LETTER

Sexual selection protects against extinction

Alyson J. Lumley[1]*, Łukasz Michalczyk[2]*, James J. N. Kitson[1], Lewis G. Spurgin[1], Catriona A. Morrison[1], Joanne L. Godwin[1], Matthew E. Dickinson[1], Oliver Y. Martin[3], Brent C. Emerson[4], Tracey Chapman[1] & Matthew J. G. Gage[1]

Reproduction through sex carries substantial costs, mainly because only half of sexual adults produce offspring[1]. It has been theorized that these costs could be countered if sex allows sexual selection to clear the universal fitness constraint of mutation load[2–4]. Under sexual selection, competition between (usually) males and mate choice by (usually) females create important intraspecific filters for reproductive success, so that only a subset of males gains paternity. If reproductive success under sexual selection is dependent on individual condition, which is contingent to mutation load, then sexually selected filtering through 'genic capture'[5] could offset the costs of sex because it provides genetic benefits to populations. Here we test this theory experimentally by comparing whether populations with histories of strong versus weak sexual selection purge mutation load and resist extinction differently. After evolving replicate populations of the flour beetle *Tribolium castaneum* for 6 to 7 years under conditions that differed solely in the strengths of sexual selection, we revealed mutation load using inbreeding. Lineages from populations that had previously experienced strong sexual selection were resilient to extinction and maintained fitness under inbreeding, with some families continuing to survive after 20 generations of sib × sib mating. By contrast, lineages derived from populations that experienced weak or nonexistent sexual selection showed rapid fitness declines under inbreeding, and all were extinct after generation 10. Multiple mutations across the genome with individually small effects can be difficult to clear, yet sum to a significant fitness load; our findings reveal that sexual selection reduces this load, improving population viability in the face of genetic stress.

Sexual selection is a widespread evolutionary force giving rise to a striking diversity of sights, sounds and smells that filter reproductive success away from less competitive or attractive individuals, frequently at the expense of survival[6]. Sexual selection will operate to varying degrees whenever sexual reproduction exists, and its significance as a potent force profoundly influencing reproductive fitness of individuals is long established[6]. In contrast, limited empirical work has been directed at measuring the consequences of sexual selection for the fitness of populations. This lack of attention is surprising for two reasons: first, because population viability is vital for biodiversity maintenance and ecosystem stability, especially under modern anthropogenic stress[7,8]; and second, because it is predicted that the maintenance of costly sex as the dominant mode of reproduction might only be possible if it allows sexual selection to operate, reducing the universal handicap of mutation load[3,4].

Population or lineage health will always suffer at some level from mutation load—the difference in fitness between a (usually theoretical) mutation-free lineage and one that carries a load of deleterious mutations that are segregating in mutation-selection balance[7,9]. This load exists because new mutations continually arise in all populations every generation, most of which will be deleterious[10]. Haldane calculated that mutation load would be unexpectedly high[11], owing to large numbers of loci across each genome presenting multiple targets to unavoidable

mutation rates. Empirical estimates of mutation rate and load are hotly debated[7], but we know, for example, that the average human lineage carries hundreds of deleterious loss-of-function mutations[12], possibly thousands[7]. Natural selection will quickly remove mutations with large effects, but load persists through accumulation of mutations that have small individual effects, and/or exist as recessive alleles where their deleterious phenotypes are less frequently exposed to selection. Mutation load can therefore sum to a significant fitness constraint for a population which, because of its dispersed or concealed nature, is difficult to clear[7,11]. Sex could allow more effective purging of mutation load than asexual reproduction, if there are synergistic epistatic interactions between mutations[7], so that their negative fitness impact is greater than strictly additive. However, the evidence that sufficient levels of synergistic epistasis exist remains equivocal[7], so alternative explanations for the maintenance of costly sexual reproduction are sought.

Sexual selection could be a key filter against mutation load if, as Darwin acknowledged, "sexual selection will have given its aid to ordinary selection"[13]. This recognition that sexual selection places an additional, intraspecific filter on adaptive gene flow has been formalized by the idea of 'genic capture'[5], which proposes that reproductive success in the face of competition and choice depends on most or all aspects of an individual's condition. Thus, sexual selection will act on most loci across the genome, purifying deleterious alleles from individuals within a lineage, and promoting fixation of advantageous ones, via three potentially connected routes. First, if competition or choice promotes non-random mating or fertilization success as an inverse consequence of mutation load, then deleterious dominant and recessive alleles will be under stronger purifying selection in lineages that experience sexual selection[3]. Second, if mutations are more deleterious in the competing sex than the choosing sex, which would occur if the competing sex suffers amplified reproductive success variance as a result of mutation load, then the equilibrium frequency of both dominant and recessive deleterious mutations will be lower within sexually selected populations, even accounting for synergistic epistasis in the competing sex[2]. Put simply, fathers should carry fewer mutations than all males[3]. Third, if female ability to maintain condition for reproductive fitness is under selection from interlocus sexual conflict, then sexual selection could purge mutations through females[14,15]. These routes for purging load via sexual selection provide theoretical explanations for how costly sex can persist as a dominant mode of reproduction[2,3].

Few experiments have tested whether sexual selection removes deleterious mutations, yielding inconsistent findings. After introduction of mutations, heightened sexual selection rescued fitness faster in *Rhizoglyphus robini* mites[16], *Drosophila melanogaster* flies[17] and *Onthophagus taurus* beetles[18]. However, further work showed that sexual selection reduced mutation loads only for a subset of fitness traits in *Drosophila*[19]. More detailed studies found no evidence that sexual selection could purge deleterious alleles from experimentally evolving lineages[20], or restore fitness after mutation load had been induced, even after 60 generations[21]; in fact, reproductive fitness within

mutated lines became higher when sexual selection had been minimized[21]. These inconsistent results could be explained by interlocus sexual conflict[22]: while sexual selection might play a beneficial role in purging mutation load, direct short-term constraints on population productivity may also arise when female fitness is constrained by conflicting adaptations that promote only male reproductive potential[22]. Thus, short-term fitness costs arising simultaneously from interlocus sexual conflict could confound the measurement of longer-term fitness benefits arising from sexual selection[7]. To avoid this problem, we assayed mutation load after populations (from the same ancestry) had experienced almost 7 years of sole variation in the intensity of sexual selection (Extended Data Figs 1 and 2). Although no evidence for carry-over effects of interlocus sexual conflict exists in our model under these conditions[15], we also removed potential trans-generational effects by enforcing two generations of monogamous reproduction before beginning to assay fitness. To expose mutation load in each population, we enforced inbreeding using sib × sib pairings, tracking fitness changes down multiple family lines (Extended Data Fig. 3). Because mutations will more likely persist as recessive alleles that are less frequently exposed to selection, compared with dominant wild types, fitness depression is a normal consequence of inbreeding as homozygosity increases the expression frequency of deleterious recessives through partial dominance[23]. By tracking extinction rates and fitness declines down inbreeding lineages perpetuated by monogamous sibling pairings, we tested the hypothesis that sex allows sexual selection to generate significant benefits through the purging of mutation load[2,3], while avoiding the concurrent confound of interlocus sexual conflict[22].

Experimental evolution lines began in 2005, with two different regimes (A and B) that both exposed replicate lineages to either 'strong' or 'weak' treatments of sexual selection, providing parallel independent experiments to measure whether sexual selection purges mutation load. Regime A exposed populations to 54 generations of divergent adult operational sex ratios that were either male-biased (10 females to 90 males) versus female-biased (90 females to 10 males), while regime B allowed polyandry (1 female to 5 males) versus enforced monogamy (1 female to 1 male) for 45 generations. All other conditions among lines and treatments within a regime were kept identical, including equalizing theoretical effective population sizes (see Methods, Extended Data Figs 2 and 4 and Supplementary Information for details confirming that heterozygosity was identical between treatments within either regime). The only difference between 'strong' and 'weak' treatments was therefore intensity of sexual selection during adult reproduction.

