

数量遗传学理论框架下的进化生态学研究： 以动物模型的运用为例

王代平^① 夏灿玮^② 刘阳^③ 孙悦华^④

① 德国马普鸟类学研究所, 行为生态与遗传进化学院 Seewiesen 德国 82319; ② 北京师范大学生命科学院, 生物多样性与生态工程教育部重点实验室 北京 100875; ③ 中山大学生命科学院, 有害生物控制与资源利用国家重点实验室 广州 510275; ④ 中国科学院动物研究所 北京 100101

摘要: 目前, 生态学家越来越关注深入的生物学问题, 例如, 1) 理解生态和进化过程的互作和关系; 2) 种群中一个重要的表型特征, 受遗传基因影响多大? 即其可遗传程度, 表示该性状的进化潜能; 3) 基因是怎样影响表型性状, 及其对应的个体适合度以及种群动态? 4) 决定多个重要表型性状的基因之间关系和互作如何? 随着生物统计软件尤其是线性混合模型的发展, 结合经典数量遗传学的理论, 发展出了针对上述问题的动物模型 (Animal Model), 使得我们可以对野外种群进行上述研究。本文首先介绍了经典数量遗传学的重要概念, 随后在其理论框架下, 举例介绍了动物模型的操作和使用, 最后探讨和展望了利用数量遗传学方法进行进化生态学研究的前景。

关键词: 遗传力; 遗传相关; 混合模型; 数量遗传学; 进化生态学; 自然选择; 适合度

中图分类号: Q958 **文献标识码:** A **文章编号:** 0250-3263 (2017) 04-690-12

The Theory of Quantitative Genetics into Evolutionary Ecological Studies: the Applications of Animal Model

WANG Dai-Ping^① XIA Can-Wei^② LIU Yang^③ SUN Yue-Hua^④

① *Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard-Gwinner-Str. 7, 82319 Seewiesen, Germany;* ② *Ministry of Education Key Laboratory for Biodiversity and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875;* ③ *School of Life Sciences, Sun Yat-sen University, Guangzhou 510275;* ④ *Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China*

Abstract: Over the last decades field ecologists are more and more interested in understanding deeper biological questions, for instance: 1) the links between evolutionary and ecological dynamics. 2) For an important trait, to what extent determined by genes (as the degree of heritable reflecting the trait's evolutionary potential)? 3) How do genes influence phenotypes, fitness and population dynamics? 4) How is the genetic correlation between important traits? As the development of biological statistics, especially the mixed effect model, combined classical theory of quantitative genetics, a professional model (Animal Model)

第一作者简介 王代平, 男, 博士研究生; 研究方向: 进化生物学, 数量遗传学, 行为生态学, 分子生态学; E-mail: dwang@om.mpg.de。

收稿日期: 2016-11-26, 修回日期: 2017-02-26 DOI: 10.13859/j.cjz.201704020

provided us opportunity to carry out robust studies on wild populations. In this review, we introduce some basic conceptions of classical quantitative genetics, and then, by using typical examples, we focus on how to use the Animal Model. Finally, we discuss the future of integrating quantitative genetics and evolutionary ecology in the wild.

Key words: Heritability; Genetic correlation; Mixed effects model; Quantitative genetics; Evolutionary ecology; Natural selection; Fitness

生态学家一直以来关注种群内(间)个体间的变异,正如达尔文所述:“这些个体变异(individual differences)为自然选择的作用提供了基质模板”(Darwin 1859)。传统的生态学往往只涉及到这些差异的表型(表型变异: phenotypic variance)层面,例如,测量一个重要的适合度表型性状在种群内的分布,或者寻找多个重要表型性状之间的相关性。然而,进化学家不仅仅关注某一性状在个体间的表型变异,还要深入到其基因层面,即更加侧重这个性状的可遗传性(可遗传变异: genetic variance)。因为可遗传变异是决定一个性状在选择(包括自然和人工选择)下是否进化,以及进化速度和潜能的关键,而且只有理清影响种群生活史、适合度等各种性状的可遗传性,才能真正理解该种群的进化过程(Ellegren et al. 2008, Kruuk et al. 2008, Johnston et al. 2013)。近年来,研究种群动态中自然选择的作用以及小尺度的进化成为热点(Hairston et al. 2005, Saccheri et al. 2006, Kasada 2014)。

数量遗传学是一门定量研究个体可遗传变异的学科。自 Fisher (1918, 1958) 建立该学科理论框架,再经 Haldane (1932) 和 Wright (1921) 等的总结和完善的完善以来,该学科在进化生物学和动植物育种研究中发挥了重要作用。它不仅能够建立种群性状的遗传结构,还能预测表型性状的进化潜能(Falconer et al. 1996, Lynch et al. 1998, Wilson et al. 2016)。例如,如果想研究某种鸟类的一个重要生活史性状“窝卵数”,在确认该种群遗传家系背景下,运用数量遗传学方法可以定量计算出这个性状变异在多大程度上是可遗传的(遗传力:

