Sexual swelling in mandrills (*Mandrillus sphinx*): a test of the reliable indicator hypothesis

Joanna M. Setchell\(^a\,b\) and E. Jean Wickings\(^b\)

\(^a\)University of Surrey Roehampton, London, UK, and \(^b\)Centre International de Recherches Médicales, B.P. 769, Franceville, Gabon

Various hypotheses have been proposed to account for the function of sexual swellings in female primates, but few empirical data exist to test predictions arising from these hypotheses. Controversy has recently arisen over a field study that appeared to support the predictions of the reliable indicator hypothesis. This hypothesis proposes that females compete for males or matings, that differences in swelling size between females reliably advertise female quality, and that males use swelling characteristics to differentially allocate mating effort to females with certain swelling characteristics, hence to females of higher quality. To provide an independent test of this hypothesis, we collected data concerning the size and coloration of 40 sexual swellings for 29 semi-free-ranging female mandrills, varying in age and parity, along with data concerning the behavior of males toward the females, and compared these with the long-term reproductive history of the females. We examined the following predictions: (1) swelling characteristics are consistent across subsequent cycles for individual females, (2) swelling characteristics indicate aspects of female reproductive quality, and (3) males prefer to mate with females that show particular swelling characteristics. Our results support prediction 1; we found little change in swelling characteristics across swellings for individual females. However, we found no significant relationships between female reproductive history and swelling characteristics and, thus, no support for prediction 2. Finally, we found only limited support for prediction 3; females with larger (wider) sexual swellings were more likely to have a sperm plug when maximally swollen. However, male mate-guarding was not significantly related to female swelling characteristics. Furthermore, in situations in which more than one female was maximally swollen, the alpha male (who has “free” sperm plug when maximally swollen) did not show the most interest in the female with the largest swelling. We conclude that the reliable indicator hypothesis does not explain variation in sexual swellings in female mandrills. **Key words:** mandrills, prediction, primates, reproduction, sexual swelling.  *Behav Ecol* 15:438–445 (2004)

Females of many Old World primate species develop prominent and conspicuous swellings of the perineal skin that reach maximum size around the time of ovulation. These highly visible “sexual swellings” have attracted a great deal of attention from evolutionary biologists. They have evolved independently in at least three separate primate lineages and are associated with multimale-multifemale mating systems, in which females mate with more than one male per receptive period (Dixson, 1983). Sexual swellings are hormone dependent (Dixson, 1998), attract males for mating independently of olfactory or behavioral cues (Bielert and Anderson, 1985; Bielert and Girolami, 1986; Bielert and van der Walt, 1982; Bielert et al., 1989; for summary, see Snowdon, 2003), and are likely to be costly to the females (for review, see Nunn et al., 2001). A number of hypotheses, not necessarily mutually exclusive, have been proposed to account for the evolution of sexual swellings (for review, see Nunn, 1999; Stallmann and Froehlich, 2000). These include indirect female choice through incitement of male-male competition, allowing the female to mate with the “best male” (Clutton-Brock and Harvey, 1976); paternity confusion (Hrdy, 1981); increased paternity confidence (Hamilton, 1984); promotion of postcopulatory sexual selection (Harvey and May, 1989); and honest signaling of female quality (the “reliable indicator” hypothesis; Pagel, 1994).

Few empirical studies have addressed the predictions of these hypotheses, and the exact function of sexual swellings remains unclear and controversial. A recent field study on olive baboons (*Papio cynocephalus anubis*) in Gombe National Park, Tanzania (Domb and Pagel, 2001) found support for the reliable indicator hypothesis, which proposes that sexual swellings can be regarded as a costly handicap that honestly signals female quality (Pagel, 1994). In the first detailed study concerning male mate choice for sexual swelling characteristics, Domb and Pagel (2001) showed that sexual swellings in wild female baboons reliably advertise a female’s reproductive value over her lifetime, and that males use swelling size to determine their reproductive effort. Female baboons with larger (in one of three dimensions measured) swellings attained sexual maturity earlier, produced more offspring, and produced more surviving offspring, whereas male baboons showed greater interest in, and competed more intensely for, females with larger swellings. However, Zinner et al. (2002) have highlighted several flaws in the design of this study. In particular, the research ignored intraindividual variability of swelling characteristics, the possibility that the covariate body size could explain the significant results, and variation in demography and food availability between the five baboon groups studied. Synchronously mating females were also excluded from the analysis, although investigation of overlapping estrus cycles might be expected to show the strongest patterns of male mate choice for large swellings (Zinner et al., 2002). Domb and Pagel (2002) have replied to these criticisms, but the findings remain controversial.

Zinner et al. (2002) list three lines of evidence that are required to support the reliable indicator hypothesis and to distinguish it from alternative hypotheses: (1) differences in swelling size are enduring characteristics of individual females, and swelling size does not vary over subsequent cycles for individuals; (2) between-female variation in swelling size is related to differences in female reproductive quality, independently of covariates; and (3) males prefer to mate with females that demonstrate high reproductive quality with larger...
swellings. Here we provide an independent test of these predictions, using long-term reproductive data combined with data regarding sexual behavior and measurements of female sexual swelling characteristics, body size, and condition for a semi-free-ranging population of mandrills (*Mandrillus sphinx*). 