Our findings clearly showed for both regimes that lineages derived from populations experiencing evolutionary histories of strong sexual selection resisted extinction and maintained fitness more effectively when mutation load was exposed (Figs 1 and 2 and Extended Data Figs 5–7; all means presented in this article are arithmetic means ± standard errors). Over the 3-year extinction assay, families derived from male-biased populations survived 44% longer than families from female-biased populations (mean number of generations to extinction = 9.24 (±1.29 s.e.m.) versus 6.46 (±0.15); Fig. 1a). Families from polyandrous histories survived 37% longer on average than those from monogamous treatments (8.50 (±1.02) generations versus 6.21 (±0.46); Fig. 1b). When we combined data from regimes A and B into one analysis (incorporating sexual selection as a fixed factor variable), the history of sexual selection remained a significant predictor of number of generations to extinction ($z = -3.43$, $P < 0.001$), and there was no difference in extinction rates between regimes A and B ($z = -0.51$, $P = 0.611$), nor a significant interaction ($z = 0.44$, $P = 0.660$), revealing a consistent effect of sexual selection on extinction rates (Extended Data Fig. 5c). Overall, families derived from populations evolving under histories of strong sexual selection survived for 40% longer than those derived from weak sexual selection histories, giving an average survival time under mutation load exposure of 8.87
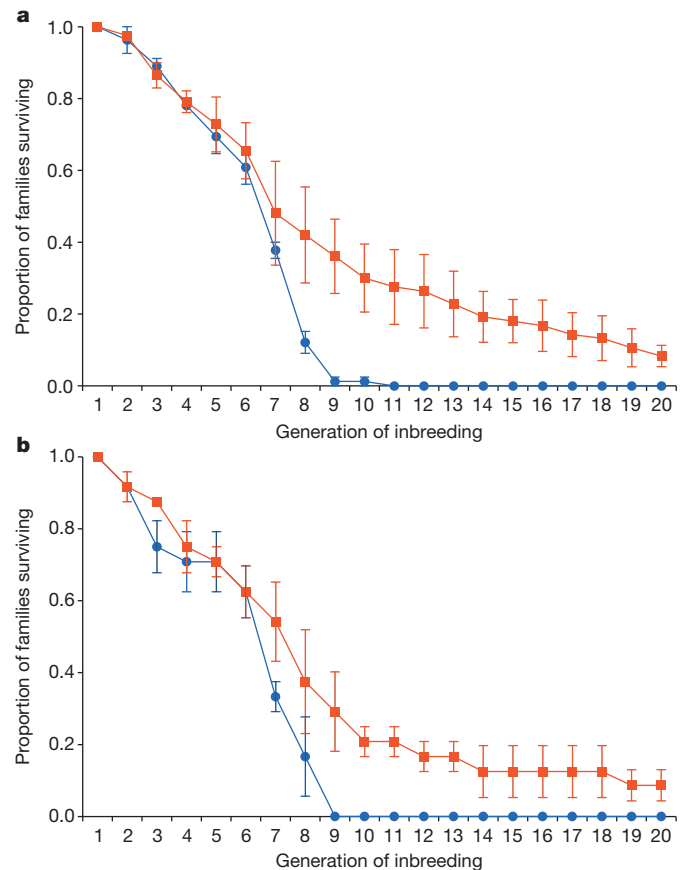
Figure 1 | Extinction trajectories under increasing inbreeding differ between family lines derived from strong (red squares) versus weak (blue circles) sexual selection histories. Each generation mean presents the average (±s.e.m.) proportion of surviving families for three independent lines per treatment. **a**, Regime A: male-biased (red) versus female-biased (blue) sexual selection histories. Each line is represented by 28 initial families ($n = 84$ total families for either treatment). **b**, Regime B: polyandrous (red) versus monogamous (blue) selection histories. Each line is represented by eight initial families ($n = 24$ total families for either treatment). Using parametric accelerated failure-time survival models with sexual selection treatment as a fixed effect, and incorporating correlated data within lines using a generalized estimating equation approach, we identified significantly lower extinction rates in populations that had previously experienced strong histories of sexual selection (Extended Data Fig. 5a, regime A: $z = 3.40$, $P < 0.001$; Extended Data Fig. 5b, regime B: $z = 2.81$, $P = 0.005$).

(±0.37) versus 6.33 (±0.13) generations respectively. All 108 initial families derived from weak sexual selection histories ceased to produce offspring beyond the 10th generation of inbreeding, whereas 8 of the 108 families from the strong sexual selection histories were still producing offspring after 20 generations of inbreeding.

Declines in reproductive fitness exhibited similar patterns to the extinction rates they underpinned. Having removed sexual selection history via enforced monogamy in regime B, baseline reproductive fitness even without inbreeding is substantially reduced (Fig. 2b shows that a polyandrous history improves baseline fitness by ~30%). Most importantly, significant interactions between inbreeding and fitness decline revealed that fitness declines were much faster for families derived from treatments experiencing weak sexual selection histories in both regimes ($z = 13.82$ and 10.56 for A and B respectively, $P < 0.001$ for both; Fig. 2, Extended Data Table 1 and Extended Data Figs 6 and 7). To control for the possibility that failure-to-mate could have created differences between treatments, we repeated these analyses including only pairs that produced offspring: identical patterns of significant interactions between sexual selection history,
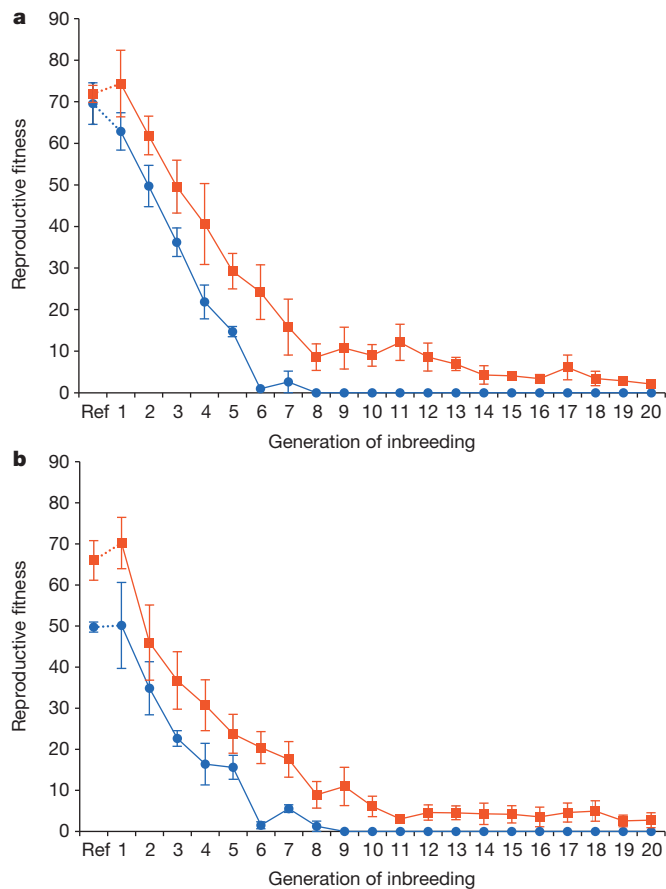
Figure 2 | Reproductive fitness declines under increasing inbreeding of families derived from strong (red squares) versus weak (blue circles) sexual selection histories differ in magnitude and rate. a, Regime A: male-biased (red) versus female-biased (blue) selection histories. **b**, Regime B: polyandrous (red) versus monogamous (blue) histories. Each generation mean presents the average ($\pm$s.e.m.) number of offspring produced under standardized conditions for three independent lines per treatment. Lines are represented by eight initial families ($n = 24$ total families per treatment), and two breeding pairs per family. Average fitness under identical conditions but without inbreeding ('Ref') plotted for reference. Using generalized linear mixed models, and accounting for overdispersion as well as nesting non-independent replicate families within lines as random effects, we found in both regimes that inbreeding was a highly significant predictor of fitness (regime A: $z = -21.34$, $P < 0.001$; regime B: $z = -21.17$, $P < 0.001$), and that sexual selection treatment history was also a significant predictor of overall differences in fitness (regime A: $z = -3.75$, $P < 0.001$; regime B: $z = -1.97$, $P = 0.048$; see Extended Data Table 1 and Extended Data Figs 6 and 7).

inbreeding and fitness decline remained ($z = 5.73$ and 6.36 for A and B respectively, $P < 0.001$ for both; Extended Data Figs 6c, d and 7c, d).

Results from this 10-year experiment provide compelling empirical support for the complementary models given in refs 2 and 3, which argued that costs of sex[1] could be offset by population genetic benefits derived from sexual selection. The most obvious mechanism explaining these differences is that heightened sexual selection, via 'genic capture'[4,5], more effectively strips out, or prevents, mutation load from becoming fixed in a population, strengthening its ability to withstand stress. Hemizygous selection could make sex-linked loci especially prone to this process[24], but since the *T. castaneum* X chromosome makes up only ~6% of the genome[25], we expect purging of load to have occurred on the autosomes too. There is emerging evidence that sexual selection can profoundly shape the genome, with feminized patterns of sex-linked expression in the transcriptomes of both sexes after experimental evolution under monogamy in *Drosophila*[26]. Our findings indicate that sexual selection also acts to purge mutation load,

even in populations that we expect to be adapting close to their natural fitness peaks[27]. Stronger sexual selection will drive greater variance in reproductive success, so that the average father should carry fewer deleterious mutations than the average male[3], and perhaps also the average mother if interlocus sexual conflict constrains female fitness[15].