heritability),又有多少是受母性作用(maternal effect)的,以及环境作用在决定这个性状表型中占有的分量。

数量遗传学有着近 100 年的历史,但由于性状的遗传参数计算必须在确认该种群的遗传家系后才能完成,所以过去一直局限在动植物育种相关研究中。进入 21 世纪以来,随着分子生态学的发展,特别是亲本鉴定技术,加上计算机庞大的统计和模拟计算能力(如贝叶斯统计),使得数量遗传学方法在野外进化生态学的运用成为可能。近年来,国外一些拥有长期积累数据的课题组已经成功地运用数量遗传学,解决了在自然条件下种群进化动态的一些难题。经典的例子有:欧洲马鹿(*Cervus elaphus*)研究(Kruuk et al. 2000, Kruuk 2004, Kruuk et al. 2008, Stopher et al. 2012)及白颈姬鹀(*Ficedula albicollis*)研究(Merila et al. 2000, 2001a, b, Garant et al. 2004, Brommer et al. 2005)。这些研究突破了传统进化生态学的研究瓶颈,使得宏观生态学深入到基因遗传层次。野外数量遗传学在进化生态学中的应用从此成为新的研究热点,并广泛扩展到其他物种:如绵羊(*Ovis aries*)(Hayward et al. 2009, Robinson et al. 2009, Wilson et al. 2009, Hayward et al. 2013)、青山雀(*Cyanistes caeruleus*)(Hadfield et al. 2006a, b, 2007, 2013a, b)、歌带鸫(*Melospiza melodia*)(Reid et al. 2011a, b, 2012, 2014)、斑胸草雀(*Taeniopygia guttata*)(Forstmeier et al. 2009, Schielzeth et al. 2010, Forstmeier et al. 2011, Husby et al. 2013)。上述研究有一个共同点,大都以野外种群为研究对象,运用一个特殊的混合统计模型,即动物模

型 (animal model), 利用已知的遗传家系, 将一个性状的表型变异分解成遗传 (genetic) 和环境 (environmental) 变异两大部分, 进而这两大部分各自又可以分解成相应的小部分, 如, 遗传变异接着可以分解成加性遗传变异 (additive genetic component, 其定义和解释见本文后文)、母系遗传变异 (maternal effect)、基因显性变异 (dominance) 等。上述研究与传统的动植物育种有着很大的差别, 首先, 动植物育种者主要关注其研究对象的经济价值, 利用人工选择的方法改良品种, 而上述研究着重于自然选择条件下种群的进化动态; 其次, 动植物育种研究是在稳定的环境条件下, 种群的遗传家系很容易得到, 而对于野外进化学生态学研究, 获得一个研究种群的遗传家系充满着挑战, 且其所在环境也在不停地变动。以上这些差异, 使得野外生态学家发展了自己的理论和研究工具, 产生一个新的学科, 称之为野外数量遗传学 (wild quantitative genetics)

本文将重点介绍动物模型的研究方法, 包括介绍数量遗传学的基本理论和重要的概念, 以及应用动物模型所需的数据, 即野外数据收集类型。我们将以模拟数据为例, 介绍几款代表软件的操作, 包括数据的输入, 软件代码, 以及结果分析, 并讨论如何在实际中利用动物模型去验证假说, 同时探讨该模型的延伸和未来的发展方向。

1 数量遗传学基本理论和重要概念

1.1 基本理论

经典数量遗传学内容主要是在已知种群个体间遗传关系 (家系) 和某个表型性状 (直接测量获得) 的情况下, 对该性状的遗传强度和进化潜能进行推算 (Wilson et al. 2010)。具体来讲, 对于某一个表型特征, 如果发现具有亲缘关系的个体 (共有很多基因) 的相似度比一些没有亲缘关系的个体 (几乎不共有基因) 明显大, 则认为在这个表型特征的表达中, 基因可能起着很强的作用, 也就是说这个表型性状

在某种程度上是可遗传的。孟德尔的豌豆实验其实就是最简单的例子。豌豆种子的形状 (圆形还是皱形) 是二项分布 (仅仅只有两种表型), 而且决定该表型性状的基因也很简单 (单基因调控)。然而, 自然界中大部分的表型性状是连续分布的, 而决定这些连续性状的又往往是很多基因的共同作用。针对以上情形, 数量遗传学应运而生 (Falconer et al. 1996, Lynch et al. 1998), 其本质是孟德尔遗传定律的延伸和扩展 (从单个基因决定的离散变量性状扩展到多个基因决定的连续变量性状), 在其基础上再整合自然种群的进化动态 (自然选择), 可以说, 数量遗传学是经典进化生物学 (达尔文) 和经典遗传学 (孟德尔) 的完美结合。

1.2 遗传力 (heritability, h^2)

对于单个表型性状, 我们能计算出相对于整个表型变异 (phenotypic variance, V_P) 中由遗传决定的部分 (V_G) (Falconer et al. 1996)。继而, V_G 可以接着分解为加性遗传部分 (additive genetic variance, V_A , 决定该性状的所有基因所起的线性作用总和, 称为加性遗传变异), 基因显性效应部分 (dominance, V_D , 除了线性作用, 在一个座位子内, 存在基因间显性作用, 这些显性作用的总和即 V_D), 基因相互效应部分 (epistatic, V_I , 座位基因型之间存在的相互作用总和)。在实际情况中, 后两部分很难计算出来, 其需要庞大而且完整平衡的家系和人工育种操作, 而且一般占 V_P 的比重不大, 因此, 不管是生态学家还是动植物育种者, 都关注于通过测量具有亲缘关系的个体表型来推算 V_A 的大小比重 (Falconer et al. 1996, Kruuk 2004)。为了简单起见, 一般将 V_D 、 V_I 和环境决定的部分 (V_E) 归为残差 (residual, V_R), 因而, 整个表型变异简单的分解方程为 $V_P = V_A + V_R$ 。狭义的遗传力 (heritability, h^2) 即是指 V_A 占总体表型变异 V_P 的比重,

$$h^2 = V_A / V_P \quad \text{公式 1}$$

这是数量遗传学的一个最重要的概念, 遗传力描述了具有亲缘关系个体之间的相似程度, 表

示一个性状的可遗传强度, 也反映出该性状的进化潜能 (在自然选择作用下的反应强度)。

1.3 遗传相关 (genetic correlation, r_G)

