

研究报告

A Letter

地方稻月亮谷花期控制基因 *Hd3a*、*Hd1*、*Edh1* 和 *Ghd7* 的多态性分析

高东¹, 毛如志¹, 魏富刚², 何霞红¹

1 云南农业大学农业生物多样性应用技术国家工程中心, 教育部农作物多样性与病害控制重点实验室, 植物病理重点实验室, 昆明, 650201

2 文山州植保站, 文山, 663000

✉ 通讯作者: gaodong521@yahoo.com.cn ✉ 作者

分子植物育种, 2012 年, 第 10 卷, 第 23 篇 doi: 10.5376/mpb.cn.2012.10.0023

收稿日期: 2012 年 04 月 20 日

接受日期: 2013 年 04 月 26 日

发表日期: 2012 年 05 月 23 日

这是一篇采用 Creative Commons Attribution License 进行授权的开放取阅论文。只要对本原作有恰当的引用, 版权所有人允许并同意第三方无条件的使用与传播。

引用格式(中文):

高东等, 2012, 地方稻月亮谷花期控制基因 *Hd3a*、*Hd1*、*Edh1* 和 *Ghd7* 的多态性分析, 分子植物育种(online) Vol.10 No.23 pp.1164-1170 (doi: 10.5376/mpb.cn.2012.10.0023)

引用格式(英文):

Gao et al., 2012, Analysis of Functional Nucleotide Polymorphisms in *Hd3a*, *Hd1*, *Edh1* and *Ghd7* Genes Controlling Heading Date in Yuelianggu Landrace, Fenzi Zhiwu Yuzhong (online) (Molecular Plant Breeding) Vol.10 No.23 pp.1164-1170 (doi: 10.5376/mpb.cn.2012.10.0023)

摘要 水稻抽穗期控制的关键基因已基本清楚, 但是栽培稻花期多样性的分子机理仍然不清楚。本文利用重测序月亮谷的全基因组序列、Takahashi 等(2009)在 PNAS 上发表的部分数据和 Xue 等(2008)在 Nature genetics 发表的部分数据, 研究了 *Hd3a*、*Hd1*、*Edh1* 和 *Ghd7* 等 4 个基因的核苷酸多态性。结果显示, 月亮谷 *Hd3a* 启动子区核苷酸多态性分布与前人报道的有所不同, 发现 5 个特有的 SNP 位点, 分别位于 -1 851 bp、-1 720 bp、-700 bp、-681 bp 和 -1 607 bp 处; 1 个特有的 InDel, 位于 -523 bp 处。月亮谷 *Hd1* 编码区其具有很高的多态性, 发现 6 个特有的 SNP/InDel 位点, 其中包括了所有 4 个移码位点, 1 个插入位点和 1 个缺失位点, 月亮谷 *Hd1* 属无功能的等位基因类型。月亮谷 *Edh1* 编码区 SNP/InDel 与前人研究结果相似。月亮谷 *Ghd7* 属于前人报道的有功能的等位基因 *Ghd7-1*。月亮谷抽穗期相关基因复杂的多态性, 一方面与其生长环境垂直海拔落差导致的“一山分四季, 十里不同天”的复杂小生境有关, 另一方面可能是其耐病、耐寒, 适应性广的分子基础。

关键词 地方稻; 基因; 抽穗期; 单核苷酸多态性; 哈尼梯田

Analysis of Functional Nucleotide Polymorphisms in *Hd3a*, *Hd1*, *Edh1* and *Ghd7* Genes Controlling Heading Date in Yuelianggu Landrace

Gao Dong¹, Mao Ruzhi¹, Wei Fugang², He Xiaohong^{1,2}

1. The National Center for Agricultural Biodiversity, Key Laboratory of Ministry of Education for Agricultural Biodiversity and Plant Disease Control, Key Laboratory of Plant Pathology, Yunnan Agricultural University, Kunming, 650201, P.R. China;

