



植物遗传资源学报

Journal of Plant Genetic Resources

ISSN 1672-1810, CN 11-4996/S

《植物遗传资源学报》网络首发论文

题目: 基于 QTL-Seq 的水稻籽粒蛋白含量性状的 QTL 定位
作者: 林添资, 孙立亭, 景德道, 余波, 曾生元, 李闯, 钱华飞, 杜灿灿, 胡庆峰, 杨军, 周义文, 巫章平, 龚红兵
DOI: 10.13430/j.cnki.jpgr.20240812001
收稿日期: 2024-08-12
网络首发日期: 2025-01-03
引用格式: 林添资, 孙立亭, 景德道, 余波, 曾生元, 李闯, 钱华飞, 杜灿灿, 胡庆峰, 杨军, 周义文, 巫章平, 龚红兵. 基于 QTL-Seq 的水稻籽粒蛋白含量性状的 QTL 定位[J/OL]. 植物遗传资源学报.
<https://doi.org/10.13430/j.cnki.jpgr.20240812001>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于 QTL-Seq 的水稻籽粒蛋白含量性状的 QTL 定位

林添资, 孙立亭, 景德道, 余波, 曾生元, 李闯, 钱华飞, 杜灿灿, 胡庆峰, 杨军, 周义文,
巫章平, 龚红兵

(江苏丘陵地区镇江农业科学研究所, 句容 212400)

摘要: 稻米蛋白质含量是影响稻米品质的重要因素之一, 解析稻米蛋白质含量的遗传机制对培育优质食味稻米至关重要。本研究以稻米蛋白质含量有显著差异的镇粳 2400 和嘉禾 218 构建的 RIL 群体为材料, 利用 QTL-Seq 方法对 RIL 群体中的高低蛋白质含量极端家系进行混池重测序以定位稻米蛋白含量 QTL 位点, 之后构建局部遗传连锁图谱, 通过 QTL Ici Mapping 4.1 软件验证结果并精细定位。F_{8:9} 代采用 Δ SNP 指数法分析得到 4 个 QTL, 分别位于第 1、2 和 12 染色体; F_{9:10} 代使用 ED 方法分析得到分布于第 1、2、6、7、8、9 和 11 染色体上共 8 个 QTL, 其中位于第 1 染色体的 35.9~37.8 Mb 区间的 QTL 与 F_{8:9} 代检测到的第 1 染色体 34.4~38.0 Mb 区间的 QTL 重合, 该 QTL 命名为 *qGPC1*, 位于标记 1-3782~1-3834 之间, 物理距离为 516 kb, 表型贡献率和 LOD 值分别为 13.20% 和 3.91。通过测序分析发现, 双亲在已报道的调控籽粒蛋白质含量基因 *OsAAP6* 上没检测到差异, 因此该区间可能有一个新的基因调控稻米蛋白质含量, 同时筛选出低蛋白型优质食味种质镇稻 1818, 米飯食味分显著高于对照品种, 株高、结实率和千粒重比对照品种显著增加。本研究为进一步克隆稻米蛋白质含量基因及解析遗传调控机制奠定了基础, 且提供了可资育种利用的优异种质资源。

关键词: 水稻; 食味品质; 稻米蛋白质含量; QTL 定位

QTL-Seq Analysis for Identification of Protein Content Traits in Rice Grains

LIN Tianzi, SUN Liting, JING Dedao, YU Bo, ZENG Shengyuan, LI Chuang, QIAN Huafei, DU Cancan,
HU Qingfeng, YANG Jun, ZHOU Yiwen, WU Zhangping, GONG Hongbing

(Zhenjiang Institute of Agricultural Sciences Hilly Region of Jiangsu Province, Jurong 212400)

Abstract: Grain protein content is one of the important factors affecting rice eating quality. It is of great significance for rice eating quality improvement to analyze the genetic regulation mechanism of grain protein content. Using an recombinant inbred lines (RILs) population from two japonica rice cultivars, Zhengeng 2400 and Jiahe 218, with significant different protein content of rice grain, BSA-Seq (bulk segregating analysis) method was used to locate the QTLs controlling grain protein content in rice. Subsequently the QTL locus was further verified by QTL mapping using Ici Mapping 4.1 software and narrowed down by fine mapping. Four QTLs were found on chromosomes 1, 2, and 12 in F_{8:9} RIL by Δ SNP index method analysis, and eight QTLs distributed on chromosomes 1, 2, 6, 7, 8, 9, and 11 in F_{9:10} RIL were detected by ED method analysis. A major QTL locus designated as *qGPC1* accounting for 13.2% of the phenotypic variation with a LOD score of 3.91, which was detected both in F_{8:9} and F_{9:10} RIL, was verified by ICIM and fine mapped. Finally, the *qGPC1* locus was delimited in the region of 516 kb between markers 1-3782 and 1-3834. Of them, a reported gene *OsAAP6* regulating grain protein content was no differences in sequencing between zhengeng 2400 and Jiahe 218. Maybe a new gene in this region regulates protein content of rice grain. Moreover, Zhendao 1818 with excellent eating quality and higher plant height, seed setting rate and thousand grain weight compared with the control variety was finally screened.

收稿日期: 2024-08-12

第一作者研究方向为水稻遗传育种, E-mail: 1668432530@qq.com

通信作者: 龚红兵, 研究方向为水稻遗传育种, E-mail: 1179809265@qq.com

基金项目: 江苏省重点研发计划 (BE2021374); 生物育种钟山实验室项目 (ZSBBL-KY2023-01-03); 江苏省种业振兴项目 (JBG2021037, JBG2021038);

Foundation projects: Jiangsu Science and Technology Development Program (BE2021374); Zhongshan Biological Breeding Laboratory Project (ZSBBL-KY2023-01-03); Seed Industry Revitalization Project of Jiangsu Province (JBG2021037, JBG2021038)

This study lays the foundation for further cloning of grain protein content genes and analyzing genetic regulatory mechanisms, and provides excellent germplasm resources for improving rice eating quality.

