

高 阳,侯非凡,熊 雄,等. 基于候选基因开发 gSSR 标记加密黄花菜和萱草种间遗传图谱[J].江苏农业学报,2021,37(1):139-149.  
doi:10.3969/j.issn.1000-4440.2021.01.018

## 基于候选基因开发 gSSR 标记加密黄花菜和萱草种间遗传图谱

高 阳<sup>1,2</sup>, 侯非凡<sup>1,2,3</sup>, 熊 雄<sup>1,2</sup>, 公菲菲<sup>1,2</sup>, 王金耀<sup>1,2,3</sup>, 邢国明<sup>1,2,3</sup>, 李 森<sup>1,2,3</sup>

(1.山西农业大学园艺学院,山西 太谷 030801; 2.大同黄花产业发展研究院,山西 大同 037004; 3.山西省设施蔬菜产业提质增效协同创新中心,山西 太谷 030801)

**摘要:** 为了更加精准地开展黄花菜基因组组装和性状定位,在课题组前期构建的大同黄花×摇篮曲种间遗传图谱的基础上,基于大同黄花全基因组和转录组数据库,针对秋水仙碱及昼夜节律 2 个关键性状所涉及的候选 Unigene 开发基因组 SSR 标记,进一步通过 55 个 F<sub>1</sub> 后代加密图谱。结果表明,新开发的 95 个 gSSR 标记中有 46 个标记在亲本及子代间具有多态性,多态性比例为 48.42%,有 22 个 gSSR 标记到图谱上,同时有 12 个前期未上图的 EST-SSR 标记加密到新图谱。加密后的遗传图谱包含 255 个标记,由 9 个连锁群组成,总图距为 3 142.64 cM,平均图距为 11.88 cM,各连锁群标记数量为 5~58 个。研究结果为进一步研究黄花菜秋水仙碱含量及昼夜节律的功能基因定位奠定了重要基础。

**关键词:** 黄花菜; gSSR; 标记开发; 图谱加密

**中图分类号:** Q755 **文献标识码:** A **文章编号:** 1000-4440(2021)01-0139-11

## Enhancement of the interspecific genetic map density for *Hemerocallis citrina* and *H. fulva* by developing gSSR markers based on candidate genes

GAO Yang<sup>1,2</sup>, HOU Fei-fan<sup>1,2,3</sup>, XIONG Xiong<sup>1,2</sup>, GONG Fei-fei<sup>1,2</sup>, WANG Jin-yao<sup>1,2,3</sup>,  
XING Guo-ming<sup>1,2,3</sup>, LI Sen<sup>1,2,3</sup>

(1. College of Horticulture, Shanxi Agricultural University, Taigu 030801, China; 2. Datong Huanghua Industry Development Research Institute, Datong 037004, China; 3. Collaborative Innovation Center of Quality and Profit Improvement for the Facility Vegetables of Shanxi Province, Taigu 030801, China)

**Abstract:** To carry out genome assembly and character mapping of *Hemerocallis citrina* more accurately, genomic simple sequence repeats (SSR) markers were developed for the candidate Unigenes involved in two key characters (colchicine and circadian rhythm), on the basis of the interspecific genetic map of DatongHuanghua×Lullaby Baby constructed by our research group previously and the whole genome and transcriptome database of DatongHuanghua. The map density was further enhanced through 55 F<sub>1</sub> progenies. The results showed that 46 of the 95 newly developed gSSR markers were polymorphic between parents and offsprings, with a polymorphic ratio of 48.42%. 22 gSSR markers were added to the map, while 12 EST-SSR markers not mapped earlier were added to the new map. The updated genetic map contained 255 markers which was composed of nine linkage groups. The total map distance was 3 142.64 cM, and the average map distance was 11.88 cM. The number of markers in each linkage group ranged from 5 to 58. The results can make important foundation for further study on colchicine content and functional gene mapping of circadian rhythm in *H. citrina*.

**Key words:** *Hemerocallis citrina*; gSSR; marker development; density enhancement of the genetic linkage map

收稿日期:2020-06-26

基金项目:山西农业大学青年拔尖创新人才支持计划项目(BJRC20-1601);山西省农业重点研发计划重点项目(20170-3D211001-04-05,201903D211011-01);山西省高等学校科技创新项目(2019L0375)

作者简介:高 阳(1995-),男,山西朔州人,硕士研究生,主要从事园艺植物生物技术与遗传改良方面的研究。(E-mail) sau\_gaoyang@163.com

通讯作者:李 森,(E-mail) saulisen@163.com

萱草属 (*Hemerocallis* spp.) 植物是世界著名的宿根花卉之一, Li 等<sup>[1]</sup>将 155 个萱草属植物品种分为黄花菜 (*Hemerocallis citrina*) 等夜间开花类群和萱草 (*Hemerocallis fulva*) 等白天开花类群。黄花菜花蕾营养价值丰富, 又名“金针菜”, 是一种重要的药食兼用型功能蔬菜<sup>[2-4]</sup>, 产量、品质和营养成分含量是其重要的育种目标, 黄花菜中富含多种次生代谢物, 这些物质在现代医学及生物学研究中发挥着重要作用<sup>[5]</sup>。黄花菜新鲜花蕾中含有次生代谢物秋水仙碱<sup>[6-7]</sup>, 有报道显示鲜食黄花菜可能产生中毒现象, 导致腹泻等不良反应<sup>[8]</sup>。萱草花色、花型丰富, 耐干旱瘠薄, 是重要的园林绿化材料, 其育种主要围绕改良观赏性状、提升观赏价值来进行<sup>[9]</sup>, 萱草的观赏性主要体现在花器官, 而萱草属植物花器官存在明显的昼夜节律现象, 该现象成为影响其观赏性的重要因素。综上, 在黄花菜与萱草育种过程中, 次生代谢物含量和花器官昼夜节律现象是 2 个值得重点关注和研究的性状。由于萱草属植物为宿根植物, 新品种多通过自然发现新优性状种质后无性扩繁获得, 育种存在盲目性和随机性。分子标记辅助选择育种是开展多年生植物新品种选育的有效手段, 性状选择有针对性, 育种年限可大幅缩短。其中, 高密度遗传图谱是开展重要农艺性状分子标记辅助育种的重要基础。

萱草属植物遗传背景较为复杂<sup>[10]</sup>, 童期较长, 制作高代自交系难度较大, 截至目前仅发表了 2 张本属植物的遗传图谱。侯非凡<sup>[11]</sup>于 2017 年基于大同黄花 (*Hemerocallis citrina* cv. DatongHuanghua) 转录组数据, 利用“拟测交”策略构建了 1 张包含 199 个 EST-SSR (表达序列标签微卫星) 分子标记、11 个连锁群的黄花菜种内遗传图谱, 总图距为 2 522.34 cM, 标记间平均图距为 12.68 cM; 同时构建了 1 张黄花菜与萱草种间遗传图谱, 包含 222 个 EST-SSR 分子标记、11 个连锁群, 总图距为 2 485.59 cM, 标记间平均图距为 11.20 cM。2 张图谱标记数量有限, 标记开发缺乏一定的针对性。