2. Plant Protection Station of Wensan Autonomous Prefecture, Wensan, 663000, P.R. China

✉ Corresponding author, gaodong521@yahoo.com.cn; ✉ Authors

Abstract The major genes involved in rice flowering have been identified. However, the molecular mechanism for promoting the diversity of flowering time in cultivated rice is still unknown. In this paper, the genome re-sequencing of Yuelianggu, the partial data quoted from papers written by Takahashi et al. in PNAS in 2009 and by Xue et al. in Nature Genetics in 2008, are used to study the nucleotide polymorphisms of 4 genes (*Hd3a*, *Hd1*, *Edh1* and *Ghd7*) in the rice flowering pathway. The results indicate that nucleotide polymorphisms in the *Hd3a* promoter region are different from what was reported previously. 5 specific SNPs for Yuelianggu are distributed at the locations of -1 851 bp, -1 720 bp, -700 bp, -681 bp and -1 607 bp. 1 specific InDel for Yuelianggu was found at the location of -523 bp. A high polymorphism in the *Hd1* coding region was indentified. There are 6 specific SNP/InDels for Yuelianggu, including 4 frame shift 1 insert and 1 delet. *Hd1* is a nonfunctional allele in Yuelianggu. The nucleotide polymorphisms in the *Edh1* coding sequence are similar to previously reported data. *Ghd7* is a functional allele in Yuelianggu and belongs to previously reported *Ghd7-1*. The complicated polymorphisms of the genes related to the flowering time in Yuelianggu, may have some relations with its habitat, Ailao mountain region, which has sharp elevation and is diversiform in climate. On the other hand, this could be the molecular basis for its tolerance to disease, pest and cold, and its wide adaptation.

Keywords *Oryza sativa* L. landrace; Gene; Heading date; Single nucleotide polymorphisms (SNP); Hani terrace

研究背景

抽穗期是决定作物品种生育期及地区适应性

的重要性状, 受自身遗传因子和外界环境因素两方面决定, 是开花基因时空顺序表达的结果(Yano et

al., 2001)。水稻的全基因组序列测定已经完成(Goff et al., 2002; Yu et al., 2002), 重测序项目发展迅速。序列数据库为候选同源基因搜索提供了方便(Izawa et al., 2003)。SNP/InDel 是一种在基因组中广泛存在的多态性(岳兵和邢永忠, 2005)。

在水稻上与农艺性状自然变异有关的功能核苷酸多态性(functional nucleotide polymorphisms, FNP)已经得到了鉴定, 其中包括花期相关基因 *Hd1*、*Edh1* 和 *Ghd7* (Yano et al., 2000; Doi et al., 2004; Xue et al., 2008; 薛为亚等, 2009)。根据 *GHD7* 蛋白序列可以把 *Ghd7* 分为 5 个等位基因, 各等位基因间仅有几个氨基酸的不同, 从而导致 *Ghd7* 的 5 个等位基因在地理分布上表现出明显的规律性。效应较大的 *Ghd7-1* 和 *Ghd7-3* 主要存在于栽种在亚热带、热带或夏季温度高且持续时间长的品种里; *Ghd7-2* 分布于栽种在温带地区的粳稻品种里; 完全缺失的 *Ghd7-0* 分布于华中和华南二季稻区的早稻品种里; 蛋白质提前终止的 *Ghd7-0a* 则存在于东北黑龙江省的品种中(薛为亚等, 2009)。最近(Takahashi et al., 2009)等通过收集来自世界不同地区 332 份水稻品种(代表了水稻种质 91% 的多样性),

来研究水稻花期的多样性, 从开花时间上看所有收集的种质同样存在着极大的多样性, 花期 45~153 d 不等(播种后到开花的时间)。同时在前人研究的基础上, 选择涉及到影响水稻开花的相关基因来分析那些基因导致了这种开花时间的差异。研究发现 *Hd3a* 启动子序列的变异, *Hd1* 核心编码序列的变异, *Ehd1* 的表达量与对应的表型即开花时间的变异有着极显著的相关性。

本文以哀牢山哈尼梯田的传统代表品种(月亮谷)全基因组重测序为材料, 借助目前在水稻花期基因中的研究结果, 通过序列的搜索比对分析其 SNP/InDel 的数量与分布; 同时针对已报道的与开花适应性, 花期多样性有关的基因的变异与类型与月亮谷的比较分析, 为从分子水平上来解释当地传统品种能够持续种植上百年的依据。

