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]

[1]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

[2]Department of Entomology, Institute of Zoology, Jagiellonian University, Gronostajowa 9, 30-387 Kraków, Poland

[3]ETH Zurich, Institute of Integrative Biology, D-USYS, Universitatsstrasse 16, CHN J 11, 8092 Zürich, Switzerland

[4]Island Ecology and Evolution Research Group (CSIC), C/Astrofísico Francisco Sánchez 3, La Laguna, Tenerife, Canary Islands, 38206, Spain

*These authors contributed equally to this work.

Correspondence: m.gage@uea.ac.uk

SUMMARY

Reproduction through sex carries substantial costs, mainly because only half of sexual adults produce offspring[1]. It has been theorised 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 depends on 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 ~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 non-existent 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 which 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], due 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 formalised

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 fertilisation 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 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. Following 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 minimised[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♀:90♂) *versus* female-biased (90♀:10♂), while Regime B allowed polyandry (1♀:5♂) *versus* enforced monogamy (1♀:1♂) for 45 generations. All other conditions among lines and treatments within a Regime were kept identical, including equalising theoretical effective population sizes (see Methods and Extended Data Figs 2 and 4, and Supplementary Information S1, Table 1 and 2, 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, 6 and 7; all means presented in this article are arithmetic means ± standard errors). Over the three-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 SE) vs 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 vs 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, $P < 0.001$), but 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 (± 0.37) vs 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; 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, 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 and d, and 7c and d).

Results from this 10-year experiment provide compelling empirical support for the complementary models of Agrawal[2] and Siller[3], who 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

177 especially prone to this process[24], but since the *T. castaneum* X chromosome makes up only
178 ~6% of the genome[25], we expect purging of load to have occurred on the autosomes too.
179 There is emerging evidence that sexual selection can profoundly shape the genome, with
180 feminised patterns of sex-linked expression in the transcriptomes of both sexes after
181 experimental evolution under monogamy in *Drosophila*[26]. Our findings indicate that sexual
182 selection also acts to purge mutation load, even in populations that we expect to be adapting
183 close to their natural fitness peaks[27]. Stronger sexual selection will drive greater variance in
184 reproductive success, so that the average father should carry fewer deleterious mutations than
185 the average male[3], and perhaps also the average mother if interlocus sexual conflict
186 constrains female fitness[15].
187
188 Within the promiscuous *Tribolium* model[28] (Extended Data Fig. 8, Methods), both pre- and
189 post-copulatory processes for winning or controlling fertilisations will operate through broad
190 behaviours and physiology within the whole organism, down to competition and choice at the
191 gamete level. Genic capture predicts that individual success across this continuum will
192 depend on overall condition and genes at many loci[5]. Under monogamy, where sexual
193 selection has been removed, populations suffer the constraints of sexual reproduction, but
194 none of the benefits of sexual selection, explaining the lowered base-line fitness even without
195 inbreeding, as well as more rapid fitness declines and heightened extinction rates under
196 inbreeding (Figs 1 and 2, Extended Data Figs 5, 6 and 7). Although we exposed mutation
197 load for experimental measurement through inbreeding, the intense and diverse demands
198 from selection across multiple generations in the natural environment, where populations are
199 more likely to be displaced from their fitness peaks[27], are likely to expose even greater fitness
200 differentials due to load variance. After only one generation of inbreeding, for example, mean
201 fitness between strong *versus* weak sexual selection histories differs by 20% and 40% in
202 Regimes A and B respectively (Fig. 2). Our results indicate, as demands rise upon
203 populations that have been depleted and fragmented, and displaced further from their fitness
204 peaks[27] via increasing anthropogenic stress[29], that sexual selection could be an important
205 force protecting species or populations from the extinction vortex[30].

References

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 in Ecology and Evolution* **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. *Annual Reviews in Ecology and Systematics* **43,** 115-135 (2012).
8. Holman, L. & Kokko, H. The consequences of polyandry for population viability, extinction risk and conservation. *Philosophical Transactions of the Royal Society series B* **368,** doi: 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. *The American Naturalist* **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, London, 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. *Ecology Letters* **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 *Drosphila 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. *Journal of Evolutionary Biology* **24,** 1918-1926 (2011).
22. Holland, B. & Rice, W.R. Experimental removal of sexual selection reverses intersexual antagonistic coevolution and removes a reproductive load. *Proceedings of the National Academy of Sciences of the United States of America* **96,** 5083-5088 (1999).
23. Charlesworth, B. & Charlesworth, D. The genetic basis of inbreeding depression. *Genetical Research* **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 Research* **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 Communications* **5,** doi: 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. *Current Biology* **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. Minimum viable populations: processes of extinction. In: *Conservation Biology: The Science of Scarcity and Diversity*, ed. M.E. Soulé, 19-34. (Sinauer Associates, Sunderland, Massachusetts, 1986).

Figure legends

Figure 1: Extinction trajectories under increasing inbreeding differed between family lines derived from strong (red squares) *versus* weak (blue circles) sexual selection histories. Each generation mean presents the average (± SE) 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 generalised estimating equation (GEE) 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).

Figure 2: Reproductive fitness declines under increasing inbreeding of families derived from strong (red squares) *versus* weak (blue circles) sexual selection histories differed in magnitude and rate. a) Regime A: male-biased (red) *versus* female-biased (blue) selection histories, and **b**) Regime B: polyandrous (red) *versus* monogamous (blue) histories. Each generation mean presents the average (± SE) number of offspring produced under standardised 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 generalised 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; and 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; and Regime B: z = -1.97, P = 0.048; see Extended Data Table 1 and Figs 6 and 7).

Extended Data Figure Legends

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*).

Extended Data Figure 2: Experimental evolution protocols for Regime A and Regime B. Contrasting intensities of strong (red) *versus* weak (blue) sexual selection were imposed upon each generation of adult reproduction, while equalising 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 (a) and fitness decline (b) 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). In both Regime 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. 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 Regime 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 (± SE) 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 to that of female-biased (F: H_{est} = 0.318, t = 9.295, P = 0.863), but is significantly different to 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% CIs). 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% CIs 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 1st and 3rd quartile, respectively). **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 (i.e. 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% CIs 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 1st and 3rd quartile, respectively). **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 (i.e. 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 seven days of mating opportunity, males successfully inseminated 50 females on average. Six virgin females were allocated to individual GA1 control stock males (n = 11) every 12 hours for seven days, providing males with 84 potential mates. Over this one week period (replicating that applied within the experimental evolution protocols, Extended Data Fig. 2), males successfully inseminated and generated offspring from on 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.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank NERC and the Leverhulme Trust for financial support, Dominic Edward for statistical advice, and colleagues at the 2013 Biology of Sperm meeting for comments that improved analytical design and interpretation.

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 reported in this paper are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.TBC. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. 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 Lab (United States Department of Agriculture, Biological Research Unit, Grain Marketing & Production Research Centre, 1515 College Avenue, Manhattan, KS 66502), 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[30,31]. 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[31,32]. After 7 days, adults were removed and the eggs and larvae (typically ca. 70 offspring per female) were left to develop under standardised conditions with equal offspring densities by maintaining *ad libitum* levels of food in proportion to the number of offspring. These conditions were maintained for *ca*. 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 equalising 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 simultaneous competition between the 90 males to fertilise the 10 females, whereas in female-biased treatments the reverse scenarios applied. Six independent lines (3 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 5 males under polyandry (12 replicate groups per line). Thus, polyandry provided each female with a choice among 5 males, and competition between 5 males for just 1 female, whereas monogamy completely removed all female choice and male-male competition. Six independent lines (3 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 standardised 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; 24h before their trial they were identified with a small dot of white correction fluid on their thorax[32]. 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 hours for seven days (12 × 7 = 84 females given to each male over the week). Following each 12 hour mating period, females were placed singly into petri dishes with standard fodder for seven 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 fertilised females across this one week period per male was 50 (± 3 SE); 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 ten females at each adult generation in the male-biased treatment were potentially exposed to 90 × 50 = 4500 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 polyandrous treatment generated a 5-fold increase in potential for pre- and post-copulatory male-male competition and female choice through the availability of five males, all of whom could mate with the female.

Effective population sizes. In order to avoid differential inbreeding that could have subsequently influenced our extinction assay, we equalised the theoretical N_e for mixed adult sex ratios ($N_e = \frac{4N_fN_m}{N_{f+}N_m}$)[33] 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

478 treatments with male-biased or polyandrous structures may translate into reduced realised N_e
479 due to male success skew under strong sexual selection (where competition and choice could
480 allow reproductive success by a smaller subset of males in the breeding population). Thus,
481 our experimental design is conservative to any influence of N_e, because we test the prediction
482 that lineages with histories of strong sexual selection (and therefore potentially lower N_e)
483 should demonstrate resistance to extinction and reduced rates of fitness decline under
484 inbreeding (because they have more effectively purged mutation load). If N_e is lower in these
485 strong sexual selection populations, we would expect this to result in reduced heterozygosity
486 at the start of inbreeding, resulting in lower initial fitness and therefore faster extinction rates
487 (yet we observe the converse in Figs 1 and 2). As an additional, third, check for N_e confounds,
488 we also directly measured heterozygosity and allelic richness in our lines to establish that
489 differential inbreeding had not occurred under different sexual selection treatments (methods
490 outlined below). By screening 628 individuals representing the different sexual selection
491 treatments in both Regime A and B at 13 microsatellite loci (Supplementary Information S1),
492 we were able to confirm that heterozygosity and allelic richness showed no differences
493 between 'weak' or 'strong' sexual selection histories (Extended Data Fig. 4).
494
495 **Exposing mutation load.** Having been subjected to experimental evolution that applied
496 strong *versus* weak sexual selection under equal effective population sizes, we then exposed
497 the mutation load carried in these populations by inbreeding down multiple replicate family
498 lines to expose deleterious recessives, and tracking fitness decline to extinction as a result of
499 inbreeding depression through partial dominance[26]. At the end of experimental evolution, we
500 created multiple full-sib families using monogamous crosses of randomly chosen unmated
501 males and females within each independent selection line, which simultaneously removed
502 any transgenerational interlocus sexual conflict effects, and from which full-sib offspring
503 were then used to continue a total of 216 family lines down increasing inbreeding coefficients
504 and homozygosity (Extended Data Fig. 3). There were thus two generations of monogamy
505 applied before fitness assays began, eliminating any carry-over effects of sexual selection or
506 interlocus conflict. As levels of inbreeding increased, we tracked extinction rates and fitness
507 declines for multiple families within each of the 12 independent lines from the original sexual
508 selection treatments. In Regime A, we created 28 full-sib families in each independent line
509 ($N_{total} = 28_{families} \times 3_{independent\ lines} \times 2_{sexual\ selection\ treatments} = 168$ families), allowing extinction
510 rates to be measured within each of the 3 independent lines in either of the two sexual
511 selection treatments. When sibling pairs failed to produce female and male offspring that

