水稻杂种优势的转录组基础

何光明*, 何航, 邓兴旺*

北京大学现代农业学院, 北京大学生命科学学院, 蛋白质与植物基因研究国家重点实验室(北京大学), 北大-清华生命科学联合中心, 北京 100871

* 联系人, E-mail: heguangming@pku.edu.cn; deng@pku.edu.cn

2016-09-19 收稿, 2016-09-19 修回, 2016-09-20 接受, 2016-10-19 网络版发表

国家重大科学研究计划(2012CB910900)、蛋白质与植物基因研究国家重点实验室和北大-清华生命科学联合中心项目资助

摘要 在杂交种中, 亲本等位基因的表达将会发生变化, 导致杂交种转录组活性不同于亲本. 通过比较杂交种与亲本之间的转录组活性差异, 鉴定出这些差异的调控因子并建立起其与表型差异的联系, 就可能从分子水平上解析杂种优势形成的机理. 在水稻杂交种全基因组基因差异表达分析的不同研究中, 由于所采用转录组分析平台和实验取材组织器官等的差异, 所鉴别出来的差异表达基因及其在杂交种中的表达变化模式各不相同. 然而, 某些研究也揭示了一些共性的结果, 主要体现在水稻杂交种中差异表达基因的生物学功能在光合作用、碳水化合物代谢和能量代谢途径中富集. 对于导致水稻杂交种转录组活性变化的原因, 可以从基于基因启动子区顺式调控元件和与其相结合的反式作用因子的遗传调控机制, 以及从基于DNA甲基化、组蛋白修饰和小RNA的表观遗传调控机制两方面来进行解释. 今后在水稻杂交种转录组活性变化的分析中, 需要针对特定的生物学性状来进行, 并提高基因差异表达分析的精确性和特异性. 此外, 由于杂种优势是多个数量性状的综合体现, 可以将全基因组的基因差异表达分析与杂种优势的QTL (quantitative trait loci, 数量性状位点)定位及GWAS (genome-wide association studies, 全基因组关联分析)等数量遗传学方法相结合, 以对水稻杂种优势分子机理进行深入解析.

关键词 杂交稻, 杂种优势, 转录组活性

两个遗传背景不同的亲本进行杂交, 其杂交种在多个性状上优于双亲, 这被称为杂种优势现象. 该现象自被发现以来, 人们就一直在试图揭示其遗传和分子机制. 经典数量遗传学对杂种优势机理的解释主要包括“显性假说”^[1-4]、“超显性假说”^[5,6]和“上位性假说”^[7], 它们分别强调亲本中有利等位基因在杂交种中的显性互补、双亲异质的等位基因在杂交种中的互作以及亲本中不同位点上非等位基因在杂交种中的互作对杂种优势的贡献. 然而, 这些假说都是在现代分子生物学和分子遗传学理论形成之前提出的, 并不包含任何分子概念和基础, 因而它们都不能很好地用于解释杂种优势现象形成的分子机理^[8]. 但

是, 这些假说都包含了杂种优势最本质的遗传基础, 即它们都要求杂交亲本之间存在遗传上的异质性.

对于杂交亲本间遗传差异的评价先后经历了从形态性状遗传距离, 到生化标记(例如同工酶)差异, 再到DNA水平上的各种分子标记多态性, 如限制性片段长度多态性(restriction fragment length polymorphism, RFLP)、简单序列重复(simple sequence repeat, SSR)、单核苷酸多态性(single nucleotide polymorphism, SNP)等, 都是基于对亲本基因型之间真实遗传差异的间接反映, 并不能从本质上构成亲本之间以及杂交种与亲本间表型差异的遗传基础. 现代分子遗传学的发展已使人们在分子水平上揭示了遗传

引用格式: 何光明, 何航, 邓兴旺. 水稻杂种优势的转录组基础. 科学通报, 2016, 61: 3850-3857

He G M, He H, Deng X W. The transcriptomic basis of heterosis in rice (in Chinese). Chin Sci Bull, 2016, 61: 3850-3857, doi: 10.1360/N972016-01046

物质的本质(即基因组DNA序列),并解析了基因组在结构和功能上的差异及其产生的原因。因而,对于杂交亲本之间遗传差异的分子评价及其与杂种优势关系的研究,必然要涉及基因组DNA碱基序列及其编码的功能基因与其他活性功能位点,以及由DNA序列差异或由不基于DNA序列的表观遗传差异导致的对基因表达水平的调控和对基因产物功能的影响等。两个基因型不同的亲本杂交,其杂交种在许多性状上不同于双亲,这必然涉及到亲本基因组在杂合遗传背景下活性的变化以及亲本等位基因在杂交种中的表达变化。通过比较杂交种与亲本之间的转录组活性和基因表达的差异,鉴定出这些差异的调控因子并建立起其与表型差异的联系,就可能从分子水平上解析杂种优势形成的遗传机理。

