Current Biology, Volume *22*

Supplemental Information

Phenotypic Engineering Unveils the Function of Genital Morphology

Cosima Hotzy, Michal Polak, Johanna L. Rönn, and Göran Arnqvist

Supplemental Experimental Procedures

We generated males with relatively short and long genital spines using two complementary but distinct approaches. First, we created replicated selection lines (population size: $N = 100$ beetles per line) where we selected for long ($N = 3$ lines) or short ($N = 3$ lines) genital spines during five consecutive generations. In each generation, we selected those 33% of the males with the longest/shortest spines to seed the next generation. Second, we used micro-scale laser ablation [1] to reduce the length of genital spines to about ½ or their original length. In one group of males, we ablated 30 ventral spines. In a second group, serving as an internal control for laser ablation *per se*, we only ablated 10 ventral spines. Post-treatment assays of male life span and food consumption showed no difference between these groups of males (see below). We then conducted two main assays using males with relatively long and short genital spines as focal males. (A) Relative male competitive fertilization success was assayed in standard double-mating experiments, where stock females were each first mated to a sterile stock male and then to a focal fertile male. Here, the proportion of hatched eggs laid subsequent to the 2nd mating measures the relative fertilization success of the 2nd (focal) male. (B) Focal males were radiolabelled by feeding them with ^{14}C-Arginine and were then mated to previously once-mated stock females. Females were frozen 18 h after their 2nd copulation and the amount of radio isotopic label present in their reproductive tract (bursa copulatrix and copulatory canal) and the rest of their body were measured in a scintillation analyser. In experimental matings, we also recorded potential confounding variables, such as male and female body size, mating duration and ejaculate weight. Below, we describe the experimental procedures in more detail.

Stock and Maintenance

We used the South Indian (SI) population of the seed beetle *Callosobruchus maculatus* (Coleoptera, Bruchidae) for all our experiments, maintained at a large population size ($N > 500$). Beetles were reared on mung beans (*Vigna radiata*) in climate chambers at 30°C, 55% RH and a 12:12 diurnal light cycle. In order to generate virgin individuals, single beans with larvae were isolated in cell culture well plates prior to emergence of adult beetles. Experiments were conducted at the Uppsala University (work involving selection lines and other experiments) and at the University of Cincinnati (all work involving spine ablated males).

We tested the key hypotheses that genital morphology affects (i) male competitive fertilization success, (ii) copulation duration, (iii) female egg production and (iv) the dispersal of male seminal fluid inside the female after mating, using two independent and replicated full sets of experiments involving males with genetically (1st set) or phenotypically (2nd set) manipulated genital spine length.

Genital Spine Selection Lines

We established replicated artificial selection lines (population size per line N = 100; see below) for the length of genital spines (3 lines selected for long and 3 for short spines) from the base population using the following selection protocol. Batches of 12 unmated males at a time were inspected simultaneously under a dissecting microscope (Leica® MZ7 5) and first divided into two halves, those deemed to have the longest (L_a) and the shortest (S_a) relative spine length, based on the simultaneous impression of both lateral and ventral genital spines. From L_a, the four males with the longest relative spine length were then selected (L_b) and from S_a the four males with the shortest relative spine length were selected (S_b). During the first generation, L_b males were used to seed the long spine selection lines and S_b males to seed the short spine selection lines. During subsequent generations, only L_b males were retained in the long lines and S_b males in the short lines. Thus, the selection protocol aimed at selecting the 33% tail of the distribution (4 out of 12 males in each batch) in each generation. In each selection episode and line, we repeated the selection process until 50 long/short spined males had been selected from each line. We note, however, that the selection protocol was no doubt less than perfect, as selection was based on visual estimation of relative spine length within groups of 12 males at a time. Hence, our estimate of the heritability of spine length that was subsequently calculated from the response to selection represents a conservative estimate, as realized selection intensity was likely lower than that assumed in these calculations [2].

