

研究报告

Research Report

观赏樱属植物基因组 SSR 分子标记开发及鉴定

李水根¹ 关媛² 李秀芬¹ 朱继军³ 李记开^{1*} 朱建军^{1*}

1 上海市农业科学院林木果树研究所, 上海, 201403; 2 上海市农业科学院作物育种栽培研究所, 上海, 201403; 3 上海植物园, 上海, 200231

* 通信作者, jkli@sibs.ac.cn; zhujianjun@saas.sh.cn

摘要 为了开发樱属植物的基因组 SSR (简单重复序列) 分子标记, 本研究在 IlluminaNovaSeq PE 平台上进行喜马拉雅樱花(*Cerasus cerasoides*) 基因组序列勘测。利用 Perl 脚本 MISA 识别微卫星序列, 并随机合成 100 对引物进行引物多态性筛选。选择易扩增、多态性高的引物, 采用荧光标记毛细管电泳法在 45 种樱属植物中进行检测和遗传多样性分析。结果表明, 喜马拉雅樱花基因组共存在 113 571 个 SSR 位点(≥ 5 个重复), 其中包含 53.31% 单核苷酸基序, 36.64% 二核苷酸基序, 其余基序类型占比 10%。在此基础上共筛选出 21 个 SSR 标记, 并将其用于观赏樱属植物的遗传多态性评估, 结果共检测到 215 个等位基因, 平均每个位点 10.24 个等位基因, 平均杂合度为 0.38。采用 SSR 标记构建的进化树较好地反应了 45 种供试樱属植物之间的亲缘关系。本研究开发了全新的樱花基因组 SSR 分子标记, 可用于樱属植物的遗传多样性分析和分子标记辅助育种。

关键词 喜马拉雅樱花; 基因组 SSR; 微卫星; 分子标记; 遗传多样性

Development and Characterization of Genome-SSR Markers in Flowering Cherries

Li Shuigen¹ Guan Yuan² Li Xiufen¹ Zhu Jijun³ Li Jikai^{1*} Zhu Jianjun^{1*}

1 Forest & Fruit Tree Research Institute, Shanghai Academy of Agricultural Sciences, Shanghai, 201403; 2 Crop Breeding & Cultivation Research Institute, Shanghai Academy of Agricultural Sciences, Shanghai, 201403; 3 Shanghai Botanical Garden, Shanghai, 200231

* Corresponding author, jkli@sibs.ac.cn; zhujianjun@saas.sh.cn

DOI: 10.5376/mpb.cn.2021.19.0010

Abstract In order to develop genome-SSR (simple sequence repeat) molecular markers of *Cerasus*, the genome sequence survey of *Cerasus cerasoides* was done on IlluminaNovaSeq PE in this study. Microsatellite sequences were identified by using Perl script MISA, and 100 pairs of primers were synthesized randomly for primer polymorphism screening. Then the primers with easy amplification and high polymorphism were selected to detect and analyze the genetic diversity of 45 flowering cherries by fluorescence labeled capillary electrophoresis. The results showed that there were 113 571 SSR loci (≥ 5 repeats) including 53.31% single nucleotide motifs, 36.64% dinucleotide motifs and 10% other motifs. 21 SSRs were developed and their polymorphisms were evaluated in flowering cherries. A total of 215 alleles were detected. The mean number of alleles per locus was 10.24 and the mean heterozygosity was 0.38. The phylogenetic tree constructed by SSR markers could well reflect the genetic relationship among 45 tested *Cerasus* species. In this study, new SSR molecular markers were developed, which

本文首次发表在《分子与植物育种》上, 现依据版权所有人授权的许可协议, 采用 Creative Commons Attribution License, 协议对其进行授权, 再次发表与传播

收稿日期: 2021 年 3 月 5 日; 接受日期: 2021 年 3 月 12 日; 发表日期: 2021 年 3 月 17 日

引用格式: 李水根, 关媛, 李秀芬, 朱继军, 李记开, 朱建军, 2021, 观赏樱属植物基因组 SSR 分子标记开发及鉴定, 分子植物育种 (网络版), 19(10): 1-10 (doi: 10.5376/mpb.cn.2021.19.0010) (Li S.G., Guan Y.A., Li X.F., Zhu J.J., Li J.K., and Zhu J.J., 2021, Development and characterization of genome-SSR markers in flowering cherries, *Fengzi Zhiwu Yuzhong* (Molecular Plant Breeding (online)), 19(10): 1-10 (doi: 10.5376/mpb.cn.2021.19.0010))

can be used for genetic diversity analysis and molecular marker assisted breeding of *Cerasus*.

Keywords *Cerasus cerasoides*; Genome-SSR; Microsatellites; Molecular marker; Genetic diversity

樱花(*Cerasus* sp.)为蔷薇科(Rosaceae)樱属(*Cerasus*)的一类具有重要观赏价值的园林树种,在全世界范围内广泛栽种,由于其艳丽的花朵,在商业和住宅区景观设计中,经常被用于行道树。野生樱花品种极其丰富,主要分布于北半球的温带,特别是在东亚地区,包括中国,日本和韩国。进行樱花的育种工作,培育具有卓越观赏特性和抗性强的樱花品种,可以满足不同地区人们的赏花需求、丰富园林设计的树种选择。然而野生樱花存在自然杂交和种间杂交的发生,从而导致名称,起源和分类单元之间的界限混乱,经常出现同种异名或同名异种的情况,给引种栽培和育种工作带来困难(Katoet al., 2012; Ogawaet al., 2012)。