水稻是杂种优势现象在农作物生产实践中应用最成功的作物之一。1966年,袁隆平在《科学通报》发表“水稻的雄性不孕性”论文,提出了将雄性不育用于制备杂交水稻的设想^[9]。此后50年来,水稻杂交育种在中国和世界范围内取得了巨大的成功,为我国和世界粮食安全提供了充分的保障。在此期间,研究者也从数量遗传学、分子遗传学、基因组学和表观遗传学等方面对水稻杂种优势现象形成的遗传机理进行了不断探索。由于杂种优势的衡量是与亲本相比较,杂交种各性状是在量上的优势,即表现为数量

性状的相对优势。而数量性状一般由多个基因控制。因此,对杂种优势分子机理的研究必然要涉及到许多和各性状相关的功能基因或基因组活性位点以及它们在杂交种中的变化模式和调控机制及其与杂交种表型优势之间的关系。本文介绍近年来在水稻杂交种转录组活性变化及其调控方面研究进展,为深入理解水稻杂种优势形成的分子机理提供思考。

1 水稻杂交种转录组活性变化

杂交种中转录组活性变化的一个最主要和最直接的体现是全基因组范围基因表达的变化。近10多年来,研究者运用各种低通量和高通量的基因组学技术,从全基因组水平上检测了水稻杂交种与其亲本间的基因表达差异,揭示了杂交种中基因表达变化的一些模式和规律,并根据这些差异表达基因所涉及的生物学功能推测了它们在水稻杂种优势形成中的可能机制(表1)。

基因表达系列分析(serial analysis of gene expression, SAGE)是一种基于短序列标签、可以同时数千个转录本进行研究的低通量基因表达分析技术^[10]。Bao等人^[11]于2005年利用SAGE技术,比较了超级稻“两优培九”与其亲本“培矮64s”(母本)和“93-11”(父本)之间特定发育时期穗、叶片和根中的转录组差异,发现595个SAGE标签(基因)的表达在杂

表1 水稻杂交种转录组分析

Table 1 Transcriptome analysis in rice hybrids

转录测定方法	杂交组合	测定部位	杂交种中差异表达基因功能富集	参考文献
SAGE	两优培九	穗、叶片、根	碳和氮素同化, 信号转导, 蛋白质加工	[10~12]
SAGE	两优2186	旗叶	光合作用, 碳固定	[13]
cDNA微阵列	汕优63	幼穗	DNA复制与修复, 碳水化合物、能量和脂代谢, 翻译, 蛋白质降解	[14,15]
cDNA微阵列	两优培九	地上部、根	基础代谢, 遗传信息加工, 细胞生长和死亡, 信号转导, 运输, 胁迫反应	[16]
EST测序	两优培九	成熟胚	胚发育相关	[17]
全基因组微阵列	两优培九, 日本晴×91-11	苗期地上部	碳水化合物、维生素、氨基酸代谢, 次级代谢产物的合成	[18]
全基因组微阵列	两优培九	叶片、穗	能量代谢, 运输	[19]
全基因组微阵列	两优2163, 两优2186, 两优培九	旗叶	碳水化合物和能量代谢	[20]
RNA-seq	日本晴、91-11间正反交	苗期地上部	能量代谢	[21,22]
RNA-seq	广陆矮、特青、93-11间正反交	叶片	NA	[23]
RNA-seq	协优9308	根	碳水化合物代谢, 植物激素信号转导	[24]
RNA-seq	协优9308	根	胁迫反应	[25]
RNA-seq	日本晴、91-11间正反交	叶片	碳水化合物代谢, 氮素同化	[26]

杂交种中上调, 而仅25个基因在杂交种中的表达下调. 进一步分析显示, 在杂交种中表达上调的基因与碳和氮素同化(包括叶片中的光合作用和根中氮的吸收)的增强相关; 在杂交种中表达下调的基因包括一个编码光呼吸中关键酶的基因. Song等人^[12]于2007年采用改进的策略, 重新分析了以上SAGE数据, 鉴别出了406个在杂交种“两优培九”中表达上调的基因, 其功能多与叶和根中碳同化相关. 与Bao等人^[11]的研究不同的是, Song等人也同时鉴别许多(469个)在杂交种中表达下调的基因, 其中442个基因的功能与信号转导及蛋白质加工相关. Song等人^[13]于2010年采用SAGE技术比较超级稻“两优2186”与其亲本间旗叶的转录组, 鉴别到1183个差异表达基因, 这些基因显著富集在光合作用和碳固定生物学路径上, 其中大部分碳固定路径中的关键基因在杂交种中表达上调. 同时, 分析也发现杂交种中的差异表达基因与水稻产量性状QTL相关联.

