



作物驯化规律解析与未来作物设计实践

马新¹, 张敏¹, 王建林², 刘羽诚^{1*}

1. 中国科学院遗传与发育生物学研究所, 种子创新全国重点实验室, 北京 100101

2. 中科东营分子设计育种研究中心, 东营 257509

* 联系人, E-mail: ychliu@genetics.ac.cn

收稿日期: 2025-06-03; 接受日期: 2025-10-14; 网络版发表日期: 2026-04-09

摘要 作物驯化是人类文明发展的重要里程碑事件, 驯化过程中的基因组演变规律潜藏着作物驯化改良的经验启示. 由于异常气候频发及人口急剧增长对粮食安全构成重大挑战, 可利用作物的扩展与作物改良是应对这一挑战的重要策略. 本文系统梳理了作物驯化的基本表型变化特征及基因组变化模式, 针对驯化综合征相关表型的遗传基础进行总结; 概述作物驯化研究的基因组研究思路与方法, 并对从头驯化的育种策略进行总结与展望. 总体而言, 本文综述了作物驯化的基因组规律解析与应用相关研究, 以期对未来作物设计育种提供重要的理论参考.

关键词 作物, 驯化综合征, 基因组学, 从头驯化, 育种

人类作为独立的分类单元存在于地球上已有近 300 万年历史. 文明的起源与发展, 包含诸多复杂的历史事件及偶然性, 在自然和人文科学层面都留下大量值得探讨研究的学术问题. 伴随着更新世向全新世的全球气候过渡, 人类在距今 1 万年左右由旧石器时代进入新石器时代, 而人类农耕出现的时间与这一过渡时期吻合^[1]. 生物资源的驯化, 是依照人类需求定向改造生物的过程, 其本质是一种人工选择驱动的生物演化. 人类的驯化深刻地改造与自身相关动植物的遗传背景, 动植物的驯化过程也反映出人类农耕文明的发展变迁^[2].

作物驯化过程, 可以总结为三个阶段^[3]: 前驯化阶段、驯化阶段和改良阶段. 前驯化阶段始于人类定点种植植物的行为. 种植行为克服采集的风险和收益的不确定性, 使得收获更加安全稳定. 在此阶段人类积

累大量可供驯化的植物资源. 在植物的人工驯化过程中, 人类未必主观意识到选择与繁种的作用, 但是在人工种植条件下相对具有高收获指数、易栽植、低毒性等特征的植物个体有更高的适合度, 这使得遗传选择发挥效力, 也造成栽培作物相对于野生祖先显著的遗传多态性丢失. 经过长年的农耕生产, 人类开始理解并有意识地进行作物改良, 开始应用定向选择、无性繁殖和杂交等手段. 改良事件可能多地多次发生, 主要是基于驯化后作物的基本特征并且根据需求在一个或若干性状上进行改良, 例如产量、风味、口感等方面. 由于剧烈的瓶颈效应已经结束, 改良后的作物遗传多态性并没有显著下降, 这在一些作物的研究中得到证明^[4,5]. 改良阶段仍可能保持着高水平的正选择, 但强选择效力可能仅作用于小范围的遗传背景. 因此, 尽管改良后的种系间可能表型上存在明显差异,

引用格式: 马新, 张敏, 王建林, 等. 作物驯化规律解析与未来作物设计实践. 中国科学: 生命科学, 2026, 56: 1001–1014

Ma X, Zhang M, Wang J L, et al. Deciphering of domestication principles enlightens future crop breeding (in Chinese). Sci Sin Vitae, 2026, 56: 1001–1014, doi: [10.1360/SSV-2025-0178](https://doi.org/10.1360/SSV-2025-0178)

但是仍可能具有相似的遗传背景。

目前人类已驯化的植物有200余种^[6]。尽管对不同作物利用的方向、器官各不相同, 驯化进程也有所差异^[3], 但总体而言, 驯化后的作物往往出现一系列相似的特征, 称为“驯化综合征”^[7](图1和表1)。在形态上, 食用作物总体有果实/种子增大的特点, 这些器官生物量的增加, 是作物驯化的直接动力; 顶端优势增强, 株型直立紧凑也是作物驯化的重要方向, 保证其适合农田系统的栽培管理及收获; 落粒性的丢失尽管与植物在自然界中种子扩散的繁衍需求相悖, 但对于农业生产至关重要。落粒性的丢失降低收获器官的损失, 奠定“耕作-收获”的行为闭环。在生理生化上, 则有种子休眠丧失, 光周期适应性变化、花期同步等驯化趋同特征, 这些特征使作物在农业生产中的种植、生长、收获过程更为同步, 适宜栽培管理。此外, 还具有有毒物质与苦味丧失等特点, 这些确保食用作物的安全与口感。总体而言, 驯化综合征是一系列满足农业生产和收获需求的表型变化。这些变化在将来可能会随着生产条件、生产模式及需求的变化而发生改变^[8]。这也是研究者思考未来如何改造作物, 驯化新作物的重要切入点。

驯化综合征是作物的整体改变趋势, 但是不同性状在作物中的发展可能截然不同。在器官大小和落粒性方面, 人类的选择倾向基本是一致的。而光周期适应性在植物中的驯化状态可能并不一致。部分作物经过驯化形成较强的光周期广适性, 有的作物则广适性

弱, 只具有明显的局部适应性。由于原产地和栽培地点不同, 不同植物具体的适应历程可能存在差异^[9]。尽管栽培作物种子休眠期普遍短于野生种, 在一些谷类作物中, 休眠期缩短造成的穗发芽现在反而成为全球农业生产问题, 造成经济损失。这使得人工改良在这些作物中对种子休眠进行反向调和^[10,11]。同时研究者需要认知到, 在超过2000余种人类利用的植物中, 仅仅一小部分受到完整的驯化^[12]。不同植物由于自身特征, 以及变异多样性, 遗传网络关系, 也表现出不同的驯化程度, 这也表明未来作物驯化存在诸多挑战。

作物驯化是农业的起点, 一直以来是学界研究的重点领域^[12~16]。“驯化综合征”相关的遗传变异经过人类长时间利用和筛选, 一定程度上潜藏着作物改造的“参考路径”^[7,17,18]。因此, 研究驯化的基因组演化机制, 解析驯化关键基因和变异, 是深入理解作物基因功能, 挖掘种质资源遗传价值, 评估作物改造潜力的重要途径, 能够为作物的遗传改良提供理论指导。

由于人工驯化在时间及性状尺度的复杂性, 解析作物驯化分子遗传规律是极具挑战性的工作。作物驯化是否可以重现, 是否由关键基因主导, 需要涉及多少基因的改造, 选择变异的类型有何种偏好, 这些基因间的关联是什么, 是领域研究的关键问题^[12,15,19,20]。鉴定作物人工驯化的核心基因组成分, 精准挖掘驯化的关键目标变异, 评估相关基因在遗传功能网络中的位置及育种应用价值, 能够为作物的遗传改良及从头驯化提供更新的理论指导和基因资源。

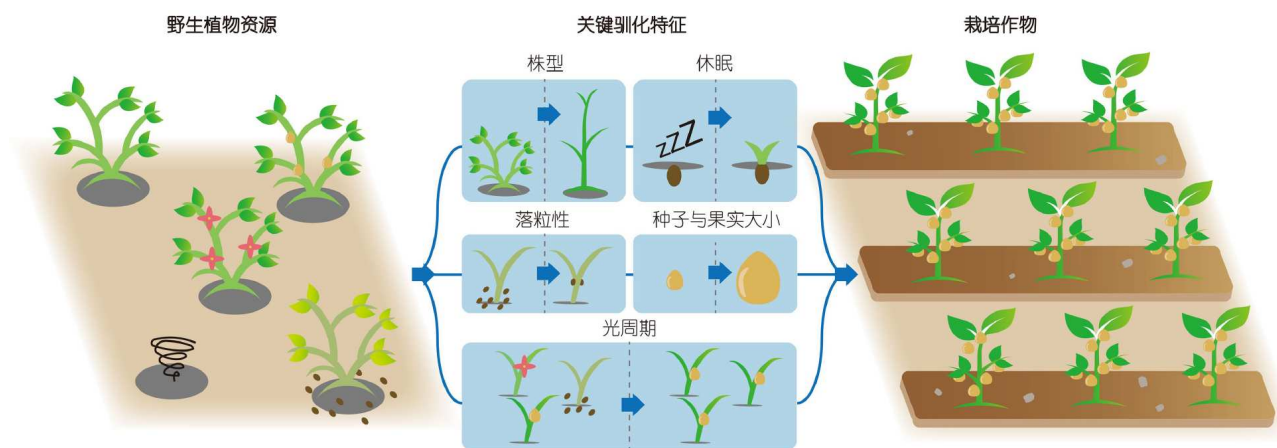


图1 作物驯化的实践过程及重要性状演变规律

Figure 1 The practices and evolutionary patterns of key traits in crop domestication

