

种质资源学与基因组学相结合-破解基因发掘 与育种利用的难题

张学勇, 郝晨阳, 焦成智, 李甜, 毛龙, 刘旭
(中国农业科学院作物科学研究所, 北京 100081)

摘要: 2010年前后, 随着各大作物模式品种基因组测序的完成, 拟南芥、水稻、玉米等的重测序研究, 突破了分子标记数量的限制, 带动作物科学研究全面进入基因组时代, 大量代表性品种、种质资源完成了重测序工作, 数以百万甚至千万计的 SNP 标记, 全基因组关联分析 (GWAS, genome-wide association study) 广泛应用于遗传资源研究, 使近 10 年成为种质资源研究的黄金期, 通过 GWAS 解析一些复杂重要农艺性状的遗传基础, 成为 Cell、Nature 和 Science 及其子刊这一时期的主要内容, 推动种质资源学迈入一个崭新时代。

20 世纪, 遗传育种学的发展和不断完善推动了种质资源学科的萌芽和初步形成, 而 21 世纪基因组学的发展和广泛应用, 逐步形成了种质资源学推动育种学发展的新局面, 一些长期困扰育种家的问题, 通过 GWAS 分析得到了重要启示或答案 (如番茄的驯化、育种史、品质与产量矛盾问题, 小麦骨干亲本等)。而泛基因组研究的迅速发展, 突破了单一参考基因组的局限性, 使我们认识到品种间基因组结构变异的普遍性, 为深度解析重大品种、骨干亲本的形成及突破性资源的创制提供了更加宽广的视野。巢式关联作图群体 (NAM, nested association mapping)、多亲本互交群体 (MAGIC, multi-parent advanced generation inter-cross) 及以此为基础衍生的构建多亲本遗传群体的思路和实践, 使研究群体的遗传背景水平达到同期育种要求, 不仅加快了基因的精细定位, 并为组装育种提供了平台, 推动着种质资源学、基因组学和育种学的融合与互动, 开启了以基因组信息为支撑的基因资源和分子设计育种新时代, 也预示着大学科融合与调整时代的到来。

关键词: 种质资源学; 基因组学; 基因发掘; 作物育种

Integration of Germplasmics and Genomics: Bridging up Crop Gene Discovery and Breeding

ZHANG Xue-yong, HAO Chen-yang, JIAO Cheng-zhi, LI Tian, MAO Long, LIU Xu
(Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081)

Abstract: Since 1990s, scientists have mapped many quantitative trait loci through bi-parental populations such as NILs, DH, and IBD in plants through co-segregation between traits and markers. A set of important genes have been isolated by fine mapping, which are relating to biotic and abiotic stress resistance, environment adaptation, quality, and yield. These bring a lot of new biological knowledge. However, most of these genes were selected successfully in domestication and breeding.

By 2010, with the establishment of reference genome sequences for most crops, genome re-sequencing has broken the limitation of the number of molecular markers, driving crop genetic research fully entering the genome era. Millions of SNP markers have made genome-wide association study (GWAS) widely used in genetic resources research, making it the most important part of germplasm resources research in the past 10 years. In this golden period, dissecting genetic basis of important agronomic traits through GWAS has become the

收稿日期: 2023-01-04 修回日期: 2023-01-05 网络出版日期: 2023-01-11

URL: <https://doi.org/10.13430/j.cnki.jpgr.20230104001>

第一作者研究方向为小麦种质资源与基因组学, E-mail: zhangxueyong@caas.cn

基金项目: 国家重点研发计划 (2016YFD0100302); 中国农科院基本科研业务费 (Y2017PTC9)

Foundation projects: The Key Research and Development Program of China (2016YFD0100302); The Central Public-interest Scientific Institution Basal Research Fund (Y2017PTC9)

main content of *Cell*, *Nature* and *Science* and their sub-publishers.

In the last century, the establishment and development of crop breeding science promoted establishment of germplasm resource discipline. The development and wide application of genomics in this century has gradually formed a new situation in which research of germplasm resources promotes the development of breeding. Many problems that plague breeders have received important enlightenment or answers through GWAS analysis (such as tomato domestication, breeding history, contradiction between quality and yield, founder genotype, etc.). Further, the pan-genome research breaks through the limitations of a single reference genome, makes researchers realize the universality of genomic structural variations within a species, and provides a broader perspective for analyzing the formation of landmark cultivars, founder genotypes and the creation new super-cultivars. NAM, MAGIC, and the ideas and practice of constructing new multi-parents genetic mapping populations based on these can make the background level of the mapping population basically reaching the level of the breeding population at the same time, speeding up the fine mapping of agronomically important genes and pre-breeding. This has driven integration and interaction of germplasmics, genomics and breeding. Furthermore, in assistance of speed breeding and gene editing technology, the integration and interaction will surely promote the three disciplines to enter a new time.

Key words: germplasmics; genomics; gene discovery; crop breeding

遗传多样性是一个物种在自然界生存的基础,也是作物遗传改良的基石。经过近 100 年的考察、收集、评价、编目、入库,全球收集种质资源约 740 万份,存放在 1750 个种质库中;我国收集各类作物种质资源总量达 50 万份^[1]。丰富的资源一方面为作物改良奠定了基础,另一方面也带来了困扰,即如何进行有效利用^[2]。全球科学家对此进行了长期的思考与探索^[3-4],提出通过建立核心种质^[5]提高种质资源的管理和利用效率,我国也启动了主要农作物“核心种质”的重点基础研究项目,对各大作物遗传多样性的认识和浓缩起到了积极地推动作用^[6-8]。如何将建立的核心种质有效融入到育种和基因组学研究,并发现育种急需的关键基因,成为过去 10 年及未来很长一段时间作物种质资源学研究的核心问题之一,也是实现学科融合与发展的关键所在^[4,9-10]。在国家重点基础研究发展计划(973 项目)和十三五国家重点研发计划七大农作物育种试点专项和中国农科院基本科研业务费项目的支持下,对此进行了 20 余年的思考和探索,本研究既是对过去一些实践的总结,也是对未来工作的思考。

1 建立核心种质、浓缩遗传变异:突破巨量种质资源对研究人员财力、物力和心理的压力

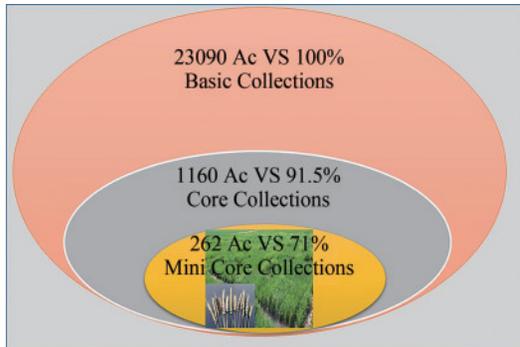
“核心种质”一词最早由澳大利亚的科学家 Frankel 和 Brown 提出,是在应对大量作物种质资源收集样本的整理、繁殖、储存、发放和应用难题中提

出的,即如何利用最小的样本数代表最大的遗传多样性,以实现便于管理、便于应用、便于研究的目的,一般认为其总量不应超过基础种质的 10%^[5]。核心种质的建立主要依据作物种质资源的总收集量(份数)、来源地、产生或形成的方式(地方品种、引进品种、育成品种)、主要编目性状等。常用的取样方法主要包括简单按比例取样(资源地区分布比较均衡)、平方根法、对数法(资源地区分布不均衡),入选样本应考虑代表性和优先性,对育种或生产做出突出贡献的品种、品系优先入选。

2000 年以前,世界范围内建立的核心种质主要依靠田间表型数据,如植株高矮、叶片大小、成熟期、抗病表现、籽粒或果实大小、颜色等,因不同的国家、地区关注或感兴趣的性状不尽相同,收集数据的一致性、可比性和兼容性存在比较大的困难,因此只在一些小宗作物或资源数量不是太多的作物中尝试建立核心种质,并未引起广泛关注。20 世纪 90 年代,DNA 分子标记技术的迅速发展和应用,推动了大作物(如小麦、水稻、玉米、大豆)核心种质建立技术和方法的探索,以样品的 DNA 分子指纹数据为基础,通过聚类分析,筛选具有更高代表性的样品入选核心种质,受到学术界的广泛接受。

