



水稻重要农艺性状自然变异研究进展及其应用策略

张林, 何祖华*

中国科学院上海生命科学研究院植物生理生态研究所, 植物分子遗传国家重点实验室, 上海 200032

* 联系人, E-mail: zhhe@sibs.ac.cn

2015-01-28 收稿, 2015-02-10 接受, 2015-03-06 网络版发表

国家自然科学基金(31330061)资助

摘要 水稻在长期的自然进化和人工选择过程中积累了大量自然变异, 这些自然变异是许多重要农艺性状的遗传基础, 也是水稻育种筛选的主要目标. 近些年, 随着分子技术的发展, 水稻自然变异研究取得了一系列重要突破. 本文结合最新研究进展, 系统介绍了水稻自然变异的特点及相应分析方法, 并强调了高通量测序技术在今后水稻自然变异挖掘中的意义; 对最近10年来报道的抽穗期、产量元件及抗逆和驯化相关性状自然变异的分子机制及驯化特征进行了详细说明; 结合这些自然变异的分子特征, 总结了利用不同变异类型改良现有品种的方法, 为今后分子育种提供理论依据.

关键词

水稻
自然变异
数量性状
分子机制
育种

水稻(*Oryza sativa* L.)是亚洲及部分非洲国家的主要粮食作物, 而人口问题也是这些国家的主要问题. 尽管绿色革命基因*sd1*大大增加了水稻产量, 但随着人口增加, 这种效益将被逐渐吞噬, 未来的粮食需求对水稻育种提出了更高要求^[1]. 水稻种质资源丰富, 拥有很多地方品种及野生种质, 由于特定生长环境和选择历史, 这些种质中包含许多优良自然变异, 包括产量、抗性基因等, 挖掘这些有利自然变异将为培育更好的品种打下基础. 许多重要农艺性状都属于数量性状, 会受到多个自然变异位点调控, 称为数量性状基因座(quantitative trait locus, QTL), 不同位点对表型效应的贡献不同, 且有的位点可能存在与遗传背景和环境的互作效应, 因此对这些位点进行遗传分析和基因克隆相对困难. 随着水稻参考基因组的完善及分子标记技术的发展, 结合相应统计模型的提出, 现在研究者已经能够对不同种质自然变异的遗传基础进行分析, 并挖掘其中的有利位点. 而且随着高通量测序技术的成熟及成本降低, 高密度基因型鉴定和关联分析也成为水稻自然变异发掘的新手段. 尽管大量的QTL分析结果使人们对复杂农

艺性状的遗传规律有了较全面认识, 但阐明这些QTL的分子机制才能更有效地利用这些位点. 目前, 水稻多个重要农艺性状QTL已经被成功克隆, 包括抽穗期、种子大小、穗粒数等, 相应基因的分子机制也逐渐解析清楚, 从而使对这些性状进行分子改造成为可能. 另外, 作为单子叶植物研究的模式植物, 水稻自然变异的发掘能够与突变体研究相辅相成, 加强人们对一些基础生物过程(如开花时间的调控)的理解, 因此, 自然变异的分子机制研究在理论和育种实践中都有重要价值.

1 水稻自然变异的特点及研究策略

1.1 自然变异的复杂性及分析方法

大部分农艺性状由多个QTL位点控制, 根据对表型的贡献不同, 可分为主效位点和微效位点. 一般认为贡献率达到15%即为主效位点. 但主效和微效是相对的, 取决于一个性状由多少位点调控, 例如, 通过回交将主效位点排除后, 则效应较小的位点的贡献率会增加. 另外, 基因型×基因型(上位性)互作

引用格式: 张林, 何祖华. 水稻重要农艺性状自然变异研究进展及其应用策略. 科学通报, 2015, 60: 1066–1078

Zhang L, He Z H. Understanding natural variations: The source of elite agronomic traits for rice breeding (in Chinese). Chin Sci Bull, 2015, 60: 1066–1078, doi: 10.1360/N972015-00052

和基因型×环境互作也是表型变化的重要原因,但大部分数量性状统计模型只能进行双位点互作分析,而基因型×环境互作分析则需要将群体种植于不同生长环境或年份.自然变异位置和遗传效应的探索依赖于初级定位群体,常用的群体有重组自交系(recombinant inbred lines, RIL)、回交自交系(backcross inbred lines, BIL)、双单倍体系(double haploid, DH)和 $F_{2:3}$ 家系4种. $F_{2:3}$ 群体由于基因型中存在杂合子, F_3 代会出现表型分离,因此为非永久性群体,但可以进行显性效应分析,该类型群体还可以演变成 VF_2 (vegetatively replicated F_2)群体,即将 F_2 单株通过分蘖移栽形成株系,用于表型判断;其他3种都是通过自交固定或单倍体加倍形成的永久性群体,基因型不分离,可进行加性效应分析、上位性分析和基因型×环境的互作分析,其中,重组自交系群体株系间两两随机杂交可形成永久性 F_2 (immortalized F_2 , IF_2)群体,从而进行显性效应分析.Xing等人^[2]通过对籼稻杂交组合ZS97和MH63衍生的重组自交系进行不同年份的产量性状分析,共计检测到29个加性QTL和35个上位性QTL,其中13个加性QTL存在环境互作效应,而上位性QTL没有检测到环境互作效应.贡献率分析显示,加性位点贡献率大于上位性位点,而环境效应对表型变化的影响最小.Wan等人^[3]利用重组自交系对粒宽性状在4个环境下进行分析发现,基因型变异与表型变化相关显著,而基因型与环境互作效应微弱,只占3.5%的表型变异,尽管如此,他们在上位性QTL中也检测到了环境互作效应.另外,环境对粒宽性状的影响主要体现在微效位点上,而对主效位点影响不大.这些研究暗示,水稻品种具有复杂的自然变异构成,而主效加性位点表现最稳定,在育种中的应用价值更大.群体特征对自然变异的检测存在影响,玉米(*Zea mays* L.)中的研究发现,随着群体中个体数目的增加,检测到的QTL数目也增多,且QTL估计的正确率和精度也增加,如果群体过小则不能检测到所有QTL,且检测到的QTL效应会被放大,与实际贡献不符^[4].因此,利用一个群体进行QTL分析往往不能发掘所有自然变异,以水稻抽穗性状的研究情况为例,用 F_2 进行QTL分析,只检测到5个QTL,用 BC_1F_5 群体又检测到3个新的QTL,而更多的QTL则只能在高世代的 BC_3F_2 或 BC_4F_2 群体中检测出^[5].不同初级群体的QTL检测效率也存在差异,张启发研究组^[6]对4种初级群体($F_{2:3}$, VF_2 , IF_2 , RIL) 2

年的QTL分析结果进行总结发现,尽管来自相同的亲本组合,不同群体检测出的QTL数目差异较大,且很多位点只能在1年中检测到,表明通过初级群体得到的结果需经过进一步确认.

1.2 近等基因系在自然变异分析中的作用

构建QTL近等基因系(near isogenic line, NIL)可以有效评价对应位点的真实效应,Ding等人^[7]根据重组自交系定位结果分别构建了17个抗旱相关QTL的近等基因系,通过对这些株系在正常和干旱条件下的表型分析来判断QTL效应,在排除抽穗期对表型的干扰后,发现只有4个株系的表现与初级定位结果一致.NIL能够排除其他位点对表型的影响,Zhang等人^[8]将1个主效QTL和3个微效QTL分别构建近等基因系进行分析.在重组自交系群体分析中,有的微效QTL只能在1年中被检测到,LOD(log of odds)值和表型贡献率也较小;而在近等基因系分析中,所有QTL的效应都得到了确认,且LOD值和表型贡献率都显著提高.因此他们认为,高质量NIL能够有效鉴定微效QTL,并将其分离成单个孟德尔因子.另外,对于一些特殊表型的QTL分析,NIL的作用也得到充分利用,如水稻组织培养能力、抗紫外线辐射能力^[9,10]等.NIL还可以用来明确QTL间的上位效应,Lin等人^[11]构建了*Hd1*, *Hd2*和*Hd3*的近等基因系,并将它们进行两两组合,发现*Hd1*和*Hd2*相对于*Hd3*的功能位点存在上位效应,*Hd3*的功能发挥需要*Hd1*和*Hd2*功能位点存在.近等基因系除了用于确定QTL效应外,还可直接用于构建高级定位群体进行QTL基因克隆.

