

网络出版时间:2020-06-01 10:45 DOI:10.13207/j.cnki.jnwafu.2020.12.002
网络出版地址: <http://kns.cnki.net/kcms/detail/61.1390.s.20200529.0830.002.html>

蒙农红豆草花瓣转录组微卫星 SSR 特征分析

乔雨¹, 石凤翎¹, 张雨桐¹, 钱亚斯¹, 闫伟¹, 伊风艳²

(1 内蒙古农业大学 草原与资源环境学院, 草地资源教育部重点实验室, 内蒙古 呼和浩特 010019;
2 内蒙古自治区农牧业科学院, 中国科学院内蒙古草原研究中心, 内蒙古 呼和浩特 010031)

【摘要】【目的】对蒙农红豆草花瓣转录组微卫星 SSR 特征进行分析, 为开发新的分子标记及探明各类性状的遗传机制研究奠定基础。【方法】以蒙农红豆草突变群体中浅色花瓣和紫红色花瓣(对照)为试验材料, 通过 Illumina HiSeq™ 2000 对其进行高通量测序和组装, 利用 MISA 软件对其所有 Unigene 进行 SSR 搜索, 并对 SSR 分布及其序列特征进行分析。【结果】在蒙农红豆草转录组数据中共获得 53 009 条 Unigene, 其中 6 202 条 Unigene 上分布了 8 175 个 SSR 位点, 每 5.86 kb 就有 1 个 SSR 位点; 重复基元数量以三核苷酸为主(49.13%), 二核苷酸次之(22.07%); 在 284 种重复基元中以 AG/CT 类型最高(16.16%), 其次为 AAG/CTT(14.30%); 三核苷酸基元重复次数种类最多(11 种), 二核苷酸次之(10 种)。随着基元重复次数的增加, 核苷酸数量减少。蒙农红豆草 SSR 长度为 12~441 bp, 主要集中在 12~24 bp, 其中具有高度多态性潜能的 SSR(长度≥20 bp)有 3 130 条(38.29%)。【结论】蒙农红豆草 SSR 位点出现频率高、基元类型丰富、分布密度大、多态性潜能高, 在红豆草基因定位和分子标记辅助育种等方面具有较大的应用价值。

【关键词】 蒙农红豆草; 花瓣转录组; 微卫星 SSR

【中图分类号】 S541⁺.4 **【文献标志码】** A **【文章编号】** 1671-9387(2020)12-0009-09

SSR characteristics of transcriptome of Mengnong sainfoin

QIAO Yu¹, SHI Fengling¹, ZHANG Yutong¹, QIAN Yasi¹, YAN Wei¹, YI Fengyan²

(1 College of Grassland, Resources and Environment, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia 010019, China; 2 Inner Mongolia Academy of Agricultural & Animal Husbandry Sciences, Inner Mongolia Prataculture Research Center, Chinese Academy of Sciences, Hohhot, Inner Mongolia 010031, China)

Abstract: 【Objective】 This study analyzed SSR characteristics of petal transcriptome of Mengnong sainfoin to provide basis for development of new molecular markers and genetic mechanism of various traits. 【Method】 Light color flowers from mutant population and purplish-red flowers in the control group of Mengnong sainfoin were selected and high-throughput sequencing of petals was conducted with Illumina HiSeq™ 2000 and de novo. Then, SSR from all Unigenes were searched using software MISA and SSR characteristics were analyzed. 【Result】 A total of 53 009 Unigene were obtained from the transcriptome of Mengnong sainfoin petals, and 8 175 SSR loci were distributed on 6 202 Unigene with a distribution density of 5.86 kb. The numbers of repeat motifs were mainly tri-nucleotide (49.13%), followed by di-nucleotide (22.07%). AG/CT type was the highest (16.16%) among 284 repeat motifs, followed by AAG/CTT (14.30%). Tri-nucleotide type of motifs repeated the most, followed by di-nucleotide. The number of nu-

〔收稿日期〕 2019-11-14
〔基金项目〕 内蒙古自治区草原英才创新团队项目“内蒙古高原乡土草种质资源遗传改良与利用”; 内蒙古科技计划项目“抗逆乡土草品种培育及良繁技术研究与示范”
〔作者简介〕 乔雨(1992—), 女, 内蒙古丰镇人, 在读博士, 主要从事牧草栽培与遗传育种研究。E-mail: 649338050@qq.com
〔通信作者〕 石凤翎(1964—), 女, 内蒙古奈曼人, 教授, 博士, 博士生导师, 主要从事牧草栽培与遗传育种研究。
E-mail: sfl0000@126.com

cleotides decreased with the increase of motif repeats. The length of SSRs of Mengnong sainfoin was 12—441 bp with the most in 12—24 bp. There were 3 130 SSR showing high polymorphism potential with length greater than or equal to 20 bp,accounting for 38.29% of total SSRs.【Conclusion】The SSR locus in the petal transcriptome of Mengnong sainfoin had high occurring frequency,motif type richness,high distribution density and high polymorphism potential,and these characteristics can be applied in future studies.