cDNA微阵列是以分子杂交为基本原理、检测基因表达水平和分离差异表达基因的一种较高通量基因组分析技术^[14]. Huang等人^[15]于2006年采用包含9198个惟一表达序列标签(expressed sequence tags, EST)的cDNA微阵列, 测定和比较了优良杂交稻“汕优63”与其亲本“珍汕97”(母本)和“明恢63”(父本)之间3个发育阶段幼穗的基因表达谱. 分析发现, 与中亲值相比, 有141个EST的表达在杂交种发生了显著变化(原文中以表达的中亲杂种优势表示), 其中仅有7个EST在杂交种的3个发育阶段均表现出上调(原文中以表达的正向杂种优势表示), 而有50个EST在杂交种的3个发育阶段均表现出下调(原文中以表达的负向杂种优势表示). 对这些差异表达EST所对应的基因进行功能分类分析发现, 在DNA复制与修复中起作用的基因倾向于上调, 而与碳水化合物、能量和脂代谢、翻译、蛋白质降解相关的基因倾向于下调. 与氨基酸代谢、转录、信号转导、植物防御以及运输相关的基因则表现出上调和下调两个方向的变化. Huang等人^[16]还于同年应用同样的cDNA微阵列测定与比较了超级稻“两优培九”与其亲本间萌发72 h后地上部和根部的基因表达谱, 鉴别出了214个与中亲值相比, 在杂交种中发生差异表达的EST, 其中117个EST与功能已知的基因同源, 它们涉及基础代谢、遗传信息加工、细胞生长和死亡、信号转导、运输和胁迫反应生物学过程. Ge等人^[17]则采用EST测序的方

法测定了“两优培九”及其亲本成熟胚的基因表达谱, 鉴别出了191个在杂交种和亲本间差异表达的基因. 这些基因在杂交种中表现出各种表达变化模式(包括正负超显性、低亲和高亲显性以及加性表达). 大多数与胚发育相关的基因在杂交种中特定地富集且显著上调.

水稻基因组序列及其注释的完善为在全基因组范围进行基因差异表达分析提供了条件. Zhang等人^[18]于2008年首次采用全基因组微阵列的方法对“两优培九”杂交组合以及籼粳亚种间杂交组合“日本晴”×“91-11”的苗期地上部进行了高通量全基因组水平基因差异表达分析. 由于分析的通量得到显著提高, 大量在杂交种与亲本间发生表达差异的基因被鉴别出来, 其中数以千计基因的表达呈现出非加性模式(即与亲本的平均值相比差异显著). 例如, 在“两优培九”和“日本晴”×“91-11”杂交种中, 分别检测到1171个和1361个基因发生非加性表达变化. 功能分类分析发现, 在两个杂交种中, 差异表达基因均富集在碳水化合物、维生素、氨基酸代谢以及次级代谢产物的合成路径中. 值得注意的是, 碳代谢中的关键限速步骤卡尔文循环、和赤霉素生物合成中相关基因在杂交种中显著上调. Wei等人^[19]于2009年也采用全基因组微阵列的方法, 测定和比较了“两优培九”与其亲本不同发育时期的叶片和穗的转录组, 共检测到3926个基因在杂交组合中发生了表达差异. 功能分析发现, 于多个发育阶段特定在杂交种与亲本之间差异表达的基因最显著富集在能量代谢和运输类生物学途径, 在某些发育阶段差异表达基因也富集在氨基酸、蛋白质和碳水化合物代谢以及信号转导路径中. 令人感兴趣的是, 他们还发现一些在杂交种与亲本之间差异表达的基因与水稻中已知的产量性状QTL相关联, 由此推测这些基因可能是水稻产量杂种优势的潜在候选基因. Peng等人^[20]于2014年采用全基因组微阵列的方法测定和比较了超级稻“两优2163”、“两优2186”和“两优培九”及其亲本开花期和灌浆期的转录本, 在这3个杂交组合中分别鉴别出了1193, 1630和1046个差异表达基因. 功能分析发现, 这些差异表达基因富集在碳水化合物和能量代谢尤其是碳固定路径中. 并且, 在3个杂交组合中, 80%的差异表达基因定位于Gramene数据库中的水稻QTL上, 而其中超过90%的差异表达基因定位于产量相关QTL上.

