Human sperm competition: testis size, sperm production and rates of extrapair copulations

LEIGH W. SIMMONS*, RENÉE C. FIRMAN*, GILLIAN RHODES† & MARIANNE PETERS†

*School of Animal Biology, University of Western Australia
†School of Psychology, University of Western Australia

(Received 10 September 2003; initial acceptance 13 October 2003; final acceptance 20 November 2003; MS. number: 7854)

We examined the claim that sperm competition is an important selection pressure operating in human populations. We recruited 222 men and 194 women to complete a survey of their sexual behaviour. Of these, 28% of men and 22% of women reported engaging in extrapair copulations (EPCs). A review of the literature suggests that rates of extrapair paternity are in the region of 2%. These values suggest that the risk of sperm competition in humans is relatively low, in line with comparative studies of relative testis sizes of humans and other primates. Testis volume was positively correlated with the number of sperm ejaculated. However, we found no support for a recent controversial claim that the within-population frequency distribution of testis size reflects a balanced polymorphism between men who specialize in sperm competition through EPCs and men who are monogamous.

When sperm from two or more males are present within the reproductive tract of a single female, there will be competition between them to fertilize available ova (Parker 1998). Sperm competition is widely recognized as a pervasive force in evolution, favouring adaptations in males for gaining fertilizations (reviewed in Birkhead & Møller 1998). Across species, increased risk (the average probability that females will copulate with more than one male) and intensity (the average number of different males with which females copulate) of sperm competition are expected to favour increased male expenditure on sperm production (Parker 1998), a pattern found in studies of a variety of taxa (Gage 1994; Møller & Briskie 1995; Hosken 1997; Stockley et al. 1997; Byrne et al. 2002). Within species, males are expected to respond to increased risk and intensity of sperm competition by adjusting the numbers of sperm ejaculated at a given mating (reviewed in Wedell et al. 2002). Furthermore, in species where males adopt alternative mating tactics, sneaks have a greater investment in their testes, and in the quantity or quality of sperm they produce (Parker 1990; Simmons et al. 1999; Vladic & Järvi 2001).

There is evidence that sperm competition might have been important in the evolutionary lineage leading to humans. Across primates, testis size is associated with mating system, with socially monogamous species or species with single-male groups having smaller testes for their body size than species with multimale groups (Harcourt et al. 1995). Sperm motility is higher in species with relatively larger testes (Anderson & Dixon 2002), and there has been rapid evolution of genes involved in sperm and seminal fluid production (Wyckoff et al. 2000). Humans have moderately sized testes for their body size, midway between those of the socially monogamous gorilla, Gorilla gorilla, and the promiscuous chimpanzee, Pan troglodytes (Harcourt et al. 1995), and rates of nucleotide changes in genes coding for sperm and seminal fluid proteins are greater in chimpanzees and humans than in gorillas (Wyckoff et al. 2000). These findings suggest that, although humans may not be subject to the levels of sperm competition experienced by chimpanzees, neither have they been free of the selection pressures that favour increased expenditure on sperm production.

In a controversial series of publications, Baker & Bellis (1993a, b, 1995) claimed that sperm competition is more important in human populations than the available data might suggest. They claimed that women seek extrapair copulations (EPCs) when they are at greatest risk of conception, that they manipulate ejaculates to ensure the success of EPC males, and, based on a sample of just 14 men, that the within-population frequency distribution of testis size reflects a balanced polymorphism between men who specialize in sperm competition through EPCs and men who are monogamous. Baker & Bellis’s (1995) claims have received much criticism (Birkhead 1995, 2000;
Barrett 1996; Birkhead et al. 1997) but there has been little attempt to test them empirically (but see Harcourt 1991; Moore et al. 1999). Baker & Bellis’s (1995) claims also attempt to test them empirically (but see Harcourt 1991; Barrett 1996; Birkhead et al. 1997) but there has been little

ANIMAL BEHAVIOUR, and examined the relation between EPC behaviour and looked for variation between men in their EPC behaviour and examined the relation between EPC behaviour and expenditure on sperm production.

