

基于 RAD-seq 技术的异型花 SSR 信息分析

王久利^{1,2} 朱明星³ 徐明行³ 陈世龙¹ 张发起^{1*}

(1. 中国科学院高原生物适应与进化重点实验室, 中国科学院西北高原生物研究所, 西北生态环境资源研究院, 西宁 810001; 2. 中国科学院大学, 北京 100039; 3. 青海大学生态环境工程学院, 西宁 810016)

摘要 用 RAD-seq (restriction-site associated DNA sequencing) 对异型花 (*Sinoswertia tetraptera* (Maxiowicz) T. N. Ho, S. W. Liu & J. Q. Liu) 进行简化基因组测序, 并借此分析异型花的 SSR (simple sequence repeats) 信息。利用 SR search 软件甄别所得序列中的 SSR, 得到了两端各有至少 100 bp 的 SSR 位点 5 844 个, 其中 5 339 个 (91.38%) 成功设计引物, 而三核苷酸 SSR 位点最多 (3 323 个); 在能成功设计引物的 SSR 位点中, 重复序列长度包括 17 种 (12 ~ 36 bp); 重复序列的基序共 277 种, 其中五核苷酸基序种类最多 (106 种); 随机挑选 10 对 SSR 引物, 用 4 个异型花居群的 32 个个体检测其可用性和多态性。经 PCR 和聚丙烯酰胺凝胶电泳检测, 有 4 对 (ST2、ST3、ST6 和 ST10) 成功扩增并表现出多态性; 经 GENEPOP 4.4 对 4 个位点分析, 显示其等位基因数量均值为 6, 多态性较高且互不连锁 ($P < 0.01$); 4 个位点在多数居群中偏离哈迪温伯格平衡 ($P < 0.01$) 且存在较高的纯合子数量 (观测杂合度均值 0.023), 该结果归因于异型花主要进行自花授粉, 在自然界中很难形成进行自由交配的居群; 此外, ST2 和 ST6 可在椭圆叶花锚 (*Halenia elliptica* D. Don) 中成功扩增, 具有潜在通用性。本研究将为日后基于异型花 SSR 标记的相关研究提供数据库支持。

关键词 RAD-seq; 简化基因组测序; 异型花; SSR; 獐牙菜亚族

中图分类号: Q949.776.4 文献标志码: A doi: 10.7525/j.issn.1673-5102.2017.03.016

Analysis on SSR in *Sinoswertia tetraptera* Base on RAD-seq

WANG Jiu-Li^{1,2} ZHU Ming-Xing³ XU Ming-Hang³ CHEN Shi-Long¹ ZHANG Fa-Qi^{1*}

(1. Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Xining 810008; 2. University of the Chinese Academy of Sciences, Beijing 100039; 3. College of Ecology-Environment Engineering, Qinghai University, Xining 810016)

Abstract We used the restriction-site associated DNA sequencing (RAD-seq) technology to analyze simple sequence repeats (SSR) information of *Sinoswertia tetraptera* (Maxiowicz) T. N. Ho, S. W. Liu & J. Q. Liu. The 5844 SSR loci, with at least 100 bp at two ends, were identified using SR search software. The 5339 (91.38%) loci's primers were designed successfully. Among which, amount of tri-nucleotide SSR loci is the most (3323); repeat sequence length type number is 17, while repeat motif type number is 227, and type number of penta-nucleotide motif is the most (106). We employed 32 individuals from 4 natural populations of *S. tetraptera* to estimate usability and polymorphism of 10 pairs of SSR primers selected randomly from the 5339 pairs of primers. According to the result of PCR and Polyacrylamide gel electrophoresis, four pairs of primers (ST2, ST3, ST6 and ST10) amplified favorably and showed polymorphism. By the GENEPOP 4.4, the mean number of alleles of the four loci is 6; these loci do not link to each other ($P < 0.01$); these loci deviate from

基金项目: 国家自然科学基金 (31400322); 青海省应用基础研究计划 (2016-ZJ-761)

第一作者简介: 王久利 (1991—) 男, 博士研究生, 主要从事植物学方面的研究。

* 通信作者: E-mail: fqzhang@nwipb.cas.cn

收稿日期: 2016-12-26

Foundation item: National Natural Science Foundation of China (31400322); Applied Basic Research Programs of Qinghai Province (2016-ZJ-761)

First author introduction: WANG Jiu-Li (1991—) male, doctoral candidate, Botany major.