Mandrills are found only in the dense rainforest of central Africa (Gabon, Republic of Congo, Equatorial Guinea and Cameroon; Grubb, 1973). They are a particularly useful model to test the reliable indicator hypothesis because they are thought to travel in large, semidisaggregated bands in which males may not have direct knowledge of a female’s previous reproductive success or history (Abernethy et al., 2002). Thus, a quality indicator may be more important for this species than for other species, such as macaques or baboons, in which males and females have more stable, longer-term relationships. However, forest conditions mean that mandrills are extremely rare and females have more stable, longer-term relationships. Studies of this colony show that mandrills have typical cercopithecid matrilineal inheritance of rank, and that matrilineal ranks show very little change across time (Setchell, 1999). Breeding is moderately seasonal (with the majority of mating occurring from June–October), and females mate with more than one male per estrus cycle (Setchell, 1999). Male–male competition for access to peri-ovulatory females is intense, and the mandrill is one of the most sexually dimorphic primates. Adult males (31 kg) are 3.4 times the mass of females (Setchell et al., 2001), have upper canines measuring 44 mm (Setchell and Dixson, 2002), and possess showy secondary sexual ornamentation, including brightly colored skin on the face, rump, and genitalia. Top-ranking males mate-guard only maximally swollen females, and the alpha male sires the majority of offspring (Dixson et al., 1993; Wickings, 1995), whereas subordinate males mate sneakily, often with females that are less likely to be fertile (Setchell, 1999).

**METHODS**

**Study animals**

The CIRMF mandrill colony was established in 1983/1984, when 15 animals (seven males, eight females) were released into a 6.5-ha forest enclosure. There have been no subsequent additions to the colony, other than by breeding, although animals are occasionally removed, and in 1994 a second semifree-ranging group was established in a smaller enclosure (3.5 ha) by transferring animals from the first enclosure. During the mating season 2002, the two enclosures housed 158 animals (109 in enclosure 1, 49 in enclosure 2). The animals forage freely and receive daily supplements of monkey chow, fruit, and vegetables. Water is always available from a stream, which runs through both enclosures. This study involved the 29 females that showed estrus cycles with swellings during the 2002 mating season (June–November 2002). Subjects had an age range of 2.98–24 years (at June 2002), and a parity of 0–17 previous pregnancies.

Data on the size and shape of the maximal swelling for each cycle for each female were collected by using still photographs of females from behind and from the side, taken at close range during daily provisioning. Only photographs taken during the peri-ovulatory period (last 5 days of maximal swelling) were used for analysis. Suitable photographs were scanned into a computer and analyzed using Scion Image for Windows (available from www.scion.com). Images were scaled using one or (preferably) more of the following: the height and width of the ischial callosities, the length and width of the tail, and standing height (all measured at capture). For each swelling we recorded the length (top to bottom, rear or side view), maximum width (rear view), and maximum depth (side view). The accuracy of this method was determined by analyzing more than one photograph for each subject (maximum = 9, mean = 4 for length, mean = 2 for width and depth). Measures for an individual female did not differ by more than 5%. We also took advantage of four occasions when females were captured at maximal swelling to obtain accurate measures of swelling length, width, and depth. In each case these measures were within 5% of those obtained from photographs. Capture did not affect the swelling parameters (swellings did not deflate after capture).

As the brightness of the red-pink coloration of the sexual swelling had previously been observed to vary between females (data not shown), we also used graduated color charts (published by the Royal Horticultural Society, London) to obtain a quantitative measure of the brightness of each maximal swelling (for more details of color measurement, see Setchell, 1999; Setchell and Dixson, 2001). Color was measured by J.M.S. halfway between base of tail and the top of the vaginal opening when females were close to the observer (within 2 m).

**Female “quality”**

All females photographed were captured and anesthetized for an annual health check within 3 months of cycling (anaesthesia was accomplished by using Telinject blowpipe to deliver a syringe containing Imalgene1000 10 mg/kg). We took this opportunity to measure body mass (nearest 100 g), crown–rump length and standing height (using a metal ruler, nearest 0.5 cm), the length and maximum width of the tail, and the length and width of the ischial callosities (using callipers, nearest mm). Body condition was calculated as the body mass index (BMI; kilogram mass per squared meter crown–rump length).

Ad libitum records of agonistic and approach-avoid interactions between females during observation periods were used to confirm previous records of dominance rank (Setchell, 1999). The age and reproductive history (age at first birth, parity, mean interbirth interval, and mean number of cycles to conception) of the subject females were collated from colony records. An early age at first birth contributes to a longer reproductive life-span in females, and the mean interbirth interval measures the time between conceptions and is inversely related to the rate of offspring birth (if the infant survives). The shorter the interbirth interval, the more infants a female will produce in her lifetime and the greater her...
reproductive quality. Finally, the mean number of cycles to conception is the inverse of the probability of conceiving each cycle. Thus, a low number of cycles to conception indicates a higher probability of conception per cycle. An important additional aspect of female fitness is infant survival. However, it was not possible to investigate the relationship between swelling size and infant survival in this study, as mortality rates are very low in this semi-free-ranging population (e.g., 99% infant survival to 6 months; Setchell et al., 2002).