表 1 作物关键驯化性状相关基因及功能变异

Table 1 Genes and functional variations associated with key domestication traits in crops

作物	基因/QTL	变异区域	变异类型	引起的变化	表型	参考文献
水稻	<i>RPAD</i>	基因所在的基因组区域	结构变异(约110 kb的DNA片段缺失)	基因丢失	株型	[27]
水稻	<i>OsSH3</i>	基因所在的基因组区域	结构变异(45 kb DNA片段缺失)	基因丢失	落粒性	[29]
水稻	<i>GIF1</i>	上游调控区域	核苷酸变异	可能影响基因表达	果实/种子大小	[39]
水稻	<i>GAD1</i>	编码区	插入缺失	移码突变	芒、果实/种子大小	[40,41]
水稻	<i>LABA1/An-2</i>	编码区	单核苷酸变异	移码突变、提前终止	芒、果实/种子大小	[42,43]
水稻	<i>An-1</i>	启动子区域、编码区	单核苷酸变异、插入缺失	可能影响基因表达、移码突变、提前终止	芒、果实/种子大小	[44]
水稻	<i>PROG1</i>	上游调控区域、编码区	单核苷酸变异、插入缺失	影响基因表达、非同义突变	株型	[46,48]
水稻	<i>qSH1</i>	上游调控区域	单核苷酸变异	可能影响基因表达	落粒性	[53]
水稻	<i>SH4/SHA1</i>	编码区	单核苷酸变异	非同义突变	落粒性	[55,56]
水稻	<i>OsSh1</i>	—	—	—	落粒性	[57]
水稻	<i>Rc</i>	编码区	插入缺失(14 bp DNA片段)	提前终止	果皮颜色、种子休眠	[82]
玉米	<i>zag1l</i>	—	—	—	光周期适应	[23]
玉米	<i>tb1</i>	上游60 kb	插入缺失(Hopscotch转座子)	影响基因表达	株型	[25,26]
玉米	<i>Bt2</i>	内含子	插入缺失	影响基因表达	果实/种子大小	[45]
玉米	<i>PRR7</i>	—	—	—	光周期适应	[65]
玉米	<i>ZmCCT9</i>	基因上游57 kb处	结构变异(Harbinger-like transposable element)	影响基因表达	光周期适应	[71]
玉米	<i>ZmCCT</i>	启动子区域	结构变异 (CACTA-like transposable element)	影响基因表达	光周期适应	[72]
玉米	<i>ZmMADS69</i>	—	—	—	光周期适应	[74]
大豆	<i>SHAT1-5</i>	启动子	插入缺失	影响基因表达	落粒性	[24]
大豆	<i>Dt1</i>	编码区	单核苷酸变异	非同义突变	株型	[50,51]
大豆	<i>Phd1</i>	编码区	单核苷酸变异	提前终止	落粒性	[59]
大豆	<i>Sh1</i>	编码区	单核苷酸变异、插入缺失	氨基酸缺失插入、非同义突变	落粒性	[59]
大豆	<i>E5-E11</i>	—	—	—	光周期适应	[75]
大豆	<i>E1</i>	编码区、基因所在的基因组区域	单核苷酸变异、结构变异	非同义突变、提前终止; 基因丢失	光周期适应	[76]
大豆	<i>E2</i>	编码区	单核苷酸变异	提前终止	光周期适应	[78,79]
大豆	<i>E3</i>	编码区	单核苷酸变异、结构变异	非同义突变、基因部分丢失	光周期适应	[80]
大豆	<i>E4</i>	编码区	结构变异	基因功能紊乱	光周期适应	[81]
大豆	<i>G</i>	编码区	单核苷酸变异	提前终止	种子休眠	[84]
大豆	<i>GmHSL-1</i>	编码区	单核苷酸变异	非同义突变	种子休眠	[85]
番茄	<i>fw1.1</i>	—	—	—	果实/种子大小	[5]
番茄	<i>fw12.1</i>	—	—	—	果实/种子大小	[5]
番茄	<i>fw5.2</i>	—	—	—	果实/种子大小	[5]
番茄	<i>fw7.2</i>	—	—	—	果实/种子大小	[5]
番茄	<i>lcn12.1</i>	—	—	—	果实/种子大小	[5]
番茄	<i>SP</i>	编码区	单核苷酸变异	非同义突变	株型	[22]
番茄	<i>FTL1</i>	编码区	插入缺失(2 bp)	提前终止	光周期适应	[68]
番茄	<i>SP5G</i>	上游调控区域	—	—	光周期适应	[68,69]
番茄	<i>SFT</i>	编码区	单核苷酸变异、插入缺失(2 bp)	非同义突变、移码突变	光周期适应	[70]
高粱	<i>SbTB1</i>	上游调控区域、编码区	单核苷酸变异、插入缺失	可能影响基因表达、非同义突变	株型	[49]

1 作物驯化的遗传规律

1.1 人工选择的基因组演化规律

遗传变异是生物演化与作物人工驯化的基础。基因组变异——从单核苷酸变异到染色体结构变异; 从编码区与调控区域序列改变到整个基因缺失等等——在人类对作物的驯化过程中都扮演重要角色。单核苷酸变异看似最为简单, 但也能造成基因功能显著的改变。水稻中*COLD1*基因第4外显子上单核苷酸的复等位突变, 赋予不同等位基因间耐寒性差异, 塑造了栽培稻驯化过程中因分布区迁移导致的适应性分化^[21]; 番茄*SP*基因上的单核苷酸突变, 造成氨基酸序列76位变化, 从而形成紧凑株型并且在现代栽培番茄中被广泛使用^[22]; 玉米中*zag1*基因影响开花期和籽粒数量, 其第6外显子上的一个单核苷酸突变在大刍草及现代玉米中存在明显的分化, 被认为与驯化高度相关^[23]。序列变异发生在非编码序列, 则有可能通过调控基因表达造成性状变化。大豆*SHAT1-5*基因启动子区23 bp的序列插入缺失变异, 被证明对基因表达造成影响, 进而影响裂荚性状^[24]; 玉米*tb1*上游约60 kb处的Hops-cotch转座子插入与在栽培玉米中一段开放染色质区域重合, 并且造成甲基化水平差异^[25,26]。该变异引起基因表达改变, 造成株型变化^[25]。更复杂的染色体结构变异也在作物驯化起到重要作用。在亚洲栽培稻与非洲栽培稻的驯化过程中, 独立的基因组大片段缺失促进水稻由匍匐向直立生长的转变^[27,28]。非洲栽培稻中包含*OsSH3*在内的45 kb DNA片段缺失, 导致落粒性的丧失^[29]。大豆控制种子颜色的*I*位点上, 染色体倒位变异以及染色体双交换造成查尔斯合成酶*CHS*基因簇的加倍, 发生同源依赖的基因沉默(homology-dependent gene silencing, HDGS), 导致*CHS*以及下游黄酮代谢途径相关基因的沉默^[30]。在陆地棉中, A08染色体上出现的大于4 Mb的倒位事件显著地抑制染色体间的重组和多态性, 造成育种过程中棉花品种的显著分化^[31]。不同类型的变异在驯化中都能起到重要的作用, 在育种中的应用价值也需要具体衡量。

从分子遗传学的角度而言, 驯化的过程主要是依据植物的表型, 筛选出其中造成目标表型的功能变异。通常情况下, 自然界中野生植物包含有更丰富的遗传多样性, 以适应复杂的环境。而驯化过程往往在连续的世代间高强度地选择最优的少数植物个体, 导致基

因组范围内的遗传多样性的大规模下降, 称为遗传瓶颈效应^[7,32]。此外, 对于驯化相关目标变异的选择会导致其以及周围连锁区域的遗传多样性剧烈下降, 这类现象称为选择性清除^[7,32]。当周围固定的这类变异对于作物驯化存在不利效应却由于紧密连锁难以通过重组交换清除时, 则认为存在着连锁累赘或遗传负荷^[14,33]。人类对作物的遗传改造, 则是不断发掘与聚合有利变异, 同时减少不利变异的一个过程。然而自然界中产生的变异是有限的, 当选择过程无法满足人类的需求时, 人类会加速变异产生的效率。其中手段包括加快世代繁衍、加强筛选压力、主动诱导变异等, 以获得满足需求的作物性状。这通常发生在改良等有意识的人类作用阶段。而随着遗传理论及手段发展到一定程度, 杂交、诱变等手段才能够被有效地用于植物的定向改造。