Within the promiscuous *Tribolium* model[28] (Extended Data Fig. 8 and Methods), both pre- and post-copulatory processes for winning or controlling fertilizations will operate through broad behaviours and physiology within the whole organism, down to competition and choice at the gamete level. Genic capture predicts that individual success across this continuum will depend on overall condition and genes at many loci[5]. Under monogamy, where sexual selection has been removed, populations suffer the constraints of sexual reproduction, but none of the benefits of sexual selection, explaining the lowered baseline fitness even without inbreeding, as well as more rapid fitness declines and heightened extinction rates under inbreeding (Figs 1 and 2 and Extended Data Figs 5–7). Although we exposed mutation load for experimental measurement through inbreeding, the intense and diverse demands from selection across multiple generations in the natural environment, where populations are more likely to be displaced from their fitness peaks[27], are likely to expose even greater fitness differentials due to load variance. After only one generation of inbreeding, for example, mean fitness between strong versus weak sexual selection histories differs by 20% and 40% in regimes A and B respectively (Fig. 2). Our results indicate, as demands rise upon populations that have been depleted and fragmented, and displaced further from their fitness peaks[27] via increasing anthropogenic stress[29], that sexual selection could be an important force protecting species or populations from the extinction vortex[30].

1. Maynard Smith, J. *The Evolution of Sex* (Cambridge Univ. Press, 1978).
2. Agrawal, A. F. Sexual selection and the maintenance of sexual reproduction. *Nature* **411,** 692–695 (2001).
3. Siller, S. Sexual selection and the maintenance of sex. *Nature* **411,** 689–692 (2001).
4. Whitlock, M. C. & Agrawal, A. F. Purging the genome with sexual selection: reducing mutation load through selection on males. *Evolution* **63,** 569–582 (2009).
5. Tomkins, J. L., Radwan, J., Kotiaho, J. S. & Tregenza, T. Genic capture and resolving the lek paradox. *Trends Ecol. Evol.* **19,** 323–328 (2004).
6. Andersson, M. B. *Sexual Selection* (Princeton Univ. Press, 1994).
7. Agrawal, A. F. & Whitlock, M. C. Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. *Annu. Rev. Ecol. Syst.* **43,** 115–135 (2012).
8. Holman, L. & Kokko, H. The consequences of polyandry for population viability, extinction risk and conservation. *Phil. Trans. R. Soc. B* 368, http://dx.doi.org/10.1098/rstb.2012.0053 (2013).
9. Crow, J. F. Mutation, mean fitness, and genetic load. *Oxf. Surv. Evol. Biol.* **9,** 3–42 (1993).
10. Keightley, P. D. & Lynch, M. Toward a realistic model of mutations affecting fitness. *Evolution* **57,** 683–685 (2003).
11. Haldane, J. B. S. The effect of variation of fitness. *Am. Nat.* **71,** 337–349 (1937).
12. McVean, G. A. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491,** 56–65 (2012).
13. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
14. Wigby, S. & Chapman, T. Female resistance to male harm evolves in response to manipulation of sexual conflict. *Evolution* **58,** 1028–1037 (2004).
15. Michalczyk, Ł. *et al.* Experimental evolution exposes female and male responses to sexual selection and conflict in *Tribolium castaneum*. *Evolution* **65,** 713–724 (2011).
16. Radwan, J. Effectiveness of sexual selection in removing mutations induced with ionizing radiation. *Ecol. Lett.* **7,** 1149–1154 (2004).
17. Hollis, B., Fierst, J. L. & Houle, D. Sexual selection accelerates the elimination of a deleterious mutant in *Drosophila melanogaster*. *Evolution* **63,** 324–333 (2009).
18. Almbro, M. & Simmons, L. W. Sexual selection can remove an experimentally induced mutation load. *Evolution* **68,** 295–300 (2014).
19. Sharp, N. P. & Agrawal, A. F. Mating density and the strength of sexual selection against deleterious alleles in *Drosophila melanogaster*. *Evolution* **62,** 857–867 (2008).

20. Arbuthnott, D. & Rundle, H. D. Sexual selection is ineffectual or inhibits the purging of deleterious mutations in *Drosophila melanogaster*. *Evolution* **66,** 2127–2137 (2012).
21. Hollis, B. & Houle, D. Populations with elevated mutation load do not benefit from the operation of sexual selection. *J. Evol. Biol.* **24,** 1918–1926 (2011).
22. Holland, B. & Rice, W. R. Experimental removal of sexual selection reverses intersexual antagonistic coevolution and removes a reproductive load. *Proc. Natl Acad. Sci. USA* **96,** 5083–5088 (1999).
23. Charlesworth, B. & Charlesworth, D. The genetic basis of inbreeding depression. *Genet. Res.* **74,** 329–340 (1999).
24. Rice, W. R. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38,** 735–742 (1984).
25. Kim, H. S. *et al.* BeetleBase in 2010: revisions to provide comprehensive genomic information for *Tribolium castaneum. Nucleic Acids Res.* **38,** D437–D442 (2010).
26. Hollis, B., Houle, D., Yan, Z., Kawecki, T. J. & Keller, L. Evolution under monogamy feminizes gene expression in *Drosophila melanogaster. Nature Commun.* 5, http://dx.doi.org/10.1038/ncomms4482 (2014).
27. Long, A. F. T., Agrawal, A. F. & Rowe, L. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Curr. Biol.* **22,** 204–208 (2012).
28. Michalczyk, Ł. *et al.* Inbreeding promotes female promiscuity. *Science* **333,** 1739–1742 (2011).
29. Lande, R. Genetics and demography in biological conservation. *Science* **241,** 1455–1460 (1988).
30. Gilpin, M. E. & Soulé, M. E. in *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soulé, M. E.) 19–34 (Sinauer Associates, 1986).

Supplementary Information is available in the online version of the paper.

Author Contributions Ł.M., O.Y.M. and M.J.G.G. initiated the experimental evolution lines used in this work in 2005 and, with A.J.L., have maintained them since. M.J.G.G., Ł.M. and A.J.L. conceived, designed, conducted and analysed the study, with input from B.C.E. and T.C. J.J.N.K. and L.G.S. ran the microsatellite analyses. J.L.G., M.E.D. and O.Y.M. helped with line maintenance and experimental data collection. C.A.M. performed the fitness analyses. M.J.G.G. and A.J.L. wrote the paper, with contributions from all authors.

Author Information Data sets for all experiments and assays have been deposited in the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.86750. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.J.G.G. (m.gage@uea.ac.uk).

METHODS

Experimental evolution lines. Beetles were of the widely used Georgia-1 (GA1) 'wild type' strain, originally collected from stored corn in 1980, and since cultured by the Beeman laboratory (United States Department of Agriculture), maintained under standard conditions. Adult virgin beetles from the same ancestral GA1 population were randomly allocated to begin their respective treatments within both regimes. Each generation, male and female pupae were separated and placed in fresh fodder (organic flour, yeast (10:1) and oats) for 10 days to allow adult emergence and sexual maturation[15,28,31]. Then, using the controlled sex ratios in the different sexual selection treatments, mature adults were placed in *ad libitum* fodder for 7 days to compete, choose, mate, oviposit and therefore reproduce under divergent intensities of sexual selection[15,28,31]. After 7 days, adults were removed and the eggs and larvae (typically about 70 offspring per female) were left to develop under standardized conditions with equal offspring densities by maintaining *ad libitum* levels of food in proportion to the number of offspring. These conditions were maintained for about 7 years, allowing the application of strong versus weak sexual selection in a total of 12 independent lines through sole variation in the adult operational sex ratio, while equalizing effective population size within regimes. Two parallel experimental evolution regimes were run, regimes A and B, which both applied treatments that created contrasting divergences in sexual selection intensities (see Extended Data Fig. 2 for visual details). Under regime A, the male-biased treatment provided 10 females with choice among 90 males, and simultaneously competition between the 90 males to fertilize the 10 females, whereas in female-biased treatments the reverse scenarios applied. Six independent lines (three per treatment) were maintained for 54 generations under regime A. Regime B generated divergence by enforcing either monogamy (20 replicate pairs per line), or allowing reproduction to be achieved through competition between five males under polyandry (12 replicate groups per line). Thus, polyandry provided each female with a choice among five males, and competition between five males for just one female, whereas monogamy completely removed all female choice and male–male competition. Six independent lines (three per treatment) were maintained for 45 generations under regime B. After each adult reproduction period and sexual selection treatment, offspring from replicate families within each independent line were pooled under standardized densities for larval development and genetic mixing to the next generation.

