



**HAL**  
open science

## Measuring fitness and inferring natural selection from long-term field studies: different measures lead to nuanced conclusions

Vincent Viblanc, Claire Saraux, Anouch Tamian, François Criscuolo, David Coltman, Shirley Raveh, Jan Murie, F. Stephen Dobson

### ► To cite this version:

Vincent Viblanc, Claire Saraux, Anouch Tamian, François Criscuolo, David Coltman, et al.. Measuring fitness and inferring natural selection from long-term field studies: different measures lead to nuanced conclusions. *Behavioral Ecology and Sociobiology*, 2022, 76 (79), 10.1007/s00265-022-03176-8 . hal-03663338

**HAL Id: hal-03663338**

**<https://hal.science/hal-03663338v1>**

Submitted on 10 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

1 This article is a contribution to the Topical Collection Measuring individual reproductive success in the wild -  
2 Guest Editors: Marco Festa-Bianchet, Janet Mann  
3 In: Behavioral Ecology and Sociobiology  
4

5 **Measuring fitness and inferring natural selection from long-term**  
6 **field studies: different measures lead to nuanced conclusions**

7 Vincent A Viblanc<sup>1\*</sup>, Claire Saraux<sup>1\*</sup>, Anouch Tamian<sup>1</sup>, François Criscuolo<sup>1</sup>, David W Coltman<sup>2</sup>,  
8 Shirley Raveh<sup>3</sup>, Jan O Murie<sup>4</sup> & F Stephen Dobson<sup>1,5</sup>  
9

10 <sup>1</sup>Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France

11 <sup>2</sup>Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

12 <sup>3</sup>Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical,  
13 Veterinary and Life Sciences, University of Glasgow, Glasgow, UK

14 <sup>4</sup>Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada

15 <sup>5</sup>Department of Biological Sciences, Auburn University, Auburn, AL, USA  
16

17 § Correspondence: [vincent.viblanc@iphc.cnrs.fr](mailto:vincent.viblanc@iphc.cnrs.fr)

18 \* These authors contributed equally to the study

## 19 ABSTRACT

20 Measuring individual reproductive success in the wild is often achieved by counting the  
21 number of descendants produced by individuals. In seeking to understand how reproductive  
22 success can inform us about natural selection, however, we are faced with a conundrum. In terms  
23 of timing, what is the most relevant measure for examining selection? We might count the number  
24 of offspring born, surviving to the termination of parental care, surviving to adulthood, or only  
25 those surviving to themselves reproduce. Clearly, only the latter are passing on genes and traits to  
26 future generations, but this estimate may not always be available. So, are different estimates of  
27 fitness consistent? Do they provide us with similar inferences of selection on phenotypic traits?  
28 We examined these questions on a 29-year long-term study of individually monitored male and  
29 female Columbian ground squirrels (*Urocitellus columbianus*). We used the long-term data to  
30 calculate male and female fitness based on an annual measure of adult survival and the yearly  
31 production of offspring counted at the stages of birth, weaning, and yearling age. We then  
32 decomposed fitness into its constitutive elements including (1) adult survival to the next spring,  
33 and (2) the yearly production of offspring counted at the stages previously mentioned. We then  
34 compared fitness metrics to evaluate if they provided similar or contrasting information in the wild.  
35 Next, we used those fitness metrics to test for selection on the date of emergence from annual  
36 hibernation, a phenotypic trait previously shown to be highly variable, heritable, and associated  
37 with reproduction. Finally, we partitioned selection on emergence date into additive episodes of  
38 selection by looking at how selection changed from reproduction measured at birth, weaning and  
39 when offspring reached yearling age. Overall, fitness metrics were well correlated, but correlations  
40 weakened the further offspring were counted from birth. We generally found directional selection  
41 for earlier emergence dates both in males and females. The strength of selection depended on

42 which fitness metric was used. Most of the selection gradient on emergence date was explained by  
43 offspring born, and the selection differential was stronger in males than females. We evaluate how  
44 the choice of fitness metrics in life-history studies may nuance our inferences about natural  
45 selection.

46 **Keywords:** fitness, life history, lifetime reproductive success, mating, natural selection, survival

## 47 **SIGNIFICANCE STATEMENT**

48 This study explores how our inferences about natural selection acting on organismal traits vary  
49 depending on our choice of fitness metrics. Focusing on the timing of emergence from hibernation  
50 in Columbian ground squirrels, we show that directional selection for earlier emergence dates  
51 occurs, but the strength of selection depends on whether fitness is evaluated from offspring counted  
52 at birth, at weaning or later in life. These results show that the choice of timing for fitness  
53 measurements may nuance inferences about natural selection in life-history studies.

