硬脂酸酯中的至少一种。

[0024] 在本发明的一些实施方式中,所述第二水相为步骤S2中所述乳液PCR的反应液。

[0025] 在本发明的一些实施方式中,所述第二扩增产物经PCR扩增得到。

[0026] 在本发明的一些实施方式中,所述PCR结束后,需采用磁珠分离纯化法对扩增产物进行纯化,得到干净的所述第二扩增产物。

附图说明

[0027] 下面结合附图和实施例对本发明做进一步的说明,其中:

[0028] 图1为本发明实施例制备的磁珠-N1-N2-UMI-N3的结构示意图;

- [0029] 图2为本发明实施例第一引物对的结构示意图；
- [0030] 图3为本发明实施例形成的第二乳液液滴的示意图；
- [0031] 图4为本发明实施例引入UMI标签序列的乳液PCR过程示意图；
- [0032] 图5为本发明实施例第一扩增产物结构示意图；
- [0033] 图6为本发明实施例第二引物对的结构示意图；
- [0034] 图7为本发明实施例以第二引物对进行PCR扩增的示意图；
- [0035] 图8为本发明实施例第二扩增产物结构示意图。
- [0036] 附图标记：
- [0037] 11-磁珠；12-N1-N2-UMI-N3核酸片段；13-DNA样本；14-第一上游引物；15-第一下游引物；16-第二乳液液滴；21-N1-N2-UMI-N3-Insert-N2'-N4'核酸片段；31-第二上游引物；32-第二下游引物。

具体实施方式

[0038] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能理解为对本发明的限制。

[0039] 如果有描述到第一、第二只是用于区分技术特征为目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量或者隐含指明所指示的技术特征的先后关系。

[0040] 本发明的描述中,除非另有明确的限定,扩增、连接等词语应做广义理解,所属技术领域技术人员可以结合技术方案的具体内容合理确定上述词语在本发明中的具体含义。

[0041] 本发明的描述中,参考术语“一个实施例”、“一些实施例”等的描述意指结合该实施例描述的具体特征或材料包含于本发明的至少一个实施例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例。而且,描述的具体特征或材料可以在任何一个或多个实施例中以合适的方式结合。

[0042] 实施例中所使用的试验方法如无特殊说明,均为常规方法;所使用的材料、试剂等,如无特殊说明,均可从商业途径得到的试剂和材料。

[0043] 实施例1

[0044] 本实施例对待测样本引入了标签序列,具体过程为:

[0045] (1) 第一油相配制:按吐温80:TritonX-100:三聚甘油单硬脂酸酯=75:4:920的体积比配制;

[0046] 第一水相配制:按表1进行配制;

[0047] 表1

[0048]

模板：核酸片段 N2-UMI-N3 (50fM)	2 μ L
2 $\times$ Kapa Hifi Mix	50 μ L
上游引物：磁珠-N1-N2	50 μ g
下游引物：N3' (0.5 μ M)	2.5 μ L
H ₂ O	100 μ L-上述试剂的总体积量
总体积	100 μ L

[0049] N1为固定序列,序列信息如SEQ ID N0:1所示,具体为acactcttttcctacacgac;N2为固定序列,序列信息如SEQ ID N0:2所示,具体为gctcttccgatct;N3为固定序列,序列信息如SEQ ID N0:3所示,具体为ggtcttaggaagacaa;UMI为8碱基的随机序列,表示为NNNNNNNN;N3'与N3反向互补配对,序列信息如SEQ ID N0:4所示,具体为ttgtcttcctaagacc,N3'的5'端进行了生物素修饰;

[0050] 第一乳液液滴形成及扩增:在每100 μ L第一水相体系中加入320 μ L第一油相,涡旋混匀仪最大转速振荡5min,制成第一乳液液滴体系;每50 μ L一管进行乳液PCR,PCR结束后即得到包括UMI标签序列的核酸片段,其中每个第一乳液液滴中包括一种UMI标签序列,各第一乳液液滴包含的UMI标签序列不同。扩增程序如表2所示;

[0051] 表2

[0052]

阶段	温度	时间	循环数
1	98°C	1min	1
2	98°C	15s	16
	68°C	30s	
	72°C	30s	
3	72°C	5min	1
4	4°C	保持	1

[0053] 扩增产物纯化:取乳液PCR扩增产物,加入洗涤后的M270磁珠(为修饰有链霉亲和素的磁珠,可与带有生物素的扩增产物结合),室温孵育45min,上磁力架,去除反应液和未吸附的磁珠,采用1 $\times$ Beads Wash Buffer清洗两次,ddH₂O清洗一次,回溶至ddH₂O中,得到磁珠-N1-N2-UMI-N3-M270磁珠(其中磁珠与N1的5'端相连,M270磁珠与互补链中N3'的5'端相连);取纯化产物,90°C变性5min,迅速转移至磁力架上,将液体和未吸附的磁珠转至新的离

