

李嘉琦,刘有春,魏鑫,等.全基因组关联分析在果树品质及抗性性状中的研究进展[J].江苏农业科学,2024,52(11):10-19.
doi:10.15889/j.issn.1002-1302.2024.11.002

全基因组关联分析在果树品质及抗性性状中的研究进展

李嘉琦,刘有春,魏鑫,杨艳敏,杨玉春,王升,高树清,林佳琦,徐艺格,
孙斌,张舵,王兴东,王宏光,刘成
(辽宁省果树科学研究所,辽宁营口 115009)

摘要:全基因组关联分析(genome-wide association analysis, GWAS)是一种基于连锁不平衡理论的成本低、精度高的基因分型方法,用于影响复杂性状的基因或基因组区域定位。随着果树各树种基因组的陆续公布,近年来 GWAS 在果树研究中得到广泛应用。本文介绍了 GWAS 的基本原理和影响因素,综述了 GWAS 在仁果类、核果类、柑橘类、浆果类和其他经济类果树中的研究进展,关联性状主要集中在果实品质性状评估(风味、香气、质地、果实外观、色泽、单果重等)和抗性基因鉴定(苹果褐斑病、李痘病毒、褐腐病、草莓枯萎病、葡萄裂果等)等方面,梳理了利用结构变异、转座子开展关联分析的研究报道,基于 GWAS + TWAS 等衍生应用在作物中的研究进展分析了对果树相关研究的借鉴和参考意义,最后归纳了果树 GWAS 研究的不足,并提出未来研究方向和建议。

关键词:全基因组关联分析(GWAS);果树;品质性状;抗性性状

中图分类号:S660.1 **文献标志码:**A **文章编号:**1002-1302(2024)11-0010-10

我国是目前世界上最大的水果生产国和消费国,种植面积和产量均居世界第一。全国共有 21 个国家果树种质资源圃,保存果树种质资源 2.3 万余份,是多种果树的起源演化中心之一,同时也是世界上最大的果树资源国^[1]。然而,我国的果树资源仍然存在数量不足、覆盖面积不足、缺少外国种质等问题^[2]。据调查发现,自 1950 年以来,我国育成了约 2 000 多个果树品种^[1],但在实际生产中这些品种的应用十分有限,这主要是由于在种质资源利用方面,骨干亲本的重复利用导致品种间基因同质

性高、遗传背景狭窄。此外,我国的果树育种以传统的杂交育种为主,现代育种技术与理论基础结合不够紧密,未能实现分子标记辅助育种、基因编辑、全基因组选择等实用化技术突破。我国的果树产业与先进国家之间存在一定差距,尤其是突破性优良品种方面,需加强对优异野生资源的挖掘利用,加强高效育种技术理论研究和实践应用,以高效培育出遗传背景丰富、品质优异的特色品种,调整产业品种结构、引领产业发展。

全基因组关联分析(genome-wide association study, GWAS)是利用连锁不平衡(linkage disequilibrium, LD)为理论基础,通过大样本群体中的表型变异和多态性标记展开统计分析,定位影响复杂性状的基因或基因组区域,是一种成本低、精度高的基因分型方法^[3],通过高密度单核苷酸多态性序列标记(single nucleotide polymorphism, SNP)在全基因组的分布,至少找到 1 个能够影响目标性状

收稿日期:2023-08-01

基金项目:辽宁省重点实验室建设项目(编号:2020JH13/10200051);

辽宁省农业科学院学科建设项目(编号:2019DD164924)。

作者简介:李嘉琦(1994—),女,辽宁凌源人,硕士,研究实习员,主要从事蓝莓遗传与育种研究。E-mail:755569357@qq.com。

通信作者:刘成,博士,研究员,主要从事小浆果种质资源、遗传育种及栽培技术研究。E-mail:stevecliu@hotmail.com。

[91] Liu Z W, Zhu M T, Wang J M, et al. The responses of soil organic carbon mineralization and microbial communities to fresh and aged biochar soil amendments [J]. Global Change Biology Bioenergy, 2019, 11(12): 1408-1420.

[92] Nan Q, Wang C, Wang H, et al. Mitigating methane emission via annual biochar amendment pyrolyzed with rice straw from the same paddy field [J]. The Science of the Total Environment, 2020, 746: 141351.

[93] Wu Z, Zhang Q Q, Zhang X, et al. Biochar-enriched soil mitigated N₂O and NO emissions similarly as fresh biochar for wheat production [J]. The Science of the Total Environment, 2020, 701: 134943.

[94] Yang Y, Sun K, Liu J, et al. Changes in soil properties and CO₂ emissions after biochar addition; role of pyrolysis temperature and aging [J]. The Science of the Total Environment, 2022, 839: 156333.

的数量性状位点 (quantitative trait locus, QTL), 从而实现目标性状的定位、新标记开发和新基因挖掘。GWAS 已成为研究植物种质资源和育种的重要手段和方法^[4]。在作物领域, 关联分析主要应用于水稻稻瘟病^[5]、小麦穗粒数^[6]、玉米穗部性状^[7]、油菜角果粒数^[8]、棉花产量^[9]、大豆耐低磷性^[10]等。近年来, 随着全基因组重测序成本大幅度降低, 果树基因组序列陆续公布, GWAS 已应用于包含樱桃^[11]、桃^[12]、苹果^[13]、梨^[14]、葡萄^[15]、柑橘^[16]等果树研究中。本文基于全基因组关联分析的基本原理和影响因素概述, 重点阐述了 GWAS 在果树重要性状定位的最新研究进展, 探讨了果树 GWAS 未来研究趋势。

1 全基因组关联分析的基本原理

连锁不平衡 (LD) 是指在一段基因序列中, 2 个不同基因座之间的 2 种等位基因非随机地联合存在的现象^[17]。LD 的存在会影响全基因组关联分析的精度和可靠性, 其程度的高低受多种因素影响, 包括基因座的物理距离、遗传多样性、种群历史等。在关联分析中, 通常使用 D' (standardized disequilibrium coefficients, D') 和 r^2 (squared allele - frequency correlation, r^2) 2 个参数来衡量基因座间的 LD 程度。 D' 值范围在 0 ~ 1 之间, 表示 2 个基因座间的非随机联合存在程度; r^2 值也在 0 ~ 1 之间, 表示 2 个基因座在不同群体中的等位基因频率的协方差, 值越高表示 LD 程度越高^[18]。全基因组关联分析利用 LD 的特性, 对大规模遗传数据进行统计分析, 寻找基因座与性状之间的关联, 从而找到可能影响性状的基因或基因组区域具有重要意义。

在 GWAS 分析中, r^2 被认定为是衡量 LD 的首选指标, 用于评估标记和遗传变异性状之间的相关性^[19]。LD 衰减是指随着物理距离的增加, 基因座之间的 r^2 值逐渐减小的过程, 表示从连锁不平衡状态到平衡状态的转变。一般情况下, 驯化程度较高的群体, 其 LD 衰减距离较大。

