Can fertility signals lead to quality signals? Insights from the evolution of primate sexual swellings

Elise Huchard¹,²,* , Alexandre Courtiol¹, Julio A. Benavides¹, Leslie A. Knapp³, Michel Raymond¹ and Guy Cowlishaw²

¹CNRS-Institut des Sciences de l’Evolution, Université Montpellier 2, Place Eugène Bataillon, CC 065, 34 095 Montpellier Cedex 5, France
²Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK
³Department of Biological Anthropology, University of Cambridge, Downing Street, Cambridge CB2 3DZ, UK

The sexual swellings of female primates have generated a great deal of interest in evolutionary biology. Two hypotheses recently proposed to elucidate their functional significance argue that maximal swelling size advertises either female fertility within a cycle or female quality across cycles. Published evidence favours the first hypothesis, and further indicates that larger swellings advertise higher fertility between cycles. If so, a male preference for large swellings might evolve, driving females to use swellings as quality indicators, as proposed by the second hypothesis. In this paper, we explore this possibility using a combination of empirical field data and mathematical modelling. We first test and find support for three key predictions of the female-quality hypothesis in wild chacma baboons (Papio ursinus): (i) inter-individual differences in swelling size are maintained across consecutive cycles, (ii) females in better condition have larger swellings and higher reproductive success, and (iii) males preferentially choose females with large swellings. We then develop an individual-based simulation model that indicates that females producing larger swellings can achieve higher mating success even when female–female competition is low and within-female variance in the trait is high. Taken together, our findings show that once sexual swellings have evolved as fertility signals, they might, in certain socio-sexual systems, be further selected to act as quality signals. These results, by reconciling two hypotheses, help to clarify the processes underlying sexual swelling evolution. More generally, our findings suggest that mate choice for direct benefits (fertility) can lead to indirect benefits (good genes).

Keywords: sexual swellings; honest signals; mating synchrony; primates; fertility indicators; mate choice

1. INTRODUCTION
The sexual swellings of female primates are a classic signal in sexual selection. Exaggerated sexual swellings are found in a number of old world monkeys and apes (Rowell 1972; Dixson 1983). They are produced when the anogenital skin gradually swells during the oestrous cycle, reaching its maximal size around the time of ovulation before subsequently returning to its non-swollen state. The enduring fascination of sexual swellings has led to several plausible hypotheses about how they might have evolved (reviewed by Nunn 1999; Zinner et al. 2004). Yet the functional significance of this signal still remains highly contentious.

The current controversy revolves around two competing explanations. On the one hand, the graded-signal hypothesis proposes that sexual swellings signal fertility variation within females (larger swellings indicate higher probability of ovulation). Swellings thereby encourage dominant males to mate-guard females when they are at their most fertile period, increasing paternity certainty and thus the payoffs of future paternal care, but also allow other males to mate at other times (when females are not guarded), creating paternity confusion and thus reducing the risk of infanticide (Nunn 1999). On the other hand, the reliable-indicator hypothesis states that sexual swellings signal heritable quality variation between females (larger swellings indicate higher quality). Swellings are therefore found under conditions of reversed sexual selection, where males have become the choosier sex due to the high mating costs experienced in multimale social groups (where swellings are characteristically found; Clutton-Brock & Harvey 1976) and females compete for the best males by honestly advertising their quality through costly signals (Pagel 1994). This hypothesis thus brings sexual swellings into the mainstream paradigm of condition-dependent signals and good-genes theories of sexual selection (e. g. Zahavi 1975; Grafen 1990; Johnstone & Grafen 1992; Cotton et al. 2006).

