

岳庆春,傅迦得,章辰飞,等. 植物关联分析应用研究进展[J]. 江苏农业科学,2019,47(18):24-30.

doi:10.15889/j.issn.1002-1302.2019.18.004

植物关联分析应用研究进展

岳庆春¹, 傅迦得², 章辰飞², 吴月燕²

(1. 水产科学国家级实验教学示范中心/水产动物遗传育种中心上海市协同创新中心/上海海洋大学, 上海 201306;

2. 浙江万里学院生物与环境学院, 浙江宁波 315100)

摘要:随着分子生物学和基因组学的快速发展,关联分析成为近些年来在植物数量性状研究和植物良种选育中行之有效的分析方法。利用关联分析在分子水平阐明植物表型性状的遗传变异规律和机制,从而为植物的农艺性状改良及新品种选育提供新思路。系统详实综述了关联分析基本原理、关联作图的基本策略、关联分析的应用以及各种分子标记在关联分析中的应用,并讨论了关联分析在未来研究中的发展前景。

关键词:关联分析;植物;表型性状;遗传变异;农艺性状改良;新品种选育;连锁不平衡;研究发展趋势;应用前景

中图分类号: S184 **文献标志码:** A **文章编号:** 1002-1302(2019)18-0024-06

关联分析(association analysis)也称连锁不平衡作图(LD mapping)或者关联作图(association mapping),该方法通常以自然群体为研究对象,以连锁不平衡(linkage disequilibrium, LD)为基础,将目的性状表型的多样性以及遗传标记或候选基因的多态性联结起来分析,鉴定某一群体内目的性状与遗传标记或候选基因之间的关系^[1]。从而在分子水平解释植物表型性状的遗传变异规律和机制,为植物表型性状的标记辅助选择以及目的基因的分离、检测、利用提供依据,进而为植物性状遗传改良研究提供理论基础,为植物杂交育种和性状改良寻求新途径^[2]。

到目前为止,关联分析已经在部分植物性状研究中取得进展,如玉米的开花期^[3]、小麦的籽粒大小和研磨品质^[4]、水稻的柱头^[5]、葡萄的果穗长度^[6]等,关联分析已经成为当前植物遗传育种研究的热点。

本文经过系统全面的介绍关联分析基本原理以及分析策略,详细论述关联分析在目前植物遗传学研究中的应用进展以及各类分子标记技术在关联分析中的应用,探讨关联分析在今后的研究发展趋势和在植物遗传研究中的应用前景。

1 关联分析基本原理

1.1 连锁不平衡

关联分析是以连锁不平衡(linkage disequilibrium, LD)为基础,也可称为配子相不平衡(gametic phase disequilibrium)、配子不平衡(gametic disequilibrium)、等位基因关联(allelic association)等,是指群体内不同座位等位基因(可以是标记,也可以是基因/QTL间与标记)间的非随机关联^[7]。也就是说假设2个不同位点的等位基因一同出现的频率比理论上同

时出现频率高时,那么称这2个位点处于连锁不平衡状态^[8]。LD的基本定义式为 $D_{ij} = f_{ij} - PA_i \cdot PB_j$,其中 f_{ij} 是 $A_i B_j$ 基因型的频率, PA_i 和 PB_j 分别是等位基因 A_i 和 B_j 的频率。由于 D_{ij} 可以假定的最大值是所观察到的等位基因频率的函数,因此对于双等位基因和多等位基因,LD的强度有多种标准化度量,其中2种最常见的LD强度测量方法是:(1)单个LD值的标准化度量, $D_{ij}' = D_{ij}/D_{\max}$; (2)双等位基因数据的相关系数 r ,常用定义为 $r^2 = D_{ij}^2 / (PA_1 \cdot PA_2 \cdot PB_1 \cdot PB_2)$ ^[9]。同一染色体或者不同染色体的基因座之间均可出现连锁不平衡状态,群体内存在的LD都是由突变造成的等位基因出现后座位间所有重组响应累积的结果,位点间连锁越紧密,其LD程度越高^[10]。

1.2 影响连锁不平衡的因素

遗传因素和非遗传因素综合作用影响群体的LD水平^[11]。一般情况下,在随机匹配群体里没有突变、迁移或选择因素的影响时,多态性位点则处于连锁平衡状态;与此相反,连锁、群体混合和选择将增加LD水平^[12]。影响LD程度最重要的2个要素是突变和重组,突变是造成LD的一个重要因素,新突变的发生可冲破原有LD,进而导致新的多态性产生;然而重组则是经过重新组合序列变异,进而减弱染色体内部的LD。无连锁和自由交配的重组使位点间等位基因处于连锁平衡状态,因此LD的水平与重组率成反比^[10]。群体中的LD是突变、重组和其他因素影响共同累积的结果^[13]。

此外,其他非生物要素和生物要素也影响LD程度,例如物种之间的交配体系、染色体位置、群体大小以及自然与人工选择^[10]。基因转换或染色体片段所受的选择强度、遗传漂变^[14-15]等也是影响LD水平的因素。

1.3 连锁不平衡与关联分析

在自然群体中,表型差异的根本原因主要是个体等位基因间的差异。连锁分析则是采用标记位点与引起表型差异位点之间的重组来定位数量性状基因座(quantitative trait locus, QTL),而关联分析利用引起表型差异的位点与标记之间的LD来定位QTL^[10]。因此,进行关联分析的前提和基础是了解群体基因组LD的构造和规律。往往因为群体的基因组中

收稿日期:2018-06-01

基金项目:浙江省宁波市重大科技专项(编号:2015C110016)。

作者简介:岳庆春(1994—),男,甘肃定西人,硕士研究生,研究方向为果树种质资源开发与利用。E-mail: yueqingchun118@163.com。

通信作者:吴月燕,硕士,教授,研究方向为果树种质资源。E-mail: wyybn2009@163.com。

存在数目巨大的多态性,因此多态位点的等位基因间存在广泛的非随机关联,亦称为 LD 状态。多个基因座等位基因间的连锁不平衡结构会产生一系列的单倍型,单倍型的大小则受 LD 衰减程度的影响。不同物种的连锁不平衡衰减距离不同,同一物种不同群体、同一群体不同座位的 LD 衰减距离也不同^[16]。染色体上不同位置的连锁不平衡程度也不相同,研究发现位于着丝粒附近片段的重组率比较低,LD 水平则较高;然而染色体臂上的片段区域重组率相对较高,相应 LD 水平则较低^[17]。连锁不平衡的衰减程度越高,则形成的单倍型越小^[10]。