2 全基因组关联分析的影响因素

2.1 群体大小和结构

群体大小直接影响表型和基因型之间的关联关系, 大型群体具有增强友系关联的作用^[20]。然而, 特定的育种亲本及其后代可能会影响群体结构, 从而影响群体中 SNP 与目标性状之间的关联, 是因为估计的育种值会显著高于其他亲本材料, 并

且在亲本后代组成的亚群中等位基因频率更高, 因此在特定亚群中可能存在假阳性结果^[21]。为了平衡遗传多样性和等位基因频率, 选择全球范围内的多样性个体组成大群体是最佳选择^[20]。

2.2 表型变异

表型变异是关联分析的重要组成部分, 为了保证分析的准确性, 表型数据应符合正态分布。通常情况下, 绘制箱线图可以排除极端异常值。为了减少误差, 经常采用采集多年多点表型数据的方式, 但是这种方法可能会受到环境的影响, 从而导致遗传力下降。为了解决这个问题, 可以使用最佳线性无偏预测 (BLUP) 和最佳线性无偏估计 (BLUE) 等方法进行数据归一化处理, 调整基因型 × 环境间的相互作用, 上述方法可以有效地减少误差并提高分析的准确性^[22]。

2.3 统计模型

根据基因控制的数量不同, 可以将性状分为质量性状和数量性状。对于质量性状, 通常采用 Logistic 回归模型进行关联分析。而对于数量性状, 关联分析多采用一般线性模型 (GLM) 和混合线性模型 (MLM) 2 种方法^[23], GLM 模型通过引入协变量来解释基因型在群体中的方差^[24], 而 MLM 则在 GLM 的基础上加入群体结构和个体间关系, 以减少假阳性的影响^[25]。

3 GWAS 在果树中的应用

对于果树, 果实品质 (内在品质和外观品质) 和抗性是种质资源和育种研究经久不衰的方向, 以下综述了 GWAS 在仁果类、核果类、柑橘类、浆果类及其他果树上的研究应用, 分析现有文献, 关联分析相关研究也主要集中在品质性状评估和抗性基因鉴定等方面, 详见表 1。

3.1 仁果类果树

苹果是仁果类果树中的代表性树种, 我国苹果栽培面积和产量居世界首位^[26], 果实品质、病害、胁迫响应机制等一直是苹果研究的重点, GWAS 也不例外, 但相关报道较少。果实香气受糖酸含量、着色、重量、外界环境等多因素影响, 是苹果重要的品质因子^[27]。研究人员基于气相色谱 - 质谱法 (gas chromatography - mass spectrometry, GC - MS) 从 149 个苹果样本中鉴定出 49 种挥发性有机化合物 (volatile organic compounds, VOCs), 经关联分析发现, 在 2 号染色体上有和乙酸丁酯、乙酸己酯相关的

表 1 全基因组关联分析在果树中的应用

| 物种 | 材料数量 | 表型 | 研究结果 | 参考文献 |
|-----|--|---|---|------|
| 苹果 | 149 份丹麦特色栽培苹果 | 49 种挥发性物质(包含醇类、醛类、酯类、萜类、酮类、烯类) | 筛选到 chr1:30221387、chr2:1730413、chr3:31409362 等显著性位点 | [28] |
| | 192 份苹果种质 | 叶片感染褐斑病指数 | 发现 5 个显著性位点,其中 LOC103442381 和 LOC103443382 的候选基因与 AMB 抗性密切相关 | [29] |
| | 123 份苹果种质 | 17 苹果叶蜡成分(包含醇、烷烃、脂肪酸和萜烯) | 通过 GO 富集分析得到 <i>MdSHN1</i> 、 <i>MdWSD1</i> 、 <i>MdLTP4</i> 、 <i>MdRDR1</i> 、 <i>MdACBP6</i> 、 <i>MdNLE</i> 和 <i>MdABCG21</i> 等 7 个叶蜡合成和运输相关基因 | [33] |
| 梨 | 312 份沙梨种质 | 3 个物候期性状:初花期、果实发育期、营养生长天数 | 筛选到 Chr9_2479073、Chr13_11586174 等显著性位点,并发现 <i>Pbr-STONE</i> 、 <i>Pbr001726</i> 、 <i>Pbr013897</i> 等候选基因 | [36] |
| | 2 种上饶早梨、花厅六月雪和花厅黄皮消 | 转录组数据 | 1 561 个基因在上饶梨中表达,198 个转录本差异显著, <i>PCP028004</i> 基因表达差异达到 628 倍(花厅六月雪相对于花厅黄皮消) | [37] |
| | 亲本群体:86 育种群体:765 | 14 个品质性状:果肉硬度、果实重量、果皮颜色、糖含量、酸含量等 2 个抗病性状:抗黑斑病 1、抗黑斑病 2 2 个生长性状:成熟期、采前果实落果 | 在收获时间 果实重量、酸含量 果实表皮颜色、霉心病、采前果实落果 6 个性状中均鉴定到显著性位点 | [38] |
| 桃 | 129 份桃种质 | 9 个果实性状:果实硬度、果实形状、果肉质地、果肉颜色等 1 个种子性状:果仁味道 2 个花瓣性状:重瓣花型、蔷薇花型 | 发现 <i>PpCAD1</i> (<i>ppa003772m</i>) 参与决定果实形状, <i>ppa006339m</i> 为控制果实酸度的候选基因等 | [40] |
| 甜樱桃 | 26 个品种/品系 138 个育种后代 | 褐腐病感病指数 | 确定鉴定了 4 个与果肉和 10 个与果皮耐腐相关的 SNP 位点 | [46] |
| | 235 份甜樱桃种质 | 11 个性状:收获时间、果实颜色、果皮颜色、果皮硬度、果肉硬度、渗透率、果实瘀伤、果实长度、果实宽度、果实厚度、果实重量 | 候选基因 <i>MYB10.1</i> 存在单核苷酸缺失,与樱桃果实颜色和果肉颜色相关 | [42] |
| 杏 | 72 份杏种质 | 感染李痘病毒情况 | 发现 <i>pp022195</i> 、 <i>ppa012234m</i> 等基因,与李痘病毒抗性相关 | [43] |
| 柑橘 | 111 个柑橘品种(杂交品种:8,育种/选择品系:41,本土品种:22), 35 个全同胞家族(676 个 F ₁ 代个体) | 17 个性状:果实重量、外观、果实形状、果实硬度、果皮颜色、果皮光滑度、易剥程度、香气强度、果肉颜色、果肉硬度、多汁性、室膜硬度、种子数量、苦味、口感、含糖量、酸度 | 发现果肉硬度和多汁性有共同的显著性位点,重量、果皮颜色、苦味、酸度和含糖量有共同的显著峰 | [48] |
| | 110 份柑橘种质 | 8 个性状:果实重量、含糖量、含酸量、果皮颜色、果皮光滑度、可剥度、果肉硬度、果瓣硬度 | 鉴定了 7 个显著 QTL,4 个与果实重量相关,其他 3 个分别与果皮颜色、果肉硬度和果瓣硬度相关 | [49] |
| 葡萄 | 240 份宽皮柑橘 | 10 个农艺性状:单果重、纵径、果皮厚度、果皮光滑度、可溶性固形物含量、可滴定酸含量、种子胚型、花药颜色、囊瓣数、果形指数 | 共鉴定到 35 个显著性 SNP 标记,其中与花药颜色性状相关的显著性位点最多,共 10 个 | [50] |
| | 279 个葡萄品种 | 25 个浆果果形性状:投影周长、投影面积、横轴线长度、果实高度、纵轴线长度、果实宽度、沿果形曲线长度、果形指数外观 I、果形指数外观 II、果形弯曲指数、果端块状、果未块状、三角果形指数、果肩高度、近端微角度、近端宏角度、近端压痕面积、远端微角度、远端宏角度、最宽位置宽度(出现最大宽度的高度与最大高度的比率)、偏心率、近端偏心率、远端偏心率、果实形状指数、偏心面积指数 | 鉴定到 2 个与果实形态相关基因 <i>VIT_12s0134g000230</i> 和 <i>VIT_02s0025g01360</i> ,以及 2 个与多种浆果形状相关的泛素连接酶基因 <i>VIT_03s0088g01090</i> 和 <i>VIT_10s00003g04300</i> | [52] |