The current empirical evidence favours the graded-signal hypothesis. A growing number of studies have shown that sexual signals are a probabilistic signal of ovulation in a wide range of species (barbary macaques: Brauch et al. 2007; chimpanzees: Emery & Whitten 2003, Deschner et al. 2004; yellow baboons: Gesquiere et al. 2007; olive baboons: Higham et al. 2008) and that the
proximity of ovulation is a primary cue in male mating decisions. Thus, alpha males mostly consort with females in the maximal-fertility window of their cycle (e.g. chimpanzees: Deschner et al. 2004; yellow baboons: Gesquiere et al. 2007). By contrast, the empirical evidence for the reliable-indicator hypothesis is limited. Only one study has tested all the main predictions in a wild primate population (Domb & Pagel 2001). This study found support for sexual swellings as reliable indicators in wild olive baboons, but has subsequently been criticized on methodological grounds (Zinner et al. 2002). More recently, results from additional studies examining some predictions of the reliable-indicator hypothesis did not find support for it (e.g. chimpanzees: Deschner et al. 2004; olive baboons: Higham et al. 2008; mandrills: Setchell & Wickings 2004; barbary macaques: Möhle et al. 2005).

Nevertheless, recent studies have also indicated that larger swellings advertise higher fertility between different cycles of the same female (chimpanzees: Emery & Whitten 2003, Deschner et al. 2004; yellow baboons: Gesquiere et al. 2007; olive baboons: Higham et al. 2008). This may have important implications for the reliable-indicator hypothesis. Specifically, it raises the possibility that males may use this cue to make individual or cycle-to-cycle mating decisions when two or more females are close to ovulation (e.g. Gesquiere et al. 2007). Thus, once males have used the graded nature of the signal to identify that two or more females are equally close to ovulation, they could then choose the female with the largest swelling to maximize the probability of conception. Alternatively, males might use previous experience with a given female to judge how differences in her swelling size might relate to differences in her fertility. However, the critical period in the reproductive life of most male primates is when they achieve alpha-male status, which in species showing male-biased dispersal (including macaques, baboons and mandrills) typically occurs immediately after immigration into a new troop (van Noordwijk & van Schaik 2004). In this case, prior knowledge of females is not available and males only have access to information on current swelling size. Crucially, for the reliable-indicator hypothesis, a male preference for large swellings should in turn select females to produce the largest possible swellings. Otherwise, females may risk losingcopulations with the dominant male to other females, and failure to mate with him would not only lead to the loss of several possible direct and indirect benefits (e.g. good genes, infant protection services) but also increase the risk of subsequent infanticidal attacks on the offspring he has not fathered. Under these conditions, females should produce the largest swellings they can afford to, leading, if large swellings are costly and heritable, to the evolution of condition-dependent signals as predicted by the reliable-indicator hypothesis.

Beyond the empirical evidence outlined above, support for the graded-signal hypothesis also reflects comparative patterns across species and wider theoretical concerns about the reliable-indicator hypothesis (Nunn 1999; van Schaik et al. 1999; Nunn et al. 2001; Zinner et al. 2004). Two key issues have been raised. First, high within-female variance in swelling size, relative to between-female variance, may impede the evolution of sexual swellings as reliable indicators of female quality. It has been argued that reliable indicators could only evolve where inter-individual variance in swelling size exceeds intra-individual variance; otherwise, males might choose a low-quality female that produces a larger swelling because she is in a late cycle or a high-quality female producing a smaller swelling in an early cycle (Zinner et al. 2004). However, it is possible that as long as inter-individual variance exists, selection could discriminate between individuals regardless of the extent of intra-individual variance. Second, the distribution of sexual swellings across species is not associated with simple indices of female–female competition (Nunn et al. 2001). Thus, sexual swellings occur even where female competition is apparently low, contrary to the reliable-indicator hypothesis. However, if, as we propose, sexual swellings might secondarily evolve as reliable indicators from graded signals, we might rather expect the comparative evidence to support the graded-signal hypothesis. Clearly, both of these issues are important and need to be resolved, but they are difficult to address on the basis of field data or comparative analysis alone.