Key words: rice; eating quality; protein content of rice grain; QTL mapping

水稻是我国最重要的粮食作物之一，随着人民生活水平的提高，人们对优质食味稻米的需求越来越大。但目前国内稻米食味品质改良仍以直链淀粉含量的调控为研究重点，我国的行业标准（NYC593-2013）也仅从稻米的胶稠度、直链淀粉含量和糊化温度来量化稻米蒸煮食味品质。蛋白质作为稻米中第二大营养物质，仅次于淀粉含量，通过影响淀粉的糊化特性及与淀粉分子链的相互作用而影响食味品质^[1-2]。许多研究认为，蛋白质含量与食味品质呈显著负相关，一般总蛋白含量大于 9% 的品种其食味口感通常较差^[3]，蛋白质含量与食味品质之间不是简单的线性关系，在低蛋白含量范围内（蛋白质含量在 3.88%~6.03% 之间），蛋白质含量增高尤其是醇溶蛋白和谷蛋白含量增高的稻米，可提升米饭的综合口感^[4]，蛋白质含量过高或过低都会降低食味品质，只有当蛋白质与直链淀粉相互平衡才会形成优良品质^[5]。因此，稻米蛋白质含量研究有助于解析蛋白质遗传调控机制，对培育优质食味稻米具有重要的指导意义和参考价值。稻米蛋白质含量主要受遗传因子控制，且是由多基因控制的数量性状。目前为止，人们对控制稻米蛋白质含量的 QTL 定位已有大量报道，在水稻 12 条染色体上均有分布。为了得到较为准确的表型，定位群体主要有三类：第一类是重组自交系群体（RIL, recombinant inbred line）：钟明等^[6]和 Wang 等^[7]以珍汕 97 和南洋占杂交构建的 RIL 群体定位到 6 个控制蛋白质含量的 QTL，其中位于第 1 染色体的 *qpc1* 和位于第 2 染色体的 *qpc2* 对糙米和精米蛋白质含量均具有较大的作用；Kepiro 等^[8]利用 Cypress 和 Panda 的 RIL 群体在糙米中定位到 2 个 QTL，精米中定位到 3 个 QTL，其中第 1 和第 4 染色体上的 QTL 同时控制糙米和精米蛋白质含量；张涛等^[9]利用中优早和丰锦的 RIL 群体检测到控制糙米蛋白质含量的 QTL 位点有 6 个，分别位于第 3、6、7、8 和 11 染色体上，联合贡献率为 61.07%，其中位于第 8 染色体的 *qPC-8-1* 对糙米蛋白质含量具有主效作用；黄覃等^[10]利用明恢 63 和优质泰国香米 KDML105 的 RIL 群体分别于 2009 年检测到蛋白质含量相关 QTL 有 2 个，分别位于第 6 和第 11 染色体上，2010 年检测到 3 个，分别位于第 3 染色体和第 12 染色体上；鄢宝等^[11]以华恢 3 号与中国香稻的 RIL 群体为材料，定位到 3 个控制糙米蛋白质含量的 QTL，即 *qpbca*、*qpbcb*、*qpbcc*，分别位于第 4、6、8 染色体上，其中只有 *qpbcc* 连续两年被重复检测到；杨亚春等^[12]用 9311 和日本晴的 RIL 群体定位到一个稳定性较高的 QTL：*qBRPC1.3*，位于第 1 染色体上，同时控制糙米蛋白质含量和精米蛋白质含量；张杰等^[13]利用 Sasanishiki/Habataki 回交 RIL 群体检测到与蛋白质含量相关的 5 个稳定表达的 QTL，分别位于第 1、8、9 和 12 染色体；Park 等^[14]利用 Hwayeong 和 Wandoaengmi6 的 RIL 群体在第 9 染色体上定位到 1 个 QTL：*qPro9*。第二类是染色体片段置换系（CSSL, chromosome segment substitution line）：Yang 等^[15]利用 Sasanishiki 和 Habataki 的 CSSL 鉴定到两个主效 QTL：*qPC-1* 和 *qPC-10*，分别位于第 1 和 10 染色体；赵琳琳等^[16]以云南元江普通野生稻和特青构建的高代回交群体定位到 6 个低蛋白 QTL，分别位于第 1、7 和 10 染色体；Takayuki 等^[17]利用 Koshihikari 和 Nona Bokra 的 CSSL 鉴定到一个稳定的 QTL：*TGP12*，该 QTL 能特异降低蛋白质含量而不影响食味品质；王小雷等^[18]利用昌恢 121 和越光构建 CSSL 群体，鉴定出两个稳定表达的 QTL：*qPGWC1* 和 *qPaT12*，分别位于第 1 和 12 染色体。第三类是自然群体：

有研究者利用水稻自然群体进行全基因组关联分析,挖掘控制稻米蛋白质含量的 QTL,但检测到的 QTL 数量极少^[19-20]。由于稻米蛋白质含量影响因素众多,如环境和栽培措施等,因此对检测到的 QTL 难以进行更深入的研究,以致 QTL 克隆鲜为报道。Peng 等^[21]首次克隆了调控籽粒蛋白质含量的基因 *OsAAP6/qPC1*,该基因编码一个假定的氨基酸通透酶,能够正向调控籽粒蛋白质含量,在 *OsAAP6* 的 5' 非翻译区的顺式调控元件中两个自然变异可能与籼稻中籽粒蛋白质含量的多样性相关;Yang 等^[22]克隆了 *qGPC10* 基因,该基因编码谷蛋白 A2 前体 (*OsGluA2*),该基因启动子上的序列差异导致籼粳亚种蛋白质含量存在差异。因此,进一步挖掘控制稻米蛋白质含量的 QTL 位点或关键基因,将为稻米品质育种提供更多有利基因资源,加快稻米品质育种效率。前人应用的定位群体大多基于双亲分子标记的多态性考虑,采用籼粳杂交策略构建定位群体,尽管鉴定出了一系列的控制稻米蛋白质含量的 QTL,但在育种上有利用价值的 QTL 和高效的分子标记极少。QTL-Seq 是第二代测序技术和混合池分组分析法 (BSA-Seq, bulked segregated sequencing, 又称 QTL-Seq) 综合利用,该方法具有快速、准确和廉价的特点,本研究针对长江中下游粳稻稻米蛋白质含量高且稻米食味欠佳的现状,利用可资育种的低蛋白质含量种质镇粳 2400 和优质品种嘉禾 218 构建的 RIL 群体,采用 QTL-Seq 技术挖掘稻米蛋白质含量相关的 QTL,为稻米食味品质改良提供优异种质和高效分子标记。

1 材料与方法

1.1 试验材料

水稻种质镇粳 2400 是由镇稻育种中间材料经 MNU (甲基亚硝基尿) 诱变而来,属于优质中粳稻,抽穗期 96.6 ± 1.52 天;嘉禾 218 是嘉兴市农业科学研究院与中国水稻研究所联合培育的优质早熟晚粳稻,抽穗期 105.2 ± 1.10 天。以镇粳 2400/嘉禾 218 通过单籽传法获得一套包含 268 个株系的 RIL 群体 ($F_{8,9}$ 代)。南粳 5055 是江苏省农业科学院粮食作物研究所选育的优质食味品种,属于半糯型粳稻;武运粳 23 号是常州市武进区农业科学研究所选育的早熟晚粳稻区试对照品种,达到国标三级优质稻谷标准。

1.2 田间试验

实验材料于 2022 年和 2023 年正季分别种植于江苏丘陵地区镇江农业科学研究所行香试验基地。5 月 15 日播种,6 月 15 日移栽,单苗栽插,每份材料种植四行,每行 12 株,株行距 $16.5\text{cm} \times 19.8\text{cm}$,常规水肥管理。待成熟后收获每个株系的第二行和第三行,边棵不收,共计 20 株,脱粒后室温下储存三个月备用。

1.3 籽粒蛋白质含量测定

成熟的水稻种子经出糙机 (FC2K, Otake, Japan) 脱壳制成糙米,糙米由精米机 (VP-32, Yamamoto, Japan) 研磨制成精米。蛋白质含量测定采用米粒食味计 (RLTA10B-K, SATAKE) 进行测定,每个株系测定三次,取均值作为该株系的表型值进行分析。

1.4 农艺性状、稻米食味分及理化指标测定

镇稻 1818、武运粳 23 号和南粳 5055 在成熟期调查株高、有效穗、每穗粒数、结实率等农艺性状指标，每个样品调查 5 株。参照米饭食味仪（STA1A, SATAKE）使用说明书测定米饭外观、口感和综合评分，每个样品测定四次。垩白粒率和垩白度使用 SC-E 型大米外观品质检测分析系统（万深），每个样品测定 4 次。胶稠度、蛋白质含量和直链淀粉含量委托华智生物技术有限公司测定。镇粳 2400 和嘉禾 218 蛋白质含量分析采用 *t*-测验，镇稻 1818、武运粳 23 号和南粳 5055 农艺性状及稻米品质指标分析采用邓肯测验。

1.5 全基因组重测序 BSA-Seq 关联分析

CTAB 法提取极端家系的单株 DNA 后，等量混合成 F_{8:9} 代的低蛋白极端池、高蛋白极端池和 F_{9:10} 代的低蛋白极端池、高蛋白极端池。通过 Illumina HiSeq 对 F_{8:9} 代两个极端池、F_{9:10} 代两个极端池和两个亲本进行全基因组重测序，亲本和极端池的平均测序深度为 30×。测序数据经过筛选过滤去除低质量数据，得到 Clean Reads；使用 BWA 软件将数据与水稻参考基因组（http://plants.ensembl.org/Oryza_sativa/Info/Index）序列进行比对，通过比对定位 Clean Reads 在参考基因组上的位置，结合 GATK 软件对双亲及混池的序列进行比对和 SNP 读取^[23-24]，QTL 位点搜索采用 Δ SNP 指数法和欧氏距离法（Euclidean distance, ED）两种分析方法进行计算^[25]。全基因组重测序由南京英德尔生物技术有限公司完成。