表型变异的分解可以扩展到多个性状, 我们可以分解这些性状间的表型协变异 (COV_P), 得出加性遗传协变异 (COV_A) 所占比重, 即性状间遗传相关 (genetic correlation, r_G), 这是数量遗传学的另一个重要概念。遗传相关表示几个性状间表型下的基因联系, 正的遗传相关可以认为这些性状间有一部分是由相同的基因决定的 (基因的多重作用: pleiotropy) 或者决定这些性状的基因连锁在一起 (linkage)。

遗传相关可以验证很多的生物学问题, 例如, 以性二型的雄性孔雀鱼 (*Poecilia reticulata*) 为研究对象, 验证性选择和自然选择之间的冲突和权衡 (conflict and trade-off) 的研究中发现, 雄性孔雀鱼的性选择特征如雄性吸引力 (sexual attractiveness) 与存活寿命存在一个负遗传相关关系 (Brooks 2000)。传统的种群生态学往往热衷于寻找重要适合度性状间的相关性, 或者一个表型性状与适合度特征的相关性。这种方法其实存在较大的缺陷, 这种表型的相关关系会因为环境因子而造成很大的偏差。例如性状 X 与 Y 存在一个表型相关 r_P , 有公式 (Willis 1991):

$$r_P = r_G h_X h_Y + r_E e_X e_Y \quad \text{公式 2}$$

其中, r_G 为 X 与 Y 之间的遗传相关, h_X 和 h_Y 分别为性状 X 和 Y 的遗传力 h_X^2 和 h_Y^2 的平方根, r_E 为由于环境因子而造成 X 与 Y 之间的相关性 (可正可负), $e_X = \sqrt{1 - h_X^2}$, $e_Y = \sqrt{1 - h_Y^2}$ 。Kruuk 等 (2008) 总结和比较了来自 24 个研究的 282 个表型相关值 (r_P) 和遗传相关值 (r_G), 得出其相差值 D ($|r_P - r_G|$) 平均值为 0.245 (SD = 0.222)。由此可见, 由环境引起的表型相关是影响较大的干扰因素, 数量遗传学家建议应该对性状的表型相关表示警惕 (Willis 1991, Hadfield et al. 2007), 研究中需

想办法排除。因此, 在进化生态学研究中, 不应只简单测量几个性状之间的表型相关, 而是应该尽可能得出这些性状之间的遗传相关关系。

1.4 重复力 (repeatability, R)

重复力是一个很简单但是很重要的概念, 主要是针对种群中每个个体有多次测量的情况。在数量遗传学中, 重复力的计算公式为 (Falconer et al. 1996):

$$R = \frac{V_G + V_{Eg}}{V_P} = \frac{V_A + V_D + V_I + V_{Eg}}{V_P} \quad \text{公式 3}$$

其中, V_G 表示个体遗传变异部分 (包括 V_A 、 V_D 、 V_I), V_{Eg} 为个体的永久环境效应 (也可理解为种群一般环境效应), V_P 为表型值。简单地说, 对于每个个体, 对其表型值多次测量, 我们希望搞清这几次测量值中不变的部分 (即可重复部分) 的比重, 而这种可重复的部分包括其遗传变异部分 (V_G) 和永久的环境效应部分 V_{Eg} , 后者这部分不是由基因决定, 但是也不会随着环境的改变而改变, 因而称其为永久环境效应。

最简单的重复力计算方法即对种群中每个个体测量 2 次, 然后计算第一次测量值与第二次测量值的相关系数即可, 该相关系数即为重复力值。重复力是一个极其重要的概念, 其意义可概括为两方面: 第一, 从数量遗传学角度来说, 重复力可以视为遗传力的一个上限值, 从公式 3 中可以看出, 遗传力 $h^2 = V_A / V_P$ 不可能大于重复力。对于一个重要的性状, 计算其遗传力可能比较困难 (需要提供遗传家系信息), 但是计算重复力却非常简单, 只需多次测量即可, 因此, 野外生态学家尤其行为生态学家非常关注每一个性状的重复力, 因为一旦知道该性状重复力大小, 即可近似估算出其遗传力的大小。第二, 从研究的可重复角度来说, 重复力也是一个非常重要的概念。近期以来, 研究的可重复性一直受到科学家的广泛关注, 以生物医药研究为例, 美国每年大概有 250 亿美元的生物医药研究是不可重复的 (Freedman et al. 2015)。在生态学和进化生物学领域, 研

究统计发现,效应量(effect size)一般都偏小(Jennions et al. 2002),推断该领域研究的可重复性较低。因此应该重视生态进化学研究的可重复性,对于某一个重要的研究表型性状,第一步需确认的是其重复力大小如何。如果重复力很低,即意味着该性状受环境影响很大,可遗传度不高(低遗传力),也就意味着其进化潜能较低。

2 遗传力和遗传相关计算方法,动物模型(Animal Model)

2.1 经典计算方法

计算遗传力 h^2 和遗传相关 r_G 的方法有多种,根据遗传家系和孟德尔分离定律,来推算这些参数。这里只介绍最简单直接的一种,用于了解动物模型的计算原理:父母表型平均值-子代表型值回归(Falconer et al. 1996)。根据孟德尔分离定律,对于二倍体生物,父本和母本各传给子代 1/2 基因。因此,在一个种群中只需测量每一个家庭父母本的表型平均值[(父亲表型值 + 母亲表型值) / 2]和相应子代的表型值,然后进行线性回归,即可知遗传力 h^2 。因为在线性回归过程中,环境因子等造成的偏差会分离掉。对于这种计算方法,笔者认为其实可以用一个极端的例子来理解:在符合哈-温平衡的种群中,其中具体某一代的某个性状,设其表型平均值 V_{P1} 为 P_1 ,其加性遗传变异值 V_{A1} 为 A_1 ,其环境部分变异平均值相对该表型性状平均值 V_{E1} 为 0(因为环境对于某一个个体表型值影响可正可负,对于该代的所有个体来讲,总平均值 $V_{E1} = 0$)。如果这一代子代的表型值平均值 V_{P2} (P_2) 等于其亲本一代 V_{P1} (P_1),即 $P_2 = P_1$ (也就是回归斜率等于 1),则该性状 $h^2 = 1$, 表示该性状的可遗传强度达 100%。