We explore our proposed scenario for the sequential evolution of sexual swellings, from fertility signals to quality signals, in two ways. First, we present a new field test of the reliable-indicator hypothesis, focusing on three critical predictions (Zinner et al. 2002): that between-female differences in swelling size are maintained across consecutive cycles; that good-quality females produce larger swellings; and that males preferentially choose females with larger swellings. Second, we develop an individual-based simulation model to explore whether a signal of fertility can lead to a signal of quality through male preferences for absolute swelling size. Our model investigates conditions of high within-female variance (relative to between-female variance) in the trait, and low female–female competition (either because females live in small groups and/or do not cycle synchronously). We conduct our field test in chacma baboons (P. ursinus) and parametrize our model using values of reproductive traits that are representative from several baboon populations in the Papio species complex. Sexual swellings are already known to advertise probability of ovulation as a graded signal in two of these taxa (yellow baboons: Gesquiere et al. 2007; olive baboons: Higham et al. 2008).

2. MATERIAL AND METHODS
(a) Study site and subjects
Data were collected from a wild population of chacma baboons living on the edge of the Namib Desert in Namibia, at Tsaoibis Leopard Park (for details of the site and population, see Cowlishaw 1999). Eleven focal females (six multiparous, two primiparous and three nulliparous) from two habituated multimale troops (respectively 32 and 57 individuals) were followed over one to three oestrous cycles, with a total of 21 cycles recorded, to obtain data on sexual swellings and related patterns of behaviour. The cycle number was available for 18 cycles (11 females), including 9 first cycles, 7 second cycles and 2 third cycles. Data on age, body condition and reproductive success were also collected from a further 49 adult females, comprising another 13 females (all the remaining non-cycling females) in the two study troops, plus 36 females drawn from another four neighbouring troops.

Proc. R. Soc. B (2009)
(b) Swelling size estimation

Data on sexual swellings and behaviour were collected during dawn-to-dusk follows of both study troops carried out on foot from July 2006 to January 2007. Swelling status was assessed on a daily basis for all cycling females. The maximally swollen period was assessed visually by comparing high-resolution digital photographs taken throughout the cycle (using a Canon Eos 20D camera from behind the animal). The distance separating the baboon from the photographer was measured using a rangefinder (Bushnell) with 1 m accuracy. An average of 47 photographs was taken per cycle per female (within the same day of the maximally swollen period) using a wide range of distances (8–40 m). Four or five suitable photos (i.e. with no angle and no obstruction by the tail) per cycle were selected for subsequent analysis. The contour of the swelling was manually extracted using a semi-automatic image processor implemented in MATLAB v. 7.0. Swelling area was determined through a calibration equation, which was empirically established for the camera using a constant zoom (for further details, see Caillaud et al. 2007). The mean standard error of swelling area between pictures from the same cycle was 11.3 cm$^2$, which corresponds to a measurement error of 7.6 per cent. Maximal swelling area was then estimated for each cycle of each female by averaging across these images. In addition, all females photographed were captured and anaesthetized within three months of cycling (in October 2006), which provided an opportunity to compare swelling sizes estimated through the rangefinder with direct measurements for four females which were swollen during capture. In each case, these measures ranged within 4 per cent of those obtained from field photographs. Swelling depth, highly correlated with swelling area (Spearman’s test: $r_s = 0.80$, $n=11$, $p<0.01$), was not separately analysed.

(c) Individual traits

Individual dominance rank was established for all focal females according to ad libitum or focal observations of agonistic and approach–avoid interactions among adult individuals in a troop. Dominance rank calculations indicated a high degree of linearity in females of both study troops using Landau’s linearity index corrected for unknown relationships ($k'$; de Vries 1995) and tested by a $\chi^2$-test (Appleye 1983) using MATMAN software (de Vries et al. 1993): $k' = 0.95$, $\chi^2 = 63.63$, d.f. = 20, $p<10^{-4}$ for the smaller troop and $k' = 0.67$, $\chi^2 = 93.55$, d.f. = 25, $p<10^{-4}$ for the larger troop. Age and body condition were determined for the 11 focal females, plus further 49 adult females, during troop captures. Age was estimated (as a continuous variable in years) through dental records: tooth eruption schedules for wild baboons were used to assign age up to the eruption of the molars (Kahumbu & Eley 1991), while age beyond this point was estimated on the basis of molar wear. Subsequent validation of this approach, using individuals that had been captured on multiple occasions ($n=19$ over periods of 1–5 years) for a comparison of estimated versus known age differences between captures, indicates that these estimates are robust (the mean difference between observed and estimated elapsed time periods does not differ significantly from zero: one-sample $t$-test, $p>0.05$; G. Cowlishaw 2006, unpublished data). On this basis, the focal females had an age range of 6–16 years in October 2006.