表 1(续)

| 物种 | 材料数量 | 表型 | 研究结果 | 参考文献 |
|------|--|--|---|------|
| 葡萄 | 179 份葡萄种质 | 8 个性状: 浆果发育期、果簇大小、果簇密度、浆果重量、浆果肉质质地、浆果颜色、浆果形状、浆果风味 | 筛选出多个显著性标记, 包括与葡萄果皮颜色相关的 <i>mybA</i> 转录因子, 以及与果实发育期相关的显著性位点 | [15] |
| | 287 个葡萄品种 | 浆果开裂率、浆果开裂指数和浆果开裂类型 | 鉴定到 5 个与浆果开裂指数相关的 SNP, 2 个与浆果开裂类型相关的 SNP, 与裂果基因相关 | [53] |
| 蓝莓 | 886 份杂交后代(92 个南高丛组合) | 17 种挥发性有机化合物(包含脂肪酸衍生物、萜类、苯类) | 发现 519 个显著性 SNP 与属于不同代谢途径的 11 个挥发性有机化合物 | [54] |
| 树莓 | 亲本材料: 2 个 F ₁ 后代: 90 个 | 树莓皮刺性状 | 发现 SNP 4_35148226 被鉴定为对无刺性状有积极贡献 | [55] |
| 栽培草莓 | 565 份草莓种质 | 草莓枯萎病 | 在 2 号染色体检测到与草莓枯萎病基因 <i>Fmd</i> 连锁不平衡的 14 个 SNP | [56] |
| 东方甜瓜 | 48 份东方甜瓜 | 46 种形态性状: 植物生长各个阶段的叶、叶柄、花序、果实和果肉的形态特征 | 发现 OM1_30331998 和 OM8_8159753 这 2 个 SNP 位点与性别表达相关, 4 个候选基因 <i>MELO3C015898</i> 、 <i>MELO3C015904</i> 、 <i>MELO3C024563</i> 和 <i>MELO3C024565</i> 与东方甜瓜的性别表达相关 | [57] |
| 中国枣 | 150 份中国枣核心种质 15 份极耐开裂材料 15 份极易开裂材料 | 9 个果实质量性状: 果实横径、果实纵径、果实重量、果形指数、果核横径、果粒纵径、开裂率、开裂水平、开裂指数 | 检测到 33 个 SNP 与果实大小相关性性状有关, 9 个与果核大小相关性性状有关, 3 个与果实开裂相关性性状有关 | [58] |
| 橄榄 | 183 份橄榄种质 | 5 个性状: 叶长、叶宽、果实重量、果核重量、果肉果核比 | 发现 52 个显著性标记与 4 种性状相关, 预测到 19 个候选基因, 大部分与生理发育相关 | [59] |

显著性标记,同时检测到 3 个与收获时期显著关联的 SNP 位点,其中 1 个 SNP(chr3:31409362)位于果实成熟转录因子 NAC18.1 的编码区^[28]。苹果褐斑病(apple marssonina blotch, AMB)是由苹果双壳菌引起的病害,可在叶片和果实上发生,导致早期落叶、降低果实质量、影响树木活力和生长。通过全基因组关联分析探索影响 AMB 疾病感染的基因,发现 5 个极显著 SNP 位点与 AMB 抗性相关, *LOC103442381* 和 *LOC103443382* 这 2 个 SNP 位点相关的候选基因可能与 AMB 抗性密切相关^[29]。叶片蜡质能够阻止非气孔水分流失,从而增强植株的抗寒、抗旱能力^[30-32],CAO 等基于 GC-MS 方法分析 123 份苹果新鲜叶片中蜡成分并进行 GWAS,筛选得到了大量显著性位点,并结合 GO 富集分析,研究发现 7 个与叶蜡合成和运输相关的基因,分别是 *MdSHN1*、*MdWSD1*、*MdLTP4*、*MdRDR1*、*MdACBP6*、*MdNLE* 和 *MdABCG21*,以上研究及结果为苹果育种中品质、抗病性基因挖掘及分析标记开发提供了重要参考^[33]。

我国梨栽培面积和产量占世界一半以上^[34],但存在着果实品质较差、经济效益低的问题^[35],开展梨品质资源鉴定、性状遗传基础研究从而进行性状改良遗传育种,解决梨口味淡、品质差的弊端具有重要意义。研究者基于 312 个沙梨初花期、果实发育期和营养生长天数等果实物候期性状、单果重、石细胞含量、果皮颜色等果实品质性状的 GWAS,发现了与物候期相关的 5 个基因座和与果实品质相关的 37 个候选基因座,其中候选基因 *PbrSTONE* 的功能被证实参与了梨石细胞形成的调控,这项研究为利用 GWAS 研究果实复杂生物学性状提供了新思路^[36]。尹明华等对 2 种上饶早梨品种进行全基因组测序和转录组测序,将获得的基因组数据和转录组数据进行关联分析,发现共有 1 561 个基因在上饶梨中表达,其中 198 个转录本差异显著,部分基因具有高表达差异,为上饶梨品种资源鉴定和品种选育提供了理论基础^[37]。Minamikawa 等也通过 GWAS 鉴定到与果实质量、酸含量、果实表皮颜色、霉心病、采前果实落果等性状显著关联的位点,对于探寻位点上控制上述性状的关键基因提供了参考^[38]。

综上所述,全基因组关联分析在有效鉴定果实发育中重要品质性状的遗传位点、挖掘调控果实发育过程相关基因等中发挥了重要作用,也为其他性

状的相关研究提供了参考和思路。

3.2 核果类果树

核果类果树包含桃、李、杏、樱桃等。纵观产业发展,传统桃、李、杏、樱桃品种普遍因果肉软、不耐运贮、货架期短及综合品质差等因素而面临淘汰,产业亟需品质好、硬度高的优良品种^[39]。此外,李痘病和褐腐病是危害核果类果树的主要病害,李痘病甚至会造成毁树毁园,褐腐病主要危害果实,严重时造成绝产,所以加强种质资源发掘利用、研究果实品质性状遗传基础、开发与性状关联的标记开发等,对于核果类果树品质及抗性选育具有重要意义。

