

水稻紫色种皮基因 *Pb* 的精细定位与候选基因分析

王彩霞 舒庆尧*

(浙江大学原子核农业科学研究所, 农业部及浙江省核农学重点实验室, 杭州 310029. * 联系人, E-mail: qyshu@zju.edu.cn)

摘要 紫米因其种皮内花色素苷的沉积而着色. 控制种皮着色的 *Pb* 基因位于水稻第 4 染色体, 紫色种皮对白色种皮呈显性. 本研究利用培矮 64S(白米) × 豫南黑粳糯(紫米)和培矮 64S × 川黑糯(紫米)两个 F_2 分离群体对 *Pb* 基因进行了精细定位. 首先用 SSR 标记将 *Pb* 基因初步定位在距微卫星位点 RM3820 下游 0.79 cM 的位置. 然后, 通过在候选区域内发展高密度的 InDel 标记和 CAPS 标记, 将 *Pb* 基因限定在两个 InDel 标记 RID3 与 RID4 之间 25 kb 范围以内. 在该区段内, TIGR 水稻基因组(R.5)标有两个注释基因: 一个为与玉米 *Lc* 基因同源的 *Ra*, 为控制花色素苷代谢的 Myc 类转录因子; 另一个为与拟南芥 *TT8* 同源的 *bhlh16*, 也与植物色素代谢有关. 根据两者的性质和前人相关研究结果, 推测 *Pb* 与 *Ra* 实为同一个基因. 序列分析结果表明, 与白米品种培矮 64S, 9311 (粳)和日本晴(粳)相比, 豫南黑粳糯和川黑糯中 *Ra* 基因第 7 外显子存在一个 GT 缺失. 根据 GT 缺失发展了一个 CAPS 标记 CAPSRa, 并对两个 F_2 分离群体和 100 余份水稻品种进行分析. 结果发现, 所有 F_2 白米单株以及所有白米($n = 63$)和红米($n = 23$)品种的 CAPSRa 基因型与培矮 64S 相同, 而所有紫米品种($n = 20$)与豫南黑粳糯和川黑糯相同. 据此推测, 水稻紫色种皮性状可能由 *Ra* 基因第 7 外显子内的 GT 缺失引起.

关键词 水稻 紫色种皮 *Pb* 基因 精细定位 *Ra* 基因

现今全世界范围内栽培和食用的水稻大多为白米. 有色米在少数地区有少量栽培, 以红米与紫米最为常见. 紫米由于水稻种子种皮内花色素苷的积累而呈现紫色, 其花色素苷的主要成分为矢车菊素-3-葡萄糖和芍药素-3-葡萄糖^[1,2]. 研究人员对水稻花色素苷着色已经开展了大量遗传学研究, 花色素苷着色至少涉及色素原基因 *C*, 激活基因 *A* 和器官特异性表达有关的基因 *Pl*^[3]. 紫色种皮是由基因 *Pb* (又称 *Prp-b*)与其互补基因 *Pp* (又称 *Prp-a*)共同作用的结果, 两者分别位于第 4 染色体和第 1 染色体上^[4,5]. 当 *Pb* 基因不存在时, 种皮呈白色; 当 *Pb* 基因单独存在时, 种皮呈棕色; 当基因 *Pb* 和 *Pp* 同时存在时, 种皮呈紫色.

Hu 等人^[6]以控制玉米花色素苷合成的调节基因 *Lc* 的 cDNA 为探针, 从紫叶水稻品种 Purple 522 的叶片 cDNA 库中分离出具有 bHLH 结构、与 *Lc* 同源的 *Ra* 和 *Rb* 基因, 并将 *Ra* 和 *Rb* 基因分别定位在第 4 和第 1 染色体上. 此后, 他们又发现 *Ra* 实际由 *Ra1* 和 *Ra2* 两个基因组成^[7]. Sakamoto 等人^[8]利用相似的技术, 用玉米 *B-Peru* cDNA 确定了水稻紫叶等位基因 *Pl^w* 位点由至少两个基因 *OSB1* 和 *OSB2* 组成. 虽然 Hu 等人^[6]认为 *Ra* 与 *Pl* 并非等位基因, 但 Sakamoto

等人^[8]的结果认定 *OSB1* 与 Hu 等人发现的 *Ra* 基因为同一个基因. 庄杰云等人^[9]已将 *Pb* 基因定位在水稻第 4 染色体距 RG214 26.3 cM, RG329 18.9 cM 的位置, 但未见对 *Pb* 基因有进一步研究. 为了明确 *Pb* 基因在染色体上的确切位置, 为克隆和研究 *Pb* 基因在水稻紫米形成以及花色素苷生物合成中的作用奠定基础, 本研究利用两个杂交群体对控制紫米种皮颜色性状的 *Pb* 基因进行了精细定位, 并对候选基因进行了深入分析.