2.2 动物模型(animal model)

动物模型作为一个混合模型,最先在动物育种中使用(Henderson 1953, 1984, Thompson 2008),因而称其为动物模型。进化生态学在野外种群研究中引进动物模型比较晚(Shaw 1987),进入 21 世纪才开始大量关注和使用动

物模型。动物模型能利用种群中所有的亲缘关系,还允许种群中一些亲缘关系不明的个体存在,而且还允许加入其他随机变量和固定变量(例如年龄、性别、生境类型等),在验证一些特定假说的同时也排除了一些相对不重要的因子,使遗传参数计算更为准确(Wilson et al. 2010),因此动物模型成为在野外进化生态学研究中最普及的方法。

动物模型的算法和统计与普通的线性混合模型一样,如限制性最大似然法(restricted maximum likelihood, REML)、贝叶斯统计(Markov chain Monte Carlo, MCMC),为避免繁琐的数学算法,这里不详细论述。其计算遗传参数(遗传力、遗传相关等)的统计原理,其实与上述经典的计算方法相同,而且更为强大:上述经典的计算方法需利用特定的家系关系,如父母亲平均值-子代,而动物模型能整合各种亲缘信息,例如,祖父母与第三代关系、全胞(full-sibling)兄弟姐妹关系、半胞(half-sibling)关系、表亲关系等等,该模型自动将研究种群的所有家系信息整合,存入到一个独立的空间,再和另一个所需数据文件(即该种群某个性状的表型值)整合。最后算出该种群某一个或多个性状的遗传参数。

动物模型分为单变量和多变量动物模型。单变量动物模型主要为了测量种群某个性状(作为因变量)的遗传力和其他参数,例如,除了该模型自动整合遗传家系后将每个个体 ID(个体的编号)作为随机自变量,还可以添加其他的随机变量,例如母亲的 ID 来测量该性状母系遗传作用(见电子版附件 1)。而多变量动物模型是指多个表型性状作为因变量的动物模型,它不仅计算得出其中每一个表型性状的遗传力,还可以计算出这些表型性状间的遗传相关(见电子版附件 2)。

3 应用动物模型所需的数据

动物模型运行所需要的数据分两部分,第一部分是研究种群的遗传家系,举个简单例子,假如研究的某个种群有如图 1 家系。

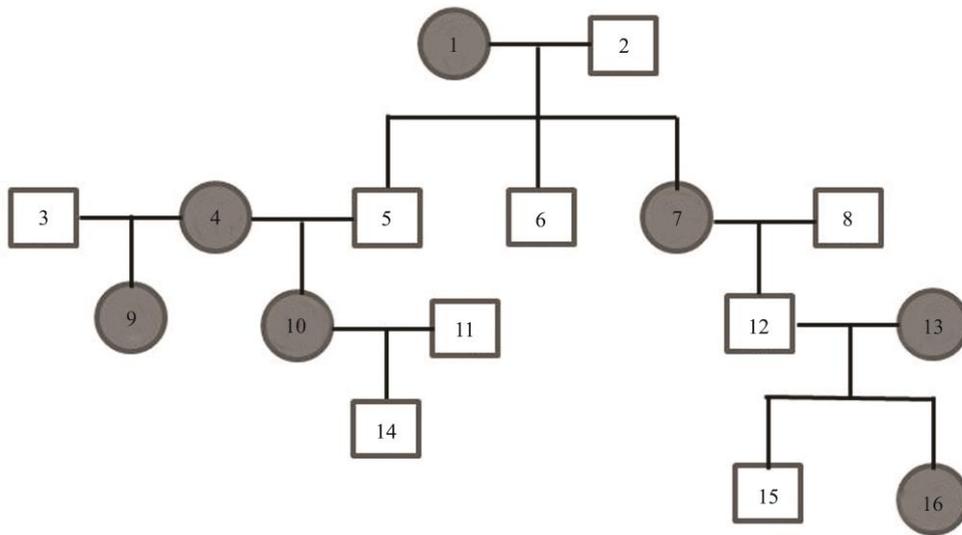


图 1 某个种群的家系图谱示意图

Fig. 1 The pedigree information of one population

圆形表示雌性个体，方形为雄性个体。Females as circles and males as squares.

