(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105713984 B
(45)授权公告日 2020.02.07

(21)申请号 201610251776.X

(22)申请日 2016.04.21

(65)同一申请的已公布的文献号
申请公布号 CN 105713984 A

(43)申请公布日 2016.06.29

(73)专利权人 中国科学院西北高原生物研究所
地址 810008 青海省西宁市西关大街59号

(72)发明人 张发起 田尊哲 陈世龙 高庆波

(74)专利代理机构 北京市卓华知识产权代理有限公司 11299
代理人 陈子英

(51)Int.Cl.
C12Q 1/6895(2018.01)
C12N 15/11(2006.01)
C12N 15/10(2006.01)

(56)对比文件
Zunzhe Tian et al..Development of SSR Markers for a Tibetan Medicinal Plant, Lancea tibetica (Phrymaceae), Based on RAD Sequencing.《Applications in Plant Sciences》.2016,第4卷(第11期),apps1600076.
刘小利.优良的野生阔叶草坪植物—肉果草.《青海农林科技》.2006,第2006年卷(第4期),52-53.

审查员 马艳林

权利要求书2页 说明书6页
序列表7页 附图4页

(54)发明名称

肉果草微卫星分子标记

(57)摘要

本发明涉及一种肉果草微卫星分子标记,属DNA分子标记技术领域,其通过RAD-seq技术从肉果草基因组DNA中分离微卫星分子标记,根据每个微卫星分子标记位点两端的侧翼序列设计该分子标记的特异性引物,对来自不同居群的肉果草个体进行PCR扩增并检测扩增结果的稳定性和多态性,获得了10个有效的微卫星分子标记。本发明开发出有效的肉果草微卫星分子标记,建立了制备肉果草微卫星分子标记的方法,并可用这些分子标记进行遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析、进化与亲缘关系研究等。

CN 105713984 B

1.一种肉果草微卫星分子标记，其特征在于具有序列表SEQ.ID.No.1至SEQ.ID.No.10中任意一种或多种的核苷酸序列。

2.一种适于权利要求1所述分子标记的特异性引物，其特征在于采用下列任意一对或多对引物：

1)序列表中SEQ ID NO.11和12组成的引物对；

2)序列表中SEQ ID NO.13和14组成的引物对；

3)序列表中SEQ ID NO.15和16组成的引物对；

4)序列表中SEQ ID NO.17和18组成的引物对；

5)序列表中SEQ ID NO.19和20组成的引物对；

6)序列表中SEQ ID NO.21和22组成的引物对；

7)序列表中SEQ ID NO.23和24组成的引物对；

8)序列表中SEQ ID NO.25和26组成的引物对；

9)序列表中SEQ ID NO.27和28组成的引物对；

10)序列表中SEQ ID NO.29和30组成的引物对。

3.一种基于RAD-seq技术获取权利要求1所述的肉果草微卫星分子标记的方法，其特征在于包括以下步骤：

(1)肉果草DNA文库的构建：用CTAB法提取肉果草基因组DNA、通过限制性内切酶进行酶切、连接接头p1、随机打断、末端修复、加A尾、连接接头p2、PCR扩增、回收300-700bp序列和纯化，在Illumina HiSeq测序平台上进行双末端测序；

(2)测序数据质量控制：每端测序125bp，用Illumina Casava 1.8进行碱基识别，获得原始测序序列raw base，对其测序质量分布、测序错误率分布以及GC含量分布进行检查，得到clean base；

(3)RAD-Tag捕获率统计：统计clean base去重后reads中酶Tag开头的reads数，以及酶捕获的reads数目与去重后reads数目的比值，得到RAD-Tag相关统计信息；

(4)聚类与组装：对于样本中的含有酶识别位点reads，用cd-hit-est软件进行聚类，并从中选取reads支持数为10-400之间的类，根据聚类结果，用Velvetopt进行参数设置，对筛选后的类进行组装，组装后，对125bp以下的contig进行过滤；

(5)SSR检测：对于已经组装好的contig，用rimmomatic v.0.32软件进行SSR检测，SSR检测标准如下：SSR重复单元的最小长度为2bp、SSR重复单元的最大长度为6bp、SSR序列的最小长度为12bp、SSR上下游序列长度为100bp、两个SSR的最小距离为12bp；

(6)分离SSR引物：用Primer 3软件批量设计SSR引物，合成SSR引物，鉴定其在不同肉果草中的多态性，筛选出具有稳定性和多态性的引物。

4.一种利用权利要求2所述的引物分析肉果草遗传多样性的方法，其特征在于包括以下步骤：

(1)实验材料的准备：肉果草采样，并分别提取用于检测分析的每个个体的基因组DNA；

(2)PCR扩增：利用权利要求2所述的引物，以提取的不同个体的基因组DNA为模板进行PCR扩增反应，94℃预变性4分钟，94℃变性45秒，退火30秒，72℃延伸35秒，变性至退火三个步骤重复35次，最后72℃充分延伸8分钟，4℃保存；

(3)电泳检测：对PCR扩增产物进行1％的琼脂糖凝胶电泳检测，对于条带清晰的PCR产

物用聚丙烯酰氨凝胶电泳进行检测；

(4)结果分析：用genepop软件计算期望杂合度和观察杂合度，并计算近交系数，以此来描述肉果草相关微卫星DNA多态性的特征。

5.权利要求1所述的分子标记或权利要求2所述的引物在肉果草遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析和/或进化与亲缘关系研究中的应用。

肉果草微卫星分子标记

技术领域

[0001] 本发明涉及一种生物科学中的DNA分子标记技术，尤其涉及一种肉果草微卫星分子标记、引物对及其制备方法和应用。

背景技术

[0002] 微卫星分子标记又称作简单序列重复 (Simple Sequence Repeats, SSRs)，是指基因组中以1-6个核苷酸为单位串联组成的核苷酸序列。由于重复次数的不同或重复程度的不完全相同，造成了微卫星序列长度的高度变异性，由此产生了微卫星分子标记。虽然微卫星位点核心序列的重复次数在个体间呈高度变异性，但是其两端侧翼序列多是保守的单拷贝序列，因此可以用两端的侧翼序列设计成对引物，通过PCR技术，经聚丙烯酰胺凝胶电泳(PAGE)或其他适宜方式显示这些分子标记在不同个体间的多态性。微卫星分子标记具有以下特点：广泛分布于真核生物的基因组中；多态性信息容量高；呈共显性；遵循孟德尔遗传法则、中性选择等。基于以上特点，微卫星分子标记被广泛用于遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析、进化与亲缘关系等方面的研究。