1 材料与与方法

() 定位群体构建. 用于本研究的两份种皮呈紫色的水稻材料分别为粳稻品系“豫南黑粳糯”及粳稻品系“川黑糯”, 2004 年冬至 2005 年春在海南以两系不育系培矮 64S 为母本, 于花期将其与上述材料杂交获得培矮 64S × 豫南黑粳糯和培矮 64S × 川黑糯 F_1 种子. 2005 年夏将 F_1 种子单本种植, 单株收获得 F_2 种子. F_2 种子按株系于 2006 年夏单本种植, 于抽穗期单株挂牌并取倒二叶用于基因组 DNA 提取. 单株收获 F_3 种子, 编号与叶片相对应. 所有世代的种子待完全成熟后, 晒干, 经去壳辨认种皮颜色.

() DNA 提取. 基因组 DNA 的提取以叶片为

2007-07-18 收稿, 2007-09-18 接受

材料,参照卢杨江和郑康乐^[10]的方法进行.用紫外分光光度计测定 DNA 样品的浓度,并将终浓度调整至 25 ng/ μ L.

() 分子标记.先用均匀分布于第 4 染色体的 27 个 SSR 标记^[11]进行 *Pb* 基因的初步定位.经初步定位后,合成候选区域内的新 SSR 标记(www.gramene.org)用于进一步分析.在此基础上,开发该区域内的 InDel 和 CAPS 标记用于 *Pb* 基因的精细定位.对于 InDel 标记,从 Shen 等人^[12]开发的水稻 DNA 多态性数据库(<http://shenghuan.shnu.edu.cn/genefunction/ricemarker.htm>)获取候选区域内的 InDels 数据.挑选其中差异大于 7 bp 的 InDel,在最新的日本晴和 9311 序列数据之间用 BioEdit 进行比对确认.选择大小合适(InDel 差异大小及产物大小的比例足以保证 InDel 差异能以 3%琼脂糖凝胶电泳或 8%聚丙烯酰胺凝胶电泳分离显示)且分布均匀的 InDel,以其左右各 500 bp 的日本晴基因组序列为基础设计引物开发 InDel 标记. CAPS 标记则是根据更小候选区域内的日本晴和培矮 64S 序列(由北京华大基因组研究所提供),分析两者之间的酶切位点差异,以差异位点左右各 800 bp 的日本晴基因组序列为基础设计引物开发 CAPS 标记,并保证 PCR 产物和酶切产物大小合适以利于电泳检测.引物采用 Primer Premier 5 进行设计并在 TIGR 数据库(<http://tigrblast.tigr.org/euk-blast/index.cgi?project=osa1>)比对检查引物的特异性.所用引物均由上海生工生物工程技术服务有限公司合成.

PCR 扩增体系总体积为 20 μ L,包括 50 ng 模板 DNA, 1 \times PCR 缓冲液, 0.2 μ mol/L 引物, 0.2 mmol/L dNTP, 2 mmol/L MgCl₂ 和 1 U *Taq* 酶. SSR 标记的扩增程序参照 Temnykh 等人^[11]的方法. InDel 和 CAPS 标记的扩增程序为: 94 预变性 5 min; 94 变性 45 s, 55 退火 45 s, 72 延伸 90 s, 37 个循环; 72 延伸 10 min. 对于个别引物需根据具体情况调节退火温度和延伸时间.对于 CAPS 标记,采用相应的限制性内切酶对 PCR 产物进行酶切.限制性内切酶购自上海生工生物工程技术服务有限公司,酶切反应按照说明书进行.

SSR 标记和 InDel 标记的 PCR 扩增产物根据差异大小分别用 3%琼脂糖凝胶电泳或 8%聚丙烯酰胺凝胶电泳进行分离. CAPS 标记的 PCR 产物和酶切产物用 1.2%琼脂糖凝胶电泳分离检测.

() 初步定位. 选用 F₂ 群体中种皮呈白色的

10 个单株建 DNA 池,并与双亲一起用均匀分布于第 4 染色体的 27 个 SSR 标记筛选与 *Pb* 基因连锁的候选 SSR 标记.然后利用培矮 64S \times 豫南黑粳和培矮 64S \times 川黑糯 F₂ 群体中的白色种皮单株分析候选 SSR 标记与种皮颜色性状的共分离情况.在连锁标记附近合成新的 SSR 标记并分析其多态性及连锁分离情况.

() 精细定位. 经初步定位后,用候选区域内在双亲间具有差异 InDel 标记和 CAPS 标记进一步分析培矮 64S \times 豫南黑粳和培矮 64S \times 川黑糯 F₂ 群体中白色种皮单株的重组情况.与豫南黑粳和川黑糯相同大小的带记为“1”,与培矮 64S 相同大小的带记为“0”.对于 F₂ 群体白色种皮单株,存在 3 种可能的带型,即未发生交换“0/0”,单交换“0/1”以及双交换“1/1”.

