

大菱鲆选育家系体重性状的分形特征

王新安¹ 马爱军^{1*} 雷霁霖¹ 杨志²

(¹ 农业部海洋渔业可持续发展重点实验室 青岛市海水鱼类种子工程与生物技术重点实验室
中国水产科学研究院黄海水产研究所, 266071)
(² 烟台开发区天源水产有限公司, 264003)

摘要 分析了2007~2010年度18月龄体重性状的信息维数和关联维数。分析结果显示, 2007~2010年度所构建的大菱鲆选育家系群体18月龄体重性状的信息维数介于0.942 2~0.958 6之间, 不同年度之间的信息维数变异不显著, 这说明不同年度各选育家系群体18月龄体重性状的信息维数均为较大值, 所统计群体的遗传变异都十分丰富, 统计性状的变异范围也较大, 显示出良好的育种潜力。相应数据的关联维数分析结果表明, 除2008年度小尺度范围的关联维数是0.428 4外, 其余各小尺度范围的关联维数均高于0.73, 这表明, 从分析结果的总体来看, 大菱鲆各选育家系群体内个体间的遗传结构的相关性较高; 信息维数和关联维数作为研究分形特征的指标, 从两个角度阐明了选育性状遗传结构的分形特征, 在实际的良种选育过程中, 应考虑这两方面的因素, 确定适当的留种率。

关键词 非线性 分形理论 体重 大菱鲆

中图分类号 Q346 **文献识别码** A **文章编号** 1000-7075(2013)06-0037-07

Fractal properties of body weight of turbot selective breeding families

WANG Xin-an¹ MA Ai-jun^{1*} LEI Ji-lin¹ YANG Zhi²

(¹ Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture,
Qingdao Key Laboratory for Marine Fish Breeding and Biotechnology, Yellow Sea Fisheries Research Institute,
Chinese Academy of Fishery Sciences, 266071)
(² Yantai Tianyuan Aquatic Ltd Corp, 264003)

ABSTRACT Fractal theory of nonlinear theory is an ideal method for studying the genetic structure and trait genetic variation of animal population. For animal breeding, it is very important to clarify genetic structure of the selected quantitative traits. In this study, the body weight of 18-month-old turbot in four selective breeding families during 2007-2010 was analyzed using fractal theory. The information dimension and correlation dimension were obtained. The results revealed that the information dimension of 18-month-old turbot body weight ranged in 0.942 2-0.958 6 from 2007 to 2010, which showed that all the four populations had high dimension, high variation, and high breeding potential. The correlation dimensions were higher than 0.73

现代农业产业技术体系建设专项资金(CARS-50-G01)资助

* 通讯作者。E-mail: maaj@ysfri.ac.cn

收稿日期: 2012-07-25; 接受日期: 2012-11-15

作者简介: 王新安(1970-), 男, 副研究员, 主要从事鱼类繁育与育种研究。E-mail: wangxa@ysfri.ac.cn

except for that of 2008, which showed that the correlations of the traits in the four populations were all relatively high; the information and correlation dimensions have illuminated the fractal properties of the traits in the four breeding populations from different points of view, which can provide a theoretic basis for determining appropriate selection proportion during selective breeding of turbot.

KEY WORDS Nonlinear Fractal theory Body weight Turbot

分形理论是现代非线性科学的研究的前沿和重要分支,作为一种全新的科学概念和研究方法,目前正在诸多研究领域开展应用探索。“分形热”的研究始于国外20世纪80年代。1975年,Mandelbrot(1982)创造性地提出了分形概念,5年后,出版了《自然界的分形几何》一书,这标志着分形理论的正式创立,并逐渐在生物学、天文学、地理学、化学、计算机科学、社会学、经济学等众多自然科学和社会科学中得到非常广泛的应用(刘莹等2006)。

分形维数(也称分维数)是分形理论的核心概念,是定量刻画分形特征的最主要的参数。物体的复杂程度及其动态过程随分形维数的变化而发生改变,它可以很好地用于刻画不规则事物的不同组成部分之间的相似性(Falconer *et al.* 1991)。显然,在动物的遗传改良过程中,可利用分形维数去研究选育目标的遗传结构和性状变异。

