

1个水稻产量 QTL 区段及品系特异性 重复序列的比较分析

夏 鹏 林海艳 罗美中

华中农业大学作物遗传改良国家重点实验室, 武汉 430070

摘要 为克隆 1 个水稻产量 QTL, 对遗传作图亲本珍汕 97 和明恢 63 物理图谱上与该 QTL 相对应的区段进行了比较分析, 关闭了 2 个物理图谱在该区段的物理缺口, 对该区段珍汕 97 和明恢 63 BAC 末端序列分别与日本晴参考序列之间的预测单核苷酸多态性 SNP 位点进行了检验, 发现大部分预测 SNP 序列正确, 而有一部分预测 SNP 为 BAC 末端测序错误。验证了的 SNP 将是基因和 QTL 克隆的重要分子标记。同时对代表性的推测水稻中花 11 基因组特有重复序列进行了验证。

关键词 水稻; 物理图谱; 比较基因组学; BAC 末端序列; 单核苷酸多态性

中图分类号 S 511.032 **文献标识码** A **文章编号** 1000-2421(2011)03-0261-05

水稻(*Oryza sativa* L.)是我国最广泛种植的粮食作物,也是一种十分重要的模式植物。亚洲栽培稻主要有 2 个亚种:粳稻(*Oryza sativa* L. ssp. *japonica*)与籼稻(*Oryza sativa* L. ssp. *indica*)。利用比较基因组学方法对稻属不同种、亚种或品种间进行比较分析是开发水稻潜能的一种新方法,而物理图谱是进行比较基因组学研究的一种重要工具。自 2002 年 4 月粳稻和籼稻的基因组序列发表在 Science 上以来^[1-2],水稻序列与图谱的相关研究突飞猛进。水稻作为一种模式植物,其基因组序列的完成,为我们了解植物基因组成、功能以及基因组起源和进化等提供了大量的信息,水稻亚种或品种间的比较研究更是为水稻进化与起源提供了直接的证据, Ma 等^[3]比较了粳稻品种和籼稻品种之间基因组的差异,揭示了粳稻和籼稻基因组中同源基因区段的异常重组现象。同时,水稻物理图谱与遗传图谱的发展也极大地推动了水稻亚种或品种间的比较分析,从而促进了水稻遗传、进化、基因差异以及性状差异的研究^[4-5]。

细菌人工染色体(bacterial artificial chromosome, BAC)文库技术作为基因组学研究不可或缺的工具已经受到广泛的关注,随着大规模测序技术与 BAC 文库构建技术的发展,利用 BAC 末端序列

发掘的信息量也越来越大。通过对 BAC 末端序列的分析,可以对基因组内部的组成与结构作出初步的了解,并且可以通过与不同物种基因组的比较研究,揭示基因组序列的共性,探讨物种的起源和进化。

笔者所在实验室构建了籼稻品种珍汕 97、明恢 63 和粳稻品种中花 11 的物理图谱。本研究通过对珍汕 97 和明恢 63 物理图谱分别与日本晴参考序列的比对,利用日本晴序列设计探针筛选这 2 个品种的 BAC 文库膜,完成了 1 个产量数量性状位点(quantitative trait locus, QTL)区段珍汕 97 和明恢 63 物理图谱的补缺,并对该区段 BAC 末端序列与日本晴序列之间的单核苷酸多态性(single nucleotide polymorphisms, SNP)位点进行了检验,同时对代表性的推测中花 11 基因组特有重复序列进行了比较分析。

1 材料与方法

1.1 试验材料

研究中所用到的籼稻明恢 63、珍汕 97 和粳稻中花 11 种子由华中农业大学作物遗传改良国家重点实验室提供,粳稻日本晴种子由笔者所在实验室提供,此种子与日本晴基因组测序所用的种子来源

收稿日期: 2010-12-15

基金项目: 国家自然科学基金项目(30971748)

