Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Human Evolution 49 (2005) 618–642

Do homoiologies impede phylogenetic analyses of the fossil hominids? An assessment based on extant papionin craniodental morphology

Stephen J. Lycett [a], Mark Collard [b,c,*]

[a] *Leverhulme Centre for Human Evolutionary Studies, Department of Biological Anthropology, University of Cambridge, Cambridge, United Kingdom*
[b] *Department of Anthropology and Sociology, University of British Columbia, Vancouver, British Columbia, Canada*
[c] *AHRB Centre for the Evolutionary Analysis of Cultural Behaviour, University College London, London, United Kingdom*

Received 8 June 2004; accepted 8 July 2005

Abstract

Homoiologies are phylogenetically misleading resemblances among taxa that can be attributed to phenotypic plasticity. Recently, it has been claimed that homoiologies are widespread in the hominid skull, especially in those regions affected by mastication-related strain, and that their prevalence is a major reason why researchers have so far been unable to obtain a reliable estimate of hominid phylogeny. To evaluate this ''homoiology hypothesis,'' we carried out analyses of a group of extant primates for which a robust molecular phylogeny is available—the papionins.

We compiled a craniometric dataset from measurements that differ in their susceptibility to mastication-related strain according to developmental considerations and experimental evidence. We used the coefficient of variation and analysis of variance with post hoc least significant difference comparisons in order to evaluate the variability of the measurements. The prediction from the homoiology hypothesis was that dental measurements, which do not remodel in response to strain, should be less variable than low-to-moderate-strain measurements, and that the latter should be less variable than high-strain measurements. We then performed phylogenetic analyses using characters derived from the measurements and compared the resulting phylogenetic hypotheses to the group's consensus molecular phylogeny. The prediction was that, if the homoiology hypothesis is correct, the agreement between the craniometric and molecular phylogenies would be best in the analyses of dental characters, intermediate in the analyses of low-to-moderate-strain characters, and least in the analyses of high-strain characters.

The results of this study support the suggestion that mastication-related mechanical loading can result in variation in hominid cranial characters. However, they do not support the hypothesis that homoiology is a major reason why phylogenetic analyses of hominid crania have so far yielded conflicting and weakly supported hypotheses of

* Corresponding author. Department of Anthropology and Sociology, University of British Columbia, 6303 NW Marine Drive, Vancouver, British Columbia, Canada, V6T 1Z1. Tel.: +1 604 822 4845; fax: +1 604 822 616.
E-mail address: mark.collard@ubc.ca (M. Collard).

doi:10.1016/j.jhevol.2005.07.004

relationship. These findings are consistent with a recent test of the homoiology hypothesis using craniodental data from extant hominoids, and cast doubt on the validity of the homoiology hypothesis, as originally formulated.

Keywords: Phylogeny; Homoplasy; Phenotypic plasticity; Phenotypic variation; Homoiology; Strain; Mastication; Cladistics; Hominid; Papionin

Introduction

Phylogenetic analysis is central to human evolutionary research. A reliable phylogeny is required to establish ancestor–descendent relationships, to evaluate hypotheses concerning the nature and number of adaptive changes in human evolution, and to test evolutionary scenarios that link events in human evolution with wider patterns of faunal evolution and with changes in the environment (Eldredge and Tattersall, 1975). Unfortunately, we currently do not have such a phylogeny for the fossil hominids (Lieberman, 1995; Collard and Wood, 2000; Curnoe, 2003; Hawks, 2004). Despite the availability of numerous well-dated specimens and sophisticated methods of phylogenetic reconstruction, the phylogenetic relationships of many fossil hominid species remain uncertain (Corruccini, 1994; Lieberman, 1995; Lieberman et al., 1996; Wood and Collard, 1999). This problem is illustrated by the small increases in length required to alter the topologies of the cladograms favored in recent analyses. For instance, Skelton et al.'s (1986) most parsimonious cladogram, in which *Homo habilis* formed a sister group with *Paranthropus* to the exclusion of *Australopithecus africanus*, was supported by only one more character than the next most parsimonious cladogram, which linked *Paranthropus* with *A. africanus* to the exclusion of *H. habilis*. Similarly, although the cladograms favored by Wood (1991) and Strait et al. (1997) suggested that *Homo* is monophyletic, these cladograms are only slightly shorter than ones in which *Homo* is paraphyletic (Wood and Collard, 1999). The problems we face in relation to hominid phylogeny are further illustrated by the recent work of Strait and Grine (2004). Their bootstrap analyses not only returned insignificant levels of support for many fossil hominid phylogenetic relationships, but also failed to support the widely accepted relationships among the extant hominoids at the 70% level, which is commonly used to classify clades as statistically significant in biological applications of the phylogenetic bootstrap (Hillis and Bull, 1993).

Over the last few years it has become clear that our inability to reliably reconstruct hominid phylogenetic relationships is due primarily to the presence of numerous homoplasies in available datasets. Homoplasies are resemblances between taxa that can be ascribed to processes other than descent from a common ancestor, and that imply relationships that conflict with the best estimate of phylogeny for the taxa (Willey, 1911; Simpson, 1961; Hennig, 1966; Cain, 1982; Patterson, 1982; Sober, 1988; McHenry, 1996; Sanderson and Hufford, 1996; Lockwood and Fleagle, 1999). Homoplasies are problematic because they can be mistaken for shared derived similarities (i.e., synapomorphies), which are the principal evidence for phylogeny. Ideally, a character state data matrix should contain a small number of homoplasies in relation to the number of synapomorphies. In such circumstances, it is possible to obtain an unambiguous estimate of phylogeny using parsimony analysis, which favors the hypothesis of relationship requiring the least number of changes to account for the distribution of character states among a group of taxa (Wiley et al., 1991; Quicke, 1993; Kitching et al., 1998; Schuh, 2000). However, in phylogenetic studies of the hominids, the ratio of putative homoplasies to inferred synapomorphies has generally been high, around 1:2 (e.g., Skelton et al., 1986; Chamberlain and Wood, 1987; Wood, 1991; Skelton and McHenry, 1992; Lieberman et al., 1996; Strait et al., 1997). When homoplasy reaches such high levels,

parsimony analysis tends to yield multiple phylogenies that are essentially equally plausible (Lieberman et al., 1996).

Recently, it has been suggested that many hominid homoplasies are likely to be homoiologies (Lieberman, 1995, 1997, 1999, 2000; Lieberman et al., 1996; Collard and Wood, 2000; Gibbs et al., 2000). Homoiologies are phylogenetically misleading resemblances among a group of taxa that can be ascribed to phenotypic plasticity. That is, homoiologies are homoplasies that result from the expression by a genotype of different phenotypes in response to different environmental conditions (Reidl, 1978; Lieberman, 1995, 1997, 1999, 2000; Lieberman et al., 1996). The suggestion that homoiologies may be an important impediment to phylogenetic analyses of the fossil hominids is based on work examining how mechanical loading affects bone. This work suggests that interactions between the skeleton and its mechanical environment greatly influence bone size and shape (Currey, 1984; Lanyon and Rubin, 1985; Frost, 1986, 1998; Herring, 1993; Lieberman, 1995, 1997, 1999, 2000; Lieberman et al., 1996; Lieberman and Crompton, 1998; Martin et al., 1998; Skerry, 2000). For example, mechanical loading experienced during development has been found to affect both the growth of cortical bone in diaphyses and the growth of trabecular bone in epiphyses (Currey, 1984; Lanyon and Rubin, 1985; Frost, 1986; Lieberman and Crompton, 1998; Martin et al., 1998). Likewise, studies of individuals experiencing lower than normal mechanical strains (e.g., following denervation, bed-rest, or exposure to gravity-free environments) indicate that bone may resorb at rapid rates in many regions of the skeleton (Martin et al., 1998). According to proponents of the ''homoiology hypothesis,'' the responsiveness of bony morphology to environmental stimuli means that individuals that behave in similar ways are likely to develop osteological similarities that are phylogenetically misleading. This idea was first outlined by Lieberman (1995: 165), who argued that ''nonheritable, nonhomologous similarities often occur in functionally important regions that experience a high level of strain.'' Other studies that have suggested that phenotypic plasticity may be a major cause of homoplasy among hominids and other primates include Lieberman et al. (1996), Lieberman (1997, 1999, 2000), Collard and Wood (2000, 2001), and Gibbs et al. (2000, 2002).

Collard et al. (in press) used extant hominoid craniodental data to assess the suggestion that phenotypic plasticity is likely to have given rise to many of the homoplasies that are encountered in hominid phylogenetic analyses. Collard et al. divided their cranial characters into ''high-strain characters'' and ''low-to-moderate-strain characters'' on the basis of their likely susceptibility to mastication-related strain. Dental characters were also included in the analysis because, although they are subject to high levels of strain during mastication, they do not remodel. In their first set of analyses, Collard et al. used the coefficient of variation (CV) and the *t*-test to evaluate the phenotypic plasticity of the three sets of measurements. This approach, which was also used by Wood and Lieberman (2001), assumes that differences in variation among characters are likely to reflect differences in phenotypic plasticity. Specifically, characters that are under relatively close genetic control are expected to be less variable than characters that are more heavily affected by environmental factors such as strain. In their second set of analyses, Collard et al. (in press) used cladistic methods and the widely accepted molecular phylogeny for the extant hominoids to test the hypothesis that phenotypic plasticity is a cause of phylogeny-obscuring homoplasy. The results of the CV/*t*-test-based analyses were in line with the prediction from the hypothesis to the extent that high-strain measurements exhibited significantly more variation than either dental measurements or low-to-moderate-strain measurements. In contrast, the results of the phylogenetic analysis were not those predicted by the hypothesis. The phylogeny derived from high-strain characters fitted the consensus molecular phylogeny considerably better than the phylogenies obtained using the low-to-moderate-strain and dental characters. Thus, the results of Collard et al.'s analyses were not in keeping with the suggestion that homoiology is an important form of homoplasy in hominid phylogenetic analyses.

The primary aim of our study was to further test the homoiology hypothesis by repeating Collard et al.'s (in press) analyses with data from another extant primate group for which a robust molecular phylogeny is available, the papionins (Fig. 1) (Disotell et al., 1992; Disotell, 1994, 1996, 2000; Harris and Disotell, 1998; Harris, 2000; Page and Goodman, 2001; Tosi et al., 2003). In addition, we wanted to address some of the shortcomings of Collard et al.'s (in press) study. One of these is that the dataset they employed was relatively small, comprising values for just 36 measurements. A second is that the dataset was not collected with testing the homoiology hypothesis in mind. Rather, it was taken from a previous study that was designed to assess sexual dimorphism in a range of primates (Wood, 1975). While Collard et al. (in press) employed a review of published in vivo strain-gauge analyses to assign characters to specific groups on the basis of strain, there is a possibility that craniofacial regions subject to strain were inadequately sampled. This could have led to an underestimation of load-induced plasticity and/or a false assessment of the relative reliability of such traits when employed in cladistic analyses. The final shortcoming of Collard et al.'s (in press) study that we address here is that they did not assess the potential impact of sexual dimorphism or allometry on their results.

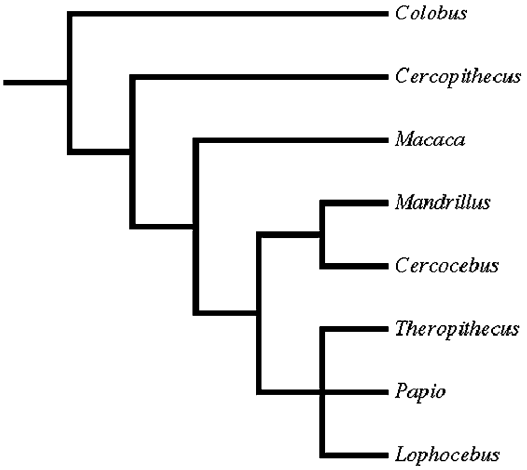

Fig. 1. Consensus molecular phylogeny for the taxa used in this study.