METHODS

Our subjects were 194 women and 222 men recruited through advertisements displayed on the campus of the University of Western Australia; most were students. The majority of subjects attended a session in the School of Psychology where they selected a personal identification number (PIN), and their weight and height were documented before they completed a questionnaire. The questionnaire ascertained a number of basic details such as age, sex, sexual orientation (only heterosexuals were considered in our analyses) and ethnicity. The questionnaire then asked them about their lifetime sexual behaviour. Relevant to this study, subjects were asked: ‘Have you ever had sexual intercourse with a third party while in a relationship with your partner? If yes, with how many different nonpartner people (totalled across all such occasions)?’ These questions provided details on the occurrence and number of EPCs. Finally, we asked subjects to rate a series of statements from 1 (strongly agree) to 9 (strongly disagree): ‘sex without love is OK; casual sex outside of existing relationships is OK; sex on the first date is OK; I would need to know my partner emotionally and psychologically before having sex’. The last statement was reverse scored and all scores summed to provide a composite index of permissive attitude, low scores reflecting a relatively permissive attitude. This composite attitude score is highly repeatable and correlates with sexual behaviours such as lifetime number of sexual partners, numbers of one-night stands and numbers of EPCs (G. Rhodes, L. W. Simmons & M. Peters, unpublished data).

Completed questionnaires were kept in a locked box. After the session, male subjects received another questionnaire to take away with them, together with a stamped, addressed envelope for its return. In this questionnaire, men were asked to provide their PIN and measurements of the width and length of both the left and right testis. To facilitate measurement, subjects were provided with a pair of callipers. We minimized variance in measurement protocol across individuals by providing an explicit set of instructions on how to measure the length and width of their left and right testes. Testis volume was estimated as the volume of an ovoid. A sample of nine men were recruited to assess the repeatability of testis measurements. These men were asked to measure their testes twice, without reference to their previous measurements. Repeated measures ANOVAs revealed that the between-subject variance was significantly greater than within-subject variance for both left and right testis volumes (left: \(F_{(1,12)} = 49.59, P < 0.0001\); right: \(F_{(1,12)} = 36.16, P < 0.0001\)) and the repeatability estimates were high (left: 0.980; right: 0.972). The values for testis measurements were comparable to those reported from medical examinations. (Farkas 1971: \(\bar{X} \pm S.E\) left testis length: 45.2 ± 0.4 mm; width: 28.0 ± 0.2 mm; this study: self-reported left testis length: 47.6 ± 0.7 mm; width: 29.4 ± 0.6 mm). Thus, self-reported testis measurements appeared to be reliable.

Of the 222 men recruited, 50 were recruited specifically to provide a semen sample for analysis. These men attended a session in the School of Animal Biology where they completed their questionnaire and received instructions on how to provide their semen sample. Subjects were required to collect by masturbation into a sterile vial after a sexual abstinence minimum of 48 h, but no longer than 7 days (subjects were asked to note the actual duration of abstinence on the vial). Samples were delivered to the School of Animal Biology within 2.5 h of collection and were analysed according to the World Health Organization (1999) protocol. After liquefaction, the number of sperm contained within the ejaculate was determined by volumetric dilution and counting using a Neubauer haemocytometer. To assess the reliability of sperm counts, we counted the sperm in a single ejaculate, provided by each of seven men, on four occasions, twice on 2 successive days (\(F_{(6,21)} = 261.8, P < 0.001\), repeatability estimate 0.996). All means are presented ± 1SE.

The study was conducted with the approval of the University of Western Australia Human Research Ethics Committee.

RESULTS

There were no significant differences in the proportion of men and women who reported having engaged in EPCs (men: 27.9%; women: 22.2%); chi-square test: \(x^2 = 1.83, P = 0.176\) or in their total number of extrapair partners (men: 1.33 ± 0.31, range 0–50; women: 0.48 ± 0.10, range 0–15; Mann–Whitney U test, normal approximation: Z = 1.64, N1 = 222, N2 = 194, P = 0.100). As might be expected, the lifetime number of extrapair partners increased with age for both men and women (Spearman correlation: men: \(r_5 = 0.375, N = 222, P < 0.001\); women: \(r_5 = 0.362, N = 194, P < 0.001\) as did the probability that they had engaged in an EPC (logistic regression: men: Wald \(x^2 = 21.66, P < 0.001\); women: Wald \(x^2 = 20.06, P < 0.001\)). Age ranged from 17 to 51 years in our sample and did not differ significantly between men (23.7 ± 0.4 years) and women (22.9 ± 0.4); t test: \(t_{444} = 1.29, P = 0.19\).