1.6 局部的遗传图谱构建及 QTL 分析验证

根据重测序结果获得双亲具有差异的基因组结构，开发出特异性 SNP 标记，筛选亲本间有多态的 SNP 标记（表 1），利用这些标记检测 RIL 群体家系的基因型，通过 QTL Ici Mapping 4.1 软件，分析多态性标记之间的遗传连锁关系，构建目标基因所在区间的连锁遗传图谱。SNP 基因型检测委托景肽生物有限公司完成。*OsAAP6* 基因测序委托南京思普金生物技术有限公司完成。

表 1 用于本研究基因定位的 SNP 标记及 *OsAAP6* 基因测序引物

Table 1 SNP markers for gene mapping and *OsAAP6* gene sequencing primers in this study

标记	物理位置(bp)	引物序列 (5'-3')
Marker	Position	Primer sequence (5'-3')
1-3764 F	37643035	GTATTCGTGATCCATTCATGTG
1-3764 Rg	37643035	GAAGGTGACCAAGTTCATGCTGTATGTAACGGTCAGGTATAAGTTCAG
1-3764 Rc	37643035	GAAGGTGCGGAGTCAACGGATTGTATGTAACGGTCAGGTATAAGTTCAC
1-3782 F	37825493	GTGAGCTGCCATGACGCC
1-3782 Ra	37825493	GAAGGTGACCAAGTTCATGCTCGGCGTGGAAGAAGACGA
1-3782 Rg	37825493	GAAGGTGCGGAGTCAACGGATTTCGGCGTGGAAGAAGACGG
1-3806 Fc	38060795	GAAGGTGACCAAGTTCATGCTGGTTTTGTTAATTATCCCGTGACTAC
1-3806 Ft	38060795	GAAGGTGCGGAGTCAACGGATTGGTTTTGTTAATTATCCCGTGACTAT
1-3806 R	38060795	CAAACCATTCCAATACTTCGCA
1-3834 F	38341969	CGTTGTTGTTCCCCCGG

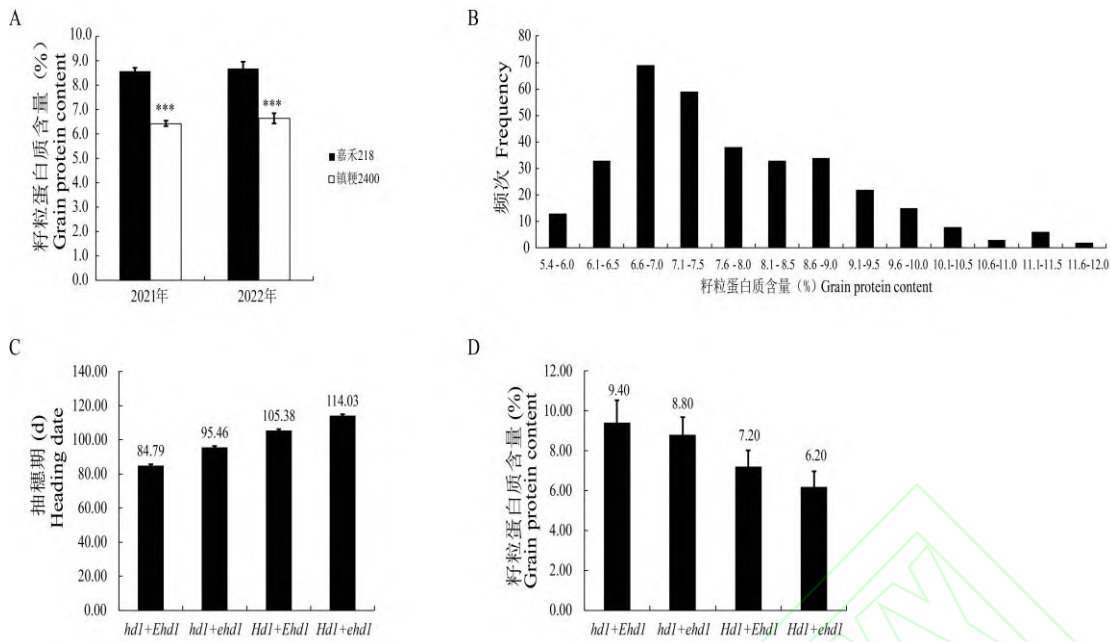
1-3834 Rg	38341969	GAAGGTGACCAAGTTCATGCTGAAGGTCGTGCGCCCTG
1-3834 Ra	38341969	GAAGGTCGGAGTCAACGGATTAGAAGGTCGTGCGCCCTA
1-3876 Fa	38762634	GAAGGTGACCAAGTTCATGCTACATATCTATATAGTATCGGCCTTCAGCA
1-3876 Ft	38762634	GAAGGTCGGAGTCAACGGATTACATATCTATATAGTATCGGCCTTCAGCT
1-3876 R	38762634	AACCGAGAAAGCAGAGCGAA
1-3960 Fg	39602780	GAAGGTGACCAAGTTCATGCTCTGCAAATAGGCGTGATCCATTTG
1-3960 Fa	39602780	GAAGGTCGGAGTCAACGGATTCTGCAAATAGGCGTGATCCATTTA
1-3960 R	39602780	ATCTGCATTGAAAGGGGCCA
OsAAP6-F1	38136496	CTCAAATCCTTCCGCTCAG
OsAAP6-R1	38137749	CCACTTGGACTTGGTAGGGTAT
OsAAP6-F2	38137597	TTACTAGGACAACGTGAAGGATAG
OsAAP6-R2	38138955	AACAAAATTCAGTTTCGTCAGTG
OsAAP6-F4	38139928	GCGTCACGACCGTCTTCTAC
OsAAP6-R4	38140674	ACCATACTCCTCAAGTCCCAT
OsAAP6-F5	38138660	GATGGAATCCTGGAGATCCTTTTG
OsAAP6-R5	38139382	TCGTAGCGGTCGTGATGGTGT
OsAAP6-F6	38138943	ACGCAAGAATGGCACTGACGAAA
OsAAP6-R6	38140201	CCAGCTTGCACAGCGTGAACC
OsAAP6-1-f	38134303	GGTGAAGGAAAACCACAGCCG
OsAAP6-1-r	38135699	CCGCCACGAAGCACACATTG
OsAAP6-2-f	38135363	ATATGACGACACGTGAGGCTGT
OsAAP6-2-r	38136838	TCAACGGTATGTGGTGACAATG

2 结果与分析

2.1 双亲及 RIL 群体的籽粒蛋白质含量表型

镇粳 2400 连续两年的籽粒蛋白质含量都在 6.5% 左右, 而嘉禾 218 的籽粒蛋白质含量连续两年都在 8.6% 左右, 双亲的籽粒蛋白质含量在年间表现较为稳定, 且双亲连续两年都存在极显著差异 (图 1A)。重组自交系 (RIL) 的籽粒蛋白质含量分布于 5.4% ~ 12.0%, 中值偏向低蛋白质含量, 表现出连续分布和显著的超亲分离 (图 1B), 表明该性状是由多基因控制。

