

椭圆叶花锚简化基因组的 SSR 信息分析及 SSR 引物开发

王久利^{1,2} 陈世龙¹ 邢睿¹ 宋相杰³ 朱明星³ 张发起^{1,4*}

(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 810008; 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; 4. 青海省作物分子育种重点实验室, 西宁 810008)

摘要 基于 RAD-seq 简化基因组测序技术获得椭圆叶花锚(*Halenia elliptica* D. Don) 简化基因组水平上的序列信息并开发 SSR 分子标记。利用 SR search 软件检测而得到双端各有至少 100 bp 的五种类型的 SSR(二、三、四、五、六核苷酸) 位点共 6 201 个, 并成功设计其中 3 865 个 SSR 引物。在能成功设计引物的 SSR 位点中, 三核苷酸 SSR 位点最多; 重复序列长度包括 17 种(12~36 bp); 重复序列的基序共达 316 种, 其中五核苷酸基序种类最多(91 种)。从中挑选 65 对可对应五种 SSR 类型的引物, 经梯度 PCR 检验后, 利用椭圆叶花锚的 4 个居群 32 个个对可扩增的引物进行 PCR 和聚丙烯酰胺凝胶电泳检测, 其中 14 对引物能在绝大多数个体中扩增, 且 13 对具多态性。13 个多态性位点的等位基因数量均值为 5.462, 多态性较高且不连锁($P < 0.05$); 其中 10 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 H_o 均值 0.226); 近交系数在 $-0.443 \sim 1$, 均值为 0.656; 基因流 N_m 为 0.474。

关键词 椭圆叶花锚; RAD-seq; SSR 分子标记; 简化基因组测序

中图分类号: Q949.776.4 文献标志码: A doi: 10.7525/j.issn.1673-5102.2018.02.018

Simplified Genome SSR Information and Development of SSR Primers of *Halenia elliptica*(Gentianaceae)

WANG Jiu-Li^{1,2} CHEN Shi-Long¹ XING Rui¹ SONG Xiang-Jie³ ZHU Ming-Xing³ ZHANG Fa-Qi^{1,4*}

(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; 4. Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Xining 810008)

Abstract The restriction-site associated DNA sequencing (RAD-seq) technology was used to analyze simple sequence repeats (SSR) information of *Halenia elliptica* D. Don; SSR markers for this species were developed based on RAD-seq data. We identified 6 201 SSR loci with at least 100 bp at two ends, including five types (di-nucleotide, tri-nucleotide, tetra-nucleotide, penta-nucleotide, and hexa-nucleotide). Primers of 3 865 loci were designed successfully. Among which, amount of tri-nucleotide SSR loci is the most. Repeat motif type number is 316, and type number of penta-nucleotide motif is the most (91). Afterward, 65 designed primers were randomly selected to represent the five SSR types, and tested using annealing temperature gradient PCR. Subsequently, 32 individuals from four natural populations of *H. elliptica* were employed to estimate the SSR primers that can be amplified. One monomorphic and 13 polymorphic SSR loci were developed according to the

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); 青海省自然科学基金青年基金(2015-ZJ-925Q); 青海省作物分子育种重点实验室(2017-ZJ-Y14)

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: 2017-09-21

Foundation item: National Natural Science Foundation of China(31400322); Applied Basic Research Programs of Qinghai Province(2016-ZJ-761); The Natural Science Foundation of Qinghai Province(2015-ZJ-925Q); Qinghai Provincial Key Laboratory of Crop Molecular Breeding (2017-ZJ-Y14)

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

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

Received date: 2017-09-21

result of PCR and Polyacrylamide gel electrophoresis. The number of alleles of the 13 polymorphic SSR loci is high (mean 5.462), and no linkage relation was found ($P < 0.01$); ten loci deviate from HWE ($P < 0.01$) and have many homozygotes (observed heterozygosity mean of 0.226); the range of inbreeding coefficient is in $-0.443 - 1$ (mean of 0.656); the gene flow N_m is 0.474.

