ORIGINAL PAPER

Sources of variance in a female fertility signal: exaggerated estrous swellings in a natural population of baboons

Courtney L. Fitzpatrick · Jeanne Altmann · Susan C. Alberts

Abstract Signals of fertility in female animals are of increasing interest to evolutionary biologists, a development that coincides with increasing interest in male mate choice and the potential for female traits to evolve under sexual selection. We characterized variation in size of an exaggerated female fertility signal in baboons and investigated the sources of that variance. The number of sexual cycles that a female had experienced after her most recent pregnancy ("cycles since resumption") was the strongest predictor of swelling size. Furthermore, the relationship between cycles since resumption and swelling size was most evident during rainy periods and was not evident during times of drought. Finally, we found significant differences in swelling size between individual females; these differences endured across cycles (i.e., were not explained by variation within individuals) and persisted in spite of ecological effects. This study is the first to provide conclusive evidence of significant variation in swelling size between female primates (controlling for cycles since resumption) and to demonstrate that ecological constraints influence variation in this signal of fertility.

Keywords Baboon · Estrous swelling · Fertility

C. L. Fitzpatrick · S. C. Alberts
Department of Biology, Duke University, Durham, NC, USA

J. Altmann
Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

J. Altmann · S. C. Alberts
Institute of Primate Research, National Museums of Kenya, Nairobi, Kenya

C. L. Fitzpatrick (✉)
National Evolutionary Synthesis Center, 2024 W Main St. Suite A200, Durham, NC 27705-4667, USA
e-mail: courtney.fitzpatrick@duke.edu

Introduction

Signals of fertility are common among females across animal taxa, from insects (Liebig et al. 2000; Endler et al. 2004) to fish (Rowland et al. 1991; Massironi et al. 2005; Benson 2007) to reptiles (Baird 2004; Weiss 2006) to mammals (Brauch et al. 2007; Charlton et al. 2010). These signals are generally understood to have evolved because they increase the probability of fertilization (i.e., they reduce wasted reproductive opportunities) or because they focus mating activity around the time that a female can conceive (i.e., they reduce the costs of harassment and non-conceptive mating). Recently, evolutionary biologists have demonstrated a growing interest in the potential for these signals to both influence male mate choice decisions and to evolve under sexual selection. Indeed, many of the female traits that are proposed to be under sexual selection are considered signals of female fertility (Amundsen and Forsgren 2001; Huchard et al. 2009a, b). In order to understand the extent to which a fertility signal is shaped by forces *beyond* the most basic need for a female to become pregnant, an understanding of the sources of variation in the female trait is required. Despite the growing interest in fertility signals and sexual selection on females, relatively little research has been conducted on variation in female fertility signals.

One of the most conspicuous signals of female reproductive status in vertebrates has evolved multiple times in the primate lineage (Dixson 1983; Nunn 1999): the exaggerated estrous swelling displayed by many female cercopithecine primates. These swellings of the anogenital region appear

during the follicular phase of the sexual cycle and increase in size until some period of maximal swelling, which generally occurs near ovulation (but this relationship varies across species; e.g., for species with good prediction of ovulation, see Wildt et al. 1977; Emery and Whitten 2003; Higham et al. 2012; for species with looser concordance, see Reichert et al. 2002; Engelhardt et al. 2005). This variation within a given sexual cycle has been well-characterized and its proximate hormonal drivers are well understood (characterization of within-cycle variation, Wildt et al. 1977; Reichert et al. 2002; Emery and Whitten 2003; Deschner et al. 2004; Engelhardt et al. 2005; hormonal mechanisms, Gillman 1941; Kato et al. 1980; Onouchi and Kato 1983; Bentley et al. 1986; West et al. 1990). A key consequence of exaggerated swellings is that they are one mechanism by which females might extend their period of sexual receptivity beyond what is required only for fertilization. In other words, exaggerated swellings signal the *possibility* of conception beyond that period of time when it may be most *probable*, even for those species in which ovulation and maximal swelling are tightly correlated.

Males respond to the appearance of these swellings unambiguously; decades of studies in natural primate populations demonstrate that swellings motivate males to initiate mating and mate-guarding behaviors (Hausfater 1975; Tutin 1979; Packer 1979; van Noordwijk 1985). Additionally, experimental studies confirm that swellings are a primarily visual stimulus (Bielert and Girolami 1986; Girolami and Bielert 1987). Although some studies suggest that shape, color, and odor may also provide information to males, the key importance of swelling size has been consistently upheld (shape, Huchard et al. 2009a, b; color, Higham et al. 2008a, b; odor, Clarke et al. 2009; size, Bielert and Anderson 1985; Emery and Whitten 2003; Deschner et al. 2004; Gesquiere et al. 2007; Brauch et al. 2007; Higham et al. 2008a, b, 2012; Huchard et al. 2009a, b; Nitsch et al. 2011; Breaux et al. 2012).

Indeed, this established functional significance of swellings (they signal something about ovulation and males respond to them) has led to one widely accepted hypothesis for their evolution. The "graded signal" hypothesis posits that, by extending the period of receptivity and thus the probability of multiple mating, swellings function to *both* confuse paternity (by facilitating multiple mating by females, thereby increasing the number of males that are potential fathers and thus protectors of the infant) *and* bias paternity towards high quality males (by inciting the heaviest competition when conception is most probable). This hypothesis, which seeks to explain variation in swelling size within each cycle, is supported by multiple lines of evidence (Nunn 1999).

However, swellings exhibit two additional types of variation that may also have functional significance but remain largely uncharacterized. (1) Swelling size can vary *between cycles within an individual*; females in most primate species

cycle repeatedly before conception occurs, and the few studies that have examined swelling size during consecutive cycles suggest that swelling size increases from one cycle to the next until conception (Emery and Whitten 2003; Deschner et al. 2004; Higham et al. 2008a, b; Huchard et al. 2009a, b). (2) Swelling size can vary *between individuals*. That is, differences in maximal swelling size may be an intrinsic characteristic of an individual such that some always have large swellings while others always have relatively small swellings. Demonstrating between-individual variation requires controlling for within- and between-cycle variation, and this has not been done before. A couple of studies have reported such variance (Domb and Pagel 2001; Huchard et al. 2009a, b) but the first study was confounded by differences in foraging regimes (Zinner et al. 2002) and the second had a limited sample size. Importantly, no study to date has been able to examine individual variation in swelling size specifically on conceptive cycles, which is the ultimate test of the hypothesis that individuals differ in swelling size even controlling for conceptive status.

Despite the relative dearth of studies on *within-individual* and *between-individual* variation, each has the potential to reveal important information to males, information beyond just the likelihood of ovulation. The size of a sexual swelling may, like many sexual signals in males, reveal information about access to resources. The expression of such an "indicator trait" is influenced by the condition of the individual displaying it. In the classical "handicap principle" form (Zahavi 1975), some genotypes can more easily bear the cost of displaying the trait. In these cases, environmental stress increases phenotypic variation and reveals genetic quality, much like the patterns observed in stalk-eyed flies (David et al. 2000). Condition-dependent sexual signals may also be important irrespective of additive genetic variance if, for example, superior access to resources has meaningful fitness benefits for the receiver. Indeed, evidence suggests that exaggerated swellings may be just this kind of signal; they may be larger in provisioned populations of primates and larger in females that are in good condition (Domb 2000; see "Discussion" in Zinner et al. 2002; Huchard et al. 2009a, b) That is, condition-dependence in sexual swellings may result in the signaling of information about both within-individual differences (if a female's access to resources changes throughout her lifetime) and between-individual differences (if some females have better access than others). Alternatively or in addition to information about condition, some researchers have proposed that swellings reveal heritable differences in female fitness (Pagel 1994; Domb and Pagel 2001; Clutton-Brock 2007). In this scenario, swellings might signal differences in individual female quality that persist despite variation in condition.

Here, we describe variation within and between individuals in swelling size in a natural population of baboons (*Papio*

cynocephalus). We used longitudinal data to investigate potentially key sources of this variance in swelling size and we present a field method that is precise enough to capture fine scale variation. This method involves not only an optimized method for collecting images, but also a thorough analysis of error in measuring the images; we argue that in order to obtain valid morphological data with non-invasive methods, the biologically meaningful variation must exceed the variation that results from error intrinsic to the method (see Online Resource 1 for more elaboration on this point). In addition, because the study period spanned both a period of severe, unprecedented drought and a period of extreme rainfall, we had a unique opportunity to examine the effects of natural ecological variation on swelling size. We discuss both the implications of our findings for understanding exaggerated signals of fertility in females and the future potential for this field method.