本研究基于课题组前期已获得的大同黄花全基因组数据及转录组数据, 使用秋水仙碱及昼夜节律所涉及的代谢通路上的候选 Unigene 与基因组数据进行比对, 开发与这 2 个性状相关的基因组 gSSR (基因组 SSR) 标记, 对侯非凡<sup>[11]</sup>构建的黄花菜与萱草种间遗传图谱进行加密, 旨在获得 1 张密度更高的黄花菜与萱草种间遗传图谱, 为黄花菜和萱草的关键性状基

因定位和 QTL (数量性状座位) 分析奠定基础。

## 1 材料与方法

### 1.1 gSSR 标记开发

1.1.1 基于转录组数据的候选 Unigene 筛选 结合课题组已有的转录组数据库<sup>[1]</sup> (NCBI 项目编号为 PRJNA628147; 登录号为 SRR11610941、SRR11610942 和 SRR11610943), 根据测序数据所得 Unigene 在 KEGG 数据库中的功能注释结果, 筛选功能注释于秋水仙碱 (编号: ko00950) 和昼夜节律 (编号: ko04712) 2 条代谢通路中的 Unigene 作为候选基因。

1.1.2 基于基因组的候选 Scaffold 序列获取 结合课题组大同黄花全基因组 De Novo 测序结果 (基因组序列未发布), 使用 Blastn 2.2.31 软件将候选 Unigene 与基因组序列进行比对, 设软件中 *E*-value 等参数为默认值, 筛选其中最佳的 Scaffold 序列。根据筛选结果, 使用 faSomeRecords 脚本 ([http://hg-download.cse.ucsc.edu/admin/exe/linux.x86\\_64/](http://hg-download.cse.ucsc.edu/admin/exe/linux.x86_64/)) 提取基因组中目标 Scaffold 序列, 统计各候选 Unigene 比对到基因组 Scaffold 上的区间位置信息, 将每个比对位点上游、下游各扩展 3 kb 序列, 使用 BioEdit 软件将此区间序列导出。

1.1.3 SSR 识别及引物设计 本研究采用 SSRHunter1.3 软件识别候选 Scaffold 序列中的 SSR 位点<sup>[12]</sup>。设置构成重复元件的核苷酸数最多为 6 个, 重复次数最少为 5 次, 进行 SSR 位点识别。SSR 基序分析参考 Zhang 等<sup>[13]</sup>的方法。使用 Primer Premier 5 进行引物设计, 主要参数: 引物长度为 20~26 bp, 退火温度 ( $T_m$ ) 为 58~62 °C, (G+C) 含量为 45%~55%, PCR 扩增产物大小为 150~300 bp, 引物均由南京钟鼎生物技术有限公司合成。

### 1.2 萱草属植物种间遗传图谱的加密

1.2.1 作图亲本及作图群体 作图群体引自课题组前期以黄花菜地方品种大同黄花作为杂交母本, 萱草园艺品种摇篮曲 (*Hemerocallis fulva* Lullaby Baby) 作为杂交父本 (图 1) 构建的 55 株  $F_1$  代群体<sup>[11]</sup>, 亲本材料在开花时间、花色与花型上具有较大差别, 且在所有的种间杂交组合中, 大同黄花×摇篮曲组合的坐果率与萌发率较高, 收获了最多的  $F_1$  杂交子代, 因此使用大同黄花、摇篮曲作为构建黄花菜和萱草种间遗传图谱的亲本, 亲本及杂交  $F_1$  代群体现保存于山西农业大学萱草种质资源圃内。



左侧为母本大同黄花,右侧为父本摇篮曲。

图1 杂交亲本花部照片

Fig.1 Photograph of parents for crossbreeding

**1.2.2 DNA 提取** 采取改良 CTAB(十六烷基三甲基溴化铵)法<sup>[11]</sup>提取亲本及 55 个  $F_1$  子代 DNA,通过 1%琼脂糖凝胶电泳检测 DNA 完整度,使用 NanoDrop Spectrometer 检测 DNA 浓度与纯度,合格样品于  $-20\text{ }^{\circ}\text{C}$  冰箱存放备用。

**1.2.3 SSR 引物筛选及扩增** 本试验中所有 gSSR 引物均来自 1.1.3 中 Primer Premier 5 的设计结果,标记名称延续冀芳芳<sup>[14]</sup>的命名方式,以 *sau* 开头,顺次编号为 *sau01169*、*sau01170* 等。利用母本大同黄花和父本摇篮曲及从  $F_1$  代中随机选取的 4 个子代共 6 个样品来筛选多态性 SSR 引物。PCR 扩增、PCR 扩增产物凝胶电泳检测、SSR 标记分离检测、读带数据处理等参照侯非凡<sup>[11]</sup>的方法。

**1.2.4 图谱加密** 利用 Joinmap 4.1 软件对新开发的 gSSR 标记及原始图谱构建时所使用的标记进行分析和作图,从而加密原始种间遗传图谱。

## 2 结果与分析

### 2.1 候选 Unigene 的筛选

秋水仙碱属于异喹啉类生物碱,共有 20 个 Unigene 注释到秋水仙碱生物合成所在的异喹啉类生物碱代谢通路中,其在 KEGG 数据库中的通路编号为 ko00950。结合转录组数据中基因功能注释结果,20 个 Unigene 所注释到的酶均为秋水仙碱合成通路中所涉及的关键酶(表 1)。

基于植物昼夜节律代谢通路图(KEGG 数据库编号为 ko04712),其中心振荡环上共有 6 个基因,分别为 *CCA1*、*LHY*、*TOC1*、*APRR1*、*ZTL*、*GI*。结合转录组数据中基因功能注释结果,共 18 个 Unigene 注释为这 6 个基因。由于本试验基于拟南芥 *LHY* 基因的 CDS(编码序列)对 Unigene 库进行 BLAST 分析,未直接注释到 *LHY* 基因的 c56440.graph\_c0 与拟南芥中的 *LHY* 基因序列相似度较高,所以将此 Unigene 也加入到候选 Unigene 范围(表 1)。

综上,共有 39 个 Unigene 成为本试验候选 Unigene。

表 1 注释到秋水仙碱合成和昼夜节律通路上的 Unigene

Table 1 Unigenes annotated in the pathway of colchicine synthesis and circadian rhythm

Unigene 编号	代谢通路名称	代谢通路编号	节点编号	功能注释
c57714.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00276	Copper amine oxidase
c86988.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00276	Copper amine oxidase
c96096.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00276	Copper amine oxidase
c100846.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c14678.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c50506.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c56295.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c90004.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c90621.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c98491.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00813	Aspartate aminotransferase
c13585.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Tyrosine aminotransferase
c1751.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Tyrosine aminotransferase
c33600.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Nicotianamine aminotransferase
c54449.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Nicotianamine aminotransferase
c65364.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Aminotransferase TAT2
c84507.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K00815	Aminotransferase TAT2 isoform X1
c41726.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K01592	Tyrosine decarboxylase
c65389.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K01592	Dopa decarboxylase