Key words *Halenia elliptica*; RAD-seq; SSR molecular marker; simplified genome sequencing

椭圆叶花锚 (*Halenia elliptica* D. Don) 是龙胆科 (Gentianaceae) 的草本植物, 主产于青藏高原及其周边, 分布较广^[1]。椭圆叶花锚在藏药、蒙药、普米药、佉药等传统民族医药中均有应用, 主要用于治疗肝胆疾病, 同时具有清热祛湿等功效^[2~3]。目前已有多种使用椭圆叶花锚治疗肝胆疾病的准字成药, 资源需求量较大^[4], 当前该植物药材资源主要为野生资源, 因而合理地利用和保护椭圆叶花锚野生资源以及准确地鉴定该物种已经成为亟待考虑的议题。基于分子生物学的生物遗传多样性研究可以为生物保护提供一定的理论依据, 同时基于分子生物学——尤其是基于 SSR 的指纹图谱也为物种的准确鉴定提供快速而有效的工具^[5]。

目前, 对椭圆叶花锚的研究主要集中在资源学、化学成分提取与分析、有效成分的含量测定及药理作用等方面^[3~4, 6~9]; 而遗传多样性等从分子生物学角度分析的研究工作, 主要有 Zhang 等^[10]利用磁珠富集法 (FIASCO) 开发了 12 个椭圆叶花锚的多态性微卫星 DNA (microsatellite DNA, 亦称 SSR), 但是并没有对椭圆叶花锚较全面的 SSR 信息分析的研究。鉴于花锚属植物尚无足够的核酸序列数据可供参考 (没有公布的基因组或转录组数据), 因而本研究将利用限制性酶切位点相关的 DNA 测序 (Restriction-site associated DNA sequence, RAD-seq)^[11~13] 这一较理想的 SSR 信息研究技术, 对椭圆叶花锚的 SSR 信息进行较全面的分析, 并据此开发其多态性分子标记。

1 材料与方法

1.1 样品采集与 DNA 提取

椭圆叶花锚的 DNA 材料采自分别分布于青海、甘肃、四川和云南 4 个省份的 4 个居群 (表 1)。每个居群选取 8 个个体, 居群内个体之间至少相隔 15 m。DNA 材料的采集、凭证标本的保存以及总 DNA 的提取与检测方法同王久利等^[14]。

1.2 文库构建、测序与序列数据处理

将不同居群的 4 个个体的基因组 DNA 等量混合, 由北京诺禾致源生物信息科技有限公司进

行建库测序。检测合格的 DNA 样品通过酶切、随机打断、经末端修复、加 A 尾、加测序接头、纯化、PCR 扩增等步骤完成整个文库制备。库检合格后, 在 Illumina HiSeqTM 2000 (USA) 平台上双向测序, 获得原始测序序列 (raw data), 用序列分析软件 Illumina Casava 1.8^[15] 进行碱基识别, 然后对 raw data 进行过滤, 接着用 cd-hit-est 软件^[16] 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 Velvet 对筛选后的类进行组装^[17], 组装后, 过滤 125 bp 以下的 contig。

表 1 椭圆叶花锚的居群信息

Table 1 Information on the populations of *H. elliptica*

居群 Population	区域 Location	纬度/北纬 Latitude/N	经度/东经 Longitude/E	海拔 Altitude (m)	个体数 Number of individuals
chen2013577	云南香格里拉 Yunnan Shangri-La	27°53'47"	99°43'37"	3 331	8
chen2013014	甘肃渭源 Gansu Weiyuan	35°00'00"	103°59'17"	2 550	8
zhang2014351	青海门源 Qinghai Menyuan	37°46'50"	101°11'13"	3 424	8
chen2013453	西藏亚东 Tibet Yadong	27°44'38"	88°58'01"	3 065	8