Materials and methods

An extensive review of published in vivo mastication-induced strain-gauge analyses was undertaken (Hylander, 1975, 1977, 1979a,b, 1984, 1986, 1988; Hylander and Bays, 1979; Brehnan et al., 1981; Demes, 1984; Hylander and Crompton, 1986; Hylander et al., 1987, 1991a,b, 1992, 1998, 2000; Herring and Mucci, 1991; Daegling, 1993; Hylander and Johnson, 1994, 1997, 2002; Ross and Hylander, 1996, 2000; Herring et al., 1996; Daegling and Hylander, 1997, 1998, 2000; Spencer, 1998; Rafferty and Herring, 1999; Wall, 1999; Dechow and Hylander, 2000; Herring and Teng, 2000; Ravosa and Profant, 2000; Ravosa et al., 2000a,b; Ross, 2001; Meyer et al., 2002). The purpose of this review was to identify features of the primate skull that experience different levels of strain during mastication. Particular attention was paid to regions that routinely experience strain gradients in the order of $\geq 1000\ \mu\epsilon$ during incision, biting, and mastication, as strains of this magnitude are known to induce bone growth (Currey, 1984; Martin and Burr, 1989; Martin et al., 1998). Based on the information recovered during the literature review, a list of 60 inter-landmark measurements was compiled (Table 1). Twenty-two of the measurements were included because they relate to features that strain-gauge analyses indicate experience high levels of strain during mastication. These high-strain measurements are located on the mandible, mandibular fossa, zygomatic bone, and zygomatic arch. A further 22 measurements were included because they are associated with features of the primate skull that experience low-to-moderate levels of strain during mastication according to the available strain-gauge data. These low-to-moderate-strain measurements are located on the viscerocranium, neurocranium, and basicranium. The remaining 16 measurements are labiolingual and buccolingual diameters of teeth. These were included because teeth, unlike osseous features, do not remodel in response to mechanical loading. Labiolingual and buccolingual diameters were employed instead of mesiodistal diameters to avoid the potentially confounding effects of interstitial wear (Hinton, 1982).

Table 1
Measurements employed

Measurement	Description	Source, with original code in parentheses
High-strain characters ($n = 22$)		
1. Mandibular corpus height at M_1	Minimum distance between the most inferior point on the base and the lingual alveolar margin at the midpoint of M_1	Wood, 1991 (#150)
2. Mandibular corpus width at M_1	Maximum width at right angles to Measurement 1, taken at midpoint of M_1	Wood, 1991 (#151)
3. Height of mandibular symphysis	Minimum distance between the base of the symphysis and infradentale	Wood, 1991 (#141)
4. Depth of mandibular symphysis	Maximum depth at right angles to symphyseal height	Wood, 1991 (#142)
5. Condylar height	Maximum distance between base of ramus and superior point of condyle	Wood, 1975 (#36)
6. Coronoid height	Maximum distance between base of ramus and superiormost point of coronoid process	Wood, 1975 (#38)
7. Ramus breadth	Maximum width in the (anterior–posterior) body of ramus	Wood, 1975 (#42)
8. Mandibular condyle head length	Maximum length in anterior–posterior plane	Wood, 1975 (#41)
9. Mandibular condyle head width	Maximum width in medial–lateral plane	Wood, 1975 (#40)
10. Bigonial width	Minimum distance between the inner margins of left gonion and right gonion	Wood, 1975 (#44)
11. Inner alveolar breadth at M_3	Minimum chord distance between the walls of the lingual mandibular alveoli at the midpoint of M_3	Wood, 1975 (#49)
12. Height of zygomatic arch	Maximum height at zygomatico-temporal suture	This study
13. Thickness of zygomatic arch	Maximum width at zygomatico-temporal suture	This study
14. Mandibular fossa length	Minimum chord distance between the tympanic plate and the inferiormost projection of the articular eminence; taken midway along breadth measurement (see below)	Wood, 1991 (#80)
15. Mandibular fossa breadth	Minimum chord distance in the coronal plane between the tip of the entoglenoid process and the lateralmost extent of the articular eminence	Wood, 1991 (#82)
16. Orbitale to zygomaxillare	Chord distance between orbitale and zygomaxillare	Wood, 1991 (#58)
17. Mandibular corpus thickness at M_3	Minimum distance between the inferiormost point on the base and the lingual alveolar margin at the midpoint of M_3	Wood, 1991 (#157)
18. Mandibular corpus height at M_3	Maximum width at right angles to Measurement 17, taken at midpoint of M_3	Wood, 1991 (#158)
19. Lower inter-canine distance	Minimum chord distance between the walls of the mandibular canine alveoli	Wood, 1991 (#166)
20. Upper ramus breadth	Distance between midpoint of the articular surface of the condyle (instrumentally determined; see Measurements 8 and 9) and the superiormost point of coronoid process	This study
21. Bicondylar breadth	Right condylion laterale to left condylion laterale	Wood, 1975 (#37)
22. Height of ramus to sigmoid notch	Maximum distance between base of ramus and inferiormost point of sigmoid notch	This study
Low-to-moderate strain ($n = 22$)		
23. Orbital breadth	Distance between maxillofrontale and ektoconchion	Wood, 1991 (#56)
24. Orbital height	Maximum distance between the superior and inferior orbital margins in a direction perpendicular to orbital breadth	Wood, 1991 (#57)
25. Interorbital breadth	Chord distance between maxillofrontale	Wood, 1991 (#55)
26. Biorbital breadth	Chord distance between ektoconchion	Wood, 1991 (#50)
27. Glabella to rhinion	Chord distance between glabella and rhinion	This study
28. Rhinion to nasospinale	Chord distance between rhinion and nasospinale	Wood, 1991 (#70)
29. Nasion to inion	Chord distance between nasion and inion	This study

Table 1 (*continued*)

Measurement	Description	Source, with original code in parentheses
30. Basion to bregma	Chord distance between basion and bregma (in specimens with a sagittal crest, "bregma" was taken to be the plane of the surrounding vault surface)	Wood, 1991 (#4)
31. Biparietal breadth	Maximum breadth across homologous points on the left and right parietal bones	Wood, 1991 (#9)
32. Biporionic breadth	Chord distance between left porion and right porion	Wood, 1991 (#11)
33. Opisthion to lambda	Chord distance between opisthion and lambda	This study
34. Hormion to basion	Chord distance between hormion and basion	This study
35. Opisthion to inion	Chord distance between opisthion and inion	Wood, 1991 (#37)
36. Porion to basion	Chord distance between porion and basion	This study
37. Pterion to bregma	Chord distance between pterion and bregma	This study
38. Basion to opisthion	Minimum distance between basion and opisthion	Wood, 1991 (#76)
39. Width of foramen magnum	Maximum distance in the coronal plane between the inner margins of the foramen magnum	Wood, 1991 (#77)
40. Pterion to lambda	Chord distance between pterion and lambda	This study
41. Porion to opisthion	Chord distance between porion and opisthion	This study
42. Staphilion to hormion	Chord distance between staphilion and hormion	This study
43. Pterion to pterion	Chord distance between left pterion and right pterion	This study
44. Hormion to porion	Chord distance between hormion and porion	This study
Dental ($n = 16$)		
45. I_1 labiolingual diameter	Maximum crown diameter perpendicular to the basal part of the labial enamel surface	Wood, 1991 (#248)
46. I_2 labiolingual diameter	Maximum crown diameter perpendicular to the basal part of the labial enamel surface	Wood, 1991 (#251)
47. C_1 labiolingual diameter	Maximum diameter of the crown in the labiolingual axis of the tooth	Wood, 1991 (#254)
48. P_3 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#258)
49. P_4 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#272)
50. M_1 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#286)
51. M_2 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#314)
52. M_3 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#342)
53. I^1 labiolingual diameter	Maximum crown diameter perpendicular to the basal part of the labial enamel surface	Wood, 1991 (#187)
54. I^2 labiolingual diameter	Maximum crown diameter perpendicular to the basal part of the labial enamel surface	Wood, 1991 (#189)
55. C^1 labiolingual diameter	Maximum diameter of the crown in the labiolingual axis of the tooth	Wood, 1991 (#191)
56. P^3 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#194)
57. P^4 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#203)
58. M^1 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#212)
59. M^2 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#224)
60. M^3 buccolingual diameter	Maximum distance between the buccal and lingual borders taken at right angles to the longitudinal axis of the crown	Wood, 1991 (#236)

Values for the 60 measurements were obtained from specimens belonging to the six extant papionin genera, along with two outgroup taxa. The species sampled were *Cercocebus torquatus*, *Cercopithecus mitis*, *Colobus polykomos*, *Lophocebus albigena*, *Macaca fascicularis*, *Mandrillus leucophaeus*, *Papio anubis*, and *Theropithecus gelada*. The measurements were taken on 10 males and 10 females of each species. All the specimens were wild-shot adults. A specimen was judged to be adult if its third molars were erupted. Specimens were deemed to be male or female on the basis of museum records. Cranial and mandibular measurements were recorded to the nearest 1 mm, and dental measurements to the nearest 0.1 mm. All data were collected by SJL with sliding digital calipers and analogue spreading calipers.

Four sets of analyses were carried out to evaluate the homoiology hypothesis. The first evaluated the prediction that measurements of osseous structures subject to high levels of mastication-related strain should be more variable than measurements of osseous structures that are subject to low-to-moderate levels of mastication-related strain, and that the latter should in turn be more variable than dental measurements. We recognize that the variation of many cranial traits will often be affected by environmental factors other than strain. However, in contrast to certain other sources of epigenetic variation, fluctuation in strain levels can potentially cause significant bone remodeling throughout ontogeny and beyond the normal phase of somatic growth (Lanyon and Rubin, 1985; Martin and Burr, 1989; Herring, 1993). More importantly, we did not assume that traits outside those we have designated "high strain" would not vary. Rather, in line with Wood and Lieberman (2001) and Collard et al. (in press), we predicted that, on average, high-strain traits would be significantly more variable than traits that experience low-to-moderate levels of strain and dental traits, which do not remodel. Following Wood and Lieberman (2001) and Collard et al. (in press), phenotypic variation was assessed using the coefficient of variation (CV). Coefficients of variation were calculated for each trait, and mean CVs for each group of measurements (i.e., high-strain, low-to-moderate-strain, and dental traits) were computed. In order to test for statistically significant differences between the mean CVs of each trait group, analysis of variance (ANOVA) with post-hoc least significant difference pairwise comparisons was employed. With the latter test, there is no need to reduce the critical p-value below 0.05 for pairwise comparisons when the ANOVA is significant (Dytham, 2003), which was the case here. Since ANOVA assumes data are normally distributed (Sokal and Rohlf, 1995), the CVs were logarithmically transformed (log e) prior to analysis. We predicted that the CVs for the high-strain measurements would be significantly higher than the CVs for the low-to-moderate-strain measurements, and that the CVs for the latter would be significantly higher than the CVs for the dental measurements.

The second set of analyses evaluated the prediction that characters from regions of the cranium that are subject to high levels of strain will yield phylogenies that are less compatible with the papionin consensus phylogeny than either low-to-moderately strained osseous characters or dental characters. In order to employ metric data in a cladistic analysis, it is necessary to adjust the data to counter the confounding effects of the body-size differences among the taxa, and then to convert the resulting values into discrete character states (Simon, 1983; Almeida and Bisby, 1984; Thorpe, 1984; Archie, 1985; Chamberlain and Wood, 1987; Baum, 1988; Wood, 1991; Thiele, 1993; Strait et al., 1996; Rae, 1998; Collard and Wood, 2000, 2001). Size adjustment was accomplished by dividing each specimen value by the geometric mean of the specimen's values (Mosimann, 1970; Jungers et al., 1995). The geometric mean was computed as the *n*th root of the product of all *n* variables (Jungers et al., 1995). Testing for skewness and kurtosis indicated that the size-corrected data were normally distributed.

After size adjustment, the size-corrected data were converted into discrete character states using divergence coding (Thorpe, 1984). This technique proceeds by calculating the mean values for the taxa, and then testing the differences among them for statistical significance on a pairwise basis. The means are then ranked in ascending order, and a taxon-by-taxon matrix is compiled. Each cell in the

top row of the matrix is filled with a taxon name such that the rank of the taxa decreases from left to right. The cells in the first column of the matrix are also filled with the names of the taxa on the basis of their rank, with the highest-ranked taxon being placed in the top cell, and the lowest-ranked taxon in the bottom cell. Thereafter, each cell in the matrix is assigned a score of -1, $+1$, or 0. A cell is scored as $+1$ if the mean of the taxon in the column is significantly greater than the mean of the taxon in the row. A cell is scored with a -1 if the mean of the taxon in the column is significantly lower than the mean of the taxon in the row. If the difference between the mean of the taxon in the column and the mean of the taxon in the row is not significant, the cell is filled with a 0. Once the matrix is completely filled, the total score of each column (i.e., the sum of every 0, -1, and $+1$ for a given taxon) is calculated. Lastly, an integer is added to each taxon's total to ensure that every score is positive. In converting the dataset, Student's *t*-test (two-tailed) was used to test for statistical significance ($p \leq 0.05$), and five was added to the taxon totals. The Bonferroni correction was not employed because it heightens the risk of making type II errors (Perneger, 1998; Nakagawa, 2004). An elevated type II error rate is likely to be especially problematic in a phylogenetic study because fewer differences among the taxa will be recognized, and therefore more false similarities will be incorporated into the character state data matrix.

