

## COMMENTARY

## Cranial Anatomy and Baboon Diversity

CLIFFORD JOLLY\*

Department of Anthropology, New York University, New York, New York

The following article by Frost et al. (2003) exemplifies an exciting application of new ways of using anatomical evidence to reconstruct evolutionary history and embodying significant innovations on both the practical and the theoretical fronts. Their chosen subject is cranial variation in the large, extant African monkeys known as baboons, which, together with the mangabeys (their close and probably primitive relatives), comprise a zoological subtribe, the Papionina. The baboons are a particularly interesting example of mammalian evolution in tropical Africa. They have been extensively watched in the wild, so their natural history is well known. They include one genus that is very widespread and diverse, and two others that are much more restricted in diversity and distribution, thus representing different patterns of evolutionary change. Finally, for those concerned with human evolution, the study of baboons has the added interest that the place, time-frame, and ecological setting of their evolutionary emergence and diversification coincide quite closely with those of the human species and its close, extinct relatives (the Tribe Hominini). For this reason, it has been suggested that the baboons provide a uniquely apt set of analogies to help us understand human evolutionary history (Jolly, 2001).

Another great advantage of the baboons as a subject of study is that they have been quite extensively “collected”; that is, they have been the object—some might say the victims—of the curious custom of taking heads as trophies. This practice, a favorite pastime of the Celts of Caesar’s day, among other human groups, appears to be underlain by a more or less implicit belief that a head (or a skull) embodies an individual’s essence. And if the skin also is collected, this completes the trophy by capturing its outward appearance too. Whatever the murky cultural roots of the practice of decapitation and flaying, it is fortunate for science that early zoological collectors followed it so enthusiastically. The scientific headhunters of a bygone era achieved a comprehensive geographical coverage that is now documented by the skin and skull collections of traditional museums of natural history. It could never be duplicated in the pitifully impoverished world ecosystem of today. All systematists agree that these collections are a treasure chest of evidence about biological diversity, but how are they best exploited?

Some researchers delight in shocking their colleagues by suggesting that the best use for such museum materials would be to grind them up as a source of DNA, an organism’s quintessence in today’s popular mythology. Those of a more traditional bent see scientific justifica-

tions, not unlike the headhunter’s mystical ones, for valuing skulls as structures. For an evolutionary biologist, the “essence” that a cranium embodies is information about phylogenetic relationships, encoded in an anatomical complexity that results from evolutionary adaptation to the skull’s unique multifunctionality. A mammalian cranium, of course, houses the front end of the digestive and respiratory systems, the major component of the central nervous system, and most of the receptor organs for the special senses. Its form is also influenced by locomotion, social communication, and even reproduction. Each species’—indeed each individual’s—cranial form is a compromise between the demands of multiple functional roles, mediated by developmental patterns from its evolutionary heritage. Because the exact configuration of the compromise is unlikely to be duplicated in another evolutionary lineage, detailed resemblance in craniodental anatomy is a likely indicator of a phylogenetic relationship. Some remarkable similarities in cranial form can certainly be explained by functional parallels, similarity in absolute size, or chance, rather than by shared evolutionary ancestry, but it is most unlikely that cranial anatomy is so dominated by such homoplasies that no evidence of evolutionary history can be found. Extracting this evidence in useable form presents three challenges: how to select measures that capture the essential features of complex anatomical shapes, how to make these measurements accurately and reproducibly, and how to process the data to yield evolutionary information.

For the task of identifying critical morphological variables, the human eye is a wonderfully sensitive and subtle instrument, but it is also idiosyncratic, and notoriously easy to mislead when judging similarities and differences among the shapes of objects of different sizes and (as occurs especially in paleontology) varied superficial textures and states of preservation. One practical solution to the problem of achieving some quantification and reproducibility, while not losing the value of the informed eye, is to combine observation and measurement. Characters

\*Correspondence to: Dr. Clifford Jolly, Professor of Anthropology, Department of Anthropology, New York University, New York, NY 10003. E-mail: clifford.jolly@nyu.edu