To determine body condition, the baboons were weighed and measured. Among the focal females, body mass ranged from 11.1 to 17.6 kg (mean = 13.8 ± 0.2 kg) and crown–rump length ranged from 56.0 to 66.3 cm (mean = 62.9 ± 4.0 cm). Body condition was calculated as the residuals of body mass plotted against crown–rump length (i.e. Schulte-Hostedde et al. 2005) in a general linear model that included all captured females and controlled for the effects of reproductive state at the time of capture. As such, this body condition index reflects variations in total body fat and muscle for a given body size.

Female reproductive performance was assessed by genotyping all animals across the six captured troops ($n=199$ individuals) for 16 microsatellite loci and the major histocompatibility complex (MHC) class II DR region. Maternities were inferred for 97 offspring using a combination of behavioural observations and of parentage inference based on microsatellites and Mhc-DRB using a strict (95%) and a relaxed (80%) confidence level (see the electronic supplementary material for details regarding genotyping and maternity analysis). This allowed an estimate of the number of surviving offspring for a total of 60 adult females, including the 11 females with swelling data, with the number of such offspring ranging from 0 to 6 (mean = 1.60 ± 0.02).

(d) Male preference assessment

Male preferences were assessed by the expression of mate-guarding behaviour (consortship): a costly male strategy (Alberts et al. 1996). Mate guarding was monitored through focal observations, which were carried out on a daily basis for swollen females. Any changes in patterns of consortship were recorded ad libitum if they occurred outside focal sampling (visibility conditions are excellent at Tsaobis). Male–male aggression over females was not recorded because, during periods of social stability (as seen in both groups during the present study), the dominant male rarely encounters aggression from other males when copulating or mate guarding. Our analysis of male preferences followed the priority-of-access model (Altman 1962; Alberts et al. 2006). Thus, when two cycling females were maximally swollen simultaneously, we assessed the choice of the alpha male, but if he was already in consortship and there were two further swollen females, we assessed the choice of the beta male, and so on.

(e) Statistical analysis

To partition within- and between-female variance in swelling size, a one-way analysis of variance (ANOVA) was conducted on swelling area (including all cycles of each female) with female identity as a factor. In this ANOVA, the $F$-value represents the ratio of the between-female variance in swelling size on the mean within-female variance. Additional analyses to establish within-female consistency in swelling size across two consecutive cycles used a Spearman rank correlation test.

To analyse whether female quality influenced swelling size, we ran a generalized linear mixed model (GLMM) with swelling size as the response variable. This model contained a series of fixed effects both categorical (troop membership) and continuous (age, dominance rank, body size, body condition and cycle number) as well as one random factor, ‘female identity’, to account for repeated cycles within females. The introduction of female parity (as a continuous variable including three levels: nulliparous, primiparous and multiparous) in the model did not affect the significance of the variables mentioned above. Given that parity was not significant itself, and strongly correlated with age (Pearson’s correlation: $r_p = 0.86$, d.f. = 9, $p<10^{-3}$), it was not included
We devised a simple individual-based simulation model of primate groups, in which females differed in the maximal size of their sexual swellings according to the range of variation observed in our sample. The model simulations tracked the reproductive state of each female and recorded which female produced larger swellings, body size (bigger females produce larger swellings), body size using the parameter estimates of the model. Female condition is measured as the residuals of body mass against body size (crown–rump length). Each circle represents an estimate of the swelling area for one cycle. Estimates for the same female are linked by dotted lines. The line represents the predicted value of the model when the categorical variables are set as described above and the continuous variables are set to the sample average.