RNA测序(RNA-sequencing, RNA-seq)是于2008年出现的高通量转录组分析方法。该方法利用第二代高通量测序技术对从细胞中分离的所有mRNA转录本进行测序,从而更精确地测定基因组中每一基因的表达水平^[21]。He等人^[22]于2010年首次利用RNA-seq技术测定和比较了“日本晴”与“91-11”间正反杂交种及其亲本苗期地上部的转录组,检测到2800个基因在正反杂交种中均表现出与中亲值相比变化方向一致的非加性表达模式。功能分类分析显示,这些差异表达基因显著富集在能量代谢途径中,其中编码卡尔文循环中6个关键组分的基因表达均上调。此外,由于可以利用亲本之间基因转录区中的SNP、基于RNA-seq的测序序列来对杂交种转录本的亲本来源进行追踪,He等人^[22]还分析了杂交种中亲本等位基因之间的表达差异,鉴别出了一些发生了等位基因特异性表达的基因(即亲本等位基因在杂交种中表达水平的比例显著偏离1:1),并发现杂交种中亲本等位基因表达水平之比与亲本间该基因表达水平之比呈正相关性。Song等人^[23]于2013年通过采用RNA-seq的方法对“广陆矮”、“特青”和“93-11”之间的6个正反杂交组合中杂交种及其亲本的叶片进行转录组分析,也在全基因组水平上揭示了水稻杂交种中等位基因特异性表达及其与亲本之间基因差异表达的关系。他们还发现,杂交种中的等位基因特异性表达依赖于亲本的基因型,且在杂交种的基因表达差异中起重要贡献。例如,等位基因特异性表达可以解释大部分(79.8%)在杂交种与亲本间的显著基因差异表达(10倍以上)和几乎全部(97.3%)基因在杂交种中表达而在亲本之一中不表达的现象。Zhai等人^[24]于2013年采用RNA-seq技术测定和比较了超级稻“协优9308”及其亲本分蘖期和抽穗期根部的转录组。他们于分蘖期检测到829个基因在杂交种与亲本之间发生了差异表达,而在抽穗期检测到4186个差异表达基因,暗示杂交种中基因差异表达具有发育时期特异性。功能分析发现,于抽穗期的差异表达基因富集在碳水化合物代谢和植物激素信号转导途径中,且这两个途径中的大部分关键基因在杂交种中表达上调。Zhai等人^[25]还利用这些转录组数据鉴别出480个基因在分蘖期或抽穗期的杂交种中发生了等位基因特异性表达,而其中大部分(74%, 355个基因)仅在一个发育时期发生,暗示杂交种中等位基因特异性表达也具有发育时期特异性。E等人^[26]通过对已发表的

数据进行分析,也发现杂交种与其亲本间的基因差异表达具有发育时期特异性,然而,分析结果也显示在杂交种中发生了等位基因特异性表达的基因在不同发育时期具有稳定的表达模式。

在以上研究中所鉴别出来的差异表达基因及其在杂交种中的表达变化模式各不相同,但值得注意的是,在其中一些研究中,水稻杂交种中差异表达基因的生物学功能在碳同化、碳水化合物和能量代谢途径中富集^[11-13,18-20,22,24,26]。在这些途径中,一些对植物生长起关键正调控作用的基因在杂交种中表达上调,可导致杂交种中碳固定和能量利用效率增强,从而表现出生长势和生物量的杂种优势效应。然而,需要采用分子遗传学实验直接证明这些基因在水稻杂种优势形成中的作用并进一步追溯导致这类基因在杂交种中表达变化的上游调控因子。

2 水稻杂交种转录组活性变化的调控

与亲本相比,在杂交种中,转录组活性发生了显著变化,大量的基因呈现出各种表达变化模式,包括加性表达和非加性表达。一些研究基于影响基因表达的顺式调控元件和与其结合的反式作用因子的变化对杂交种中基因差异表达产生的原因进行了探讨。Zhang等人^[18]的研究发现,亲本之间基因启动子区indel(插入/删除)多态性与水稻杂交种中基因差异表达模式相关。加性表达、正/负超显性表达模式的基因所对应的亲本间启动子indel多态性显著高于非差异表达基因,而高亲显性表达模式的基因所对应的亲本间启动子indel多态性显著低于非差异表达基因。因为启动子区的indel可能会导致顺式调控元件的变化,作者由此提出基于基因启动子区顺式调控元件与该元件所结合的反式作用因子(即转录激活因子或转录抑制因子)相互作用来解释杂交种中各种基因表达模式的遗传模型。在该模型中,基因启动子区的indel导致的顺式调控元件的差异与激活型反式作用因子相互作用将导致杂交种中基因的加性、趋高亲和超高亲表达;而顺式调控元件的差异与抑制型反式作用因子的相互作用将导致杂交种中基因的加性、趋低亲和超低亲表达。该模式能解释杂交种中基因表达变化的各种模式,但迄今还待以具体基因为实例来进行证实。