Quantifying sexual selection intensity. Divergence in the strength of sexual selection between the different treatments was estimated using a male mating potential assay, in which males from the ancestral GA1 control stock population were each provided with a series of 84 unmated females. Individual were sexed as pupae and matings took place after 10 days of adult maturation (as per the sexual selection treatments). Males were placed individually in microcentrifuge tubes with approximately 1 ml of flour topped with oats; 24 h before their trial they were identified with a small dot of white correction fluid on their thorax[15,28,31]. Females were placed in groups of six per Petri dish with 7 g of flour topped with oats. Mating assays were performed by placing individual males into Petri dishes containing standard fodder and six females, moving males to a new group of six females every 12 h for 7 days (12 × 7 = 84 females given to each male over the week). After each 12 h mating period, females were placed singly into Petri dishes with standard fodder for 7 days to lay eggs. Any offspring were then allowed to develop for 35 days before being frozen, and their presence noted to score successful matings. Eleven males were assayed, and the average number of successfully fertilized females across this 1 week period per male was 50 (±3 s.e.m.; Extended Data Fig. 8), providing extreme divergence in potential levels of female polyandry (and therefore sexual selection) between strong and weak sexual selection treatments. Thus, in regime A the 10 females at each adult generation in the male-biased treatment were potentially exposed to 90 × 50 = 4,500 successful matings, or 450 matings per female per week involving all 90 males. Whereas the 90 females in the female-biased treatment potentially experienced 500 matings (10 × 50), or 5.6 per female per week, enabling each female to be mated, on average, by 5.6 of the males available. In the monogamous treatment, we removed sexual selection altogether as only one male was available for mating, while the poly-androus treatment generated a fivefold increase in potential for pre- and post-copulatory male–male competition and female choice through the availability of five males, all of which could mate with the female.

Effective population sizes. To avoid differential inbreeding that could have subsequently influenced our extinction assay, we equalized the theoretical effective population size, N_e, for mixed adult sex ratios ($N_e = \frac{4N_fN_m}{N_f + N_m}$) (ref. 32) in our divergent sexual selection treatments; thus, in regime A $N_e = 36$ for both male- and female-biased treatments, and in regime B $N_e = 40$ for both monogamous and polyandrous treatments. It is important to note that strong sexual selection treatments with male-biased or polyandrous structures may translate into reduced realized N_e due to male success skew under strong sexual selection (where competition and choice could allow reproductive success by a smaller subset of males

in the breeding population). Thus, our experimental design is conservative to any influence of N_e, because we test the prediction that lineages with histories of strong sexual selection (and therefore potentially lower N_e) should demonstrate resistance to extinction and reduced rates of fitness decline under inbreeding (because they have more effectively purged mutation load). If N_e is lower in these strong sexual selection populations, we would expect this to result in reduced heterozygosity at the start of inbreeding, resulting in lower initial fitness and therefore faster extinction rates (yet we observe the converse in Figs 1 and 2). As an additional, third, check for N_e confounds, we also directly measured heterozygosity and allelic richness in our lines to establish that differential inbreeding had not occurred under different sexual selection treatments (methods outlined below). By screening 628 individuals representing the different sexual selection treatments in both regimes A and B at 13 microsatellite loci (Supplementary Information), we were able to confirm that heterozygosity and allelic richness showed no differences between 'weak' or 'strong' sexual selection histories (Extended Data Fig. 4).

Exposing mutation load. Having been subjected to experimental evolution that applied strong versus weak sexual selection under equal effective population sizes, we then exposed the mutation load carried in these populations by inbreeding down multiple replicate family lines to expose deleterious recessives, and tracking fitness decline to extinction as a result of inbreeding depression through partial dominance[23]. No statistical methods were used to predetermine sample size, which was fundamentally controlled by the number of independent selection lines that had been maintained for 6 to 7 years. At the end of experimental evolution, we created multiple full-sib families using monogamous crosses of randomly chosen unmated males and females within each independent selection line, which simultaneously removed any transgenerational interlocus sexual conflict effects, and from which full-sib offspring were then used to continue a total of 216 family lines down increasing inbreeding coefficients and homozygosity (Extended Data Fig. 3). There were thus two generations of monogamy applied before fitness assays began, eliminating any carry-over effects of sexual selection or interlocus conflict. As levels of inbreeding increased, we tracked extinction rates and fitness declines for multiple families within each of the 12 independent lines from the original sexual selection treatments. In regime A, we created 28 full-sib families in each independent line ($N_{total} = 28_{families} \times 3_{independent\ lines} \times 2_{sexual\ selection\ treatments} = 168$ families), allowing extinction rates to be measured within each of the three independent lines in either of the two sexual selection treatments. When sibling pairs failed to produce female and male offspring that survived to adulthood to enable continuation of the line, it was recorded as extinct (see Extended Data Fig. 3). In addition to extinction, we also measured reproductive fitness in a subset of these families, assayed as the average number of offspring produced by two randomly chosen sib × sib pairs within eight of the families per line (Extended Data Fig. 3). Identical protocols for regime B were followed, except that both extinction rates and average fitness were measured across eight families per independent line ($N_{total} = 8_{families} \times 3_{independent\ lines} \times 2_{sexual\ selection\ treatments} = 48$ families). Blinding was not performed as the protocol did not permit biasing to affect results. Inbreeding was continued for 20 generations of sib × sib matings (which operated over 3 years), by which time Wright's inbreeding coefficient F had increased to 0.986. After 20 generations of inbreeding, 208 of the 216 initial lines had gone extinct, with the eight survivors all derived from strong sexual selection histories.

Baseline fitness. At four time-points throughout the extinction assay (parental, F_7, F_{15} and F_{21} generations), we assayed 'baseline' fitness (without inbreeding) of standard monogamous crosses from the different sexual selection treatments, using identical protocols to fitness measures applied for the inbred crosses. Pupae from each of the three independent lines within each of the four different selection treatments were separated into single-sex groups and isolated for 10 days to allow adult eclosion and sexual maturation. Twenty randomly chosen male–female pairs were then established for each of the three independent lines within each of the four sexual selection treatments. Each pair was placed into a 7 ml plastic vial, with *ad libitum* food and oats, and allowed to mate and oviposit for 7 days. After this, the flour from each vial, containing the eggs laid during the 7 day mating period, was transferred to a Petri dish containing a further 10 g of fodder, and the eggs allowed to develop to adult eclosion (~35 days). Fitness was scored as the total number of adults produced from each cross across 7 days of mating and oviposition. Data from previous work[15] where male–female pairs were allowed to interact for 7 days shows (a) that the first week of oviposition produces >25% of the total female reproductive fitness for such male–female interactions, and (b) that female fitness over the first week significantly correlates with total female reproductive fitness ($r = 0.48$, $P = 0.008$, $n = 29$ pairs). Average baseline fitness values are presented for reference in Fig. 2a, b ('Ref').

Extinction analyses. Extinction rates were analysed using the 'survival' package[33] in R version 3.1.0 (ref. 34). To assess whether sexual selection history influenced extinction rate after 20 generations of sib × sib inbreeding for either regimes A or B, we analysed generation to extinction using a parametric accelerated failure-time

survival model with Weibull baseline hazard distribution, taking into account right-censored data from families that were still alive at the end of the experiment. Generation of extinction was modelled with sexual selection treatment as a fixed effect, and the shape and scale of the underlying Weibull hazard distribution was allowed to vary by treatment. We also ran a combined analysis where both experimental evolution regimes were incorporated into the model, with level of sexual selection (strong versus weak), comparison (regime A or B) and their interaction as fixed effects. In all survival analyses, correlated data within lines were incorporated using a generalized estimating equation approach. Akaike's information criteria and graphical interpretation of the complementary log–log survival plots were used to confirm appropriateness of the specified Weibull distribution; no violations of model assumptions were detected. Model fits are presented in Extended Data Fig. 5a–c.

Reproductive fitness analyses. We tested for a relationship between fitness and generation of inbreeding by fitting generalized linear mixed models (GLMMs) using the glmmADMB package[35] in R 3.0.3 (ref. 34). We assumed a negative binomial error structure which added the parameter (k) to the variance mean relationship, allowing us to account for the overdispersion in our data introduced by a large number of zero fitness observations. GlmmADMB has two options for fitting the relationship between the mean and the variance. These are family = 'nbinom1', which assumes the variance = $k \times$ mean, and family = 'nbinom2', which assumes the variance = mean($1 + $ mean/k). We fitted models with each of these options in turn and then selected the parameterization that provided the lowest AIC score and therefore accounted for the greatest amount of variance in our data set. To account for the nested nature of the experimental design and the non-independence between replicate families, random effects were included in the model as replicate nested within family nested within line. Separate GLMMs were fitted to the regime A (male-biased versus female-biased) comparison and the regime B (polyandry versus monogamy) comparison (Extended Data Figs 6 and 7). To ensure that failure-to-mate was not a reason for differences in fitness declines between treatments, we repeated the analyses including only those families that produced some offspring each generation (Extended Data Figs 6c, d and 7c, d).

