

上位性及其在遗传育种研究中的应用

张文英,程君奇,朱军*,吴为人

(浙江大学生物信息学研究所,浙江 杭州 310029)

摘要: 上位性引入遗传学已有一个多世纪,直到近些年才受到广泛关注,成为复杂性状遗传研究体系的一个重要组成部分。上位性可分为统计上位性和功能上位性两类,前者具有群体特性,后者属于基因型现象。分子标记技术是研究上位性的一个有力工具,理论与实验研究证实上位性在动植物数量性状的表现中具有重要作用。上位性在作物育种中的应用因作物的繁殖方式,育种方法等不同而异,上位性是杂种优势形成的重要遗传基础。

关键词: 上位性;遗传学;育种;数量性状基因座

中图分类号: Q943 **文献标识码:** A **文章编号:** 1672-5565(2004)-02-0009-04

Epistasis and its application in genetics and breeding

ZHANG Wen - Ying, CHENG Jun - Qi, ZHU Jun, WU Wei - Ren

Institute of Bioinformatics, Zhejiang University, Hangzhou 310029, China

Abstract: Conception and analytical methods of epistasis and its application in genetics and breeding were reviewed. Epistasis was introduced into genetics in 1909, however, it was not received enough attention until recent years, and it became an indisceptible part of genetic research system upon complex traits. Epistasis includes statistical epistasis and functional epistasis, the former is a population phenomenon whereas the latter is a genotypic phenomenon. Molecular marker technology is a powerful tool in studying epistasis, evidence from both theoretic researches and empirical experiments have proved that epistasis is important for performance of main quantitative traits. Implications of epistasis for plant breeding methodology should depend on seed production and the methods for breeding. Epistasis is an important genetic basis for basis for the yielding of heterosis crops.

Key words: epistasis; genetics; breeding; Quantitative trait locus

生物的大多数形态性状和行为性状是具有复杂遗传体系的数量性状。控制生物复杂性状的遗传效应体系包括基因主效应、基因间互作效应(上位性)以及基因与环境的互作效应^[1],上位性以及基因与环境互作对复杂遗传体系的性状要比对简单遗传体系的性状更为重要^[2]。当前遗传学面临的一个主要任务和挑战就是阐明这些复杂性状的遗传效应体系^[3,4],而基因间互作(上位性)是复杂性状遗传体系的一个重要部分。

1 上位性概念的由来

上位性(epistasis)一词最早可追溯到1807年,当时在医学上用于描述尿样静置后表面产生的一种物质形态—浮泡(scum)^[5]。1909年, Bateson 将上位性一词引入遗传学,用以描述一个基因座的表现型效应掩盖另一个基因座表现型效应的基因间互作现

象^[6]。1918年, Fisher 将多基因座间非加性的互作效应称为上位性(epistacy)^[7]。在提出后的七十多年,上位性一直没有引起人们的广泛关注和足够认识^[6]。直到上世纪八十年代,随着分子技术的大量应用、遗传学以及复杂性状研究方法的发展,上位性才逐渐成为研究热点^[8]。目前孟德尔经典遗传学和分子遗传学倾向于使用 Bateson 给出的上位性的严格定义,而进化理论和数量遗传学则将所有形式的基因座间互作统称为上位性。

2 上位性种类及分析方法

遗传学由关于上位性的定义有很多种^[6,9,10],根据所采用的分析方法和适用范围可将其归纳为两类:一类为统计上位性(statistical epistasis),另一类为功能上位性(functional epistasis)。

2.1 统计上位性

收稿日期:2004-4-10;修回日期:

作者简介:张文英(1972-),男,湖北天门,博士研究生,长江大学农学院讲师,研究方向:统计遗传学。

* 通讯作者:朱军(1949-),Tel:0571-86971731, E-mail: jzhu@zju.edu.cn

Fisher(1918)首次提出统计上位性,用以描述非等位基因座累加效应的统计偏差,并将群体的遗传变异分解为加性、显性和上位性方差分量^[11]。Cockerham(1954)和 Kempthorne 对统计上位性概念作出详细的阐述,并将其引入现代数量遗传学^[12]。同时,Cokerham(1954)进一步将上位性方差分解为加性×加性(AA)、加性×显性(AD)、以及显性×显性(DD)3种分量^[12],从而使上位性更具群体特性而不是非等位基因座功能的互作。由此,可以定义统计上位性(基因间互作)为多基因座总基因型值与各单基因座累积基因型值间的离差。统计上位性是一种群体现象,它依赖于特定群体中的等位基因的频率,上位性互作离差仅与互作遗传方差分量有关。

Hayman 和 Mather(1955)、Mather(1967)也先后提出了统计上位性模型,归纳起来,二基因上位性遗传模型分为3类:第一类,纯系度量模型(pure-line-metric model)或称为 F_{∞} 度量模型(F_{∞} -metric model, F_{∞} 表示从 F_2 代起自交繁殖无穷代的群体),模型中的参数与连锁不平衡下的 F_{∞} 群体中的基因型频率呈正交性。第二类: F_2 度量模型(F_2 -metric model,即 Cockerham 模型),模型中参数与连锁不平衡下 F_2 群体中基因频率呈正交性。第三类,混合度量模型(mixed-meric model,即 Hayman-Mather 模型),是 F_2 度量模型与 F_{∞} 度量模型的混合,它可由 F_{∞} -metric 模型通过均值消减转化而来。Kao 和 Zeng(2002)比较分析发现,(Cockerham 模型)在方差分量分解、遗传效应的估计和解释、QTL 定位等方面都明显优于其他两个模型,该模型可简化基因间上位性的研究,能很好地应用于 QTL 定位。

运用 Cockerham 提出的广义遗传模型,朱军(1989,1992)提出了估算加性×加性上位性(AA)遗传方差分量的遗传分析方法。运用该方法,分析包括杂种一代和杂种二代或回交一代的双列杂交遗传资料,可以无偏地估算加性×加性上位性(AA)遗传方差分量。上位性的数量遗传学模型可利用表现型数据估计遗传参数,将其与遗传率、选择响应等联系起来,用于遗传理论研究和指导育种实践,但是难以解释分子水平上基因与基因产物间的互作关系。

2.2 功能上位性

1995年,Cheverud 和 Routman 提出生理上位性(physiological epistasis),用以表示在一个基因座上具有不同基因型的个体间的表现型差异依赖于其它基因座上的基因型。Hansen 和 Wangner(2001)提出功能上位性一词取代生理上位性,他们认为基因间的

互作可能来自没有生理意义的机制^[11]。这里成对的上位性可以理解为一个基因替换另一个基因后产生的直接效应,并可从低阶推广到高阶上位性。功能上位性是一种基因型现象,独立于等位基因的频率,对加性以及显性遗传方差还有一定的贡献,因此,它在进化研究中的应用要明显多于统计上位性。

在功能上位性分析方面,Kauffman(1993)提出了能够分析多基因互作的布尔调控网络 NK 模型,Gibson(1996)发展了转录调控的热动力模型。Cheverud 和 Routman(1995)提出了分析生理上位性的数学方法,定义上位性值为基因型值非加权回归的残差,该方法将上位性视为线性模型的离差,虽然基于非加权而非群体均值,但其实质类似于统计方法。Hansen 和 Wangner(2001)提出了适用于各阶次的基因间互作的多元线性模型(multilinear model)^[11],该模型可用于估计遗传试验中的功能上位性的效应大小。

