SUPPLEMENTARY INFORMATION

Table S1: Mixed model results for litter size and wean weight from breeding cages.

Litter Size	GLMM with Poisson distribution and logarithmic link (groups = 30, observations = 65)			
Random effects	*Variance*	*Std. Dev.*		
Cage (Intercept)	0.0173	0.1316		
Fixed effects	*Estimate*	*Std. Error*	*z value*	*Pr(>|z|)*
Intercept (Litter 1)	1.4721	0.1047	14.0590	<0.0001***
Treatment (Rofecoxib)	-0.3432	0.2020	-1.6990	0.0894
Parity	0.0267	0.0903	0.2950	0.7678
Treatment (Rofecoxib)×Parity	0.2085	0.1929	1.0800	0.2800
Female Wean Weight	LMM (groups = 23, observations = 96)			
Random effects	*Variance*	*Std. Dev.*		
Cage (Intercept)	0.8188	0.9049		
Fixed effects	*Estimate*	*Std. Error*	*t value*	*Pr(>|t |)*
Intercept (Litter 1)	11.1311	0.3242	34.3330	<0.0001***
Treatment (Rofecoxib)	-0.5296	0.6608	-0.8010	0.4270
Parity	-0.0937	0.2064	-0.4540	0.6510
Treatment (Rofecoxib)×Parity	0.3591	0.4697	0.7640	0.4470
Male Wean Weight	LMM (groups = 25, observations = 97)			
Random effects	*Variance*	*Std. Dev.*		
Cage (Intercept)	1.0400	1.0200		
Fixed effects	*Estimate*	*Std. Error*	*t value*	*Pr(>|t |)*
Intercept (Litter 1)	12.0867	0.3715	32.5300	<0.0001***
Treatment (Rofecoxib)	-1.0360	0.8101	-1.2800	0.2048
Parity	0.6496	0.2277	2.8500	0.0054**
Treatment (Rofecoxib)×Parity	0.0250	0.5889	0.0400	0.9663

** Indicates a p value < 0.01, ***< 0.001.

Table S2: Linear mixed model results for body weight over time within OPAs.

Female Body Weight	LMM (groups = 74, observations = 337)			
Random effects	*Variance*	*Std. Dev.*		
Individual (Intercept)	9.2891	3.0478		
Population (Intercept)	1.9916	1.4112		
Fixed effects	*Estimate*	*Std. Error*	*t value*	*Pr(>|t |)*
Intercept (Week 0)	19.4696	0.9668	20.1380	<0.0001***
Treatment (Rofecoxib)	0.2144	1.0208	0.2100	0.8340
Time	0.3916	0.0326	12.0220	<0.0001***
Treatment (Rofecoxib)×Time	-0.0357	0.0442	-0.8080	0.4200
Male Body Weight	LMM (groups = 39, observations = 142)			
Random effects	*Variance*	*Std. Dev.*		
Individual (Intercept)	2.3450	1.5314		
Fixed effects	*Estimate*	*Std. Error*	*t value*	*Pr(>|t |)*
Intercept (Week 0)	21.4308	0.5111	41.9270	<0.0001***
Treatment (Rofecoxib)	-1.3016	0.7323	-1.7770	0.0793
Time	0.1116	0.0246	4.5420	<0.0001***
Treatment (Rofecoxib)×Time	0.0715	0.0361	1.9830	0.0495*

* Indicates a p value < 0.05, *** < 0.001.

Table S3: Generalized linear mixed model results for reproduction and male competitive ability over time within OPAs.

Female Reproduction	GLMM with Poisson distribution and logarithmic link (groups = 4, observations = 34)			
Random effects	*Variance*	*Std. Dev.*		
Population (Intercept)	0.1755	0.4189		
Population (Slope)	0.0008	0.0283		
Fixed effects	*Estimate*	*Std. Error*	*z value*	*Pr(>|z|)*
Intercept (Week 8)	2.8937	0.2283	12.6760	<0.0001***
Treatment (Rofecoxib)	0.4393	0.1129	3.8910	<0.0001***
Time	0.0063	0.1616	0.3900	0.6960
Treatment (Rofecoxib)×Time	0.0116	0.0097	1.2040	0.2290
Male Reproduction	GLMM with Poisson distribution and logarithmic link (groups = 5, observations = 44)			
Random effects	*Variance*	*Std. Dev.*		
Population (Intercept)	0.1036	0.3219		
Population (Slope)	0.0002	0.0150		
Fixed effects	*Estimate*	*Std. Error*	*z value*	*Pr(>|z|)*
Intercept (Week 8)	2.3056	0.1786	12.9100	<0.0001***
Treatment (Rofecoxib)	0.0273	0.1468	0.1860	0.8530
Time	0.0156	0.0113	1.3830	0.1670
Treatment (Rofecoxib)×Time	-0.0006	0.0117	-0.0520	0.0959
Male Competitive Ability	GLMM with binomial distribution and logit link (groups = 5, observations = 104)			
Random effects	*Variance*	*Std. Dev.*		
Population (Intercept week 3)	0.1287	0.3587		
Fixed effects	*Estimate*	*Std. Error*	*z value*	*Pr(>|z|)*
Intercept (Week 3)	-0.8640	0.2525	-3.4220	0.0006***
Treatment (Rofecoxib)	0.0208	0.2767	0.0750	0.9401
Time	0.0272	0.0158	1.7210	0.0853
Treatment (Rofecoxib)×Time	-0.0059	0.0221	-0.2670	0.7895