该 RIL 家系抽穗期变化幅度较大, 前期研究表明控制抽穗期的主效基因为 *Hd1* 和 *Ehd1*^[26], 而家系蛋白质含量在 *Hd1* 和 *Ehd1* 不同基因型组合下也有差异 (图 1D), 具体表现为 *hd1+Ehd1* 抽穗期 84.79 天, 家系蛋白质含量均值为 9.40%, *hd1+ehd1* 抽穗期为 95.46 天, 家系蛋白质含量均值为 8.80%, *Hd1+Ehd1* 抽穗期为 105.38 天, 家系蛋白质含量均值为 7.20%, *Hd1+ehd1* 抽穗期为 114.03 天, 家系蛋白质含量均值为 6.20% (图 1C、D), 表现出随抽穗期的增加蛋白质含量降低。



A: 镇梗 2400 和嘉禾 218 的籽粒蛋白质含量, 平均值±标准差, n=3, *** P<0.001 (t-测验); B: RIL 家系的籽粒蛋白质含量分布; C: RIL 家系抽穗期, 平均值±标准差; D: 不同抽穗期类型的 RIL 家系蛋白质含量, 平均值±标准差

A: Grain protein content in Zhengeng 2400 and Jiahe 218, Means ±SD, n=3, *** P<0.001 (t-test); B: Grain protein content distribution in RIL; C: Heading date of RIL, Means ±SD; D: Grain protein content of RIL with different heading date, Means ±SD

图 1 亲本和 RIL 的籽粒蛋白质含量及抽穗期分布

Fig.1 Grain protein content in parent, RIL and the heading date distribution

2.2 RIL 家系蛋白质含量极端选取及 QTL-Seq 混池重测序

由于 RIL 家系的蛋白质含量随抽穗期的变化而变化(图 1C、D), 当蛋白质含量在 7.5% 时, 在 *hd1+Ehd1* 条件下, 属于低蛋白质含量, 而在 *Hd1+ehd1* 条件下, 属于高蛋白质含量(图 2), 为排除抽穗期对定位蛋白质含量相关 QTL 的影响, 挑选的家系高低蛋白极端应保证每个抽穗期类型里都有分布, 因此根据抽穗期制定不同标准(表 2), 根据此标准, 从 $F_{8,9}$ 代的 RIL 中挑选出低蛋白质含量的极端株系 28 个, 高蛋白质含量的极端株系 25 个; 从 $F_{9,10}$ 代的 RIL 中挑选出低蛋白质含量的极端株系 22 个, 高蛋白质含量的极端株系 25 个。

对亲本和 4 个混池进行全基因组重测序, 亲本和极端池的测序深度为 30x, 获得的高质量碱基量如下: 镇梗 2400 为 11.82Gb, 嘉禾 218 为 10.67 Gb, $F_{8,9}$ 代高低混池分别为 12.51 Gb 和 11.57 Gb, $F_{9,10}$ 代高低混池分别 13.33 Gb 和 12.90 Gb, GC 含量在 41%~43% 之间, Q20 均大于 96%, Q30 均大于 91%, 测序数据充足, 质量较高, 将过滤后的序列比对到水稻日本晴参考基因组上, 6 个样品与参考基因组的比对率在 98.5%~99% 之间, 比对率高, 可用于后续目标 QTL 的定位(表 3)。经过序列比对和 SNP 位点提取, 将 $F_{8,9}$ 代和 $F_{9,10}$ 代极端混池共得到的 SNP 和 InDel 位点与两亲本进行比较, 过滤不一致的位点后, $F_{8,9}$ 代共有 978999 个 SNP 和 InDel 位点, $F_{9,10}$ 代共有 938767 个 SNP 和 InDel 位点用于 QTL-Seq 分析。

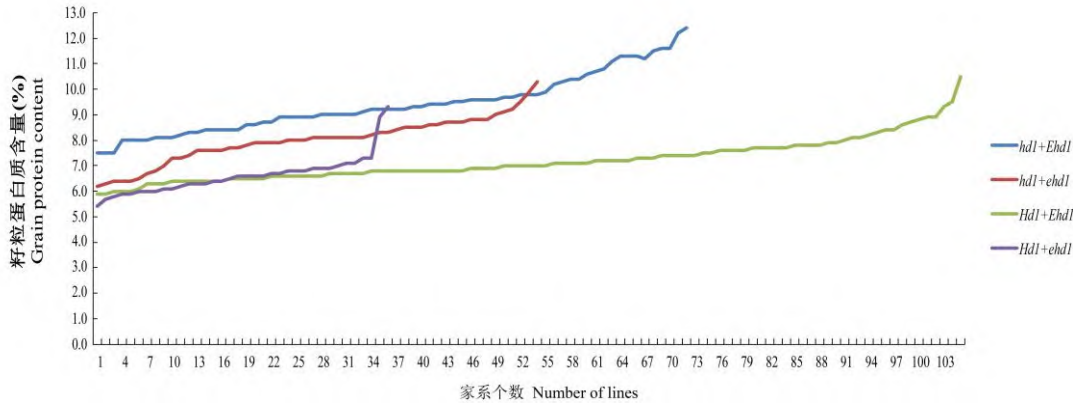


图 2 不同抽穗期类型的家系蛋白质含量分布

Fig.2 Distribution of protein content in RILs with different heading date types

表 2 蛋白质含量极端家系挑选标准

Table 2 Selection criteria for lines with extreme protein content

抽穗期类型	高蛋白极端家系含量(%)	低蛋白极端家系含量(%)
Genotype of heading date	Extreme lines with high protein content	Extreme lines with low protein content
<i>hd1+Ehd1</i>	≥11.0	≤7.5
<i>hd1+ehd1</i>	≥9.0	≤7.0
<i>Hd1+Ehd1</i>	≥8.5	≤6.5
<i>Hd1+ehd1</i>	≥8.0	≤6.0

表 3 全基因组重测序结果

Table 3 Result of whole genome resequencing

样品名	镇粳 2400	嘉禾 218	F _{8,9} 低蛋白池	F _{8,9} 高蛋白池	F _{9,10} 低蛋白池	F _{9,10} 高蛋白池
Sample	Zhengeng	Jiahe 218	F _{8,9} DP-Pool	F _{8,9} HP-Pool	F _{9,10} DP-Pool	F _{9,10} HP-Pool
原始测序读数	79381740	71597982	77642488	83963220	86657424	89637122
Raw reads	79381740	71597982	77642488	83963220	86657424	89637122
原始数据的总碱基数	11907261000	10739697300	11646373200	12594483000	12998613600	13445568300
Raw bases	11907261000	10739697300	11646373200	12594483000	12998613600	13445568300
质量清理后读数	78924050	71234480	77269480	83530060	86085918	88936900
Cleaned reads	78924050	71234480	77269480	83530060	86085918	88936900
过滤后数据的总碱基数	11817033900	10667141054	11571778444	12509203080	12899711423	13327540729
Cleaned bases	11817033900	10667141054	11571778444	12509203080	12899711423	13327540729
碱基质量值	97.08	97.32	97.42	97.21	97.19	96.96
Q20 (%)	97.08	97.32	97.42	97.21	97.19	96.96

碱基质量值	91.89	92.40	92.58	92.12	92.22	91.81
Q30 (%)						
GC 含量 (%)	42	41	41	41	43	43
GC content						
与参考基因组的比对率 (%)	99.20	98.84	98.78	98.54	99.00	99.00
Comparison rate with reference genome						