研究人员利用由野生品种、地方品种、育成品种组成的 129 份桃种质进行高通量测序,并对果肉质、果实特性、种子特性等进行 GWAS,筛选得到的关于果肉质、果实形状和果肉黏度等性状的峰值信号与连锁分析结果相一致,并发现在桃的驯化和改良过程中,一些与果实重量和可溶性固形物含量相关的基因组区域与预测的选择性清除相重叠^[40]。对于樱桃而言,果皮颜色丰富,可分为黄色、红黄色、红色、紫红色、红紫色、紫色樱桃,樱桃果皮颜色呈红色与花青苷的含量与组成有关^[41]。为了提高甜樱桃标记密度,以 235 份甜樱桃种质为试验材料进行 GWAS,发现樱桃果实颜色和果肉颜色相关的候选基因 *MYB10.1* 存在单核苷酸缺失,位于 chr3_23995550 位点上,与最显著位点 chr3_23939472 相距约 55 kb,对黄色樱桃果实进行测序后,发现 3 号染色体上存在一段 90 kb 的缺失与果实颜色有关,其中包括 5 个 *MYB10* 转录因子,这些结果对于了解樱桃的遗传机制和 DNA 辅助育种具有重要意义^[42]。

李痘病毒(plum pox virus, PPV),别名莎卡(sharka),是李属(*Prunus* L.)植物中危害最大的病毒之一。为进一步了解杏树抗 PPV 的遗传特性,Marquette 等利用包括 22 份抗性材料在内的 72 份杏育种资源进行全基因组 SNP 分型,通过对 56 708 个 SNP 进行 GWAS,共鉴定到 38 个显著性 SNP,其中有 34 个 SNP 位点的候选基因与植物-病毒的相互作用有关,基因 *pp022195* 编码 BTB/POZ-MATH-TRAF 样蛋白, *ppa012234m* 推测编码丝裂原活化蛋白激酶(mitogen-activated protein kinase, MAPK)双特异性(丝氨酸/苏氨酸和酪氨酸)磷酸酶^[43]。这些候选蛋白表明植物降解途径存在与杏抗 PPV 相

关的潜在抗病毒机制^[44-45]。褐腐病(*Monilinia* spp.)是一种核果病害,会导致果实腐烂,造成产量严重损失。为了确定与褐腐病相关的基因组区域,研究人员选用了 26 个具有 Bolinha 品种抗性的桃栽培品种及其子代作为试验材料,在 2 年中对损伤和非损伤果实的褐腐病严重程度指数进行 GWAS,鉴定出了 4 个与果肉和 10 个与果皮耐腐相关的 SNP 位点,确定了 25 个与抗病性相关的候选基因,该报道研究结果有助于深入了解褐腐病的遗传机制,为 DNA 辅助育种提供了重要基础^[46]。

3.3 柑橘类果树

相比较于其他果树,柑橘已进入分子育种时代,国际上已经公布 14 个柑橘基因组,进入功能基因组和泛基因组时代^[47]。目前针对柑橘的农艺性状已经展开多项全基因组关联分析研究,获得了与果实重量、果皮颜色、果肉硬度、花药颜色等多种表型相关的显著性位点。国内外均有相关报道,根据 Minamikawa 等的研究,基于 111 个柑橘品种组成的亲本群体和 676 个 F_1 代个体组成的育种群体,获得 1 841 个 SNP 的基因分型,对包括果实重量、果皮颜色、果实硬度和种子数量等的 17 个农艺性状进行 GWAS,发现最显著的 SNP 位于 8 号染色体,并与果实重量性状相关联,且该 SNP 位于 QTL 作图研究中的 *FW8* 区域内^[48]。Imai 等对 110 份柑橘材料的 8 种果实品质性状进行全基因组关联分析,在 2 309 个 SNP 中鉴定到 7 个显著相关的 QTL 区域,包含与果实重量相关的 QTL 4 个,与果皮颜色、果肉硬度和囊硬度相关的 QTL 各 1 个^[49]。孙珍珠利用 240 份宽皮柑橘与 112 768 个 SNPs 进行 GWAS,以果形指数、果皮光滑度、花药颜色等 10 个重要性状为表型指标,共鉴定到 35 个显著性 SNP 标记,其中与花药颜色性状相关的显著性位点最多,共 10 个。此外,还找到了 55 个候选基因,分布在 NW006262339.1 和 NW006262022.1 等 2 条染色体上。这些 QTL 的发现和候选基因的发掘有望应用于柑橘杂交种的培育中^[50]。

3.4 浆果类果树

分析文献报道发现,浆果类果树的全基因组关联分析主要集中于葡萄的研究中,其他浆果类果树的 GWAS 处于起步阶段。鲜食葡萄育种的目标主要包括有香气、自然大粒、耐贮运^[51],因而育种家在品种选育过程中更多地关注果实形态、果实重量、果实开裂程度等果实重要农艺性状。研究表明浆

果果形性状可能受多基因协同调节,基因 *VIT_12s0134g00230* 和 *VIT_02s0025g01360* 通过调节植物激素影响果实形态,泛素连接酶基因 *VIT_03s0088g01090* 和 *VIT_10s0003g04300* 与多种浆果形状相关^[52]。Guo 等对 179 份葡萄种质进行了全基因组测序,并将 32 311 个 SNP 与果实发育期、果实重量、果实形状等 8 个性状进行 GWAS,筛选出多个显著性标记。其中,在 2 号染色体上检测到与葡萄果皮颜色相关的 *mybA* 转录因子,以及在 16 号染色体上检测到 2 个与果实发育期相关的显著性标记^[15]。这些与浆果性状密切相关的标记有助于葡萄育种。裂果是果实生长发育过程中常见的现象,严重影响果实产量和品质。Zhang 等通过 2019 年和 2020 年连续调查,对浆果开裂率、浆果开裂指数和浆果开裂类型进行关联分析,鉴定到 5 个与浆果开裂指数相关的 SNP,2 个与浆果开裂类型相关的 SNP,这些位点分布于 4 条染色体上,与 10 个候选裂果基因相关^[53]。

蓝莓和树莓作为新兴水果,相较于其他果树历史较短,参考基因组公布时间也比较晚,全基因组关联分析相关研究仍处于起步阶段。在南高丛蓝莓品种中,Ferrão 等利用源自 92 个南高丛组合的 886 份杂交后代为试验材料,以 10 种脂肪酸衍生物、5 种萜类化合物、2 种苯类化合物为表性指标,利用 GWAS 分析成功鉴定到蓝莓上有 519 个显著性 SNPs 与属于不同代谢途径的 11 个挥发性有机化合物(volatile organic compounds, VOCs)相关。他们证明了 VOCs 受几个主要的基因组区域控制,其中一些区域含有生物合成酶编码基因,部分 VOCs 调节消费者偏好,为育种提供了方向^[54]。为简化生产管理难度,树莓品种的无刺性状也备受关注。Khadgi 和 Weber 以 92 个树莓样本为材料进行测序,鉴定了 4 个极显著 SNP,其中 SNP 4_35148226 被鉴定为对无刺性状有积极贡献,为研究树莓无刺性状基因组区域奠定了基础^[55]。