Selection was carried out by first anaesthetizing males with CO_2, keeping them under a constant light CO_2 flow (FlowBuddy, Genesee Scientific®) under the microscope. To enable comparison of the genital spine lengths, the aedegus of each male was first individually everted using an adjustable vacuum pump connected to a fine tube. We chose to base the selection on batches of 12 males simultaneously to minimize the time males spent anaesthetized while retaining selection efficiency. The 50 long/short spined males that had been selected from each line were then introduced with 50 randomly selected virgin females from the same line to seed the next generation. Thus, the population size of each line was 100 individuals. Each line was kept in a 1 litre glass jar provided with 150 g of mung beans.

During generation 1, three sets of replicate lines (L1 & S1, L2 & S2 and L3 & S3) were started subsequent to one another, using three consecutive cohorts of the base line (spaced two – three days apart) which were kept in temporal succession (i.e., time-staggered cohorts) throughout both the selection procedure and the experiments described below. Cohort was then used as a factorial variable in those analyses which were perfomed at an individual level (see below). Artificial spine selection was performed for five generations. Virgin males from the sixth generation were used for the experiments described below.

Genital Spine Ablation

We used a laser ablation system to shorten ventral genital spines, because correlational data have suggested that these spines are related to success in sperm competition [3]. We used two treatment groups, one which had 30 ventral spines of the aedegus ablated and a second which only had 10 ventral spines ablated. Males were randomly assigned to one of these two groups. Following anaesthetization and eversion of the male genitalia (see above), males were transferred to a surgical chamber mounted on a motorized stage (Prior® H117, Rockland, MA, USA) of an inverted light microscope (IX71, Olympus®, Center Valley, PA, USA) integrated with a Q-switched laser (Vector 532-1000-20, Coherent, Santa Clara, CA, USA) [see 1]. We used pulsed laser light (at 40% power) to ablate individual spines. We aimed at cutting a single spine to approximately half its original length with each laser shot (Fig. 1). In both spine ablation treatments, half of the ventral spines were cut from one lateral side and half from the other. Males were allowed 2-4 days to recover post-treatment, during which they were held in Petri dishes (Ø 3 cm) provided with 20% sugar solution, prior to participating in the mating experiments described below.

We aimed to ablate the sclerotized genital spines at a point where they do not contain any hemolymph and to ensure that the laser shots did not hit any other male tissue. Nevertheless, to verify that males in our two treatment groups were not differentially affected by the surgical laser treatment *per se*, we compared the amount of water/food consumed after the treatment as well as the life span of all males used in the experiment (see below).

Assay I: Male Fertilization Success

We used a standard double-mating experiment [4], where randomly selected base population females were first mated to a virgin sterilized base population control male and then to a focal virgin experimental male. We confirmed that control males were sterile. The proportion of eggs laid after the second mating that hatched are thus fertilized by the second male (i.e., P2) and forms our measure of the relative fertilization success of the focal males. We estimated the fertilization success of 22-25 replicate males per line/treatment.

Base population males were sterilized by exposures to gamma radiation using either a caesium source at the Rudbeck laboratory in Uppsala (selection line experiment; dose 80 Gy) or a cobalt source at the Department of Nuclear & Radiological Engineering at the University of Cincinnati (spine ablation experiment; dose 87 Gy). The dosage was set to 80-90 Gy since this dose leads to complete and lasting sterilization but does not notably affect sperm competitive ability of sterilized males [5,6]. During irradiation, 10 to 20 males were kept together in a single Petri dish to avoid crowding and unequal doses of radiation. Following irradiation, males were separated and were kept isolated in Eppendorf tubes to recover from the sterilization until the mating experiments.

One to two days after sterilization, 24 – 48 hours old virgin females were each mated to a virgin sterilized base population male. Here, one male and one female were introduced into a small Petri dish (Ø 3 cm) and mating was visually verified and recorded (this applies to all experimental matings). After copulation was completed, each female was transferred into an Eppendorf tube containing mung beans as oviposition substrate. The females were kept under rearing conditions until the second mating.