* Corresponding author: E-mail: fqzhang@nwipb.cas.cn

Received date: 2016-12-26

HWE ($P < 0.01$) in most populations and have many homozygotes (observed heterozygosity mean 0.023), which due to the cleistogamous pollination mode of *S. tetraptera*. ST2 and ST6 were successfully amplified in *Halenia elliptica*. Our study will offer a SSR dataset in the future based on SSR markers of *S. tetraptera*.

Key words RAD-seq; reduced-representation sequencing; *Sinoswertia tetraptera*; SSR; subtribe Swertiinae

异型花 (*Sinoswertia tetraptera* (Maxiowicz) T. N. Ho, S. W. Liu & J. Q. Liu) 又名四数獐牙菜, 是龙胆科 (Gentianaceae) 异型花属 (*Sinoswertia*) 的青藏高原特有的一年生草本植物^[1-3]。特有种的 DNA 记录了其对特殊环境的响应等信息, 这些信息对物种演化的研究具有重要意义^[4-6]; 异型花也是一种重要的藏药, 主要应用于肝胆疾病的治疗^[7]。随着近年来人们对传统医药的重视以及加工技术的发展, 市场对传统药材的需求量增大, 异型花也因此受到一定的威胁, 因此有必要研究其 DNA 水平上的遗传多样性^[8]; 目前对异型花的研究多集中在植物化学分析领域^[8], 然而在 DNA 水平上研究异型花遗传多样性的研究, 已见报道的仅有 Yang 等^[9]利用 ISSR 指纹图谱分析了异型花的居群遗传结构, 并提出了对异型花的遗传多样性高的区域进行保护的策略。

简单重复序列 (simple sequence repeats, SSR) 又称微卫星 (microsatellite), 一般认为是由 2 至 6 个碱基对 (base pair, bp) 为重复单元的序列, 大量分布于真核生物和原核生物基因组中。其共显性和高重复性比 RAPD、ISSR 和 AFLP 等分子标记更具优势^[10], 因而被用于指纹图谱、遗传变异分析、物种起源进化、基因定型、遗传图谱构建、品种评估和微生物鉴定等研究中^[11-16]。

简化基因组测序 (Reduced-representation sequencing) 是一类利用不同方法得到目的基因组的部分区域, 并对这些区域进行测序, 从而反映基因组一部分序列结构信息的一种新兴的测序技术, 其与全基因组测序相比具有低成本优势^[17]。紧随技术的发展, 简化基因组测序有产生了多种不同的方法, 而利用限制性酶切位点相关的 DNA 测序 (Restriction-site associated DNA sequence, RAD-seq)^[18] 是其中具有代表性的方法。RAD-seq 具有通量高、准确性高、数据利用率高、实验周期短、不受基因组序列的限制以及性价比高等诸多优点^[19], 因而其广泛应用于分子标记开发、系统发育研究、种群结构分析、基因功能研究、物种起源研究等众多领域^[20-23]。

为了在利用异型花资源的同时能更好地保护

这一珍贵物种, 本研究中将基于 RAD-seq 技术分析异型花的 SSR 信息并讨论基于 RAD-seq 技术开发异型花的多态性 SSR 标记的可行性, 为异型花的 SSR 多态性位点引物的设计提供便利的数据库, 进一步为基于异型花 SSR 多态性位点相关研究奠定基础。