512 survived to adulthood to enable continuation of the line, it was recorded as extinct (see
513 Extended Data Fig. 3). In addition to extinction, we also measured reproductive fitness in a
514 subset of these families, assayed as the average number of offspring produced by two
515 randomly chosen sib × sib pairs within 8 of the families per line (Extended Data Fig. 3).
516 Identical protocols for Regime B were followed, except that both extinction rates and average
517 fitness were measured across 8 families per independent line ($N_{total} = 8_{families} \times 3_{independent\ lines} \times$
518 $2_{sexual\ selection\ treatments} = 48$ families). Blinding was not performed as the protocol did not permit
519 biasing to affect results. Inbreeding was continued for 20 generations of sib × sib matings
520 (which operated over three years), by which time Wright's inbreeding coefficient *F* had
521 increased to 0.986. After 20 generations of inbreeding 208 of the 216 initial lines had gone
522 extinct, with the 8 survivors all derived from strong sexual selection histories.
523
524 **Baseline fitness.** At four time-points throughout the extinction assay (Parental, F_7, F_{15} and
525 F_{21} generations), we assayed 'baseline' fitness (without inbreeding) of standard monogamous
526 crosses from the different sexual selection treatments, using identical protocols to fitness
527 measures applied for the inbred crosses. Pupae from each of the three independent lines
528 within each of the four different selection treatments were separated into single-sex groups
529 and isolated for 10 days to allow adult eclosion and sexual maturation. $n = 20$ randomly
530 chosen male-female pairs were then established for each of the three independent lines within
531 each of the four sexual selection treatments. Each pair was placed into a 7 ml plastic vial,
532 with *ad libitum* food and oats, and allowed to mate and oviposit for 7 days. After this, the
533 flour from each vial, containing the eggs laid during the 7 day mating period, was transferred
534 to a petri dish containing a further 10 g of fodder, and the eggs allowed to develop to adult
535 eclosion (~ 35 days). Fitness was scored as the total number of adults produced from each
536 cross across 7 days of mating and oviposition. Data from previous work[31] where male-female
537 pairs were allowed to interact for 7 days shows (a) that the first week of oviposition produces >
538 25% of the total female reproductive fitness for such male:female interactions, and (b) that
539 female fitness over the first week significantly correlates with total female reproductive
540 fitness ($r = 0.48$, $P = 0.008$, $n = 29$ pairs). Average baseline fitness values are presented for
541 reference in Figs 2a and 2b ('ref').
542
543 **Extinction analyses.** Extinction rates were analysed using the 'survival' package[34] in R
544 version 3.1.0[35]. To assess whether sexual selection history influenced extinction rate after 20
545 generations of sib-sib inbreeding for either Regimes A or B, we analysed generation to

546 extinction using a parametric accelerated failure time survival model with Weibull baseline
547 hazard distribution, taking into account right-censored data from families that were still alive
548 at the end of the experiment. Generation of extinction was modelled with sexual selection
549 treatment as a fixed effect, and the shape and scale of the underlying Weibull hazard
550 distribution was allowed to vary by treatment. We also ran a combined analysis where both
551 experimental evolution regimes were incorporated into the model, with level of sexual
552 selection (strong *versus* weak), comparison (Regime A or B), and their interaction as fixed
553 effects. In all survival analyses, correlated data within lines were incorporated using a
554 generalised estimating equation (GEE) approach. Akaike Information Criteria and graphical
555 interpretation of the complimentary log-log survival plots were used to confirm
556 appropriateness of the specified Weibull distribution; no violations of model assumptions
557 were detected. Model fits are presented in Extended Data Fig. 5a, b and c.
558
559 **Reproductive fitness analyses.** We tested for a relationship between fitness and generation
560 of inbreeding by fitting generalised linear mixed models (GLMMs) using the glmmADMB
561 package[36] in R 3.0.3[35]. We assumed a negative binomial error structure which adds the
562 parameter (k) to the variance mean relationship, allowing us to account for the overdispersion
563 in our data introduced by a large number of zero fitness observations. GlmmADMB has two
564 options for fitting the relationship between the mean and the variance. These are
565 family='nbinom1' which assumes the variance = k*mean and family='nbinom2' which
566 assumes the variance = mean(1 + mean/k). We fitted models with each of these options in
567 turn and then selected the parameterisation which provided the lowest AIC score and
568 therefore accounts for the greatest amount of variance in our dataset. In order to account for
569 the nested nature of the experimental design and the non-independence between replicate
570 families, random effects were included in the model as replicate nested within family nested
571 within line. Separate GLMMs were fitted to the Regime A (male-biased *versus* female-biased)
572 comparison and the Regime B (polyandry *versus* monogamy) comparison (Extended Data
573 Figs 6 and 7). To ensure that failure-to-mate was not a reason for differences in fitness
574 declines between treatments, we repeated the analyses including only those families that
575 produced some offspring each generation (Extended Data Figs 6c and d, and 7c and d).
576
577 **Microsatellite analyses.** A total of 628 individuals from the 12 independent lines
578 representing all four sexual selection treatments (strong *versus* weak in Regimes A and B)
579 were genotyped at 13 loci conforming to Hardy Weinberg equilibrium (Supplementary

Information S1 and Tables 1 and 2). Scored genotypes for all individuals were analysed in Arlequin version 3.5[37] to obtain allelic richness and observed and expected heterozygosity. We tested for differences in heterozygosity across selection treatments using linear mixed models, implemented in the lme4 version 1.1-6[38] package in R[35], with significance testing performed using the package lmerTest version 2.0-6[39]. 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 ID had a significant effect on genetic diversity beyond that of selection treatment, a second mixed model was tested with the same variables plus line ID 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: http://dx.doi.org/10.5061/dryad.[NNNN].

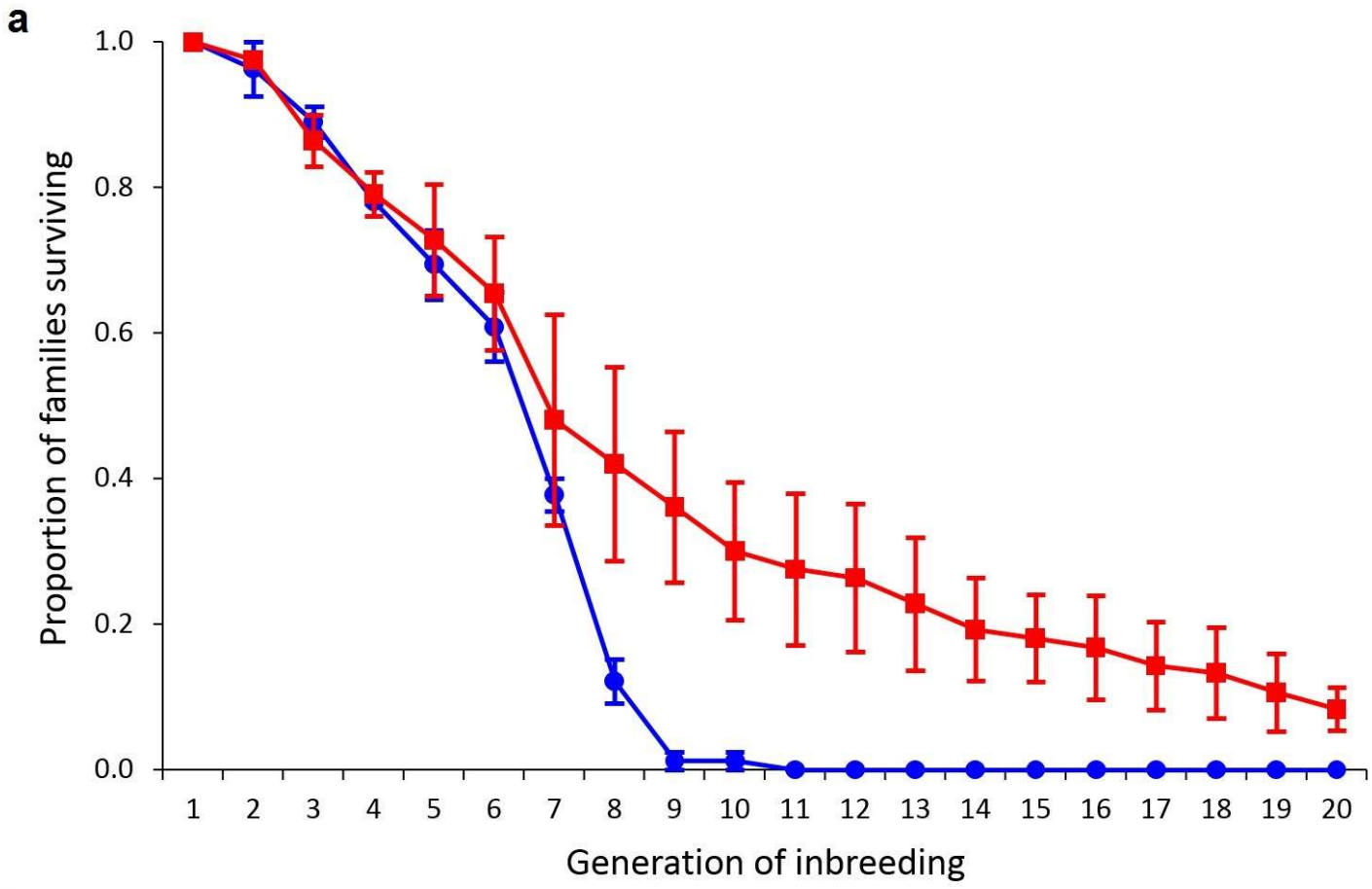
Code availability. Scripts of analyses and code used for figure production are available upon request to the corresponding author, M.J.G.G. (m.gage@uea.ac.uk).

