

太行山猕猴跖骨性差的逻辑回归分析

胡海洋^a,王凤产^b,胡凤霞^c,田华香^c,毛晓静^c,赵哲^a,赵晓进^c

(河南师范大学 a.生命科学学院;b.学报编辑部;c.水产学院,河南新乡 453007)

摘要:用猕猴跖骨变量建立性别判别逻辑回归模型,用 ROC 分析来比较跖骨不同变量、不同侧别和每根跖骨的性差大小.成年跖骨标本 45 例(31 雌,14 雄).选择跖骨的 7 个线性变量.数据分析采用 SPSS20.0 软件.结果表明:跖骨长度变量是最好的性别标识变量,然后是跖骨头和跖骨底变量,最小是跖骨骨干变量.总体上性别正确判别率较高(84.4%~95.6%).ROC 分析显示左右两侧之间的差异很小(AUC 分别是 0.922 和 0.956).5 根跖骨逐步逻辑回归结果的是 MT3,MT4,MT5 判别率较高(88.9%~95.6%).MT1 和 MT2 判别率较低(84.4%~86.7%).结果提示:猕猴跖骨性差显著,跖骨的长度变化主要受基因控制,骨干宽度主要受到环境因素的影响.跖骨的性差模式主要反映了遗传、环境与行为因素的交互作用.

关键词:猕猴;跖骨;性差;逻辑回归;ROC 分析

中图分类号:Q954

文献标志码:A

在进行骨骼的性别鉴定时,选择骨骼测量方法与选择 DNA 和染色体性别检测方法相比,很明显前者是省时省力,经济实惠^[1-3].利用骨骼变量来判别个体性别的技术日趋成熟,已经广泛应用到人类骨骼系统的各个部位,例如锁骨^[4]、跟骨和距骨^[5]、掌骨和跖骨^[3].

一些资料报道了用猕猴骨骼可以进行性别鉴定,包括颅骨^[6]、髌骨^[7]、肱骨^[8]、掌骨和跖骨^[9-10].目前尚未见到用跖骨变量进行逻辑回归和 ROC 分析.本文研究目的是用猕猴跖骨变量建立用于判别性别二元逻辑回归模型;用 ROC 分析来比较跖骨不同变量、不同侧别和每根跖骨的性差大小.

1 材料和方法

实验标本来自河南省北部太行山区.成年跖骨标本 45 例(31 雌,14 雄).标本的性别是已知的,年龄根据牙齿萌发情况确认.选择健康无病、骨骼发育正常的跖骨,用数显游标卡尺测量,精确到 0.01 mm.变量的选择依据 Khanpetch 等^[1].

跖骨的缩写字母是 MT,根据他们的位置排序,字母标识依次为 MT1,MT2,MT3,MT4 和 MT5.选择跖骨的 7 个变量,字母标识分别是跖骨长(LG),跖骨头宽(HW)、跖骨头高(HH),跖骨中间宽(MW)、跖骨中间高(MH)、跖骨底宽(BW)和跖骨底高(BH).这里的跖骨宽表示跖骨的内侧与外侧之间的线性距离,跖骨高表示跖骨背面与腹面之间的线性距离^[1].

所有测量数据分析采用 SPSS 20.0.

性别判别分析采用二元逻辑回归分析.因变量用 P (概率)表示,需要满足条件: $0 \leq P \leq 1$.为此对 P 作对数转换,即 $\text{logit } P = \ln(P/1-P)$,可以得到逻辑回归方程:

$$P = \frac{\exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_m X_m)}{1 + \exp(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_m X_m)}$$

这里 P 是因变量,为通过逻辑回归分析判定为某一性别的概率, X_1, X_2, \dots, X_m 为自变量,分别代表跖骨

收稿日期:2014-08-20

基金项目:河南省基础与前沿技术 Research 计划项目(132300410136)

作者简介:胡海洋(1990-),男,河南新乡人,河南师范大学在读硕士研究生,研究方向:动物学.

通信作者:赵晓进(1957-),男,山东聊城人,博士,主要从事灵长类生物学研究,E-mail:zqxj_61119@163.com.