动物模型要求的数据储存文件格式在 R 语言中.txt 即可（可以先在 excel 中输入数据，再转为.txt 格式输入到 R）。在这个例子中，excel 的数据表格为 3 列，第一列为个体 ID，第二列为父本 ID，第三列为母本 ID，不确定其父母的个体视为数据缺失 NA。上述的家系转换为可输入数据表格形式见表 1。

第二部分即为该种群某个性状的表型测量值，这部分数据跟一般的野外研究所需数据一样，没有太大的差异。值得一提的是，每个个体的表型值应该多次测量，因为个体 ID 将作为动物模型的随机变量，另外，多次测量可以使动物模型运算更加精确，也可以从这个性状的重复力预估其遗传力的大小。

4 可用软件

野外数量遗传学家们已经编码了使用动物模型专门的软件，如 ASReml 和 VCE 等，或者依托强大的 R 语言平台，设计了专用的软件包（如 MCMCglmm 和 ASReml_R 等）。参照 Wilson 等（2009）在表 2 中列出了能运行动物模型的常用软件。该表列出的所有软件对于动

表 1 图 1 家系图谱信息相应的表格数据形式

Table 1 data structure as required by software packages in R (corresponding to Fig. 1)

ID	Father	Mother
1	NA	NA
2	NA	NA
3	NA	NA
4	NA	NA
5	2	1
6	2	1
7	2	1
8	NA	NA
9	3	4
10	5	4
11	NA	NA
12	8	7
13	NA	NA
14	11	10
15	12	13
16	12	13

在此表格中，ID 为个体标号，Father 为该个体的父本标号，Mother 为该个体母亲标号。NA. 不确定其父母的个体。

Here, ID means the identity of individual, father referred the individual's father and mother referred its mother.

表 2 动物模型分析的常用的软件

Table 2 List of software packages that can be used to perform animal models

软件 Software packages	是否免费 Free or not	软件算法 Inference	软件网站 Website
ASReml	否 no	REML	http://www.vsnl.co.uk/software/asrem/
ASReml-R	否 no	REML	http://www.vsnl.co.uk/software/asrem/
DMU	是 yes	REML or MCMC	http://www.dmu.agrsci.dk/
MCMCglmm	是 yes	MCMC	http://cran.rproject.org/web/packages/MCMCglmm/index.html
WOMBAT	是 yes	REML	http://agbu.une.edu.au/~kmeyer/wombat.html
VCE	是 yes	REML or MCMC	http://vce.tzv.fal.de/software

REML. 限制性最大似然估计; MCMC. 贝叶斯统计。

REML. Restricted maximum likelihood; MCMC. Markov chain Monte Carlo sample.

物模型来讲,原则上计算得出的结果应该相似。这些软件内部统计原理各有优劣,例如限制性最大似然法的优点在于运行速度快,但是结果误差可能较大;而贝叶斯方法统计运行时间比较长,且对于其先验 prior 的设置需要比较专业的知识,但是得出的结果更加精确。本文将采用代表性的 MCMCglmm (贝叶斯) 为例(此软件免费使用,且相比于 VCE 等更容易设置模型参数),结合具体的例子,来介绍动物模型的使用。其他软件可在其网站下载和查询使用教程。另外,还推荐一个维基网站(名称为动物模型维基) <http://www.wildanimalmodels.org/tiki-index.php>, 这个网站由野外数量遗传学家 Alastair J. Wilson 等建立,详细介绍了表格中所有软件的使用教程,同时提供了具体的数据和详细的代码。

4.1 单变量动物模型

这里直接采用 Wilson 等(2010)的一个虚拟例子:一个狮鹫(西方世界传说中的一种神兽)种群,假设已经测量了该种群每个个体的出生体重,也了解该种群的遗传家系信息(这两个数据表格详见电子版附件 1)。根据其数据,首先,可以计算该性状(出生体重)的遗传力;第二,因为也知道每个个体的母亲,可以在该模型(混合模型)中加入另一个随机因变量,即母亲的 ID,来计算该性状的母系遗传;第三,整合该种群的性别信息,将其作为一个

固定的因变量,这样不仅可以得出该性状在两性间是否有差异,同时在控制性别对出生体重的影响下,还能更加精确地计算以上所述遗传参数。软件具体使用和代码以及计算结果详见电子版附录 1。

4.2 多变量动物模型

例如,进化学家希望了解动物生长速率和免疫功能之间的关系,这两者是否存在一个权衡(trade-off)?以野外一个田鼠种群作为研究对象,预测这两个性状都是可遗传的,而且受到自然选择的正驱动作用,但是种群在快速的生长率和更好的免疫功能这两者间可能有一个权衡,因此会有一个负遗传相关关系。

为了验证这个假说,通过调查该种群的遗传家系(可以利用亲本分析加以辅助),同时测量每个个体这两个性状的表型值,利用 MCMCglmm 来运行动物模型。软件具体使用和代码以及计算结果详见电子版附件 2。

4.3 模型参数设置,随机还是固定?

动物模型归根结底是一个特殊的混合模型,其特殊性在于同时整合了另外一个专门的家系数据,并自动将表型数据中的个体 ID 作为随机变量(这个随机变量是动物模型所必须的,动物模型就是基于这个随机变量同时整合该种群的家系结构来计算遗传力等参数)。除了这个特殊的随机变量外,既然是混合模型,还可以根据需要或者结合要验证的假说来相应地设置

或者增加其他的因变量。例如，上述单变量模型例子中增加了另一个随机变量母亲 ID，来计算母系遗传变异，同时增加了性别作为固定变量来检验该性状性别间的差异。将一个因变量视为随机变量还是固定变量，应该视具体情况而定，没有固定的原则，一般认为，如果数据包含该因子的所有水平，而且研究目的是比较该因子不同水平下平均值的差异，则视为固定因变量（例如性别、不同的实验施加水平），如果该因子只是从总体中取样而获得（例如，个体 ID），则视为随机因变量（Pinheiro et al. 2000, Galwey 2014）。但是，有些因子变量则很灵活，例如年份，如果要比较不同年份间某个性状的平均值差异，则可以视为固定变量，而要了解该性状年际间的波动情况时，则应该将其视为随机变量。

5 动物模型的应用及延伸

5.1 动物模型不仅仅是“动物模型”

