



(21) 申请号 202111199578.0

(22) 申请日 2021.10.14

(65) 同一申请的已公布的文献号  
申请公布号 CN 114807302 A

(43) 申请公布日 2022.07.29

(73) 专利权人 深圳联合医学科技有限公司  
地址 518055 广东省深圳市南山区西丽街  
道留仙洞村留仙文化园(原‘桑泰科技  
园’)第6栋2楼202室

(72) 发明人 雷湘华 叶苑青 郭永超 徐仲尧  
王艳平 蔡锦刚

(74) 专利代理机构 上海助之鑫知识产权代理有  
限公司 31328  
专利代理师 陈详

(51) Int.Cl.

C12Q 1/6806 (2018.01)

C12N 15/10 (2006.01)

C40B 50/06 (2006.01)

C12Q 1/6883 (2018.01)

C12N 15/11 (2006.01)

(56) 对比文件

CN 108796054 A, 2018.11.13

CN 108998508 A, 2018.12.14

审查员 李紫嘉

权利要求书1页 说明书19页  
序列表11页 附图2页

(54) 发明名称

扩增子文库构建方法及用于地中海贫血突  
变型与缺失型基因检测的试剂盒

(57) 摘要

本发明提供了扩增子文库构建方法及用于  
地中海贫血突变型与缺失型基因检测的试剂盒,  
具体的,本发明公开了一种多重PCR扩增子文库  
构建方法,实验操作简便快捷。基于该文库构建  
方法,本发明构建了可以同时检测 $\alpha$ 地贫和 $\beta$ 地  
贫的突变型和缺失型的文库并建立了缺失型分  
析方法。该方法检测范围广,可以检测新发突变,  
实验操作简便快捷,通量高成本低,很好地弥补  
现有技术缺点。

1. 一种能同时检测 $\alpha$ 地中海贫血和 $\beta$ 地中海贫血缺失型和突变型的引物对集,其特征在于,所述引物对集包括:如SEQ ID No.1-51、64所示的引物;所述引物对集包括第一引物组、第二引物组、和第三引物组;其中,所述第一引物组包括如SEQ ID NO.1、2、3、4、5、6、30、31、36、37、38、39、40、41、42、43所示的引物;所述第二引物组包括如SEQ ID NO.7、8、9、10、11、12、13、22、23、24、25、44、45、46、47、48、49、50、51、64所示的引物;所述第三引物组包括如SEQ ID NO.14、15、16、17、18、19、20、21、26、27、28、29、32、33、34、35所示的引物。

2. 如权利要求1所述的引物对集,其特征在于,所述引物对集中,各正向引物的5'端还包括第一测序标签序列;所述引物对集中,各反向引物的5'端还包括第二测序标签序列。

3. 如权利要求2所述的引物对集,其特征在于,所述第一测序标签序列如SEQ ID NO.54所示。

4. 如权利要求2所述的引物对集,其特征在于,所述第二测序标签序列如SEQ ID NO.55所示。

5. 一种试剂盒,其特征在于,所述试剂盒包含权利要求1所述引物对集;

所述试剂盒包括第一容器、第二容器和第三容器,所述第一容器包括第一引物组,所述第二容器包括第二引物组,所述第三容器包括第三引物组;其中,所述第一引物组包括如SEQ ID NO.1、2、3、4、5、6、30、31、36、37、38、39、40、41、42、43所示的引物;所述第二引物组包括如SEQ ID NO.7、8、9、10、11、12、13、22、23、24、25、44、45、46、47、48、49、50、51、64所示的引物;所述第三引物组包括如SEQ ID NO.14、15、16、17、18、19、20、21、26、27、28、29、32、33、34、35所示的引物。

6. 如权利要求5所述的试剂盒,其特征在于,所述第一引物组、第二引物组、第三引物组中,各正向引物的5'端包括第一测序标签序列;所述第一引物组、第二引物组、第三引物组中,各反向引物的5'端包括第二测序标签序列。

7. 如权利要求6所述的试剂盒,其特征在于,所述第一测序标签序列如SEQ ID NO.54所示。

8. 如权利要求6所述的试剂盒,其特征在于,所述第二测序标签序列如SEQ ID NO.55所示。

9. 权利要求1所述引物对集在制备检测地中海贫血具体基因型的筛选试剂盒中的应用。

## 扩增子文库构建方法及用于地中海贫血突变型与缺失型基因检测的试剂盒

### 技术领域

[0001] 本发明属于基因检测技术领域,具体的,涉及一种扩增子文库构建方法及用于地中海贫血突变型与缺失型基因检测的试剂盒。

### 背景技术

[0002] 地中海贫血(简称地贫)是一种常见的溶血性单基因遗传病。该病是由于珠蛋白基因的缺陷使血红蛋白中的珠蛋白肽链有一种或几种合成减少或不能合成,致使血红蛋白的组成成分比例失衡,进而导致血红蛋白不稳定,红细胞容易破裂从而引发溶血性贫血。根据缺陷的珠蛋白基因不同,地贫有多种类型,主要类型为 $\alpha$ 地中海贫血(简称 $\alpha$ 地贫)和 $\beta$ 地中海贫血(简称 $\beta$ 地贫)。珠蛋白基因在染色体上以基因簇的形式存在, $\alpha$ 珠蛋白基因簇定位于16号染色体16p13.3, $\alpha$ 地贫主要由 $\alpha$ -珠蛋白基因(HBA1、HBA2基因)缺失引起,少数由 $\alpha$ -珠蛋白基因突变引起; $\beta$ 珠蛋白基因簇定位于11号染色体11p15.3, $\beta$ 地贫主要由 $\beta$ -珠蛋白基因(HBB基因)突变引起,包括点突变及微小插入/缺失,少数的 $\beta$ 地贫由 $\beta$ -珠蛋白基因簇上的大片段缺失引起。

[0003] 目前地贫有多种检测方法,而基因检测是地贫诊断的金标准。常见的地贫基因检测技术有Sanger测序(双脱氧核苷酸末端终止法)、MLPA(多重连接探针扩增技术)、Gap-PCR(跨越断裂点PCR)、PCR-RDB(PCR-反向点杂交)和高通量测序。然而,现有技术各有不足。

[0004] 现有技术使用Gap-PCR结合高通量测序技术进行地贫检测,先使用样本DNA进行多重PCR,所得的PCR产物经纯化后作为建库起始模板,经过打断-末端修复-加“A”-接头连接-片段筛选等步骤完成建库后,进行高通量测序分析,实验操作繁琐耗时长。Sanger测序技术检测地贫通量低,成本较高。MLPA不适用于新发突变检测。Gap-PCR技术也只能检测已知的缺失型,不能检测突变型。PCR-RDB技术(PCR-反向点杂交技术)目前主要应用于 $\alpha$ 地贫或 $\beta$ 地贫的17-23种已知突变型或缺失型检测,无法检测新发突变。

[0005] 现有技术分别存在检测型别少、无法检测新发突变、无法同时检测缺失型和突变型、通量小成本高、操作繁琐耗时长的缺点。另外目前市面上的地贫商用试剂盒仅检测 $\alpha$ 地贫缺失型、 $\alpha$ 地贫突变型及 $\beta$ 地贫突变型,无 $\beta$ 地贫缺失型检测试剂盒。市面上 $\beta$ 地贫缺失型检测能力不足。因此,目前用于地贫检测的技术方法仍有待改进。

### 发明内容

[0006] 本发明的目的在于提供一种检测地中海贫血多种基因型的试剂盒及其应用。

[0007] 在本发明的第一方面,提供了一种扩增子文库的构建方法,所述方法包括步骤:

[0008] S1:第一轮PCR

[0009] 所述第一轮PCR过程中,使用第一引物、第二引物、和第三引物对目标序列进行PCR扩增,获得第一轮PCR扩增产物;

[0010] 其中,所述第一引物从5'端到3'端依次包括第一测序标签序列和目标序列特异性

的正向引物序列;所述第二引物从5'端到3'端依次包括第二测序标签序列和目标序列特异性的反向引物序列;所述第三引物从5'端到3'端依次包括第一测序接头序列、样本标签序列、和第二测序标签序列;

[0011] S2:第二轮PCR

[0012] 所述第二轮PCR过程中,使用第四引物和第五引物对所述第一轮PCR扩增产物进行PCR扩增,获得第二轮PCR扩增产物,经纯化后即得所述扩增子文库;

[0013] 其中,所述第四引物从5'端到3'端依次包括第二测序接头序列和第一测序标签序列;所述第五引物为第一测序接头序列。

[0014] 在另一优选例中,所述步骤S1中,对多个样本来源的目标序列进行PCR扩增,并且针对不同样本来源的所述第三引物的样本标签序列各不相同,从而实现根据不同的样本标签序列区别不同样本。

[0015] 在另一优选例中,所述步骤S1中,所述第一轮PCR为单重或多重PCR。

[0016] 在另一优选例中,所述步骤S2中,所述第二轮PCR在同一PCR反应体系中进行。

[0017] 在另一优选例中,所述第一轮PCR扩增产物的片段长度与高通量测序读长匹配。

[0018] 在本发明的第二方面,提供了一种判断地中海贫血- $\alpha$ 3.7型缺失方法,所述方法包括步骤:

[0019] (1) 构建HBA2和HBA1基因扩增子文库

[0020] 设计引物对,所述引物对扩增片段包含HBA2基因Chr16:223447位置和HBA1基因Chr6:16:227251-227258位置;使用所述引物对样本核酸进行PCR扩增,获得HBA2和HBA1基因扩增子文库;

[0021] (2) 测序并根据测序reads比例判断所述地中海贫血- $\alpha$ 3.7型缺失。

[0022] 在另一优选例中,所述引物对针对HBA2基因和HBA1基因的扩增效率一致。

[0023] 在另一优选例中,所述步骤(2)中,判断方法为:

[0024] 当测序reads比接近2:2时,判断为野生型;

[0025] 当测序reads比接近1:2时,判断为- $\alpha$ 3.7型杂合缺失;