1.4 关联分析与传统连锁分析的差异

关联分析与传统连锁分析相比具有以下优势:(1)关联分析不必构建专门作图群体,而是运用自然群体的遗传多样性,将复杂的性状变异进行分解。利用关联分析构建的群体不须要管制研究对象的交配方式,而传统的连锁分析以父母本杂交产生的子代群体为研究对象。相比而言,关联分析可应用的种质材料更加广泛。(2)关联分析所研究的材料有较为宽泛的遗传基础,因此可同时针对同一基因座的多个等位基因进行检测分析,相比绝大部分传统连锁分析,其所研究群体通常为 2 个亲本杂交重组的后代,所以基因座一般只触及 2 个等位基因。关于具备更小效应的基因,关联分析的发掘能力显著高于传统连锁分析^[13]。(3)关联分析定位更精准,能够抵达单基因程度,由于关联分析应用在长期进化进程中自然群体所积累的重组信息,因而可到达更高的分辨率,从而达到对 QTL 的精准定位,甚至可直接定位到基因本身^[10]。而传统连锁分析往往受到重组发生率的影响,进而导致分辨率较低,一般认为初级群体只能将 QTL 定位到 10 ~ 20 cM 的基因组区间内,而次级群体可达到单基因水平^[18-19]。(4)运用的统计分析方法不同,传统的连锁作图措施包含了单标记分析、区间作图、复合区间作图以及贝叶斯区间作图^[13]。与此相比,适用于关联分析作图的统计方法较为匮乏。

2 关联分析的基本策略

2.1 基于全基因组扫描的关联分析

全基因组关联分析 (genome-wide association study, GWAS) 是采用自然变异群体,联合高密度分子标记图谱进行扫描,进而分析表型性状与分子标记之间关联关系的有效方法,现已发展成为发掘复杂农艺性状遗传变异的有效手段^[20]。在以全基因组扫描为基础的关联分析中,须要用散布于全基因组的高通量分子标记对某物种大群体的全部基因进行同时检测^[8]。GWAS 以群体中 LD 水平为基础,借助成百上千的个体组成的定位群体,采用一定数量的 SNP 标记构建的高密度遗传图谱,从而与表型数据进行关联分析。近年来,基于 GWAS 技术已在多种植物表型研究中取得一定的进展。代力强等以 80 份玉米核心自交系为关联作图群体,通过全基因组测序,筛选出 16 个与玉米粒长紧密关联的显著性 SNP 标记和 3 个候选基因^[21]。刘静利用高密度的小麦 90K 单核苷酸多态性 (SNP) 芯片对西南麦区 192 份小麦品种进行株高性状的全基因组关联分析,发现 57 个与株高显著相关的 SNP 位点^[22]。Feng 等利用全基因组测序的 472 份油菜种质,在染色体 A03、A05、A07 和 C07 上鉴定出 8 个 QTL 与株高显著相关,在染色体 A01、A03、A07 和 C07 上的 5 个 QTL 被鉴定为与主枝

数显著相关^[23]。目前,基于全基因组关联分析已在各类植物物种深入研究,但在园艺植物中应用报道较为匮乏。

GWAS 一般采用 5 步进行:(1)关联群体的选择。应选择遗传变异丰富、表型差异较大、遗传基础较宽泛且应尽量包含某物种全部的遗传变异。(2)样本基因分型。基于常用的分子标记主要包括 RFLP、AFLP、SSR 及 SNP 等,随着全基因组测序技术的不断发展,SNP 标记方法得到广泛运用。除了使用基因芯片进行基因分型以外,还可直接重新测序获得研究样本个体的基因型,进而更加全面地挖掘样本基因组变异^[20]。(3)群体构造与个体亲缘关系分析。GWAS 通常以自然变异群体为研究对象,存在一定的遗传结构,其个体间也存在一定的亲缘关系,因而有可能导致染色体间的 LD 水平提高,使得目标性状与不相关的标记产生伪关联。因此,检测分析并矫正种质材料的群体结构有一定的必要。(4)目标性状的鉴定。目标性状评价的准确性对于关联分析的结果有重要影响,应反复对种质材料进行多重表型分析鉴定。(5)关联统计分析模型的选择。随着生物统计学的不断发展,关联统计分析模型不断得到完善,主要包含一般线性模型 (GLM) 和混合线性模型 (MLM),通常可利用 TASSEL 软件或 ANOVA 计算方法进行关联分析^[24]。

随着第 3 代测序技术即单分子测序技术的发展,植物中主要物种全基因组测序逐步完成,物种的基因组信息越来越丰富,进而开发出大量的 SNP 标记。全基因组关联分析将成为今后植物数量性状研究的有利工具^[25]。

2.2 基于候选基因的关联分析

基于候选基因关联分析主要针对目标 QTL 区段内候选基因进行生物信息学分析,推定其生物学功能是否与数量性状表型位于同一调控网络,或是辅以生理生化分析,从而快速确定 QTL 区间内的候选基因,最终只针对筛选后的少数候选基因开展关联分析^[26]。早在 2001 年,Thornberry 等初次将关联分析方法引入植物领域研究^[27]。根据前人研究发现,dwarf8 基因是一个与赤霉素的代谢相关且显著影响玉米株高的基因^[28],而后 Thornberry 等选用 92 份玉米自交系种质对 dwarf8 基因的多态性进行验证,研究表明 dwarf8 基因不仅影响玉米的株高,而且首次发现其中几个多态性位点与玉米开花期的变异性状显著相关^[27]。此项研究发现意味着基于连锁不平衡的关联分析可能是进行基因功能验证以及基因发掘的一种行之有效的办法,为植物表型性状研究提供了新思路^[16]。近年来,基于候选基因的关联分析已经成功应用于部分植物研究。Yu 等以 295 份水稻材料在苗期进行水稻耐盐相关表型的全基因组关联研究,获得了 93 个候选基因,其中有 6 个与耐盐表型具有高关联^[29]。Perez 等以 315 份不同高粱材料,利用候选基因关联作图的方法,检测油菜素内酯生物合成和信号传导基因与植物结构性状之前的标记-性状关联,共检测出 26 个油菜素内酯基因的 73 个 SNPs 与目标表型显著相关^[30]。于永涛利用 94 份玉米自交系验证了 rab17 基因与玉米籽粒产量相关联^[31]。Andersen 等利用 SNP 分子标记验证了 PAL 基因与玉米饲用品质间相互关联^[32]。刘翠霞以 150 份葡萄杂交后代为试验材料,联合 RNA-seq 技术初步筛选了 32 个候选基因参与单萜代谢^[33]。国内外研究结果均表明,基于候选基因关联分析是一个进行鉴定候选基因功