Male mate choice

Opportunities for behavioral data collection were limited owing to the dense nature of the forested enclosures and to variation in observability between individual females. Lengthy focal follows of individual females were not possible, precluding detailed examination of male-male aggression relating directly to tumescent females, and investigation of whether males endure higher costs to mate with females that show larger swellings. We therefore chose two simple indicators of male sexual interest in a female: mate-guarding and the presence of copulatory plugs. Mate-guarding is a readily observed behavior in mandrills, in which a male follows a female closely, maintaining proximity with her, and attempts to prevent other males from interacting with her sexually. In a classic example of priority of access (Altmann, 1962), the dominant male in a mandrill group is responsible for the majority of mate-guarding behavior, but when female swelling periods overlap, other males mate-guard females not guarded by the dominant male (Setchell, 1999; Setchell JM, unpublished observations). As there are more sexually active males than there are females cycling on any day (Setchell 1999), the occurrence of mate-guarding can be regarded as a useful measure of male sexual interest.

We noted daily whether a female was mate-guarded, and by which male, and used this information to calculate the percentage of days at maximum swelling on which a female was mate-guarded for each swelling cycle (termed mate-guarding). The presence or absence of a sperm plug (coagulated semen in the vagina) was also noted daily for each female as an indicator of a successful mating (by any male). Sperm plugs represent a zero-one indicator that any male found the female attractive and succeeded in mating with her. However, the absence of a sperm plug cannot be taken as absence of copulation, as females may remove the sperm plug after mating (Setchell JM, unpublished observations). For each swelling cycle, we calculated the percentage of days at maximum swelling during which a sperm plug was present (sperm plug presence). These measures of male interest allowed us to investigate male behavior when more than one female cycled on the same day, and to examine male choice between different females cycling simultaneously (perhaps the best time to test male preferences; Zinner et al., 2002). Mate-guarding and sperm plug presence data for the 4 days on which females were captured were discarded, as capture may have affected the quantification of male-male competition.

Statistical analyses

Each cycle was treated as an independent sample, described by female identification number, cycle number, female age, parity, social rank, body mass, body size, condition, reproductive history (details above), swelling characteristics (details above), conception, and male interest (details above). Where data were available for multiple cycles for individual females, we used Pearson’s correlations to examine intraindividual variation in swelling characteristics. This tested whether the measures of two consecutive cycles for an individual female were more similar to each other than random pairings. Correlations were positive and significant for all cycle characteristics (see Results), indicating that the swelling characteristics of an individual female were consistent over consecutive swellings, and we therefore used the mean swelling characteristics over all cycles for each female in all further analyses. Insufficient data were available to compare conception versus nonconceptive cycles within individual females, and an unpaired t-test was therefore used to compare swelling characteristics of conception versus nonconception cycles.

Partial correlations were used to examine relationships between swelling characteristics, female characteristics, female reproductive history, and male sexual interest, controlling for potentially confounding variables (female standing height, age, parity, and rank). As the covariates female standing height, age, and parity (but not rank) were collinear, we constructed a derived composite measure by using principal components analysis. This measure accounted for 79% of the variance (factor loadings: age = 0.969, parity = 0.942, height = 0.744) and was then used to control for these possible confounding variables simultaneously, without the model instability caused by collinearity in multivariate models.

Statistical tests were two-tailed, except in cases in which the reliable indicator hypothesis makes a priori predictions concerning the expected direction of a relationship, when directed tests (Rice and Gaines, 1994) were used. Directed tests provide much of the statistical power of one-sided tests but with the advantage that they can also be used to detect strong deviations in the unanticipated direction (Rice and Gaines, 1994). The critical values for each directed test were obtained from the p values from the corresponding one-tail test in SPSS for Windows release 11.00 (SPSS Inc.), using the method detailed in Rice and Gaines (1994), with γ / α = 0.8.

To test the hypothesis that the dominant male was more likely to guard the female with the largest (in any dimension, absolute swelling size, or corrected for female standing height) or brightest sexual swelling than other females on days when more than one female was peri-ovulatory, we determined the number of days on which the male guarded the largest (brightest) female available (score +), the number of days on which he guarded another female (although a female with a larger or brighter swelling was available; score −), and the number of days on which he guarded neither (matched pair). A one-tailed sign test was used to test whether significantly more + scores were obtained than − scores.

The probability of a type I statistical error (rejecting the null hypothesis when it is true) increases monotonically with the number of tests made of a hypothesis or prediction, and the significance level should accordingly be adjusted downwards to correct for the probability of observing at least one individually significant test by chance (Rice, 1989). However, applying the sequential Bonferroni correction (Rice, 1989) to our tables of results did not change our conclusions, and we therefore present the unadjusted significance levels.

