

数量性状发育遗传模型及其分析方法的研究进展

叶子弘 朱 军

(浙江大学农学系 杭州 华家池校区 310029)

摘要 发育遗传模型是同时反映性状遗传和发育本质、提供影响遗传变异及调整发育进程的有关因素的信息的模型。建立在群体遗传学基础上的直接效应模型适用于单一基因控制的简单性状。渐成模型将遗传变异分解成直接分量和渐成分量(母体效应和互作效应),能更好地反映有机体遗传和发育的生物学机制。生长轨迹模型有效地综合了复杂性状各分量的发育动态,可获得连续的、综合的、详细的、动态的发育信息。条件遗传分析方法不仅可以估算特定时间段的净效应,且可将净效应分解为不同遗传分量,了解各效应分量的相对贡献。

关键词 数量性状;发育遗传模型;分析方法

中图分类号:Q348;Q344; 文献标识码:A 文章编号:0253-977X(2001)01-0065-04

Advance in Developmental Genetic Models and Analysis Methods for Quantitative Traits

YE Zi-hong, ZHU Jun

(Agronomy Department, Zhejiang University, Hangzhou, Huajiachi Campus 310029, China)

Abstract Developmental genetic models and analysis methods for quantitative traits are presented. Developmental genetic models should reflect the genetic and developmental essence and provide the information of the factors influencing the genetic variation and the developmental process. Direct effect models, which based on the population genetics, may be suitable to analyze simple traits with single gene. Epigenetic models can decompose the whole genetic variation into direct and epigenetic components (maternal effects and epigenetic interaction effects), so that biological mechanism can be better understood. Growth trace models effectively synthesize the developmental dynamics of components of complex traits. With them, continuous, composite, detailed and dynamic information of development is available. Conditional analysis method can not only estimate the net effects in a specific time interval, but also decompose them into genetic components and help to appreciate the contributions of different effects.

Key words quantitative traits; developmental genetic models; analysis methods

了解复杂数量性状随时空动态变化的遗传机制是备受生物学家重视的一个问题。20世纪80年代以来通过突变体分离早期胚胎发育基因的研究,揭示了果蝇形态发生的分子机制,成为1995年诺贝尔生理或医学奖的主要内容,同时也标志着发育生物学已成为生物学的带头学科之一^[1]。迄今为止,遗传学家们已发展了大量的遗传模型及分析方法,用于研究发育的变化规律。随着发育生物学的发展,数量性状的发育遗传模型从简单的直接效应模型发展到包含渐成效应的渐成模型,及反映动态发育过程的生长轨迹模型。运用

求偏导-积分的数学方法可分析发育遗传模型,但无法处理复杂模型,无法获得任意两个发育阶段间的净效应。Zhu (1995)提出的条件遗传分析方法可有效解决这些问题^[2]。本文将系统地介绍数量性状的发育遗传模型及其分析方法的概况。

1 发育遗传模型的内涵

性状的发育是其各分量发育过程的综合^[3,4]。模型是对特定系统及其特征的一种简洁的描述和概括^[5]。从静态现象的简单描述到复杂动态过程的详细的数学描述,模型广泛地运

收稿日期 2000-01-06;修回日期 2000-03-18

基金项目 国家自然科学基金项目(39670390)资助

作者简介 叶子弘(1975-),女,汉,浙江人,博士研究生,专业方向 数量遗传。通讯作者 朱 军(1949-),男,汉,上海人,博士学位,教授,博导,专业方向 数量遗传。电话 0571-6971444, E-mail: zql@zju.edu.cn.