动物模型不仅适用于动物生态学研究，它完全适合于植物种群生态学，对于一个植物种群，如果知道该种群遗传家系结构，同样可以测量该种群某些重要特征的遗传参数（Lamy et al. 2012, Hernández-Serrano et al. 2014, Zenni et al. 2014）。当然，植物的繁殖系统很复杂，要理清一个野生植物种群可能更复杂。

5.2 直接的假说验证

动物模型还可以帮助直接的假说验证。对于一个研究种群，要测量该种群某个表型性状的遗传力，其零假说则可认为：该表型性状遗传力 $h^2 = 0$ 。如果得出的遗传力显著大于 0，就说明该表型性状的加性遗传部分很显著，有较强的进化潜能（参见选择预测方程）。对于多个表型性状，则其零假说为：这些性状之间不存在遗传相关， $r_G = 0$ 。如果 r_G 显著大于或者小于 0，则认为这些性状之间存在显著的遗传相关，其生物学意义可以根据具体的表型性状来解释。一般认为，负遗传相关表示个体在其生活中不能都兼顾，存在一个权衡，而正遗传

相关则意味着，决定这些表型性状的基因具有很强的连锁性，或者这些表型性状本身就是由相同的基因决定的（基因的多重作用：pleiotropy）。

5.3 遗传力的时空动态（种群的进化动态）

由于一个表型变异的遗传部分 V_A 是自然选择作用的基质和模板，而自然选择或者进化又是一个随时发生的过程。因此，从时间维度上，可以测量一个重要表型性状（例如某生活史或者适合度特征）的遗传力在个体水平或者种群水平上随时间的变化度。从空间的维度，也可以比较同一个物种同一表型性状在不同种群间的遗传力差异。如果种群间某一个表型性状的遗传力相差较大，可以尝试从环境差异、自然选择差异以及环境和基因的相互作用等角度去解释，这样就可以从基因水平上更好地了解种群的进化动态（Charmantier et al. 2005, Blumstein et al. 2013）。

5.4 性选择方面的应用

我们知道，在性选择理论中有一个重要而尚未得到完全解决的问题，即为什么有些物种（尤其是鸟类）中雌性与雄性一样拥有绚丽多彩的羽色？目前主要有两种假说，第一种是“进化副产品假说”（by-product hypothesis），该假说认为，雌性鲜艳的羽色是由于雄性的羽色基因受到强大的正选择驱动，这些基因最开始由父亲传给女儿，即使在雌性中没有进化生物学意义，但是也会在雌性个体中表达。第二种假说认为，雌性鲜艳的羽色也受到正选择驱动，有其独立的进化生物学意义，因为双向选择在动物中也是比较普遍的，即不仅仅只有雌性选择雄性，雌性鲜艳的羽色是雌性个体质量和适合度的一个指标，雄性个体可能偏向于选择鲜艳羽色的雌性个体作为繁殖配偶。这两种假说利用动物模型能很好地加以验证。我们可以分别把雄性和雌性的羽色作为两个独立的表型性状，利用多变量动物模型，如果得出这两个性状存在很强的正遗传相关（例如 $r_G > 0.9$ ），则可认为这原则上属于同个性状，即支持上

述第一种假说。如果这两性状的遗传相关程度不高,则可以认为它们确实是由不同的基因决定的两个独立性状,甚至受到不同的选择驱动,支持第二种假说。

5.5 系统比较学研究 MCMCglmm: 进化谱系遗传力 phylogenetic heritability

在系统比较学研究中,可以巧妙地利用动物模型,来进行更大尺度(时间和空间)的进化学研究,一个种群的遗传家系数据,可以用某一个分类单元谱系树来代替。例如,可以利用已知的鸡形目系统发生树作为动物模型的“种群家系”,将每一个物种视为“种群中的个体”,来计算鸡形目中某个共有性状的变异在多大程度上是由谱系关系决定的(类似遗传力的概念: phylogenetic heritability ‘进化谱系遗传力’)。具体可以参考 Hadfield (2010) 研发的 R 软件程序包 MCMCglmm 的使用指南,以及网上公开的关于该程序包的课堂笔记。MCMCglmm 程序包不仅仅用来运行动物模型,而且也用来进行系统比较学研究,二者其实本质上一致,只是尺度不同。具体例子可参照 Cooney 等 (2016) 和 Benson-Amrama 等 (2016)。

6 注意事项

动物模型对于数据的需求永远是“饥饿”的,因此,使用动物模型最大的注意事项是数据量不足而导致参数估算的偏差。例如,想准确地测量某个性状的母系遗传(母系效应),则要求数据提供每一个母亲都有多个子代,而且最好是比较多的同母异父这样的半同胞关系存在 (Kruuk et al. 2007)。对于动物模型具体的样本量需求(能够提供较为可靠的结论),这里举两个具体成功的例子: 1) 歌带鹀种群 (Reid et al. 2011a, b), 该种群(一个小型岛屿)每年约有 11 ~ 49 繁殖对,研究者自 1975 年开始研究和监测该种群,该种群确切完整的家系信息是从 1993 年开始; 2) 青山雀种群 (Hadfield et al. 2006a, b), 该种群当时在运用动物模型的样本

量条件为: 3 年(2002 ~ 2004 年)共 450 繁殖对的数据。另外增加或者减少某一个因变量可能会导致遗传参数的偏差,例如,在测量某一个表型性状的遗传力时,通常会根据生物学背景,相应地增加随机因变量和固定变量,以剔除影响遗传力计算的偏差因子。如果没有考虑到这些因素,会使遗传力相应偏大 (Wilson et al. 2010)。