能的强有力的工具。

2.3 关联分析中的假阳性及其消除

在进行关联分析时,群体结构往往是最先考虑的一个重要影响要素,因为它往往会导致假阳性结果出现。为此研究者通常选择生长区域较远的种质资源作为研究对象,目的是增加群体材料之间的遗传多样性^[34]。除此之外,群体中的连锁不平衡还受到遗传漂变、自然选择、群体分层以及奠基效应等诸多因素影响,进而使得一些非关联的等位基因与目标性状基因座形成连锁不平衡,从而导致表现出关联性,这种现象称为伪关联或者假阳性^[35]。研究者们为了减轻或者消除关联分析时形成的假阳性,开发了诸多试验方法来控制,例如结构关联法^[36]、结构关联(Q) + 亲缘关系(K)混合模型法^[37]、传递不平衡法(transmission disequilibrium test, TDT)^[38]、基因组对照法(genome control)^[39]、主成分分析法(principal component analysis, PCA)^[40-41]、多维标度法(multidimensional scaling, MDS)^[42]和非计量多维标度法(nonmetric multidimensional scaling, NMDS)^[43]等。Stich 等研究认为, Q + K 混合模型法是目前解决关联分析时减轻假阳性可能的最好方法,可有效控制假阳性的发生^[44]。随着统计学研究的不断发展,今后研究者将会开发出更加高效、准确的减轻或消除假阳性的分析方法。

关联分析作图的基本步骤主要包括种质的选择、群体结构的估计、表型性状的观测、候选基因位点多样性的鉴定以及统计分析。随着关联分析在动植物遗传研究中的运用日益广泛,更加准确和便捷的统计方法以及分析软件不断地被开发出来,都为关联分析研究提供了空前的便利。冯建英等对目前关联分析研究中常用的统计方法和软件做了系统的调查,对复杂性状关联分析方法及其适用软件做了详细的阐述^[45]。常用的关联作图软件有 PLINK 软件^[46]、TASSEL 软件^[47]、QTXNetwork 软件^[48]、mrMLM 软件^[49]等,应用者可根据所研究材料不同,选择合适的统计方法和软件。

3 关联分析在植物研究中的应用

3.1 关联分析与植物 QTL 定位的研究

到目前为止,关联分析已在多种植物研究中广泛应用。在这些研究中,常应用连锁不平衡作图达到挖掘和定位与目标性状相互关联的 QTL 以及等位基因。杨晓旭以杂交育种获得的 149 份葡萄材料为试验群体,结合 SSR 分子标记和 SRAP 分子标记技术,最终构建欧亚种葡萄的分子遗传图谱,进而对葡萄果实中主要香气成分进行 QTL 定位分析,发现 DXS 基因与萜类物质代谢路径相关^[50]。武玉国等采用 SSR 分子标记对黄淮麦区 175 份小麦种质进行研究,结果表明标记 wmc128(1B)和 wmc236(3B)与小穗数极显著相关^[51]。金亮利用 QTL 定位结合关联分析 2 种方法对稻米淀粉品质以及营养品质表型性状进行分子标记,结果验证了 Wx 和 SSI(淀粉酶 1)基因与淀粉特性相关^[52]。郝德荣通过鉴定大豆基因组中与产量性状紧密关联的功能位点,结合已知的大豆基因组序列信息进行了关联定位,最终在染色体 7 上筛选到与大豆产量性状紧密相关的候选基因 *GmGA3ox*^[53]。赖勇等采用 86 个 SSR 标记对 113 份大麦亲本材料进行遗传多样性分析以及农艺性状关联分析,结果发现了与株高、芒长和小

穗着生密度相关联的 6 个标记,这些标记分别位于 1H、2H、3H、4H 和 7H 染色体上^[54]。申聪聪利用多亲本高代互交系、群体及其复合群体作为遗传定位群体,结合高通量数据进行关联分析定位,挖掘影响水稻抽穗期和株高的 QTL,共检测出 3 个影响抽穗期的主效 QTL,即 qHD3、qHD6 和 qHD8,分别位于第 3、第 6 和第 8 号染色体上^[55]。近年来,越来越多的研究证实了关联分析可有效地鉴定 QTL,充分说明关联分析在植物遗传学研究中扮演着不可或缺的重要角色。

3.2 关联分析与基因功能的发掘与验证

基于关联分析中候选基因法与基于全基因组扫描分析对比而言,前者所需工作量小、标记数量较少、成本低,还可对目的基因进行功能鉴定,是植物遗传学中基因功能鉴定中的常用方法^[56]。赵曦以 95 份玉米核心自交系为试验材料,构建关联作图群体,采用关联分析方法共检测得到 7 个多态性位点与表型性状显著相关,结果发现,3 个 SNP 均为 C/G 突变,其中 1 个含有碱基 G 的等位基因可能与叶绿素含量增加有关,另外 2 个 3 bp 片段插入的等位基因可能与雄穗长度的增加有关,1 bp 片段插入的等位基因可能与穗轴粗增加有关,1 个 9 bp 的片段缺失的等位基因可能与粒长增加有关^[57]。杨胜先等研究了利用覆盖大豆基因组的 135 对 SSR 引物对六大生态区的 257 份栽培大豆种质进行农艺性状关联分析,发掘了一批农艺性状的优异等位基因,其中 satt669 - 145bp、satt102 - 154bp、satt382 - 395bp 和 satt534 - 161bp 分别是主茎节数、分枝数、茎粗和株高的增效效应最大的等位基因^[58]。孙晓棠等利用苗期微室接种鉴定法对 456 份水稻材料组成的自然群体进行纹枯病抗性关联分析,结果发现 RM1036、RM5371 和 RM7585 标记位点位于连锁定位的抗纹枯病 QTL 附近^[59]。运用关联分析进一步挖掘和验证基因功能,成为阐明候选基因或基因位点与目标性状之间相关性的重要方法。

