

功能标记的开发、特点和应用研究进展

贺道华^a,雷忠萍^b,邢宏宜^a

(西北农林科技大学 a. 农学院, b. 生命科学学院, 陕西 杨凌 712100)

【摘要】【目的】从基于分子标记研究所遭遇的困境入手,总结标记开发的最新进展,为未来 DNA 分子标记的开发提供新的研究思路。【方法】结合作者的研究经历,综合分析相关文献资料,从功能标记的开发手段、特点及应用前景等方面进行了阐述。【结果】与早期从基因组中随机 DNA 序列开发出来的匿名性 DNA 标记相比,从控制生物表型的序列或序列模体,开发可检测功能基因多态性的 DNA 标记(称为功能标记),具有多方面的优点,使其成为标记开发的新阶段。分子生物学的快速发展、特别是功能基因组学的发展和研究成果,为功能性 DNA 标记(功能标记)的开发提供了坚实的基础;功能标记因其独特的优点而具有广阔的应用前景。【结论】根据影响表型的等位基因序列差异开发简便低廉的功能标记,是 DNA 标记开发的新方向和重点,也将为基因资源的挖掘和研究、有利基因的跟踪、标记辅助选择、分子设计育种等提供高效的工具。

【关键词】DNA 标记; 功能标记; 数量性状; 分子设计育种

【中图分类号】Q78 【文献标识码】A 【文章编号】1671-9387(2009)01-0110-07

Development progress, characteristics and application of functional marker

HE Dao-hua^a, LEI Zhong-ping^b, XING Hong-yi^a

(a. College of Agronomy; b. College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: 【Objective】The purpose of the article was to sum up the limitations derived from past DNA markers and latest advances of markers development, so as to facilitate the future development and application of desirable DNA markers. 【Method】The research method was based on the authors research experience and overall analysis of the related documents. 【Result】Compared with DNA markers previously derived from random region of genome (termed anonymous DNA markers, RDM), the latest markers are derived from sequence motifs affecting phenotypic variation and facilitate detecting polymorphism of functional genes controlling phenotype (termed functional markers, FMs). FMs are characterized with many advantages, symbolizing a new stage in the development of DNA markers. With the rapid development of molecular biology, especially the development and production of functional genomics, sufficient foundation and cheerful prospects have been provided for the development of FMs. Due to merits of FMs, the research activity involving the development and use of FMs will certainly gain momentum in the future. 【Conclusion】Based on the sequences of different allele on the same locus affecting agronomic and/or economic traits, the development of FMs (characterized with simplicity and cheapness) will predominate in all kinds of DNA marker systems, and will provide efficient tools for discovering desirable allele resources, for detection and pyramiding desirable alleles, and for “Molecular Aid Selection (MAS)” and “Molecular Design Breeding (MDB)”.

* 【收稿日期】 2007-12-11
 【基金项目】 西北农林科技大学人才基金项目(01140403);西北农林科技大学“青年学术骨干支持计划”项目(01140308)
 【作者简介】 贺道华(1975—),男,湖北随州人,讲师,博士,主要从事棉花生物技术育种研究。
 【通信作者】 邢宏宜(1956—),男,陕西临潼人,副教授,主要从事棉花遗传育种研究。E-mail: xinghongyi1169@163.com

Key words: DNA marker; functional marker; quantitative trait; molecular design breeding

基于 DNA 序列多样性的分子标记,是研究生物多样性、数量性状位点(QTL)作图/基因定位、图位克隆、标记辅助选择以及分子设计育种等的重要工具。以各种标记(如 RFLPs、AFLPs、RAPDs 和 SSRs 等)为基础,已有大量关于生物多样性和 QTL/基因作图的研究报道^[1]。作者在棉花纤维产量、品质等性状的 QTL 定位方面,也获得了一系列的结果^[2-5]。

然而,以生物多样性、QTL 作图/基因定位的研究结果为基础的后续研究和应用仍然存在困难。如,根据建立在分子标记基础上的生物多样性研究结果,在育种初始阶段进行亲本选择,最后并不能获得预期的结果;在植物中 QTL 作图的研究结果也很多,但克隆的 QTL 基因很少,绝大多数 QTL 都未能定位到可克隆区间内;分子标记辅助选择和分子设计育种在实践上的成功例子也极少^[6]。