种群的配对方式也可能导致遗传参数的偏差,例如,选型互配 (positive assortative mating) 会使性状的遗传力偏大,原因是由于父母亲之间的这种选型互配会“人为”地增加具有亲缘关系个体间的相似度,从而导致遗传力的计算误差偏大。关于选型互配影响遗传参数的估算以及解决方法可详见 Falconer 等 (1996)。

目前国内野外生态学、进化生物学研究中还很少涉及野外数量遗传学的相关内容。动物模型实现的一个前提条件就是必须提供该种群的家系信息,而国内生态学研究起步较晚,很难拥有一个稳定研究种群的长期数据积累,特别是还要清楚该种群基本的遗传家系。随着中国生态学的发展和研究的深入,加上国家对基础研究的重视和投入的增加,我们建议国内生态学家重视对稳定研究种群的长期研究,并注意积累相关的遗传家系数据(见本文举例家系),为应用数量遗传学方法(动物模型)来研究种群的进化动态提供条件,这样可以突破传统野外生态学研究瓶颈,将其上升到基因层次(如遗传力、母系遗传效应、环境比重等)。

致谢 非常感谢德国马普鸟类学研究所 Wolfgang Forstmeier 教授对本文的建议和 Niels Dingemanse 教授组织的数量遗传学研讨会,本文附录中的例子二来自于该研讨会的材料; 特别感谢 Alastair J. Wilson 教授在软件使用方面的指导。非常感谢马普研究所裴一凡博士研究生对本文的建议和评论。

参 考 文 献

- Benson-Amrama S, Dantzer B, Stricker G, et al. 2016. Brain size predicts problem-solving ability in mammalian carnivores. *Proceedings of the National Academy of Sciences of the United States of America*, 113(9): 2532–2537.
- Blumstein D T, Nguyen K T, Martin J G A. 2013. Ontogenetic variation of heritability and maternal effects in yellow-bellied marmot alarm calls. *Proceedings of the Royal Society B: Biological Sciences*, 280(1758): 20130176.
- Brommer J E, Merila J, Sheldon B C, et al. 2005. Natural selection and genetic variation for reproductive reaction norms in a wild bird population. *Evolution*, 59(6): 1362–1371.
- Brooks R. 2000. Negative genetic correlation between male sexual attractiveness and survival. *Nature*, 406(6791): 67–70.
- Charmantier A, Garant D. 2005. Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences*, 272(1571): 1415–1425.
- Cooney C R, Seddon N, Tobias J A. 2016. Widespread correlations between climatic niche evolution and species diversification in birds. *Journal of Animal Ecology*, 85(4): 869–878.
- Darwin C. 1859. *On the origin of species*. London: Murray, 45.
- Ellegren H, Sheldon B C. 2008. Genetic basis of fitness differences in natural populations. *Nature*, 452(7184): 169–175.
- Falconer D S, Mackay T F C. 1996. *Introduction to Quantitative Genetics*. 4th ed. London: Prentice Hall, 122–377.
- Fisher R A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Edinburgh: Philosophical Transactions of the Royal Society of London*, 52, 399–433.
- Fisher R A. 1958. *The genetical theory of natural selection*. New York: Dover Publications.
- Forstmeier W, Burger C, Temnow K, et al. 2009. The genetic basis of zebra finch vocalizations. *Evolution*, 63(8): 2114–2130.
- Forstmeier W, Martin K, Bolund E, et al. 2011. Female extra-pair mating behavior can evolve via indirect selection on males. *Proceedings of the National Academy of Sciences of the United States of America*, 108(26): 10608–10613.
- Freedman L P, Cockburn I M, Simcoe T S. 2015. The economics of reproducibility in preclinical research. *Plos Biology*, 13(6): e1002165.
- Galwey N W. 2014. *Introduction to Mixed Modelling: Beyond Regression and Analysis of Variance*. Chichester: Wiley.
- Garant D, Sheldon B C, Gustafsson L. 2004. Climatic and temporal effects on the expression of secondary sexual characters: genetic and environmental components. *Evolution*, 58(3): 634–644.
- Hadfield J D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, 33(2): 1–22.
- Hadfield J D, Burgess M D, Lord A, et al. 2006a. Direct versus indirect sexual selection: genetic basis of colour, size and recruitment in a wild bird. *Proceedings of the Royal Society B: Biological Sciences*, 273(1592): 1347–1353.
- Hadfield J D, Heap E A, Bayer F, et al. 2013a. Disentangling genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine. *Evolution*, 67(9): 2701–2713.
- Hadfield J D, Heap E A, Bayer F, et al. 2013b. Intra-clutch differences in egg characteristics mitigate the consequences of age-related hierarchies in a wild passerine. *Evolution*, 67(9): 2688–2700.
- Hadfield J D, Nakagawa S. 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, 23(3): 494–508.
- Hadfield J D, Nutall A, Osorio D, et al. 2007. Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of Evolutionary Biology*, 20(2): 549–557.
- Hadfield J D, Owens I P F. 2006b. Strong environmental determination of a carotenoid-based plumage trait is not mediated by carotenoid availability. *Journal of Evolutionary Biology*, 19(4): 1104–1114.
- Haldane J B S. 1932. *The Causes of Evolution*. London: Longmans, 34–62.
- Hairston N G, Ellner S P, Geber M A, et al. 2005. Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*, 8(10): 1114–1127.
- Hayward A D, Wilson A J, Pilkington J G, et al. 2013. Reproductive senescence in female Soay sheep: variation across traits and contributions of individual ageing and selective disappearance. *Functional Ecology*, 27(1): 184–195.