Key words: Mengnong sainfoin;petals transcriptome;microsatellite SSR

红豆草(*Onobrychis viciifolia* Scop)是一种多年生豆科牧草,其茎叶中含有丰富的蛋白质、氨基酸和粗脂肪等,被誉为“牧草皇后”。研究发现,红豆草内原花青素(缩合单宁)浓度较高,这些原花青素已被证明可以在提高蛋白质利用率的情况下避免家畜患膨胀病^[1-2]。蒙农红豆草(*Onobrychis viciifolia* Scop ‘Mengnong’)是乌云飞等^[3]在 1994 年培育成功的高产、优质、抗寒性强的牧草品种。其花冠呈粉红到紫红色,可美化庭院,也可作为蜜源植物,每朵小花日泌蜜 0.1~0.9 mg,每公顷泌蜜 190~220 kg,其蔗糖含量为 0.42~0.70 g/kg^[4-6]。

微卫星 SSR 也称微卫星 DNA 或简单重复序列,是由 2~6 个核苷酸为重复单位构成的、长度为几十到几百 bp 的重复序列,其序列多态性是因同一物质不同基因型的寡聚核苷酸重复次数不同,以及重复序列在染色体复制时滑动或染色单体的不等交换造成的^[7],因此可以利用重复序列的重复次数和重复程度不同揭示同一物种不同基因型间的差异,以证明物种的多态性^[8-9]。按照来源将 SSR 分为转录组 SSR 和 EST-SSR,前者利用 cDNA 文库构建与克隆测序方法获得,后者通过在公共数据库搜索的表达序列标签和转录组测序结果开发获得^[10]。与其他分子标记相比,SSR 标记具有多态性高、重复性好、共显性遗传等优点,被广泛应用于构建高密度

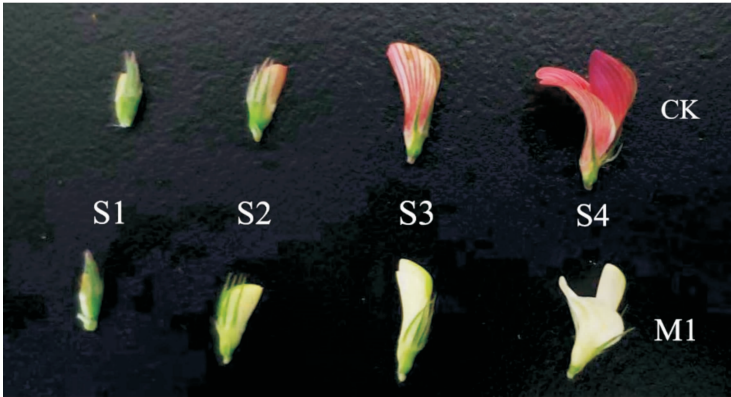
基因组图谱、标记辅助育种、建立遗传和进化关系等研究^[11]。目前,利用转录组测序技术开发 SSR 引物在褐沙蒿^[12]、鬼针草^[13]和虎耳草^[14]等牧草中均有报道,关于红豆草转录组的研究主要集中在分子标记开发^[15]和单宁生物合成途径相关基因的筛选^[16]上。但尚未见关于利用转录组测序技术在红豆草上开发 SSR 引物的研究报道。

本研究通过对蒙农红豆草突变群体中花色变浅的花瓣和紫红色花瓣(对照)的高通量测序,建立转录组数据库,并对 SSR 分布及其序列特征进行分析,以期为红豆草遗传多样性分析、遗传图谱构建和分子辅助育种等提供更多分子标记。

1 材料与方法

1.1 试验材料

根据蒙农红豆草花瓣生长状态,将其划分为 4 个不同发育时期(图 1),分别为 S1(花瓣尚未露出花萼)、S2(花瓣高于花萼)、S3(花瓣包裹在一起但未开放)和 S4(花朵盛开)。本研究以花色变浅突变体 M1 代和紫红色植株(对照,CK)的 S1、S3 时期花瓣为试验材料(图 1)。为了更好地体现整朵花的花色,将同一时期同一种花色小花的旗瓣和翼瓣(因龙骨瓣较小而舍弃,各花瓣颜色一致)混合,然后放入液氮速冻,并保存在-80℃冰箱,用于后续转录组测序。



S1—S4. 不同发育时期;M1. 花色变浅突变体;CK. 紫红色花瓣
S1—S4. Different stages;M1. Light color flower from mutant population;CK. Purplish-red

图 1 供试蒙农红豆草花瓣材料
Fig. 1 Tested Mengnong sainfoin petals

1.2 试验方法

1.2.1 文库构建与测序 提取蒙农红豆草花瓣总 RNA,检测合格后,用带有 Oligo(dT)的磁珠富集 mRNA,向得到的 mRNA 中加入 fragmentation buffer 使其成为短片段,再以短片段的 mRNA 为模板,用六碱基随机引物(random hexamers)合成 cDNA 第一链,并加入缓冲液、dNTPs、RNase H 和 DNA polymerase I 合成 cDNA 第二链,用 Qia-Quick PCR 试剂盒纯化并加 EB 缓冲液洗脱,经末端修复、加碱基 A 和测序接头,再经琼脂糖凝胶电泳回收目的大小片段,进行 PCR 扩增,完成整个文库制备,构建好的文库用 Illumina HiSeq™ 2000 进行测序。

1.2.2 质量评估与拼接 对从测序仪获得的基因片段,去除含接头的序列(reads)、无法确定碱基信息(N)比例大于 10%的序列,以及低质量序列(质量值 Q≤20 的碱基数占 40%以上的序列)的基因片段后,获得高质量的过滤(clean)基因片段。使用短序

列组装软件 Trinity 5 进行组装得到蒙农红豆草花瓣 Unigene,并进行后续分析。

1.2.3 SSR 的筛选和统计分析 使用软件 MISA (<http://pgrc.ipk-gatersleben.de/misa/>)对组装的所有 Unigene 进行搜索,寻找 Unigene 中的 SSR。用 Microsoft Excel 2010 对数据进行整理和绘图,并用 SAS 进行相关性分析。

2 结果与分析

2.1 蒙农红豆草花瓣转录组中 SSR 分析

由表 1 可知,对蒙农红豆草花瓣进行高通量测序后得到 53 009 条 Unigene,平均长度为 904 bp,总长度为 47 910 205 bp,其中包括 8 175 个 SSR 位点,分布在 6 202 条 Unigene 上。SSR 的出现频率和发生频率分别为 15.42%和 11.70%。SSR 的分布密度为 170.63 个/Mb。此外,含复合型 SSR 的序列数有 755 条,含 1 个以上 SSR 位点的序列有 1 373 条,占 Unigene 序列总数的 2.59%。