54

## 55 **INTRODUCTION**

56 Understanding natural selection and the selective processes that shape the evolution of phenotypic  
57 traits in living organisms hinges upon our ability to measure and aptly capture the adaptive  
58 influence of traits with respect to specific environments (Endler 1986). Studies of natural selection  
59 can be broadly separated into those that model the evolutionary process and those that attempt to  
60 measure the action of natural selection in populations in nature. Models of natural selection seek  
61 to reveal how the process works through mathematical expressions of change from one generation  
62 to the next (Wright 1931; Crow and Kimura 1970; Orr 2009). These models often assume that  
63 changes in trait frequencies are the result of natural selection, without the occurrence of other  
64 processes such as migration, mutation, drift, or epigenetic modifications of gene expression.

65 In contrast, evolutionary change in real populations may actually reflect a combination of  
66 these different processes. In addition, empirical studies of natural selection, in action in real  
67 populations, generally measure change in trait frequencies from short-term measures of survival  
68 or reproduction. This is problematic for inferences on natural selection, since fitness advantages  
69 of particular trait values are often measured over a shorter time period than a complete generation

70 (Lande and Arnold 1983; Kruuk et al. 2002; Dobson et al. 2017). Such studies measure fitness as  
71 success at surviving a storm or at producing juvenile offspring, and then compare the fitness metric  
72 to trait values, rather than using success at producing reproducing offspring that themselves  
73 reproduce in the next generation (but see Boyce and Perrins 1987).

74 Fitness is the complex metric by which one typically measures the adaptive value of  
75 organismal traits in evolutionary ecology. Studies that refine the definition of fitness or seek  
76 alternative measures of fitness are legion (Fisher 1930; Hamilton 1964; Grafen 1982, 2015; Endler  
77 1986; Lucas et al. 1996; Queller 1996; Oli 2003; Qvarnström et al. 2006; van de Pol et al. 2006;  
78 Hunt and Hodgson 2010; Dobson et al. 2012; Zhang et al. 2015; Scranton et al. 2016; Harris et al.  
79 2017; Rubach et al. 2020; Levin and Grafen 2021). These studies usually propose some practical  
80 means of measuring the propensity of a given gene, group of genes, trait, or group of traits, to  
81 spread through populations over generations (Sæther and Engen 2015). In the present paper, we  
82 use a trait-based definition of fitness: the change in frequencies of alternative forms of an  
83 organismal trait over time in a population (Dobson and Viblanc 2019). This definition presents the  
84 advantage that trait variation integrates sources of both genetic and non-genetic (e.g., early  
85 environmental or maternal effects affecting trait expression) variation. It further has the advantage  
86 of practicality since it is the phenotypic traits (e.g., individual mass, size, age) that are measured  
87 in field studies, with no assumption made about their underlying genetic or non-genetic  
88 architecture (e.g., the genetic variance-covariance matrix; Lande 1979; Arnold et al. 2008).

89 Regardless of the metric used, measuring fitness requires understanding how traits are  
90 transmitted through the population over time. In terms of evolution, it is the fitness of a trait form  
91 relative to others that counts (Ayala and Campbell 1974). For sexual organisms, such as many  
92 animals and plants, transmission of traits is achieved through reproduction. Thus, most studies

93 evaluate fitness by measuring the propensity of organisms to transmit traits through the number of  
94 offspring they produce. For iteroparous organisms, this includes successive reproductions, so that  
95 adult survival and longevity, as well as age at first reproduction and reproductive senescence, are  
96 also important components of fitness (Brommer et al. 2002; Oli et al. 2002; Bouwhuis et al. 2012).  
97 However, when measuring fitness, the most important issues are what ultimately influences  
98 changes in trait frequencies, and how our conclusions about trait evolution are shaped by the nature  
99 of our fitness measures.