1.3 替换系是自然变异分析的有效方法

除了用初级群体进行QTL分析外,替换系是一种不需要借助复杂统计学分析模型就可直接找到QTL的方法,该方法通过不断回交将供体亲本的染色体片段导入到受体亲本中,结合分子标记筛选得到的一套覆盖供体亲本不同染色体片段的株系群,通过与受体亲本比较即可明确哪些染色体片段参与相应表型的调控.该方法具有高通量的特点,且遗传效应稳定,找到相关位点后回交一代即可形成精细定位群体并快速产生近等基因系,构建一套替换系可以对多种复杂性状进行分析,特别是对一些测量方法复杂的表型十分有效.Garcia-Oliveira等人^[12]用85个野生稻染色体片段置换系寻找种子矿质元素含量QTL,共检测到31个

QTL, 26个有利位点来自于野生稻, 暗示野生稻对相关元素有更强的代谢或积累能力, 其中, 在8号染色体发现一个控制锌元素含量的主效QTL. Takai等人^[13]用39个替换系寻找叶光合指标碳同位素分辨率QTL, 分别在2, 3, 6, 7, 12号染色体检测到相关QTL, 并发现3号染色体QTL能够同时增加二氧化碳的气孔导度和碳同位素分辨率, 但不影响光合速率.

2 高通量技术在自然变异发掘中的应用

2.1 高通量测序在连锁分析中的应用

2005年, 水稻品种日本晴高质量基因组序列拼接完成^[14], 而籼稻品种9311基因组也得到了进一步完善^[15], 为此后的水稻研究提供了参考基因组. 随着高通量技术的发展, 个体基因组测序成本大为降低, 使群体高通量基因型鉴定成为可能. Huang等人^[16]利用二代测序技术对日本晴和9311产生的150个重组自交系进行0.02×覆盖度的测序, 并结合测序错误率发展了一种基于窗口移动的基因型判断方法, 最终得到群体的遗传连锁图, 该方法可以将交换位点限定在40 kb范围内, 比PCR方法的分子标记精确度高35倍, 而基因型鉴定速度则提高20倍; 对株高性状进行QTL分析可以将绿色革命基因*sd1*定位到100 kb范围内, 大大超出常规QTL分析能够达到的精度. 随后, 利用该群体和连锁图对14个农艺性状进行检测共得到49个QTL位点, 其中5个主效QTL被定位到很小的范围, 并找到了它们的候选基因^[17]. 与此同时, Xie等人^[18]也采用了低覆盖度的测序方法对重组自交系进行基因型鉴定, 由于没有相应的亲本参考序列, 他们利用最大简约法通过比较子代单核苷酸多态性(single nucleotide polymorphism, SNP)来推测亲本基因型, 然后利用隐马可夫模型对群体基因型判断, 构建连锁图. 利用该群体可将粒宽基因*GW5*定位到200 kb的范围内, 与采用RFLPs(restricted fragment length polymorphisms)/SSRs(simple sequence repeats)标记构建的连锁图相比, 该方法可以检测到更多的QTL, 特别是粒重QTL, 且定位精度更高^[19]. 高通量测序技术除了用于初级群体定位, 还可用于替换系的发展. Xu等人^[20]用常规方法构建了9311背景下的日本晴染色体片段置换系, 在128个株系中检测到了142个替换片段, 然后它们对这些株系进行了0.13×的高通量测序和基因型鉴定, 又检测到

117个新的片段置换, 其中大部分为小片段置换. Zhang等人^[21]则构建了日本晴背景下的9311片段置换系, 它们对初步筛选出的57个置换系进行0.13×基因组重测序并进行基因型判断, 最终挑出38个株系构建基因型图谱, 这些替换系覆盖了87.4%的水稻基因组, 平均每个株系包含2.5个替换片段.

2.2 高通量测序在关联分析中的应用

高通量技术除了在QTL连锁分析中得到应用外, 在关联分析中也展现出巨大优势. 与连锁分析不同, 关联分析是利用自然群体的历史重组事件和连锁不平衡进行基因组位点和表型的关联, 对群体中驯化相关位点能够很好地检出, 另外, 由于该方法利用不同地区的自然种质, 因此能够检测多个种质中的有利变异, 而基于连锁分析的QTL定位则一次只能寻找2个种质间的优良变异. Huang等人^[22]从50000份中国水稻品种中选出517个地方品种进行1×覆盖度的基因组重测序, 共检测到360万个SNP, 由于1×覆盖度不能覆盖所有SNP型, 因此他们采用KNN算法来推测空缺区的基因型, 最终得到500多个品种的饱和SNP图, 利用该SNP图, 他们对其中373个籼稻品种的14个农艺性状进行关联分析, 在混合线性模型下共检测到37个显著性关联, 其中7个位点与已知调控位点相同, 精确度可达到26 kb. 后来, 他们又对额外330个种质进行测序, 包括100个中国粳稻品种和230个世界各国品种. 结合原来的500多个种质信息, 对抽穗期及10个谷粒相关性状分别进行籼稻亚种、粳稻亚种及混合群体的关联分析, 检测到32个新的关联位点. 此外, 他们利用基因附近的SNP将种质分类, 然后在同一类中利用该类所有品种测序结果进行基因序列拼接, 从而得到不同分类中基因的完整序列, 来预测候选基因的变异位点. 他们还结合表达数据对关联位点的候选基因进行预测, 共得到18个候选基因^[23]. 基于高通量技术的关联分析研究能够快速发掘水稻自然种质中的有利位点, 然而这些位点的效应还需要进一步通过连锁分析或遗传互补进行确认, 并最终应用于水稻育种研究.

3 水稻自然变异分子克隆及其功能

3.1 水稻抽穗期性状调控机制

大量的QTL分析和关联分析结果虽然提供了自

然变异位点的位置信息和遗传效应,但只有了解相应的分子机制才能够有效地操作自然变异,并应用到生产中去。水稻可以在不同纬度地区种植,但其生长周期需适应当地光温条件,抽穗期是决定这种适应性的关键性状。*Hd1*是水稻中第一个克隆的QTL,同时也是参与水稻抽穗期调控的QTL。定位结果显示,*Hd1*是拟南芥*CO*基因的同族体,与拟南芥不同,*Hd1*在短日照下促进开花,而在长日照下抑制开花^[24]。*Hd3a*是另一个抽穗期QTL,该位点被精细定位到20 kb范围内,序列比较发现,其中有一个拟南芥成花素*FT*的同源基因。在短日照条件下,品种Kasalath的*Hd3a*位点促进开花,且在*Hd1*位点近等基因系中表达上调,证明*Hd3a*位于*Hd1*的下游,形成与拟南芥*CO-FT*类似的保守通路;但在长日照条件下,*Hd1*抑制*Hd3a*表达^[25]。后续研究发现,*Hd3a*蛋白可以从叶片移动到顶端分生组织中,证明*Hd3a*为成花素基因^[26]。*Ehd1*位点是*Hd1*调控途径的平行通路,在*Hd1*功能缺失时,*Ehd1*能够在短日照下促进开花,该位点编码一个B类反应调控因子,基因表达受短日照诱导,近等基因系和转基因植株表达分析证明*Hd3a*和*RFT1*位于该位点下游^[27]。*RFT1*是*Hd3a*的同源基因,不同的是该位点在长日照条件开花途径中也发挥作用,形成OsMADS50-Ehd1-RFT1通路。在长日照条件下,*RFT1*基因突变使植株开花推迟约100 d,而其上游基因*OsMADS50*突变则导致植株无法开花^[28]。*Ghd7*编码一个长日照开花调控因子,包含CCT功能域,表现为长日照下延迟开花。在长日照条件下,*Ghd7*表达量提高并抑制*Ehd1*及其下游*Hd3a*表达从而延迟开花,而对*Hd1*表达无影响,短日照条件则对相关基因没有抑制效应,表型也没有变化^[29]。进一步研究显示,这种差异是由响应蓝光诱导的*Ehd1*和响应红光诱导的*Ghd7*在不同日照条件下反应不同造成的,短日照条件下,*Ghd7*不能被及时诱导,从而释放了Ehd1-Hd3a通路^[30]。*Ghd8/DTH8*表现为类似的长日照抑制功能。图位克隆显示,该位点编码一个潜在的HAP3亚基蛋白;转录分析表明,*Ghd8*下调*Ehd1*和*Hd3a*表达,且这种抑制效应不依赖于*Ghd7*和*Hd1*^[31,32]。最近研究显示,一个微效位点*DTH2*能够在长日照下促进开花,该位点编码类CO蛋白,通过上调*Hd3a*和*RFT1*促进开花,而不依赖于*Hd1*和*Ehd1*,因此是一个新的开花调节因子^[33]。而*OsPRR37/DTH7/Ghd7.1*则是最新发现的长日照开花抑制因