[0003] 肉果草(*Lancea tibetica*)，藏名巴丫巴，隶属于玄参科肉果草属，分布于青藏高原及其毗邻山区(西藏、青海、甘肃、四川、云南及印属喜马拉雅地区等)，生于海拔2000-4500米的草地、疏林中或沟谷旁。肉果草为重要的藏药植物，全草有养肺排脓、清热止咳功效，花、果能治疗心脏病、血性肿瘤(血癌)、哮喘、痛肿疮疡等。肉果草药理作用主要有如下几个方面：(1)有抗肿瘤作用；(2)有抗氧化作用；(3)有抗疟作用；(4)有抗真菌作用；(5)有杀菌作用；(6)具有降糖的作用；(7)具有保肝作用。由于其具有很重要的药用价值，肉果草已经吸引了很多国内外学者的关注。在与玉树等地的藏医交流中发现，肉果草作为一味十分重要的藏药材，有着较为广泛的应用。但由于对该物种的遗传背景等研究较少，难以开展人工种植，仅依靠野外采挖获取，加剧了该物种的濒危进程。

发明内容

[0004] 本发明的目的是针对现有技术尚未开发肉果草微卫星标记的现状，利用限制性酶切位点相关DNA的简化基因组测序(RAD-seq)技术开发出肉果草微卫星分子标记并获得对应的引物，建立一种制备肉果草微卫星分子标记的方法，并由此可以进一步用这些分子标记进行遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析、进化与亲缘关系研究等。

[0005] 为了实现上述目的，本发明的技术方案为：

[0006] 一种肉果草微卫星分子标记，具有序列表SEQ.ID.No.1至SEQ.ID.No.10中任一种或多种的核苷酸序列，这些分子标记所对应的位点分别编号为LT4、LT7、LT9、LT10、LT12、LT15、LT16、LT18、LT25、LT28。

[0007] 适于本发明中任意一种或多种肉果草微卫星分子标记的特异性引物，具有序列表

SEQ ID NO.11至SEQ ID NO.30中任一种或多种的核苷酸序列，各引物的特点及与微卫星分子标记(位点)的对应关系见表1。

[0008] 表1 微卫星片段引物的特点及与微卫星分子标记(位点)的对应关系

[0009]

位点	引物序列(5'-3')	重复单元	产物大小(bp)	退火温度
LT4	F: ATTGATTGATTCACGTTCCAAAT	$(TA)_6$	132	54℃
	R: TGAAAATGAATAACTTGGGGATCT			
LT7	F: TTTGGAAAGCATGATCTACCACT	$(AAT)_5$	151	56℃
	R: TTTCTGGACTGTTGTAATCTTGAAA			
LT9	F: GGATTTCTAAGTGCAATCCTCAA	$(GA)_7$	140	51℃
	R: CATCACTCACCAAATGAAAGACA			
LT10	F: AATTGTTCCAGGTATGCAGTGTT	$(TG)_6$	155	51℃
	R: CTATTCTGCAAGTTAATGCAGGG			
LT12	F: GTAGACATTTTTGCAGCACCTCT	$(CAT)_5$	150	51℃
	R: ATGAGGACTCAAAGACAGCTCAG			
LT15	F: CTTATAACCTATCGTTCTCCGGC	$(AG)_7$	138	54℃
	R: ATTTCGCTCTCTCTTTCACACAC			
LT16	F: TGTATTGTCAATGGAAGAGGCAT	$(AAG)_4$	155	51℃
	R: GAATGAGATGCTCCACTAACCAC			
LT18	F: AACAAGTTTATGCAAGGAGGAGA	$(TCT)_5$	160	51℃
	R: CCCAAGTCCCAAATGATATAAAA			
LT25	F: GATGCCAAGGAATTGTTATATGC	$(TA)_7$	153	51℃
	R: TTTCTAGAAGTCGGAGCTGTCC			
LT28	F: AACAGCAATGGCAATATGGTATC	$(TA)_9$	158	56℃
	R: AACTGTTCAAGTTGGCAAAACAT			

[0010] 一种基于RAD-seq技术获取本发明所述肉果草微卫星分子标记的方法，其包括以下步骤：

[0011] (1)肉果草DNA文库的构建：用CTAB法提取肉果草基因组DNA、通过限制性内切酶进行酶切、连接接头p1、随机打断、末端修复、加A尾、连接接头p2、PCR扩增、回收300-700bp序列和纯化，在Illumina HiSeq测序平台上进行双末端测序；

[0012] (2)测序数据质量控制：每端测序125bp，用Illumina Casava 1.8 进行碱基识别(Base Calling)，获得原始测序序列(Sequenced Reads)，称之为raw base，对其测序质量分布、测序错误率分布以及GC含量分布进行检查，得到clean base；

[0013] (3)RAD-Tag捕获率统计:统计clean base 去重后reads中酶Tag开头的reads数,以及酶捕获的reads数目与去重后reads数目的比值,得到RAD-Tag相关统计信息(本发明RAD-Tag捕获率为97.65%,所有样本的数据量足够,测序质量合格,GC含量正常);

[0014] (4)聚类与组装:对于样本中含有酶识别位点reads,用cd-hit-est软件进行聚类,并从中选取reads支持数为10-400之间的类,根据聚类结果,用Velvetopt 进行参数设置,对筛选后的类进行组装,组装后,并对125bp以下的contig进行过滤;

[0015] (5)SSR检测:对于已经组装好的contig,用rimmomatic v.0.32软件进行SSR检测,SSR检测标准如下:SSR重复单元的最小长度为2bp、SSR重复单元的最大长度为6bp、SSR序列的最小长度为12bp、SSR上下游序列长度为100bp、两个SSR的最小距离为12bp;

[0016] (6)分离SSR引物:用Primer 3软件批量设计SSR引物,合成SSR引物,并鉴定其在不同肉果草中的多态性。

[0017] 一种利用本发明所述引物分析肉果草遗传多样性的方法,包括以下步骤:

[0018] (1)实验材料的准备:肉果草采样,采集野外自然状态下肉果草生长正常的嫩叶,用硅胶迅速干燥后带回实验室,并分别提取用于检测分析的每个个体的基因组DNA;

[0019] (2)PCR扩增:利用本发明所述的任意一对或多对引物,以提取的不同个体的基因组DNA为模板进行PCR扩增反应,94℃预变性4分钟,94℃变性45秒,退火30秒(退火温度见表1),72℃延伸35秒,变性至退火三个步骤重复35次,最后72℃充分延伸8分钟,4℃保存;

[0020] (3)电泳检测:对PCR扩增产物进行1%的琼脂糖凝胶电泳检测,对于条带清晰的PCR产物用聚丙烯酰氨凝胶电泳进行检测;

[0021] (4)结果分析:用genepop软件计算期望杂合度和观察杂合度,并计算近交系数,以此来描述相关肉果草微卫星DNA多态性的特征。

[0022] 本发明的有益效果:开发出肉果草微卫星分子标记并分离出对应的引物,并建立了行之有效的获取肉果草微卫星分子标记的方法,这些分子标记及其引物可进行遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析、进化与亲缘关系研究等。根据实际需要,可以采用本发明公开的任意一种或多种标记,或与相关标记对应的任意一对或多对特异性引物进行分析研究。

附图说明

[0023] 图1是肉果草的DNA文库构建及库检示意图;

[0024] 图2是测序质量分布检测;

[0025] 图3是测序错误率分布检测;

[0026] 图4是G、C含量分布检测;

[0027] 图5是肉果草SSR不同重复单元数量统计;

[0028] 图6是LT4位点对所有个体扩增后经PAGE电泳检测图。

具体实施方式

[0029] 为了详细说明本发明肉果草微卫星分子标记的技术内容、构造特征,以下结合实施方式并配合附图作进一步说明。

[0030] 实施例一:肉果草微卫星分子标记的制备方法

[0031] 1.肉果草DNA文库的构建

[0032] 1.1基因组DNA的提取：用CTAB法提取肉果草基因组，然后通过限制性内切酶进行酶切，酶切后的DNA 片段连接接头p1。

[0033] 1.2混池：打断的DNA，经过末端修复后，在打断的DNA另一端加上A尾，然后加上p2接头。

[0034] 1.3测序：对以上的DNA片段进行PCR扩增，回收300-700bp的序列，纯化后，在Illumina HiSeq 测序平台上进行双末端测序。肉果草DNA文库构建及库检示意图见图1。

[0035] 2.测序数据质量控制

[0036] 2.1数据质量检查：每端测序125bp，用Illumina Casava 1.8 进行碱基识别(Base Calling)，获得原始测序序列(Sequenced Reads)，称之为raw base、raw data或raw reads，对其测序质量分布进行检测，结果见图2；对测序错误率分布进行检查，结果见图3；对GC含量分布进行检查，结果见图4。

[0037] 2.2测序数据过滤：测序得到的2,800,948,250bp的raw base中，错误分布率为0.04%，Q20为93.24%，Q30为87.66%；GC含量为35.33%，共得到2,764,204,500bp的clean base。综上所述，所有样本的数据量足够，测序质量合格，GC含量正常，建库成功。

[0038] 2.3 RAD-Tag捕获率统计：统计clean base去重后reads中酶Tag开头的reads数。酶捕获率的reads数目与去重后reads数目的比值为97.65%，从11056818条clean reads中共得到8074787条去重后的reads。

[0039] 3.聚类与组装

[0040] 对于样本中含有酶识别位点reads，用cd-hit-est软件进行聚类，并从中选取reads支持数为10-400之间的类，对过滤的类数的统计见表2。根据聚类结果，用Velvetopt进行参数设置，对筛选后的类进行组装，组装后，对125bp以下的contig进行过滤，结果统计见表3。

[0041] 表2 类数统计结果

[0042]

样品	数据聚类后的类数	聚类的reads数	支持数在10-400的类数	过滤后的类数	过滤后的reads数/参加聚类reads数（%）
L7	1417277	8074786	222628	4972900	61.59

[0043] 表3 RAD局部组装结果统计

[0044]

样品	组装结果总长度（bp）	组装结果总个数	组装contig的平均长度（bp）	N50长度（bp）	GC（%）
L7	106,439,694	401,203	265	361	34.22

[0045] 4.SSR检测

[0046] 对于已经组装好的contig，用rimmomatic v.0.32软件进行SSR检测。SSR检测标准如下：(1)SSR重复单元的最小长度为2；(2)SSR重复单元的最大长度为6；(3)SSR序列的最小

长度为12;(4)SSR上下游序列长度为100bp;(5)两个SSR的最小距离为12bp;其统计结果见图5;随机挑选出100对SSR序列用Primer 3软件进行引物设计。

[0047] 5.分离SSR引物

[0048] 用设计的100对SSR引物,对来自3个居群(YD,QML,MY)的56个个体进行PCR扩增,肉果草居群位置信息见表4。反应程序:94℃预变性4分钟,94℃变性45秒,退火30秒(退火温度见表1),72℃延伸35秒,变性至退火三个步骤重复35次,最后72℃充分延伸8分钟,4℃保存,1%的琼脂糖凝胶电泳检测。对于条带清晰的PCR产物用PAGE电泳进行检测,得到10个多态性的微卫星标记。用LT4位点对应引物对所有的个体进行扩增后,在PAGE电泳进行检测的电泳图见图6。

[0049] 表4 肉果草居群位置信息

[0050]

居群编号	采样地点	个体数	采集号	地理坐标	海拔(m)
YD	西藏亚东	17	Chen2014498	27°47'N,99°17'E	4350
QML	青海曲麻莱	20	Chen2014684	33°58'N,96°34'E	4570
MY	青海门源	19	Zhang2014341	37°51'N,101°05'E	3636

[0051] 实施例二:用所述10对引物进行肉果草遗传多样性分析

[0052] 用genepop软件计算期望杂合度和观察杂合度,并计算近交系数,以此来描述10个肉果草微卫星DNA多态性的特征。结果见表5。

[0053] 由表5可以看出本发明的10个微卫星序列在3个居群的56个个体中都有多样性。由此可见,本发明的10个微卫星标记可以用于遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析、进化与亲缘关系研究等,具有重复性好的特点,是一种可靠有效的分子标记。