出现这种困境的原因很多,其中标记的局限性(即匿名性)是一个重要方面。在开发上述标记(即 RFLPs、AFLPs、RAPDs 和 SSRs 等)时,只重视了物种内不同个体间在某同源序列段上存在的片段长度多型性,而对该段序列的功能未加以关注。Lübberstedt 等^[7]称这些标记为匿名性标记(Anonymous genetic markers)。与匿名性标记相对应,Andersen 等^[8]提出了功能标记(Functional marker, FM)的概念。功能标记来源于控制表型的基因序列内部;在鉴别基因序列的表型功能后,挖掘该序列中的多态性信息及对应序列的表型效应,从而开发出能够区分和预测(复)等位基因及相对性状的 DNA 标记,即功能标记。

本文从基于匿名性标记研究所产生的窘境入手,总结了功能标记的开发手段、特点及应用前景,为今后 DNA 标记的开发提供新的思路,以期利用功能基因组学的研究成果,推动基于分子标记的应用研究,使标记辅助选择、分子设计育种等获得较大突破。

1 功能标记的开发

功能标记开发的先决条件是:明确基因位点(Locus)等位基因(Alleles)序列的多态性信息,以及该基因位点控制的性状;从控制同一性状的若干个等位基因中发现影响表型的多态性序列信息,并以简单易行的检测技术进行检测。因此,功能标记开

发的首要工作,是对物种内多个个体的基因组或表达序列标签(EST)进行测序。在某个特定的基因位点,由于平均核苷酸多样性所产生的序列多样性在各物种内不同,一般情况下,自花授精物种的多样性高于异花授粉物种。

1.1 以基因序列开发功能标记

公共数据库中已经积累了大量的核苷酸序列信息,截至到 2006-09,GenBank 中已经收录的序列达 6 100 万条^[9],这在一定程度上促进了 FMs(Functional markers)的开发。但在这些核苷酸序列中,只对其中很少一部分进行了功能研究,例如,模式植物拟南芥约有 25 000 个基因,进行了功能研究的基因尚不足 10%;在其他物种中,功能明确的核苷酸序列数目更少,从而制约了 FMs 的大规模开发。尽管根据核苷酸序列同源性(相似性),可以推测出任何物种 30%~50%表达序列的功能,但其中只有小部分序列的功能与植物农艺性状有关。而要挖掘控制农艺性状的基因(序列),候选基因法和同线性关系(植物基因组间的)是很好的途径。另外,还有其他一些方法可用来确定特定核苷酸序列的功能,如差异显示表达谱、RNA 干涉、T-DNA 以及转座子基因标签(插入突变体)、QTL 作图(包括遗传基因组学、表达 QTL(Expression QTL, eQTL)、蛋白质数量位点(Protein quantity loci, PQLs)、定向诱变(TILLING)等^[10]。因此,对于大多数农艺性状和作物品种而言,可以利用这些方法鉴别候选基因的表型功能,从而开发出大量的功能标记^[8]。

1.2 由基因序列的特定区域开发功能标记

一些研究表明,基因序列的特定区域与功能有关,且这些特定区域对基因最终产物(如蛋白质)的结构具有决定性作用。因此,这些区域可以用于功能标记的开发^[11-13]

如在作物中,可从控制株高及开花期基因序列的特定区域开发功能标记^[10]。Ellis 等^[11]从小麦基因 *Rht1* 中开发出了二等位的功能标记,并用于小麦高产矮化品种的选育。Peng 等^[12]在公共数据库中搜索(tBLASTN,蛋白质序列对核苷酸序列搜索)拟南芥的赤霉素不敏感基因(突变基因 *gai*)编码多肽链时,发现水稻的 EST 序列 D39460 与其直系同源(Orthologues);随后以 D39460 搜索数据库,发现了一系列的基因同系(同源)物,如玉米的 *Dwarf8*、小麦的 *Rht1*、水稻的 *SLR1*、大麦的 *sln1* 以及拟南芥

的 *GAI*。这些同源基因的突变可降低植株的高度,由此开发出二等位的功能标记,可用于选育高产矮化的谷物类品种。Thornsberry 等^[14]研究表明,*Dwarf8* 基因不但控制玉米株高,还影响开花期,这是因为在玉米 *Dwarf8* 中存在 9 个二等位多态性位点,其对开花期存在影响。以上研究表明,理论上存在 $512(=2^9)$ 种单倍型,均可用来开发影响开花期的功能标记。若单倍型所涉及的多个多态性位点都处于单个基因中,则可以预测该基因位点存在较多的(复)等位基因;若要同时改良玉米的株高和开花期,则必须精心选择 *Dwarf8* 功能标记的单倍型。