续表1 Continued1

Unigene 编号	代谢通路名称	代谢通路编号	节点编号	功能注释
c89227.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K01592	Tyrosine decarboxylase
c27305.graph_c0	Isoquinoline alkaloid biosynthesis	ko00950	K01593	Tyrosine decarboxylase
c34670.graph_c0	Circadian rhythm - plant	ko04712	K12134	<i>Circadian Clock Associated 1</i>
c38019.graph_c0	Circadian rhythm - plant	ko04712	K12134	<i>Circadian Clock Associated 1</i>
c34855.graph_c0	Circadian rhythm - plant	ko04712	K12133	<i>Late Elongated Hypocoty l</i>
c14137.graph_c0	Circadian rhythm - plant	ko04712	K12133	<i>Late Elongated Hypocoty l</i>
c69447.graph_c0	Circadian rhythm - plant	ko04712	K12133	<i>Late Elongated Hypocoty l</i>
c33185.graph_c0	Circadian rhythm - plant	ko04712	K12133	<i>Late Elongated Hypocoty l</i>
c33464.graph_c0	Circadian rhythm - plant	ko04712	K12133	<i>Late Elongated Hypocoty l</i>
c56440.graph_c0				<i>Late Elongated Hypocoty l-like</i>
c56403.graph_c0	Circadian rhythm - plant	ko04712	K12127	<i>Timing Of Cab Experssion 1</i>
c63847.graph_c0	Circadian rhythm - plant	ko04712	K12127	<i>Timing Of Cab Experssion 1</i>
c59239.graph_c0	Circadian rhythm - plant	ko04712	K12127	<i>Timing Of Cab Experssion 1</i>
c65537.graph_c0	Circadian rhythm - plant	ko04712	K12127	<i>Arabidopsis Pseudo-Response Regulator 1</i>
c5401.graph_c0	Circadian rhythm - plant	ko04712	K12127	<i>Arabidopsis Pseudo-Response Regulator 1</i>
c82874.graph_c0	Circadian rhythm - plant	ko04712	K12115	<i>Zeitlupe</i>
c56932.graph_c0	Circadian rhythm - plant	ko04712	K12115	<i>Zeitlupe</i>
c102994.graph_c0	Circadian rhythm - plant	ko04712	K12124	<i>Gigantea</i>
c95972.graph_c0	Circadian rhythm - plant	ko04712	K12124	<i>Gigantea</i>
c93745.graph_c0	Circadian rhythm - plant	ko04712	K12124	<i>Gigantea</i>
c58223.graph_c0	Circadian rhythm - plant	ko04712	K12124	<i>Gigantea</i>

## 2.2 候选 Scaffold 序列比对结果

39 条候选 Unigene 的 BLAST 报告筛选结果表明,与秋水仙碱合成通路相关的 20 条 Unigene 中共有 9 条比对到了基因组序列上,涉及 16 条 Scaffold 序列;与植物昼夜节律代谢通路的中心振荡环相关的 19 条 Unigene 中共有 11 条 Unigene 比对到了基因组序列上,涉及 35 条 Scaffold 序列(表 2)。综上,共得到 51 条候选 Scaffold 序列。其中与昼夜节律相

关的 c58223.graph\_c0 比对到的 Scaffold 序列最多,为 8 条。c56932.graph\_c0 和 mergeScaf2663 序列相似性最高,比对分数为 3 781 分,c56403.graph\_c0 的比对分数最低。

根据各候选 Unigene 比对到基因组 Scaffold 上的区间位置信息,将每个位点上下游各扩展 3 kb 序列,即区间共扩展 6 kb 序列,候选序列在各 Scaffold 自 5'端到 3'端的位置信息如表 2 所示。

表 2 基因组数据 BLAST 比对结果及相应序列位置信息

Table 2 BLAST alignment results and corresponding sequence location information

Unigene 编号	Scaffold 编号	BLAST 分值	E 值	Unigene 长度 (bp)	Scaffold 长度 (bp)	候选 Scaffold 5'端起始位点(bp)	候选 Scaffold 3'端终止位点 (bp)
c27305.graph_c0	mergeScaf13882	819	0	478	446 837	70 313	76 798
c33600.graph_c0	mergeScaf6547	754	0	516	304 350	62 301	68 711
	mergeScaf2284	754	0	516	213 768	64 629	71 039
c41726.graph_c0	mergeScaf19542	808	0	2 139	408 636	159 066	165 502
c50506.graph_c0	mergeScaf6820	826	0	1 834	289 639	57 621	64 076
c54449.graph_c0	mergeScaf14244	717	0	1 640	459 219	324 147	330 553
	Scaf32862	712	0	1 640	199 859	143 262	149 668
	mergeScaf1415	686	0	1 640	327 540	186 585	192 987
c56295.graph_c0	mergeScaf3901	745	0	1 769	548 301	232 116	238 527