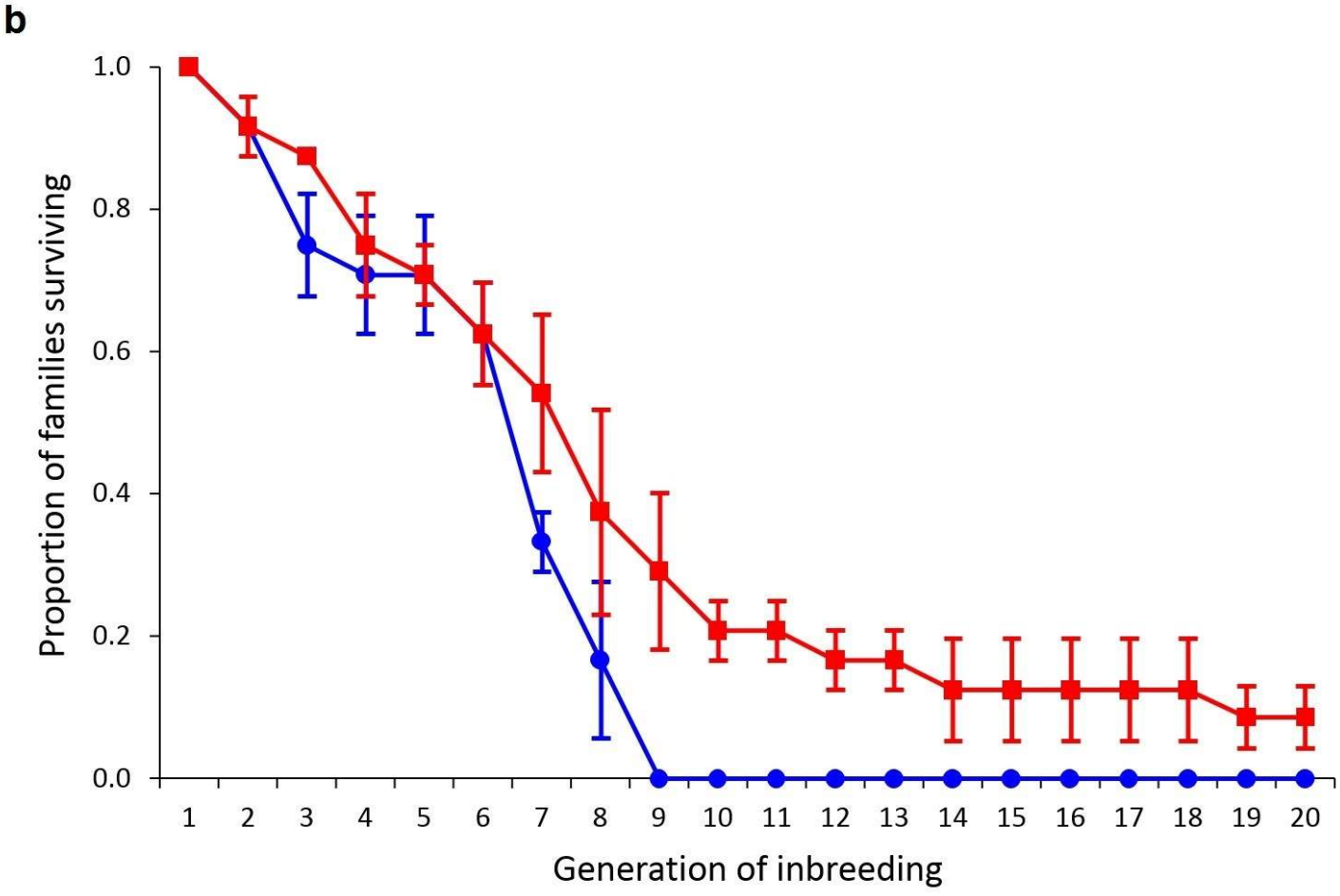
Methods References

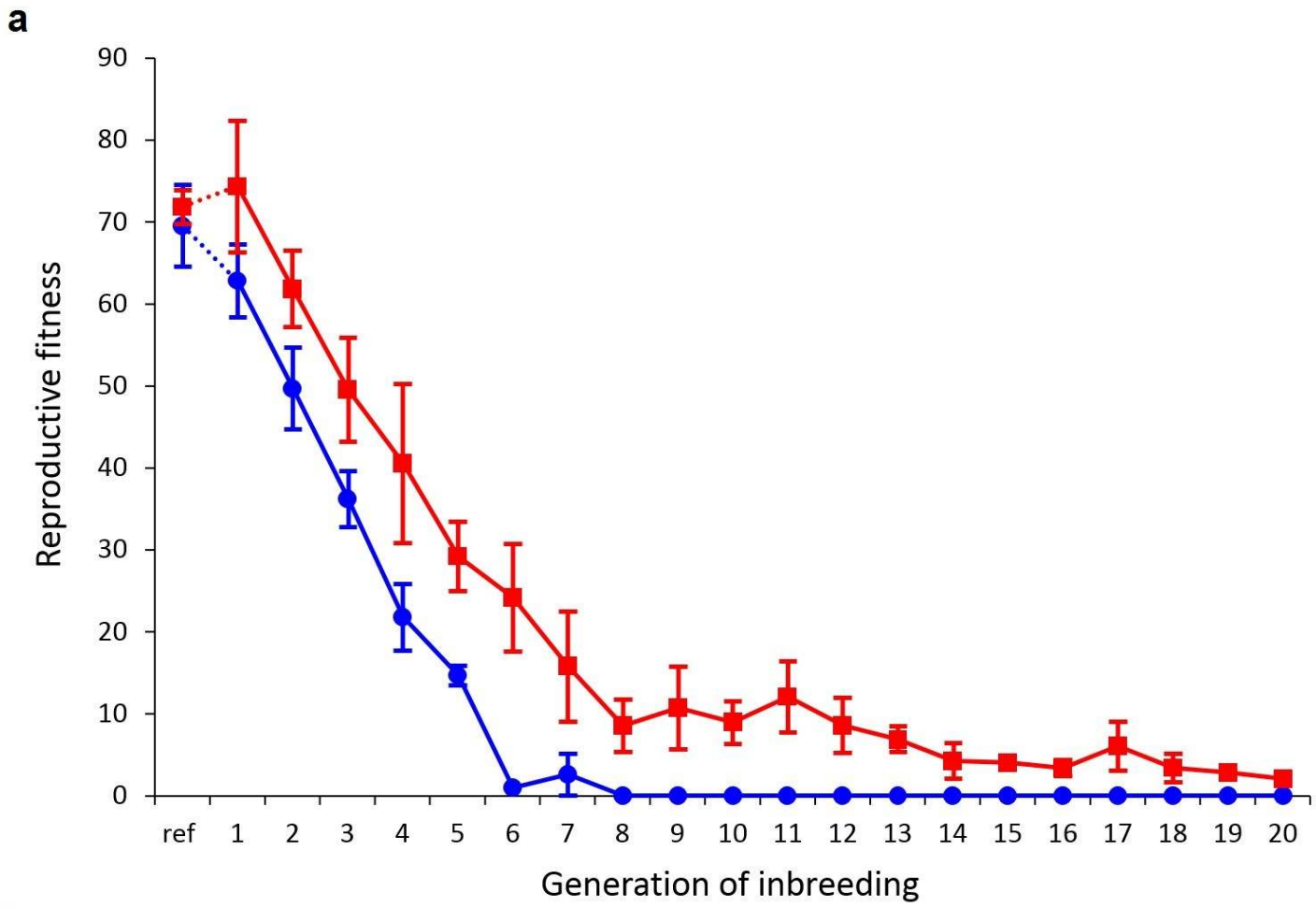
31. Demont, M. *et al*. Experimental removal of sexual selection reveals adaptations to polyandry in both sexes. *Evolutionary Biology* **41,** 62-70 (2014).
32. Falconer, D.S. & Mackay, T.F.C. *Introduction to Quantitative Genetics*, 4th edn. (Pearson Education Limited, Harlow, UK, 1996).
33. Therneau T. *A Package for Survival Analysis in S*. R package version 2.37-7. http://CRAN.R-project.org/package=survival (2014).
34. R Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria (2013).
35. Fournier, D.A., *et al*. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and 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. *Molecular Ecology Resources* **10,** 564-567 (2010).
37. Bates, D., Maechler, M., Bolker, B. & Walker, S. *lme4: Linear mixed-effects models*