基于电子标记(PIT)辅助的大规模家系选育技术对大菱鲆进行遗传改良始于国家“十一五”期间(马爱军等2010)。从2007年开始至今,根据养殖厂的实验条件,每年持续构建了部分选育家系。了解所构建家系群体的遗传结构和育种性状的遗传变异是大菱鲆育种工作的基础。通过对大菱鲆遗传结构和性状遗传变异的研究,可以分析不同家系群体的遗传变异程度,评估家系群体的育种潜力,并为确定不同家系群体的留种率提供一种新理论依据,从而为大菱鲆良种选育提供准确的指导。目前,在动物遗传育种领域,有关群体遗传结构和性状遗传变异的研究主要利用分子生物学(全迎春等2006)、生化遗传学(徐成等2001)等方法。基于动物的遗传结构和性状变异现象属于典型的非线性系统,因此,在动物遗传育种过程中,可运用非线性理论研究选育目标的遗传结构和性状变异。非线性理论中的分形理论的创立和发展,为动物育种改良性状的遗传结构和性状变异的研究提供了新的方法和手段(吴辉生等2008)。在本研究中,作者运用非线性理论中的分形理论方法,以大菱鲆18月龄体重性状为研究对象,通过对其信息维数和关联维数的计算,分析了大菱鲆体重性状的变异特征,为大菱鲆良种选育育种规划的制定和选育利用提供新的理论依据。

1 材料与方法

1.1 材料来源

本研究采用的数据来源于山东烟台天源水产有限公司在2007~2010年连续4年所构建的大菱鲆良种场的全同胞家系。每年度家系达到18月龄时进行数据采集。考虑到所有家系发育到18月龄时的生长发育状况,对于成活率极低或生长状况极差的家系本研究没有采用。因此,在2007、2008、2009和2010年所采用的家系分别为30、25、30、20个。每个家系随机抽取40~50尾鱼用电子天平测量体重(g),并计算每年度体重均重和变异系数。各年度大菱鲆选育家系在早期培育阶段的数量标准化和环境标准化方法及其成鱼阶段的养殖条件均采用王新安等(2011)的方法。

1.2 数据分析

分析各年度选育家系群体18月龄体重的变异系数。各年度测度范围以每一年度中所采集数据的最大值和最小值为其上、下限。采用人工判定法来确定无标度区间,利用信息维数和关联维数这两种指标对目标性状进行分形特征的研究。对于信息维数,依据Dong(1991)描述的信息维数的定义,其具体的方法如下所述:将各年度所采集的分析数据按照从小到大的顺序分别排序,并对每一年度的排序数列进行逐级等分(2→20),从而获

得不同的有序分布的数据点集合。统计每一等分组中的个体数量。信息维数的计算公式为:

$$D = -\lim_{\epsilon \rightarrow 0} \frac{\sum [(-N_i/N) \times \ln N_i/N]}{\ln \epsilon}$$

$$\text{其中, } I(\epsilon) = \sum [(-N_i/N) \times \ln N_i/N]$$

式中, D 为信息维数, ϵ 为等分距离, N_i 为各等分组中的测量值数目, N 为所分析年度的总测量值数目, $I(\epsilon)$ 为所求得的总信息量。将计算获得的 $I(\epsilon)$ 与其相对应的 ϵ 在平面坐标系中进行线性拟合, 若数据点集呈现显著不同的线性区间, 则应对其分别进行线性拟合。通过目测判定呈现显著线性特征的数据点区间, 对这一区间的数据点进行线性拟合, 所得直线斜率的绝对值就是所要求得的信息维数。

计算不同年度选育家系群内所有个体之间的欧氏距离值, 设定欧氏距离值, 找出小于给定欧氏距离值的数据个数; 然后, 变换不同的欧氏距离值, 进而得到一系列的欧氏距离。关联维数的计算公式为:

$$D_c = \lim_{\epsilon \rightarrow 0} \frac{\lg(N_i/N)}{\lg \epsilon} \quad \text{其中, } C_{(\epsilon)} = N_i/N$$