[0054] 表5 肉果草的10个引物在3个居群中结果

[0055]

位点	居群 YD (N=17)					居群 QML (N=20)					居群 MY (N=19)				
	A	H_E	H_O	NA	F_{IS}	A	H_E	H_O	NA	F_{IS}	A	H_E	H_O	NA	F_{IS}
LT4	6	0.800	0.529	0.131	0.346	3	0.683*	0.200	0.307	0.713	5	0.758	0.474	0.141	0.382
LT7	7	0.838	0.706	0.054	0.162	3	0.719	0.600	0.160	0.169	5	0.807*	0.579	0.116	0.288
LT9	7	0.848	0.706	0.062	0.172	5	0.745	0.800	0.024	-0.076	8	0.865*	0.789	0.110	0.089
LT 10	7	0.816	0.882	0.038	-0.084	7	0.814	0.950	0.000	-0.173	7	0.859	1.000	0.000	-0.169
LT 12	5	0.779	0.765	0.000	0.019	5	0.768	0.900	0.000	-0.177	6	0.831	0.895	0.000	-0.079
LT 15	5	0.779	0.882	0.000	-0.137	6	0.814	0.750	0.028	0.081	4	0.708	0.842	0.012	-0.195
LT 16	8	0.850	0.765	0.199	0.103	8	0.879	0.850	0.030	0.034	6	0.812	0.632	0.095	0.223
LT 18	6	0.820	0.706	0.058	0.143	7	0.821	0.850	0.000	-0.035	5	0.700	0.895	0.000	-0.289
LT 25	7	0.872*	0.824	0.165	0.057	4	0.781	0.850	0.034	-0.091	6	0.797	0.737	0.046	0.077
LT 28	6	0.804	0.824	0.000	-0.025	6	0.833	0.850	0.036	-0.020	6	0.835	0.895	0.000	-0.074

[0056] 注:H_E:期望杂合度;H_O:观察杂合度;NA:零等位基因频率;F_{IS}:近交系数,*代表偏离哈德温伯格平衡的$p<0.01$。

[0057] 从上述分析结果可以看出,本发明公开的任意一种微卫星分子标记及其相应的引物(对)均可以用于肉果草遗传多态性的分析研究,可以根据需要和实验,选定所需的标记、

标记组合及相关引物，可以为其中的某一种、某几种，也可以是全部。

[0058] 本发明的标记物和/或相应的引物可以应用于肉果草遗传多样性分析、植物遗传图谱的构建、基因定位、品种鉴定、种质保存、数量性状基因的分析和/或进化与亲缘关系等方面的研究中。

[0059] 以上揭示的仅为本发明的较佳实施例而已，不能以此来限定本发明之权利范围，依本发明权利所做的等同变化，仍属于本发明所涵盖的范围。

[0001] SEQUENCE LISTING
[0002] <110> 中国科学院西北高原生物研究所
[0003] <120> 肉果草微卫星分子标记
[0004] <160> 30
[0005] <210> 1
[0006] <211> 212
[0007] <212> DNA
[0008] <213> 肉果草(Lancea tibetica)
[0009] <223> LT4分子标记核苷酸序列
[0010] <400> 1
[0011] gtccattaat tccattaatt gattgattca cgttccaaat agatcataat tccttgatca 60
[0012] gtggaactaa atatgataag actctttgat agcgagggtc tatatatata taaaggatta 120
[0013] cagttagatc cccaagttat tcattttcat acggtattcc tacaccatct agtaagattg 180
[0014] agatagttta taaacaccgc aacaattgac tc 212
[0015] <210> 2
[0016] <211> 215
[0017] <212> DNA
[0018] <213> 肉果草(Lancea tibetica)
[0019] <223> LT7分子标记核苷酸序列
[0020] <400> 2
[0021] atgcataagc tagtacaggg aattgaaaca aatttggaaa gcatgatcta ccactgctct 60
[0022] ggatgagtat aaatgaaaaa ccttcaaaca attaccaaaa aataataata ataatttcag 120
[0023] aatagtatat aaccagcagc accatatata tatagctgtt tcaagattac aacagtccag 180
[0024] aaatacataa ttaatttaga ttcctccaag ataca 215
[0025] <210> 3
[0026] <211> 214
[0027] <212> DNA
[0028] <213> 肉果草(Lancea tibetica)
[0029] <223> LT9分子标记核苷酸序列
[0030] <400> 3
[0031] actttgtttc ttgattctcc ttcaattctt ggtgtaggat ttctaagtgc aatcctcaat 60
[0032] agcttctaag tgatatgaaa attgaggaaa atattgttgg gagagagaga gagaggtgag 120
[0033] gggaggatat atactacttg ttgtcttcaa ttttgtcttt catttggtga gtgatgaaga 180
[0034] agcatctaca tctccctccc aaaatagtaa tttt 214
[0035] <210> 4
[0036] <211> 212
[0037] <212> DNA
[0038] <213> 肉果草(Lancea tibetica)

[0039] <223> LT10分子标记核苷酸序列

[0040] <400> 4

[0041] tctcaaagga tggaaactct cctctggaat tgttccaggt atgcagtgtt tcataaattc 60

[0042] ctttatattt gtgaaatatc agtattttct aatgactctc tgtgtgtgtg tgcgtctctc 120

[0043] tcactctctc tgtgaccata actggatgaa acctacgccc cctgcattaa cttgcagaat 180

[0044] agtaactgaa aacactgtaa acaaggaaaa gt 212

[0045] <210> 5

[0046] <211> 215

[0047] <212> DNA

[0048] <213> 肉果草(Lancea tibetica)

[0049] <223> LT12分子标记核苷酸序列

[0050] <400>5

[0051] atctggttgt ttggttctgc tggtgcttga gtatcctttc tcttcctttg ataagtagac 60

[0052] atttttgcag cacctctagt agggccatga tcatcattta catcatcatc atcatgactt 120

[0053] ttcttaggat gatgctttga gggatgatta tctgactatg atttctaatt tttcgacctt 180

[0054] gctgagctgt ctttgagtcc tcattgtttg atcct 215

[0055] <210> 6

[0056] <211> 214

[0057] <212> DNA

[0058] <213> 肉果草(Lancea tibetica)