We found considerable variation in testis size across our population (\(\bar{X} \pm S.E\) and interquartile ranges: left testis: 23.9 ± 1.3 cm³, 14.8–28.2; right testis: 24.3 ± 1.2 cm³, 16.0–28.2). General linear modelling revealed no significant variation in combined testis volume caused by male age, height, weight or ethnic grouping (our sample for
which we had testis measurements contained 103 Caucasians, 11 Asians and two Indians; whole model $F_{(5,110)} = 0.922, P = 0.469$.

We were able to account for about 40% of the variation in sperm numbers per ejaculate (whole model: $F_{(5,44)} = 5.70, P < 0.001$). Sperm numbers increased with combined testis volume (partial estimate: $1.33 \pm 0.39$; $t_{44} = 3.33, P = 0.002$; Fig. 1), increased with male age (partial estimate: $4.49 \pm 1.90$; $t_{44} = 2.36, P = 0.023$) and height (partial estimate: $2.93 \pm 1.38$; $t_{44} = 2.12, P = 0.040$) and decreased with male weight (partial estimate: $-2.26 \pm 0.99$, $t_{44} = 2.29, P = 0.027$). Sperm numbers tended to increase with the period of sexual abstinence before sampling, but this was not significant because of the 2–7-day limit imposed for this variable (partial estimate: $6.73 \pm 3.94$; $t_{44} = 1.71, P = 0.095$).

There was no significant difference in combined testis volumes between groups of men who reported engaging in EPCs and those who reported not to have engaged in such activity (men with EPCs: $45.95 \pm 4.66$ cm$^3$; men without EPCs: $49.37 \pm 3.12$ cm$^3$) regardless of whether we controlled for variables such as height, weight, age and ethnic group (raw effect: $t_{114} = 0.61, P = 0.544$; partial effect: $t_{109} = 0.56, P = 0.580$). Neither was there a relation between the number of EPC partners and testis size in men who reported EPC behaviour (Spearman correlation: $r_S = -0.096, N = 36, P = 0.306$, the highly skewed nature of the behavioural data prevented a parametric analysis of partial effects). Finally, testis size was not related to a man’s permissive attitude (raw effect: $t_{114} = -0.18, P = 0.85$; partial effect: $t_{109} = -0.39, P = 0.70$).

Men recruited to provide semen samples all also provided measures of testis size. However, only 39% of men recruited to provide data on sexual behaviour returned their testis measurements, so there is a risk that this self-selecting population of men could bias our conclusions regarding the relation between testis size and sexual behaviour. However, men who returned their testis measurements did not differ from those who did not in their attitudes towards sexual behaviour (composite attitude score for responders: $20.25 \pm 0.96$; nonresponders: $20.55 \pm 0.76$; $t_{68} = 0.24, P = 0.80$) or in their rates of EPC activity (23 of 66 responders had engaged in EPCs versus 25 of 105 nonresponders; $\chi^2 = 2.41, P = 0.120$), suggesting that men who provided testis measurements were a representative sample of our population with respect to the variables of interest.