RESULTS

Intraindividual variation in swelling characteristics

Female sexual swellings measured an average of 12.7 ± 0.3 cm long (n = 38), by 9.3 ± 0.3 cm wide (n = 38) and 9.3 ± 0.3 cm deep (n = 34). Of the 29 subjects, seven females underwent multiple swelling cycles during the cycling period. Three females showed two swelling periods, three females showed three swelling periods, and one female showed six swelling periods. The size and brightness of one sexual swelling were highly significantly correlated with the characteristics of the subsequent swelling (length: r = .88, p = .008, n = 7; width:
Results of unpaired t tests comparing swelling characteristics between conceptive and nonconceptive cycles

<table>
<thead>
<tr>
<th>Swelling character</th>
<th>Conceptive cycle</th>
<th>Nonconceptive cycle</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height (cm)</td>
<td>12.2 ± 0.4 (23)</td>
<td>13.0 ± 0.6 (8)</td>
<td>0.99</td>
</tr>
<tr>
<td>Width (cm)</td>
<td>9.1 ± 0.3 (20)</td>
<td>10.0 ± 0.5 (8)</td>
<td>1.54</td>
</tr>
<tr>
<td>Depth (cm)</td>
<td>7.3 ± 0.4 (23)</td>
<td>8.1 ± 0.7 (8)</td>
<td>0.90</td>
</tr>
<tr>
<td>Brightness</td>
<td>87.5 ± 2.0 (19)</td>
<td>89.4 ± 1.0 (10)</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Values are mean ± SEM (n). All two-tailed comparisons were nonsignificant.

$r = .761$, $p = .047$, $n = 7$; depth: $r = .915$, $p = .011$, $n = 6$; brightness: $r = .967$, $p < .001$, $n = 7$; using the first two swelling cycles for females that cycled more than twice, implying that swelling characteristics were consistent characteristics of individual females, at least during one mating season. Furthermore, swelling characteristics did not vary by more than 6% between cycles for individual females (the mean magnitude of change in swelling characteristics from swelling to swelling within a female was 4% for swelling height, 5% for depth, 3% for width, and 4% for brightness).

Of the seven females that cycled more than once, four failed to conceive at all. Of the other 22 females sampled, 20 conceived and two did not. Although there were insufficient data to allow a paired t test between conceptive and nonconceptive cycles in the same female, an unpaired t test showed no significant difference in swelling characteristics between conceptive and nonconceptive cycles in general (Table 1).

Swelling characteristics and female age, rank, parity, standing height, and BMI

The three measures of swelling size were highly correlated with one another (length versus width: $r = .564$, $p < .001$, $n = 36$; length versus depth: $r = .553$, $p < .001$, $n = 38$; width versus depth: $r = .560$, $p < .001$, $n = 35$). Standing height varied from 36.0–48.5 cm in this sample of females (mean ± SEM = 41.6 ± 0.6 cm), and was positively and significantly linearly related to all three measures of swelling size (Table 2). Correlations between swelling characteristics and female age, rank, parity, and BMI are shown in Table 2. None of the three measures of swelling size was significantly related to female age or condition (BMI). Higher-ranking females had shallower swellings than did lower-ranking females, but neither swelling length nor width was significantly related to female rank. Nulliparous females had shallower swellings than did parous females, in all three dimensions. Swelling brightness decreased significantly with increasing female age and also decreased with increasing body condition, but was not related to female rank or parity. When standing height was controlled, neither swelling length nor width was significantly related to female age, rank, parity, or BMI (Table 2). Swelling depth was significantly negatively related to female rank, with lower-ranking females having deeper swellings for their height than did higher-ranking females.

Correlation coefficients between swelling characteristics and female standing height, age, parity, rank, and BMI

<table>
<thead>
<tr>
<th>Swelling character</th>
<th>Standing height</th>
<th>Age</th>
<th>Parity*</th>
<th>Rank</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length (n = 29)</td>
<td>0.373*</td>
<td>0.049</td>
<td>0.373*</td>
<td>0.195</td>
<td>0.195</td>
</tr>
<tr>
<td>Width (n = 29)</td>
<td>0.440*</td>
<td>0.101</td>
<td>0.452*</td>
<td>0.308</td>
<td>0.311</td>
</tr>
<tr>
<td>Depth (n = 26)</td>
<td>0.478**</td>
<td>-0.219</td>
<td>0.445*</td>
<td>0.369*</td>
<td>0.338</td>
</tr>
<tr>
<td>Brightness (n = 29)</td>
<td>0.270</td>
<td>0.494*</td>
<td>0.116</td>
<td>0.140</td>
<td>-0.411*</td>
</tr>
<tr>
<td>Controlling for standing height</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length (n = 29)</td>
<td>-</td>
<td>0.163</td>
<td>0.117</td>
<td>-0.263</td>
<td>0.034</td>
</tr>
<tr>
<td>Width (n = 29)</td>
<td>-0.102</td>
<td>0.026</td>
<td>-0.109</td>
<td>0.081</td>
<td></td>
</tr>
<tr>
<td>Depth (n = 26)</td>
<td>-0.162</td>
<td>-0.237</td>
<td>-0.436*</td>
<td>0.126</td>
<td></td>
</tr>
</tbody>
</table>

* Nulliparous females versus parous females.