子^[34~36],编码一个带有CCT功能域的PRR蛋白,能够抑制*Ehd1*及其下游成花素基因。该基因在*phyb*突变体中表达下调,暗示其位于*PHYB*下游^[35]。上述研究显示,水稻抽穗期由多个途径参与,且形成复杂的调控网络。

3.2 水稻产量元件调控机制

作为育种的重要目标,产量性状的分子机制研究具有重要应用价值。产量的主要构成因素包括3部分:单株穗数、每穗粒数和粒重,现有的自然变异研究结果主要集中在穗粒数和粒重2个方面。抽穗期相关自然变异往往对穗粒数也有影响,3个长日照延迟开花因子(*Ghd7*, *Ghd8*和*Ghd7.1*)都表现出增产潜力,表现为穗分支和每穗粒数增加,同时增加株高^[29,31,32,35,36],推测这种促进效应是由于穗分化持续期延长导致的。例如,包含*Ghd7*的MH63穗发育持续期为30 d,而无*Ghd7*的ZS97穗发育持续期只有23 d^[6]。最近报道显示,*Hd1*的非功能型位点也能够促进穗粒数和株高增加,暗示延长生育期增加水稻产量潜力具有普遍性^[37]。除了抽穗期基因,许多直接参与穗发育调控的QTL已被克隆。*Gn1a*是水稻中第一个克隆的穗粒数QTL,其编码一个细胞分裂素氧化酶,能够降解细胞分裂素,该位点突变或表达下调可以增加二级分支数目,从而增加产量^[38],而水稻细胞分裂素激活酶基因*LOG*突变则会导致穗分支及穗粒数剧烈减少^[39],说明细胞分裂素在调控水稻穗分化过程起重要作用。*dep1*为一个显性抑制突变位点,编码一个类磷脂酰乙醇胺结合蛋白,该突变位点能够促进细胞分裂,增加穗原基大小,导致植株产生直立穗,且穗一级分支和二级分支数目增加,穗粒数增多,最终产量增加。表达分析显示,*dep1*中*Gn1a*基因表达量下调,暗示细胞分裂素参与了*dep1*对穗发育的调控途径^[40]。*IPA1/FWP*位点编码一个含有SBP(SQUAMOSA-promoter binding protein)功能域的转录因子*OsSPL14*,该基因受miR156调控,*IPA1*中miR156靶向位点中产生一个点突变导致该基因不受小RNA调控,从而上调表达,表现为穗粒数增加、植株变高、分蘖变少、茎秆变粗的理想株型表型,将该位点导入推广品种中可以显著提高产量^[41]。定位结果显示,*FWP*位于*OsSPL14*启动子区,暗示*FWP*属于另一种突变类型,推测该位点通过表观遗传方式上调*OsSPL14*表达^[42]。*AP01*是拟南芥中*UFO*的同源基因,突变体研

究表明,其通过抑制花序分生组织到小穗分生组织的转化来调控分支数目,另外它还通过影响C类基因的表达来调控花器官发育^[43]. *APOL*自然优势等位表现为多种效应,能够同时增加水稻产量潜力和抗倒伏能力,因此Ookawa等人^[44]提出可以利用QTL多效性进行品种快速改良的新策略. Terao等人^[45]通过精细定位发现, *APOL*位点的不同区域调控不同表型,其中编码区和靠近编码区的一段启动子只影响一级分支数,而不影响穗粒数,远端的启动子区则影响产量和收获指数. 转基因和显性突变体研究也证明,在*APOL*编码区上游约3500 bp处存在1个负调控因子,该位点缺失或突变即可表现为穗分支增多的过表达表型^[43,46]. *NAL1*是另一个具有多效性的产量基因,该基因首先在窄叶突变体中被发现,编码一个植物特有蛋白,主要在维管组织中富集表达,突变体表现为叶片变窄、维管束发育畸形,另外,生长素的极性运输也受到抑制^[47]. 后续研究发现,该基因的自然变异还能够影响叶片叶绿素含量(*qLSCHL4*)、叶片光合速率(*GPS*)、穗型及穗粒数(*SPIKE*),表现出良好的增产潜力^[48-50]. 近等基因系分析显示,包含*SPIKE*位点的植株产量能够增加13%~36%,且对种子形态没有负效应. 其功能位点来自热带粳稻品种,导入籼稻品种IRRI146可使产量增加约18%^[50]. 结实率也是决定每穗粒数的关键因素, *PTBI*是目前唯一被克隆的结实率基因,编码一个RING类型的E3泛素化连接酶. 该位点能够刺激花粉管穿过柱头传送管道,提高子房受精效率,从而维持正常结实率. 在转基因株系和自然品种中,该基因表达量与结实率表现出明显相关性,暗示该位点是决定结实率的主要自然变异^[51].

粒重是产量提高的另一个关键元件,包括粒长、粒宽和粒厚3个元素,其中粒长和粒宽主要取决于小穗发育过程,而粒厚则取决于品种灌浆能力. *GS3*是一个主要影响种子粒长的QTL,精细定位发现其编码一个跨膜蛋白,其中第二外显子中一个C到A的突变导致该基因翻译提前终止而产生长粒表型^[52]. 对不同种质分析发现,该突变在长粒品种和短粒品种中具有代表性,可以作为功能型分子标记^[53]. 该基因包括4个功能域,分别是OSR区、跨膜区、TNFR/NGFR家族半胱氨酸富集区和VWFC区, C到A突变发生在OSR区,利用转基因实验对3个区域进行功能验证发现, OSR区为粒长的负调控因子,而TNFR/NGFR和VWFC区则抑制OSR的负调作用,如果这2