**DISCUSSION**

A limitation of our study was that our subjects were a nonrandom sample of the general population, comprising predominantly students. However, this limitation is true of all studies that have attempted to examine human sexual behaviour in an evolutionary context so that our data are directly comparable with those of previous studies (cf. Baker & Bellis 1993a, b; Thornhill & Gangestad 1994; Baker 1997; Gangestad & Thornhill 1997). Our figures for the numbers of extrapair partners reported by men and women are remarkably similar to those reported by Gangestad & Thornhill (1997) in their survey of student couples in the U.S.A. Furthermore, even though our subjects were drawn from a student population, the proportion of subjects reporting extrapair copulations was broadly consistent with the results of previous surveys in which subjects were drawn at random from the general population (Table 1). The data in Table 1 give little indication that our student sample differed in its extrapair activity from the general population. Across populations, for the age group 18–24, 7–41% of men reported engaging in EPCs compared with 5–27% of women. In our sample the corresponding figures were 20 and 13% (Table 1). The surveys reviewed in Table 1 used slightly different methodologies. Like our study, the survey from France provided lifetime activity and, like us, showed that the probability of having engaged in EPC activity increased across increasing age groups. The remaining studies reported activity during the 12 months or 5 years before the survey. These studies suggest that EPC activity is at its highest levels in the younger age groups.

The important data from the perspective of sperm competition risk is the probability that women will engage in EPCs. The surveys from France, the U.K., the U.S.A. and now Australia suggest that across these populations, on average, about 14% of women under 30 years report one or more EPCs. Baker & Bellis’s (1995) nationwide U.K. survey reported that 6–9% of women copulated with an extrapair partner within 5 days of copulating with their main partner. However, this behaviour provides only an upper estimate of the risk of sperm competition. Sperm competition will be a selective force only where women’s behaviour has the potential to generate extrapair paternity. EPC activity outside the fertile period or the use of contraception is likely to circumvent selection arising from sperm competition. The question arises, then, as to whether EPC results in extrapair paternity. Comparative data for birds suggest a positive relation between rates of EPC and extrapair paternity but the relation has little predictive power (Birkhead & Møller 1995). The incidence

---

**Figure 1.** Relation between combined testis volume and number of sperm contained in the ejaculate of 50 men.
of sperm competition in human populations is evident from double paternity in dizygotic twins (Girela et al. 1997), which James (1993) estimated to occur in around 1 in 400 twin births in the U.S.A. Wilcox et al. (2001) estimated a likelihood of conception with a single act of intercourse ranging from 1% at the start of the menstrual cycle to 6% at the fertile period. On average, across the menstrual cycle, the probability of conception is around 3%. Assuming that the timing of copulations for partners and extrapair males occurs at random, a population-wide rate of EPC of about 20% is expected to generate a population-wide rate of extrapair paternity of about 0.6%. If, however, women time their EPC activity to coincide precisely with their most fertile period to increase the probability of extrapair paternity, as suggested by Bellis & Baker (1990), then extrapair paternity might be closer to 1.2%.

Data on extrapair paternity (or nonpaternity) in human populations are rare, and rates are often quoted that are based on hearsay, anecdote, or unpublished findings (Macintyre & Sooman 1991). Published data that can be evaluated are given in Table 2. The data come from a variety of sampling methods. Many use blood group comparisons which can exclude paternity, but cannot confirm it. Furthermore, data from families seeking genetic counselling represent nonrandom samples of the population, which may influence the result. For example, Le Roux et al. (1993) noted that their estimate may be higher than that for the general population because the subjects tested (X-fragile carriers) were reported to have more sexual partners. More recent studies have used DNA fingerprinting and were designed specifically to examine rates of extrapair paternity in the general population (Sasse et al. 1994). The rates vary across populations and are highly skewed towards low values. In general, the median rate of extrapair paternity, 1.82%, is much lower than is often cited (Macintyre & Sooman 1991), and although this value is much lower than the probability of EPC, it is expected given the likelihood of conception with a single act of intercourse around the fertile period (Wilcox et al. 2001). Traditional cultures have higher rates of extrapair paternity so that the low median value may be influenced by modern contraceptive practices. The data for the Sykes lineage in the U.K. are of interest in this regard because the study used a per generation average based on discrepancies between Y chromosome haplotypes and surnames that date back about 700 years (Sykes & Irven 2000). The rate of just 1.3% suggests that contraception may not have as great an influence as one might expect.

