

大麦糯性相关基因 *Wx* 单核苷酸多态性分析

朱彩梅, 张 京

(中国农业科学院作物科学研究所/农作物基因资源与基因改良国家重大科学工程/农业部作物种质资源与生物技术重点开放实验室, 北京 100081)

摘要: 【目的】大麦 *Wx* 是控制直链淀粉合成相关的糯性基因, 研究大麦糯性相关基因 *Wx* 单核苷酸多态性, 并分析其与籽粒直链淀粉含量的关系。【方法】以 2 个国外糯大麦品种为对照, 对 30 份高、中、低直链淀粉含量的中国大麦进行 *Wx* 的克隆和测序, 分析 *Wx* 的单核苷酸多态性 (SNP) 及其与籽粒直链淀粉含量之间的关系。【结果】在对 32 个大麦品种的核苷酸序列多态性鉴定中, 共检测到了 169 个多态性位点, 平均每 26 bp 检测到一个多态性位点。在所有检测到的多态性位点中, 包括 143 个 SNP 和 26 个 InDel, 二者的频率分别为 1/310 和 1/169。*Wx* 的内含子 1、3、5、8 区, 外显子 2、5 和 5'-UTR 及 3'-UTR 区域为变异富集区, 其它区域变异较小。外显子 2 和内含子 1 区域所承受的选择压力较小。单倍型分析表明, 第 1 种单倍型中包括所有低直链淀粉含量的材料。【结论】大麦 *Wx* 的多态性与直链淀粉含量之间存在明显的对应关系。

关键词: 大麦; *Wx*; 单核苷酸多态性; 单倍型; 直链淀粉含量

Single Nucleotide Polymorphism of *Wx* Gene Associated with Amylose Content in Barley Germplasm

ZHU Cai-mei, ZHANG Jing

(National Key Facility for Crop Gene Resources and Genetic Improvement/Key Laboratory of Crop Germplasm & Biotechnology, Ministry of Agriculture/Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract: 【Objective】 *Waxy* gene (*Wx* in short) is a key gene responsible for the barley amylose synthesis. The objective of the present study was to detect single nucleotide polymorphisms (SNP) of *Wx* and determine the relationship between SNP and amylose content of seed in barley. 【Method】 After PCR amplification and gene clone, 30 Chinese normal and waxy barley cultivars, which had high, medium and low amylose content respectively, were sequenced and polymorphism was analyzed for *waxy* gene with 2 Canadian waxy barley as controls. 【Result】 By polymorphism identification of *waxy* gene nucleotide sequences for the 32 barley cultivars, 169 polymorphism sites were identified, which included 143 SNP and 26 InDel with frequencies of 1/310 and 1/196, respectively. It was found that the introns 1,3,5,8, exons 2,5 and 5'-UTR, 3'-UTR were the rich nucleotide variation regions. Exon 2 and intron 1 suffered stronger selection pressure. Four haplotypes were identified from 32 plant materials. Among them, haplotype 1 included all low amylose content accessions. 【Conclusion】 The results of research showed that there is a clear correlation between SNPs in *waxy* gene and amylose content in barley.

Key words: barley; *Wx*; single nucleotide polymorphism; haplotype; amylose content

0 引言

【研究意义】大麦 (*Hordeum vulgare* L.) 糯性相关基因即 *waxy* (简称 *Wx*) 主要编码颗粒结合型 ADP-葡萄糖转移酶 (E.C.2.4.1.11), 该酶又被称为颗粒结

合型淀粉合成酶 (granule-bound starch synthase, GBSS), 是负责籽粒胚乳和花粉中直链淀粉合成的关键酶^[1-3], 位于大麦 7H 染色体的短臂^[4-6]。普通大麦携带 *Wx* 等位基因的胚乳中淀粉由 27%直链淀粉和 73%支链淀粉组成, 而纯合的 *wx* 胚乳通常含 2%—98%直

收稿日期: 2009-08-24; 接受日期: 2009-11-12

基金项目: 农业部引进国际先进农业科学技术“948”项目 (2006-G9)、农业部公益行业科研项目 (nyhyzx07-010)