The statistical procedure used to test mate choice is presented in the electronic supplementary material. All statistical analyses were run using software R v. 2.5.1 (R Development Core Team 2008).

3. RESULTS
(a) Testing the reliable-indicator hypothesis
The maximal size (surface area) of individual sexual swellings varied twofold between females (95.6–194.8 cm²; n = 11 females; mean ± s.e. = 145.8 ± 34.6 cm²). Variance in maximal swelling size between females (1917 cm³) is higher than the average variance within females (771.8 cm³), but this trend is only marginally significant (one-way ANOVA: F₁,₁₀ = 2.48, p = 0.08). Nevertheless, maximal size in the first and second cycle post-partum was highly correlated within females (n = 6, rₛ = 0.94, p < 0.02), suggesting that inter-individual differences in swelling size for a given cycle are maintained across consecutive cycles.

Maximal swelling size was greater among females in better physical condition (table 1; figure 1) when controlling for the effects of age (younger females produce larger swellings), body size (bigger females produce larger swellings) troop membership (females in smaller troops produce larger swellings) and cycle number (swelling size increases with cycle number; table 1). There was no effect of female dominance rank on swelling size (table 1). In addition, females in better physical condition...
Primate sexual swellings as quality signals  E. Huchard et al.  1893

4. DISCUSSION

Our empirical findings support three critical predictions of the reliable-indicator hypothesis in a wild primate population. Our first prediction addressed the relative importance of inter- and intra-individual variability in maximal swelling size. Previous work has already shown that there is important inter-individual variability in swelling size (olive baboons: Domb & Pagel 2001, Higham et al. 2008; chimpanzees: Deschner et al. 2004; mandrills: Setchell & Wickings 2004). On the other hand, previous studies have also reported important intra-individual variability, by showing that individual swelling size increases across consecutive cycles (within an interbirth interval) in a pattern linked to hormonal levels and believed to reflect differences in between-cycle fertility (yellow baboons: Gesquiere et al. 2007; olive baboons: Higham et al. 2008; chimpanzees: Emery & Whitten 2003, Deschner et al. 2004). Our study connects these two lines of work. By demonstrating in combination that inter-individual differences in swelling size are maintained while individual swelling size increases across cycles, our findings indicate how inter-individual variability can still be substantial despite high intra-individual variability. While our data suggest that inter-individual

For age ($\chi^2_1 = 18.63, p < 10^{-3}$) and troop membership ($\chi^2_1 = 16.84, p < 0.01$) within a larger sample of 58 adult females across six troops.

We observed that on 11 occasions when a male could choose freely between two or more maximally swollen females, he chose the female with the largest swelling in 9 cases. A randomization test taking into account non-independence between female dyads (the sample comprised 6 males, including 2 alpha males, and 12 females) shows a significant association between male choice and female swelling size (Kendall’s correlation: $\tau = 0.46, p < 0.05$).

(b) Modelling the mating benefits of large sexual swellings

Our modelling results indicate that the use of a simple mate-choice rule by males, based on absolute swelling size, consistently favours the production of large swellings by females (i.e. the intensity of selection, indexed by the Pearson correlation coefficient, was always positive). This was true in the range of values explored for all three parameters tested (the ratio of within- to between-female variance, female group size and number of cycles to conception) across all possible values of breeding synchrony (figure 2). Most importantly, the correlation coefficient between mating success and swelling size remained positive in the conjunction of low female competition, indexed by total breeding asynchrony, and high within-female variance in swelling size (figure 2a). Repeated cycles to conception increased the relative magnitude of intra-individual variance compared with inter-individual variance (see fig. S2A of the electronic supplementary material) but also the intensity of female–female competition (see fig. S2B of the electronic supplementary material). Such contrasting effects seemed to compensate each other when considering the effect of the number of cycles on the intensity of selection for large swellings (figure 2c).