我国在国家重点基础研究发展计划(简称 973 项目)支持下,以表型数据为基础,DNA 分子标记为主导,于 2003 年在全球率先建立了水稻、小麦、大豆三大作物的核心种质,分别用 5% 的核心样品(水稻:3100 份,小麦:1160 份,大豆:1400 份)囊括了基础

样品(水稻61500份,小麦23100份,大豆28800份)约90%的遗传变异,达到浓缩遗传变异,明确多样性地理分布特点,实现方便管理、研究与应用,在此基础上压缩形成的微核心种质(Mini core collection),已成为我国三大作物深入研究的重要材料平台(图1),为其他作物建立核心种质起到了很好的引领作用,推动了全国乃至全球的作物种质资源研究。



Ac: 种质资源; Basic Collections: 基础种质; Core Collections: 核心种质; Mini Core Collections: 微核心种质^[7]
Ac: Accession^[7]

图1 中国小麦核心种质、微核心种质及其代表性
Fig.1 Chinese wheat core collections, mini core collections and their genetic representation

随着高通量分子标记技术体系(SNP芯片、GBS)和二代测序技术的广泛利用^[11],基于长期积累的自然变异与人工变异和重组,通过全基因组关联分析(GWAS, genome-wide association study),发掘重要农艺性状基因,揭示作物重要性状的演变规律,已成为作物种质资源学和应用基因组学研究的前沿。

核心种质因浓缩了大量的自然和人工(育种)变异和重组,逐步成为通过GWAS方法发现重要基因的材料平台^[12]。但不可否认,以此为基础所获得的信息(强关联信号)大多是主效基因,对基础生物研究很重要,但对育种价值有待商榷,因为在早期的驯化和育种中对这些基因的选择已基本完成,甚至在一些生态区域内的品种或资源中发生固定现象(Fixed)。如何发掘潜在的中等作用基因,客观评价其效应,开发有效标记并应用于育种,是应用基因组学研究的重要内容。国内外提出通过构建多亲本遗传群体,如巢式关联作图群体(NAM, nested association mapping)^[13]、高世代回交-巢式关联作图群体(AB-NAM, advanced backcross-nested association mapping)^[14]、多亲本高级世代互交群体(MAGIC, multi-parent advanced generation inter-cross)^[15-16]和基于双列杂交且包含育种选择的多亲本高世代自

交群体(CUBIC, complete-diallel design plus unbalanced breeding-like inter-cross)^[17]来解决上述难题^[18]。

2 以模式品种参考基因组为基础,通过核心种质重测序和高密度SNP芯片分析,规模化发掘重要基因

果实大小是作物驯化和品种改良的重要性状之一,也是作物产量三要素中遗传力最强的性状。早期对果实(籽粒)大小的研究主要通过双亲群体的QTL精细作图,克隆候选基因,进而通过转基因及基因干扰等技术获得关键基因,我国及日本科学家对水稻的籽粒大小进行了比较系统地研究^[19],并推动了小麦、玉米等作物产量基因的研究。

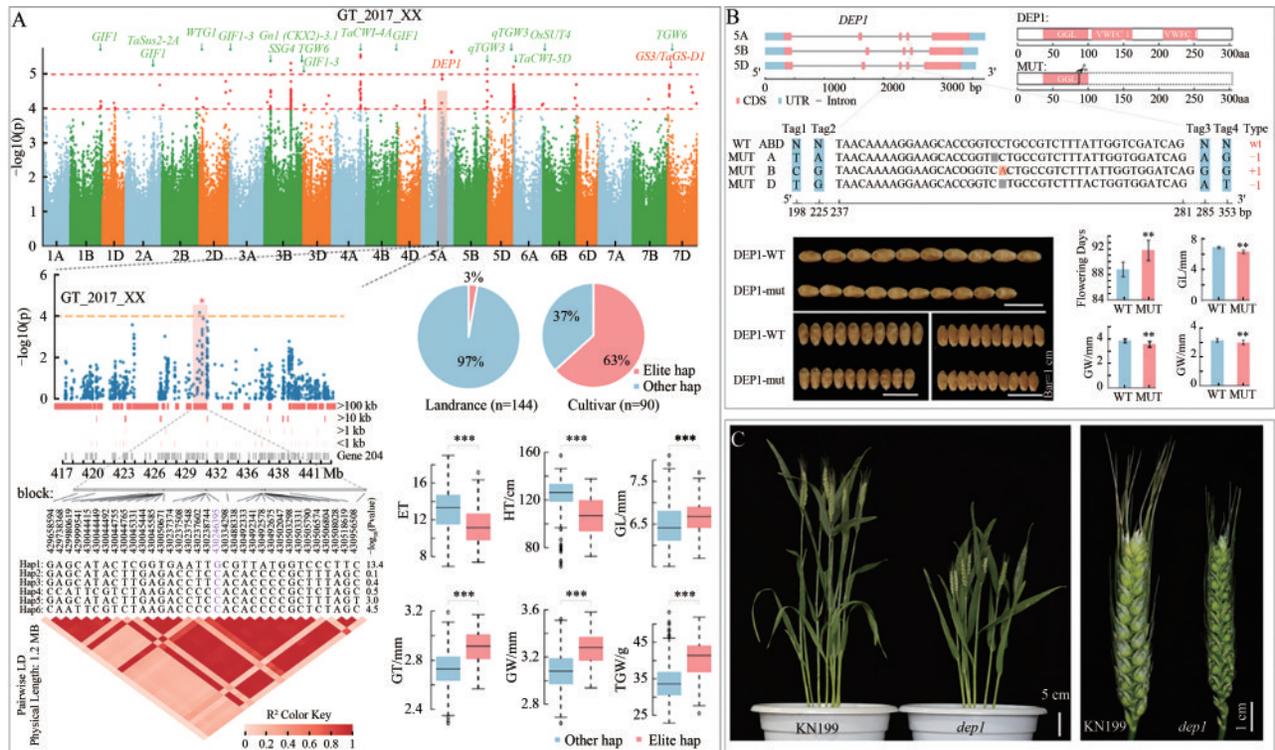
中国农业科学院作物科学研究所毛龙和张学勇团队,以中国春基因组序列为参考,通过对小麦微核心种质的外显子捕获,获得大量基因的SNP标记,结合多年的表型数据分析和基因编辑技术,克隆了控制小麦籽粒大小的重要基因*DEP1*,其紧邻小麦5A染色体着丝粒。研究发现该基因是重要的育种选择基因,*TaDEP1-Hap1*不仅增加粒长、粒宽、粒厚、千粒重,同时降低株高,表现出位点或基因的多效性,进一步分析发现其与小麦中的赤霉素信号传递系统互作,调控氮素的吸收和利用,为庞大、复杂基因组物种的基因克隆摸索出一条新的思路^[12](图2)。西北农林科技大学毛虎德利用小麦微核心种质,发现ABA受体基因(*PYLs-1B*)启动子区一个442 bp的插入,显著提升了该基因的表达,提高了小麦品种的耐旱性和千粒重,是抗旱育种的重要选择基因^[20]。

中国农业科学院蔬菜花卉研究所黄三文课题组通过基因组重测序,发现番茄从野生种到现代加工型品种,浆果平均重量大幅度提高是通过两个大的历史阶段完成,第一阶段(驯化)在番茄基因组中186个区段出现选择谷,选择的主要基因包括*fw1.1*、*fw5.2*、*fw7.2*、*fw12.1*、*icn12.1*等;第二阶段(现代育种)在133个区段形成选择谷,重点选择*fw2.1*、*fw2.2*、*fw2.3*、*fw11.1*、*fw11.2*、*fw11.3*、*inc2.2*等13个目标基因,研究还发现第一阶段形成的186个选择区段中有21%在第二阶段持续受到选择^[21],这为通过关联分析,获得候选基因探索出一条新的思路(图3)。