式中, D_c 为关联维数, ϵ 为距离值, N_i 为小于设定欧氏距离的距离值的数目, N 为所分析年度家系群中的总欧氏距离的数目。将计算获得的低于特定欧氏距离值(ϵ)的数目的 $C_{(\epsilon)}$ 与其相对应值在平面坐标系中进行线性拟合(分区间), 所得直线斜率的绝对值即是所求的关联维数值。在本研究中, 基于实际情况, 确定 0g 到所分析年度的最大欧氏距离值为距离值, 步长变换值为 8g。本研究采集的数据经初步处理后, 利用 Excel 软件进行统计分析。

2 结果与分析

2.1 体重变异程度的分析

18月龄体重变异程度的分析结果如表1所示。从表1可以看出, 不同年度间选育家系的18月龄体重的差异很明显, 经检验, 年度效应对18月龄体重的影响达到显著水平($P < 0.05$)。在4个年度中, 2009年度的家系群体均重最大, 为 (1005.64 ± 195.88) g, 2008年度的均重最小, 仅为 (764.24 ± 126.94) g。体重的变异系数及测度范围可看作家系群体体重遗传变异程度的参考指标。表1显示, 不同年度家系体重的变异系数是不同的, 这可认为不同年度家系群体间存在着不同的变异。4个年度家系群体体重变异系数的大小排序为 $2007 > 2009 > 2010 > 2008$ 。各年度体重均有较大的测度范围, 达到280g以上。

表1 不同年度大菱鲆选育家系的体重比较

Table 1 Mean body weight of the four turbot selective breeding families in different year

年度 Year	体重均值 Body weight (g)	变异系数 C. V. (%)	测度范围 Scale size (g)
2007	837.99 ± 185.28^a	22.11	539 ~ 1 343
2008	764.24 ± 126.94^b	16.60	363 ~ 957
2009	1005.64 ± 195.88^c	19.47	603 ~ 1 441
2010	910.76 ± 153.54^d	16.85	125 ~ 405

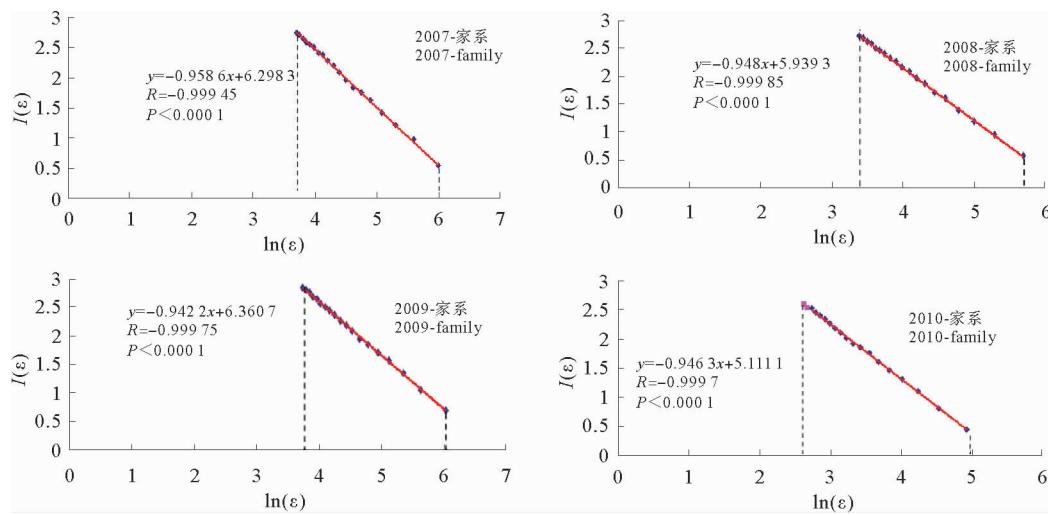
注:实验数值上标表示多重比较结果, 同列字母相同表示差异不显著($P > 0.05$); 同列字母不同表示差异显著($P < 0.05$)

Note: Data within the same column with same superscripts are not significantly different ($P > 0.05$); different superscripts mean significant difference ($P < 0.05$)