作者简介: 朱彩梅, 博士。通信作者张 京, 研究员, 博士。Tel: 010-62186653; E-mail: zhangjing@caas.net.cn

链淀粉和 98%—90%支链淀粉^[7-8]，以拥有粘性胚乳而被称为糯大麦。在大麦胚乳中贮藏淀粉组成是影响籽粒制麦、食用和饲用品质的性质之一。研究大麦 *Wx* 对于揭示 *Wx* 起源及资源的有效利用具有现实意义。中国大麦栽培历史悠久，是栽培大麦的原产地之一，已收集保存近 2 万份大麦遗传资源，这是大麦品质改良育种的重要基因资源^[9]。【前人研究进展】大麦糯性基因已经被克隆和鉴定，大麦 *Wx* 位点的精细结构图也已经被报道^[10]。Domon 等^[11]分析了来自日本和朝鲜的大麦种质资源，发现在 *Wx* 的 5'端前导序列区存在 InDel（插入/删除）突变。Patron 等^[12]在日本的低直链淀粉含量大麦地方品种中发现了 *Wx* 的启动子和 5'非翻译区存在一个 413 bp 的删除序列，该删除序列的出现与否和 *Wx* 的表达直接相关。Domon 等^[13]还对比分析了日本大麦品种的一个非糯大麦品种、一个自发突变的糯大麦品种和一个人工突变的糯大麦的品种的 *Wx* 序列不同位点的碱基序列。单核苷酸多态性（single nucleotide polymorphism, SNP），主要是指在基因组水平上由单个核苷酸的变异所引起的 DNA 序列多态性。SNP 所表现的多态性只涉及到单个碱基的变异，这种变异可由单个碱基的转换（transition）或颠换（transversion）所引起，也可由碱基的插入或缺失所致。具有数量多、分布广、稳定性强、易于分型等优点^[14]。目前，人类疾病的 SNP 研究已取得一定进展^[15]，在玉米、大豆、小麦及其它植物中也发现了 SNP 位点，并表现高度的种内核苷酸多态性^[16]。日本冈山大学佐藤和广教授研究比较了酿造啤酒用的“春名二条”、食用的“赤神力”及里海周边的一种野生大麦，发现这 3 种大麦的染色体上存在着 1 000 多个 SNP 位点^[17]。Bundock 等^[18]的研究结果表明在大麦的 *P450* 基因中也存在 SNP 位点，每 131 个碱基就有一个 SNP 位点。Rostoks 等^[19]的进一步研究表明在大麦的微卫星侧翼序列也有丰富的 SNP 位点的存在。随着基因组测序研究的发展，SNP 已成为生命科学领域中的一个研究热点，并且成为遗传作图、系统进化等研究的有力工具。【本研究切入点】迄今为止，还没有关于中国大麦遗传资源 *Wx* 的核苷酸多态性研究报道，遗传资源材料背景的限制性直接导致了中国在糯大麦育种水平上的滞后性。【拟解决的关键问题】本文以直链淀粉含量不同的普通大麦为材料，研究大麦 *Wx* 的多态性，分析 *Wx* 的多态性与直链淀粉含量之间的关系，为发掘利用优异基因资源、提高大麦品质奠定基础。

1 材料与方法

1.1 材料

共有 32 份试验材料，其中 30 份来源于中国农业科学院国家农作物种质资源保存中心，另有 2 份加拿大糯大麦材料 CDC Alamo 和 CDC Candle 作为对照。材料的基本情况见表 1。

1.2 方法

1.2.1 材料的糯性鉴定和直链淀粉含量（Amylose content, AC）测定 根据普通大麦和糯性大麦胚乳直链淀粉含量的不同，采用 I₂-KI 染色法对每份材料随机选取的 100 粒成熟籽粒横切面染色鉴定，染色后选取颜色表现不同的单粒纯化繁殖，第二年收获单株后，采用近红外法^[20]进行单株直链淀粉含量测定。

1.2.2 目标片段的获得和测序 以 NCBI 中的 *waxy* 基因序列（X07931）为参照，利用 Primer Premier 5.0 软件设计部分引物，另外部分引物来自于前人研究^[21]，所用引物序列见表 2。以供试材料基因组 DNA 为模板，进行 PCR 扩增。将每个目标片段部分连接于 pMD18-T 载体上，转化后挑取阳性克隆再进行测序，另外将部分目标片段回收纯化后直接送生物公司测序。每个目标片段至少检测 2—3 个克隆。根据扩增产物预期片段大小和序列比对结果，把两端大约 50—90 bp 的序列去除，只保留比对结果比较好的部分，每个材料得到约为 4 400 bp 片段。

1.2.3 目标片段 DNA 序列单核苷酸多态性检测 用 DNASTar 软件包内的 SeqMan 程序对测序数据进行质量评价，通过设定组装参数（最小匹配长度为 50 bp，最小匹配率为 95%），对来自同一材料不同克隆的测序结果进行组装，得到一致序列（consensus sequence）。将同一材料测出的不同区域的 Contig，用 DNASTar 软件包内的 SeqMan 程序进行序列拼接。利用 DNA Star 软件包内的 Megalign 软件分别对来自不同材料的序列进行多序列联配（multialignment）（Clustal W），用 DnaSP 4.0 和 MEGA 3.1 软件包对来自不同材料的序列进行单核苷酸多态性及单倍型分析。

2 结果

2.1 品种间的直链淀粉含量差异

糯大麦胚乳由于不含或极少含有支链淀粉，染色后表现为棕红色，而普通大麦胚乳则染为蓝黑色。根据不同品种对 I₂-KI 染色的表现不同（图 1）选取单株，采用近红外法^[20]对所选单株进行直链淀粉含量