Received 23 September 2003; Accepted 23 September 2003  
DOI 10.1002/ar.a.10121

suspected of being functionally or phylogenetically significant are identified qualitatively by eye. These impressions are then tested by devising ad hoc but repeatable measures and indices to standardize and quantify the observations. The classic example is the cranial index (braincase width  $\times$  100/ braincase length), which was devised by early craniometrists to express the fact that some human skulls are observably longer and narrower than others. One hazard of this approach is reification: the index comes to be regarded as an entity in itself, rather than simply a way of expressing an aspect of shape. However, if indices and angles are treated strictly as heuristic devices, each designed specifically to express a particular aspect of shape, the observe-then-measure method has advantages, not least of which is that it is explicitly hypothetico-deductive, and that its results are intuitively comprehensible. Deservedly, it continues to be a mainstay of comparative morphology and paleontology, for both the cranium and the postcranial skeleton.

A second approach is more inductive, and in some ways, perhaps, more objective. Rather than using metrics to test the reality of differences already perceived by eye, or predicted on functional grounds, it extracts differences and similarities statistically from a mass of metric data collected by standardized methods. One way to do this is to record all the traditional, point-to-point metrics devised to express simple measures and proportions (cranial length and breadth, nasal breadth and height, etc.), and then reduce them statistically to fewer dimensions or components that can then be used to compare individuals and clusters of specimens. Though the meaning of the output is somewhat less easy to interpret intuitively than simple metrics, the influence of particular parameters on extracted dimensions can be judged to some extent. One strength of such methods is that they can do what the eye cannot easily do—compensate for the effects of absolute size on proportions. Furthermore, when we attempt to assess the meaning of similarity, they allow for correlation among dimensions. Specimens can be characterized and compared in terms of attributes that are composites and are thus impossible to comprehend by simple inspection. These attributes include overall shape and its statistically extracted components, overall size independent of shape, and shape as correlated with size.

For their craniological study, Frost et al. (2003) followed the second approach, and moved even further in the inductive direction by abandoning the traditional set of point-to-point dimensions. Instead they collected data consisting of the relative positions, in three dimensions, of 45 standardized points on the skull surface, selected to capture the major features of cranial shape. They were able to perform this sophisticated procedure because instead of the traditional callipers they used a digitizer—a mechanical device that relates the standard points spatially to a few constant points of reference (Microscribe 3DX, Immersion Corp., San Jose, CA). They then used a comprehensive multivariate statistical treatment to extract variables expressing shape, size, and size-related shape to compare them among taxa and between sexes, and plot their geographical distribution.

Considering the very extensive technical, statistical, and computational power expended on the data collection and analysis, it is relevant to ask what it achieves that would be unattainable by the more traditional, hypothetico-deductive, eye-and-calliper approach. The cur-

mudgeonly traditionalist is moved to point out that in some respects, the results confirm the obvious. No experienced primatologist needs such complex analysis to appreciate that the skull of a mandrill or gelada differs from that of a *Papio* baboon, or that males differ from females of the same species. The same experienced primatologist could easily devise simple measures to test whether the differences were real or illusory. For this commentator, at least, it is actually easier to see, and to document metrically, that a gelada baboon has a deeper face, shorter nasal bones, and less arched brows than a common baboon by simply comparing the skulls side by side, and designing critical measures, than by looking at abstractions such as landmark deviations and splines. However, this may well be an effect of my unfamiliarity with the new technique.

That said, however, one must recognize a highly significant advance represented by the new approach: the reproducibility it achieves in data collection and analysis. Some element of subjectivity and experience must be involved in deciding which digitized points will best represent the skull's shape, but from that stage onward the procedure is standardized. This makes the recognition, representation, and measurement of features of resemblance and difference much more amenable to duplication by independent observers and less dependent upon the educated eye, and thus in that sense more completely "scientific." Once the unfamiliar images are mastered, spline and landmark deviation analysis should enable investigators to pinpoint the cranial regions that are most susceptible to evolutionary change. And if the results appear to ignore or misrepresent a difference that seems obvious to the eye, there is no prohibition against repeating the analysis with a fresh set of data points, until the investigator is satisfied that the apparent difference has been captured, or is truly illusory. Another major advantage accrues from sophisticated multivariate analysis. Even informed eyeballing could not suggest the hypothesis that differences in skull shape between males and females of a species can be partitioned into those that are simple correlates of size, and those that are not, and thus require a functional explanation. When comparisons are made between species within a genus, rather than between genera, variation among populations is more subtle, the effects of sex and absolute size difference are relatively more obtrusive, and the value of multivariate methods is correspondingly more obvious. In this context, it is interesting that a higher-level principal component separates the two species of *Mandrillus*, which to the experienced eye appear similar, though not identical.