1.2 关键驯化性状的遗传基础

(1) 种实大小。作物种子大小是关键性状, 其遗传机制也受到较广泛的研究。主要涉及激素信号传导、蛋白质泛素化、胚发育等生物学过程^[34~36]。在水稻中, 以往研究鉴定一系列与籽粒大小、粒长/粒宽/粒重相关的QTL, 其中部分被证实存在人工驯化^[37,38]。水稻驯化过程中*GIF1*调控区域上积累的变异使其呈现组织特异性表达, 有利于籽粒灌浆, 导致粒重增加^[39]。此外, 水稻中控制芒长的驯化基因*GAD1*, *An-1*及*An-2/LABAI*, 也能够造成种子大小的变化^[40~44]。在番茄中*fw2.2*, *fw1.1*, *fw5.2*, *fw7.2*, *fw12.1*和*lcn12.1*的驯化选择导致果实的增大^[5]。在玉米中, 控制籽粒灌浆的一些基因, 如*Bt2*等基因的驯化可能是籽粒大小变化的原因^[45]。作物种子大小属于数量性状, 受到众多基因的调控, 在驯化过程中并非少数基因发生作用。

(2) 株型。顶端优势形成造成植物株型的改变是作物中普遍发生的现象。研究表明这类表型通常存在主效的驯化位点。在水稻中, 编码一个锌指蛋白转录因子的*PROG1*以及邻近的锌指蛋白转录因子簇受到强烈人工选择, 使得水稻植株从匍匐变为直立^[27,28,46~48]。玉米中一个编码TCP转录因子的*tb1*, 其驯化导致玉米的叶夹角变小, 分蘖数减少, 顶端优势增强^[25]。同时, *SbTBI*也是高粱株型驯化的重要基因^[49]。大豆中株型相关位点*Dt1*, 在栽培大豆中存在若干不同突变位点^[50,51], 这些突变位点组成的单倍型均

造成大豆的有限生长习性, 且在驯化过程中受到选择^[50]. *Dt1*编码一类TFL蛋白, 位于光周期调控途径的重要节点. 其同源基因在绿豆与番茄中也受到驯化, 改变茎生长习性^[22,52].

(3) 落粒性. 在谷物中, 落粒性主要是小穗与枝梗连接处的离层控制^[53,54]. 水稻中*SH4/SHA1*, *qSH1*和*OsSh1*是报道与落粒性驯化相关的基因. 其中*SH4/SHA1*为未知功能转录因子^[55,56]; *qSH1*编码一个BEL1型同源异型蛋白^[53]; *OsSh1*编码一类YABBY转录因子, 其同源基因在玉米和谷子中也存在人工选择的迹象^[57]. 这些基因驯化后的变异造成离层的不连续或缺失, 增强种子的附着能力^[55,57]. 大豆的裂荚性并非简单由果荚腹缝线的离层细胞决定, 而是果荚及腹缝线组织机械特性共同影响. *SHAT1-5*启动子区的20 bp插入缺失造成表达量差异且受到人工选择, 导致腹缝线纤维冠细胞次生细胞壁增厚, 减弱了裂荚性^[24]. 此外, *Pdh1*影响大豆果荚中木质素分布的对称性, 导致果荚扭转^[58]. 其临近位点*Sh1*通过抑制*SHAT1-5*表达来影响裂荚性状. 这两个位点的协同驯化对大豆的裂荚起到重要的作用^[59]. 值得注意的是, *Phd1*在暖季型豆科植物的裂荚习性中也存在平行选择的适应性机制^[60].

(4) 光周期适应. 作物光周期适应性的驯化在不同作物中存在明显的相似和差异. *FT*与*PRR7*在许多作物中都报道受到驯化^[61~65], 是植物适应光周期变化的重要功能基因^[66,67]. 番茄由短日照植物变为日中性植物, 使广适性大幅上升, 主要为一系列*FT*基因, 包括*SP5G*, *FTL1*, *SFT*等位点的人工选择^[68~70]. 而玉米中, 许多编码CCT保守结构域和MADS-box转录因子的基因在驯化中发挥重要作用^[23,71~74]. 大豆中*E1~E11*等一系列位点为控制光周期的重要位点^[75,76], 其中如*E2* (*GIGANTEA*), *E3* (*phyB*)等位点在驯化过程中并未有一种主要类型的固定, 而是在不同纬度地带表现出基因型与地理区域的关联分化^[77~81]. 在表现出特异纬度适应性的同时, 也没有形成很好的广适性.

(5) 种子休眠. 种子休眠主要有生理休眠和机械休眠, 其中驯化的成因也各不一样. 在生理性休眠方面, 休眠主要与ABA, GA, IAA等激素的传递相关. 在水稻, 一个编码bHLH转录因子的驯化基因*Rc*控制果皮颜色, 同时亦控制ABA积累, 最终影响种子休眠^[82,83]. 大豆种子颜色相关基因*G*被发现与种子休眠相关^[84]. 该基因可能参与ABA的合成途径, 也被证实

在水稻、番茄中受到平行选择^[84]. 在机械休眠方面, 通常是种子休眠主要取决于外层包裹组织的机械强度, 以及相应的透气透水效率. 有研究表明大豆中对于*GmHSL-1*等位点的驯化, 造成种皮硬度的下降, 减弱休眠^[85]. 而向日葵在驯化过程中果皮木质化、纤维化组织孔径的增加及硬度的下降, 也可能是休眠打破的主要成因^[86]. 目前相关基因未见明确报道.

总结不同作物驯化综合征的研究结果, 可以发现趋同的性状背后存在着一定程度遗传基础的相似性. 这些相似性往往被归纳为变异位点、基因及代谢通路的平行选择^[87]. 目前关于作物平行选择变异位点的报道较少. 目前关于作物平行选择的基因, 大多由总结归纳不同作物领域各自的研究而得出^[88~90]. 以驯化的平行选择机制为直接驱动的研究则较少^[52,57,84,91]. 而代谢通路或功能网络水平的平行选择的定义, 不仅需要考虑基因的同源性, 同时涉及基因功能间的关联性. 有观点认为, 代谢途径的单一性, 基因在功能网络中关键性等特征, 能够增加基因在物种间发生平行选择的概率^[88]. 作物间驯化的共性规律潜在反应出对于作物的驯化改良而言哪些路径是可以参考的, 而哪些路径可能需要回避. 这些理论基础对于未来育种而言可能是重要的启示(图2).

2 作物驯化相关的基因组研究手段

2.1 人工选择的基因组特征鉴定

DNA测序技术的发展对基因组学领域起到极大的推进作用. 大数据的涌现不仅拓展遗传学、数学与统计学的交叉范畴, 而且促进它和计算机、信息技术的密切融合. 伴随着基因组学的数据规范日渐成熟, 处理算法蓬勃发展, 测序技术不断突破, 生物信息学等诸多交叉学科应运而生. 许多作物的基因组都获得基因组层面的遗传解析. 高通量测序技术极大程度地降低解析生物遗传信息的成本, 使大规模群体水平的全基因组DNA变异鉴定成为可能. 对作物驯化的研究也由性状的表型调查向分子遗传机制的探究深入(图2).