表 1 分析 Wx 多态性的大麦材料

Table 1 Barley accessions for detecting Wx gene polymorphism

统一编号 Accession number	品种名称 Cultivar	直链淀粉含量 Amylose content (%)	原产地 Origin	皮/裸 Hulled/Naked	棱形 Row type
ZDM00063	保定钩芒大麦 Baodinggoumangdamai	18.61	河北保定 Baoding,Hebei	裸 Naked	6
ZDM00110	没皮大麦 Meipidamai	23.56	山西高平 Gaoping,Shanxi	裸 Naked	6
ZDM00155	长粒大麦 Changlidamai	16.44	山西阳曲 Yangqu,Shanxi	皮 Hulled	6
ZDM00208	大麦 2 号 Damai2	17.96	山东泰安 Taian,Shandong	皮 Hulled	6
ZDM00418	宁阳三月黄 Ningyangsanyuehuang	23.65	山东宁阳 Ningyang,Shandong	裸 Naked	6
ZDM00454	笨大麦 Bendamai	13.09	山东胶南 Jiaonan,Shandong	裸 Naked	6
ZDM00467	灰巩大麦 Huigongdamai	11.90	山东胶南 Jiaonan,Shandong	裸 Naked	6
ZDM00500	米大麦 Midamai	15.24	山东胶南 Jiaonan,Shandong	裸 Naked	6
ZDM00512	皮大麦 Pidamai	36.00	河南宜阳 Yiyang,Henan	皮 Hulled	6
ZDM00524	长芒 Changmang	33.67	河南灵宝 Lingbao,Henan	皮 Hulled	6
ZDM00576	大麦 Damai	37.84	河南南召 Nanzhao,Henan	皮 Hulled	6
ZDM00604	芒大麦 Mangdamai	33.07	河南泌阳 Biyang,Henan	皮 Hulled	6
ZDM00652	白大麦 Baidamai	33.83	河南封邱 Fengqiu,Henan	皮 Hulled	6
ZDM00690	芒大麦 Mangdamai	26.26	河南鹿邑 Luyi,Henan	皮 Hulled	6
ZDM00693	芒大麦 Mangdamai	8.53	河南项城 Xiangcheng,Henan	皮 Hulled	6
ZDM01130	露仁大麦 Lourendamai	12.83	陕西旬阳 Xunyang,Shaanxi	裸 Naked	6
ZDM01874	寒大麦 Handamai	13.48	江苏睢宁 Suining,Jiangsu	皮 Hulled	6
ZDM02008	埠芒大麦 Fumangdamai	22.80	江苏宜兴 Yixing,Jiangsu	皮 Hulled	6
ZDM02866	盘安红筋大麦 Pananhongjindamai	21.94	浙江东阳 Dongyang,Zhejiang	皮 Hulled	6
ZDM03338	无芒裸麦 Wumangluomai	24.44	浙江丽水 Lishui,Zhejiang	裸 Naked	6
ZDM05845	白青裸 Baiqingke	27.97	西藏贡嘎 Gongga,Tibet	裸 Naked	6
ZDM06712	钩芒青裸 Goumangqingke	24.88	西藏曲松 Qusong,Tibet	裸 Naked	6
ZDM07715	麻阳黄 Mayanghuang	24.68	西藏阿里 Ali,Tibet	裸 Naked	6
ZDM08677	钜野米大麦 Juyemidamai	9.70	山东 Shandong	裸 Naked	6
ZDM09028	三月黄大麦 Sanyuehuangdamai	9.81	贵州施秉 Shibing,Guizhou	裸 Naked	6
ZDM09036	新场大麦 Xinchangdamai	24.27	贵州台江 Taijiang,Guizhou	裸 Naked	6
ZDM09098	短钩裸大麦 Duangouluodamai	13.17	云南禄劝 Luquan,Yunnan	裸 Naked	2
ZDM09113	光芒黄大麦 Guangmanghuangdamai	9.97	云南禄劝 Luquan,Yunnan	裸 Naked	2
ZDM09167	温泉青裸 Wenquanqingke	16.84	云南宁浪 Ninglang,Yunnan	裸 Naked	6
ZDM09222	鉴 24 Jian24	8.59	北京 Beijing	裸 Naked	6
	CDC Alamo	0	加拿大 Canada	裸 Naked	2
	CDC Candle	4	加拿大 Canada	裸 Naked	2

测定, 结果见表 1。近红外检测的 32 份材料的光谱扫描图结果见图 2。结果表明, 所有被检测的 32 份材料间直链淀粉含量差异明显, 其中来自加拿大的 2 份材料 CDC Alamo 和 CDC Candle 的直链淀粉含量分别为 0 和 4%, 另外来自北京的鉴 24、云南的光芒黄大

麦、贵州的三月黄大麦、山东的钜野米大麦、河南项城的芒大麦等 5 份材料的直链淀粉含量均低于 10%, 其它 25 份材料的直链淀粉含量则大于 10%, 最高的为来自河南南召的大麦品种, 其直链淀粉含量达 37.84%。

表 2 Wx 扩增测序用引物

Table 2 The primers for Wx gene sequence

引物名称 Primer	引物序列(5'-3') Primer sequence	扩增区域 Amplified region (bp)	PCR 片段大小 Amplified length (bp)
GBSS11	F:CGCAGAAACGAACAAACAGA R:TTCAGTCAGGTTTGTGCGACTTT	543—1500	958
GBSS12	F:CCGACAAAACCGTGTACTT R:GATCGATCGACGACGAGAAT	1434—2291	857
GBSS13	F:CCTTGCGGATGTCCTCGGC R:GTTGTGGATGCAGAAGGCCACC	2046—2943	897
GBSS14	F:AGGACGCCTGGGACACCAGC R:CGAAAGTCGTCGAAGGAGAAG	2701—3586	885
GBSS15	F:ACGACTGGCACACGGGCCTTC R:TCACTCAATCATCCATCCATCC	3338—4525	1188
GBSS16	F:GGCCAAGGCGCTGAACAAG R:GAGCAAGCGGCACAACAAAAC	4268—5126	859