1.3 SSR 序列检测、引物设计、扩增与检测

SSR 序列检测、引物设计、扩增与检测过程中, 软件参数设置及实验策略参照王久利等^[14]的方法。利用 SR search 软件 (Novogene, 北京) 对检测经组装好的 contig 以寻找 SSR 序列, 然后利用 SR search 软件中的 Primer 3^[18] 模块设计 SSR 引物。从所设计的引物中随机挑选 65 对可代表 2 - 6 核苷酸五种基序 (motif) 类型的引物 (对应位点序列的 NCBI 接收号: KY315063-KY315127), 交由生工生物工程 (上海) 股份有限公司 (中国) 合成, 利用梯度退火温度 PCR 反应选定成功扩增的引物并确定其最佳退火温度: 在 20 μ L 的 PCR 反应体系中, 含 30 $\text{ng} \cdot \mu\text{L}^{-1}$ 的 DNA 模板 0.9 μL , 10 \times PCR Buffer 2.0 μL , 10 $\mu\text{mol} \cdot \text{L}^{-1}$ 正反向引物各 0.4 μL , 10 $\text{mmol} \cdot \text{L}^{-1}$ dNTPs 0.7 μL , Taq DNA 聚合酶 (5 U $\cdot \mu\text{L}^{-1}$) 0.1 μL , ddH₂O 15.5 μL 。PCR 扩增在 Mastercycler pro PCR 仪 (Eppendorf, 德国)

上进行扩增程序为 [95°C 5 min; (94°C 30 s; T_m , 60 s; 72°C ,60 s) × 25; 72°C ,7 min; 16°C ,∞] (T_m 为各引物的梯度退火温度,50 ~ 60°C)。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,然后拍照。

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,然后染色并拍照。

2 数据分析与结果

2.1 RAD-seq 测序数据质量评估

椭圆叶花锚 DNA 的 RAD 测序共产生 Raw data 2.754G,过滤后得到 Clean data 2.722G。若测序错误率用 e 表示,illumina HiSeq™ 2500/Miseq™ 的碱基质量值用 Qphred 表示,则有: $Qphred = -10\log_{10}(e)$,Phred 分值大于 20(Q_{20})碱基占总体碱基的百分比不低于 93.72%;GC 含量为 34.10%;RAD-Tag 捕获率为 98.21%。可见样本的数据量足够,测序质量合格,GC 分布正常,建库与测序均成功。

2.2 聚类与组装

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,并对每一类的 reads 支持数进行初步的过滤,数据聚类后共得到 1 155 618 类,用于聚类的所有 reads 为 8 635 429,统计类中的 reads 支持数在 10 ~ 400 之间的类数目有 174 113,对 reads 支持深度进行过滤之后所剩下的 pair reads 数为 5 558 782,为过滤之后的 reads 数与所有参加聚类 reads 的比例是 64.37%。

本次组装总长度为 79 426 526 bp,组装总个数 284 558,其平均长度为 279 bp,序列从大到小排列,当长度达到组装总长度一半时,contig 的长度是 383 bp。组装结果中 GC 含量为 34.16%,而所有 reads 的 GC 含量为 34.10%。GC 含量与测序下机数据的 GC 含量高度一致,反应组装结果能够代表部分基因组。

将去重复之后的 reads 通过 Burrows-Wheeler alignment tool(BWA)软件^[19]比对到组装结果中,利用 SAMtools 进行变异检测^[20]。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 91.45%;每个位点的平均深度为

21.55;reads 能够覆盖组装基因组的 98.39%;reads 能够覆盖组装基因组且深度不低于 4× 的比例是 80%;自身 reads 比对组装结果所检测出的 SNP 数为 208 010,其中杂合 SNP 为 196 243,杂合比例为 94.34%。检测结果中出现大部分的杂合 SNP、少量的纯合 SNP,则说明组装结果可靠。样本比对率反映了样本 Clean data 的利用率情况,覆盖深度和覆盖度直接反应测序数据的均一性及与组装基因组的一致性。

2.3 SSR 检测结果

利用 SR search 软件对样本组装 contig 进行过滤共得到 6 201 个双端含有 100 bp(便于设计引物)SSR 片段。成功设计引物的片段数量为 3 865 个,成功设计率为 62.33%,其中 SSR 数量较多的基序类型是二、三、四核苷酸重复,SSR 数量分别是 833、1 666、1 133,分别占总数的 21.55%、43.11%、29.31%。而五、六核苷酸重复类型所占比例较少,分别为 3.49%、2.54%。在获得的五种 SSR 类型中,SSR 的基序类型达 316 种,其中五、六核苷酸重复类型的种类数居多,分别为 91、83 种;最少的是二核苷酸重复,仅 10 种(表 2)。