分子遗传学方法使阐明植物物种之间的遗传关系和促进分子标记辅助育种成为可能。特别是简单序列重复(SSR)标记具有丰度高,基因组内随机分布,高多态性信息含量(PIC)和稳定的共显性(Fraser et al., 2004; Chen et al., 2006),已被证明对植物基因型和栽培品种鉴定极为有效。前人已经采用 SSR 标记对李属物种之间的系统发育关系进行了研究。SSR 被用来鉴定西班牙东北部欧洲李(*Prunus domestica*)的本地品种与其它变种之间的遗传关系(Urrestarazu et al., 2018)。Zhao 等(2015)对中国野生樱桃李(*Prunus cerasifera* Ehrhart) 的遗传多样进行了 SSR 分析。Dirlewanger 等(2002)从富含 CT 的桃基因组文库中筛选出 41 个 SSR 标记,并应用于酸樱桃(*Prunus cerasus*)、扁桃(*Eugenia uniflora*)、草莓(*Fragaria xananassa* Duch.)等的遗传多样性分析。在观赏樱属植物中也有 SSR 标记的研究,并用于区分樱花的种类、杂种及品种(Hongmei et al., 2009; Katoet al., 2012)。伊贤贵等(2018)对山樱花群体的遗传多样性采用 SSR 分析。聂超仁等(2018, 2018 年中国观赏园艺学术研讨会)采用荧光标记电泳法对 54 种国内栽培的观赏樱属品种进行了 SSR 分析。

由于樱花参考基因组缺乏,现有的 SSR 标记大多为 EST (表达序列标签)来源的,而没有樱属基因组 SSR 的报道。公开可用的 EST 和转录组数据库为 EST-SSR 的开发提供大量的参考数据(Thielet et al., 2003; Park et al., 2013; Zhanget al., 2018),然而由于 EST-SSR 位于编码区,序列相对保守,导致多态性水平偏低。而基因组 SSR 具有高多态性,在评估植物遗

传多样性研究中更有优势。对于缺乏参考基因组序列的物种,基因组序列勘测(GSS)可用于开发基因组 SSR (Chen et al., 2009; Wei et al., 2014; Liu et al., 2016; Zhou et al., 2016)。本研究基于喜马拉雅樱花(*Cerasus cerasoides*)的基因组勘测序列,鉴定并开发了新的基因组 SSR。对筛选的 21 个 SSR 标记在 45 种观赏樱属植物中进行鉴定,为樱属植物的遗传标记开发、亲缘关系鉴定、分子标记辅助育种提供参考。

1 结果与分析

1.1 喜马拉雅樱花基因组勘测序列获得

对插入片段大小为 350 bp 的文库进行测序,共获得 32 514 513 900 bp 的原始数据。在对原始数据进行过滤和校正后,获得有效读段数据(clean reads) 32 437 856 100 bp,其中 Q20 和 Q30 分别达到 96.85%和 91.61%,GC 含量为 36.91%。通过 K-mer 分析表明,喜马拉雅樱花基因组大小为 369.68 Mbp,基因组杂合率为 0.37%。对 clean reads 进行组装,共获得 279 319 个重叠群(Contig)和 190 282 个基本骨架(scaffold)(表 1)。原始数据上传至 NCBI 公共数据库获得登录号 SRR8403066 和 SRR8403067。

1.2 SSR 标记的开发与鉴定

使用 MISA 脚本对组装好的 scaffold 进行 SSR 位点的筛选和鉴定,结果共发现 113 571 个 SSR (≥ 5 个重复序列)。在这些 SSR 标记中,单核苷酸是最丰富的基序类型,数量为 60 545, 占有 SSR 的 53.31%。二核苷酸基序的数量仅次于单核苷酸,占有 SSR 的 36.64%。其余基序类型(从三核苷酸到六核苷酸)占有 SSR 的 10%。对 SSR 的重复数类型进行分析表明,最常见为十个串联重复的 SSR (数量 29 108; 占比 25.63%),其次是具有 11 个串联重复的 SSR (数量 14 333; 占比 12.62%) (图 1)。

1.3 SSR 标记的验证和多态性检测

为了验证基因组 SSR 的有效性,我们随机选择了 100 个 SSR (SSR1-100),在供试材料中选择 3 个具有代表性的樱花品种进行多态性检测。结果 100 个 SSR 位点均能扩增出清晰的条带(数据未显示)。然后我们从中筛选出 21 个条带清晰、扩增稳定、多态性好的 SSR 位点,其中包含 10 种二核苷酸基序,6 种

表 1 基因组勘测序列组装结果

Table 1 The results of genome survey sequence assembly

组装类型	总长	总数	最大长度	N50 长度
Assembly type	Total length	Total number	Max length	N50 length
重叠群	218884209 bp	279319	67382 bp	3710 bp
Contig				
基本骨架	228838592 bp	190282	119843 bp	6933 bp
Scaffold				

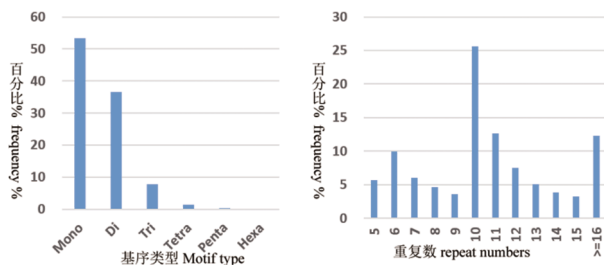


图 1 不同基序类型和不同串联重复类型的 SSR 比例分布

注: Mono: 单核苷酸; Di: 二核苷酸; Tri: 三核苷酸; Tetra: 四核苷酸; Penta: 五核苷酸; Hexa: 六核苷酸

Figure 1 Frequency distributions of different motif type and repeat numbers type

Note: Mono: Mononucleotide; Di: Dinucleotide; Tri: Trinucleotide; Tetra: Tetranucleotide; Penta: Pentanucleotide; Hexa: Hexanucleotide