100         If adaptive traits are those that spread in the population relative to other trait forms, then  
101 surely offspring that die before reproducing themselves do not contribute to the fitness of a given  
102 trait. Yet, for studies in nature, our ability to monitor individuals over time is often limited: many  
103 animal species disperse at a young age (Greenwood 1980; Dobson 1982), and for many species  
104 age at maturity may be delayed (Cole 1954; Oli and Dobson 2003), so that the contribution of  
105 offspring to trait fitness through their own future reproduction is difficult to measure. Thus, as  
106 second-best proxies, studies of animals often rely on metrics such as clutch size, brood size, and  
107 litter size at birth or at offspring independence, to quantify the number of offspring produced that  
108 potentially can contribute to the next generation. Yet, because numerous offspring die at a young  
109 age, it is unclear how trait fitness is influenced by further selection or changed by environmental  
110 stochasticity as time goes by (Hadfield 2008). On one hand, trait fitness may be influenced by  
111 genetic correlations and selection on other traits than the one of interest. On the other hand, the  
112 further fitness is measured from the production of reproductive offspring, the more environmental  
113 variation may change the association between a phenotypic trait and realized fitness (the fitness of  
114 alternative traits when offspring begin reproduction).

115           The purpose of our study was to assess how the evaluation of fitness changed depending  
116 on the timing of measurement using a 29-year long-term data set of male and female Columbian  
117 ground squirrels (*Urocitellus columbianus*). Specifically, we were interested in understanding (1)  
118 how individual fitness varied depending on whether we considered the offspring born, weaned, at  
119 yearling age, and at age at first reproduction; and (2) whether the association of a phenotypic trait  
120 and fitness changed depending on the timing of the fitness measurement. To do this, we measured  
121 individual survival *and* reproduction, and estimated the number of gene copies that a breeding  
122 parent passes on to the next year, by adding the parent's survival from the current year to the next  
123 (1 for surviving, 0 for not) to half the number of offspring it produced. We then compared this  
124 measure to other individuals in the population to obtain an estimate of annual fitness (Qvarnström  
125 et al. 2006). We also estimated fitness from adult annual survival and reproduction separately, and  
126 here the number of offspring was not halved (both annual fitness and these latter measures were  
127 used, for example, by Dobson et al. 2017). Offspring production may be measured at different time  
128 points, such as birth, weaning, or survival to a later period. We evaluated how important the choice  
129 of when to measure offspring production was with an examination of repeatability (the intraclass  
130 correlation coefficient "ICC," used as a measure of consistency) of annual and reproductive fitness  
131 measures when using number of offspring measured at birth, weaning, and yearling age and  
132 investigated paired correlations between these.

133           We also evaluated how measuring fitness at different time periods, and using different  
134 metrics, affects inferences about natural selection. For illustrative purposes, we focused on a single  
135 trait, and examined the strength of selection on emergence date from hibernation, i.e., the date of  
136 spring emergence above ground from hibernation. For ground squirrels, this occurs close to the  
137 termination of torpor (Williams et al. 2014). We did not know the genetic variance-covariance



138 matrix for traits associated with emergence from hibernation, and thus could not differentiate  
139 selection among genetically correlated traits (Lande and Arnold 1983). However, adult female  
140 Columbian ground squirrels go through a single day of estrus and mating each year, about 3-5 days  
141 after emergence from hibernation (Murie and Harris 1982). The genetic correlation between  
142 emergence and estrus dates was  $0.98 \pm 0.01$  SE (Lane et al. 2011), indicating selection on one of  
143 these traits would undoubtedly influence selection on the other. Since our sample sizes were  
144 greatest for emergence date from hibernation, we chose this trait for analyses. Columbian ground  
145 squirrels are hibernating rodents, with a short and intense reproductive season lasting only a few  
146 months (Murie and Harris 1982; Dobson et al. 1992). The timing of emergence from hibernation  
147 is highly variable (Tamian et al. 2022), and previous studies have shown it to be both significantly  
148 heritable ( $h^2 = 0.22-0.34$ ; Lane et al. 2011), and negatively associated with annual fitness when  
149 measured using offspring at emergence from their first hibernation, at about one year old (for adult  
150 females; Lane et al. 2012), prerequisite conditions for responding to selection. However, nothing  
151 is known about whether the strength or direction of selection gradients differs when the fitness for  
152 this trait is measured at different time periods, nor whether selection gradients exist for males.  
153 Thus, emergence date appeared to be a good candidate trait to investigate its association with  
154 different estimates of fitness measured from different life stages.