Two days after their first mating, females were remated to focal virgin males. In the selection line experiment, males from the selection lines were randomly selected from virgin chambers one day before the mating trials. In the laser ablation experiment, the focal males were selected from one of the two treatment groups. Focal second matings were performed as with the first matings, but we placed mung beans into the mating Petri dish as this facilitates female remating. In a few cases where females did not remate within one hour, males (selection line experiment) or females (spine ablation experiment) were replaced. In all copulations involving spine ablated males, we recorded copulation duration (grand mean copulation duration for spine ablated males was 8.41 minutes; 95% CI: 7.77 – 9.15). For selection line males, copulation duration was instead assessed in the experiment described below. Laser ablated males were isolated after mating and kept under rearing conditions. Their lifespan was subsequently determined by daily inspections.

Twice-mated females were isolated individually in Petri dishes (Ø 6cm) containing an *ad-libitum* supply of mung beans for oviposition, to avoid larval cannibalism. Under these conditions, females of the SI population very rarely deposit more than a single egg per mung bean. Females were allowed to oviposit for seven days following their second mating. The proportion of eggs fertilized by the first (sterile) and second (fertile) male was determined by counting the number of unhatched and hatched eggs, respectively, seven days after the removal of the female. All eggs laid by females between the first and the second mating were also counted, separately, as this is known to markedly affect the relative fertilization success of the second male [3,7]. Six eggs out of a total of 2279 eggs laid by females in between matings hatched, confirming that first males were indeed sterile (fertility rate 0.3%). As a measure of body size, we measured the elytra length of all individuals used in the experimental matings (see below).

In addition to our factorial variable (i.e., male spine length), we recorded a number of potential confounding factors and subsequently assessed whether these significantly explained residual variation in male fertilization success by including them in statistical models based on individual matings (see below). These were copulation duration, female body size, body size of both first and focal males, number of eggs laid between first and second mating, age of first male at sterilization, age of first male at mating, time between sterilization and first mating, female age at second mating, age of focal male at mating, time between laser ablation and mating and the time between first and second mating.

Assay II: Ejaculate Dispersal within Females

To trace the passage of seminal fluid substances of long and short spined males within females, from the reproductive tract into the body, we performed experimental matings where once-mated females were re-mated to radiolabelled focal males. We then measured the proportion of radiolabel within different female body parts. These assays were replicated in 14-19 females per selection line and in 24 females per treatment group for spine ablated males. In addition, we ran 20 blank scintillation controls for selection line matings and 18 blank controls for laser ablated matings.

One day old virgin focal males were mated to virgin base population females to stimulate ejaculate production. For selection line males, copulation duration was recorded during these matings to assess the effects of genital spine length on copulation duration (grand

mean copulation duration for selection line males was 9.36 minutes; 95% CI: 8.59 – 10.13). After this initial mating, males were radiolabelled by immediately transferring them to a Ø 3 cm Petri dish provided with a feeding vial filled with 340 µl of a 1:1 mixture of ^{14}C-Arginine : 20% sugar solution (Arginine L-[^{14}C(U)], 250µCi, PerkinElmer®), or pure 20% sugar solution (blank controls), and were kept there for 24 hours. In the spine ablation experiment, males were ablated between the first copulation and the feeding period. The focal males and blank control males were then all mated a second time to another base population female one day after their first mating and were then reintroduced into their feeding dish to gain more radiolabel. They were used for mating experiments the next day and were hence all three days old at their focal (and third) mating. The experiment involving spine ablated males was conducted in two consecutive blocks.

The focal females in this experiment were first mated to an un-labelled base population male each at age one-day post-eclosion, and were isolated in Petri dishes provided with mung beans for oviposition. Three days later, they were re-mated to a focal radiolabelled male (or a blank control male) and were then again isolated with a new set of mung beans for oviposition. Females were frozen 18 hours (±10 Min) after their second and focal copulation. In the experiment involving selection line males, both the first and second males were weighed to the nearest 10^{-5}g using an electronic microbalance (Sartorius® Genius Series ME235P-0CE) immediately before and after copulation to estimate ejaculate weight and male body size. Here, each reading was taken twice in immediate succession and the mean of the two measurements was used. In the experiment involving laser ablated males, males were not weighed. We instead used elytra length as a measure of male body size. Elytra length was used as a measure of female size in both experiments. Radio labelled males and blank control males were all frozen immediately after mating, for subsequent scintillation analyses. Eggs laid by females after the first and second mating were counted separately.