草莓枯萎病(*Fusarium wilt*)是一种由真菌引起的突然传播性疾病,严重影响全球草莓生产。Pincot 等通过对 565 份重要草莓种质的 14 408 个 SNP 进行全基因组关联分析,在 2 号染色体的 2.3 Mb 区间内检测到与草莓枯萎病基因 *Fw1* 连锁不平衡的 14 个 SNP,最显著 SNP 解释了 85% 的表型变异,预测了 97% 的抗性,广义遗传力为 0.96^[56]。

综上,研究发现的 SNP 位点有望在品种改良中

发挥重要作用,为推进通过分子标记辅助育种和基因选择来提高核果类果树的品质、抗病性等奠定了基础。

3.5 其他经济类果树

除上述大宗类果树外,在甜瓜、枣、橄榄等经济类果树中也逐步开展了全基因组关联分析研究。其中在甜瓜中,研究者基于 48 个商业东方甜瓜品种的叶片、叶柄、花序、果实等 48 个表型进行了 GWAS,新发现了 2 个分别位于 1 号染色体和 8 号染色体上的 OM1_30331998 和 OM8_81597532 个 SNP 位点与性别表达相关,并预测到 4 个潜在候选基因 *MELO3C015898*、*MELO3C015904*、*MELO3C024563* 和 *MELO3C024565* 与东方甜瓜的性别表达相关^[57],这些结果为研究甜瓜性别表达的分子遗传机制提供了基础。在枣中,研究人员基于 180 份中国枣的 9 个品质性状开展关联分析,成功鉴定出 45 个与果实大小、果核大小和果实开裂相关的 SNP 标记,并通过基因功能注释,鉴定出了 21 个候选基因,这些基因参与细胞扩增、非生物胁迫反应、激素信号和生长发育等过程^[58]。在橄榄中,Kaya 等对来自美国和土耳其共 183 份橄榄材料,利用 GWAS 技术探究了叶片长度、叶片宽度、果实重量、果核重量和果肉果核比 5 个性状与 SNP 的相关性,共发现 52 个显著性标记与叶片长度、果实重量、果核重量和果肉果核比相关联,预测到 19 个候选基因,大部分的基因注释是参与生理发育过程的蛋白质^[59]。

综上,通过全基因组关联分析,可以找到与目标性状密切相关的位点,挖掘基因功能,从而在解析果树品质形成机制、遗传驯化规律、病害发生机理及鉴定种质资源、分析基因调控网络等方面发挥重要作用。上述研究为理解核果类果树李痘病、褐腐病等重要病害的遗传机制及品质等农艺性状候选基因的筛选研究提供了重要基础,为推动分子标记辅助选择(marker-assisted selection, MAS)育种的进程提供了重要参考。

4 GWAS 的发展应用

GWAS 作为定位性状相关位点、基因的有效工具,在植物中最早被应用于糖料作物海甜菜的研究^[60],距今已有 20 多年,随着对植物基因组认识的不断深入,GWAS 相关的衍生技术也已被广泛应用,为研究提供了更多的思路 and 选择。目前在植物中应用比较广泛的 GWAS 衍生技术主要集中于对基

因型数据进行改变,以及关联分析和其他方法的结合运用 2 方面,从而更好地研究性状的遗传变异趋势,弥补 GWAS 的不足。以下综述结构变异(structural variant,SV)、转座子(transposable element,TE)及转录组关联分析(TWAS)与 GWAS 联合运用的研究进展。

结构变异(SV)包括插入、缺失、重复和倒位^[61],最初被定义为超过 1 kb 大小的插入、缺失和反转,但随着人类基因组测序的常规化,结构变异已经包括长度大于 50 bp 的事件^[62]。随着泛基因组研究的增多,基于 SV 开展 GWAS,已成功应用于大豆^[63]、玉米^[65]、油菜^[66]、桃^[67-69]、水稻^[70-71]、番茄^[72]等植物性状的研究中,证明与传统的 SNP 相比,SV 具有更大的表型效应,基因或调控序列中出现的 SV 基因可能会影响表达,进而影响农艺性状的变化。

转座子(TE)活性通过转位、插入、切除、染色体断裂和异位重组产生基因和基因组结构的变化,通常伴随着基因活性的改变^[73]。有研究证明,使用 TE 标记进行 GWAS 比使用 SNP 标记连锁不平衡的假阳性低,基因组扫描显示了与农艺性状相关的 TE 阳性选择^[74]。在番茄中用 TE 插入多态性(TE insertions polymorphisms, TIP)进行关联分析,确定了至少 40 个与主要农艺性状或次级代谢产物的极端变异密切相关的 TIP,大部分 TE 插入都是低频变体,很少被 SNP 标记^[75]。玉米 *ZmCCT* 基因编码含有 CCT 结构域的光周期相关蛋白质,通过 GWAS 在启动子中检测到 CACAT 样转座子,显著缩短开花时间,TE 插入发生在驯化之后,在相关的连锁不平衡区块内产生了强烈的选择性清除^[76]。

全转录组关联分析(transcriptome-wide association study, TWAS)是基于 GWAS 的统计分析,将基因表达量与表型性状相关联,识别与复杂性状相关联的顺式调节表达基因^[77],补充研究 GWAS 定位到的基因与关联性状背后的生物学机制^[78]。GWAS + TWAS 联合运用的方法已经在多种植物中得到应用,包括玉米^[79-80]、棉花^[81]、水稻^[82]、油菜^[83-84]等,首先通过 GWAS 与功能基因组学相结合鉴定调节基因,其次利用 TWAS 研究调节基因的表达量与复杂性状之间的关联,对鉴定到的基因进行贡献率排序,从而更全面地了解基因调控机制和复杂性状的遗传基础^[85]。

综上,利用 SV、TE 开展关联分析以及 GWAS +

TWAS 等衍生应用在扩大表型效应、降低连锁不平衡的假阳性、关联基因表达量与复杂性状的遗传基础等方面具有优势,弥补了 GWAS 单独分析的不足。

5 展望

GWAS 是一种挖掘功能基因/位点的有效研究手段,其优势在于不需要任何先验信息来衡量遗传变异和表型多样性之间的关联便可以挖掘种质资源中尚未发现的有价值的遗传变异,为今后的育种工作提供新的遗传变异,为实现品种改良,开发具有理想性状的新品种奠定了前期基础。