对一些育种选择的重要基因,通过基因单倍型在育成品种和地方品种中的分布比较,就可对一个

基因的主要功能做出初步判断,再通过全球育成品种的比较分析,即可更进一步验证判断的准确性,为基因功能的分子解析提供基本的理论支撑(图4)。从图4中可以清晰的看出,小麦 *VRN-5A* 基

因与品种的耐寒性密切相关,单倍型 *Hap3* 作为春化对低温需求比较强的类型,主要分布在我国和全球冬季比较寒冷的冬麦品种中,而在澳大利亚和南美品种中则以 *Hap8* 为主。

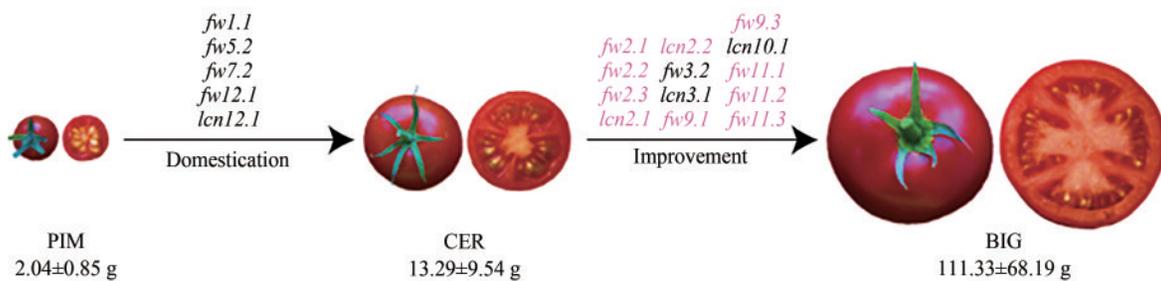


A: 5A 染色体长臂上与 GT 相关的区段覆盖了 1.3 Mb 的 LD 区段,该区段包含 15 个基因,其中包括水稻直立密穗基因 (*DEP1*); *TaDEP1* 所在区段形成 6 种单倍型,其中单倍型 1 (Hap1) 与粒厚显著关联,其分布频率从地方品种的 3% 增加到栽培品种的 63%,表明该单倍型在小麦现代育种中受到了强烈的正向选择; B, C: 敲除 *DEP1* 导致籽粒长度,宽度和厚度三个维度均显著下降,同时突变体表现为株高降低,穗变细^[12]

A: The GT-related genomic region on the long arm of chromosome 5A covers a 1.3 Mb LD block, and this region contains 15 candidate genes including the rice *DENSE AND ERECT PANICLE 1* ortholog *TaDEP1*. Polymorphic loci of *TaDEP1* form six haplotypes, and the haplotype Hap1 is significantly associated with grain length, and its frequency increased from 3% in Chinese landraces to 63% in modern Chinese cultivars, indicating it was selected positively in modern wheat breeding; B, C: Knocking down of *DEP1* resulted in a significant decrease in grain length, width and thickness, as well as plant height reduction and spike thinning^[12]

图2 小麦5A染色体上 *TaDEP1* 基因的单倍型遗传效应及验证

Fig.2 Genetic effect and verification of haplotypes at *TaDEP1* on chromosome 5A



PIM: 醋栗番茄; CER: 樱桃番茄; BIG: 栽培番茄^[21-22]

PIM: *S. pimpinellifolium*; CER: *S. lycopersicum* var. *cerasiforme*; BIG: big-fruited *S. lycopersicum*^[21-22]

图3 番茄驯化和改良过程果实从小变大选择的关键基因

Fig.3 Key genes selected for fruit size in tomato domestication and improvement

20世纪60-80年代,小麦、水稻育种曾发生了一场举世公认的、对粮食生产和人类生存产生巨大影

响的事件,被称为绿色革命,即通过矮化育种,解决作物的田间倒伏问题。然而,化肥的大量使用一方

面显著提高了作物的产量,但残余的氮、磷元素对地下水、江河湖泊的污染也是有目共睹的,如何在保证产量的前提下,提高品种的肥料利用效率,减少化肥的施用量已成为我国乃至全球发达国家的一个生态难题。我国科学家储成才、傅向东等对其进行了深入而卓有成效地研究。储成才研究团队克隆了关键基因 *OsTCP19*,其通过调节分蘖促进基因的表达,介导氮触发的发育过程。发现不同等位变异对土壤中N的吸收利用率显著不同,在土壤贫瘠的地方, *OsTCP19* 氮高效变异为主,并随着土壤氮含量的增加,氮高效类型品种逐步减少,我国现代水稻品种中氮高效变异几乎全部丢失,但野生稻中 *OsTCP19-H* 的等位基因频率却很高,表明 *OsTCP19-H* 在氮含量较低的自然土壤中经历了正向选择。将这一氮高效变异重新导入现代水稻品种,在减少氮素的条件下,氮肥利用效率可提高20%~30%,使用较少的化肥,也能达到相同的产量^[23]。

3 以当家品种为受体、核心种质为供体,构建新的人工群体,促进库存资源在基因组学研究和育种中的应用

在建立小麦核心种质的基础上,中国农业科学院作物科学研究所小麦基因资源挖掘与利用创新团队以生产上正在大面积推广应用的周麦18、郑麦366、邯鄯6172为受体,核心种质为供体,通过杂交、回交,在早期分离世代适度的选择,并在亚群间进行优良选系的互交,以实现三大亲本优良区段及来自核心种质优良区段和基因的充分整合,建立了一套综合NAM和MAGIC群体优点,与育种研究相衔接的小麦回交-互交-巢式关联作图群体(AB-NAMIC, advanced backcross-nested association mapping plus inter-crossed)群体,并对其进行了多年的表型鉴定,发现AB-NAMIC群体的产量相关性状明显提高,生育期性状明显缩短(图5)。与自然群体相比,AB-NAMIC群体显著提高了GWAS发掘基因的效率(图6),并从中选育出中麦66、中优69、中优88、中麦110等优良品种(系),为种质资源的有效利用探索出一条比较高效的途径(图7)。

4 建立代表性品种的泛基因组,解析基因组结构变异与重大品种形成之关系

随着测序技术的迅速发展和成本的不断降低,

越来越多的种质资源被重测序,人们发现许多重要的基因组区段或基因并不存在于最初测定的单个参考基因组中,于是提出了泛基因组的概念^[25],也就是一个物种内存在于不同个体中的全部功能基因的总和。这些基因既包括存在于所有个体的核心基因(Core gene),也包括在不同个体分散存在的非核心基因(Dispensable gene)(图8)。由于驯化和育种充分利用的是物种内部的变异类型,对于一个物种来说,非核心基因(Dispensable gene)恰恰提供了驯化和育种的重要变异来源,是重要性状的形成的基础。开展对作物泛基因组的研究,探讨非核心基因对作物重要性状的影响已成为基因组研究的一个新热点^[26]。