表 1 蒙农红豆草花瓣转录组中 SSR 搜索分析结果

Table 1 Searching and analysis of SSRs in transcriptome of Mengnong sainfoin petals			
指标 Index	测定值 Value	指标 Index	测定值 Value
Unigene 序列总数 Total number of sequences examined	53 009	SSR 分布密度/(个·Mb ⁻¹) Density	170.63
Unigene 序列总长度/bp Total size of examined sequences	47 910 205	含 SSR 位点的序列数目 Number of SSR containing sequences	6 202
SSR 位点总数 Total number of identified SSRs	8 175	含 1 个以上 SSR 位点的序列数目 Number of sequences containing more than 1 SSR	1 373
SSR 发生频率/% Frequency of SSR incidence	11.70	含复合型 SSR 的序列数目 Number of SSRs present in compound formation	755
SSR 出现频率/% Frequency of SSR occurrence	15.42		

2.2 蒙农红豆草花瓣转录组 SSR 重复类型

由表 2 可知,从蒙农红豆草花瓣转录组获得的 SSR 重复类型中,所占比例和出现频率最高的是三核苷酸,分别为 49.13%和 7.58%;其次为二核苷酸;最少的是六核苷酸,所占比例和出现频率分别为

4.99%和 0.77%;三核苷酸的 SSR 数量约是六核苷酸 SSR 的 10 倍。从分布情况来看,每 11.93 kb 出现 1 个三核苷酸 SSR 位点,每 117.43 kb 出现 1 个六核苷酸位点,平均每 5.86 kb 就出现 1 个 SSR 位点。

表 2 蒙农红豆草花瓣转录组 SSR 重复类型的数量和分布情况

Table 2 Number and distribution characteristics of SSRs repeating types in transcriptome of Mengnong sainfoin petals				
重复类型 Repeat type	SSR 数量 Number	所占比例/% Rate	出现频率/% Frequency	平均距离/kb Mean distance
二核苷酸 Di-nucleotide	1 804	22.07	3.40	26.56
三核苷酸 Tri-nucleotide	4 016	49.13	7.58	11.93
四核苷酸 Tetra-nucleotide	1 332	16.29	2.51	35.97
五核苷酸 Penta-nucleotide	615	7.52	1.16	77.90
六核苷酸 Hexa-nucleotide	408	4.99	0.77	117.43
总计 Total	8 175	100.00	15.42	—

2.3 蒙农红豆草花瓣转录组 SSR 重复基元类型和比例

在蒙农红豆草花瓣转录组中检测到由 284 种重

复基元构成的 6 202 条 SSR,其中二、三、四、五、六核苷酸重复基元种类分别为 4,10,32,82,156,其总频率为 15.42%,约 15 条 Unigene 出现 1 个 SSR 位

点。蒙农红豆草花瓣转录组中不同 SSR 重复基元类型的数量和分布情况见表 3,因篇幅所限,每种核苷酸重复基元最多只列出前 10 个。由表 3 可知,二核苷酸中的 AG/CT 在所有重复基元中数量最多,为 1 321 个,占二核苷酸位点的 73.22%,占总位点数的 16.16%;其次为 AC/GT (4.37%)、AT/AT (1.48%)和 CG/CG(0.06%)。三核苷酸中的主要重复基元为 AAG/CTT,占三核苷酸总数的 14.30%,其余重复基元数量为 ATC/ATG(6.56%)>AAC/GTT(6.04%)>AGG/CCT(5.82%)。六核苷酸中 AAATTG/AAT TTC 为优势重复基元,

占总位点数的 0.16%,其次为 AAGCAG/CTGCTT 和 AAAGAG/CTCTTT,均占总位点数的 0.12%。
二、三、四、五、六核苷酸的分布频率为 0.01%~2.49%,其中 AG/CT 的分布频率最高,为 2.49%;AAG/CTT 的分布频率居次,为 2.21%。而植物转录组中出现频率较低的 CG/CG 重复基元占 SSR 总数的 0.06%,以及在单子叶植物中出现频率较高而在双子叶植物中出现频率较低的 CCG/CGG 重复基元占 SSR 总数的 0.48%。总体而言,六核苷酸重复基元种类最多,三核苷酸重复基元数量最多。

表 3 蒙农红豆草花瓣转录组中不同 SSR 重复基元类型的数量和分布情况

Table 3 Number and distribution of different SSRs repeat types in transcriptome of Mengnong sainfoin petals