三核苷酸基序和 5 种四核苷酸基序。设计了相应的 21 对引物(表 2), 并基于荧光毛细管电泳的最高峰位置记录目标 PCR 产物(图 2)。

随后, 在 45 个樱花品种中对这 21 个 SSR 基因座的多态性进行检测评估(表 3)。结果总共检测到 215 个等位基因, 平均每个位点有 10.24 个等位基因。其中 SSR19 位点产生的等位基因数量最多, 为 24 个, 而在 SSR7、SSR20、SSR50、SSR58、SSR64 上仅观察到 5 个等位基因。有效性数值范围为 78%~100%。基因多样性(预期杂合度)范围为 0.38~0.92(平均为 0.71)。杂合度数值的范围为 0.05~0.75(平均值为 0.38)。多态性信息含量 PIC 值的范围为 0.37~0.92(平均值为 0.68)(表 3)。我们发现, 这 21 个 SSR 标记的有效性最低为 78%(SSR58)。其中 7 对 SSR 引物可以在所有供试樱属品种中扩增出条带, 包括 SSR69、SSR74、SSR81、SSR88、SSR89、SSR90 和 SSR93(表 3)。

1.4 观赏樱属植物的亲缘关系分析

对荧光标记毛细管电泳结果进行解析, 并采用 UPGMA 法进行聚类分析, 获得 45 种樱属观赏植物的亲缘关系树状图(图 3)。从聚类图上可知, 在相似系

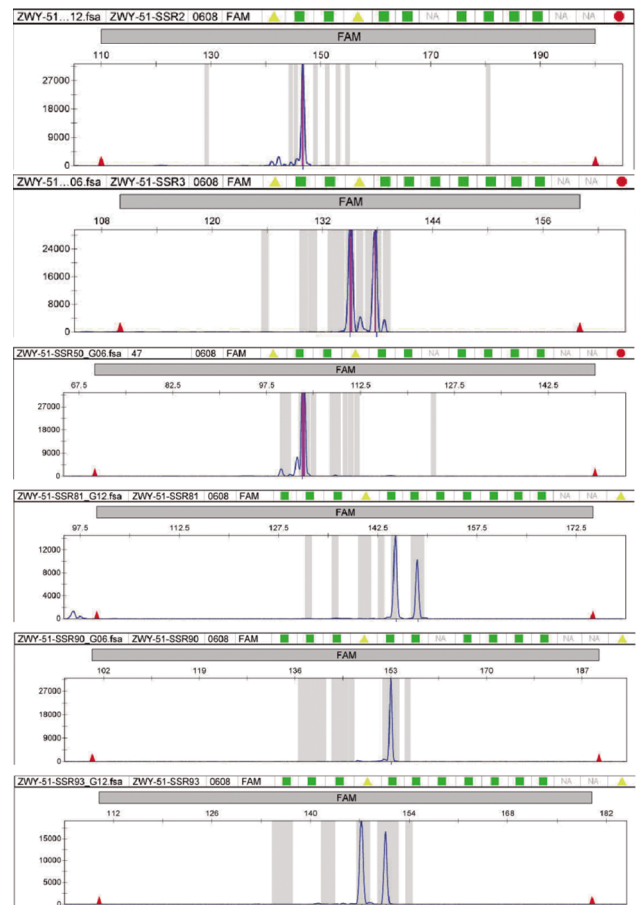


图 2 不同 SSR 引物的荧光毛细管电泳结果(以广州樱为例)

注: 用序列分析仪记录荧光毛细管电泳峰值; x 轴代表 SSR 扩增的片段大小, y 轴代表扩增产物的荧光强度

Figure 2 Fluorescence capillary electrophoresis results of different SSR primers (a case study of *C. campanulata* 'GUANGZHOU')

Note: Capillary electrophoresis peaks were detected using a sequence analyzer; The x-axis represents fragment size of an amplified microsatellite and the y-axis represents the fluorescence intensity of amplified products

数 0.79 处, 可将 45 种樱属植物分为 2 大组(用 A 和 B 表示)。其中 B 组仅包含华中樱和崖樱。其余大部分种类则分到 A 组。A 组又可分成 3 个亚组(A1, A2, A3)。其中 A1 亚组聚集了一些野生品种, 包含亲缘关系较近的高盆樱和喜马拉雅樱。钟花樱及其地

表 2 喜马拉雅樱花 21 个基因组 SSR 开发及引物设计

Table 2 21 genome-SSRs development in *Cerasus cerasoides* 'Himalaya' and design of primers