Methods

Study species and population

We collected morphological, reproductive, and ecological data from a natural population of baboons that has been under continuous study by the Amboseli Baboon Research Project (ABRP) for over four decades (Hausfater 1975; Alberts et al. 2006; Gesquiere et al. 2007; Altmann et al. 2010; Alberts and Altmann 2012). Baboons are large, diurnal, semi-terrestrial monkeys that are highly flexible foragers and have successfully adapted to habitats ranging from moist evergreen forests to deserts. Most baboon taxa, including the one studied here, live in stable social groups of 20 to 100 members; females remain in their natal groups throughout their lives while males disperse, first as they approach adulthood and then often repeatedly throughout their lives. Groups include multiple adults and juveniles of both sexes.

Female baboons reach menarche at a median of 4.5 years (Charpentier et al. 2008; Onyango et al. 2013). During adulthood, they exhibit the highly visible sexual swellings that are the subject of this study, and that correspond to ovarian cycle phase (follicular vs luteal; Wildt et al. 1977). Mating occurs year-round and in the context of mate-guarding episodes, typically called consortships, which occur during the follicular phase of the female cycle when females have swellings. Males compete intensively for access to fertile females; dominance rank is an important predictor of male mating success in most baboon populations, as in many other primates (Cowlishaw and Dunbar 1991; Bulger 1993; Alberts 2012).

The Amboseli baboons occupy a short-grass savannah habitat that has undergone dramatic ecological change (from acacia woodland to semi-arid savannah) and is subject to extreme variation in both intra-annual and inter-annual rainfall (Altmann et al. 2002; Alberts et al. 2005). The study population consisted of over 300 individuals of both sexes and all ages, all of which were habituated to human observers, individually identifiable by sight, and distributed across five different social groups during the period of this study.

Swelling size measures

Swelling sizes were measured from digital images of individual females, collected opportunistically during one field season of 12 months (November 2008 through November 2009) and a second field season of 5 months (February 2010 through June 2010). Because we were especially interested in capturing the maximal swelling size for each female, we attempted to follow and collect size estimates from all cycling females as they approached the period of maximal swelling until deturgescence (reduction of swelling) was observed; these beginning and end points of data collection were assessed by experienced observers (Gesquiere et al. 2007). We excluded one female who appeared to have reached complete age-related infertility prior to our study period. We also excluded nulliparous females from our data set to avoid confounding our results with patterns of variation that might be specific to the transition from adolescence to adulthood. In other words, we only collected swelling size measures from females who had completed at least one pregnancy that resulted in a live birth.

We designated the day of deturgescence as "d-day" and then retrospectively assigned each preceding day relative to d-day such that one day prior to d-day was "d-1", two days prior to d-day was "d-2," and so on. When more females in the population were cycling than we could follow simultaneously on a given day, we prioritized those turgescent females that were expected to be closer to deturgescence on that day, estimated by experienced observers and based on swelling size and the date of onset of turgescence (we assumed that females that had been swollen longer were closer to deturgescence). We thus obtained digital images from multiple days for a given female's cycle, but only occasionally obtained images for multiple *consecutive* days leading up to d-day.

Photoscale-2 and Nikon D70

We used a Nikon D70 digital camera body with a 300-mm fixed manual lens, fitted with a Photoscale-2 to generate images that allow conversion of a direct measurement of pixels to an estimate of size in millimeters (Jacobsen 1991). The Photoscale-2 comprises a lens ring, which the observer attaches to the camera lens, and a digital reader with a sliding scale bar, which the observer attaches to the camera body (Online Resource 2) (Lee and Moss 1995; Emery and Whitten 2003). In this way, the Photoscale-2 exploits the lens extension (LE), a mechanical contraction and extension that

occurs as the photographer focuses on a subject, and records a digital measurement of the LE associated with any possible "focus distance," the distance from digital sensor plane to subject. The closer the subject is to the photographer, the further the lens will extend. Therefore, a relatively close subject will produce a higher LE than a subject that is further away. For this method to be consistent and effective, the size of the lens aperture must be taken into consideration because aperture size ("f-stop") has a marked influence on the depth of field, which, in turn, affects the ability of LE to capture subtle changes in focus distance. Therefore, we maintained the shallowest depth of field by fixing the lens aperture to its widest possible setting. Use of the wide aperture in combination with the high-light conditions of the East African savannah allowed for fast shutter speeds (usually >1/1,000 s), which prevented blur in the images. This combination of wide aperture and fast shutter speed maximized both precision and photographic clarity.

Pixels/mm conversion

In order to convert pixels to millimeters, we photographed a two-dimensional calibration image (a circle, diameter of 119 mm) from serial distances and recorded the associated lens extension. We then uploaded the images and measured the diameter of the calibration image in pixels using Image J (Schneider et al. 2012). Using the lens extensions, known diameter, and pixel measurements, we generated a conversion equation of pixels/mm=(LE)×0.4213+0.0454. This equation combined with the lens extension for a given photographic image allowed measurement in pixels to be converted to measurements in millimeters. The accuracy of this method is entirely dependent upon the successful alignment of the focus plane to the desired plane of measurement. Using this conversion, we estimated swelling size from every digital image that was included in the sample set.

Individual measures

We only included a digital image in the sample set if it met the following four criteria: (1) the image was shot from squarely behind the subject, assessed by sight on the computer screen after the images were downloaded, (2) the desired plane of measurement (posterior view of swelling) was clearly in focus, also assessed by sight on the computer screen after image download, (3) the image had an associated lens extension higher than 3 mm (because error in the conversion increased markedly with very small lens extensions), and (4) there was a change in focus distance between consecutive images (either the subject moved, the photographer moved, or the lens was retracted to infinity and then refocused; e.g., Fig. 1a; see Table 1 for sample sizes).

Initial measurements (pixels) were recorded using Image J software. After rotating the image so that a line bisecting the swelling was vertical, a Wacom Bamboo stylus and pen tablet was used to collect measurements both lengthwise and widthwise. To measure swelling width, we traced a line horizontally across the widest part of the swelling. The boundaries of the swelling width were discrete and therefore relatively easy to identify. Nonetheless, because the widest part of the swelling could not be precisely identified by the naked eye, we traced 10 measurements across the apparent widest place and retained the largest measurement. To measure swelling length, we traced a line vertically bisecting the swelling. In contrast to the discrete boundaries of the swelling width, the boundary between the anogenital tissue at the bottom of the swelling and the adjacent skin was more gradual and indiscrete. Therefore, rather than selecting only the largest of several measurements, we *averaged* the five length measurements collected from each image. These pixel measurements were then converted to a size estimate in millimeters using the previously generated pixel/mm conversion.

We collected as many images for cycling females as possible, given daily conditions. Multiple images for a given female on the same day contributed to a single mean estimate of size for that day. That is, each day's value for swelling size was the mean of all size estimates in a given day. For example, a female for whom we collected eight usable images on a given day would receive one size estimate for that day, averaged across the eight images (Fig. 1b). In cases where it was only possible to collect one usable image from a female, the value for her swelling size on that day was equal to the measurement taken from the sole image. For the remainder of this article, references to measurements of swelling "size" indicate the mean of several measures from one day (in cases where multiple images were collected) or the measurement from the only available image (in cases where conditions prevented the collection of more than one image). We were able to score more images for swelling width than for swelling length because the tail of the animal often fell exactly vertical, obscuring the bottom of the swelling (see Table 1 for sample sizes). Finally, we restricted our data set for this study to those individual females for whom we were able to determine *maximal* swelling width and length for a given cycle.