2.2 大菱鲆体重性状的信息维数

2.2.1 无标度区

基于大菱鲆18月龄体重测量值的分析, 获得 $I(\epsilon)$ 与 $\ln(\epsilon)$ 的关系(图1)。从图1可以看出, 2007 ~ 2010年度的 $I(\epsilon)$ 与 $\ln(\epsilon)$ 在各自的观测尺度区间均显示出良好的线性特征, 因此, 4个观测尺度区间均可视为无标度区, 在其区间拟合直线得到的斜率的绝对值即为体重的信息维数。

图 1 2007~2010 年 18 月龄大菱鲆体重 $I(\epsilon)$ - $\ln(\epsilon)$ 分布函数Fig. 1 $I(\epsilon)$ - $\ln(\epsilon)$ distribution function of body weight of turbot families from 2007 to 2010

2.2.2 信息维数

根据所拟合直线的斜率估计得到的 2007~2010 年度 18 月龄体重的信息维数见表 2。从表 2 可以看出, 各年度线性方程拟合时, 其两变量数据间的相关性都很高, 相关系数的绝对值都高于 0.999, 经检验, 线性相关达到极其显著水平 ($P < 0.01$)。这表明, 据此方法得到的信息维数是准确的。从表 2 还可以看出, 各年度大菱鲆 18 月龄体重的信息维数的范围是 0.942 2~0.958 6, 变异不显著, 不同年度信息维数值的大小排序依次为 2007 > 2008 > 2010 > 2009, 最大值与最小值之间的差异仅为 0.016 4。进一步观察图 1, 可见各年度的信息维数均有较大的所适尺度范围。

表 2 18 月龄大菱鲆体重的信息维数

Table 2 The information dimension of body weight of 18-month-old turbot

年度 Year	信息维数 Information dimension	相关系数 Correlation coefficient	无标度区 No-scale interval	显著性水平 Significant level
2007	0.958 6	-0.999 45	3.69~5.99*	$P < 0.01$
2008	0.948 0	-0.999 85	3.39~5.69*	$P < 0.01$
2009	0.942 2	-0.999 7	3.73~6.03*	$P < 0.01$
2010	0.946 3	-0.999 7	2.74~4.94*	$P < 0.01$

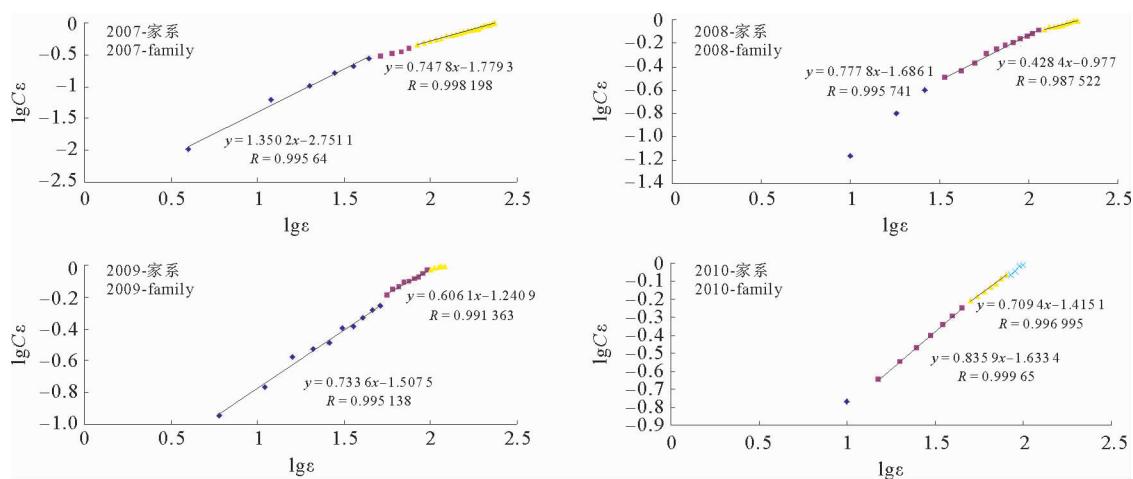
注: * 代表所分析年度信息维数的无标度区间是整个测度尺度范围

Note: * denotes that the no-scale interval of information dimension is distributed in the whole range of the scale size