夏 鹏, 硕士研究生, 研究方向: 植物基因组学. E-mail: xiapengx@gmail.com

通讯作者: 罗美中, 博士, 教授, 研究方向: 植物基因组和叶绿体功能. E-mail: mzl原因@mail.hzau.edu.cn

一致。材料分别种植在大田或培养室,收集幼嫩叶片于-80℃保存备用。

1.2 物理图谱补缺

物理图谱补缺采用了 2 种方式,对大的缺口采用杂交筛选 BAC 克隆的方法,以日本晴序列为参考,用 Primer Premier 5 设计合适的探针扩增引物,用日本晴基因组 DNA 为模板扩增探针序列,每个物理图谱缺口设计 3 个探针,扩增序列经过纯化后作为探针对珍汕 97 和明恢 63 的 BAC 文库膜进行杂交筛选,杂交温度为 65℃,然后对筛选出的阳性克隆进行指纹分析。对小的缺口则直接采用 PCR 方法,根据缺口两端的 BAC 末端序列设计 PCR 引物,用相关基因组 DNA 作为模板进行 PCR 扩增。

1.3 BAC 末端序列与日本晴序列之间预测的 SNP 位点的检验

采用 2 种方法对 BAC 末端序列与日本晴序列之间预测的 SNP 进行检验。首先采用创造酶切位点 PCR 的方法,通过对预测 SNP 位点进行分析,根据碱基差别,用 Primer Premier 5 利用引物 3' 末端的错配形成新的酶切位点设计分型引物,通过对 PCR 产物的酶切鉴定 SNP。其次是直接重新测定 BAC 末端序列,对原有 BAC 末端序列与日本晴序列的碱基差异进行验证。

1.4 中花 11 特异性重复序列的分析

采用 CTAB 法提取各个水稻品种的基因组 DNA,用 *Hind*Ⅲ 在 37℃ 酶切 16 h,电泳后转移到 Hybond N+ 膜上,80℃ 烘烤 2 h 固定备用。利用推测的中花 11 特异性重复序列设计引物,并用 PCR 的方法扩增出相关序列作为探针,对制备的杂交膜杂交过夜,杂交温度为 55℃,洗膜后放射自显影检测杂交结果。

2 结果与分析

2.1 与 1 个产量 QTL 区段对应的明恢 63 和珍汕 97 的物理图谱补缺

笔者所在的实验室分别构建了水稻珍汕 97 和明恢 63 的物理图谱,并将这 2 个物理图谱与由这 2 个品种为亲本制作的遗传图谱进行了整合(未发表资料),从整合的图谱上发现第 3 染色体长臂末端的 1 个区段在珍汕 97 和明恢 63 中存在结构差异,而且与大量已发表的 QTL 及华中农业大学作物遗传改良国家重点实验室余四斌教授未发表的影响水稻千粒重、穗重和枝梗数的 QTL 相对应,笔者所在的

实验室正与余四斌教授合作对这一区段的功能进行研究。珍汕 97 在该区段存在 2 个大缺口(均在 100 kb 以上)和 1 个小缺口,明恢 63 在该区段存在 1 个大缺口(100 kb 以上)。对于珍汕 97 的小缺口,笔者根据缺口两端的 BAC 末端序列设计了 1 对引物,以珍汕 97 基因组 DNA 为模板用 PCR 的方法进行了扩增,获得 1 条专一的与预测缺口大小相符的 DNA 片段,测序结果表明,此 DNA 片段序列与对应的日本晴序列完全一致。对于珍汕 97 的 2 个和明恢 63 的 1 个大缺口,为每个缺口设计了 3 个杂交探针,对这 2 个品种的 BAC 文库进行了筛选,明恢 63 筛选得到 24 个阳性克隆,珍汕 97 得到 31 个阳性克隆,这些克隆关闭了所有 3 个缺口,图 1 显示明恢 63 三个探针混合筛选明恢 63 BAC 文库高密度杂交膜的结果。