The internal systematics of *Mandrillus* and (living) *Theropithecus* are, however, relatively unproblematic. Each includes one or two well defined, geographically limited units that by any definition are "good" species (that is, until drill-mandrill hybrids are discovered in the wild, causing conniptions among purist adherents of the biological species concept (BSC). Variation in extant *Papio*, however, is quite another (and more interesting) story, which is worth outlining in the present context, since Frost et al. (2003) use "taxonomy" as a variable against which to compare cranial variation. However, the complex geographical variation of *Papio* baboons fits less readily into a conventional taxonomic scheme than their study might suggest. In fact, among those who care about baboon systematics there is a bewildering variety of opinions about the number (5, 6, 14, 18?) and rank (species, subspecies,

semispecies, superspecies?) of baboon taxa that should be recognized. (The circle of interested parties is wider than one might imagine, since it includes many medical researchers who use baboons and want a consistent name to put on them, even though they may know little and care less about the natural history or evolution of their subjects). Much of the disagreement stems not from the facts of baboon biology, but from divergent philosophies of classification, and especially the quest for that holy grail of taxonomy, the perfect species definition (Hey, 2001). At present, most biologists (especially those whose subject is living, rather than extinct, vertebrate animals) adhere to one of two rival definitions: the tried-and-true BSC (Mayr, 1942), which essentially defines a “good” species by genetic isolation, and the phylogenetic species concept (PSC) (Eldredge and Cracraft, 1980), which defines it as the smallest natural population whose members all share at least one evolutionary novelty.

It has been suggested (Jolly, 1993, 2001) that preoccupation with species definitions distracts attention from the real biological issues, and that it is more productive to shift the discussion to a less abstract level, focusing not on generating taxonomies but on gathering and interpreting the data on which all taxonomies are based, under any species definition. These essential facts of nature fall quite straightforwardly into two conceptually distinct categories: information about interbreeding (“zygostructure”), and information about the distribution of heritable characters, including genes (“phenostructure”). The two are functionally intertwined, in that present phenostructure is the product of past zygostructure (and therefore can be used to reconstruct it), and present zygostructure determines future phenostructure (which it can therefore predict). The BSC is essentially zygo-based, and therefore mostly looks *forward* in time. In conversation, its proponents tend to defend it in general terms, such as “because they are reproductively isolated, different species will follow distinct evolutionary trajectories,” or, more narrowly, “populations A and B should be called the same species because an advantageous gene arising in population A could spread throughout A and B.” The PSC is pheno-based (the key attribute of a species is “diagnosability”) and therefore deals with species as products of past events. Its proponents regard it as more factually based than the BSC, and, especially, as a more appropriate unit for cladistic reconstruction of phylogeny (a sequence of species origins *in the past*). One unfortunate by-product of the quasi-theological debate about the “real” nature of species is that both sides tend to deemphasize the continuity in nature that, as Darwin recognized, is one of the most potent pieces of evidence for the evolutionary origin of species. Total reproductive isolation and complete diagnosability each lie at the end of a continuum, and many less extreme states along the two continua are found in natural populations. These intermediate states of differentiation or reproductive isolation are not necessarily either ephemeral or transitional—subspecies are not necessarily either short-lived or species in the making. On the other hand, neither complete genetic isolation nor complete diagnosability is in fact irreversible.