Brightness was not corrected for female height, as the two variables were not significantly correlated.

$p < .05$, ** $p < .01$, two-tailed tests.

Swelling characteristics and female “fitness”

Female “fitness” estimators were very variable in this population. The age at first birth for the females studied varied between 2.70–5.65 years (mean ± SEM = 4.51 ± 0.15 years), the interbirth interval varied between 427–648 days (mean ± SEM = 526 ± 14 days), and the mean number of cycles to conception from 1–5 (mean ± SEM = 1.4 ± 0.1). Correlations between measures of female fitness and swelling characteristics are shown in Table 3. Females with larger swellings did not begin to produce offspring at an earlier age, did not have shorter mean interbirth intervals, and did not show fewer cycles to conception. Controlling for the influence of female height, age, rank, and parity (using the composite measure derived using principal components analysis) did not change the significance of these results (Table 3). Swelling brightness was not significantly related to female fitness estimators.

Male sexual interest and female swelling characteristics

Females were maximally swollen for 10.8 ± 1.1 days (mean ± SEM, $n = 45$ cycles) during a cycle. Males mate-guarded on at least one of these days during 38 of 45 female cycles (84%). Males did not mate-guard nonmaximally swollen females. The mean percentage of days at maximal swelling on which a female was mate-guarded was 47 ± 5% (range = 0–100%), and females were guarded for 56 ± 5% of peri-ovulatory days (range = 0–
Table 4
Correlation coefficients between male sexual interest and swelling characteristics

<table>
<thead>
<tr>
<th>Male interest measure</th>
<th>Length (n = 27)</th>
<th>Width (n = 24)</th>
<th>Depth (n = 27)</th>
<th>Brightness (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% days mate-guarded</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>-.241</td>
<td>-.327</td>
<td>-.345</td>
<td>-.003</td>
</tr>
<tr>
<td>r_p</td>
<td>-.239</td>
<td>-.452</td>
<td>-.359</td>
<td>.112</td>
</tr>
<tr>
<td>% days with sperm plug</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>.308</td>
<td>.438*</td>
<td>.155</td>
<td>-.101</td>
</tr>
<tr>
<td>r_p</td>
<td>.345</td>
<td>.360</td>
<td>.111</td>
<td>.141</td>
</tr>
</tbody>
</table>

* r_p is partial correlation coefficient controlling for covariates female age, parity, standing height, and rank using a composite measure derived using principal components analysis.

* p < .05, using directed tests.

100%). Correlations between mate-guarding and female swelling characteristics are shown in Table 4. Mate-guarding was not significantly related to any swelling characteristics, but when female age, parity, standing height, and rank were controlled (using the composite measure derived using principal components analysis), males showed a (nonsignificant) tendency to mate-guard females with narrower and shallower swellings more than those with wider and deeper swellings. Mate-guarding was not significantly related to swelling length or brightness when female age, parity, standing height, and rank were controlled (Table 4), but all swelling measures were negatively correlated with percentage of days guarded.

Sperm plugs were observed for maximally swollen females on 137 of a total 384 days (36%), and during 40 of 45 cycles (89%). Maximally swollen females had sperm plugs on a mean of 50 ± 6% of days (range = 0–100%), and on 51 ± 6% of peri-ovulatory days (range = 0–100%). Sperm plugs were occasionally (n = 3) observed in females that were swelling but not yet maximally swollen, but never occurred in nontumescent females. Correlations between sperm plug presence and female swelling characteristics are shown in Table 4. Sperm plug presence correlated significantly positively with swelling width; females with wider swellings were more likely to have a sperm plug. The positive correlation between the presence of a sperm plug and swelling length also approached significance; there was a tendency for females with longer swellings to have more sperm plugs. Sperm plug presence was not significantly related to swelling depth or brightness, and when female age, parity, standing height, and rank were controlled (using the composite measure derived using principal components analysis), there were no significant relationships between swelling characteristics and sperm plug presence (Table 4).

Neither measure of male interest was significantly related to female age, rank, or parity, nor were measures of male sexual interest directly related to female fitness estimators, height, or BMI (Table 5).

Male sexual interest did not vary predictably across cycles for females that showed more than one swelling period, either when measured as mate-guarding (first swelling period: 34 ± 13%, versus subsequent swelling periods: 49 ± 12%, t0 = 0.84, p = .425), or as sperm plug presence (first swelling period: 46 ± 9%, versus subsequent swelling periods: 47 ± 12%, t0 = 0.98, p = .924). Nor was male sexual interest related to the conceptive nature of a cycle, when measured as mate-guarding (conceptive cycles: 50 ± 1% [n = 18], versus nonconceptive cycles: 61 ± 13% [n = 10], t0 = 0.94, p = .358), or as sperm plug presence (conceptive cycles: 62 ± 12% [n = 18], versus nonconceptive cycles: 68 ± 1% [n = 10], t0 = 0.57, p = .575).