() *Ra* 基因序列分析. 根据 Gramene 数据库中 *Ra* 基因序列(LOC_Os04g47080),用 Primer Premier 5 软件设计 9 对引物对豫南黑粳和川黑糯基因组 DNA 进行扩增,并使相邻扩增片段两端重叠 120~200 bp,以便于产物的拼接. *Taq*Plus 酶(上海生工生物工程技术服务有限公司)扩增产物用 1.2%琼脂糖凝胶电泳分离,经纯化试剂盒(AP-GX-250, Axygen)纯化后送往上海英骏生物技术有限公司测序.利用软件 ContigExpress 对测序片段进行序列拼接,并用 ClustalX 和 BioEdit 进行序列比对分析.

() CAPSRa 标记分析. 根据 *Ra* 基因序列设计引物:正向, 5'-CGTCCATTACAGGGTA-3';反向, 5'-CAGCAGATGAGGCAAAC-3'.其基因组扩增产物为大小约 858 bp 的片段,位于 *Ra* 基因第 7 外显子内.这段序列内包含 2 bp (GT)的 InDel, GT 的缺失产生 *Bam*H 酶切位点,经酶切后生成两段大小分别约 653 和 203 bp 的片段.该标记命名为 CAPSRa.

分析 CAPSRa 在培矮 64S \times 豫南黑粳和培矮 64S \times 川黑糯双亲以及 F₂ 群体白色种皮单株中的 PCR 产物和酶切产物大小差异.同时,分析 CAPSRa 在 106 份水稻品种间的差异情况,其中 63 份材料的种皮为白色, 23 份材料的种皮为红色, 20 份材料的种皮为紫色.

2 结果与分析

2.1 初步定位

经基因池与双亲之间的多态性筛选后发现,在位于第 4 染色体上的 27 个 SSR 标记中,分别有 6 个和 9

个 SSR 标记在培矮 64S/豫南黑粳糯和培矮 64S/川黑糯之间具有差异. 在培矮 64S × 豫南黑粳糯 F₂ 分离群体的 192 个白色种皮单株中, RM252 与 *Pb* 基因呈共分离, 两者的遗传距离约为 6.51 cM. 在培矮 64S × 川黑糯 F₂ 分离群体的 190 个白色种皮单株中, RM252 与 *Pb* 基因也表现共分离, 遗传距离为 8.68 cM.

Gramene 数据库 (<http://filetta.cshl.org/multi/blastview>)对 RM252 在染色体上的具体位置并未标注. 为此, 从 NCBI 数据库获取 RM252 的序列(<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=AF344072>), 并用该序列比对后得知, RM252 的物理位置在第 4 染色体 25144688~25145368 bp 之间 (TIGR, psuedomolecule v4). 在 RM252 附近合成新的 SSR 标记并分析了其在培矮 64S 与两个紫米品系间的多态性, 结果未找到在培矮 64S 与豫南黑粳糯间存在差异的标记. RM252 下游标记 RM17231, RM1354, RM3288 和 RM3820 在培矮 64S 与川黑糯间具有多态性; 单株分析结果表明, 它们与 *Pb* 基因间的重组值分别为 7.11%, 4.74%, 1.58%和 0.79%. 根据以上结果, 初步确定 *Pb* 基因位于第 4 染色体 RM3820 (TIGR, psuedomolecule v4: 27589809~27589903)下游附近.

2.2 InDel 标记与 CAPS 标记开发

根据初步定位的结果, 在 RM3820 下游附近继续

寻找其他 SSR 标记. 但该区域已开发的 SSR 标记较少, 且未找到在两个定位群体双亲间具有差异的 SSR 标记. 为了将 *Pb* 基因确定在更小的范围内, 在 27.62~28.02 Mb 范围内合成了 22 个 InDel 标记和 5 个 CAPS 标记. 在这 22 个 InDel 标记中, 其中 17 个标记在培矮 64S, 豫南黑粳糯及川黑糯三者之间表现出片段大小的多态性. 通过电泳呈现的条带, 这些多态性表现为, 所有 17 个具有多态性的 InDel 标记, 豫南黑粳糯和川黑糯具有相同大小的片段, 而培矮 64S 呈现大小异于两者的片段(表 1).

此外, 在 27.72~27.73 Mb 范围内合成的 5 个 CAPS 标记中, 只有 2 个标记(CAPS1 和 CAPS2)在培矮 64S、豫南黑粳糯及川黑糯之间具有差异. 三者的 PCR 产物大小相同, 其中培矮 64S 的 PCR 产物能被相应的酶切开, 而豫南黑粳糯和川黑糯的 PCR 产物不能被切开(表 1).