人工选择往往在野生和栽培群体的基因上留下特征痕迹, 这些特征可以通过群体水平的变异数据进行全基因组扫描进行检测. 人工选择基因组特征的检测主要有以下三个方面: 基于核苷酸多态性的检测, 基

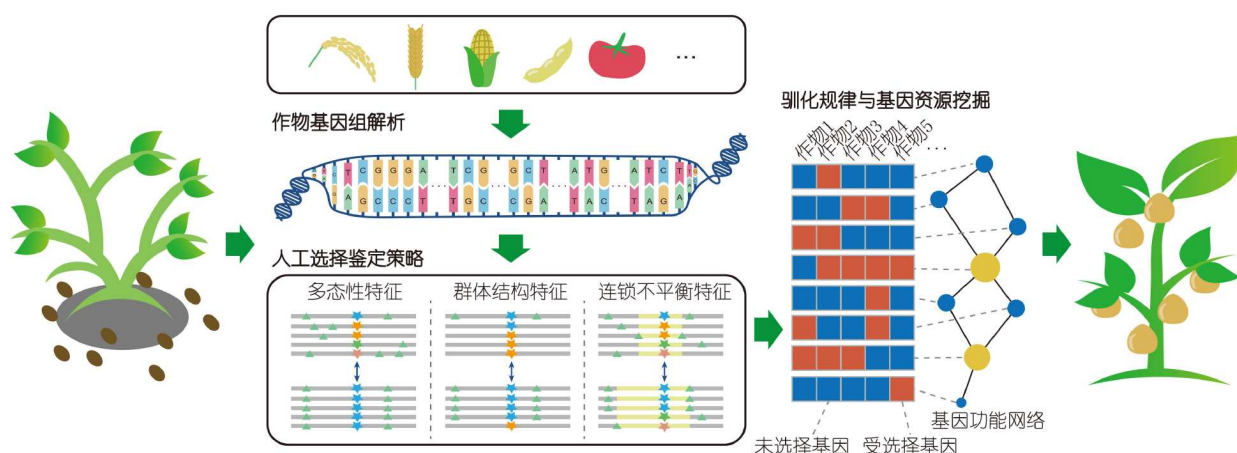


图 2 作物驯化规律的基因组解析与应用研究模式. 其包括研究对象选择、基因组解析、人工选择特征鉴定、驯化规律总结与基因资源挖掘等部分

Figure 2 Genomic dissection and application of crop domestication. Pipelines for study include target crop selection, genome dissection, artificial selective feature detection, domestication law conclusion, and gene resource discovery

于群体结构的检测和基于连锁不平衡的检测(图2). 作物在人工驯化过程中面临瓶颈效应, 在受选择的区域及周边会产生明显的遗传多态性下降. 在群体水平, 可以经由变异的等位基因频率基于几何分布或超几何分布计算变异位点及基因组区段的核苷酸多态性. 另外, 基于核酸多态性的Tajima's D 也是一种常用的检验方式^[92]. 当检测值趋近0时为选择近中性; 取值接近-2及2时分别代表正选择及歧化选择, 前者通常被用于描述人工选择信号. 需要注意的是, Tajima's D 的检测只需要单独的检测群体, 而核苷酸多态性需要在野生亲缘和栽培种质群体间进行比较分析, 其适用场景有所差异. 这些方法在植物中的使用相当普遍, 被应用在如芸薹属植物亚基因组平行选择、水稻驯化、大豆等作物休眠基因的平行选择等研究中^[84,93,94]. 在一些情况下, 野生群体可能自身的核苷酸多态性不高, 但是仍然发生显著的等位基因变化. 这种情况通常可以借由群体结构分析进行描述, 常见方法有Fixation-index(F_{ST})及XP-CLR. F_{ST} 的计算方式为核苷酸多样性的方差来源分析, 将研究对象分为野生和栽培群体两组, 计算核苷酸多样性组间方差的比例, 表示群体间分化的程度^[95]. XP-CLR的关键原理是通过比较两个种群的等位基因频率差异, 结合遗传距离和重组率信息, 识别超出中性进化预期的区域^[96]. 相较于 F_{ST} 方法, 该计算方式考虑位点连锁信息, 可以降低群体历史干扰. 该方法在大豆、木薯、小麦等作物的人工选择中

得到应用, 鉴定一系列与农艺性状相关的基因组候选区间^[4,33,97]. 经历自然或人工选择后, 基因上可以观察到部分区段的连锁不平衡显著增强. 单倍型延展同质性(EHH)方法可以基于基因组位点间连锁不平衡的衰减程度, 估计位点与周围变异在基因组上组成单倍型的延伸范围^[98]. 并且基于野生-栽培分群的群体间比较, 估计人工选择强度. 该方法在人类群体遗传学研究中提出, 在芸薹属作物的人工驯化研究中也受到应用^[93,98]. 上述研究方法通常在基因组上以滑动窗口方式进行覆盖全基因组的参数检测, 并且依照群体参数的分布范围划定显著性阈值, 以及界定驯化选择区间和驯化相关基因. 由于选择压力作用于基因组会呈现出复杂的变化模式, 其他一些工具则会对多种类型参数进行计算整合, 鉴定人工选择区间. 如RAiSD方法整合了多态性、等位基因频谱(SFS)、连锁不平衡等方法, 与中性进化背景进行对比, 综合考虑基因组区间的选择特征^[99]. 此外, 一些机器学习的方法可以更为灵活有效地识别基因组上的差异特征, 也开发一系列基于此的选择性检验方法, 如S/HIC, SIA等^[100]. 这些方法在作物人工选择分析中的效力有待进一步验证.

2.2 作物驯化的多维组学研究视角与手段

基于群体遗传学的选择性检验是研究作物驯化基因组规律的基本手段, 而不断发展的多维组学研究技

术极大拓展人工选择的研究视角. 如何获得全面完整的DNA变异信息, 这些变异在转录、翻译、生理生化水平发挥作用, 最终塑造表型, 是领域当下及未来研究的重点方向.

随研究的深入, 学者们认识到单一参考基因组涵盖的遗传多样性十分有限, 难以满足作物遗传资源挖掘及农艺性状相关位点克隆的需求. 因此, 比较与整合多个体基因组的泛基因组研究是作物驯化规律及遗传育种的重要研究领域^[101]. 2011年, Gan等人^[102]对拟南芥自然材料的基因组比较是植物泛基因组研究的雏形. 此后泛基因组研究逐渐在植物中兴起, 在许多作物及其近缘物种中, 如水稻、小麦、玉米、大豆、番茄、油菜等, 都有了泛基因组工作的开展^[103~113]. 泛基因组极大地补充作物遗传多样性的解析, 扩充对变异, 尤其是染色体结构变异的认知, 挖掘到一系列驯化演化相关的位点. 在油菜的泛基因组研究中, Song等人^[107]使用存在/缺失变异(PAV)进行关联分析, 其信号区间和单核苷酸多态性(SNP)的计算结果共定位, 而其中3.6 kb的PAV位于信号峰值. 该变异为一个转座子(TE)插入, 统计NAM群体的表型发现对于这个变异的存在/缺失, 角果长度和种子重量都呈现显著差异. 而该TE下游为一个*CYP78A9*基因, 推测变异影响此基因的表达, 造成性状的变化^[107]. 大豆的泛基因组研究发现, 光周期调控基因*E3*上存在多种结构变异类型, 其中一类造成编码序列的丢失及基因融合^[108]. 而大豆中与种皮颜色相关的*I*位点更为复杂. 通过泛基因组的解析, 发现该位点存在更为多样的结构变异类型. 除以往发现的110 kb左右包含*CHS*基因单元的倒位外, 另一个涉及*CHS*基因单元的23 kb倒位也可能造成种皮颜色的变化^[108]. 一项非洲栽培稻及亚洲栽培稻、野生稻泛基因组, 发现直系同源的*PROG1*和*PROG7*位点在亚洲和非洲栽培稻中分别受到驯化. 研究鉴定到一个在亚洲栽培稻种与*PROG1*紧密连锁的110 kb基因组片段缺失, 以及一个非洲栽培稻中*PROG7*附近113 kb的基因组片段缺失, 在驯化中分别被固定, 造成株型由匍匐到直立的表型^[27,28]. 以上研究表明, 通过泛基因组研究不仅能够辅助新基因的定位, 对于已知的驯化基因也能挖掘到新的功能变异. 因此, 泛基因组是开展作物驯化演化规律研究和关键变异挖掘的重要手段.

在多维组学层面, 包括基因表达、表观修饰、代

谢组、三维结构等方面, 驯化造成的影响同样能被观察到. 与去驯化过程相反, 水稻驯化过程中DNA甲基化明显降低, 可能通过影响邻近以及远距离基因的表达来调控表型变异^[114]. 大豆群体水平DNA甲基化的研究发现, 野生和栽培大豆间存在许多差异甲基化区间^[115]. 这些区间中富集到许多与碳水化合物代谢相关的基因, 与基因组驯化区间表现出较大的差别^[115]. 对番茄驯化的代谢组与基因组联合研究发现, 影响代谢物积累的基因与果实大小基因存在连锁, 对大果实相关基因的选择造成代谢物谱的改变. 此外还发现对5个主要关键基因位点的选择显著降低成熟果实中抗营养甾体糖生物碱的含量, 从而提升现代番茄果实的可食性^[116]. 棉花中13个棉属物种泛基因组研究发现, 在棉属物种演化过程中发生的结构变异与三维基因组区间关联的事件中, 2.6%的插入变异及2.2%的缺失变异造成染色体拓扑结构域的改变^[117].