Overlapping swelling cycles

If swelling size (or brightness) indicates quality, then we predicted that where multiple females are peri-ovulatory and thus potentially fertile, the dominant male should choose to mate-guard the female with the largest (or brightest) swelling. To test this prediction, we examined days on which more than one female was peri-ovulatory (days of overlap). Overall, in the two groups, two females were peri-ovulatory on the same day for 26 days, three females were peri-ovulatory on 15 days, and four females were peri-ovulatory on four days. Whether a cycle overlapped with that of another female did not influence whether the female was mate-guarded by males or whether she had a sperm plug. There was no significant difference between females whose peri-ovulatory periods overlapped and those whose peri-ovulatory periods did not overlap in percentage of mate-guarding (overlapped females: 56 ± 1% [n = 28], nonoverlapped females: 51 ± 1% [n = 10], t0 = 0.51, p = .613) or sperm plug presence (overlapped: 48 ± 1%, nonoverlapped: 60 ± 1%, t0 = 1.05, p = .300).

On days when more than one female was peri-ovulatory, the dominant male was not more likely to guard the female with the largest (in any dimension, absolute swelling size, or corrected for female standing height) or brightest sexual swelling than he was to guard any other female (Table 6). These data suffer from pseudo-replication, when the same pair of females overlapped on consecutive days, but using only one data point from such dyads yielded the same result: males did not guard females with larger or brighter swellings significantly more than other females.

DISCUSSION

The aim of the present study was to test the hypothesis that differences in sexual swelling characteristics between female mandrills are reliable indicators of female reproductive quality, and that males prefer to mate with females possessing swellings that advertise higher reproductive quality (the reliable indicator hypothesis; Pagel, 1994). Here we examine whether our

Table 5
Correlation coefficients between male sexual interest and female characteristics

<table>
<thead>
<tr>
<th>Male interest measure</th>
<th>Age (n = 27)</th>
<th>Parity (n = 27)</th>
<th>Standing height (n = 27)</th>
<th>Rank (n = 27)</th>
<th>BMI (n = 27)</th>
<th>Age first birth (n = 27)</th>
<th>Mean inter-birth interval (n = 27)</th>
<th>Mean cycles to conception (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% days mate-guarded</td>
<td>0.246</td>
<td>0.192</td>
<td>-.044</td>
<td>-.509</td>
<td>0.177</td>
<td>-0.002</td>
<td>-0.048</td>
<td>-0.276</td>
</tr>
<tr>
<td>% days with sperm plug</td>
<td>0.308</td>
<td>0.283</td>
<td>0.178</td>
<td>0.134</td>
<td>0.137</td>
<td>-0.079</td>
<td>-0.037</td>
<td>0.042</td>
</tr>
</tbody>
</table>

All correlations were nonsignificant (p > .05) using two-tailed tests.
Table 6  
Female swelling characteristics and mate-guarding by the dominant male

<table>
<thead>
<tr>
<th>Swelling character</th>
<th>Number of days male mate-guarded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female with greatest measure (score +)</td>
</tr>
<tr>
<td>Height</td>
<td>12</td>
</tr>
<tr>
<td>Width</td>
<td>15</td>
</tr>
<tr>
<td>Depth</td>
<td>16</td>
</tr>
<tr>
<td>Brightness</td>
<td>14</td>
</tr>
</tbody>
</table>

- Results of one-tailed sign tests comparing the number of days on which the male guarded the largest (brightest) female available, and the number of days on which he guarded another female, with a zero which the male guarded the largest (brightest) female available, and with a zero if he guarded neither. All tests were nonsignificant.

- Controlling for female size results support the three predictions listed by Zinner et al. (2002), which are required to support the reliable indicator hypothesis and to distinguish it from alternative hypotheses.

Do differences between swellings endure characteristics of individual females?

If swelling characteristics advertise the reproductive quality of a female, then swelling characteristics should remain consistent over subsequent cycles for that female. Our results suggest that swelling characteristics were enduring features of individual female mandrills, at least during one mating season, supporting the prediction of the reliable indicator hypothesis. These data are contrary to the prediction of the alternative hypothesis that swelling size indicates cycle-to-cycle variability in the probability that an individual female conceives (Zinner et al., 2002). However, most females in the present study conceived first cycle, and we had few repeated measures of cycle characteristics for individual females, precluding an analysis of whether within-female variation in swelling characteristics is smaller than between-female differences, which is necessary to reject this prediction conclusively. Our data should therefore be regarded as preliminary.

Do swelling characteristics indicate aspects of female reproductive quality?

If sexual swellings act as reliable indicators of female reproductive quality, then swelling size (or color) should be related to fitness estimators. Our results do not support this prediction. Neither age at first birth, mean interbirth interval, nor the number of cycles to conception was significantly correlated with swelling characteristics. Controlling for female age, rank, parity, and standing height did not change the significance of these results. These results differ with respect to those reported for wild baboons. Domb and Pagel (2001) showed a positive correlation between swelling height (but not width or depth) and female reproductive history in their study of baboons. Females with larger (longer) swellings had an earlier age at first birth and produced more offspring per year, more surviving offspring per year, and more surviving offspring overall. These correlations were still significant when the effects of female age and rank were controlled. However, Zinner et al. (2002) have shown that when female standing height and group membership (Domb and Pagel pooled data for five baboon groups) are controlled, swelling length is no longer significantly related to female fitness estimators in the baboons studied. Contra the reliable indicator hypothesis, a review of the literature suggests that the least fertile females often have larger swellings (adolescent females: Anderson and Bie!ert, 1984; females that cycle repeatedly; Nunn et al., 2001), although this was not the case in the mandrills studied here.