[0026] 当HBA2测序reads为0,而HBA1测序reads正常时,则判断为- $\alpha$ 3.7型纯合缺失。

[0027] 在另一优选例中,所述引物对为:

[0028] 正向引物:GGGAGCGATCTGGGTCGA (SEQ ID NO.1);

[0029] 反向引物:GTGCTCACAGAAGCCAGG (SEQ ID NO.2)。

[0030] 在另一优选例中,所述正向引物的5'端还包括第一测序标签序列。

[0031] 在另一优选例中,所述反向引物的5'端还包括第二测序标签序列。

[0032] 在本发明的第三方面,提供了一种能同时检测 $\alpha$ 地中海贫血和 $\beta$ 地中海贫血缺失型的引物对集,所述引物对集包括:SEQ ID No.1-21所示的引物。

[0033] 在另一优选例中,所述引物对集中的各引物对按照说明书表1所述分组分为三组(三个多重PCR反应体系)。

[0034] 在另一优选例中,所述引物对集包括第一引物组、第二引物组、和第三引物组;其中,所述第一引物组包括SEQ ID NO.1、2、3、4、5、6所示的引物;所述第二引物组包括SEQ ID NO.7、8、9、64、10、11、12、13所示的引物;所述第三引物组包括SEQ ID NO.14、15、16、17、18、19、20、21所示的引物。

- [0035] 在另一优选例中,所述引物对集中,各正向引物的5'端还包括第一测序标签序列。
- [0036] 在另一优选例中,所述引物对集中,各反向引物的5'端还包括第二测序标签序列。
- [0037] 在本发明的第四方面,提供了一种能同时检测 $\alpha$ 地中海贫血和 $\beta$ 地中海贫血突变型的引物对集,所述引物对集包括:SEQ ID No.22-51所示的引物。
- [0038] 在另一优选例中,所述引物对集中的各引物对按照说明书表2所述分组分为三组(三个多重PCR反应体系)。
- [0039] 在另一优选例中,所述引物对集包括第一引物组、第二引物组、和第三引物组;其中,所述第一引物组包括SEQ ID NO.30、31、36、37、38、39、40、41、42、43所示的引物;所述第二引物组包括SEQ ID NO.22、23、24、25、44、45、46、47、48、49、50、51所示的引物;所述第三引物组包括SEQ ID NO.26、27、28、29、32、33、34、35所示的引物。
- [0040] 在另一优选例中,所述引物对集中,各正向引物的5'端还包括第一测序标签序列。
- [0041] 在另一优选例中,所述引物对集中,各反向引物的5'端还包括第二测序标签序列。
- [0042] 在本发明的第五方面,提供了一种能同时检测 $\alpha$ 地中海贫血和 $\beta$ 地中海贫血缺失型和突变型的引物对集,所述引物对集包括:SEQ ID No.1-51所示的引物。
- [0043] 在另一优选例中,所述引物对集中的各引物对按照说明书表1和表2所述分组分为三组(三个多重PCR反应体系)。
- [0044] 在另一优选例中,所述引物对集包括第一引物组、第二引物组、和第三引物组;其中,所述第一引物组包括SEQ ID NO.1、2、3、4、5、6、30、31、36、37、38、39、40、41、42、43所示的引物;所述第二引物组包括SEQ ID NO.7、8、9、64、10、11、12、13、22、23、24、25、44、45、46、47、48、49、50、51所示的引物;所述第三引物组包括SEQ ID NO.14、15、16、17、18、19、20、21、26、27、28、29、32、33、34、35所示的引物。
- [0045] 在另一优选例中,所述引物对集中,各正向引物的5'端还包括第一测序标签序列。
- [0046] 在另一优选例中,所述引物对集中,各反向引物的5'端还包括第二测序标签序列。
- [0047] 在另一优选例中,所述第一测序标签序列如SEQ ID NO.54所示。
- [0048] 在另一优选例中,所述第二测序标签序列如SEQ ID NO.55所示。
- [0049] 在本发明的第六方面,提供了一种试剂盒,所述试剂盒包含本发明第三方面、第四方面、或第五方面所述引物对集。
- [0050] 在另一优选例中,所述试剂盒包括第一容器、第二容器和第三容器,所述第一容器包括第一引物组,所述第二容器包括第二引物组,所述第三容器包括第三引物组;其中,所述第一引物组包括SEQ ID NO.1、2、3、4、5、6、30、31、36、37、38、39、40、41、42、43所示的引物;所述第二引物组包括SEQ ID NO.7、8、9、64、10、11、12、13、22、23、24、25、44、45、46、47、48、49、50、51所示的引物;所述第三引物组包括SEQ ID NO.14、15、16、17、18、19、20、21、26、27、28、29、32、33、34、35所示的引物。
- [0051] 在另一优选例中,所述第一引物组、第二引物组、第三引物组中,各正向引物的5'端包括第一测序标签序列。
- [0052] 在另一优选例中,所述第一引物组、第二引物组、第三引物组中,各反向引物的5'端包括第二测序标签序列。
- [0053] 在另一优选例中,所述第一测序标签序列如SEQ ID NO.54所示。
- [0054] 在另一优选例中,所述第二测序标签序列如SEQ ID NO.55所示。

[0055] 在另一优选例中,所述第一引物组,和/或所述第二引物组,和/或所述第三引物组还包括样本标签引物,所述样本标签引物的5'端到3'端依次包括第一测序接头序列、样本标签序列、和第二测序标签序列,其中所述样本标签序列被设置为根据不同的样本标签序列区别不同样本。

[0056] 在另一优选例中,所述第一测序接头序列如SEQ ID NO.53所示。

[0057] 在另一优选例中,所述试剂盒还包括第四容器,所述第四容器包括第四引物组,所述第四引物组包括正向测序引物和反向测序引物,所述正向测序引物从5'端到3'端依次包括第二测序接头序列和第一测序标签序列;所述反向测序引物为第一测序接头序列。

[0058] 在另一优选例中,所述第二测序接头序列如SEQ ID NO.52所示。

[0059] 在本发明的第七方面,提供了本发明第三方面、第四方面、或第五方面所述引物对集在制备检测地中海贫血具体基因型的试剂盒中的应用。

[0060] 在本发明的第八方面,提供了一种利用本发明第三方面、第四方面、或第五方面所述引物对集或本发明的第六方面所述试剂盒构建地中海贫血基因文库的方法,所述方法包括如下步骤:

[0061] S1:第一轮PCR

[0062] 所述第一轮PCR过程中,分别使用所述第一引物组、所述第二引物组、所述第三引物组配制多重PCR反应体系,对目标序列进行PCR扩增,获得第一轮PCR扩增产物;

[0063] S2:第二轮PCR

[0064] 所述第二轮PCR过程中,使用所述第四引物组对所述第一轮PCR扩增产物进行PCR扩增,获得第二轮PCR扩增产物,经纯化后即得所述地中海贫血基因文库。

[0065] 在本发明的第九方面,提供了一种缺失型地中海贫血的基因分型方法,所述方法包括步骤:

[0066] 1)  $-\alpha 4.2$ 型缺失、 $--SEA$ 型缺失、 $--THAI$ 型缺失、 $\beta-G\gamma+(A\gamma\delta\beta)0$ 缺失、及 $\beta-SEA-HPFH$ 缺失分析:

[0067] 针对各缺失型基因,设计引物对F1-R1和引物对F2-R2;所述引物对F1-R1用于特异性扩增上断裂点上下游基因区段,所述引物对F1-R1中的正向引物F1和上断裂点上游基因区段匹配,所述引物对F1-R1中的反向引物R1和上断裂点下游基因区段匹配;所述引物对F2-R2用于特异性扩增下断裂点上下游基因区段,所述引物对F2-R2中的正向引物F2和下断裂点上游基因区段匹配,所述引物对F1-R1中的反向引物R2和下断裂点下游基因区段匹配;

[0068] 经PCR反应,正向引物F1和反向引物R2扩增出断点相连序列,无正向引物F1和反向引物R1扩增产物,并且无正向引物F2和反向引物R2扩增产物时,判断为纯合缺失型;

[0069] 正向引物F1和反向引物R2扩增出断点相连序列,正向引物F1和反向引物R1扩增出跨越上断裂点的野生型序列,并且正向引物F2和反向引物R2扩增出跨越下断裂点的野生型序列时,判断为杂合缺失型;

[0070] 无正向引物F1和反向引物R2扩增产物,正向引物F1和反向引物R1扩增出跨越上断裂点的野生型序列,并且正向引物F2和反向引物R2扩增出跨越下断裂点的野生型序列时,判断为野生型。

[0071] 在另一优选例中,所述步骤1)中PCR反应被设置为针对不存在地中海贫血缺失型基因(如野生型)的核酸样本,无正向引物F1和反向引物R2扩增产物。

[0072] 在另一优选例中,针对--THAI型缺失:设计的引物对F1-R1为SEQ ID NO.3和SEQ ID NO.5,引物对F2-R2为SEQ ID NO.:6和SEQ ID NO.:4。

[0073] 在另一优选例中,针对- $\alpha$ 4.2型缺失:设计的引物对F1-R1为SEQ ID NO.7和SEQ ID NO.9,引物对F2-R2为SEQ ID NO.:64和SEQ ID NO.:8。

[0074] 在另一优选例中,针对--SEA型缺失:设计的引物对F1-R1为SEQ ID NO.10和SEQ ID NO.12,引物对F2-R2为SEQ ID NO.13和SEQ ID NO.:11。

[0075] 在另一优选例中,针对 $\beta$ -G  $\gamma$  + (A  $\gamma$   $\delta\beta$ ) 0缺失:设计的引物对F1-R1为SEQ ID NO.14和SEQ ID NO.16,引物对F2-R2为SEQ ID NO.17和SEQ ID NO.:15。

[0076] 在另一优选例中,针对 $\beta$ -SEA-HPFH缺失分析:设计的引物对F1-R1为SEQ ID NO.18和SEQ ID NO.20,引物对F2-R2为SEQ ID NO.21和SEQ ID NO.:19。