另一些研究则从基因表达的表观遗传调控机制来解释杂交种中的基因表达差异^[22,27,28]。表观遗传

调控是一种不依赖于基因组序列变化来影响基因表达和转录组活性的机制. 真核生物中表观遗传调控机制主要包括DNA甲基化、组蛋白修饰和小RNA等方面^[29,30]. DNA甲基化是通过DNA甲基转移酶在胞嘧啶的5'位置上引入一个甲基. 在植物中, DNA甲基化的最主要的功能是抑制转座子的活性从而保持基因组的稳定性, 部分位于基因启动子和基因区的DNA甲基化可能与基因本身的转录活性相关^[31,32]. He等人^[22]首次在水稻杂交组合中进行了全基因组范围的DNA甲基化分析, 在杂交种中鉴别出了许多差异DNA甲基化基因, 且它们在杂交种中主要在呈现出DNA甲基化水平上调的趋势, 并观察到杂交种与亲本间基因DNA甲基化的差异与这些基因的表达差异呈弱负相关性. 此外, He等人^[22]还在杂交种中观察到等位基因特异性DNA甲基化, 其与亲本之间的DNA甲基化差异相关. Chodavarapu等人^[33]的研究发现, 在水稻杂交种中, 亲本来源染色体的DNA甲基化水平与其对应的亲本染色体的DNA甲基化水平类似, 进一步分析也发现在杂交种中发生了等位基因特异性DNA甲基化, 并推测其可能与等位基因DNA序列差异一起在杂交种中等位基因特异性表达中起作用.

组蛋白修饰是在组成核小体的核心组蛋白上发生的各种转录后共价修饰, 起到调控染色质状态和基因组活性的作用. 组蛋白赖氨酸的乙酰化和甲基化是目前研究得较为清楚的组蛋白修饰, 分别由组蛋白乙酰基转移酶和组蛋白甲基转移酶催化形成. 其中组蛋白乙酰化通常与基因的转录激活相关(如H3K9ac), 而组蛋白甲基化既可与转录激活(如H3K4me3)也可与转录抑制(如H3K27me3)相关^[34-36]. He等人^[22]首次采用染色体免疫共沉淀结合高通量测序方法(chromatin immunoprecipitation followed by next-generation sequencing, ChIP-seq), 在水稻杂交组合中进行了全基因组水平上组蛋白修饰谱的测定与比较, 结果表明, 亲本组蛋白修饰能够较为稳定地传递到杂交种中. 值得注意的是, 在杂交种与其亲本之间, 激活型组蛋白修饰H3K4me3强度的差异与基因表达水平的差异显著正相关. 此外, 在杂交种中也观察到了等位基因特异性组蛋白修饰现象, 其与亲本之间的组蛋白修饰差异相关. Guo等人^[37]利用ChIP-seq技术对水稻杂交种中等位基因特异修饰进行了进一步详细分析, 发现在杂交种中H3K36me3等

位基因特异性修饰在杂交种与亲本之间组蛋白修饰差异中起主要作用, 且与杂交种中等位基因特异性表达显著正相关. Li等人^[38]则通过在水稻杂交种中分别过量表达和失活组蛋白脱乙酰化酶编码基因, 揭示了组蛋白修饰在杂交种中一些非加性表达变化模式中起作用.

小RNA是生物体内天然存在的各类不编码蛋白质产物的小分子RNA. 目前, 在植物中研究得比较深入的两类小RNA包括miRNAs (microRNAs)和siRNAs (small interfering RNAs), 前者主要通过转录和翻译水平上对其靶标mRNA的活性进行调控, 后者主要包括rasiRNA (repeat-associated siRNA), 其通过RNA介导的DNA甲基化(RNA-directed DNA methylation, RdDM)控制转座子的活性和维持基因组稳定性^[39]. He等人^[22]采用高通量测序方法, 测定和比较了水稻籼粳亚种间杂交组合的小RNA转录组, 发现在杂交种与亲本之间, miRNA的差异表达与其靶基因的差异表达呈负相关. 随后, Chen等人^[40]通过进一步采用miRNA微阵列的方法, 在水稻杂交组合中对上述研究中的miRNA和高丰度小RNA的表达水平进行了比较分析, 进一步证实了该发现, 由此表明miRNA可能参与调控杂交种中的基因差异表达. 此外, He等人^[22]发现, siRNA在水稻杂交种中主要呈现加性表达变化模式, 一部分在杂交种中非加性表达的siRNA表现为表达下调. siRNA的表达变化在杂交种转录组活性变化中作用还有待进一步研究. Fang等人^[41]发现, miRNA在杂交种中主要呈现非加性表达模式, 且具有很强的发育时期和组织特异性, 暗示miRNA可能参与调控其靶基因在杂交种中的组织和发育时期特异性表达变化. Zhang等人^[42]通过在“两优培九”杂交组合中测定与比较小RNA转录组, 鉴别出了一些在杂交种中非加性表达的小RNA, 发现在不同组织中, 杂交种中差异表达的小RNA与其靶基因的表达变化负相关, 且靶基因功能富集在包括生长素信号通路在内的重要生物学过程中. 而且, 部分小RNA与其靶基因显著富集在重要农艺性状QTL中, 这些结果暗示小RNA通过调控其靶基因在杂交种中的表达变化而对杂种优势的形成具有重要贡献.