近几年,已经分离出一些控制农艺性状的数量性状位点或基因(QTL/gene),如玉米株高 *tb1*^[14]、番茄果重 *fw2.2*^[14]、水稻株高 *Hd6*^[12] 和食味品质 *GBSS*^[10]、小麦春化作用(Vernalization)基因 *VRN1*^[15] 和抗叶锈基因 *Lr10/Lr21*^[13,16],这些数量性状位点或基因都可用于开发功能标记。

1.3 用异种探针开发功能标记

利用控制一个物种某性状基因衍生的探针,在另一相关物种中可以针对对应的基因开发功能标记。根据同线性关系和比较基因组学,Yan 等^[15]用小麦 *VRN1* 基因附近的 RFLP 标记,筛选了水稻、高粱和小麦的 BACs 库;Han 等^[17]借用水稻标记,使得小麦 1 号染色体的着丝粒区域(含有控制麦芽品质的两个 QTLs)和 7 号染色体长臂次末端区(含有杆锈抗性基因 *Rpg4*)的标记密度更加饱和;Kato 等^[18-19]用水稻第 3 号染色体中控制抽穗期的 QTL(*Hd6*)区间的 cDNA 标记,成功检测出小麦 5A 染色体上春化作用基因 *Vrn-A1* 处的序列区域与 *Hd6* 存在共线性(*Vrn-A1* 与 *Hd6* 是两个不同的基因);Tikhonov 等^[20]用 *adh* 基因作探针研究了高粱与玉米之间的共线性;Simko 等^[21]用番茄抗黄萎病基因 *Vel* 衍生的探针,在土豆中检测出了其同源性序列 *StVel1*;Li 等^[22]用 PCR 引物和探针,研究了玉米、水稻和高粱中 *Sh2/A1* 的直系同源关系。同线性和比较基因组学的研究表明,用异物种的探针开发功能标记是可行性的。

1.4 诱导突变开发功能标记

利用乙基甲磺酸(Ethylmethane sulfonate, EMS)对基因组进行局部的定向诱变或随机诱变,可产生较大的突变群体(该群体中含有一系列的错义突变)。利用获得的突变材料与原始野生型回交获得等基因系(仅存在单个序列模体的差异),然后筛查点突变,比较等基因系的表型差异,获知差异序

列模体的功能(提供关于其功能的直接证据)^[23],再从这些经功能鉴定的序列模体中开发出功能标记。Andersen 等^[8]将通过 EMS 诱变或 TILLING 技术开发的功能标记称为“直接功能标记(Direct FMs, DFMs)”。如 Colbert 等^[24]利用高通量的 TILLING 技术对 *Sir2B* 基因进行点突变检测,发现其中的 5 个 G/C 向 A/T 的转换,并获得了这些点突变(SNPs)位点的功能标记;Till 等^[25]在 100 多个基因序列中检测出了 >1 000 个 EMS 诱导的可引起表型变异的点突变;Somers 等^[26]开发出能检测这些点 SNPs 的功能标记。

但建立在化学诱变基础上,通过 TILLING 技术来开发功能标记,存在两个显著的局限性,一是点突变很难被检测出来,二是化学诱变引入了大量的点突变,使得表型分析变得相对困难^[24]。因此,通过诱导突变,每次只能开发极少基因的功能标记,效率较低。

1.5 通过关联分析开发功能标记

关联分析多依赖于候选基因中位点多态性与表型变异关联的程度。以一些相互间无亲缘关系的种质资源、纯系品种或自交系为材料(不需要人工杂交自交构建作图群体),针对候选基因区段进行序列多样性分析,结合生物材料的表型数据,采用统计学的方法,从而检测表型变异与候选基因序列多样性之间的关联,也是一种验证候选基因表型功能的途径^[27-28]。若二者之间存在显著关联,则来源于该位点的可检测该序列多样性的分子标记,即相应的功能标记。此方法充分利用自然进化和人工进化过程中所积累的遗传重组,从而可高分辨率地检测现存物种中候选基因的遗传多型性与表型变异之间的关系^[29],最终识别功能性的多态性及互作^[30],获得功能标记。在人类分子生物学研究中,用高通量的基因型分型技术,充分检测单个基因中连锁不平衡(linkage disequilibrium, LD)的方式,选择能代表同一基因位点不同等位基因合适的标签 SNPs(tSNPs),找出与功能性序列模体紧密关联的 tSNPs,再把这些 tSNPs 组装为单倍型,那么可将这些 tSNPs 称为单倍型标签 SNPs(htSNPs)^[31]。起源于人类遗传学的关联分析,通过 htSNPs 进行基因分型,然后结合表型数据鉴别或确认控制表型的基因或其内部的序列模体。关联分析不仅需要对等位基因进行广泛测序,而且还需要具有代表性的群体以便进行表型研究。由基因内多样性所导致的表型差异,其区分与剖析需要高的遗传分辨率(即低水