- Hayward A D, Wilson A J, Pilkington J G, et al. 2009. Ageing in a variable habitat: environmental stress affects senescence in parasite resistance in St Kilda Soay sheep. *Proceedings of the Royal Society B: Biological Sciences*, 276(1672): 3477–3485.
- Henderson C R. 1953. Estimation of variance and covariance components. *Biometrics*, 9(2): 226–252.
- Henderson C R. 1984. *Applications of Linear Models in Animal Breeding*. Guelph: University of Guelph Press, 285–300.
- Hernandez-Serrano A, Verdu M, Santos-del-Blanco L, et al. 2014. Heritability and quantitative genetic divergence of serotiny, a fire-persistence plant trait. *Annals of Botany*, 114(3): 571–577.
- Husby A, Schielzeth H, Forstmeier W, et al. 2013. Sex chromosome linked genetic variance and the evolution of sexual dimorphism of quantitative traits. *Evolution*, 67(3): 609–619.
- Jennions M D, Moller A P. 2002. Publication bias in ecology and evolution: an empirical assessment using the ‘trim and fill’ method. *Biological Reviews*, 77(2): 211–222.
- Johnston S E, Gratten J, Berenos C, et al. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*, 502(7469): 93–95.
- Kasada M, Yamamichi M, Yoshida T. 2014. Form of an evolutionary tradeoff affects eco-evolutionary dynamics in a predator-prey system. *Proceedings of the National Academy of Sciences*, 111(45): 16035–16040.
- Kruuk L E B. 2004. Estimating genetic parameters in natural populations using the “animal model”. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(359): 873–890.
- Kruuk L E B, Clutton-Brock T H, Slate J, et al. 2000. Heritability of fitness in a wild mammal population. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2): 698–703.
- Kruuk L E B, Hadfield J D. 2007. How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology*, 20(5): 1890–1903.
- Kruuk L E B, Slate J, Wilson A J. 2008. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology Evolution & Systematics*, 39(1): 525–548.
- Lamy J B, Lagane F, Plomion C, et al. 2012. Micro-evolutionary patterns of juvenile wood density in a pine species. *Plant Ecology*, 213(11): 1781–1792.
- Lynch M, Walsh B. 1998. *Genetics and Analysis of Quantitative Traits*. Sunderland: Sinauer, 537–800.
- Merila J, Kruuk L E B, Sheldon B C. 2001a. Cryptic evolution in a wild bird population. *Nature*, 412(6842): 76–79.
- Merila J, Sheldon B C. 2000. Lifetime reproductive success and heritability in nature. *American Naturalist*, 155(3): 301–310.
- Merila J, Sheldon B C, Kruuk L E B. 2001b. Explaining stasis: micro-evolutionary studies in natural populations. *Genetica*, 112/113: 199–222.
- Pinheiro J C, Bates D M. 2000. *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag, 3–52.
- Reid J M, Arcese P, Keller L F, et al. 2014. Female and male genetic effects on offspring paternity: additive genetic (co)variances in female extra-pair reproduction and male paternity success in song sparrows (*Melospiza Melodia*). *Evolution*, 68(8): 2357–2370.
- Reid J M, Arcese P, Sardell R J, et al. 2011a. Heritability of female extra-pair paternity rate in song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B: Biological Sciences*, 278(1708): 1114–1120.
- Reid J M, Arcese P, Sardell R J, et al. 2011b. Additive genetic variance, heritability, and inbreeding depression in male extra-pair reproductive success. *American Naturalist*, 177(2): 177–187.
- Reid J M, Sardell R J. 2012. Indirect selection on female extra-pair reproduction? Comparing the additive genetic value of maternal half-sib extra-pair and within-pair offspring. *Proceedings of the Royal Society B: Biological Sciences*, 279(1734): 1700–1708.
- Robinson M R, Wilson A J, Pilkington J G, et al. 2009. The impact of environmental heterogeneity on genetic architecture in a wild population of Soay sheep. *Genetics*, 181(4): 1639–1648.
- Saccheri I, Hanski I. 2006. Natural selection and population dynamics. *Trends in Ecology & Evolution*, 21(6): 341–347.
- Schielzeth H, Bolund E, Forstmeier W. 2010. Heritability of and early environment effects on variation in mating preferences. *Evolution*, 64(4): 998–1006.

- Shaw R G. 1987. Maximum-likelihood approaches applied to quantitative genetics of natural populations. *Evolution*, 41(4): 812–826.
- Stopher K V, Walling C A, Morris A, et al. 2012. Shared spatial effects on quantitative genetic parameters: accounting for spatial autocorrelation and home range overlap reduces estimates of heritability in wild red deer. *Evolution*, 66(8): 2411–2426.
- Thompson R. 2008. Estimation of quantitative genetic parameters. *Proceedings of the Royal Society B: Biological Sciences*, 275(1635): 679–686.
- Willis J H, Coyne J A, Kirkpatrick M. 1991. Can one predict the evolution of quantitative characters without genetics. *Evolution*, 45(2): 441–444.
- Wilson A J, Pemberton J M, Pilkington J G, et al. 2009. Trading offspring size for number in a variable environment: selection on reproductive investment in female Soay sheep. *Journal of Animal Ecology*, 78(2): 354–364.
- Wilson A J, Poissant J. 2016. Quantitative Genetics in Natural Populations. In 'Encyclopedia of Evolutionary Biology'. Elsevier.
- Wilson A J, Reale D, Clements M N, et al. 2010. An ecologist's guide to the animal model. *Journal of Animal Ecology*, 79(1): 13–26.
- Wright S. 1921. Systems of mating. *Genetics*, 6(2): 111–178.
- Zenni R D, Lamy J B, Lamarque L J, et al. 2014. Adaptive evolution and phenotypic plasticity during naturalization and spread of invasive species: implications for tree invasion biology. *Biological Invasions*, 16(3): 635–644.