由于亲本之间的基因组序列差异是导致杂交种杂种优势形成的本质前提, 这种遗传上的差异可以直接、或通过顺式或反式的作用方式引起杂交种基因组表观遗传状态的变化, 来调控杂交种中转录组活

性的变化。

3 小结与展望

水稻杂种优势现象是许多数量性状的综合体现,每一性状通常由多个基因或生物学路径共同控制,各基因产物之间或生物学路径之间还会相互作用,组成一个有大量基因参与的复杂调控网络。因此,对水稻杂种优势分子机理的解析需要在全基因组水平上进行,从杂交种转录组活性变化(主要是杂交种与亲本之间全基因组基因差异表达)分析入手,一方面采用基因组学和生物信息学方法分析导致基因差异表达的上游遗传和表观遗传调控机制,另一方面采用分子遗传学手段调查和验证差异表达基因是否直接或通过调控其下游生物学路径而确实参与了杂种优势的形,从而提出对水稻杂种优势形成分子机理的完整解释。在此过程中有两点值得特别关注:(i) 提高基因差异表达的精确性和特异性。基因的活性(基因表达)受细胞内外环境因子的调控,在植物不同组织器官之间及不同生长发育阶段中都会表现出明显差异。杂交种转录组活性也会受到这些因素的影响,是一个不断变化的动态过程,并且还会依据亲本基因型的不同而发生变化。在水稻杂交种转录组活性测定及其与亲本的比较中,由于在所采

用的杂交组合、实验取材的组织器官、转录组测定技术平台和分析方法上的不同,导致不同研究之间杂交种中基因差异表达分析结果存在很大差异,所鉴别到的具体差异表达基因在数量、表达变化模式以及分子和生物学功能上相差较大。因此,有必要针对特定性状,利用多个水稻杂交组合和同一转录组分析平台,在单一器官、组织甚至细胞水平上进行基因差异表达分析。(ii) 将基因差异表达分析与数量遗传学方法结合。目前,通过采用QTL定位方法,已经鉴别出了大量控制水稻不同数量性状的基因或基因组位点,在此基础上,一些研究在水稻杂交种中差异表达基因与水稻产量性状QTL之间建立起了一定的相关性^[19,20],有助于对差异表达基因生物学功能的理解。然而,对某性状杂种优势形成的遗传调控可能并不完全同于该性状本身的调控,因此,有必要针对某特定性状,在该性状杂种优势QTL与杂交种中基因差异表达之间进行相关性分析。此外,最近Huang等人^[43,44]首次在水稻中将GWAS的方法应用于对杂种优势遗传基础的分析,鉴别出了一些对水稻产量杂种优势具有贡献的候选基因。如果将水稻各产量相关性状杂种优势的GWAS分析与杂交种中全基因组基因差异表达分析相结合,将有利于解析水稻产量杂种优势形成的分子机理。

致谢 感谢Abigail Coplin对英文摘要的修改。

参考文献

- 1 Davenport C B. Degeneration, albinism and inbreeding. *Science*, 1908, 28: 454–455
- 2 Bruce A B. The mendelian theory of heredity and the augmentation of vigor. *Science*, 1910, 32: 627–628
- 3 Keeble F, Pellew C. The mode of inheritance of stature and of time of flowering in peas (*Pisum sativum*). *J Genetics*, 1910, 1: 47–56
- 4 Jones D F. Dominance of linked factors as a means of accounting for heterosis. *Proc Natl Acad Sci USA*, 1917, 3: 310–312
- 5 East E M. Inbreeding in corn. In: Reports of the Connecticut Agricultural Experiment Station for Years 1907–1908, Connecticut Agricultural Experiment Station, 1908. 419–428
- 6 Shull G H. The composition of a field of maize. *J Hered*, 1908, 4: 296–301
- 7 Powers L. An expansion of Jones's theory for the explanation of heterosis. *Am Nat*, 1944, 78: 275–280
- 8 Birchler J A, Auger D L, Riddle N C. In search of the molecular basis of heterosis. *Plant Cell*, 2003, 15: 2236–2239
- 9 Yuan L P. A preliminary report on male sterility in rice, *Oryza sativa* L. *Chin Sci Bull*, 1966, 11: 322 [袁隆平. 水稻的雄性不孕性. 科学通报, 1966, 17: 185–188]
- 10 Velculescu V E, Zhang L, Vogelstein B, et al. Serial analysis of gene expression. *Science*, 1995, 270: 484–487
- 11 Bao J, Lee S, Chen C, et al. Serial analysis of gene expression study of a hybrid rice strain (LYP9) and its parental cultivars. *Plant Physiol*, 2005, 138: 1216–1231
- 12 Song S, Qu H, Chen C, et al. Differential gene expression in an elite hybrid rice cultivar (*Oryza sativa* L.) and its parental lines based on SAGE data. *BMC Plant Biol*, 2007, 7: 49
- 13 Song G S, Zhai H L, Peng Y G, et al. Comparative transcriptional profiling and preliminary study on heterosis mechanism of super-hybrid