个区域缺失会产生更小的种子^[54]. Wang等人^[55]利用该特点又发展了2个功能型分子标记,可以用来进一步区分非长粒品种中的中粒位点和短粒位点,从而有利于培育各种粒型品种. *GL3.1/qGL3/qGL3-1*是最近发现的粒长QTL,位于三号染色体GS3位点下游,该位点编码PPKL家族的丝氨酸/苏氨酸磷酸酶^[56-58]. 酶活分析显示,长粒品种GL3.1的去磷酸化酶活性降低;互作分析显示, Cyclin-T1;3是GL3.1的直接底物,而RNAi下调Cyclin-T1;3表达使粒长变短,暗示GL3.1通过改变Cyclin-T1;3蛋白的磷酸化特征调节粒长^[57]. *GW5*参与水稻粒宽的调控,精细定位显示一段1.2 kb的缺失导致水稻颖壳外颖细胞数目增多,从而增加粒宽^[59]. 该基因的具体功能未知,酵母双杂交实验发现,该蛋白可以与多聚泛素化蛋白互作,暗示其可能通过泛素化蛋白降解途径来调控细胞分裂^[60]. *GW2*是另一个参与粒宽调控的QTL,该基因编码一个RING类型的E3泛素化连接酶蛋白,同*GW5*类似,也通过泛素化蛋白降解途径来调节细胞分裂, *GW2*功能缺失后,颖壳变大,同时灌浆速度加快,最终表现为粒重和产量增加^[61]. *GS5*和*GW8*都是粒重的正调控因子^[62,63]. *GS5*编码一个丝氨酸羧肽酶,转基因实验证明启动子区的序列差异是导致表型变化的主要原因,粒宽随基因表达量升高而增大. 超表达和突变体植株中调控细胞分裂相关基因表达分析显示, *GS5*位于这些调控基因的上游,通过调控细胞分裂来增加细胞数目^[62]. *GW8*编码与IPA1同家族的OsSPL16,与*GS5*的功能等位类似, *GW8*启动子区的变异导致该基因表达上调,并促进细胞分裂相关基因表达,增加粒宽^[63]. *TGW6*是最近克隆的粒重QTL,与上述研究不同,该位点对颖壳大小无影响,而直接影响胚乳长度,从而增加粒重,该位点编码一个IAA-葡萄糖水解酶, *kasalath*品种该位点突变,导致IAA在种子中的积累降低,并延迟多核期到细胞期的转化过程,最终增加胚乳细胞数目和长度,说明野生型*TGW6*是一个粒重负调控因子. 更有意义的是, *TGW6*能够增加源组织供应碳水化合物能力,从而满足库容增加^[64]. 另一个粒重位点*GW6a*被定位在*TGW6*的临近区域,也表现为*kasalath*位点增加粒重,不同的是该位点是一个功能增强型突变,该基因在*kasalath*中表达量上调,其编码的OsgIHAT1是一个组蛋白H4乙酰转移酶,能够增加组蛋白H4整体的乙酰化水平,并导致许多细胞分裂相关基因及*GS5*, *GW8*和*GW5*等粒形相

关基因表达上调,最终促进颖壳细胞分裂增加并增强籽粒灌浆^[65]。通过上述研究可知,粒重性状的改变主要依赖相关因子对细胞分裂的正向或负向调控来实现,且染色质修饰因子对调节相关基因表达起重要作用。

3.3 水稻抗逆相关基因分子机制

抗逆性状是水稻自然变异发掘的另一个热点,*SKC1*是水稻中第一个克隆的抗逆QTL,该基因编码一个HKT类转运蛋白,膜片钳实验证明SKC1介导对钠离子的运输,在盐胁迫条件下调节钠离子和钾离子的动态平衡,从而提高抗盐性^[66]。水淹是南亚和东南亚水稻生产的重要限制因素,*Sub1*是一个抗水淹的主效QTL,精细定位显示,3个包含乙烯反应因子(ethylene response factor, ERF)功能域的蛋白可能参与抗水淹的调控,分别命名为Sub1A, Sub1B, Sub1C,不抗水淹品种中*Sub1A*基因缺失。表达分析和转基因实验发现,*Sub1A*在水淹条件下被诱导表达,并抑制*Sub1C*表达,同时乙醇脱氢酶基因*Adh1*表达上调,证明Sub1A是抗水淹的正调控因子^[67]。进一步研究表明,*Sub1A*响应乙烯信号,在水淹条件下乙烯促进*Sub1A*表达,而*Sub1A*又可以增加赤霉素(gibberellic acid, GA)信号抑制子*SLR1*和*SLRL1*的表达,导致植株对GA不敏感,无法伸长,从而减少碳水化合物代谢,增加无氧呼吸能力^[68]。*SNORKEL1(SK1)*和*SNORKEL2(SK2)*是另一对参与深水反应的基因,与*Sub1A*基因不同,*SK1*和*SK2*通过促进茎秆伸长来避免水淹。该基因也包含ERF功能域蛋白,表达受水淹条件和乙烯处理诱导,在水淹条件下植株体内活性GA的含量升高,对应功能缺失品种中则无反应,证明*SK1*和*SK2*通过响应水淹条件下的乙烯信号来直接或间接影响GA途径,从而促进茎秆和叶片伸长,避免水淹^[69]。上述结果显示GA相关途径在耐水淹过程中有重要作用。低温萌发能力是温带地区直播水稻品种培育的重要选择目标,*qLTG3-1*是一个促进低温萌发的主效QTL,该位点编码未知蛋白,主要在种子萌发时的胚中表达,通过使包被胚的组织细胞空泡化,减少胚芽鞘生长时的机械阻碍,从而促进发芽^[70]。大部分水稻根系较浅,在干旱时不能吸收土壤深处的水分,*Dro1*是一个最近从旱稻品种Kinandang Patong中克隆的控制根生长角的QTL,该位点增加根系与地面夹角,干旱时更多根系能够进入深层土壤,从而

达到避旱的目的。该位点编码未知蛋白,对根系细胞伸长有促进作用,且表达受生长素负调控,在水平方向的根系中,由于重力作用生长素在根尖下侧积累,抑制对应区域*Dro1*表达,导致下部细胞生长慢于上部细胞,最终促进根系向地生长^[71]。*Ghd7*除了延迟开花和增加产量潜力外,还表现出抗逆效应,*Ghd7*表达受到干旱、脱落酸(abscisic acid, ABA)、茉莉酸(jasmonic acid, JA)及高温抑制,而被低温诱导,*Ghd7*超表达植株对干旱敏感而将*Ghd7*基因表达下调则增加植株抗旱能力,这些结果显示,*Ghd7*能够整合多重环境信号,从而使植株在最佳时期进入生殖生长^[72]。综合利用这些抗逆位点能够有效增强水稻对不同环境的适应能力,达到逆境下稳产的目的。

3.4 水稻自然变异驯化的分子特征

在水稻栽培历史上,随着人类的不断筛选,许多优良的自然变异被固定下来并广泛使用,因此目前许多克隆的QTL都是经过人类驯化的自然变异。*GS3*第二外显子中C到A的突变导致粒长增加,而长粒品种该位置的连锁不平衡范围增大,且核苷酸多态性比野生稻低,证明长粒品种受到正向选择^[73]。粒重基因*GW5*在籼稻和粳稻中表现出不同的分布模式,籼稻中的*GW5*位点都是野生型的,而粳稻中功能型和缺失型*GW5*位点表现出特定的地域分布特征,暗示籼稻和粳稻经历了不同的驯化过程,进一步结合另外2个驯化基因的功能多态性位点分布,能够推测得到不同地区水稻品种的驯化历史^[38]。*Gif1*是一个基于突变体表型定位的水稻灌浆基因,该基因编码一个胞壁转化酶,其启动子使该基因在维管束中特异表达,从而促进水稻灌浆并增加产量。核酸多态性分析显示,与野生稻相比,栽培稻在该基因启动子区核酸多态性很低,表现为强烈的选择痕迹,而基因编码区选择痕迹则相对较弱,由此可知,对调控区的选择也是驯化的重要途径^[74]。落粒性状的选择是作物驯化历史上的重要事件,*qSH1*是水稻中一个控制落粒性的主效QTL,该基因编码一个BEL1类的同源异形盒(homeobox)蛋白,为拟南芥RPL的同族体。精细定位显示,5'调控区的一个SNP差异导致离层不能正常形成,落粒性降低;关联分析显示,该SNP和一个AT重复序列多态性与温带粳稻中的落粒表型相关;对粳稻中不同品种进行单倍型(haplotype)分析发现,该SNP可能在驯化早期发生,但却晚于*GW5*基因^[75]。