重复类型 Repeat type	重复基元 Motif	数量 Number	比例/% Rate	频率/% Frequency	重复类型 Repeat type	重复基元 Motif	数量 Number	比例/% Rate	频率/% Frequency
二核苷酸 Di-nucleotide	AG/CT	1 321	16.16	2.49	五核苷酸 Penta-nucleotide	AAGAG/CTCTT	43	0.53	0.08
	AC/GT	357	4.37	0.67		AAAAG/CTTTT	29	0.35	0.05
	AT/AT	121	1.48	0.23		AAATC/ATTTG	26	0.32	0.05
	CG/CG	5	0.06	0.01		AACAG/CTGTT	23	0.28	0.04
三核苷酸 Tri-nucleotide	AAG/CTT	1 169	14.30	2.21		AAAAC/GTTTT	23	0.28	0.04
	ATC/ATG	536	6.56	1.01		AATTC/AATTG	23	0.28	0.04
	AAC/GTT	494	6.04	0.93		AACAC/GTGTT	22	0.27	0.04
	AGG/CCT	476	5.82	0.90		AAAGC/CTTTG	21	0.26	0.04
	ACC/GGT	369	4.51	0.70		AACTC/AGTTG	20	0.24	0.04
	ACT/AGT	350	4.28	0.66		AGAGG/CCTCT	16	0.20	0.03
	AGC/CTG	354	4.33	0.67	六核苷酸 Hexa-nucleotide	AAATTG/AAT TTC	13	0.16	0.02
	AAT/ATT	172	2.10	0.32		AAGCAG/CTGCTT	10	0.12	0.02
	ACG/CGT	57	0.70	0.11		AAAGAG/CTCTTT	10	0.12	0.02
	CCG/CGG	39	0.48	0.07		AATTCC/AATTGG	8	0.10	0.02
四核苷酸 Tetra-nucleotide	AAAG/CTTT	236	2.89	0.45		AAGAGG/CCTCTT	7	0.09	0.01
	ACTC/AGTG	127	1.55	0.24		AATGGC/ATTGCC	7	0.09	0.01
	AGAT/ATCT	98	1.20	0.18		AAACTC/AGTTTG	7	0.09	0.01
	AAAC/GTTT	93	1.14	0.18		AAACTG/AGTTTC	7	0.09	0.01
	AAGG/CCTT	82	1.00	0.15		AACTCC/AGTTGG	6	0.07	0.01
	AATG/ATTC	74	0.91	0.14		AAGGAG/CCTTCT	6	0.07	0.01
	ACAT/ATGT	68	0.83	0.13	总计 Total		8 175	100.00	15.42
	AGCT/AGCT	66	0.81	0.12					
	AATC/ATTG	53	0.65	0.10					
	AACT/AGTT	44	0.54	0.08					

2.4 蒙农红豆草花瓣转录组 SSR 重复次数和分布

由图 2 可知,对蒙农红豆草 8 175 个 SSR 位点的分析发现,SSR 重复次数主要集中在 4~7 次,共 7 353 条,占 SSR 总数的 89.94%,是蒙农红豆草 SSR 重复次数的主要组成部分;重复次数在 8~10 的核苷酸数量共 720 条,占总 SSR 的 8.81%;重复次数≥11 核苷酸数量的共 102 条,占总 SSR 的 1.25%。总体而言,随着基元重复次数的增加,核苷酸的数量在减少。

由表 4 可知,三核苷酸基元重复次数的种类最

多(11 种),主要集中在 5~7 次,其核苷酸数量为 3 886 条,占三核苷酸 SSR 总数的 96.76%,占总 SSR 数量的 47.54%;其次是重复次数种类为 10 的二核苷酸,其核苷酸重复次数主要集中在 6~10 次,数量为 1 732 条,占二核苷酸 SSR 总数的 96.01%,占总 SSR 数量的 21.19%。四、五、六核苷酸基元重复次数的种类分别为 8,6,8 种,其中 4~5 次重复基元为重要组成部分,其核苷酸数量分别为 1 256,578 和 381 条,分别占各自核苷酸 SSR 总数的 94.29%,93.98%和 93.38%,占总 SSR 的 15.36%,7.07%,

4.66%。

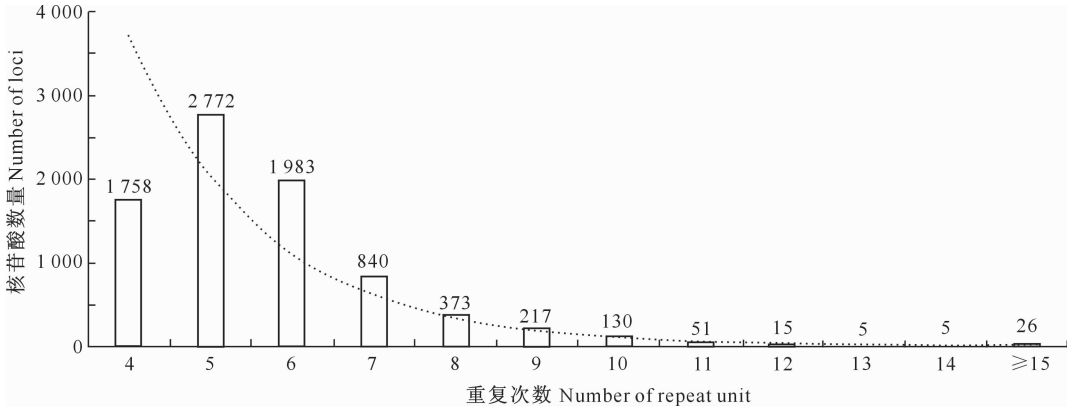


图 2 蒙农红豆草花瓣转录组 SSR 重复次数的分布
Fig. 2 Distribution of SSRs repeating types in transcriptome of Mengnong sainfoin petals

表 4 蒙农红豆草花瓣转录组 SSR 不同基元的重复次数及分布

Table 4 Repeat times and distribution of different motifs of SSRs in petal transcriptome of Mengnong saifoin																		
重复类型 Repeat type	重复次数 Number of repeat unit													总计 Total number	种类 Types	主要重复基元 Primary motif repetition		占总 SSR 比例/% Percentage of total SSR
																数目 Number	比例/% Proportion	
	4	5	6	7	8	9	10	11	12	13	14	≥15						
二核苷酸 Di-nucleotide	0	0	742	396	288	206	100	45	9	1	1	16	1 804	10	1 732	96.01	21.19	
三核苷酸 Tri-nucleotide	0	2 315	1 151	420	70	7	30	3	5	2	3	10	4 016	11	3 886	96.76	47.54	
四核苷酸 Tetra-nucleotide	965	291	49	15	7	2	0	0	0	2	1	0	1 332	8	1 256	94.29	15.36	
五核苷酸 Penta-nucleotide	492	86	30	2	4	1	0	0	0	0	0	0	615	6	578	93.98	7.07	
六核苷酸 Hexa-nucleotide	301	80	11	7	4	1	0	3	1	0	0	0	408	8	381	93.38	4.66	