Focal females were thawed and rinsed thoroughly three times in 70% ethanol to remove any external radio label residue they might have acquired during copulation. Females were then dissected under a dissecting microscope (Leica® MZ7 5) on a small square of gelatine (1.2 cm^2) placed in a Petri dish (Ø 12cm), using micro-scissors, fine forceps and insect needles. The gelatine square was placed on a drop of water (10µl) to fix it to the Petri dish. In *C. maculatus*, the reproductive tract (here defined as the external genitalia, copulatory duct and bursa copulatrix) forms a well defined and integrated unit which is well separated from the narrow common oviduct. The reproductive tract of each female was carefully removed, to avoid potential leakage, and placed in an Eppendorf tube. The rest of the female's body (including the gelatine) was placed in a separate Eppendorf tube. All radio labelled males were also rinsed thoroughly three times in 70% ethanol and then placed in separate Eppendorf tubes and lightly macerated using a steel pestle. Dissecting instruments and pestles were thoroughly rinsed in 70% ethanol between each individual. All Eppendorf tubes were stored at -18 °C until scintillation analyses. Two days prior to scintillation analysis, a tissue solubilizer was added to the samples. We used TS-2 (Koch-Light Research Laboratories®, England) for material involving selection line males and Solvable (PerkinElmer®, USA) for material involving spine ablated males. The amounts of tissue solubilizer added to each Eppendorf tube were 100 µl for focal males, 200 µl for female body and 50 µl for female reproductive tract. Two days after adding the tissue solubilizer, samples were centrifuged for 1 Min at 10.7x1000 rpm to ensure that the

tissues were covered by the tissue solubilizer. Twenty-four hours later, 1.5 ml scintillation cocktail (Optiphase 'Hisafe' 2, PerkinElmer®, USA) was added to each Eppendorf tube.

Twenty-five hours after adding the scintillation cocktail, the samples were analysed using a liquid scintillation analyser (Packard®, TRI-CARB 2100TR) to measure the radioactivity of the different samples. All samples were analyzed twice, for ten minutes each, and the mean value of CPM during these two runs was used for subsequent analyses. The repeatability of the two CPM readings was very high indeed (all $R > 0.98$). Each run contained between 6 and 9 blank controls in addition to the labelled samples to measure background radioactivity. We compensated for the background by subtracting the mean reading of the blank controls in each run from the mean of each sample readings.

In these experiments, we recorded a number of potential confounding factors (apart from our factorial variable - male spine length) which were assessed by inclusion in models based on individual matings (see below). These were copulation duration of focal mating, ejaculate weight in the first mating, ejaculate weight in the second (focal) mating, amount of radio label in focal male after mating, female body size, body size of both first and focal males, number of eggs laid between first and second mating, number of eggs laid after the second mating and age of first male at mating.

Female Remating Rate Assay

To assess whether female re-mating rate is affected by the length of the genital spines of her mate, we conducted an additional mating experiment using base population females and males from the artificial selection lines. In each replicate, we introduced 8 virgin base population females with 8 males deriving from one of the selection lines into a Ø 6 cm Petri dish. We then recorded all observed copulations during 30 minute interval spot-checks for three consecutive days (11-14 spot checks per day). Thus, each observation represents the mean observed remating rate of 8 females. We set up five such replicates per selection line (total $N = 30$).

Morphometrics

To measure the response to artificial selection, we measured the genital spine length and body size of random samples of males from our selection lines in generation six ($N =$ 40 per selection lines and $N = 111$ for the base population). Genitalia were inflated (see above) and males were carefully placed into 95^{o}C water for 20 seconds. This resulted in primary fixation of the genitalia due to instant coagulation of proteins. Each male was then transferred to Bouin's solution for final fixation. The males were subsequently stored in Bouin's solution at room temperature for approximately 3 weeks.