3 上位性的 QTL 定位研究

数量性状遗传基础复杂,易受环境影响,表现型与基因型间没有明确的对应关系,因此,对数量性状的研究十分困难。分子标记技术的出现,为深入研究数量性状的遗传基础提供了可能。利用分子标记进行遗传连锁分析,可以检测出 QTLs 及其间的相互关系,即 QTL 定位(QTL mapping)。从 20 世纪 80 年代末期开始,一系列基于分子标记的 QTL 定位统计方法相继提出,为数量性状的 QTL 定位提供了有效的统计工具,但早期提出的方法并没有考虑上位性效应。实际上,DNA 标记研究可用来检测作物中 QTLs 之间上位性互作的存在、数量和性质等。Li 等(1997)应用限制性片断长度多态性(RFLP)标记研究了水稻产量的 3 个组分性状的遗传基础,检测出 19 个有关的 QTLs,发现上位性是形成诸如产量组分这类复杂性状的重要遗传基础,尤其是是像每穗粒数、每穗粒重这种遗传力较低的性状。Kao 等(1999)提出了基于多元回归模型的多元区间作图(multiple interval mapping)方法,可以估算多基因座之间的上位性。Wang 等(1999)采用混合线性模型(mixed linear model)方法,提出了考虑上位性效应和 QTL 与环境互作的 QTL 定位方法,可以检测全基因组范围的上位性效应。Cao 等(2001)利用条件 QTL 定位方法研究上位性、QTL 与环境互作对水稻株高发育的影响,发现上位性对水稻株高性状很重要。

如今 QTL 的研究已成为沟通群体内连续变异

性与产生这些偏差的内在遗传机制的有力工具,相比传统数量遗传学方法, QTL 试验研究提供了更多的关于上位性的证据,因为它能估计特定染色体区间的效应而不是全基因组的平均或总体基因效应,能估计基因效应而不仅仅是统计遗传效应。当有更多的 QTL 被检测出来时,基因间互作将成为统计遗传学的研究热点,就像它们在孟德尔遗传学中作为功能分析的工具一样^[6]。使用分子标记鉴别农业上重要数量性状的 QTLs 已经成为洞悉这些性状的遗传基础和规划动植物遗传改良方案的一个重要手段,在今后的理论和实际应用中必将起到越来越大的作用。

4 上位性在遗传育种中的应用

4.1 常规选育

在常规选育中,一般认为只有加性效应可以稳定遗传,选择育种中得到的可稳定遗传的变异——育种值就是加性效应值。事实上,对于自花授粉作物,加 \times 加上位性效应可以通过选择稳定遗传,如果加性 \times 加性上位效应存在,那么在选择育种中将会产生额外的遗传收益。由于计算的复杂性以及估计的不准确等原因,非加性遗传变异在遗传效应估计体系以及许多育种过程中被忽略了。忽略非加性遗传因素的协方差将导致加性遗传方差估计发生偏差,使加性遗传效应的预测不准确,而准确估计非加性遗传方差有助于评价育种过程中特殊配合力的潜在效益,因此,对非加性遗传方差分量的估计十分重要。虽然上位性并非只发现对自花授粉作物起重要作用,但是在自花授粉作物中发现上位性的事例确实要比在异花授粉作物中多,这一现象可归于两个方面的原因,一方面是在自花授粉作物中有更完善的试验设计可以检测出上位性,另一方面是自花授粉作物往往显示出较强的非等位基因互作效应,异花授粉作物则表现出较强的显性效应。

4.2 分子标记辅助选择

分子标记辅助选择 (Molecular marker - assisted selection, MAS) 自提出之日起就为育种工作展示了广泛的前景,现已成为分子育种的重要组成部分。如今 MAS 在动植物育种质量性状的改良方面已取得可喜的成绩,而对数量性状研究主要局限在理论上的一个重要原因就是上位性的存在。最近, Liu 等 (2003) 提出了考虑上位性效应、基于 QTLs 的标记辅

助选择 (marker - assisted selection, MAS) 新方法,通过计算机模拟研究发现,在上位性存在的情况下,该方法比仅考虑加性或考虑加性和显性时的选择方法具有更持久的选择响应,而忽略上位性会给高世代的选择响应带来相当大的损失。Jannink (2003) 通过模拟试验比较了在加性效应和在加性 \times 加性上位性效应单独作用下有限群体的选择效应,结果也表明后者比前者具有更大和更持久的响应。