平的连锁不平衡)。在 QTL/基因作图时,应用 htSNPs 进行基因型分型,可以降低 SNPs 分析的工作量。如 Thornsberry 等^[14]应用控制开花期的位点 *FRI* 的分子标记,对 92 个玉米自交系进行关联分析发现,开花时间与株高之间存在高度的相关性,这推动了株高基因 *Dwarf8* 的测序,从而发现了 41 个与不同开花期紧密关联的单倍型;同样, Nordborg 等^[32]在拟南芥中也有相似的研究。这些研究均用单倍型标签 SNPs(也包括 InDels)来分析标记-性状间的关联。

通过关联分析开发功能标记时,需要首先研究清楚基因中的 LD 方式。在拟南芥中,LD 区域达 10 kb 以上,而玉米中 LD 区域不到 1 kb^[33]。因此,在 LD 水平低的物种(如玉米)中,关联分析有助于鉴别控制表型的序列模体(一些核苷酸或核苷酸插入/缺失),还能确定这些序列模体、基因的遗传效应。但对于 LD 水平高、单倍型序列(/结构)较长的物种,若要鉴别多态性是否与表型变异相关,由于遗传背景可能影响关联分析的结果,因此仅依靠关联分析是不够的,还需要采用一些统计方法用于控制未知的群体结构。在鉴别序列模体的功能时,关联分析仅提供统计学上的(间接的)证据,故 Andersen 等^[8]将通过关联分析而开发出的功能标记称为“间接功能标记(Indirect FMs, IFMs)”。

2 功能标记的特点

2.1 功能标记的优点

随机标记(Random DNA markers, RDMs)来源于全基因组 DNA 序列中某个任意位点的多态性,并不考虑该位点是否存在于基因中,即随机标记可能来源于基因序列(包括外显子 exon 和内含子 intron)的多态性,也可能来源于基因间隔区(spacer)序列的多态性;基因标记(Gene targeted markers, GTMs)来源于基因序列内部的多态性,主要是外显子区域的多态性。RDMs 与 GTMs 的开发均不考虑标记位点的多态性与表型变异之间是否存在关联,因此属于匿名性标记。相反,功能标记(FMs)来源于基因序列内部的多态性,并且还要求该基因的多样性必须与表型变异相关联(即 FMs 标记来源于功能已经明确的且具有多态性的 DNA 序列)。由于功能标记与功能性序列模体完全连锁(不存在交换重组),因此功能标记(FMs)主要应用于以下方面:(1)在群体中有效跟踪目标等位基因;(2)控制性平衡选择(Controlled balancing selection);(3)在自

然群体或育种群体中,筛选有利基因;(4)在育种群体中,对影响相同或不同性状的 FMs 有利等位基因进行组合,进行“分子设计育种”;(5)构建连锁的 FMs 单倍型,消除连锁累赘。

与 GTMs 和 RDMs 相比, FMs 具有以下优点。

2.1.1 功能标记的遗传效应值具有普适性,且可靠性高 RDMs 和 GTMs 遗传效应值的估计,依赖于该标记与目标等位基因的连锁相^[34]及连锁的紧密程度。在以 RDMs 为基础的 QTL 作图中,即使 2 个作图群体的亲本完全一样,能在 2 个群体中同时检测出的 QTL 只有 60%;若 2 个作图群体只有 1 个共同亲本,则能在 2 个群体中同时检测出的 QTL 只有 38%;若 2 个作图群体无共同亲本,能在 2 个群体中同时检测出的 QTL 仅有 30%^[34]。在 $A \times B$ 的后代群体中,经检测为具有正效应的 RDMs 基因,在 $C \times D$ 的后代群体中可能具有负效应。以 RDMs 为基础的 QTL 作图,对每个新构建的群体而言都是必要的(RDMs 的遗传效应值在其他群体中缺乏可推广性和可演绎性)^[34],因为不同群体具有不同的多态性 QTLs 组合,即使作图群体的亲本关系很近, RDMs 与 QTL 的连锁相及连锁紧密程度也可能不一致,减数分裂过程中人工不可控制的交换重组会打破这种连锁,导致 RDMs 和 GTMs 的遗传效应值需要重新估计。