Designation of "maximal" swelling size

Our identification of maximal swelling sizes for each cycle followed three steps. First, we examined our measurement error, which we defined as the set of standard deviations for every daily size estimate in our sample that was averaged across two or more images. We produced a distribution of these standard deviations (measurement errors) and generated one error distribution for width measurements and one error distribution for length measurements.
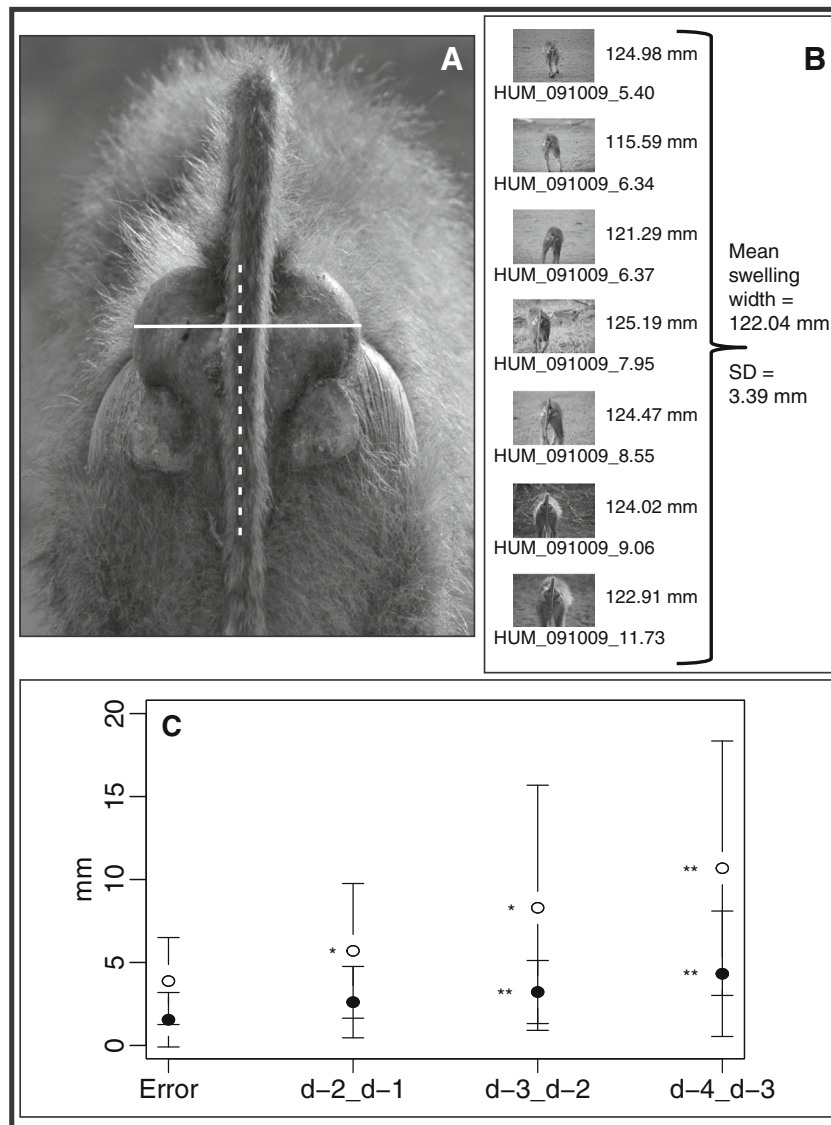
Fig. 1 Examples of digital images used for data collection and scoring, completed from high resolution color versions. **a** Images were first subjected to four explicit criteria before they were added to the data set (see Methods). Next, they were rotated until the swelling could be bisected vertically (represented by *a dashed white line*). Finally, width was measured at the largest point across the swelling (represented by *a solid white line*). **b** Multiple images for the same individual on the same day were scored individually, and all contributed to one final measurement of swelling width for a given individual on a given day. Using the pixels/mm conversion associated with each lens extension, a width measurement was generated for each image (displayed at immediate right of image). The mean of these measurements constituted the daily swelling width and the standard deviation was considered the measurement error. A similar procedure was followed for swelling length. **c** Comparison of measurement errors to detected changes in swelling sizes from d-4 to d-3, d-3 to d-2, d-2 to d-1 for swelling width (*closed circles*) and swelling length (*open circles*). Average detectable change in swelling length from d-4 to d-3 and from d-3 to d-2 and from d-2 to d-1 was significantly larger than the average measurement error. Likewise, average detectable change in swelling width from d-4 to d-3 and d-3 to d-2 was statistically larger the average measurement error ($*p<0.05$; $**p<0.01$). *Error bars* represent one standard deviation from the mean

Second, we examined whether the changes in swelling size that we measured from one day to the next were larger than our measurement error, in order to confirm that our method captured real day-to-day change in swelling size. To do this, we identified all individuals for whom we had swelling size measures on at least two consecutive days between day d-4 and day d-1 of the cycle (see "Error analysis: *matched photographic pairs*, Table 1). We had one set of swelling size differences for swelling length and one set for swelling width

between d-4 and d-3, a second set of swelling size differences between d-3 and d-2, and a third set of swelling size differences between d-2 and d-1. Swelling widths ranged from 105.36 to 164.83 mm with an average standard deviation of 1.94 mm within female-days (1.18–1.84 % error); swelling lengths ranged from 100.90 to 300.90 mm with an average standard deviation of 3.88 mm within female-days (2.36–3.68 % error). For both width and length, we found a highly significant difference between the measurement error and the

Table 1 Sample sizes for full data set, error analysis, multi-model inference, and restricted post hoc analyses

	Width	Length
Full data set: number of photographs	1,609	746
Full data set: number of unique females	61	50
Full data set: number of sexual cycles	143	89
Error analysis: number of photographs	248	182
Error analysis: matched photograph pairs (d-4 to d-3)	14	13
Error analysis: matched photograph pairs (d-3 to d-2)	20	17
Error analysis: matched photograph pairs (d-2 to d-1)	23	22
Multi-model inference: number of unique females	50	48
Multi-model inference: number of sexual cycles	86	84
Restricted post hoc analyses	50	48

detected change in swelling size from day d-4 to d-3 (Wilcoxon rank sum test: width: $W=2{,}486$, $p=0.01$; length: $W=1{,}839$, $p<0.01$), and from day d-3 to d-2 (Wilcoxon rank sum test: width: $W=3{,}804$, $p<0.01$; length: $W=2{,}073$, $p=0.02$). In contrast, from day d-2 to d-1, we found a significant difference between the measurement error and the change in swelling length (Wilcoxon rank sum test: $W=2{,}555$, $p=0.03$), but not in swelling width, although it trended toward being larger than the error (Wilcoxon rank sum test: $W=3{,}452$, $p=0.10$) (Fig. 1c).

Third, because our second step demonstrated that changes in swelling size from day d-4 to d-3, and from d-3 to d-2, were large enough to be reliably measured, but changes from d-2 to d-1 sometimes were and sometimes were not, we sought confirmation that swelling sizes *increased* each day approaching d day. We found, as expected, that swellings were more likely to increase than decrease in both width and length from d-4 to d-3 and from d-3 to d-2 (Exact Binomial Test for d-4 to d-3: width, $p=0.03$; length $p=0.02$; Exact Binomial Test for d-3 to d-3: width, $p=0.01$; length, $p=0.14$). However, between d-2 and d-1, swellings were just as likely to decrease as they were to increase in both width and length (Exact Binomial Test; width, $p=1$; length, $p=0.5$). This inability to predict the direction of change in size between d-2 and d-1 suggests that some swellings began to slowly decrease before observers could visually detect deturgescence whereas others were still increasing; this heterogeneity would make it more difficult to pinpoint the exact d-day by sight. In other words, multiple lines of evidence showed that swelling sizes (both width and length) predictably increased from d-4 to d-2, but *did not predictably increase* during the window of time between d-2 and d-day.

Therefore, we considered measurements from either d-2 or d-1 as measurements of "maximal swelling size". We then restricted the remainder of the analyses only to those subjects for whom we had size estimates on one or both of days d-1 and d-2 (see Multi-model inference, Table 1). When we had

size estimates for both days for a given subject, we designated the larger of the two as the subject's "maximal swelling size" for that cycle.

Ecological data

From November through May each year, the Amboseli ecosystem receives highly variable amounts of rain; the months of June-October constitute a long and predictably dry season. The yearly average is ~350 mm, but the range is 141–757 mm, and any given month or set of months between November and May may experience no rain. (Altmann et al. 2002; Alberts et al. 2005). Indeed, very little rain fell between Nov 2007 and Nov 2009, resulting in the most severe drought that Amboseli had experienced in over 50 years (Alberts and Altmann, unpublished data). The first of our study's two time periods occurred during the end of this period of extreme drought (June 2009–October 2009). The second of our study's two time periods (February through June 2010) occurred during a period of heavy rains, which were probably due to an El Nino effect (Kiage and Obuoyo 2011).