Figure 2. Selection intensity for large swellings in relation to breeding synchrony across a range of conditions. The intensity of selection for large swellings is measured as the Pearson correlation coefficient between female mating success and mean maximal swelling size. Breeding synchrony follows a beta distribution, described by two parameters $\alpha$ and $\beta$, ranging from total breeding asynchrony ($\alpha = \beta = 1$) to total breeding synchrony ($\alpha = \beta = \infty$). The selection intensity for any given level of breeding synchrony was investigated in relation to (a) the relative magnitude of intra-individual variance (denoted $\text{var}_{\text{intra}}$) compared with inter-individual variance (denoted $\text{var}_{\text{inter}}$) in swelling size (white bars, $\text{var}_{\text{intra}}/\text{var}_{\text{inter}} = 5$; light grey bars, $\text{var}_{\text{intra}}/\text{var}_{\text{inter}} = 1$; dark grey bars, $\text{var}_{\text{intra}}/\text{var}_{\text{inter}} = 1/5$), (b) the number of females per troop (white bars, 5 females; light grey bars, 15 females; dark grey bars, 40 females) and (c) the number of cycles to conception (white bars, two cycles; light grey bars, four cycles; dark grey bars, six cycles). The means and standard errors of the correlation coefficients are shown.

Enjoy a higher reproductive success, measured by a greater number of surviving offspring ($\chi^2 = 6.97, p < 0.01$) than females in poorer condition when controlling

Proc. R. Soc. B (2009)
differences in sexual swelling size are maintained across consecutive intervals, further research is required to establish whether this pattern extends across different interbirth intervals (our sample only includes consecutive swellings within a single interval). Notably, our finding that the swelling size of the first cycle post-partum was smaller than in subsequent cycles is also consistent with the fact that swellings advertise fertility between cycles, given that female baboons typically undergo several cycles before conception (e.g. Altmann & Alberts 2003).

Second, our data further show a positive correlation between female body condition and swelling size, which is independent of individual variation in fertility both between cycles (which is statistically controlled by the presence of cycle number in the GLMM) and within cycles (since only maximal swelling size was compared between females). Other aspects of female condition (which our condition index does not capture) might also mediate the observed effects of cycle, age and troop on swelling size. Specifically, reproduction is costly, so females that have reproduced more often (the age effect) or that have had less recovery time since the last event (the cycle effect) are likely to be in poorer condition (similarly, females take fewer cycles to conceive again when lactation has been cut short due to infant mortality; Altmann et al. 1978; Altmann & Alberts 2003). Likewise, females in larger troops (the troop effect) may be in poorer condition because feeding competition in these troops is likely to be higher (Altmann & Alberts 2003). Although the number of swollen females that were multiparous (n = 6) precluded the analysis of a direct link between swelling size and reproductive success, further analysis based on a larger sample indicated that females in better physical condition enjoyed a higher reproductive success (measured by the number of surviving offspring) in our population. This confirms that our condition index captures meaningful aspects of female quality and suggests that female baboons that produce larger swellings achieve higher reproductive success. A relationship between swelling size and reproductive success has also been reported in a previous study (n = 1–3 decisions per male across the six males), so these findings must be interpreted with caution (see the electronic supplementary material). Nevertheless, this pattern is consistent with results from a study of wild yellow baboons showing that females displaying larger maximal swellings had a higher chance of being consorted by males (Gesquiere et al. 2007).