*** Indicates a p value < 0.001.

Reproductive success

In one of five populations, female reproductive success was determined by parentage analysis using multiple microsatellite loci to gain more knowledge on individual founder reproductive success for another study. Microsatellite data were converted to population

levels readouts and combined with the mitochondrial and Y-chromosomal data for analysis. Between six and 11 autosomal microsatellite loci were amplified, scored and analyzed in a stepwise fashion. Loci used were: d1mit251, d1mit449, d3mit22, d3mit312, d3mit333, d6mit138, d9mit232, d9mit251, d12mit277, d14mit128 and d19mit110. Primer sequences were obtained from the Mouse Genome Informatics website.[1] Fluorescently tagged primers were used in PCR reactions (CY-5 or CY-3). Tagged PCR products were run on 14" x 17", 6.25% denaturing acrylamide gels at 40 W for three to seven hours (locus dependent). Gels were imaged on a Typhoon Scanner 8600 and ImageQuant software (Amersham Biosciences, Piscataway, NJ).

Parentage was assigned by using Cervus 3.0, [2] a program that uses a likelihood based statistical approach. Allele frequencies were calculated using the genotypes of all candidate mothers and fathers and all offspring within the population. Simulations were run 10,000 cycles with an error rate of 1% to derive a delta score. Assigned parents were accepted when the trio confidence of mother, father and offspring was 95%. With this rule, 86% (49/57) of the population was genotyped.

Statistical analyses

Wean weight of offspring was analyzed with a linear mixed model (LMM). This model assessed the effects of treatment, time and the interaction of time and treatment. The model intercept was set to litter one. Treatment, parity and their interaction were treated as fixed effects while cage, was modeled as a random effect with a random intercept generated for each. Sexes were analyzed separately. A normal distribution was assumed because weight data are continuous. Sample sizes used to assess weight differences

include: 21 daughters from seven rofecoxib-exposed breeding pairs, 75 daughters from 16 control breeding pairs and 26 sons from eight rofecoxib-exposed breeding pairs compared to 71 sons from 17 control breeding pairs.

Litter size data are discrete counts and therefore were analyzed with a generalized linear mixed model (GLMM) with a Poisson distribution and logarithmic link. This model assessed the effects of treatment, time and the interaction of time and treatment. The model intercept was set to litter one. Treatment, parity and their interaction were treated as fixed effects while cage, was modeled as a random effect with a random intercept generated for each. There were 21 rofecoxib-exposed litters and 44 control litters for a total of 65 observations from 30 cages this analysis.

Body weight in semi-natural enclosures was analyzed with a LMM. This model assessed the effects of treatment, time and the interaction of time and treatment on the 116 population founders (females = 72, males = 42). Sexes were analyzed separately. A normal distribution was assumed because weight data are continuous. Treatment, time and their interaction were modeled as fixed effects and individual and population were modeled as random effects with random slopes and intercepts generated for each. The intercept was set at week zero, as this was when founders were release into the enclosures and at which collected of weight data from OPAs began. Collection of founder weight data continued on surviving individuals at each pup sweep. There were a total of 337 female observations and 142 male observations collected throughout the experiment.

Reproductive outputs were in terms of total offspring and thus are discrete data. These data were analyzed with a GLMM with a Poisson distribution and logarithmic link. The model assessed the effects of treatment, time and the interaction of treatment and time on population-level reproduction. These effects were set as fixed effects in the model and population was set as a random effect with random intercept calculated for each. The intercept was set at week eight, as that was when the first collection period or pup sweep occurred. Male reproductive output for each treatment was measured five times over the course of the 28-week study in each of the five independent populations (except for one population that was only measured two times) for a total of 44 observations. Female reproductive output for each treatment was measured five times over the course of the 28-week study in four independent populations (except for one population that was only measured two times) for a total of 34 observations. Female reproductive output was analyzed in terms of total offspring and male reproductive success was analyzed in terms of male offspring.

To assess the probability of territorial ownership, a GLMM was used. As a territory can be defended or not, a binomial distribution was used with a logit link. There were six territories within a population and were either occupied by rofecoxib-exposed males, control males or unoccupied. A total of 104 observations were collected and analyzed throughout the study. The model assessed the effect of treatment, time and their interaction. These effects were set as fixed effects and population was set as a random effect with a random intercept generated for each. The model intercept was set to week three as that was when data existed for each population.

Survival of founders was analyzed by a multivariate Cox proportional hazard model. Impacts of treatment and population were examined in the model. In the male analyses, the interaction of treatment and population was also assessed. This was not possible for females because of low mortality (n = 2). Sexes were analyzed separately due to differences in mortality rates. Individuals that survived the length of the study or that were intentionally removed from the study were censored. A total of 74 females were analyzed; two events and 72 censorings; and 42 males were analyzed; 17 events and 25 censorings.

For all mixed models, several candidate models were fit to the data. These models varied in terms of random effects that estimated both intercept and/or slope. For each analysis, the model that explained some of the variance and with the lowest Akaike information criterion (AIC) score was selected and reported. Neither the significance of a fixed effects nor the magnitude of the significance varied between models.

REFERENCES

1) The Jackson Laboratory. [Last accessed 2016 February 10]. Available from: http://www.informatics.jax.org

2) Kalinowski, ST, Taper, ML, Marshall, TC. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol. 2007; 16:1099–1106.