After coding, the three groups of traits were independently subjected to parsimony analysis using the phylogenetic reconstruction program PAUP* 4 (Swofford, 1998). In all three analyses, the traits were treated as linearly ordered and freely reversing (Chamberlain and Wood, 1987; Slowinski, 1993; Rae, 1997), and the minimum-length cladogram was identified using the branch-and-bound algorithm. The most parsimonious cladograms recovered from the three groups of morphological traits were then compared to the consensus molecular phylogeny for the papionins in order to determine the amount of homoplasy exhibited by each morphological trait group. Some researchers reject this approach because it assumes that molecular phylogenies are more reliable than phylogenies derived from the morphological data (e.g., Smith, 1994; Kluge, 1998; Wiens, 2004). We understand why these workers take this view, but believe that they are mistaken. There are several reasons why, when a conflict occurs between molecular and hard-tissue-based phylogenies, the former should be favored, at least at the low taxonomic levels considered here. First, biological phylogenetic relationships are genetic relationships. It is genes that are passed between generations, not morphological characters. Thus, in phylogenetics, morphology can never be more than a proxy for molecular data. Second, it is well documented that many reproductively defined species are genetically distinct but dentally and osteologically indistinguishable. Since speciation events create phylogenetic relationships, there is thus an expectation that dental and skeletal characters will be less useful for phylogeny estimation than genetic characters. Third, some of the techniques of molecular phylogenetics have been tested successfully on taxa of known phylogeny (Fitch and Atchley, 1987; Atchley and Fitch, 1991; Hillis et al., 1992), whereas comparable analyses of morphological data have not been successful (Fitch and Atchley, 1987). In addition to the foregoing general points, the molecular cladogram used in this study is supported by multiple lines of independent molecular and karyological evidence. Given that congruence among multiple lines of evidence is the strongest possible support for a phylogenetic hypothesis, the notion of evaluating morphological phylogenies in the light of the molecular ones is strongly supported. Overall, therefore, we believe that this part of the research protocol is justified.

To assess the fit between the most parsimonious cladogram recovered from each group of traits and the consensus molecular phylogeny for the papionins, both topologies were imposed on the relevant part of the dataset in MacClade 4 (Maddison and Maddison, 1998), and the percentage difference in length between the cladograms calculated. Based on the homoiology hypothesis, our expectation was that the dental traits should exhibit the smallest increase in length between the most parsimonious cladogram and the molecular phylogeny, the low-to-moderate-strain traits

should exhibit the an intermediate increase in length, and the high-strain traits should exhibit the greatest increase in length. The matrices used in this analysis are presented in Appendix 1.

The third set of analyses examined the possibility that sexual dimorphism has confounded attempts to test the homoiology hypothesis. This was accomplished by repeating the above-described analyses for males and females separately, and then comparing the results of the sex-specific analyses with each other and with the results of the combined-sex analyses. We assumed that, if sexual dimorphism is not a confounding factor, then the results for the male-only and female-only analyses should be congruent, whereas, if sexual dimorphism is a confounding factor, then the results of the male-only and female-only analyses should differ in important respects. The matrices used in this analysis are shown in Appendices 2 and 3.

The last set of analyses investigated a further potentially problematic aspect of the research protocol—the use of the geometric mean method of size correction. This method equalizes the volumes of the specimens while maintaining their original shapes (Jungers et al., 1995). Unfortunately, as Jolly (2001) has recently reiterated, the geometric mean method does not remove size-related shape differences among taxa. We consider this to be a less serious drawback than those associated with the main alternative method, regression-based size adjustment. The latter is heavily dependent both on the line-fitting technique and the dataset employed to generate the regression equation (Aiello, 1992; Falsetti et al., 1993; Martin, 1993; Jungers et al., 1995). In addition, Jungers et al. (1995) showed that allometric methods of size adjustment can fail to correctly identify specimens of the same shape. Nevertheless, it is possible that some of the homoplastic similarities in the dataset could be allometric in nature. Depending on the number and especially the distribution of such "allometric homoplasies," the results of the phylogenetic analyses may or may not be valid. To evaluate this possibility, we carried out a series of Pearson correlation analyses in which the statistical association between each size-corrected character and the relevant geometric mean was measured, and then re-ran the cladistic analyses after excluding the characters that the correlation analyses suggested were significantly correlated with the geometric means at the $p \leq 0.05$ level.

Results

Mechanical loading and phenotypic plasticity

The results of the first set of analyses were mixed with regard to the prediction that measurements based on morphology subject to high levels of mastication-related strain should be more variable than measurements subject to low-to-moderate levels of mastication-related strain, and that these should be more variable than dental measurements (Table 2). As predicted, the CVs derived from the high-strain measurements were significantly higher than the CVs obtained from the low-to-moderately strained regions in all eight taxa. However, contrary to expectation, the low-to-moderate-strain measurements were not significantly more variable than the dental measurements. Rather, in seven of the eight taxa (*Cercocebus*, *Colobus*, *Lophocebus*, *Macaca*, *Mandrillus*, *Papio*, and *Theropithecus*) the low-to-moderate-strain measurements were actually less variable than the dental measurements, although the differences were not significant. Also contrary to expectation, the high-strain measurements were significantly more variable than the dental measurements in only five of the eight taxa. In *Lophocebus*, *Macaca*, and *Papio*, the CVs of the low-to-moderate and dental measurements were statistically indistinguishable according to the ANOVA.

Phenotypic plasticity and homoplasy

In the second set of analyses maximum parsimony analysis was used to test the predictions of the homoiology hypothesis concerning the relative phylogenetic utility of each trait group. According to the homoiology hypothesis, non-remodeling (i.e., dental) characters and those drawn from low-to-moderately strained regions of the cranium should yield phylogenies that are

Table 2
Results of mixed sex CV/ANOVA analyses

Taxon	CV			ANOVA		
	High	LM[1]	Dental	High vs. LM	High vs. dental	LM vs. dental
Macaca	13.3	9.3	10.5	0.004***	0.052	0.450
Lophocebus	9.7	6.7	7.8	0.001***	0.069	0.184
Cercocebus	13.0	8.4	8.7	0.000***	0.002***	0.884
Theropithecus	13.3	9.0	10.4	0.005***	0.035*	0.602
Papio	13.6	9.2	11.5	0.000***	0.107	0.083
Mandrillus	20.1	11.5	12.9	0.000***	0.001***	0.416
Cercopithecus	11.7	9.2	8.4	0.013*	0.014*	0.844
Colobus	10.0	6.6	6.2	0.001***	0.001***	0.873

* = significant at $p \leq 0.05$.
** = significant at $p \leq 0.01$.
*** = significant at $p \leq 0.005$.
[1] LM = measurements subject to low-to-moderate mastication-related strains.

more compatible with the papionin consensus molecular phylogeny than high-strain characters. None of the analyses suggested relationships that were wholly congruent with the papionin consensus molecular phylogeny (Figs. 2–4). In fact, the morphological datasets performed so poorly in these analyses that relationships of ingroup taxa could not even be resolved monophyletically when rooted with the outgroup taxa, *Colobus* and *Cercopithecus*.

Following parsimony analyses, the relative phylogenetic utility of each trait group was assessed by imposing with morphological topologies and the consensus molecular topology upon the datasets in MacClade. Table 3 describes the number of informative characters, cladogram lengths, consistency indices, and retention indices for all the most parsimonious morphological trees, and those obtained when the molecular topology was imposed on the datasets. The percentage increases in cladogram length when the papionin consensus molecular phylogeny was imposed on each dataset can be regarded as a relative measure of the amount of homoplasy in each dataset, whereby a higher percentage indicates more homoplasy, while a lower percentage indicates less homoplasy. Contrary to the predictions of the homoiology hypothesis, the high-strain characters exhibited a markedly better fit with the molecular phylogeny than the low-to-moderate-strain characters. When the molecular topology was imposed on the high-strain dataset, the length of the cladogram increased by 11%, from 171 to 190. When the molecular topology was imposed on the low-to-moderate-strain dataset, the cladogram length increased by 30%, from 122 to 158. When this procedure was repeated for dental characters, the length of the cladogram increased by just 10%, from 125 to 138. Hence, the low-to-moderate-strain characters exhibited a markedly worse fit with the molecular phylogeny than did the high-strain characters or dental characters. Goodness-of-fit indices (i.e., consistency indices and retention indices) obtained from the morphological and molecular topology comparisons in MacClade display the same pattern of difference between the various datasets as those obtained from tree-length comparisons (Table 3). That is, when the molecular topology is imposed on the morphological datasets, reduction of RI and CI is greatest in the low-to-moderate-strain traits, intermediate in the high-strain traits, and least in the dental traits. This pattern is consistent in all analyses, and reflects a higher level of homoplasy in the low-to-moderate-strain characters compared to the high-strain and dental datasets. Hence, the homoiology hypothesis is not supported by this analysis.

Impact of sexual dimorphism

The third set of analyses examined the possibility that sexual dimorphism has confounded

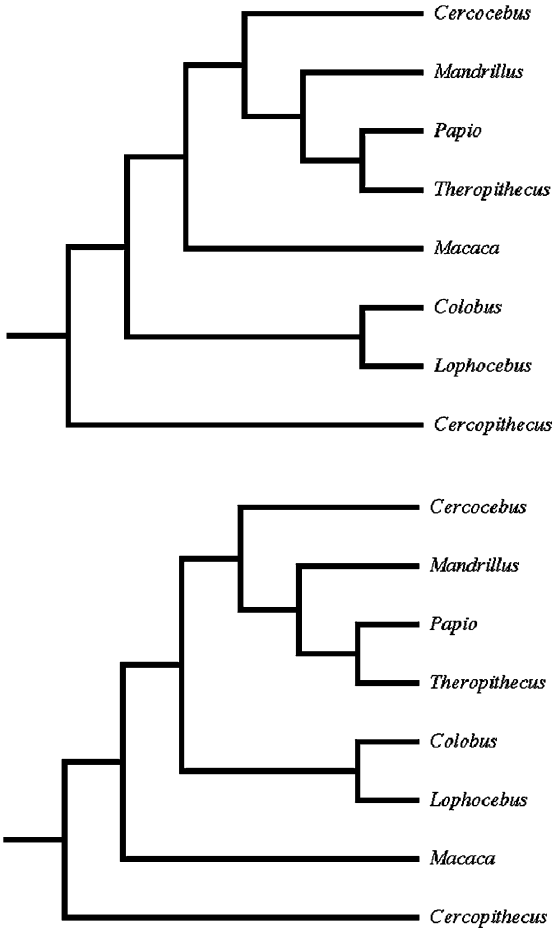

Fig. 2. Most parsimonious cladograms derived from mixed-sex high-strain characters.

attempts to test the homoiology hypothesis. The CV/ANOVA analyses and the cladistic analyses were repeated for males and females separately, and then the results of the sex-specific analyses were compared with the results of the mixed-sex analyses. It was predicted that, if sexual dimorphism is a confounding factor, then results of sex-specific analyses should differ in important respects from the results of the mixed-sex analyses.

Table 4 summarizes the results of the sex-specific CV/ANOVA analyses. In both the male analysis and the female analysis, the CVs of the high-strain measurements were consistently, and in most cases significantly, higher than the CVs

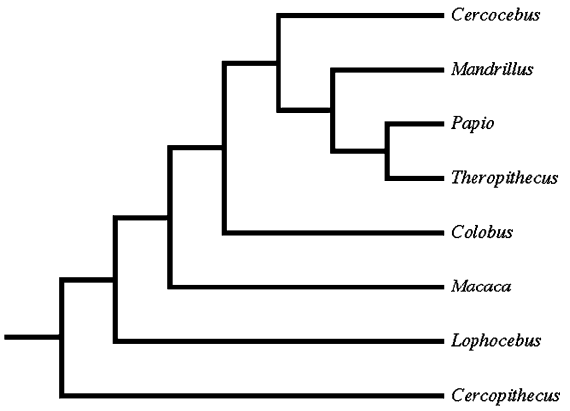

Fig. 3. Most parsimonious cladogram derived from mixed-sex low-to-moderate-strain characters.

associated with the low-to-moderately strained measurements; the CVs of the dental measurements were generally higher than the CVs for the low-to-moderately strained measurements, although the differences were only significant in one taxon; and the high-strain measurements were more variable than the dental measurements, but only significantly so in some taxa. Thus, the sex-specific CV/ANOVA analyses were consistent with their mixed-sex counterparts regarding the predictions of the homoiology hypothesis. They

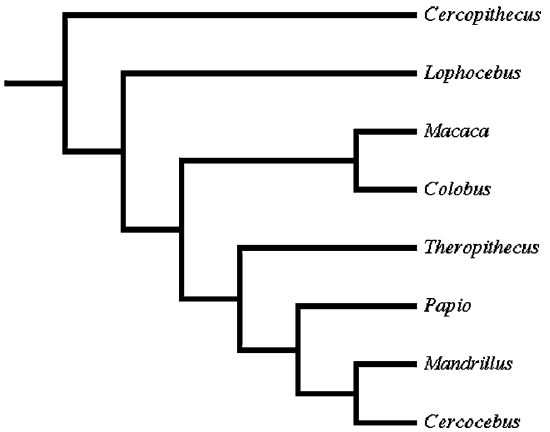

Fig. 4. Most parsimonious cladogram derived from mixed-sex dental characters.