1 材料与方法

1.1 样品的采集与 DNA 提取

本研究所用异型花样品分别采自定西、果洛、临夏和甘孜的 4 个自然居群, 所用椭圆叶花锚的样品采自 5 个自然居群, 各采样点的坐标和海拔信息在采集地中心使用 GPS 全球定位系统 Etre GIS 采集器记录 (Garmin, 中国台湾) (表 1)。凭证标本保存于中国科学院西北高原生物研究所青藏高原植物标本馆 (HNWP)。从每个居群随机选取的一个个体的叶片用于提取 DNA, 叶片的采集、处理的方法以及 DNA 提取的方法同 Zhang 等^[24], 并用 1% 的琼脂糖凝胶电泳和分光光度计 NanoDrop 2000C (Thermo Scientific, USA) 检测所提取的 DNA, 以保证 DNA 的总量和纯度。

1.2 文库构建与测序

将来自 4 个不同居群异型花的合格 DNA 样品等量混合之后, 交由北京诺禾致源生物信息科技有限公司进行建库测序。建库过程具体参见 Davey 和 Blaxter^[25]。本研究回收 300 ~ 700 bp 的序列, 通过 HiSeq 2000 基因分析系统 (Illumina, USA) 进行双向测序。

1.3 SSR 位点甄别与引物设计

利用 SR search 软件 (Novogene, 中国北京) 以重复单元长度 2 ~ 6 bp、SSR 最小长度 12 为标准, 检测 DNA 序列所有 SSR; 过滤掉其中距离过近的 SSR 序列, 两个 SSR 的最小距离为 12 bp; 为方便引物设计, SSR 的重复序列的上下游各有长度为不短于 100 bp 的序列; 然后利用 SR search 软件中的 primer3 模块设计 SSR 引物^[26]; 参数设置: 引物长度 20 ~ 28 bp, 最适长度 24 bp, 引物退火温度 60 ~ 65°C 且每对引物退火温度差最大值为 1, 最佳退火温度 63°C。

表 1 异型花和椭圆叶花锚样品的居群信息
Table 1 Information of populations of *S. tetraptera* and *Halenia elliptica*

物种 Species	标本号 Specimen code	地区 Locality	纬度 Latitude(N)	经度 Longitude(E)	海拔 Altitude(m)	个体数 Number of individuals
异型花 <i>S. tetraptera</i>	chen2013015	定西 Dingxi	35°00'00"	103°59'17"	2 550	8
异型花 <i>S. tetraptera</i>	chen2013161	果洛 Guoluo	32°51'19"	100°52'54"	3 670	8
异型花 <i>S. tetraptera</i>	chen2014008	临夏 Linxia	35°34'31"	102°45'50"	1 910	8
异型花 <i>S. tetraptera</i>	chen2014166	甘孜 Ganzi	32°14'21"	100°06'24"	3 840	8
椭圆叶花锚 <i>H. elliptica</i>	zhang2014351	海北 Haibei	37°46'50"	101°11'13"	3 424	2
椭圆叶花锚 <i>H. elliptica</i>	zhang2014367	西宁 Xining	37°06'25"	101°42'22"	2 650	2
椭圆叶花锚 <i>H. elliptica</i>	chen2014320	林芝 Linzhi	29°14'00"	94°14'37"	2 970	2
椭圆叶花锚 <i>H. elliptica</i>	chen2014435	错那 Cuona	27°48'28"	91°59'30"	3 720	2
椭圆叶花锚 <i>H. elliptica</i>	zhang2014244	临夏 Linxia	35°34'24"	102°46'51"	3 139	2