3 从头驯化驱动的未来作物设计

3.1 从头驯化的概念与思路

从头驯化是将野生或半野生植物资源, 依据驯化规律进行高效改造, 使其成为栽培作物的一种全新育种策略^[13,20,90]. 现今地球人口数量不断上升, 对于主粮、经济等各类作物的需求也达到新的高度. 然而, 人类能够利用的植物资源有限, 获得充分驯化的作物更是少数, 不论是可供栽培的作物数量还是作物多样性层面, 都面临着潜在的危机. 因此, 从头驯化是未来应对这一危机的重要探索方向.

从头驯化研究的整体思路主要有三个阶段: 野生资源的收集与评估, 驯化规律解析与育种策略制定, 遗传改造体系构建与验证(图3). 从头驯化育种, 首先要明确育种的目标, 例如提供主粮、蛋白、油脂, 或者适应特定农业生态体系等. 基于这些目标, 选择合适的野生类群. 这些野生类群往往具备上述方面的优势, 同时与现有作物存在差异. 其中一些在有限地区种植、对低投入条件具有较好适应性、交易量不大且很少受到研究人员关注的作物, 通常被称为孤儿作物. 这些植物类群中可能包含有营养或抗性方面的重要利用价值, 在从头驯化中是值得着重挖掘的^[20,118]. 其次是设计野生资源从头驯化的路径. 其中面对核心问题包括: 人工驯化过程是否可以重演, 驯化的重演可以

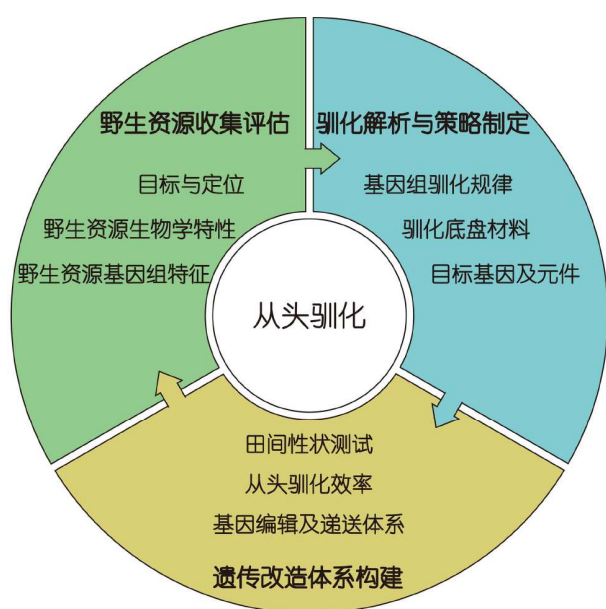


图3 野生资源从头驯化的一般研究思路

Figure 3 A general strategy for *de novo* domestication of wild resources

多快的速度发生, 已发生过驯化的物种是否具有新的驯化可能. 一方面, 针对已有驯化种的野生物种, 其已发生的驯化历史是很重要的参考, 对于关键驯化基因/基因组元件, 及突变模式的解析能够极大地辅助从头驯化设计; 对于孤儿植物, 可能需要解析其驯化性状在生理、生化上的成因, 并且利用作物驯化中更为普遍的遗传改造路径, 进行设计探索. 与此同时, 基因编辑等技术的日趋成熟, 包括碱基定向替换、大片段序列的定向敲入敲除等, 为快速驯化奠定坚实的技术基础. 最后, 遗传改造体系的构建是实践植物从头驯化的重要步骤. 遗传改造体系包括野生资源详尽的功能基因组信息, 稳定、高效率的基因递送系统, 以及相应的培育及表型鉴定平台. 此外, 合成生物学的方法对于作物育种范式的转变也可能发挥重要作用^[119-121].

3.2 作物从头驯化实践案例

(1) 野生番茄从头驯化. 番茄(*Solanum lycopersicum*)是世界重要的果蔬作物, 最早于7000年前左右由南美洲先民驯化, 其野生祖先为醋栗番茄(*S. pimpinellifolium*). 经过人工驯化, 樱桃番茄(*S. lycopersicum* var. *cerasiforme*)物种得以出现. 而现代的栽培番茄, 即大果番茄, 是在16世纪左右由欧洲人自美洲引种樱桃番

茄后改造获得, 经历了近期强烈的遗传瓶颈效应. 栽培番茄驯化性状, 通常被描述为紧凑的株型(*SP*)、削弱的光周期敏感性(*SP5G*)、变大的果实(*fw2.2*, *FAS*)、果实形状(*OVA TE*)、改变的营养物质含量(*CycB*)^[22,69,122-125]. 醋栗番茄具有极强的耐逆性及更丰富的风味, 受到学者的关注. Zsögön^[17]及Li等人^[126]依据这些已知基因的功能, 对醋栗番茄进行基因定向改造尝试. 借助CRISPR技术, 在醋栗番茄中进行基因敲除, 得到如株型、果实大小/房室数、果实数目接近栽培番茄的遗传材料, 证实了通过对驯化关键基因的基因编辑可以实践野生资源的快速驯化^[17,126]. 这些工作为从头驯化提供重要的启示意义.

(2) 多倍体野生稻从头驯化. 世界人口增长为主粮产量带来前所未有的挑战, 而气候快速变化、极端天气频发等问题加剧主粮创新的需求^[127,128]. 水稻(*Oryza sativa*)是世界重要的主粮作物, 被认为最早于9000年前左右被驯化, 主要来源为东亚的普通野生稻(*O. rufipogon*). 作为主粮的亚洲栽培稻为二倍体植物, 包括籼稻(*O. sativa* ssp. *indica/xian*)和粳稻(*O. sativa* ssp. *japonica/geng*)两个亚种. 而多倍体植物通常认为在生物量及环境变化的适应性上具有更明显的优势, 具有开发的潜能. 基于此, Yu等人^[18]聚焦于多倍体野生稻, 进行种质资源挖掘和从头驯化探索. 研究人员在5种异源四倍体稻属物种中, 锁定了CCDD基因组物种作为研究对象. 在28个该类型种系中进行愈伤组织诱导再生能力检测, 筛选出具有高转化效率的*O. alta*物种品系, 作为从头驯化基础材料. 通过三代测序与Hi-C测序的组装策略, 完成*O. alta*的高质量参考基因组. 比较基因组分析发现*O. alta*相对于栽培稻具有更多的生物抗性基因拷贝, 与其表型吻合, 也表明其遗传资源的利用前景.

针对栽培水稻的特征, 包括直立生长习性、落粒性、颖壳颜色、芒形态、种皮颜色和籽粒大小等, 研究人员总结一系列栽培稻中相关的基因, 评估这些基因在*O. alta*两个亚基因组中序列相似性情况. 首先对落粒性基因*qSH1*及芒长基因*An-1*的同源基因在*O. alta*中进行敲除, 发现能够改变*O. alta*的相关表型, 证明了基于同源基因对多倍体野生稻进行快速改造的可行性. 之后, 针对矮秆、籽粒大小等生产重要性状, 理想株型、光周期一致性等改良性状, 研究人员对*SD1*, *GS3*, *IPA1*, *DTH7*与*Ghd7*的同源基因在*O. alta*中进行

编辑并也得到理想的表型. 该工作提出野生资源从头驯化的核心思路及实践方案, 表明深度的基因组解析结合高效的基因编辑系统, 是未来作物育种的重要方向.

4 总结与展望

基因组学和分子生物学技术的发展, 打开生物遗传机制解析的大门. 如今人类已经从归纳生物演化现象, 到总结遗传规律, 乃至编辑遗传信息. 这不仅是认知维度的提升, 也是面向未来生物多样性安全、粮食安全等重要命题的理论与实践依据. 作物驯化是农业乃至人类文明发展的基石, 是人类行为参与的自然生物演化, 二者合一的典型事件. 作物驯化的遗传演化规律隐藏着人类既往改造作物的“参考答案”, 而将其运用于未来作物的种质创新, 不仅需要对遗传规律、分子机制的总结与还原, 还需要有逆向工程及整体设计的思路. 现阶段植物的基因组及群体遗传学研究呈现出深度和广度并举的现状, 主粮及重要作物从单一基因组到泛基因组, 从功能基因到全景多维组学的研

究积累丰盛, 许多经济植物、野生植物类群也有自身的参考基因组及遗传图谱解析. 在此基础上, 作物的从头驯化也在部分植物类群中得到实践, 未来也会有更多的植物资源得到发掘并且获得驯化.