对于 FMs 标记而言,若某个序列模体在表型上的遗传效应得以确定,则由这个序列模体开发的 FMs,就可用于在很广的遗传背景中跟踪目标等位基因,而无需对该标记的功能效应值进行重新检测和验证。特别是在植物育种中,构建分离群体(杂交)前的亲本选择以及随后的纯系或自交系的选择,功能标记均具有极大的辅助作用。利用功能标记的上述优点,在进行杂交育种、聚合育种时,可以用 FMs 进行有目的的等位基因组合(分子设计育种);在群体改良和轮回选择时, FMs 可以避免目标位点的遗传漂变;在品种检测鉴定时,可根据 FMs 特征带的有无区分品种。

2.1.2 功能标记可准确地检测、跟踪功能位点的目标基因 QTL 作图群体中所检测出的与目标位点等位基因连锁的 RDMs/GTMs,其在该群体后续世代的有效性(应用价值),取决于 RDMs/GTMs 标记与目标位点基因连锁的持续性(连锁不平衡的程度)。这种持续性与 RDMs-基因间的遗传距离、所跟踪的基因数目以及自 QTL 作图后群体所经历的世代数 3 个因素有关。因此,在 RDMs、GTM 辅助

选择过程中,至少需要 1 次表型考查以确认目标基因的存在。而在利用 FMs 进行前景选择(在回交进程中)时,无需表型考查。

由于遗传重组较少,较长的染色体片段向下一代传递,不利基因可能与目标基因连锁,特别是用不合适的材料作供体亲本时,更易出现连锁累赘。而 FMs 源于控制表型的序列模体,与目标基因紧密连锁,因此可以高效地用来在较大的群体和广泛的世代中跟踪基因^[7],或者监测与表型变异相关的多态性。所以,在进行前景选择时,FMs 比 RDMs 更为有效。

2.1.3 功能标记的多态性信息含量(Polymorphism information content, PIC)能反应功能性等位基因的遗传变异 一个基因位点的等位基因数目,同与之连锁的 RDMs 位点等位基因数目可能不同。例如,二等位的 AFLP 在读带时被记作有或无,而等位基因的数目可能多于 2 个。用 AFLP 标记来研究生物多样性时,通常认为具有相同分子量带的个体在该标记位点的核酸序列相同。事实上,该标记位点的核酸序列组成并不一定相同。尽管(复)等位性的 RDMs(如 SSR)也能描述等位序列多样性,但并不能证明其紧邻的基因序列也具有同等程度的多样性。然而,功能标记来源于具有多态性的、控制生态/农艺性状的序列模体内部,因此在植物育种与多样性研究中,FMs 能更准确地描述基因的多样性;在育种群体和自然群体中,更能准确地筛选和追踪已知基因。此外,在 FMs 开发过程中,积累了大量关于控制表型的序列模体特征和位置信息,这有助于在种质资源中发掘有利基因。

2.2 功能标记的局限性

对一些基因,如玉米的矮化基因 *Dwarf8* 及其同源基因 *Rht1*(小麦)、*SLR1*(水稻)、*sln1*(大麦)等,均可用于功能标记的开发。在拟南芥和一些作物中,等位基因测序工作已经展开,这都有助于功能标记的开发。然而,到目前为止,即使在模式植物中进行了功能研究的基因也只有 10% 左右,而在非模式植物中可能不到 5%,且在生物学上功能明确的基因并不一定具有农艺学意义上的功能^[35]。因此,在进行 FMs 开发以前,不仅必须对基因及序列模体的功能进行深入研究,而且对所有控制生态和农艺性状的基因,需要检测、鉴定同一位点所存在的全部等位基因。待基因功能明确以后,须选择和创建适宜的植物材料进行性状考查,从而辨别其遗传效应的大小。

利用一个大群体进行关联分析所开发的间接功能标记(IFMs),其遗传效应具有普适性,可以直接应用于育种。利用等基因系的比较开发的直接功能标记(DFMs),其开发成本较高,且相关研究结果仅在特定的遗传背景中是可靠的。为了降低开发成本,首先可以通过关联分析来筛选序列模体,然后通过等基因系比较,最终开发出 DFMs。DFMs 需要在不同的遗传背景中进行广泛评价,从而更准确地估算其遗传效应。而 FMs 在表型效应上的稳定性尚需作充分研究,并要与植物材料的性状稳定度一致。因此,需要应用试验、统计和生物信息学等多种手段,积累不同来源的数据并进行综合评价,从而确证 FMs 遗传效应的估算值。