表 2 椭圆叶花锚的 SSR 序列信息

Table 2 Information of SSR sequences in *H. elliptica*

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	833	21.55%	14.51	10
三核苷酸 Tri-nucleotide	1666	43.10%	13.27	58
四核苷酸 Tetra-nucleotide	1133	29.31%	17.71	74
五核苷酸 Penta-nucleotide	135	3.49%	21.15	91
六核苷酸 Hexa-nucleotide	98	2.54%	25.22	83

所得 SSR 的长度为 12 ~ 36 bp,平均长度为 15.42 bp,长度为 12 bp 的位点数量最多,最少的为 27 和 36 bp;不同重复序列的长度分布存在较大差异,而不同核苷酸数目的基序的重复次数亦各不相同(表 3)。二核苷酸基序的 SSR 主要为 6、7 次重复(68.79%);三核苷酸基序的 SSR 主要为 4、5、6 次重复(96.70%);四核苷酸基序组成的 SSR 有 5 种主要为 4、5 次重复(91.62%);五、六核苷酸基序组成的 SSR 主要为 4 次重复(80.26%)。

总体而言, 在所有基序重复次数中, 以 4 次重复占多数, 共 2 189 个 (56.64%)。其次是五次重复, 共 585 个 (15.14%)。而 15 次重复最少, 仅有 2 个 (0.05%)。在二、三、四、五、六核苷酸五类重复类

型中, 重复序列最多的种类分别为 AT(380 个)、AAT(166 个)、AAAT(191 个)、AAAAT(7 个) 和 CCTAAA(6 个), 它们分别占全部 SSR 序列数量的 9.83%、4.29%、4.94%、0.18% 和 0.15%。

表 3 椭圆叶花锚 SSR 位点不同重复序列的长度分布

Table 3 Distribution of the length of repeat sequences of SSR loci of the *H. elliptica*

长度 Length(bp)	二核苷酸 Di-nucleotide	三核苷酸 Tri-nucleotide	四核苷酸 Tetra-nucleotide	五核苷酸 Penta-nucleotide	六核苷酸 Hexa-nucleotide	SSR 位点数 Number of SSR loci	分布百分比 Distribution percentage(%)
12	396	1211	—	—	—	1 607	41.58
14	177	—	—	—	—	177	4.58
15	—	297	—	—	—	297	7.68
16	98	—	791	—	—	889	23.00
18	69	103	—	—	—	172	4.45
20	41	—	247	108	—	396	10.25
21	—	36	—	—	—	36	0.93
22	27	—	—	—	—	27	0.70
24	13	11	62	—	79	165	4.27
25	—	—	—	23	—	23	0.60
26	9	—	—	—	—	9	0.23
27	—	1	—	—	—	1	0.03
28	1	—	20	—	—	21	0.54
30	2	4	—	4	18	28	0.72
32	—	—	13	—	—	13	0.34
33	—	3	—	—	—	3	0.08
36	—	—	—	—	1	1	0.03