2.3 精细定位

利用 17 个 InDel 标记和 2 个 CAPS 标记对 *Pb* 基因进行了精细定位. RID16 和 RID17 为上述标记中最靠近染色体末端的标记, 因此先分析了两个 F₂ 分离群体中的所有白色种皮单株上这两个标记的分离情况. 结果表明, 在培矮 64S × 豫南黑粳糯的 192 个 F₂

表 1 用于 *Pb* 基因精细定位的 InDel 标记和 CAPS 标记

标记	位置 ^{a)}	引物序列 (5' 3')		片段大小 ^{b)/bp}		
		上游	下游	豫南黑粳糯	川黑糯	培矮 64S
RID1	27623211	GTGGCTCTTTCCGTTTC	TGTGCTCCTCATACAAAC			
RID2	27654641	CATCCAAGATCATATTACA	TCTCCCAGTCCCAGGTC	271	271	254
RID3	27712299	GCTCAAGCACGCCACTC	CGTCCCTTATGCCTCCA	360	360	403
RID4	27737750	ATTGAAACGGAGGGAGT	CTGTACGTGATGGCAA	180	180	158
RID5	27761483	GTGTATTTAAGGTTTCGAA	TTGTTTCATCTGTAGCTTTGT	153	153	145
RID6	27764264	GGCTGAGTCAAAGAAGGG	TTTTCTACGTGGCGACC	143	143	136
RID7	27772442	CAGCATCTGACACTGGA	CTACTGCAACCTAATCTAA	160	160	171
RID8	27787666	CCGCCAAATACGTTCTG	TTCTCCCGTGCTTCTG	131	131	117
RID9	27793566	TCGGGCAACTCACCTCC	GCGGGCCTCCTCCTTTA	179	179	170
RID10	27821277	ATGGGCTTGATCGGTAG	TTGGGTTATTTGAGTGCTAA	198	198	210
RID11	27827101	CACATCTCGCTAAGGAAA	CCTATGCAATTCGGTCA	299	299	279
RID12	27851717	CCGCTGAGCTTGAAAGT	CGGAGGGAGTAGAGTAATG	290	290	266
RID13	27879300	GAGATAGGCTCACTTCG	ACATACATACGTGTATAAACT	295	295	279
RID14	27883558	CAAGTTACATAGGGCTCG	GGTGGAGGAGTTCATTCT	488	488	424
RID15	27932108	GAGTGAACAGAAAGGGAC	CAACTACAAGCGTACAAAC	280	280	302
RID16	27947572	CAGGGTTGCGTCGAGAT	TTTCAAAGGTGCTTATAGAG	393	393	435
RID17	28020277	GTTCAAGACCTTCCCATCA	CCTACAAGAGCCCAGCA	283	283	256
CAPS1 (<i>Sma</i>)	27722111	CGCATCGTCAATTTCTG	GGCAACAAGGTAGGCACT	399	399	151/248
CAPS2 (<i>Xba</i>)	27728667	AGCATTATCCCTGAGTGT	GACTGTAGAGGGCAAGAC	840	840	562/278

a) InDels 或 SNPs 差异起始位点在第 4 染色体上的位置(TIGR, psuedomolecule v4); b) 片段大小根据日本晴和 9311 基因组序列以及电泳条带推测

白色种皮单株中, RID17 和 RID16 均有 4 个重组型的单株. 在培矮 64S × 川黑糯的 190 个 F₂ 白色种皮单株中, RID17 为重组型的有 7 个, RID16 处则为 6 个. 据此推测, *Pb* 基因位于 RM3820 和 RID16 之间(图 1(a)).

接下来对两个 F₂ 群体白色种皮单株在位于 RM3820 和 RID16 之间的 15 个 InDel 标记的重组情况进行了分析. 如表 2 所示, 在培矮 64S × 豫南黑糯 F₂ 群体中, 有 1 个单株(Y628)的 RID1 发生单交换, 而 RID15 发生单交换的有 4 个单株(Y214, Y508, Y586, Y712). 逐步向中间推移, 下游有 1 个单株(Y508)在标记 RID10 发生了单交换, 但未发现更近的发生重组的上游标记. 在培矮 64S × 川黑糯 F₂ 群体中, 有 2 个单株(C104, C176)的 RID1 发生了单交换, 有 6 个单株(C217, C271, C325, C513, C830, C913)的 RID15 发生了单交换. 逐步向中间推移, 有 1 个单株(C176)在上游标记 RID3 发生了单交换, 1 个单株

(C325)在下游标记 RID4 发生了单交换(表 2). 根据上述结果, 将 *Pb* 基因定位于第 4 染色体 RID3 和 RID4 之间 25451 bp 范围以内(TIGR, psuedomolecule v4: chr4: 27712299~27737750; 图 1(b), 表 2).

在水稻第 4 染色体 TIGR, psuedomolecule v4: 27712299~27737750 范围内涉及两个已有注释的基因, 分别为 *bhlh16* (LOC_Os04g47059, 27696505~27720731) 和 *Ra* (LOC_Os04g47080, 27730292~27736548). TIGR 对于两者的注释均为 Myc 类转录因子, *bhlh16* 编码与 TRANSPARENT TESTA 8(TT8)相似的蛋白, TT8 是一个参与拟南芥种皮中类黄酮物质代谢的调节蛋白^[13]; *Ra* 编码是与 Lc 相似的蛋白, Lc 参与玉米花色素苷的代谢调节^[14]. 由于两者都为可能的候选基因, 我们试图用 27.72~27.73 Mb 范围内的标记 CAPS1 和 CAPS2 排除其中之一, 但在两个分离群体中, 均未找到在 CAPS1 和 CAPS2 标记处发生重组的单株.