心管中,离心,去除上清,用ddH₂O清洗两次,回溶至ddH₂O中,得到核酸片段磁珠-N1-N2-UMI-N3,如图1所示,N1-N2-UMI-N3核酸片段12为单链DNA,通过N1的5'端与磁珠11相连,每个第一乳液液滴中含有一种磁珠-N1-N2-UMI-N3,各第一乳液液滴中的这一核酸片段因UMI标签序列不同而不同。

[0054] (2) 基因组DNA的提取:取200μL人源全血,使用磁珠法通用型基因组DNA提取试剂盒(TIANGEN,DP705),提取全血DNA,取提取产物用ddH₂O稀释至DNA样本浓度为0.88ng/μL。

[0055] (3) 第二油相配制:按吐温80:TritonX-100:三聚甘油单硬脂酸酯=75:4:920的体积比配制;

[0056] 第二水相配制:按表3进行配制;

[0057] 表3

[0058]	模板: 稀释后的 DNA 样本 (0.88 ng/μL)	2μL
	2×Kapa Hifi Mix	50μL
	磁珠-N1-N2-UMI-N3	1μg
	上游引物 F1: N3-A (0.5μM)	2.5μL
	下游引物 R1: N4-N2-B (0.5μM)	2.5μL
	总体积	100μL

[0059] 上游引物F1和下游引物R1的结构示意图如图2所示,其中A为扩增DNA样本所需的上游引物,B为扩增DNA样本所需的下游引物;N1、N2和N3的序列信息如(1)所述,N4为固定序列,序列信息如SEQ ID NO:5所示,具体为gtgactggagttcagacgtgt;磁珠-N1-N2-UMI-N3用于在扩增时引入UMI标签序列;

[0060] 第二乳液液滴形成及扩增:在每100μL第二水相体系中加入320μL第二油相,涡旋混匀仪最大转速振荡5min,制成第二乳液液滴16(如图3和图4所示),包括磁珠11,N1-N2-UMI-N3核酸片段12,DNA样本13,第一上游引物14和第一下游引物15;每100μL一管进行乳液PCR,PCR结束后即完成对DNA样本的扩增,且给扩增产物中引入了UMI标签序列。扩增程序如表4所示;

[0061] 表4

[0062]

阶段	温度	时间	循环数
1	98°C	1min	1
2	98°C	15s	15
	60°C	30s	
	72°C	30s	
3	72°C	5min	1
4	4°C	保持	1

[0063] 扩增产物纯化:取扩增产物,离心,去除反应液,用80%乙醇清洗两次,室温晾干1~2min至表面无光泽,回溶至20 μ L ddH₂O中,得到第一扩增产物(如图5所示):磁珠-N1-N2-UMI-N3-Insert-N2'-N4'(5'-3'),由磁珠11和N1-N2-UMI-N3-Insert-N2'-N4'核酸片段21构成;其中Insert为扩增引入的插入片段,来源于DNA样本,N2'与N2反向互补配对,N4'与N4反向互补配对,此时第一扩增产物上带有UMI标签序列,且位于同一第二乳液液滴中的第一扩增产物携带的UMI标签序列相同,位于不同第二乳液液滴中的第一扩增产物携带的UMI标签序列不同,这样就可根据UMI标签序列来判断扩增产物是否来源于同一模板,即携带有相同UMI标签序列的第一扩增产物是由同一DNA模板扩增得到的。

[0064] (4)按表5配制PCR反应体系;

[0065] 表5

[0066]

模板:纯化的第一扩增产物	20 μ L
2 $\times$ Kapa Hifi Mix	25 μ L
上游引物F2:P1-Index1-N1(5 μ M)	2.5 μ L
下游引物R2:P2-Index2-N4(5 μ M)	2.5 μ L
总体积	50 μ L

[0067] 上游引物F2和下游引物R2的结构示意图如图6所示,P1为固定序列,序列信息如SEQ ID NO:6所示,具体为aatgatacggcgaccaccgagatctacac;Index1为区分样本的标签序列1;P2为固定序列,序列信息如SEQ ID NO:7所示,具体为caagcagaagacggcatagagat,Index2为区分样本的标签序列2;

[0068] 进行PCR(如图7所示),其中以磁珠11连接的N1-N2-UMI-N3-Insert-N2'-N4'核酸片段21为模板,以第二上游引物31和第二下游引物32为引物对进行扩增,扩增结束后即在扩增产物的两端各引入一个样本标签序列;按表6扩增程序进行;

[0069] 表6

[0070]