Swelling size and color did (at least partly) yield information concerning female rank, parity, and body condition (BMI), providing a potential indicator of female reproductive status. Fecundity in female mandrills increases with rank and parity (Seitchell et al., 2002), and higher-ranking females also produce heavier offspring (Seitchell et al., 2001). If swelling characteristics indicate quality, therefore, high-ranking, multiparous females should have the largest and/or brightest swellings. Such information might be useful to wild male mandrills, who may not necessarily be familiar with a female’s age or rank. However, only swelling depth was significantly related to female rank, and this relationship was negative, with high-ranking females having smaller swellings than did low-ranking females. All swelling dimensions were smaller in nulliparous females than in parous females, but this relationship was explained by the shorter standing height of nulliparous females, who are still growing to adult size (it is possible, therefore, that males might use swelling size rather than female size to judge parity and preferentially allocate mating effort to parous females). Finally, contrary to prediction, females that were in better condition (higher BMI) had duller colored swellings.

Do males prefer to mate with females with larger (or brighter) swellings?

If swelling characteristics advertise female reproductive quality, then males should show more interest in larger (or brighter) swellings. Our data provided limited support for this prediction. The presence of sperm plugs (but not mate-guarding) was related to swelling size; females with wider swellings were more likely to have a sperm plug, and there was a tendency for females with longer swellings to have sperm plugs on a greater proportion of maximal swelling days. However, these relationships were not significant when female standing height was controlled. Moreover, relationships between mate-guarding and swelling size were negative when female age, rank, parity, and standing height were controlled, contrary to the prediction of the reliable indicator hypothesis. Finally, when more than one female was maximally swollen, the alpha male (who has free choice) did not mate-guard females with larger or brighter swellings more than he guarded other females.

Two studies of baboons provide more support for the prediction that males prefer larger swellings than this study of mandrills. Captive experiments show that male baboons masturbate more in response to females with larger, “supranormal” swellings than they do to females with normal swellings (Bie!ert and Anderson, 1985; Girolami and Bie!ert, 1987). Secondly, Domb and Pagel (2001) found that male baboons use female swelling size to determine mating effort, competing more to mate with females possessing longer swellings. However, in the present study of mandrills, only swelling width (one of three measures of size) explained aspects of male competition.
We found no direct relationship between male sexual interest and female reproductive quality, age, rank, or parity. Under natural circumstances males are unlikely to have detailed knowledge of the reproductive history of each individual female on which to base mating decisions. Domb and Pagel (2001) suggest that the immigrant nature of male baboons selects for signals that reliably advertise a female’s fitness in the absence of long-term knowledge of individual females. This does not appear to be the case, however, for mandrills.

The reliable indicator hypothesis

Sexual swelling characteristics were maintained across consecutive cycles in individual female mandrills, and there was some evidence that males mated selectively as a function of swelling characteristics; females with wider swellings were more likely to have sperm plugs than females with narrower swellings. However, the occurrence of mate-guarding was, if anything, negatively related to swelling size, and we found no support for the prediction that swelling characteristics advertise female fitness; the size and color of a female’s swelling was not an indicator of her reproductive quality. On the basis of this evidence, we conclude that the reliable indicator hypothesis does not explain variation in mandrill sexual swellings in this semi-free-ranging colony. Furthermore, the reliable indicator hypothesis, as proposed by Pagel (1994), implies that females compete for mates. There is little evidence for female-female competition in mandrills. Breeding is only moderately seasonal (Setchell et al., 2002) and female peri-ovulatory periods do not overlap extensively, males are not limiting, and females do not interfere in mating attempts by other females (unpublished observations). Finally, comparative tests also fail to support the reliable indicator hypothesis, as sexual swellings are not associated with increased female mating competition (measured using the adult sex ratio, female canine size, and expected female mating synchrony; Nunn et al., 2001).

It could be argued that the potential differences in the demography and feeding ecology of this semiwild population, compared with wild mandrills, limit the interpretation of evolutionary patterns. Provisioning leads to access to better and/or more food, which may be more evenly distributed among individual females and may result in the artificial equalizing of differences in female fecundity and swelling characteristics. However, females in this colony do show large variation in reproductive fitness (this study, Setchell et al., 2002). Furthermore, provisioning is provided as a clustered resource, and food is (if anything) less likely to be evenly distributed between individuals than in the wild. However, the possibility remains that in this captive situation, poor-quality females may receive sufficient resources to produce large swellings, thus confounding any relationships between swelling size, female fitness estimators, and male sexual interest.