*SH4*是另一个落粒性QTL, 该基因编码一个功能未知蛋白, 其DNA结合功能域的一个氨基酸替换导致离层不能正常发育, 引起落粒性降低. 与野生落粒种质相比, 该替换在栽培种中特有, 表明该位点是驯化过程中特异选择的多态性位点. 除了该替换外, 表达模式不同也暗示对调控区的选择可能也参与了落粒性的驯化过程^[76]. *Spr3*是一个控制散穗性状的QTL, 该性状是野生稻的特有特征, 而大部分栽培稻穗型紧凑. 该位点被定位到约4.6 kb的物理区间, 然而其中没有任何编码基因^[77]. 后续研究发现, 该位点调控下游*OsGLI*基因的表达, 与*IPA1*和*GW8*类似, *OsGLI*编码SBP功能域转录因子*OsSPL8*, 暗示SPL家族在多个发育过程中起作用; 序列分析显示, *Spr3*及其下游*OsGLI*核酸多态性降低, 证明该位点受到驯化选择, 该研究认为, 紧密穗能够降低落粒性和异交率, 满足人类种植需求^[78]. Zhu等人^[79]通过对158份紧密穗栽培稻和21份散穗野生稻进行测序并进行关联分析, 结果显示, *OsGLI*上游11198 bp处的一个SNP与表型紧密连锁, 是*Spr3*的功能性多态位点. 所有品种被分为19种单倍体型, 而大部分栽培稻集中于一种类型, 进一步证明该位点受到强烈选择. 水稻匍匐生长到直立生长是另一个关键的驯化过程, 2个实验室同时克隆了野生稻中控制匍匐生长习性的基因*PROG1*, 该基因编码一个C₂H₂锌指蛋白, 亚细胞定位和转录活性分析证明该蛋白是转录因子. 转基因实验则证明编码区一个氨基酸替换导致该基因功能缺失, 使植株从匍匐生长变为直立生长^[80]. 另一实验室对不同品种及野生稻进行序列比较发现, 所有驯化水稻品种都有相同的缺失位点, 证明直立生长习性是从同一起源驯化得来. 而野生稻中有一部分包含缺失型位点, 但却表现为匍匐生长表型, 暗示野生稻可能有其他途径调控匍匐生长^[81]. 最近研究表明, 绿色革命基因*sd1*中存在2个功能型多态性位点与粳稻矮秆表型相关, 与野生稻和籼稻相比, 粳稻*sd1*位点附近的核酸多态性降低, 表明在粳稻驯化过程中, *sd1*受到古代人类的人工选择^[82]. 种子休眠可以帮助水稻度过不利环境, 同时也可避免穗发芽, 该性状也表现出一定的驯化特征. *Sdr4*为水稻控制休眠QTL, 其编码基因功能未知, 蛋白定位在核中, 介导种子成熟过程中休眠特性的调控. 对不同品种及野生种序列分析发现, 所有粳稻都包含抑制休眠位点, 而籼稻中则表现为促进和抑制位点同时存在; 对周围染色体

进行单倍型分析发现, 籼稻中的抑制位点来自粳稻, 而野生稻中则只有促进休眠位点. 由此可知, 粳稻和籼稻中*Sdr4*位点经历了不同程度的驯化过程^[83].

3.5 表观遗传变异参与水稻重要农艺性状调控

通过上述研究可以发现, 人类对优良自然变异的选择不仅集中在蛋白功能缺失型突变, 还包括功能获得型突变及调控区的改变. 而最近几篇报道显示, 表观遗传、非编码RNA也在水稻重要农艺性状的调控中起重要作用. *Epi-d1*是一个亚稳定矮秆突变体, 这种亚稳定性与DNA序列无关, 而是由*D1*基因启动子区的表观遗传状态决定, 包括组蛋白和DNA甲基化修饰, 该突变体具有双向回复突变能力, 推测有利于植株快速适应环境变化^[84]. 光敏雄性不育材料的发现使水稻两系法杂交育种成为可能, 了解其分子机制有重要意义. 农垦58S从农垦58N突变而来, 在长日照下表现为光敏雄性不育, 精细定位显示, 一个SNP导致两者表型的差异; 进一步功能分析发现, 非编码RNA参与这一过程的调控, 长日照下58S和58N中该基因表达量升高, 且58S表达量低于58N; 甲基化分析显示, 58S中该基因启动子甲基化水平提高, 推测SNP导致非编码RNA二级结构改变, 影响其对自身位点的甲基化修饰, 从而导致表达水平变化, 最终改变58S的育性^[85]. Zhou等人^[86]的研究则显示, 该位点能够产生1个21 nt的小RNA, 而SNP位于小RNA上, 认为SNP降低了小RNA与靶基因位点的结合能力, 使相关基因表达失控, 从而产生表型.

4 水稻自然变异在育种中的应用前景

4.1 转基因技术实现品种的快速改良

水稻QTL研究的迅猛发展为育种家利用这些优良的自然变异提供了理论基础, 采用生物技术对现有品种进行分子改良, 从而加速育种进程. 根据现有自然变异序列信息, 育种家能够对一个品种的相关性状进行遗传构成分析, 通过测序了解不同位点的自然变异类型, 然后有目的地改良劣势位点. 转基因技术的进步能够进一步促进这一过程的实现, 最近发展的TALEN(transcription activator-like effector nuclease)和CRISPER(clustered regularly interspaced short palindromic repeat)/Cas9技术能够定点敲除目的基因, 该方法一方面能够对多个靶位点进行敲除, 另

一方面能够通过自交筛选将外源转基因序列清除, 其在水稻中应用的可行性和稳定性已被多个研究组证实^[87~90]。现有研究的自然变异主要分为基因功能缺失后产生有利性状和功能获得后产生有利性状两种类型。符合第一种情况的基因包括 *GS3*, *GL3.1*, *GW2*, *GW5*, *TGW6*, *Gn1a*, *dep1* 和 *sd1*, 利用基因敲除技术可以产生类似的效果, 而不产生负效应。针对 *GS3* 还可以有选择地敲除不同的功能域, 从而形成不同的粒长组合。更重要的是, 这些位点大部分是直接参与产量相关的元件, 能够快速提高品种库容, 增加产量潜力。最近, Shan 等人^[91] 利用TALEN技术成功敲除了影响水稻香味形成的基因 *OsBADH2*, 转基因纯和突变植株产生的芳香物质达到该基因自然变异品种的水平, 证明转基因敲除技术可用于水稻重要农艺性状改良。功能获得型位点特别是基因表达上调位点的引入则需要转入全部基因序列及其上部调控元件, 相关位点包括 *GS5*, *GW6a*, *GW8*, *AP01*, *Ghd7*, *Ghd8*, *Ghd7.1*, *SPIKE*, *PTB1*, *Dro1* 等, 然而这种类型的序列构建往往操作难度大, 不容易实现, 而且有些基因表达上调存在组织特异性和染色体位置依赖性, 外源转基因不能完全模拟相应表型, 这一情况在 *qSH1* 和 *Spr3* 位点的研究中都有体现^[75,78]。另外, 虽然一个性状在QTL分析中显示为一个位点控制, 精细定位却显示多个基因参与其中, 如 *SK1* 和 *SK2* 同时参与水稻耐水淹性状的调控, 增加了转基因的操作难度^[69]。

4.2 利用分子标记辅助选择聚合育种

对于那些转基因操作不容易实现或不能达到预期效果的自然变异, 可以采用分子标记辅助选择 (marker assisted selection, MAS) 的方法将特定优势位点导入推广品种中, 从而提高原有品种品质。例如, 在 *Sub1A* 的研究中, 该位点被导入到印度推广品种 *Swarna* 中, 从而产生抗水淹的栽培品种, 且该位点的存在不影响正常产量^[67]。后续研究也将该位点导入泰国香米品种 *KDML105* 中, 提高其耐水淹能力^[92]。类似地, 将多个控制不同优良性状的基因或QTL导入相同品种成为现代分子育种中的一个新策略, 称为QTL聚合育种, 不同位点组合的有效性已经被多项研究证实。QTL聚合的概念在 *Gn1a* 的研究中被首次提出, 为了解决包含 *Gn1a* 位点的植株易倒伏的问题, 研究者将 *Gn1a* 与绿色革命基因 *sd1* 聚合到同一个遗传背景中, 结果显示, 虽然 *sd1* 位点会削弱 *Gn1a* 的