Table 3
Goodness-of-fit statistics associated with the most parsimonious cladograms recovered from the mixed-sex datasets, and those obtained when the consensus molecular topology for the papionins was imposed on the same datasets[1]

Character group	Most parsimonious				Molecular		
	IC	CL	CI	RI	CL	CI	RI
High-strain	22	171	0.60	0.47	190	0.54	0.33
Low-to-moderate-strain	21	122	0.75	0.71	158	0.58	0.38
Dental	16	125	0.58	0.50	138	0.53	0.38

[1] Abbreviations as follows: IC = number of informative characters; CL = cladogram length; CI = consistency index; RI = retention index.

support the prediction that high-strain measurements should be significantly more variable than low-to-moderate-strain measurements, but do not support the prediction that low-to-moderate-strain measurements should be more variable than dental measurements. They also only partially support the prediction that high-strain measurements should be significantly more variable than dental measurements.

The sex-specific cladistic analyses were also consistent with their mixed-sex counterparts (Table 5). When the molecular topology was imposed on the male high-strain characters, cladogram length increased by 12%. Cladogram length also increased by 12% when the molecular topology was imposed on the male dental characters. When the same was done to the male low-to-moderate characters, cladogram length increased by 36%. The results of the analyses of the female dataset were similar. When the molecular topology was imposed on the female dataset, the high-strain and dental characters were found to be markedly less homoplastic than the low-to-moderate-strain characters. The molecular cladograms for the high-strain and dental characters were 8% and 10% longer, respectively, than the most parsimonious cladograms, whereas the molecular cladogram for the low-to-moderate-strain characters was 23%

Table 4
Results of single-sex CV/ANOVA analyses

Taxon	CV			ANOVA		
	High	LM[1]	Dental	High vs. LM	High vs. dental	LM vs. dental
Males						
Macaca	10.3	7.6	8.2	0.006	0.089	0.366
Lophocebus	8.7	5.5	6.9	0.000***	0.090	0.083
Cercocebus	8.3	5.3	5.7	0.001***	0.025*	0.324
Theropithecus	10.2	7.3	8.2	0.016*	0.261	0.258
Papio	11.7	8.3	7.5	0.001***	0.001***	0.740
Mandrillus	12.6	6.8	6.7	0.000***	0.000***	0.876
Cercopithecus	9.0	7.5	5.1	0.024*	0.001***	0.204
Colobus	10.2	5.7	5.1	0.000***	0.000***	0.626
Females						
Macaca	9.6	7.2	7.4	0.017*	0.120	0.506
Lophocebus	7.9	5.5	6.3	0.006**	0.074	0.415
Cercocebus	10.3	7.4	6.9	0.000***	0.002***	0.884
Theropithecus	9.3	6.4	6.6	0.004***	0.039*	0.545
Papio	9.4	5.9	8.1	0.000***	0.161	0.000***
Mandrillus	7.8	5.5	5.4	0.001***	0.008***	0.681
Cercopithecus	8.3	5.2	6.0	0.000***	0.057	0.068
Colobus	8.2	6.2	6.1	0.017*	0.089	0.609

* = significant at $p \leq 0.05$.
** = significant at $p \leq 0.01$.
*** = significant at $p \leq 0.005$.
[1] LM = measurements subject to low-to-moderate mastication-related strains.

Table 5
Goodness-of-fit statistics associated with the most parsimonious cladograms recovered from the single-sex datasets, and those obtained when the consensus molecular topology for the papionins was imposed on the same datasets[1]

Character group	Most parsimonious				Molecular		
	IC	CL	CI	RI	CL	CI	RI
Males							
High-strain	22	141	0.60	0.27	158	0.53	0.30
Low-to-moderate-strain	21	122	0.74	0.74	166	0.54	0.38
Dental	16	107	0.55	0.42	120	0.55	0.27
Females							
High-strain	22	194	0.60	0.43	209	0.56	0.33
Low-to-moderate-strain	21	158	0.68	0.62	194	0.55	0.36
Dental	16	115	0.63	0.55	126	0.58	0.43

[1] Abbreviations as follows: IC = number of informative characters; CL = cladogram length; CI = consistency index; RI = retention index.

longer than the most parsimonious cladogram. Thus, the sex-specific cladistic analyses, like the mixed-sex cladistic analyses, contradict the homoiology hypothesis. They do not support the prediction that the dental characters will yield phylogenies that are more compatible with the papionin consensus molecular phylogeny than the low-to-moderately strained and highly strained characters. They also do not support the prediction that the low-to-moderately strained characters will yield phylogenies that are more compatible with the molecular phylogeny than the highly strained characters. In sum, there is no evidence that sexual dimorphism has confounded tests of the homoiology hypothesis.

Impact of "allometric homoplasy"

The last set of analyses investigated the possibility that the results of the aforementioned phylogenetic analyses may have been confounded by allometry. We carried out a series of Pearson correlation analyses in which the statistical association between each putatively size-corrected character and the relevant geometric mean was measured, and then re-ran the cladistic analyses after excluding the characters that the correlation analyses suggested were significantly correlated with the geometric means. According to the correlation analyses, six of the 22 high-strain characters were significantly correlated with the geometric mean of the high-strain characters, 12 of the 22 low-to-moderate-strain characters were significantly correlated with the geometric mean of the low-to-moderate-strain characters, and five of the 16 dental characters were significantly correlated with the geometric mean of the dental characters. The most parsimonious cladograms produced for the three character groups following removal of the significantly correlated characters

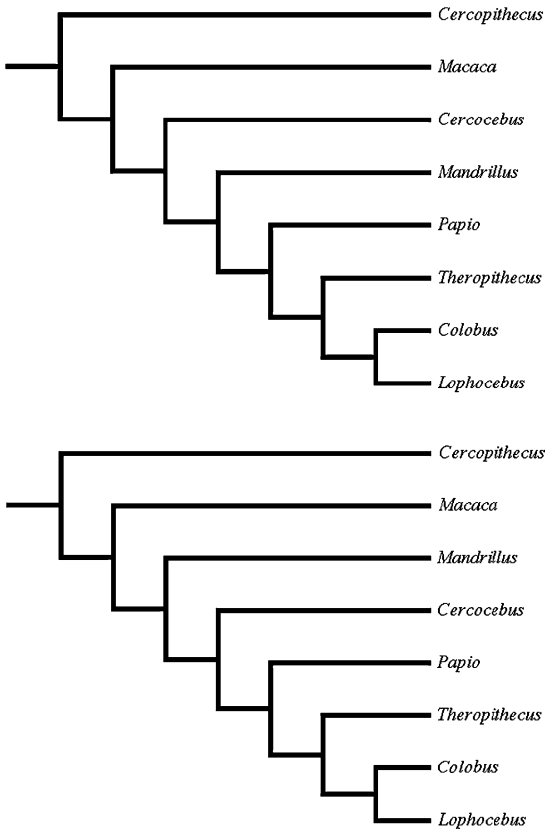

Fig. 5. Most parsimonious cladograms derived from mixed-sex high-strain characters after removal of characters that remain significantly correlated with the geometric mean after size-correction.

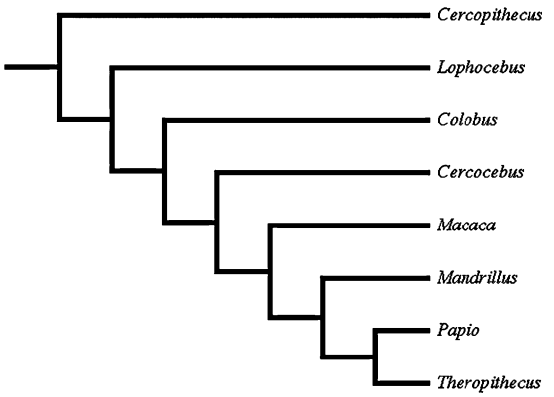

Fig. 6. Most parsimonious cladogram derived from mixed-sex low-to-moderate-strain characters after removal of characters that remain significantly correlated with the geometric mean after size-correction.

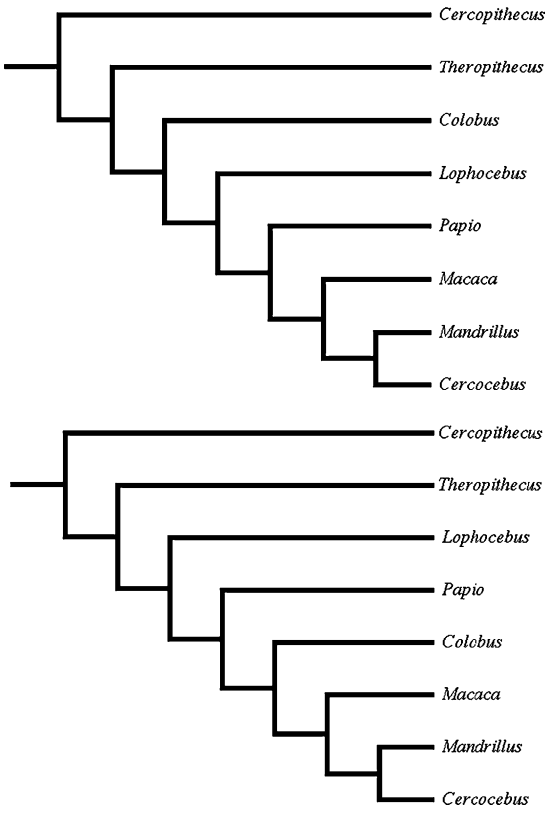

Fig. 7. Most parsimonious cladograms derived from mixed-sex dental characters that remain significantly correlated with the geometric mean after size-correction.

are shown in Figs. 5–7. Table 6 shows the goodness-of-fit statistics associated with these cladograms, as well as the goodness-of-fit statistics obtained when the topology of the papionin consensus molecular phylogeny was fitted to each character group. The percentage increases in cladogram length required to change the most parsimonious topologies into the molecular topology indicate that, after removal of potential allometric homoplasies, fitting the molecular topology to the high-strain characters involves a 13% increase in cladogram length relative to the most parsimonious cladogram; fitting the molecular topology to the dental characters involves a 14% increase in cladogram length relative to the most parsimonious cladogram; and fitting the molecular topology to the low-to-moderate-strain characters involves an 18% increase in cladogram length relative to the most parsimonious cladogram. These figures indicate that residual allometry likely accounts for some of the homoplasies in the papionin dataset. They also suggest that allometric homoplasies are especially prevalent in the low-to-moderate-strain character set. However, the percentage increases in cladogram length do not support the notion that our results are solely an artifact of the size-correction method we employed. Even with potential allometric homoplasies removed, the dataset does not support the homoiology hypothesis. Contrary to the predictions of the homoiology hypothesis, the high-strain characters do not differ from the dental characters in terms of homoplasy. Also contrary to the homoiology hypothesis predictions, the low-to-moderate-strain characters are considerably more homoplastic than the high-strain characters. Thus, it seems unlikely that allometric homoplasy explains the failure of our phylogenetic analyses to support the homoiology hypothesis.