纵观水稻、玉米和大豆等大田作物,除了鉴定显著性位点,更多的是关注性状的遗传调控机制,挖掘功能基因所处的信号调节通路。如利用大豆 GWAS 得到的显著性关联区域构建遗传调控网络,发现有部分性状可能共同遗传,受同一个显著性区域调控,为一因多效作用的结果^[86];在水稻研究中发现, *OsGSK2* 基因是参与油菜素甾体信号(brassinosteroids, BR)传导的保守激酶,通过协调独脚金内酯和油菜素甾醇调节中胚轴长^[87]; *MRG702* 基因、粒宽基因 *GW5* 在籼稻、粳稻 2 个栽培稻群体中受选择,参与 BR 信号通路的调控^[88];对 8 个植物代表性性状进行 PCA,利用 PC1 作为因变量进行 GWAS 研究,发现 *SPINDLY* 基因被鉴定为通过抑制赤霉素(GA)信号传导来调节水稻结构^[89]。与之相比在果树中的研究应用还相对滞后,可借鉴、参考其思路和方法。

果树开展 GWAS 研究的瓶颈因素包括以下内容:(1)样本数量和种质资源:GWAS 研究需要大量的样本和种质资源,而果树因保存资源受试验地面积影响,多数野生资源、地方品种以原生态就地保存为主,分布广泛,集中获取难度大。(2)基因组数据:GWAS 研究需要高质量完整的基因组数据进行分析,但是果树基因组数据的获取、分析和注释仍存在困难,这也制约了 GWAS 研究的深入开展。(3)环境因素的影响:果树对于环境的适应性很强,而环境因素对果树性状的影响比其他作物更为复杂多变,环境因素可能会掩盖或干扰 GWAS 的结果。(4)遗传背景和多倍体物种分析:果树多数为多倍体物种,造成基因型分析和遗传背景建立存在难度,同时多倍体物种还需要考虑亲代效应、杂种优势、杂合子丧失等问题。这些因素都增加了果树 GWAS 研究的复杂性。针对这些问题,需要通过多

种手段解决,如建设多样化的果树种质资源库,改进数据测序技术和生物信息学分析方法,开展多个作物种类的 GWAS 研究以提高 GWAS 的可靠性和推广应用。

综上所述,随着苹果^[90]、桃^[91]、梨^[92]、蓝莓^[93]等果树基因组的公布,在果实品质性状评估、抗性基因鉴定等方面发挥了重要作用。此外,果树上 GWAS 研究目前相对集中在核果类和仁果类果树中,而对于树莓、蓝莓等树种的 GWAS 研究领域仍有大量空白需要填补。因此,在今后的研究中,开展结构变异、转座子层面的关联分析,以及将 GWAS 研究与 QTL 定位、转录组、基因选择、基因编辑、功能基因解析等多种方法结合在一起,将更有利于实现基因组学辅助育种和基于基因组学的技术对果树进行遗传改良。

致谢:感谢辽宁省果树科学研究所刘硕副研究员对本文的修改和提出的宝贵意见。

参考文献:

- [1] 陈学森,王楠,张宗营,等. 关于果树种质资源与遗传育种若干问题的理解与思考[J]. 中国农业科学,2022,55(17):3395–3410.
- [2] 王力荣,吴金龙. 中国果树种质资源研究与新品种选育 70 年[J]. 园艺学报,2021,48(4):749–758.
- [3] Zahid G, Aka Kaçar Y, Dönmez D, et al. Perspectives and recent progress of genome – wide association studies (GWAS) in fruits[J]. Molecular Biology Reports,2022,49(6):5341–5352.
- [4] 赵惠,房玉林,张颖,等. 关联分析及其在果树育种研究中的应用[J]. 分子植物育种,2018,16(7):2291–2299.
- [5] 周海平,张帆,陈凯,等. 水稻种质资源稻瘟病抗性全基因组关联分析[J]. 作物学报,2023,49(5):1170–1183.
- [6] 徐鑫,张德华,赵吉顺,等. 普通小麦小穗粒数性状全基因组关联分析[J]. 植物遗传资源学报,2022,23(4):1098–1110.
- [7] 李婷,董远,张君,等. 玉米杂交种穗部性状的全基因组关联分析[J]. 中国农业科学,2022,55(13):2485–2499.
- [8] 唐玉凤,姚敏,何昕,等. 甘蓝型油菜 SGR 基因家族的全基因组鉴定与功能分析[J]. 作物学报,2023,49(7):1829–1842.
- [9] 王娟,马晓梅,周小凤,等. 棉花产量构成因素性状的全基因组关联分析[J]. 中国农业科学,2022,55(12):2265–2277.
- [10] 梁腾月,谷勇哲,马英杰,等. 大豆耐低磷性全基因组关联分析[J]. 植物遗传资源学报,2023,24(1):237–251.
- [11] 彭芳芳,龙治坚,魏召新,等. 樱桃种质 SCot 分子标记与叶片表型性状关联分析[J]. 园艺学报,2021,48(2):325–335.
- [12] 李雄伟,贾惠娟,高中山. 桃基因组及全基因组关联分析研究进展[J]. 遗传,2013,35(10):1167–1178.
- [13] 储宝华,曹富国,卞宁宁,等. 84 个苹果栽培品种对斑点落叶病的抗性评价和全基因组关联分析[J]. 中国农业科学,2022,55(18):3613–3628.
- [14] 胡哲辉,刘园,王江波,等. 3 个品种梨香气感官品质与挥发性物质关联分析[J]. 华中农业大学学报,2022,41(4):217–225.
- [15] Guo D L, Zhao H L, Li Q, et al. Genome – wide association study of berry – related traits in grape (*Vitis vinifera* L.) based on genotyping – by – sequencing markers[J]. Horticulture Research, 2019,6(1):1655–1667.
- [16] Minamikawa M F, Nonaka K, Kaminuma E, et al. Genome – wide association study and genomic prediction in citrus: potential of genomics – assisted breeding for fruit quality traits[J]. Scientific Reports,2017,7(1):1–13.
- [17] Flint – Garcia S A, Thornsberry J M, Th B E. Structure of linkage disequilibrium in plants. [J]. Annual Review of Plant Biology, 2003,54(1):357–374.
- [18] Flint – Garcia S A, Thornsberry J M, Buckler Iv E S. Structure of linkage disequilibrium in plants [J]. Annual Review of Plant Biology,2003,54(1):357–374.
- [19] Abdallah J M, Goffinet B, Cierco – Ayrolles C, et al. Linkage disequilibrium fine mapping of quantitative trait loci: a simulation study[J]. Genetics Selection Evolution,2003,35(5):513–532.
- [20] Alqudah A M, Sallam A, Baenziger P S, et al. GWAS: fast – forwarding gene identification and characterization in temperate cereals: lessons from barley – a review[J]. Journal of Advanced Research,2020,22:119–135.
- [21] Hayes B. Overview of statistical methods for genome – wide association studies (GWAS) [J]. Methods in Molecular Biology, 2013,1019:149–169.
- [22] Aoun M, Carter A H, Ward B P, et al. Genome – wide association mapping of the ‘Super – Soft’ kernel texture in white winter wheat [J]. Theoretical and Applied Genetics,2021,134:2547–2559.
- [23] 赵宇慧,李秀秀,陈倬,等. 生物信息学分析方法 I : 全基因组关联分析概述[J]. 植物学报,2020,55(6):715–732.
- [24] Pritchard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data [J]. Genetics,2000,155(2):945–959.
- [25] Yu J, Pressoir G, Briggs W H, et al. A unified mixed – model method for association mapping that accounts for multiple levels of relatedness[J]. Nature Genetics,2006,38(2):203–208.
- [26] 丛佩华,张彩,韩晓,等. 我国苹果育种研究现状及展望[J]. 中国果树,2018(6):1–5.
- [27] 王森,何平,王海波,等. 鲁丽苹果在不同产地果实风味及品质差异分析[J]. 果树学报,2023,40(6):1135–1145.
- [28] Larsen B, Migicovsky Z, Jeppesen A A, et al. Genome – wide association studies in apple reveal loci for aroma volatiles, sugar composition, and harvest date [J]. The Plant Genome,2019,12(2):180104.
- [29] Noh J, Do Y S, Kim G H, et al. A genome – wide association study for the detection of genes related to apple *Marssonina* blotch disease resistance in apples[J]. Scientia Horticulturae,2020,262:108986.
- [30] Liao P, Chen Q F, Chye M L. Transgenic *Arabidopsis* flowers overexpressing acyl – CoA – binding protein ACBP6 are freezing