a: 棕红色; b: 蓝黑色
a: Brownish red; b: Blueness black

图 1 大麦籽粒横切面 I₂-KI 染色对照

Fig. 1 The kernel transverse section of barley with I₂-KI staining

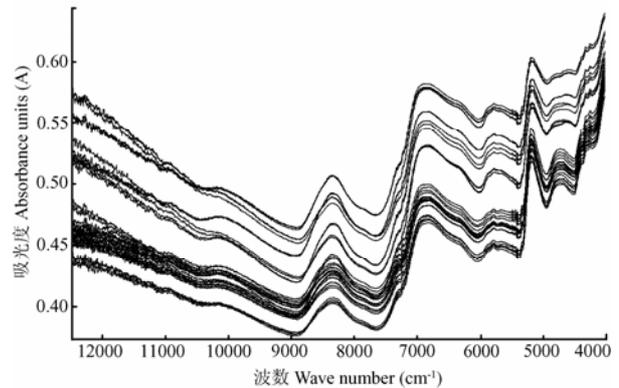


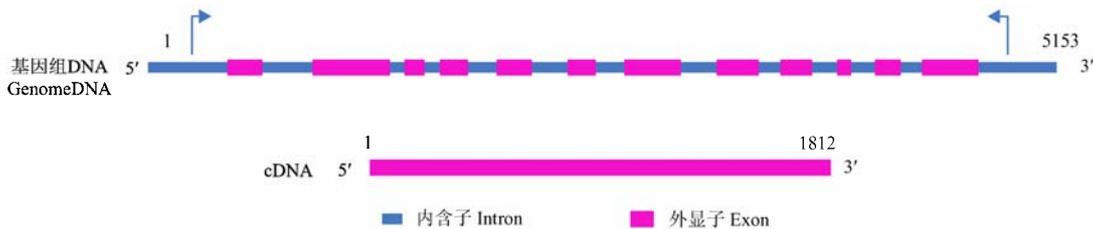
图 2 32 份材料的近红外光谱扫描图

Fig. 2 Raw spectra of 32 accessions

2.2 Wx 序列结构

以 X07931 为对照, 大麦 Wx 全长为 5 153 bp, 包含 12 个外显子 (exon) 和 11 个内含子 (intron) 及 5' 和 3' 端的非编码区 (图 3)。990—1 548 bp, 1 908—1 996 bp, 2 078—2 161 bp, 2 261—2 386 bp, 2 541—

2 676 bp, 2 778—2 883 bp, 3 238—3 329 bp, 3 510—3 603 bp, 3 796—3 877 bp, 3 965—4 061 bp, 4 191—



本试验所得到的 4400 bp 的序列对应于序列 X07931 的 601—5001 bp, 因 InDel 的存在, 对应长度略有不同。图上两个箭头之间所示部分为本试验所测序列
The 4400 bp sequence, which is of different length owing to the existed InDel, corresponds to the 601 to 5001 bp in X07931. The sequence between the two arrows shown in the figure is the sequence determined from this study

图 3 大麦 Wx 结构示意图

Fig. 3 Schematic representation of barley Wx gene

4 266 bp 为该基因的内含子区段; 989 bp 上游部分序列, 1 549—1 907 bp, 1 997—2 077 bp, 2 162—2 260 bp, 2 387—2 540 bp, 2 677—2 777 bp, 2 884—3 237 bp, 3 330—3 509 bp, 3 604—3 795 bp, 3 878—3 964 bp, 4 062—4 190 bp, 4 267—4 380 bp 是该基因的外显子区域。从 1 590 bp 开始为 cDNA 编码, 到 4 383 bp 编码终止, cDNA 区段全长为 1 812 bp。764—769 bp 为该基因的 TATA 盒区域。

2.3 *Wx* 序列多态性

本试验以 X07931 为对照, 克隆了该基因 601—5 001 bp 间的 4 400 bp 碱基序列。对 32 份供试材料的核苷酸多态性分析结果见表 3。在总计 32 个核苷酸序列中, 共检测到了 169 个多态性位点, 平均每 26 bp 检测到一个多态性位点。所有检测到的多态性位点包括 143 个 SNP 和 26 个 InDel, 二者的频率分别为 1/310 和 1/169。其中 25 个 SNP 来自编码区, 118 个来自非编码区, 编码区和非编码区 SNP 的出现频率分别为 1/75 和 1/22, 非编码区 SNP 变异频率约为编码区的 3.4 倍。在所检测出来的 143 个 SNP 中, 有 56 个转换

(transition) 和 86 个颠换 (transversion), 转换和颠换的比值为 0.65。在发现的 26 个 InDel 中, 1 个来自编码区, 25 个来自非编码区, 编码区和非编码区的 InDel 出现频率分别为 1/1 812 和 1/103, 非编码区 InDel 变异频率约为编码区的 17.6 倍。在编码区仅仅出现的 1 个 InDel, 发生在第 2 个外显子, 是大小为 15 bp 的插入片段。该插入片段共在 25 份材料中检测到, 而在其余 7 份材料中未能检测到, 这 7 份材料分别是来自陕西的露仁大麦 (ZDM01130)、江苏的寒大麦 (ZDM01874)、山东的钜野米大麦 (ZDM08677)、云南的光芒黄大麦 (ZDM09113)、北京的鉴 24 (ZDM09222) 和 2 份国外的材料 CDC Alamo、CDC Candle。进一步分析发现, 不含该插入片段的 7 份材料, 直链淀粉含量均表现较低或无 AC。故此推断, 这一插入片段可能与 *Wx* 的表达有关, 进而影响直链淀粉的合成。广义的单核苷酸多态性即包括单碱基突变, 也包括少数几个碱基的插入和缺失, 如果将 InDel 归为 SNP, 则在 *Wx* 整个分析区域的单核苷酸变异频率为 1/26。

表 3 *Wx* 序列多态性分布

Table 3 Polymorphism of *Wx* sequence in 32 barley accessions

	大小 Size (bp)	等位变异数目(个) No. of allele mutation	备注 Comments
非编码区 Noncoding region	2588		
SNP		118	1 个 SNP/22 bp
InDel		25	1 个 InDel/103 bp
总计 Total		143	1 个 polymorphic/18 bp
编码区 Coding region	1812		
SNP		25	1 个 SNP/75 bp
InDel		1	1 个 InDel/1812 bp
总计 Total		26	1 个 polymorphic/72 bp
整个分析区 Complete region	4400		
SNP		143	1 个 SNP/31 bp
Transitions/Transversion ratio			0.65
InDel		26	1 个 InDel/169 bp
总计 Total		169	1 个 polymorphic/26 bp

2.4 *Wx* 序列多态性位点分布

基因序列多样性研究通常以核苷酸多样性 π 值和核苷酸多态性 θ 值为指标, π 值则是衡量同一位点不同序列两两之间的差异的一个参数指标, 而 θ 值是一个衡量群体突变率的参数指标, 与核苷酸变异占序列位点数的比率有关。本研究对 32 份材料的 *Wx* 序列进