位点 Locus	引物序列(5'-3') Primer sequences (5'-3')	重复序列 Repeat motif	产物大小(bp) Size (bp)	退火温度(°C) Tm (°C)	登录号 GenBank accession No.
SSR1	F:AAGACAGAATTGGTTTTTATTCCA R:TATGGAGAAGTAATGGAAGCCAG	(TA) ₆	143	59	MN243543
SSR2	F:AGCAGAAGAAAAGGAGGAAGAAC R:TTTAATGCCCAAGCTAGCAGTAG	(TA) ₇	148	59	MN243544
SSR3	F:CTTAGGCTTTCTCCCTCACTCAC R:TCGAGACAAGAGATTGAGAGATTG	(CT) ₈	144	60	MN243545
SSR4	F:CCAATATGCAATAAGCAGGATTC R:TTTTTCAATTTTCCTTGACGTG	(GT) ₆	96	59	MN243546
SSR5	F:GCGTGAATGCTAAATGTAGAGGT R:TCGAAGACTCTCGGAAATCAAT	(GA) ₆	107	59	MN243547
SSR7	F:AAGGATATGATGACCAGCAAAAA R:AGATCTAGGACAAGGCAGAGCTT	(AG) ₆	109	59	MN243548
SSR8	F:GCAGAAGGGGACAAATAGTACCT R:TAGCCAAGGAAATTGTTTCCTGTA	(AT) ₁₀	105	59	MN243549
SSR19	F:TCCTTCTAATTTTCCTTTTCCCA R:GCTAAGCCTTCACAAGTTTAGCA	(TC) ₁₆	120	60	MN243550
SSR20	F:GAGATGAGACAGTGGAGACCAAC R:CACTTCTTCATCATCATCAGTGC	(AT) ₆	135	59	MN243551
SSR25	F:CCCCAGACGACACTCATAAACTA R:ATGTAAAGTTGTGGATCCGAGTG	(TC) ₁₅	104	60	MN243552
SSR50	F:GGTCTTTTGGGTGAATAAGAAGG R:AACCAAATGAATCTGAGATGGTG	(AAG) ₄	108	60	MN243553
SSR51	F:AAAGATGACTCAAACAAACCCAA R:TGCATACTAAAGGGGTGTCTTTT	(TTC) ₄	98	59	MN243554
SSR58	F:AGCTTTGATGACTGAGGTTGAAG R:ACTTCTGTGTCGTTACGCAAAAT	(AAT) ₄	156	59	MN243555
SSR64	F:TCATAGACATGAAAGGAAAAGGC R:CGTCATACCAACTGCCTGTAGT	(TGG) ₅	111	59	MN243556
SSR69	F:ATTCCTCTACAACCTAGGCTGGCT R:TGAAAATTCAAAATGAAAAGCAAA	(TGC) ₄	126	59	MN243557
SSR74	F:TGGTTACAGAGAAGCAGCACATA R:TTTCATTTTCATGCCAAAATCTCT	(CAA) ₄	128	59	MN243558
SSR81	F:TCTTGCATGATTCAACAAACATC R:CCCATGGGAGTAAGTTCCTACTATT	(AGCT) ₄	144	59	MN243559
SSR88	F:ATAGGCCCTCTCTGTACAAGTC R:CATCCACCAATTCATTTTCGTAT	(AATT) ₄	156	60	MN243560
SSR89	F:TTAAGGTTGACGAAATGATAATGC R:CACGGCTTTGAAAAATAAGAAAA	(TATT) ₄	156	59	MN243561
SSR90	F:TTCAGAACATTGCACGTATCATC R:TTTTTAATGCGAAAAATGTACCAC	(TAAA) ₄	148	59	MN243562
SSR93	F:CAAAGTGCTGAGTTTGATTGACA R:CAACAAACAAAATTGCACAACAT	(TGAT) ₄	151	59	MN243563

表 3 21 个 SSR 位点在 45 种观赏樱属植物中的多态性分析
Table 3 Summary statistics of 21 SSR loci among 45 flowering cherry accessions

位点名称	样本数	基因型数量	观测数量	等位基因数	有效性	基因多样性	多态信息含量
Locus	Sample Size	GenotypeNo.	Obs.No.	Allele No.	Availability	GeneDiversity	PIC
SSR1	45	14	43	14	0.96	0.77	0.74
SSR2	45	12	44	8	0.98	0.73	0.69
SSR3	45	15	40	9	0.89	0.83	0.81
SSR4	45	16	45	10	1.00	0.82	0.80
SSR5	45	25	42	11	0.93	0.87	0.86
SSR7	45	6	45	5	1.00	0.62	0.57
SSR8	45	23	43	16	0.96	0.82	0.81
SSR19	45	36	44	24	0.98	0.92	0.92
SSR20	45	9	45	5	1.00	0.58	0.53
SSR25	45	33	44	21	0.98	0.92	0.92
SSR50	45	9	45	5	1.00	0.56	0.51
SSR51	45	15	43	11	0.96	0.73	0.70
SSR58	45	8	35	5	0.78	0.76	0.72
SSR64	45	9	45	5	1.00	0.63	0.57
SSR69	45	8	45	9	1.00	0.38	0.37
SSR74	45	15	42	12	0.93	0.74	0.70
SSR81	45	14	39	8	0.87	0.77	0.74
SSR88	45	9	37	9	0.82	0.52	0.51
SSR89	45	19	38	14	0.84	0.78	0.77
SSR90	45	8	44	7	0.98	0.51	0.48
SSR93	45	10	45	7	1.00	0.65	0.59
Mean	45	15	42.5	10.24	0.94	0.71	0.68

理位置的变种(寒绯樱, 琉球寒绯樱)聚在 A1 亚组的一个分支, 并且与云南樱和牡丹樱的亲缘关系也较近。迎春樱、毛樱桃和刺毛樱桃也聚在一起。A2 和 A3 亚组则聚集了绝大多数的栽培品种, 其中亲缘关系最近的飞寒樱和广州樱, 相似系数达到了 0.95, 两者和阳光樱聚在了一起, 存在较近的亲缘关系。一些栽培品种及其杂交子代或亲本大多也聚集在一起, 例如阳春樱、美国樱、染井吉野樱和江户彼岸樱聚在一支; 寒樱、冬樱和大岛樱等聚为一类。郁金樱和御衣黄樱的相似系数很高, 与太白樱聚在了一起, 属于晚樱品种, 并且与它们的亲本山樱花遗传距离较近。樱桃和椿寒樱聚在了 A2 亚组的一个小分支, 印证了两者的亲缘关系。A3 亚组中包含小彼岸樱、越之彼岸樱和八重红枝垂, 三者具有相近的亲缘关系。

综上, 采用这 21 个基因组 SSR 标记构建的樱属亲缘关系聚类图, 与采用传统的形态分类学构建的进化关系相似, 并且较好的地将野生种和具有亲缘关系的栽培种进行了区分和关联, 能够有效鉴定樱属观赏植物的基因型和进化关系。