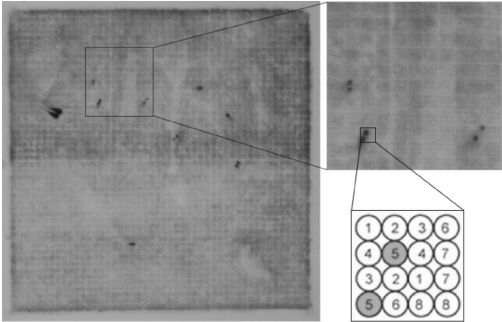
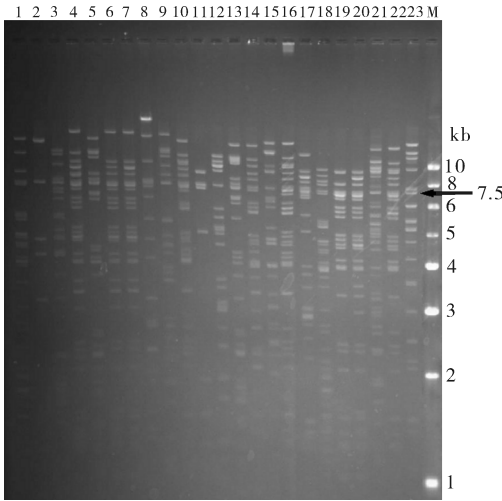


图 1 明恢 63 BAC 文库膜筛选结果

Fig. 1 Screening results of the Minghui 63 BAC membranes



箭头显示载体条带 The black arrow shows the vector bands.

图 2 明恢 63 阳性克隆指纹分析

Fig. 2 Fingerprint analysis of the Minghui 63 positive clones

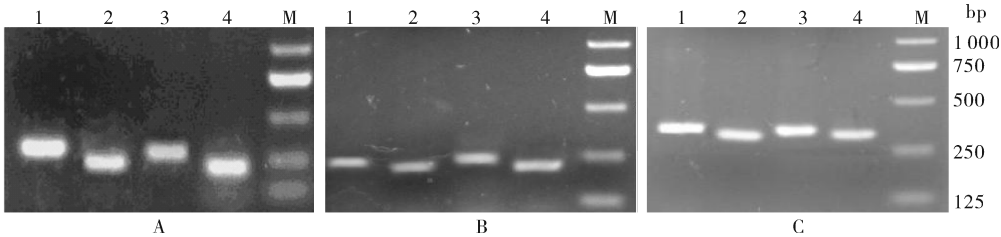
图 2 中显示的是明恢 63 筛选得到的 24 个阳性克隆中的 23 个克隆酶切结果,由图 2 可以看出,这

些克隆有大量的一致条带,证明这些克隆能够拼装成为克隆重叠群。

2.2 上述 QTL 区段珍汕 97 和明恢 63 BAC 末端序列分别与日本晴序列之间预测 SNP 的检验

SNP 是 1 种重要的分子标记。为了检验该 QTL 区段原珍汕 97 和明恢 63 BAC 末端序列分别与日本晴序列之间预测 SNP 位点的准确性,首先采用了创造酶切位点 PCR 的方法,以珍汕 97、明恢 63