线性测量的 7 个变量。 b_0 是常数, b_1, b_2, \dots, b_m 为各个自变量(X) 的系数。

逻辑回归模型描述性统计参数采用偏回归系数, 标准误, Wald, P 值和 $\text{Exp}(B)$ 。Wald 统计量类似于 χ^2 分布检验。Wald 检验的 P 值为每个变量在回归模型中的统计学检验($\alpha=0.05$)^[11]。逻辑回归模型中的因变量 P 的判别临界值是 0.5, 将自变量带入模型, 若 $P < 0.5$, 判别为雄性; 若 $P > 0.5$, 判别为雌性。采用 ROC 曲线分析是为了比较不同变量、不同侧别和不同跖骨的判别能力, ROC 结合二元逻辑回归一起使用。

2 结果

两侧跖骨所选择出来的变量基本一致。左右侧 10 根跖骨共得到 10 个逻辑回归模型, 每根跖骨的 7 个变量选择出的最好性别识别变量是 LG。可以用来直接进行性别判别, 正确判别性别的判别率的范围是 77.8%~88.9%。两侧跖骨逐步逻辑回归结果见表 1。

结果显示两侧跖骨的 MT1 只选择出来 LG 一个变量, 其他 8 根跖骨除了选择 LG 外, 还分别选择出来了其他一些变量, 这些变量是 HH(3/8)、HW(2/8)、BH(2/8) 和 BW(1/8)。注意这里所选择出来的变量主要是跖骨头和跖骨底变量, 没有跖骨粗壮度变量。除了 RMT2 外, 其他的回归模型都显示出比单变量 LG 具有更高的判别率。提示通过逐步回归选择出来的一些变量组合可以提高性别的判别率, 正确判别率的范围是 84.4%~95.6%。

表 1 两侧跖骨逐步逻辑回归分析

Bone	Step	Left					Right				
		Var.	Wald	Sig.	Exp(B)	A/%	Var.	Wald	Sig.	Exp(B)	A/%
MT1	Step 1	LG	13.706	0.000	0.301	86.7	LG	13.506	0.000	0.278	84.4
		Constant	14.021	0.000	1.437E+18		Constant	13.793	0.000	2.318E+19	
MT2	Step 1	LG	11.295	0.001	0.494	77.8	LG	12.199	0.000	0.458	86.7
		Constant	11.597	0.001	6.661E+14		Constant	12.463	0.000	2.500E+16	
	Step 2	LG	8.825	0.003	0.468	84.8	LG	9.418	0.002	0.469	84.4
		HW	4.817	0.028	0.059		HW	4.676	0.031	0.055	
		Constant	10.735	0.001	3.716E+23		Constant	11.817	0.001	7.286E+23	
	MT3	Step 1	LG	11.957	0.001	0.445	82.2	LG	10.927	0.001	0.381
Constant			12.175	0.000	6.244E+17		Constant	11.044	0.001	1.668E+21	
Step 2		LG	8.529	0.003	0.427	88.9	LG	7.961	0.005	0.362	95.6
		BW	5.221	0.022	0.043		HH	5.576	0.018	0.024	
Constant	9.646	0.002	1.733E+31		Constant	10.155	0.001	5.589E+35			
MT4	Step1	LG	11.848	0.001	0.508	82.2	LG	12.033	0.001	0.435	84.4
		Constant	12.134	0.000	8.174E+14		Constant	12.278	0.000	1.720E+18	
	Step2	LG	7.598	0.006	0.546		LG	8.230	0.004	0.448	91.1
		HH	6.087	0.014	0.058	88.9	HH	5.154	0.023	0.079	
Constant	10.780	0.001	3.907E+23		Constant	9.834	0.002	4.523E+26			
MT5	Step1	LG	9.159	0.002	0.407	88.9	LG	10.101	0.001	0.422	84.4
		Constant	9.369	0.002	7.881E+18		Constant	10.352	0.001	1.372E+18	
	Step2	LG	4.957	0.026	0.478	91.1	LG	5.945	0.015	0.449	91.1
		BH	5.738	0.017	0.072		BH	6.142	0.013	0.045	
Constant	11.016	0.001	9.489E+26		Constant	10.271	0.001	8.432E+27			

* : Accuracy

不同性别组的正确判别率明显不同, 性别组中雌性的性别判别率高于雄性(9/10), 见表 2。

两侧判别率比较结果显示右侧的判别率略微大于左侧, 分别是 89.32% 和 88.08%。卡方检验结果显示

两侧没有显著性差异($P=0.656$). 5 根跖骨左右两侧的 ROC 分析比较, 同样显示左右两侧之间的差异很小. ROC 曲线下的面积值(AUC)分别是 0.938 和 0.941. 说明两侧变量对性别的识别能力很好(表 3).