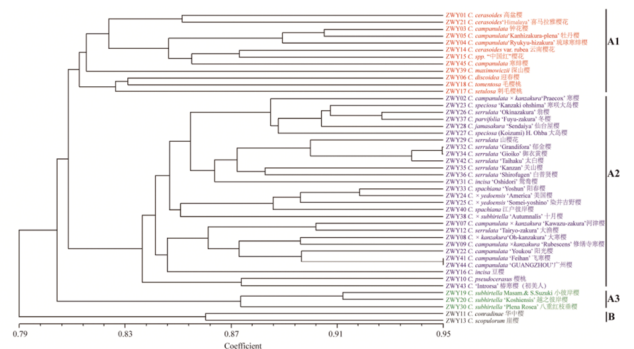


图 3 45 份观赏樱属种质的 UPGMA 分析树状
Figure 3 Dendrogram for 45 Cerasus accessions derived from UPGMA analysis

2 讨论

SSR 由于其高度多态, 共显性和易复制等特征而被广泛应用于遗传多样性分析和分子标记辅助育种(Yanget al., 2015)。与 EST-SSR 相比, 基因组 SSR 标记具有高度多态性, 能够更高效地进行遗传多样性分析(Parthibanet al., 2018)。与琼脂糖或聚丙烯酰胺凝胶上的常规手动读取 DNA 片段相比, 使用荧光

标记毛细管电泳可准确读取 PCR 产物片段大小,并且数据更准确,具有高效和自动化等优点(Lianget al., 2018)。本研究通过筛选喜马拉雅樱花的基因组勘测序列,开发了全新的樱属植物基因组 SSR 标记,并采用荧光标记毛细管电泳法进行分析,为观赏樱属植物的分子标记研究提供参考。

本研究通过对喜马拉雅樱花基因组勘测序列进行分析,共发现 113 571 个 SSR 位点,并且具有相对丰富的二核苷酸和三核苷酸基序。在其他一些植物基因组中也发现了高频率的二核苷酸和三核苷酸基序类型(Chenget al., 2009; Weiet al., 2014)。其次本研究鉴定出的 SSR 中,具有 10 个串联重复序列的 SSR 最常见,与甘蓝型油菜(*Brassica napus*)基因组勘测的研究结果一致(Chenget al., 2009),但与芝麻(*Sesamum*)和疫霉(*Phytophthora*) (常见的重复基序分别为六个和四个重复类型)的研究结果不同(Schenaet al., 2008; Weiet al., 2014)。可能不同植物的 SSR 类型存在差异,SSR 具有丰富的序列多样性。

我们使用 SSR 标记对 45 种樱属植物进行遗传多样性分析,平均 PIC 值为 0.68,其中有 19 个 SSR 的 PIC 值大于 0.5(表 3),这表明通过基因组勘测序列开发的 SSR 标记具有很高的信息量。先前研究中开发的樱属植物 SSR 分子标记,几乎都来自李属,尤其是桃,且大多数为 EST-SSR(Hongmeiet al., 2009; Tsudaet al., 2009; Katoet al., 2012)。Kato 等(2012)使用 SSR 鉴定了日本观赏樱属品种的遗传多样性,期望杂合度为 0.58。Tsuda 等(2009)设计了 13 个 EST-SSR 分子标记,在 6 种观赏樱花品种中进行扩增,并采用荧光毛细管电泳法进行分析,结果获得平均等位基因数量为 3.9。而我们开发的基因组 SSR 标记,扩增的等位基因数量和 PIC 值均高于 Tsuda 等的研究,表明新开发的樱属植物 SSR 标记有效性更高。

在本研究中,我们选取的 45 个供试品种涵盖范围广,包含主要栽培品种和一些重要的野生种质,结果表明新开发的基因组 SSR 标记的有效性高达百分之八十,其中 7 对引物可以在所有品种中扩增出条带。在之前的一些研究中,来自李属的 SSR 被用于分析樱属品种的遗传关系(Hongmeiet al., 2009; Tsudaet al., 2009; Katoet al., 2012)。由于 SSR 基因座可以跨物种扩增,我们开发的 21 个 SSR 分子标记或许也能应用于樱属和李属植物之间的遗传多样性分析,但是需要进一步的实验来验证。

樱花品种繁多,分类方法不尽统一。根据我们采用的遗传标记分析,华中樱和崖樱聚在一起,并且与

其他品种遗传距离较远,两者均为原产于中国的樱属野生种,基于形态特征的分类学也将其放在一起(俞德浚等,编著,1986,科学出版社,中国,北京,pp. 76)。其余种类基本上按照野生系列(A1)和栽培系列(A2 和 A3)进行聚类。野生系列包括高盆樱种系和钟花樱种系。来自日本的品种,包括牡丹樱,琉球寒绯樱,寒绯樱和深山樱也聚集在这个小分支中,表明该分支的种质可能具有共同的祖先。在现有研究中也记录了高盆樱和深山樱之间具有密切的亲缘关系(Hongmeiet al., 2009)。在聂超仁等(2018, 2018 年中国观赏园艺学术研讨会)的研究中,也将高盆樱聚在了钟花樱组,并认为钟花樱种系是较为原始的高盆樱种系向中国东南及南部地区迁移过程中逐步形成的变种。栽培系列聚集了绝大多数的樱花栽培品种,并且基本上按照杂交品种的共同亲本进行聚类。河津樱,大渔樱,飞寒樱和广州樱基于他们已知的共同祖先寒绯樱而聚在一起。椿寒樱与它的亲本樱桃聚集在一起。江户彼岸樱是阳春樱、美国樱、染井吉野樱的亲本,它们聚为一小类。由于山樱花和樱桃是现有樱花栽培种的野生原始祖先,也被分到了栽培系列的大分枝里。这些研究结果表明 SSR 标记较好地反应了观赏樱属植物的亲缘关系,能够有效鉴定樱属植物的野生和栽培品种的遗传多样性。