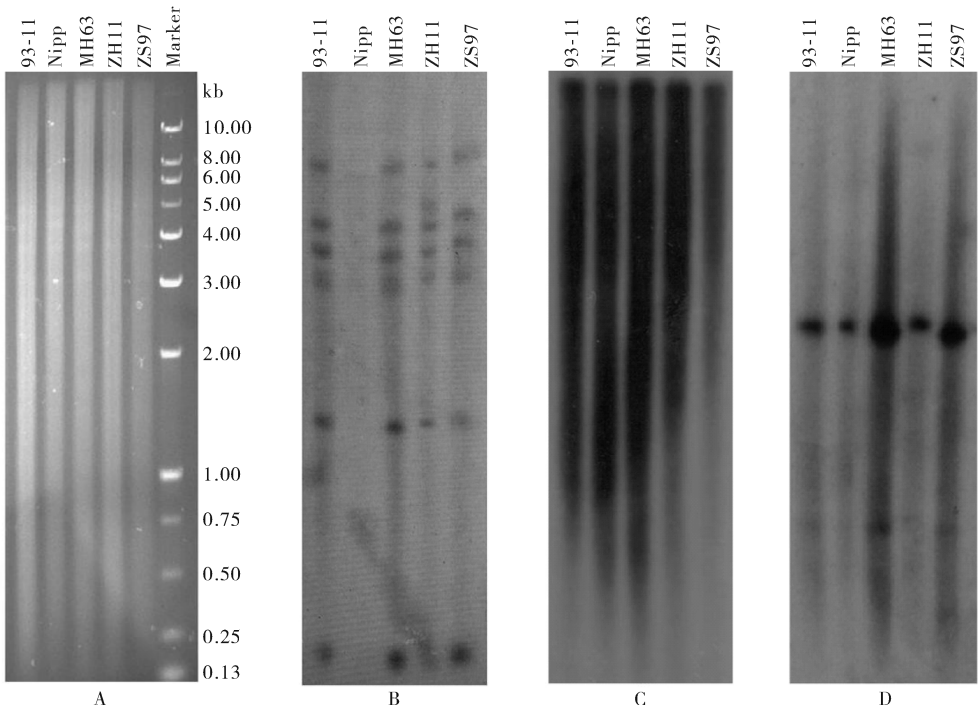
和日本晴材料中提取的 DNA 为模板进行 PCR 扩增并对 PCR 产物进行酶切。结果表明,该区段明恢 63 末端序列与日本晴序列之间被检测的 SNP 都得到验证,而该区段珍汕 97 末端序列与日本晴序列之间被检测的 7 个 SNP 都没有得到验证。图 3 显示以珍汕 97 和日本晴为模板的创造酶切位点 PCR 产物酶切后没有差异,也即是在设计被检测的位点处日本晴和珍汕 97 没有碱基差异。



A,B,C 分别为 SNPS2、SNPS3、SNPS31 3 对引物的 PCR 产物酶切结果; 1. 以珍汕 97 为模板扩增的 PCR 产物; 2. 其酶切产物; 3. 以日本晴为模板扩增的 PCR 产物; 4. 其酶切产物。A,B and C stand for the digestion results of the PCR products from three primer pairs, SNPS2, SNPS3 and SNPS31. 1. PCR product amplified from Zhenshan 97; 2. Digestion of the PCR product amplified from Zhenshan 97; 3. PCR product amplified from Nipponbare; 4. Digestion of the PCR product amplified from Nipponbare.

图 3 以珍汕 97 和日本晴为模板的创造酶切位点 PCR 产物酶切结果

Fig. 3 Digestion of PCR products from Zhenshan 97 and Nipponbare templates



A 为 93-11、日本晴(NIP)、明恢 63(MH63)、中花 11(ZH11)和珍汕 97(ZS97)5 个水稻品种的基因组 DNA *Hind*III 酶切结果, B、C 和 D 分别为 3 个中花 11 重复序列探针杂交结果。A shows the *Hind*III digestion results of 93-11, Nipponbare (NIP), Minghui 63 (MH63), Zhonghua 11 (ZH11) and Zhenshan 97 (ZS97) genomic DNA. B, C and D are hybridization results of three Zhonghua 11 repetitive element probes.