Microsatellite analyses. A total of 628 individuals from the 12 independent lines representing all four sexual selection treatments (strong versus weak in regimes A and B) were genotyped at 13 loci conforming to Hardy–Weinberg equilibrium (Supplementary Information). Scored genotypes for all individuals were analysed in Arlequin version 3.5 (ref. 36) to obtain allelic richness and observed and expected heterozygosity. We tested for differences in heterozygosity across selection treatments using linear mixed models, implemented in lme4 version 1.1-6 (ref. 37) package in R[34], with significance testing performed using the package lmerTest version 2.0-6 (ref. 38). Heterozygosity was included as a response variable with selection treatment as a fixed factor and locus as a random factor. To assess whether independent line identity had a significant effect on genetic diversity beyond that of selection treatment, a second mixed model was tested with the same variables plus line identity as an additional random factor. The two models were then compared using likelihood ratio tests. All model residuals were tested for normality and no violations of model assumptions were found. Estimated heterozygosity from the linear mixed effect models and the associated standard error is plotted in Extended Data Fig. 4. Microsatellite data are available from the Dryad Digital Repository at http://dx.doi.org/10.5061/dryad.86750.
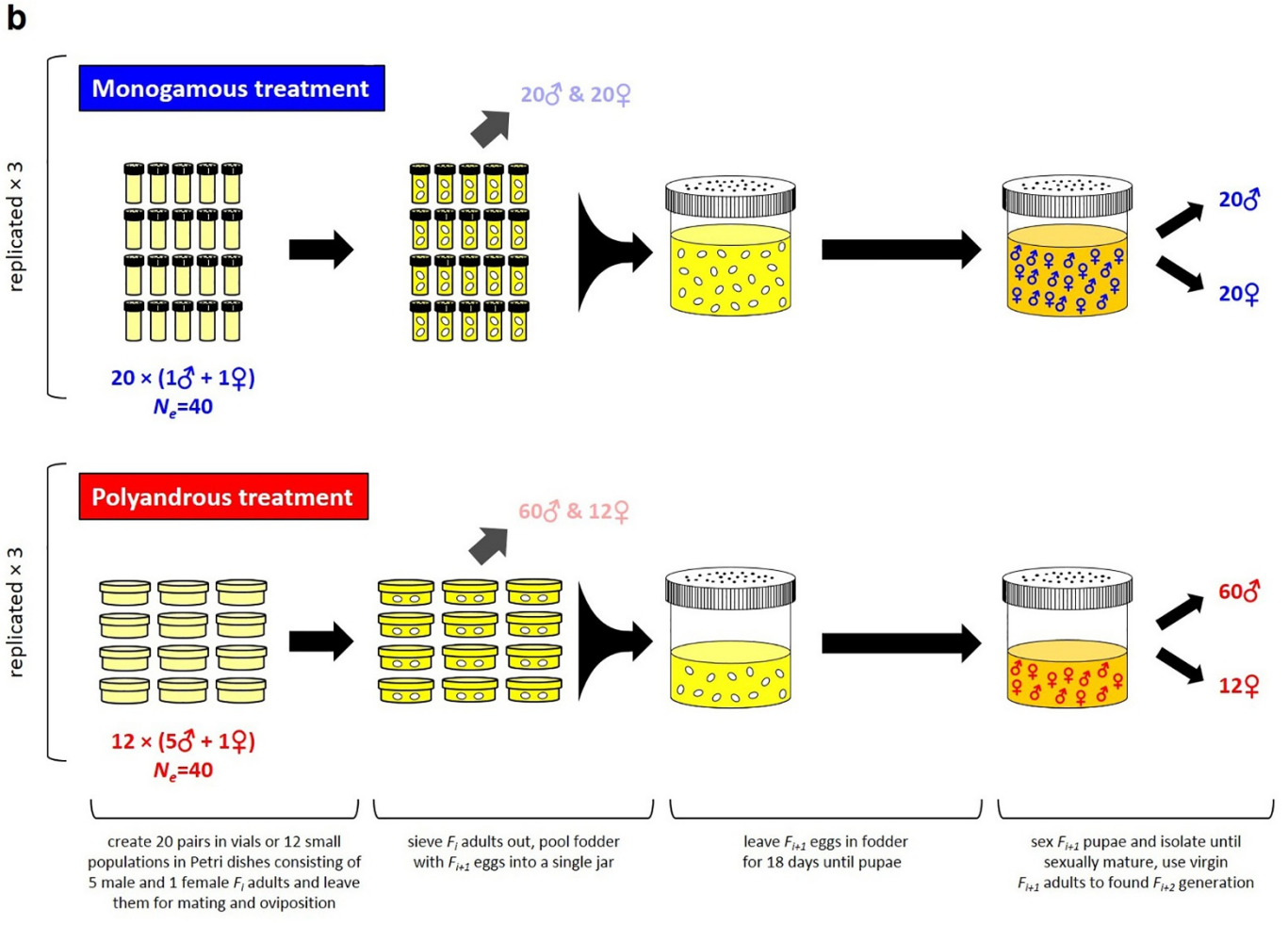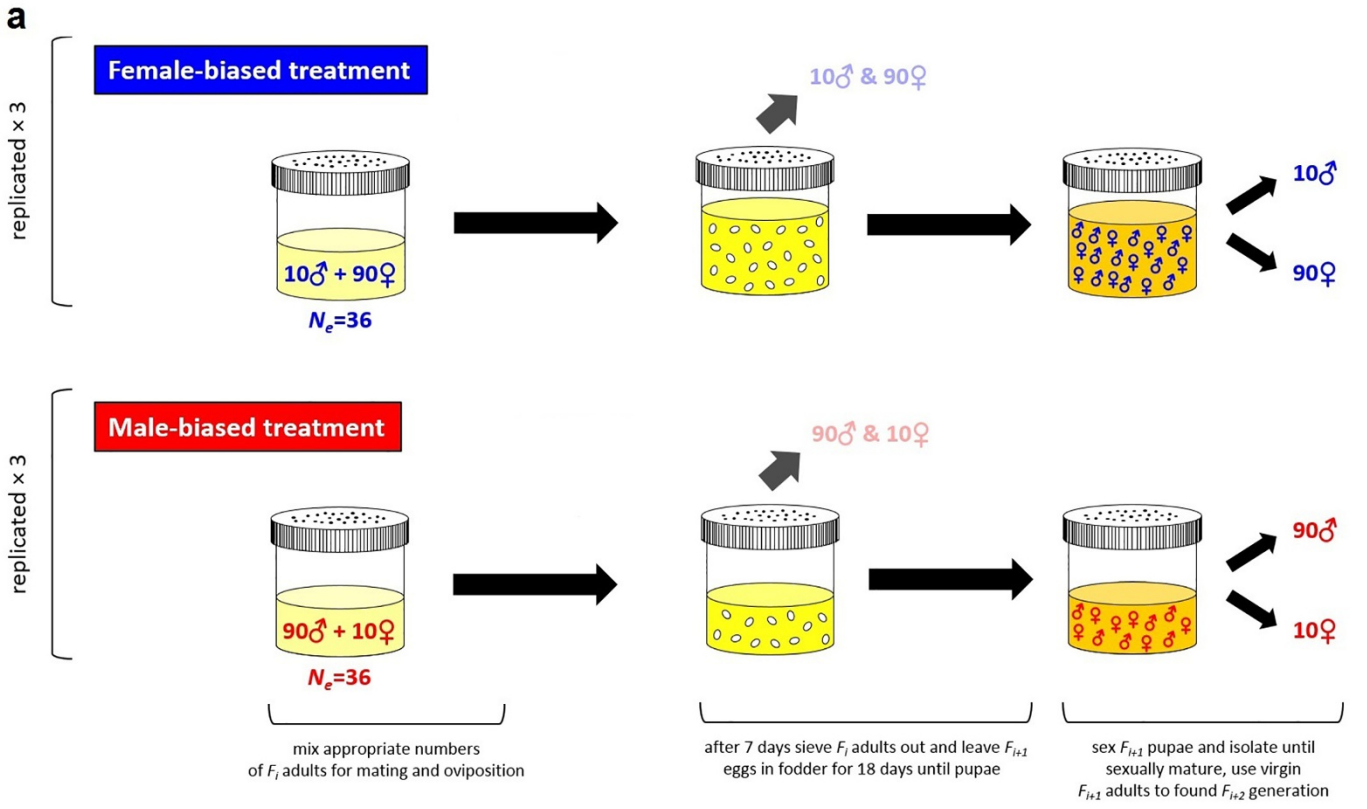
Code availability. Scripts of analyses and code used for figure production are available upon request from M.J.G.G.