2.5 蒙农红豆草花瓣转录组 SSR 长度分布

由图 3 可知,蒙农红豆草 SSR 长度为 12~441 bp,其中长度为 233,296,441 bp 的 SSR 为复合型,数量分别为 3,4,7。整体而言,蒙农红豆草花瓣转录组的 SSR 长度主要集中在 12~24 bp,其数量为

6 495 条,占总 SSR 的 79.45%。在所有 SSR 中数量最多的是长 15 bp 的 SSR (共 1 935 条,占比 23.67%),且都是基元重复次数为 5 的三核苷酸。据统计,具有高度多态性潜能 SSR (长度≥20 bp) 有 3 130 条,占 38.29%。

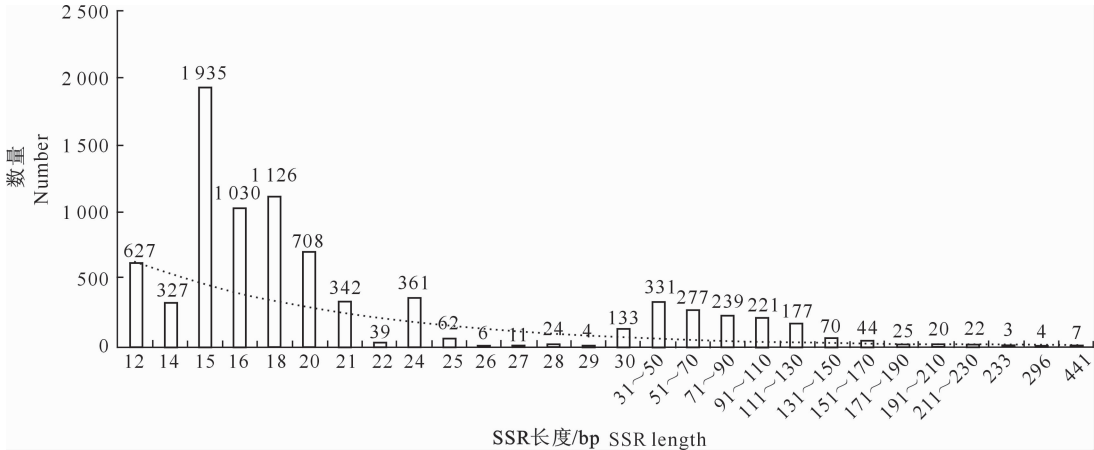


图 3 蒙农红豆草花瓣转录组 SSR 长度分布
Fig. 3 Length distribution of SSRs in transcriptome of Mengnong sainfoin petals

由表 5 可知,二核苷酸 SSR 长度主要集中在 12~24 bp,其 SSR 数量为 1 489 条,占二核苷酸 SSR 总数的 82.54%;其中长度为 12 bp 的 SSR 最多(627 个),占二核苷酸 SSR 总数的 34.76%,约是长度为 14 bp SSR 的 2 倍。三核苷酸中 SSR 长度为 15 bp 的核苷酸最多,为 1 935 个,占三核苷酸

SSR 总数的 48.18%;三核苷酸的长度主要集中在 15~24 bp,其数量为 3 300 条,占三核苷酸 SSR 总数的 82.17%。四、五、六核苷酸 SSR 长度分别主要集中在 16~20,20~25,24~30 bp,其数量分别为 1 023,450,330 条,分别占各自核苷酸 SSR 总数的 76.80%,73.17%,80.88%。

表 5 蒙农红豆草花瓣转录组不同重复类型的 SSR 长度分布

Table 5 Length distribution of SSRs different repeating types in transcriptome of Mengnong sainfoin petals							
重复类型 Repeat type	长度 Length	SSR 数量 SSR number	比例/% Rate	重复类型 Repeat type	长度 Length	SSR 数量 SSR number	比例/% Rate
二核苷酸 Di-nucleotide	12	627	34.76	四核苷酸 Tetra-nucleotide	16	795	59.68
	14	326	18.07		20	228	17.12
	16	235	13.03		24	33	2.48
	18	161	8.92		25	1	0.08
	20	90	4.99		27	2	0.15
	22	39	2.16		28	14	1.05
	24	11	0.61		29	2	0.15
	25	1	0.06		31~50	50	3.75
	26	6	0.33		51~70	65	4.88
	27	3	0.17		71~90	42	3.15
	28	9	0.50		91~110	38	2.85
	29	1	0.06		111~130	26	1.95
	30	6	0.33		131~150	12	0.90
	31~50	71	3.94		151~170	9	0.68
	51~70	44	2.44		171~190	6	0.45
	71~90	53	2.94		191~210	2	0.15
	91~110	48	2.66		211~230	2	0.15
	111~130	35	1.94		233	1	0.08
	131~150	14	0.78		441	4	0.30
	151~170	8	0.44	五核苷酸 Penta-nucleotide	20	390	63.41
	171~190	5	0.28		25	60	9.76
	191~210	3	0.17		30	24	3.90
	211~230	7	0.39		31~50	18	2.93
	441	1	0.06		51~70	23	3.74
					71~90	28	4.55
					91~110	26	4.23
					111~130	21	3.41
三核苷酸 Tri-nucleotide	15	1 935	48.18		131~150	9	1.46
	18	966	24.05		151~170	5	0.81
	21	342	8.52		171~190	2	0.33
	24	57	1.42		191~210	5	0.81
	27	6	0.15		211~230	4	0.65
	28	1	0.02	六核苷酸 Hexa-nucleotide	24	260	63.73
	29	1	0.02		30	70	17.16
	30	33	0.82		31~50	19	4.66
	31~50	173	4.31		51~70	17	4.17
	51~70	128	3.19		71~90	14	3.43
	71~90	102	2.54		91~110	13	3.19
	91~110	96	2.39		111~130	2	0.49
	111~130	93	2.32		131~150	7	1.72
	131~150	28	0.70		151~170	4	0.98
	151~170	18	0.45		171~190	1	0.25
	171~190	11	0.27		191~210	1	0.25
	191~210	9	0.22				
	211~230	9	0.22				
	233	2	0.05				
	296	4	0.10				
	441	2	0.05				