用不同的测验种(Tester)进行多个测交,通过关联分析和纯合的等基因系比较,从而确定序列模体的加性和显性遗传效应。但加性和显性效应值的估计是相对的(如在一个特定的研究中,遗传效应值的估计不仅与基因有关,而且与基因的遗传背景有关)。因此,不同研究应当使用一个公用的参照材料,以便对不同来源的研究结果(数据)进行整合。

生物材料的自交能力,是进行等基因系创建和关联研究的限制条件。对于自交不亲和或自交衰退严重的物种,需要对含有不同 FMs 等位基因的杂合基因型群体进行研究,其中遗传效应小的 FMs 检测,会受检测对象内其他遗传成分的影响。因此,FMs 的开发,对于自花授粉物种、自交亲和物种及可自交的异交物种可行性更高。然而,自花授粉物种通常有较长的单倍型。对于这些物种,需要比较等基因系以确定序列模体的功能,并需构建(积累了多代的遗传重组的)群体以进行关联分析。

分子标记辅助选择有助于提高选择效率,特别是在性状遗传率较低之时。但是,在对大量遗传效应较小的基因进行选择时,标记选择的效率会有所降低。因此,在对某个数量性状进行 FMs 开发时,应该考虑最佳的基因数目,这是十分重要的。利用“关键基因”可以确保对目标性状进行理想控制,避免一因多效的负面影响。然而,仅使用较少的 FMs 将有助于降低开发成本,减少不利基因的连锁累赘。QTL 作图表明,QTL 的遗传效应是偏分布的,大效应的 QTL 极少,多数 QTL 具有适中或很小的效应。鉴于 FMs 开发的特性和限制因素(费力),初始的 FMs 开发应着眼于对表型变异有根本性影响的基因。因此,“关键基因”的选择,是 FMs 开发的重要限制步骤。

3 功能标记的应用和未来研究方向

3.1 应用

来自基因组表达区的功能标记,对于作物改良具有重要的价值。例如,利用转录图谱和功能图谱,可以分离控制农艺性状的基因及 QTL。充分利用 SNPs,对一些重要基因进行单倍型分析(Haplotyping),从而鉴别出该位点的所有等位基因(每个基因具有一些特征性的 SNPs),有助于对控制农艺性状位点上的所有等位基因进行挖掘和了解,从而为“设计育种”提供信息^[36]。对于仅在杂种中表达而不在亲本中表达的序列,可用 FMs 鉴别并进行作图^[37],在杂交组配时,选择“杂种特异表达位点”和“杂种优势位点”最多的材料作亲本,能获得最佳的杂种优势。

功能标记的开发,有助于推动育种中的分子标记辅助选择的应用和基于候选基因策略(candidate-gene approach)的^[38]研究。在候选基因策略中,可利用 FMs 研究生物分子的代谢途径。例如,Finlay 等^[39]开创了使用候选基因策略进行低温糖化机理(代谢途径)研究的新途径。用这种标记进行多样性研究,可以对遗传多样性和遗传距离进行更准确地估计,从而指导杂交前的亲本选择工作,这对整个育种进程具有很重要的意义。因此,功能标记的应用将促进植物育种,利用重要基因进行遗传转化从而产生转基因作物,并通过标记辅助选择提高了传统育种的效率。将区间作图产生的 QTL 和功能标记结合,可能会使“以基因为媒介的育种途径(Gene-mediated breeding)”代替标记辅助选择方法^[40]成为可能。

功能标记的开发将推动种质资源的挖掘与研究。在丰富多样的种质资源中,控制表型的基因位点,在进化(包括自然进化和人工进化)和驯化过程中,积累了丰富的序列变异,来自该区段的功能标记,将有助于在该物种内充分地挖掘该位点的各种等位基因,并且结合表型考查识别出(复)等位基因中的优势等位基因及相应的种质材料,为杂交育种中亲本的选择、优异性状的转育奠定基础。

3.2 未来的研究方向

从基因组的表达序列中开发功能标记,以及此类标记的研究与应用,在结构基因组学与功能基因组学进展的推动下将会成为分子标记研究的热点。目前,正在对拟南芥、水稻和毛果杨的基因组序列,及其他一些植物基因组的部分序列进行功能注释,