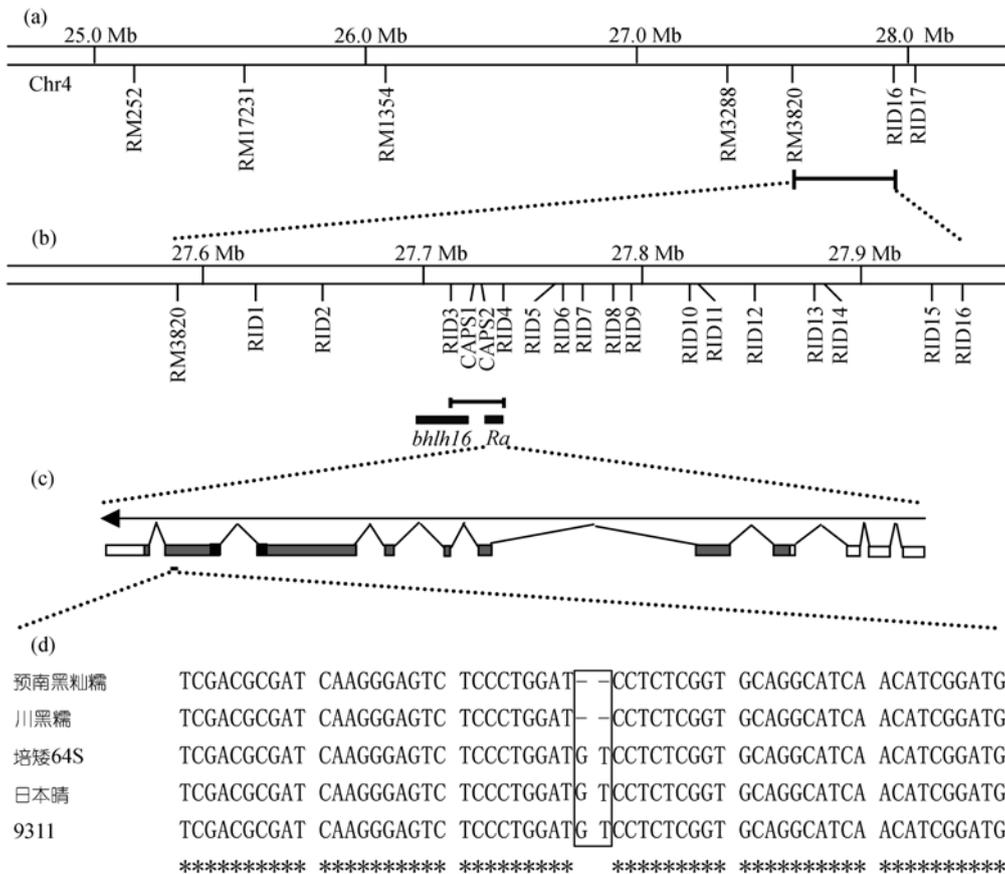


图 1 *Pb* 基因精细定位与候选基因 *Ra* 的序列分析

(a) 将 *Pb* 基因定位在 RM3820 和 RID16 之间; (b) 将 *Pb* 基因限定在 RID3 和 RID4 之间; (c) *Ra* 基因, 灰色部分为外显子, 第 6 外显子末端和第 7 外显子前端的黑色部分为 bHLH 结构域; (d) *Ra* 基因序列分析, 与培矮 64S (Pei'ai 64S)、日本晴(Nipponbare)和 9311 相比, 紫米材料豫南黑糯(YNHXN)和川黑糯(CHN)在靠近第 7 外显子末端处具有 2 bp(GT)缺失

表 2 用于 *Pb* 基因精细定位的重组 F₂ 白色种皮单株不同标记的基因型^{a)}

标记	培矮 64S×豫南黑粳糯 ^{b)}					培矮 64S×川黑糯 ^{b)}								
	Y628	Y214	Y508	Y586	Y712	C104	C176	C217	C271	C325	C513	C830	C913	
RID1	0/1	0/0	0/0	0/0	0/0	0/1	0/1	0/0	0/0	0/0	0/0	0/0	0/0	
RID2	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	
RID3	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	
CAPS1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	
CAPS2	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	
RID4	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	
RID5	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	
RID6	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/0	0/0	
RID7	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/0	0/0	0/0	
RID8	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0	
RID9	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/0	0/0	
RID10	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/1	0/0	
RID11	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/1	0/0	0/1	0/1	0/1	0/1	
RID12	0/0	0/1	0/1	0/1	0/0	0/0	0/0	0/1	0/1	0/1	0/1	0/1	0/1	
RID13	0/0	0/1	0/1	0/1	0/1	0/0	0/0	0/1	0/1	0/1	0/1	0/1	0/1	
RID14	0/0	0/1	0/1	0/1	0/1	0/0	0/0	0/1	0/1	0/1	0/1	0/1	0/1	
RID15	0/0	0/1	0/1	0/1	0/1	0/0	0/0	0/1	0/1	0/1	0/1	0/1	0/1	