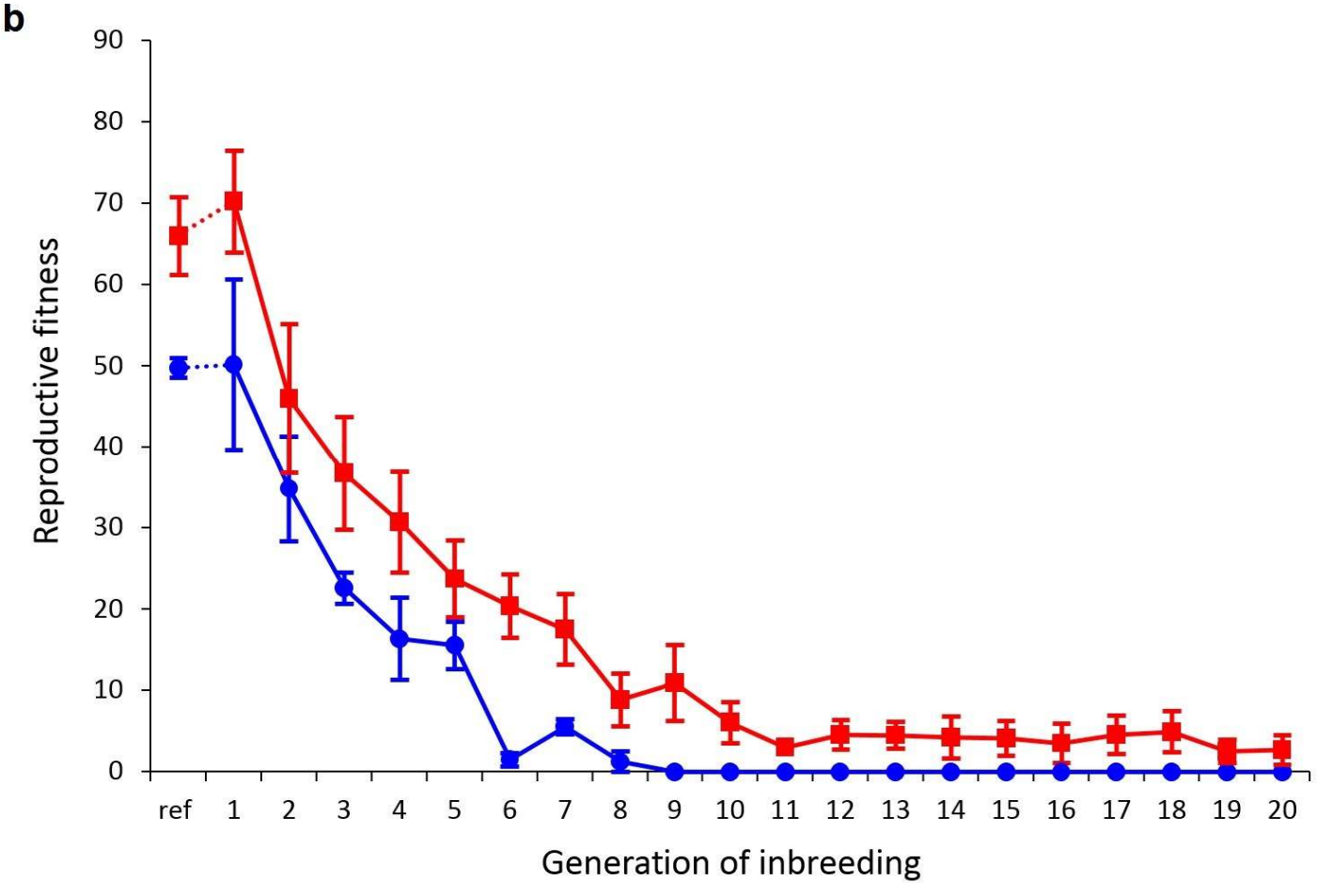
using Eigen and S4. R package version 1.1-6. http://CRAN.R-project.org/package=lme4 (2014).

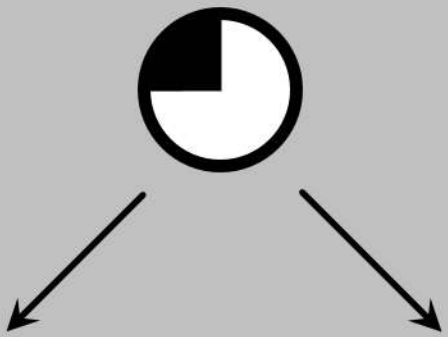
38. Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. *lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)*. R packageversion 2.0-6. http://CRAN.R-project.org/package=lmerTest (2014).


a
1.0
0.8
0.6
0.4
0.2
0.0
Proportion of families surviving
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Generation of inbreeding


b
1.0
0.8
0.6
0.4
0.2
0.0
Proportion of families surviving
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Generation of inbreeding


a
Reproductive fitness
90
80
70
60
50
40
30
20
10
0
ref 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Generation of inbreeding

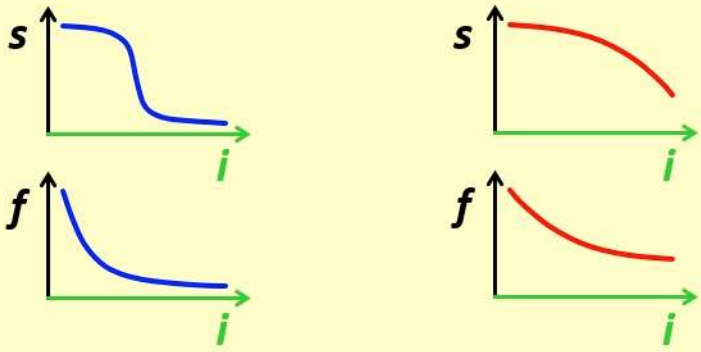

b
Reproductive fitness
90
80
70
60
50
40
30
20
10
0
ref 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Generation of inbreeding



the original population *carries a mutation load of deleterious alleles (a theoretical proportion marked in black)*

NS-SS **NS+SS**

experimental evolution *with contrasting strengths of sexual selection is predicted to purge more mutation load under strong sexual selection*



sib×sib inbreeding *exposes mutation load measured at each generation through declines in survival and fitness to extinction*

a

replicated × 3

Female-biased treatment

10♂ + 90♀

N_e=36

10♂ & 90♀

10♂

90♀

replicated × 3

Male-biased treatment

90♂ + 10♀

N_e=36

90♂ & 10♀

90♂

10♀

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

replicated × 3

Monogamous treatment

20 × (1♂ + 1♀)

N_e=40

20♂ & 20♀

20♂

20♀

replicated × 3

Polyandrous treatment

12 × (5♂ + 1♀)