155

## 156 **Materials and Methods**

### 157 *Study species and long-term population monitoring*

158 Columbian ground squirrels were studied from 1992 to 2020 in the Sheep River Provincial Park,  
159 Alberta, Canada (50° 38' 10.73" N; 114° 39' 56.52" W; 1524 m; 2.3 ha). Individuals were fitted  
160 with permanent numbered ear tags (#1-Monel metal tag; National Band and Tag Company,

161 Newport, KY) when weaned (or at first capture for immigrants). Thus, it was not possible to record  
162 data blind because our study involved focal animals in the field. In each year of the study, the  
163 entire population was trapped at spring emergence using  $13 \times 13 \times 40 \text{ cm}^3$  live-traps (Tomahawk  
164 Live Trap, Hazelhurst, WI, USA) baited with a small amount of peanut butter. Each individual  
165 was then dyed with a temporary unique dorsal mark (Clairol® Hydrience black hair dye N°52  
166 Black Pearl, Clairol Inc., New York, USA) for identification during field observations. We  
167 followed ground squirrels daily throughout the breeding season to assess breeding phenology and  
168 success. Females copulated with multiple males within 3-5 days following emergence from  
169 hibernation, typically during a single day of estrus (Murie and Harris 1982; Raveh et al. 2010).  
170 We determined female mating date through behavioral observations and by inspecting female  
171 genitalia (presence of a copulatory plug or plug material in abdominal fur, or sperm in vaginal  
172 smears; Murie and Harris 1982). Following mating events, we identified female single-entrance  
173 nest burrows during gestation by visual observations of females stocking them with dry grass  
174 (Murie et al. 1998), and marked them with colored and flagged metal pins (1 m in length).

175         Females in the wild gave birth to an average of three (one to seven) blind and hairless  
176 offspring in a specially constructed nest burrow, after some 24-days of gestation (Dobson and  
177 Murie 1987; Murie 1995). From 2000 to 2016, we caught pregnant females within 2-3 days of  
178 expected parturition, about 21-22 days after mating, and brought them to an on-site field laboratory  
179 (Hare and Murie 1992). Females were housed indoor in polycarbonate microvent rat cages (483 x  
180 267 x 200 mm; Allentown Caging Equipment Company, Allentown, NJ), supplied with wood  
181 shavings and newspaper as nesting material. Females were provided with apple, lettuce, and horse  
182 feed (Equisine sweet show horse ration; Unifeed, Okotoks, Alberta, Canada) *ad libitum*. After  
183 parturition, offspring were sexed and a small tissue biopsy was acquired for paternity analyses (see

184 below) by clipping a toenail bud as previously described by Hare and Murie (1992). We returned  
185 mothers and their offspring to flagged nest burrows, usually within a day of birth.

186         After a lactation period of approximately 27-28 days, offspring first emerged above ground  
187 around the time of weaning (Murie and Harris 1982). We trapped females and their entire litters  
188 the day the young first emerged. Mothers were determined by observation of the single lactating  
189 female that associated with the natal burrow from which young emerged. Juvenile ground squirrels  
190 hibernate within their colony of origin for the winter, and those juvenile males that immigrate  
191 typically do so towards the end of the subsequent spring. Thus, in each year, we were able to  
192 recapture all yearling males and females that survived their first hibernation.

193

#### 194 *Paternity analyses*

195 From 2001 to 2017, paternity was estimated following the methods of Raveh et al. (2010). Briefly,  
196 DNA was extracted from offspring tissue biopsies using DNeasy Tissue extraction kits (Qiagen,  
197 Venlo, The Netherlands). We amplified 13 microsatellites with polymerase chain reaction (PCR),  
198 using primer pairs previously developed for *U. columbianus* GS12, GS14, GS17, GS20, GS22,  
199 GS25, and GS26 (Stevens et al. 1997); *Marmota marmota* BIBL18 (Goossens et al. 1998); MS41  
200 and MS53; (Hanslik and Kruckenhauser 2000); and *Marmota caligata* 2g4, 2h6 (Kyle et al. 2004)  
201 and 2h4 (GenBank accession no. GQ294553) amplified polymorphic microsatellite loci. We used  
202 similar PCR conditions and cycling parameters as Kyle et al. (2004), but with an annealing  
203 temperature of 54 °C. We tested for deviation from Hardy–Weinberg equilibrium (HWE) at each  
204 locus within cohorts, and for linkage disequilibrium between pairs of loci within cohorts using  
205 exact tests.

206 Paternity assignment was done using CERVUS 3.0 (Marshall et al. 1998; Kalinowski et al.  
207 2007). Paternity was assigned with 95–99% trio confidence (assumed dam–sire–offspring  
208 relationship). Analyses were conducted for each year separately. The input parameters for the  
209 simulation step of CERVUS were 10 000 cycles, 70 candidate fathers, 90% of the population  
210 sampled and 1% genotyping error. Parental assignments were accepted when the offspring had no  
211 more than 2 mismatches with both parents.