增产比例, 但双位点植株依然表现为增产^[38]。随后, Steele 等人^[93] 成功将 *Azucena* 品种中4个控制根部性状的QTL及1个控制稻米香味的隐性位点聚合到印度早稻品种 *Kalinga III* 中, 试图以此提高该品种的抗旱能力。整个过程从亲本杂交形成 F_1 代到最终获得理想单株一共经历了8代, 且每一代都经过不同程度的分子标记筛选, 这些标记能够快速鉴定目标位点并排除无效位点。而 Ando 等人^[94] 则直接利用替换系对2个穗结构QTL进行聚合, 结果显示, 2个QTL对产量的贡献是独立的, 两者聚合能够产生更多的穗粒数, 因此可以利用更多相关性状的QTL聚合来提高水稻产量。而将 *gw8* 与 *gs3* 聚合能在保证产量的同时也改良米质, 采用该方法育成了新品种 *Huabiao1*^[63]。类似的尝试也在 Wang 等人^[95] 的研究中得到实现, 他们将长剑叶位点 *qFL1* 与 *GS3*, *Wx*, *Ghd8* 位点进行多重聚合, 产生携带不同位点组合的 *ZS97*, 其中 *qFL1*, *GS3* 和 *Wx* 的聚合植株剑叶和籽粒增长, 直链淀粉含量降低, 产量增加52.5%, 而 *qFL1*, *Ghd8* 和 *Wx* 的聚合植株则增产高达64.6%。这些研究显示, QTL聚合育种可以提高品种的多重性状, 将是未来育种的重要手段。另外, 该方法不局限于功能基因已知的自然变异, 具有更广泛的适用性。

5 展望

培育高产稳产的水稻品种是育种的主要目标, 然而传统育种主要依赖于育种工作者对性状的肉眼观察和个人经验, 有些性状表型难于观察且表现不稳定, 在选择过程中容易丢失。现代分子技术的发展已经使人们能够对调控相应性状的自然变异进行挖掘, 并进行后续基因克隆和功能分析, 从而能够从分子水平对相应性状进行鉴定和改造。水稻QTL在线注释数据库Q-TRAO使人们能够方便地查询调控不同农艺性状的QTL位置信息及其候选基因^[96]。而水稻3000份核心种质的全基因组测序项目也已完成, 平均测序深度高达14x, 相应序列可以免费从GigaDB数据库下载^[97], 对应的SNP信息也可以从IRIC网站^[98]上下载, 这使人们能够轻易了解不同品种的遗传构成, 也为进行QTL连锁分析和关联分析提供了数据支撑, 从而挖掘更多的自然变异。转基因技术的发展使生物技术育种成为可能, 但只有了解不同自然变异的分子机制才能对相关性状进行操作, 许多产量性状的自然变异都是功能缺失型变异, 利

用基因敲除技术可以方便地对这些性状进行改造,然而功能获得型变异则可对相应调控元件进行改造或敲除其抑制的下游基因,但这一过程的实现需要更为精细的分子机制探索.除了转基因改造品种,利用分子标记辅助选择将优良自然变异引入也是今后育种的重要方向,由此中国科学院战略性先导科技专项提出了分子模块育种的概念,即将不同自然变异位点作为不同的模块,通过不同单位分工协作,将这些位点分别导入到现有当家品种中,然后再根据需要将不同位点聚合.随着更多自然变异位点的发

掘,综合采用不同技术手段进行品种培育成为可能.由此张启发教授^[99]提出了绿色超级稻的概念,在保证产量和谷粒质量的前提下,绿色超级稻还应该具有抗多种虫害和病害、高营养吸收率和抗旱性等特点,从而减少农药、化肥和水的使用.该策略的执行需要结合多种资源和技术,包括种质资源和突变体筛选,结合基因组和转录组研究进行基因定位、克隆及功能分析,然后利用分子标记辅助选择或转基因的方法逐渐将所有优良性状导入栽培品种中,形成绿色超级稻.届时,世界粮食生产将掀起又一次绿色革命.

参考文献

- 1 Khush G S. Green revolution: Preparing for the 21st century. *Genome*, 1999, 42: 646–655
- 2 Xing Y Z, Tan Y F, Hua J P, et al. Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. *Theor Appl Genet*, 2002, 105: 248–257
- 3 Wan X Y, Weng J F, Zhai H Q, et al. Quantitative trait loci (QTL) analysis for rice grain width and fine mapping of an identified QTL allele *gw-5* in a recombination hotspot region on chromosome 5. *Genetics*, 2008, 179: 2239–2252
- 4 Schon C C, Utz H F, Groh S, et al. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics*, 2004, 167: 485–498
- 5 Yano M, Kojima S, Takahashi Y, et al. Genetic control of flowering time in rice, a short-day plant. *Plant Physiol*, 2001, 127: 1425–1429
- 6 Xing Y Z, Zhang Q F. Genetic and molecular bases of rice yield. *Annu Rev Plant Biol*, 2010, 61: 421–442
- 7 Ding X P, Li X K, Xiong L Z. Evaluation of near-isogenic lines for drought resistance QTL and fine mapping of a locus affecting flag leaf width, spikelet number, and root volume in rice. *Theor Appl Genet*, 2011, 123: 815–826
- 8 Zhang Y S, Luo L J, Liu T M, et al. Four rice QTL controlling number of spikelets per panicle expressed the characteristics of single Mendelian gene in near isogenic backgrounds. *Theor Appl Genet*, 2009, 118: 1035–1044
- 9 Taguchi-Shiobara F, Yamamoto T, Yano M, et al. Mapping QTLs that control the performance of rice tissue culture and evaluation of derived near-isogenic lines. *Theor Appl Genet*, 2006, 112: 968–976
- 10 Sato T, Ueda T, Fukuta Y, et al. Mapping of quantitative trait loci associated with ultraviolet-B resistance in rice (*Oryza sativa* L.). *Theor Appl Genet*, 2003, 107: 1003–1008
- 11 Lin H X, Yamamoto T, Sasaki T, et al. Characterization and detection of epistatic interactions of 3 QTLs, *Hd1*, *Hd2*, and *Hd3*, controlling heading date in rice using nearly isogenic lines. *Theor Appl Genet*, 2000, 101: 1021–1028
- 12 Garcia-Oliveira A L, Tan L B, Fu Y C, et al. Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J Integr Plant Biol*, 2009, 51: 84–92
- 13 Takai T, Ohsumi A, San-oh Y, et al. Detection of a quantitative trait locus controlling carbon isotope discrimination and its contribution to stomatal conductance in *japonica* rice. *Theor Appl Genet*, 2009, 118: 1401–1410
- 14 Matsumoto T, Wu J Z, Kanamori H, et al. The map-based sequence of the rice genome. *Nature*, 2005, 436: 793–800
- 15 Yu J, Wang J, Lin W, et al. The genomes of *Oryza sativa*: A history of duplications. *PLoS Biol*, 2005, 3: e38
- 16 Huang X H, Feng Q, Qian Q, et al. High-throughput genotyping by whole-genome resequencing. *Genome Res*, 2009, 19: 1068–1076
- 17 Wang L, Wang A H, Huang X H, et al. Mapping 49 quantitative trait loci at high resolution through sequencing-based genotyping of rice recombinant inbred lines. *Theor Appl Genet*, 2011, 122: 327–340
- 18 Xie W B, Feng Q, Yu H H, et al. Parent-independent genotyping for constructing an ultrahigh-density linkage map based on population sequencing. *Proc Natl Acad Sci USA*, 2010, 107: 10578–10583
- 19 Yu H H, Xie W B, Wang J, et al. Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS One*, 2011, 6: e17595
- 20 Xu J J, Zhao Q A, Du P N, et al. Developing high throughput genotyped chromosome segment substitution lines based on population whole-genome re-sequencing in rice (*Oryza sativa* L.). *BMC Genomics*, 2010, 11: 656