Discussion

The results of the CV/ANOVA analyses were mixed. As predicted, the high-strain measurements exhibited significantly more variation than the

Table 6
Goodness-of-fit statistics associated with the most parsimonious cladograms recovered from the mixed-sex datasets, and those obtained when the consensus molecular topology for the papionins was imposed on the same datasets, after removal of characters that remain significantly correlated with the geometric mean after size-correction[1]

Character group	Most parsimonious				Molecular		
	IC	CL	CI	RI	CL	CI	RI
High-strain	16	124	0.58	0.43	140	0.51	0.26
Low-to-moderate-strain	9	61	0.69	0.60	72	0.58	0.38
Dental	11	86	0.60	0.51	98	0.53	0.33

[1] Abbreviations as follows: IC = number of informative characters; CL = cladogram length; CI = consistency index; RI = retention index.

low-to-moderate-strain measurements in all the taxa. However, the results of the CV/ANOVA analyses do not support the prediction that low-to-moderately strained characters should be more variable than characters that do not remodel in response to mechanical loading. The CVs for the low-to-moderately strained measurements and the dental measurements were statistically indistinguishable. Furthermore, the analyses do not unambiguously support the prediction that the high-strain characters should be more variable than characters that do not remodel in response to strain. The high-strain measurements were consistently more variable than the dental measurements, but this difference was not significant in all taxa. Thus, the CV/ANOVA analyses provide only partial support for the homoiology hypothesis. When these results are combined with the results of the comparable analysis carried out by Collard et al. (in press), it seems reasonable to conclude that the hypothesis that strain leads to intraspecific phenotypic variation in primates is overly simplistic. Together, the analyses suggest that, at least as far as mastication-related strain is concerned, strain can be a cause of significant variation in the primate skull, but strain-induced variation will not always exceed variation caused by other environmental factors (e.g., activity-stimulated circulation of hormones) and/or by genetic factors (e.g., epistasis).

It is worth noting in this regard that recent work by Daegling (2004) also suggests that the hypothesis that high levels of strain result in increased variation in primate skull bones is too simple. Daegling tested the hypothesis with a sample of macaque mandibles. He found no evidence that regions of the mandible that experience high strain during mastication are more variable than mandibular regions that experience lower levels of mastication-related strain. He concluded that the hypothesis is context-specific, holding at certain levels of analysis but not at others.

The three sets of cladistic analyses presented here are unequivocal in their lack of support for the idea that homoiology is a major form of homoplasy in primates. The prediction that the agreement between the morphological and molecular phylogenies would be best in the analyses of the dental characters, intermediate in the analyses of the low-to-moderate-strain characters, and least in the analyses of the high-strain characters was not fulfilled in the primary analysis. It also was not fulfilled in the analyses designed to control for the effects of sexual dimorphism and allometry. In all three analyses, the dental characters *and* the high strain characters were found to have a markedly better fit with the molecular phylogeny than the low-to-moderate-strain characters. Thus, the results of our cladistic analyses are consistent with those of Collard et al.'s (in press) in failing to support the homoiology hypothesis.

In light of the failure of the phylogenetic analyses reported here and those presented by Collard et al. (in press) to support the homoiology hypothesis, it seems reasonable to conclude that the hypothesis is incorrect, at least as currently formulated. With regard to revising it, one possibility that might be worth considering is that homoiologous resemblances are primarily a problem in intraspecific phylogenetic analyses and do not affect interspecific analyses to any great extent. That is, phenotypic plasticity may be a major source of homoplasy, but only in analyses of the relationships among subspecific taxa. When analysing the relationships among superspecific taxa, the situation may be more complicated because of morphological integration. As has been widely noted, few features of the skull are likely to be

totally independent. Rather, such features are integrated at numerous hierarchical levels of development (Olsen and Miller, 1958; Cheverud, 1982; Lieberman, 1999; Lovejoy et al., 1999, 2000; McCollum, 1999; Lieberman et al., 2000; McCollum and Sharpe, 2001; Strait, 2001). Thus, while the mechanisms by which bone tissue responds to strain may be conservative across species, the morphological effects of such responses may differ markedly depending on a wide variety of other developmental and structural factors. Given this possibility, it is perhaps unrealistic to expect a simple correspondence between the phenotypic plasticity of characters and their phylogenetic valence in interspecific studies. The situation may be further complicated by the fact that, in this and other measurement-based interspecific cladistic analyses, character states are based on species means. Given that the proportion of a given measurement that can be attributed to phenotypic plasticity can be expected to vary randomly among the members of a species, the use of species means to generate character states is likely to greatly reduce the impact of phenotypic plasticity on species-level phylogenetic analyses. Testing the possibility that phenotypic plasticity is a major source of homoplasy in intraspecific phylogenetic analyses is particularly important given that phylogeny-oriented analyses of population-level samples and individual specimens are becoming increasingly popular in hominid palaeontology (e.g., Brauer and Rimbach, 1990; Caparros, 1997; Hawks et al., 2000; Brace et al., 2001; Kramer et al., 2001; Wolpoff et al., 2001; Cameron et al., 2004).

The other issue that seems worth discussing here is the relationship between the results of our phylogenetic analyses and previous morphology-based work on the phylogenetic relationships of the papionins. Two taxa are of particular concern here. One is *Colobus*; the other is *Macaca*. To reiterate, in the most parsimonious cladograms derived from the high-strain characters, *Colobus* was positioned as the sister group of *Lophocebus*, while *Macaca* was positioned as either the sister group of a clade comprising *Cercocebus*, *Mandrillus*, *Papio*, and *Theropithecus*, or the sister group of a clade comprising the papionins plus *Colobus* (Fig. 2). In the most parsimonious cladogram yielded by the low-to-moderate-strain characters, *Colobus* was suggested to be the sister group of a clade comprising *Cercocebus*, *Mandrillus*, *Papio*, and *Theropithecus*, while *Macaca* was positioned as the sister group of a clade consisting of *Cercocebus*, *Colobus*, *Mandrillus*, *Papio*, and *Theropithecus* (Fig. 3). In the most parsimonious cladogram derived from the dental characters, *Colobus* was positioned as the sister group of *Macaca*, and this clade was suggested to be the sister group of clade comprising *Cercocebus*, *Mandrillus*, *Papio*, and *Theropithecus* (Fig. 4).

At first glance, these results are striking. This is because the general understanding is that, like the molecular data, the morphological data support a monophyletic papionin group and suggest that macaques are the sister group to the African papionins. The difference between the molecular and morphological data, according to this view, concerns the relationships within the African papionin clade, with the molecular data supporting a *Cercocebus–Mandrillus* clade and a *Lophocebus–Papio–Theropithecus* clade, and the morphological data supporting a *Cercocebus–Lophocebus* clade and a *Mandrillus–Papio–Theropithecus* clade. However, the disagreement between our results and those of previous morphology-based studies of papionin phylogeny is actually more apparent than real. A reasonably thorough literature search covering the last 20 years identified only 13 morphology-based studies that deal with papionin phylogeny (Strasser and Delson, 1987; Strasser, 1988; Cheverud, 1989; Delson, 1993; Delson and Dean, 1993; Jablonski, 1993; Fleagle and McGraw, 1999, 2002; Groves, 2000; Collard and Wood, 2000; Collard and O'Higgins, 2001; Singleton, 2002; Leigh et al., 2003). Seven of these studies involved interpreting morphological data in the light of a consensus phylogeny for the papionins rather than reconstructing their relationships from morphological data (Strasser, 1988; Cheverud, 1989; Fleagle and McGraw, 1999, 2002; Collard and O'Higgins, 2001; Singleton, 2002; Leigh et al., 2003). Three of the other studies focused on the relationships of species assigned to *Papio* and *Theropithecus*, and did not examine the relationships of either the

macaques or the mangabeys (Delson, 1993; Delson and Dean, 1993; Jablonski, 1993). Thus, there have been just three studies published since 1985 in which the phylogenetic relationships of the six currently recognized extant papionin genera were reconstructed from morphological data: Strasser and Delson (1987), Collard and Wood (2000), and Groves (2000).

Significantly, the studies of Strasser and Delson (1987), Collard and Wood (2000), and Groves (2000) do not consistently support papionin monophyly or a sister group relationship between *Macaca* and the African papionins. Strasser and Delson (1987) used 37 cranial and postcranial characters to reconstruct the relationships of 20 cercopithecid taxa, including the six extant papionin genera. Their preferred phylogeny suggested that the papionins are monophyletic, that *Macaca* is the sister group of the African papionins, and that *Theropithecus* is the sister group of a *Cercocebus–Lophocebus–Mandrillus–Papio* clade. It also suggested that *Cercocebus* and *Lophocebus* form a clade, while *Mandrillus* and *Papio* form a second. Collard and Wood (2000) reconstructed the relationships of the six extant papionin genera from 62 cranial and dental characters. The most parsimonious cladogram suggested that *Lophocebus* is the sister group of *Cercocebus*, *Macaca*, *Mandrillus*, *Papio*, and *Theropithecus*, and that *Cercocebus* is the sister group of *Macaca*, *Mandrillus*, *Papio*, and *Theropithecus*. It also suggested that *Macaca* is the sister group of *Mandrillus*, *Papio*, and *Theropithecus*, and that *Papio* and *Mandrillus* form a clade to the exclusion of *Theropithecus*. Groves (2000) used 46 cranial, postcranial, and soft-tissue characters to reconstruct the relationships among 16 Old World monkey genera, including the six extant papionins. His analyses returned 56 equally parsimonious cladograms. The strict consensus of these contained only one clade, which linked together *Cercocebus* and *Mandrillus*.

Thus, while Strasser and Delson's (1987) study supports papionin monophyly and a sister group relationship between *Macaca* and the African papionins, the other two studies do not. Moreover, even the phylogeny favored by Strasser and Delson (1987) is at odds with the general understanding of papionin phylogeny as derived from morphological data in that it suggests *Theropithecus* is the sister group of *Cercocebus*, *Lophocebus*, *Mandrillus*, and *Papio*. As such, the parsimony analyses reported in this study are not particularly remarkable with regard to the sister-group relationships they reconstruct among the papionin genera. It is worth noting that the situation does not alter if we consider some older, classic publications. For example, in their widely cited volume on primate systematics, Szalay and Delson (1979) identified three subtribes within the papionins—the macaques, baboons and mangabeys, and geladas—and suggested that their relationships are best viewed as a trichotomy. Likewise, in his well-known monographic treatment of primate anatomy, Hill (1974) endorsed Jolly's (1966) division of the papionins into a macaque and mangabey tribe, a savannah baboon and mandrill tribe, and a gelada baboon tribe. Perhaps the most important implication of the variability in the results of the studies discussed in the last three paragraphs is that there is considerably more morphological homoplasy among papionins and other Old World monkeys than is usually recognized, especially in the skull and dentition.

Conclusions

The study reported here was stimulated by Collard et al.'s (in press) study, in which craniodental data from the extant hominoids were used to test the homoiology hypothesis. Their analyses supported the idea that mastication-related strain results in greater phenotypic plasticity in cranial characters, but did not support the notion that cranial characters that are more phenotypically plastic are more likely to be homoplastic. The primary goal of the present study was to further evaluate the homoiology hypothesis by replicating Collard et al.'s analyses with cranial data derived from another group of extant primates, the papionins. We also wanted to address some factors that may have confounded Collard et al.'s attempt to test the homoiology hypothesis.

In order to evaluate the homoiology hypothesis, we compiled a craniodental dataset for the extant papionin primates using three series of measurements that differ in their susceptibility to remodeling according to experimental evidence and developmental considerations. We then carried out three analyses. The first sought to determine whether measurements that are subject to high levels of masticatory strain are significantly more variable (i.e., more phenotypically plastic) than measurements that are subject to low-to-moderate strains, and whether low-to-moderate-strain measurements are significantly more variable than dental measurements, which are not phenotypically plastic. The second analysis investigated whether non-remodeling, low-to-moderate-strain, and high-strain characters differ in their ability to recover the phylogenetic relationships of the extant papionins. The third analysis examined the possibility that sexual dimorphism has confounded attempts to test the homoiology hypothesis.

The results of this study partially support the notion that mechanical loading can result in significant phenotypic variation in the highly strained osseous characters of the primate cranium. As predicted, the characters subject to high levels of strain were found to be significantly more variable than the characters subject to low-to-moderate levels of strain. However, contrary to expectation, the low-to-moderate-strain characters were not more variable than the non-remodeling dental characters. Moreover, in several taxa, the high-strain characters were not consistently more variable than the dental characters. When these results are combined with the results of the comparable analysis carried out by Collard et al. (in press) it seems reasonable to conclude that the first part of the homoiology hypothesis—the notion that strain leads to exaggerated intraspecific phenotypic variability—needs to be amended. Our combined analyses show that, at least as far as mastication-related strain is concerned, strain can be a cause of significant variation in primate skull bones, but strain-induced variation will not always exceed variation caused by other factors.