用于生物研究中。通常选用特定的模型探讨各发育过程的发生、速率、方向及对其他时期数量性状的影响。对于描述发育数量性状的遗传模型而言,它必须能同时反映性状遗传和发育的本质,必须注重性状的发育现象,提供有关影响遗传变异、调整发育进程的因素的信息。早期的发育遗传研究大部分只在某时间点上或在粗略划分的几个时间段、空间段上进行,且多数只考虑最终性状。现代发育遗传学已探明在生物体的不同发育阶段,基因是按一定的时空秩序有选择地表达的^[6,7]。控制复杂性状的遗传机制在个体发育过程中会发生显著的变化^[8,9]。由于基因表达的时空特性、细胞群体之间复杂的互作及随时间变化的渐成互作的存在,复杂数量性状也呈现出依时空变化的动态表现过程^[10]。国内外许多研究者以番茄、水稻、棉花等为遗传材料^[11-13]进行的不同发育时期遗传动态的研究证实了这一论点。由于控制数量性状的基因通常很多,各不同基因座对表现型的贡献不同,基因之间还可能存在着相互作用,且基因的表达易受环境的影响。所以,仅研究最终性状的遗传模型难以探明多基因体系在性状建成过程中的作用规律^[14]。因此,数量性状发育遗传模型必须能反映性状的遗传复杂性本质,体现发育为一个调整的、互作的、连续的、可遗传的、渐成的过程^[4]。它必须能探讨数量性状各构成因素的个体发育起源、发育轨迹、互作方式及综合成最终复合结构的机制。该模型还必须能反映发育过程的动态本质,包括不同控制因素的变异动态、发育遗传的动态变化。一个适宜的遗传模型不仅要求能有效地反映性状发育遗传的生物学本质,而且模型所包含的参数应比较简单,使模型可缩减且能进行假设检验^[10]。

生理和生态研究领域通常采用连续方程(如指数或对数生长模型)来检测群体及群体之间的相互作用随时间的动态变化,描述个体形态特征^[15]及研究共进化或捕食者-被捕食者之间的关系^[16]。常用的生长模型多为数学模型,通常是在总结群体生长规律的基础上得到的时间与检测目标(如体重、身高)之间的关系等式: $W = f_1(t)$ 或 $dW/dt = f_2(W)$ 。以个体为基本单位研究某性状的整体变化,预测群体发展趋势,能有效分析生长的遗传选择响应及环境变化响应。早期生长方程的参数是从一群检测对象的平均数中估算得到,忽略了参数会因检测对象的变化而变化,即使同一检测对象的不同生长过程,其参数也可不同^[17,18]。现代混合模型方法^[19]在模型中加入检测对象之间的变异性,可在同一个模型中同时获得固定效应和随机效应的估算。但生长模型具有一定的适用范围,如 Gompertz 方程不适于研究鼯鼠或其他具有延伸生长的哺乳动物^[20],且无法进一步研究引起这些生长变化的可能因素、各因素的相对贡献及引起生长发育的基因数目、表达水平、互作关系等。

区别于生长模型,发育遗传模型包含各遗传效应分量。从基因水平分析,发育遗传模型可获知控制发育数量性状的可能基因数目、等位基因频率、不同基因的贡献率。从基因表达的总效应分析,发育遗传模型可用于研究遗传和非遗传因素及其相互作用对表型变异的影响、个体发育过程中这些因素的变化规律、性状间遗传相关的根本原因、遗传选择响应的大小、方向及维数。不同阶段遗传方差的大小及遗传协

方差的大小和方向决定了性状发育的方向和程度^[21]。因此,数量性状的发育遗传模型可进一步分析复杂数量性状各构成因素的遗传本质及综合成最终复杂生物有机体的机制。

2 发育遗传模型

2.1 直接效应模型

早期的一些模型建立在传统的群体遗传模型的基础上,适用于单一基因控制的简单性状。但多数数量性状由一系列受环境因素影响并具有不同效应的基因控制。早在 1989 年, Falconer 就认为,各性状的表型值可分解为遗传和环境两个分量,基因型值可进一步分解为加性、显性和上位性。这种直接效应模型认为遗传信息完全来自于个体的基因组,是传统的孟德尔遗传模型,并假设(1)所有遗传因素以孟德尔遗传方式遗传(2)只包括一个世代(3)所有发育信息仅限于直接遗传分量(4)不存在遗传的渐成因素(5)基因型与环境互作及协方差的期望值为 0(6)性别之间为完全相关^[22]。生态学家和进化生物学家大量地运用直接效应模型来估算多基因控制性状发育变化的原因分量^[23-26]。简单的直接效应模型可预测性状的整体发育变化,但忽略了真实生命的遗传复杂性,无法洞悉复杂性状发育的内在机制;无法给出遗传变异和协同变异机制及各变异在发育过程中的变化方式。简单直接效应模型无法正确了解复杂生命体尤其哺乳动物的发育变化^[27-30]。此外,在个体发育过程中会产生包括母体效应和互作效应的渐成现象,它会影响到控制性状发育的基因的表达。因此,要描述复杂的发育过程,必须建立更全面的模型。