5 根跖骨逐步逻辑回归结果是 MT3、MT4 和 MT5 判别率较高(88.9%~95.6%). MT1 和 MT2 判别率较低(84.4%~86.7%). 左侧判别率最高的回归模型是 MT5(91.9%), 最低的是 MT2(84.8%); 右侧判别率最高的是 MT3(95.6%), 最低的是 MT2 和 MT1(84.8%).

表 2 两侧跖骨逐步逻辑回归分析的性别判别率(%)

	MT1			MT2			MT3			MT4			MT5		
	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T
Left	78.6	90.3	86.7	63.3	93.5	84.8	85.7	90.3	88.9	85.7	90.3	88.9	85.7	93.5	91.1
Right	71.4	90.3	84.4	71.4	90.3	84.4	92.9	96.8	95.6	92.9	90.3	91.1	85.7	93.5	91.1

The cut value is 0.50. Overall predictive value is 88.08% for the left foot and 89.32% for the right foot. Overall predictive value is 81.43% for the males and 91.91% for the females. M=Male, F=Female, T=Total (Male+Female).

表 3 两侧跖骨 ROC 分析曲线下面积

	MT1		MT2		MT3		MT4		MT5	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Area	0.924	0.922	0.924	0.945	0.956	0.956	0.947	0.947	0.938	0.933
SE*	0.041	0.041	0.039	0.034	0.028	0.041	0.046	0.042	0.037	0.042
P**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

*. Under the nonparametric assumption; **. Null hypothesis: true area=0.5.

3 讨 论

以前的研究资料证实非人灵长类跖骨性差的确存在, 这一性差特征一般不会受到个体形态大小的影响^[10,12]. 本结果表明: 太行山猕猴跖骨变量性差显著.

Tague^[13]报道了 17 种灵长类的掌跖骨的性差存在. McFadden 等报道了大猩猩和黑猩猩掌跖骨长度比值存在性差, 提出尽管它们主要是受遗传基因的控制, 但是同时也会受到出生前的胚胎发育环境中性激素释放的影响^[12]. 结果表明猕猴跖骨的 7 个变量中 LG 仍然是识别个体性别的最佳变量, 其次是跖骨头和跖骨底部的变量. Scheuer 等^[14]对现代英国人掌骨不同变量的性差大小进行研究, 结果显示掌骨骨干中间径的性差最大, 其次是掌骨底宽、头宽、头高和底高, 掌骨 LG 的性差最小. 其他一些研究也证实了掌骨 LG 性差相对较低, 趋向是掌骨骨干中间径性差较大, 其次是底宽和头宽^[1].

本结果与以前其他资料报道的结果不同. 推测原因如下: (1) 人类和非人灵长类的研究结果都证实跖骨 LG 存在明显性差, 而且常常用来作为性别鉴定的主要变量和最佳选择变量. 这主要是由于跖骨 LG 受遗传因素影响, 而掌跖骨骨干宽度变量很容易受行为习性和机械刺激因素的影响; 同一块骨骼的变量性差大小不同, 可能更好地反映了适当的机械负荷刺激引起骨骼重构和不同发育阶段的发育程度之间的交互作用. 实际上, 跖骨 LG 与其他变量的生长发育模式和影响因素是不同的; (2) 骨骼的生长发育是一个非常复杂的由激素控制的生物学现象, 他们的形态特征主要依赖基因和环境的影响. 不同的种属之间具有不同的遗传概貌^[15]. (3) 猕猴与人类跖骨变量性差大小不同, 很可能还与骨骼生长周期有关. 猕猴总体上生长周期较短, 所以对跖骨骨干粗壮度影响相对较小. 人类的生命活动周期相对较长, 所以对跖骨骨干粗壮度影响相对较大. (4) 可能与种属差异和生物多样性有关, 因为即使在人类的不同种群之间跖骨长度和宽度的性差大小是不同的^[16]. 虽然掌骨 LG 也是性差最为显著的变量, 但是是一些掌骨最先选择出来的变量不是 LG 而是掌骨的 HW 和 MW. 说明两个问题, LG 并不一定是性差最大或是鉴定性别识别的最佳变量, 就像在人类掌骨上那样^[17]; 二是掌骨和跖骨的性差大小或模式可能不同. 这一结果还需要进一步检验.