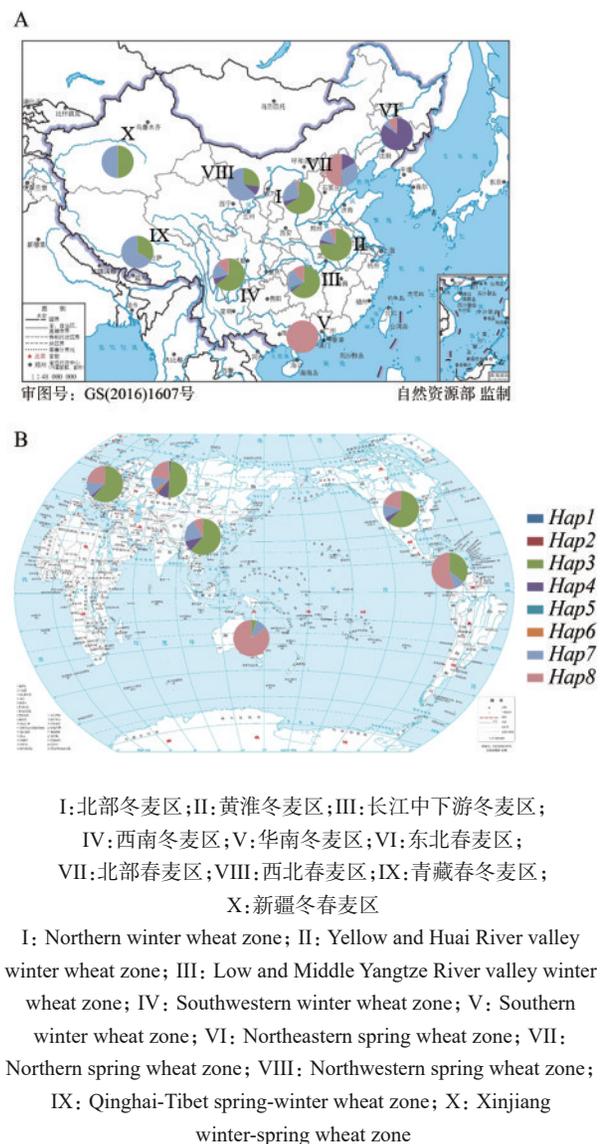
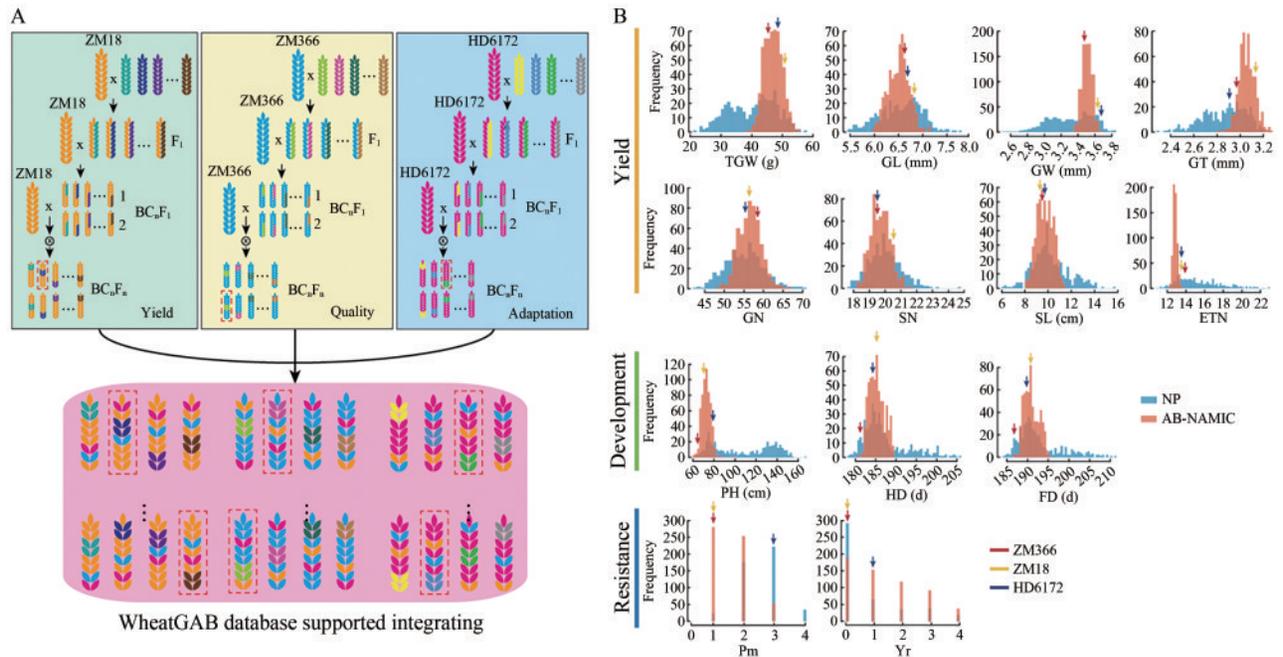


图4 小麦 *VRN1-5A* 单倍型在我国(A)及全球(B)主要产区育成品种中的分布

Fig.4 Haplotype distribution of *TaVRN1-5A* in Chinese (A) and global (B) wheat modern cultivars

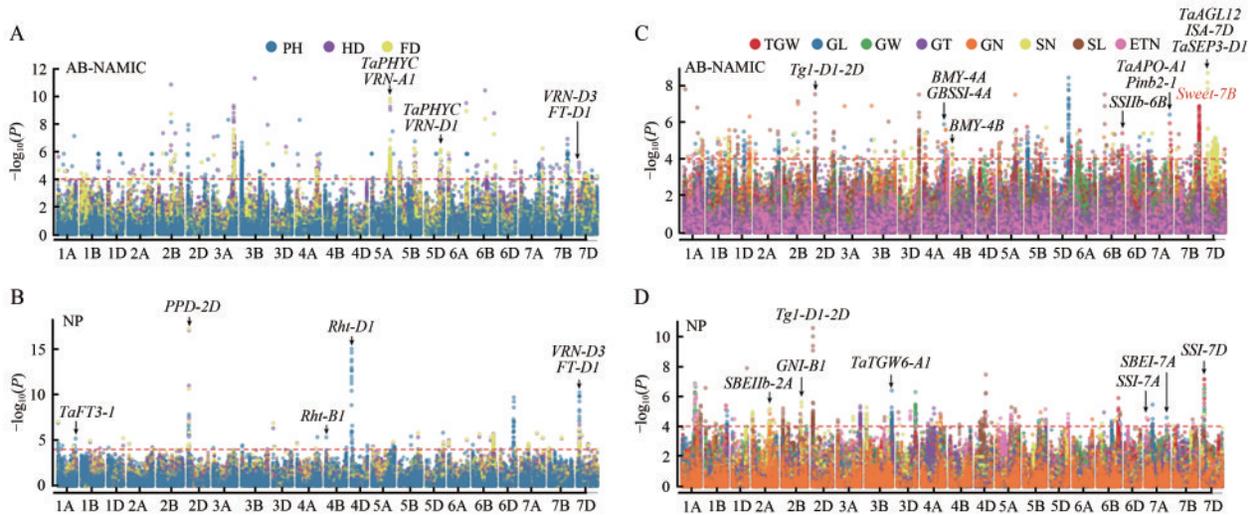


ZM18:周麦18;ZM366:郑麦366;HD6172:邯鄲6172;NP:自然群体;TGW:千粒重;GL:粒长;GW:粒宽;GT:粒厚;GN:穗粒数;SN:每穗小穗数;SL:穗长;ETN:有效分蘖数;PH:株高;HD:抽穗期;FD:开花期;Pm:白粉病;Yr:条锈病^[24];下同

ZM18: Zhoumai 18; ZM366: Zhengmai 366; HD6172: Handan 6172; NP: Natural population; TGW: Thousand-grain weight; GL: Grain length; GW: Grain width; GT: Grain thickness; GN: Grain number per spike; SN: Spikelet number per spike; SL: Spike length; ETN: Effective tiller number; PH: Plant height; HD: Heading date; FD: Flowering date; Pm: Powdery mildew; Yr: Yellow rust^[24]; The same as below

图5 AB-NAMIC群体构建流程(A)及6年表型鉴定评价性状BLUP分布(B)

Fig.5 AB-NAMIC population construction (A) and the traits distribution based on the BLUP values from six environments (B)



A、B: 基于BLUP值的AB-NAMIC群体(A)和自然群体(B)的发育性状全基因组关联分析;

C、D: 两套群体产量相关性状的基因组关联分析^[24]