3 讨 论

目前转录组研究广泛用于物种新基因发现、基因功能注释以及分子标记。而物种的微卫星序列 (SSR) 对其基因功能有很大影响,研究某物种的 SSR 特征是了解其基因组差异的重要手段^[17]。本研究通过对蒙农红豆草花瓣的高通量测序,得到 53 009 条、长度为 47 910 205 bp 的 Unigene,得到 SSR 位点 8 175 个,含 SSR 位点的序列 6 202 条。蒙农红豆草 SSR 平均分布距离为 5.86 kb,其高于换锦花 (3.73 kb)^[18]、白沙蒿 (4.38 kb) 和乌丹蒿 (4.37 kb)^[19],低于蒙古黄芪 (6.51 kb)^[20]、山地虎耳草 (10 kb) 和帮腺虎耳草 (10.4 kb)^[14]。蒙农红豆草花瓣转录组 SSR 出现频率为 15.42%,高于褐沙蒿 (13.02%)^[12]、鬼针草 (12.04%)^[13]、帮腺虎耳草 (7.45%)^[14],低于换锦花 (18.59%)^[18] 和蒙古黄芪 (31.26%)^[20]。上述差异可能与物种、SSR 搜索标准和数据库大小等因素有关。

研究发现,SSR 重复基元长度的大小与物种的选择压力密切相关,随着长度的增加其选择压力也会变大,重复次数变少,所以某物种的基因组中 SSR 长度越短,其变异速率也越快,相反较长的 SSR 变异速率相对较慢^[21]。因此,基因序列中的 SSR 增加了基因的不稳定性,但同时也加大了其基因变异的可能性。大多数植物 SSR 以二核苷酸和三核苷酸重复基元为主,但是主导重复基元的类型有所不同。本试验中,蒙农红豆草花瓣转录组 SSR 的重复基元类型主要是三核苷酸 (49.13%),其次为二核苷酸 (22.07%),这与褐沙蒿^[12]、鬼针草^[13] 和虎耳草^[14] 等物种的研究结果一致。但川西獐牙菜^[22]、蒙药冷蒿^[23] 以二核苷酸和三核苷酸为主,而芝麻^[24]、寒兰^[25] 以二核苷酸为主,这种优势重复基元的差异可能与物种不同有关。

本研究发现,蒙农红豆草花瓣转录组中二核苷酸以 AG/CT 为主,三核苷酸以 AAG/CTT 为主,与鬼针草 (AG/CT、AAT/ATT)^[13]、乌丹蒿 (AC/GT、ATC/ATG)^[19]、蒙古黄芪 (AG/CT、AAT/CTT)^[20] 中二、三核苷酸的优势基元不同,这可能与不同物种中相应编码蛋白的使用频率存在差异相关^[26]。本研究中,蒙农红豆草花瓣转录组中二核苷酸 CG/CG 重复基元的频率很低,仅为 0.01%,这与蒙药冷蒿^[23] 和白沙蒿^[19] 等物种的研究结果一致。导致 CG/CG 频率低的原因可能是由于该物种基因组 DNA 中的 CpG 岛发生甲基化,甲基化后的胞苷

酸 C 容易通过脱氨基转化为胸腺嘧啶 T,同时该 CpG 的甲基化成为该物种基因突变的热点,但是一定含量的 GC 是维持物种 DNA 热力学稳定性所必需的^[27]。

SSR 位点的多态性主要是指其序列长度的多态性,而碱基数和基元重复次数不同是造成 SSR 长度多态性的主要原因^[28]。有研究发现,物种 SSR 长度的长短与多态性有关,当 SSR 长度 ≥ 20 bp 时其多态性较高,当 SSR 长度在 12~20 bp 时多态性中等,当 SSR 长度 < 12 bp 时多态性较低^[29]。因此本研究将蒙农红豆草花瓣转录组 SSR 长度小于 12 bp 的序列提前进行了过滤,因为这些 SSR 序列多态潜能很低。过滤后发现,蒙农红豆草花瓣转录组 SSR 长度集中在 12~230 bp,长度 233,296,441 bp 的 SSR 数量分别为 3,4,7,具有高度多态性的 SSR (≥ 20 bp) 占 38.29%,这些 SSR 在蒙农红豆草上可能具有较高的应用价值。

蒙农红豆草花瓣微卫星序列特征的分析结果表明,挖掘的 8 175 个 SSR 位点出现频率较高、分布密度大、基元重复次数和种类较为丰富,并具有较高的多态性。因转录组测序获得的红豆草微卫星序列较为完整,由此开发的 SSR 标记对于红豆草及其近源种的研究将有一定的参考价值。

[参考文献]

- [1] Tanner G, Moate P, Davis L, et al. Proanthocyanidins (condensed tannin) destabilise plant protein foams in a dose dependent manner [J]. Australian Journal of Agricultural Research, 1995, 46(6): 1101-1109.
- [2] Escaray F J, Passeri V, Babuin F M, et al. *Lotus tenuis* L. corniculatus inter specific hybridization as a means to breed bloat-safe pastures and gain insight into the genetic control of proanthocyanidin biosynthesis in legumes [J]. BMC Plant Biology, 2014, 14(1): 40.
- [3] 乌云飞, 石凤翎, 崔志明. 蒙农红豆草的选育及其特征特性的研究 [J]. 中国草地, 1995(4): 24-28.
Wu Y F, Shi F L, Cui Z M. Breeding of *Onobrychis viciifolia* CV. 'Mengnong' and studies on its characteristic [J]. Grassland of China, 1995(4): 24-28.
- [4] 陈光耀. 红豆草: 理想的蜜源植物 [J]. 中国草业科学, 1987(2): 42-43.
Chen G Y. Sainfoin: ideal honey plant [J]. Pratacultural Science of China, 1987(2): 42-43.
- [5] 陈宝书. 红豆草 [M]. 兰州: 甘肃科学技术出版社, 1992.
Chen B S. Sainfoin [M]. Lanzhou: Gansu Science & Technology Press, 1992.
- [6] Kells A. Sainfoin: an alternative forage crop for bees [J]. Bee World, 2001, 82(4): 3.