[0059] <223> LT15分子标记核苷酸序列

[0060] <400> 6

[0061] cttataacct atcgttctcc ggcggccgca agaggtggcg gtagaggagg gcggtcatcg 60

[0062] gcagaggctg gtggacggcg tggtggcggg tagtggctgt agagagagag agagcgtgtg 120

[0063] tgaaagagag agcgaaatag agaaagaaga gagagagaag ggacttaccc atgtggcggt 180

[0064] gactggtggc gacgagcgga gactggtgga tggc 214

[0065] <210> 7

[0066] <211> 212

[0067] <212> DNA

[0068] <213> 肉果草(Lancea tibetica)

[0069] <223> LT16分子标记核苷酸序列

[0070] <400> 7

[0071] acacgagtcc tataactgtc aaaataacaa atcagagggg aaagggaaga catgtgtatt 60

[0072] gtcaatggaa gaggcatgga aggggtggaa tgagagatat aagaagaaga aggaggacaa 120

[0073] caagaagagg tgaatgagac aatgataagt atgatatctc tactcgatgc acctacctta 180

[0074] atacttgtgg ttagtggagc atctcattca tt 212

[0075] <210> 8

[0076] <211> 215

[0077] <212> DNA

[0078] <213> 肉果草(Lancea tibetica)

[0079] <223> LT18分子标记核苷酸序列

[0080] <400> 8

[0081] aaaatccttt tttcgtatta aattcataaa aacaagttta tgcaaggagg agattaacct 60

[0082] ctgtttcata ggatttgatc cattgtttta atttattaaa tcttcttctt cttcttttcc 120

[0083] atatttatta ggtttactga tgtgaatata gtaaggatgc ataccaattt tatatcattt 180

[0084] gggacttggg agtggacgat tgggttgagt caaaa 215

[0085] <210> 9

[0086] <211> 214

[0087] <212> DNA

[0088] <213> 肉果草(Lancea tibetica)

[0089] <223> LT25分子标记核苷酸序列

[0090] <400> 9

[0091] tataatcaaa ttcatttcca tgatgccaag gaattgttat atgcacagga taccatggac 60

[0092] gaggcaattg ccaataaaag aaaaaaatac atctaatatt tatatatata tatatgttgt 120

[0093] gggtgtgata ttttttgatc ttgtcccatg tgggacagct ccgacttcta gaaacttttt 180

[0094] attttaataa gatagacgtc acagaaacaa ttaa 214

[0095] <210> 10

[0096] <211> 218

[0097] <212> DNA

[0098] <213> 肉果草(Lancea tibetica)