N_e=40

60♂ & 12♀

60♂

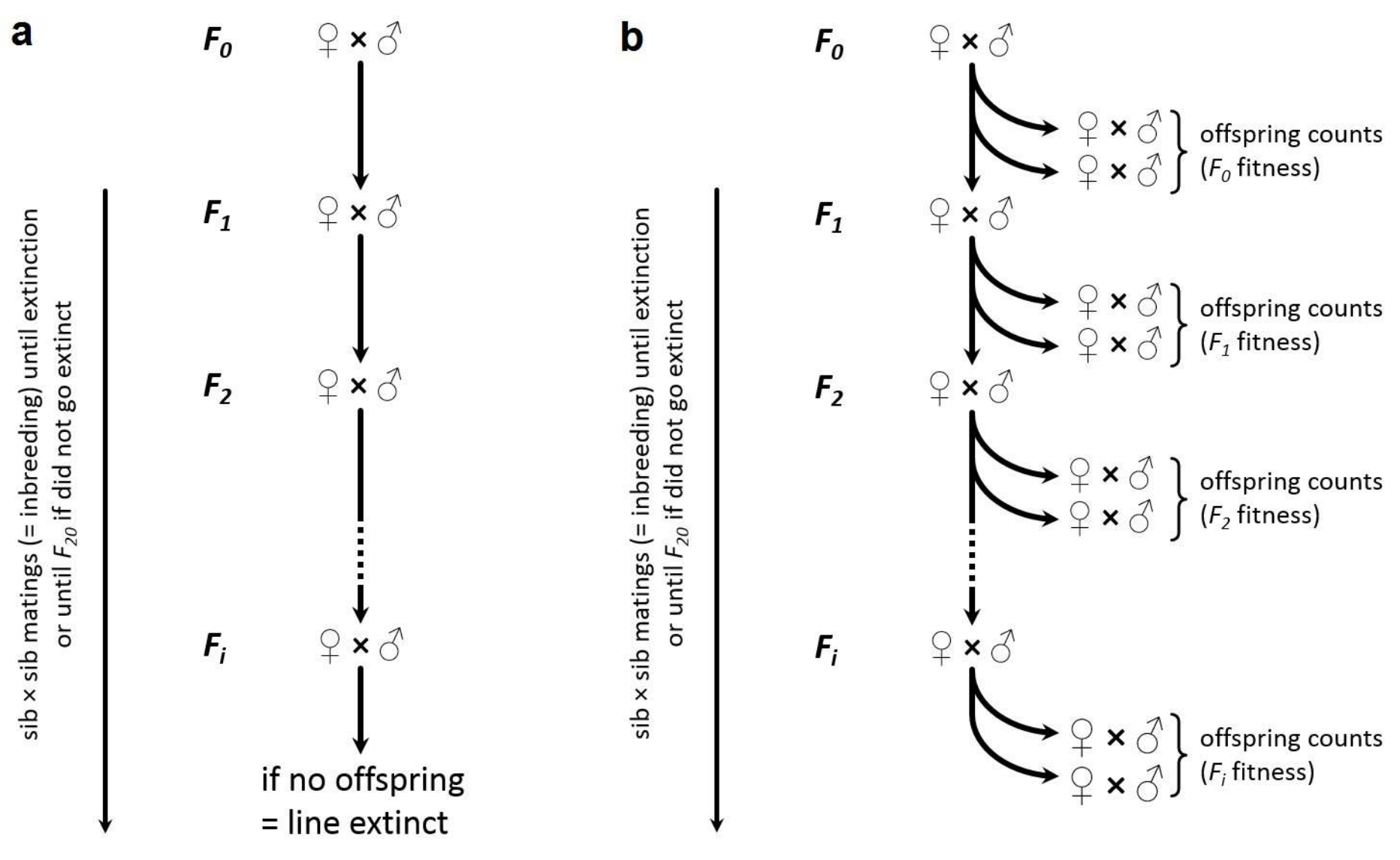
12♀

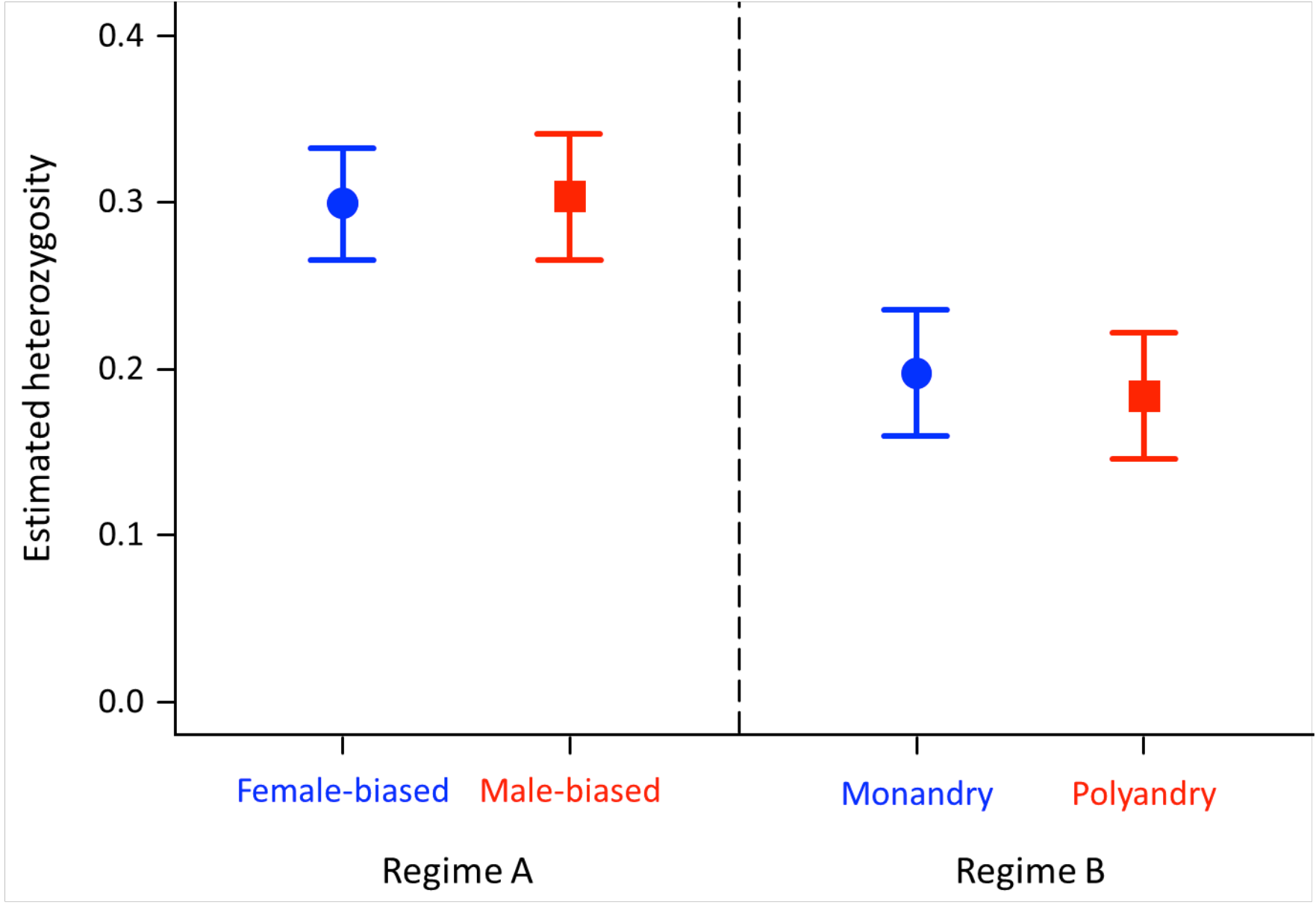
create 20 pairs in vials or 12 small populations in Petri dishes consisting of 5 male and 1 female F_i adults and leave them for mating and oviposition

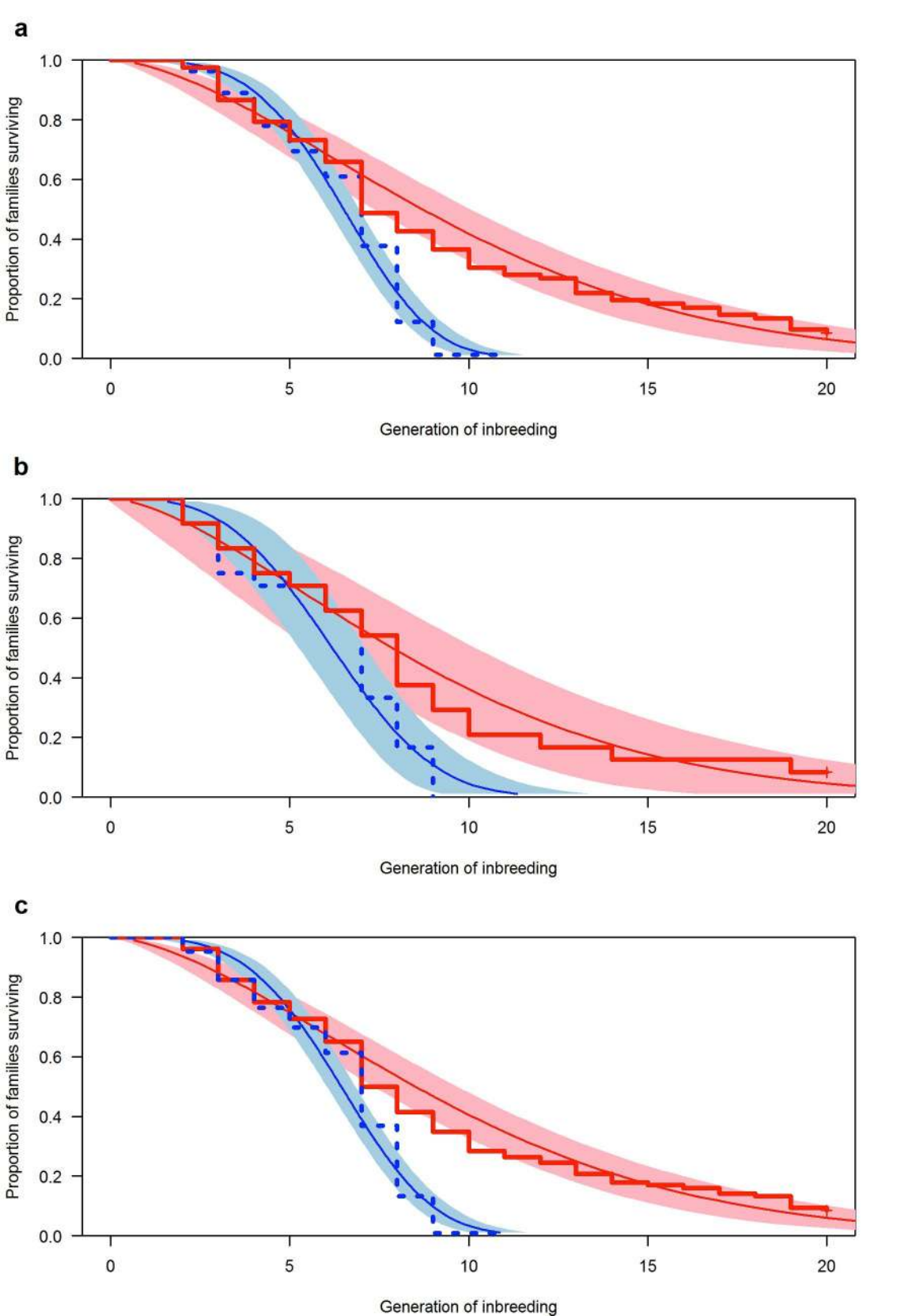
sieve F_i adults out, pool fodder with F_{i+1} eggs into a single jar

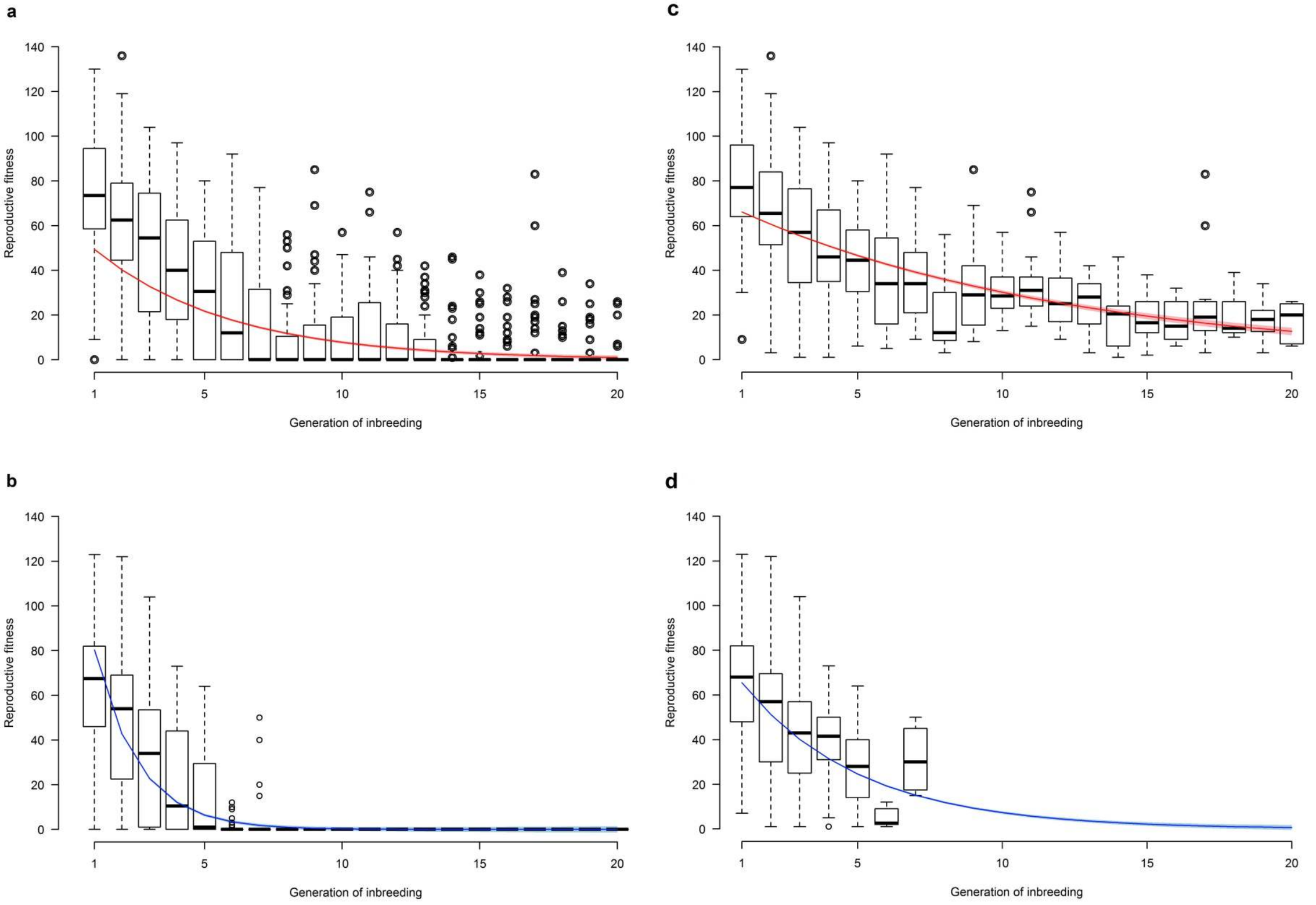
leave F_{i+1} eggs in fodder for 18 days until pupae