Our individual-based model addressed two key challenges to the reliable-indicator hypothesis, both of which may reduce the strength of selection for females to produce the largest swelling. The first challenge is that inter-individual variance may be insufficient to overcome the effects of intra-individual variance if males use a simple mate-choice rule based on absolute swelling size. Nevertheless, our model shows that females that produce larger swellings than other females achieve, on average, higher mating success. This remains true even when intra-individual variance is high relative to inter-individual variance and when there are a large number of cycles to conception. The second challenge is that low female–female competition, reflecting low numbers of co-cycling females, may reduce opportunities for male mate choice such that females do not benefit from producing larger swellings. Again, our simulations show that females that produce larger swellings experience a higher mating success even under conditions of weak female competition, namely small group size and low breeding synchrony. Overall, this suggests that sexual swellings could secondarily evolve as reliable indicators from graded signals even where within-female variance is high or where female competition is low.

Inevitably, some simplifying assumptions had to be made in our model. For instance, we postulated a window of maximal fertility that lasts several days, following most primate studies (e.g. Wildt et al. 1977; Shaikh & Celaya 1982), but recent research suggests that male chimpanzees (Deschner et al. 2004) and olive baboons (Higham et al. 2009) might further discriminate the proximity of ovulation within this maximally swollen period. This might act to reduce female–female competition and thus weaken selection for large swellings. Further model development is needed to explore the implications of such complexities, which might lead to different results from those obtained here, as well as working towards the development of more realistic models more generally. Such future models would also provide a valuable opportunity to consider the evolution of sexual swellings in a wider perspective, over multiple generations, and incorporating heritability and signalling costs. Unfortunately, the data necessary to parametrize such models are currently unavailable, but it is hoped that this information will become available in the future.

Previous empirical findings may appear contradictory to our hypothesis. Although some field observations have suggested that female nutritional status may influence the production of sexual swellings (Mori et al. 1997), this is the first study to detect a direct effect of condition on swelling size. Sexual swelling size was found to be independent of body condition in captive mandrills (Setchell & Wickings 2004), captive chimpanzees (Emery & Whitten 2003), free-ranging barbary macaques (MobLe et al. 2005) and wild olive baboons (Higham et al. 2008). Moreover, there was no preference for large swellings in captive mandrills (Setchell & Wickings 2004) or wild chimpanzees (Deschner et al. 2004). So why does our result contrast with these previous findings? The pattern reported here may not extend to every primate population with exaggerated swellings, given that a number of ecological and social factors probably mediate the strength of selection on sexual swellings to become quality indicators. In the case of ecological factors, the relationship between sexual swelling size and body condition would be expected to vary with the abundance and monopolizability of resources.
Where resources are abundant or non-monopolizable, the variance in female body condition is likely to be lower. This may explain why earlier studies dealing with provisioned animals did not detect any effect of body condition on swelling size (Emery & Whitten 2003; Setchell & Wickings 2004; Möhle et al. 2005). By contrast, the Tsobis baboons live in a demanding desert environment, with occasional food shortages, which may maintain variance in female body condition.

In the case of social factors, the evolution of sexual swellings as quality indicators should be favoured where female–female competition is intense and male mating costs are elevated. Such factors are influenced by a number of specific (or even subspecific) traits, including the operational sex ratio, the length and stability of both consortships and alpha-male tenures, the frequency of male–male coalitions and infanticides, and female life history. Chacma baboons are characterized by long and stable consortships, which can last up to 13 consecutive days in our population (see also Bulger 1993; Weingrill et al. 2003), and these may increase both female–female competition and male mating costs compared with those species and populations where consortships last for only a few hours. Moreover, a high frequency of infanticide, which has been reported in several chacma baboon populations (e.g. Palombit 2003), may also increase female–female competition through two distinct mechanisms. First, females may compete more intensely to mate with the alpha male if failure to mate with him increases the risk of infanticide, thereby reducing the apparent operational sex ratio. Second, high infant mortality from infanticidal attacks following the successful male takeover of a group may result in an increase in the number of co-cycling females. By contrast, other species with different social and mating systems, such as chimpanzees and mandrills, may not share these same pressures with chacma baboons. This may be particularly true when comparing chimpanzees with monkeys, since chimpanzees exhibit an unusual pattern of male philopatry and female dispersal (such that new alpha males should have prior knowledge of individual quality for the majority of females), and the evolution of sexual swellings is believed to have occurred through independent events in chimpanzees and cercopithecines (Nunn 1999).