[0099] <223> LT28分子标记核苷酸序列

[0100] <400> 10

[0101] atgagtgcgc cgagggagga gcttctacaa gcaacagcaa tggcaatatg gtatcctgtt 60

[0102] aaatttgcta tacttgctta attttccact gtgatattgc tatatatata tatatatagt 120

[0103] taataattaa taatgtaata atttctatac acttgaattg aatagtgatg ttttgccaac 180

[0104] ttgaacagtt gattgtgctg tgctttgtga tgtttatt 218

[0105] <210> 11

[0106] <211> 23

[0107] <212> DNA

[0108] <213> 人工序列

[0109] <223> LT4引物正向序列

[0110] <400> 11

[0111] attgattgat tcacgttcca aat 23

[0112] <210> 12

[0113] <211> 24

[0114] <212> DNA

[0115] <213> 人工序列

[0116] <223> LT4引物反向序列

[0117] <400> 12

[0118] tgaaaatgaa taacttgggga tct 24

[0119] <210> 13

[0120] <211> 23

[0121] <212> DNA

[0122] <213> 人工序列

[0123] <223> LT7引物正向序列

[0124] <400> 13

[0125] tttggaaagc atgatctacc act 23

[0126] <210> 14

[0127] <211> 25

[0128] <212> DNA

[0129] <213> 人工序列

[0130] <223> LT7引物反向序列

[0131] <400> 14

[0132] tttctggact gttgtaatct tgaaa 25

[0133] <210> 15

[0134] <211> 23

[0135] <212> DNA

[0136] <213> 人工序列

[0137] <223> LT9引物正向序列

[0138] <400> 15

[0139] ggatttctaa gtgcaatcct caa 23

[0140] <210> 16

[0141] <211> 23

[0142] <212> DNA

[0143] <213> 人工序列

[0144] <223> LT9引物反向序列

[0145] <400> 16

[0146] catcactcac caaatgaaag aca 23

[0147] <210> 17

[0148] <211> 23

[0149] <212> DNA

[0150] <213> 人工序列

[0151] <223> LT10引物正向序列

[0152] <400> 17

[0153] aattgttcca ggtatgcagt gtt 23

[0154] <210> 18

[0155] <211> 23

[0156] <212> DNA

[0157] <213> 人工序列

[0158] <223> LT10引物反向序列

[0159] <400> 18

[0160] ctattctgca agttaatgca ggg 23

[0161] <210> 19

[0162] <211> 23

[0163] <212> DNA

[0164] <213> 人工序列

[0165] <223> LT12引物正向序列

[0166] <400> 19

[0167] gtagacattt ttgcagcacc tct 23

[0168] <210> 20

[0169] <211> 23

[0170] <212> DNA

[0171] <213> 人工序列

[0172] <223> LT12引物反向序列

[0173] <400> 20

[0174] atgaggactc aaagacagct cag 23

[0175] <210> 21

[0176] <211> 23

[0177] <212> DNA

[0178] <213> 人工序列

[0179] <223> LT15引物正向序列

[0180] <400> 21

[0181] cttataacct atcgttctcc ggc 23

[0182] <210> 22

[0183] <211> 23

[0184] <212> DNA

[0185] <213> 人工序列

[0186] <223> LT15引物反向序列

[0187] <400> 22

[0188] atttcgctct ctctttcaca cac 23

[0189] <210> 23

[0190] <211> 23

[0191] <212> DNA

[0192] <213> 人工序列

[0193] <223> LT16引物正向序列

[0194] <400> 23

[0195] tgtattgtca atggaagagg cat 23

[0196] <210> 24

[0197] <211> 23

[0198] <212> DNA

[0199] <213> 人工序列

[0200] <223> LT16引物反向序列

[0201] <400> 24

[0202] gaatgagatg ctccactaac cac 23

[0203] <210> 25

[0204] <211> 23

[0205] <212> DNA

[0206] <213> 人工序列

[0207] <223> LT18引物正向序列

[0208] <400> 25

[0209] aacaagttta tgcaaggagg aga 23

[0210] <210> 26

[0211] <211> 23

[0212] <212> DNA

[0213] <213> 人工序列

[0214] <223> LT18引物反向序列

[0215] <400> 26

[0216] cccaagtccc aaatgatata aaa 23

[0217] <210> 27

[0218] <211> 23

[0219] <212> DNA

[0220] <213> 人工序列

[0221] <223> LT25引物正向序列

[0222] <400> 27

[0223] gatgccaagg aattgttata tgc 23

[0224] <210> 28

[0225] <211> 22

[0226] <212> DNA

[0227] <213> 人工序列

[0228] <223> LT25引物反向序列

[0229] <400> 28

[0230] tttctagaag tcggagctg tcc 22

[0231] <210> 29

[0232] <211> 23

[0233] <212> DNA

[0234] <213> 人工序列

[0235] <223> LT28引物正向序列

[0236] <400> 29

[0237] aacagcaatg gcaatatggt atc 23

[0238] <210> 30

[0239] <211> 23

[0240] <212> DNA

[0241] <213> 人工序列

[0242] <223> LT28引物反向序列

[0243] <400> 30

[0244] aactgttcaa gttggcaaaa cat 23

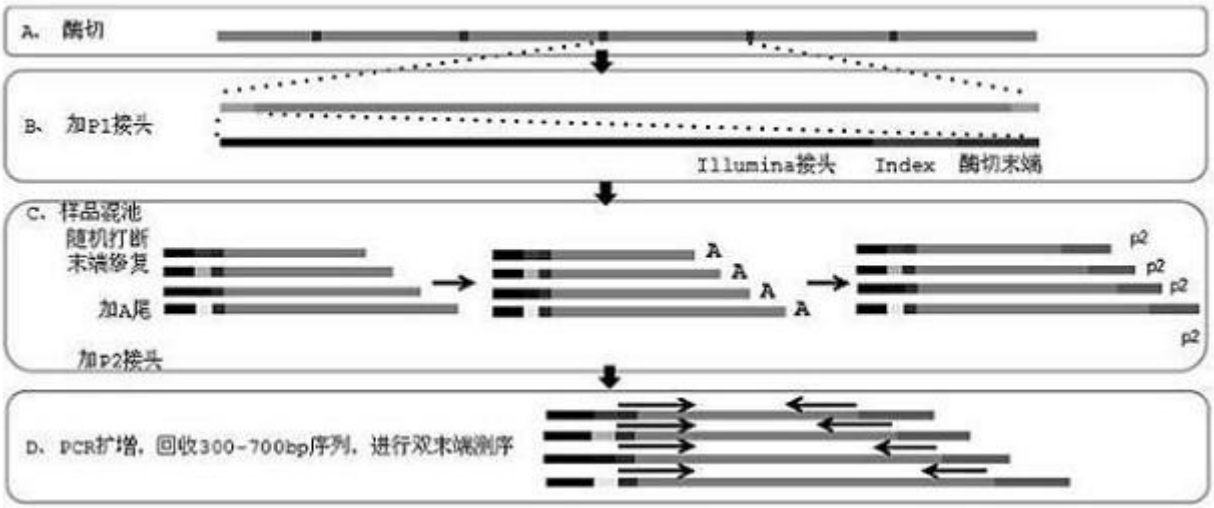
A、酶切
B、加P1接头
Illumina接头
Index
酶切末端
C、样品混池
随机打断
末端修复
加A尾
加P2接头
A
p2
D、PCR扩增，回收300~700bp序列，进行双末端测序

图1

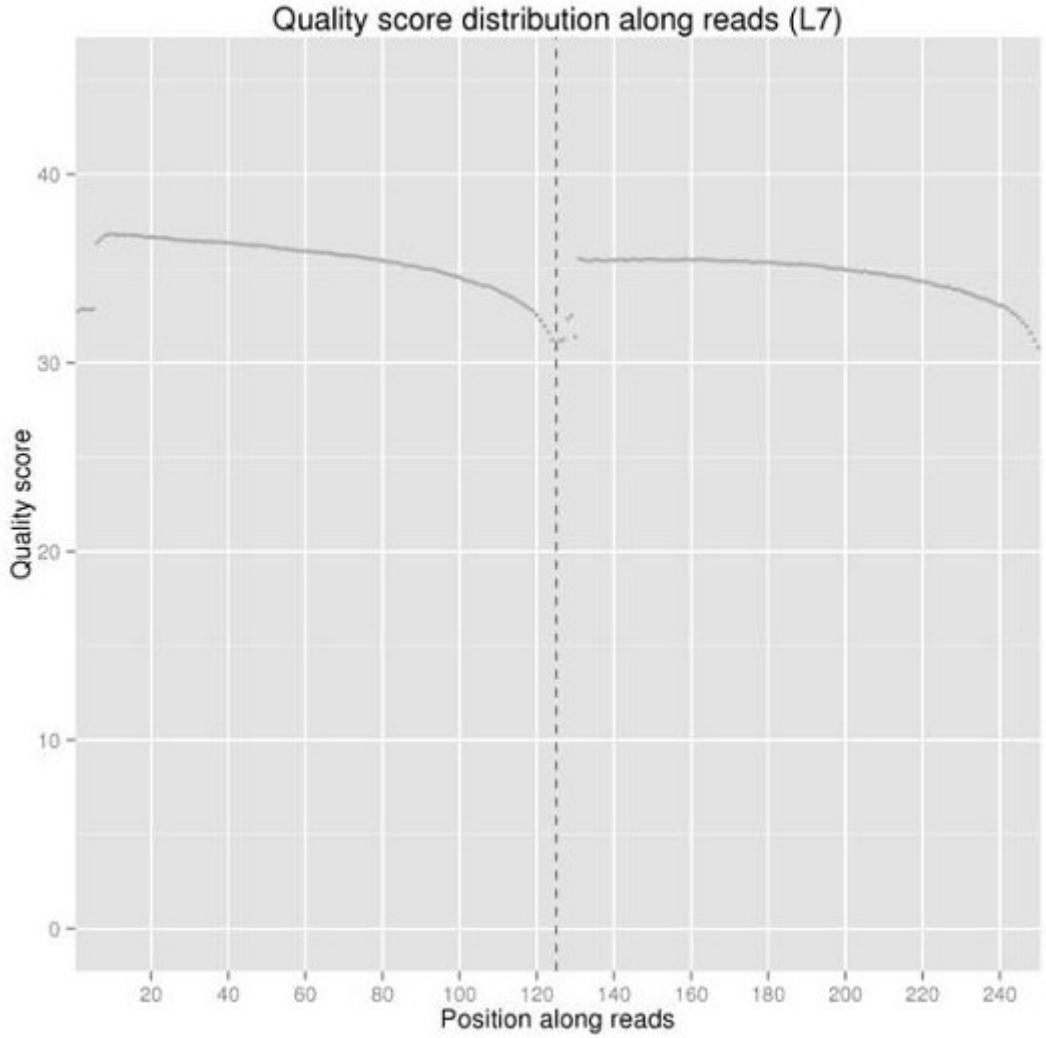
Quality score distribution along reads (L7)
Quality score
Position along reads
0
10
20
30
40
20
40
60
80
100
120
140
160
180
200
220
240