1.4 PCR 扩增与检测

随机抽取 10 对 SSR 引物(对应的位点囊括二、三、四、五、六核苷酸重复类型 表 4) 利用梯度退火温度 PCR 反应选定引物的最佳退火温度: 在 20 μ L 的 PCR 反应总体积中, 含 30 $\text{ng} \cdot \mu\text{L}^{-1}$ 的 DNA 模板(4 个来自不同居群的个体的混合 DNA) 0.9 μ L, 10 \times PCR Buffer 2.0 μ L, 10 μ M 正反向引物各 0.4 μ L, 10 $\text{mmol} \cdot \text{L}^{-1}$ dNTPs 0.7 μ L, Taq DNA 聚合酶(5 U $\cdot \mu\text{L}^{-1}$) 0.1 μ L, ddH₂O 15.5 μ L。PCR 扩增在 Mastercycler pro PCR 仪(Eppendorf 德国) 上进行, 扩增程序为【95 $^{\circ}$ C, 5 min; (94 $^{\circ}$ C, 30 s; T_m , 60 s; 72 $^{\circ}$ C, 60 s) \times 25; 72 $^{\circ}$ C, 7 min; 16 $^{\circ}$ C, ∞ 】(T_m 为各引物的退火温度)。接着, 用 1.5% 的琼脂糖凝胶对 PCR 产物进行电泳检测, 以溴化乙锭(ethidium bromide, EB) 染色, 在 GelDoc 2000 凝胶成像系统(Bio-Rad, 美国) 上拍照。根据琼脂糖凝胶电泳结果选定可成功 PCR 扩增的引物及其最佳退火温度。

然后用 4 个自然居群总共 32 个异型花个体(表 1) 对能成功扩增引物按照上述体系和程序进行 PCR, 其中退火温度换成相应引物的最佳退火温度。然后将 PCR 产物进一步用 20% 的聚丙烯酰胺凝胶进行电泳检测, 然后同样以 EB 染色并在 GelDoc 2000 凝胶成像系统上拍照。

2 数据分析与结果

2.1 RAD-seq 测序数据质量评估

异型花混池 DNA 的 RAD 测序共产生 Raw data 2.76G, 过滤后得到 Clean data 2.71G。Phred 数值大于 20、30 的碱基占总体碱基的百分比分别不低于 93.35% 和 88.0%, GC 含量为 36.12%;

RAD-Tag 捕获率为 97.68%。可见 样本的数据量足够, 测序质量合格, GC 分布正常, 建库测序成功。

2.2 聚类

对样本含有酶识别位点 reads 用 Cd-hit-est 把 RAD-tag 相近的 reads 聚集到一个类中^[27]。对每一类的 reads 支持数进行初步的过滤, 过滤标准为每类中 reads 支持数为 10 ~ 400, 数据聚类后共得到 1 275 817 类, 用于聚类的所有 reads 数为 7 818 523, 对 reads 支持深度进行过滤之后所剩下的 pair reads 数为 5 127 265, 过滤之后的 reads 数与所有参加聚类 reads 的比例(Cut/Pair) 为 65.58%。

2.3 组装

根据聚类结果用 Velvetopt 对筛选后的每一类进行局部组装^[28], 得到组装序列后过滤掉 125 bp 以下的 contig。本次组装总长度为 93 655 566 bp, 组装总数为 301 821, contig 平均长度为 310, 序列从大到小排列, 当长度达到组装总长度一半时, contig 的长度 $N_{50} = 425$ bp。组装结果中的 GC 含量为 35.83% 而所有 reads 的 GC 含量为 36.12%。

将去重复之后的 reads 通过 Burrows-Wheeler alignment tool(BWA) 软件比对到组装结果中^[29], 利用 SAMtools 检测变异情况^[30]。检测结果显示, 所有能比对上组装基因组的 reads 占有参加比对 reads 的 88.8%; 每个位点深度为 16.04; reads 能够覆盖组装基因组的 98.56%; reads 能够覆盖组装基因组且深度不低于 4 \times 的比例是 79.63%; 自身 reads 比对组装结果所检测出的 SNP 数为 252 614, 其中杂合 SNP 为 241 335, 杂合比例为 95.54%。样本比对率反映了样本 Clean data 的利用情况, 覆盖深度和覆盖度足以直接反应测序数