3 材料与方法

3.1 植物材料

用于基因组序列勘测的材料为上海市农业科学院奉浦院区引种的 5 年生喜马拉雅樱花(朱建军等, 2017),于 2018 年 3 月采集新发生的幼嫩叶片,迅速植于液氮中冷冻,并保存于 -80°C 。用于 SSR 标记评估的 45 份观赏樱属品种(表 4)分别采自上海植物园和上海农业科学院奉浦院区,于 2019 年春季采集各樱花品种的幼嫩叶片,于液氮中冷冻,并保存于 -80°C 中用于基因组 DNA 的提取。

3.2 基因组 DNA 提取、Illumina 测序和序列拼接

采用 CTAB (十六烷基三甲基溴化铵)法提取樱花叶片的基因组 DNA。通过 Nanodrop 分光光度计(Thermo Fisher Scientific Inc, 美国)检测 DNA 的质量和浓度。将高质量的喜马拉雅樱花基因组 DNA 通过超声处理打成片断,然后筛选出 350 bp 大小的片段用于测序文库构建。将材料送至诺禾基因有限公司(中国北京),在 IlluminaNovaSeq PE 平台上进行测序。测序完成后,对原始数据(Raw data)进行过滤处理

表 4 供试的 45 份观赏樱属品种及其特征

Table 4 List of 45 *Cerasus* accessions used in the polymorphism analysis

编号	品种名	起源	类型
No.	Accession	Origin	Type
ZWY-01	高盆樱 <i>C. cerasoides</i>	中国 China	野生 Wild material
ZWY-02	寒樱 <i>C. campanulata</i> × <i>kanzakura</i> 'Praecox'	日本 Japan	栽培 Cultivar
ZWY-03	钟花樱 <i>C. campanulata</i>	中国 China	野生 Wild material
ZWY-04	琉球寒绯樱 <i>C. campanulata</i> 'Ryukyu-hizakura'	日本 Japan	栽培 Cultivar
ZWY-05	牡丹樱 <i>C. campanulata</i> 'Kanhizakura-plena'	日本 Japan	栽培 Cultivar
ZWY-06	迎春樱 <i>C. discoidea</i>	中国 China	野生 Wild material
ZWY-07	河津樱 <i>C. campanulata</i> × <i>kanzakura</i> 'Kawazu-zakura'	日本 Japan	栽培 Cultivar
ZWY-08	大寒樱 <i>C. × kanzakura</i> 'Oh-kanzakura'	日本 Japan	栽培 Cultivar
ZWY-09	修缮寺寒樱 <i>C. campanulata</i> × <i>kanzakura</i> 'Rubescens'	日本 Japan	栽培 Cultivar
ZWY-10	樱桃 <i>C. pseudocerasus</i>	中国 China	野生 Wild material
ZWY-11	华中樱 <i>C. conradinae</i>	中国 China	野生 Wild material
ZWY-12	大渔樱 <i>C. serrulata</i> 'Tairyo-zakura'	日本 Japan	栽培 Cultivar
ZWY-13	崖樱 <i>C. scopulorum</i>	中国 China	野生 Wild material
ZWY-14	云南樱花 <i>C. cerasoides</i> var. <i>rubea</i>	中国 China	栽培 Cultivar
ZWY-15	"中国红" 樱花 <i>C. spp.</i>	中国 China	栽培 Cultivar
ZWY-16	豆樱 <i>C. incisa</i>	日本 Japan	野生 Wild material
ZWY-17	刺毛樱桃 <i>C. setulosa</i>	中国 China	野生 Wild material
ZWY-18	毛樱桃 <i>C. tomentosa</i>	中国 China	野生 Wild material
ZWY-19	小彼岸樱 <i>C. subhirtella</i> Masam.& S.Suzuki	日本 Japan	栽培 Cultivar
ZWY-20	越之彼岸樱 <i>C. subhirtella</i> 'Koshiensis'	日本 Japan	栽培 Cultivar
ZWY-21	喜马拉雅樱花 <i>C. cerasoides</i> 'Himalaya'	尼泊尔 Nepal	野生 Wild material
ZWY-22	阳光樱 <i>C. campanulata</i> 'Youkou'	日本 Japan	栽培 Cultivar