The genitalia were then removed and submerged in a drop of glycerol on a microscope slide and placed under a dissecting microscope (Leica® MZ8). Each aedegus was oriented in two different positions to measure the length of the lateral and the ventral genital spines [see 3]. In each position, the 5 longest spines were measured by means of a digitizing tablet (SummaSketch III, Summagraphics®) placed under a dissecting microscope provided with a camera lucida (Leica® MZ8). Spine length was measured as the distance between the tip and base of each spine and we used the mean lateral and the mean ventral spine length per male for subsequent analyses. Sample identities were blind to the person taking measurements and were performed at the same magnification and with the genitalia oriented in the same manner.

We note here that two different groups of males, from each of the six selection lines, were measured at different times: approximately half of the measured males derived from the fertilization success experiment and half were males not involved in any experiment. We used "group" as a factor in subsequent analyses of morphology based on individual males.

Elytra length was measured of all individuals used in the mating experiments, where data on body weight was absent, by either (i) the use of dial callipers (0.02 mm resolution) (focal females from the ejaculate dispersal experiments and focal males from the ejaculate dispersal experiments using laser ablated males) or (ii) by means of a digitizing tablet placed under a dissecting microscope provided with a camera lucida (all other beetles). Both left and right elytra were measured twice and the mean of these four measurements was used.

Statistical Analyses and Modeling

Our main inferential tests of the effects of artificial selection of genital morphology were based on mean values per line [8], using conventional two-sample *t*-tests. Analyses at the individual level are only presented here in the supplementary information for completeness and comparison. We note, however, that analyses at these two levels yielded results that were identical in terms of our ability or inability to reject null hypotheses.

Male competitive fertilization success (i.e., P2) in *C. maculatus* is strongly affected by the time between matings [3,7], which varied across replicate females at an individual level (Table S2). To enable a test based on mean male fertilization success per selection line, we used the following inferential route. We first validated that the relationship between male fertilization success and time between matings did not differ among lines in an analysis of covariance of square-root arc-sine transformed P2, including line as a factor, time between matings as a covariate and their interaction. Here, the interaction was not significant ($F_{5,95}$ = 0.85, P = 0.518), suggesting that slopes are parallel. We then regressed square-root arc-sine transformed P2 values on the time between matings in a single global model including all individuals and saved the residuals from this model. We then tested whether mean residual P2 per line differed between long and short spine selection regimes.

Because a major determinant of the radiolabel content in the body of females is the amount of radiolabel content in their reproductive tract (Table S3), reflecting the fact that individual females received different amount of total radiolabel from males, we analysed mean radiolabel content in the body of females per selection line in a similar manner. We first validated that the relationship between radiolabel content in the body and radiolabel content in the reproductive tract did not differ among lines in an analysis of covariance of radiolabel content in the body, including line as a factor, radiolabel content in the reproductive tract as a covariate and their interaction. The interaction was not significant ($F_{5,99}$ = 0.75, P = 0.584), suggesting that slopes are parallel. We then regressed the amount of radiolabel in the body on the amount of radiolabel in the reproductive tract in a single global model including all individuals and saved the residuals from this model. We then tested whether mean residual label in the body per population differed between long and short spine selection regimes.

Complimentary analyses of selection lines and all analyses of spine ablated males involved fitting either general linear models, for continuously distributed response variables, in SYSTAT® v. 13 or generalized linear models using JMP® v. 9. The former were fitted using conventional LS minimization with Type III SS. The latter was employed to analyse variation in

male fertilization success, using a binomial error distribution and a logit link function, where the number of fertile eggs laid after the second mating was the response and the total number of eggs laid after the second mating was the binomial denominator. These models were fitted with Firth's penalized maximum likelihood estimation and used Pearson χ^2 adjustment for overdispersion. Model fit to data was assessed by residual plots and for general linear models also by tests for homogeneity of variance (Levene's test) and normality of residuals (Kolmogorov-Smirnov test). Data were found to meet the assumptions made for all inferential models, following the exclusion of a few deviant observations. These were strictly defined as having an absolute value of the standardized residual > 2 (standardized deviance residual for generalized linear models) (in total for all data, $N = 4$ for data on male fertilization success and $N = 7$ for data on ejaculate dispersal within females). We present those models (among larger sets of alternative models containing combinations of covariates [standardized prior to modelling]; see above) that minimized the Akaike information criterion.