The evolution of sexual swellings: alternative hypotheses

Why do female mandrills (and females of many other Old World primate species) advertise impending ovulation with exaggerated sexual swellings? Morphological advertisement of receptivity is not necessary for females to attract mates; male-male competition over matings is inevitable in multimale, multifemale species, not all of which show exaggerated sexual swellings (Pagel, 1994). Hypotheses proposed to account for the function of sexual swellings include both obvious ovulation and paternity confidence (Hamilton, 1984) and, conversely, concealed ovulation and paternity confusion (Hrdy, 1981). The obvious ovulation hypothesis (Hamilton, 1984) holds that swellings advertise the exact timing of ovulation females, in order to mate with a single high-quality male, leading to paternal investment from a male that is certain of paternity. The “best-male” hypothesis (Clutton-Brock and Harvey, 1976; Hrdy and Whitten, 1987)—that females advertise ovulation to incite competition between males, and mate with the “winner,” thus gaining “good genes” for their offspring—makes the same prediction, that females should mate with only one male. By contrast, the concealed ovulation hypothesis predicts that females advertise in order to mate with many males, distributing paternity chances across many, otherwise potentially infanticidal, males (Hrdy, 1981). The postcopulatory selection hypothesis (Harvey and May, 1989) also predicts that female swellings attract many males, and that females mate polyandrously, in order to promote sperm competition and female choice. Stallmann and Froehlich (2000) draw attention to the fact that only in the reliable indicator hypotheses do the benefits of female sexual swellings accrue to males, via choice for reproductively superior females. The remaining hypotheses involve benefits to female fitness and are therefore not necessarily alternatives to the reliable indicator hypothesis. Moreover, other hypotheses do not necessarily account for the variation in swelling size observed between females. Finally, it is important to note that possession of a swelling may have evolved for one purpose, whereas variation in swelling measures may have a different function.

The evidence in support of the reliable indicator hypothesis presented by Domb and Pagel (2001), that female’s swelling characteristics correlate positively with female life-history variables, assumes lifetime consistency of female fitness and signaling. Thus, a male mating with a female that possesses a large swelling has chosen a female that had an early age at first reproduction, and has short interbirth intervals. Such fitness variables are important to the male if they are inherited by his offspring. However, from the male’s perspective, a female’s current fitness, namely, whether she will conceive now, seems more important than her lifetime measures. Nunn (1999) has reviewed features of sexual swellings, observing that they increase gradually in size and that ovulation is most likely at peak swelling, but this association is not perfect. Species with sexual swellings tend to have longer follicular phases and dominant males mate at peak swelling when females are most likely to ovulate, but other males also mate at times when ovulation is less likely but still possible. Showing that although several of the existing hypotheses explained some of these features, no one of the hypotheses could account for all of them, Nunn (1999) combined elements of paternity confidence and paternity confusion hypotheses “graded signal” hypothesis (Martin, 1992), whereby female sexual swellings act to both bias and confuse paternity, representing the probability of ovulation, without advertising its exact timing. Females thus manipulate male behavior, such that dominant males mate-guard at the time when the female is most likely to ovulate, but other males mate at suboptimal times, when there is a smaller (but greater than zero) probability of sirehood. In this way females protect their infants against infanticidal males, by giving each male that mates a possibility of paternity while concentrating paternity to desirable mates, and solve the “female dilemma” (van Schaik et al., 1999).

The graded signal hypothesis (Nunn, 1999) is a potential explanation for the function of sexual swellings in mandrills. Female mandrills mate with multiple males during a swelling cycle (Setchell, 1999), potentially confusing paternity. Genetic analyses of paternity in the CIRMF colony show that the dominant male sires 80–100% of the offspring in a mating season (Dixson et al., 1993; Wickings, 1995). Thus, dominant male mandrills achieve a high concentration of paternity, but the probability of paternity for other males is greater than zero, as predicted by the graded-signal hypothesis. This hypothesis...
also predicts that female preferences should change over the swelling cycle, and preliminary evidence shows that female mandrills do indeed show mate choice for the dominant male when peri-ovulatory (Setchell, 2002). Further studies, combining the study of mating behavior, swelling morphology, and endocrine analysis to determine whether swellings are indeed a probabilistic signal of ovulation will be needed to elucidate the function of sexual swellings (see Aujard et al., 1998; Reichert et al., 2002). Finally, a limitation of any semi-free-ranging or free-ranging study of sexual swellings are the multiple signals given by a swollen female that potentially influence male behavior. For example, in addition to visual signals, male mandrills both touch and smell sexual swellings, receiving both tactile (turgidity) and olfactory information. Our observations show that males may show great sexual interest in a female one day, but not even approach the same animal on the following day, although no difference in the sexual swelling is discernible to the observer. Captive mate choice experiments may be required to resolve the question of the function of sexual swellings (see Snowdon, 2003).

We are grateful to the CIRMF for permission to conduct this study and to past and present staff at the Centre de Primatologie for maintaining the colony and keeping records of female reproductive history. We thank Phyllis Lee and Dietmar Zinner for useful comments on the design of the study, and Dietmar Zinner, Ann Maclarnon, Melissa Gerald, and two anonymous reviewers for helpful comments on the manuscript.

REFERENCES