A, B: GWAS Manhattan plot of traits in AB-NAMIC and NP panels using BLUP data for development traits; C, D: Yield traits^[24]

图6 AB-NAMIC群体显著提高了GWAS的解析率和精度(与自然群体相比)

Fig.6 AB-NAMIC population-based GWAS significantly improved the resolution rate and accuracy when compared with natural population

稀有等位变异(重要基因)的发现和利用是育种取得突破的关键,但这些基因往往不是一个物种生存所必须的,甚至是不利的,很容易在作物驯化

和育种的长河中被“遗失”。作物泛基因组分析发现,玉米自交系共同拥有的核心基因组(Core genome)只占62%,38%的序列则是非必须的

(Dispensable genome), 这些非必须序列和基因在品种间以存在/缺失变异 (PAV, presence/absence variation) 为主^[28], 类似的大量变异也广泛存在于水稻^[29]、小麦^[30-31]、大豆^[32]、大麦^[33]、向日葵^[34]等作物的不同品种中。因此, 用任何单一品种的基因组作为参考基因组进行新基因发掘, 都不可避免地遗落

了一些基因, 甚至大的基因组片段。为了更全面、更有效地发掘重要基因, 我国有必要整合和构建包含国际公认的模式品种、代表性地方品种及现代品种、甚至野生种的泛基因组, 作为基因发掘的新参考基因组, 发掘有重要的育种价值的基因, 服务于育种。

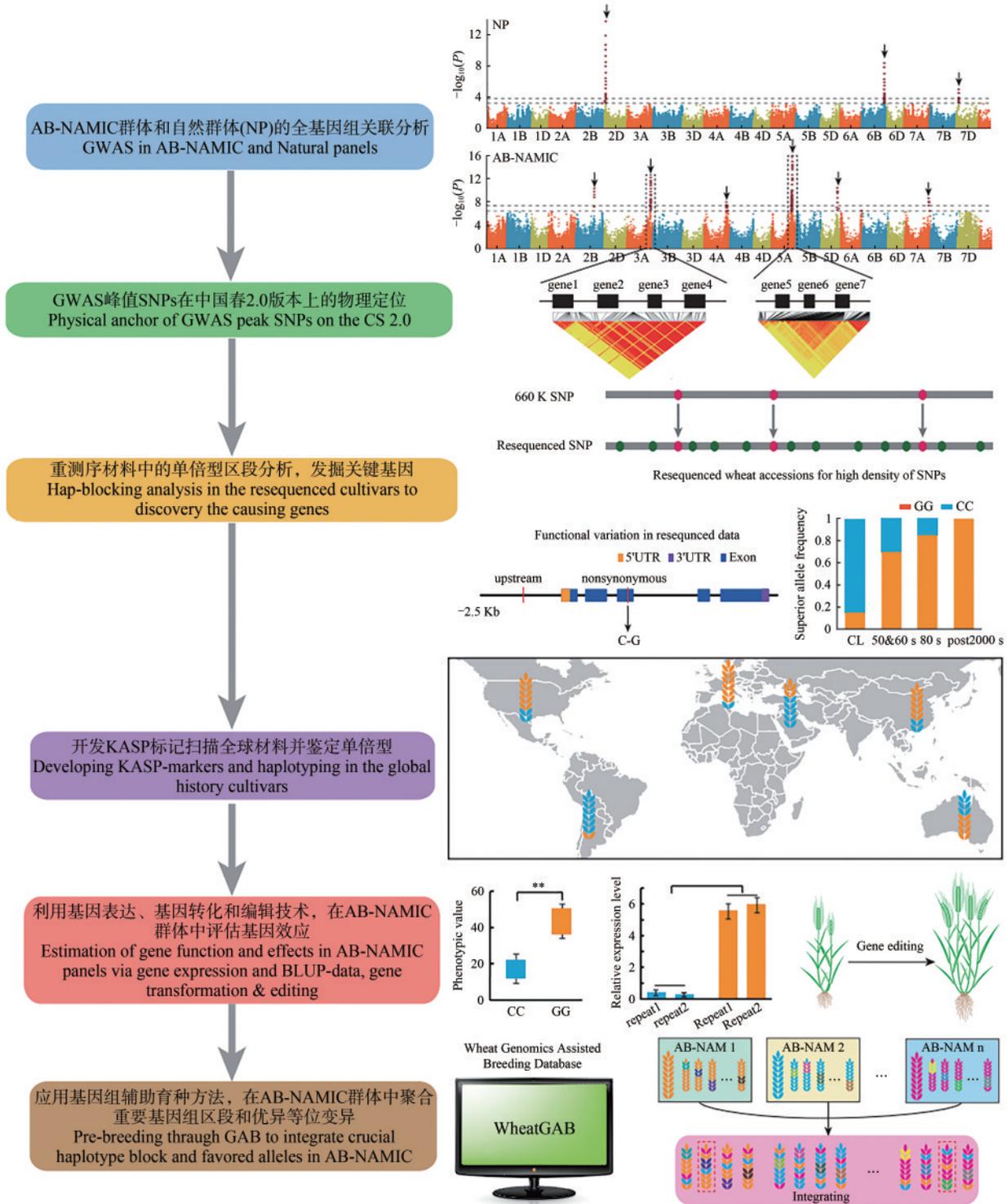
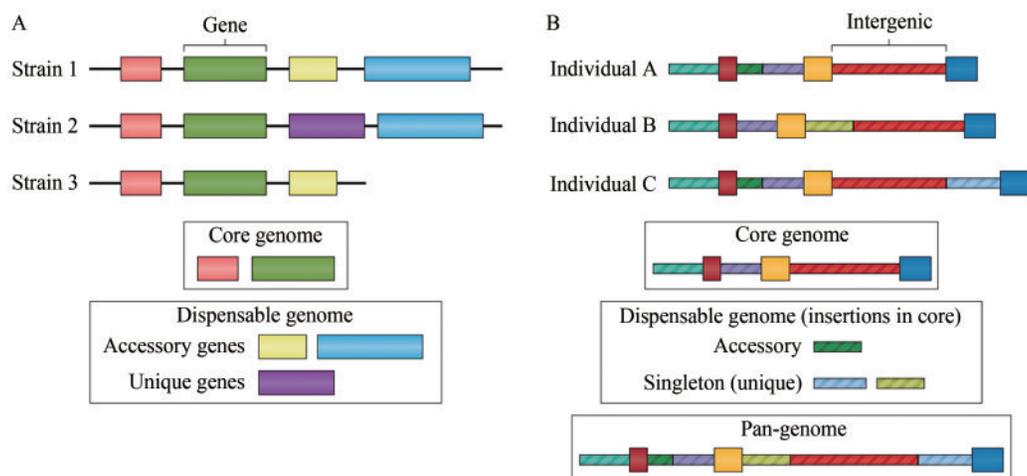


图7 基于新型人工群体发掘和利用库存种质资源的基本思路和流程^[24]

Fig.7 Basic ideas and processes for discovering and utilizing germplasm resources based on new multi-parents populations^[24]



A: 细菌和其他原核生物的基因组主要由基因组成,很少有基因间序列。核心基因组和可变基因组构成了泛基因组;B: 真核生物基因组的基因含量变化不大。泛基因组既考虑基因序列,也考虑基因间序列,形成所有序列的泛基因组^[27]

A: Bacterial and other prokaryotic genomes consist predominantly of genes, with little intergenic sequence. The core and dispensable genomes make up the pan-genome; B: Eukaryotic genomes are not highly variable in their genic content. Pan-genomes consider intergenic sequence as well as genes, resulting in an ordered pan-genome of all sequence present in at least one individual^[27]