这些经功能注释的序列,包括特定基因的序列,均可用于开发 GTMs 和 FMs。对于分子生物学研究相对滞后的棉花而言,cDNA 序列信息或/和候选基因序列信息也在不断累积,新近修订出的第七版棉花基因索引(Cotton Gene Index version 7,CGI7),包含了 55 673 条非冗余基因序列,研究这些序列的多态性及相关表型的多样性,均可推动棉花功能标记的开发。GenBank 中的 EST 序列与日俱增,非常有利于物种间、物种内序列多态性的发现和物种间表型效应预测。这些基因组学的研究结果,为功能标记的开发提供了大量信息。

分子标记的出现和发展,在很大程度上推动了种质资源遗传多样性的研究、基因定位、QTL 作图、品种指纹图谱的构建,也在一定程度上推动了标记辅助选择^[41-44]和基因的图位克隆。然而,这与研究的最终目标——随心所欲地改造生物并获得期望的表现型还相距甚远^[1,6],很多基于基因定位/QTL 作图的标记辅助选择未获得理想的结果,图位克隆也未获得相应的基因。这说明以前的标记开发思路过于简单,今后必须针对基因序列或者基因序列的特定区域,首先进行表型效应的识别和验证,然后结合该序列碱基组成的多样性信息,开发出能鉴别不同表型的、简便易行的分子标记,即功能标记,如功能性的 SNP 标记、功能性的 SSR 标记等。通过功能标记的开发,识别功能基因及其各种等位基因的标签,可以更准确地检测和跟踪优势等位基因,最终推动标记辅助选择和分子设计育种在育种实践中的实施,实现快速、准确改良生物的终极目标。

[参考文献]

- [1] Salvi S, Tuberosa R. To clone or not to clone plant QTLs: Present and future challenges [J]. Trends Plant Sci, 2005, 10(6): 297-304.
- [2] 贺道华, 林忠旭, 张献龙, 等. 陆地棉纤维品质遗传基础的分子剖析 [J]. 棉花学报, 2004, 16(3): 131-136.
He D H, Lin Z X, Zhang X L, et al. Dissection of genetic basis of fiber quality in *Gossypium hirsutum* with molecular markers [J]. Cotton Science, 2004, 16(3): 131-136. (in Chinese)
- [3] He D H, Lin Z X, Zhang X L, et al. Mapping QTLs of traits contributing to yield and analysis of genetic effects in tetraploid cotton [J]. Euphytica, 2005, 144(1): 141-149.
- [4] He D H, Lin Z X, Zhang X L, et al. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* × *Gossypium barbadense* [J]. Euphytica, 2007, 153(1): 181-197.
- [5] He D H, Zhang X L, Lin Z X, et al. Dissection of genetic variance of fiber quality in advanced generation from interspecific