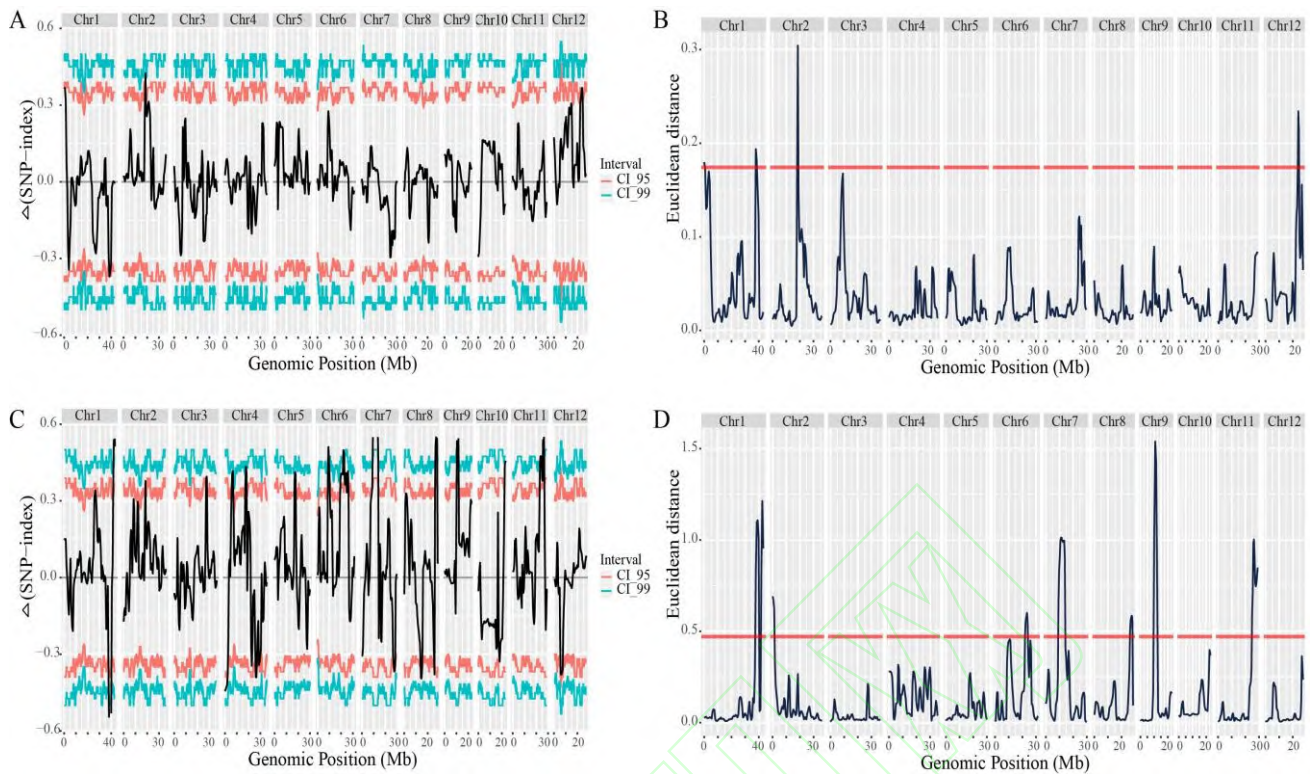
2.3 籽粒蛋白质含量 QTL 定位

分别利用 ED 和 SNP 指数两种关联算法对重测序数据进行关联分析。F_{8,9} 代采用 Δ SNP 指数法分析后在第 1、2 和 12 染色体上共得到 4 个 QTL，第 1 染色体检测到 2 个 QTL，第 2 和 12 染色体分别检测到 1 个 QTL（表 4，图 3A、B）。进一步将 F_{9,10} 代进行 QTL-Seq 分析发现，采用 Δ SNP 指数法分析在 1-12 染色体上均有显著的 QTL，而使用 ED 分析方法分析后，在第 1、2、6、7、8、9 和 11 染色体上共得到 8 个 QTL（表 4，图 3C、D），第 1 染色体检测到 2 个 QTL，其中位于 35.9 ~ 37.8 Mb 区间的 QTL 被重复检测到。

表 4 QTL-Seq 检测到的控制稻米蛋白质含量的 QTL 区间

Table 4 Grain protein content QTLs detected by QTL-Seq

世代	染色体	起始 (bp)Start	终止 (bp)End	Δ SNP 峰值	Δ SNP 峰值位置 (bp)	片段大小(Mb)
Generations	Chromosome			Δ SNP peak value	Δ SNP peak position	Size
F _{8,9}	1	21547	1975210	0.83	304376	1.9
	1	36072210	39871484	-0.89	38762634	3.6
	2	16395716	19999826	-0.86	19801825	3.6
	12	21143708	24994311	0.77	23162720	3.9
F _{9,10}	1	37643035	39609020	-0.96	38974500	1.9
	1	41135157	43226829	-0.95	42529885	2.1
	2	17321	2481618	-0.96	1585992	2.5
	6	21023225	24999963	0.96	21359081	4.0
	7	10000071	14998576	0.97	10572722	5.0
	8	25022875	28433855	0.96	26870574	3.4
	9	9000387	11997888	-0.96	11300356	3.0
	11	24066710	29018426	0.97	26390493	4.9



A 和 C 分别是 RIL 家系 F_{8,9} 和 F_{9,10} 基于 SNP/InDel 的 $\Delta(\text{SNP-index})$ 关联分析，红线代表 95% 的阈值线，蓝线代表 99% 的阈值线；B 和 D 分别是 RIL 家系 F_{8,9} 和 F_{9,10} 基于 SNP/InDel 的 ED 关联分析，红线代表 99% 的阈值线

A and C: $\Delta(\text{SNP-index})$ association analysis diagram based on SNP/InDel for RIL F_{8,9} and F_{9,10}, respectively, the red line indicates the 95% threshold line, the blue line indicates the 99% threshold line; B and D: ED association analysis figures based on SNP/InDel for RIL F_{8,9} and F_{9,10}, respectively, the red line indicates the 99% threshold line.

图 3 QTL-Seq 分析结果

Fig.3 Result of the QTL-Seq analysis

2.4 QTL 位点的验证及精细定位

利用全基因组重测序数据获得的亲本间基因组差异数据，在第 1 染色体 35.9~37.8 Mb 的区间内设计 SNP 标记，共获得 6 对有多态的引物（表 1）。利用这些多态性 SNP 标记对 RIL 群体进行基因型鉴定，通过 QTL IciMapping 4.1，构建了控制籽粒蛋白质含量基因的局部遗传连锁图谱，并结合各 RIL 家系蛋白质含量进行 QTL 精细定位，最终在第 1 染色体检测到一个调控籽粒蛋白质含量的 QTL 位于标记 1-3782~1-3834 之间，物理距离为 516kb，表型贡献率和 LOD 值分别为 13.20% 和 3.91（图 4），我们将其命名为 *qGPC1*。