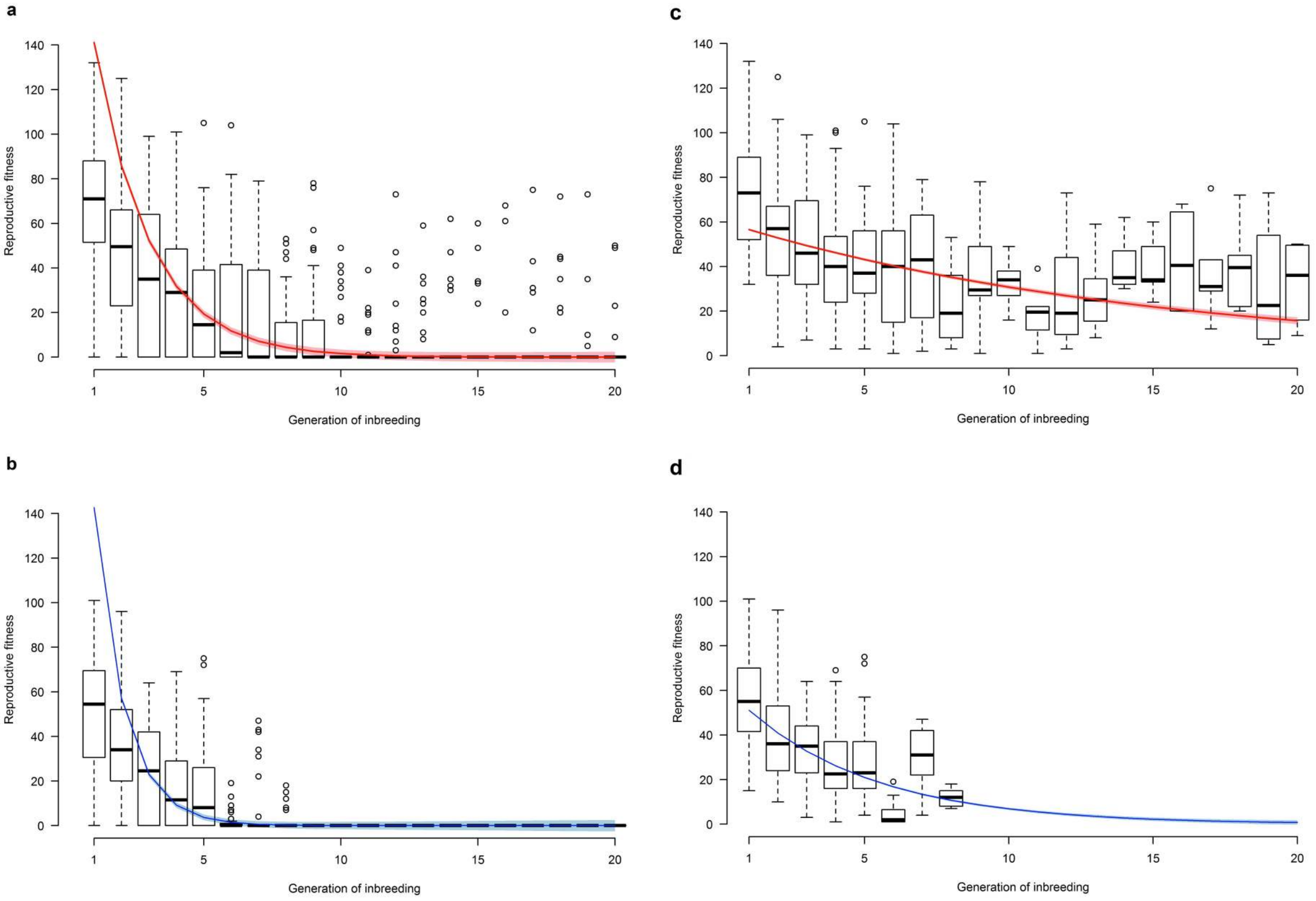
sex F_{i+1} pupae and isolate until sexually mature, use virgin F_{i+1} adults to found F_{i+2} generation

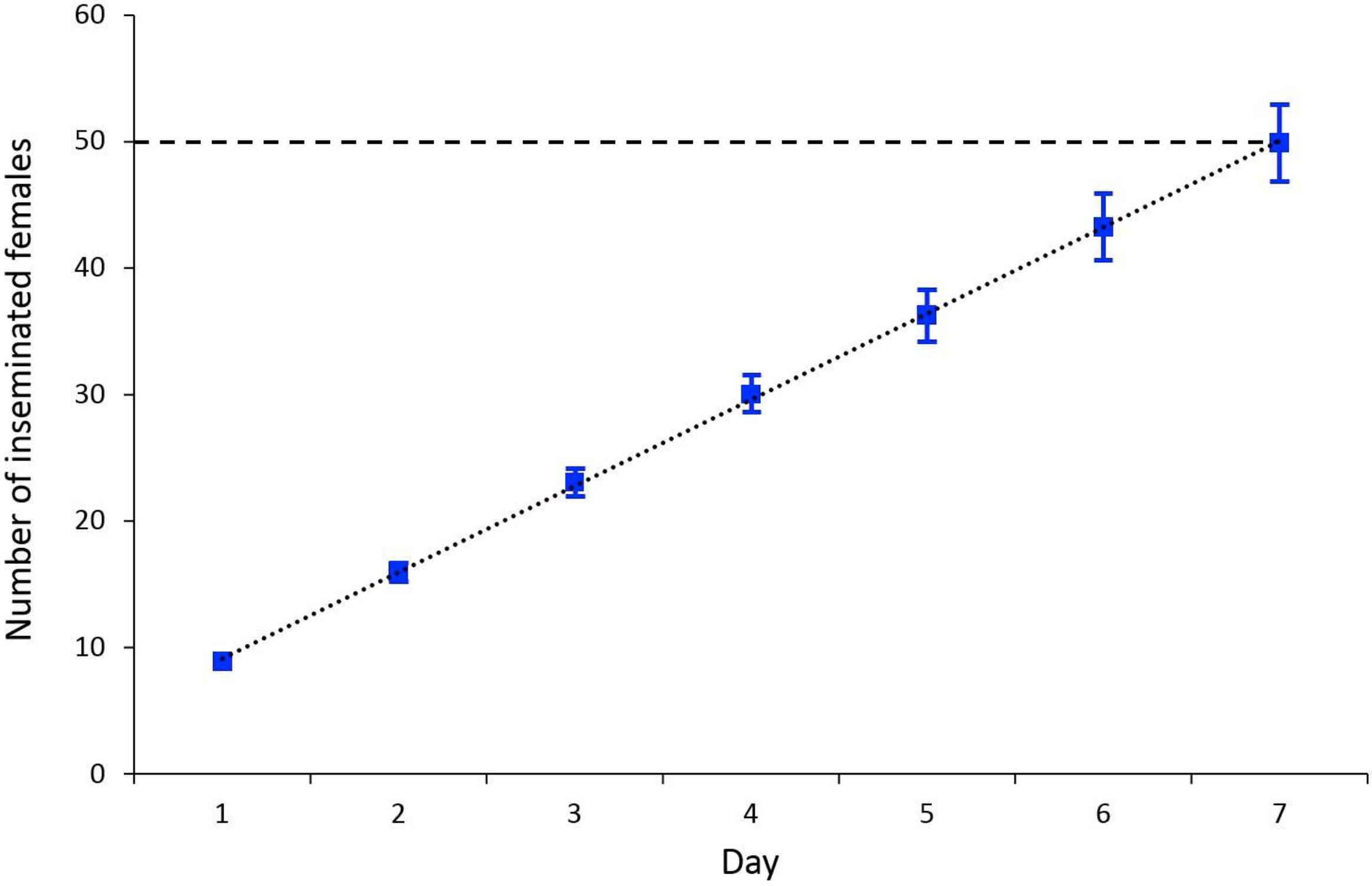

a
F0
F1
F2
Fi
♀ × ♂
sib × sib matings (= inbreeding) until extinction
or until F20 if did not go extinct
if no offspring
= line extinct
b
F0
F1
F2
Fi
♀ × ♂
sib × sib matings (= inbreeding) until extinction
or until F20 if did not go extinct
offspring counts
(F0 fitness)
offspring counts
(F1 fitness)
offspring counts
(F2 fitness)
offspring counts
(Fi fitness)


Estimated heterozygosity
0.4
0.3
0.2
0.1
0.0
Female-biased
Male-biased
Monandry
Polyandry
Regime A
Regime B


a
b
c
Proportion of families surviving
Generation of inbreeding
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
20


a
c
b
d
Reproductive fitness
Generation of inbreeding
0
20
40
60
80
100
120
140
1
5
10
15
20


a
b
c
d
Reproductive fitness
Generation of inbreeding
0
20
40
60
80
100
120
140
1
5
10
15
20


60
50
40
30
20
10
0
Number of inseminated females
1
2
3
4
5
6
7
Day

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

SUPPLEMENTARY INFORMATION

Sexual selection protects against extinction

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

correspondence: m.gage@uea.ac.uk

This supplementary PDF file includes:

S1. Molecular analyses

Samples used

For each replicate of each selection regime, the parents of all eight sibling pair lines were chosen for extraction. An additional 36 beetles (18 male and 18 female) were randomly chosen from the general populations of the same generation for each of the selection regime replicates.