3.3 关联分析与功能性标记的开发

2003 年,Andersen 等初次提出功能性标记的概念,其方法是从影响性状变异基因的功能区域来开发出多态性标记^[60]。功能标记的开发必须符合 2 个条件:(1)已有确定功能的候选基因且等位基因的序列信息也是已知的;(2)对多个材料中的目标性状进行分析,进而对目的基因进行序列分析,联结性状以及基因序列信息进行基于连锁不平衡的关联分析^[25]。胡明建选取调控水稻千粒质量的 *TGW6* 为候选基因,分析其等位变异与千粒质量间的关系,随后开发功能标记并在重组自交系群体和自然群体中进行分析 and 验证,试验结果开发了 CAPS 标记 *TGW7*,最终鉴定出 2 种等位基因(*TaTGW-7Aa* 和 *TaTGW-7Ab*)与粒质量相关^[61]。张照贵克隆研究了 *TaSnRK2.10* 基因为候选基因,在小麦中进行单倍型分析以及功能标记开发与农艺性状相关关联分析,根据 SNP 开发的 CAPS 功能标记,最终将该标记定位在 4A 染色体的 D - 1092101、D - 100014232 标记之间,遗传距离分别为 0.8、0.5 cM^[62]。Kumar 等利用 74 份玉米自交系研究筛选了 4 个候选基因 *Rtcl*、*Rth3*、*Rum1* 和 *Rull*,结果发现,4 个候选基因均与玉米幼苗根系性状相关^[63]。

运用关联分析与功能基因标记进行分子标记辅助育种(molecular marker-assisted selection, MAS),有以下 2 个方面的优势:(1)由于选择的基因就是基因本身,因而可以保证选择

的准确性,从而提高选择效率;(2)能够采用关联分析针对定向的影响性状变异的功能区域进行特定选择,从而保障选择效果。关联分析联合功能基因标记将是今后分子辅助育种发展的目标。采用功能性分子标记的办法进行 MAS 以及植物种质鉴定,对加速育种世代选择、缩短育种年限、保护植物知识产权具有极其重要的意义^[56]。

3.4 关联分析与数量性状的研究

植物中许多重要的表型性状如产量、生理特征、营养品质以及抗逆性等多属于数量性状,往往采用传统的常规育种时效率不高,基于目前的科技发展,采用分子育种技术育种将是未来的发展方向。关联分析作为一种高效的 QTL 鉴定工具,未来可在分子育种中起到一定的重要作用^[10]。连锁分析以及关联分析均在数量性状的研究中起关键作用,由于它们对数量性状基因座定位的精确和宽泛、提供的信息量丰富以及统计方法准确等方面具备显著的互补效应,因此研究者通常可先利用连锁分析初步定位控制目标表型性状等位基因的位置,进而使用关联分析则可快速对目标基因进行精细定位,且根据特定候选基因提供丰富信息,从而验证候选基因功能。因此,应当联结连锁分析和关联分析各自的优点,对数量性状开展纵向和横向的分析,以达到加快数量性状基因的鉴定、分离以及克隆,最终为深入了解数量性状的遗传学和分子生物学基础,更重要的是为植物数量性状的遗传改良提供新的契机^[16]。

4 各类分子标记在关联分析中的应用

4.1 SRAP 分子标记

相关序列扩增多态性(sequence-related amplified polymorphism, SRAP)标记是基于 PCR 技术的第 2 代分子标记,目前已广泛运用于植物的遗传多样性分析、遗传连锁图谱的构建、测序以及基因克隆等方面^[64]。李仁伟利用 SRAP 分子标记对 58 份大菊种质的数量性状进行关联研究,结果表明有 6 个标记位点与花部性状显著相关^[65]。丁泽婷等采用 20 份柱花草种质进行表型性状与 SRAP 标记关联分析,结果发现有 15 个 SRAP 标记与鲜质量相关联,1 个 SRAP 标记与株高显著相关^[66]。羊杏平等利用 46 对 SRAP 引物组合对 35 份西瓜核心种质进行关联分析,将枯萎病抗性与 SRAP 多态性标记进行回归分析,检测到 1 个与抗病性显著关联的 SRAP 位点^[67]。随着第 3 代分子标记的发展,基于 SRAP 分子标记的关联分析研究应用越来越少。

4.2 SSR 分子标记

简单序列重复(simple sequence repeats, SSR)原理为基因组中一般由 1~6 个核苷酸组成的基本单位反复多次构成的一段 DNA,其广泛分布在基因组中的不同位置。谭文丽采用 80 对 SSR 引物对 149 份木薯种质进行全基因组多态性位点扫描,并通过关联分析方法,结果表明,有 151 个多态性位点与 21 个性状相关联,最终发现 m224-c 位点与株高极显著关联^[68]。吴静等利用经过筛选得到的 11 对 SSR 标记对 99 份紫斑牡丹种质进行多态性扫描,进而分析其遗传多样性和群体结构,采用一般线性模型进行标记与表型性状关联分析,结果发现 5 个标记位点与 6 个性状显著关联^[69]。潘磊等利用 156 对多态性 SSR 引物对 83 份豇豆种质材料进行产量性状与分子标记全基因组关联分析,结果表明,有 10 个 SSR 标

记位点与 8 个表型性状关联,主要散布在 LG2、LG3、LG4、LG7、LG11 连锁群上^[70]。Yu 等以 99 份多年生黑麦草种质为材料,利用 109 对 SSR 标记进行关联分析,发现其中 15 个 SSR 标记与淹水胁迫下叶色、叶绿素荧光、最大株高以及相对生长速率的降低显著相关^[71]。近年来,利用 SSR 分子标记进行关联分析以及挖掘基因功能位点的研究越来越多,由此可见,关联分析对于找寻与植物表型性状紧密连锁的 SSR 标记有显著作用,从而可有效应用于分子标记辅助育种。