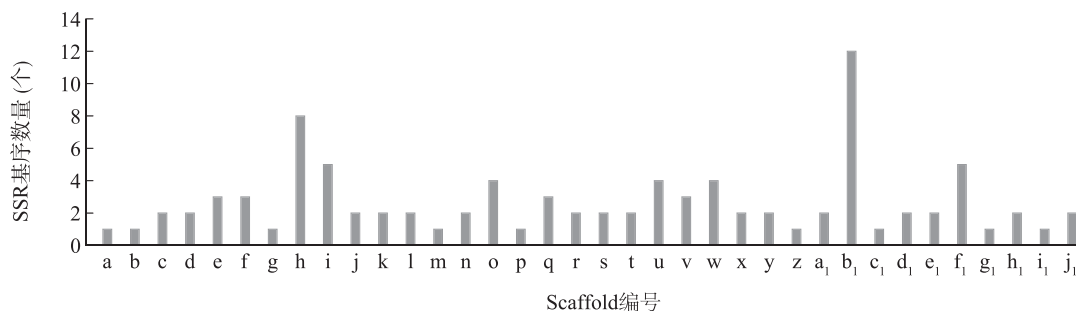
续表2 Continued2

Unigene 编号	Scaffold 编号	BLAST 分值	E 值	Unigene 长度 (bp)	Scaffold 长度 (bp)	候选 Scaffold 5'端 起始位点(bp)	候选 Scaffold 3'端 终止位点(bp)
c57714.graph_c0	mergeScaf306	745	0	1 769	473 828	154 115	160 526
	mergeScaf1256	1 125	0	3 036	483 881	471 559	478 173
	mergeScaf4261	1 085	0	3 036	259 030	140 294	146 907
	mergeScaf18125	776	0	3 036	351 385	149 785	156 222
	mergeScaf14141	765	0	3 036	295 418	199 596	206 033
c65389.graph_c0	mergeScaf14640	689	0	3 036	365 049	270 763	277 175
	mergeScaf13882	518	$2 \times 10^{-144}$	286	446 837	68 412	75 974
	mergeScaf17683	472	$1 \times 10^{-130}$	258	209 149	95 227	101 484
c89227.graph_c0	mergeScaf13882	472	$1 \times 10^{-130}$	258	446 837	68 748	76 282
c33185.graph_c0	Scaffold40991	2 440	0	2 561	199 302	181 810	190 001
c33464.graph_c0	mergeScaf19802	2 148	0	2 983	327 701	230 799	239 710
	mergeScaf20216	750	0	2 983	309 468	282 519	289 276
	mergeScaf4427	708	0	2 983	323 214	184 960	191 868
c34670.graph_c0	mergeScaf1362	553	$6 \times 10^{-155}$	330	323 670	257 539	263 837
c38019.graph_c0	mergeScaf17780	1 142	0	1 473	535 294	23 061	30 187
	mergeScaf17754	1 090	0	1 473	428 398	262 857	269 494
	mergeScaf14839	1 085	0	1 473	353 627	150 450	157 087
	mergeScaf1978	1 085	0	1 473	371 664	174 866	181 503
c5401.graph_c0	mergeScaf9187	702	0	442	438 166	433 788	440 619
	mergeScaf2893	702	0	442	303 924	58 510	65 341
	mergeScaf1566	702	0	442	200 689	46 531	53 362
c56403.graph_c0	mergeScaf15491	315	$4 \times 10^{-82}$	4 069	592 064	523 741	530 005
c56440.graph_c0	mergeScaf664	1 018	0	936	571 437	49 040	57 438
	mergeScaf17072	647	0	936	279 517	31 660	37 945
	mergeScaf7711	647	0	936	312 328	40 884	47 244
c56932.graph_c0	mergeScaf2663	3 781	0	2 557	247 717	91 186	99 295
	mergeScaf17449	1 408	0	2 557	437 531	290 246	305 318
	mergeScaf582	1 391	0	2 557	404 541	64 735	72 333
	mergeScaf19449	1 380	0	2 557	143 501	111 312	118 910
	mergeScaf19271	1 127	0	2 557	260 642	256 268	263 642
c58223.graph_c0	mergeScaf20123	664	0	2 557	414 633	359 996	366 435
	mergeScaf3154	2 763	0	4 644	285 912	68 058	79 812
	mergeScaf8083	2 121	0	4 644	230 067	28 605	35 773
	mergeScaf1943	1 980	0	4 644	368 772	110 492	128 323
	mergeScaf3246	1 700	0	4 644	398 453	250 456	257 882
	mergeScaf2588	965	0	4 644	163 507	126 357	141 114
	mergeScaf11145	789	0	4 644	460 462	43 788	52 490
	mergeScaf9853	701	0	4 644	340 860	10 430	16 808
	mergeScaf12713	673	0	4 644	434 257	253 134	259 521
	mergeScaf17882	889	0	484	227 685	65 228	71 711
c63847.graph_c0	mergeScaf16053	889	0	484	325 315	88 848	95 331
	mergeScaf10442	889	0	484	702 422	289 358	295 841
c65537.graph_c0	mergeScaf13017	571	$2 \times 10^{-160}$	315	503 026	444 453	450 767
	mergeScaf9197	571	$2 \times 10^{-160}$	315	362 727	105 182	111 496

### 2.3 候选 Scaffold 序列中 SSR 基序的鉴定

SSRHunter 1.3 检测结果表明,有 36 条 Scaffold 检测到了 SSR 位点,其中与秋水仙碱相关的 Scaffold 为 11 条,与昼夜节律相关的 Scaffold 为 25 条。在上

述 36 条 Scaffold 序列中共有 95 个 gSSR 位点,其中 mergeScaf17449 包含最多的 gSSR (12 个),其次是 mergeScaf1943 (8 个),有 9 条 Scaffold 只能检测到 1 个 gSSR 位点(图 2)。



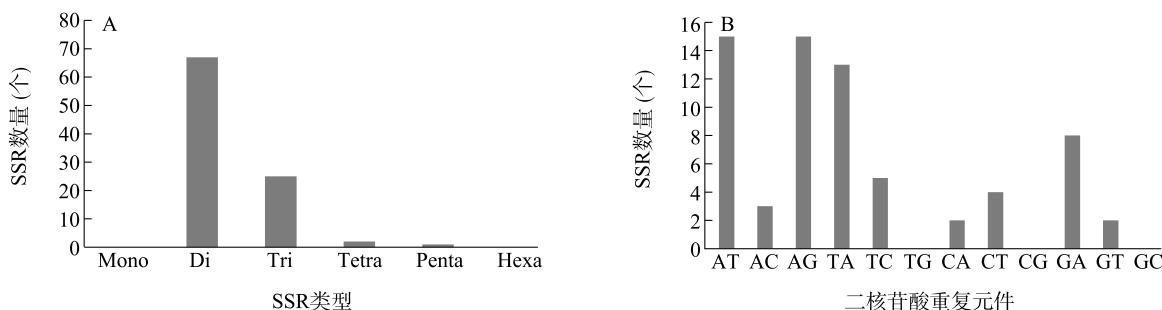
a: mergeScaf306; b: mergeScaf582; c: mergeScaf664; d: mergeScaf1256; e: mergeScaf1362; f: mergeScaf1415; g: mergeScaf1566; h: mergeScaf1943; i: mergeScaf2284; j: mergeScaf2663; k: mergeScaf2893; l: mergeScaf3154; m: mergeScaf3901; n: mergeScaf4261; o: mergeScaf4427; p: mergeScaf6547; q: mergeScaf7711; r: mergeScaf8083; s: mergeScaf9187; t: mergeScaf9197; u: mergeScaf11145; v: mergeScaf12713; w: mergeScaf13017; x: mergeScaf13882; y: mergeScaf14640; z: mergeScaf14839; a<sub>1</sub>: mergeScaf15491; b<sub>1</sub>: mergeScaf17449; c<sub>1</sub>: mergeScaf17754; d<sub>1</sub>: mergeScaf18125; e<sub>1</sub>: mergeScaf19271; f<sub>1</sub>: mergeScaf19449; g<sub>1</sub>: mergeScaf19542; h<sub>1</sub>: mergeScaf19802; i<sub>1</sub>: mergeScaf20216; j<sub>1</sub>: Scaffold40991。

图 2 36 条 Scaffold 中 SSR 基序的数量

Fig.2 Number of simple sequence repeats (SSR) motifs in 36 scaffolds

此外,在检测到的 95 个 gSSR 位点中,基序类型为二核苷酸重复 (Di) 的有 67 个,占 gSSR 位点总数的 70.53%;基序类型为三核苷酸重复 (Tri) 的共有 25 个,占比 26.32% (图 3A);基序类型为四核苷酸重复 (Tetra)、五核苷酸重复 (Penta)、六核苷酸重复 (Hexa) 的 gSSR 位点数很少。在 36 条 Scaffold 中统