DNA extractions

DNA was extracted using an ammonium acetate protocol adapted from Bruford *et al.*[39]. Beetles were broken open using pipette tips and placed in 250 µl of digestion buffer (consisting of 20 mM EDTA, 12 mM NaCl, 50 mM Tris-HCl and 5% SDS (sodium lauryl sulphate)) to which was added 9 U of Proteinase K (Roche Diagnostics). Samples were digested at 55 °C over night. Proteins were precipitated in 300 µl of 4 M ammonium acetate and pelleted by centrifugation at 4000 rpm for 15 minutes (4 °C). The supernatant was placed in new tubes and DNA was precipitated in 500 µl of ice cold absolute ethanol and pelleted by centrifugation at 4000 rpm for 15 minutes (4 °C). The supernatant was discarded and the pellet washed in 500 µl of 70% ethanol. The tubes were dried to remove residual ethanol and the DNA pellet was re-suspended in 100 µl of TE (10 mM Tris-HCl and 0.1 mM EDTA).

Testing of microsatellite loci and multiplex reactions

Twenty-three individuals representing all replicates of all the selection regimes were chosen for microsatellite locus testing. Fifty loci were selected for testing[40-42] and arranged into duplex reactions containing one FAM and one HEX labelled locus for testing. Once monomorphic and failing loci were discarded, the remaining loci were entered into Multiplex Manager version 1.2[43] in order to predict possible multiplex PCR reactions. Multiplex reactions were then tested and primer concentrations were optimised to find the best set of loci for genotyping.

PCR protocol and genotyping

All samples were genotyped at 14 microsatellite loci, using a protocol based on Kenta *et al.*[44]. One microlitre of template DNA was added to each well on a plate and the liquid was evaporated. To each tube we then added 1 µl of primer mix (see Supplementary Information Table 1 for primer details) and 1ul of 2× Qiagen Multiplex PCR Master Mix. Reaction samples were covered in mineral oil (Sigma) to prevent evaporation during heating. The PCR thermal profile was as follows: 95 °C for 15 minutes to activate the hot-start enzymes followed by 40 cycles of 94 °C for 30 seconds, 56 °C for 90 seconds and 72 °C for 60 seconds. Finally, the mixture was held at 60 °C for 30 minutes to complete the reaction. After PCR, all samples were serially diluted with H_2O to 1/125×. One microlitre of the diluted PCR product was denatured in 9 µl of Hi-Di formamide (Applied Biosystems) and GeneScan500 ROX size standard (5 µl per 1 ml of formamide). Genotypes were then read on a 3730XL sequencer (Applied Biosystems) and scored using GeneMapper version 4.0 (Applied Biosystems).

A total of 628 individuals were genotyped at 14 microsatellite loci. Details of loci tested and those contained in each multiplex are available in Supplementary Table 2. Twelve loci were in Hardy-Weinberg disequilibrium in no more than three populations. Loc11 was either monomorphic or not in Hardy-Weinberg equilibrium in all replicates and was discarded. Loc9 was in Hardy Weinberg disequilibrium in 5 populations; removing Loc 9 did not change the overall pattern of the data (analyses not shown), so was retained, leaving a dataset of 13 loci. Results are included in Extended Data Figure 8 legend. Including selection regime replicate as a random factor did not significantly improve model fit (likelihood ratio = 0, $p > 0.99$). Removing Loc9 did not change the overall pattern in the data (data not shown).

Supplementary Table 1: Details and amplification results of all microsatellite loci tested.

Locus	Source	Multiplex	Dye	Result	Primer concentration in multiplex (ul)	min size	max size	Fwd primer	Rev Primer
Tca5.44	Demuth 2007	Not Used	HEX	Fail	n/a	n/a	n/a	TTGGAGTAGCTCCGGCTAAC	TGACATCCCGATGGGTAAAT
Tca5.6	Demuth 2007	Not Used	HEX	Fail	n/a	n/a	n/a	AGCCGTATTCGCAGTGTTTT	AATCTGCAAAAATGGCAATG
Tca8.40	Demuth 2007	Not Used	HEX	Fail	n/a	n/a	n/a	GTTGGCAACAGTATTTGATTTTC	TTTCATCGTTTAATTTTGGGAAA
Tca10.2	Demuth 2007	Not Used	HEX	Monomorphic	n/a	n/a	n/a	TGAATTCAGGCATAAAACAAACA	ACAGTGATTTGATTAAGGATTTCAA
Tca5.14	Demuth 2007	Not Used	HEX	Monomorphic	n/a	n/a	n/a	CGAATTCAGTAAACCTGCCCTA	AAAACCCACGCTTGAAAAAT
Tca5.20	Demuth 2007	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	GAAACTTGCCTTGAAACATGC	ATGCCTTAATAGCCGGAACC
Tca5.7	Demuth 2007	Not Used	HEX	Monomorphic	n/a	n/a	n/a	CGTGTATGTGTTCGACAGCAA	TTGGGCATCCTATGTGTTGT
Tca5.8	Demuth 2007	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	AGCACTGAACTGTGGTACATTC	GGTGTGAACACAAACAAGGG
Tca6.19	Demuth 2007	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	TGTTGCCAAATTAAAATTAAAAG.	AATCAACATCTCGGCTACGC
Tca9.2	Demuth 2007	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	TCCCAAGTTATCGGTTTTGG	ATTGTTCCCGAACACAATGAG
Tca10.5	Demuth 2007	Not Used	6-FAM	Polymorphic	n/a	154	157	GTGGATGCGCCGGTAAAATA	GCATCCACCATTTCTGCTTT
Tca2.13	Demuth 2007	Not Used	HEX	Polymorphic	n/a	204	208	CCAAATCCGATTCAGGACAT	AACTTCCGTTTGACCCAAAAT
Tca3.1	Demuth 2007	Not Used	6-FAM	Polymorphic	n/a	205	208	CCGGCCAAACATACACATTA	CGCCTCCCGAGTTGTATTTA
Tca3.2	Demuth 2007	Not Used	6-FAM	Polymorphic	n/a	199	231	TATGTTTCCGGGTTTTGAGG	TTTCTCATACTTTTGCCGGG
Tca8.4	Demuth 2007	Not Used	HEX	Polymorphic	n/a	197	200	GGTTTGAGTGGAAGAGCAGA	TCTAGCAAACTTCAGTTGTCAAAAT
Tca9.4	Demuth 2007	Not Used	6-FAM	Polymorphic	n/a	214	273	TGTTTTCCCTTGAATGCAGA	TGCAAATTTTAGATGAGACACCC
32C7	Drury and Wade 2009	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	TCCTAAAGTCGCGGAAATTG	TTATTCACCCCGGGTAGTGT
32F3	Drury and Wade 2009	Not Used	HEX	Monomorphic	n/a	n/a	n/a	GTGCAATATTCGAAGCAAAACA	CACAGACCAGTGTTATTTGGACA
34H3	Drury and Wade 2009	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	TTCTTCAGGATGTTGCTTCC	CCAATGATGATGTGGTCGAA
LG9B7	Drury and Wade 2009	Not Used	HEX	Monomorphic	n/a	n/a	n/a	CAGAAAGCTATCAAGCTATTGGA	CACGACGTTGTAAAACGAC
LG9F3	Drury and Wade 2009	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	CGTCAAAATAGCCAAATTGTGT	CACGACGTTGTAAAACGAC
32C3	Drury and Wade 2009	Not Used	HEX	Polymorphic	n/a	195	222	AAACAATTGAGAATTTTTGTTGT	ATTTTGCGCCAACCGTATAA
32E7	Drury and Wade 2009	Not Used	6-FAM	Polymorphic	n/a	199	231	TCGGTTTGTTTCCGTAAAGG	GTGACCTGGTATTTCCATTGC
32H7	Drury and Wade 2009	Not Used	HEX	Polymorphic	n/a	212	214	CATGAGGAGACCGAGAGGAG	CGTCAGACGTTTGGACATTC
Loc13	Lagisz and Wolff 2011	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	AGGTCGAAGGCAGGACAAT	ACCAGAGAGGGATGTGCAGT
TCUB16	Pai 2003	Not Used	HEX	Fail	n/a	n/a	n/a	TTATTCGCATTTTGCGACAG	GCCAGTTTGCAGAACCAAAT
TCUB17	Pai 2003	Not Used	HEX	Fail	n/a	n/a	n/a	GATTGACATTTCGCGACCTT	ACAGTTCAGCTTCGCAACAA
TCUB1	Pai 2003	Not Used	HEX	Monomorphic	n/a	n/a	n/a	CACTTGTGCTTGGGCTTCT	AACGACTGGGAGGATTACGA
TCUB19	Pai 2003	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	GTGCTGCTGCTGTTGATGAT	ATGCACCAGCGTGAACAAT
TCUB22	Pai 2003	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	CCAAGCCCAAATCTTCGTAA	AACAAAAACCGGGCACTCTT
TCUB3	Pai 2003	Not Used	6-FAM	Monomorphic	n/a	n/a	n/a	CTCTTGTGTCCGCCCTACAT	TCCCCATCAACGTTTTTGTC
TCUB5	Pai 2003	Not Used	HEX	Polymorphic	n/a	284	300	GCCTGAAGCACCGAAACAAA	TCATCACCGAAGCATATCAAAGAG
TCUB6	Pai 2003	Not Used	6-FAM	Polymorphic	n/a	111	130	GCTGCAGCAGTATCATCAGC	GGGAAGGTAGATGGACCGTA
Tca8.1	Demuth 2007	Tribolium 1	HEX	Polymorphic	0.4	289	307	CAATTCCTGTCATTTGGTTCAA	GACAAAAGGCAAAAACAGCA
Tca8.3	Demuth 2007	Tribolium 1	6-FAM	Polymorphic	0.1	205	209	ACAACCTGCCGACATTCATC	TACTCGAGACCGGAGAATCC
34D3	Drury and Wade 2009	Tribolium 1	6-FAM	Polymorphic	0.2	193	201	TGACATAAACCCACCCCTTG	GACGAACGAAAAGGACGAAA
34E3	Drury and Wade 2009	Tribolium 1	HEX	Polymorphic	0.4	207	211	GCACAGTCAGTGTCCTTGTCA	GTCCAGTGTTGCCTGGATAAA
Loc1	Lagisz and Wolff 2011	Tribolium 1	6-FAM	Polymorphic	0.4	151	164	GGAAATTTTGCTAAATAGAACTA	AAACACGTACTTCGATTCTGATACC
Loc11	Lagisz and Wolff 2011	Tribolium 1	HEX	Polymorphic	0.2	181	193	GTCGTTCTGCATCACCTTGA	GGAAAGTACCAACAACTTGGGTAT
Tca4.3	Demuth 2007	Tribolium 2	HEX	Polymorphic	0.2	279	295	CAAAATTGGGTCTGCCTCTG	GGTCGATTGCACTTGTGATG
Tca5.9	Demuth 2007	Tribolium 2	HEX	Polymorphic	0.4	111	115	TCAACTCCTGGTCCAACTCC	TGTCATTGGACAAAAGCAAAA
Tca8.2	Demuth 2007	Tribolium 2	6-FAM	Polymorphic	0.8	220	239	TTTTTGAACGCACCGTATGA	GGAGTTAGGTGAAGTTATGCCG
32D7	Drury and Wade 2009	Tribolium 2	HEX	Polymorphic	0.8	231	240	GTATATTGTATTGCTACTTGTCC	TCCTTAGCAACGGTATCGATTT
32F7	Drury and Wade 2009	Tribolium 2	6-FAM	Polymorphic	0.2	172	182	TCTGTGGTCTGCGCTTGTAG	TTTGAACTCCGCCTGTTTGT