据的均一性及与组装基因组的一致性。

2.4 SSR 位点及其引物

利用 SR search 软件对样本组装 contig 进行过滤,最终获得双端各有 100 bp 的 SSR 位点共 5 844 个,并成功设计 5 339 对 SSR 引物,成功设计率为 91.38%。已成功设计引物的 SSR 位点的部分信息列于表 2,我们获得的 SSR 最多的是三核苷酸(3 323 条),其次是二核苷酸(1 167 条),最少的是六核苷酸(89 条);但是,重复序列的基序种类共 277 种,其中基序种类数最多的是五核苷酸(106 种),其次是六核苷酸(89 种),最少的是二核苷酸(11 种)。从能成功设计引物的 5 339 个 SSR 位点的重复序列长度包括 17 种(12~36 bp),从其分布(表 3)来看,长度在 12 bp 的 SSR 位点多达 3 028 个,占总 SSR 位点数的 56.71%;其次是 16 bp 的 SSR 位点,为 594 个(11.13%);而不短于 25 bp 的 SSR 位点有 127 个占总数的 12.38%(成功设计引物的 5 339 个 SSR 位点详细信息详见附件 1)。

2.5 异型花 SSR 引物的有效性分析

通过梯度退火温度 PCR 从 10 对备选 SSR 引物中筛选出可在 *S. tetraptera* 中成功扩增的引物分别是 ST2、ST3、ST6 和 ST10,并确定这 4 对引物的最佳退火温度。聚丙烯酰胺凝胶电泳显示这 4 对引物都表现出多态性,等位基因的大小变化范围见表 4。

表 2 异型花 SSR 序列的信息

Table 2 Information of SSR sequences in *S. tetraptera*

SSR 类型 SSR type	数量 Number	百分比 Percentage	平均长度 Average length(bp)	重复序列的 基序的种类数 Motif type number
二核苷酸 Di-nucleotide	1 167	21.86	14.25	11
三核苷酸 Tri-nucleotide	3 323	62.24	13.31	59
四核苷酸 Tetra-nucleotide	560	10.49	17.24	12
五核苷酸 Penta-nucleotide	200	3.75	17.25	106
六核苷酸 Hexa-nucleotide	89	1.67	25.69	89
总计 Total	5 339	100	—	277

表 3 异型花 SSR 位点重复序列的长度分布

Table 3 Distribution of the length of SSR sequences loci of *S. tetraptera*

长度 Length(bp)	SSR 位点数 Number of SSR loci	分布百分比 Distribution percentage
12	3 028	56.71
14	242	4.53
15	587	10.99
16	594	11.13
18	260	4.87
20	280	5.24
21	62	1.16
22	25	0.47
24	134	2.51
≥25	127	2.38

表 4 异型花的 10 对引物特点

Table 4 Characteristics of 10 SSR primers developed for *S. tetraptera*

位点 Locus	重复基序 Repeat motif	引物序列 Primer sequences (5'-3')	退火温度 T_a ($^{\circ}\text{C}$)	观测杂合度 H_o	等位基因大小 Allele size range(bp)	GenBank 接收号 GenBank accession No.
ST1	(TCT) 6	F: AAGAAAATAAGGTCCTCCCTCCT R: ATAACCGGTCACACACATTTAG	—	—	—	KY315128
ST2	(TGG) 6	F: AGGAAGGGAATCCTCCTATTTTC R: ACCATTATAAGTGGTGCATGTC	57	0	129~138	KY315129
ST3	(TA) 10	F: TTGAACATGCACCTTGACACTTTT R: CATGTTTTGAAAACCTTGAATTT	54	0	120~128	KY315130
ST4	(GACCCG) 6	F: CCTGAGTTATTTGATCAGCAACC R: GATCAAGTAAAATTCCAAGCAGC	—	—	—	KY315131
ST5	(AT) 16	F: TGGATTGTAATTTTAGGGATGTCA R: ATATTGCAATTGTTTTGCTTTGC	—	—	—	KY315132
ST6	(TACA) 8	F: GACTACTGCGAGTAAGCTCGAT R: TCAAAGTGCATAATCCAAGAGA	55	0.093	149~189	KY315133
ST7	(CA) 16	F: TGAAGGCAAACAGAGAGCTAGAG R: TCATTGTAATCCCTCTCTGATCG	—	—	—	KY315134
ST8	(AAAAC) 6	F: AATGCTTTGCCACTTGA AAAATTA R: AGTTCCCTGCTATGAACTTGCTG	—	—	—	KY315135
ST9	(TTA) 11	F: TTGCCTTATCATGGTACGAATCT R: GCACACATAGCACATCCTTCATA	—	—	—	KY315136
ST10	(TTGTA) 6	F: TTCGTATTCTCATTTGTCCATCA R: TAATGTCGTATTAACAATGCCCC	55	0	101~131	KY315137