1 结果分析

1.1 *Hd3a* 启动子区的核酸多态性

Hd3a 启动子区超过 2 kb 序列(包括 5'末端非转录序列)的核苷酸多态性的分析结果显示(图 1), 月

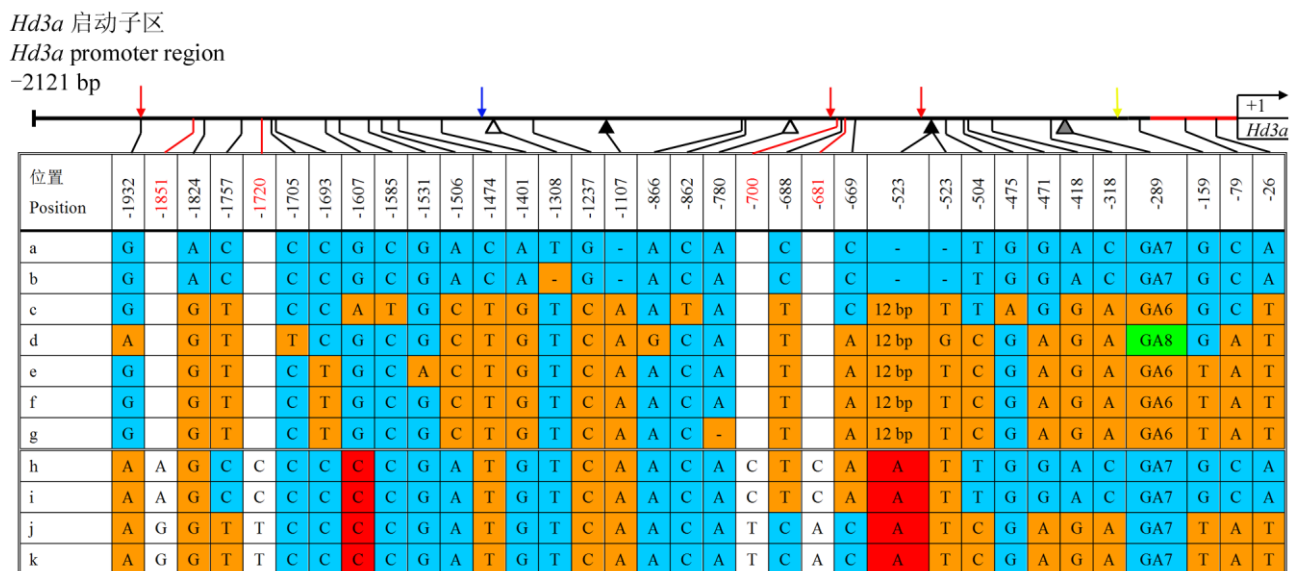


图 1 *Hd3a* 启动子区核苷酸多态性

注: a-g: 类型 1 至类型 7; h: 9311; i: 日本晴; j: 月亮谷-1; k: 月亮谷-2; 红色, 蓝色和黄色箭头所指分别为 *ARR1* 结合元件, *CCAAT* 盒和 *TATA* 盒; 缺失和插入位点分别用空心 and 实心的三角箭头指示; *AGA* 重复用灰色三角箭头指示; 5'末端非转录序列用红色所示; 核苷酸多态性用不同的颜色表示; 9311 前数据引自 Takahashia 等(2009)

Figure 1 Nucleotide polymorphisms in the *Hd3a* promoter region

Note: a-g: Type 1 to type 7; h: 9311; i: Nipponbare; j: Yuelianggu-1; k: Yuelianggu-2; Red arrows indicate *ARR1* binding elements, the blue arrow indicates the *CCAAT* box, and the yellow arrow indicates the *TATA* box; Deletion and insertion sites are shown by open and closed arrowheads, respectively; *AGA* repeat sequence is indicated by gray arrowhead; The 5'-untranslated region is colored in red; Polymorphic nucleotides are indicated by different colors; The data before 9311 is quoted from Takahashia et al. (2009)

亮谷启动子区核苷酸多态性分布与前人报道的明显不同。核苷酸多态性变异介于两组(类型 1, 类型 2; 类型 3, 类型 4, 类型 5, 类型 6, 类型 7)之间, 与 7 种类型的任何一种都不相似, 而且发现 5 个特有的 SNP 位点, 分别位于-1 851 bp, -1 720 bp, -700 bp, -681 bp 和-1 607 bp 处; 1 个特有的 InDel, 位于-523 bp 处。在启动子区的一些顺式作用元件(cis-elements)如 CCAAT 盒以及功能域如 ARR 没有核苷酸变异, 这与文献报道一致。

1.2 Hd1 编码区的核酸多态性

Hd1 编码区的核酸多态性分析结果显示(图 2), *Hd1* 其具有很高的多态性。发现 6 个特有的 SNP/InDel 位点, 其中包括了所有 4 个移码位点, 1 个插入位点和 1 个缺失位点。月亮谷 *Hd1* 基因包含

有前人报道的 871 位基因位点上的移码突变, 从而导致 *Hd1* 基因 CCT 结构域的部分失活, 属无功能的等位基因类型。而对前人报道的非移码突变及提前终止位点, 月亮谷 *Hd1* 基因类型介于上述类型之间; 9311 属于无功能的类型 7, 日本晴则属于有功能的类型 5。

1.3 Ehd1 编码区的核酸多态性

Ehd1 编码区的核酸多态性分析的结果显示(图 3), 月亮谷 *Ehd1* 编码区没有特有的 SNP/InDel 位点, 与前人研究结果相似, 可以归为类型 3、类型 4、类型 5、类型 6 和类型 7 组成的组。