- tolerant[J]. *Plant and Cell Physiology*,2014,55(6):1055–1071.
- [31] Kerstiens G. Water transport in plant cuticles; an update[J]. *Journal of Experimental Botany*,2006,57(11):2493–2499.
- [32] Kunst L,Samuels A L. Biosynthesis and secretion of plant cuticular wax[J]. *Progress in Lipid Research*,2003,42(1):51–80.
- [33] Cao F, Li Z, Jiang L, et al. Genome – wide association study (GWAS) of leaf wax components of apple[J]. *Stress Biology*, 2021,1(1):1–14.
- [34] 陈学森,王楠,张宗营,等. 仁果类果树资源育种研究进展 I: 我国梨种质资源、品质发育及遗传育种研究进展[J]. *植物遗传资源学报*,2019,20(4):791–800.
- [35] 张绍铃. 当前我国梨产业发展面临的重大问题和对策措施[J]. *中国果业信息*,2016,33(12):12–14.
- [36] Zhang M Y,Xue C,Hu H,et al. Genome – wide association studies provide insights into the genetic determination of fruit traits of pear[J]. *Nature Communications*,2021,12(1):1144.
- [37] 尹明华,吴平华,刘邦旺,等. 上饶早梨全基因组重测序和转录组的关联分析[J]. *分子植物育种*,2019,17(11):3490–3496.
- [38] Minamikawa M F,Takada N,Terakami S, et al. Genome – wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia* Nakai)[J]. *Scientific Reports*,2018,8(1):1–12.
- [39] 姜全. 当前我国桃产业发展面临的重大问题和对策措施[J]. *中国果业信息*,2017,4(1):5–6,10.
- [40] Cao K,Zhou Z,Wang Q,et al. Genome – wide association study of 12 agronomic traits in peach[J]. *Nature Communications*,2016,7(1):13246.
- [41] Wei H,Chen X,Zong X,et al. Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in the red and yellow fruits of sweet cherry (*Prunus avium* L.)[J]. *PLoS One*,2015,10(3):e0121164.
- [42] Holušová K,Čmejlová J,Suran P,et al. High – resolution genome – wide association study of a large Czech collection of sweet cherry (*Prunus avium* L.) on fruit maturity and quality traits[J]. *Horticulture Research*,2023,10(1):uhac233.
- [43] Mariette S, Wong Jun Tai F, Roch G, et al. Genome – wide association links candidate genes to resistance to plum pox virus in apricot (*Prunus armeniaca*) [J]. *New Phytologist*,2016,209(2):773–784.
- [44] Zuriaga E,Soriano J M,Zhebentyayeva T, et al. Genomic analysis reveals MATH gene(s) as candidate(s) for plum pox virus (PPV) resistance in apricot (*Prunus armeniaca* L.) [J]. *Molecular Plant Pathology*,2013,14(7):663–677.
- [45] Gupta R, Huang Y, Kieber J, et al. Identification of a dual – specificity protein phosphatase that inactivates a MAP kinase from *Arabidopsis*[J]. *The Plant Journal*,1998,16(5):581–589.
- [46] Fu W, da Silva Linge C,Gasic K. Genome – wide association study of brown rot (*Monilinia* spp.) tolerance in peach[J]. *Frontiers in Plant Science*,2021,12:635914.
- [47] 邓秀新. 中国柑橘育种 60 年回顾与展望[J]. *园艺学报*,2022,49(10):2063–2074.
- [48] Minamikawa M F,Nonaka K,Kaminuma E,et al. Genome – wide association study and genomic prediction in citrus: potential of genomics – assisted breeding for fruit quality traits[J]. *Scientific Reports*,2017,7(1):4721.
- [49] Imai A,Nonaka K,Kuniga T, et al. Genome – wide association mapping of fruit – quality traits using genotyping – by – sequencing approach in citrus landraces,modern cultivars,and breeding lines in Japan[J]. *Tree Genetics & Genomes*,2018,14(2):1–15.
- [50] 孙珍珠. 宽皮柑橘种质资源的多样性研究及重要农艺性状的全基因组关联分析[D]. 重庆:西南大学,2018:53–54.
- [51] 段长青. 当前我国葡萄产业发展面临的重大问题和对策措施[J]. *中国果业信息*,2017,34(1):3–4.
- [52] Zhang C,Cui L,Fang J. Genome – wide association study of the candidate genes for grape berry shape – related traits[J]. *BMC Plant Biology*,2022,22(1):42.
- [53] Zhang C,Wu J,Cui L,et al. Mining of candidate genes for grape berry cracking using a genome – wide association study[J]. *Journal of Integrative Agriculture*,2022,21(8):2291–2304.
- [54] Ferrão L F V,Johnson T S,Benevenuto J, et al. Genome – wide association of volatiles reveals candidate loci for blueberry flavor[J]. *New Phytologist*,2020,226(6):1725–1737.
- [55] Khadgi A,Weber C A. Genome – wide association study (GWAS) for examining the genomics controlling prickly production in red raspberry (*Rubus idaeus* L.) [J]. *Agronomy*,2020,11(1):27.
- [56] Pincot D D A,Poorten T J,Hardigan M A, et al Genome – wide association mapping uncovers *Fu1*, a dominant gene conferring resistance to *Fusarium* wilt in strawberry[J]. *G3:Genes,Genomes,Genetics*,2018,8(5):1817–1828.
- [57] Kishor D S,Noh Y,Song W H, et al. SNP marker assay and candidate gene identification for sex expression via genotyping – by – sequencing – based genome – wide associations (GWAS) analyses in oriental melon (*Cucumis melo* L. var. *makuwa*) [J]. *Scientia Horticulturae*,2021,276:109711.
- [58] Hou L,Chen W,Zhang Z,et al. Genome – wide association studies of fruit quality traits in jujube germplasm collections using genotyping – by – sequencing[J]. *The Plant Genome*,2020,13(3):e20036.
- [59] Kaya H B,Akdemir D,Lozano R, et al. Genome wide association study of 5 agronomic traits in olive (*Olea europaea* L.) [J]. *Scientific Reports*,2019,9(1):1–14.
- [60] Hansen M,Kraft T,Ganestam S, et al. Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers[J]. *Genetics Research*,2001,77(1):61–66.
- [61] Abel H J,Larson D E,Regier A A, et al. Mapping and characterization of structural variation in 17,795 human genomes[J]. *Nature*,2020,583(7814):83–89.
- [62] Alkan C,Coe B P,Eichler E E. Genome structural variation discovery and genotyping[J]. *Nature Reviews Genetics*,2011,12(5):363–376.
- [63] Liu Y,Du H,Li P,et al. Pan – genome of wild and cultivated soybeans[J]. *Cell*,2020,182(1):162–176.