a) “0”表示与白米培矮 64S 相同的带型, “1”表示与紫米豫南黑粳糯或川黑糯相同的带型; b) 用于精细定位的培矮 64S×豫南黑粳糯和培矮 64S×川黑糯的 F₂ 白米单株总数分别为 192 和 190 株

2.4 *Ra* 基因序列分析

我们起初试图通过对黑米的这两个候选基因的测序, 找到可能与紫色种皮相关的 DNA 差异。但是, 在水稻基因组中存在大量与 *bhlh16* 序列高度同源的区域, 因此, 无法设计合适的引物, 直接用基因组 DNA 通过 PCR 扩增获得完整的目标产物用于测序。此外, 由于 *Ra* 是一个与花色素苷代谢相关的基因, 而 *bhlh16* 主要与原花色素合成有关, 因此, 我们认为 *Ra* 与水稻种皮中花色素苷合成的关系更为密切。为此我们对 *Ra* 全长进行了测序, 豫南黑粳糯和川黑糯 GenBank 序列号分别为 EU095985 和 EU095986。序列比对发现, 豫南黑粳糯和川黑糯的 *Ra* 基因序列与培矮 64S、日本晴和 9311 的相应序列存在诸多差异, 但在紫米和白米之间只有位于第 7 外显子内的 2 bp (GT) 的插入/缺失存在一定规律, 即白色种皮的培矮 64S、日本晴和 9311 序列含有 GT, 而紫色种皮的豫南黑粳糯和川黑糯序列缺失 GT(图 1(c)和(d))。

2.5 CAPSRa 标记分析

为分析上述 GT 差异在杂交后代单株和其他品种中是否与种皮颜色有关, 根据 GT 的缺失产生 *Bam*H 酶切位点的特性设计了一个 CAPS 标记 CAPSRa。分析结果为: 所有材料该标记的 PCR 产物为 858 bp 左

右; 豫南黑粳糯和川黑糯的 PCR 产物均能被 *Bam*H 酶切成两段, 酶切产物分别约 653 和 203 bp, 而培矮 64S 的 PCR 产物不能为 *Bam*H 酶切(图 2(a)); 培矮 64S×豫南黑粳糯和培矮 64S×川黑糯 F₂ 分离群体中, 白色种皮单株均不能被切动, 呈一条带; 紫色种皮单株可以被切开(图 2(b))。

在此基础上, 我们还对 106 份水稻材料进行了 CAPSRa 标记的差异分析。结果显示, 所有材料的 PCR 扩增产物电泳条带大小相似, 但 63 份白米材料

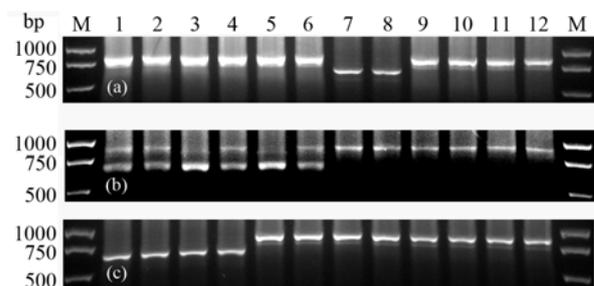


图 2 CAPSRa 标记分析

(a) 1 和 7 为豫南黑粳糯, 2 和 8 为川黑糯, 3 和 9 为培矮 64S, 4 和 10 为日本晴, 5 和 11 为舟山红米, 6 和 12 为红花米; 1~6 为 PCR 产物, 7~12 为酶切产物。(b) F₂ 分离群体。1~6 为纯合或杂合紫色种皮单株, 7~12 为白色种皮单株; 均为酶切产物。(c) 1~4 为紫米品种; 5~8 为白米品种; 9~12 为红米品种; 均为酶切产物。M 为 DNA 片段大小标记, 左边为相应条带的碱基对数

和 23 份红米材料均不能被 *Bam*H 酶切, 而 20 份紫米材料均能被切开(图 2(c)).

3 讨论

虽然前人已经对不同植物中参与花色素苷生物合成的基因进行了广泛研究, 但对水稻紫米种皮色素的遗传机制尚缺乏了解. 本研究将控制该特性的基因 *Pb* 限定在仅有两个注释基因的基因组区段内, 并发现该基因极可能与 Sakamoto 等人^[8]报道的紫叶等位基因 *Pl^w* 位点上的其中一个基因(*OSB1*)及 Hu 等人^[7]报道的与玉米 *Lc* 基因同源的水稻基因之一 *Ra* 为同一基因. 在此基础上, 我们还发现紫色种皮特性可能与该基因的一个 GT 缺失有关. 因此, 本研究结果将对植物花色素苷的合成与调控, 特别是对水稻紫米种皮颜色性状的分子遗传学基础的认识起到重要作用.