As might be expected, in practice the difference between the species definitions emerges most clearly in cases in which zygostructure and phenostructure do not coincide, and in which a widespread taxon includes a patchwork of physically distinct populations that interbreed where they

meet. The PSC calls such populations species; the BSC calls them subspecies of a single polytypic species. Such a situation is seen in the *Papio* baboons. Frost et al. (2003) are proponents of the BSC, although, mercifully, they do not repeat all the arguments for and against this concept. True to the zygo-definition of the BSC, for them it is the fact that natural hybrid zones have been found in the wild that definitively demotes the recognizable baboon taxa to subspecies of a single species. However, they also cite features of the geographical distribution of cranial variation that are consistent with this view, providing additional, indirect support for the recent occurrence of gene flow between all geographically adjacent populations. Neighboring populations tend to resemble each other, and when individual specimens are plotted on a map of trans-Saharan Africa, the main axis of variation in cranial shape forms a single, uncomplicated gradient or cline, running from northwest to southeast.

The authors’ conclusion that all living *Papio* baboons belong to one *biological* species is indisputable, and in fact is undisputed. More debatable, however, is how best to express the patterning of variation within this species. After all, calling this variation “intraspecific” does not cause it to disappear, or make it any less significant as information about the evolutionary history of the genus. Moreover, Frost et al. (2003) use the taxonomic divisions within *Papio* as units of comparison in some of their analyses, so the definition of those divisions obviously affects the results they obtain. The taxonomy that Frost et al. (2003) adopt is a modification of a widely used scheme that can be called the “traditional five” (Hill, 1967; Groves, 2001). Like all such classifications it is based primarily on external features of coat color and texture, body build, and tail carriage. It recognizes five kinds of *Papio* baboons (Guinea, anubis, hamadryas, chacma, and yellow) as full species. Frost et al. (2003) (see their Fig. 1) modify this scheme by simply demoting the first four to subspecies, and dividing the fifth (yellow baboons) between two subspecies that differ in size and some other features (yellow baboons (*sensu stricto*) and Kinda baboons).

This six-subspecies scheme may seem like an obvious and admirably conservative compromise, but it obscures much geographical diversity. For example, chacma baboons, which Frost et al. (2003) call subspecies *Papio hamadryas ursinus*, comprise at least four, and probably more, physically distinct, geographically circumscribed populations (gray-footed, Kalahari, Transvaal, and Cape chacmas). If chacmas are recognized as a full species, each of these can be a subspecies. When chacmas as a group are demoted to a subspecies, the physical distinctiveness of these populations remains a fact of nature, but the neo-Linnaean system does not allow it to be expressed by naming “sub-subspecies.” Anubis and yellow baboons (*P. h. anubis* and *P. h. cynocephalus*, respectively, in the scheme of Frost et al., 2003) are similarly polytypic, and so, probably, are hamadryas baboons (*P. h. hamadryas*.) A recent survey (Jolly and Burrell, unpublished result) recognized 18 probably diagnosable, geographically-replacing forms, or allotaxa (Grubb, 1999), among extant *Papio* baboons, and there may well be more. In formal taxonomy, each of these could be named a (BSC) subspecies of *P. hamadryas* (*P. h. ibeanus*, *P. h. griseipes*, *P. h. anubis*, etc.), or a full (PSC) species of *Papio* (*Papio ibeanus*, *P. griseipes*, etc.). The difference in naming (I would argue) has no biological significance. The important feature is

that this scheme eliminates the “traditional five” as formal taxa. This would no doubt cause widespread confusion and dismay among primatologists (which was why I did not propose it in 1993), but in some ways it makes for a more accurate representation of biological realities. For example, the “gray-footed chacma” population whose range adjoins that of yellow baboons in east-central Africa has as much in common, phenotypically, with yellow baboons to its north as it does with chacmas from further south, being intermediate between the two. In fact, the whole series of baboon allotaxa form a single chain, with a lateral branch that links Kinda baboons to yellow baboons. In almost all cases (the single obvious exception is mentioned below), the features of pelage and other external features that define the 18 allotaxa change in a step-wise pattern, so that each allotaxon is physically intermediate between its neighbors on either side. Since this distribution of phenotypic characters appears to match the geographical cline that Frost et al. (2003) find in cranial form, it would be most interesting to see the cranial data analyzed with all 18 distinct allotaxa treated separately. Though small sample sizes would be a problem in some cases, the steps in the resulting gradients would probably be lower, and the case for continuity within a single biological species would be correspondingly stronger.