逻辑回归和 ROC 分析结果显示侧别之间的判别率没有显著性差异. 在不同的性别组中, 雌性的正确判别率较高, 雄性的正确判别率较低. 推测原因可能是(1)雌雄标本数量不对称, 雌性较多, 使得统计结果更加

稳定,这与猕猴种群结构(少雄多雌群体)组成有关,是自然选择的结果^[18]; (2) 雄性跖骨变量的标准差相对较大,说明雄性组内个体差异较大,结果95%置信区间增大,导致部分雄性标本与部分雌性标本之间出现重叠; Özer等^[19]根据古人类肩胛骨变量来进行判别分析. 结果显示雌性的性别正确判别率大于雄性,但是他们没有进一步解释原因. (3) Lazenby等^[17]采用古代人类第二掌骨进行性别判别分析, 结果发现无论是左侧还是右侧,男性的判别率都是较高的(>90.0%),女性的判别率只有65.0%. 他们解释这些差异的原因可能反映了古代女性的掌骨可能比现代人的掌骨具有更大的粗壮度^[17]. 他们的结果与本文结果相反. 作者推测还是与行为习惯、种属差异和掌跖骨位置差异有关.

本文研究结果表明5根跖骨无论是左侧还是右侧都是MT3、MT4和MT5的判别率相对最高,意味着这3根跖骨可能性差较大. 有关非人灵长类跖骨性差的研究较少. McFadden等^[12]报道了大猩猩和黑猩猩跖骨长度均值,大猩猩跖骨性差最大的是MT5,其次是MT4,最小的是MT1. 黑猩猩的性差明显小于大猩猩,但是跖骨LG均值性差最大的是MT4,其次是MT5,最小的是MT1. 同时指出无论是大猩猩还是黑猩猩可能和人类一样,跖骨在早期生长发育时明显受到性激素的调控^[12]. 他们的研究结果总体上是支持这里的猕猴跖骨性差的研究结果. Case等^[20]对现代美国人跖骨进行了研究. 结果显示: 跖骨LG性差较大的是MT5,其次是MT2,最小的是MT1. 他们的结果提示了现代人类的跖骨,特别是MT4和MT5,对个体的性别识别起到一定的作用^[20].