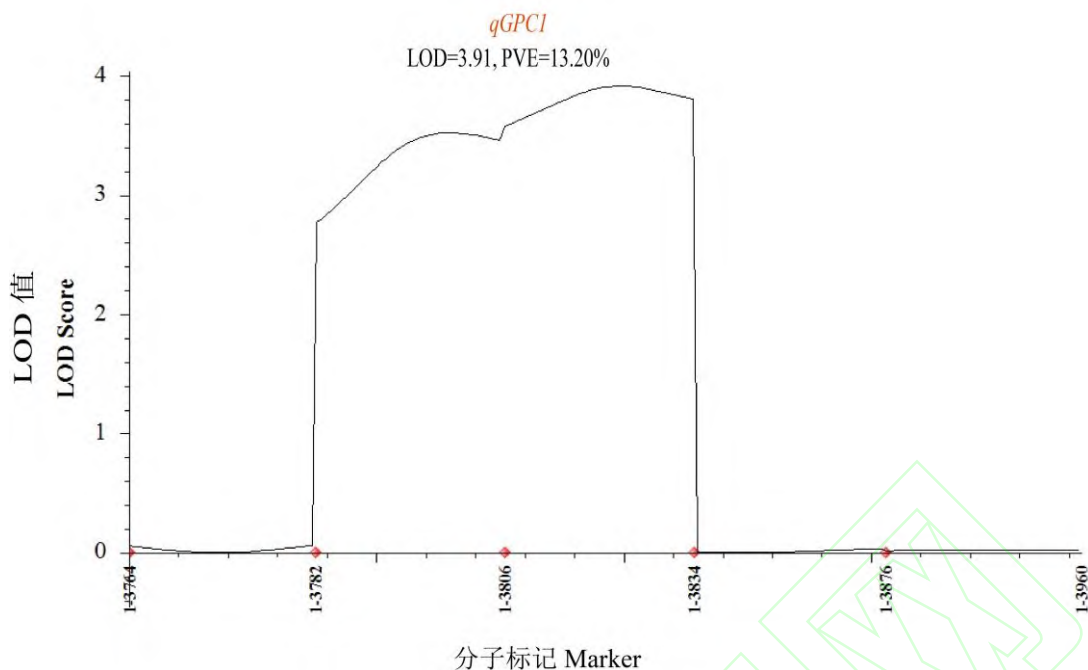


图 4 *qGPC1* 在染色体上的位置及遗传效应

Fig.4 Location and genetic effect of *qGPC1*

2.5 候选基因分析及测序验证

通过与 Peng 等^[21]的相关研究结果比较,发现 *qGPC1* 位点含有已克隆的籽粒蛋白质含量调控基因 *qPC1*, 该基因编码一个假定的氨基酸通透酶 *OsAAP6*, 能够正向调控籽粒蛋白质含量, 在 *OsAAP6* 的 5'UTR 区的顺式调控元件中存在两个自然变异可能主要与籼稻中籽粒蛋白质含量的多样性相关。为检测该基因是否在双亲之间存在差异, 本研究对镇粳 2400 和嘉禾 218 的 *OsAAP6* 基因进行测序, 结果发现该基因的 5'UTR 区已报道的两个自然变异在双亲间无差异 (图 5A), 且该基因的启动子区和编码区在双亲间也没检测到差异 (图 5A、B), 这说明 *qGPC1* 位点存在一个新的基因调控籽粒蛋白质含量。

利用重测序数据对双亲在 516 kb 区间内进行序列比对分析发现, 排除发生在基因间、基因上下游、内含子区突变及同义突变, 共有 12 个基因发生 SNP 或 InDel 变异, 利用水稻表达谱数据库 (<http://ricexpro.dna.affrc.go.jp/index.html>) 进行基因功能预测, 结果表明这 12 个基因中有 4 个为保守假定蛋白基因, 1 个表达蛋白基因, 7 个有注释功能的基因 (表 5), 这些突变可能与稻米蛋白质含量相关。

LOC_Os01g65560	SNP	保守假定蛋白
LOC_Os01g65690	SNP	类 4,5-DOPA 双加氧酶外二醇蛋白
LOC_Os01g65720	SNP	糖苷水解酶
LOC_Os01g65770	SNP	保守假定蛋白
LOC_Os01g65780	SNP	糖基转移酶
LOC_Os01g65810	SNP	表达蛋白
LOC_Os01g65920	SNP/InDel	富含亮氨酸重复、含半胱氨酸的蛋白
LOC_Os01g66060	SNP	保守假定蛋白

2.6 优异种质镇稻 1818 的创制

通过前期对 RIL 各家系农艺性状及稻米食味品质及理化特性调查^[26], 筛选出优异种质镇稻 1818, *qGPC1* 位点来自镇粳 2400, 大米外观与对照品种武运粳 23 号和南粳 5055 相比, 表现为更透明(图 6), 镇稻 1818 株高、结实率和千粒重比两者显著增加, 每穗粒数与南粳 5055 无显著差异, 生育期和有效穗与二者无显著差异(表 6)。通过稻米食味品质和理化特性分析发现, 镇稻 1818 的米饭食味分、胶稠度和直链淀粉均显著高于两对照品种, 而垩白粒率、垩白度和蛋白质含量均显著低于武运粳 23 号和南粳 5055(表 7)。



图 6 镇稻 1818 与对照品种大米外观分析

Fig.6 Rice appearance analysis of Zhendao 1818 and control varieties

表 6 新种质镇稻 1818 的农艺性状

Table 6 The agronomic traits of the new germplasm Zhendao 1818

品种名称	株高(cm)	每公顷有效穗 ($\times 10^4$)	每穗粒数	结实率(%)	千粒重(g)	生育期(d)
Variety	Plant height	Effective panicles per hm^2	Spikelets per panicle	Seed setting rate	Thousand-grain weight	Growth period
镇稻 1818	101.2 \pm 0.84 ^a	325.7 \pm 32.58 ^a	113.6 \pm 11.19 ^b	92.8 \pm 2.32 ^a	26.7 \pm 0.21 ^a	160.8 \pm 1.92 ^a
武运粳 23 号	97.6 \pm 1.52 ^b	321.0 \pm 20.10 ^a	135.2 \pm 16.75 ^a	89.8 \pm 0.47 ^b	26.0 \pm 0.40 ^b	159.2 \pm 1.30 ^a
南粳 5055	96.0 \pm 1.41 ^b	309.0 \pm 24.60 ^a	130.8 \pm 15.22 ^{ab}	90.2 \pm 0.75 ^b	24.6 \pm 0.29 ^c	159.4 \pm 1.67 ^a

小写字母表示不同品种间差异在(P<0.05)的显著性水平; 下同

Lower case letters indicated the level of significance at (P<0.05) of differences between different varieties; The same as below

表 7 新种质镇稻 1818 的米饭食味特性及稻米理化指标分析

Table 7 Eating quality of rice and physicochemical indicators of the new germplasm Zhendao 1818

品种名称	米饭食味分	垩白粒率 (%)	垩白度(%)	蛋白质含量 (%)	胶稠度 (mm)	直链淀粉含量(%)
Variety	Overall eating quality	Rate of chalky kernel	Chalkiness	Protein content	Gel consistence	Amylose content
镇稻 1818	84.85 \pm 0.39 ^a	14.21 \pm 0.85 ^c	3.73 \pm 0.22 ^c	5.63 \pm 0.04 ^c	81.50 \pm 0.41 ^a	17.94 \pm 0.25 ^a
Zhendao 1818						
武运粳 23 号	79.70 \pm 0.65 ^c	27.63 \pm 0.48 ^b	9.93 \pm 0.46 ^a	6.50 \pm 0.04 ^b	65.63 \pm 0.31 ^c	17.30 \pm 0.13 ^b
Wuyungeng23hao						
南粳 5055	82.60 \pm 0.25 ^b	34.60 \pm 2.06 ^a	8.50 \pm 1.29 ^b	7.27 \pm 0.06 ^a	76.33 \pm 0.43 ^b	8.61 \pm 0.14 ^c
Nangeng5055						

3 讨论

光温对稻米蛋白质含量的积累有显著影响, 特别在抽穗至成熟期的有效积温和光照时间对稻米蛋白质含量积累存在显著影响, 表现为随有效积温减少蛋白质含量提高^[27]。本研究中 RIL 家系的抽穗期变异幅度较大, 最大相差 32.09 天(图 1C), 蛋白质含量与抽穗期呈负相关, 抽穗期越早, 蛋白质含量越高(图 1D), 但在相同抽穗期内, 各家系的蛋白质含量仍表现出多样化(图 2), 这说明该 RIL 家系的蛋白质含量存在稳定遗传。

本研究中, 利用镇粳 2400 和嘉禾 218 的 RIL 群体, 在 F_{8,9} 代检测到 4 个 QTL, 分别位于第 1、2 和 12 染色体, F_{9,10} 代检测到 8 个 QTL, 分布于第 1、2、6、7、8、9 和 11 染色体上。与前人研究结果相比较,