[0077] 在另一优选例中,所述方法还包括步骤:

[0078] 2) - $\alpha$ 3.7型缺失分析:

[0079] 设计引物对,所述引物对扩增片段包含HBA2基因Chr16:223447位置和HBA1基因Chr6:16:227251-227258位置;使用所述引物对样本核酸进行PCR扩增,获得HBA2和HBA1基因扩增子文库;测序并根据测序reads比例判断所述地中海贫血- $\alpha$ 3.7型缺失;其中,所述引物对针对HBA2基因和HBA1基因的扩增效率一致;当测序reads比接近2:2时,判断为野生型;当测序reads比接近1:2时,判断为- $\alpha$ 3.7型杂合缺失;当HBA2测序reads为0,而HBA1测序reads正常时,则判断为- $\alpha$ 3.7型纯合缺失。

[0080] 在另一优选例中,所述步骤2) 引物对为:

[0081] 正向引物:GGGAGCGATCTGGGTCGA (SEQ ID NO.1);

[0082] 反向引物:GTGCTCACAGAAGCCAGG (SEQ ID NO.2)。

[0083] 在另一优选例中,所述正向引物的5'端还包括第一测序标签序列。

[0084] 在另一优选例中,所述反向引物的5'端还包括第二测序标签序列。

[0085] 在另一优选例中,所述方法中,判断缺失型时,先判断是否存在- $\alpha$ 4.2型缺失、-SEA型缺失、-THAI型缺失,再判断是否存在- $\alpha$ 3.7型缺失。

[0086] 应理解,在本发明范围内中,本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合,从而构成新的或优选的技术方案。限于篇幅,在此不再一一累述。

## 附图说明

[0087] 图1为多重PCR构建测序文库示意图。

[0088] 图2为GAP-PCR检测缺失型原理图。

[0089] 图3为根据碱基序列差异判断- $\alpha$ 3.7型缺失类型原理示意图。

[0090] 图4为对照引物的电泳检测结果。

## 具体实施方式

[0091] 本发明人通过广泛而深入的研究,获得一种多重PCR扩增子文库构建方法,多重PCR的扩增产物片段长度与NGS测序读长匹配。多重PCR后,仅需经过纯化、二轮PCR、纯化三个步骤即可完成建库,实验操作简便快捷。基于该文库构建方法,本发明构建了可以同时检

测 $\alpha$ 地贫和 $\beta$ 地贫的突变型和缺失型的文库并建立了缺失型分析方法。该方法检测范围广,可以检测新发突变,实验操作简便快捷,通量高成本低,很好地弥补现有技术缺点。

[0092] 在描述本发明之前,应当理解本发明不限于所述的具体方法和实验条件,因为这类方法和条件可以变动。还应当理解本文所用的术语其目的仅在于描述具体实施方案,并且不意图是限制性的,本发明的范围将仅由所附的权利要求书限制。

[0093] 除非另外定义,否则本文中所用的全部技术与科学术语均具有如本发明所属领域的普通技术人员通常理解的含义。如本文所用,在提到具体列举的数值中使用术语“约”意指该值可以从列举的值变动不多于1%。例如,如本文所用,表述“约100”包括99和101和之间的全部值(例如,99.1、99.2、99.3、99.4等)。

[0094] 虽然在本发明的实施或测试中可以使用与本发明中所述相似或等价的任何方法和材料,本文在此处例举优选的方法和材料。

[0095] 多重PCR

[0096] 多重PCR(multiplex PCR),又称多重引物PCR或复合PCR,它是在同一PCR反应体系里加上二对以上引物,同时扩增出多个核酸片段的PCR反应,其反应原理,反应试剂和操作过程与一般PCR相同。

[0097] 影响多重PCR反应的因素有很多,比如:

[0098] (1) 反应体系不平衡,反应体系的不平衡导致在前期的几轮反应中某些优势引物及其模板迅速扩增,获得大量的扩增产物,而这些扩增产物同时又是DNA聚合酶的良好抑制剂。所以,随着扩增产物的大量出现,聚合酶的聚合能力被越来越强烈的抑制,因此,前期处于劣势的引物及其模板,这时就更加难以反应,最终导致扩增产物量非常之小,以至于无法检测。

[0099] (2) 引物特异性,如果引物与体系中其他非目的基因片段结合力更强,那么目的基因结合引物的能力就会受到竞争,从而导致扩增效率下降。

[0100] (3) 最佳退火温度不一致,将多对引物放置入一个体系中扩增,由于进行PCR反应的退火温度相同,所以要求每一对引物的最佳退火温度接近。

[0101] (4) 引物二聚体,包括引物间的二聚体以及引物自身所形成的发卡结构,还有一类是第三方DNA介导的聚体,这些二聚体和非特异引物一样都会干扰引物与目标结合位点的竞争,影响扩增效率。

[0102] 虽然上述提及了影响扩增效率的几个因素,但更多的因素尚不清楚。到目前为止,还没有一个可以明确预测扩增效率的有效方法。

[0103] 本发明的一个目的在于提供一种地贫检测的方法,可以同时检测 $\alpha$ 地贫突变型和缺失型以及 $\beta$ 地贫突变型和缺失型。突变型除可以检测常见突变外,还可以检测较为罕见的突变以及新发突变;缺失型除常见的 $\alpha$ 地贫缺失外,还可以检测 $\beta$ 地贫缺失,同时该方法实验操作简便快捷,一次可检测数百个或千个样本,检测通量高成本低,可以弥补现有技术的缺点。

[0104] 本发明的技术方案能够扩大检测地中海贫血的具体基因型种类,提高通量,降低成本。本发明公开了覆盖HBA1、HBA2、HBB基因全外显子区域及内含子区等非编码区的重点区域(包含已知致病性变异)的引物对,可以检测覆盖区域内的所有已知突变型及未知突变型,同时设计引物对中国人群4种 $\alpha$ 地贫缺失型及2种 $\beta$ 地贫缺失型进行检测。4种 $\alpha$ 地贫缺失

型包括- $\alpha 3.7$ 型缺失、- $\alpha 4.2$ 型缺失、--SEA型缺失及--THAI型缺失,2种 $\beta$ 地贫缺失型包括 $\beta$ -G $\gamma$  + (A $\gamma$   $\delta\beta$ ) 0缺失及 $\beta$ -SEA-HPFH缺失。

[0105] 具体的,本发明提供了以下技术方案:

[0106] 第一方面,本申请公开涉及32个扩增子的31对引物。如果分开PCR反应一个样本需要进行31个PCR反应,而不同的引物间扩增区域有重叠,如果混合一起PCR,会有交叉反应,干扰检测。因此,本发明采用多重PCR,并将31对引物分为3组,分组原则为扩增区域有重叠的引物对不能在同一个组,互相影响扩增效率的引物也不能在同一个组。经过上述设计,一个样本仅需3个PCR反应即可捕获目标区域,极大地减少工作量。

[0107] 第二方面,本发明公开了一种利用多重PCR引物的精心设计构建测序文库以及测序的方法。多重PCR的扩增产物片段长度与高通量测序读长匹配,3个PCR反应可使用同样的反应程序,且多重PCR后不需针对片段长度再做处理。多重PCR产物经纯化后,仅需使用一对引物进行第二轮PCR及一次纯化,即可完成建库,实验操作简便快捷,极大地减少时间成本。

[0108] 本发明在多重PCR中,使用引物1、引物2和引物3,在引物3中加入样本标签(Index),因此扩增产物带上了Index。同个样本的扩增产物带上同样的Index,不同样本的扩增产物带上不同的Index,多重PCR后,将所有样本的扩增产物混合为1管,后续的建库实验由多管操作简化为1管,极大地减少工作量(参见图1)。

[0109] 建库完成后的文库根据高通量测序仪步骤要求进行上机测序。测序完成后,根据样本标签将测序reads归到不同的样本,从而得到单个样本的测序数据。不同的样本标签设计数量可达上万条,而有多少条不同样本标签即可混合多少个样本进行检测。基于高通量测序的特性,可以实现单次检测数百个样本或千个样本,并且单个样本检测成本低。

[0110] 获得单个样本测序数据后,经序列比对分析即可获得样本检测结果。

[0111] 第三方面,本申请公开了不同缺失型地贫的分析方法:

[0112] 1) - $\alpha 4.2$ 型缺失、--SEA型缺失、--THAI型缺失及 $\beta$ -G $\gamma$  + (A $\gamma$   $\delta\beta$ ) 0缺失、 $\beta$ -SEA-HPFH缺失检测。采用GAP-PCR原理,分别在上、下断裂点附近设计引物(参见图2)。当发生杂合缺失时,会扩增断点相连序列以及野生型序列,而发生纯合缺失时,只会扩增出断点相连序列。本发明根据不同缺失型断裂点创建对应的断点相连序列,测序得到的序列,与参考基因组及创建的缺失型断点相连序列比对。通过测序序列比对及测序reads分析即可判断是否存在缺失。

[0113]

变异类型	F1/R1	F2/R2	F1/R2
野生型	√	√	×
杂合缺失	√	√	√
纯合缺失	×	×	√

[0114] 2) 对于- $\alpha 3.7$ 型缺失,由于HBA2和HBA1基因的高度同源性,无法通过GAP-PCR原理进行检测。本发明采用的方法为基于HBA2基因与HBA1基因差异碱基测序reads比例差异进行判定。通过基因序列比对,发现HBA2基因Chr16:223447位置碱基为G(位于缺失断裂点下游,标记为HBA2'),而HBA1基因同一位置Chr6:16:227251-227258碱基为CTCGGCC(位于缺失断裂点下游,标记为HBA1')。基于这种差异性,设计扩增两个区域的引物,为使两个区域扩增效率一致,设计一对引物同时扩增两个区域,对应产生两个扩增子。野生型HBA2'和HBA1'的测序reads比接近2:2,当发生- $\alpha 3.7$ 型杂合缺失时,HBA2'和HBA1'的测序reads比接