- rice. *Mol Plant*, 2010, 3: 1012–1025
- 14 Schena M, Shalon D, Davis RW, et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 1995, 270: 467–470
- 15 Huang Y, Zhang L, Zhang J, et al. Heterosis and polymorphisms of gene expression in an elite rice hybrid as revealed by a microarray analysis of 9198 unique ESTs. *Plant Mol Biol*, 2006, 62: 579–591
- 16 Huang Y, Li L, Chen Y, et al. Comparative analysis of gene expression at early seedling stage between a rice hybrid and its parents using a cDNA microarray of 9198 uni-sequences. *Sci China Ser C: Life Sci*, 2006, 49: 519–529
- 17 Ge X, Chen W, Song S, et al. Transcriptomic profiling of mature embryo from an elite super-hybrid rice LYP9 and its parental lines. *BMC Plant Biol*, 2008, 8: 114
- 18 Zhang H Y, He H, Chen L B, et al. A genome-wide transcription analysis reveals a close correlation of promoter INDEL polymorphism and heterotic gene expression in rice hybrids. *Mol Plant*, 2008, 1: 720–731
- 19 Wei G, Tao Y, Liu G, et al. A transcriptomic analysis of superhybrid rice LYP9 and its parents. *Proc Natl Acad Sci USA*, 2009, 106: 7695–7701
- 20 Peng Y, Wei G, Zhang L, et al. Comparative transcriptional profiling of three super-hybrid rice combinations. *Int J Mol Sci*, 2014, 15: 3799–3815
- 21 Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, 2009, 10: 57–63
- 22 He G, Zhu X, Elling A A, et al. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell*, 2010, 22: 17–33
- 23 Song G, Guo Z, Liu Z, et al. Global RNA sequencing reveals that genotype-dependent allele-specific expression contributes to differential expression in rice F1 hybrids. *BMC Plant Biol*, 2013, 13: 221
- 24 Zhai R, Feng Y, Wang H, et al. Transcriptome analysis of rice root heterosis by RNA-Seq. *BMC Genomics*, 2013, 14: 19
- 25 Zhai R, Feng Y, Zhan X, et al. Identification of transcriptome SNPs for assessing allele-specific gene expression in a super-hybrid rice Xieyou9308. *PLoS One*, 2013, 8: e60668
- 26 E Z, Huang S, Zhang Y, et al. Genome-wide transcriptome profiles of rice hybrids and their parents. *Int J Mol Sci*, 2014, 15: 20833–20845
- 27 Zhao Y, Zhou D X. Epigenomic modification and epigenetic regulation in rice. *J Genet Genomics*, 2012, 39: 307–315
- 28 He G, He H, Deng XW. Epigenetic variations in plant hybrids and their potential roles in heterosis. *J Genet Genomics*, 2013, 40: 205–210
- 29 Henderson I R, Jacobsen S E. Epigenetic inheritance in plants. *Nature*, 2007, 447: 418–424
- 30 Zhang X. The epigenetic landscape of plants. *Science*, 2008, 320: 489–492
- 31 Chan S W, Henderson I R, Jacobsen S E. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat Rev Genet*, 2005, 6: 351–360
- 32 Law J A, Jacobsen S E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*, 2010, 11: 204–220
- 33 Chodavarapu R K, Feng S, Ding B, et al. Transcriptome and methylome interactions in rice hybrids. *Proc Natl Acad Sci USA*, 2012, 109: 12040–12045
- 34 Shahbazian M D, Grunstein M. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem*, 2007, 76: 75–100
- 35 Liu C, Lu F, Cui X, et al. Histone methylation in higher plants. *Annu Rev Plant Biol*, 2010, 61: 395–420
- 36 He G, Elling A A, Deng X W. The epigenome and plant development. *Annu Rev Plant Biol*, 2011, 62: 411–435
- 37 Guo Z, Song G, Liu Z, et al. Global epigenomic analysis indicates that epialleles contribute to Allele-specific expression via Allele-specific histone modifications in hybrid rice. *BMC Genomics*, 2015, 16: 232
- 38 Li C, Huang L, Xu C, et al. Altered levels of histone deacetylase OsHDT1 affect differential gene expression patterns in hybrid rice. *PLoS One*, 2011, 6: e21789
- 39 Chen X. Small RNAs and their roles in plant development. *Annu Rev Cell Dev Biol*, 2009, 25: 21–44
- 40 Chen F, He G, He H, et al. Expression analysis of miRNAs and highly-expressed small RNAs in two rice subspecies and their reciprocal hybrids. *J Integr Plant Biol*, 2010, 52: 971–980
- 41 Fang R, Li L, Li J. Spatial and temporal expression modes of microRNAs in an elite rice hybrid and its parental lines. *Planta*, 2013, 238: 259–269
- 42 Zhang L, Peng Y, Wei X, et al. Small RNAs as important regulators for the hybrid vigour of super-hybrid rice. *J Exp Bot*, 2014, 65: 5989–6002
- 43 Huang X, Yang S, Gong J, et al. Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. *Nat Commun*, 2015, 6: 6258
- 44 Huang X, Yang S, Gong J, et al. Genomic architecture of heterosis for yield traits in rice. *Nature*, 2016, doi: 10.1038/nature19760