阶段	温度	时间	循环数
1	98°C	1min	1
2	98°C	15s	25
	60°C	30s	
	72°C	30s	
3	72°C	2min	1
4	4°C	保持	1

[0071] 扩增产物纯化:取扩增产物50 μ L,加入60 μ L XP Beads,室温孵育5min,上磁力架,去除反应液及未吸附的磁珠,用80%乙醇清洗两次,室温晾干1~2min至表面无光泽,回溶至20 μ L ddH₂O中,重新上磁力架,得到第二扩增产物(如图8所示):P1-Index1-N1-N2-UMI-N3-Insert-N2'-N4'-Index2'-P2'(5'-3'),其中Index2'与Index2反向互补配对,P2'与P2反向互补配对。

[0072] 此时即在扩增产物中同时携带有样本标签序列Index和UMI标签序列,这样就可以给扩增产物加上样本标签,每个扩增产物就带上属于它的“专属号码牌”,在多样本混合测序后就能通过这个号码牌来获取来源于各样本的测序数据;且扩增于同一模板的扩增产物中携带有同一种UMI标签,通过这一标签既可以判断扩增产物的来源,又可以使得在后续的测序和聚类分析中,同一种扩增产物不需要对应大量聚类,大大减少了数据浪费。

[0073] 上面结合附图对本发明实施例作了详细说明,但是本发明不限于上述实施例,在所属技术领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化。此外,在不冲突的情况下,本发明的实施例及实施例中的特征可以相互组合。

[0001] SEQUENCE LISTING
[0002] <110> 深圳市核子基因科技有限公司
[0003] <120> 一种引入标签序列的方法
[0004] <130> 1
[0005] <160> 7
[0006] <170> PatentIn version 3.5
[0007] <210> 1
[0008] <211> 20
[0009] <212> DNA
[0010] <213> 人工序列
[0011] <400> 1
[0012] acactctttc cctacacgac 20
[0013] <210> 2
[0014] <211> 13
[0015] <212> DNA
[0016] <213> 人工序列
[0017] <400> 2
[0018] gctcttccga tct 13
[0019] <210> 3
[0020] <211> 16
[0021] <212> DNA
[0022] <213> 人工序列
[0023] <400> 3
[0024] ggtcttagga agacaa 16
[0025] <210> 4
[0026] <211> 16
[0027] <212> DNA
[0028] <213> 人工序列
[0029] <400> 4
[0030] ttgtcttcct aagacc 16
[0031] <210> 5
[0032] <211> 21
[0033] <212> DNA
[0034] <213> 人工序列
[0035] <400> 5
[0036] gtgactggag ttcagacgtg t 21
[0037] <210> 6
[0038] <211> 29

-
- [0039] <212> DNA
[0040] <213> 人工序列
[0041] <400> 6
[0042]aatgatacgg cgaccaccga gatctacac 29
[0043] <210> 7
[0044] <211> 24
[0045] <212> DNA
[0046] <213> 人工序列
[0047] <400> 7
[0048]caagcagaag acggcatacg agat 24