4.3 SNP 分子标记

植物基因组中的核苷酸多态性主要有 3 种,包括插入、缺失以及单核苷酸多态性(single nucleotide polymorphism, SNP)。其主要工作原理是在基因组水平上,由单个核苷酸碱基转换或颠倒而引起的 DNA 序列多态性,又因为基因组 DNA 的任何碱基均有可能发生变异,因此单核苷酸多态性几乎遍布整个基因组,SNP 是植物中普遍存在的一种分子标记^[72]。Zhou 等利用 533 份水稻种质中的 650 万个 SNP 进行柱头外露和相关花性状的全基因组关联分析,共鉴定了 23 个与柱头外露和相关花性状显著相关的基因组位点^[5]。Wang 等对 144 份玉米自交系种质中的 45 868 个 SNP 位点进行关联分析,结果发现,18 个 SNP 与玉米中黑穗病抗性显著相关^[73]。刘翔攀采用 238 份玉米自交系种质材料,利用关联分析找到相应的 SNP,最终发现与单株产量降幅相关的 SNP 有 15 个,分别位于 2、3、4、5、6、7、8 号染色体上^[74]。蒋俊利用 593 份杂交桃后代为试验材料,对亲本及杂交群体进行 SNP 位点检测,从而构建遗传连锁图谱以及针对简单性状的全基因组关联分析,最后得到与果实酸度关联的 SNP 标记,为 SNPIGA545261^[75]。许理文等采用 240 份 DH 系为关联作图群体,利用 SNP 芯片进行基因型分析,构建连锁图谱,对 DH 群体进行容重性状鉴定,结果表明,共检测到 5 个控制玉米容重的 QTL^[76]。随着分子标记技术的快速发展,SNP 标记技术在遗传多样性、遗传基础分析、针对重要性状的基因定位、进行分子辅助育种与 SNP 标记开发、种质鉴定等植物育种方面的研究进展中起到了不可替代的作用^[77]。

4.4 SCoT 分子标记

目标起始密码子多态性分子标记(start codon targeted polymorphism, SCoT)则是在 2009 年由 Collard 和 Mackill 共同开发的一种新型分子标记方法^[78],该方法具备操作简单、可有效产生与性状连锁的标记、重复性好、引物设计简单并具有通用性等一系列优点。Rahimi 等利用 24 个 SCoT 引物对 39 份车前草种质进行全基因组关联作图,使用 Q 和 K 因子的 MLM 模型进行关联分析,结果得到 16 个与穗质量、株高、穗长以及千粒质量等性状相关的 SCoT 标记^[79]。Yan 等使用 18 个 EST-SSR 和 21 个 SCoT 标记检测了 75 份果园草种质中的锈病性状,利用 TASSEL 软件进行关联分析,发现了 20 个标记与锈病性状显著相关^[80]。杨旭旭利用 59 份大菊种质,通过测量其表型性状,进而结合 SCoT 分子标记进行关联分析,共发现 23 个标记位点与 7 个性状关联,其中有 18 个位点与花部性状显著相关^[81]。袁王俊等采用 10 条 SCoT 引物对 99 份菊花种质的基因组变异进行扫描,从而分析种质的群体结构,运用一般线性模型和混合线性模型方法对 10 个表型性状进行关联分析,共发现 3 个 SCoT 位点与花序直径相关

联^[82]。韩国辉利用 20 条 SCoT 引物对 92 株 Wan2 橘橙 × Li2 甜橙杂交群体进行多样性分析,最终得到 30 个 SCoT 标记被定位在已构建的 9 个连锁群上^[83]。SCoT 作为一种新型的目的基因分子标记,目前已在部分植物中得到应用,包括种质资源遗传多样性与亲缘关系分析、种质鉴定与指纹图谱等,但绝大多数只停留在对研究材料 PCR 体系的优化,而在基因差异表达和高密度分子遗传连锁图谱以及重要性状精细定位和功能基因克隆等方面应用相对较少^[84]。

5 展望

植物关联分析正处在快速发展的时期,随着关联分析方法的日益完善和更加便利准确的分析软件不断开发,特别是高通量测序技术以及基因芯片技术研究的不断发展,基于全基因组扫描的关联分析方法将会帮助研究者在更多的植物物种中、更大的范围内完成关联分析。到目前为止,关联分析只在少数经济作物如油菜、大豆、油菜等中研究较为广泛,而在一些园艺植物中,应用关联分析进行功能基因定位研究报道极其匮乏,尤其基于候选基因关联研究只扫描到植物基因组的一小部分,且基于候选基因关联作图的成功的研究成果很少。因此,具有高密度的 SNP 覆盖度、大的样本、小的群体结构的关联分析在复杂性状的分解上具有非常重要的发展前景^[13]。值得关注的是,关联分析对于遗传多样性偏低的群体作图效果不如传统 QTL 作图,因此对于一些植物复杂数量性状的解析中仍然不能撇弃传统的连锁作图。在研究中应当将关联分析和传统 QTL 作图优势互补,可先用连锁分析来对 QTL 进行初步定位,然后采用关联作图来对其进行精细定位, QTL 作图与关联分析的综合利用将会是未来遗传连锁分析的发展方向。伴随着植物基因组学的飞速发展和不断完善,基于关联分析挖掘新的功能基因、开发新标记以及定位优异性状将成为未来植物种质资源研究以及良种选育的重要方法和手段。

参考文献:

- [1] Flint S A, Thornsberry J M, Buckler E S. Structure of linkage disequilibrium in plants[J]. Annual Review of Plant Biology, 2003, 54:357–374.
- [2] 彭 波. SSR 标记与桑树表型性状的关联分析[C]. 中南五省区蚕桑育种协作组, 国家蚕桑产业技术体系遗传育种研究室. 中南五省区蚕桑育种协作研讨会论文集, 2009:5.
- [3] Salvi S, Sponza G, Morgante M, et al. Conserved noncoding genomic sequences associated with a flowering – time quantitative trait locus in maize[J]. Proc Natl Acad Sci USA, 2007, 104:11376–11381.
- [4] Breseghello F, Sorrells M E. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars [J]. Genetics, 2006, 172:1165–1177.
- [5] Zhou H, Li P B, Xie W B, et al. Genome – wide association analyses reveal the genetic basis of stigma exertion in rice [J]. Molecular Plant, 2017, 10(4):634–644.
- [6] Vafaei Y, Ghaderi N, Khadivi A. Morphological variation and marker – fruit trait associations in a collection of grape (*Vitis vinifera* L.) [J]. Scientia Horticulturae, 2017, 225:771–782.
- [7] Gaut B S, Long A D. The lowdown on linkage disequilibrium[J]. Plant Cell, 2003, 15(7):1502–1506.
- [8] 王荣煊, 王天宇, 黎 裕. 关联分析在作物种质资源分子评价中的应用[J]. 植物遗传资源学报, 2007, 8(3):366–372.
- [9] Single R M, Strayer N, Thomson G, et al. Martin asymmetric linkage disequilibrium: tools for assessing multiallelic LD [J]. Human Immunology, 2016, 77(3):288–294.
- [10] 谭贤杰, 吴子恺, 程伟东, 等. 关联分析及其在植物遗传学研究中的应用[J]. 植物学报, 2011, 46(1):108–118.
- [11] 丁 霞, 王林海, 张艳欣, 等. 芝麻核心种质株高构成相关性状的遗传变异及关联定位[J]. 中国油料作物学报, 2013, 35(3):262–270.
- [12] 蒋思霞, 倪正斌, 印志同, 等. 糯玉米自交系遗传多样性及其产量、农艺性状与 SSR 分子标记的关联研究[J]. 安徽农业科学, 2012, 40(6):3212–3217, 3283.
- [13] 王彩洁, 徐 冉, 于明飞. 关联分析在植物育种中的应用现状[J]. 山东农业科学, 2012, 44(6):32–38, 62.
- [14] Gupta P K, Rustgi S, Kulwal P L. Linkage disequilibrium and association in higher plants: present status and future prospects[J]. Plant Mol Biol, 2005, 57:461–485.
- [15] Oraguzie N C, Wilcox P L, Rikkerink E H A, et al. Linkage disequilibrium[M]// Oraguzie N C, Rikkerink E H A, Gardiner S E, et al. Association mapping in Plants. New York: Springer Verlag, 2007:11–39.
- [16] 杨小红, 严建兵, 郑艳萍, 等. 植物数量性状关联分析研究进展[J]. 作物学报, 2007, 33(4):523–530.
- [17] 倪正斌. 糯玉米自交系遗传多样性及产量、农艺性状与 SSR 分子标记的关联研究[D]. 扬州: 扬州大学, 2010.
- [18] Doerge R W. Mapping and analysis of quantitative trait loci in experimental populations[J]. Nat Rev Genet, 2002, 3:43–52.
- [19] Holland J B. Genetic architecture of complex traits in plants[J]. Curr Opin Plant Biol, 2007, 10:156–161.
- [20] 鲁 清. 水稻种质资源重要农艺性状的全基因组关联分析[D]. 北京: 中国农业科学院, 2016.
- [21] 代力强, 吴 律, 董青松, 等. 玉米籽粒长度的全基因组关联分析[J]. 西北农林科技大学学报(自然科学版), 2018, 46(6):1–8.
- [22] 刘 静. 小麦株高性状的全基因组关联分析[C]. 第八届全国小麦基因组学及分子育种大会摘要集, 2017:1.
- [23] Feng L, Chen B Y, Xu K, et al. A genome – wide association study of plant height and primary branch number in rapeseed (*Brassica napus*) [J]. Plant Science, 2016, 242:169–177.
- [24] 李海权, 刁现民. 关联分析及其在植物研究中的应用[J]. 河北农业科学, 2010, 14(11):157–160.
- [25] 于海霞, 肖 静, 田迎春, 等. 关联分析及其在植物中的应用[J]. 基因组学与应用生物学, 2009, 28(1):187–194.
- [26] 霍冬敏. 玉米穗粒数相关性状 QTL 定位与候选基因关联分析[D]. 武汉: 华中农业大学, 2016.
- [27] Thornsberry J M, Goodman M M, Doebley J, et al. Dwarf8 polymorphisms associate with variation in flowering time[J]. Nature Genetics, 2001, 28(3):286–289.
- [28] Fujioka S, Yamane H, Spray C R, et al. The dominant non – gibberellin – responding dwarf mutant (D8) of maize accumulates native gibberellins [J]. Proceedings of the National Academy of Sciences of the United States of America, 1988, 85(23):9031–9035.
- [29] Yu X, Bai G H, Luo N, et al. Association of simple sequence repeat