利用 GENEPOP 4.4^[31] 分析聚丙烯酰胺凝胶电泳结果。所检测的 4 对引物所对应的位点的等位基因数为 4~9, 均值为 6, 多态性较高; 连锁不平衡 (Linkage Disequilibrium) 分析显示 4 个 SSR 位点不连锁 ($P < 0.01$); 而 ST2、ST3、ST6 和 ST10 的观测杂合度分别为 0.0、0.093 和 0, 均值 0.023。

异型花属是单型属, 其近缘类群为花锚属 (*Halenia* Borkh.) 和獐牙菜属 (*Swertia* Linn.)^[2~3], 因而用椭圆叶花锚 (*Halenia elliptica* D. Don) 对异型花多态性引物进行近缘物种交叉检验, 发现 ST2 和 ST6 可以在椭圆叶花锚中成功扩增。

3 讨论

目前, 异型花尚未获得全基因组或转录组数据, 甚至也没有龙胆科内近缘物种的可参考全基因组序列。因而我们采用 RAD-seq 技术对异型花进行简化基因组测序, 测序结果成功, 获得了异型花基因组水平上的序列信息。所得的序列经聚类 and 组装后, 也呈较高的组装准确性, 从侧面反映组装结果能够代表部分基因组。样本比对率 (88.80%) 反映了样本 Clean data 的高利用率, 覆盖深度和覆盖度足以直接反映测序数据的均一性及与组装基因组的一致性。

利用 SR search 软件对获得的序列进行 SSR 甄别, 获得双端各有 100 bp 的 SSR 位点多达 5 844 个, 并成功设计 5 339 对 SSR 引物, 成功设计率高达 91.38%。在能成功设计引物的 SSR 位点中, 重复序列的基序种类共 277 种, 而且重复序列长度包括 17 种, 可见已成功设计引物的异型花的 SSR 多样性较高。随机挑选的 10 对 SSR 引物中, 经 PCR 和凝胶电泳检测, 有 4 对可在 *S. tetraptera* 中成功扩增, 并且表现出多态性。经 GENEPOP 4.4 对已获得的位点分析, 显示其等位基因的均值为 6, 多态性较高, 并且相互之间不连锁; 4 个位点偏离哈迪温伯格平衡, 这是因为异型花主要进行自花授粉, 在自然界中很难形成进行自由交配的居群^[2~3], 较高的纯合子数量 (观测杂合度均值 0.023) 可能也归因于其授粉方式; 4 对引物中有 2 对可以在其近缘异属中成功扩增, 表现出了一定的通用性。可见, 本研究将为后续的基于 SSR 标记的异型花居群遗传和指纹图谱等研究提供足够的 SSR 标记数据库。