行多样性统计分析后发现, *Wx* 的序列多样性在基因各区域中呈现不均匀分布 (图 4, 图 5)。

由图 4 和图 5 可以看出, π 值和 θ 值的变化趋势基本上较为一致。对于不同区段内 π 值与 θ 值变化幅度都比较明显, 在整个测序区段的前 1 800 bp 区段内 π 值和 θ 值均最大, π 值在 0.09—0.26 之间变化, 而 θ

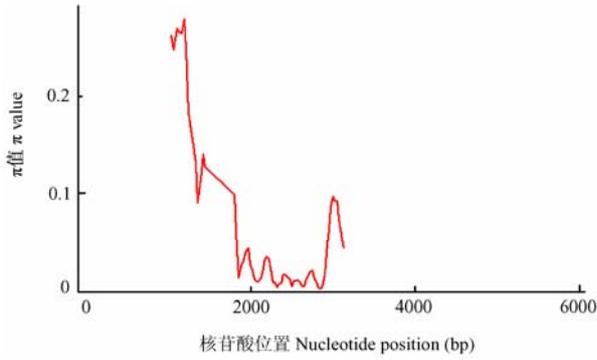


图 4 Wx 区域内的序列多样性 (π)
Fig. 4 The sequence diversity of Wx (π)

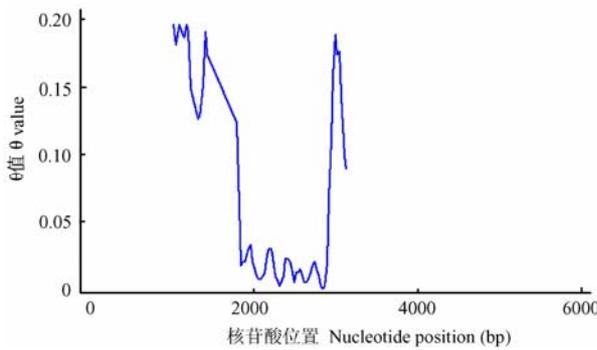


图 5 Wx 区域内的序列多样性 (θ)
Fig. 5 The sequence diversity of Wx (θ)

值在 0.126—0.196 之间变化；在 1 800—2 900 bp 区段 π 值和 θ 值均最低，π 值在 0.000—0.042 之间，θ 值在 0.000—0.029 之间变化；在后 2 600 bp 区段，π 值在 0.042—0.061，而 θ 值在 0.089—0.188 之间变化。

从图中多态性位点变化分布来看，该基因在内含子 1、3、5、8 区，外显子 2、5 和 5'-UTR 及 3'-UTR 区域为变异富集区，其 π 值均超过 0.04，其中内含子 1、2 的核苷酸多样性 π 值超过 0.06，θ 值超过 0.05，而其它内含子区段的核苷酸多样性相对较低；外显子 9、10、11、12 区段核苷酸多样性 π 值均低于 0.03，θ 值则均低于 0.029，其中第 9 外显子区核苷酸多样性最低，2 个指标值均为 0。从分析的结果来看，编码区的遗传变异程度小于非编码区的遗传变异程度。Wx 的多态性位点分布模式可能与各区段所承受的选择压力有关，非编码区承受的选择压力相对较小，因而变异频率较高，而编码区编码有生物学功能的蛋白质，承受的选择压力较大，所以只有很少的核苷酸发生变异，

变异频率较低。

2.5 Wx 编码区多态性分析

对 Wx 编码区核苷酸多态性进行的检测结果如表 4，整个编码区共有 26 个碱基位点发生了变化，主要发生在外显子 2、5、8 和 10 内，而且在不同材料中均有碱基变化，但不同碱基位点发生变化的材料数不同。外显子 2 的多态性碱基位点数最多，共有 18 个，其中有 25 份材料均在第 1 654 碱基处有一个 15 bp 的插入序列，而在其它碱基位点均是单核苷酸变化，其中有 10 个为碱基转换，7 个为碱基颠换，7 个为同义突变，10 个为非同义突变；外显子 5 有 5 个突变位点，均发生为碱基转换，3 个为同义突变，2 个为非同义突变，在第 2 523 bp 处检测到了一处非同义突变，在 CDC Alamo 中出现，碱基由 T 代替了原来的 A，对应的氨基酸则由缬氨酸代替了原来的天冬氨酸，该位点位于蛋白酶活性中心，碱基变化的结果可能是直接导致基因编码的蛋白酶活性丧失的重要原因；外显子 6、8、10 各有一个碱基变化，外显子 8 为同义突变，而外显子 6、10 两个则为非同义突变；外显子 10 有一处非同义突变，为碱基颠换，仅在 CDC Alamo 中检测到，很可能这个位点的变化也是造成 CDC Alamo 发生基因突变导致该品种直链淀粉含量为零的原因之一。

2.6 Wx 序列单倍型分析

利用建树工具软件对 32 份材料的基因序列比对结果进行系统发育树构建和单倍型分析（图 6）。发现 32 个材料被分成了 4 种单倍型。其中第 1 种单倍型又被分成 2 种，第 1 种共包括 10 个品种，其中有 3 个品种，分别为来自西藏的白青稞（ZDM05845）、钩芒青稞（ZDM06712）和浙江的无芒裸麦（ZDM03338）距离较近，另外的 7 个品种与其它品种关系最远，其中 2 个加拿大品种 CDC Alamo 和 CDC Candle 均包括在内，另外江苏的寒大麦（ZDM01874）、山东的钜野米大麦（ZDM08677）、陕西的露仁大麦（ZDM01130）、云南的光芒黄大麦（ZDM09113）、河南的皮大麦（ZDM00512）等 5 个品种也包括在内；除皮大麦外，这 5 个品种的直链淀粉含量均偏低，都在 15% 以下。第 2 种仅包括一个品种鉴 24（ZDM09222），其直链淀粉含量低于 10%。第 2 种单倍型仅包括 2 个品种，均为来自山东的笨大麦（ZDM00454）和灰巩大麦（ZDM00467）。第 3 种单倍型包括的品种较多，共 18 个品种，分为 5 个种，从遗传距离上看，这 5 种关系均较近。第 4 种单倍型仅包括一个品种，为山东的米大麦（ZDM00500）。从