计了每种 SSR 基序重复元件数量的分布,在二核苷酸重复类型中,含有 AT & TA、AG & GA 单元的二核苷酸 SSR 数量要明显多于其他类型的二核苷酸重复;在其余核苷酸重复类型中,并未出现以 TG、CG 或 GC 作为重复单元的基序(图 3B)。



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

图 3 每种简单重复序列 (SSR) 基序的数量 (A) 及各二核苷酸重复元件的数量 (B)

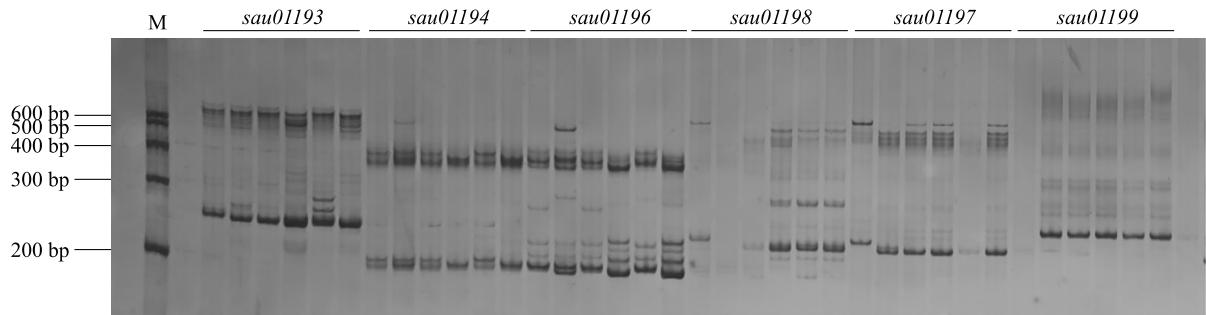
Fig.3 Number of each simple sequence repeats (SSR) motif (A) and dinucleotide repeat element (B)

### 2.4 gSSR 引物筛选及多态性分析

对检测到的 95 个 gSSR 位点设计相应的引物并进行多态性筛选,部分结果如图 4 所示。结果表明,共有 69 对 gSSR 引物能够扩增出清晰稳定的条带,占引物总数的 72.63%;共有 46 对 gSSR 引物在亲本及子代间具有多态性,多态性 gSSR 引物占引物总数的 48.42%。

使用多态性引物在群体中进行扩增,共有 29 对 gSSR 引物扩增产物条带可以判读,经过卡方 ( $\chi^2$ ) 测验检测这 29 对多态性 SSR 引物扩增产物在子代中的分离情况,其中 22 对引物所得产物在子代中分离情况符合孟德尔分离比例 (1:1, 1:2:1, 1:1:1:1:1), 7 对引物所得产物存在不同程度偏分离现象

( $P<0.05$ ), 偏分离标记占多态性引物的比例为 24.13%。29 对引物具体分离情况见表 3。



M 为 600 bp 标准 marker, sau01193、sau01194、sau01196、sau01197、sau01198、sau01199 为部分 gSSR 标记。

图 4 6 对 gSSR 引物在亲本及 4 个  $F_1$  子代中多态性位点的聚丙烯酰胺凝胶电泳结果

Fig.4 Polyacrylamide gel electrophoresis result of polymorphic loci in parental lines and four  $F_1$  progenies using six pairs of gSSR primers

表 3 29 个 gSSR 标记在  $F_1$  代中的分离情况

Table 3 Segregation of 29 gSSR markers in  $F_1$  generation

引物	分离类型	分离比例	卡方值 ( $\chi^2$ )	自由度 ( $Df$ )
sau01170	nnxnp	26 : 16	2.38	1
sau01172	nnxnp	18 : 24	0.86	1
sau01173	nnxnp	32 : 18	3.92	1
sau01174	nnxnp	30 : 23	0.92	1
sau01179	abxcd	17 : 7 : 9 : 13	5.13	3
sau01182	efxeg	9 : 22 : 9 : 12	8.77	3
sau01183	nnxnp	24 : 19	0.58	1
sau01186	lmxll	29 : 26	0.16	1
sau01195	nnxnp	30 : 24	0.67	1
sau01196	efxeg	8 : 12 : 14 : 17	3.35	3
sau01197	nnxnp	29 : 24	0.47	1
sau01201	abxcd	15 : 14 : 10 : 14	1.11	3
sau01202	nnxnp	25 : 30	0.45	1
sau01203	nnxnp	28 : 27	0.02	1
sau01209	lmxll	23 : 27	0.32	1
sau01214	nnxnp	20 : 23	0.21	1
sau01217	nnxnp	24 : 28	0.31	1
sau01218	lmxll	29 : 25	0.30	1
sau01220	lmxll	23 : 26	0.18	1
sau01221	nnxnp	22 : 27	0.51	1
sau01222	nnxnp	35 : 17	6.23	1
sau01225	nnxnp	33 : 22	2.20	1
sau01226	nnxnp	20 : 33	3.19	1
sau01228	nnxnp	32 : 21	2.28	1
sau01229	nnxnp	32 : 16	5.33	1
sau01233	abxcd	3 : 2 : 18 : 30	40.36	3
sau01254	nnxnp	26 : 28	0.07	1
sau01260	efxeg	6 : 19 : 13 : 16	6.89	3
sau01263	nnxnp	26 : 29	0.16	1

$P<0.05$ ; 检验标准:  $df=1, \chi^2=3.84$ ;  $df=3, \chi^2=7.81$ 。

2.5 种间遗传图谱加密

利用 Joinmap 4.1 软件加密黄花菜种间遗传图谱, 共有 291 个 SSR 标记参与运算, 其中包括课题组前期开发的 262 个 EST-SSR 标记以及 29 个新开发的 gSSR 标记。在  $LOD$  值为 3.0 时划分为 9 个连锁群。制图过程中有 13 个标记 (12 个 EST-SSR 标记和 1 个 gSSR 标记) 由于对连锁群结构和标记分布均匀性产生较大影响而舍去, 有 23 个标记 (17 个 EST-SSR 标记和 6 个 gSSR 标记) 由于形成单标记或者双标记连锁群而未能进入最终的图谱。连锁图谱总图距为 3 142.64 cM, 图谱长度较原始种间图谱增加了 657.05 cM, 连锁群 (LG) 长度为 28.26~879.50 cM (表 4)。

表 4 加密后黄花菜与萱草种间遗传图谱各连锁群信息

Table 4 Information of linkage groups of interspecific genetic map for *Hemerocallis citrina* and *H. fulva* after density enhancement