- [7] 袁媛,黄璐琦.中药资源转录组分析操作指南[M].上海:上海科学技术出版社,2016.
Yuan Y, Huang L Q. Guide to the analysis and operation of the transcriptional group of traditional Chinese medicine [M]. Shanghai: Shanghai Scientific & Technical Publishers, 2016.
- [8] Litt M, Luty J A. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene [J]. The American Journal of Human Genetics, 1989, 44(3): 397-401.
- [9] Tautz D, Trick M, Dover G A. Cryptic simplicity in DNA is a major source of genetic variation [J]. Nature, 1986, 322(6080): 652-656.
- [10] 陈怀琼,隋春,魏建和.植物 SSR 引物开发策略简述[J].分子植物育种, 2009, 7(4): 845-851.
Chen H Q, Sui C, Wei J H. Summary of strategies for developing SSR primer [J]. Molecular Plant Breeding, 2009, 7(4): 845-851.
- [11] 云娜. EMS 对蒙农红豆草的诱导效应研究[D]. 呼和浩特: 内蒙古农业大学, 2014.
Yun N. The research on EMS mutagenic effects of Mengnong sainfoin [D]. Hohhot: Inner Mongolia Agricultural University, 2014.
- [12] 杜永婷. 基于转录组的褐沙蒿 EST-SSR 开发及遗传多样性研究[D]. 呼和浩特: 内蒙古师范大学, 2019.
Du Y T. Study on the EST-SSR development and genetic diversity of *Artemisia intramongolica* H. C. Fu based on transcriptome [D]. Hohhot: Inner Mongolia Normal University, 2019.
- [13] 张璐. 镉胁迫下鬼针草的转录组分析及 SSR 分子标记开发[D]. 沈阳: 沈阳师范大学, 2019.
Zhang L. Transcriptome analysis and development of SSR molecular markers for *Bidens bipinnata* L. under cadmium stress [D]. Shenyang: Shenyang Normal University, 2019.
- [14] 李彦, 焦秀洁, 更吉卓玛, 等. 山地虎耳草和棒腺虎耳草转录组 SSR 和 SNP 分析[J]. 西北植物学报, 2018, 38(7): 1244-1253.
Li Y, Jiao X J, Gengji Z M, et al. Analysis of SSR and SNP in transcriptome of *Saxifraga sinomontana* and *Saxifraga consanguinea* [J]. Acta Botanica Boreali-Occidentalia Sinica, 2018, 38(7): 1244-1253.
- [15] 申书恒. 红豆草 EST-SSR 分子标记开发与遗传多样性分析[D]. 兰州: 兰州大学, 2019.
Shen S H. Development of EST-SSR markers and analyses of genetic diversity in sainfoin (*Onobrychis viciifolia*) [D]. Lanzhou: Lanzhou University, 2019.
- [16] Moraortiz M, Swain M T, Vickers M J, et al. De-novo transcriptome assembly for gene identification, analysis, annotation, and molecular marker discovery in *Onobrychis viciifolia* [J]. BMC Genomics, 2016, 17(1): 1-13.
- [17] 祁云霞, 刘永斌, 荣威恒. 转录组研究新技术: RNA-Seq 及其应用[J]. 遗传, 2011(11): 42-53.
Qi Y X, Liu Y B, Rong W H. RNA-Seq and its applications; a new technology for transcriptomics [J]. Hereditas, 2011(11): 42-53.
- [18] 石艳, 童再康, 高燕会. 换锦花 EST-SSR 标记开发及遗传多样性分析[J]. 核农学报, 2018, 32(6): 1089-1096.
Shi Y, Tong Z K, Gao Y H. Development of EST-SSR markers and genetic diversity analysis in *Lycoris sprengeri* [J]. Journal of Nuclear Agricultural Sciences, 2018, 32(6): 1089-1096.
- [19] 阿娜尔. 两种蒿属沙生半灌木的 SSR 开发及遗传分化的研究[D]. 呼和浩特: 内蒙古师范大学, 2018.
A N E. Study on the SSR development and genetic differentiation of two kinds of sandy semi-shrub of *Artemisia* L. [D]. Hohhot: Inner Mongolia Normal University, 2018.
- [20] 贺润丽, 韩毅丽, 王芳, 等. 蒙古黄芪转录组 SSR 标记开发及遗传多样性研究[J]. 中国中药杂志, 2018, 43(9): 1838-1843.
He R L, Han Y L, Wang F, et al. Analysis on SSR information in transcriptome of *Astragalus membranaceus* var. *mongolicus* and its polymorphism [J]. China Journal of Chinese Materia Medica, 2018, 43(9): 1838-1843.
- [21] Samadi S, Artiguelbelle E, Estoup A, et al. Density and variability of dinucleotide microsatellites in the parthenogenetic polyploid snail *Melanoides tuberculata* [J]. Molecular Ecology, 1998, 7(9): 1233-1236.
- [22] 刘越, 岳春江, 王翊, 等. 藏茵陈川西獐牙菜转录组 SSR 信息分析[J]. 中国中药杂志, 2015, 40(11): 2068-2076.
Liu Y, Yue C J, Wang Y, et al. Data mining of simple sequence repeats in transcriptome sequences of Tibetan medicinal plant *Zangyinchen Swertia musotii* [J]. China Journal of Chinese Materia Medica, 2015, 40(11): 2068-2076.
- [23] 岳春江, 陈川川, 郭凤仙, 等. 蒙药冷蒿转录组 SSR 信息分析[J]. 中国农业科技导报, 2016, 18(6): 31-43.
Yue C J, Chen C C, Guo F X, et al. Data mining of simple sequence repeats in transcriptome sequences of Mongolia medicinal plant *Artemisia frigida* Willd [J]. Journal of Agricultural Science and Technology, 2016, 18(6): 31-43.
- [24] Lv H X, Li D H, Hua W, et al. Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers [J]. BMC Genomics, 2011, 12(1): 451.
- [25] 张亚楠, 杨柏云, 熊冬金, 等. 寒兰转录组 SSR 信息分析及其分子标记开发[J]. 南昌大学学报(理科版), 2017, 41(3): 249-254.
Zhang Y N, Yang B Y, Xiong D J, et al. Analysis on SSR information based on transcriptome and development of molecular markers in *Cymbidium kanran* [J]. Journal of Nanchang University (Natural Science), 2017, 41(3): 249-254.
- [26] 袁阳阳, 王青锋, 陈进明. 基于转录组测序信息的水生植物荇菜 SSR 标记开发[J]. 植物科学学报, 2013, 31(5): 485-492.
Yuan Y Y, Wang Q F, Chen J M. Development of SSR markers in aquatic plant *Nymphoides peltata* (Menyanthaceae) based on information from transcriptome sequencing [J].