续表 4

Continuing table 4

编号 No.	品种名 Accession	起源 Origin	类型 Type
ZWY-23	寒咲大岛樱 <i>C. speciosa</i> 'Kanzaki ohshima'	日本 Japan	栽培 Cultivar
ZWY-24	美国樱 <i>C. × yedoensis</i> 'America'	日本 Japan	栽培 Cultivar
ZWY-25	染井吉野樱 <i>C. × yedoensis</i> 'Somei-yoshino'	日本 Japan	栽培 Cultivar
ZWY-26	翁樱 <i>C. serrulata</i> 'Okinazakura'	日本 Japan	栽培 Cultivar
ZWY-27	大岛樱 <i>C. speciosa</i> (Koizumi) H. Ohba	日本 Japan	野生 Wild material
ZWY-28	仙台屋樱 <i>C. jamasakura</i> 'Sendaiya'	日本 Japan	栽培 Cultivar
ZWY-29	山樱花 <i>C. serrulata</i>	中国 China	野生 Wild material
Zwy-30	八重红枝垂樱 <i>C. subhirtella</i> 'Plena Rosea'	日本 Japan	栽培 Cultivar
ZWY-31	鸳鸯樱 <i>C. incisa</i> 'Oshidori'	日本 Japan	栽培 Cultivar
ZWY-32	郁金樱 <i>C. serrulata</i> 'Grandifora'	日本 Japan	栽培 Cultivar
ZWY-33	阳春樱 <i>C. spachiana</i> 'Yoshun'	日本 Japan	栽培 Cultivar
ZWY-34	御衣黄樱 <i>C. serrulata</i> 'Gioiko'	日本 Japan	栽培 Cultivar
ZWY-35	关山樱 <i>C. serrulata</i> 'Kanzan'	日本 Japan	栽培 Cultivar
ZWY-36	白普贤樱 <i>C. serrulata</i> 'Shirofugen'	日本 Japan	栽培 Cultivar
ZWY-37	冬樱 <i>C. parvifolia</i> 'Fuyu-zakura'	日本 Japan	栽培 Cultivar
ZWY-38	十月樱 <i>C. × subhirtella</i> 'Autumnalis'	日本 Japan	栽培 Cultivar
ZWY-39	深山樱 <i>C. maximowiczii</i>	日本 Japan	野生 Wild material
ZWY-40	江戸彼岸樱 <i>C. spachiana</i>	日本 Japan	野生 Wild material
ZWY-41	飞寒樱 <i>C. campanulata</i> 'Feihan'	中国 China	栽培 Cultivar
ZWY-42	太白樱 <i>C. serrulata</i> 'Taihaku'	日本 Japan	栽培 Cultivar
ZWY-43	椿寒樱(初美人) <i>C. 'Introrsa'</i>	日本 Japan	栽培 Cultivar
ZWY-44	广州樱 <i>C. campanulata</i> 'GUANGZHOU'	中国 China	栽培 Cultivar
ZWY-45	寒绯樱 <i>C. campanulata</i>	日本 Japan	野生 Wild material

并获得 clean reads。然后通过 De Bruijn Graph 法 (Liet al., 2010) 获得 contigs, 并根据 clean reads 确定 contigs 的方向、减少并缩短空缺 (gap), 最终形成 scaffold。原始数据上传至 NCBI 公共数据库。

3.3 SSR 位点鉴定及引物设计

基于喜马拉雅樱花基因组勘测序列, 采用 Perl 脚本 MicroSATellite (MISA) 进行微卫星 DNA 序列的识别。然后使用 Primer Premier 3.0 在 SSR 位点的侧翼区域设计引物。由擎科生物(中国上海)进行荧光 (FAM) 标记引物的合成。

3.4 PCR 及荧光标记毛细管电泳

PCR 反应采用 50 μL 体系, 其中包含 DNA 模板 1 μL (50 ng/ μL)、上下游引物(10 $\mu\text{mol/L}$)各 2 μL 、Green Mix (上海生工) 45 μL 。PCR 反应程序为: 预变性 98 $^{\circ}\text{C}$, 2 min; 变性 98 $^{\circ}\text{C}$, 10 s; 退火 10 s (温度根据不同的 SSR 引物而设定); 延伸 72 $^{\circ}\text{C}$, 10 s; 循环 30 次; 最后在 72 $^{\circ}\text{C}$ 下再延伸 2 min, 并保持在 4 $^{\circ}\text{C}$ 。

取 0.5 μL 的 PCR 产物用于毛细管电泳上样, 上样液包含 0.5 μL 的 500 LIZ 大小的标准液 (Applied Biosystems Inc, 美国) 和 9 μL 的二甲酰胺 (Applied Biosystems Inc)。上样液采用如下处理: 98 $^{\circ}\text{C}$, 5 min, 置于冰上 2 min, 4000 rpm 离心 1 min。使用 ABI 3730XL 遗传分析仪 (Applied Biosystems Inc) 对上样反应过程进行分析。在毛细管电泳过程中, PCR 产物的分离过程会自动记录到单独的 GeneScan 文件中, 然后使用 GeneMapper 4.0 软件 (Applied Biosystems Inc) 对收集的原始数据进行分析。

3.5 数据统计分析

根据荧光标记毛细管电泳结果, 采用 PowerMarker 3.25 计算基因型个数、等位基因个数、基因多样性、杂合性、多态性信息指数。利用 NTSYSpc-2.11 软件包 (Exeter software Setauket, 美国) 构建基于 UP-GMA 聚类算法的树状图。

作者贡献

李水根是本研究的执行人, 完成数据分析和论文的写作; 关媛和李秀芬参与基因组勘测、荧光毛细管电泳及数据分析; 朱继军提供遗传多样性分析的植物材料。李记开完成生物信息学分析、SSR 引物设计和荧光毛细管电泳; 朱建军是试验的构思者, 指导实验设计, 数据分析, 论文写作与修改。全体作者都

阅读并同意最终的文本。

致谢

本研究由上海市种业发展项目 (沪农科种字 (2017) 第 1-6 号) 和上海市市级农口系统青年人才成长计划项目 (沪农青字 (2018) 第 1-9 号) 共同资助。