图8 核心基因组和可变基因组

Fig.8 Core genome and dispensable genome in crop pan-genome

对作物育种和种质资源研究而言,泛基因组是继模式品种基因组完成后,基因组学与育种研究深度融合的必然,也是在新的层面深度解析遗传多样

性形成和品种演变的必要支撑,是未来10年应用基因组学和种质资源学的前沿热点^[32,35-36]。迄今为止已完成了超过40多项植物泛基因组的研究^[26,37](图9)。

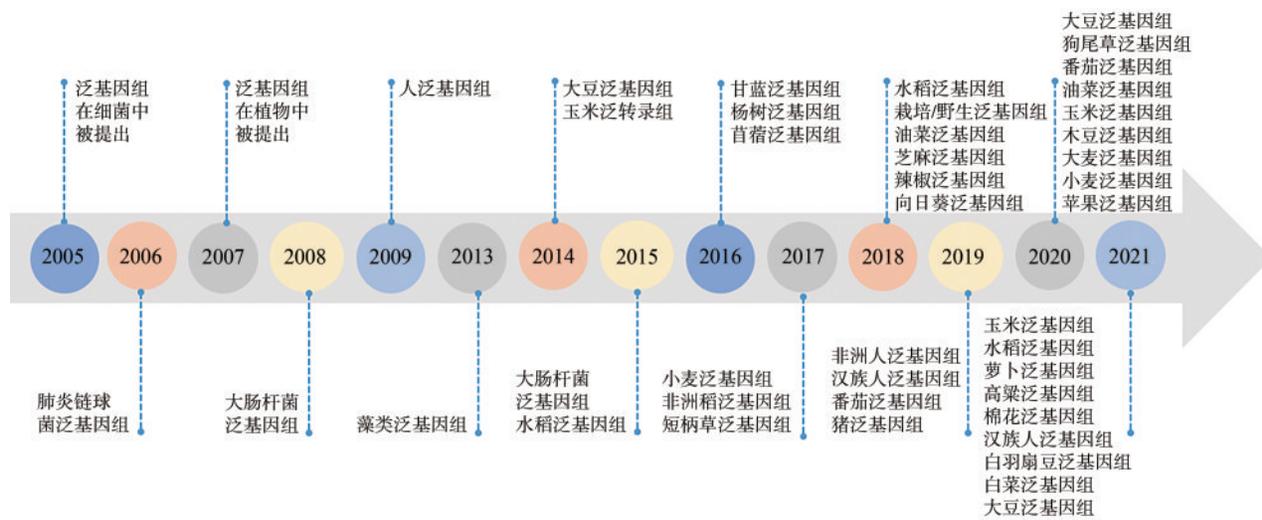


图9 泛基因组的研究进展

Fig.9 Research progress on pan-genome

5 基因编辑与基因发掘相结合,实现性状之定向突变和修饰

基因编辑是对生物基因组进行靶向修饰的一项新型生物技术,可以实现对目标基因的定点敲除、基因片段置换以及基因定点插入等基因定向编辑。基因编辑技术的开发及应用使得生物体的遗传改造达到前所未有的深度与广度,将在大基因组

作物基因的功能解析中发挥重要作用。目前基因编辑技术已与作物基因发掘相结合,在作物基因功能解析和遗传改良研究中得到广泛应用。中科院遗传与发育生物学研究所储成才研究组对全球110份早期水稻农家种进行了全面的农艺性状鉴定,发现不同氮肥条件下,水稻分蘖(分枝)氮响应能力与氮肥利用效率变异间存在高度关联。研究组利用全基因组关联分析技术鉴定到一个水稻氮高效基因

OsTCP19, 其作为转录因子调控水稻分蘖。进一步研究发现, *OsTCP19* 上游调控区一个小片段(29 bp) 的缺失与否是不同水稻品种分蘖氮响应差异的主要原因, 在氮高效品种中 *OsTCP19* 调控区缺失该 29 bp 核酸序列, 通过基因编辑技术对 *OsTCP19* 调控区该 29 bp 核酸序列进行靶向敲除, 实现了水稻品种分蘖氮响应能力的提升, 进而提高氮素利用效率^[23]。此外, 中科院遗传与发育生物学研究所李家洋团队通过组装异源四倍体高秆野生稻 (*Oryza alta*) 基因组, 优化遗传转化体系, 应用基因组编辑技术, 使其落粒性、芒性、株型、籽粒大小及抽穗期等决定作物驯化成功与否的重要性状发生改变, 实现了异源四倍体高秆野生稻的从头定向驯化, 为创制新作物提供了可能^[38]。

中国农业科学院作物科学研究所夏兰琴和马有志团队利用 CRISPR/Cas9 介导的多基因编辑技术, 以穗发芽抗性相关 (*TaQsd1*)、氮吸收利用 (*TaARE1*)、株型 (*TaNPT1*、*TaIPA1*)、支链淀粉合成 (*TaSBEIIa*) 和磷转运 (*TaSPDT*) 6 个基因作为靶基因, 在冬小麦品种郑麦 7698 中实现了同时靶向 2 个、3 个、4 个和 5 个基因的定点敲除编辑, 一代实现了多个优异等位基因聚合, 并通过胚拯救和分离后代筛选, 成功创制了无转基因、聚合多个优异等位基因的小麦新种质, 为小麦和其他多倍体农作物开展多基因聚合育种提供了重要的技术支撑^[39]。

6 以重大品种基因组为参考, 实现基因组、基因的设计和组装育种

优良基因组架构是重要育种亲本形成的基础, 一些重大有影响力的品种, 其基因组组成的独特性应是未来十年作物基因组学研究的重点。可在更深层次揭示骨干亲本、杂种优势群、一般配合力、特殊配合力等形成的基因组学和遗传学基础, 推动种质资源和育种研究迈上一个新台阶。

作物泛基因组研究揭示了同一物种不同品种基因组间存在的巨大差异, 使我们意识到骨干亲本间基因组结构上可能存在的差异, 这些差异不是永恒的, 会随着品种间的杂交、育种家的选择发生变化和调整, 即骨干亲本不是永恒的, 新的骨干亲本可能拥有上一代骨干亲本不具备的特性, 因此一定是与时俱进的。因此, 研究一些最新骨干亲本基因组组成特点, 在此基础上实现重要区段的替换 (Haplotype-based breeding)、优良基因的整合应是基因组学时代种质资源创新和育种的重要内

容^[40-41]; 而重要区段优良单倍型的发现和标记与快速育种技术的结合, 将加快新品种选育的速度, 在较短周期内创制一批突破性新材料^[42]。

参考文献

- [1] 卢新雄, 辛霞, 尹广鹏, 张金梅, 陈晓玲, 王述民, 方涛, 何娟娟. 中国作物种质资源安全保存理论与实践. 植物遗传资源学报, 2019, 20 (1): 1-10
Lu X X, Xin X, Yin G K, Zhang J M, Chen X L, Wang S M, Fang W, He J J. Theory and practice of the safe conservation of crop germplasm resources in China. Journal of Plant Genetic Resources, 2019, 20 (1): 1-10
- [2] 董玉琛. 作物种质资源学科的发展和展望. 中国工程科学, 2001, 3 (1): 1-5
Dong Y C. The development and prospect of crop germplasm science. Engineering Science, 2001, 3 (1): 1-5
- [3] Tanksley S D, McCouch S R. Seed banks and molecular maps: Unlocking genetic potential from the wild. Science, 1997, 277: 1063-1066
- [4] Mascher M, Schreiber M, Scholz U, Graner A, Reif J C, Stein N. Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding. Nature Genetics, 2019, 51: 1076-1081
- [5] Frankel O H, Brown A H D. Plant genetic resources today: A critical appraisal//Holden J H W, Williams J T. Crop genetic resources: Conservation & evaluation. London: George Allen & Urwin Ltd, 1984: 249-257
- [6] 李自超, 张洪亮, 曹永生, 裘宗恩, 魏兴华, 汤圣祥, 余萍, 王象坤. 中国地方稻种资源初级核心种质取样策略研究. 作物学报, 2003, 29 (1): 20-24
Li Z C, Zhang H L, Cao Y S, Qiu Z E, Wei X H, Tang S X, Yu P, Wang X K. Studies on the sampling strategy for primary core collection of Chinese indigenous rice. Acta Agronomica Sinica, 2003, 29 (1): 20-24
- [7] 郝晨阳, 董玉琛, 王兰芬, 游光霞, 张洪娜, 盖红梅, 贾继增, 张学勇. 我国普通小麦核心种质的构建及遗传多样性分析. 科学通报, 2008, 59 (8): 908-915
Hao C Y, Dong Y C, Wang L F, You G X, Zhang H N, Ge H M, Jia J Z, Zhang X Y. Genetic diversity and construction of core collection in Chinese wheat genetic resources. Chinese Science Bulletin, 2008, 59 (8): 908-915
- [8] 邱丽娟, 李英慧, 关荣霞, 刘章雄, 王丽侠, 常汝镇. 大豆核心种质和微核心种质的构建、验证与研究进展. 作物学报, 2009, 35 (4): 571-579
Qiu L J, Li Y H, Guan R X, Liu Z X, Wang L X, Chang R Z. Establishment, representative testing and research progress of soybean core collection and mini core collection. Acta Agronomica Sinica, 2009, 35 (4): 571-579
- [9] 贾继增, 高丽锋, 赵光耀, 周文斌, 张卫健. 作物基因组学与作物科学革命. 中国农业科学, 2015, 48 (17): 3316-3332
Jia J Z, Gao L F, Zhao G Y, Zhou W B, Zhang W J. Crop