2.3 大菱鲆体重的关联维数

2.3.1 无标度区

图 2 显示, 2007~2010 年度的 $\lg e$ 和 $\lg C_{(\epsilon)}$ 之间均呈现两个直线特征区间。这里, 拐点的存在对于关联维数的分析具有非常显著的意义。由于在不同尺度范围所拟合线性方程的斜率是不同的, 因而不同尺度区域内的个体的关联度是不同的。在本研究中, 针对各年度的两个线性特征区域均分析其关联维数。

图2 2007~2010年18月龄大菱鲆体重性状的 $\lg \varepsilon-\lg C_\varepsilon$ 分布函数Fig. 2 $\lg \varepsilon-\lg C_\varepsilon$ distribution function of body weight of turbot families from 2007 to 2010

2.3.2 关联维数

根据所拟合线性方程的斜率估计得到的2007~2010年度18月龄体重的关联维数见表3。从表3可以看出,各年度不同的尺度区域均有不同的关联维数,且各年度线性方程拟合时,两变量间的相关性都很高,相关系数都大于0.987,线性相关也都达到极其显著水平($P < 0.01$)。显然,据此方法得到的关联维数均是准确的。从表3还可以看出,从2007~2010年,各年度18月龄体重的关联维数值均是在较大尺度区域高于在较小尺度区域,其中,2007年度大尺度范围的关联维数最高,达1.3502。尽管如此,除2008年度小尺度范围的关联维数是0.4284外,其余各小尺度范围的关联维数均高于0.73,这是一个较高的关联维数值。本研究显示不同年度家系个体的遗传结构的相关性是较高的。

表3 18月龄大菱鲆体重的关联维数

Table 3 The correlation dimension of body weight of 18-month-old turbot

年度 Year	关联维数 Correlation dimension	相关系数 Correlation coefficient	显著性水平 Significant level	测度范围 Scale size(g)
2007	1.3502	0.99564	$P < 0.01$	0.61~1.64
	0.7478	0.998198	$P < 0.01$	2.03~2.37
2008	0.7778	0.995741	$P < 0.01$	1.53~2.06
	0.4284	0.987522	$P < 0.01$	2.09~2.27
2009	0.7336	0.995138	$P < 0.01$	0.78~1.71
	0.6061	0.991363	$P < 0.01$	1.75~1.98
2010	0.8359	0.99965	$P < 0.01$	1.18~1.65
	0.7094	0.996995	$P < 0.01$	1.70~1.90

3 讨论

研究表明,生物体具有分形的特征(夏晓智 1996)。Havlin等(1995)、West等(1999)、Williams(1997)和Weibel(1991)的研究结果显示出生命中分形特征的普遍性,并进一步阐明了分形在生命科学领域中的重要性。分形维数作为分形理论的核心内容,是研究物种空间分布特征的有力工具(Kunin 1998; Cannas *et al.* 2006),在生态学研究中应用较为普遍(Halley *et al.* 2004; Ma *et al.* 2003)。在群体遗传学中,分形的作用同样也非常显著(Paetkau *et al.* 1999; Wingen *et al.* 2007),Vlad等(2007)系统阐明了生命的遗传结构