近1:2;当发生- $\alpha$ 3.7型纯合缺失时,HBA2' 测序reads为0,而HBA1' 正常;如图3所示。

[0115] 3) 因为存在- $\alpha$ 4.2型缺失、-SEA型缺失、-THAI型缺失时,HBA1和HBA2基因中其一或两个基因会缺失。因此,判断缺失型时,先判断是否存在- $\alpha$ 4.2型缺失、-SEA型缺失、-THAI型缺失,再判断是否存在- $\alpha$ 3.7型缺失,判断依据如下表:

基因型	非- $\alpha$ 3.7 缺失型对应断点相连序列	HBA2' 与 HBA1' Reads 比例
( $\alpha \alpha / \alpha \alpha$ )	×	2:2
( $\alpha \alpha / -\alpha 3.7$ )	×	1:2
( $-\alpha 3.7 / -\alpha 3.7$ )	×	0:2
( $\alpha \alpha / -\alpha 4.2$ )	√	1:2
( $-\alpha 4.2 / -\alpha 3.7$ )	√, 且有上、下断裂点野生型序列	0:2
( $-\alpha 4.2 / -\alpha 4.2$ )	√, 且无上、下断裂点野生型序列	0:2
( $\alpha \alpha / --SEA$ )	√	1:1
( $--SEA / -\alpha 3.7$ )	√	0:1
( $--SEA / --SEA$ )	√	0:0
( $\alpha \alpha / --THAI$ )	√	1:1
( $--THAI / -\alpha 3.7$ )	√	0:1
( $--THAI / --THAI /$ )	√	0:0

[0118] 另外,针对突变型地贫的检测:

[0119] 针对HBA1、HBA2、HBB基因,设计引物覆盖基因全外显子区域及内含子区等非编码区的重点区域(包含已知致病性变异)。测序序列比对参考基因组目标区域,经过分析注释,找到其中点突变和微小插入/缺失,从而得到地贫突变型检测结果。

[0120] 本发明的有益效果包括:

[0121] 1) 本发明突变型检测区域包括HBA1、HBA2、HBB基因所有外显子区域及内含子区等非编码区的重点区域(包含已知致病性变异),除能检测 $\alpha$ 地贫及 $\beta$ 地贫的常见突变外,还可以检测较罕见突变及新发突变。缺失型检测包括4种 $\alpha$ 地贫缺失型:- $\alpha$ 3.7型缺失、- $\alpha$ 4.2型缺失、-SEA型缺失及-THAI型缺失,2种 $\beta$ 地贫缺失型: $\beta$ -G $\gamma$ +(A $\gamma$  $\delta$  $\beta$ )0缺失及 $\beta$ -SEA-HPFH缺失。本发明可以同时检测 $\alpha$ 地贫和 $\beta$ 地贫的突变型和缺失型,检测范围广,能检测新发突变,可以弥补市面上商业试剂盒 $\beta$ 地贫缺失型检测能力不足的缺点。

[0122] 2) 本发明为多重PCR扩增子建库,多重PCR的扩增产物片段长度与NGS测序读长匹配。多重PCR后,仅需经过纯化、二轮PCR、纯化三个步骤即可完成建库,实验操作简便快捷。

[0123] 3) 本发明通过设计样本标签,在多重PCR中引入样本标签,多重PCR后所有样本的PCR产物可以混合为1管,因此后续的实验由多管操作简化为1管,极大地减少工作量。

[0124] 4) 本发明的样本标签的设计数量可达上万条,有多少条样本标签即可混合多少个样本进行测序。结合高通量测序技术,本发明单次可检测数百个或千个样本,检测通量高,单个样本检测成本低。

[0125] 综合而言,本发明可以同时检测 $\alpha$ 地贫和 $\beta$ 地贫的突变型和缺失型,检测范围广,可以检测新发突变,实验操作简便快捷,通量高成本低,很好地弥补现有技术缺点。

[0126] 本发明适用于检测 $\alpha$ 地贫和 $\beta$ 地贫的突变型和缺失型,有利于地贫早期诊断。另外,

本发明的方法同样适用于非诊断的目的,例如,在新药研发过程中,使用本发明的检测方法获得用作中间结果的基因突变信息,这些基因突变信息可以作为公众健康管理之需,也可以用于地贫基因突变相关研究及靶向新药研发。

[0127] 下面结合具体实施例,进一步详陈本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明详细条件的实验方法,通常按照常规条件如美国Sambrook.J等著《分子克隆实验室指南》(黄培堂等译,北京:科学出版社,2002年)中所述的条件,或按照制造厂商所建议的条件。除非另外说明,否则百分比和份数按重量计算。以下实施例中所用的实验材料和试剂如无特别说明均可从市售渠道获得。

[0128] 实施例1引物设计及合成

[0129]  $-\alpha 4.2$ 型缺失、 $-\text{SEA}$ 型缺失、 $-\text{THAI}$ 型缺失及 $\beta\text{-G}\gamma + (\text{A}\gamma\delta\beta) 0$ 缺失和 $\beta\text{-SEA-HPFH}$ 缺失,采用GAP-PCR原理,分别在上、下断点附近设计特异性引物,共15对引物(15个扩增子); $-\alpha 3.7$ 型缺失根据HBA1和HBA2差异碱基,设计1对特异性引物同时扩增HBA1和HBA2的差异碱基区域,产生2个扩增子,该2个扩增子可以同时检测突变型。根据HBA1、HBA2、HBB基因的外显子区及内含子区等非编码区的重点区域(包含已知致病性变异)设计特异性PCR引物,共15对引物(15个扩增子)。在扩增子正向引物5'端加上测序标签1(Rd1)序列即为引物1;扩增子反向引物5'端加上测序标签2(Rd2)即为引物2;由测序接头P7、样本标签(Index)、Rd2序列组合为引物3;引物4由测序接头P5与Rd1组成;引物5为测序接头P7序列。引物设计完成后,由生工生物工程(上海)有限公司合成。

[0130] 共合成31对引物,分为3个引物组,引物组1共9对引物,引物组2共12对引物,引物组3共10对引物。每对引物根据扩增效率差异设定不同的终浓度。

[0131] 本发明中的接头序列P5、P7为Illumina测序平台的接头序列;Rd1、Rd2为Illumina测序平台的测序标签,序列如下:

[0132] P5序列:AATGATACGGCGACCACCGAGATCTACAC (SEQ ID NO.52);

[0133] P7序列:CAAGCAGAAGACGGCATAACGAGAT (SEQ ID NO.53);

[0134] Rd1序列:ACACTCTTTCCCTACACGACGCTCTTCCGAT (SEQ ID NO.54);

[0135] Rd2序列:GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC (SEQ ID NO.55)。

[0136] 表1.缺失型引物信息表

[0137]

检测类型	扩增子名称及描述	特异性引物名称	引物序列 (5' -3' )	SEQ ID NO.	分组
-α 3.7型缺失	Amp1: HBA2', 可同时检测HBA2突变型	3.7-F	GGGAGCGATCTGGGTCGA	1	引物组1
		3.7-R	GTGCTCACAGAAGCCAGG	2	
	Amp2: HBA1', 可同时检测HBA1突变型	3.7-F	GGGAGCGATCTGGGTCGA	1	引物组1
		3.7-R	GTGCTCACAGAAGCCAGG	2	
--THAI型缺失	Amp3: 断点相连序列	THAI-u-F	AAGAAAACATTTGCCTTTGA	3	引物组1
		THAI-d-R	GGATCTGCACCTCTGGGTA	4	
	Amp4: 跨越上断裂点	THAI-u-F	AAGAAAACATTTGCCTTTGA	3	引物组1
		THAI-u-R	TGGGTGAACCGTATGGTAT	5	
	Amp5: 跨越下断裂点	THAI-d-F	GAGCATCACTGAAGGTGGA	6	引物组1
		THAI-d-R	GGATCTGCACCTCTGGGTA	4	
-α 4.2型缺失	Amp6: 断点相连序列	4.2-u-F	TGAATGAAGTCCGAGTAGGC	7	引物组2
		4.2-d-R	ACCCTCACTCCAGCCAC	8	
	Amp7: 跨越上断裂点	4.2-u-F	TGAATGAAGTCCGAGTAGGC	7	引物组2
		4.2-u-R	CACCTCACTCCAGCCACC	9	
	Amp8: 跨越下断裂点	4.2-d-F	TGGTCTTTGAATAAAGTCTGAGT	64	引物组2
		4.2-d-R	ACCCTCACTCCAGCCAC	8	
--SEA型缺失	Amp9: 断点相连序列	SEA-u-F	GCCCTTCACCCTCCACAGT	10	引物组2
		SEA-d-R	CCATGCTGGAGTGCAGTGTT	11	
	Amp10: 跨越上断裂点	SEA-u-F	GCCCTTCACCCTCCACAGT	10	引物组2
		SEA-u-R	AAACGCCCGTCCGACTCAA	12	
	Amp11: 跨越下断裂点	SEA-d-F	CCAAGAATATCCTGAGCAAC	13	引物组2

[0138]

	点	SEA-d-R	CCATGCTGGAGTGCAGTGTT	11	
β-Gγ+(Aγδβ)0 缺失	Amp12: 断点相连序列	β-G-u-F	CATCAGAGGTAAGCAGGG	14	引物组3
		β-G-d-R	TCTGGAGGCAGGACAAGT	15	
	Amp13: 跨越上断裂点	β-G-u-F	CATCAGAGGTAAGCAGGG	14	引物组3
		β-G-u-R	TTGTGGGAGGAGGCATAGA	16	
Amp14: 跨越下断裂点	β-G-d-F	GCTATCTTCTTGCCACCAT	17	引物组3	
	β-G-d-R	TCTGGAGGCAGGACAAGT	15		
β-SEA-HPFH缺失	Amp15: 断点相连序列	β-SEA-u-F	TCATGGTAGCAGAATCA	18	引物组3
		β-SEA-d-R	CATAAGAAGTTGTAGGGTA	19	
	Amp16: 跨越上断裂点	β-SEA-u-F	TCATGGTAGCAGAATCA	18	引物组3
		β-SEA-u-R	ACATAAAGGGCAGCCAAAT	20	
	Amp17: 跨越下断裂点	β-SEA-d-F	CAATGAGGATCTAGGCACA	21	引物组3
β-SEA-d-R	CATAAGAAGTTGTAGGGTA	19			