1.4 Ghd7 等位基因的蛋白序列多态性

Ghd7 是影响每穗粒数、抽穗期的一个数量性

*Hd1*编码区

Hd1 coding region

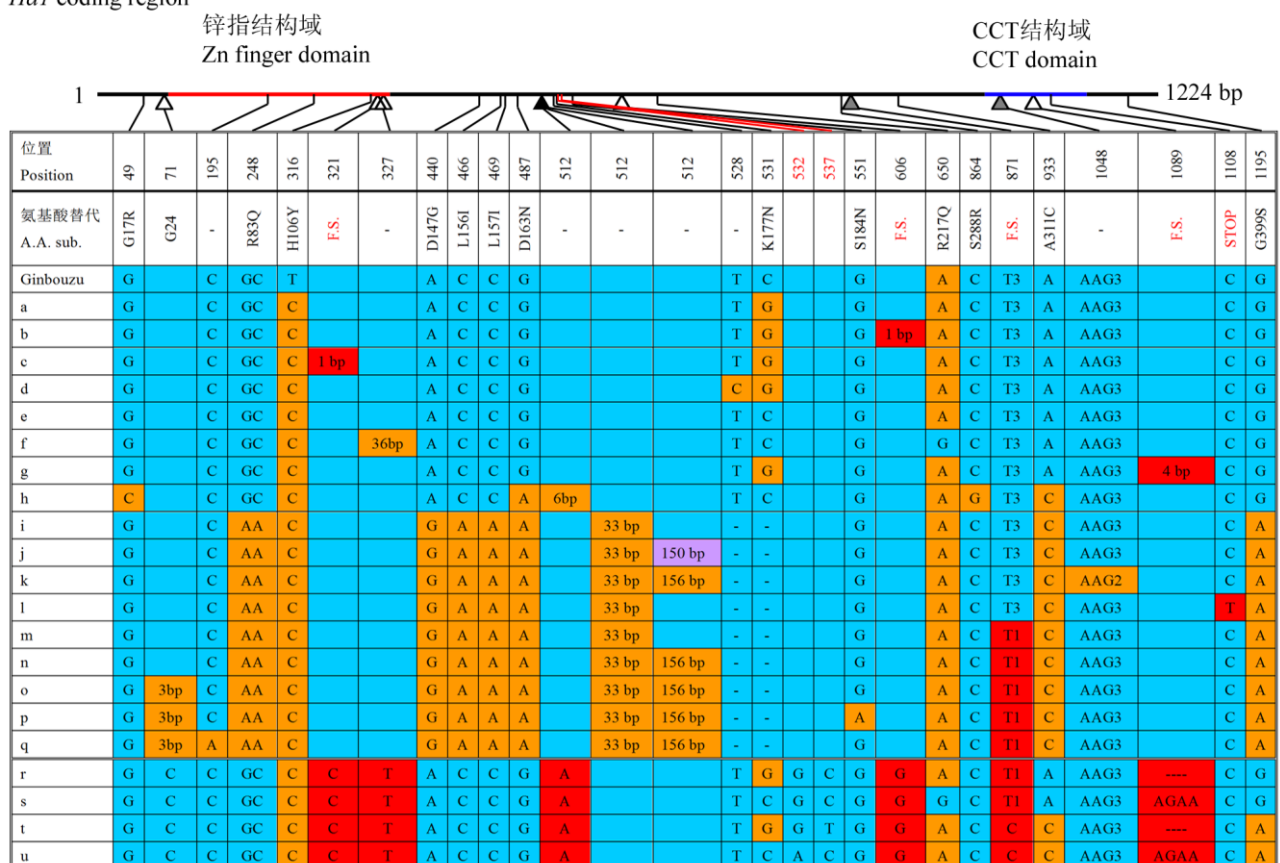


图 2 *Hd1* 核心编码序列区的多态性

注: a-q: 类型 1 至类型 17; r: 9311; s: 日本晴; t: 月亮谷-1; u: 月亮谷-2; 缺失和插入位点分别用空心和实心的三角箭头指示; F.S 表示阅读框架移码; 核苷酸多态性用不同的颜色表示; 9311 前数据引自 Takahashia 等(2009)

Figure 2 A high polymorphism in the *Hd1* coding sequence

Note: a-q: Type 1 to type 17; r: 9311; s: Nipponbare; t: Yuelianggu-1; u: Yuelianggu-2; Deletion and insertion sites are shown by open and closed arrowheads, respectively; F.S. is frame shift; Polymorphic nucleotides are indicated by different colors; The data before 9311 is quote from Takahashia et al. (2009)

