

MATTERS ARISING

Sexual dimorphism in early anthropoids

FLEAGLE *ET AL.*¹ provide an estimate of sexual dimorphism in extinct anthropoid primates based on a positive relationship between the degree of sexual dimorphism in body weight and the coefficient of variation (c.v.) of mandible depth in extant anthropoids. A cursory glance at Table 1 shows that their data on the degree of body weight dimorphism are incorrect. As a striking example, capuchin monkeys (*Cebus apella*) are supposedly more dimorphic than gorillas (*Gorilla gorilla*) which, as any visitor to the zoo knows, is simply not the case. Gorillas are far more dimorphic than capuchin monkeys.

A closer examination of the data reveals that 9 (those taken from our study on sexual dimorphism²) out of 13 species do not express a mean male weight/mean female weight ratio as intended by Fleagle *et al.*, but instead show a ratio between the cube roots of male and female weights. In our study we used the cube root of body weight to obtain a linear dimension comparable to a linear canine dimension. Using corrected data for body weight, I recalculated its correlation with the c.v. of mandible depth and obtained a correlation coefficient, $r = 0.597$, which is substantially lower than $r = 0.940$ as calculated by Fleagle *et al.* This leaves >60% of the variance unexplained, which in turn indicates that the coefficient of variation is a much poorer predictor of the degree of sexual dimorphism than they suggest. A solid quantitative method for estimating the degree of sexual dimorphism in fossils based on the variability within a species and independent of prior sex determination of individuals still eludes us.

WALTER LEUTENEGGER

Department of Anthropology,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Fleagle, J. G., Kay, R. F. & Simons, E. L. *Nature* **287**, 328–330 (1980).

2. Leutenegger, W. & Kelly, J. T. *Primates* **18**, 117–136 (1977).

FLEAGLE *ET AL.* REPLY—Leutenegger correctly points out that we inadvertently mixed the body size data in Table 1 of our paper¹ and therefore overestimated the correlation coefficient. In rechecking our original data, it turns out that we incorrectly transcribed the c.v. of mandibular depth for *Cebus apella* as well; the correct value is 14.5. When the

correlation coefficient for the entire sample was recalculated, using the cubed roots of body weight as suggested by Leutenegger, the correct value for r was 0.755. Although this value is lower than that previously reported, it is still highly significant ($P < 0.01$) and demonstrates a positive correlation between c.v. of mandible depth and sexual dimorphism in body weight. This correction in no way affects our major points: that the coefficient of variation for mandible depth and the amount of canine dimorphism in the Fayum anthropoids are considerably higher than that found in any of the monogamous taxa.

JOHN G. FLEAGLE

Department of Anatomical Sciences,
Health Sciences Center,
State University of New York
at Stony Brook,
Long Island, New York 11794, USA

RICHARD F. KAY

Department of Anatomy,
Duke University Medical Center,
Durham, North Carolina 27710, USA

ELWYN L. SIMONS

Duke University Center for the
Study of Primate Biology and History,
3705 Erwin Road,
Durham, North Carolina 27705, USA

1. Fleagle, J. G., Kay, R. F. & Simons, E. L. *Nature* **287**, 328–330 (1980).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

Enzyme heterozygosity and morphological variance

HANDFORD¹ recently reported a lack of association between enzyme heterozygosity and measures of morphological

variance in the Rufous-collared sparrow. He suggested that this negative case, which was in apparent contrast to the findings of myself² and Mitton³, could be explained by an inherent difference, with regard to this relationship, between homoiotherms and poikilotherms. Although Handford makes several statistical points worth consideration, he ignores the fundamental problem of statistical power in his analysis. Essentially, his data set is so small that unless we assume the putative relationship between heterozygosity and morphological variability to be extremely strong, the ability to reject the null hypothesis of equal variances will be very restricted.

If enzyme heterozygosity does exert some influence on morphological variability, our biological intuition suggests that the contribution of individual loci will be subtle. Examination of the data sets for the monarch butterfly² and the killifish³ clearly support this notion. For example, the ratio of homozygote to heterozygote variances in the monarch is only 1.17 for the entire set of loci. This does not diminish the evolutionary importance of such a relationship, but rather places realistic expectations on the magnitude of the effect. Furthermore, because we are comparing variances, the ability to make strong statistical inferences about such associations will be constrained without large sample sizes and many independent observations.

The statistical power for the variability measures and sample sizes reported by Handford² is difficult to quantify. Regardless of the drawbacks of the F -test, it can be used in a simple univariate case to illustrate the limitations of these sample sizes. For a normally distributed character, the homozygote/heterozygote F ratios associated with Handford's larger sample sizes which would be statistically significant ($P < 0.05$) are 1.8 (d.f. = 130, 22) and 1.5 (d.f. = 76, 49). These are unrealistically large ratios given the subtlety of the relationship expected and observed in other organisms. The problem of type II error does not disappear by using Handford's particular treatment or by applying the sign test. In addition, Handford uses multivariate measures of variability which further reduces the statistical power with the inclusion of each new character. Indeed, three of the five tests for matrix homogeneity in females cannot be carried out because the sample sizes are so small that there are insufficient degrees of freedom.

His cautionary statement that one nonsignificant result does not negate the relationship is true, but a nonsignificant result for a truly powerful data set would have permitted some strong inference to