- genomics and crop science revolutions. *Scientia Agricultura Sinica*, 2015, 48 (17): 3316-3332
- [10] Varshney R K, Bohra A, Yu J M, Graner A, Zhang Q F, Sorrells M E. Design future crops: Genomics-assisted breeding comes of age. *Trends in Plant Science*, 2021, 26: 632-648
- [11] Balfourier F, Bouchet S, Robert S, De Oliveira R, Rimbart H, Kitt J, Choulet F, International Wheat Genome Sequencing Consortium, BreedWheat Consortium, Paux E. Worldwide phylogeography and history of wheat genetic diversity. *Science Advances*, 2019, 5: eaav0536
- [12] Li A L, Hao C Y, Wang Z Y, Geng S F, Jia M L, Wang F, Han X, Kong X C, Yin L J, Tao S, Deng Z Y, Liao R Y, Sun G L, Wang K, Ye X G, Jiao C Z, Lu H F, Zhou Y, Liu D C, Fu X D, Zhang X Y, Mao L. Wheat breeding history reveals synergistic selection of pleiotropic genomic sites for plant architecture and grain yield. *Molecular Plant*, 2022, 15: 504-519
- [13] McMullen M D, Kresovich S, Villeda H S, Bradbury P, Li H H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell S E, Peterson B, Pressoir G, Romero S, Rosas M O, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz J C, Goodman M, Ware D, Holland J B, Buckler E S. Genetic properties of the maize nested association mapping population. *Science*, 2009, 325: 737-740
- [14] Nice L M, Steffenson B J, Brown-Guedira G L, Akhunov E D, Liu C, Kono T J, Morrell P L, Blake T K, Horsley R D, Smith K P, Muehlbauer G J. Development and genetic characterization of an advanced backcross-nested association mapping (AB-NAM) population of wild \times cultivated barley. *Genetics*, 2016, 203: 1453-1467
- [15] Huang B E, George A W, Forrest K L, Kilian A, Hayden M J, Morell M K, Cavanagh C R. A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnology Journal*, 2012, 10: 826-839
- [16] Gardner K A, Wittern L M, Mackay I J. A highly recombined, high-density, eight-founder wheat MAGIC map reveals extensive segregation distortion and genomic locations of introgression segments. *Plant Biotechnology Journal*, 2016, 14: 1406-1417
- [17] Liu H J, Wang X Q, Xiao Y J, Luo J Y, Qiao F, Yang W Y, Zhang R Y, Meng Y J, Sun J M, Yan S J, Peng Y, Niu L Y, Jian L M, Song W, Yan J L, Li C H, Zhao Y X, Liu Y, Warburton M L, Zhao J R, Yan J B. CUBIC: An atlas of genetic architecture promises directed maize improvement. *Genome Biology*, 2020, 21 (1): 20
- [18] Scott M F, Ladejobi O, Amer S, Bentley A R, Biernaskie J, Boden S A, Clark M, Dell'Acqua M, Dixon L E, Filippi C V, Fradgley N, Gardner K A, Mackay I J, O'Sullivan D, Percival-Alwyn L, Roorkiwal M, Singh R K, Thudi M, Varshney R K, Venturini L, Whan A, Cockram J, Mott R. Multi-parent populations in crops: A toolbox integrating genomics and genetic mapping with breeding. *Heredity*, 2020, 125: 396-416
- [19] Zuo J R, Li J Y. Molecular dissection of complex agronomic traits of rice: A team effort by Chinese scientists in recent years. *National Science Review*, 2014, 1: 253-276
- [20] Mao H D, Jian C, Cheng X X, Chen B, Mei F M, Li F F, Zhang Y F, Li S M, Du L Y, Li T, Hao C Y, Wang X J, Zhang X Y, Kang Z S. The wheat ABA receptor gene *TaPYL1-1B* contributes to drought tolerance and grain yield by increasing water-use efficiency. *Plant Biotechnology Journal*, 2022, 20: 846-861
- [21] Lin T, Zhu G T, Zhang J H, Xu X Y, Yu Q H, Zheng Z, Zhang Z H, Lun Y Y, Li S, Wang X X, Huang Z J, Li J M, Zhang C Z, Wang T T, Zhang Y Y, Wang A X, Zhang Y C, Lin K, Li C Y, Xiong G S, Xue Y B, Mazzucato A, Causse M, Fei Z J, Giovannoni J J, Chetelat R T, Zamir D, Städler T, Li J F, Ye Z B, Du Y C, Huang S W. Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics*, 2014, 46: 1220-1226
- [22] Tomato Genome Consortium. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 2012, 485: 635-641
- [23] Liu Y Q, Wang H R, Jiang Z M, Wang W, Xu R N, Wang Q H, Zhang Z H, Li A F, Liang Y, Ou S J, Liu X J, Cao S Y, Tong H N, Wang Y H, Zhou F, Liao H, Hu B, Chu C C. Genomic basis of geographical adaptation to soil nitrogen in rice. *Nature*, 2021, 590: 600-605
- [24] Jiao C Z, Hao C Y, Li T, Bohra A, Wang L F, Hou J, Liu H X, Liu H, Zhao J, Wang Y M, Liu Y C, Wang Z W, Jing X, Wang X E, Varshney R K, Fu J J, Zhang X Y. Fast integration and accumulation of breeding beneficial alleles through AB-NAMIC strategy in wheat. *Plant Communications*, 2023, on line
- [25] Tettelin H, Massignani V, Cieslewicz M J, Donati C, Medini D, Ward N L, Angiuoli S V, Crabtree J, Jones A L, Durkin A S, DeBoy R T, Davidsen T M, Mora M, Scarselli M, Margarit y Ros I, Peterson J D, Hauser C R, Sundaram J P, Nelson W C, Madupu R, Brinkac L M, Dodson R J, Rosovitz M J, Sullivan S A, Daugherty S C, Haft D H, Selengut J, Gwinn M L, Zhou L W, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor K J B, Smith S, Utterback T R, White O, Rubens C E, Grandi G, Madoff L C, Kasper D L, Telford J L, Wessels M R, Rappuoli R, Fraser C M. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proceedings of the National Academy of Sciences*, 2005, 102: 13950-13955
- [26] Bayer P E, Golicz A A, Scheben A, Batley J, Edwards D. Plant pan-genomes are the new reference. *Nature Plants*, 2020, 6: 914-920
- [27] Sherman R M, Salzberg S L. Pan-genomics in the human genome era. *Nature Reviews Genetics*, 2020, 21: 243-254
- [28] Hirsch C N, Foerster J M, Johnson J M, Sekhon R S, Muttoni G, Vaillancourt B, Peñagaricano F, Lindquist E, Pedraza M A, Barry K, de Leon N, Kaeppler S M, Buell C R.