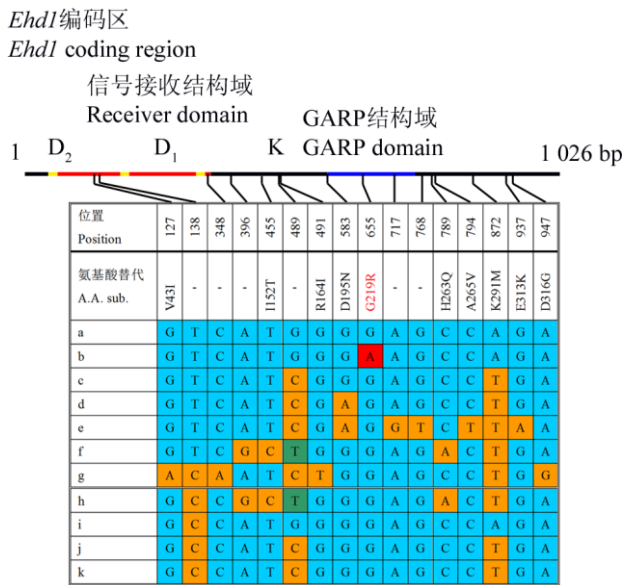


图 3 *Ehd1* 编码区核苷酸多态性
注: a-g: 类型 1 至类型 17; h: 9311; i: 日本晴; j: 月亮谷-1; k: 月亮谷-2; 天冬氨酸(D1, D2)和赖氨酸(K)残基用黄色线条表示; 在品种单玉的 GARP 结构域发现一个单氨基酸替换, 用红色表示; 核苷酸多态性用不同的颜色表示; 9311 前数据引自 Takahashia 等(2009)

Figure 3 Nucleotide polymorphisms in the *Ehd1* coding sequence
Note: a-g: Type 1 to type 17; h: 9311; i: Nipponbare; j: Yuelianggu-1; k: Yuelianggu-2; Asp (D1, D2) and Lys (K) residues are indicated by yellow lines; A single amino acid substitution in the GARP domain, found in cv; DANYU, is shown in red; Polymorphic nucleotides are indicated by different colors; The data before 9311 is quoted from Takahashia et al. (2009)

状位点(QTL), 对水稻的一系列性状包括单穗粒数、株高和抽穗期都产生重要影响。通过对元阳传统水稻品种月亮谷 *Ghd7* 基因蛋白序列分析(图 4), 结果显示, 月亮谷 *Ghd7* 属于前人报道的有功能的等位基因 *Ghd7-1*。前人报道在野生稻种也发现具有功能的 *Ghd7-1*, 说明功能型的 *Ghd7-1* 可能是其他等位基因的祖先, 随着水稻种植区域的扩大为了适应不同的地域环境, 而衍生出其他等位基因, 同时也导致开花时间对不同地区的适应性及多样性。

2 讨论

植物对自然环境的适应一个很重要的机制就是通过对开花时间的调控进而在适当的季节开花, 繁衍后代。水稻(*Oryza sativa* L.)从过去的 8000 至 1 万年前开始被驯化和繁育, 有着很长的进化史(Doebley et al., 2006; Khush, 1997)。能够找到其适合

的生态位和对自然环境的适应能力, 是水稻生态繁荣的关键策略。水稻种植地理区域的扩展和广泛分布, 产量的增加, 一个很重要的因素就是水稻开花时间(花期)的多样性(Khush, 1997)。在特定环境下的开花时间调控, 长期以来一直被理解为是一种很重要的适应性(Izawa, 2002)。

Takahashi 等(2009)通过收集来自世界不同地区 332 份水稻品种(代表了水稻种质 91%的多样性)来研究水稻花期的多样性, 从开花时间上看所有收集的种质同样存在着极大的多样性, 花期 45 d~153 d 不等(播种后到开花的时间)。同时在前人研究的基础上, 选择涉及到影响水稻开花的相关基因来分析那些基因导致了这种开花时间的差异。研究发现 *Hd3a* 启动子序列的变异, *Hdl* 核心编码序列的变异, *Ehd1* 的表达量与对应的表型即开花时间的变异有着极显著的相关性。

Hd3a 启动子 2 kb 的区域被认为是 *Hd3a* 表达所必须的, 依据启动子以及 5' UTR 的核酸多态性分析, 将收集的核心种可分为 7 种类型, 对应于其开花时间多样性表型可将其分为两组; 类型 1 类型 2 归为 A 组, 类型 3-7 归为 B 组, 表达分析表明 B 组 *Hd3a* 的表达显著高于 A 组, 说明 *Hd3a* 启动子类型与开花时间有很显著的相关性。