籼稻 9311 和粳稻日本晴全基因组序列测定的完成, 特别是基于籼稻 9311 和粳稻日本晴基因组序列的 InDels 和 SNPs 被大量地用于设计分子标记, 为通过基因定位最终克隆目标基因打下了良好的基础^[12,15,16]. 庄杰云等人^[9]报道, *Pb* 基因与 RFLP 标记 RG329 相距 18.9 cM; 目前已知 RG329 在第 4 染色体的位置为 30819286~30819457, 以 244 kb/cM 计算^[17], 控制黑珍珠米种皮颜色的基因大概在 26.21 Mb 的位置. 本研究的定位结果与其相差约 1.5 Mb, 造成这一差异的原因可能有二: () 可能是本研究所用的紫米材料与庄杰云等人^[9]所用的紫米品种“黑珍珠”控制着色的基因本身不同. “黑珍珠”系组培诱发的体细胞突变^[18], 而本研究所用的两个品系为传统水稻品系; () 可能是由于定位时精度不高以及不同区段的物理距离/遗传距离的比例因存在重组热点/冷点差异而导致由重组值推算的物理距离与实际情况存在较大偏差. 如在本研究中, 对 *Pb* 基因进行初步定位时, 与 RM252 的遗传距离在所用两个群体间相差达 2 cM. 因此, 本研究与庄杰云等人^[9]的结果的这种差异也可能并不代表控制黑珍珠米种皮颜色的基因与本研究的两个紫米品系存在差异.

对于花色素苷的代谢, 已在玉米、矮牵牛、金鱼草等植物中展开了大量研究. 参与花色素苷代谢的调节因子主要有 3 类, 分别为 Myb, Myc 和 WD40. Myc 类转录因子具有碱性螺旋-环-螺旋(basic helix-loop-helix, bHLH)结构, 该结构与 Myb 相互作用而调节花色素苷的代谢^[19,20]. 作物基因组测序工作完成

后, 在一个物种中被克隆的重要性状基因, 就可能可以运用比较基因组学技术或其他相关技术在其他物种中寻找目标基因. 在水稻中, Hu 等人^[6,7]根据玉米 cDNA 找到了 3 个与 *Lc* 基因同源的水稻花色素合成调控基因 *Ra1*, *Ra2* 和 *Rb* 基因; Sakamoto 等人^[8]发现了与玉米 *B-Peru* 同源的两个基因 *OSB1* 和 *OSB2*. *Lc* 基因和 *B-Peru* 基因均属于玉米 *R/B* 基因家族成员, 属于 Myc 类转录因子. 同时, Saitoh 等人^[21]在水稻中发现了与玉米 *C1* 基因同源的水稻基因 *OsC1*, 该基因属于 Myb 类转录因子. 在本研究中, 将 *Pb* 基因定位在水稻第 4 染色体约 25 kb 的范围内, 而在该范围内刚好存在同属 Myc 家族的 *Ra* 和 *bhlh16* 两个基因. 由于 *Ra* 基因在玉米中参与花色素合成的调控, 并已在水稻中证明可以激活花色素苷合成^[6,8]; 对于 *bhlh16*, 与其相似的 *TT8* 在拟南芥中与原花色素(proanthocyanidins)合成相关, 可以调节类黄酮结构基因的特异性表达^[22]. 因此, 如果只从其一般功能考察, 很难排除这两个基因中的任何一个.

Pl 位点是控制水稻叶片组织花色素合成的 6 个调节因子之一, 其中等位基因 *Pl^w* 不但可以使除了节间之外所有气生组织着色, 而且还表现紫色种皮^[8]. 本研究所用的两份紫米材料虽然与 Sakamoto 等人^[8]使用的携带 *Pl^w* 的材料 T65-Plw 不同, 其叶片不表现紫色. 但是, 由于两个候选基因之一 *Ra* 基因与 Sakamoto 等人^[8]鉴定出的 *Pl^w* 的其中一个基因 *OSB1* 实为同一基因. 因此, 我们倾向 *Pb* 基因更有可能为 *Ra* 基因. 为此, 我们对 *Ra* 基因进行了测序分析, 发现两个紫米亲本与白米品种(培矮 64S, 9311 和日本晴)之间存在一个 2 bp (GT) 的差异; 并根据测序结果发展了 CAPS 标记 CAPSRa. 虽然在两个杂交群体中 CAPSRa 与种皮特性完全共分离在意料之中, 因为该标记本身就在基因之内, 但对 100 余份其他品种所做分析得到的结果充分说明, 该基因 GT 的缺失与紫米的产生存在密切关系, 因为所有紫米存在这一特性, 而所有白米和红米没有. 因此, 这一结果也说明, *Ra* 较 *bhlh16* 更有可能是调控水稻紫色种皮这一特性的基因.