212

### 213 *Annual survival, reproduction and fitness*

214 We calculated annual survival (*Surv*), as adult survival from the spring mating period to the time  
215 of emergence in the next spring (1 for surviving, 0 for not surviving), separately for males and  
216 females. Second, we quantified annual reproduction (*R*), as the number of offspring measured at  
217 birth, weaning, and yearling age, separately for males and females. All immature individuals (i.e.,  
218 females not observed mating or males with testes in abdominal position) were excluded from these  
219 calculations. However, individuals considered mature (females observed mating and males with  
220 testes in scrotal position) were all kept, even when their reproductive success was zero.

221 We calculated annual contributions to lifetime fitness, viz. annual fitness ( $\lambda_{an}$ ), following  
222 the method of Qvarnström et al. (2006), as applied to ground squirrels by Lane et al. (2011, 2012)  
223 and Dobson et al. (2016, 2020):

$$224 \lambda_{an} = Surv + \frac{1}{2} R$$

225 For any given individual, *R* is halved since only half of an individual's genetic contribution is  
226 passed-on to offspring.

227 The objective of our study was to determine whether fitness metrics were comparable when  
228 reproduction was measured at different time points. Thus, annual fitness was calculated based on

229  $R$  being measured as the number of offspring either produced at birth ( $\lambda_{an_b}$ ), or weaning ( $\lambda_{an_w}$ ),  
230 or surviving up to yearling age ( $\lambda_{an_y}$ ).

231

### 232 *Selection analyses*

233 We tested for directional, stabilizing or disruptive selection on emergence date using linear and  
234 quadratic regression on annual fitness  $\lambda_{an}$  (Lande and Arnold 1983). We further ran selection  
235 analyses decomposing  $\lambda_{an}$  into its constitutive elements including annual reproduction (fecundity  
236 selection) and annual survival (viability selection). Those three fitness metrics were calculated  
237 relative to the population in a given year by dividing them by the annual population mean for  
238 annual fitness, reproduction and survival, respectively, and for males and females separately  
239 (Lande and Arnold 1983). As selection operates among individuals, we first centered emergence  
240 dates per year by subtracting from individual emergence dates the mean emergence date of the  
241 population in a given year, hence translating how early or late (number of days) an individual was  
242 compared to others in that same year.

243 We ran three sets of analyses where we estimated (1) directional selection gradients ( $\beta$ )  
244 using univariate linear models, followed by (2) quadratic selection gradients ( $\gamma$ ) from models that  
245 included both a linear and quadratic term (Lande and Arnold 1983; Arnold and Wade 1984a). The  
246 general form for these models is:

$$247 \quad \omega = \alpha + \beta z + \varepsilon \quad (1)$$

$$248 \quad \omega = \alpha + \beta' z + \frac{1}{2} \gamma z^2 + \varepsilon \quad (2)$$

249 where  $\omega$  is the considered measure of fitness (i.e., either  $\lambda_{an}$ ,  $R$  or  $Surv$ ),  $\alpha$  is the intercept,  $z$  the  
250 phenotypic trait (here centered emergence date), and  $\varepsilon$  an error term. Note that  $\gamma$  coefficients are  
251 reported as Lande and Arnold (1983)'s original formulation and do not require doubling to be

252 interpreted as stabilizing or disruptive selection gradients (Stinchcombe et al. 2008). We used  
253 linear mixed models to account for individual ID as a repeated random term, and individual age as  
254 a random factor because of known fitness differences that occur with age (Broussard et al. 2003;  
255 Raveh et al. 2010). Statistical significance of the selection gradients for viability was estimated  
256 with generalized linear-mixed models (GLMMs) with a binomial error structure (Garant et al.  
257 2007; Dobson et al. 2017). Directional selection is indicated by significant linear coefficients ( $\beta$ ),  
258 the sign of the coefficients indicating the direction of selection. Stabilizing or disruptive selection  
259 occurs when  $\gamma$  is significantly  $<0$  or  $>0$ , respectively (Lande and Arnold 1983; McGraw and  
260 Caswell 1996). We ran the analyses separately in males and females because the variance of fitness  
261 metrics was far greater in males than females (Jones et al. 2012).