Supplementary Table 2: Summary statistics from microsatellite analyses for each selection regime replicate used. Loci labelled "mono" are monomorphic in that replicate. Values highlighted in green represent loci not in Hardy-Weinberg equilibrium and values labelled "ND" indicate there were insufficient heterozygotes in the replicate to test for Hardy-Weinberg equilibrium.

	FA				FB				FC				MA				MB				MC			
Locus	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-W
34D3	2	0.3585	0.3430	1.0000	2	0.0179	0.0179	ND	2	0.3962	0.4013	1.0000	2	0.6154	0.5041	0.1655	2	0.1250	0.1182	1.0000	2	0.0357	0.0354	1.0
34E3	1	mono	mono	mono	1	mono	mono	mono	Fail	Fail	Fail	Fail	2	0.2885	0.2493	0.5739	2	0.4643	0.3958	0.3000	1	mono	mono	m
Loc1	3	0.2800	0.5826	<0.001	4	0.6607	0.6433	0.3956	3	0.5849	0.5461	0.2497	2	0.3269	0.3882	0.2855	3	0.5185	0.6032	0.3321	3	0.5926	0.6225	0.6
Loc11	3	0.0189	0.1256	<0.001	1	mono	mono	mono	1	mono	mono	mono	2	0.0192	0.0192	ND	2	0.0816	0.3518	<0.001	2	0.0000	0.1931	<0
Tca8.1	3	0.5962	0.5579	0.4084	2	0.1964	0.1787	1.0000	2	0.4151	0.3321	0.0945	2	0.5000	0.4571	0.5550	2	0.4464	0.4194	0.7524	2	0.1964	0.2071	0.5
Tca8.3	2	0.5094	0.5040	1.0000	2	0.4464	0.4336	1.0000	2	0.4528	0.4744	0.7729	2	0.6154	0.4929	0.0913	2	0.5000	0.4730	0.7755	2	0.5714	0.4813	0.1
Loc4	3	0.3889	0.3197	0.3285	4	0.3929	0.3676	1.0000	2	0.5472	0.5048	0.5916	2	0.5962	0.5002	0.2593	4	0.5000	0.5056	0.0136	2	0.5000	0.4524	0.5
32D7	2	0.5094	0.5048	1.0000	2	0.3818	0.3118	0.1837	2	0.3774	0.3925	0.7371	1	mono	mono	mono	2	0.4546	0.5039	0.5873	3	0.0357	0.0356	1.0
32F7	3	0.1482	0.1712	0.3797	2	0.4182	0.4085	1.0000	2	0.3962	0.4329	0.5317	2	0.3269	0.3478	0.6916	1	mono	mono	mono	2	0.2857	0.2471	0.5
Tca4.3	2	0.5660	0.4643	0.1400	3	0.5091	0.5086	1.0000	2	0.3019	0.3536	0.4261	2	0.1731	0.3015	0.0073	2	0.3273	0.3443	0.7030	2	0.2679	0.3499	0.1
Tca8.2	5	0.4039	0.4033	0.8232	3	0.2407	0.2439	1.0000	5	0.4717	0.3998	0.8380	2	0.2308	0.2061	1.0000	3	0.4364	0.4943	0.0552	2	0.0727	0.0707	1.0
Loc9	2	0.0192	0.0192	ND	3	0.1250	0.1836	0.0032	3	0.1961	0.2270	0.0665	2	0.2115	0.2209	0.5723	2	0.1818	0.2762	0.0250	2	0.1143	0.2882	0.0
Tca4.5	2	0.1154	0.1434	0.2551	2	0.0364	0.0707	0.0549	3	0.3519	0.4027	0.4619	3	0.2500	0.2265	1.0000	2	0.1455	0.1361	1.0000	2	0.1607	0.1491	1.0
Tca5.4	2	0.1321	0.1569	0.3069	3	0.0000	0.1680	<0.001	3	0.2407	0.2459	0.0139	2	0.4231	0.3368	0.0921	2	0.2546	0.2242	0.5776	2	0.4643	0.4884	0.7

	MoA				MoB				MoC				PA				PB				PC			
Locus	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-Weinberg P	Allelic richness	ObsHet	ExpHet	Hardy-W
34D3	1	mono	mono	mono	2	0.2449	0.2760	0.5935	1	mono	mono	mono	2	0.4600	0.4160	0.5095	1	mono	mono	mono	2	0.4200	0.4467	0.7
34E3	2	0.3600	0.4073	0.4871	2	0.2600	0.2576	1.0000	2	0.4694	0.4513	1.0000	1	mono	mono	mono	2	0.0600	0.0588	1.0000	2	0.5400	0.4596	0.3
Loc1	1	mono	mono	mono	2	0.3333	0.3333	ND	1	mono	mono	mono	2	0.0000	0.5455	0.0229	2	0.2000	0.1818	1.0000	2	0.0400	0.0396	1.0
Loc11	1	mono	mono	mono	1	mono	mono	mono	2	0.0000	0.4966	<0.001	1	mono	mono	mono	1	mono	mono	mono	1	mono	mono	m
Tca8.1	3	0.4000	0.4459	0.6627	2	0.2400	0.2432	1.0000	3	0.4694	0.4778	0.0145	2	0.1200	0.1139	1.0000	2	0.0600	0.0588	1.0000	3	0.5600	0.4600	0.2
Tca8.3	2	0.3000	0.3109	1.0000	1	mono	mono	mono	2	0.0000	0.0404	0.0112	2	0.2000	0.1818	1.0000	1	mono	mono	mono	2	0.3400	0.2851	0.3
Loc4	3	0.6000	0.4721	0.1111	3	0.4000	0.3384	0.4796	3	0.5200	0.4459	0.5290	2	0.0200	0.0200	ND	2	0.7200	0.4760	0.0002	2	0.4400	0.3467	0.0
32D7	1	mono	mono	mono	1	mono	mono	mono	2	0.4000	0.4921	0.2462	2	0.2245	0.2014	1.0000	2	0.4000	0.4242	0.7414	1	mono	mono	m
32F7	1	mono	mono	mono	1	mono	mono	mono	2	0.0400	0.0396	1.0000	1	mono	mono	mono	1	mono	mono	mono	1	mono	mono	m
Tca4.3	1	mono	mono	mono	2	0.4082	0.3518	0.4160	1	mono	mono	mono	3	0.2708	0.3090	0.0135	2	0.0600	0.0588	1.0000	2	0.4800	0.5042	0.1
Tca8.2	2	0.3800	0.5000	0.1003	2	0.1400	0.1315	1.0000	2	0.1000	0.0960	1.0000	2	0.0408	0.0404	1.0000	2	0.5800	0.4806	0.2278	3	0.5800	0.5428	0.5
Loc9	2	0.0244	0.0244	ND	2	0.2200	0.1978	1.0000	3	0.0357	0.1981	<0.001	2	0.0357	0.0357	ND	2	0.1667	0.3333	0.0021	2	0.0200	0.0200	N
Tca4.5	1	mono	mono	mono	1	mono	mono	mono	4	0.3200	0.2855	1.0000	1	mono	mono	mono	1	mono	mono	mono	1	mono	mono	m
Tca5.4	2	0.0000	0.0396	0.0110	2	0.5000	0.4160	0.1796	2	0.5600	0.5042	0.5757	1	mono	mono	mono	2	0.1600	0.1487	1.0000	2	0.5200	0.5018	1.0

Supplementary Information References

39. Bruford, M., Hanotte, O., Brookfield, J. & Burke, T. *Molecular Genetic Analysis of Populations: A Practical Approach.* (IRL Press, 1998).

40. Demuth, J. P. *et al*. Genome-wide survey of *Tribolium castaneum* microsatellites and description of 509 polymorphic markers. *Molecular Ecology Notes* **7,** 1189-1195 (2007).

41. Drury, D. W., Siniard, A. L. & Wade, M. J. Genetic differentiation among wild populations of *Tribolium castaneum* estimated using microsatellite markers. *J. Hered*. **100,** 732-741 (2009).

42. Pai, A., Sharakhov, I. V., Braginets, O., Costa, C. & Yan, G. Identification of microsatellite markers in the red flour beetle, *Tribolium castaneum*. *Molecular Ecology Notes* **3,** 425-427 (2003).

43. Holleley, C. E. & Geerts, P. G. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *Biotechniques* **46,** 511 (2009).

44. Kenta, T. *et al*. Multiplex SNP-SCALE: a cost-effective medium-throughput single nucleotide polymorphism genotyping method. *Molecular Ecology Resources* **8,** 1230-1238 (2008).