钟明等^[6]检测到的位于第 1 染色体的 *qpc1* 位于标记 RM472~RM104, 相对于日本晴物理距离为 37889084 bp ~ 40167060 bp; Wang 等^[7]检测到一个控制多个氨基酸含量的主效 QTL, 位于标记 RM315~ RM104, 物理距离为 36734135 bp ~ 40167060 bp, 这两个 QTL 与本研究中的位于第 1 染色体的 *qGPC1* (36072210 bp ~ 39871484 bp) 位点属于同一位点, 这说明该 QTL 确实存在且稳定遗传。王小雷等^[18]定位到的位于第 12 染色体的 *qPaT12* 位于分子标记 RM3331~RM5479, 物理距离为 23460827 bp ~ 24379166 bp, 与 F_{8,9} 代检测到的位于第 12 染色体的 QTL (物理位置 21143708 bp ~ 24994311 bp) 位点重合。杨亚春等^[12]检测到的控制糙米蛋白质含量的 QTL 位点 *qBRPC6.2* 位于分子标记 RM20076~RM3827 (物理位置: 16811995 bp ~ 22297320 bp) 与 F_{9,10} 代检测到的位于第 6 染色体的 QTL (物理位置: 21023225 bp ~ 24999963 bp) 部分重合, 本研究中 F_{8,9} 代检测到 4 个 QTL 和 F_{9,10} 代检测到 8 个 QTL 中, 只有第 1 染色体的 *qGPC1* 被重复检测到, 其他位点均不相同, 这说明稻米蛋白质含量的遗传变异受主效基因调控的同时, 还受到微效基因和环境因素的影响。有研究表明, 蛋白质含量与直链淀粉存在相关性^[28], 本研究中 *qGPC1* 靠近垩白基因 *WCR1*^[29]; F_{9,10} 代检测到的位于第 6 染色体的 QTL 靠近 *Du13* 基因, *Du13* 基因编码 C₂H₂ 锌指蛋白, 通过影响 Wx^b 的剪接效率进一步调控稻米胚乳中直链淀粉含量^[30]。F_{9,10} 代检测到的第 2 染色体 QTL 与 *OsBE11b* 基因紧密连锁, *OsBE11b* 基因编码淀粉分支酶 IIb, 对水稻胚乳中淀粉结构起重要作用^[31]; F_{9,10} 代检测到的第 8 染色体 QTL 与淀粉去分支酶基因 *OsISA1* 连锁, *ISA1* 影响淀粉合成和胚乳发育^[32], 这些结果验证蛋白质含量与淀粉合成相关基因的遗传关系密切, 其遗传调控机制需进一步研究。

通过与前人研究结果相比较发现, *qGPC1* 位点含有已克隆的基因 *OsAAP6/qPC1*, 该基因编码一个假定的氨基酸通透酶, 能够正向调控籽粒蛋白质含量, 该基因 5' 端潜在的顺式作用元件的 2 个自然变异导致了籼稻蛋白质含量的多样性^[21]。通过对镇粳 2400 和嘉禾 218 的 *OsAAP6* 基因测序, 结果发现该基因的 5' 端已报道的两个自然变异在双亲间无差异, 且该基因的编码区和编码区上游 2kb 也没检测到差异, 这暗示着, 该 QTL 可能存在一个新的基因调控粳稻籽粒蛋白质含量。进一步将 *qGPC1* 定位到 516 kb 区间, 通过双亲全基因组重测序分析发现, 共有 12 个基因发生了 SNP 或 InDel 变异, 其中 LOC_Os01g65200 编码硝酸盐转运蛋白, 该基因过表达会增加水稻叶鞘中的硝酸盐含量和总氮量^[33], 该基因对籽粒蛋白质含量的调控作用未知, 因此候选基因的筛选与功能验证还需进一步的实验证明。本研究从 RIL 家系中筛选出优异种质镇稻 1818, 与生产上大面积推广的品种武运粳 23 号 (推广面积 1000 万亩以上) 和南粳 5055 (推广面积 2000 万亩以上) 相比表现出更高的食味品质, 可直接用于育种, 结合分子标记辅助选择可达到基因优化配置的目的。

参考文献

- [1] Xie L H, Chen N, Duan B W, Zhu Z W, Liao X Y. Impact of proteins on pasting and cooking properties of waxy and non-waxy rice. *Cereal Science*, 2008, 47 (2):372-379
- [2] Saleh M I. Protein-starch matrix microstructure during rice flour pastes formation. *Cereal Science*, 2017, 74:183-186
- [3] Ning H F, Qiao J F, Liu Z H, Lin Z M, Li G H, Wang Q S, Wang S H, Ding Y F. Distribution of proteins and amino acids in milled and brown rice as affected by nitrogen fertilization and genotype. *Journal of Cereal Science*, 2010, 52(1):90-95

- [4] 王琦. 粳稻蒸煮食味品质形成的理化基础研究. 南京: 南京农业大学, 2016
- Wang Q. Physical and chemical foundation for cooking and eating quality of Japonica rice. Nanjing: Nanjing Agricultural University, 2016
- [5] 钱春荣,冯延江,杨静,刘海英,金正勋. 水稻籽粒蛋白质含量选择对杂种早代蒸煮食味品质的影响.中国水稻科学, 2007, 21(3): 323-326
- Qian C R, Feng Y J, Yang J, Liu H Y, Jin Z X. Effects of protein content selection on cooking and eating properties of rice in early-generation of crosses. Chinese Journal of Rice Science, 2007, 21(3): 323-326
- [6] 钟明, 王令强 罗利军, 何予卿. 利用 RIL 群体比较定位糙米和精米蛋白质含量的 QTL.分子植物育种, 2007, 5(5): 631-638
- Zhong M, Wang L Q, Luo L J, He Y Q. Comparison of quantitative trait loci controlling the protein content of brown and milled rice using a recombinant inbred line population. Molecular Plant Breeding, 2007, 5(5): 631-638
- [7] Wang L Q, Zhong M, Li X H, Yuan D J, Xu Y B, Liu H F, He Y Q, Luo L J, Zhang Q F. The QTL controlling amino acid content in grains of rice (*Oryza sativa*) are co-localized with the regions involved in the amino acid metabolism pathway. Molecular Breeding, 2008, 21:127-137
- [8] Kepiro J L, McClung A M, Chen M H, Yeater K M, Fjellstrom R G. Mapping QTLs for milling yield and grain characteristics in a tropical japonica long grain cross. Journal of Cereal Science, 2008, 28 (2): 477-485
- [9] 张涛, 郑家奎, 吴先军, 蒋开锋, 杨乾华, 陈温福, 杨莉. 水稻糙米蛋白质含量的QTL定位. 分子植物育种, 2009, 7(1): 67-72
- Zhang T, Zheng J K, Wu X J, Jang K F, Yang Q H, Chen W F, Yang L. QTL mapping of brown rice protein content in a RIL population of rice. Molecular Plant Breeding, 2009, 7(1): 67-72
- [10] 黄覃, 于波, NASSIROU Tondi-yacouba, 高冠军, 张庆路, 何予卿. 水稻糙米蛋白质和粗脂肪含量的 QTLs 分析. 湖北农业科学, 2012, 51(21): 4709-4713
- Huang Q, Yu B, Nassirou T Y, Gao G J, Zhang Q L, He Y Q. QTLs mapping of protein content and crude fat content in rice. Hubei Agricultural Sciences, 2012, 51(21): 4709-4713
- [11] 鄢宝,王岩,高冠军,张庆路,刘鑫, 何予卿. 水稻糙米蛋白质含量 QTL 定位及上位性分析.分子植物育种, 2012, 10(5): 594-599
- Yan B, Wang Y, Gao G J, Zhang Q L, Liu X, He Y Q. QTL mapping and epistasis analysis of the protein content in brown rice. Molecular Plant Breeding, 2012,10 (5): 594-599
- [12] 杨亚春,倪大虎,宋丰顺,李莉,冯光,李泽福, 杨剑波. 两个环境下糙米和精米蛋白质含量的 QTL 分析.中国水稻科学, 2012,26 (3): 351-355
- Yang Y C, Ni D H, Song F S, Li L, Feng G, Li Z F, Yang J B. Identification of QTL for protein content in brown and milled rice in two environments. Chinese Journal of Rice Science, 2012, 26(3): 351-355
- [13] 张杰, 郑蕾娜, 蔡跃, 尤小满, 孔飞, 汪国湘, 燕海刚, 金洁, 王亮, 张文伟, 江玲. 稻米淀粉 RVA 谱特征值与直链淀粉、蛋白含量的相关性及其 QTL 定位分析. 中国水稻科学, 2017, 31 (1): 31-39
- Zhang J, Zheng L N, Cai Y, You X M, Kong F, Wang G X, Yan H G, Jin J, Wang L, Zhang W W, Jiang L. Correlation analysis and QTL mapping for starch RVA profile properties and amylose and protein contents in rice. Chinese Journal of Rice Science, 2017, 31 (1): 31-39
- [14] Park S G, Park H S, Baek M K, Jeong J M, Cho Y C, Lee G M, Lee C M, Suh J P, Kim C S, Kim S M. Improving the glossiness of cooked rice, an important component of visual rice grain quality. Rice, 2019, 12(1):87
- [15] Yang Y, Guo M, Li R, Shen L, Wang W, Liu M, Zhu Q, Hu Z, He Q, Xue Y, Tang S, Gu M, Yan C. Identification of quantitative trait loci responsible for rice grain protein content using chromosome segment substitution lines and fine mapping of *qPC-1* in rice (*Oryza sativa* L.). Molecular Breeding, 2015, 35: 130
- [16] 赵琳琳, 李楠, 吕志伟, 吴姗姗, 赵志超, 张文会. 野稗渗入系水稻籽粒储藏蛋白质含量的 QTL 遗传解析.江苏农业科学, 2015, 43(3):50-53