值得补充的是, 我们用同样的方法开发了肉

果草 (*Lancea tibetica* Hook. F. et Thoms)^[23] 和椭圆叶花锚 (待发表数据) 的多态性 SSR 位点, 通过与我们以往的窄叶鲜卑花 (*Sibiraea angustata* (Rehder) K. S. Hao)^[32]、绣线菊属 (*Spiraea* Linn.)^[33]、西川红景天 (*Rhodiola alsia* (Frod.) S. H. Fu)^[34] 和黄绿蜜环菌 (*Armillaria luteo-virens* (Aalb. et Schw. Fr.) Sacc.)^[35] 等物种的 SSR 标记开发研究的比较, 我们发现基于 RAD-seq 技术进行 SSR 标记研究表现出了 RAD-seq 技术本身所具备的优势, 包括通量高、准确性高、数据利用率高、实验周期短、不受基因组序列的限制以及性价比高^[18~19], 在 SSR 位点开发方面明显优于传统的 FIASCO (Fast Isolation by AFLP of Sequences Containing Repeats)^[36] 等方法。所以, 对于尚无足够核酸序列数据参考的物种, RAD-seq 是目前的一种比较理想的 SSR 信息研究技术。

参 考 文 献

1. He T N, Liu S W, Chen S L. Nomenclatural novelties in *Swertia* (Gentianaceae) [J]. *Plant Diversity and Resources* 2013, 35(3): 386-392.
2. He T N, Liu S W, Liu J Q. A new Qinghai-Tibet Plateau endemic genus *Sinoswertia* and its pollination mode [J]. *Plant Diversity and Resources* 2013, 35(3): 393-400.
3. Ho T N, Liu S W. A worldwide monograph of *Swertia* and its allies [M]. Beijing: Science Press 2015: 301-308.
4. Fu P C, Gao Q B, Zhang F Q, et al. Responses of plants to changes in Qinghai-Tibetan Plateau and glaciations: Evidence from phylogeography of a *Sibiraea* (Rosaceae) complex [J]. *Biochemical Systematics and Ecology* 2016, 65: 72-82.
5. Gao Q B, Zhang D J, Duan Y Z, et al. Intraspecific divergences of *Rhodiola alsia* (Crassulaceae) based on plastid DNA and internal transcribed spacer fragments [J]. *Botanical Journal of the Linnean Society* 2012, 168(2): 204-215.
6. Wang L Y, Abbott R J, Zhang W, et al. History and evolution of alpine plants endemic to the Qinghai-Tibetan Plateau: *Aconitum gymnandrum* (Ranunculaceae) [J]. *Molecular Ecology* 2009, 18(4): 709-721.
7. 中国科学院西北高原生物研究所. 藏药志 [M]. 青海: 青海人民出版社, 1991: 111-112.
8. 杨路存, 陈桂琛, 周国英, 等. 四数獐牙菜 ISSR-PCR 反