中具有广泛的分形特征,这说明运用非线性理论中的分形理论对性状的遗传结构和遗传变异进行研究是完全可行的(吴辉生等 2008)。

信息维数(Information dimension)反应几何形态不均匀程度量的特征,是描述事物分形特征的一种分维数。信息,即负熵,是系统结构有序化、组织化、复杂化状态的度量。一个系统结构的复杂性是指其随着测度的变化而发生变化的程度(祖元刚等 2004)。系统结构的信息量与其结构的复杂程度呈现正相关,即信息量大的系统,结构复杂;而信息量少的系统,结构简单(马克明等 2000)。信息维数值能够揭示随观测尺度的减小系统复杂性的变化程度(角媛梅等 2007)。在用信息维数这一指标来研究育种目标的遗传变异时,信息维数越大,表明系统结构复杂性的尺度变化越强烈、结构越复杂(马克明等 2000),个体分布越不均匀,群体内个体间的差异越显著,选育的潜力也就越大。因此,在良种选育过程中,可在高信息维数的育种群体中选留较多的变异个体。另外,考虑到变异系数以及变异范围在某种程度上也揭示育种目标的变异程度,因此,在本研究中采用信息维数、变异系数和变异范围来综合研究选育目标的遗传变异,可以提高判断变异程度的准确度。本研究对 2007~2010 年度 4 年间大菱鲆 18 月龄体重进行信息维数的研究表明,4 个年度的家系群体的体重信息维数均集中在 0.942 2~0.958 6 之间,均接近于 1,这表明大菱鲆 18 月龄体重性状在观测尺度变化时,测量值的分布变化较为强烈。随着观测尺度的逐渐减小,信息量随之较为迅速地增加。综合信息维数及表 1 中较大的变异系数和测度范围表明,每年度家系群体 18 月龄体重均存在丰富的变异,从中可选留优良变异个体的机率高,显示出良好的育种潜力。

不同于信息维数,关联维数是另一种描述研究对象分形特征的分维数,是一个生物群体内部个体之间相关性的反应,揭示群体内个体相关的尺度变化规律,表明了群体内个体的相关程度(吴辉生等 2008)。当对某一群体内的关联维数进行分析时,在一定的测度范围内,如仅有一个线性无标度区,这说明群体内个体的分布均服从相同的规律,显示出分布均匀的特征(宋萍等 2005)。在本研究中,大菱鲆 2007~2010 年度连续 4 年的家系群体 18 月龄体重均含有两个有统计学意义的无标度区域及其关联维数,这显示出各家系群体内的个体分布呈现非均匀的特征。这一研究结论基本上是与凉山半细毛羊初生重关联维数的统计分析结果相一致的(吴辉生等 2008)。尽管这两个关联维数都具有统计学上的意义,但较高的关联维数更能很好地揭示个体之间的相关性。因此,在分析某一特定群体的关联维数时,如出现不同的、有意义的无标度区,得到不同的关联维数,通常用较高的关联维数来表示群体内个体分布的关联维数(吴辉生等 2008)。本研究结果显示,大菱鲆 2007~2010 年度 18 月龄体重均是大尺度关联维数值高于小尺度关联维数,但即使较小的关联维数,除 2008 年度外,其余关联维数均高于 0.73。2008 年度存在一个较低关联维数的原因,可能是由于当年度家系群体转移导致养殖环境变化稍大(其他年度家系均不存在转移养殖),同时还发生过一次疾病引起的。一个群体的关联维数值越大,其内部个体的关联度也就越高,遗传结构的相关程度就越高(吴辉生等 2008)。对于大菱鲆家系群体体重的关联维数(表 3),从分析结果的总体上看是很高的,这表明大菱鲆家系群体内个体间体重的遗传结构存在较高的相关性。

综合比较大菱鲆 2007~2010 年度体重的关联维数和信息维数发现,2007 年度家系体重的信息维数和关联维数均为最高,这表明,2007 年度家系群体内体重性状存在着最为丰富的遗传变异(2007~2010 年度),体重性状的变异潜力较大,对这个家系群体进行个体选留时,可供选择的个体应最多,但此家系群内个体的遗传结构的相关性也最高。2008、2009 这两个年度,相比较而言,体重的信息维数和关联维数均较低一些,这表明,这两个年度内,家系群体内体重的遗传变异相对较少一些,可供选择的个体应相对较少,但这两个家系群内个体体重的遗传结构的相关性也相对较弱。因此,基于分形理论,依据分形参数对于不同年度家系进行选留个体时,应充分从信息维数和关联维数两个角度进行综合考虑。对于连续 4 年所构建的大菱鲆育种家系,从体重的信息维数和关联维数的总体上看,群体的体重变异性较大,以体重为选育目标进行遗传改良时,显示出良好的育种潜力,但个体遗传结构的相关性也较强,因此,在实际选育过程中,应考虑这两方面的因素,确定适当的留种率。