262

### 263 *Partitioning selection into additive episodes*

264 Following Arnold and Wade (1984a, b), we partitioned selection on emergence date into additive  
265 episodes of selection by looking at how selection changed from reproduction measured at birth,  
266 weaning and when offspring reached yearling age. To do so, we separated reproduction in three  
267 distinct biological episodes: production of offspring at birth, offspring survival from birth to  
268 weaning, and offspring survival from weaning to yearling age. To better understand how each of  
269 these episodes contributed to total selection, we estimated selection differentials for each of them.  
270 Selection differentials represent the change in the mean value of a phenotypic character (here  
271 emergence date) produced by selection. Because reproductive success at yearling age can be  
272 decomposed in the product of the number of offspring born and the number surviving in the two  
273 following episodes, selection differentials corresponding to each of the episode should sum to the  
274 total selection differential. Selection differentials are thus presented for each episode in absolute

275 values (shift of the emergence date in number of days), but also as a % of contribution to total  
276 selection.

277

## 278 *Statistics*

279 All analyses were done in R v. 4.0.2 (R Core Team 2020). The 95% confidence intervals for male  
280 and female survival were obtained by parametric bootstrapping (10,000 simulations, 50% of the  
281 individuals resampled each time). The consistency of fitness metrics (repeatability, or intraclass  
282 correlation coefficient (ICC)) was calculated at different stages using the ‘rptR’ package in R  
283 (Stoffel et al. 2017),  $ICC = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_R}$ , where  $V_G$  is the among-individual variance,  $V_R$  is the  
284 within-individual (or residual) variance and  $V_P$  is the total variance in fitness. The ICC for annual  
285 fecundity (number of offspring) was estimated using a Generalized Linear Model with a Poisson  
286 distribution (Stoffel et al. 2017), as appropriate when working with count data (we added +1 to  
287 offspring numbers to avoid zero values). Correlation between fitness metrics measured at different  
288 time points are Spearman rank correlation tests, as not all distributions were Gaussian. Because  
289 information at birth and paternity analyses were not available in all years, sample sizes vary and  
290 are indicated as n the number of reproductive events,  $N_1$  the number of individuals and  $N_2$  the  
291 number of years in the results.

292

## 293 **RESULTS**

### 294 **Male and female annual survival**

295 Sexually mature (scrotal) male survival from one breeding season to the next was on average 65%  
296 ( $CI_{95} = [59\% - 71\%]$ ,  $n = 209$  reproductive events,  $N_1 = 78$ ,  $N_2 = 21$ ). In contrast, sexually mature

297 females (i.e., those that estrus cycled and mated) had survival rates of 73% (CI<sub>95</sub> = [70% – 76%],  
298 n = 732 reproductive events, N<sub>1</sub> = 223, N<sub>2</sub> = 28), on average.

299

### 300 **Male and female annual fecundity**

301 Adult male annual reproductive success varied from 0 to 29 offspring born (mean ± SD =  
302 6.2 ± 5.7, n = 154 reproductive events, N<sub>1</sub> = 61 fathers, N<sub>2</sub> = 16 years), 0 to 26 offspring weaned  
303 (4.2 ± 4.8, n = 190, N<sub>1</sub> = 73, N<sub>2</sub> = 19), and 0 to 16 offspring surviving to yearling age (1.7 ± 2.6, n  
304 = 190, N<sub>1</sub> = 73, N<sub>2</sub> = 19). Adult female annual reproductive success was lower, and varied from 0  
305 to 7 offspring born (3.0 ± 1.2, n = 454 reproductive events, N<sub>1</sub> = 167 mothers, N<sub>2</sub> = 21 years), 0 to  
306 7 offspring weaned (2.1 ± 1.5, n = 759, N<sub>1</sub> = 228, N<sub>2</sub> = 29), and 0 to 5 offspring surviving to  
307 yearling age (0.9 ± 1.1, n = 732, N<sub>1</sub> = 223, N<sub>2</sub> = 28). Overall, offspring survival from birth to  
308 weaning was 75% (CI<sub>95</sub> = [72% – 78%], N = 684) for male offspring and 77% (CI<sub>95</sub> = [72% –  
309 78%], N = 655) for female offspring; and from weaning to yearling age 43% (CI<sub>95</sub> = [40% – 47%],  
310 N = 775) for male offspring and 42% (CI<sub>95</sub> = [39% – 46%], N = 742) for female offspring. Annual  
311 fecundity was strongly consistent in males (ICC = 0.71, CI<sub>95</sub> = [0.60 – 0.78]), but less so in females  
312 (ICC = 0.08, CI<sub>95</sub> = [0.03 – 0.13]).