Swelling size is likely to be influenced by food availability. One study (Domb and Pagel 2001) demonstrated swelling size differences that were strongly correlated with differences in food availability (discussed in Zinner et al. 2002), and Huchard et al. (2009a, b) showed that females with better body condition have larger swellings. Our data set presented a rare and unanticipated opportunity to examine direct ecological effects on sexual swelling size in a completely unprovisioned population. Baboons feed primarily on plants whose growth is determined by a combination of rainfall and temperature. Thus, we created a metric of drought status using rainfall data (mm) and temperature data (°C), both of which are collected continuously by the Amboseli Baboon Research Project (Alberts and Altmann 2011). Following Le Houerou (1989) and Beehner et al. (2006), we used a metric that is based on the combined temperature and rainfall requirements for most African crops (see also Bronikowski and Altmann 1996). We defined a "wet month" as any 30-day period that received more than twice the mean annual temperature; in Amboseli, this threshold corresponds to 50 mm of rain (Beehner 2006). In other words, rather than characterizing given calendar months as wet or not, this method classifies each 30-day period in a sliding window. To categorize the ecological conditions that each female was under when a given swelling measurement was collected, we counted the number of days since the end of the last "wet" 30-day period. This created a "drought score", *days since last wet month (d)* a continuous variable that reflected the distribution of ecological conditions. For example, a day falling within a 30-day period of over 50 mm of rain—lush ecological conditions in the Amboseli basin—would receive a drought score of 0. In contrast, a day that was 25 days after there had been a

cumulative rainfall of 50 mm within a 30-day period received a score of 25. Because a typical year in Amboseli includes a 5-month dry season, a drought score greater than 150 (corresponding to more than five continuous dry months) can be thought of as reflecting drought conditions, with the severity of the drought increasing as the drought score increases. Because our study spanned two periods (see above), our drought score data were distributed bimodally, with most of the swelling size data either having been collected under very wet conditions (for Amboseli) or under extreme drought conditions (Online Resource 3 for distribution of swelling size data across rainfall, https://amboselibaboons.nd.edu/downloads/ for Amboseli rainfall data from 1976–2012 calendar years).

Age, dominance rank, and reproductive status

Because female age, dominance rank, and aspects of reproductive status are all likely to be associated with female baboon fertility and fecundability, we included these variables in our analysis. Using behavioral, reproductive, and other longitudinal data that were collected in accordance with the ABRP protocol, we included female age, female rank (ordinal rank number), and cycles since resumption (number of cycles since last pregnancy) as predictor variables in the model. We did not examine the effect of parity on variation in swelling size because age and parity are highly collinear (female age accounted for 83 % of the variance in parity).

Age (*a*)

If sexual swelling size signals reproductive capacity or reproductive value, the signal may change with age (because fertility changes with age; Alberts and Altmann 2003; Beehner 2006; Altmann et al. 2010). Each individual female in the study population had a closely known birthdate, from which age (rounded to the nearest tenth of a year) was calculated.

Dominance rank (*r*)

Dominance rank influences access to resources (Barton 1993; Barton and Whiten 1993), which influences fertility (Bercovitch and Strum 1993; Altmann and Alberts 2003). Furthermore, previous studies of sexual swellings have noted that provisioned primate populations display especially large sexual swellings (Fa and Southwick 1988) suggesting that access to resources may be a contributor to variation in swelling size. As part of the ongoing Amboseli baboon research, females are assigned an ordinal dominance rank each month on the basis of decided agonistic interactions between pairs of individuals (Hausfater 1975; Alberts et al. 2003). The highest

ranking female in the group is assigned rank 1, the next highest ranking female is assigned rank 2, and so on.

Cycles since resumption (*c*)

In order to examine within-individual variation, we included as a predictor variable an ordinal number indicating the number of "cycles since resumption" represented by that cycle. Females do not cycle when pregnant and, after parturition, generally undergo a period of post-partum amenorrhea, ranging from 6 to 16 months (Altmann et al. 1977). All post-menarcheal females were monitored regularly for cycling status, and each cycle was counted relative to the first cycle since the last pregnancy, so that the first ("resume") cycle was 1, the second cycle was 2, and so on.

Data analysis

We used an information theoretic approach to investigate the potential sources of variance in maximal swelling size, because, rather than testing one alternative hypothesis against a null, our inquiry entertained a family of hypotheses that were, a priori, equally plausible (reviewed by Burnham et al. 2011; Garamszegi 2011; Symonds and Moussalli 2011). This method assumes that the researcher has used previous biological knowledge to select appropriate predictors and then evaluates a set of candidate models. That is, while the variables that we considered as potential predictors of variance in swelling size were informed by our biological knowledge, several different combinations of those predictors were equally plausible at the outset, making this study ideal for an information theoretic approach.

We evaluated a candidate model set that included *age*, *female rank*, *number of cycles since resumption*, and *days since last wet month* as fixed effects in a linear mixed-effects model using a maximum likelihood method (Table 2). To both account for the repeated sampling of subjects that were represented in the data set more than once (for multiple cycles) and to test for individual differences in swelling size, we included female identity as a random effect in all of the models. We calculated an adjusted measure of Akaike's Information Criterion (AIC) (Akaike 1973) for combinations of fixed effects and interactions, up to five terms, using the "dredge" function in the MuMIn package of the statistical software, R. This adjusted measure, AIC_C, accounted for our limited sample sizes (Burnham et al. 2011; Symonds and Moussalli 2011). We then calculated the difference between each model's AIC_C and the model with the lowest AIC_C (ΔAIC_C) and the Akaike weight value (w_i). As is the convention, we considered those models with a ΔAIC_C of less than 2 to be equivalent in terms of their likelihood to capture meaningful variation in the response variable. We performed this

Table 2 Predictor variables used in analysis of variation in swelling size

Variable	Abbreviation	Description
Cycles since resumption	c	Ordinal number representing, for a given cycle, how many cycles have passed since completion of most recent episode of post-partum ammeonorrhea
Days since last wet month	d	Number of days that have passed since a single 30-day period that contained 50 mm of rain or more
Rank	r	A female's dominance rank relative to other females in the social group. 1=highest rank
Age	a	A female's age at time of measurement of swellings, to the nearest tenth of a year

analysis twice: once for swelling width, and once for swelling length.

After the initial information theoretic analysis, we used a standard model averaging technique to estimate the effect sizes for each relevant parameter. In order to estimate the relative effect sizes of each term that appeared in any of the top models, we averaged the models in each of the 95 % confidence sets (i.e. $\Delta AIC_C < 10$). Model averaging with this threshold of confidence provides an additional and conservative method of estimating the effects of a given predictor (Burnham and Anderson 2002). To create these parameter estimates, we re-calculated Akaike weights for one subset of models per term for the width confidence set and one subset of models per term for the length confidence set. Weighted averages were created by multiplying these revised weights by the term of interest for each model respectively and then summing these products.

In order to elucidate how each of the important predictors contributed to variation in maximal swelling width and length, we examined the individual effects of each in post hoc analyses. We controlled for pseudoreplication in the post hoc analysis of *age* and *dominance rank* by further restricting the data set so that each female was only represented once (selecting, for each female, the cycle with the highest *cycles since resumption*). We chose this approach—rather than additional mixed-effects models, which evaluate only the means (and not the variances) of fixed effects—because it allowed us to estimate the variance in swelling size that was explained by each of the relevant predictors. Because female identity is independent of *cycles since resumption* and *days since last wet month* and therefore should not impose a random effect on the relationship between these two predictors and swelling size, we did not restrict those post hoc analyses. In other words, we allowed females to be represented in the data set more than once and we did not include female identity as a random variable for post hoc analyses of *cycles since resumption* and *days since last wet month*.

Finally, in order to estimate the *between individual* variation in swelling size, we compared our best model for swelling width and our best model for swelling length to identical models that excluded individual female identity as a random variable. We then performed a variance component analysis by partitioning the variance that remained unaccounted for by the fixed effects in a mixed-model, using the identical data set as in the multi-model inference. We performed this variance component analysis for every model in each of the two 95 % confidence sets, producing a range of estimates for the proportion of unexplained variance that is attributable to female identity.