It is quite possible, however, that such an analysis might reinstate some or all of the “traditional five” as clusters of allotaxa with shared cranial growth patterns. Frost et al. (2003) find that each of the six taxa they recognize differs from the others in size-corrected cranial shape, with a single, interesting exception. Kinda baboons differ only in size, not in size-corrected cranial shape, from “regular” yellow baboons, with which they are traditionally grouped. There is an assumption prevalent among morphologists who study allometry (changes of proportion correlated with size) that this is an important distinction. A resemblance (or difference) between species in shape-determining trajectories of growth is considered more significant in some way than a “mere” difference in overall size. This assumption, in turn, rests on the belief that a shape change is more complex, and is thus less likely than a size change to be reversed in the course of evolution, or duplicated in separate evolutionary lines. (This seems reasonable, but in fact it is something of a leap of faith. Since we know nothing whatever about the genetics underlying either of these aspects of variation, we simply have no way of estimating whether it is actually “easier” for evolution to change the shape of a baboon’s skull than to alter its size.)

Be that as it may, the finding that Kinda baboons are, cranially, simply miniaturized versions of “regular” yellow baboons is still interesting. To complete the picture, however, we now need to learn whether this is also true of other small-skulled allotaxa. For example, are the small anubis baboons found in the Saharan massif of Tibesti simply scaled-down versions of the large anubis baboons of west-central Africa, and the small chacmas found in the Kalahari just miniaturized versions of the large chacmas of the Cape? If this should prove to be the case, it would support recognition of the “traditional five”—perhaps as informal clusters of subspecies (or, in the PSC system, clusters of species). In terms of evolutionary history, it might imply that each of the five species represents a population that was geographically isolated and became

distinct, perhaps during a severe arid period corresponding to a high-latitude glaciation.

It might also be instructive to look at the association between aspects of cranial variation and states of the individual external characters (coat color, texture, tail carriage, and so on) that are bundled into the variable labeled “taxonomy” by Frost et al. (2003). One such association is already apparent. Size- and sex-corrected cranial features document a break in the overall northwest to southeast cline, separating the northern cluster (anubis, hamadryas, and Guinea baboons) from the southern cluster (yellow, chacma, and Kinda baboons), and the same dichotomy is seen in pelage features. The northern forms share wavy, bushy shoulder-manes and cheek-tufts, as well as broad skulls, low faces, and less-downturned muzzles. The southern forms mostly share straight, silky mane hairs, untufted cheeks, and light facial patches, as well as narrow skulls with high faces that are more klinorhynch. The twofold, north–south division in pelage characters has long been recognized, along with its association with a contrast in skull form between anubis (northern) and chacmas (southern), both of which are large in body size, and therefore extreme in cranial form (Jolly, 1965). However, it took the more sophisticated analysis by Frost et al. (2003), with its analytical separation of size and shape, to show clearly that the smaller-sized baboons (hamadryas, Guinea, yellow, and Kinda) also fit the north–south dichotomy in cranial form. An analysis recognizing all allotaxa might produce further insights. It would be interesting to know, for example, whether baboons from coastal Kenya and Somalia (“ibean” baboons), which have “southern” yellow baboon coloration but “northern” wavy mane hair, also share some cranial features with northern baboons, and if so, whether they are closer to anubis (their neighbors to the west) or hamadryas (to the north).

Appropriately, because their data are purely phenetic and not readily converted to a form amenable to cladistic analysis, Frost et al. (2003) do not attempt to push their inferences about phylogeography back very far, beyond using current phenostructure to infer the zygostructure of the immediate past. This approach, of course, is also in accord with their objective of diagnosing BSC species and subspecies—they are more concerned with present and future gene-flow than with past events of population differentiation. Perhaps for this reason, and because they treat “taxonomy” as something of a black box—a unitary variable, stipulated in advance—they do not apply their data to one of the most interesting and potentially informative features of contemporary baboon phenostructure. This is the fact that although hamadryas baboons in the northeast of sub-Saharan Africa, and Guinea baboons in the northwest are widely separated geographically, they share several, unrelated, features of external appearance that are not seen in the anubis baboons whose range lies between theirs. This pattern is significant because, unlike most phenotypic variation among living baboons, it can be explained only by a somewhat complex phylogeographic history. It implies that an intrusive expansion of anubis baboons displaced an ancestral and continuous hamadryas-Guinea population. It is unclear (to this reviewer, at least) whether the interpopulational differences in size- and sex-controlled cranial shape, as presented by Frost et al. (2003), tend to support this hypothesis, but it would be worthwhile to apply the cranial data directly to test it.