313

### 314 **Male and female annual fitness**

315 When calculated using all of the offspring born ( $\lambda_{anb}$ ), weaned ( $\lambda_{anw}$ ) or surviving to yearling  
316 age ( $\lambda_{any}$ ), annual fitness was strongly consistent in males (ICC = 0.67, CI<sub>95</sub> = [0.60 – 0.73]), but  
317 less so in females (ICC = 0.49, CI<sub>95</sub> = [0.45 – 0.54]). The correlation between annual fitness metrics  
318 generally weakened as the period between life stages increased, from 0.95 between birth and



319 weaning to 0.73 between birth and yearling age in males (Fig. 1A) and from 0.74 to 0.65 in females  
320 (Fig. 1B).

321

## 322 **Selection on emergence date**

### 323 *Annual fitness ( $\lambda_{an}$ ) selection*

324 In males, we found directional selection for earlier relative emergence date when annual  
325 fitness was calculated based on offspring born ( $\beta = -0.024$ ; Fig. 2A), and on offspring surviving to  
326 weaning ( $\beta = -0.018$ ; Fig. 2C), but not when calculated based on offspring that survived to yearling  
327 age ( $\beta = -0.007$ ; Fig. 2E) (Table 1). The strength of selection appeared to decrease from  $|\beta\lambda_{an_b}|$   
328  $> |\beta\lambda_{an_w}| > |\beta\lambda_{an_y}|$ . In females, we also found directional selection for earlier relative emergence  
329 date when annual fitness was calculated based on offspring born ( $\beta = -0.007$ ; Fig. 2B), and on  
330 offspring surviving to yearling age ( $\beta = -0.013$ ; Fig. 2F), but not when calculated based on  
331 offspring that survived to weaning ( $\beta = -0.004$ ; Fig. 2D). Note, that the directional selection  
332 coefficient for relative emergence date almost doubled for females between annual fitness  
333 calculated based on offspring born compared to when fitness was based on offspring surviving to  
334 yearling age ( $-0.013/-0.007 = 1.86$ ), but decreased in males ( $-0.007/-0.024 = 0.29$ ).

335 Both in adult males and adult females, we found no evidence of substantial disruptive or  
336 stabilizing selection on relative emergence date, regardless of whether annual fitness was measured  
337 based on offspring born, or on offspring surviving to weaning, or yearling age (Table 1).

338

### 339 *Viability selection*

340 In males, we found directional viability selection for later emergence date ( $\beta = +0.022$ ), but no  
341 evidence of non-linear viability selection on emergence date (Table 1). In females, there was no

342 evidence of significant directional viability selection, but suggestion of weak stabilizing viability  
343 selection on emergence date that approached significance (Table 1).

344

#### 345 *Fecundity selection*

346 In males, we found directional selection for earlier relative emergence date when fecundity was  
347 calculated from the number of offspring born ( $\beta = -0.032$ , Fig. 3A), weaned ( $\beta = -0.033$ , Fig. 3C),  
348 and surviving to yearling age ( $\beta = -0.033$ , Fig. 3E) (Table 1). Selection gradients were of similar  
349 magnitude regardless of the period considered. We found no evidence of disruptive or stabilizing  
350 selection regardless of whether fecundity was calculated from the number of offspring born,  
351 weaned, or surviving to yearling age (Table 1).

352 In females, there was mild directional selection for earlier relative emergence date when  
353 fecundity was calculated from the number of offspring born ( $\beta = -0.007$ , Fig. 2B), but not when  
354 calculated from offspring weaned ( $\beta = +0.002$ , Fig. 3C), or surviving to yearling age ( $\beta = -0.012$ )  
355 (Table 1). However, we found weak stabilizing selection for emergence date based on offspring  
356 surviving to yearling age that approached significance ( $\gamma = -0.005$ ; Table 1, Fig. 3E).

357

#### 358 *Partitioning selection into additive episodes*

359 In both males and females, selection on the number of offspring born provided a major selective  
360 advantage favoring earlier emergence date (Table 2). For males, selection on the number of  
361 offspring born accounted for 93% of the total selection differential and shifted the mean emergence  
362 date by 2.5 days, while selection acting between birth and weaning or between weaning and  
363 yearling age were minor, accounting for only 5 and 2% of the total selection differentials. For  
364 females, selection differentials were much lower than in males (0.4 days vs. 2.6 days shift due to

365 total selection). While selection on the number of offspring born, as well as from weaning to  
366 yearling, also shifted emergence dates earlier (by 0.6 and 0.3 days respectively), selection from  
367 birth to weaning counteracted this by shifting emergence dates later by 0.5 days.