Results

Multi-model inference

Multiple models in our candidate set produced a ΔAIC_C that was within two units of the model with the lowest AIC_C, indicating that they had substantial and nearly equivalent support (Table 3). For swelling width, the top three models all indicated that some combination of *age, cycles since resumption*, and our drought variable (*days since last wet month*) played a role in determining swelling width. For swelling length, female *dominance rank* appeared in six of the seven top models, while *age* appeared only twice. As with swelling width, *cycles since resumption* and *days since last wet month* were important for swelling length; they appeared in every model in the confidence set. That is, our analysis indicated that *cycles since resumption* and *days since last wet month* affected both aspects of swelling size, and that age and dominance rank each influenced one aspect of size more than the other. Finally, two interaction terms appeared. An interaction between *cycles since resumption* and *days since last wet month* appeared in both confidence sets, while an interaction between *days since last wet month* and female *dominance rank* appeared in one of the top models for length. The model-averaged parameter estimates provide additional evidence that *cycles since resumption* affected both swelling width and length, that *age* influenced swelling width, and that *dominance rank* influenced swelling length (Table 4).

Female age

Age appeared in all the top models for swelling width but in only two of the best models for swelling length. Indeed, in a post hoc simple linear model with a restricted data set so that each female was only represented once (see "Restricted post

Table 3 Models of fixed effects on variation in maximal swelling size (width, N=96 cycles; length, N=85 cycles; see Table 2 for descriptions of variables. Interaction terms are indicated with an asterisk. All models include female identity as a random effect, only the first 20 models of full candidate set are shown, and italicized rows should be considered equivalent to best model (i.e., Δi<2)

Fixed Effects	AIC_C	ΔAIC_C	w_i
SWELLING WIDTH			
a + c + d + c*d	*653.62*	*0.00*	*0.15*
a + c + d	*654.99*	*1.37*	*0.07*
a + c	*655.09*	*1.47*	*0.07*
c + d + c*d	655.67	2.05	0.05
a + c + d + a*d + c*d	655.72	2.09	0.05
a + c + a*c	655.76	2.14	0.05
a + c + d + a*c	655.90	2.28	0.05
a + c + d + r + c*d	655.96	2.33	0.05
a + c + d + a*c + c*d	656.00	2.37	0.05
c	656.23	2.61	0.04
a + c + r + a*r	656.51	2.89	0.04
c + d	656.83	3.20	0.03
a + c + d + a*d	656.85	3.23	0.03
a + c + r	657.28	3.65	0.02
a + c + d + r	657.31	3.69	0.02
a + c + d + r + a*r	657.43	3.81	0.02
c + d + r + c*d	657.67	4.04	0.02
a + c + d + a*c + a*d	657.81	4.19	0.02
a + c + r + a*c	657.92	4.29	0.02
c + r	658.11	4.49	0.02
SWELLING LENGTH			
c + d + r + c*d	*700.40*	*0.00*	*0.12*
c + d + c*d	*700.64*	*0.24*	*0.10*
a + c + d + r + c*d	*701.47*	*1.08*	*0.07*
c + d + r	*701.65*	*1.25*	*0.06*
a + c + d + r + d*r	*702.00*	*1.60*	*0.05*
c + d + r + c*d + c*r	*702.12*	*1.72*	*0.05*
c + d + r + d*r	*702.31*	*1.91*	*0.04*
a + c + d + c*d	702.49	2.09	0.04
c + d	702.50	2.10	0.04
c + d + r + c*d + c*r	702.51	2.11	0.04
c + r	702.77	2.37	0.04
a + c + d + r	703.01	2.62	0.03
c	703.25	2.85	0.03
a + c + d + r + a*c	703.42	3.02	0.03
c + d + r + c*r	703.96	3.56	0.02
a + c + r + a*c	704.06	3.67	0.02
a + c + d + a*d + c*d	704.35	3.95	0.02
a + c + d + r + a*r	704.46	4.06	0.02
c + d + r + c*r + d*r	704.50	4.11	0.01
a + c + d	704.55	4.16	0.01

hoc analyses", Table 1), age explained 4 % of the variance in swelling width but none of the variance in swelling length. However, the relationship between age and swelling size strengthened when we excluded the oldest females in our data set (>20 years). We chose 20 years as a conservative threshold because Amboseli females older than ~18 years have declining rates of fertility (Beehner et al. 2006; Altmann et al. 2010). When we included only females younger than 20 years of age, we found that both maximal swelling width and length increased for females as they aged *from early to middle adulthood* (Fig. 2a.). Although the maximum lifespan for baboons in Amboseli is 27–28 years (Bronikowski et al. 2011), only 6 % of the females in our total sample (3 of 46 females) were over the age of 20. Understanding the way in which maximal swelling size changes with age in late adulthood will require further study with a data set that includes more females that are older.

Dominance rank

Although female dominance rank did not appear in any of the best top models for swelling width, it appeared in all but one of the seven top models for swelling length. This difference between width and length remained when we examined the bivariate relationships in two simple linear models; dominance rank explained none of the variance in swelling width ($R^2=-0.005$), but females of higher dominance rank tended to have longer swellings ($R^2=0.03$) (Fig. 2b). This relationship persisted when we removed the oldest females from our analysis.

Cycles since resumption

Cycles since resumption appeared in every model in the confidence set for both swelling width and length. Specifically, maximal swelling width and length were larger with increasing numbers of cycles since resumption, indicating a pattern of progressively increasing maximal swelling size as a female continued to cycle without becoming pregnant (Fig. 2c). This population-level pattern was also evident in individual trajectories of maximal swelling size across consecutive cycles. All

Table 4 Model-averaged parameter estimates for fixed effects in multimodel inference

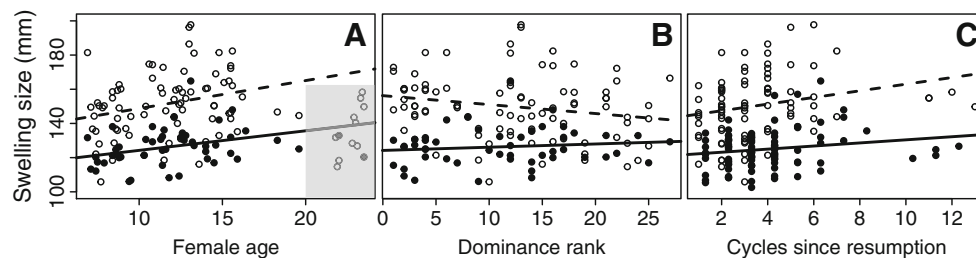| Variable	Parameter	estimate
	Width (mm)	Length (mm)
Age (a)	1.059	1.051
Cycles since resumption (c)	1.587	3.342
Days since last wet month (d)	0.009	0.013
Dominance rank (r)	NA	−0.473
c*d	−0.006	0.174
d*r	NA	0.004

Fig. 2 Bivariate relationships between swelling width (*closed circles, solid lines*) or swelling length (*open circles, dashed lines*) and **a** female age, **b** female dominance rank, **c** cycles since resumption. Trend lines are least squares regressions and are not a substitute for the full multivariate analysis. *Shaded portion* in **a** represents data points for the oldest females in our data set, which were excluded during the calculation of linear regression. Each female is represented only once for **a** and **b**, but may appear multiple times in **c**

10 individual females that we observed across at least two cycles in the wetter 2010 study period showed this pattern of increased size across cycles (Online Resource 4).

Days since last wet month

Days since last wet month appeared in all of the top models for length and the top two models for width. We therefore compared the distributions of maximal swelling size during the most extreme drought period (when females had experienced 150 days or more since a 30-day period of over 50-mm rain) with the distribution of maximal swelling sizes during a wet period (i.e., during a 30 day period when there had been at least 50 mm of rain.) Mean maximal swelling widths and lengths during drought periods were smaller than during "wet" periods (Fig. 3).