表 4 *Wx* 编码区核苷酸变异Table 4 Nucleotide mutation in coding region of *Wx*

位点 Site	核苷酸变异 Nucleotide mutation (x/y)	区域 Region	氨基酸变化 Amino acid mutation (X/Y)	突变类型 Mutation type	材料份数 Number of accession
1652	a/g	外元 2 Exon 2	R/R	S	9
1654	Insertion 15 bp	外元 2 Exon 2			26
1660	g/a	外元 2 Exon 2	R/H	N	27
1662	c/g	外元 2 Exon 2	P/A	N	27
1679	c/g	外元 2 Exon 2	L/L	S	27
1691	c/g	外元 2 Exon 2	N/K	N	27
1694	g/c	外元 2 Exon 2	P/P	S	27
1697	g/a	外元 2 Exon 2	A/A	S	27
1702	c/g	外元 2 Exon 2	A/G	N	27
1704	g/a	外元 2 Exon 2	A/T	N	27
1707	c/t	外元 2 Exon 2	L/F	N	27
1722	a/g	外元 2 Exon 2	I/V	N	27
1771	g/a	外元 2 Exon 2	S/N	N	27
1797	a/c	外元 2 Exon 2	S/R	N	28
1877	t/c	外元 2 Exon 2	D/D	S	27
1892	c/t	外元 2 Exon 2	L/L	S	27
1895	c/a	外元 2 Exon 2	P/P	S	27
1898	g/a	外元 2 Exon 2	P/P	S	27
2434	a/g	外元 2 Exon 5	T/T	S	28
2440	t/c	外元 2 Exon 5	Y/Y	S	21
2463	g/a	外元 2 Exon 5	S/N	N	5
2482	a/g	外元 2 Exon 5	A/A	S	32
2523	t/a	外元 2 Exon 5	V/D	N	1
2744	c/a	外元 2 Exon 6	N/K	N	5
3362	g/a	外元 2 Exon 8	A/A	S	24
3935	g/t	外元 2 Exon 10	G/W	N	1

x/y 表示碱基由“x”变异为“y”，X/Y 表示氨基酸由“X”变异为“Y”。N 表示非同义突变；S 表示同义突变

x/y means that base x changed to y, X/Y means that amino acid X changed to Y. N: Non-synonymous; S: Synonymous mutation

供试材料聚类结果与其直链淀粉含量之间存在一定程度的对应关系来看，直链淀粉含量相对较低的材料基本上聚在第一大单倍型当中，而且距离较近，说明低直链淀粉含量大麦材料之间可能存在相似的 *Wx* 序列变异。

3 讨论

DNA 序列的变异是物种遗传多样性的前提因素，*Wx* 序列的分析结果为通过遗传转化途径，在分子水平上对其进行有目的改造创造了条件。较普遍的 DNA 序列变异是单个碱基的多态性差异变化，其次为碱基

的插入和缺失片段。越来越多的研究证明，SNP 是基因组中发生频率最高、数量最多而且密度最高的遗传多态性，而 InDel 则是引起基因发生突变并对基因表达产物具有较大影响的遗传多态性。本试验通过对 32 份大麦材料的 *Wx* 核苷酸多态性进行分析，在总计 32 个核苷酸序列中，共检测到了 168 个多态性位点，平均每 26 个 bp 检测到一个多态性位点。所有检测到的多态性位点包括了 143 个 SNP 和 26 个 InDel，二者的频率分别为 1/310 和 1/169。在非编码区的碱基变异分析中，在加拿大糯大麦 CDC Candle 的 *Wx* 序列第 623 碱基处，发现一个 401 bp 的缺失片段，位于 *Wx* 的前