参 考 文 献

- 王新安,马爱军,黄智慧,刘庆明,郭黎,杨志,曲江波. 2011. 大菱鲆(*Scophthalmus maximus*)大规模家系选育优良家系的生长发育规律. 海洋与湖沼, 42(2): 266-273
- 马克明,祖元刚. 2000. 兴安落叶松种群格局的分布特征 - 信息维数. 生态学报, 20(2): 187-192
- 马爱军,王新安,薛宝贵,黄智慧,杨志,曲江波. 2010. 大菱鲆(*Scophthalmus maximus*)选育家系的构建和培育技术研究. 海洋与湖沼, 41(3): 301-306
- 全迎春,李大宇,曹鼎辰,孙效文,梁利群. 2006. 微卫星DNA标记探讨镜鲤的种群结构与遗传变异. 遗传, 28(12): 1541-1548
- 刘莹,胡敏,余桂英,李小兵,刘晓林. 2006. 分形理论及应用. 江西科学, 24(2): 201-209
- 吴辉生,吴登俊,周明亮,张翔宇. 2008. 凉山半细毛羊初生重性状分形特征分析. 遗传, 30(10): 1319-1325
- 角媛梅,杨丽萍. 2007. 哀牢山区哈尼梯田的分布特征. 生态学报, 27(11): 4583-4589
- 宋萍,洪伟,吴承祯,封磊,蓝斌. 2005. 天然黄山松种群空间格局的分形特征 - 关联维数. 武汉植物学研究, 23(1): 32-36
- 祖元刚,赵则海,于景华,杨逢建. 2004. 非线性生态模型. 北京:科学出版社, 222-226
- 徐成,王可玲,张培军. 2001. 鲈鱼群体生化遗传学研究 II. 种群生化遗传结构及变异. 海洋与湖沼, 32(3): 248-254
- 夏晓智. 1996. 分形在生命中的作用. 大自然探索, 15(4): 57-61
- Cannas SA, Marco DE, Montemurro MA. 2006. Long range dispersal and spatial pattern formation in biological invasions. Math Biosci 203(2): 155-170
- Dong LK. 1991. Fractal geometry and its application. Shenyang: Liaoning Science and Technology Press, 25-30
- Falconer KJ (Zeng WQ, et al. translation). 1991. Fractal geometry: Mathematical foundation and applications. Shenyang: Northeast China University of Technology Press
- Halley JM, Hartley S, Kallimanis AS and 3 others. 2004. Uses and abuses of fractal methodology in ecology. Ecol Lett 7(3): 254-271
- Havlin S, Buldyrev SV, Goldberger AL and 5 others. 1995. Fractals in biology and medicine. Chaos Solitons Fractals 6(1): 171-201
- Kunin WE. 1998. Extrapolating species abundance across spatial scales. Science 281(5382): 1513-1515
- Ma KM, Zu YG. 2003. Fractal properties of the spatial pattern of *Larix gmelini* population information dimension. Acta Ecol Sin 20(2): 187-192
- Mandelbrot BB. 1982. The fractal geometry of nature. New York: WH Freeman
- Paetkau D, Amstrup SC, Born EW and 8 others. 1999. Genetic structure of the world's polar bear populations. Mol Ecol 8(10): 1571-1584
- Vlad MO, Morán F, Popa VT and 2 others. 2007. Functional, fractal nonlinear response with application to rate processes with memory, allometry, and population genetics. Proc Natl Acad Sci USA 104(12): 4798-4803
- Weibel ER. 1991. Fractal geometry: A design principle for living organisms. Am J Physiol Lung Cell Mol Physiol 261(6): 361-369
- West GB, Brown JH, Enquist BJ. 1999. The fourth dimension of life: Fractal geometry and allometric scaling of organisms. Science 284(5420): 1677-1679
- Williams N. 1997. Biology: Fractal geometry gets the measure of life's scales. Science 276(5309): 34-40
- Wingen LU, Brown JKM, Shaw MW. 2007. The population genetic structure of clonal organisms generated by exponentially bounded and fat-tailed dispersal. Genet 177(1): 435-448