2.2 渐成模型

Atchley 和 Hall 于 1991 年认为渐成因子是来自发育细胞之外但影响细胞发育的可遗传的控制因子^[4]。不同基因、细胞、组织或器官之间的互作作用是生长、分化及形态渐成的基础,效应之间的相互关系即渐成互作。基因对一个或一群细胞进行调控时产生的对另一个或一群细胞的影响即渐成控制^[12]。通常渐成因子包括性状之间的互作作用和母体效应两部分。

Willham 于 1963 和 1972 年提出了一些包含母体效应的数量遗传模型^[27-28]。Kirkpatrick 和 Lande 于 1989 年及 Cowley 和 Atchley 于 1992 年提出在早期直接效应模型的基础上增加母体遗传效应,认为 $t+1$ 世代表现型值受本身加性和环境效应作用的同时还受 t 世代母体表现型的影响^[29,30]。鉴于性状之间存在随时空动态变化的互作作用, Cowley 和 Atchley 提出多变量渐成模型: $z = Dy = D(a + e)$, D 为渐成效应矩阵。通过对直接效应模型进行线性转换, Wagner 于 1989 年发展了多变量的发育遗传新模型,认为基因产物决定基因型值^[31]。Wagner 模型与 Cowley-Atchley 渐成模型均是从直接效应模型发展而来,具有一定的相似性,但 Wagner 模型是从基因水平上进行分析,而 Cowley-Atchley 渐成模型是在表型水平上的应用,因此 Wagner 模型又称为 Wagner 发育作图模型。此外, Wagner 模型除普通的环境效应向量 e 外,增加了一个新的环境效应 e^* (直接影响内在变异的环境效应),该效应与内在遗传效应(如加性效应)以相同的途径影响表型变异。

上述模型的直接效应仅包含加性效应,但显性和性连锁效应(尤其在动物遗传中)是不容忽视的重要因素。针对这一不足,Zhu 等于 1996 年在混合线性模型的基础上,提出包含加性、显性、性连锁和母体效应及基因型 \times 环境互作的遗传模型,不仅能估算加性、显性、性连锁及母体效应,且能进一步预测它们与环境的互作效应^[32]。

渐成模型将遗传变异分解成内在的、直接的分量和渐成分量(母体效应和互作效应),能更好地反映有机体遗传的和发育的生物学机制,更精确地预测发育变化的大小和方向^[2]。但渐成模型的应用包含了精确的、复杂的试验设计和繁琐的计算,难以对模型进行简化,且渐成模型对遗传变异的分解仍很粗略,未包含上位性,也未对母体效应及互作效应进行进一步的分解。

2.3 生长轨迹模型

组成复杂性状的各分量具有各自的个体发育历史或称“发育轨迹”(即连续发育时间段内任意两点之间的不同特征)^[3,33]。不同发育时间段里,各分量的“发育轨迹”呈现不同“形状”,性状的遗传和非遗传控制因子有着不同的贡献^[3,6]。因此,各分量之间的遗传协方差结构在个体发育过程中会发生变化^[10]。各分量在不同阶段的遗传方差的大小及遗传协方差的大小和方向决定了性状发育的方向和程度^[21]。