图2

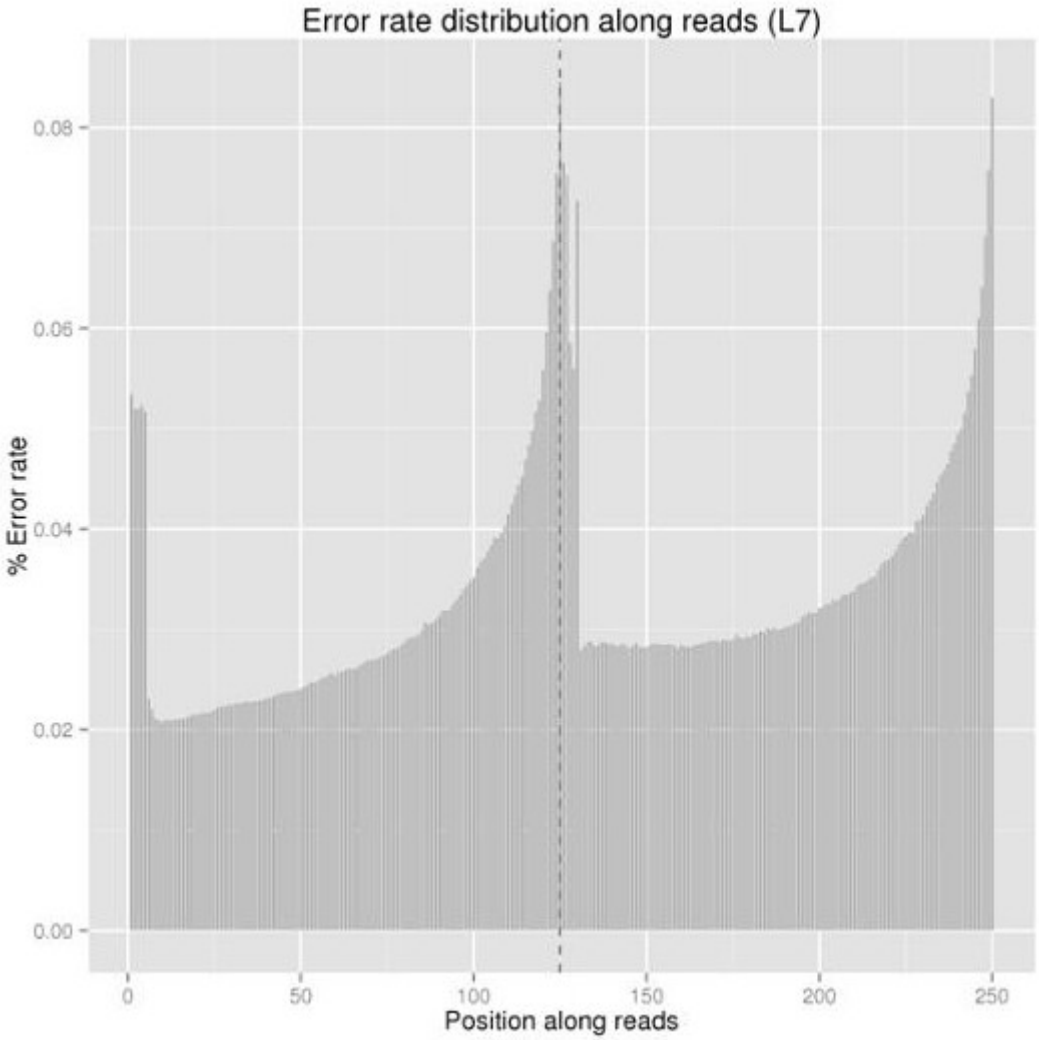
Error rate distribution along reads (L7)
% Error rate
0.08
0.06
0.04
0.02
0.00
0
50
100
150
200
250
Position along reads