- 应体系的正交优化 [J]. 生物技术通报, 2010, 20(9): 123 - 128.
- Yang L C, Chen G C, Zhou G Y, et al. Optimization for IS-SR-PCR reaction system in *Swertia tetraptera* using orthogonal design [J]. Biotechnology Bulletin, 2010, 20(9): 123 - 128.
9. Yang L C, Zhou G Y, Chen G C. Genetic diversity and population structure of *Swertia tetraptera* (Gentianaceae), an endemic species of Qinghai-Tibetan Plateau [J]. Biochemical Systematics and Ecology, 2011, 39(4-6): 302 - 308.
10. Wang X R, Szmidt A E. Molecular markers in population genetics of forest trees [J]. Scandinavian Journal of Forest Research, 2001, 16(3): 199 - 220.
11. 何平. 真核生物中的微卫星及其应用 [J]. 遗传, 1998, 20(4): 42 - 47.
- He P. Abundance, Polymorphism and applications of microsatellite in eukaryote [J]. Hereditas, 1998, 20(4): 42 - 47.
12. Doerge R W. Mapping and analysis of quantitative trait loci in experimental populations [J]. Nature Reviews Genetics, 2002, 3(1): 43 - 52.
13. Liu W J, Nie H, Wang S B, et al. Mapping a resistance gene in wheat cultivar Yangfu 9311 to yellow mosaic virus, using microsatellite markers [J]. Theoretical and Applied Genetics, 2005, 111(4): 651 - 657.
14. Yoshimoto K, Yoshida J, Ishii G, et al. Two lung adenocarcinomas in the same lobe: multiple primaries or intrapulmonary metastasis? [J]. Annals of Thoracic and Cardiovascular Surgery, 2011, 17(6): 584 - 587.
15. 陶星辰, 尚秋菊, 刘晓英, 等. 云南省部分地区鸽粪中新型隐球菌的分离与鉴定 [J]. 中国病原生物学杂志, 2014, 9(10): 873 - 876.
- Tao X C, Shang Q J, Liu X Y, et al. Isolation and identification of *Cryptococcus neoformans* from pigeon droppings in parts of Yunnan Province [J]. Journal of Pathogen Biology, 2014, 9(10): 873 - 876.
16. Conti S, Gallo E, Sioletic S, et al. Molecular genetic alterations in egfr CA-SSR-1 microsatellite and egfr copy number changes are associated with aggressiveness in thymoma [J]. Journal of Thoracic Disease, 2016, 8(3): 386 - 395.
17. 石璇, 王茹媛, 唐君, 等. 利用简化基因组技术分析甘薯种间单核苷酸多态性 [J]. 作物学报, 42(5): 641 - 647.
- Shi X, Wang R Y, Tang J, et al. Analysis of interspecific SNPs in sweetpotato using a Reduced-Representation genotyping technology [J]. Acta Agronomica Sinica, 2016, 42(5): 641 - 647.
18. Miller M R, Dunham J P, Amores A, et al. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers [J]. Genome Research, 2016, 17(2): 240 - 248.
19. 王洋坤, 胡艳, 张天真. RAD-seq 技术在基因组研究中的现状及展望 [J]. 遗传, 2014, 36(1): 41 - 49.
- Wang Y K, Hu Y, Zhang T Z. Current status and perspective of RAD-seq in genomic research [J]. Hereditas, 2014, 36(1): 41 - 49.
20. Willing E M, Hoffmann M, Klein J D, et al. Paired-end RAD-seq for de novo assembly and marker design without available reference [J]. Bioinformatics, 2011, 27(16): 2187 - 2193.
21. Hou Y, Nowak M D, Mirr V, et al. RAD-seq data point to a northern origin of the arctic-alpine genus *Cassiope* (Ericaceae) [J]. Molecular Phylogenetics and Evolution, 2016, 95: 152 - 160.
22. Gamble T, Coryell J, Ezaz T, et al. Data from: Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems [J]. Molecular Biology and Evolution, 2015, 32(5): 1296 - 1309.
23. Tian Z Z, Zhang F Q, Liu H R, et al. Development of SSR markers for a Tibetan medicinal plant, *Lancea tibetica* (Phrymaceae), based on RAD sequencing [J]. Applications in Plant Sciences, 2016, 4(11): 1600076.
24. Zhang F Q, Lei S Y, Gao Q B, et al. Isolation of microsatellite loci for *Rhodiola alsia* (Crassulaceae): an important ethno-medicinal herb endemic to the Qinghai-Tibetan plateau [J]. Genetics and Molecular Research, 2015, 14(2): 5266 - 5269.
25. Davey J W, Blaxter M L. RADSeq: next-generation population genetics [J]. Briefings in functional genomics, 2010, 9(5-6): 416 - 423.
26. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers [M]. // Misener S, Krawetz S A, eds. Bioinformatics Methods and Protocols. Totowa, NJ: Humana Press, 1999: 365 - 386.
27. Li W Z, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences [J]. Bioinformatics, 2006, 22(13): 1658 - 1659.
28. Zerbino D R, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs [J]. Genome Research, 2008, 18(5): 821 - 829.
29. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform [J]. Bioinformatics, 2010, 25(5): 589 - 595.
30. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools [J]. Bioinformatics, 2009, 25(16): 2078 - 2079.

(下转 460 页)