[0139] 表2. 突变型引物信息表

[0140]

检测基因	扩增子名称及描述	特异性引物名称	引物序列 (5' -3' )	SEQ NO.	ID	分组
HBA2	Amp18	HBA2-F1	GTGAGGCTCCCTCCCCT	22		引物组2
		HBA2-R1	CGTTGGGCATGTGCTCC	23		
HBA2	Amp19	HBA2-F2	TCTCAGGGCAGAGGATCA	24		引物组2
		HBA2-R2	GGCAGGAGGAACGGCTA	25		
HBA2	Amp20	HBA2-F3	TTCTGTCTTCCCCACCA	26		引物组3
		HBA2-R3	CAACCCGCGTGATCCTCT	27		
HBA2	Amp21	HBA2-F4	GCTGACCTCCAAATACCGTT	28		引物组3
		HBA2-R4	CAGTTCTCCCTCCCCAGC	29		
HBA1	Amp22	HBA1-F1	GCCCAAGCATAAACCCCT	30		引物组1
		HBA1-R1	AGGACGGTTGAGGGTGG	31		
HBA1	Amp23	HBA1-F2	TCCTGTCTTCCCCACCA	32		引物组3
		HBA1-R2	AACCCGCGTGATCCTCTG	33		
HBA1	Amp24	HBA1-F3	CTCTTCTCTGCACAGCTC	34		引物组3
		HBA1-R3	CACGTTTGCTGAGGAAA	35		
HBB	Amp25	HBB-F1	GCTGTTGCCAATGTGCATT	36		引物组1
		HBB-R1	TGTTCCCTAAGTCCAACTACT	37		
HBB	Amp26	HBB-F2	CAGGAGCTGTGGGAGGAA	38		引物组1
		HBB-R2	CCTAATCTCTTTCTTCAGGG	39		
HBB	Amp27	HBB-F3	CCCATAGACTCACCTGAAG	40		引物组1
		HBB-R3	TGCCTATTGGTCTATTTTCCC	41		
HBB	Amp28	HBB-F4	TTTCTATTGGTCTCCTTAAACCTG	42		引物组1
		HBB-R4	CTAGGGTTGCCAATCTACT	43		
HBB	Amp29	HBB-F5	AGATGCTCAAGGCCCTTCAT	44		引物组2
		HBB-R5	CTTTTGCTAATCATGTTTCATACCTC	45		
HBB	Amp30	HBB-F6	AAACTAAAACGATCCTGAGAC	46		引物组2
		HBB-R6	GAAAGTGCTCGGTGCCTTTA	47		
HBB	Amp31	HBB-F7	CAAGGGTAGACCACCAGCA	48		引物组2
		HBB-R7	GAAGTCTGCCGTTACTGC	49		
HBB	Amp32	HBB-F8	GCTCTGCCCTGACTTTTATG	50		引物组2
		HBB-R8	TGCTTACCAAGCTGTGATTC	51		

[0141] 2、样本DNA提取

[0142] 来自血液、唾液和组织等样本DNA均可用于扩增。用核酸自动提取仪提取DNA后，Qubit Flurometer 3.0定量，检测一个样本需要DNA模板约100-150ng。

[0143] 3、一轮PCR反应

[0144] 采用以下PCR反应体系及PCR程序对获得的每份基因组DNA样本做多重PCR扩增，分3个反应体系进行，其中多重PCR反应体系1对应引物组1，多重PCR反应体系2对应引物组2，多重PCR反应体系3对应引物组3。配制96个样本所需扩增子上游引物(引物1)混合池、下游引物(引物2)混合池，其中每个引物组的引物总终浓度为10 $\mu$ M；96种不同样本标签的引物3终浓度定为10 $\mu$ M。，同一个样本的3个多重PCR反应体系中加入相同样本标签的引物3，具体反应体系如下：

[0145] 多重PCR反应体系1(25 $\mu$ l)：

	组份	体积 ( $\mu$ l )
	PCR buffer	12.5
	上游引物混合池1 (10 $\mu$ M)	0.5
[0146]	下游引物混合池1 (10 $\mu$ M)	0.2
	引物3 (10 $\mu$ M)	0.5
	DNA polymerase (5U/ $\mu$ l)	0.6
	DNA模板	3
	H <sub>2</sub> O	7.7

[0147] 多重PCR反应体系2(25 $\mu$ l) :

	组份	体积 ( $\mu$ l )
	PCR buffer	12.5
	上游引物混合池2 (10 $\mu$ M)	0.5
[0148]	下游引物混合池2 (10 $\mu$ M)	0.2
	引物3 (10 $\mu$ M)	0.5
	DNA polymerase (5U/ $\mu$ l)	0.6
	DNA模板	3
	H <sub>2</sub> O	7.7

[0149] 多重PCR反应体系3(25 $\mu$ l) :

	组份	体积 ( $\mu$ l )
	PCR buffer	12.5
[0150]	上游引物混合池2 (10 $\mu$ M)	0.5
	下游引物混合池2 (10 $\mu$ M)	0.2
	引物3 (10 $\mu$ M)	0.5
	DNA polymerase (5U/ $\mu$ l)	0.6
[0151]	DNA模板	3
	H <sub>2</sub> O	7.7

[0152] 将PCR管置于PCR仪中,PCR程序如下:

	温度	时间	
	95°C	3 min	
	95°C	30s	} 20 cycle
[0153]	58°C	30s	
	72°C	1min	
	72°C	3 min	
	4°C	∞	

[0154] 4、一轮PCR产物纯化

[0155] 反应完成后,将96个样本的PCR产物,每管取10 $\mu$ l合并为1管,震荡混匀后取500 $\mu$ l用0.8倍体积的AMPure XP Beads纯化2次,溶于50 $\mu$ l的TE。

[0156] 5、二轮PCR反应

[0157] 添加另一测序接头P5序列,此轮PCR只需两条引物,引物4为测序接头P5序列与Rd1组合,引物5为测序接头P7序列。PCR反应50 $\mu$ l体系:

	组份	体积 ( $\mu$ l)
	2 X KAPA HiFi HotStart ReadyMix	25
[0158]	引物4 (25 $\mu$ M)	1.5
	引物5 (25 $\mu$ M)	1.5
	一轮PCR纯化产物	10
	H <sub>2</sub> O	12

[0159] 将PCR管置于PCR仪中,PCR程序如下:

	温度	时间	
	95°C	3 min	
	95°C	30s	} 15 cycle
[0160]	60°C	30s	
	72°C	40s	
	72°C	3 min	
	4°C	∞	

[0161] 6、二轮PCR产物纯化

[0162] 反应完成后,PCR产物用0.8倍体积的AMPure XP Beads纯化1次,溶于50 $\mu$ l的TE,完成测序文库构建。

[0163] 7、文库定量、上机测序

[0164] 参考Qubit Flurometer 3.0说明书,对二轮PCR后的扩增子文库进行准确定量。产物经文库检测合格后,使用Illumina测序平台进行PE150上机测序,步骤严格按照供应商要求进行。

[0165] 8、测序数据信息分析

[0166] 对测序所得数据进行低质量序列以及接头序列过滤后,利用比对软件BWA将其比对到参考基因组 (GRCh37/hg19) 上的目标区域及创建的断点相连序列。

[0167] 9、测序结果

[0168] 96个已知基因型样本,检测结果与已知结果100%相符。

[0169] 表3 96个已知基因型样本检测结果

[0170]

样本编号	样本基因型	检测基因型	是否相符
S1	$\alpha \alpha / -\alpha 3.7$	$\alpha \alpha / -\alpha 3.7$	是
S2	$\beta$ -SEA-HPFH 杂合缺失	$\beta$ -SEA-HPFH 杂合缺失	是
S3	N	N	是
S4	$-\alpha 3.7 / -\alpha 4.2$	$-\alpha 3.7 / -\alpha 4.2$	是
S5	$-\alpha 4.2 / -\alpha 4.2$	$-\alpha 4.2 / -\alpha 4.2$	是
S6	Codons 27/28 (+C) beta0 杂合	Codons 27/28 (+C) beta0 杂合	是
S7	N	N	是
S8	Codons 27/28 (+C) beta0 杂合	Codons 27/28 (+C) beta0 杂合	是
S9	N	N	是
S10	$\alpha \alpha / --SEA$	$\alpha \alpha / --SEA$	是
S11	Codon 17 (A->T) beta0 杂合	Codon 17 (A->T) beta0 杂合	是
S12	$\alpha \alpha / --SEA$	$\alpha \alpha / --SEA$	是
S13	$\alpha \alpha / --SEA$	$\alpha \alpha / --SEA$	是
S14	-28 (A->G) beta+ 杂合	-28 (A->G) beta+ 杂合	是
S15	$\alpha \alpha / --SEA$	$\alpha \alpha / --SEA$	是
S16	$\alpha \alpha / --SEA$	$\alpha \alpha / --SEA$	是
S17	$\alpha \alpha / --SEA$ 与-29 (A->G) beta+杂合	$\alpha \alpha / --SEA$ 与-29 (A->G) beta+杂合	是
S18	$\alpha \alpha / -\alpha 3.7$	$\alpha \alpha / -\alpha 3.7$	是
S19	$\alpha \alpha / -\alpha 3.7$	$\alpha \alpha / -\alpha 3.7$	是
S20	IVS-II-654 (C->T) beta+杂合	IVS-II-654 (C->T) beta+杂合	是
S21	$\alpha \alpha / -\alpha 4.2$	$\alpha \alpha / -\alpha 4.2$	是

[0171]