The results of the study presented here do not support the second part of the homoiology hypothesis—the idea that phenotypic plasticity is major cause of homoplasy in the primate skull. The prediction that the agreement between the morphological and molecular phylogenies would be best in the analyses of dental characters, intermediate in the analyses of low-to-moderate-strain characters, and least in the analyses of high-strain characters was not fulfilled. Both the dental characters *and* the high-strain characters were found to have a markedly better fit with the molecular phylogeny than the low-to-moderate-strain characters. Thus, the results of our cladistic analyses are consistent with those of Collard et al.'s (in press) in failing to support the homoiology hypothesis.

In sum, the study presented here and that of Collard et al. (in press) cast serious doubt on the validity of the homoiology hypothesis, as currently formulated, and re-emphasize that fossil hominid homoplasy is likely to be a complex and multifaceted phenomenon.

Acknowledgments

Much of the research reported in this paper was carried out while SJL was a postgraduate student and MC was a faculty member in the Department of Anthropology, University College London. SJL's work on the project was funded by the Graduate School and Department of Anthropology, University College London, and by Trinity College, University of Cambridge. We are indebted to Leslie Aiello, Clare Holden, Daniel Lieberman, David Pilbeam, Daniele Serdoz, Noreen von Cramon-Taubadel, and Bernard Wood for their assistance with various aspects of the project. We are also grateful to Bill Kimbel, Charles Lockwood, Jeremiah Scott, and three anonymous referees for providing helpful comments and suggestions. Lastly, we thank the following institutions for providing access to primate skeletal material: the Natural History Museum, London, the Powell-Cotton Museum, Birchington-on-Sea, and the Anthropologisches Institut und Museum, Universität Zürich-Irchel, Zürich.

Appendix 1. Papionin character state data matrices for mixed-sex dataset (for details of characters, see Table 1)

High-strain traits	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Cercocebus	2	2	1	3	3	5	0	1	2	0	2	3	2	2	2	1	3	1	3	3	3	1
Cercopithecus	1	0	1	0	2	0	2	3	0	3	4	0	1	0	3	2	0	0	0	5	5	0
Colobus	2	1	3	0	5	6	3	0	4	5	3	4	0	3	3	0	1	4	4	5	2	3
Lophocebus	2	1	0	1	4	6	1	2	1	4	3	0	3	1	0	4	0	2	5	4	4	2
Macaca	0	1	2	2	0	1	0	2	2	1	3	2	3	3	1	1	2	0	2	5	2	0
Mandrillus	2	2	5	4	1	2	0	2	3	2	1	1	4	1	0	3	5	0	2	2	1	1
Papio	2	1	4	4	3	3	0	3	4	0	0	5	6	4	4	6	1	0	1	1	0	1
Theropithecus	3	2	5	4	6	4	1	2	5	1	0	4	5	5	5	5	4	3	6	0	1	4
Low-to-moderate-strain traits	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Cercocebus	3	2	2	2	2	1	3	2	1	1	1	2	1	2	2	0	2	3	3	1	1	3
Cercopithecus	6	4	3	4	2	2	6	5	3	3	2	5	1	4	3	1	4	6	6	5	3	4
Colobus	4	3	5	4	0	1	4	3	1	1	0	5	0	2	1	0	3	3	4	3	1	0
Lophocebus	5	2	1	3	1	1	5	4	2	2	1	4	1	3	3	1	4	4	5	3	4	3
Macaca	4	4	0	3	4	0	5	4	2	2	1	3	2	3	3	1	3	5	4	0	2	3
Mandrillus	2	2	1	1	5	1	1	1	0	0	1	0	3	2	1	0	1	1	2	1	1	2
Papio	1	1	3	1	6	2	2	0	0	0	1	1	4	0	0	0	0	2	0	4	1	1
Theropithecus	0	0	4	0	3	2	0	3	0	0	1	2	5	1	1	0	1	0	1	2	0	2
Dental traits	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60						
Cercocebus	4	5	2	3	5	5	3	2	3	2	1	6	4	5	4	2						
Cercopithecus	0	0	1	0	0	0	0	0	0	0	0	0	1	4	1	1						
Colobus	2	2	3	6	2	1	1	1	0	3	2	4	4	1	0	1						
Lophocebus	5	4	0	1	1	2	0	0	3	2	0	1	0	0	0	0						
Macaca	6	1	3	4	2	1	0	0	3	2	3	5	3	3	2	1						
Mandrillus	4	5	3	5	4	4	3	4	1	3	4	6	4	2	3	3						
Papio	3	3	4	2	3	4	3	4	2	4	5	2	2	3	2	4						
Theropithecus	1	0	2	0	3	3	2	3	0	1	3	3	2	1	2	4						

Appendix 2. Papionin character state data matrices for females (for details of characters see Table 1).

High-strain traits	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Cercocebus	3	5	1	3	2	4	1	2	1	1	3	2	1	2	3	1	5	4	2	2	4	2
Cercopithecus	1	0	1	0	2	0	5	5	0	1	6	0	1	0	5	2	1	1	0	5	6	0
Colobus	4	2	2	1	4	6	6	0	3	3	4	5	0	1	4	0	3	7	1	6	1	5
Lophocebus	4	1	0	2	3	5	4	3	1	2	5	0	4	1	2	3	0	6	4	3	5	4
Macaca	0	3	1	2	0	1	0	2	1	1	5	3	2	2	1	1	2	2	3	4	3	0
Mandrillus	2	4	3	4	1	2	2	1	2	1	2	1	2	4	0	2	7	0	1	2	1	1
Papio	4	2	1	5	3	3	3	4	3	0	1	6	5	3	5	4	4	3	2	1	0	3
Theropithecus	5	3	3	4	5	7	2	2	3	1	0	4	3	5	6	4	6	5	5	0	2	6
Low-to-moderate-strain traits	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Cercocebus	2	4	2	2	2	3	3	1	1	1	4	2	2	2	2	1	2	3	1	2	3	3
Cercopithecus	6	6	3	4	2	6	6	3	4	4	5	5	1	5	3	3	7	5	4	5	5	4
Colobus	3	5	5	4	0	4	4	1	1	0	0	6	0	0	0	0	4	2	1	3	1	0
Lophocebus	5	3	3	3	1	1	5	2	2	2	3	4	2	3	3	3	6	4	3	2	5	2
Macaca	4	6	0	3	4	0	5	2	3	3	1	3	2	4	3	3	5	5	2	0	5	2

Appendix 2 (*continued*)

Low-to-moderate-strain traits	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Mandrillus	2	2	1	1	5	2	1	1	1	0	4	0	3	2	1	2	3	2	1	1	4	1
Papio	1	1	3	0	6	7	0	0	0	0	2	1	4	1	0	1	0	0	0	4	2	1
Theropithecus	0	0	4	0	3	5	2	1	0	0	1	3	4	0	0	1	1	1	0	3	0	1

Dental traits	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Cercocebus	2	6	2	5	5	5	4	2	1	1	3	4	4	4	5	3
Cercopithecus	0	1	3	1	0	0	1	0	0	0	0	0	1	2	2	1
Colobus	0	3	6	6	2	0	2	1	0	1	4	2	4	1	1	2
Lophocebus	2	5	2	3	1	0	0	0	1	1	0	1	0	0	0	0
Macaca	3	2	4	4	2	1	1	0	1	1	4	3	3	2	3	1
Mandrillus	2	4	1	5	4	4	5	5	1	1	5	5	5	3	6	4
Papio	1	3	5	2	3	3	4	4	1	1	2	1	2	2	3	4
Theropithecus	0	0	0	0	2	2	3	3	0	0	1	1	2	0	4	5

Appendix 3. Papionin character state data matrices for males (for details of characters see Table 1)

High-strain traits	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Cercocebus	2	4	0	1	3	3	0	1	2	0	2	2	2	5	2	1	3	1	1	2	3	4
Cercopithecus	1	0	0	0	2	1	2	4	0	3	4	0	0	0	4	2	0	0	0	5	5	1
Colobus	2	1	0	0	5	5	2	0	2	3	3	1	0	2	5	0	2	3	1	4	3	7
Lophocebus	2	2	0	0	3	4	0	2	1	3	3	0	1	1	0	3	0	3	2	3	4	5
Macaca	0	2	0	1	0	1	0	5	2	1	3	1	3	4	3	2	2	0	0	5	2	0
Mandrill	2	5	3	2	1	2	0	5	2	2	1	0	3	6	1	4	3	0	1	1	2	3
Papio	2	3	1	2	2	2	0	4	3	1	1	2	3	3	5	5	2	0	0	1	1	2
Theropithecus	2	2	2	2	4	0	1	3	2	0	1	1	2	3	3	4	1	2	2	0	0	6

Low-to-moderate-strain traits	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
Cercocebus	3	3	4	2	2	1	2	2	1	1	1	2	1	1	2	0	0	1	0	1	3	2
Cercopithecus	4	6	4	5	2	4	5	4	5	4	2	5	2	4	5	1	4	2	4	3	5	4
Colobus	3	4	6	5	0	2	3	2	2	2	0	5	0	3	2	0	2	1	2	1	4	1
Lophocebus	4	4	1	4	2	2	4	3	4	4	1	4	1	4	4	2	3	2	3	2	6	2
Macaca	3	5	0	3	3	0	3	3	3	3	1	3	2	4	3	1	1	2	1	0	5	3
Mandrill	2	2	3	1	4	2	0	0	0	0	1	0	3	2	0	0	0	0	0	2	2	1
Papio	1	1	5	1	5	4	0	0	0	0	1	1	3	0	0	0	0	0	0	2	1	0
Theropithecus	0	0	2	0	1	3	1	1	0	0	1	1	2	0	1	0	0	0	0	0	0	0

Dental traits	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Cercocebus	2	3	2	1	4	5	1	2	3	2	2	5	3	5	2	3
Cercopithecus	0	0	1	0	0	0	0	0	0	0	1	0	0	4	1	1
Colobus	0	1	1	4	3	3	0	2	0	3	2	4	3	1	0	1
Lophocebus	4	3	0	0	3	4	0	0	3	1	0	1	0	2	0	0
Macaca	4	2	3	2	3	1	0	0	3	2	3	4	2	3	1	1
Mandrill	1	3	4	3	4	3	1	3	2	4	4	3	3	0	1	3
Papio	3	3	3	1	2	4	1	3	3	5	4	2	0	3	1	4
Theropithecus	0	0	2	0	1	2	0	1	1	3	4	3	1	0	0	2

References

Aiello, L.C., 1992. Allometry and the analysis of size and shape in human evolution. J. Hum. Evol. 22, 127–147.

Almeida, M.T., Bisby, F.A., 1984. A simple method for establishing taxonomic characters from measurement data. Taxon 33, 405–409.

Archie, J.W., 1985. Methods for coding variable morphological features for numerical taxonomic analysis. Syst. Zool. 34, 326–345.

Atchley, W.R., Fitch, W.M., 1991. Gene trees and the origins of inbred strains of mice. Science 254, 554–558.

Baum, B.R., 1988. A simple procedure for establishing discrete characters from measurement data, applicable to cladistics. Taxon 37, 63–70.

Brace, C.L., Nelson, A.R., Seguchi, N., Oe, H., Sering, L., Qifeng, P., Yongyii, L., Dashtseveg Tumen, D., 2001. Old World sources of the first New World human inhabitants: a comparative craniofacial view. Proc. Natl. Acad. Sci. 98, 10017–10022.

Brauer, G., Rimbach, K.W., 1990. Late archaic and modern *Homo sapiens* from Europe, Africa and southwest Asia: craniometric comparisons and phylogenetic implications. J. Hum. Evol. 19, 789–807.

Brehnan, K., Boyd, R.L., Laskin, J.L., Gibbs, C.H., Mahan, P.E., 1981. Direct measurement of loads at the temporomandibular joint in *Macaca arctoides*. J. Dent. Res. 60, 1820–1824.

Caparros, M., 1997. *Homo sapiens* archaïques: un ou plusieurs taxons (espèces)? Analyse cladistique et analyse morphométrique. Ph.D. Dissertation, Museum National D'Histoire Naturelle, France.

Cain, A.J., 1982. On homology and convergence. In: Joysey, K.A., Friday, A.E. (Eds.), Problems in Phylogenetic Reconstruction. Academic Press, London, pp. 1–19.

Cameron, D., Patnaik, R., Sahni, A., 2004. The phylogenetic significance of the middle Pleistocene Narmada hominin cranium from central India. Int. J. Osteoarcheol. 14, 419–447.