- Insights into the maize pan-genome and pan-transcriptome. *Plant Cell*, 2014, 26: 121-135
- [29] Qin P, Lu H W, Du H L, Wang H, Chen W L, Chen Z, He Q, Ou S J, Zhang H Y, Li X Z, Li X X, Li Y, Liao Y, Gao Q, Tu B, Yuan H, Ma B T, Wang Y P, Qian Y W, Fan S J, Li W T, Wang J, He M, Yin J J, Li T, Jiang N, Chen X W, Liang C Z, Li S G. Pan-genome analysis of 33 genetically diverse rice accessions reveals hidden genomic variations. *Cell*, 2021, 184: 1-17
- [30] Montenegro J D, Golicz A A, Bayer P E, Hurgobin B, Lee H, Chan C K, Visendi P, Lai K, Doležel J, Batley J, Edwards D. The pangenome of hexaploid bread wheat. *The Plant Journal*, 2017, 90: 1007-1013
- [31] Walkowiak S, Gao L L, Monat C, Haberer G, Kassa M T, Brinton J, Ramirez-Gonzalez R H, Kolodziej M C, Delorean E, Thambugala D, Klymiuk V, Byrns B, Gundlach H, Bandi V, Siri J N, Nilsen K, Aquino C, Himmelbach A, Copetti D, Ban T, Venturini L, Bevan M, Clavijo B, Koo D H, Ens J, Wiebe K, N'Diaye A, Fritz A K, Gutwin C, Fiebig A, Fosker C, Fu B X, Accinelli G G, Gardner K A, Fradgley N, Gutierrez-Gonzalez J, Halstead-Nussloch G H, Hatakeyama M, Koh C S, Deek J, Costamagna A C, Fobert P, Heavens D, Kanamori H, Kawaura K, Kobayashi F, Krasileva K, Kuo T, McKenzie N, Murata K, Nabeka Y, Paape T, Padmarasu S, Percival-Alwyn L, Kagale S, Scholz U, Sese J, Juliana P, Singh R, Shimizu-Inatsugi R, Swarbreck D, Cockram J, Budak H, Tameshige T, Tanaka T, Tsuji H, Wright J, Wu J Z, Steuernage B, Small I, Cloutier S, Keeble-Gagnère G, Muehlbauer G, Tibbets J, Nasuda S, Melonek J, Hucl P J, Sharpe A G, Clark M, Legg E, Bharti A, Langridge P, Hall A, Uauy C, Mascher M, Krattinger S G, Handa H, Shimizu K K, Distelfeld A, Chalmers K, Keller B, Mayer K F X, Poland J, Stein N, McCartney C A, Spannag M, Wicker T, Pozniak C J. Multiple wheat genomes reveal global variation in modern breeding. *Nature*, 2020, 588: 277-283
- [32] Liu Y C, Du H L, Li P C, Shen Y T, Peng H, Liu S L, Zhou G A, Zhang H K, Liu Z, Shi M, Huang X H, Li Y, Zhang M, Wang Z, Zhu B G, Han B, Liang C Z, Tian Z X. Pan-Genome of wild and cultivated soybeans. *Cell*, 2020, 182: 162-176
- [33] Jayakodi M, Padmarasu S, Haberer G, Bonthala V S, Gundlach H, Monat C, Lux T, Kamal N, Lang D, Himmelbach A, Ens J, Zhang X Q, Angessa T T, Zhou G F, Tan C, Hill C, Wang P H, Schreiber M, Boston L B, Plott C, Jenkins J, Guo Y, Fiebig A, Budak H, Xu D D, Zhang J, Wang C C, Grimwood J, Schmutz J, Guo G G, Zhang G P, Mochida K, Hirayama T, Sato K, Chalmers K J, Langridge P, Waugh R, Pozniak C J, Scholz U, Mayer K F X, Spannag M, Li C D, Mascher M, Stein N. The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature*, 2020, 588: 284-289
- [34] Hübner S, Bercovich N, Todesco M, Mandel J R, Odenheimer J, Ziegler E, Lee J S, Baute G J, Owens G L, Grassa C J, Ebert D P, Ostevik K L, Moyers B T, Yakimowski S, Masalia R R, Gao L, Čalić I, Bowers J E, Kane N C, Swanevelder D Z H, Kubach T, Muñoz S, Langlade N B, Burke J M, Rieseberg L H. Sunflower pan-genome analysis shows that hybridization altered gene content and disease resistance. *Nature Plants*, 2019, 5: 54-62
- [35] Li Y H, Zhou G, Ma J, Jiang W, Jin L G, Zhang Z, Guo Y, Zhang J, Sui Y, Zheng L, Zhang S S, Zuo Q, Shi X H, Li Y F, Zhang W K, Hu Y, Kong G, Hong H L, Tan B, Song J, Liu Z X, Wang Y, Ruan H, Yeung C K, Liu J, Wang H, Zhang L J, Guan R X, Wang K J, Li W B, Chen S Y, Chang R Z, Jiang Z, Jackson S A, Li R, Qiu L J. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits. *Nature Biotechnology*, 2014, 32: 1045-1052
- [36] Lei L, Goltsman E, Goodstein D, Wu G A, Rokhsar D S, Vogel J P. Plant pan-genomics comes of age. *Annual Review of Plant Biology*, 2021, 72: 411-435
- [37] Golicz A A, Bayer P E, Bhalla P L, Batley J, Edwards D. Pangenomics comes of age: From bacteria to plant and animal applications. *Trends in Genetics*, 2020, 36 (2): 132-145
- [38] Yu H, Lin T, Meng X B, Du H L, Zhang J K, Liu G F, Chen M J, Jing Y H, Kou L Q, Li X X, Gao Q, Liang Y, Liu X D, Fan Z L, Liang Y T, Cheng Z K, Chen M S, Tian Z X, Wang Y H, Chu C C, Zuo J R, Wan J M, Qian Q, Han B, Zuccolo A, Wing R A, Gao C X, Liang C Z, Li J Y. A route to *de novo* domestication of wild allotetraploid rice. *Cell*, 2021, 184 (5): 1156-1170
- [39] Luo J M, Li S Y, Xu J J, Yan L, Ma Y Z, Xia L Q. Pyramiding favorable alleles in an elite wheat variety in one generation by CRISPR-Cas9-mediated multiplex gene editing. *Molecular Plant*, 2021, 14: 847-850
- [40] 张学勇, 马琳, 郑军. 作物驯化和品种改良所选择的关键基因及其特点. *作物学报*, 2017, 43 (2): 157-170
- Zhang X Y, Ma L, Zheng J. Characteristics of genes selected by domestication and intensive breeding in crop plants. *Acta Agronomica Sinica*, 2017, 43 (2): 157-170
- [41] Hao C Y, Jiao C Z, Hou J, Li T, Liu H X, Wang Y Q, Zheng J, Liu H, Bi Z H, Xu F F, Zhao J, Ma L, Wang Y M, Majeed U, Liu X, Appels R, Maccaferri M, Tuberosa R, Lu H F, Zhang X Y. Resequencing of 145 landmark cultivars reveals asymmetric sub-genome selection and strong founder genotype effects on wheat breeding in China. *Molecular Plant*, 2020, 13: 1733-1751
- [42] Brinton J, Ramirez-Gonzalez R H, Simmonds J, Wingen L, Orford S, Griffiths S; 10 Wheat Genome Project, Haberer G, Spannagl M, Walkowiak S, Pozniak C, Uauy C. A haplotype-led approach to increase the precision of wheat breeding. *Communications Biology*, 2020, 3: 712