S22	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S23	$\alpha\alpha/-\alpha 4.2$	$\alpha\alpha/-\alpha 4.2$	是
S24	Codons 41/42 (-TTCT) beta0 杂合	Codons 41/42 (-TTCT) beta0 杂合	是
S25	N	N	是
S26	Codon 17 (A->T) beta0 杂合	Codon 17 (A->T) beta0 杂合	是
S27	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S28	$\alpha\alpha/-\alpha 4.2$	$\alpha\alpha/-\alpha 4.2$	是
S29	Hb Westmead 杂合	Hb Westmead 杂合	是
S30	$\alpha\alpha/-\alpha 4.2$	$\alpha\alpha/-\alpha 4.2$	是
S31	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S32	-28 (A->G) beta+ 杂合	-28 (A->G) beta+ 杂合	是
S33	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S34	Codon 17 (A->T) beta0 杂合	Codon 17 (A->T) beta0 杂合	是
S35	$-\alpha 3.7/--SEA$	$-\alpha 3.7/--SEA$	是
S36	Codon 17 (A->T) beta0 杂合	Codon 17 (A->T) beta0 杂合	是
S37	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S38	Hb Westmead 杂合	Hb Westmead 杂合	是
S39	IVS-II-654 (C->T) beta+ 杂合	IVS-II-654 (C->T) beta+ 杂合	是
S40	$-\alpha 4.2/--SEA$	$-\alpha 4.2/--SEA$	是
S41	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S42	-29 (A->G) beta+ 杂合	-29 (A->G) beta+ 杂合	是
S43	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S44	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S45	Hb Westmead 杂合	Hb Westmead 杂合	是
S46	-28 (A->G) beta+ 杂合	-28 (A->G) beta+ 杂合	是
S47	-29 (A->G) beta+ 杂合	-29 (A->G) beta+ 杂合	是
S48	$\beta$ -SEA-HPFH 杂合缺失	$\beta$ -SEA-HPFH 杂合缺失	是
S49	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S50	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S51	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S52	Hb Quong Sze 杂合	Hb Quong Sze 杂合	是
S53	Codons 27/28 (+C) beta0 杂合	Codons 27/28 (+C) beta0 杂合	是
S54	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是

[0172]

S55	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S56	$\alpha\alpha/-\alpha 4.2$	$\alpha\alpha/-\alpha 4.2$	是
S57	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S58	Codons 71/72 (+A) beta0 杂合	Codons 71/72 (+A) beta0 杂合	是
S59	$-\alpha 3.7/-\alpha 3.7$	$-\alpha 3.7/-\alpha 3.7$	是
S60	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S61	$\alpha\alpha/--SEA$ 与 Codons 41/42 (-TTCT) beta0 杂合	$\alpha\alpha/--SEA$ 与 Codons 41/42 (-TTCT) beta0 杂合	是
S62	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S63	N	N	是
S64	--THAI 杂合缺失	--THAI 杂合缺失	是
S65	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S66	$\alpha\alpha/--SEA$	$\alpha\alpha/--SEA$	是
S67	$\alpha\alpha/-\alpha 4.2$	$\alpha\alpha/-\alpha 4.2$	是
S68	IVS-II-654 (C->T) beta+杂合	IVS-II-654 (C->T) beta+杂合	是
S69	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S70	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S71	-29 (A->G) beta+杂合	-29 (A->G) beta+杂合	是
S72	$-\alpha 3.7/--SEA$	$-\alpha 3.7/--SEA$	是
S73	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S74	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S75	$\beta-G\gamma+(A\gamma\delta\beta)0$ 杂合缺失	$\beta-G\gamma+(A\gamma\delta\beta)0$ 杂合缺失	是
S76	$-\alpha 4.2/--SEA$	$-\alpha 4.2/--SEA$	是
S77	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S78	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S79	Codons 41/42 (-TTCT) beta0 杂合	Codons 41/42 (-TTCT) beta0 杂合	是
S80	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S81	$-\alpha 3.7/--SEA$	$-\alpha 3.7/--SEA$	是
S82	$\alpha\alpha/-\alpha 3.7$	$\alpha\alpha/-\alpha 3.7$	是
S83	Codon 17 (A->T) beta0 杂合	Codon 17 (A->T) beta0 杂合	是
S84	Hb Constant Spring (Hb CS) 杂合	Hb Constant Spring (Hb CS) 杂合	是
S85	$\alpha\alpha/-\alpha 4.2$ 合并 Hb Westmead 杂合	$\alpha\alpha/-\alpha 4.2$ 合并 Hb Westmead 杂合	是
S86	Codons 41/42 (-TTCT) beta0 杂合	Codons 41/42 (-TTCT) beta0 杂合	是
S87	$-\alpha 4.2/--SEA$	$-\alpha 4.2/--SEA$	是

[0173]	S88	Codons 41/42 (-TTCT) beta0 杂合	Codons 41/42 (-TTCT) beta0 杂合	是
	S89	Hb Westmead 杂合	Hb Westmead 杂合	是
	S90	$\alpha\alpha$ / - $\alpha$ 3.7 合并 Hb Quong Sze 杂合	$\alpha\alpha$ / - $\alpha$ 3.7 合并 Hb Quong Sze 杂合	是
	S91	Codons 27/28 (+C) beta0 杂合	Codons 27/28 (+C) beta0 杂合	是
	S92	$\alpha\alpha$ / - $\alpha$ 4.2	$\alpha\alpha$ / - $\alpha$ 4.2	是
	S93	N	N	是
	S94	$\alpha\alpha$ / - $\alpha$ 4.2 与 IVS-II-654 (C->T) beta+杂合	$\alpha\alpha$ / - $\alpha$ 4.2 与 IVS-II-654 (C->T) beta+杂合	是
	S95	- $\alpha$ 3.7 / --SEA	- $\alpha$ 3.7 / --SEA	是
	S96	Hb Westmead 杂合	Hb Westmead 杂合	是

[0174] 以下列出其中3个样本的结果分析,其中样本S1基因型为 $\alpha\alpha$  / -  $\alpha$  3.7,样本S2基因型为 $\beta$ -SEA-HPFH杂合缺失,样本S3为阴性样本。

[0175] 表4.1.  $\alpha\alpha$  / -  $\alpha$  3.7基因型测序结果

[0176]	样本	断点相连序列	Amp1:Amp2
	S1	×	1:2
	S2	×	2:2
	S3	×	2:2

[0177] 表4.2.  $\beta$ -SEA-HPFH杂合缺失测序结果

[0178]	样本	Amp15	Amp16	Amp17
	S1	×	√	√
	S2	√	√	√
	S3	×	√	√

[0179] 以上结果,Amp1和Amp2是- $\alpha$ 3.7型缺失检测的扩增子。Amp1为HBA2' 对应扩增子,Amp2为HBA1' 对应扩增子。三个样本均无检出- $\alpha$ 4.2型缺失、--SEA型缺失及--THAI型缺失的断点相连序列。S1样本,基因型为- $\alpha$ 3.7/ $\alpha\alpha$ ,Amp1和Amp2的测序reads比约为1:2,不存在此变异的样本S2和S3的Amp1比Amp2比例为2:2。

[0180] Amp16和Amp17为 $\beta$ -SEA-HPFH缺失的跨越上、下断裂点的野生型扩增子,Amp15为 $\beta$ -SEA-HPFH缺失断点相连的扩增子。样本S2基因型为 $\beta$ -SEA-HPFH杂合缺失,除检测到野生型扩增子Amp16和Amp17外,还检测到断点相连扩增子Amp15;而不存在此变异的样本S1和S3,则检出Amp16和Amp17,未检出Amp15。

[0181] 对比例1

[0182] 本发明人在研究过程中,针对各目标序列筛选了数对甚至数十对PCR引物,最终筛选出了特异性、扩增效率、多重PCR扩增效果能够满足检测需求的引物。

[0183] 例如,针对某一种变异,先设计3-5对引物,每一对引物进行特异性及扩增效果验证,验证方案为每一对引物进行PCR,然后进行琼脂糖凝胶电泳,得到电泳条带,如果条带单一且为目标长度或者条带亮度相对更亮的,判断为更优的引物,同时扩增产物进行Sanger测序,验证目标长度条带是否正确。选出其中更优的2-3对引物,检测不同基因型的引物设

计多种组合进行多重PCR,再进行NGS测序,通过测序reads数判断扩增效果及引物间是否存在互相干扰的情况。同等样本平均测序深度下,单对引物测序reads数越多,说明测序效果越好,引物扩增效果及引物组合效果越优。不同的引物扩增效果会有差异,测序reads数有差异。经过以上测试的引物,如果效果不理想,再重新设计引物进行测试,直至达到要求。

[0184] 例如,针对Amp11扩增子,本发明人设计的部分典型引物序列如下:

[0185] 对照上游引物1:ACCAAACAGAGCCATTTACCT (SEQ ID NO.56)。

[0186] 对照下游引物1:CCTGACCTCGTGATCCACC (SEQ ID NO.57)。

[0187] 对照上游引物2:CAAACAGAGCCATTTACCTT (SEQ ID NO.58)。

[0188] 对照下游引物2:CCATGCTGGAGTGCAGTGTTG (SEQ ID NO.59)。

[0189] 具体检测步骤、检测条件同以上实施例。使用对照引物1的检测结果如图4中泳道2所示,检测结果表明该引物对特异性较差。使用对照引物2的检测结果如图4中泳道1及表5所示,检测结果表明该引物对虽然特异性较佳,但是在多重PCR反应体系中被抑制,测序reads少,扩增效率较低。

[0190] 表5. 阴性样本对照引物对2测序结果

测试样本	Amp10测序reads数	Amp11测序reads数
Y1	1031x	30x

[0192] 对比例2

[0193] 针对- $\alpha$ 3.7型缺失的引物也是经过数十次实验获得的。由于需要使用同一对引物进行HBA2' 和HBA1' 的扩增,且需要保证扩增效率一致,以便真实反映HBA2' 和HBA1' 的拷贝数比例。使用- $\alpha$ 3.7型缺失阳性样本通过NGS测序验证来选择更优的引物,测序reads数多、HBA2' 与HBA1' 比例符合预期结果且多样本测试HBA2' 与HBA1' 比例稳定的引物为更优的引物。