Chamberlain, A.T., Wood, B.A., 1987. Early hominid phylogeny. J. Hum. Evol. 16, 119–133.

Cheverud, J.M., 1982. Phenotypic, genetic and environmental morphological integration in the cranium. Evolution 36, 499–516.

Cheverud, J.M., 1989. A comparative analysis of morphological variation patterns in the papionins. Evolution 43, 1737–1747.

Collard, M., O'Higgins, P., 2001. Ontogeny and homoplasy in the papionin monkey face. Evol. Dev. 3, 322–331.

Collard, M., Lieberman, D.E., Wood, B.A. Hominid homoiology: an assessment of the impact of phenotypic plasticity on phylogenetic analyses of humans and their fossil relatives. J. Hum. Evol., in press.

Collard, M., Wood, B.A., 2000. How reliable are human phylogenetic hypotheses? Proc. Natl. Acad. Sci. 97, 5003–5006.

Collard, M., Wood, B.A., 2001. Homoplasy and the early hominid masticatory system: inferences from analyses of extant hominoids and papionins. J. Hum. Evol. 41, 167–194.

Corruccini, R.S., 1994. How certain are hominoid phylogenies: the role of confidence intervals in cladistics. In: Corruccini, R.S., Ciochon, R.L. (Eds.), Integrative Paths to the Past. Prentice Hall, Englewood Cliffs, pp. 167–183.

Curnoe, D., 2003. Problems with the use of cladistic analysis in palaeoanthropology. Homo 53, 225–234.

Currey, J.D., 1984. The Mechanical Adaptation of Bones. Princeton University Press, Princeton.

Daegling, D.J., 1993. The relationship of in vivo bone strain to mandibular corpus morphology in *Macaca fascicularis*. J. Hum. Evol. 25, 247–269.

Daegling, D.J., 2004. Relationship of strain magnitude to morphological variation in the primate skull. Am. J. Phys. Anthropol. 124, 346–352.

Daegling, D.J., Hylander, W.L., 1997. Occlusal forces and mandibular bone strain: is the primate jaw "overdesigned"? J. Hum. Evol. 33, 705–717.

Daegling, D.J., Hylander, W.L., 1998. Biomechanics of torsion in the human mandible. Am. J. Phys. Anthropol. 105, 73–87.

Daegling, D.J., Hylander, W.L., 2000. Experimental observation, theoretical models, and biomechanical inference in the study of mandibular form. Am. J. Phys. Anthropol. 112, 541–551.

Dechow, P.C., Hylander, W.L., 2000. Elastic properties and masticatory bone stress in the macaque mandible. Am. J. Phys. Anthropol. 112, 553–574.

Delson, E., 1993. *Theropithecus* specimens from Africa and India and the taxonomy of the genus. In: Jablonski, N.G. (Ed.), *Theropithecus*: The Rise and Fall of a Primate Genus. Cambridge University Press, Cambridge, pp. 157–189.

Delson, E., Dean, D., 1993. Are *Papio baringensis* R. Leakey, 1969 and *P. quadratirostris* Iwamoto, 1982 species of *Papio* or *Theropithecus?* In: Jablonski, N.G. (Ed.), *Theropithecus*: The Rise and Fall of a Primate Genus Cambridge University Press, Cambridge, pp. 125–156.

Demes, B., 1984. Mechanical stresses at the primate skull base caused by the temporomandibular joint force. In: Chivers, D.J., Wood, B.A., Bilsborough, A. (Eds.), Food Acquisition and Processing in Primates. Plenum, London, pp. 407–413.

Disotell, T.R., 1994. Generic level relationships of the Papionini (Cercopithecoidea). Am. J. Phys. Anthropol. 94, 47–57.

Disotell, T.R., 1996. The phylogeny of Old World monkeys. Evol. Anthropol. 5, 18–24.

Disotell, T.R., 2000. Molecular systematics of the Cercopithecidae. In: Whitehed, P., Jolly, C.J. (Eds.), Old World Monkeys. Cambridge University Press, Cambridge, pp. 28–56.

Disotell, T.R., Honeycutt, R.L., Ruvolo, M., 1992. Mitochondrial DNA phylogeny of the Old World monkey tribe Papionini. Mol. Biol. Evol. 9, 1–13.

Dytham, C., 2003. Choosing and Using Statistics: A Biologist's Guide. Blackwell, Oxford.

Eldredge, N., Tattersall, I., 1975. Evolutionary models, phylogenetic reconstruction and another look at hominid

phylogeny. In: Szalay, F.S. (Ed.), Contributions to Primatology 5: Approaches to Primate Paleobiology. Karger, Basel, pp. 218–242.

Falsetti, A.B., Jungers, W.L., Cole III, T.M., 1993. Morphometrics of the callitrichid forelimb: a case study in size and shape. Int. J. Primatol. 14, 551–572.

Fitch, W.M., Atchley, W.R., 1987. Divergence in inbred strains of mice: a comparison of three different types of data. In: Patterson, C. (Ed.), Molecules and Morphology in Evolution: Conflict or Compromise? Cambridge University Press, Cambridge, pp. 203–216.

Fleagle, J.G., McGraw, W.S., 1999. Skeletal and dental morphology supports diphyletic origin of baboons and mandrills. Proc. Natl. Acad. Sci. 96, 1157–1161.

Fleagle, J.G., McGraw, W.S., 2002. Skeletal and dental morphology of African papionins: unmasking a cryptic clade. J. Hum. Evol. 42, 267–292.

Frost, H.M., 1986. Intermediary Organization of the Skeleton. CRC Press, Boca Raton.

Frost, H.M., 1998. From Wolff's law to the mechanostat: a new "face" of physiology. J. Orthop. Sci. 3, 282–286.

Gibbs, S., Collard, M., Wood, B.A., 2000. Soft-tissue characters in higher primate phylogenetics. Proc. Natl. Acad. Sci. 97, 11130–11132.

Gibbs, S., Collard, M., Wood, B.A., 2002. Soft tissue anatomy of the extant hominoids: a review and phylogenetic analysis. J. Anat. 200, 3–49.

Groves, C.P., 2000. The phylogeny of the Cercopithecoidea. In: Whitehed, P., Jolly, C.J. (Eds.), Old World Monkeys. Cambridge University Press, Cambridge, pp. 77–98.

Harris, E.E., 2000. Molecular systematics of the Old World monkey tribe Papionini: analysis of the total available genetic sequences. J. Hum. Evol. 38, 235–256.

Harris, E.E., Disotell, T.R., 1998. Nuclear gene trees and the phylogenetic relationships of the mangabeys (Primates: Papionini). Mol. Biol. Evol. 15, 892–900.

Hawks, J., 2004. How much can cladistics tell us about early hominid relationships? Am. J. Phys. Anthropol. 125, 207–219.

Hawks, J., Oh, S., Hunley, K., Dobson, S., Cabana, G., Dayalu, P., Wolpoff, M.H., 2000. An Australasian test of the recent African origin theory using the WLH-50 calvarium. J. Hum. Evol. 39, 1–22.

Hennig, W., 1966. Phylogenetic Systematics. University of Illinois Press, Urbana.

Herring, S.W., 1993. Epigenetic and functional influences on skull growth. In: Hanken, J., Hall, B.K. (Eds.), The Skull: Vol. 1. Chicago University Press, Chicago, pp. 153–206.

Herring, S.W., Mucci, R.J., 1991. In vivo strain in cranial sutures: the zygomatic arch. J. Morphol. 207, 225–239.

Herring, S.W., Teng, S., Huang, X., Mucci, R.J., Freeman, J., 1996. Patterns of bone strain in the zygomatic arch. Anat. Rec. 246, 446–457.

Herring, S.W., Teng, S., 2000. Strain in the brain case and its sutures during function. Am. J. Phys. Anthropol. 112, 575–593.

Hill, W.C.O., 1974. Primates: Comparative Anatomy and Taxonomy: VII Cynopithecinae: *Cercocebus*,*Macaca*, *Cynopithecus*. University of Edinburgh Press, Edinburgh.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Hillis, D.M., Bull, J.L., White, M.E., Badgett, M.R., Molineux, I.J., 1992. Experimental phylogenetics: generation of a known phylogeny. Science 255, 589–592.

Hinton, R.J., 1982. Differences in interproximal and occlusal tooth-wear among prehistoric Tennessee Indians: implications for masticatory function. Am. J. Phys. Anthropol. 57, 103–115.

Hylander, W.L., 1975. The human mandible: lever or link? Am. J. Phys. Anthropol. 43, 227–242.

Hylander, W.L., 1977. In vivo bone strain in the mandible of *Galago crassicaudatus*. Am. J. Phys. Anthropol. 46, 309–326.

Hylander, W.L., 1979a. An experimental analysis of temporomandibular joint reaction force in macaques. Am. J. Phys. Anthropol. 51, 433–456.

Hylander, W.L., 1979b. Mandibular function in *Galago crassicaudatus* and *Macaca fascicularis*: an in vivo approach to stress analysis of the mandible. J. Morphol. 159, 253–296.

Hylander, W.L., 1984. Stress and strain in the mandibular symphysis of primates: a test of competing hypotheses. Am. J. Phys. Anthropol. 64, 1–46.

Hylander, W.L., 1986. In vivo bone strain as an indicator of masticatory bite force in *Macaca fascicularis*. Arch. Oral Biol. 31, 149–157.

Hylander, W.L., 1988. Implications of in vivo experiments for interpreting the functional significance of "robust" australopithecine jaws. In: Grine, F.E. (Ed.), Evolutionary History of the "Robust" Australopithecines. Aldine de Gruyter, New York, pp. 55–83.

Hylander, W.L., Bays, R., 1979. An in vivo strain-gauge analysis of the squamosal-dentary joint reaction force during mastication and incisal biting in *Macaca mulatta* and *Macaca fascicularis*. Arch. Oral Biol. 24, 689–697.

Hylander, W.L., Crompton, A.W., 1986. Jaw movements and patterns of mandibular bone strain during mastication in the monkey *Macaca fascicularis*. Arch. Oral Biol. 31, 841–848.

Hylander, W.L., Johnson, K.R., 1994. Jaw muscle function and wishboning of the mandible during mastication in macaques and baboons. Am. J. Phys. Anthropol. 94, 523–547.

Hylander, W.L., Johnson, K.R., 1997. In vivo bone strain patterns in the zygomatic arch of macaques and the significance of these patterns for functional interpretations of craniofacial form. Am. J. Phys. Anthropol. 102, 203–232.

Hylander, W.L., Johnson, K.R., 2002. Functional morphology and in vivo bone strain patterns in the craniofacial region of primates: beware of biomechanical stories about fossil bones. In: Plavcan, J.M., Kay, R.F., Jungers, W.L., van Schaik, C.P. (Eds.), Reconstructing Behavior in the Primate Fossil Record. Kluwer Academic/Plenum, New York, pp. 43–72.

Hylander, W.L., Johnson, K.R., Crompton, A.W., 1987. Loading patterns and jaw movements during mastication in *Macaca fascicularis*: a bone strain, electromyographic and cinefluorographic analysis. Am. J. Phys. Anthropol. 72, 287–314.

Hylander, W.L., Johnson, K.R., Crompton, A.W., 1992. Muscle force recruitment and biomechanical modeling: an analysis of masseter muscle function during mastication in *Macaca fascicularis*. Am. J. Phys. Anthropol. 88, 365–387.

Hylander, W.L., Picq, P.G., Johnson, K.R., 1991a. Function of the supraorbital region of primates. Arch. Oral Biol. 36, 273–281.

Hylander, W.L., Picq, P.G., Johnson, K.R., 1991b. Masticatory stress hypotheses and the supraorbital region of primates. Am. J. Phys. Anthropol. 86, 1–36.

Hylander, W.L., Ravosa, M.J., Ross, C.F., Johnson, K.R., 1998. Mandibular corpus strain in primates: further evidence for a functional link between symphyseal fusion and jaw-adductor muscle force. Am. J. Phys. Anthropol. 107, 257–271.

Hylander, W.L., Ravosa, M.J., Ross, C.F., Wall, C.E., Johnson, K.R., 2000. Symphyseal fusion and jaw-adductor muscle force: an EMG study. Am. J. Phys. Anthropol. 112, 469–492.

Jablonski, N.G., 1993. The phylogeny of *Theropithecus*. In: Jablonski, N.G. (Ed.), *Theropithecus*: The Rise and Fall of a Primate Genus. Cambridge University Press, Cambridge, pp. 209–224.