Interaction between *cycles since resumption* and *days since last wet month*

An interaction between *cycles since resumption* and *days since last wet month* appeared in both of our confidence sets. In a post hoc analysis, we examined that potential interaction by partitioning the data into drought days (*days since last wet month*>150) and wet days (*days since last wet month*<30), excluding data that fell between these two partitions. We then modeled the effect of *cycles since resumption* on maximal swelling size in each partition. In a least squares regression, *cycles since resumption* was a significant predictor of swelling width during the wet months (parameter estimate=2.06, p=0.03, R^2=0.08, N=29 females), but not during the drought (parameter estimate=−0.06, p=0.91, R^2=−0.03, N=21 females) (Fig. 4a). A similar, though non-significant, pattern was found for swelling length. That is, *cycles since resumption* explained variance in swelling length during the wet months (parameter estimate=2.30, p=0.16, R^2=0.03, N=28 females), but not during the drought months (parameter estimate=1.16, p=0.23, R^2=0.01, N=21 females) (Fig. 4b). Although this analysis does not control for the multiple testing inherent in the information theoretic approach, we report the p values here

to demonstrate the considerable reduction in the predictive power of *cycles since resumption* during drought days as opposed to wet days.

Interaction between *dominance rank* and *days since last wet month*

Finally, an interaction between female *dominance rank* and *days since last wet month* appeared once in our confidence set for swelling length. Specifically, swelling lengths of low ranking females tended to be smaller than those of high ranking females during the drought (parameter estimate= −0.68, p=0.08, R^2=0.05, N=21 females), but dominance rank did not predict swelling length during the wet months (parameter estimate=−0.29, p=0.51, R^2=−0.01, N=29 females).

Unexplained variation and enduring differences between individuals

Female identity had a substantial effect on swelling size. For swelling width, excluding female identity as a random variable from the single top model in our candidate set produced an AIC_C of 716.6, compared to an AIC_C of 653.6 for a model including the random variable, yielding a Δ_i of 64. For swelling length, the AIC_C values were 734 versus 700.40
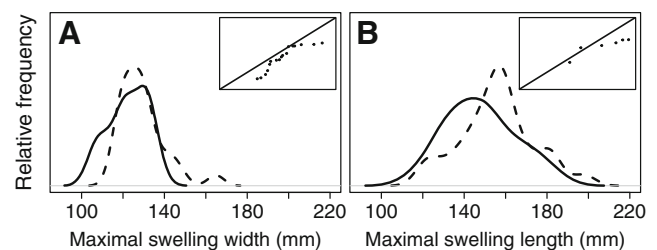

Fig. 3 Comparative distributions during drought conditions (*solid line*) and wet conditions (*dashed line*) for **a** maximal swelling widths **b** maximal swelling lengths. To reduce noise introduced by within-individual variation, females are represented only once. Maximal swelling sizes were smaller during drought conditions. Insets show quantile-quantile plots demonstrating the differentiation between the distributions
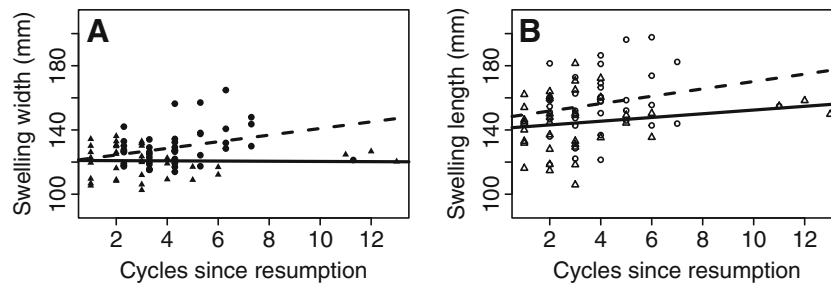
Fig. 4 Least-squared regression of **a** maximal swelling width and **b** length on *cycles since resumption* during both drought periods (*triangles, solid line*—150 days or over since last wet month) and wet periods (*circles, dashed lines*—30 days or less since last wet month). Each data point represents maximal swelling size for one cycle. During drought times, cycles since resumption has no effect on swelling size. However, during wet times, cycles since resumption predicts swelling size. Because repeated samples of individual females are represented in this visualization and it should not be considered a substitute for the full post hoc analysis

respectively, yielding a Δ_i of 33. These are unambiguously large values for Δ_i, indicating that female identity as a random variable added a large amount of information to the models. To further quantify the effect of female identity, we performed a variance component analysis for each of the models in the two 95 % confidence sets. Of the variance in swelling width that was unexplained by the fixed effects in these models, 86–88 % (depending on the model) was explained by female identity, while only 12–15 % was attributable to error (or other unidentified sources of variance). Similarly, of the unexplained variance in swelling length, 64–75 % was explained by female identity. In other words, the most conservative estimate is that female identity explained 86 % of swelling width and 64 % of swelling length. Taken together, these two analyses provide strong evidence that individuals exhibited significant differences in swelling size that were enduring across consecutive cycles (i.e., were not explained by *cycles since resumption*), that were not accounted for simply by age or dominance rank, and that persisted in spite of the effects of environmental perturbations such as droughts (Fig. 5).

Discussion

Our analysis demonstrates ecological effects on the maximal swelling size that female baboons achieve on a given sexual cycle. It also demonstrates that characteristics of both an individual female and an individual cycle influence swelling size. The most consistent effect was the positive relationship between *cycles since resumption* and swelling size: swellings got bigger as sexual cycles accumulated after a female's most recent episode of post-partum amenorrhea. This study confirms previous findings that have also found evidence for this type of *within-individual* variation in both baboons and chimpanzees, which represent independent evolutionary origins of sexual swellings (chimpanzees, Emery and Whitten 2003; Deschner et al. 2004; baboons, Higham et al. 2008a, b; Huchard et al. 2009a, b). However, sample sizes in previous

studies were limited. Because the effect of *cycles since resumption* was the strongest predictor in this analysis, and the sample size was large (for studies of wild primates), this study provides the strongest evidence to date that variation in size *within individuals across sexual cycles* signals changes in the probability of conception across cycles, in at least some species.

Likewise, the positive relationship between age and swelling size among reproductively active young to middle-aged females further suggests that swelling size tracks changes in a female's age-specific fertility. The increase in swelling size through middle age reported here mirrors an age-related increase in the probability of conception for cycling females and in birth rates in this population (Beehner 2006; Altmann et al. 2010). Older females (beyond 18 years) experience an age-related decline in fertility and we predict that when more swelling size data are available for females in that age range, they will show a gradual decline in swelling size with age (see also Fig. 2a).
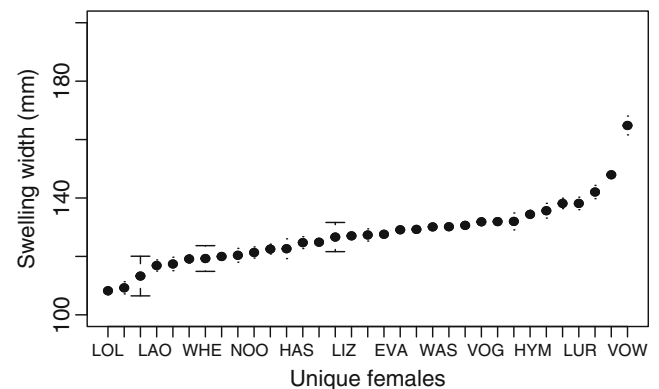

Fig. 5 Maximal swelling width for 33 unique individuals on their conceptive cycle. Each measure represents a mean calculated from multiple images collected on the same day (range=2–10 images) and error bars represent standard deviation from the mean width for that female on that cycle

Our study also provides the first evidence from a natural population (i.e. not food enhanced) that swelling size is influenced by changing environmental conditions. This study supports and extends previous observations that swelling size differences can be strongly correlated with differences in food availability between social groups (Zinner et al. 2002), and that females with better body condition may have larger swellings (Huchard et al. 2009a, b). The difference in size that we found between the drought and the period of good rainfall is likely a consequence of the interaction between *days since last wet month* and *cycles since resumption*. In particular, we found that females exhibited larger swellings with each subsequent cycle during the period of good rainfall, but not during the period of drought, suggesting that the water sequestration required to produce a swelling may present a physiological challenge to female primates.

Females of higher dominance rank tended to have longer swellings, but there was no effect of dominance rank on swelling width. In reality, males probably assess female swelling size as a whole rather in one component at a time. A signal of dominance rank in length but not width may simply reflect a somewhat weaker relationship between rank and overall swelling size than we found for *age*, *cycles since resumption*, and *days since last wet month*, all of which showed the same pattern for both length and width.