- 21 Zhang H, Zhao Q, Sun Z Z, et al. Development and high-throughput genotyping of substitution lines carrying the chromosome segments of indica 9311 in the background of *japonica* Nipponbare. *J Genet Genomics*, 2011, 38: 603–611
- 22 Huang X H, Wei X H, Sang T, et al. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat Genet*, 2010, 42: 961–976
- 23 Huang X H, Zhao Y, Wei X H, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat Genet*, 2012, 44: 32–39
- 24 Yano M, Katayose Y, Ashikari M, et al. *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell*, 2000, 12: 2473–2483
- 25 Kojima S, Takahashi Y, Kobayashi Y, et al. *Hd3a*, a rice ortholog of the *Arabidopsis* *FT* gene, promotes transition to flowering downstream of *Hd1* under short-day conditions. *Plant Cell Physiol*, 2002, 43: 1096–1105
- 26 Tamaki S, Matsuo S, Wong H L, et al. Hd3a protein is a mobile flowering signal in rice. *Science*, 2007, 316: 1033–1036
- 27 Doi K, Izawa T, Fuse T, et al. 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*, 2004, 18: 926–936
- 28 Komiya R, Yokoi S, Shimamoto K. A gene network for long-day flowering activates *RFT1* encoding a mobile flowering signal in rice. *Development*, 2009, 136: 3443–3450
- 29 Xue W Y, Xing Y Z, Weng X Y, et al. Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet*, 2008, 40: 761–767
- 30 Itoh H, Nonoue Y, Yano M, et al. A pair of floral regulators sets critical day length for *Hd3a* florigen expression in rice. *Nat Genet*, 2010, 42: 635–638
- 31 Yan W H, Wang P, Chen H X, et al. A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol Plant*, 2011, 4: 319–330
- 32 Wei X J, Xu J F, Guo H N, et al. *DTH8* suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol*, 2010, 153: 1747–1758
- 33 Wu W, Zheng X M, Lu G, et al. Association of functional nucleotide polymorphisms at *DTH2* with the northward expansion of rice cultivation in Asia. *Proc Natl Acad Sci USA*, 2013, 110: 2775–2780
- 34 Koo B H, Yoo S C, Park J W, et al. Natural variation in *OsPRR37* regulates heading date and contributes to rice cultivation at a wide range of latitudes. *Mol Plant*, 2013, 6: 1877–1888
- 35 Gao H, Jin M, Zheng X M, et al. *Days to heading 7*, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. *Proc Natl Acad Sci USA*, 2014, 111: 16337–16342
- 36 Yan W H, Liu H Y, Zhou X C, et al. Natural variation in *Ghd7.1* plays an important role in grain yield and adaptation in rice. *Cell Res*, 2013, 23: 969–971
- 37 Zhang Z H, Wang K, Guo L, et al. Pleiotropism of the photoperiod-insensitive allele of *Hd1* on heading date, plant height and yield traits in rice. *PLoS One*, 2012, 7: e52538
- 38 Ashikari M, Sakakibara H, Lin S Y, et al. Cytokinin oxidase regulates rice grain production. *Science*, 2005, 309: 741–745
- 39 Kurakawa T, Ueda N, Maekawa M, et al. Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature*, 2007, 445: 652–655
- 40 Huang X Z, Qian Q, Liu Z B, et al. Natural variation at the *DEP1* locus enhances grain yield in rice. *Nat Genet*, 2009, 41: 494–497
- 41 Jiao Y Q, Wang Y H, Xue D W, et al. Regulation of *OsSPL14* by *OsmiR156* defines ideal plant architecture in rice. *Nat Genet*, 2010, 42: 541–544
- 42 Miura K, Ikeda M, Matsubara A, et al. *OsSPL14* promotes panicle branching and higher grain productivity in rice. *Nat Genet*, 2010, 42: 545–549
- 43 Ikeda K, Ito M, Nagasawa O N, et al. Rice *ABERRANT PANICLE ORGANIZATION 1*, encoding an F-box protein, regulates meristem fate. *Plant J*, 2007, 51: 1030–1040
- 44 Ookawa T, Hobo T, Yano M, et al. New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nat Commun*, 2010, 1: 132
- 45 Terao T, Nagata K, Morino K, et al. A gene controlling the number of primary rachis branches also controls the vascular bundle formation and hence is responsible to increase the harvest index and grain yield in rice. *Theor Appl Genet*, 2010, 120: 875–893
- 46 Ikeda-Kawakatsu K, Yasuno N, Oikawa T, et al. Expression level of *ABERRANT PANICLE ORGANIZATION1* determines rice inflorescence form through control of cell proliferation in the meristem. *Plant Physiol*, 2009, 150: 736–747
- 47 Qi J, Qian Q, Bu Q, et al. Mutation of the rice *Narrow leaf1* gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiol*, 2008, 147: 1947–1959

- 48 Zhang G H, Li S Y, Wang L, et al. *LSCHL4* from *japonica* cultivar, which is allelic to *NAL1*, increases yield of indica super rice 93-11. *Mol Plant*, 2014, 7: 1350–1364
- 49 Takai T, Adachi S, Taguchi-Shiobara F, et al. A natural variant of *NAL1*, selected in high-yield rice breeding programs, pleiotropically increases photosynthesis rate. *Sci Rep*, 2013, 3: 2149
- 50 Fujita D, Trijatmiko K R, Tagle A G, et al. *NAL1* allele from a rice landrace greatly increases yield in modern indica cultivars. *Proc Natl Acad Sci USA*, 2013, 110: 20431–20436
- 51 Li S C, Li W B, Huang B, et al. Natural variation in *PTBI* regulates rice seed setting rate by controlling pollen tube growth. *Nat Commun*, 2013, 4: 2793
- 52 Fan C H, Xing Y Z, Mao H L, et al. *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor Appl Genet*, 2006, 112: 1164–1171
- 53 Fan C C, Yu S B, Wang C R, et al. A causal C-A mutation in the second exon of *GS3* highly associated with rice grain length and validated as a functional marker. *Theor Appl Genet*, 2009, 118: 465–472
- 54 Mao H L, Sun S Y, Yao J L, et al. Linking differential domain functions of the *GS3* protein to natural variation of grain size in rice. *Proc Natl Acad Sci USA*, 2010, 107: 19579–19584
- 55 Wang C R, Chen S, Yu S B. Functional markers developed from multiple loci in *GS3* for fine marker-assisted selection of grain length in rice. *Theor Appl Genet*, 2011, 122: 905–913
- 56 Hu Z J, He H H, Zhang S Y, et al. A Kelch motif-containing serine/threonine protein phosphatase determines the large grain QTL trait in rice. *J Integr Plant Biol*, 2012, 54: 979–990
- 57 Qi P, Lin Y S, Song X J, et al. The novel quantitative trait locus *GL3.1* controls rice grain size and yield by regulating *Cyclin-T1;3*. *Cell Res*, 2012, 22: 1666–1680
- 58 Zhang X J, Wang J F, Huang J, et al. Rare allele of *OsPPKL1* associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc Natl Acad Sci USA*, 2012, 109: 21534–21539
- 59 Shomura A, Izawa T, Ebana K, et al. Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet*, 2008, 40: 1023–1028
- 60 Weng J F, Gu S H, Wan X Y, et al. Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res*, 2008, 18: 1199–1209
- 61 Song X J, Huang W, Shi M, et al. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat Genet*, 2007, 39: 623–630
- 62 Li Y B, Fan C C, Xing Y Z, et al. Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nat Genet*, 2011, 43: 1266–1269
- 63 Wang S K, Wu K, Yuan Q B, et al. Control of grain size, shape and quality by *OsSPL16* in rice. *Nat Genet*, 2012, 44: 950–954
- 64 Ishimaru K, Hirotsu N, Madoka Y, et al. Loss of function of the IAA-glucose hydrolase gene *TGW6* enhances rice grain weight and increases yield. *Nat Genet*, 2013, 45: 707–711
- 65 Song X J, Kuroha T, Ayano M, et al. Rare allele of a previously unidentified histone H4 acetyltransferase enhances grain weight, yield, and plant biomass in rice. *Proc Natl Acad Sci USA*, 2015, 112: 76–81
- 66 Ren Z H, Gao J P, Li L G, et al. A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat Genet*, 2005, 37: 1141–1146
- 67 Xu K N, Xu X, Fukao T, et al. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*, 2006, 442: 705–708
- 68 Fukao T, Bailey-Serres J. Submergence tolerance conferred by *Sub1A* is mediated by SLR1 and SLRL1 restriction of gibberellin responses in rice. *Proc Natl Acad Sci USA*, 2008, 105: 16814–16819
- 69 Hattori Y, Nagai K, Furukawa S, et al. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature*, 2009, 460: 1026–1030
- 70 Fujino K, Sekiguchi H, Matsuda Y, et al. Molecular identification of a major quantitative trait locus, *qLTG3-1*, controlling low-temperature germinability in rice. *Proc Natl Acad Sci USA*, 2008, 105: 12623–12628
- 71 Uga Y, Sugimoto K, Ogawa S, et al. Control of root system architecture by *DEEPER ROOTING 1* increases rice yield under drought conditions. *Nat Genet*, 2013, 45: 1097–1102
- 72 Weng X Y, Wang L W, Wang J, et al. *Grain number, plant height, and heading date7* is a central regulator of growth, development, and stress response. *Plant Physiol*, 2014, 164: 735–747
- 73 Takano-Kai N, Jiang H, Kubo T, et al. Evolutionary history of *GS3*, a gene conferring grain length in rice. *Genetics*, 2009, 182: 1323–1334