目前所用的大部分模型为静态模型,并有一隐含假设:控制性状发育的因子在个体发育过程中不发生变化^[10]。这类模型可简化计算,减少工作量,但它只能给出性状在各时间点上的发育状态,所得到的是一系列间断的信息。为探求性状发育的动态本质,Kirkpatrick 等于 1989 和 1990 年在调查了连续选择对生长过程的影响的基础上提出连续模型^[29,34]。早期的发育数量遗传模型对于具有离散发育阶段的生物体(如:昆虫)比较适宜,而 Kirkpatrick 的连续模型对于大部分的脊椎动物均适用^[10]。但 Kirkpatrick 模型是单变量模型,只能对数量性状的整体变化或某个分量的发育状态进行分析。针对这一缺陷,1994 年,Atchley 等将具有不同发育轨迹的分量综合成复杂数量性状的发育方程:

$$dx_i(t)/dt = F_i[x(t)]x_i(t),$$

各分量的生长轨迹由模型所包含的常量 a_{ij} 、系数 b_{ij} 和初始条件决定^[10]。该模型假设 t 时刻的变化是由细胞数目的变化引起,而不是细胞大小或细胞排列方式的变化。如所有的 $i, j, b_{ij} = 0$, 则各分量相互独立,呈指数增长;如 $b_{ii} < 0$ 且 $b_{ij} = b_{ji} = 0$ (对于 $i \neq j$) 则各分量相互独立,呈对数增长,因此,该模型不仅包含了生长模型所能提供的信息,还能对生长发育进行遗传解释。生长轨迹模型有效地综合了复杂性状各分量的发育动态,可获得连续的、综合的、详细的、动态的发育信息。但在该类模型中遗传效应只分解为加性、母体及环境效应,而忽略了显性、上位性及基因型 \times 环境互作效应。

3 分析方法

3.1 非条件分析方法

发育数量性状在 t 时刻的遗传效应(G_t)是 $t-1$ 时刻的

遗传效应(G_{t-1})和从 $t-1$ 时刻到 t 时刻的额外遗传效应(G_d)的总和。发育遗传研究常用性状在不同时期的遗传方差分量及遗传协方差分量(性状在 $t-1$ 时刻的表现与在 t 时刻的表现之间的协方差)来解释群体的变异。该方法可有效估算从初始时刻到 t 时刻的基因累加效应,但无法给出从 $t-1$ 时刻到 t 时刻的净基因表达效应。为进一步探讨两时间点之间的遗传和发育变化,发育遗传研究运用求偏导的数学方法获得连续发育阶段之间的近似生长速率,或运用两时间点表型值间的差值作为新的观察值进行分析。由于 $t-1$ 时刻到 t 时刻是个连续的过程,它们之间存在相关性。

运用 $\Delta x_i(t) = x_i(t) - x_i(t-1)$ 无法求得真正的 $t-1$ 到 t 时刻的净效应^[21] 因而运用差值进行发育遗传分析是有偏的。运用求偏导的方法分析发育模型会受到复杂的数学运算的限制,很难在一个模型中包含较多的遗传参数。目前应用较多的发育数量遗传模型大部分为只包含加性、母体及环境效应的简单模型。但大量研究表明,显性、上位性及基因型 \times 环境互作效应也是影响性状发育变化的重要因素,而且在有上位性的情况下,假定上位性不存在的遗传模型可能导致研究结果有偏^[35,36]。

3.2 条件分析方法

基于单变量或多变量统计的条件分析方法可运用于不同的领域。在经济学方面,Engle 于 1982 年提出用于分析自回归条件异方差(ARCH)模型的条件方差估算方法^[37]。在动物育种中,Henderson 于 1985 和 1988 年将条件分析方法用于给定入选基础群体的遗传方差估算^[38,39]。Engle 的方法可以利用时间序列模型的机误估算出条件方差进行精确的时间参数化,但不能估算非机误的各种条件方差分量^[2]。Henderson 的方法可以估算入选基础群体的条件遗传方差,但不能估算发育数量遗传学中依时间顺序测定的条件遗传方差分量^[2]。1995 年 Zhu 提出基于混合线性模型的条件遗传分析方法,运用多元统计思路定义新变量 y^* (即 $y_i | y_{i-1}$)、 y^* 与 $t-1$ 时刻的表现型值 y_{i-1} 是相互独立的,其协方差 $Cov(y^* | y_{i-1})$ 为零矩阵。实际应用时,用方差矩阵和协方差矩阵的无偏估计值代替未知的方差矩阵和协方差矩阵。因此,运用该方法不仅可以获得两时间点之间的净效应,且可将净效应分解为加性、显性等遗传分量,可了解各效应分量的相对贡献。条件分析方法已在棉花、小鼠及水稻等的发育遗传研究中得到广泛应用^[2,9,40,41]。Zhu 运用条件遗传分析方法对陆地棉不同发育阶段平均结铃数进行研究后发现,控制发育性状表现的累加遗传效应在还未被非条件遗传分析方法检测出前,已有基因表达产生遗传效应^[2]。分子标记技术的飞速发展大大促进了发育遗传研究。通过应用分子标记连锁图谱和 QTL 作图技术,已能估算控制复杂数量性状的基因座数及其在基因组中的位置、作用方式、遗传效应及与其他 QTLs 的互作效应等特征。