What can we conclude from these data regarding testis size variation in humans? If the selection pressure from sperm competition risk lies somewhere between 2%, as reflected by the extrapair paternity data, and 22%, as reflected by our behavioural data, then we should expect male expenditure on testis size to be

<table>
<thead>
<tr>
<th>Population</th>
<th>% Extrapair paternity</th>
<th>N</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michigan, U.S.A.</td>
<td>1.4</td>
<td>1417</td>
<td>Schacht &amp; Gershowitz 1963</td>
</tr>
<tr>
<td>Detroit, U.S.A.</td>
<td>10.1</td>
<td>523</td>
<td>Potthoff &amp; Whittinghill 1965</td>
</tr>
<tr>
<td>Oakland, California</td>
<td>0.21</td>
<td>265</td>
<td>Peritz &amp; Rust 1972</td>
</tr>
<tr>
<td>Hawaii</td>
<td>0.03</td>
<td>6960</td>
<td>Ashton 1980</td>
</tr>
<tr>
<td>France</td>
<td>2.8</td>
<td>89</td>
<td>Le Roux et al. 1992</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2.7</td>
<td>1607</td>
<td>Sasse et al. 1994</td>
</tr>
<tr>
<td>West Middlesex, U.K.</td>
<td>5.9</td>
<td>2596</td>
<td>Edwards 1957</td>
</tr>
<tr>
<td>Sykes family, U.K.</td>
<td>1.3</td>
<td>269</td>
<td>Sykes &amp; Irven 2000</td>
</tr>
<tr>
<td>U.K.</td>
<td>1.4</td>
<td>521</td>
<td>Brock &amp; Shrimpton 1991</td>
</tr>
<tr>
<td>Nuevo León, Mexico</td>
<td>11.8</td>
<td>396</td>
<td>Cerda-Flores et al. 1999</td>
</tr>
<tr>
<td>South America, Yanamamo Indians</td>
<td>10.0</td>
<td>132</td>
<td>Chagon 1979</td>
</tr>
</tbody>
</table>
greater than for a monogamous species (see figure 1 in Parker et al. 1997), consistent with the position of human relative testis size lying between those of gorillas and chimpanzees (Harcourt et al. 1995). But what about variation in testis size within humans? Testis size varies widely, and consistent differences have been reported between ethnic groups (Diamond 1986). In our study, self-reported measures of testis size (mean and ranges) did not differ from those reported in studies where the authors themselves conducted the measurements (cf. Farkas 1971; Handelsman et al. 1984; Taskinen et al. 1996; Ku et al. 2002), indicating that self-measurement provided reliable data. After controlling for potentially confounding variables, we found that men with larger testes ejaculated greater numbers of sperm. Thus, larger testes could potentially bestow a selective advantage in sperm competition, as has been found for a number of nonhuman species (references in Birkhead & Möller 1998). We can therefore use our data to test Baker & Bellis's (1995) claim that the within-population frequency distribution of testis size reflects a balanced polymorphism between men who specialize in sperm competition through EPCs and men who are monogamous. They asked 20 independent ‘judges’ how likely 14 men were to pursue a sperm competition strategy. They reported a positive correlation between this index of sperm competition and testis size, which, notwithstanding the subjectivity of the sperm competition index, could have been confounded by any number of variables, including age, body weight or ethnicity (Birkhead 1995; Barrett 1996). In a follow-up study, Baker (1997) compared the testis sizes of men who were rumoured (although not known) to have engaged in EPCs during an undergraduate field course with those who were rumoured not to have engaged in EPCs, and claimed support for his hypothesis. We found no significant difference in the combined testes volume of men who reported having engaged in EPCs and those who did not. Indeed, the trend was in the direction opposite to that predicted. Neither was there a relation between the numbers of EPC partners reported by men who had engaged in EPC activity and their combined testis volume, or between testis volume and men’s attitudes towards permissive sexual behaviour.

In conclusion, the collective data suggest that the risk of sperm competition in modern human populations is relatively low compared to that for other nonhuman taxa, with rates of EPCs of about 5–27% for people under 30 years old and rates of extrapair paternity of just 2%. Our conclusions are consistent with the position of humans in comparative analyses of testis size across primates. Furthermore, we found no evidence to support the claim that human males adopt alternative reproductive strategies of sperm competition and monogamy.

Acknowledgments

We thank John Alcock for helpful comments and encouragement. This work was supported by the Australian Research Council.

References