月亮谷 *Hd3a* 启动子区的核酸多态性分布与前人报道的明显不同, 介于 A、B 两组, 与 7 种类型的任何一种都不相似, 而且有 5 个特有的 SNP 位点。这些特征可能与哀牢山“一山分四季, 十里不同天”的生态环境相适应, 其相关性有待进一步研究。

Hdl 编码区的核酸具有很高的多态性, 可分为 17 个类型和 15 种独立的蛋白(Takahashi et al., 2009)。其中有 9 种无功能类型的核苷酸变异发生在 CCT 结构域中, CCT 结构域编码一种核定位信号, 在拟南芥中 *CO* 基因 CCT 的变异可导致 *CO* 蛋白的失活; 相关性分析表明这种多态性与水稻花期有着显著的相关性; 同时有功能的和无功能的 *Hdl* 基因类型具有明显的地区分布。月亮谷 *Hdl* 编码区的核酸较高的多态性, 可能与哀牢山较高垂直海拔落差(144~2939.6 m)相关, 其相关性有待进一步研究。

Xue 等(2008)测序分析了东南亚的 19 个品种, 依据蛋白序列变异将 *Ghd7* 分为 5 个等位基因, 蛋白质编码系列表明 *Ghd7-1* 与 *Ghd7-2* 有 4 个氨基酸的不同, 导致 *Ghd7-2* 的功能较弱些; *Ghd7-3* 与 *Ghd7-1* 和 *Ghd7-2* 有 3 个氨基酸的差别, 同时有 1



Ghd7-1	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Ghd7-2	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Ghd7-3	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Ghd7-0	: -----	: 0
Ghd7-0a	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVH-----	: 52
9311	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Yueliang-1	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
NIP	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Yueliang-2	: MSMGPAAGEGCGLCGADGGGCCSRHRHDDGGFPFVFPSSACQIGAPAPPVHEFQFFGND	: 60
Ghd7-1	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGAGGLTFDVSL	: 120
Ghd7-2	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGAGGLTFDVSL	: 120
Ghd7-3	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGGGLTFDVSL	: 120
Ghd7-0	: -----	: 0
Ghd7-0a	: -----	: 52
9311	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGAGGLTFDVSL	: 120
Yueliang-1	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGAGGLTFDVSL	: 120
NIP	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGAGGLTFDVSL	: 120
Yueliang-2	: GGGDDGESVAVLFDYPPSPVAAAAGMHRQPPYDGVVAPPVSLFRRTGGGLTFDVSL	: 120
Ghd7-1	: GGRPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADVGESLNP	: 180
Ghd7-2	: GERPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADDGESLNP	: 180
Ghd7-3	: GGRPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADDGESLNP	: 180
Ghd7-0	: -----	: 0
Ghd7-0a	: -----	: 52
9311	: GGRPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADVGESLNP	: 180
Yueliang-1	: GERPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADDGESLNP	: 180
NIP	: GERPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADDGESLNP	: 180
Yueliang-2	: GGRPDL DAGLGLGGSGRHAEAAASATIMSYCGSTFTDAASSMPKEMVAAMADVGESLNP	: 180
Ghd7-1	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEADQEAVAP	: 240
Ghd7-2	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEPDQEAVAP	: 240
Ghd7-3	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEPDQEAVAP	: 240
Ghd7-0	: -----	: 0
Ghd7-0a	: -----	: 52
9311	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEADQEAVAP	: 240
Yueliang-1	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEPDQEAVAP	: 240
NIP	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEPDQEAVAP	: 240
Yueliang-2	: NTVVGAMVEREAKLMRYKEKRKKRCYEKQIRYASRKAYAEMRPRVRGRFAKEADQEAVAP	: 240
Ghd7-1	: PSTYVDPSRLELGQWFR	: 257
Ghd7-2	: PSTYVDPSRLELGQWFR	: 257
Ghd7-3	: PSTYVDPSRLELGQWFR	: 257
Ghd7-0	: -----	: 0
Ghd7-0a	: -----	: 52
9311	: PSTYVDPSRLELGQWFR	: 257
Yueliang-1	: PSTYVDPSRLELGQWFR	: 257
NIP	: PSTYVDPSRLELGQWFR	: 257
Yueliang-2	: PSTYVDPSRLELGQWFR	: 257

图 4 *Ghd7* 等位基因蛋白序列多态性

注: 图中的灰色阴影(aa189-233)是 CCD 结构域; 蓝色的部分指出的是通过 PSORT 分析而得出的两个核定位信号基序, KRKK 起始于 aa199, RKKR 起始于 aa200; NIP 前数据引自 Xue 等(2008)

Figure 4 The protein sequences polymorphisms of GHD7 for *Ghd7* alleles

Note: The area in shadow (aa189-233) is the putative CCD domain; The stretch of letters in blue indicates the two motifs of putative nuclear localization signals identified by a PSORT analysis; KRKK starts at aa199 and RKKR starts at aa200; The data before NIP is quoted from Xue et al. (2008)