表 4 椭圆叶花锚的 SSR 引物基本特征

Table 4 Basic characteristics of SSR primers of the *H. elliptica*

SSR 位点 SSR loci	引物序列 Primer sequences (5'-3')	重复序列 Repeat motif	退火温度 T_m ($^{\circ}\text{C}$)	片段大小 Fragment size(bp)	SSR 位点 SSR loci	引物序列 Primer sequences (5'-3')	重复序列 Repeat motif	退火温度 T_m ($^{\circ}\text{C}$)	片段大小 Fragment size(bp)
HE1	F: AATGCGAATATGAGACGATTGAT R: TCTATGGTTTCAATCTTTCTCTAAA	TTA	55	120	HE40	F: GACACATTTTCATTCCATTCCAT R: GTTAACTGCACCCTAATCGTTTG	ATTCC	54	120
HE13	F: ATTCTTCTTCTTCCCAGCCTTA R: CAATTCACGATGACTCTCACATC	AAT	55	109	HE42	F: CCACATCCATAATCATCCAATTT R: AGCGGTAGGGTTTATGATTTTCT	ATTTTT	53.5	133
HE15	F: CAAATCTTTGACATAAAGAAAGCG R: TCGGATGTTTTAGTTTGGTTTTT	TAA	55	160	HE44	F: AATAGTGTGAGTGGCCATAGGA R: GGGATGGAAGAAAAAGAAAAAGA	TCTTTT	55	142
HE17	F: GGGGTATATGTTTGTATTCCAGC R: AATGTGAAACCCATCAATAATCG	AAT	55	125	HE50	F: CAACCTTCACCTCATTTCCTTA R: GTGGCATTGAAGAGAAAAACATTC	CCATGT	54	121
HE21	F: GATTACATCTGTGACCCAGAAC R: AAAAGTGGCAAGCTGACAATAAC	CCA	55	157	HE56	F: TTAGCATTTGCTTTTAAACCTTTG R: GGGAACTACTCTACTGCTTCTGC	AGCAGA	55	133
HE22	F: AAACCCATACCTATCCTCAAAA R: TTTTCCGCTTATTTGTTTACTGC	AGC	55	92	HE63	F: TATTGTTGTGGACTTTGTCCCC R: TCCGTCCATGTGAAAATAGAGAT	CAG	55	145
HE36	F: AGCATGCTTCAACACTTAGTAATC R: GCATATCTCTCAGTGACTTGT	TTA	55	150	HE64	F: TCATCATCATTTTACTGACGTTCC R: TCTAAGTCTTTGTTGATGGGG	TATG	54	151

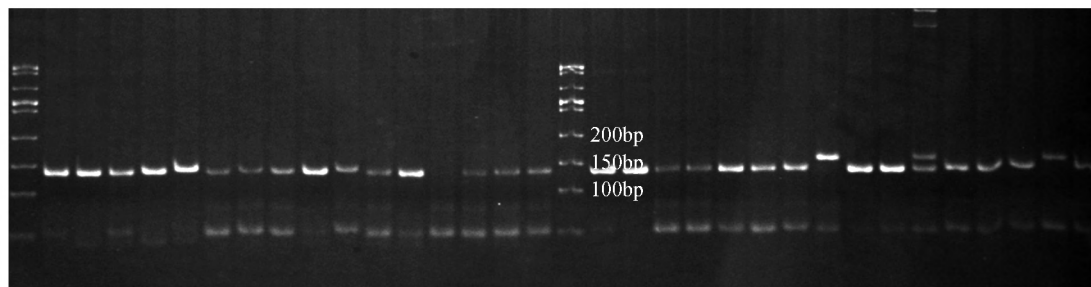


图 1 HE21 引物 PCR 产物的 EB 染色图

Fig. 1 Image of EB staining for PCR products of primers HE21

表 5 13 对多态性 SSR 位点在所研究 4 个居群中的特征
Table 5 Characteristics of 13 SSR loci in four populations studied

Locus	A	H_o	H_e	HWE	F_{is}
HE1	9	0	0.730	×	1.000
HE13	6	1	0.730	√	-0.374
HE15	7	0.344	0.821	×	0.585
HE17	10	0.563	0.864	√	0.353
HE21	3	0.031	0.347	×	0.911
HE36	4	0	0.561	×	1.000
HE40	5	0.125	0.764	×	0.839
HE42	5	0.031	0.697	×	0.956
HE44	6	0.031	0.653	×	0.953
HE50	6	0.094	0.774	×	0.881
HE56	3	0.063	0.473	×	0.870
HE63	4	0	0.617	×	1.000
HE64	3	0.656	0.458	√	-0.443

注: A. 等位基因数; H_o . 观测杂合度; H_e . 期望杂合度; HWE. 哈迪温伯格平衡; F_{is} . 近交系数

Note: A. Number of alleles per locus; H_o . Observed heterozygosity; HWE. Hardy-Weinberg's equilibrium; H_e . Expected heterozygosity; F_{is} . Inbreeding coefficient