- Plant Science Journal, 2013, 31(5): 485-492.
- [27] Schorderet D F, Gartler S M. Analysis of CpG suppression in methylated and nonmethylated species [J]. Proceedings of the National Academy of Sciences, 1992, 89(3): 957-961.
- [28] Temnykh S. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential [J]. Genome Research, 2001, 11(8): 1441-1452.
- [29] 杨 华, 陈 琪, 韦朝领, 等. 茶树转录组中 SSR 位点的信息分析 [J]. 安徽农业大学学报, 2011, 38(6): 882-886.
- Yang H, Chen Q, Wei C L, et al. Analysis on SSR information in *Camellia sinensis* transcriptome [J]. Journal of Anhui Agricultural University, 2011, 38(6): 882-886.
-
- (上接第 8 页)
- [16] 甄秀梅, 孙义民, 乔 杰, 等. 中国卵巢早衰妇女全基因组染色体拷贝数变异分析 [J]. 北京大学学报(医学版), 2013, 45(6): 841-847.
- Zhen X M, Sun Y M, Qiao J, et al. Analysis of whole genome chromosome copy number variation in Chinese women with premature ovarian failure [J]. Journal of Peking University (Medical Sciences), 2013, 45(6): 841-847.
- [17] McGuire M M, Bowden W, Engel N J, et al. Genomic analysis using high-resolution single-nucleotide polymorphism arrays reveals novel microdeletions associated with premature ovarian failure [J]. Fertility & Sterility, 2011, 95(5): 1595-1600.
- [18] 张 炜, 张思仲, 阿周存. 联合复合体: 原发无精症发病中的重要角色 [J]. 遗传, 2006, 28(2): 107-111.
- Zhang W, Zhang S Z, A Z C. Association complex: an important role in the pathogenesis of primary azoospermia [J]. Genetics, 2006, 28(2): 107-111.
- [19] Maor-Sagie E, Cinnamon Y, Yaacov B, et al. Deleterious mutation in SYCE1 is associated with non-obstructive azoospermia [J]. Journal of Assisted Reproduction & Genetics, 2015, 32(6): 1-5.
- [20] Geisinger A, Benavente R. Mutations in genes coding for synaptonemal complex proteins and their impact on human fertility [J]. Cytogenetic & Genome Research, 2016, 150(2): 77.
- [21] 罗荣松, 冯 郁, 宋明民, 等. 小鼠 Nlrp 基因家族生物信息学分析 [J]. 基因组学与应用生物学, 2018, 37(3): 1168-1176.
- Luo R S, Feng Y, Song M M, et al. Bioinformatics analysis of mouse Nlrp gene family [J]. Genomics and Applied Biology, 2018, 37(3): 1168-1176.
- [22] Harrison P M, Khachane A, Kumar M. Genomic assessment of the evolution of the prion protein gene family in vertebrates [J]. Genomics, 2010, 95(5): 268-277.
- [23] 陈 骞. 促性腺激素释放激素拮抗剂在下丘脑垂体外生殖系统中的作用及临床应用 [J]. 生殖医学杂志, 2016, 25(4): 384-388.
- Chen Q. Role and clinical application of gonadotropin releasing hormone antagonist in the inferior hypothalamic reproductive system [J]. Journal of Reproductive Medicine, 2016, 25(4): 384-388.
- [24] 张永莲. 附睾功能基因组研究进展 [J]. 中国科学院院刊, 2002, 17(1): 34-36.
- Zhang Y L. Advances in epididymal functional genomics [J]. Chinese Journal of Academy of Sciences, 2002, 17(1): 34-36.
- [25] Min-Sik K, Pinto S M, Derese G, et al. A draft map of the human proteome [J]. Nature, 2014, 509(7502): 575-581.
- [26] 阮崇美, 王相生, 张 勇, 等. 天祝白牦牛 HSP60 基因的克隆及其在下丘脑-垂体-睾丸中的表达定位研究 [J]. 核农学报, 2017, 31(7): 1290-1298.
- Ruan C M, Wang X S, Zhang Y, et al. Cloning of HSP60 gene from tianzhu white yak and its expression in the hypothalamus-pituitary-testis [J]. Chinese Journal of Nuclear Agriculture, 2017, 31(7): 1290-1298.