Jolly, C.J., 1966. Introduction to the Cercopithecoidea, with notes on their use as laboratory animals. Symp. Zool. Soc. Lond. 17, 427–457.

Jolly, C.J., 2001. A proper study for mankind: analogies from the papionin monkeys and their implications for human evolution. Yearb. Phys. Anthropol. 44, 177–204.

Jungers, W.L., Falsetti, A.B., Wall, C.E., 1995. Shape, relative size, and size adjustments in morphometrics. Yearb. Phys. Anthropol. 38, 137–161.

Kitching, I.J., Forey, P.L., Humphries, C.J., Williams, D.M., 1998. Cladistics: The Theory and Practice of Parsimony Analysis, second ed. Oxford University Press, Oxford.

Kluge, A.G., 1998. Total evidence or taxonomic congruence: cladistics or consensus classification. Cladistics 14, 151–158.

Kramer, A., Crummett, T.L., Wolpoff, M.H., 2001. Out of Africa and into the Levant: replacement or admixture in Western Asia. Quat. Int. 75, 51–63.

Lanyon, L.E., Rubin, C.T., 1985. Functional adaptation in skeletal structures. In: Hildebrand, M., Bramble, D., Liem, K., Wake, D. (Eds.), Functional Vertebrate Morphology. Harvard University Press, Cambridge, pp. 1–25.

Leigh, S.R., Shah, N.F., Buchanan, L.S., 2003. Ontogeny and phylogeny in papionin primates. J. Hum. Evol. 45, 285–316.

Lieberman, D.E., 1995. Testing hypotheses about recent human evolution from skulls: integrating morphology, function, development, and phylogeny. Curr. Anthropol. 36, 159–197.

Lieberman, D.E., 1997. Making behavioral and phylogenetic inferences from hominid fossils: considering the developmental influence of mechanical forces. Annu. Rev. Anthropol. 26, 185–210.

Lieberman, D.E., 1999. Homology and hominid phylogeny: problems and potential solutions. Evol. Anthropol. 7, 142–151.

Lieberman, D.E., 2000. Ontogeny, homology, and phylogeny in the hominid craniofacial skeleton: the problem of the browridge. In: O'Higgins, P., Cohn, M.J. (Eds.), Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton. Academic Press, London, pp. 85–122.

Lieberman, D.E., Crompton, A.W., 1998. Responses of bone to stress. In: Wiebel, E., Taylor, C., Bolis, L. (Eds.), Principles of Biological Design: The Optimization and Symmorphosis Debate. Cambridge University Press, Cambridge, pp. 78–86.

Lieberman, D.E., Ross, C., Ravosa, M.J., 2000. The primate cranial base: ontogeny, function, and integration. Yearb. Phys. Anthropol. 43, 117–169.

Lieberman, D.E., Wood, B.A., Pilbeam, D.R., 1996. Homoplasy and early *Homo*: an analysis of the evolutionary relationships of *H. habilis sensu stricto* and *H. rudolfensis*. J. Hum. Evol. 30, 97–120.

Lockwood, C.A., Fleagle, J.G., 1999. The recognition and evaluation of homoplasy in primate and human evolution. Yearb. Phys. Anthropol. 42, 189–232.

Lovejoy, C.O., Cohn, M.J., White, T.D., 1999. Morphological analysis of the mammalian postcranium: a developmental perspective. Proc. Natl. Acad. Sci. 96, 13247–13252.

Lovejoy, C.O., Cohn, M.J., White, T.D., 2000. The evolution of mammalian morphology: a developmental perspective. In: O'Higgins, P., Cohn, M.J. (Eds.), Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton. Academic Press, London, pp. 41–55.

Maddison, D.R., Maddison, W.P., 1998. MacClade 4: Analysis of Phylogeny and Character Evolution Version 4.02. Sinauer Associates, Sunderland.

Martin, R.D., 1993. Allometric aspects of skull morphology in *Theropithecus*. In: Jablonski, N.G. (Ed.), *Theropithecus*: The Rise and Fall of a Primate Genus. Cambridge University Press, Cambridge, pp. 273–298.

Martin, R., Burr, D.B., 1989. Structure, Function, and Adaptation of Compact Bone. Raven, New York.

Martin, R.B., Burr, D.B., Sharkey, N.A., 1998. Skeletal Tissue Mechanics. Springer, New York.

McCollum, M.A., 1999. The robust australopithecine face: a morphogenetic perspective. Science 284, 301–305.

McCollum, M.A., Sharpe, P.T., 2001. Developmental genetics and early hominid craniodental evolution. BioEssays 23, 481–493.

McHenry, H.M., 1996. Homoplasy, clades, and hominid phylogeny. In: Meikle, W.E., Howell, F.C., Jablonski, N.G. (Eds.), Contemporary Issues in Human Evolution. California Academy of Sciences, San Francisco, pp. 77–92.

Meyer, C., Kahn, J.L., Boutemi, P., Wilk, A., 2002. Photoelastic analysis of bone deformation in the region of the mandibular condyle during mastication. J. Craniomaxillofac. Surg. 30, 160–169.

Mosimann, J.E., 1970. Size allometry: size and shape variables with characteristics of the log normal and generalized gamma distributions. J. Am. Stat. Assoc. 65, 930–945.

Nakagawa, S., 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. Behav. Ecol. 15, 1044–1045.

Olsen, E.C., Miller, R.L., 1958. Morphological Integration. University of Chicago Press, Chicago.

Page, S.L., Goodman, M., 2001. Catarrhine phylogeny: noncoding DNA evidence for a diphyletic origin of the mangabeys and for a human–chimpanzee clade. Mol. Phylogenet. Evol. 18, 14–25.

Patterson, C., 1982. Morphological characters and homology. In: Joysey, K.A., Friday, A.E. (Eds.), Problems in Phylogenetic Reconstruction. Academic Press, London, pp. 21–74.

Perneger, T.V., 1998. What's wrong with Bonferroni adjustments. Br. Med. J. 316, 1236–1238.

Quicke, D.L.J., 1993. Principles and Techniques of Contemporary Taxonomy. Blackie, London.

Rae, T.C., 1997. The early evolution of the hominoid face. In: Begun, D.R., Ward, C.V., Rose, M. (Eds.), Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations. Plenum Press, New York, pp. 59–77.

Rae, T.C., 1998. The logical basis for the use of continuous characters in phylogenetic systematics. Cladistics 14, 221–228.

Rafferty, K.L., Herring, S.W., 1999. Craniofacial sutures: morphology, growth, and in vivo masticatory strains. J. Morphol. 242, 167–179.

Ravosa, M.J., Johnson, K.R., Hylander, W.L., 2000a. Strain in the galago facial skull. J. Morphol. 245, 51–66.

Ravosa, M.J., Noble, V.E., Hylander, W.L., Johnson, K.R., Kowalski, E.M., 2000b. Masticatory stress, orbital orientation and the evolution of the primate postorbital bar. J. Hum. Evol. 38, 667–693.

Ravosa, M.J., Profant, L.P., 2000. Evolutionary morphology of the skull in Old World monkeys. In: Whitehed, P., Jolly, C.J. (Eds.), Old World Monkeys. Cambridge University Press, Cambridge, pp. 237–268.

Reidl, R.J., 1978. Order in Living Organisms. Wiley, New York.

Ross, C.F., 2001. In vivo function of the craniofacial haft: the interorbital "pillar." Am. J. Phys. Anthropol. 116, 108–139.

Ross, C.F., Hylander, W.L., 1996. In vivo and in vitro strain in the owl monkey circumorbital region and the function of the postorbital septum. Am. J. Phys. Anthropol. 101, 183–215.

Ross, C.F., Hylander, W.L., 2000. Electromyography of the anterior temporalis and masseter muscles of owl monkeys (*Aotus trivirgatus*) and the function of the postorbital septum. Am. J. Phys. Anthropol. 112, 455–468.

Sanderson, M.J., Hufford, L., 1996. Homoplasy: The Reoccurrence of Similarity in Evolution. Academic Press, San Diego.

Schuh, R.T., 2000. Biological Systematics: Principles and Applications. Cornell University Press, London.

Simon, C., 1983. A new coding procedure for morphometric data with an example from periodical cicada wing veins. In: Felsenstein, J. (Ed.), Numerical Taxonomy. Springer-Verlag, Berlin, pp. 378–382.

Simpson, G.G., 1961. Principles of Animal Taxonomy. Columbia University Press, New York.

Singleton, M., 2002. Patterns of cranial shape variation in the Papionini (Primates: Cercopithecinae). J. Hum. Evol. 42, 547–578.

Skelton, R.R., McHenry, H.M., Drawhorn, G.M., 1986. Phylogenetic analysis of early hominids. Curr. Anthropol. 27, 21–43.

Skelton, R.R., McHenry, H.M., 1992. Evolutionary relationships among early hominids. J. Hum. Evol. 23, 309–349.

Skerry, T., 2000. Biomechanical influences on skeletal growth and development. In: O'Higgins, P., Cohn, M.J. (Eds.), Development, Growth and Evolution: Implications for the Study of the Hominid Skeleton. Academic Press, London, pp. 29–39.

Slowinski, J.B., 1993. "Unordered" versus "ordered" characters. Syst. Biol. 42, 155–165.

Smith, A.B., 1994. Systematics and the Fossil Record: Documenting Evolutionary Patterns. Blackwell, Oxford.

Sober, E., 1988. Reconstructing the Past: Parsimony, Evolution and Inference. MIT Press, Cambridge.

Sokal, R.R., Rohlf, F.J., 1995. Biometry, third ed. W.H. Freeman and Co., New York.

Spencer, M.A., 1998. Force production in the primate masticatory system: electromyographic tests of biomechanical hypotheses. J. Hum. Evol. 34, 25–54.

Strait, D.S., 2001. Integration, phylogeny, and the hominid cranial base. Am. J. Phys. Anthropol. 114, 273–297.

Strait, D.S., Grine, F.E., 2004. Inferring hominoid and early hominid phylogeny using craniodental characters: the role of fossil taxa. J. Hum. Evol. 47, 399–452.

Strait, D.S., Grine, F.E., Moniz, M.A., 1997. A reappraisal of early hominid phylogeny. J. Hum. Evol. 32, 17–82.

Strait, D.S., Moniz, M.A., Strait, P.T., 1996. Finite mixture coding: a new approach to coding continuous characters. Syst. Biol. 45, 67–78.

Strasser, E., 1988. Pedal evidence for the origin and diversification of the cercopithecid clades. J. Hum. Evol. 17, 225–245.

Strasser, E., Delson, E., 1987. Cladistic analysis of cercopithecid relationships. J. Hum. Evol. 16, 81–99.

Swofford, D.L., 1998. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods) Version 4.0. Sinauer Associates, Sunderland.

Szalay, F.S., Delson, E., 1979. Evolutionary History of the Primates. Academic Press, New York.

Thiele, K., 1993. The holy grail of the perfect character: the cladistic treatment of morphometric data. Cladistics 9, 275–304.

Thorpe, R.S., 1984. Coding morphometric characters for constructing distance Wagner networks. Evolution 38, 244–255.

Tosi, A.J., Disotell, T.R., Morales, J.C., Melnick, D.J., 2003. Cercopithecine Y-chromosome data provide a test of competing morphological evolutionary hypotheses. Mol. Phylogenet. Evol. 27, 510–521.

Wall, C.E., 1999. A model of temporomandibular joint function in anthropoid primates based on condylar movements during mastication. Am. J. Phys. Anthropol. 109, 67–88.

Wiens, J.J., 2004. The role of morphological data in phylogeny reconstruction. Syst. Biol. 53, 653–661.

Willey, A., 1911. Convergence in Evolution. John Murray, London.

Wiley, E.O., Siegel-Causey, D., Brooks, D.R., Funk, V.A., 1991. The Compleat Cladist: A Primer of Phylogenetic Procedures. The University of Kansas Museum of Natural History, Lawrence.

Wolpoff, M.H., Hawks, J., Frayer, D.W., Hunley, K., 2001. Modern human ancestry at the peripheries: a test of the replacement theory. Science 291, 293–297.

Wood, B.A., 1975. An analysis of sexual dimorphism in primates. Ph.D. Dissertation, University of London.

Wood, B.A., 1991. Koobi Fora Research Project, Vol. IV: Hominid Cranial Remains. Clarendon Press, Oxford.

Wood, B.A., Collard, M., 1999. The human genus. Science 284, 65–71.

Wood, B.A., Lieberman, D.E., 2001. Craniodental variation in *Paranthropus boisei*: a developmental and functional perspective. Am. J. Phys. Anthropol. 116, 13–25.