2.4 多态性 SSR 引物的开发

筛选结果显示,筛选所用的 65 对 SSR 引物中,有 44 对引物在梯度退火温度 PCR 之后的电泳检测结果中表现出无目标条带或模糊不清。将其余的 21 对 SSR 引物用椭圆叶花锚的 4 个居群 32 个个体扩增,并用 20% 聚丙烯酰胺凝胶电泳检测,共筛选出了 13 对可有效扩增并且具有多态性的 SSR 引物,其中 HE22 对应的位点不具多态性(表 4);而另外 7 对 SSR 引物的 PCR 产物的电泳检测结果显示条带模糊或者无目标条带。部分椭圆叶花锚 SSR 引物的扩增产物的 PAGE 电泳结果如图 1。利用 Gengpop 4.4 软件检测等位基因数(A)、哈迪温伯格平衡(HWE)偏离情况、总体期望杂合度(H_e)、总体观测杂合度(H_o)、总体近交系数(F_{is})等信息(表 5)。

Gengpop 检测结果显示:13 对多态性 SSR 引物对 4 个椭圆叶花锚居群的扩增共获得 71 个等位基因;SSR 位点的等位基因数的变化范围是 3 ~ 10,均值为 5.462;期望杂合度的变化范围为 0.347 ~ 0.864,均值 0.653;观测杂合度的变化范围为 0 ~ 1,均值 0.226;仅 H13、H17 和 H64 三个位点不明显偏离哈迪温伯格平衡;近交系数 F_{is} 范围是 -0.443 ~ 1,均值 0.656;连锁不平衡分析显示 P 值最小者为 0.194;基因流 $N_m = 0.474$ 。

3 讨论

本研究利用基于二代测序的 RAD 测序技术成功地获得了椭圆叶花锚的简化基因组序列,在此基础上分析了椭圆叶花锚的 SSR 信息,并为该物种设计了 3 865 对 SSR 引物。椭圆叶花锚的 3 865 对 SSR 引物所对应的 SSR 长度以 12 bp 最丰富;数量较多的核苷酸重复类型是三核苷酸,最少的是六核苷酸;随着 SSR 基序的核苷酸数目的增多,基序的重复次数减少;重复序列长度包括 17 种(12 ~ 36 bp);这些特征与其近缘物种异型花(*Sinoswertia tetraptera*)相同^[14]。

在从 3 865 对 SSR 引物中所挑选的 65 对引物,有 14 对在椭圆叶花锚的 4 个居群中成功扩增,并且有 13 对表现出多态性。经 Gengpop 检测,13 对多态性 SSR 位点不存在显著的连锁不平衡($P < 0.05$),可认为这些位点不存在连锁不平衡现象。SSR 位点的等位基因平均数是 5.462,多态性高于 Zhang 等开发的的位点(3.9)^[10]。基因流略偏小($N_m = 0.47$)、近交系数 F_{is} 总体偏高,多数位点(10 个)偏离哈迪温伯格平衡并且观测杂合度与期望杂合度相比存在差异明显,纯合子数量较高,这些现象可能由以下几个原因造成:椭圆叶花锚的花冠筒长于花柱^[1],不易跟在风媒作用下形成异花授粉,而本研究的样品都采自青藏高原,

受高原地带的气候影响,昆虫数量和年活动时间都较少,通过虫媒授粉的机会也低于临近的低海拔地区,从而导致该物种可能多行自花授粉;在本研究中4个居群距离较远,居群间基因流受限;一般利用SSR位点进行遗传分析,并要求在居群个体在20~40间^[21],而本研究每个居群只有8个个体,样本量不够大,造成获得的数据较少。需要补充说明的是,虽然本研究中样本量偏少,但足以筛选多态性微卫星引物,满足本研究的需求。

Zhang等^[10]已经通过FIASCO法开发了12对椭圆叶花锚的多态性SSR引物,并且这些引物在备检的一个异型花个体上也表现出通用性,所以本研究在简化基因组的水平上揭示椭圆叶花锚的SSR更多的信息,也可作为该物种的近缘类群的一个值得参考的数据库。