Our analysis revealed substantial variation in swelling size between individuals, much of which remained unaccounted for by the variables we examined. That is, significant between-individual differences in swelling size persisted even when controlling for within-cycle and within-individual variation. This is an important step toward a complete description of how primate sexual swellings vary because, while other studies have made attempts to describe inter-individual variation (Domb and Pagel 2001; Huchard et al. 2009a, b), this study is the first to control for several potentially confounding factors.

Several variables might influence this inter-individual variation, including differences in body size, genetic variation, or differences in endogenous hormones. We tried to measure body size using the same photographic technique that we used to measure swelling size, but we concluded that within-individual error in these measurements was too great enough to warrant analyzing them (see Online Resource 5). Nonetheless, if males can assess whether a female's swellings are particularly large *for her body size*, this might represent somewhat more refined information than simply absolute swelling size. Currently, we have no direct evidence to indicate whether absolute swelling size or swelling size relative to body size is the more relevant signal, but we anticipate that subsequent studies of male response to swelling size will reveal this information. If genetic variation influences variation in swelling size, important genetic differences might be found in insulin/IgF receptor genes; a recent study

hypothesizes that sexually selected are most likely to exploit the insulin/IgF pathway (Emlen et al. 2012). Indeed, one study has already indicated that swellings may sometimes reveal information about genes related to immune function (Huchard et al. 2010). Similarly, because circulating levels of estrogen drive the increase in swelling size (Ozasa and Gould 1982; Onouchi and Kato 1983), genetic variation in estrogen receptor genes and their associated regulatory regions may be important determinants of swelling size. Finally, inter-individual variation in swelling size may reflect enduring differences in circulating levels of hormones, or receptor density in the tissue involved.

Taken together, our study suggests that sexual swellings potentially signal multiple types of information relevant to fertility. They may signal the probability of conception, as it is influenced both by the cycle-distance from a female's most recent pregnancy and by access to nutritional resources. Similarly, changes in swelling size within a female or differences in swelling size between females may reflect age dependent changes in fertility. Finally, swelling size may indicate enduring differences in fertility between females.

Future research should examine the male behavioral response to this variation in swelling size in order to elucidate further the role that swellings play in signaling, or concealing, fertility information to males. Additional studies are needed that investigate the sources of variance in individual variation as well as the specific components of fitness that variation in swelling size—and fertility signals in general—might be correlated with. This is especially important as biologists begin trying to understand how these fertility signals might influence the evolution of male mate choice and how the traits themselves evolve.

References

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In: Petrov B, Csaki F (eds) Proceedings of the 2nd international symposium of information theory. Akademiai Kiado, Budapest, pp 267–281

Alberts SC (2012) Magnitude and sources of variation in male reproductive performance. In: Mitani J, Call J, Kappeler P, Palombit R, Silk J (eds) Evolution of primate societies. University of Chicago Press, USA, pp 412–431

Alberts SC, Altmann J (2003) Matrix models for primate life history analysis. In: Kappeler P, Pereira ME (eds) Primate life history and socioecology. University of Chicago Press, Chicago, pp 66–102

Alberts SC, Altmann J (2011) Monitoring guide for the Amboseli Baboon Research Project: protocols for long-term monitoring and data collection. Published online at http://princeton.edu/~baboon/monitoring_guide.html

Alberts SC, Altmann J (2012) The Amboseli Baboon Research Project: themes of continuity and change. In: Kappeler P, Watts D (eds) Long-term field studies of primates. Springer, Berlin, pp 261–288

Alberts SC, Watts HE, Altmann J (2003) Queuing and queue-jumping: long-term patterns of reproductive skew in male savannah baboons, *Papio cynocephalus*. Anim Behav 65:821–840

Alberts SC, Hollister-Smith JA, Mututua RS, Sayialel SN, Muruthi PM, Warutere JK, Altmann J (2005) Seasonality and long-term change in a savanna environment. In: Brockman DK, van Schaik CP (eds) Seasonality in primates: studies of living and extinct humans and non-human primates. Cambridge University Press, Cambridge, pp 157–196

Alberts SC, Buchan JC, Altmann J (2006) Sexual selection in wild baboons: from mating opportunities to paternity success. Anim Behav 72:1177–1196

Altmann J, Alberts SC (2003) Variability in reproductive success viewed from a life-history perspective in baboons. Am J Hum Biol 15:401–409

Altmann J, Altmann SA, Hausfater G, Mccuskey S (1977) Life history of yellow baboons: physical development, reproductive parameters, and infant mortality. Primates 18:668–673

Altmann J, Alberts SC, Altmann SA, Roy SB (2002) Dramatic change in local climate patterns in the Amboseli basin, Kenya. Afr J Ecol 40:248–251

Altmann J, Gesquiere L, Galbany J, Onyango PO, Alberts SC (2010) Life history context of reproductive aging in a wild primate model. Ann N Y Acad Sci 1204:127–138

Amundsen T, Forsgren E (2001) Male mate choice selects for female coloration in a fish. Proc Natl Acad Sci U S A 98:13155–13160

Baird T (2004) Reproductive coloration in female collared lizards, *Crotophytus collaris*, stimulates courtship by males. Herpetologica 60:337–348

Barton RA (1993) Sociospatial mechanisms of feeding competition in female olive baboons, *Papio anubis*. Anim Behav 46:791–802

Barton RA, Whiten A (1993) Feeding competition among olive baboons, *Papio anubis*. Anim Behav 46:777–789

Beehner JC (2006) The ecology of conception and pregnancy failure in wild baboons. Behav Ecol 17:741–750

Beehner JC, Nguyen N, Wango EO, Alberts SC, Altmann J (2006) The endocrinology of pregnancy and fetal loss in wild baboons. Horm Behav 49:688–699

Benson KE (2007) Enhanced female brood patch size stimulates male courtship in *Xiphophorus helleri*. Copeia 2007:212–217

Bentley JP, Brenner RM, Linstedt AD, West NB, Carlisle KS, Rokosova BC, McDonald N (1986) Increased hyaluronate and collagen biosynthesis and fibroblast estrogen receptors in Macaque sex skin. J Investig Dermotol 87:668–673

Bercovitch FB, Strum SC (1993) Dominance rank, resource availability, and reproductive maturation in female savanna baboons. Behav Ecol Sociobiol 33:313–318

Bielert C, Anderson CM (1985) Baboon sexual swellings and male response: a possible operational mammalian supernormal stimulus and response interaction. Int J Primatol 6:377–393

Bielert C, Girolami L (1986) Experimental assessments of behavioral and anatomical components of female chacma baboon (*Papio ursinus*) sexual attractiveness. Psychoneuroendocrinology 11:75–90

Brauch K, Pfefferle D, Hodges K, Möhle U, Fischer J, Heistermann M (2007) Female sexual behavior and sexual swelling size as potential cues for males to discern the female fertile phase in free-ranging Barbary macaques (*Macaca sylvanus*) of Gibraltar. Horm Behav 52:375–383

Breaux SD, Watson SL, Fontenot MB (2012) A free choice task evaluating chimpanzees' preference for photographic images of sex swellings: effects of color, size, and symmetry. Int J Comp Psychol 25:118–136

Bronikowski AM, Altmann J (1996) Foraging in a variable environment: weather patterns and the behavioral ecology of baboons. Behav Ecol Sociobiol 39:11–25

Bronikowski AM, Altmann J, Brockman DK, Cords M, Fedigan L, Pusey A, Stoinski T, Morris WF, Strier KB, Alberts SC (2011) Aging in the natural world: comparative data reveal similar mortality patterns across primates. Science 331:1325–1328

Bulger JB (1993) Dominance rank and access to estrous females in male savanna baboons. Behaviour 127:67–103

Burnham KP, Anderson DR (2002) Formal inference from more than one model: multimodel inference (MMI). Model Sel. multimodel inference; a practical information-theoretic approach, 2nd edn. Springer, New York, pp 149–153

Burnham KP, Anderson DR, Huyvaert KP (2011) AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behav Ecol Sociobiol 65:23–35

Charlton BD, Keating JL, Rengui L, Huang Y, Swaisgood R (2010) Female giant panda (*Ailuropoda melanoleuca*) chirps advertise the caller's fertile phase. Proc R Soc Lond B 277:1101–1106