[0194] 例如,针对- $\alpha$ 3.7型缺失扩增子,本发明人设计的部分典型引物序列如下:

[0195] 对照上游引物3:GGGAGCGATCTGGGTCGA (SEQ ID NO.60)。

[0196] 对照下游引物3:AGGGTCACCAGCAGGCAG (SEQ ID NO.61)。

[0197] 对照上游引物4:GGAGCGATCTGGGTCGAG (SEQ ID NO.62)。

[0198] 对照下游引物4:GGTATTTGGAGGTCAGCACG (SEQ ID NO.63)。

[0199] 具体检测步骤、检测条件同以上实施例。使用对照引物3的检测结果如表6所示,检测结果表明该引物针对HBA2' 和HBA1' 的扩增效率差异较大。使用对照引物4的检测结果如表7所示,检测结果表明该引物对虽然初筛时针对HBA2' 和HBA1' 的扩增效率较为一致,但是在多样本测试中,针对HBA2' 及HBA1' 的扩增稳定性较差,无法用于实际检测。

[0200] 表6. 阴性样本对照引物对3测序结果

测试样本	Amp1测序reads数	Amp2测序reads数
Y2	650	932

[0202] 表7. 阴性样本对照引物对4测序结果

阶段性测试	测试样本	Amp1测序reads数	Amp2测序reads数
初筛	Y3	973	982
	Y4	780	785
多样本测试	Y5	684	742
	Y6	773	1126
	Y7	992	1363

[0203]

[0204] 在本发明提及的所有文献都在本申请中引用作为参考,就如同每一篇文献被单独引用作为参考那样。此外应理解,在阅读了本发明的上述讲授内容之后,本领域技术人员可以对本发明作各种改动或修改,这些等价形式同样落于本申请所附权利要求书所限定的范围。

- [0001] 序列表
- [0002] <110> 深圳联合医学科技有限公司
- [0003] <120> 扩增子文库构建方法及用于地中海贫血突变型与缺失型基因检测的试剂盒
- [0004] <130> 00002
- [0005] <140> CN2021111995780
- [0006] <141> 2021-10-14
- [0007] <160> 64
- [0008] <170> PatentIn version 3.5
- [0009] <210> 1
- [0010] <211> 18
- [0011] <212> DNA
- [0012] <213> 人工序列
- [0013] <400> 1
- [0014] gggagcgatec tgggtcga 18
- [0015] <210> 2
- [0016] <211> 18
- [0017] <212> DNA
- [0018] <213> 人工序列
- [0019] <400> 2
- [0020] gtgctcacag aagccagg 18
- [0021] <210> 3
- [0022] <211> 20
- [0023] <212> DNA
- [0024] <213> 人工序列
- [0025] <400> 3
- [0026] aagaaaacat ttgcctttga 20
- [0027] <210> 4
- [0028] <211> 19
- [0029] <212> DNA
- [0030] <213> 人工序列
- [0031] <400> 4
- [0032] ggatctgcac ctctgggta 19
- [0033] <210> 5
- [0034] <211> 19
- [0035] <212> DNA
- [0036] <213> 人工序列
- [0037] <400> 5

[0038] tgggtgaacc gtatggtat 19  
[0039] <210> 6  
[0040] <211> 19  
[0041] <212> DNA  
[0042] <213> 人工序列  
[0043] <400> 6  
[0044] gagcatcact gaaggtgga 19  
[0045] <210> 7  
[0046] <211> 20  
[0047] <212> DNA  
[0048] <213> 人工序列  
[0049] <400> 7  
[0050] tgaatgaagt ccgagtaggc 20  
[0051] <210> 8  
[0052] <211> 17  
[0053] <212> DNA  
[0054] <213> 人工序列  
[0055] <400> 8  
[0056] accctcactc cagccac 17  
[0057] <210> 9  
[0058] <211> 19  
[0059] <212> DNA  
[0060] <213> 人工序列  
[0061] <400> 9  
[0062] caccctcact ccagccacc 19  
[0063] <210> 10  
[0064] <211> 20  
[0065] <212> DNA  
[0066] <213> 人工序列  
[0067] <400> 10  
[0068] gcccttcacc ctcccacagt 20  
[0069] <210> 11  
[0070] <211> 20  
[0071] <212> DNA  
[0072] <213> 人工序列  
[0073] <400> 11  
[0074] ccatgctgga gtgcagtgtt 20  
[0075] <210> 12  
[0076] <211> 19

[0077] <212> DNA  
[0078] <213> 人工序列  
[0079] <400> 12  
[0080] aaacgcccggt ccgactcaa 19  
[0081] <210> 13  
[0082] <211> 20  
[0083] <212> DNA  
[0084] <213> 人工序列  
[0085] <400> 13  
[0086] ccaagaatat cctgagcaac 20  
[0087] <210> 14  
[0088] <211> 18  
[0089] <212> DNA  
[0090] <213> 人工序列  
[0091] <400> 14  
[0092] catcagaggt aagcaggg 18  
[0093] <210> 15  
[0094] <211> 18  
[0095] <212> DNA  
[0096] <213> 人工序列  
[0097] <400> 15  
[0098] tctggaggca ggacaagt 18  
[0099] <210> 16  
[0100] <211> 19  
[0101] <212> DNA  
[0102] <213> 人工序列  
[0103] <400> 16  
[0104] ttgtgggagg aggcataga 19  
[0105] <210> 17  
[0106] <211> 19  
[0107] <212> DNA  
[0108] <213> 人工序列  
[0109] <400> 17  
[0110] gctatcttct tgccacat 19  
[0111] <210> 18  
[0112] <211> 17  
[0113] <212> DNA  
[0114] <213> 人工序列  
[0115] <400> 18

[0116] tcatggtagc agaatca 17  
[0117] <210> 19  
[0118] <211> 19  
[0119] <212> DNA  
[0120] <213> 人工序列  
[0121] <400> 19  
[0122] cataagaagt ttagggta 19  
[0123] <210> 20  
[0124] <211> 20  
[0125] <212> DNA  
[0126] <213> 人工序列  
[0127] <400> 20  
[0128] acataaaggg cagcccaat 20  
[0129] <210> 21  
[0130] <211> 19  
[0131] <212> DNA  
[0132] <213> 人工序列  
[0133] <400> 21  
[0134] caatgaggat ctaggcaca 19  
[0135] <210> 22  
[0136] <211> 17  
[0137] <212> DNA  
[0138] <213> 人工序列  
[0139] <400> 22  
[0140] gtgaggctcc ctcccct 17  
[0141] <210> 23  
[0142] <211> 17  
[0143] <212> DNA  
[0144] <213> 人工序列  
[0145] <400> 23  
[0146] cgttgggcat gtcgtcc 17  
[0147] <210> 24  
[0148] <211> 18  
[0149] <212> DNA  
[0150] <213> 人工序列  
[0151] <400> 24  
[0152] tctcagggca gaggatca 18  
[0153] <210> 25  
[0154] <211> 18

- [0155] <212> DNA  
[0156] <213> 人工序列  
[0157] <400> 25  
[0158] gggcaggagg aacggcta 18  
[0159] <210> 26  
[0160] <211> 19  
[0161] <212> DNA  
[0162] <213> 人工序列  
[0163] <400> 26  
[0164] ttctgtcct tccccacca 19  
[0165] <210> 27  
[0166] <211> 18  
[0167] <212> DNA  
[0168] <213> 人工序列  
[0169] <400> 27  
[0170] caaccgcgt gatcctct 18  
[0171] <210> 28  
[0172] <211> 20  
[0173] <212> DNA  
[0174] <213> 人工序列  
[0175] <400> 28  
[0176] gctgacctcc aaataccgtt 20  
[0177] <210> 29  
[0178] <211> 18  
[0179] <212> DNA  
[0180] <213> 人工序列  
[0181] <400> 29  
[0182] cagttctccc tccccagc 18  
[0183] <210> 30  
[0184] <211> 18  
[0185] <212> DNA  
[0186] <213> 人工序列  
[0187] <400> 30  
[0188] gccccaaagca taaacct 18  
[0189] <210> 31  
[0190] <211> 17  
[0191] <212> DNA  
[0192] <213> 人工序列  
[0193] <400> 31

- [0194] aggacggttg aggggtgg 17  
[0195] <210> 32  
[0196] <211> 18  
[0197] <212> DNA  
[0198] <213> 人工序列  
[0199] <400> 32  
[0200] tcctgtcctt ccccacca 18  
[0201] <210> 33  
[0202] <211> 18  
[0203] <212> DNA  
[0204] <213> 人工序列  
[0205] <400> 33  
[0206] aacccgcgtg atcctctg 18  
[0207] <210> 34  
[0208] <211> 18  
[0209] <212> DNA  
[0210] <213> 人工序列  
[0211] <400> 34  
[0212] ctcttctctg cacagctc 18  
[0213] <210> 35  
[0214] <211> 18  
[0215] <212> DNA  
[0216] <213> 人工序列  
[0217] <400> 35  
[0218] cacgtttgct gagggaaa 18  
[0219] <210> 36  
[0220] <211> 19  
[0221] <212> DNA  
[0222] <213> 人工序列  
[0223] <400> 36  
[0224] gctggtgcca atgtgcatt 19  
[0225] <210> 37  
[0226] <211> 21  
[0227] <212> DNA  
[0228] <213> 人工序列  
[0229] <400> 37  
[0230] tgttcctaa gtccaactac t 21  
[0231] <210> 38  
[0232] <211> 18