图3

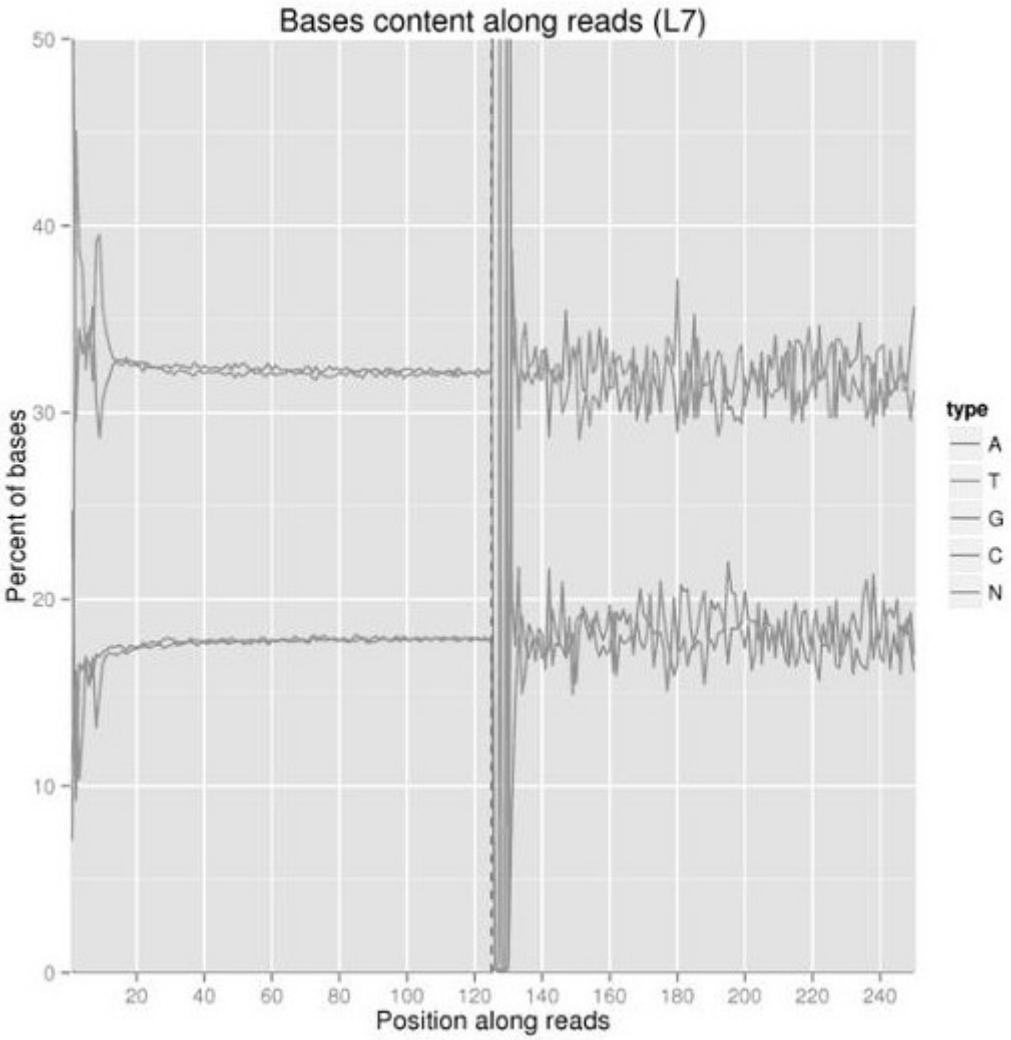
Bases content along reads (L7)
Percent of bases
Position along reads
0
10
20
30
40
50
20
40
60
80
100
120
140
160
180
200
220
240
type
A
T
G
C
N

图4

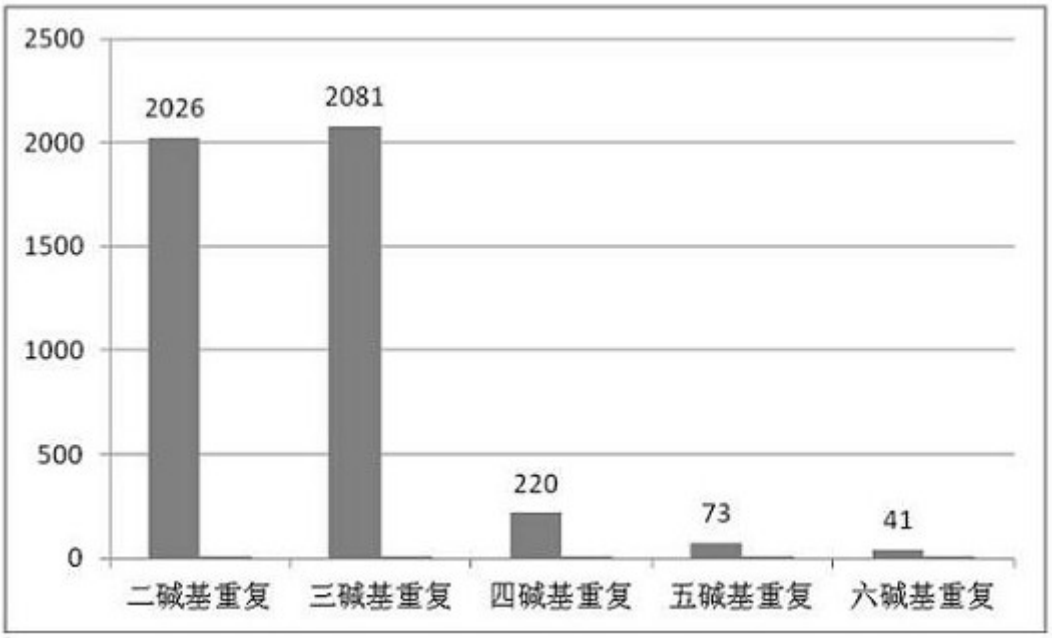
2500
2000
1500
1000
500
0
2026
2081
220
73
41
二碱基重复
三碱基重复
四碱基重复
五碱基重复
六碱基重复

图5

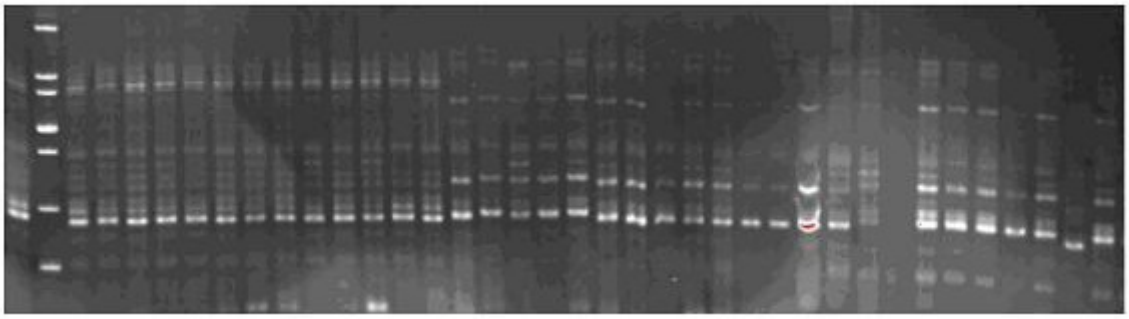

图6