- (SSR) markers with submergence tolerance in diverse populations of perennial ryegrass[J]. Plant Science, 2011, 180(2): 391–398.
- [30] Perez M B, Zhao J, Yin Y H, et al. Association mapping of brassinosteroid candidate genes and plant architecture in a diverse panel of *Sorghum bicolor* [J]. Theoretical and Applied Genetics, 2014, 127(12): 2645–2662.
- [31] 于永涛. 玉米核心自交系群体结构及耐旱相关候选基因 *rab17* 的等位基因多样性分析[D]. 北京: 中国农业科学院研究生院, 2006.
- [32] Andersen J R, Zein I, Wenzel G, et al. High levels of linkage disequilibrium and associations with forage quality at a phenylalanine ammonia-lyase locus in European maize (*Zea mays* L.) inbreds [J]. Theoretical and Applied Genetics, 2007, 114(2): 307–319.
- [33] 刘翠霞. 葡萄果实单萜化合物含量的 QTL 定位及其合成调控的候选基因筛选[D]. 武汉: 中国科学院武汉植物园, 2017.
- [34] 段继凤. 甘蓝型油菜含油量性状全基因组关联分析[D]. 武汉: 华中农业大学, 2013.
- [35] Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results [J]. Nat Genet, 1995, 11: 241–247.
- [36] Pritchard J K, Rosenberg N A. Use of unlinked genetic markers to detect population stratification in association studies [J]. Am J Hum Genet, 1999, 65: 220–228.
- [37] Yu J M, Pressoir G, Briggs W H, et al. A unified mixed model method for association mapping that accounts for multiple levels of relatedness [J]. Nat Genet, 2006, 38: 203–208.
- [38] Spielman R S, McGinnis R E, Ewens W J. Transmission test for linkage disequilibrium; the insulin gene region and insulin-dependent diabetes mellitus (IDDM) [J]. American Journal of Human Genetics, 1993, 52(3): 506–516.
- [39] Devlin B, Roeder K. Genomic control for association studies [J]. Biometrics, 1999, 55: 997–1004.
- [40] Patterson N, Price A L, Reich D. Population structure and eigenanalysis [J]. PLoS Genetics, 2006, 2(12): 2074–2093.
- [41] Price A L, Patterson N J, Plenge R M, et al. Principal components analysis corrects for stratification in genome-wide association studies [J]. Nat Genet, 2006, 38: 904–909.
- [42] Li Q Z, Yu K. Improved correction for population stratification in genome-wide association studies by identifying hidden population structures [J]. Genetic Epidemiology, 2008, 32(3): 215–226.
- [43] Zhu C, Yu J. Nonmetric multidimensional scaling corrects for population structure in whole genome association studies [J]. Genetics, 2009, 182: 875–888.
- [44] Stich B, Melchinger A E. Comparison of mixed model approaches for association mapping in rapeseed, potato, sugar beet, maize, and *Arabidopsis* [J]. BMC Genomics, 2009, 10: 94–94.
- [45] 冯建英, 温阳俊, 张瑾, 等. 植物关联分析方法的研究进展 [J]. 作物学报, 2016, 42(7): 945–956.
- [46] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses [J]. American Journal of Human Genetics, 2007, 81(3): 559–575.
- [47] Bradbury P J, Zhang Z, Kroon D E, et al. TASSEL: software for association mapping of complex traits in diverse samples [J]. Bioinformatics, 2007, 23(19): 2633–2635.
- [48] 张福涛. 遗传分析方法的 GPU 并行计算与优化研究[D]. 杭州: 浙江大学, 2014.
- [49] Wang S B, Feng J Y, Ren W L, et al. Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology [J]. Scientific Reports, 2016, 6: 19444.
- [50] 杨晓旭. 葡萄果实香气物质的 QTL 定位及关联分析研究[D]. 沈阳: 沈阳农业大学, 2017.
- [51] 武玉国, 吴承来, 秦保平, 等. 黄淮冬麦区 175 个小麦品种的遗传多样性及 SSR 标记与株高和产量相关性状的关联分析 [J]. 作物学报, 2012, 38(6): 1018–1028.
- [52] 金亮. 水稻关联定位群体的构建及若干品质性状的关联分析 [D]. 杭州: 浙江大学, 2009.
- [53] 郝德荣. 大豆产量相关性状 QTL 的关联分析及候选基因 *GmGA3ox* 单倍型鉴定[D]. 南京: 南京农业大学, 2011.
- [54] 赖勇, 王鹏喜, 范贵强, 等. 大麦 SSR 标记遗传多样性及其与农艺性状关联分析 [J]. 中国农业科学, 2013, 46(2): 233–242.
- [55] 申聪聪. 利用水稻 MAGIC 群体和种质资源关联分析定位抽穗期和株高 QTL [D]. 北京: 中国农业科学院, 2017.
- [56] 王云云. 关联分析在植物遗传学中的应用 [J]. 黑龙江科学, 2016, 7(11): 1–3.
- [57] 赵曦. 玉米耐旱相关候选基因 *dhn2* 与表型性状的关联分析 [D]. 北京: 中国农业科学院, 2008.
- [58] 杨胜先, 牛远, 李梦, 等. 栽培大豆农艺性状的关联分析及优异等位变异挖掘 [J]. 中国农业科学, 2014, 47(20): 3941–3952.
- [59] 孙晓棠, 卢冬冬, 欧阳林娟, 等. 水稻纹枯病抗性关联分析及抗性等位变异发掘 [J]. 作物学报, 2014, 40(5): 779–787.
- [60] Andersen J R, Lübberstedt T. Functional markers in plants [J]. Trends Plant Sci, 2003, 8(11): 554–560.
- [61] 胡明建. 小麦粒重相关基因 *TaTGW6*、*TaTGW-7A* 克隆及其功能标记开发 [D]. 合肥: 安徽农业大学, 2016.
- [62] 张照贵. 小麦 *TaSnRK2.10* 基因的克隆、标记开发和功能分析 [D]. 泰安: 山东农业大学, 2014.
- [63] Kumar B, Abdel-Ghani A H, Pace J A, et al. Association analysis of single nucleotide polymorphisms in candidate genes with root traits in maize (*Zea mays* L.) seedlings [J]. Plant Science, 2014, 224: 9–19.
- [64] 张飞, 房伟民, 陈发棣, 等. 切花菊花器性状的遗传变异与相关性研究 [J]. 浙江林学院学报, 2008(3): 293–297.
- [65] 李仁伟. 大菊品种遗传多样性及基于 SRAP 标记的分类研究 [D]. 北京: 北京林业大学, 2011.
- [66] 丁泽婷, 尹晓畅, 唐燕琼, 等. 桂花草重要性状与 SSR、SRAP 分子标记的关联分析 [J]. 分子植物育种, 2017, 15(5): 1839–1845.
- [67] 羊杏平, 刘广, 侯喜林, 等. 西瓜核心种质枯萎病抗性与 SRAP 分子标记的关联分析 [J]. 园艺学报, 2013, 40(7): 1298–1308.
- [68] 谭文丽. 木薯 SSR 标记与表型的关联分析 [D]. 海口: 海南大学, 2014.
- [69] 吴静, 成仿云, 庞利铮, 等. 紫斑牡丹表型性状与 SSR 分子标记的关联分析 [J]. 北京林业大学学报, 2016, 38(8): 80–87.
- [70] 潘磊, 李依, 余晓露, 等. 豇豆产量性状与 SSR 分子标记的关联分析 [J]. 湖北农业科学, 2015, 54(16): 3952–3958, 3962.
- [71] Yu J, Zao W G, He Q, et al. Genome-wide association study and gene set analysis for understanding candidate genes involved in salt tolerance at the rice seedling stage [J]. Molecular Genetics and Genomics, 2017, 292(6): 1391–1403.