We tested for an effect of spine length on ejaculate size in two distinct ways. We first tested directly for the effects of spine length on ejaculate weight (male weight before and after copulation) using data from the selection line males in the ejaculate dispersal experiment. We then also used the radio label data to estimate the amount of ejaculate transferred per copulation. For spine ablated males, this was tested in an analysis of covariance of the total signal in females after the copulation which included spine treatment as a factorial variable and total signal in the male as a covariate (test of the parallelism assumption: $F_{1,50} = 0.85$, $P = 0.360$). For our selection lines, we used the following assessment to enable a test based on mean male fertilization success per selection line (see also above). We first validated that the relationship between total signal in females after the copulation and male signal did not differ among lines in an analysis of covariance of total signal in females, including line as a factor, signal in males as a covariate and their interaction. Here, the interaction was not significant ($F_{5,94} = 0.31$, $P = 0.906$), suggesting that slopes are parallel. We then regressed total signal in females on total signal in males in a single global model including all individuals and saved the residuals from this model. These residuals, thus, quantify the amount of signal received by a female from her mate given the concentration of signal present in the male. We then tested whether mean residual signal per line differed between long and short spine selection regimes.

Supplemental Results

Genital spine length showed a strong response to artificial selection but male body size did not show any correlated response to selection on spine length (Figure S1, Table S1), suggesting that the genetic correlation between genital spine length and male body size is at most very weak. Moreover, genital spine length did not significantly correlate phenotypically with male body size (Table S1).

Among spine ablated males, mean post-mating lifespan was 15.60 (s.e.m. = 0.60) days in males that had 10 spines ablated and 14.63 (s.e.m. = 0.69) days in males that had 30 spines ablated (t_{47} = 1.068, P = 0.291). Further, the total amount of radiolabel present in males (a measure of total food and water consumption during 48 hours following the spine ablation; see below) was 66.16×10^3 (s.e.m. = 8.09×10^3) CPM in males that had 10 spines ablated and 64.97×10^3 (s.e.m. = 8.54×10^3) CPM in males that had 30 spines ablated (t_{60} = 0.101, P = 0.920). The statistical power (i.e., 1- β) of these two t - tests for α = 0.05, assuming a true effect size of d = 0.8 [9] and using the empirically derived standard deviations, were 0.81 and 0.88, respectively. We thus conclude that these two groups of males showed no marked differences in how they were affected by spine ablation. In addition, we note that the observed mean total lifespan of the laser ablated males used in our experiment (17-20 days) is rather long even when compared to unmanipulated males of this species (e.g., 16-18 days [10]), again suggesting that any adverse effects of our surgical treatment were very minor at most.

Radiolabel Analyses

Our general linear models of female uptake of male seminal fluid material (Table 2 and Table S3) implicitly assume that males with relatively short and long spines did not vary systematically in label content and that the three label content variables are linearly related. The total amount of radio label present in males with relatively long and short spines did not differ significantly, either among laser ablated males (t_{60} = 0.101, P = 0.920) or among selection line males ($F_{1,104}$ = 0.372, P = 0.543). We tested for linearity in a series of quadratic regressions based on zero-centered label content variables, in which the linear term was always highly significant (P < 0.001 in all cases). Linearity was assessed by partial F-tests of the increase in fit to data by the addition of a quadratic term to the linear model. The relationship between signal in males and in female bursa was not significantly non-linear (selection line data: $F_{1,96}$ = 1.663, P = 0.200; laser ablation data: $F_{1,78}$ = 0.244, P = 0.623), neither was the relationship between signal in males and in the rest of the female body (selection line data: $F_{1,101}$ = 3.499, P = 0.064; laser ablation data: $F_{1,79}$ = 2.248, P = 0.138) or the relationship between signal in female bursa and in the rest of the female body (selection line data: $F_{1,98}$ = 0.435, P = 0.511; laser ablation data: $F_{1,79}$ = 0.044, P = 0.834). Hence, we found no evidence for non-linearity among the three label content variables.