参考文献(References):

- [1] Li SH. Zebrafish as the best vertebrate mode for development studies [J]. Growth, Development and Aging, 1997, 6: 91~97.

- [2] Zhu J. Analysis of Conditional Genetic Effects and Variance Components in Developmental Genetics [J]. *Genetics* ,1995 ,141 :1633 ~ 1639.
- [3] Atchley WR. Developmental quantitative genetics and evolution of ontogenies [J]. *Evolution* ,1987 ,41 :316 ~ 330.
- [4] Atchley WR ,Hall BK. A model for development and evolution of complex morphological structures [J]. *Bio Re* ,1991 ,66 :101 ~ 157.
- [5] Bamhart RK. The American Heritage Dictionary of Science [M]. Boston :Houghton Mifflin Company ,1986.
- [6] Stewart AD ,Hunt DM. The genetics basis of development [M]. Blackie ,Glasgow and London ,1982.
- [7] 杨太兴 ,曾孟潜. 玉米胚乳和胚发育过程中干物质积累和同工酶、蛋白质的变化 [J]. *植物生理学报* ,1984 ,10(1) :77 ~ 86.
- [8] Atchley WR. Ontogeny ,timing of development and genetic variance-covariance structure [J]. *Am Nat* ,1984 ,124 :519 ~ 540.
- [9] Atchley WR ,*et al.* Developmental Quantitative Genetics ,Conditional Epigenetic Variability and Growth in Mice [J]. *Genetics* ,1997 ,147 :765 ~ 776.
- [10] Atchley WR ,*et al.* Developmental quantitative genetic models of evolutionary change [J]. *Developmental Genetics* ,1994 ,15 :92 ~ 103.
- [11] Kherialla AI ,Whittington AJ. Genetic analysis of growth in tomato the F_1 generation [J]. *Ann Bot* ,1962 ,26(104) :489 ~ 504.
- [12] Xu YB ,*et al.* Diallel analysis of tiller number at different growth stages in rice (*Oryza sativa* L.) [J]. *Theor Appl Genet* ,1991 ,83 :243 ~ 249.
- [13] 朱 军 等. 陆地棉(*Gossypium hirsutum* L.) 花铃动态的遗传分析 [A]. 北京国际棉花学术讨论会论文集 C [J]. 1992.
- [14] 吴国海. 玉米几个数量性状在不同发育阶段的基因效应分析 [J]. *遗传学报* ,1987 ,14(5) :363 ~ 369.
- [15] Parks JR. A Theory of Feeding and Growth in Animals [M]. Berlin :Spring-Verlag ,1982.
- [16] Murray JD. *Mathematical Biology* [M]. Berlin :Spring-Verlag ,1989.
- [17] Laird AK. Evolution of the human growth curve [J]. *Growth ,Development and Aging* ,1967 ,54 :155 ~ 165.
- [18] Laird AK ,Ware JH. Random-effects models for longitudinal data [J]. *Biometrics* ,1982 ,38 :963 ~ 974.
- [19] Lindstrom MJ ,Bates DM. Nonlinear mixed effects models for repeated measures data [J]. *Biometrics* ,1990 ,46 :673 ~ 687.
- [20] Begall S. The application of the Gompertz model to describe body growth [J]. *Growth ,Development and Aging* ,1997 ,61 :61 ~ 67.
- [21] Falconer DS. *Introduction to quantitative genetics* [M]. New York :Longman ,1981.
- [22] Falconer DS. *Introduction to quantitative genetics* [M]. New York :Longman ,1989.
- [23] Istock C. Some theoretical considerations concerning life history evolution [A]. Dingle H and Hegmann J. *Evolution and genetics of life histories* [C]. New York :Springer-Verlag ,1982 ,Pages 7 ~ 20.