参 考 文 献

1. Ho T N, Pringle J S. Gentianaceae [M]. // Wu Z Y, Raven P H. Flora of China, vol. 16. Beijing: Science Press, St. Louis: Missouri Botanical Garden Press, 1995: 98–99.
2. 中国科学院西北高原生物研究所. 藏药志 [M]. 西宁: 青海人民出版社, 1991: 113–114.
Northwest Plateau Institute of Biology, Chinese Academy of Sciences. Handbook of Tibetan medicine [M]. Xining: Qinghai People's Publishing Press, 1991: 113–114.
3. Huang B, Ban X Q, He J S et al. Hepatoprotective and antioxidant effects of the methanolic extract from *Halenia elliptica* [J]. Journal of Ethnopharmacology, 2010, 131(2): 276–281.
4. 古锐, 王昌华, 刘翔, 等. 藏药椭圆叶花锚的资源学研究 [J]. 时珍国医国药, 2011, 22(3): 720–722.
Gu R, Wang C H, Liu X, et al. Resources research of Tibetan medicinal *Halenia elliptica* [J]. Lishizhen Medicine and Materia Medica Research, 2011, 22(3): 720–722.
5. 张婉, 崔继哲, 于拴仓, 等. 白菜品种的SSR指纹图谱数据库的构建 [J]. 分子植物育种, 2013, 11(6): 843–857.
Zhang W, Cui J Z, Yu S C, et al. Construction of SSR fingerprint database of Chinese cabbage varieties (*Brassica campestris* L. ssp. *Pekinensis*) [J]. Molecular Plant Breeding, 2013, 11(6): 843–857.
6. Zhang Z F, Bian Q L, Luo P, et al. Ethnopharmacological, chemical and pharmacological aspects of *Halenia elliptica*: a comprehensive review [J]. Pharmacognosy Reviews, 2015, 9(18): 114–119.
7. Liu Y L, Wang P, Chen T, et al. One-Step isolation and purification of four Xanthone Glycosides from Tibetan medicinal plant *Halenia elliptica* by high-speed counter-current chromatography [J]. Separation Science and Technology, 2014, 49(7): 1119–1124.
8. Sun Y W, Liu G M, Huang H, et al. Chromone derivatives from *Halenia elliptica* and their anti-HBV activities [J]. Phytochemistry, 2012, 75: 169–176.
9. Huang B, Ke H B, He J S, et al. Extracts of *Halenia elliptica* exhibit antioxidant properties in vitro and in vivo [J]. Food and Chemical Toxicology, 2011, 49(1): 185–190.
10. Zhang Z R, Yang J, Sun Y, et al. A set of novel microsatellite markers developed for the traditional Tibetan medicinal plant *Halenia elliptica* (Gentianaceae) [J]. American Journal of Botany, 2011, 98(7): 173–175.
11. Davey J W, Blaxter M L. RADSeq: next-generation population genetics [J]. Briefings in Functional Genomics, 2010, 9(5–6): 416–423.
12. 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, 2007, 17(2): 240–248.
13. 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]. Beijing: Hereditas, 2014, 36(1): 41–49.
14. 王久利, 朱明星, 徐明行, 等. 基于RAD-seq技术的异型花SSR信息分析 [J]. 植物研究, 2017, 37(3): 447–452, 460.
Wang J L, Zhu M X, Xu M H, et al. Analysis on SSR in *Sinowertia tetraptera* base on RAD-seq [J]. Bulletin of Botanical Research, 2017, 37(3): 447–452, 460.
15. Hosseini P, Tremblay A, Matthews B F, et al. An efficient annotation and gene-expression derivation tool for Illumina Solexa datasets [J]. BMC Research Notes, 2010, 3: 183.
16. 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.
17. 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.
18. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers [J]. Bioinformatics Methods and Protocols, 2000, 132: 365–386.
19. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform [J]. Bioinformatics, 2010, 26(5): 589–595.
20. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools [J]. Bioinformatics, 2009, 25(16): 2078–2079.
21. Ohta T. Linkage disequilibrium with the island model [J]. Genetics, 1982, 101(1): 139–155.