Charpentier MJE, van Horn RC, Altmann J, Alberts SC (2008) Paternal effects on offspring fitness in a multimale primate society. Proc Natl Acad Sci U S A 105:1988–1992

Clarke P, Barrett L, Henzi SP (2009) What role do olfactory cues play in chacma baboon mating? Am J Primatol 71(6):493–502

Clutton-Brock T (2007) Sexual selection in males and females. Science 318:1882–1885

Cowlishaw G, Dunbar RIM (1991) Dominance rank and mating success in male primates. Anim Behav 41:1045–1056

David P, Bjorksten T, Fowler K, Pomiankowski A (2000) Condition-dependent signalling of genetic variation in stalk-eyed flies. Nature 406:186–188

Deschner T, Heistermann M, Hodges K, Boesch C (2004) Female sexual swelling size, timing of ovulation, and male behavior in wild West African chimpanzees. Horm Behav 46:204–215

Dixson A (1983) Observations on the evolution and behavioral significance of "sexual skin" in female primates. Advances in the study of behavior 13:63–106

Domb L (2000) Sexual swellings in wild baboons (*Papio cynocephalus anubis*) at Gombe National Park. Dissertation, Harvard University, Tanzania

Domb LG, Pagel M (2001) Sexual swellings advertise female quality in wild baboons. Nature 410:204–206

Emery MA, Whitten PL (2003) Size of sexual swellings reflects ovarian function in chimpanzees (*Pan troglodytes*). Behav Ecol Sociobiol 54:340–351

Emlen DJ, Warren IA, Johns A, Dworkin I, Lavine LC (2012) A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. Science 337:860–864

Endler A, Liebig J, Schmitt T, Parker JE, Jones GR, Schreier P, Hölldobler B (2004) Surface hydrocarbons of queen eggs regulate worker reproduction in a social Insect. Proc Natl Acad Sci U S A 101:2945–2950

Engelhardt A, Hodges JK, Niemitz C, Heistermann M (2005) Female sexual behavior, but not sex skin swelling, reliably indicates the timing of the fertile phase in wild long-tailed macaques (*Macaca fascicularis*). Horm Behav 47:195–204

Fa JE, Southwick CH (1988) Ecology and behavior of food-enhanced primate groups. Liss, New York

Garamszegi LZ (2011) Information-theoretic approaches to statistical analysis in behavioural ecology: an introduction. Behav Ecol Sociobiol 65:1–11

Gesquiere LR, Wango EO, Alberts SC, Altmann J (2007) Mechanisms of sexual selection: sexual swellings and estrogen concentrations as fertility indicators and cues for male consort decisions in wild baboons. Horm Behav 51:114–125

Gillman J (1941) A quantitative study of the inhibition of the ovary and of the turgescent perineum of the normal baboon produced by a single injection of estradiol benzoate. Endocrinology 29:633–638

Girolami L, Bielert C (1987) Female perineal swelling and its effects on male sexual arousal: an apparent sexual release in the Chacma Baboon (*Papio ursinus*). Int J Primatol 8:651–661

Hausfater G (1975) Dominance and reproduction in baboons (*Papio cynocephalus*)—quantitative analysis. Contrib Primatol 7:2–150

Higham JP, Heistermann M, Ross C, Semple S, McClarnon A (2008a) The timing of ovulation with respect to sexual swelling detumescence in wild olive baboons. Primates 49:295–299

Higham JP, MacLarnon AM, Ross C, Heistermann M, Semple S (2008b) Baboon sexual swellings: information content of size and color. Horm Behav 53:452–462

Higham JP, Heistermann M, Saggau C, Agil M, Perwitisari-Farajallah D, Engelhardt A (2012) Sexual signalling in female crested macaques and the evolution of primate fertility signals. BMC Evol Biol 12:89–99

Huchard E, Benavides JA, Setchell JM, Charpentier M, Alvergne A, King AJ, Knapp LA, Cowlishaw G, Raymond M (2009a) Studying shape in sexual signals: the case of primate sexual swellings. Behav Ecol Sociobiol 63:1231–1242

Huchard E, Courtiol A, Benavides JA, Knapp LA, Raymond M, Cowlishaw G (2009b) Can fertility signals lead to quality signals? Insights from the evolution of primate sexual swellings. Proc R Soc B Biol Sci 276:1889–1997

Huchard E, Raymond M, Benavides J, Marshall H, Knapp LA, Cowlishaw G (2010) A female signal reflects MHC genotype in a social primate. BMC Evol Biol 10:960106

Jacobsen J (1991) Photoscale-2. Arcata

Kato J, Onouchi T, Oshima K (1980) The presence of progesterone receptors in the sexual skin of the monkey. Steroids 36:743–749

Kiage LM, Obuoyo J (2011) The potential link between El Nino and water hyacinth blooms in Winam Gulf of Lake Victoria, East Africa: evidence from satellite imagery. Water Resour Manag 25:3931–3945

Le Houerou H (1989) The grazing land ecosystems of the African Sahel. Springer, Berlin

Lee PC, Moss CJ (1995) Statural growth in known-age African elephants (*Loxodonta africana*). J Zool 236:29–41

Liebig J, Peeters C, Oldham NJ, Markstädter C, Hölldobler B (2000) Are variations in cuticular hydrocarbons of queens and workers a reliable signal of fertility in the ant *Harpegnathos saltator*? Proc Natl Acad Sci U S A 97:4124–4131

Massironi M, Rasotto MB, Mazzoldi C (2005) A reliable indicator of female fecundity: the case of the yellow belly in *Knipowitschia panizzae* (Teleostei: Gobiidae). Mar Biol 147:71–76

Nitsch F, Stueckle S, Stahl D, Zinner D (2011) Copulation patterns in captive hamadryas baboons: a quantitative analysis. Primates 52:373–383

Nunn C (1999) The evolution of exaggerated sexual swellings in primates and the graded-signal hypothesis. Anim Behav 58:229–246

Onouchi T, Kato J (1983) Estrogen receptors and estrogen-inducible progestin receptors in the sexual skin of the monkey. J Steroid Biochem 18:145–151

Onyango PO, Gesquiere LR, Altmann J, Alberts SC (2013) Testosterone positively associated with both male mating effort and paternal behavior in Savanna baboons (*Papio cynocephalus*). Horm Behav 63:430–436

Ozasa H, Gould KG (1982) Demonstration and characterization of estrogen receptor in chimpanzee sex skin: correlation between nuclear receptor levels and degree of swelling. Endocrinology 111:125–131

Packer C (1979) Male dominance and reproductive activity in *Papio anubis*. Anim Behav 27(Pt 1):37–45

Pagel M (1994) The evolution of conspicuous oestrous advertisement in Old World monkeys. Anim Behav 47:1333–1341

Reichert KE, Heistermann M, Hodges JK, Boesch C, Hohmann G (2002) What females tell males about their reproductive status: are morphological and behavioural cues reliable signals of ovulation in bonobos (*Pan paniscus*)? Ethology 108:583–600

Rowland WJ, Baube CL, Horan TT (1991) Signalling of sexual receptivity by pigmentation pattern in female sticklebacks. Anim Behav 42:243–249

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671–675

Symonds MRE, Moussalli A (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol 65:13–21

Tutin CEG (1979) Mating patterns and reproductive strategies in a community of wild chimpanzees (*Pan troglodytes schweinfurthii*). Behav Ecol Sociobiol 6:29–38

Van Noordwijk MA (1985) Sexual behaviour of Sumatran long-tailed macaques (*Macaca fascictclaris*). Z Tierpsychol 70:277–296

Weiss SL (2006) Female-specific color is a signal of quality in the striped plateau lizard (*Sceloporus virgatus*). Behav Ecol 17:726–732

West NB, Carlisle KS, Brenner RM (1990) Progesterone treatment suppresses estrogen receptor in the sex skin of Macaca nemestrina. J Steroid Biochem 35:481–485

Wildt DE, Doyle LL, Stone SC, Harrison RM (1977) Correlation of perineal swelling with serum ovarian hormone levels, vaginal cytology, and ovarian follicular development during the baboon reproductive cycle. Primates 18:261–270

Zahavi A (1975) Mate selection—a selection for a handicap. J Theor Biol 53:205–214

Zinner D, Alberts SC, Nunn C, Altmann J (2002) Significance of primate sexual swellings. Nature 420:142–143