数量性状的分子标记辅助选择虽然难度大、理论上还不成熟,但是从长远看,该技术用于改良多基因互作(上位性)性状乃至同时改良多个复杂性状性状、多个品种等更为复杂的育种计划是未来的发展趋势。

4.3 杂种优势的遗传理论研究

杂种优势在动植物育种的应用已经十分普遍。迄今为止,人们提出了多种关于杂种优势的理论解释,其中以上世纪 Bruce 等提出的显性假说以及 Shull 和 East 等提出的超显性假说最具代表性,但这些假说大多基于单基因理论。Yu 等 (1997) 采用来自一个强优势水稻杂交组合的 250 个 $F_{2,3}$ 家系,检测了分布在全基因组的大理标记位间的二基因互作效应,分析了决定产量及其组分的遗传基础,指出上位性在杂种优势的遗传基础中起着重要作用。Li 等 (2001) 和 Luo 等 (2001) 研究指出,上位性和超显性是自交衰退和杂种优势的主要遗传基础。Omholt 等 (2000) 从基因调节网络 (gene regulatory network) 角度阐述了基因座内和基因座间各种效应间的关系,认为显性、超显性和上位性在杂种优势遗传解释中具有一定的调和性。Hua 等 (2003) 研究水稻永久 F_2 群体后发现,各种遗传效应,包括单基因座水平的部分显性、完全显性、超显性以及 3 种二基因互作效应 (加性 \times 加性互作、加性 \times 显性互作和显性 \times 显性互作) 对永久 F_2 群体中的杂种优势均有贡献,它们之间作为杂种优势的遗传基础互不排斥。上述结论有助于了结一个多世纪以来关于杂种优势遗传基础的争论。

在杂种优势预测方面, Xu 和 Zhu (1990) 基于加性 - 显性 - 加性 \times 加性上位性模型 (ADAA model) 提出了预测杂交后代潜在杂种优势的方法,并指出,上位性杂种优势 (epistatic heterosis) 在杂种 F_1 代及其后续世代中均起着重要作用。

(下转第 50 页)

个水稻 cDNA 芯片图谱中,与 EDTA-Fe 比较发现了 451 个差异 cDNA 点,其中缺铁诱导上调基因为 203 个。在缺铁诱导的 203 个上调基因中包括:膜蛋白(可能是铁吸收相关的基因)5 个,膜泡流有关的基因 9 个,细胞骨架相关的基因 4 个,物质能量代谢流有关的基因 12 个,转录翻译的调控因子有关的基因 14 个,信息流有关的基因 12 个,细胞周期调控的基因(根毛增加以被公认为是缺铁形态变形的特征)12 个,未知功能基因 22 个,其余是未知基因。不同铁处理水稻根可溶性蛋白双向 IEF-SDS-PAGE 图谱均得到多达 1500 个左右的蛋白质点。比较缺铁与柠檬酸铁培养的水稻根可溶性蛋白有 23 个差异表达蛋白点,其中缺铁时有 3 个加强、8 个减弱、9 个特异的蛋白质点。在缺铁与 EDTA 整合二价铁培养的水稻根中,也得到 23 个差异蛋白质点,其中缺铁时加强的有 4 个,减弱的有 2 个,特异的有 7 个。在对缺铁与 FeCl₃ 培养的水稻根中,未发现明显不同的蛋白质点。从缺铁处理的凝胶上选取 23 个点的蛋白质,用胰蛋白酶进行水解。继而进行了 MALDI-TOF 质谱检测,最终获得 12 种蛋白质的肽质量指纹谱的数据检索和分析结果。