图 4 部分中花 11 重复序列杂交结果

Fig. 4 Hybridization results of some Zhonghua 11 repetitive elements

珍汕 97 BAC 末端序列与日本晴序列之间的碱基差异没有得到验证可能来自于 2 个原因,一是用于创造酶切位点 PCR 模板和 BAC 文库构建的珍汕 97 种子来源不同,二是原 BAC 末端测序错误。因此,对含有待检测差异碱基的原珍汕 97 BAC 克隆重新作了末端测序,结果显示被检测的原珍汕 97 BAC 末端序列与日本晴序列之间的碱基差异来自于原 BAC 末端测序的错误。

同样对含有待检测差异碱基的明恢 63 BAC 克隆重新作了末端测序,结果显示该区段所有被检测的明恢 63 BAC 末端序列与日本晴序列之间的 23 个 SNP 均得到验证。

2.3 中花 11 特异重复序列的验证

笔者所在的实验室通过对中花 11 BAC 末端序列和日本晴序列的比较分析发现,中花 11 存在日本晴基因组没有的特异重复序列(资料未发表),为了检测这些序列是否在不同品种的水稻基因组中存在,我们利用 PCR 对部分特异重复序列进行了扩增,并利用扩增产物作为探针对不同水稻品种基因组 DNA 进行了杂交,图 4 显示其中 3 个探针杂交的结果。从图 4-B 可以看出 1 个重复序列在粳稻中花 11 和籼稻 93-11、明恢 63 和珍汕 97 中的重复度相同,而在粳稻日本晴中没有条带显示,推测中花 11 可能有一定的籼稻背景。但是另外 2 个探针杂交结果显示日本晴和中花 11 的条带没有区别,而且改变杂交严谨度并不能改变试验结果(图 4-C、4-D),推测可能是因为使用的 PCR 探针中包含有部分其他品种共有的重复序列。

3 讨 论

BAC 末端序列常用于锚定 BAC 克隆的位置,从而能够定位 BAC 克隆在基因组中的顺序,加快物理图谱构建进程,此外,利用对 BAC 末端序列的分析还能够得到大量的信息。韩月彭等^[6-7]通过分析苹果的 BAC 末端序列找到大量的 SSR 标记、SNP 标记以及转座元件。Larkin 等^[8]利用牛的 BAC 末端序列与人类和小鼠基因组序列进行比对,通过绘制比较图谱从而为哺乳动物的分类与进化研究奠定了基础。本研究利用 BAC 末端序列与参照基因组序列比对对 BAC 末端序列中的有用信息进行了分析,首先,利用水稻品种珍汕 97 和明恢 63 的 BAC 末端序列分别与日本晴基因组序列的比对,得到了品种之间的候选 SNP,然后利用各种方法对这些

SNP 进行了验证;第二,利用珍汕 97 和明恢 63 BAC 末端序列与日本晴的参考序列比对,在 1 个与产量 QTL 相对应的区段快速定位了 BAC contig 在物理图谱中的排布顺序,并预测出这些 contig 之间缺口的大小,完成了所有 4 个物理缺口的填补工作;第三,对笔者所在的实验室利用中花 11 的 BAC 末端序列经过自我比对和与日本晴基因组序列的比对得到的部分中花 11 特有重复序列,设计了实验进行验证。通过对这三方面的初步研究证实了 BAC 末端序列作为 BAC 文库更深入的应用,能够大大加快比较基因组学的研究。

水稻物理图谱的完成对水稻基因组注释以及进化、结构与功能研究等方面有着极其重要的意义^[9],一个完整的物理图谱可以极大地便利遗传研究和一些重要基因的分离^[10],但是物理图谱一般都或多或少地存在一些缺口,这些缺口可能包含一些功能基因,而由于一些特殊结构导致了该处缺口的形成^[11]。笔者所在实验室构建的明恢 63 和珍汕 97 的物理图谱在第三染色体长臂末端有结构的差异,而该区段与大量已发表的 QTL 及余四斌老师实验室发现的影响水稻千粒重、穗重和枝梗数的 QTL 对应,为了对该区段基因功能有更深入的了解,我们不仅填补了 2 个物理图谱在该区段所遗留的 4 个缺口,还验证了该区段的大量 SNP 位点,对已经得到验证的明恢 63 的 23 个 SNP 位点进行分析,发现其中有 18 个 SNP 位点比对到日本晴单一序列位置,占总数的 78.26%,这 18 个 SNP 均处于基因的编码区,可能会影响到这些基因的功能,其余 5 个 SNP 位点比对到日本晴基因组非单一序列位置,占总数的 22.74%。分子标记在分子遗传图谱制作、基因定位、基因克隆以及分子标记辅助选择育种等方面发挥了重要的作用,而本文中经验证的 SNP 均可作为重要的分子标记锚定到遗传图谱上。