- [24] Lande R. A quantitative genetic theory of life history evolution [J]. *Ecolology* ,1982 ,63 :607 ~ 615.
- [25] Lande RM ,Arnold SJ. The measurement of selection on correlated characters [J]. *Evolution* ,1983 ,37 :1210 ~ 1226.
- [26] Arnold SJ ,Wade MJ. On the measurement of natural and sexual selection theory [J]. *Evolution* ,1984 ,38 :709 ~ 719.
- [27] Willham RL. The covariance between relatives for characters composed of components contributed by related individuals [J]. *Biometrics* ,1963 ,19 :18 ~ 27.
- [28] Willham RL. The role of maternal effects in animal breeding :III. Biometrical aspects of maternal effects in animals [J]. *J Anim Sci* ,1972 ,35 :1288 ~ 1293.
- [29] Kirkpatrick M ,Lande R. Selection inheritance and evolution of maternal characters [J]. *Evolution* ,1989 ,43 :485 ~ 503.
- [30] Cowley DE ,Atchley WR. Quantitative genetic models for development , epigenetic selection and phenotypic evolution [J]. *Evolution* ,1992 ,46 :495 ~ 518.
- [31] Wagner GP. Multivariate mutation-selection balance with constrained pleiotropic effects [J]. *Genetics* ,1989 ,122 :233 ~ 234.
- [32] Zhu J ,Weir BS. Diallel analysis for sex-linked and maternal effects [J]. *Theor Appl Genet* ,1996 ,92 :1 ~ 9.
- [33] Alberch P ,*et al.* Size and shape in ontogeny and phylogen [J]. *Paleobiology* ,1979 ,5 :296 ~ 317.
- [34] Kirkpatrick M ,*et al.* Analysis of inheritance ,selection and evolution of growth trajectories [J]. *Genetics* ,1990 ,124 :979 ~ 993.
- [35] Li ZK ,*et al.* Epistasis for three grain yield components in rice (*Oryza sativa* L.) [J]. *Genetics* ,1997 ,45 :453 ~ 465.
- [36] Wang DL ,*et al.* Mapping QTLs with epistatic effects and genotype \times environment interactions by mixed linear model approaches [J]. *Theor Appl Genet* ,1999.
- [37] Engle RF. Autogressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation [J]. *Econometrica* ,1982 ,50 :987 ~ 1007.
- [38] Henderson CR. Best linear unbiased prediction using numerator relationship matrices derived for selected base populations [J]. *J Dairy Sci* ,1985 ,68 :443 ~ 448.
- [39] Henderson CR. A simple method to account for selected base populations [J]. *J Dairy Sci* ,1988 ,71 :3399 ~ 3404.
- [40] Yan ,*et al.* Molecular Dissection of Developmental Behavior of Plant Height in Rice (*Oryza Sativa* L.) [J]. *Genetics* ,1998a ,150 :1257 ~ 1265.
- [41] Yan ,*et al.* Quantitative Trait Loci Analysis for Developmental Behavior of Tiller Number in Rice (*Oryza Sativa* L.) [J]. *Theor Appl Genet* ,1998b ,97 :267 ~ 274.
- [42] 朱 军 ,段家龙. Genetic models and analysis methods for sex-linked and maternal gene effects [J]. *生物数学学报* ,1994 ,4(4) :1 ~ 9.
- [43] 朱 军. 包括基因型环境互作效应的种子遗传模型及其分析方法 [J]. *遗传学报* ,1996 ,23(1) :56 ~ 68.