与此同时, 基于遗传规律的生物育种实践存在着诸多的挑战. 首先, 一种作物驯化相关的遗传位点与基因, 在该作物的栽培种中已经获得广泛应用, 对于该类群的持续改良作用有限. 而基于多种作物的驯化基因比较, 建立驯化规律在不同作物中的可迁移知识体系, 将为作物育种提供新的思路和基因资源. 其次, 由于复杂的遗传调控网络驯化基因通常在植物中不是单独起作用, 在关键驯化基因/位点外, 基因组整体的变化特征同样值得重视. 主效的变异可能通过诸多次效及微效的变异共同发挥驯化作用, 因此解析驯化作用的基因网络, 分离核心驯化基因组元件, 可能为作物设计提供更精准的操作蓝图. 此外, 基因型-环境-表型三者的共同作用是育种面对的实际场景, 人工智能等技术手段的应用, 将提供针对高维度、高复杂度生物数据的处理方法. 智能地设计作物, 创制智能的作物品种, 是面向育种4.0乃至5.0时代的未来作物育种范式.

参考文献

- 1 Brown T A, Jones M K, Powell W, et al. The complex origins of domesticated crops in the Fertile Crescent. *Trends Ecol Evol*, 2009, 24: 103–109
- 2 Fuller D Q. Contrasting patterns in crop domestication and domestication rates: recent archaeobotanical insights from the Old World. *Ann Bot*, 2007, 100: 903–924
- 3 Meyer R S, Purugganan M D. Evolution of crop species: genetics of domestication and diversification. *Nat Rev Genet*, 2013, 14: 840–852
- 4 Zhou Z, Jiang Y, Wang Z, et al. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol*, 2015, 33: 408–414
- 5 Lin T, Zhu G, Zhang J, et al. Genomic analyses provide insights into the history of tomato breeding. *Nat Genet*, 2014, 46: 1220–1226
- 6 Duarte C M, Marbá N, Holmer M. Rapid domestication of marine species. *Science*, 2007, 316: 382–383
- 7 Doebley J F, Gaut B S, Smith B D. The molecular genetics of crop domestication. *Cell*, 2006, 127: 1309–1321
- 8 Krug A S, Drummond E B M, Van Tassel D L, et al. The next era of crop domestication starts now. *Proc Natl Acad Sci USA*, 2023, 120: e2205769120
- 9 Lin X, Fang C, Liu B, et al. Natural variation and artificial selection of photoperiodic flowering genes and their applications in crop adaptation. *aBIOTECH*, 2021, 2: 156–169
- 10 Chang C, Zhang H, Lu J, et al. Genetic improvement of wheat with pre-harvest sprouting resistance in China. *Genes*, 2023, 14: 837
- 11 Sohn S I, Pandian S, Kumar T S, et al. Seed dormancy and pre-harvest sprouting in rice—an updated overview. *Int J Mol Sci*, 2021, 22: 11804
- 12 Stetter M G. Limits and constraints to crop domestication. *Am J Bot*, 2020, 107: 1617–1621
- 13 Fernie A R, Alseekh S, Liu J, et al. Using precision phenotyping to inform *de novo* domestication. *Plant Physiol*, 2021, 186: 1397–1411
- 14 Huang X, Huang S, Han B, et al. The integrated genomics of crop domestication and breeding. *Cell*, 2022, 185: 2828–2839
- 15 Larson G, Piperno D R, Allaby R G, et al. Current perspectives and the future of domestication studies. *Proc Natl Acad Sci USA*, 2014, 111:

- 6139–6146
- 16 Paterson A H. What has QTL mapping taught us about plant domestication? *New Phytol*, 2002, 154: 591–608
- 17 Zsögön A, Čermák T, Naves E R, et al. *De novo* domestication of wild tomato using genome editing. *Nat Biotechnol*, 2018, 36: 1211–1216
- 18 Yu H, Lin T, Meng X, et al. A route to *de novo* domestication of wild allotetraploid rice. *Cell*, 2021, 184: 1156–1170.e14
- 19 Andersson L, Purugganan M. Molecular genetic variation of animals and plants under domestication. *Proc Natl Acad Sci USA*, 2022, 119: e2122150119
- 20 Tian Z, Wang J, Li J, et al. Designing future crops: challenges and strategies for sustainable agriculture. *Plant J*, 2021, 105: 1165–1178
- 21 Ma Y, Dai X, Xu Y, et al. COLD1 confers chilling tolerance in rice. *Cell*, 2015, 160: 1209–1221
- 22 Pnueli L, Carmel-Goren L, Hareven D, et al. The *SELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development*, 1998, 125: 1979–1989
- 23 Wills D M, Fang Z, York A M, et al. Defining the role of the MADS-Box gene, *Zea Agamous-like1*, a target of selection during maize domestication. *J Heredity*, 2018, 109: 333–338
- 24 Dong Y, Yang X, Liu J, et al. Pod shattering resistance associated with domestication is mediated by a *NAC* gene in soybean. *Nat Commun*, 2014, 5: 3352
- 25 Studer A, Zhao Q, Ross-Ibarra J, et al. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet*, 2011, 43: 1160–1163
- 26 Xu G, Lyu J, Li Q, et al. Evolutionary and functional genomics of DNA methylation in maize domestication and improvement. *Nat Commun*, 2020, 11: 5539
- 27 Wu Y, Zhao S, Li X, et al. Deletions linked to *PROG1* gene participate in plant architecture domestication in Asian and African rice. *Nat Commun*, 2018, 9: 4157
- 28 Shang L, Li X, He H, et al. A super pan-genomic landscape of rice. *Cell Res*, 2022, 32: 878–896
- 29 Lv S, Wu W, Wang M, et al. Genetic control of seed shattering during African rice domestication. *Nat Plants*, 2018, 4: 331–337
- 30 Tuteja J H, Clough S J, Chan W C, et al. Tissue-specific gene silencing mediated by a naturally occurring chalcone synthase gene cluster in *Glycine max*. *Plant Cell*, 2004, 16: 819–835
- 31 Yang Z, Ge X, Yang Z, et al. Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. *Nat Commun*, 2019, 10: 2989
- 32 Olsen K M, Wendel J F. A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol*, 2013, 64: 47–70
- 33 Ramu P, Esuma W, Kawuki R, et al. Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nat Genet*, 2017, 49: 959–963
- 34 Li N, Xu R, Li Y. Molecular networks of seed size control in plants. *Annu Rev Plant Biol*, 2019, 70: 435–463
- 35 Ren D, Ding C, Qian Q. Molecular bases of rice grain size and quality for optimized productivity. *Sci Bull*, 2023, 68: 314–350
- 36 Li N, Li Y. Signaling pathways of seed size control in plants. *Curr Opin Plant Biol*, 2016, 33: 23–32
- 37 Si L, Chen J, Huang X, et al. *OsSPL13* controls grain size in cultivated rice. *Nat Genet*, 2016, 48: 447–456
- 38 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
- 39 Wang E, Wang J, Zhu X, et al. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat Genet*, 2008, 40: 1370–1374
- 40 Song X, Meng X, Guo H, et al. Targeting a gene regulatory element enhances rice grain yield by decoupling panicle number and size. *Nat Biotechnol*, 2022, 40: 1403–1411
- 41 Jin J, Hua L, Zhu Z, et al. *GAD1* encodes a secreted peptide that regulates grain number, grain length, and awn development in rice domestication. *Plant Cell*, 2016, 28: 2453–2463
- 42 Luo J, Liu H, Zhou T, et al. *An-1* encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell*, 2013, 25: 3360–3376
- 43 Gu B, Zhou T, Luo J, et al. *An-2* encodes a cytokinin synthesis enzyme that regulates awn length and grain production in rice. *Mol Plant*, 2015, 8: 1635–1650
- 44 Hua L, Wang D R, Tan L, et al. *LABAI*, a domestication gene associated with long, barbed awns in wild rice. *Plant Cell*, 2015, 27: 1875–1888
- 45 Wang Y, Shi D, Zhu H, et al. Revisiting maize *Brittle endosperm-2* reveals new insights in BETL development and starchy endosperm filling.