31. Demont, M. *et al.* Experimental removal of sexual selection reveals adaptations to polyandry in both sexes. *Evol. Biol.* **41,** 62–70 (2014).
32. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* 4th edn (Pearson Education, 1996).
33. Therneau, T. A package for survival analysis in S. R package version 2.37-7. http://cran.r-project.org/web/packages/survival/index.html (2014).
34. R Core Team. R: a language and environment for statistical computing (R Foundation for Statistical Computing, 2013).
35. Fournier, D. A. *et al.* AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimiz. Meth. Software* **27,** 233–249 (2012).
36. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10,** 564–567 (2010).
37. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6. http://CRAN.R-project.org/package=lme4 (2014).
38. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-6. http://CRAN.R-project.org/package=lmerTest (2014).
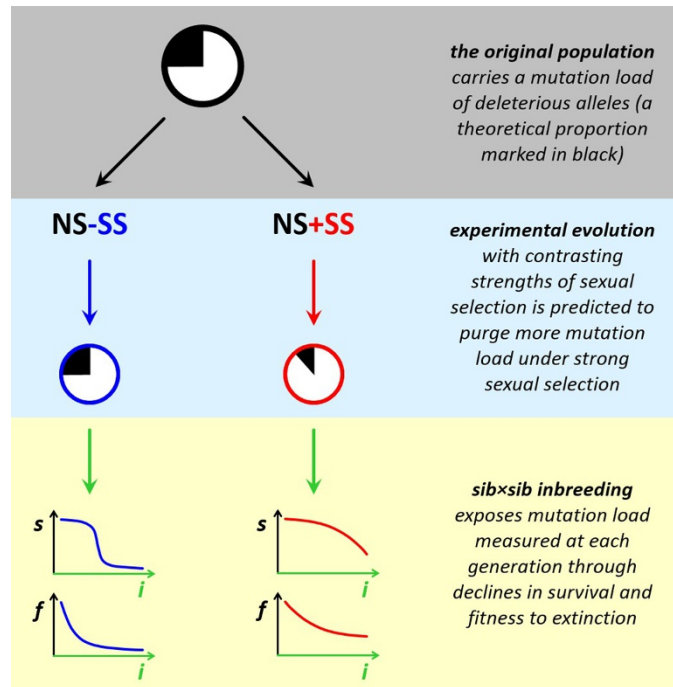
Extended Data Figure 1 | Experimental rationale for purging and then exposing mutation load. Having been changed by strong (+SS, red) versus weak (−SS, blue) histories of sexual selection, while under equal influences of natural selection (NS), variation in mutation load residing in the form of recessive alleles is exposed via inbreeding. Inbreeding was enforced through monogamous sib × sib pairings, also eliminating concurrent confounds of interlocus sexual conflict. Populations with reduced mutation load as a result of histories of strong sexual selection are predicted to resist extinction (survival, s) and maintain fitness (f) under continuous inbreeding (i).

a

Female-biased treatment

replicated × 3

$10\male + 90\female$

$N_e=36$

$10\male$ & $90\female$

$10\male$

$90\female$

Male-biased treatment

replicated × 3

$90\male + 10\female$

$N_e=36$

$90\male$ & $10\female$

$90\male$

$10\female$

mix appropriate numbers of F_i adults for mating and oviposition

after 7 days sieve F_i adults out and leave F_{i+1} eggs in fodder for 18 days until pupae

sex F_{i+1} pupae and isolate until sexually mature, use virgin F_{i+1} adults to found F_{i+2} generation

b

Monogamous treatment

replicated × 3

$20 \times (1\male + 1\female)$
$N_e=40$

$20\male$ & $20\female$

$20\male$

$20\female$

Polyandrous treatment

replicated × 3

$12 \times (5\male + 1\female)$
$N_e=40$

$60\male$ & $12\female$

$60\male$

$12\female$

create 20 pairs in vials or 12 small populations in Petri dishes consisting of 5 male and 1 female F_i adults and leave them for mating and oviposition

sieve F_i adults out, pool fodder with F_{i+1} eggs into a single jar

leave F_{i+1} eggs in fodder for 18 days until pupae

sex F_{i+1} pupae and isolate until sexually mature, use virgin F_{i+1} adults to found F_{i+2} generation

Extended Data Figure 2 | Experimental evolution protocols for regimes A and B. Contrasting intensities of strong (red) versus weak (blue) sexual selection were imposed upon each generation of adult reproduction, while equalizing effective population size within a regime, and allowing full genetic mixing within the replicate lines at the egg/larval/pupal stages. From the start, each treatment was replicated to create three independent lines. Regime A (**a**) applied contrasting sexual selection by varying adult operational sex ratio, while regime B (**b**) enforced monogamy to compare against polyandry.
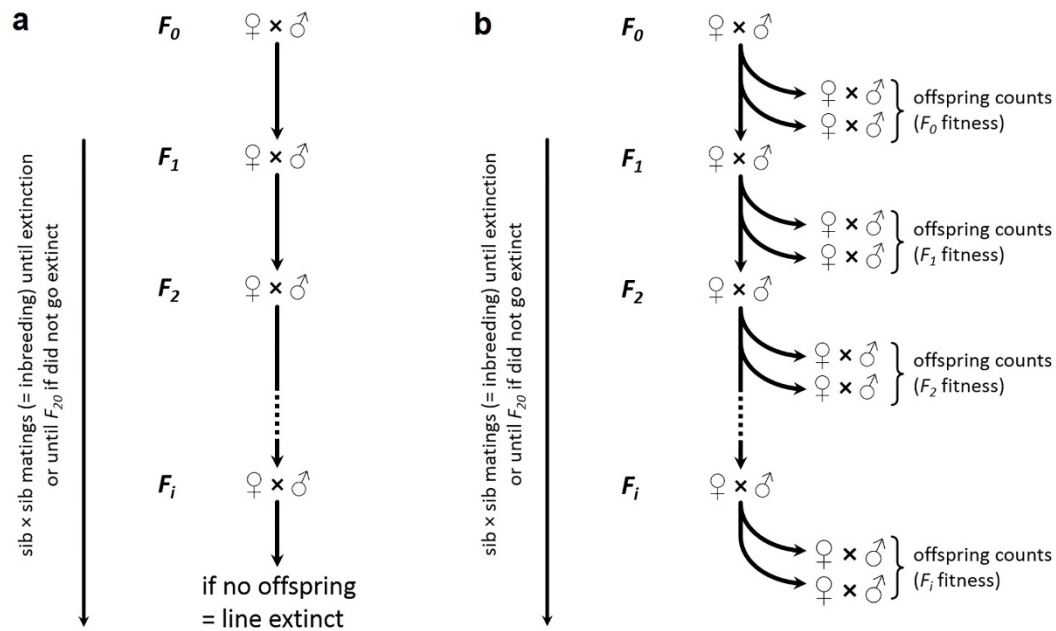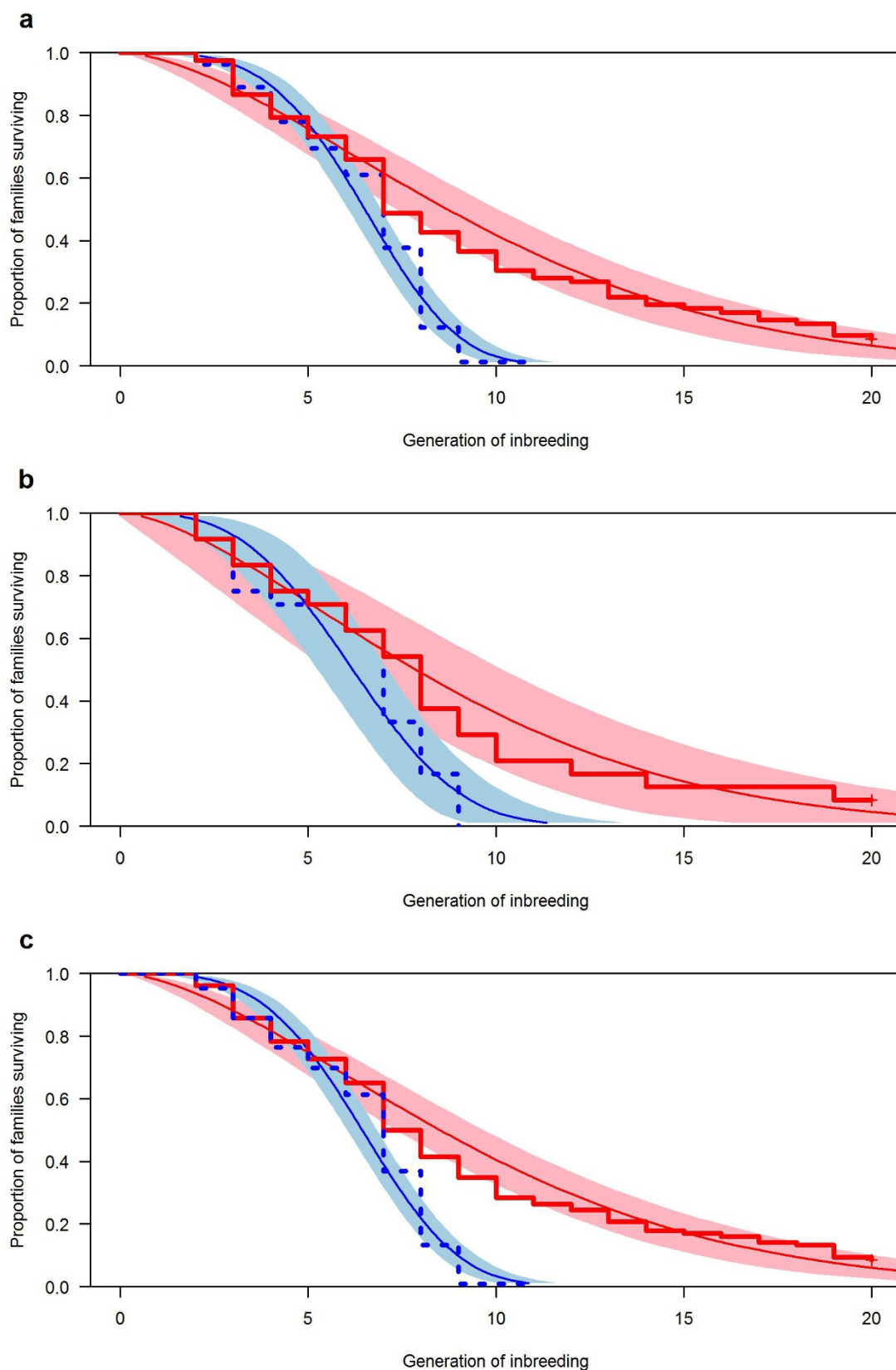
Extended Data Figure 3 | Extinction and fitness decline protocols.
Inbreeding in family lines was performed via sib × sib crosses for up to 20 generations across 3 years. To measure extinction (**a**), a family was considered extinct when it failed to produce offspring, or offspring were of the same sex (which occurred in only 9 out of 216 family lines, indicating no sex-specific pre-adult mortality by treatment). In regime A, extinction data were collected from 28 initial families per line, three lines per sexual selection treatment, comparing both strong versus weak treatments ($n = 168$ total family lines). In regime B,