水稻中花 11 和日本晴同属于粳稻亚种。经杂交验证,中花 11 的一个特异重复序列在日本晴基因组中并不存在,证明可能在驯化过程中水稻亚种内不同品种基因组仍可发生比较大的变化。

参 考 文 献

- [1] YU J, HU S N, WANG J, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*) [J]. Science, 2002, 296: 79-92.
- [2] GOFF S A, RICKE D, LAN T H, et al. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*) [J]. Science, 2002, 296: 92-100.

[3] MA J X, BENNETZEN J L. Rapid recent growth and divergence of rice nuclear genomes[J]. *Proc Natl Acad Sci USA*, 2004, 101: 12404-12410.

[4] YU J, WANG J, LIN W, et al. The genomes of *Oryza sativa*: a history of duplications[J]. *PLoS Biol*, 2005(3): 266-281.

[5] KIM H, SAN MIGUEL P, NELSON W, et al. Comparative physical mapping between *Oryza sativa* (AA genome type) and *O. punctata* (BB genome type) [J]. *Genetics*, 2007, 176: 379-390.

[6] HAN Y P, CHAGNÉ D, GASIC K, et al. BAC-end sequence-based SNPs and Bin mapping for rapid integration of physical and genetic maps in apple[J]. *Genomics*, 2009, 93: 282-288.

[7] HAN Y P, KROBAN S S. An overview of the apple genome through BAC end sequence analysis[J]. *Plant Mol Biol*, 2008, 67: 581-588.

[8] LARKIN D M, EVERTS-VAN DER WIND A, REBEIZ M, et al. A cattle-human comparative map built with cattle BAC-ends and human genome sequence[J]. *Genome Res*, 2003, 13: 1966-1972.

[9] LU F, AMMIRAJU J S, SANYAL A, et al. Comparative sequence analysis of MONOCULM1-orthologous regions in 14 *Oryza* genomes[J]. *Proc Natl Acad Sci USA*, 2009, 106: 2071-2076.

[10] ZHANG H B, WING R A. Physical mapping of the rice genome with BACs[J]. *Plant Mol Biol*, 1997, 35: 115-127.

[11] LEEM S H, KOUPRINA N, GRIMWOOD J, et al. Closing the gaps on human chromosome 19 revealed genes with a high density of repetitive tandemly arrayed elements[J]. *Genome Res*, 2004, 14: 239-246.

Comparative analyses of a rice yield QTL region and variety-specific repetitive sequences

XIA Peng LIN Hai-yan LUO Mei-zhong

*National Key Laboratory of Crop Genetic Improvement ,
Huazhong Agricultural University ,Wuhan 430070, China*

Abstract To clone a rice yield QTL, a comparative analysis on the region of the physical maps derived from rice varieties Zhenshan 97 and Minghui 63, the two parents for mapping the QTL, was performed. Physical gaps in this region of the two physical maps were closed. Putative SNPs of this region between Zhenshan 97 BAC end sequences and Nipponbare reference sequence, and that between Minghui 63 BAC end sequences and Nipponbare reference sequence were validated. Results showed that most putative SNPs are correct but some are from errors of BAC end sequence. The confirmed SNPs will be used as important markers in the positional cloning of gene or QTL. Furthermore, representative putative Zhonghua 11 specific repetitive sequences found in our laboratory through Zhonghua 11 BES analysis were comparatively analyzed *via* Southern hybridization. The results will be helpful in cloning of the rice yield QTL and realizing the evolution of rice repetitive sequences among different rice varieties.

Key words rice; physical map; comparative genomics; BAC end sequences; single nucleotide polymorphisms

(责任编辑:杨锦莲)