4 展望

蛋白质组学在大规模蛋白质水平上为基因功能的研究提供了强有力的工具。蛋白质组研究已经成

为后基因组计划的研究热点,其研究技术与方法尚未完全成熟,但是这一领域的发展十分迅猛。尤其是模式植物蛋白质组方面,近几年来已经研究的更加火热。就当前进展来看,主要在以下四个方面:(1)拟南芥、水稻等植物的测序完成成为植物蛋白质组的研究提供了保证。(2)高灵敏度、高通量分析鉴定技术的发展如 MudPIT、ICAT 以及质谱技术等,可以直接分析双向凝胶电泳结果,实现分析自动化。(3)生物信息学的发展会给蛋白质组计划提供更方便更有效的计算机分析软件。(4)国际互联网会使各领域有关蛋白质组研究的成果出现新的集成,植物蛋白质数据库不断完善,人们可以从这里获得更多的信息。

此外,亚细胞水平蛋白质组的研究把蛋白质组学推向一个更广阔的领域,植物蛋白质组学的研究正在向遗传和育种方面(钱小红等,2002)深入发展,将来会有更多研究逆境胁迫下蛋白的功能,为改良品种提供有价值的参考。基因组与蛋白质组的接轨将会有奇迹出现,可能会揭开生命的奥妙。

参考文献(References):

- [1] 胡志远,贺福初.蛋白质组研究进展[J].生物化学与生物物理进展,1999,26(3):202-205.
- [2] 孙彤,李伟,刘祥林,等.缺铁胁迫诱导的水稻根转录本组和蛋白质组分析[J].全国第四届植物基因组学大会会议论文集,陕西杨凌,2003.
- [3] 钱小红,译.蛋白质组学:从序列到功能,2002,229-239.

(上接第 42 页)

5 问题与展望

如果上位性是重要的,那么目前应用的作物育种方法在短期内还是适用的,但对长期过程会有一定制约作用。DNA 标记技术为数量性状上位性效应的估计提供了极好的方法,其它基因组技术也将为在动植物改良中更好地利用上位性提供技术支持。尽管上位性的检测和估计一直被认为极其复杂,但如果能够利用标记可靠地鉴定出上位性效应,那么就能用以选择多基因座遗传型而不是按加性效应方式选择特定的 QTL 等位基因。今后面临的挑战是如何测定更为复杂的上位性以及发展在群体内、群体间选择以更好利用加性、显性和上位性效应的新方法。最终,育种者应用这些方法把性状优良、适应强的种质库与诸多尚未利用、性状适应性较差的遗传资源有机结合在一起。

参考文献(References):

- [1] Wade M J. Epistasis, complex traits, and mapping genes[J]. Genetics, 2001, 112: 59-69.

- [2] Wade M J. A gene's view of epistasis, selection and speciation[J]. J. Evo Biol, 2002, 15: 337-346.
- [3] Mackay T F C. The genetic architecture of quantitative traits[J]. Annual Review of Genetics, 2001, 35: 303-339.
- [4] Hemmison J, Hansen T E, Wagner G P. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection[J]. Am Nat, 2003, 161: 708-734.
- [5] Wagner G P. To epistasis - and beyond[J]. E, 2002, 56: 852-855.
- [6] Phillips P C. The language of gene interaction[J]. Genetics, 1998, 149: 1167-1171.
- [7] Fisher R A. The correlations between relatives on the supposition of Mendelian inheritance[J]. Trans Roy Soc, Edin, 1918, 52: 399-433.
- [8] Avery L, Wasserman S. Ordering gene function: the interpretation of epistasis in regulatory hierarchies[J]. Trends Genet, 1992, 8: 312-316.
- [9] Wolf J B, Brodie E D III, Wade M J. Epistasis and the Evolutionary process[M]. New York: Oxford University press, 2000-1-19.
- [10] 朱军. Mixed model approaches for estimating genetic variance and covariance [J]. 生物数学学报, 1992, 7(1): 1-11. 1455-1461.
- [11] Cheverud J M, Routman E J. Epistasis and its contribution to genetic variance components[J]. Genetics, 1995, 139: 1455-1461.
- [12] Wagner G P, Laubichler M D, Bagheri-Chaichian H. Genetic measurement theory of epistatic effects[J]. Genetica, 1998, 103: 569-580.