extinction data were collected from eight initial families per line, three lines per sexual selection treatment, comparing both strong versus weak treatments ($n = 48$ total family lines). To measure fitness decline (**b**), two additional sib × sib pairs per family per generation were bred to estimate reproductive fitness in every generation by counting number of offspring produced (see Methods). In both regimes A and B, fitness data were collected from eight initial families per line, three lines per sexual selection treatment, and both strong versus weak treatment contrasts in each.
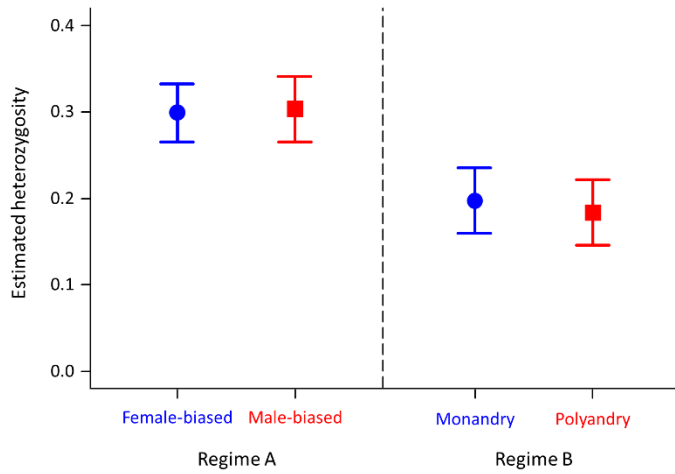
Extended Data Figure 4 | Estimated heterozygosity (±s.e.m.) does not differ between experimental evolution sexual selection treatments within regime A (left) and regime B (right). Linear mixed effect modelling showed the estimated heterozygosity of the male-biased selection treatment (M: H_{est} = 0.312, t = 9.468) is not significantly different from that of female-biased (F: H_{est} = 0.318, t = 9.295, P = 0.863), but is significantly different from monogamous and polyandrous treatments (Mo: H_{est} = 0.199, t = 6.453, P = 0.003; Po: H_{est} = 0.197, t = 6.397, P = 0.003). The estimated heterozygosities of monogamous and polyandrous treatments are not significantly different (P = 0.956) (see Methods).