- 74 Wang E T, Wang J J, Zhu X D, et al. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat Genet*, 2008, 40: 1370–1374
- 75 Konishi S, Izawa T, Lin S Y, et al. An SNP caused loss of seed shattering during rice domestication. *Science*, 2006, 312: 1392–1396
- 76 Li C B, Zhou A L, Sang T. Rice domestication by reducing shattering. *Science*, 2006, 311: 1936–1939
- 77 Luo J J, Hao W, Jin J, et al. Fine mapping of *Spr3*, a locus for spreading panicle from African cultivated rice (*Oryza glaberrima* steud.). *Mol Plant*, 2008, 1: 830–838
- 78 Ishii T, Numaguchi K, Miura K, et al. *OsLGI* regulates a closed panicle trait in domesticated rice. *Nat Genet*, 2013, 45: 462–465
- 79 Zhu Z F, Tan L B, Fu Y C, et al. Genetic control of inflorescence architecture during rice domestication. *Nat Commun*, 2013, 4: 2200
- 80 Jin J, Huang W, Gao J P, et al. Genetic control of rice plant architecture under domestication. *Nat Genet*, 2008, 40: 1365–1369
- 81 Tan L B, Li X R, Liu F X, et al. Control of a key transition from prostrate to erect growth in rice domestication. *Nat Genet*, 2008, 40: 1360–1364
- 82 Asano K, Yamasaki M, Takuno S, et al. Artificial selection for a green revolution gene during *japonica* rice domestication. *Proc Natl Acad Sci USA*, 2011, 108: 11034–11039
- 83 Sugimoto K, Takeuchi Y, Ebana K, et al. Molecular cloning of *Sdr4*, a regulator involved in seed dormancy and domestication of rice. *Proc Natl Acad Sci USA*, 2010, 107: 5792–5797
- 84 Miura K, Agetsuma M, Kitano H, et al. A metastable *DWARF1* epigenetic mutant affecting plant stature in rice. *Proc Natl Acad Sci USA*, 2009, 106: 11218–11223
- 85 Ding J H, Lu Q, Ouyang Y D, et al. A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc Natl Acad Sci USA*, 2012, 109: 2654–2659
- 86 Zhou H, Liu Q J, Li J, et al. Photoperiod- and thermo-sensitive genic male sterility in rice are caused by a point mutation in a novel noncoding RNA that produces a small RNA. *Cell Res*, 2012, 22: 649–660
- 87 Li T, Liu B, Spalding M H, et al. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat Biotechnol*, 2012, 30: 390–392
- 88 Endo M, Mikami M, Toki S. Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol*, 2015, 56: 41–47
- 89 Zhang H, Zhang J S, Wei P L, et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J*, 2014, 12: 797–807
- 90 Zhou H B, Liu B, Weeks D P, et al. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res*, 2014, 42: 10903–10914
- 91 Shan Q W, Zhang Y, Chen K L, et al. Creation of fragrant rice by targeted knockout of the *OsBADH2* gene using TALEN technology. *Plant Biotechnol J*, 2015, doi: 10.1111/pbi.12312
- 92 Siangliw M, Toojinda T, Tragoonrun S, et al. Thai jasmine rice carrying *QTLch9* (*SubQTL*) is submergence tolerant. *Ann Bot*, 2003, 91: 255–261
- 93 Steele K A, Price A H, Shashidhar H E, et al. Marker-assisted selection to introgress rice QTLs controlling root traits into an Indian upland rice variety. *Theor Appl Genet*, 2006, 112: 208–221
- 94 Ando T, Yamamoto T, Shimizu T, et al. Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theor Appl Genet*, 2008, 116: 881–890
- 95 Wang P, Zhou G L, Yu H H, et al. Fine mapping a major QTL for flag leaf size and yield-related traits in rice. *Theor Appl Genet*, 2011, 123: 1319–1330
- 96 Yonemaru J, Yamamoto T, Fukuoka S, et al. Q-TARO: QTL annotation rice online database. *Rice*, 2010, 3: 194–203
- 97 3000 rice genomes project. The 3000 rice genomes project. *Gigascience*, 2014, 3: 7
- 98 Alexandrov N, Tai S S, Wang W S, et al. SNP-seek database of SNPs derived from 3000 rice genomes. *Nucleic Acids Res*, 2015, 43: D1023–D1027
- 99 Zhang Q F. Strategies for developing Green Super Rice. *Proc Natl Acad Sci USA*, 2007, 104: 16402–16409

Understanding natural variations: The source of elite agronomic traits for rice breeding

ZHANG Lin & HE ZuHua

National Key Laboratory of Plant Genetics, Institute of Plant Physiology and Ecology, Shanghai Institute for Biological Science, Chinese Academy of Sciences, Shanghai 200032, China

As a major cereal crop worldwide, rice has undergone a long natural evolution and domestication selection, which has resulted in a large amount of sequence variations. Some variations play important roles in agronomic traits that satisfy human demands, which are usually controlled by quantitative loci. The nature of rice breeding is to select and combine beneficial variations. However, traditional breeding relies mainly on visual observation empirically, which is ineffective and laborious, and difficult to use natural variations. It is necessary to reveal the elite natural variations of important agronomic traits in modern crop breeding with molecular techniques. With the advent of molecular technology, researchers have made great progress in elucidating natural variations and the underlying molecular mechanisms in rice. In this review, we describe characteristics of natural variations and research strategies in rice, highlighting the roles of high throughput sequencing techniques. We sum the recognized molecular mechanisms controlling important agronomic rice traits, including heading date, yield elements, abiotic stress tolerance and domestication-related traits. Additionally, we discuss the strategy of using different variations in rice breeding, either by transgene technology or marker-assisted quantitative trait loci pyramiding, which would facilitate rice breeding programs with multiple superior alleles.

rice, natural variation, quantitative traits, molecular mechanism, breeding

doi: 10.1360/N972015-00052