- [Plant Sci](#), 2023, 332: 111727
- 46 Tan L, Li X, Liu F, et al. Control of a key transition from prostrate to erect growth in rice domestication. [Nat Genet](#), 2008, 40: 1360–1364
 - 47 Liu Y, Yao F, Zou J, et al. *RPAD* locus controls prostrate growth habit in *Oryza nivara*. [Plant Genome](#), 2025, 18: e70032
 - 48 Jin J, Huang W, Gao J P, et al. Genetic control of rice plant architecture under domestication. [Nat Genet](#), 2008, 40: 1365–1369
 - 49 Wu X, Liu Y, Luo H, et al. Genomic footprints of sorghum domestication and breeding selection for multiple end uses. [Mol Plant](#), 2022, 15: 537–551
 - 50 Tian Z, Wang X, Lee R, et al. Artificial selection for determinate growth habit in soybean. [Proc Natl Acad Sci USA](#), 2010, 107: 8563–8568
 - 51 Liu B, Watanabe S, Uchiyama T, et al. The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis TERMINAL FLOWER1*. [Plant Physiol](#), 2010, 153: 198–210
 - 52 Li S, Ding Y, Zhang D, et al. Parallel domestication with a broad mutational spectrum of determinate stem growth habit in leguminous crops. [Plant J](#), 2018, 96: 761–771
 - 53 Konishi S, Izawa T, Lin S Y, et al. An SNP caused loss of seed shattering during rice domestication. [Science](#), 2006, 312: 1392–1396
 - 54 Doebley J. The genetics of maize evolution. [Annu Rev Genet](#), 2004, 38: 37–59
 - 55 Li C, Zhou A, Sang T. Rice domestication by reducing shattering. [Science](#), 2006, 311: 1936–1939
 - 56 Lin Z, Griffith M E, Li X, et al. Origin of seed shattering in rice (*Oryza sativa* L.). [Planta](#), 2007, 226: 11–20
 - 57 Lin Z, Li X, Shannon L M, et al. Parallel domestication of the *Shattering1* genes in cereals. [Nat Genet](#), 2012, 44: 720–724
 - 58 Funatsuki H, Suzuki M, Hirose A, et al. Molecular basis of a shattering resistance boosting global dissemination of soybean. [Proc Natl Acad Sci USA](#), 2014, 111: 17797–17802
 - 59 Li S, Wang W, Sun L, et al. Artificial selection of mutations in two nearby genes gave rise to shattering resistance in soybean. [Nat Commun](#), 2024, 15: 7588
 - 60 Yong B, Zhu W, Wei S, et al. Parallel selection of loss-of-function alleles of *Pdh1* orthologous genes in warm-season legumes for pod indehiscence and plasticity is related to precipitation. [New Phytol](#), 2023, 240: 863–879
 - 61 Takahashi Y, Shimamoto K. Heading date 1 (*Hd1*), an ortholog of *Arabidopsis CONSTANS*, is a possible target of human selection during domestication to diversify flowering times of cultivated rice. [Genes Genet Syst](#), 2011, 86: 175–182
 - 62 Wu F, Sedivy E J, Price W B, et al. Evolutionary trajectories of duplicated *FT* homologues and their roles in soybean domestication. [Plant J](#), 2017, 90: 941–953
 - 63 Blackman B K, Strasburg J L, Raduski A R, et al. The role of recently derived *FT* paralogs in sunflower domestication. [Curr Biol](#), 2010, 20: 629–635
 - 64 Lu S, Dong L, Fang C, et al. Stepwise selection on homeologous *PRR* genes controlling flowering and maturity during soybean domestication. [Nat Genet](#), 2020, 52: 428–436
 - 65 Chen L, Luo J, Jin M, et al. Genome sequencing reveals evidence of adaptive variation in the genus *Zea*. [Nat Genet](#), 2022, 54: 1736–1745
 - 66 Farré E M, Liu T. The *PRR* family of transcriptional regulators reflects the complexity and evolution of plant circadian clocks. [Curr Opin Plant Biol](#), 2013, 16: 621–629
 - 67 Song C H, Li X P, Yang Z J, et al. Research progress on the regulation by the *FLOWERING LOCUS T(FT)* gene in plant flowering and development (in Chinese). [Sci Sin Vitae](#), 2025, 55: 1332–1346 [宋春晖, 李晓朋, 杨梓健, 等. 成花素基因*FT*调控植物开花和发育的研究进展. 中国科学: 生命科学, 2025, 55: 1332–1346]
 - 68 Song J, Zhang S, Wang X, et al. Variations in both *FTL1* and *SP5G*, two tomato *FT* paralogs, control day-neutral flowering. [Mol Plant](#), 2020, 13: 939–942
 - 69 Soyk S, Müller N A, Park S J, et al. Variation in the flowering gene *SELF PRUNING 5G* promotes day-neutrality and early yield in tomato. [Nat Genet](#), 2017, 49: 162–168
 - 70 Lifschitz E, Eviatar T, Rozman A, et al. The tomato *FT* ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. [Proc Natl Acad Sci USA](#), 2006, 103: 6398–6403
 - 71 Huang C, Sun H, Xu D, et al. *ZmCCT9* enhances maize adaptation to higher latitudes. [Proc Natl Acad Sci USA](#), 2018, 115: E334–E341
 - 72 Su H, Liang J, Abou-Elwafa S F, et al. *ZmCCT* regulates photoperiod-dependent flowering and response to stresses in maize. [BMC Plant Biol](#), 2021, 21: 453
 - 73 Li B, Wang Z, Jiang H, et al. *ZmCCT10*-relayed photoperiod sensitivity regulates natural variation in the arithmetical formation of male

- germinal cells in maize. [New Phytol](#), 2023, 237: 585–600
- 74 Liang Y, Liu Q, Wang X, et al. *ZmMADS69* functions as a flowering activator through the *ZmRap2.7-ZCN8* regulatory module and contributes to maize flowering time adaptation. [New Phytol](#), 2019, 221: 2335–2347
- 75 Lin X, Liu B, Weller J L, et al. Molecular mechanisms for the photoperiodic regulation of flowering in soybean. [J Integr Plant Biol](#), 2021, 63: 981–994
- 76 Xia Z, Watanabe S, Yamada T, et al. Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. [Proc Natl Acad Sci USA](#), 2012, 109: E2155–E2164
- 77 Xia Z, Zhai H, Wu H, et al. The synchronized efforts to decipher the molecular basis for soybean maturity loci *E1*, *E2*, and *E3* that regulate flowering and maturity. [Front Plant Sci](#), 2021, 12: 632754
- 78 Su T, Wang Y, Li S, et al. A flowering time locus dependent on *E2* in soybean. [Mol Breeding](#), 2021, 41: 35
- 79 Watanabe S, Xia Z, Hideshima R, et al. A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. [Genetics](#), 2011, 188: 395–407
- 80 Watanabe S, Hideshima R, Xia Z, et al. Map-based cloning of the gene associated with the soybean maturity locus *E3*. [Genetics](#), 2009, 182: 1251–1262
- 81 Liu B, Kanazawa A, Matsumura H, et al. Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. [Genetics](#), 2008, 180: 995–1007
- 82 Sweeney M T, Thomson M J, Pfeil B E, et al. Caught red-handed: *Rc* encodes a basic helix-loop-helix protein conditioning red pericarp in rice. [Plant Cell](#), 2006, 18: 283–294
- 83 Gu X Y, Foley M E, Horvath D P, et al. Association between seed dormancy and pericarp color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid synthesis in weedy red rice. [Genetics](#), 2011, 189: 1515–1524
- 84 Wang M, Li W, Fang C, et al. Parallel selection on a dormancy gene during domestication of crops from multiple families. [Nat Genet](#), 2018, 50: 1435–1441
- 85 Sun L, Miao Z, Cai C, et al. *GmHsI-1*, encoding a calcineurin-like protein, controls hard-seededness in soybean. [Nat Genet](#), 2015, 47: 939–943
- 86 Hernández F, Lindström L I, Parodi E, et al. The role of domestication and maternal effects on seed traits of crop-wild sunflower hybrids (*Helianthus annuus*). [Ann Appl Biol](#), 2017, 171: 237–251
- 87 Woodhouse M R, Hufford M B. Parallelism and convergence in post-domestication adaptation in cereal grasses. [Phil Trans R Soc B](#), 2019, 374: 20180245
- 88 Lenser T, Theißen G. Molecular mechanisms involved in convergent crop domestication. [Trends Plant Sci](#), 2013, 18: 704–714
- 89 Chen Q, Li W, Tan L, et al. Harnessing knowledge from maize and rice domestication for new crop breeding. [Mol Plant](#), 2021, 14: 9–26
- 90 Fernie A R, Yan J. *De novo* domestication: an alternative route toward new crops for the future. [Mol Plant](#), 2019, 12: 615–631
- 91 Chen W, Chen L, Zhang X, et al. Convergent selection of a WD40 protein that enhances grain yield in maize and rice. [Science](#), 2022, 375: eabg7985
- 92 Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. [Genetics](#), 1989, 123: 585–595
- 93 Cheng F, Sun R, Hou X, et al. Subgenome parallel selection is associated with morphotype diversification and convergent crop domestication in *Brassica rapa* and *Brassica oleracea*. [Nat Genet](#), 2016, 48: 1218–1224
- 94 Xu X, Liu X, Ge S, et al. Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. [Nat Biotechnol](#), 2011, 30: 105–111
- 95 Nei M, Chesser R K. Estimation of fixation indices and gene diversities. [Ann Hum Genet](#), 1983, 47: 253–259
- 96 Chen H, Patterson N, Reich D. Population differentiation as a test for selective sweeps. [Genome Res](#), 2010, 20: 393–402
- 97 Zhou Y, Zhao X, Li Y, et al. Triticum population sequencing provides insights into wheat adaptation. [Nat Genet](#), 2020, 52: 1412–1422
- 98 Sabeti P C, Reich D E, Higgins J M, et al. Detecting recent positive selection in the human genome from haplotype structure. [Nature](#), 2002, 419: 832–837
- 99 Alachiotis N, Pavlidis P. RAI_{SD} detects positive selection based on multiple signatures of a selective sweep and SNP vectors. [Commun Biol](#), 2018, 1: 79
- 100 Hejase H A, Mo Z, Campagna L, et al. A deep-learning approach for inference of selective sweeps from the ancestral recombination graph. [Mol Biol Evol](#), 2022, 39: msab332