参 考 文 献

- [1] Khanpetch P, Prasitwattanseree S, Case D T, et al. Determination of sex from the metacarpals in a Thai population[J]. *Forensic Sci Int*, 2012, 217(1/2/3): 229-237.
- [2] Doaa A, Adel A. Sex determination by the length of metacarpals and phalanges: X-ray study on Egyptian population[J]. *J For Leg Med*, 2013, 20: 6-13.
- [3] 赵晓进. 太行山猕猴牙齿和掌跖骨性差研究[D]. 北京: 中国地质大学, 2007.
- [4] 宋洁, 王凤产, 毛晓静, 等. 太行山猕猴锁骨性差[J]. *河南师范大学学报: 自然科学版*, 2014, 42(10): 129-132.
- [5] 赵俊杰, 赵晓进, 党晓云, 等. 太行山猕猴跟骨性别判别[J]. *解剖学杂志*, 2009, 32(5): 687-689.
- [6] 赵晓进, 罗艳蕊, 李 茨, 等. 太行山猕猴颅骨变量主成分分析[J]. *河南师范大学学报: 自然科学版*, 2001, 29(2): 54-56.
- [7] 赵晓进. 太行山猕猴髌骨变量研究初报[J]. *人类学学报*, 2003, 22(1): 69-74.
- [8] 赵晓进, 雷梦云, 张红绪, 等. 太行山猕猴肩胛骨和肱骨的性别研究[J]. *动物学杂志*, 2004, 39(3): 351-54.
- [9] 闫鑫甜, 张弯弯, 徐小解, 等. 猕猴趾骨性别逐步判别分析[J]. *解剖学杂志*, 2012, 35(5): 670-674.
- [10] 胡海洋, 赵晓进, 周 迪, 等. 太行山猕猴掌骨性别判别分析[J]. *人类学学报*, 2013, 32(2): 226-232.
- [11] 张文彤. SPSS 11 统计分析教程[M]. 北京: 北京希望电子出版社, 2002.
- [12] McFadden D, Bracht M S. Sex difference in the relative lengths of metacarpals and metatarsals in gorillas and chimpanzees[J]. *Horm Behav*, 2005, 47(1): 99-111.
- [13] Tague RG. Variability of metapodials in primates with rudimentary digits: *Ateles geoffroyi*, *Colobus guereza*, and *Perodicticus potto* [J]. *Phys Anthropol*, 2002, 117(3): 195-208.
- [14] Shreuer J L, Elkington N M. Sex determination from metacarpals and first proximal phalanx[J]. *J Forensic Sci*, 1993, 38: 769-778.
- [15] Manolis S K, Eliopoulos C, Koiliias C G, et al. Sex determination using metacarpal biometric data from the Athens Collection[J]. *Forensic Sci Int*, 2009, 193: 130-135.
- [16] Smith S L. Attribution of hand bones to sex in population groups[J]. *J Forensic Sci*, 1996, 41: 469-477.
- [17] Lazenby R A. Identification of sex from metacarpals effect of side asymmetry[J]. *J Forensic Sci*, 1994, 39: 1188-1194.
- [18] Alicioglu B, Yilmaz A, Karakas H M, et al. Sex determination by interarticular distance of metacarpals and phalanges: a digital radiologic study in contemporary Turkish people [J]. *J Anat*, 2009, 3: 14-20.
- [19] Özer I, Katayama K, Sahgir M. Sex determination using the scapula in medieval skeletons from East Anatolia[J]. *Coll Antropol*, 2006, 30(2): 415-419.
- [20] Case D T, Ross A H. Sex determination from hand and foot bone lengths[J]. *J Forensic Sci*, 2007, 52(2): 264-269.

descending order were $Zn(CH_3COO)_2$ concentration, Vc concentration, pH, Na_2SO_3 concentration; The optimal formula for preserving green were $Zn(CH_3COO)_2$ concentration of $254\text{ mg} \cdot \text{kg}^{-1}$, Vc concentration of 0.048%, Na_2SO_3 concentration of 0.498%, pH value of 8.91, blanching temperature of $100\text{ }^\circ\text{C}$, blanching time of 20 s.

Keywords: walnut tassel; response surface; preserving green

(上接第 126 页)

Sex Determination of Logistic Regression from the Metatarsal Measurements in *Macaac Mulatta*

HU Haiyang^a, WANG Fengchan^b, HU Fengxia^c, TIAN Huaxiang^c,
MAO Xiaojing^c, ZHAO Zhe^a, ZHAO Xiaojin^c

(a. College of Life Science; b. Editorial Department; c. College of Fisheries, Henan Normal University, Xinxiang 453007, China)

Abstract: Because of the dichotomous nature of the sex type, logistic regression was used to be able to predict sex based on values of predictor variables of the metatarsals and ROC analysis was used to assess sexing performance at different variables, different sides of the feet and for each of the five metatarsals. The samples selected for this study consist of 45 adult skeletons of *Macaac mulatta* (31 females, 14 males) from the Taihang Mountains. Seven variables were recorded on each bone of five metatarsals. Data were analyzed using SPSS 20.0 version. For the models, the best variable for the metatarsals was LG followed by HH, HW, BH and BW. A logistic regression model was formed to estimate sex, ranging from 84.4%–95.6% for the pooled individuals being correctly classified. There were very small bilateral differences by using ROC analysis, and the AUC values were similar between the two sides, with 0.922 and 0.956 for left and right bones, respectively. For five metatarsals of both feet, the percentage of correct classification was higher for MT3, MT4 and MT5 (88.9%–95.6%) and lower MT1 and MT2 (84.4%–86.7%). In general, results from this study suggest that the measurements of metatarsals appear to be good discriminators of sex. The lengths of metatarsals are largely genetically determined; and the measures of robusticity of feet are largely subject to environmental modification. These findings suggest that the pattern of the sex differences for metatarsals may well reflect the interplay from both genetic, environmental and behavioral factors.

Keywords: *Macaca mulatta*; metatarsal; sex dimorphism; logistic regression; ROC analysis