个氨基酸变异是 *Ghd7-3* 特有的; *Ghd7-0* 和 *Ghd7-0a* 是没功能的等位基因, *Ghd7-0* 是功能域的删除而导致的, 而 *Ghd7-0a* 是一个终止密码子的突变导致编码系列提前终止导致的。考察其地理分布, 各等位基因具有明显的规律, 效应较大的 *Ghd7-1* 和 *Ghd7-3* 主要分布在热带和亚热带以及夏季温度高且持续时间长的地区, *Ghd7-2* 分布在温带粳稻品种里; *Ghd7-0* 分布在华中和华南二季稻区的早稻里, *Ghd7-0a* 则分布在夏季凉爽, 水稻生育期短的东北黑龙江省的品种中(薛为亚等, 2009)。元阳传统水稻品种月亮谷持有功能的等位基因 *Ghd7-1*, 与其气候环境相符。但是, 鉴于哀牢山的垂直海拔落差以及“一山分四季, 十里不同天”的特征, 有必要对哀牢山区不同小生境的月亮谷的 *Ghd7* 进行比较分析, 极有可能发现其他类型的 *Ghd7*。

3 材料与方 法

3.1 哈尼族传统水稻品种月亮谷全基因组测序

哈尼族当地传统水稻品种月亮谷全基因组序列采取重测序的方法, 我课题组出资, 由华大基因(深圳)完成, 并分别以 9311 和日本晴为模板拼接成 2 套月亮谷全基因组序列, 分别记为月亮谷 1 (9311) 和月亮谷 2 (日本晴)。

3.2 基因序列的获取与比对

利用 BIOEDIT (Kostka et al., 2008)将月亮谷的重测序序列建立本地 BLAST 数据库。然后用 *Hd3a*、*Hd1*、*Edh1* 和 *Ghd7* 的序列作查询序列搜索 BLAST 数据库发现基因片段, 同时截取基因及上游 2 kb 的调控序列片段, 序列比对用 DNAMAN 软件进行, 比对结果文件保存为 .phy 文件。依据 Takahashi 等 (2009)发表的多态性位点进行汇总、做图。

作者贡献

高东是本研究的实验设计和实验研究的执行人; 毛如志和魏富刚完成实验及数据分析; 高东是论文的构思者及负责人, 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

本研究国家重点基础研究发展计划(973项目)课题(2011CB100406)资助。感谢唐有福在整个科研过程中提供的大量支持。

参考文献

Doebley J.F., Gaut B.S., Smith B.D., 2006, The molecular genetics of crop domestication, *Cell*, 127(7): 1309-1321

<http://dx.doi.org/10.1016/j.cell.2006.12.006> PMID:17190597

Doi K., Izawa T., Fuse T., Yamanouchi U., Kubo T., Shimatani Z., Yano M., and Yoshimura A., 2004, *Ehd1*, a B-type response regulator in rice, confers short-day promotion of flowering and controls *FT*-like gene expression independently of *Hd1*, *Genes Dev.*, 18(8): 926-936 <http://dx.doi.org/10.1101/gad.1189604>

Goff S.A., Ricke D., Lan T.H., Presting G., Wang R., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H., Hadley D., Hutchison D., Martin C., Katagifi F., Lange B.M., Moughamer T., Xia Y., Budworth P., Zhong J., Miguel T., Paszkowski U., Zhang S., Colbert M., Sun W.L., Chen L., Cooper B., Park S., Wood T.C., Mao L., Quail P., Wing R., Dean R., Yu Y., Zharkikh A., Shen R., Sahasrabudhe S., Thomas A., Cannings R., Gutin A., Pruss D., Reid J., Tavtigian S., Mitchell J., Eldredge G., Scholl T., Miller R.M., Bhatnagar S., Adey N., Rubano T., Tusneem N., Robinson R., Feldhaus J., Macalma T., Oliphant A., and Briggs S., 2002, A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*), *Science*, 296(5565): 92-100 <http://dx.doi.org/10.1126/science.1068275> PMID: 11935018

Izawa T., Oikawa T., Sugiyama N., Tanisaka T., Yano M., and Shimamoto K., 2002, Phytochrome mediates the external light signal to repress *FT* orthologs in photoperiodic flowering of rice, *Genes Dev.*, 16(15): 2006-2020 <http://dx.doi.org/10.1101/gad.999202>