- [0233] <212> DNA  
[0234] <213> 人工序列  
[0235] <400> 38  
[0236] caggagctgt gggaggaa 18  
[0237] <210> 39  
[0238] <211> 21  
[0239] <212> DNA  
[0240] <213> 人工序列  
[0241] <400> 39  
[0242] cctaattctct ttctttcagg g 21  
[0243] <210> 40  
[0244] <211> 20  
[0245] <212> DNA  
[0246] <213> 人工序列  
[0247] <400> 40  
[0248] cccatagact caccctgaag 20  
[0249] <210> 41  
[0250] <211> 21  
[0251] <212> DNA  
[0252] <213> 人工序列  
[0253] <400> 41  
[0254] tgcctattgg tctattttcc c 21  
[0255] <210> 42  
[0256] <211> 24  
[0257] <212> DNA  
[0258] <213> 人工序列  
[0259] <400> 42  
[0260] tttctattgg tctccttaaa cctg 24  
[0261] <210> 43  
[0262] <211> 20  
[0263] <212> DNA  
[0264] <213> 人工序列  
[0265] <400> 43  
[0266] ctagggttgg ccaatctact 20  
[0267] <210> 44  
[0268] <211> 20  
[0269] <212> DNA  
[0270] <213> 人工序列  
[0271] <400> 44

[0272] agatgctcaa ggcccttcat 20  
[0273] <210> 45  
[0274] <211> 25  
[0275] <212> DNA  
[0276] <213> 人工序列  
[0277] <400> 45  
[0278] cttttgctaa tcatgttcat acctc 25  
[0279] <210> 46  
[0280] <211> 21  
[0281] <212> DNA  
[0282] <213> 人工序列  
[0283] <400> 46  
[0284] aaactaaaac gatcctgaga c 21  
[0285] <210> 47  
[0286] <211> 20  
[0287] <212> DNA  
[0288] <213> 人工序列  
[0289] <400> 47  
[0290] gaaagtgctc ggtgccttta 20  
[0291] <210> 48  
[0292] <211> 19  
[0293] <212> DNA  
[0294] <213> 人工序列  
[0295] <400> 48  
[0296] caagggtaga ccaccagca 19  
[0297] <210> 49  
[0298] <211> 18  
[0299] <212> DNA  
[0300] <213> 人工序列  
[0301] <400> 49  
[0302] gaagtctgcc gttactgc 18  
[0303] <210> 50  
[0304] <211> 20  
[0305] <212> DNA  
[0306] <213> 人工序列  
[0307] <400> 50  
[0308] gctctgccct gacttttatg 20  
[0309] <210> 51  
[0310] <211> 20

- [0311] <212> DNA  
[0312] <213> 人工序列  
[0313] <400> 51  
[0314] tgcttaccaa gctgtgattc 20  
[0315] <210> 52  
[0316] <211> 29  
[0317] <212> DNA  
[0318] <213> 人工序列  
[0319] <400> 52  
[0320] aatgatacgg cgaccaccga gatctacac 29  
[0321] <210> 53  
[0322] <211> 24  
[0323] <212> DNA  
[0324] <213> 人工序列  
[0325] <400> 53  
[0326] caagcagaag acggcatacg agat 24  
[0327] <210> 54  
[0328] <211> 31  
[0329] <212> DNA  
[0330] <213> 人工序列  
[0331] <400> 54  
[0332] acactctttc cctacacgac gctcttccga t 31  
[0333] <210> 55  
[0334] <211> 33  
[0335] <212> DNA  
[0336] <213> 人工序列  
[0337] <400> 55  
[0338] gtgactggag ttcagacgtg tgctcttccg atc 33  
[0339] <210> 56  
[0340] <211> 21  
[0341] <212> DNA  
[0342] <213> 人工序列  
[0343] <400> 56  
[0344] accaaacaga gccatttacc t 21  
[0345] <210> 57  
[0346] <211> 19  
[0347] <212> DNA  
[0348] <213> 人工序列  
[0349] <400> 57

[0350] cctgacctcg tgatccacc 19  
[0351] <210> 58  
[0352] <211> 20  
[0353] <212> DNA  
[0354] <213> 人工序列  
[0355] <400> 58  
[0356] caaacagagc catttacctt 20  
[0357] <210> 59  
[0358] <211> 21  
[0359] <212> DNA  
[0360] <213> 人工序列  
[0361] <400> 59  
[0362] ccatgctgga gtgcagtgtt g 21  
[0363] <210> 60  
[0364] <211> 18  
[0365] <212> DNA  
[0366] <213> 人工序列  
[0367] <400> 60  
[0368] gggagcgatc tgggtcga 18  
[0369] <210> 61  
[0370] <211> 18  
[0371] <212> DNA  
[0372] <213> 人工序列  
[0373] <400> 61  
[0374] agggtcacca gcaggcag 18  
[0375] <210> 62  
[0376] <211> 18  
[0377] <212> DNA  
[0378] <213> 人工序列  
[0379] <400> 62  
[0380] ggagcgatct gggtcgag 18  
[0381] <210> 63  
[0382] <211> 20  
[0383] <212> DNA  
[0384] <213> 人工序列  
[0385] <400> 63  
[0386] ggtatttgga ggtcagcacg 20  
[0387] <210> 64  
[0388] <211> 23

- 
- [0389] <212> DNA  
[0390] <213> 人工序列  
[0391] <400> 64  
[0392] tggcttttga ataaagtctg agt 23

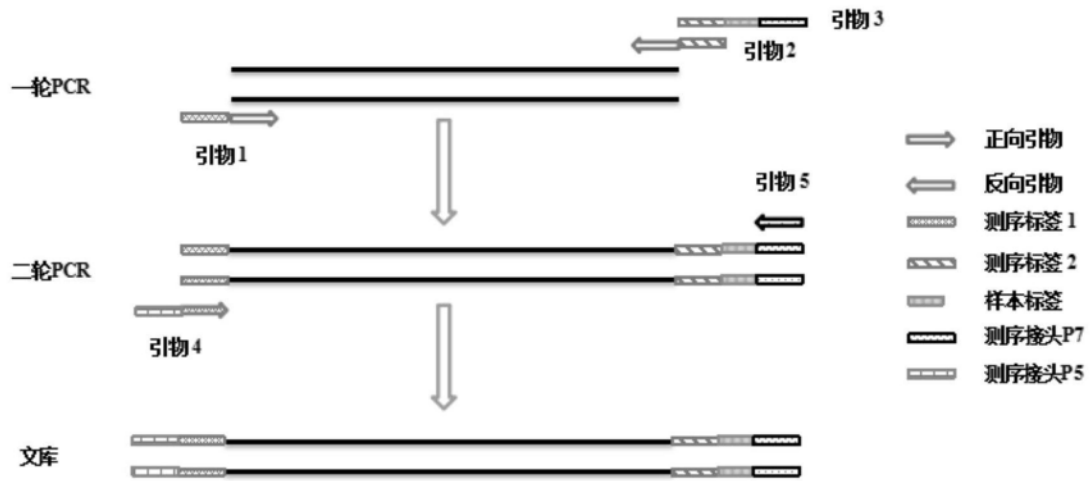


图1

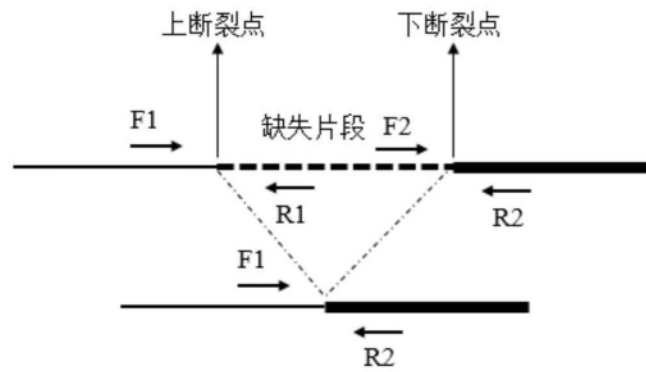


图2

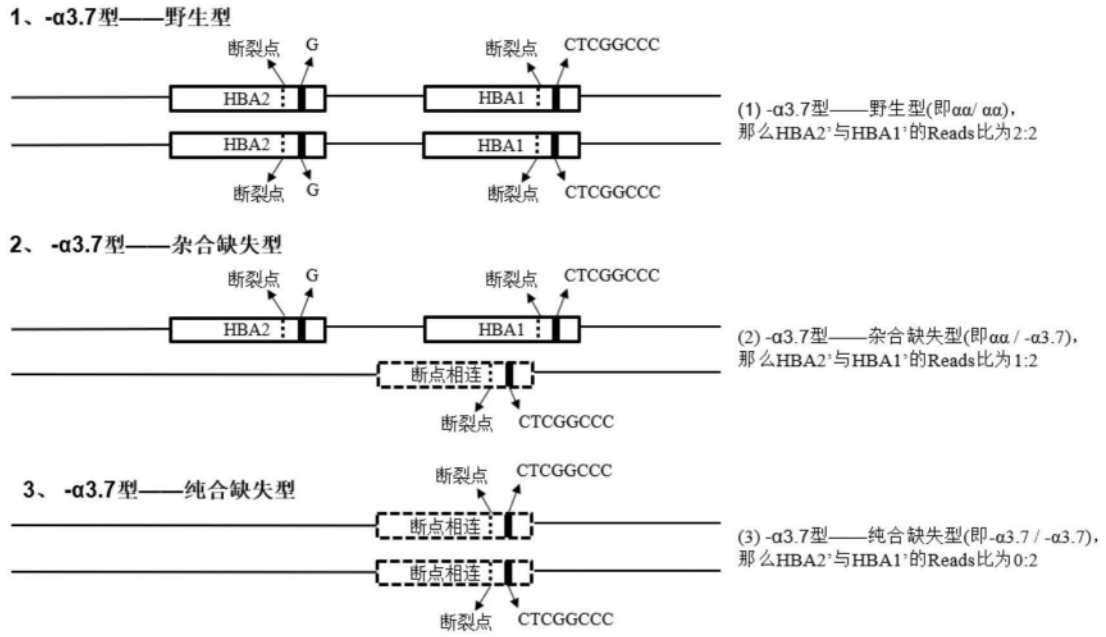


图3



图4