- [64] Yang N, Liu J, Gao Q, et al. Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement[J]. *Nature Genetics*, 2019, 51(6): 1052 – 1059.
- [65] Sun S, Zhou Y, Chen J, et al. Extensive intraspecific gene order and gene structural variations between *Moi7* and other maize genomes [J]. *Nature Genetics*, 2018, 50(9): 1289 – 1295.
- [66] Song J M, Guan Z, Hu J, et al. Eight high – quality genomes reveal pan – genome architecture and ecotype differentiation of *Brassica napus* [J]. *Nature Plants*, 2020, 6(1): 34 – 45.
- [67] Guan J, Xu Y, Yu Y, et al. Genome structure variation analyses of peach reveal population dynamics and a 1.67 Mb causal inversion for fruit shape[J]. *Genome Biology*, 2021, 22(1): 1 – 25.
- [68] Zhou H, Ma R, Gao L, et al. A 1.7 – Mb chromosomal inversion downstream of a *PpOFPI* gene is responsible for flat fruit shape in peach[J]. *Plant Biotechnology Journal*, 2021, 19(1): 192 – 205.
- [69] Guo J, Cao K, Deng C, et al. An integrated peach genome structural variation map uncovers genes associated with fruit traits [J]. *Genome Biology*, 2020, 21: 1 – 19.
- [70] Wang W, Mauleon R, Hu Z, et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice [J]. *Nature*, 2018, 557(7703): 43 – 49.
- [71] Hu Z, Wang W, Wu Z, et al. Novel sequences, structural variations and gene presence variations of Asian cultivated rice [J]. *Scientific Data*, 2018, 5(1): 1 – 7.
- [72] Alonge M, Wang X, Benoit M, et al. Major impacts of widespread structural variation on gene expression and crop improvement in tomato [J]. *Cell*, 2020, 182(1): 145 – 161. e23.
- [73] Bennetzen J L. Transposable element contributions to plant gene and genome evolution [J]. *Plant Molecular Biology*, 2000, 42(1): 251 – 269.
- [74] Yan H, Haak D C, Li S, et al. Exploring transposable element – based markers to identify allelic variations underlying agronomic traits in rice [J]. *Plant Communications*, 2022, 3(3): 100270.
- [75] Domínguez M, Dugas E, Benchouaia M, et al. The impact of transposable elements on tomato diversity [J]. *Nature Communications*, 2020, 11(1): 4058.
- [76] Yang Q, Li Z, Li W, et al. CACTA – like transposable element in *ZmCCT* attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize [J]. *Proceedings of the National Academy of Sciences*, 2013, 110(42): 16969 – 16974.
- [77] Gusev A, Ko A, Shi H, et al. Integrative approaches for large – scale transcriptome – wide association studies [J]. *Nature Genetics*, 2016, 48(3): 245 – 252.
- [78] Gamazon E R, Wheeler H E, Shah K P, et al. A gene – based association method for mapping traits using reference transcriptome data [J]. *Nature Genetics*, 2015, 47(9): 1091 – 1098.
- [79] Wu D, Li X, Tanaka R, et al. Combining GWAS and TWAS to identify candidate causal genes for tocopherol levels in maize grain [J]. *Genetics*, 2022, 221(4): iyac091.
- [80] Lin M, Qiao P, Matschi S, et al. Integrating GWAS and TWAS to elucidate the genetic architecture of maize leaf cuticular conductance [J]. *Plant Physiology*, 2022, 189(4): 2144 – 2158.
- [81] Ma Y, Min L, Wang J, et al. Combined transcriptome GWAS and TWAS reveal genetic elements leading to male sterility during high temperature stress in cotton [J]. *New Phytologist*, 2021, 231(1): 165 – 181.
- [82] Anacleto R, Badoni S, Parween S, et al. Integrating a genome – wide association study with a large – scale transcriptome analysis to predict genetic regions influencing the glycaemic index and texture in rice [J]. *Plant Biotechnology Journal*, 2019, 17(7): 1261 – 1275.
- [83] Tan Z, Xie Z, Dai L, et al. Genome – and transcriptome – wide association studies reveal the genetic basis and the breeding history of seed glucosinolate content in *Brassica napus* [J]. *Plant Biotechnology Journal*, 2022, 20(1): 211 – 225.
- [84] Zhang Y, Zhang H, Zhao H, et al. Multi – omics analysis dissects the genetic architecture of seed coat content in *Brassica napus* [J]. *Genome Biology*, 2022, 23(1): 86.
- [85] Li B, Ritchie M D. From GWAS to gene: transcriptome – wide association studies and other methods to functionally understand GWAS discoveries [J]. *Frontiers in Genetics*, 2021, 12: 713230.
- [86] Fang C, Ma Y, Wu S, et al. Genome – wide association studies dissect the genetic networks underlying agronomical traits in soybean [J]. *Genome Biology*, 2017, 18(1): 1 – 14.
- [87] Sun S, Wang T, Wang L, et al. Natural selection of a *GSK3* determines rice mesocotyl domestication by coordinating strigolactone and brassinosteroid signaling [J]. *Nature Communications*, 2018, 9(1): 2523.
- [88] 李嘉琦. 基于 GWAS 解析水稻粒型变异的遗传基础 [D]. 沈阳: 沈阳师范大学, 2020: 46 – 50.
- [89] Yano K, Morinaka Y, Wang F, et al. GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture [J]. *Proceedings of the National Academy of Sciences*, 2019, 116(42): 21262 – 21267.
- [90] Velasco R, Zharkikh A, Affourtit J, et al. The genome of the domesticated apple (*Malus × domestica* Borkh.) [J]. *Nature Genetics*, 2010, 42(10): 833 – 839.
- [91] International Peach Genome Initiative, Verde I, Abbott Ag, et al. The high – quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution [J]. *Nature Genetics*, 2013, 45(5): 487 – 494.
- [92] Wu J, Wang Z, Shi Z, et al. The genome of the pear (*Pyrus bretschneideri* Rehd.) [J]. *Genome Research*, 2013, 23(2): 396 – 408.
- [93] Colle M, Leisner C P, Wai C M, et al. Haplotype – phased genome and evolution of phytonutrient pathways of tetraploid blueberry [J]. *Gigascience*, 2019, 8: giz012.