The transcriptomic basis of heterosis in rice

HE GuangMing, HE Hang & DENG XingWang

School of Advanced Agricultural Sciences and School of Life Sciences, Peking University; State Key Laboratory of Protein and Plant Gene Research; Peking-Tsinghua Center for Life Sciences, Beijing 100871, China

Heterosis, or hybrid vigor, refers to the superior performance of hybrids relative to their parents. Although it has been utilized successfully in agriculture for many years, the genetic basis of heterosis remain largely unclear. The classical explanations of this phenomenon, including dominance, overdominance, and epistasis, ultimately tell us little about the mechanisms driving heterosis at the molecular level. Since heterosis generally manifests in a plurality of observable quantitative traits, it is believed that many genes and/or genetic loci contribute to heterosis. The continued development of genomics and bioinformatics has recently enabled us to conduct whole genome level investigations of gene action in hybrids. The expression of many parental alleles is known to be altered in hybrids, which in turn, changes the transcriptomic activity of hybrids compared to their parents. Consequently, by first comparing the divergences in transcriptomic activity observed between hybrids and their parents, and then determining the connection between this differential gene expression and the altered phenotypes observed in hybrids, it is possible to ascertain the mechanisms driving heterosis at the molecular level.

Rice has long been one of the most important crops worldwide. Heterosis in rice has been successfully utilized as a strategy for increasing both the yield and quality of this essential crop for several decades. In recent years, many studies have explored the genomic basis of heterosis in hybrid rice. To date, however, many of these genome-wide analyses of differential gene expression have exhibited a number of inconsistencies in their identification of both differentially expressed genes and patterns of gene expression. These discrepancies may stem from the different hybrid combinations, distinct tissues, and variety of genomic platforms utilized in these studies. Some studies do suggest, however, that the genes differentially expressed in rice hybrids may enhance photosynthesis, carbohydrate metabolism, energy metabolism, and other such integral biological pathways. Furthermore, some key genes required for plant growth have been shown to be unregulated in hybrids, which, in turn, suggests that increased carbon assimilation and more efficient energy utilization may be responsible for the enhanced growth observed in hybrids compared to their parents. Gene expression can be regulated by genetic components, such as *cis*-elements in the gene promoter regions and the *trans*-acting factors to which the *cis*-elements bind. Epigenetic components, such as DNA methylation, histone modification, and small RNAs, were also known to affect gene expression. Previous research has indicated that both genetic and epigenetic variations may contribute to the altered transcriptomic activity documented in rice hybrids. The mechanisms regulating these changes, however, remain unspecified, and thus merit further investigation. In order to gain a deeper understanding of the molecular mechanisms governing heterosis in rice hybrids in the future, it will be necessary to improve the accuracy and specificity of differential gene expression analyses in rice hybrid combinations. Similarly, future heterosis research should focus on the molecular basis of specific biological traits. Finally, we should also combine transcriptomic variation analysis with other methods of quantitative genetics, such as heterosis QTL (quantitative trait loci) mapping and GWAS (genome-wide association studies).

rice, heterosis, transcriptome

doi: 10.1360/N972016-01046