连锁群	标记数量 (个)	长度 (cM)	平均间距 (cM)	最大间距 (cM)
LG1	58	879.50	15.16	117.84
LG2	54	671.64	12.44	150.28
LG3	43	632.46	14.71	101.89
LG4	29	253.83	8.75	58.52
LG5	26	149.11	5.74	31.88
LG6	18	208.00	11.56	73.32
LG7	13	76.19	5.86	16.64
LG8	9	243.65	27.07	127.67
LG9	5	28.26	5.65	14.69

图谱共有 255 个标记,包括 233 个 EST-SSR 标记以及 22 个 gSSR 标记。相邻标记中最大间距大于 80 cM 的间距共有 4 个,分别位于 LG1、LG2、LG3 和 LG8,最大的间距为 150.28 cM,这个间距位于 LG2 的 *sau555*、*sau481* 和 *sau825* 之间(图 5)。图谱平均图距为 11.88 cM。加密后的黄花菜与萱草种间遗传图谱中 SSR 标记在 9 个连锁群上的分布是不均匀的,各连锁群标记数量为 5~58 个,标记数量最多的连锁群为 LG1,共有 58 个标记;LG9 的标记数量最少,为 5 个(表 4)。共有 22 个 gSSR 标记被加密到种间遗传图谱上,其中 7 个 gSSR 标记来源于秋水仙碱代谢通路相关候选 Scaffold,15 个标记来源于昼夜节律代谢通路候选 Scaffold,这些标记位于 LG1、LG2、LG3、LG4 和 LG7 上(图 5)。其中,位于 LG1 的 *sau01228* 和 *sau01182* 在遗传图谱上位于同一遗传距离,2 个标记是由不同性状的候选基因开发而来的,其对应的 Scaffold 序列是不同的,2 个标记在群体中的基因分型也是不同的,*sau01228* 为 nn×np 分型,*sau01182* 为 ef×eg 分型,是不一样的 2 个标记。类似的情况也出现在 LG3 上的 *sau01263*、*sau01260* 和 *sau01254* 这 3 个标记,它们位于不同的 Scaffold,其在群体中的基因分型也是不同的,*sau01263* 为 nn×np 分型,*sau01260* 为 ef×eg 分型,*sau01254* 为 nn×np 分型。

### 2.6 加密后图谱的优化整合

将加密后的图谱与课题组前期构建的框架种间遗传图谱<sup>[11]</sup>进行比较,结果表明,加密前黄花菜与萱草种间遗传图谱具有 11 个连锁群(HFLG1~HFLG11),新标记具有明显的连锁群整合效果,加密后为 9 个连锁群(LG1~LG9),其中,LG1 上由于标记 *sau01229* 与 *sau01183* 的加入使得之前的 HFLG2 与 HFLG10 这 2 个连锁群整合到了一起,HFLG10 整体插入到了 HFLG2 的 *sau722* 与 *sau920* 之间,*sau01179*、*sau01228*、*sau01182*、*sau01203* 和 *sau01214* 这 5 个标记也均匀地插入到了 LG1 上,并且由于这些标记的加入使得之前没有上图的单标记 *sau275* 与 3 标记 *sau677*、*sau215*、*sau979* 也加入到了 LG1 顶端,遗传距离由 HFLG2 的 401.8 cM 和 HFLG10 的 44.4 cM 合并增加到了 879.5 cM(图 6)。LG2 中共增加了 6 个标记,分别为 *sau01174*、*sau01221*、

*sau01173*、*sau01172*、*sau01201*、*sau01202*,这些标记都加入到了 LG2 的偏底端位置,原 HFLG1 上的 *sau272* 因加密后致使 LG2 长度发生较大变化而舍弃。LG3 上新加入了 *sau01263*、*sau01260*、*sau01254*、*sau01218*、*sau01226*、*sau01225* 这 6 个标记,连锁群长度增加了 59.8 cM,另外由于新标记的加入使得 LG3 上部 *sau779*~*sau1017* 片段发生了倒位,原本处于 HFLG3 底端的 *sau937* 移动到了 LG3 顶端。LG4 由于单标记 *sau01233* 的加入使之前 HFLG5 和 HFLG8 合并为 1 个连锁群,长度由 HFLG5 的 141.1 cM 和 HFLG8 的 95.4 cM 增加为 253.8 cM。对 LG5 对应的 HFLG4 上的原标记进行优化调整后,之前未上图的 *sau311*、*sau310*、*sau48*、*sau863*、*sau987*、*sau433*、*sau562*、*sau338* 这 8 个标记也加入到了 LG5 上。LG7 加入了 2 个新标记 *sau01186* 以及 *sau01196*。其余连锁群均与之前保持一致,并未发生改变(图 6)。

## 3 讨论

简单重复序列(SSR)由于重复性高、特异性强等特点已经成为基因定位和遗传多样性研究的首选标记<sup>[15]</sup>,SSR 标记广泛应用于品种鉴定<sup>[1]</sup>、图位克隆<sup>[16]</sup>、系统发育关系研究<sup>[13]</sup>、遗传多样性研究<sup>[17]</sup>、遗传图谱构建<sup>[18]</sup>等。目前,SSR 标记的类型可以大致分为 2 种,一种是通过 RNA-seq 开发的 EST-SSR 标记,EST 是基因表达序列,通过这种方式开发的标记主要对编码区进行扩增,所以在物种间具有很好的通用性,并且使用这些标记构建遗传图谱相当于定位功能已知的基因<sup>[19]</sup>。结合关联分析,如果某性状与图谱上的 EST-SSR 关联,那么这个标记有极大的可能与控制该性状的基因有关<sup>[20]</sup>。另外一种是通过基因组开发的 gSSR 标记,该标记是较为理想的构建图谱的标记,大多数存在于基因组的非编码区<sup>[21]</sup>。

基于此思路,本研究从秋水仙碱和昼夜节律相关代谢通路(ko00950、ko04712)上的候选 Unigene 入手,以期在候选 EST-SSR 附近开发 gSSR 标记来加密该区段。值得一提的是,在筛选昼夜节律相关 Unigene 时,c56440.graph\_c0 表现出与拟南芥中 *LHY* 基因极高的序列相似性,基于序列决定生物学功能的考虑,认为该 Unigene 极可能在本物种中存在 *LHY* 相关基因功能,故亦纳入候选 Unigene 筛选范畴。