- 101 Wang Y H, Yu J X, Tang H B, et al. Research status and prospect of plant complex genomes and pan-genomes (in Chinese). *Sci Sin Vitae*, 2024, 54: 233–246 [王英豪, 余嘉鑫, 唐海宝, 等. 植物复杂基因组与泛基因组研究现状与展望. 中国科学: 生命科学, 2024, 54: 233–246]
- 102 Gan X, Stegle O, Behr J, et al. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature*, 2011, 477: 419–423
- 103 Stein J C, Yu Y, Copetti D, et al. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nat Genet*, 2018, 50: 285–296
- 104 Wang W, Mauleon R, Hu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*, 2018, 557: 43–49
- 105 Zhao Q, Feng Q, Lu H, et al. Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nat Genet*, 2018, 50: 278–284
- 106 Gao L, Gonda I, Sun H, et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat Genet*, 2019, 51: 1044–1051
- 107 Song J M, Guan Z, Hu J, et al. Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat Plants*, 2020, 6: 34–45
- 108 Liu Y, Du H, Li P, et al. Pan-genome of wild and cultivated soybeans. *Cell*, 2020, 182: 162–176.e13
- 109 Zhou Y, Zhang Z, Bao Z, et al. Graph pangenome captures missing heritability and empowers tomato breeding. *Nature*, 2022, 606: 527–534
- 110 Gui S, Wei W, Jiang C, et al. A pan-*Zea* genome map for enhancing maize improvement. *Genome Biol*, 2022, 23: 178
- 111 Li N, He Q, Wang J, et al. Super-pangenome analyses highlight genomic diversity and structural variation across wild and cultivated tomato species. *Nat Genet*, 2023, 55: 852–860
- 112 Zhuang Y, Wang X, Li X, et al. Phylogenomics of the genus *Glycine* sheds light on polyploid evolution and life-strategy transition. *Nat Plants*, 2022, 8: 233–244
- 113 Guo D, Li Y, Lu H, et al. A pangenome reference of wild and cultivated rice. *Nature*, 2025, 642: 662–671
- 114 Cao S, Chen K, Lu K, et al. Asymmetric variation in DNA methylation during domestication and de-domestication of rice. *Plant Cell*, 2023, 35: 3429–3443
- 115 Shen Y, Zhang J, Liu Y, et al. DNA methylation footprints during soybean domestication and improvement. *Genome Biol*, 2018, 19: 128
- 116 Zhu G, Wang S, Huang Z, et al. Rewiring of the fruit metabolome in tomato breeding. *Cell*, 2018, 172: 249–261.e12
- 117 Wang M, Li J, Qi Z, et al. Genomic innovation and regulatory rewiring during evolution of the cotton genus *Gossypium*. *Nat Genet*, 2022, 54: 1959–1971
- 118 Wang N, Li H, Huang S. Rational redomestication for future agriculture. *Annu Rev Plant Biol*, 2025, 76: 637–662
- 119 Qu Y, Fernie A R, Liu J, et al. Doubled haploid technology and synthetic apomixis: recent advances and applications in future crop breeding. *Mol Plant*, 2024, 17: 1005–1018
- 120 Cai T, Sun H, Qiao J, et al. Cell-free chemoenzymatic starch synthesis from carbon dioxide. *Science*, 2021, 373: 1523–1527
- 121 Wang F H, Le L, Pu L. Applications of synthetic biology in genetic improvement of crops (in Chinese). *Sci Sin Vitae*, 2024, 54: 845–865 [王繁华, 乐亮, 普利. 合成生物学在作物遗传改良中的应用. 中国科学: 生命科学, 2024, 54: 845–865]
- 122 Xu C, Liberatore K L, MacAlister C A, et al. A cascade of arabinosyltransferases controls shoot meristem size in tomato. *Nat Genet*, 2015, 47: 784–792
- 123 Frary A, Nesbitt T C, Frary A, et al. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science*, 2000, 289: 85–88
- 124 Liu J, Van Eck J, Cong B, et al. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. *Proc Natl Acad Sci USA*, 2002, 99: 13302–13306
- 125 Ronen G, Carmel-Goren L, Zamir D, et al. An alternative pathway to β -carotene formation in plant chromoplasts discovered by map-based cloning of *Beta* and *old-gold* color mutations in tomato. *Proc Natl Acad Sci USA*, 2000, 97: 11102–11107
- 126 Li T, Yang X, Yu Y, et al. Domestication of wild tomato is accelerated by genome editing. *Nat Biotechnol*, 2018, 36: 1160–1163
- 127 Godfray H C J, Beddington J R, Crute I R, et al. Food security: the challenge of feeding 9 billion people. *Science*, 2010, 327: 812–818
- 128 Foley J A, Ramankutty N, Brauman K A, et al. Solutions for a cultivated planet. *Nature*, 2011, 478: 337–342

Deciphering of domestication principles enlightens future crop breeding

MA Xin¹, ZHANG Min¹, WANG JianLin² & LIU YuCheng^{1*}

¹ State Key Laboratory of Seed Innovation, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

² Zhongke-Dongying Research Center of Molecular Designed Breeding, Dongying 257509, China

* Corresponding author; E-mail: yichliu@genetics.ac.cn

Crop domestication represents a pivotal milestone in human civilization. Genomic evolutionary patterns underlying domestication harbor valuable insights into crop domestication and improvement. With increasingly frequent extreme weather events and rapid population growth posing challenges to global food security, expanding the repertoire of utilizable crops and improving crops are important strategies for addressing these challenges. This review systematically summarized fundamental phenotypic alterations and genomic variation patterns during domestication, and the genetic basis of domestication syndrome-related traits. It also outlines genomic research frameworks and methodologies related to crop domestication, while providing a comprehensive analysis and forward-looking perspective on *de novo* domestication. Collectively, this work reviews researches on deciphering genomic principles of crop domestication and their applications, aiming to lay a solid theoretical foundation for designed crop breeding in future.

crop, domestication syndrome, genomics, *de novo* domestication, breeding

doi: [10.1360/SSV-2025-0178](https://doi.org/10.1360/SSV-2025-0178)



刘羽诚, 中国科学院遗传与发育生物学研究所研究员, 博士生导师. 2021年于中国科学院遗传与发育生物学研究所获得博士学位. 2022年入选中国科学院稳定支持基础研究领域青年团队计划. 2024获得国家自然科学基金委青年科学基金(B类)资助. 致力于作物功能基因组研究, 在大豆泛基因组、生物大数据挖掘与应用平台开发领域完成一系列突破性工作. 在领域内国际核心期刊以第一/共同第一作者发表论文10余篇.