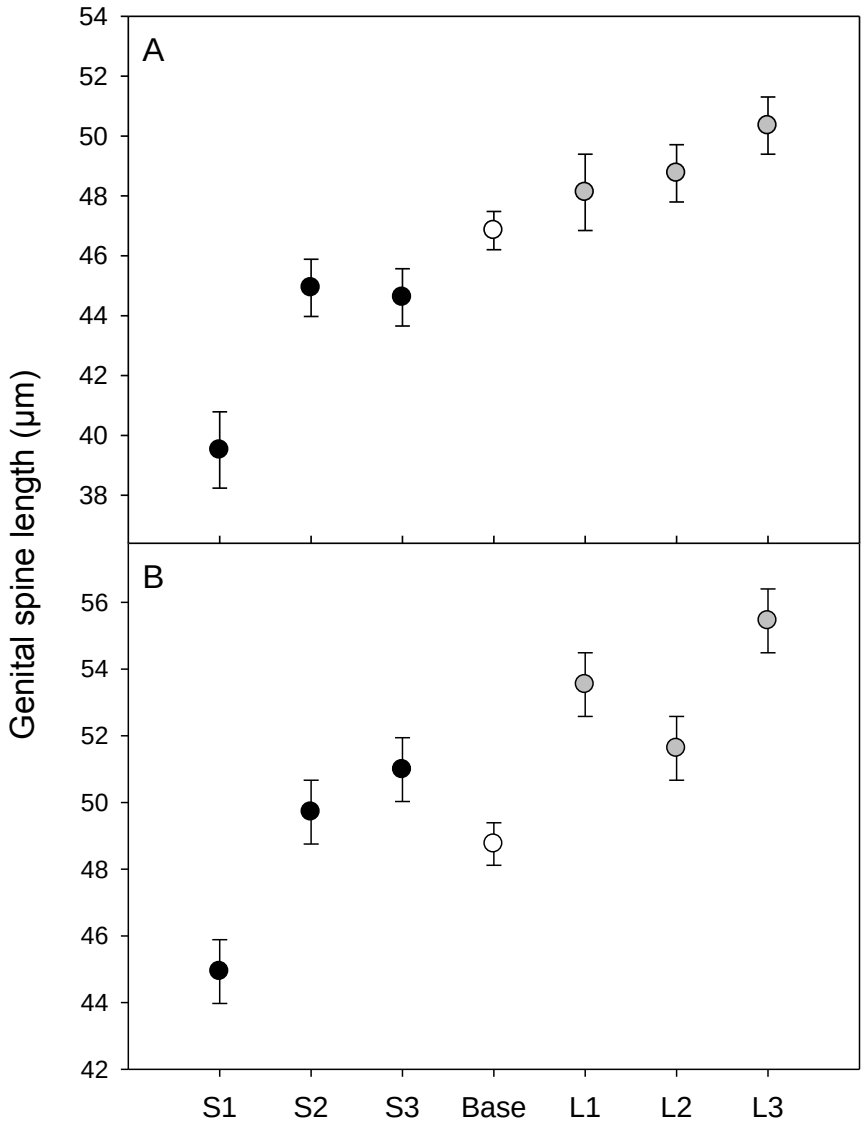

Figure S1. Effects of Artificial Selection on Genital Spine Length

The mean (± s.e.m.) length of (A) ventral and (B) lateral male genital spines in replicated lines where artificial selection was imposed for five generations for short (S1-S3) or long (L1-L3) genital spines (N = 40 per line). The length of genital spines in the base population is also shown for reference (N = 111). The base population differed significantly from the long ($F_{1,223}$ = 19.28, P < 0.001) and short ($F_{1,223}$ = 6.80, P = 0.010) selection lines in ventral spine length, and from the long ($F_{1,223}$ = 22.33, P < 0.001) but not the short lines ($F_{1,223}$ = 0.02, P = 0.91) in lateral spine length (focussed post-hoc contrasts from ANOVAs based on data on individual males).