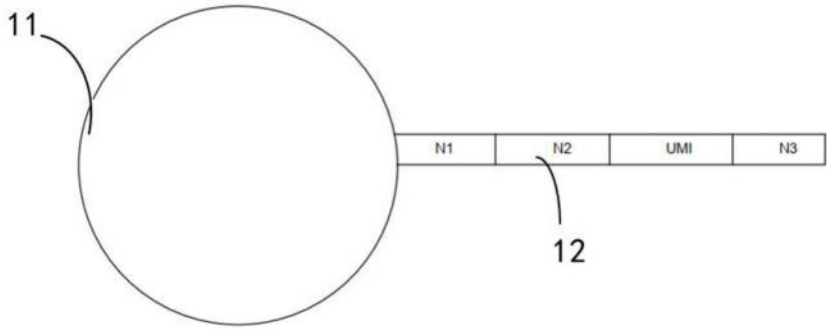


图1

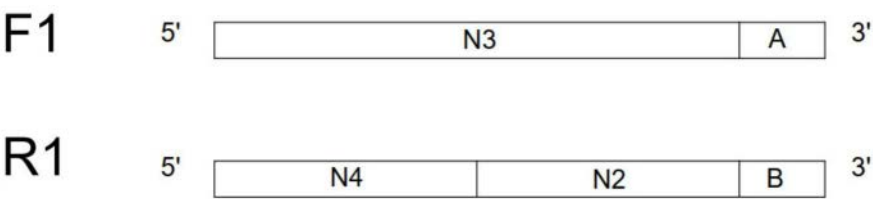


图2

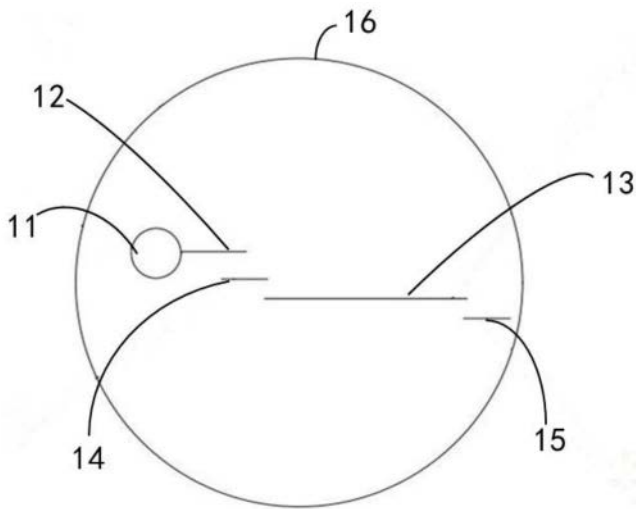


图3

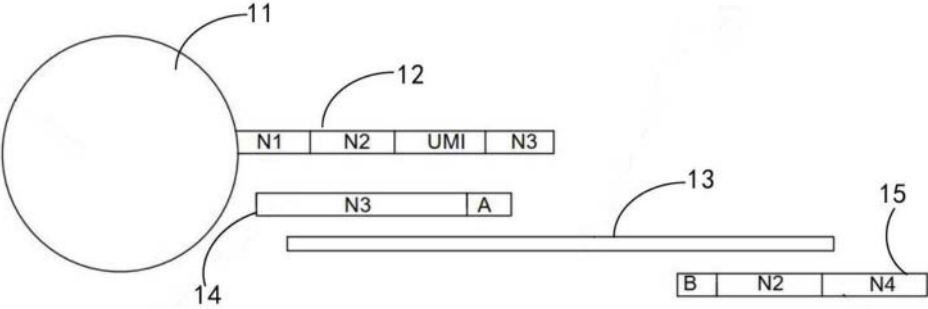


图4

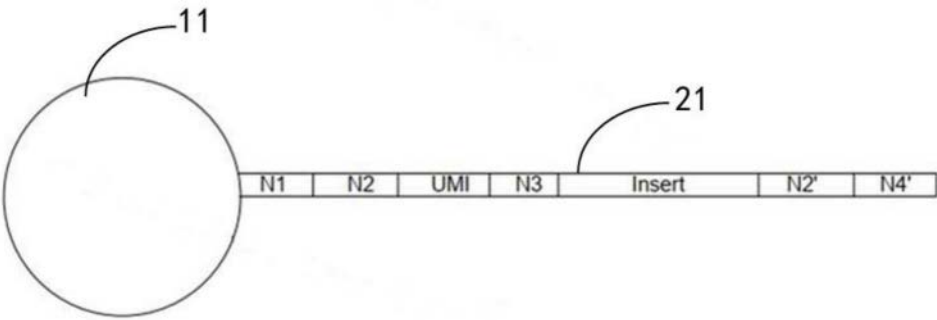


图5

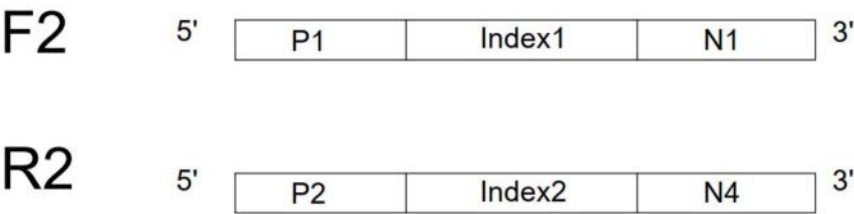


图6

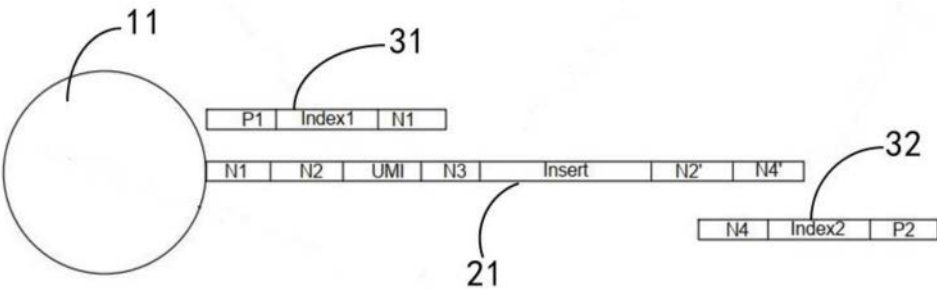


图7

P1	Index1	N1	N2	UMI	N3	Insert	N2'	N4'	Index2	P2
----	--------	----	----	-----	----	--------	-----	-----	--------	----

图8