- Zhao L L, Li N, Lv Z W, Wu S S, Zhao Z C, Zhang W H. QTL genetic analysis of storage protein content in wild planted infiltrating rice grains. *Jiangsu Agricultural Science*, 2015, 43(3): 50-53
- [17] Takayuki K, Jun M. Identification and characteristics of quantitative trait locus for grain protein content *TGPI2*, in rice (*Oryza sativa* L.). *Euphytica*, 2018, 214(9):165
- [18] 王小雷, 刘杨, 孙晓棠, 欧阳林娟, 潘锦龙, 彭小松, 陈小荣, 贺晓鹏, 傅军如, 边建民, 胡丽芳, 徐杰, 贺浩华, 朱昌兰. 不同环境下稻米品质性状 QTL 的检测及稳定性分析. *中国水稻科学*, 2020, 34 (1): 17-27
- Wang X L, Liu Y, Sun X T, OuYang L J, Pan J L, Peng X S, Chen X R, He X P, Fu J R, Bian J M, Hu L F, Xu J, He H H, Zhu C L. Identification and stability analysis of QTL for grain quality traits under multiple environments in Rice. *Chinese Journal of Rice Science*, 2020, 34 (1): 17-27
- [19] Xu F F, Bao J S, He Q, Park Y J. Genome-wide association study of eating and cooking qualities in different subpopulations of rice (*Oryza sativa* L.). *Bmc Genomics*, 2016, 17(1): 663
- [20] Wang X Q, Pang Y L, Zhang J, Wu Z C, Chen K, Ali J, Ye G Y, Xu J L, Li Z K. Genome-wide and gene-based association mapping for rice eating and cooking characteristics and protein content. *Scientific Reports*, 2017, 7(1): 17203
- [21] Peng B, Kong H, Li Y, Wang L, Zhong M, Sun L, Gao G, Zhang Q, Luo L, Wang G, Xie W, Chen J, Yao W, Peng Y, Lei L, Lian X, Xiao J, Xu C, Li X, He Y. OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. *Nature Communications*, 2014, 5:4847
- [22] Yang Y, Guo M, Sun S, Zou Y, Yin S, Liu Y, Tang S, Gu M, Yang Z, Yan C. Natural variation of *OsGluA2* is involved in grain protein content regulation in rice. *Nature Communications*, 2019, 10:1949
- [23] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 2009, 14(1): 1754-1760
- [24] Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 2010, 20(9): 1297-1303
- [25] Mansfeld B N, Grumet R. QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *Plant Genome*, 2018, 11(2): 1-5
- [26] Sun L, Lin T, Jing D, Yu B, Zeng S, Li C, Qian H, Du C, Hu Q, Yang J, Zhou Y, Wu Z, Gong H. Preponderant alleles at *Hdl* and *Ehd1* lead to photoperiod insensitivity in japonica rice varieties. *Crop Breeding and Applied Biotechnology*, 2023, 23(3): e44842338
- [27] 周年兵. 沿淮下游地区温光要素对优质水稻产量、品质及氮素吸收利用的影响. 扬州: 扬州大学, 2021
- Zhou N B. Effects of temperature and light factors on yield, quality an nitrogen uptake and utilization of high-quality rice in the lower reaches of Huai River basin[D]. Yangzhou, Yangzhou University, 2021
- [28] 张小明, 王仪春, 石春海, 鲍根良, 叶胜海. 稻米蒸煮营养品质性状的遗传研究进展. *植物遗传资源学报*, 2002, 3 (2): 51-55
- Zhang X M, Wang Y C, Shi C H, Bao G L, Ye S H. Progress of study on genetics of cooking and nutrient quality. *Journal of Plant Genetic Resources*, 2002, 3 (2): 51-55
- [29] Wu B, Yun P, Zhou H, Xia D, Gu Y, Li P B, Yao J L, Zhou Z Q, Chen J X, Liu R J, Cheng S Y, Zhang H, Zheng Y Y, Lou G M, Chen P L, Wan S S, Zhou M S, Li Y H, Gao G J, Zhang Q L, Li X H, Lian X M, He Y Q. Natural variation in WHITE-CORE RATE1 regulates redox homeostasis in rice endosperm to affect grain quality. *The Plant Cell*, 2022, 34(5): 1912-1932
- [30] Cai Y, Zhang W W, Fu Y S, Shan Z Z, Xu J H, Wang P, Kong F, Jin J, Yan H G, Ge X Y, Wang Y X, You X M, Chen J, Li X, Chen W W, Chen X G, Ma J, Tang X J, Zhang J, Bao Y Q, Jiang L, Wang H Y, Wan J M. *Du13* encodes a C₂H₂ zinc-finger protein that regulates Wx^b pre-mRNA splicing and microRNA biogenesis in rice endosperm. *Plant Biotechnology Journal*, 2022, 20(7): 1387-1401
- [31] Ying Y N, Zhang Z W, Tappiban P, Xu F F, Deng G F, Dai G X, Bao J S. Starch fine structure and functional properties during seed development in BEIIb active and deficient rice. *Carbohydrate Polymers*, 2022, 292: 119640
- [32] Chao S F, Cai Y C, Feng B B, Jiao G A, Sheng Z H, Luo J, Tang S Q, Wang J L, Hu P S, Wei X J. Editing of rice isoamylase gene *ISA1* provides insights into its function in starch formation. *Rice Science*, 2019, 26(2): 77-87

[33] Wang J, Wan R J, Nie H P, Xue S W, Fang Z M. *OsNPF5.16*, a nitrate transporter gene with natural variation, is essential for rice growth and yield. *The Crop Journal*, 2022, 10(2):397-406