Table S1. Effects of Artificial Selection on Genital Spine Length and Body Size

		Lateral spine length		Ventral spine length		Elytra length	
Source	d.f.	$F_{(216)}$	*P*	$F_{(216)}$	*P*	$F_{(217)}$	*P*
Selection	1	34.94	<0.001	47.97	<0.001	0.24	0.626
Cohort	2	7.74	0.001	5.83	0.003	4.42	0.013
Group	1	23.29	<0.001	0.04	0.848	1.21	0.272
Selection × Cohort	2	4.71	0.010	2.17	0.117	1.14	0.32
Selection × Group	1	3.01	0.084	1.70	0.194	0.0	0.823
Group × Cohort	2	8.61	<0.001	0.00	0.999	2.29	0.103
Selection × Group x Cohort	2	0.57	0.567	0.23	0.796	0.21	0.807
Elytra length	1	0.09	0.767	0.32	0.570	-	-

Given are the results of analyses of covariance across all measured males of the length of genital spines, and male body length, following five generations of artificial selection for long and short genital spines. Denominator d.f. for *F* – ratios are given within brackets.

Table S2. The Effects of Spine Length on Male Competitive Fertilization Success

Source	d.f.	χ^2	*P*
Selection	1	9.92	0.002
Cohort	2	0.98	0.613
Selection × Cohort	2	1.17	0.558
Time between matings	1	6.93	0.008

Given are the results of generalized linear models of variation in sperm precedence (P2) across individual females, using males from the selection lines (*N* = 105). Cohort refers to time-staggered sets of replicate selection lines.

Table S3. The Effects of Spine Length on Female Uptake of Male Seminal Fluid Material

Source	d.f.	MS	*F*	*P*
Selection	1	4.21×10^6	6.22	0.014
Cohort	2	3.88×10^6	5.74	0.004
Selection x Cohort	2	8.18×10^5	1.21	0.303
Label in reproductive tract	1	3.86×10^8	571.24	< 0.001
Female body size	1	2.71×10^6	4.00	0.048
Error	93	6.76×10^5		

The results of analyses of covariance of variation in the amount of radiolabeled (^{14}C) ejaculate-derived material in the female body (outside the reproductive tract) after mating to males from the selection lines, across all individual females. Cohort refers to time-staggered sets of replicate selection lines.

Supplemental References

1. Polak, M., and Rashed, A. (2010). Microscale laser surgery reveals adaptive function of male intromittent genitalia. Proc. R. Soc. Lond. B Biol. Sci. 277, 1371-1376.
2. Falconer, D.S., and Mackay, T.F.C. (1996). Introduction to Quantitative Genetics (Harlow: Longman).
3. Hotzy, C., and Arnqvist, G. (2009). Sperm competition favors harmful males in seed beetles. Curr. Biol. 19, 404-407.
4. Simmons, L.W. (2001). Sperm Competition and Its Evolutionary Consequences in the Insects (Princeton: Princeton University Press).
5. Ahmed, M.Y.Y., Younes, M.W.F., and Al-Arab, K.T. (1977). Sexual competition of gamma sterilized male of *Callosobruchus maculatus* (F.) (Col., Bruchidae). Z. Angew. Entomol. *84*, 295-298.
6. Maklakov, A.A., and Arnqvist, G. (2009). Testing for direct and indirect effects of mate choice by manipulating female choosiness. Curr. Biol. 19, 1903-1906.
7. Eady, P.E., Rugman-Jones, P., and Brown, D.V. (2004). Prior oviposition, female receptivity and last-male sperm precedence in the cosmopolitan pest *Callosobruchus maculatus* (Coleoptera : Bruchidae). Anim. Behav. *67*, 559-565.
8. Rice, W.R., and Holland, B. (2005). Experimentally enforced monogamy: inadvertent selection, inbreeding, or evidence for sexually antagonistic coevolution? Evolution *59*, 682-685.
9. Cohen, J. (1988). Statistical Power Analysis for the Behavioural Sciences (Hillsdale: Lawrence Erlbaum Associates).
10. Maklakov, A.A., Bonduriansky, R., and Brooks, R.C. (2009). Sex differences, sexual selection and ageing: an experimental evolution approach. Evolution *63*, 2491-2503.