It has also been reported that females of lower fertility produce large swellings. Hence, young females present larger swellings (Anderson & Bielert 1994), although they are less fertile than older females in several species (Anderson 1986). Moreover, infertile females, cycling repeatedly without conceiving, often produce very large swellings (reviewed in Nunn et al. 2001). Such observations would be expected if swellings are condition-dependent, because nulliparous and infertile females, by paying limited reproductive costs, are likely to be in better condition and can therefore invest more in signalling. However, the existence of such a pattern has been used to argue against the evolutionary stability of sexual swellings as reliable indicators of female quality. In fact, these observations are not incompatible with the reliable-indicator hypothesis. Indeed, in most primate species, young females are intrinsically less attractive to males (e.g. Anderson 1986; Paul 2002; Müller et al. 2006). This pattern is irrespective of swelling size, since it also appears in those species without sexual swellings. This suggests that primates are likely to use a combination of cues to choose their mate, as an increasing number of studies report for a variety of taxa (e.g. Candolin 2003), and that males may track female age using cues other than swelling size. Thus, males are likely to be already taking age into account when they compare swelling size between females. Additionally, the frequency of infertile females in any population is likely to be low. Unfortunately, estimates of infertility frequency are difficult to obtain, and published records are often based on the presence of one infertile female in a group (e.g. Anderson 1986), which is difficult to interpret, since the presence of infertile females might be reported more often than their absence. Nevertheless, it seems reasonable to assume that infertile females will be a small minority in any viable population. Male mate choice based on maximal swelling size would thus probably remain an advantageous strategy at an evolutionary scale.

Our study supports the premise that sexual swellings are condition-dependent signals in chacma baboons, although it suggests an alternative evolutionary pathway to that envisaged in the original formulation of the reliable-indicator hypothesis (Pagel 1994). This might explain why, across species, sexual swellings are not associated with strong female competition but, rather, high male control over females as predicted by the graded-signal hypothesis (Nunn et al. 2001; Zinner et al. 2004). Nevertheless, once sexual swellings have evolved as graded signals in response to high male control, female competition, in some socio-sexual systems, may drive these signals to become reliable indicators. Thus, these two hypotheses, originally conceived as alternative explanations, may in combination provide a powerful explanatory model for the processes underlying the evolution of exaggerated primate sexual swellings, including the variable patterns reported across species and populations. Finally, and more generally, our results are also consistent with the idea that mate choice for direct benefits (here, likelihood of conception) may in evolutionary time lead to indirect benefits (here, good genes; Kokko et al. 2003).

We confirm that we have adhered to the Guidelines for the treatment of animals in behavioural research and teaching (2004) and the legal requirements of the country (Namibia) in which the work was carried out.

We thank N. Camara, L. De Raad, R. Fleming, J. Kamps, H. Kelstrup, A. King, H. Marshall and H. Peck for their invaluable assistance in the field; D. Caillaud and M. Huchard for their helpful discussions; J. Osborn, C. Staples and M. Weiß for their assistance in the laboratory; and J. Bro-Jørgensen, C. Nunn, D. Zinner and two anonymous referees for their insightful comments on the manuscript. We also thank the Swart family (2000–2006) and the Ministry of Lands and Resettlement (2006–2007) for permission to work at Tsobis Leopard Park, the Gobabeb Training and Research Centre for affiliation and the Ministry of Environment and Tourism for research permission. This work was funded by a Natural Environment Research Council (UK) project grant and advanced fellowship awarded to G.C. and a Ministère de l’Education et de la Recherche (France) studentship awarded to E.H. This paper is a publication of the ZSL Institute of Zoology’s Tsobis Baboon Project. Contribution ISEM no. 2009-014.
REFERENCES

Anderson, C. M. & Bielert, C. F. 1994 Adolescent