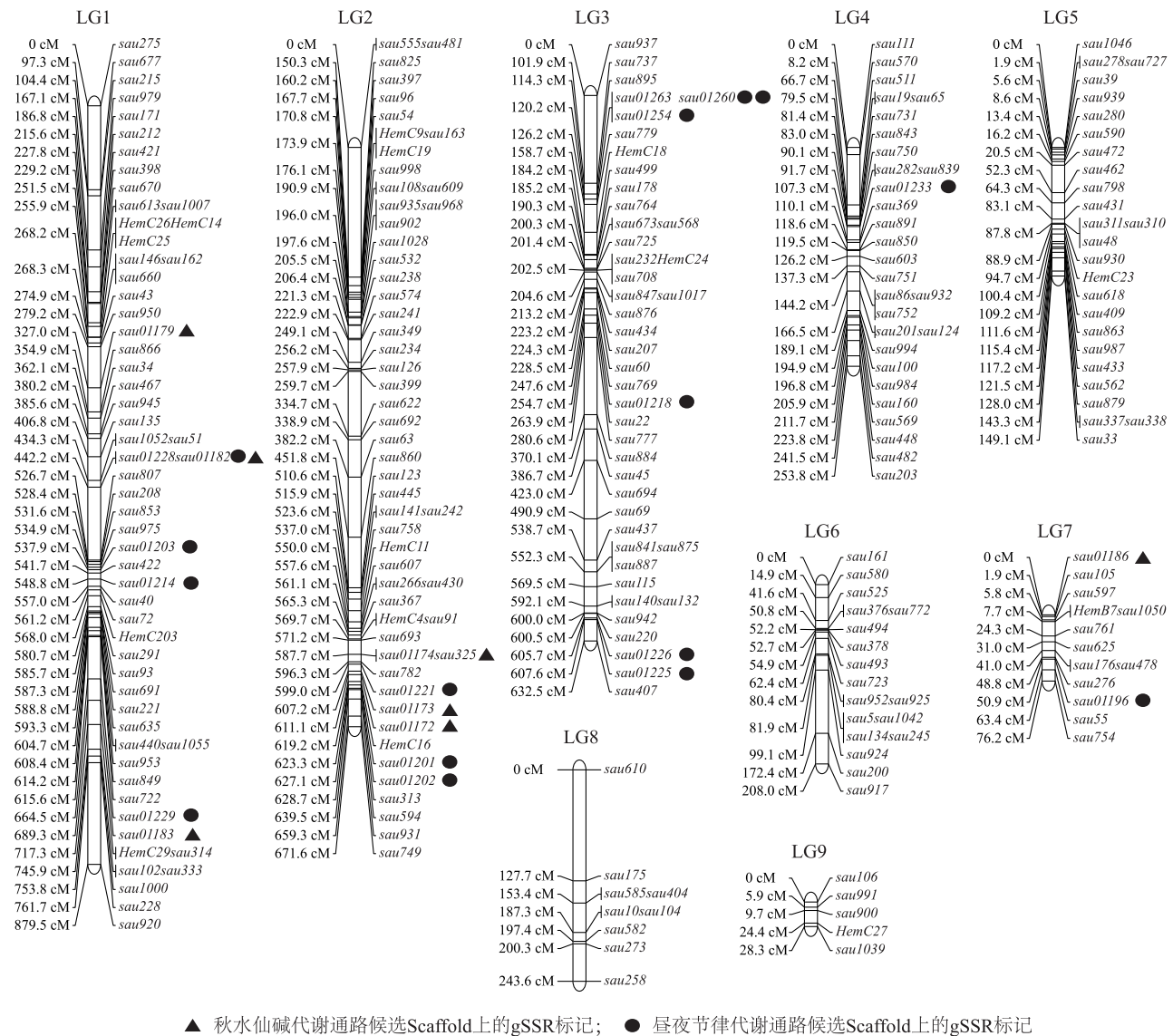


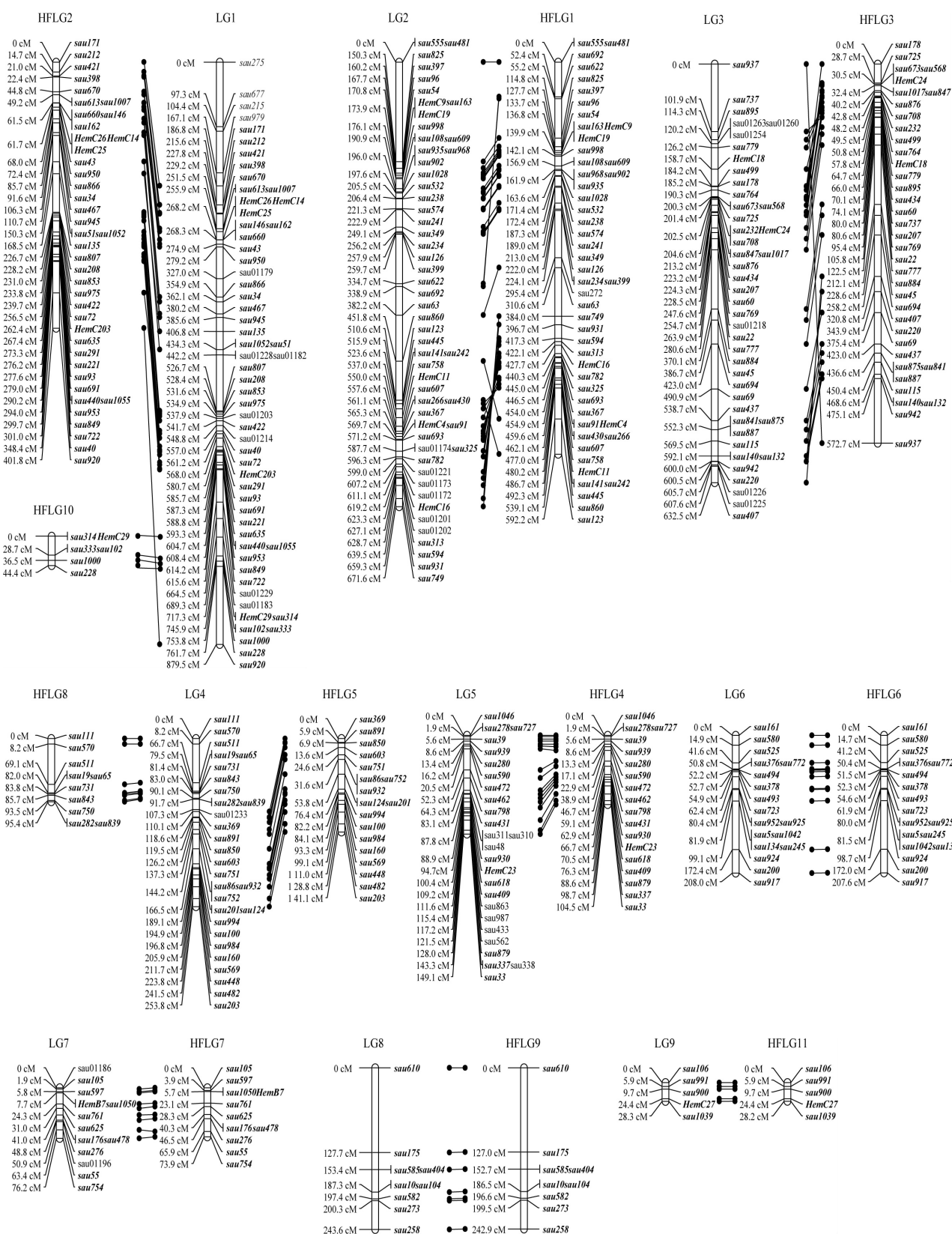
图5 加密后黄花菜与萱草种间遗传图谱

Fig.5 Interspecific genetic map for *Hemerocallis citrina* and *H. fulva* after density enhancement