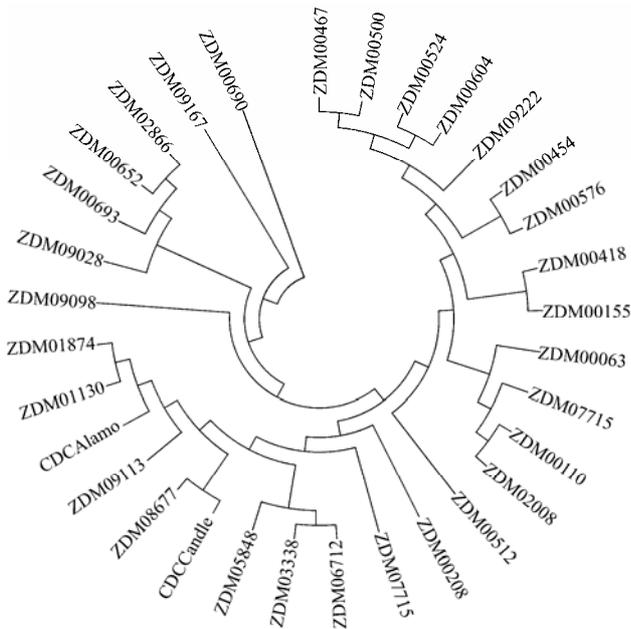


图 6 *Wx* 序列单倍型关系结构树

Fig. 6 Phylogenetic tree representing *Wx* gene haplotype relationship

导序列, 包括 TATA 盒和转录起始位点。这与 Domon 等^[11]的试验结果相同, Domon 等曾推断该片段的缺失可能是引起直链淀粉含量降低和导致胚乳发生糯性突变的主要原因。同时本试验还发现在该品种 *Wx* 序列的第 1 331 碱基处有一个 192 bp 新的插入片段, 由于该插入片段位于第一内含子当中, 可能对 *Wx* 功能表达不影响。Nicola 等^[12]认为另一份加拿大糯大麦对比材料 CDC Alamo 中, 胚乳发生糯性突变则主要是由于编码区第 2 523 和 3 935 碱基位点发生了非同义突变, 本试验同样发现这一突变现象。然而, 在中国所有的低 AC 材料中, 在这几个位点均没有发现与这两份对照材料相同的碱基变异。但是, 在中国低 AC 材料 *Wx* 的前导区却发现少量碱基的插入, 如在第 637 碱基处插入了 5 bp, 757 处插入了 4 bp, 891 处插入了 4 bp。另外在中国低 AC 大麦 *Wx* 的几个内含子中, 也检测到数个大小不等的 InDel 片段。关于在其它作物的 *Wx* 中发现短片段转录因子已有报道^[22-24]。中国的低 AC 大麦材料中出现的这些 InDel 是否影响了其 *Wx* 的功能表达, 有待于进一步研究。

在大麦整个 *Wx* 序列中, 出现了大量的 SNP。Ponomarenko 等^[25]研究指出, 植物中 SNP 的大量出现可能是自然选择的结果, 因为当基因中编码蛋白质的 SNP 随着自然选择 and 时间的推移而得以保留时, 这些

SNP 的变异就很有可能以某种方式帮助该个体成功得以进化。与 X07931 相比, 在中国低 AC 大麦 *Wx* 的编码区序列中, 发现有非同义碱基突变发生。在 1 660 bp 处出现的 SNP (G→A) 导致精氨酸转变成了组氨酸, 在 1 662 处的 SNP (C→G) 导致脯氨酸转变成丙氨酸, 在 1 691 处的 SNP (C→G) 导致天冬酰胺转变成了赖氨酸。另外在 1 702 bp、1 704 bp、1 707 bp、1 722 bp、1 771 bp、1 797 bp、2 463 bp、2 744 bp 处均发生非同义突变。这些碱基突变不仅直接导致其编码氨基酸发生变异, 而且突变氨基酸的生化性质如极性、非极性、酸性、碱性和芳香性等也与原来的有所不同, 可能也是造成直链淀粉含量差异的原因之一。在国外糯大麦 CDC Candle 中, 有 2 处发生非同义突变, 分别位于 2 523 和 3 935 bp 处, 导致其种子胚乳发生糯性突变。这些结果说明, 中国的低 AC 大麦可能与国外的糯大麦材料具有不同的突变机制。另外值得注意的是, 在中国大部分具有中、高 AC 的大麦材料中 (25 份), 1 654 bp 处共同出现了 15 bp 的插入序列, 插入的结果将导致基因编码的氨基酸序列中增加 5 个氨基酸。这一结果与 Patron 等^[12]的研究结果相一致, 但 Patron 等人的研究材料只限于日本的几个大麦品种。结合 Patron 等人的研究, 本试验结果能否说明中国和日本大麦的 *Wx* 起源于共同的祖先, 还有待于进一步扩大材料范围来加以验证。研究发现在 7 份低 AC 供试材料中, 均没有出现该 15 bp 的插入片段, 不仅表明低 AC 大麦存在某些共同的特征, 而且揭示低 AC 性状可能与该片段的插入有关, 但这还需要继续研究确定。有关这方面的试验正在进行当中。试验还发现, 大部分中 AC 与高 AC 的大麦材料之间, 碱基变异、尤其是编码区的碱基序列没有很大差别。

对大麦 *Wx* 碱基序列的多样性分析发现, 在编码区中外显子 2 的多样性指标 π 值和 θ 值均为最大, 其它位点相对较小; 而非编码区内含子 1 的多态性 π 值和 θ 值均为最大。这种多样性分布模式表明, 该基因编码区中外显子 2 区域可能承受的选择压力最小, 而非编码区则以内含子 1 区域承受的选择压力最小。

尽管中国的低 AC 大麦材料与国外的 2 份糯大麦之间, 发生核苷酸变异的位点不尽相同, 产生低 AC 性状的机理可能不同, 但基因序列进化树能够将国外的 2 份糯大麦材料和中国的低 AC 大麦材料聚在一起, 表明国外的糯大麦与中国低 AC 大麦材料具有相似的 *Wx* 序列变异, 但它们是否起源于共同的祖先还有待进一步深入研究证明。

4 结论

从 32 份大麦材料的 Wx 核苷酸序列中共检测到 169 个变异, 包括 143 个 SNP 和 26 个 InDel, 二者出现的频率分别为 1/310 和 1/169; 在编码区发现有 26 个碱基位点发生了变化, 主要发生在外显子 2、5、8 和 10 内, 而且在不同材料中均有碱基变化, 但不同碱基位点发生变化的材料数不同; Wx 的内含子 1、3、5、8 区, 外显子 2、5 和 5'-UTR 及 3'-UTR 区域为变异富集区, 其它区域变异较小。外显子 2 和内含子 1 区域所承受的选择压力较小。大麦 Wx 的单倍型与直连淀粉含量之间存在明显的对应关系。