Izawa T., Takahashi Y.J., and Yano M., 2003, Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and Arabidopsis, *Curr. Opin. Plant Biol.*, 6(2): 113-120 [http://dx.doi.org/10.1016/S1369-5266\(03\)00014-1](http://dx.doi.org/10.1016/S1369-5266(03)00014-1)

Khush G.S., 1997, Origin, dispersal, cultivation and variation of rice, *Plant Mol. Biol.*, 35(1-2): 25-34 <http://dx.doi.org/10.1023/A:1005810616885> PMID:9291957

Kostka M., Uzlikova M., Cepicka I., and Flegr J., 2008, SlowFaster, a user-friendly program for slow-fast analysis and its application on phylogeny of Blastocystis, *BMC Bioinformatics*, 9: 341 <http://dx.doi.org/10.1186/1471-2105-9-341> PMID:18702831 PMCid:2529323

Takahashi Y., Teshima K.M., Yokoi S., Innan H., and Shimamoto K., 2009, Variations in *Hd1* proteins, *Hd3a* promoters, and *Ehd1* expression levels contribute to diversity of flowering time in cultivated rice, *Proc. Natl. Acad. Sci. USA.*, 106(11): 4555-4560 <http://dx.doi.org/10.1073/pnas.0812092106> PMID:19246394 PMCid:2647979

Xue W., Xing Y., Weng X., Zhao Y., Tang W., Wang L., Zhou H., Yu S., Xu C., Li X., and Zhang Q., 2008, Natural variation in *Ghd7* is an important regulator of heading date

- and yield potential in rice, *Nat. Genet.*, 40(6): 761-767
<http://dx.doi.org/10.1038/ng.143> PMID:18454147
- Xue W.Y., Xing Y.Z., Weng X.Y., Zhao Y., Tang W.J., Wang L., Zhou H.J., Yu S.B., Xu C.G., Li X.H., and Zhang Q.F., 2009, Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice, *Zhongguo Jichu Kexue (China Basic Science)*, 2: 21-23 (薛为亚, 邢永忠, 翁小煜, 赵毓, 唐为江, 王磊, 周红菊, 余四斌, 徐才国, 李香花, 张启发, 2009, *Ghd7*自然变异是调控水稻抽穗期和产量潜力的重要因素, *中国基础科学*, 2: 21-23)
- Yano M., Katayose Y., Ashikari M., Yamanouchi U., Monna L., Fuse T., Baba T., Yamamoto K., Umehara Y., Nagamura Y. and Sasaki T., 2000, *Hdl*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*, *The Plant Cell*, 12(12):2473-2484 <http://dx.doi.org/10.2307/3871242> PMID:11148291 PMCID:102231 <http://dx.doi.org/10.1105/tpc.12.12.2473>
- Yano M., Kojima S., Takahashi Y., Lin H. and Sasaki T., 2001, Genetic control of flowering time in rice, a short-day plant, *Plant Physiol.*, 127(4): 1425-1429 <http://dx.doi.org/10.1104/pp.010710> PMID:11743085 PMCID:1540174
- Yu J., Hu S., Wang J., Wong G.K., Li S., Liu B., Deng Y., Dai L., Zhou Y., Zhang X., Cao M., Liu J., Sun J., Tang J., Chen Y., Huang X., Lin W., Ye C., Tong W., Cong L., Geng J., Han Y., Li L., Li W., Hu G., Huang X., Li W., Li J., Liu Z., Li L., Liu J., Qi Q., Liu J., Li L., Li T., Wang X., Lu H., Wu T., Zhu M., Ni P., Han H., Dong W., Ren X., Feng X., Cui P., Li X., Wang H., Xu X., Zhai W., Xu Z., Zhang J., He S., Zhang J., Xu J., Zhang K., Zheng X., Dong J., Zeng W., Tao L., Ye J., Tan J., Ren X., Chen X., He J., Liu D., Tian W., Tian C., Xia H., Bao Q., Li G., Gao H., Cao T., Wang J., Zhao W., Li P., Chen W., Wang X., Zhang Y., Hu J., Wang J., Liu S., Yang J., Zhang G., Xiong Y., Li Z., Mao L., Zhou C., Zhu Z., Chen R., Hao B., Zheng W., Chen S., Guo W., Li G., Liu S., Tao M., Wang J., Zhu L., Yuan L., and Yang H., 2002, A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Indica*), *Science*, 296(5565): 79-92 <http://dx.doi.org/10.1126/science.1068037> PMID:11935017
- Yue B., and Xing Y.Z., 2005, Progress on molecular and genetic studies of heading date in rice, *Fenzi Zhiwu Yuzhong (Molecular Plant Breeding)*, 3(2): 222-228 (岳兵, 邢永忠, 2005, 水稻抽穗期分子遗传研究进展, *分子植物育种*, 3(2): 222-228)