- cross of *G. hirsutum* and *G. barbadense* [J]. Plant Breeding, 2008, 127(3): 286-294.
- [6] Masojć P. The application of molecular markers in the process of selection [J]. Cell Mol Biol Lett, 2002, 7: 499-509.
- [7] Lübberstedt T, Zein I, Andersen J R, et al. Development and application of functional markers in maize [J]. Euphytica, 2005, 146: 101-108.
- [8] Andersen J R, Lübberstedt T. Functional markers in plants [J]. Trends Plant Sci, 2003, 8: 554-560.
- [9] Benson D A, Karsch-Mizrachi I, Lipman D J, et al. GenBank [J]. Nucleic Acids Res, 2007, 35(Database issue): D21-D25.
- [10] Gupta P K, Rustgi S. Molecular markers from the transcribed/expressed region of the genome in higher plants [J]. Funct Integr Genomics, 2004, 4: 139-162.
- [11] Ellis M H, Spielmeier W, Gale K R, et al. Perfect markers for the *Rht-B1b* and *Rht-D1b* dwarfing gene in wheat [J]. Theor Appl Genet, 2002, 105: 1038-1042.
- [12] Peng J, Richards D E, Hartley N M, et al. 'Green revolution' genes encode mutant gibberellin response modulators [J]. Nature, 1999, 400: 256-261.
- [13] Timmerman G M, Frew T J, Weeden N F, et al. Linkage analysis of *er-1*, a recessive *Pisum sativum* gene for resistance to powdery mildew fungus (*Erysiphe pisi* D. C.) [J]. Theor Appl Genet, 1994, 88: 1050-1055.
- [14] Thornsberry J M, Goodman M M, Doebley J, et al. *Dwarf8* polymorphisms associate with variation in flowering time [J]. Nat Genet, 2001, 28: 286-289.
- [15] Yan L, Loukoianov A, Tranquilli G, et al. Positional cloning of the wheat vernalization gene *VRN1* [J]. Proc Natl Acad Sci USA, 2003, 100: 6263-6268.
- [16] Feuillet C, Travella S, Stein N, et al. Map-based isolation of the leaf rust disease resistance gene *Lr10* from the hexaploid wheat (*Triticum aestivum* L.) genome [J]. Proc Natl Acad Sci USA, 2003, 100: 15253-15258.
- [17] Han F, Kleinhofs A, Ullrich S E, et al. Synteny with rice: analysis of barley malting quality QTLs and *rpg4* chromosome regions [J]. Genome, 1998, 41: 373-380.
- [18] Kato K, Miura H, Sawada S. Comparative mapping of the wheat *Vrn-A1* region with the rice *Hd-6* region [J]. Genome, 1999, 42: 204-209.
- [19] Kato K, Kidou S, Miura H, et al. Molecular cloning of the wheat *CK2alpha* gene and detection of its linkage with *Vrn-A1* on chromosome 5A [J]. Theor Appl Genet, 2002, 104: 1071-1077.
- [20] Tikhonov A P, SanMiguel P J, Nakajima Y, et al. Collinearity and its exceptions in orthologous adh regions of maize and sorghum [J]. Proc Natl Acad Sci USA, 1999, 96: 7409-7414.
- [21] Simko I, Costanzo S, Haynes K G, et al. Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach [J]. Theor Appl Genet, 2003, 108: 217-224.
- [22] Li W, Gill B S. The collinearity of the *Sh2/A1* orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the Triticeae [J]. Genetics, 2002, 160: 1153-1162.
- [23] Henikoff S, Comai L. Single-nucleotide mutations for plant functional genomics [J]. Annu Rev Plant Biol, 2003, 54: 375-401.
- [24] Colbert T, Till B J, Tompa R, et al. High-throughput screening for induced point mutations [J]. Plant Physiol, 2001, 126: 480-484.
- [25] Till B J, Reynolds S H, Greene E A, et al. Large-scale discovery of induced point mutations with high-throughput TILLING [J]. Genome Res, 2003, 13: 524-530.
- [26] Somers D J, Kirkpatrick R, Moniwa M, et al. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs [J]. Genome, 2003, 49: 431-437.
- [27] Whitt S R, Buckler E S. Using natural allelic diversity to evaluate gene function [J]. Methods in Molecular Biology, 2003, 236: 123-139.
- [28] Wilson L M, Whitt S R, Ibáñez A M, et al. Dissection of maize kernel composition and starch production by candidate gene association [J]. The Plant Cell, 2004, 16: 2719-2733.
- [29] Simko I. From association mapping to marker-assisted selection; Experimental design and application [J]. Phytopathology, 2007, 97: S149.
- [30] Bao J S, Corke H, Sun M. Nucleotide diversity in *starch synthase IIa* and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.) [J]. Theor Appl Genet, 2006, 113: 1171-1183.
- [31] Goldstein D B, Ahmadi K R, Weale M E, et al. Genome scans and candidate gene approaches in the study of common diseases and variable drug responses [J]. Trends Genet, 2003, 19: 615-622.
- [32] Nordborg M, Borevitz J O, Bergelson J, et al. The extent of linkage disequilibrium in *Arabidopsis thaliana* [J]. Nat Genet, 2002, 30(2): 190-193.
- [33] Rafalski J A. Novel genetic mapping tools in plants: SNPs and LD-based approaches [J]. Plant Science, 2002, 162: 329-333.
- [34] Lübberstedt T, Melchinger A E, Fähr S, et al. QTL mapping in testcrosses of flint lines of maize: III. Comparison across populations for forage traits [J]. Crop Sci, 1998, 38: 1278-1289.
- [35] 贺道华, 张献龙. 数量性状由表型变异到基因发现的研究进展 [J]. 遗传, 2006, 28(12): 1613-1618.
He D H, Zhang X L. Advance in quantitative traits from phenotypic variation to gene discovery [J]. Hereditas, 2006, 28(12): 1613-1618. (in Chinese)
- [36] Peleman J D, van der Voort J R. Breeding by design [J]. Trends Plant Sci, 2003, 8: 330-334.