Inevitably, the results from the cranial study invite comparison with the still sparse, but rapidly accumulating, genetic evidence about the phylogeography of *Papio* baboons. A full comparative review of the two data sets is beyond the scope of this note, and in any case would be premature; however, a few recent findings are relevant. The phylogeny of mitochondrial sequences (tentatively) suggests an ancient dichotomy between “southern” and “northern” baboons that in a general way mirrors the twofold division in pelage features and cranial shape (A. Burrell, personal communication). The geographical position of the divide in East Africa is still undocumented, but it probably does *not* coincide exactly with either a cranial or a major pelage transition. Even more striking is the mitochondrial evidence for a long history of extensive gene-flow between populations that belong to quite distinct taxa, as defined by pelage. All anubis baboons of Ethiopia, for example, appear to carry mitochondria that probably originated in hamadryas (Wildman et al., submitted), while anubis in Kenya carry mitochondrial sequences closely related to those carried by phenotypic yellow baboons in the same region (Newman et al., in press). Thus, although mitochondrial base sequences represent a completely different kind of variation, they agree with cranial shape in suggesting an evolutionary history of alternation between periods when baboon populations contracted, became isolated from each other, and differentiated physically, and other periods when they expanded, established mutual boundaries, and exchanged genes.

For those of us who see *Papio* baboons as a valuable example of primate and mammalian evolution in the tropics, the challenge now is to reconstruct the how, when, and where of this process by synthesizing the various informa-

tion streams. Like all landmark pieces of original research, the article by Frost et al. (2003) both contributes substantially to answering these questions, and provides a fruitful source of ideas and new research directions.

#### LITERATURE CITED

- Eldredge N, Cracraft J. 1980. Phylogenetic patterns and the evolutionary process. New York: Columbia University Press. 349 p.
- Frost SR, Marcus LF, Bookstein FL, Reddy DP, Delson E. 2003. Cranial allometry, phylogeography, and systematics of large-bodied papionins (primates: *Cercopithecinae*) inferred from geometric morphometric analysis of landmark data. *Anat Rec* 275A:1048–1072.
- Groves CP. 2001. Primate taxonomy. Washington, D.C.: Smithsonian Institution Press. 350 p.
- Grubb P. 1999. Evolutionary processes implicit in distribution patterns of modern African mammals. In: Bromage TG, Schrenk F, editors. African biogeography, climate change, and human evolution. New York: Oxford University Press. p 253–267.
- Hey J. 2001. The mind of the species problem. *Trends Ecol Evol* 16:150–164.
- Hill WCO. 1967. Taxonomy of the baboon. In: Vagtberg H, editor. The baboon in medical research. Vol. II. Austin, TX: University of Texas Press. p 4–11.
- Jolly CJ. 1965. The origins and specialisations of the long-faced Cercopithecoidea. Ph.D. dissertation, University of London, London, UK. 736 p.
- Jolly CJ. 1993. Species, subspecies and baboon systematics. In: Kimbel W, Martin L, editors. Species, species concepts and primate evolution. New York: Wiley. p 67–107.
- Jolly CJ. 2001. A proper study for mankind: analogies from the Papionin monkeys and their implications for human evolution. *Yearb Phys Anthropol* 33:177–204.
- Mayr E. 1942. Systematics and the origin of species. New York: Columbia University Press. 330 p.
- Newman TK, Jolly CJ, Rogers J. Mitochondrial phylogeny and systematics of baboons (*Papio*). *Am J Phys Anthropol* (in press).