致谢 定位所用材料豫南黑粳和川黑糯分别由包劲松博士和吴殿星博士(浙江大学原子核农业科学研究所)提供. CAPSRa 标记分析所用白米材料由沈圣泉博士(浙江大学原子核农业科学研究所)提供, 红米与紫米材料来自中国水稻研究所国家水稻种质资源库. 部分培矮 64S 序列由北

京华大基因研究中心刘国振博士提供。在此一并表示感谢。

参 考 文 献

- 1 Fossen T, Slimestad R, Ovstedal D O, et al. Anthocyanins of grasses. *Biochem Syst Ecol*, 2002, 30: 855—864[DOI]
- 2 Abdel-aal E S M, Young J C, Rabalski I. Anthocyanin composition in black, blue, pink, purple, and red cereal grains. *J Agric Food Chem*, 2006, 54: 4696—4704[DOI]
- 3 Nagao S, Takahashi M. Genetical studies on rice plants. Trial construction of twelve linkage groups in Japanese rice. *J Fac Agri Hokkaido Univ*, 1963, 53: 72—130
- 4 Causse M A, Fulton T M, Cho Y G, et al. Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics*, 1994, 138: 1251—1274
- 5 Yoshimura A, Ideta O, Iwata N. Linkage map of phenotype and RFLP markers in rice. *Plant Mol Biol*, 1997, 35: 49—60[DOI]
- 6 Hu J P, Anderson B, Wessler S R. Isolation and characterization of rice *R* genes: Evidence for distinct evolutionary paths in rice and maize. *Genetics*, 1996, 142: 1021—1031
- 7 Hu J P, Reddy V S, Wessler S R. The rice *R* gene family: Two distinct subfamilies containing several miniature inverted-repeat transposable elements. *Plant Mol Biol*, 2000, 42: 667—678[DOI]
- 8 Sakamoto W, Ohmori T, Kageyama K et al. The purple leaf (PL) locus of rice: *Pl^w* allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis. *Plant Cell Physiol*, 2001, 42: 982—991[DOI]
- 9 庄杰云, 杨长登, 钱惠荣, 等. 紫米基因与 RFLP 标记的连锁分析. *遗传学报*, 1996, 23: 372—375
- 10 卢杨江, 郑康乐. 提取水稻 DNA 的一种简易方法. *中国水稻科学*, 1992, 6: 47—48
- 11 Temnykh S, DeClerck G, Lukashova A, et al. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Res*, 2001, 11: 1441—1452[DOI]
- 12 Shen Y J, Jiang H, Jin J P, et al. Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiol*, 2004, 135: 1198—1205[DOI]
- 13 Nesi N, Debeaujon I, Jond C, et al. The *TT8* gene encodes a basic helix-loop-helix domain protein required for expression of *DFR* and *BAN* genes in *Arabidopsis* Siliques. *Plant Cell*, 2000, 12: 1863—1878[DOI]
- 14 Ludwig S R, Habera L F, Dellaporta S L, et al. *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the myc-homology region. *Proc Natl Acad Sci USA*, 1989, 86: 7092—7096[DOI]
- 15 Ji Q, Lu J F, Chao Q, et al. Delimiting a rice wide-compatibility gene *S5n* to a 50 kb region. *Theor Appl Genet*, 2005, 111: 1495—1503[DOI]
- 16 Chen S H, Wu J, Yang Y, et al. The *fgr* gene responsible for rice fragrance was restricted within 69 kb. *Plant Sci*, 2006, 171: 505—514[DOI]
- 17 Chen M S, Presting G, Barkazuk W B, et al. An integrated physical and genetic map of the rice genome. *Plant Cell*, 2002, 14: 537—545[DOI]
- 18 赵成章, 戚秀芳, 杨长登, 等. 应用细胞工程技术培育“黑珍珠”的研究. *农业生物技术学报*, 1993, 1: 100—105
- 19 Chandler V L, Radicella J P, Robbins T P, et al. Two regulatory genes of maize anthocyanin pathway are homologous: Isolation of *B* utilizing *R* genomic sequences. *Plant Cell*, 1989, 1: 1175—1183[DOI]
- 20 Ludwig S R, Wessler S R. Maize *R* gene family: Tissue-specific helix-loop-helix protein. *Cell*, 1990, 62: 849—851[DOI]
- 21 Saitoh K, Onishi K, Mikami I, et al. Allelic diversification at the *C* (*OsC1*) locus of wild and cultivated rice: nucleotide changes associated with phenotypes. *Genetics*, 2004, 168: 997—1007[DOI]
- 22 Baudry A, Caboche M, Lepiniec L. *TT8* controls its own expression in a feedback regulation involving *TTG1* and homologous *MYB* and *bHLHL* factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant J*, 2006, 46: 768—779[DOI]