参考文献

- Chen C., Zhou P., Choi Y.A., Huang S., and Gmitter F.G., 2006, Mining and characterizing microsatellites from citrus ESTs, *Theoretical and Applied Genetics*, 112(7):1248-1257.
- Cheng X., Xu J., Xia S., Gu J., Yang Y., Fu J., Qian X., Zhang S., Wu J., and Liu K., 2009, Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*, *TAG Theoretical and applied genetics Theoretische und angewandte Genetik*, 118(6):1121-1131.
- Dirlewanger E., Cosson P., Tavaud M., Aranzana J., Poizat C., Zanetto A., Arus P., and Laigret F., 2002, Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.), *TAG Theoretical and applied genetics Theoretische und angewandte Genetik*, 105 (1): 127-138.
- Fraser L.G., Harvey C.F., Crowhurst R.N., and De Silva H.N., 2004, EST-derived microsatellites from *Actinidia* species and their potential for mapping, *Theoretical and Applied Genetics*, 108(6):1010-1016.
- Hongmei M., Richard O., and Margaret P., 2009, Evaluation of flowering cherry species, hybrids, and cultivars using simple sequence repeat markers, *J. Amersochortsci.*, 134 (4): 435-444.
- Kato S., Matsumoto A., Yoshimura K., Katsuki T., Iwamoto K., Tsuda Y., Ishio S., Nakamura K., Moriwaki K., Shiroishi T., Gojobori T., and Yoshimaru H., 2012, Clone identification in Japanese flowering cherry (*Prunus* subgenus *Cerasus*) cultivars using nuclear SSR markers, *Breeding science*, 62(3): 248-255.
- Li R., Zhu H., Ruan J., Qian W., Fang X., Shi Z., Li Y., Li S., Shan G., Kristiansen K, Li S., Yang H., Wang J., and Wang J., 2010, De novo assembly of human genomes with massively parallel short read sequencing, *Genome research*, 20 (2):265-272
- Liang C., Wan T., Xu S., Li B., Li X., Feng Y., and Cai Y., 2018, Molecular identification and genetic analysis of cherry cultivars using capillary electrophoresis with fluorescence-labeled SSR markers, *3 Biotech*, 8(1):16.
- Liu Y.J., Zhang C.Y., Hao G., Ge X.J., and Yan H.F., 2016, Characterization of novel microsatellite loci for *Primula*

- poissonii (Primulaceae) using high-throughput sequencing technology, *Molecules*, 21(5): 536.
- Ogawa T., Kameyama Y., Kanazawa Y., Suzuki K., and Somego M., 2012, Origins of early-flowering cherry cultivars, *Prunus* × *kanzakura* cv. Atami-zakura and *Prunus* × *kanzakura* cv. Kawazu-zakura, revealed by experimental crosses and AFLP analysis, *Scientia Horticulturae*, 140:140-148.
- Park I., Kim J., Lee J., Kim S., Cho O., Yang K., Ahn J., Nahm S., and Kim H., 2013, Development of SSR markers by next-generation sequencing of Korean landraces of chamoe (*Cucumis melo* var. *makuwa*), *Molecular Biology Reports*, 40(12):6855-6862.
- Parthiban S., Govindaraj P., and Senthilkumar S., 2018, Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane, *Biotech*, 8(3):144.
- Schena L., Cardle L., and Cooke D.E., 2008, Use of genome sequence data in the design and testing of SSR markers for *Phytophthora* species, *BMC genomics*, 9:620.
- Thiel T., Michalek W., Varshney R.K., and Graner A., 2003, Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.), *TAG Theoretical and applied genetics Theoretische und angewandte Genetik*, 106(3):411-422.
- Tsuda Y., Ueno S., Kato S., Katsuki T., Mukai Y., and Tsumura Y., 2009, Development of 13 EST-SSRs for *Cerasus jamasakura* and their transferability for Japanese flowering cherries, *Conservation Genetics*, 10(3):685-688.
- Urrestarazu J., Errea P., Miranda C., Santesteban L.G., and Pina A., 2018, Genetic diversity of Spanish *Prunus domestica* L. germplasm reveals a complex genetic structure underlying, *PLoS one*, 13(4):e0195591.
- Wei X., Wang L., Zhang Y., Qi X., Wang X., Ding X., Zhang J., and Zhang X., 2014, Development of simple sequence repeat (SSR) markers of sesame (*Sesamum indicum*) from a genome survey, *Molecules*, 19(4):5150-5162.
- Yang Y., Jing Z.B., Ruan X.F., and Cheng J.M., 2015, Development of simple sequence repeat markers in persimmon (*Diospyros* L.) and their potential use in related species, *Genetics and Molecular Research*, 14(1):609-618.
- Yi X.G., Chen J., You L., Cong R., Wang H., Duan Y., and Wang X., 2018, Genetic diversity of *Cerasus serrulata* populations assessed by SSR markers, *nanjing linye daxue xuebao (Journal of Nanjing Forestry University (Natural Sciences Edition))*, 42(5): 29-35. (伊贤贵, 陈洁, 尤禄祥, 从睿, 王华辰, 段一凡, 王贤荣, 2018, 山樱花群体遗传多样性的 SSR 分析, *南京林业大学学报(自然科学版)*, 42(5): 29-35.)
- Zhang L., Yang X., Qi X., Guo C., and Jing Z., 2018, Characterizing the transcriptome and microsatellite markers for almond (*Amygdalus communis* L.) using the Illumina sequencing platform, *Hereditas*, 155:14.
- Zhao Y., Li Y., Liu Y., and Yang Y.F., 2015, Genetic diversity of wild *Prunus cerasifera* Ehrhart (wild cherry plum) in China revealed by simple-sequence repeat markers, *Genetics and Molecular Research*, 14(3):8407-8413.
- Zhou X., Dong Y., Zhao J., Huang L., Ren X., Chen Y., Huang S., Liao B., Lei Y., Yan L., and Jiang H., 2016, Genomic survey sequencing for development and validation of single-locus SSR markers in peanut (*Arachis hypogaea* L.), *BMC genomics*, 17:420.
- Zhu J.J., Li X.F., Tian Y., Li S.G., Guan Y., Fan X.F., and Yin L. Q., 2017, Introduction and ecological adaptability of Himalayan cherry, *Shanghai Nongye Xuebao (Acta Agriculturae Shanghai)*, 33(5):64-67. (朱建军, 李秀芬, 田原, 李水根, 关媛, 范晓芬, 殷丽青, 2017, 喜马拉雅樱花的引种及生态适应性研究, *上海农业学报*, 33(5): 64-67.)