References

- [1] Sano Y. Differential regulation of waxy gene expression in rice endosperm. *Theoretical and Applied Genetics*, 1984, 68: 467-473.
- [2] Nakamura T, Vrinten P, Hayakawa K, Ikeda J. Characterization of a granule-bound starch synthase isoform found in the pericarp of wheat. *Plant Physiology*, 1998, 118: 451-459.
- [3] Rosichan J, Nilan R A, Arenaz P, Kleinhofs A. Intragenic recombination at the waxy locus in *Hordeum vulgare*. *Barley Genetics Newsletter*, 1979, 9: 79-85.
- [4] Schreiber H, Habeku A. High frequencies of forward and reverse genetic changes at the waxy-locus of barley (*Hordeum vulgare* L.). *Barley Genetics Newsletter*, 1996, 25: 41-45.
- [5] Tabata M. Studies of a gametophyte factor in barley. *The Japanese Journal of Genetics*, 1961, 36: 157-167.
- [6] Kleinhofs A. Integrating barley RFLP and classical marker maps. *Barley Genetics Newsletter*, 1996, 27: 105-112.
- [7] Ishikawa N, Ishihara J, Itoh M. Artificial induction and characterization of amylose-free mutants of barley. *Barley Genetics Newsletter*, 1995, 24: 49-53.
- [8] Hang A, Obert D, Gironella A I N, Burton C S. Barley amylose and β -glucan: Their relationships to protein, agronomic traits, and environmental factors. *Crop Science*, 2007, 47: 1754-1760.
- [9] 张京, 刘旭. 大麦种质资源描述规范和数据标准. 北京: 中国农业出版社, 2005.
Zhang J, Liu X. *Descriptors and Data Standard for Barley (Hordeum vulgare L.)*. Beijing: China Agriculture Press, 2005. (in Chinese)
- [10] Rohde W, Becker D, Salamini F. Structural analysis of the waxy locus from *Hordeum vulgare*. *Nucleic Acids Research*, 1988, 16: 7185-7186.
- [11] Domon E, Fujijita M, Ishikawa N. The insertion/deletion polymorphisms in the waxy gene of barley genetic resources from East Asia. *Theoretical and Applied Genetics*, 2002, 104: 132-138.
- [12] Patron N J, Smith A M, Fahy B F, Hylton C M, Naldrett M J, Rosnagel B G, Denyer K. The altered pattern of amylose accumulation in the endosperm of low amylase barley cultivars is attributable to a single mutant allele of granule bound starch synthase I with a deletion in the 5'-noncoding region. *Plant Physiology*, 2002, 130: 190-226.
- [13] Domon E, Yanagisawa T, Saito A, Takeda K. Single nucleotide polymorphism genotyping of the barley waxy gene by polymerase chain reaction with confronting two-pair primers. *Plant Breeding*, 2004, 123: 225-228.
- [14] Brookes A J. The essence of SNPs. *Gene*, 1999, 234: 177-186.
- [15] Tsui C, Coleman L E, Griffith J L, Bennett E A, Goodson S G, Scott J D, Pittard W S, Devine S E. Single nucleotide polymorphisms (SNPs) that map to gaps in the human SNP map. *Nucleic Acids Research*, 2003, 31(16): 4910-4916.
- [16] Rafalski A. Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology*, 2002, 5(2): 94-100.
- [17] 张可喜. 日本科学家发现大麦 1000 个“单核苷酸多态性”。[2009-3-12]. <http://www2.cas.cn/html/Dir/2002/02/03/9116.htm>
Zhang K X. 1000 SNPs polymorphism were detected in barley by Japanese scientist. [2009-3-12]. <http://www2.cas.cn/html/Dir/2002/02/03/9116.htm>
- [18] Bundock P C, Christopher J T, Egger P, Ablett G, Henry R J, Holton T A. Single nucleotide polymorphisms in cytochrome P450 genes from barley. *Theoretical and Applied Genetics*, 2003, 106(4): 676-682.
- [19] Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson J T, Wanamaker S I, Walia H, Rodriguez E M, Hedley P E, Liu H, Morris J, Close T J, Marshall D F, Waugh R. Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Molecular Genetics and Genomics*, 2005, 274: 515-527.
- [20] 肖听, 谢新华, 陈奕, 罗文永, 陈建伟, 刘彦卓, 毛兴学, 李晓方. 应用近红外透射光谱法测定水稻种子直链淀粉含量的初步研究. *中国农业科学*, 2004, 37(11): 1709-1712.
Xiao X, Xie X H, Cheng Y, Luo W Y, Chen J W, Liu Y Z, Mao X X, Li X F. Application of near infrared transmittance spectroscopy(nits) to the determination of amylose content of rice grain. *Scientia Agricultura Sinica*, 2004, 37(11): 1709-1712. (in Chinese)
- [21] Domon E, Saito A, Takeda K. Comparison of the waxy locus sequence from a non-waxy strain and two waxy mutants of spontaneous and artificial origins in barley. *Genes and Genetic*

- Systems*, 2002, 77(5): 351-359.
- [22] Umeda M, Ohtsubo H, Ohtsubo E. Diversification of the waxy gene by insertion of mobile DNA elements into introns. *The Japanese Journal of Genetics*, 1991, 66: 569-586.
- [23] Bureau T E, Wessler S P. Tourist: A large family of small inverted repeat elements frequently associated with maize genes. *The Plant Cell*, 1992, 4: 1283-1294.
- [24] Fukunaga K, Kawase M, Kato K. Structural variation in the *Waxy* gene and differentiation in foxtail millet [*Setaria italica* (L.) P. Beauv.]: Implications for multiple origins of the waxy phenotype. *Molecular Genetics and Genomics*, 2002, 268(2): 214-222.
- [25] Ponomarenko J V, Merkulova T I, Vasiliev G V, Levashova Z B, Orlova G V, Lavryushev S V, Fokin O N, Ponomarenko M P, Frolov A S, Sarai A. rSNP-Guild, a database system for analysis of transcription factor binding to target sequences: Application to SNPs and site-direction mutations. *Nucleic Acids Research*, 2001, 29(1): 312-316.

(责任编辑 李 莉)