帅 良,孙 健,段振华,等. 植物非特异性磷脂酶 C 的研究进展[J]. 江苏农业科学,2019,47(18):30-37.
doi:10.15889/j.issn.1002-1302.2019.18.005

植物非特异性磷脂酶 C 的研究进展

帅 良^{1,2}, 孙 健¹, 段振华², 李 丽¹, 何雪梅¹, 李昌宝¹, 廖玲燕²

(1. 广西壮族自治区农业科学院/广西作物遗传改良生物技术重点开放实验室/广西果蔬贮藏与加工新技术重点实验室
培育基地,广西南宁 530007; 2. 贺州学院食品与生物工程学院/食品科学与工程技术研究院,广西贺州 542899)

摘要:非特异性磷脂酶 C (non-specific PLC, NPC) 能够水解磷脂胆碱和磷脂酰乙醇胺生成二酰基甘油 (diacylglycerol, DAG), 虽然非特异性磷脂酶 C 没有 PLC 家族基因的 C2、X、Y、EF 等结构域,但它仍含有 1 个磷酸酯酶结构域,能够水解磷脂。近年来研究发现,非特异性磷脂酶 C 在植物逆境胁迫和信号转导的过程中起着重要的调节作用。因此本文对非特异性磷脂酶 C 的结构、生理功能及其作用机制进行概述,并对其研究前景进行了展望。

关键词:非特异性磷脂酶 C; 二酰基甘油; 逆境胁迫; 信号转导

中图分类号:S184 **文献标志码:**A **文章编号:**1002-1302(2019)18-0030-08

磷脂酶 (phospholipases) 是一类能够水解磷脂的酶类,根据其水解磷脂位点的不同可将其分为磷脂酶 A1 (phospholipases A1, PLA1)、磷脂酶 A2 (PLA2)、磷脂酶 C

收稿日期:2018-06-11

基金项目:公益性行业(农业)科研专项(编号:201303073);国家自然科学基金(编号:31560467、31160407、31660589);2014 年国家中组部“万人计划”青年拔尖人才项目(编号:组厅字[2015]48 号);“八桂学者”工程专项经费(编号:[2016]21);广西创新驱动发展专项资金(编号:桂科 AA17204038、桂科 AA17204042);广西农业科学院博士后基金(编号:桂农科博 2018029);广西特聘专家专项经费(编号:厅发[2016]21 号);广西果蔬保鲜与深加工人才小开放课题(编号:2016XGDSHF02);现代食品加工新技术研究岗位创新人才培养示范基地建设(编号:桂科 AD17195088)。

作者简介:帅 良(1986—),男,江西新干人,博士,讲师,主要从事农产品贮藏技术研究。Tel:(0771)3240232;E-mail:shuailiang1212@163.com。

通信作者:孙 健,博士,研究员,主要从事农产品贮藏与加工技术研究。Tel:(0771)3240692;E-mail:jiansun@gsaas.net。

(PLC)和磷脂酶 D(PLD)^[1],它们均在植物生长发育、生物胁迫及非生物胁迫中起重要的调控作用。其中 PLC 介导的细胞信号通路是信号转导中最经典的通路之一,它能够水解磷脂生成 1 分子的二酰基甘油 (diacylglycerol, DAG) 和 1 个含有磷酸基团的小分子^[2]。植物中的 PLC 根据水解底物和功能的不同可以分为 3 类:第 1 类为能够水解磷脂胆碱 (phosphatidylcholine, PC) 和磷脂酰乙醇胺 (phosphatidylethanolamine, PE) 并生成 DAG 的非特异性磷脂酶 C (non-specific PLC, NPC);第 2 类为能够水解 4,5 二磷酸肌醇 (phosphatidylinositol 4,5-bisphosphate, PIP₂) 生成三磷酸肌醇 (inositol 1,4,5-trisphosphate, IP₃) 和 DAG 的磷酸肌醇特异性磷脂酶 C (phosphoinositide-specific PLC, PI-PLC);第 3 类为能够水解连接在内质网上的糖基磷脂酰肌醇锚蛋白的糖基磷脂酰肌醇特异性磷脂酶 C (glycosylphosphatidylinositol PLC, GPI-PLC)^[3]。NPC 最先在产气荚膜梭菌中被发现^[4]。1955 年,在植物质体中发现了类似 NPC 的活性^[5]。目前为止,已经从拟南芥^[6]、水稻^[7]、谷子^[8]、陆地棉^[9]等植物中克隆出 NPC 基因成员。近年来,大

[72]任凤阳. 玉米自交系萌发特性的评价及 SNP 关联分析[D]. 泰安:山东农业大学,2014.

[73]Wang M, Yan J B, Zhao J R, et al. Genome-wide association study (GWAS) of resistance to head smut in maize[J]. Plant Science, 2012, 196: 125-131.

[74]刘翔攀. 玉米自交系耐密性评价及 SNP 关联分析[D]. 泰安:山东农业大学,2017.

[75]蒋 俊. 桃高密度 SNP 图谱构建于果实质地和酸度全基因组关联分析[D]. 杭州:浙江大学,2016.

[76]许理文,段民孝,田红丽,等. 基于 SNP 标记的玉米容重 QTL 分析[J]. 玉米科学,2015,23(5):21-25.

[77]宁 洽,刘文国,杨伟光,等. SNP 标记在玉米研究上的应用进展[J]. 玉米科学,2017,25(1):57-61.

[78]Collard B C Y, Mackill D J. Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants[J]. Plant Molecular Biology

Reporter, 2008, 27(1):86-93.

[79]Rahimi M, Nazari L, Kordrostami M, et al. Scot marker diversity among Iranian Plantago ecotypes and their possible association with agronomic traits[J]. Scientia Horticulturae, 2018, 233:302-309.

[80]Yan H D, Zhang Y, Zeng B, et al. Genetic diversity and association of EST-SSR and SCoT markers with rust traits in orchardgrass (*Dactylis glomerata* L.) [J]. Molecules, 2016, 21(1):66.

[81]杨旭旭. 菊花品种表型性状与 SCoT 分子标记的关联分析[D]. 开封:河南大学,2015.

[82]袁王俊,叶 松,董美芳,等. 菊花品种表型性状与 SSR 和 SCoT 分子标记的关联分析[J]. 园艺学报,2017,44(2):364-372.

[83]韩国辉. 基于 EST-SSR、Genomic-SSR 和 SCoT 标记的柑橘连锁图谱构建及杂种和多倍体遗传分析[D]. 重庆:西南大学,2012.

[84]龙治坚,范理璋,徐 刚,等. SCoT 分子标记在植物研究中的应用进展[J]. 植物遗传资源学报,2015,16(2):336-343.