Extended Data Figure 5 | Concordance between raw data and model fit in extinction analyses. Survival curves of raw data (thick and dotted lines) overlaid on model fit (shaded areas with mean curves and 95% confidence intervals). Survival of families derived from strong (red, solid line) or weak (blue, dotted line) sexual selection treatment histories differed: (**a**) regime A, male-biased (red) versus female-biased (blue) sexual selection treatments; (**b**) regime B, polyandrous (red) versus monogamous (blue); (**c**) regimes A and B combined into a single analysis. See Fig. 1 and the main text for results of statistical analyses, and Methods and Extended Data Figs 2 and 3 for details of protocols, methods and experimental design.
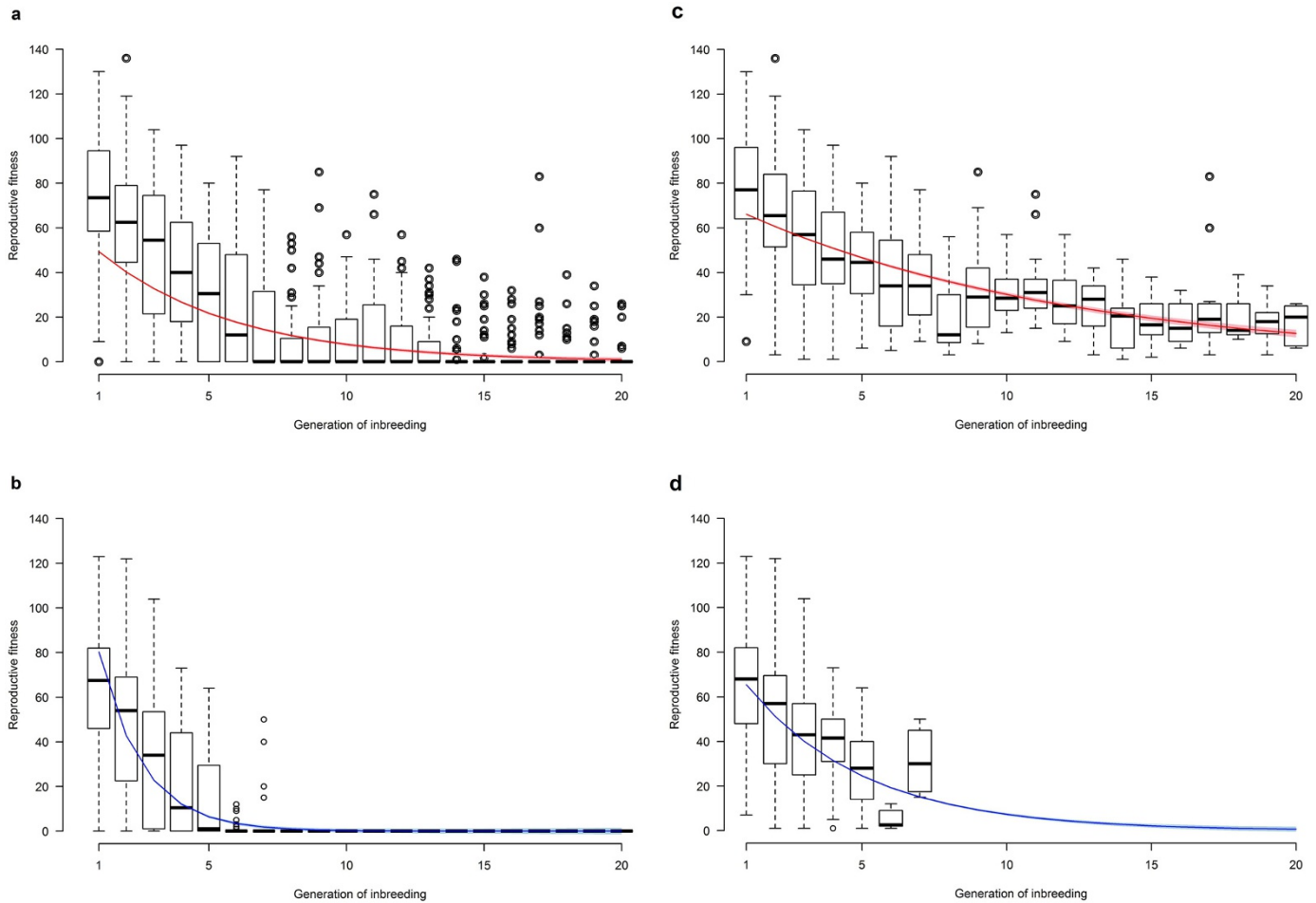
Extended Data Figure 6 | Regime A. Boxplots of the relationships between fitness and inbreeding generation for the male-biased (**a** and **c**) versus the female-biased (**b** and **d**) treatments. Curves show the predicted relationships between reproductive fitness and inbreeding generation from the GLMMs, and the narrow red and blue shadows show the 95% confidence intervals predicted from the fixed effects. Horizontal bars indicate medians, boxes indicate interquartile ranges, whiskers indicate minimum and maximum values, and circles indicate outliers (values 1.5 times higher or lower than the first and third quartiles, respectively). Comparison of **a** versus **b** identifies the difference in total fitness declines between strong versus weak sexual selection histories in regime A, while **c** versus **d** identifies the same difference in decline for fitness but only for the sibling pairs that produced at least some offspring (that is, omitting zero fitness values that may have resulted from a failure to mate). See Fig. 1, main text and Extended Data Table 1 for results of statistical analyses.
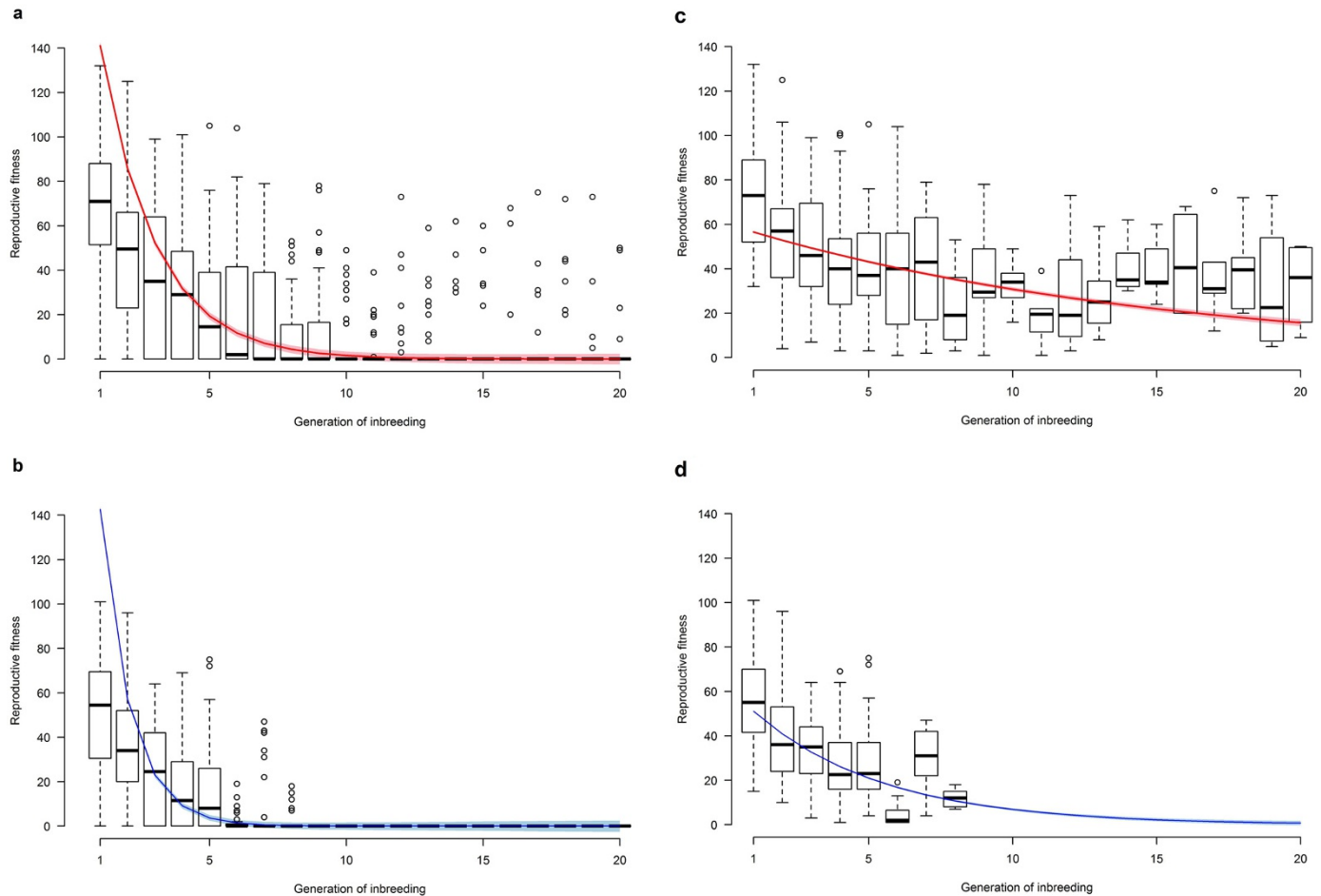
Extended Data Figure 7 | Regime B. Boxplots of the relationships between fitness and inbreeding generation for the polyandrous (**a** and **c**) versus the monogamous (**b** and **d**) treatments. Curves show the predicted relationships between reproductive fitness and inbreeding generation from the GLMMs, and the narrow red and blue shadows show the 95% confidence intervals predicted from the fixed effects. Horizontal bars indicate medians, boxes indicate interquartile ranges, whiskers indicate minimum and maximum values, and circles indicate outliers (values 1.5 times higher or lower than first and third quartiles, respectively). Comparison of **a** versus **b** identifies the difference in total fitness declines between strong versus weak sexual selection histories in regime B, while **c** versus **d** identifies the same difference in decline for fitness but only for the sibling pairs that produced at least some offspring (that is, omitting zero fitness values that may have resulted from a failure to mate). See Fig. 1, main text and Extended Data Table 1 for results of statistical analyses.
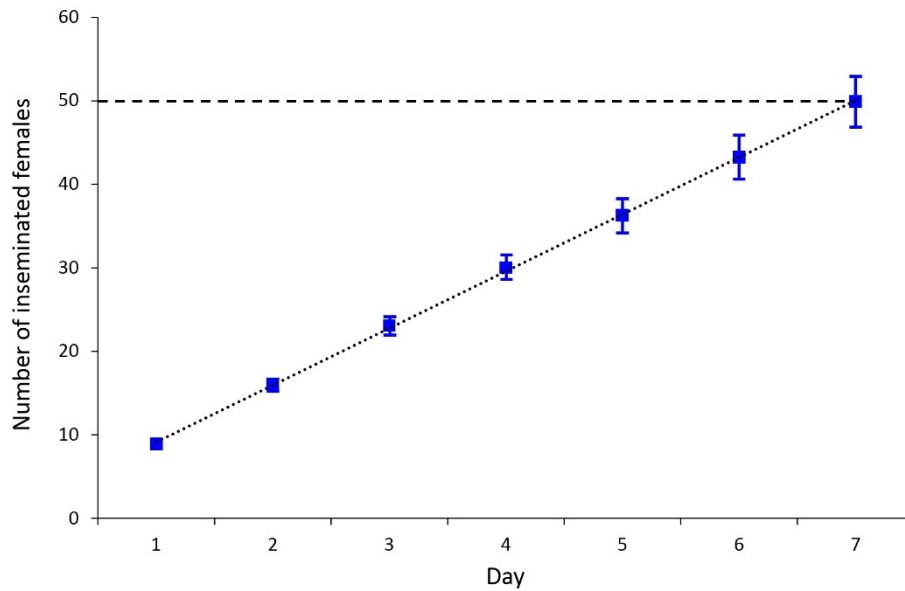
Extended Data Figure 8 | Across 7 days of mating opportunity, males successfully inseminated 50 females on average (±s.e.m.). Six virgin females were allocated to individual GA1 control stock males ($n = 11$) every 12 h for 7 days, providing males with 84 potential mates. Over this 1 week period (replicating that applied within the experimental evolution protocols, Extended Data Fig. 2), males successfully inseminated and generated offspring from an average of 50 females (see Methods).

Extended Data Table 1 | Fixed-effect parameter estimates from negative binomial GLMMs of the relationship between fitness and generation of inbreeding for male-biased and female-biased treatments (regime A), and polyandrous and monogamous treatments (regime B), and their statistical interactions. See Methods for details of replication and sample sizes.

Comparison		Estimate	Std. Error	z value	Pr(> |z|)
Regime A Male-biased *vs.* Female-biased	Intercept	5.02	0.17	28.23	<0.001
	Inbreeding	-0.63	0.03	-21.34	<0.001
	Treatment (Female- or Male-biased)	-0.91	0.24	-3.75	<0.001
	Inbreeding*Treatment (Male-biased)	0.42	0.03	13.82	<0.001
Regime B Polyandrous *vs.* Monogamous	Intercept	4.48	0.23	19.39	<0.001
	Inbreeding	-0.53	0.03	-21.17	<0.001
	Treatment (Monogamy or Polyandry)	-0.63	0.32	-1.97	0.048
	Inbreeding*Treatment (Polyandry)	0.29	0.03	10.56	<0.001