有学者的研究表明,使用单一类型标记构图会导致标记分布不均匀,标记类型的不同会影响标记在连锁群上的分布<sup>[21-22]</sup>。因此,本研究综合使用新开发的gSSR和前期的EST-SSR 2种标记来进行图谱构建。从结果来看,标记仍然是分布不均匀的,可能原因与标记数量不足以及标记自身的非随机分布有关,此次图谱加密过程仅使用了部分基因组序列来开发gSSR标记,在之后的研究中,基于全基因组SSR标记开发进行图谱加密应该是一个有效的解决方案。此外,由于EST-SSR来自比较保守的编码区,很多研究表明gSSR的多态性要高

于EST-SSR<sup>[15,23]</sup>,本研究中新开发的gSSR标记多态性比例为48.42%,高于侯非凡<sup>[11]</sup>使用的EST-SSR标记(22.49%)。

理论上,分子连锁群的数目应该与染色体数目相等<sup>[24]</sup>,加密后其连锁群数目由11个整合为9个,这可能是由于萱草属植物具有较为庞大的基因组(3.61 Gb),而图5中标记只有255个,难以覆盖整个基因组,在此基础上开发新标记,扩大作图群体,是填补图谱Gap的有效方法,这也是本研究接下来的努力方向,以期黄花菜和萱草分子标记辅助育种、关键性状基因定位及数量性状座位(QTL)分析



连锁群间连线表示图谱加密前后的同源关系,其中相同标记使用粗体表示。

图6 黄花草与萱草种间遗传图谱加密前后比较

Fig.6 Comparisons between the updated and previous interspecific genetic map for *Hemerocallis citrina* and *H. fulva*

奠定理论基础。

## 参考文献:

- [1] LI S, JI F, HOU F, et al. Characterization of *hemerocallis citrina* transcriptome and development of EST-SSR markers for evaluation of genetic diversity and population structure of *Hemerocallis* collection[J]. *Frontiers in Plant Science*, 2020, 11: 686.
- [2] XU P, WANG K Z, LU C, et al. Antidepressant-like effects and cognitive enhancement of the total phenols extract of *Hemerocallis citrina* Baroni in chronic unpredictable mild stress rats and its related mechanism[J]. *Journal of Ethnopharmacology*, 2016, 194: 819-826.
- [3] 杨利. 萱草属植物营养成分分析及品质评价[D]. 长春:吉林农业大学, 2014.
- [4] 詹利生,李贵荣,李少旦,等. 黄花菜中总黄酮的提取及其药理作用初步观察[J]. *南华大学学报(医学版)*, 2005, 33(1): 112-114.
- [5] HOU F, LI S, WANG J, et al. Identification and validation of reference genes for quantitative real-time PCR studies in long yellow daylily, *Hemerocallis citrina* Baroni[J]. *PLoS One*, 2017, 12(3): e0174933.
- [6] 张宁,李森,王金耀,等. 萱草属植物花蕾中秋水仙碱含量 HPLC 检测体系的优化[J]. *河北农业大学学报*, 2017, 40(5): 48-54.
- [7] 王金耀,党换梅. 萱草属植物花蕾中的秋水仙碱 HPLC 测定[J]. *现代园艺*, 2018(11): 33-35.
- [8] 关颖. 中毒患者血清中秋水仙碱的快速检验[J]. *中国卫生检验杂志*, 2001(5): 625-626.
- [9] CUI H, ZHANG Y, SHI X, et al. The numerical classification and grading standards of daylily (*Hemerocallis*) flower color[J]. *PLoS One*, 2019, 14(6): e216460.
- [10] 熊雄,侯非凡,崔虎亮,等. 萱草属植物种间与种内分离群体花器官主要性状的遗传分析[J]. *山西农业科学*, 2019, 47(12): 2062-2067.
- [11] 侯非凡. 黄花菜秋水仙碱合成相关基因筛选及遗传图谱构建[D]. 太谷:山西农业大学, 2017.
- [12] 冀芳芳,李森,史青青,等. 萱草属种质遗传多样性分析及初级核心种质库的构建[J]. *河北农业大学学报*, 2018, 41(2): 55-61.
- [13] ZHANG S, TANG C, ZHAO Q, et al. Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in Foxtail millet [*Setaria italica* (L.) P. Beauv][J]. *BMC Genomics*, 2014, 15(1): 78.
- [14] 冀芳芳. 萱草属植物主要观赏性状的候选基因关联分析[D]. 太谷:山西农业大学, 2018.
- [15] 唐露,金梦雅,黄琳凯,等. 基于 SSR 标记的四倍体鸭茅遗传图谱加密[J]. *中国农业科学*, 2018, 51(5): 991-998.
- [16] LI S, PAN Y, WEN C, et al. Integrated analysis in bi-parental and natural populations reveals *CsCLAVATA3* (*CsCLV3*) underlying carpel number variations in cucumber[J]. *Theoretical and Applied Genetics*, 2016, 129(5): 1007-1022.
- [17] ESER E, TOPCU H, KEFAYATI S, et al. Highly polymorphic novel simple sequence repeat markers from Class I repeats in walnut (*Juglans regia* L.) [J]. *Turkish Journal of Agriculture and Forestry*, 2019, 43(2): 174-183.
- [18] SCHLAUTMAN B, COVARRUBIAS-PAZARAN G, DIAZ-GARCIA L A, et al. Development of a high-density cranberry SSR linkage map for comparative genetic analysis and trait detection[J]. *Molecular Breeding*, 2015, 35(8): 177.
- [19] 姜春芽,廖娇,徐小彪,等. 植物 EST-SSR 技术及其应用[J]. *分子植物育种*, 2009, 7(1): 125-129.
- [20] SCHUBERT R, MUELLERSTARCK G, RIEGEL R. Development of EST-PCR markers and monitoring their intrapopulational genetic variation in *Picea abies* (L.) Karst. [J]. *Theoretical and Applied Genetics*, 2001, 103(8): 1223-1231.
- [21] 姜俊烨. 蚕豆微核心种质构建及 SSR 遗传连锁图谱加密[D]. 北京:中国农业科学院, 2014.
- [22] CASTIGLIONI P, AJMONE-MARSAN P, VAN WIJK R, et al. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution[J]. *Theoretical and Applied Genetics*, 1999, 99(3): 425-431.
- [23] 张亚东,彭婵,李振芳,等. 基因组 SSR 与 EST-SSR 标记在杨树不同种间的遗传差异[J]. *东北林业大学学报*, 2011, 39(12): 8-11.
- [24] 张俊,王荣梅,王金耀,等. 萱草属植物染色体制片技术优化及倍性鉴定[J]. *河北农业大学学报*, 2018, 41(3): 42-48.

(责任编辑:陈海霞)