368

## 369 **DISCUSSION**

370 Our results in the Columbian ground squirrels show that, overall, measures of fitness are  
371 relatively well correlated when fitness is measured from offspring counted at birth, at weaning,  
372 and at yearling age. In general, annual measures appeared quite strongly inter-correlated at  
373 different stages. Not surprisingly, the correlation between fitness measures changed, generally  
374 waning over the progression of life stages of the offspring. Overall, the strongest correlations were  
375 for fitness measured from birth and weaning offspring numbers, and associations became weaker,  
376 as time went by (from birth to yearling surviving offspring). These results might have been  
377 expected, since offspring mortality was low between birth and weaning, and much higher between  
378 weaning and yearling age.

379 Columbian ground squirrel offspring are raised over a short lactation period of 27 days  
380 (Murie and Harris 1982), during which they are secluded in specialized burrow systems and  
381 protected by their mothers (Murie and Harris 1988). As expected, mortality is relatively low during  
382 this period. After weaning, they only have a few weeks of foraging to grow and gain sufficient  
383 body mass to survive their first hibernation (Dobson 1992; Dobson et al. 1992). Offspring with  
384 insufficient body condition are unlikely to survive to the next spring (Murie and Boag 1984;  
385 Dobson et al. 1986). Thus, not surprisingly, overwinter mortality of juveniles is high (Dobson and  
386 Murie 1987; Zammuto 1987; Neuhaus and Pelletier 2001), and the correlation between fitness  
387 metrics wanes rapidly as the period between life stages of the offspring increases. Nonetheless, it

388 is noteworthy that for annual fitness measures these correlations remain relatively high ( $> 0.60$ ),  
389 so that measuring fitness from the number of offspring weaned in a given year already appears as  
390 a reasonable first approximation for estimating fitness. For instance, the correlation between  
391 annual fitness measured at weaning and when offspring were yearlings was 0.81 for adult males  
392 and 0.75 for adult females.

393 Sources of variation concerning the changes in correlation between measures between birth  
394 and yearling age of offspring remain to be determined. Such variations could include stochastic  
395 events that occur on an annual scale and accumulate over time, or other sources of biological  
396 variation known to affect early offspring survival such as changes in the social environments (e.g.,  
397 kin numbers; Viblanc et al. 2010; Barra et al. 2021). It is possible that selection on the offspring  
398 before they begin to reproduce at a later time could operate in a way that could weaken or  
399 strengthen associations of fitness estimates (whether or not an evolutionary response to selection  
400 occurs). Nonetheless, it is noteworthy that, overall, the correlations between fitness measured from  
401 offspring counted at birth, weaning, or yearling age were all high ( $> 0.65$ ) for annual fitness  
402 measures, suggesting that measuring fitness from offspring surviving to the first year is a good  
403 proxy to measuring actual fitness from offspring susceptible themselves to passing on traits (as for  
404 adult females in Lane et al. 2012; Dobson et al. 2016).

405 How did measuring fitness from an annual perspective at different time points affect our  
406 inferences on natural selection? To answer this question, we focused on the entire data set,  
407 calculating fitness from both male and female offspring, and regressed individual annual fitness  
408 on date of emergence from hibernation, a phenotypic trait known to be heritable (Lane et al. 2011),  
409 responsive to environmental fluctuation (Dobson 1988), and negatively associated with fitness  
410 (Lane et al. 2012). There appeared to be little or no stabilizing or disruptive selection (no quadratic

411 effects), but only directional selection, acting on emergence date. In general, individuals emerging  
412 earlier from hibernation had higher estimated fitness, and this was especially true for males.

413 In males, selection coefficients decreased when based on measurements from offspring  
414 counted at birth, weaning to yearling age. This likely occurred because of relatively strong  
415 mortality during the juvenile period, and supports the hypothesis that environmental stochasticity  
416 may dilute the association between a phenotypic trait and fitness due to passage of time since the  
417 environmental event (*viz.*, emergence from hibernation) increases. This is important, since  
418 although the directional selection coefficients measured at different time points are consistent in  
419 sign (*i.e.*, negative), the conclusions made would differ substantially. For instance, in males, based  
420 on fitness calculated from the number of offspring counted at weaning, significant linear ( $\beta$ )  
421 coefficients would give rise to a conclusion of directional selection (Lande and Arnold 1983)  
422 acting on emergence date, whereas non-significant coefficients based on fitness calculated from  
423 the number of offspring counted at yearling age would not (if relying on significance thresholds,  
424 but see below). Similarly, in females and based on statistical significance, one might conclude on  
425 directional selection on emergence when using offspring counted at birth, but not when using  
426 offspring counted at weaning or at a yearling age, despite the selection gradients being strongest  
427 when using number of offspring that reach a yearling age (Table 1).

428 Decomposing annual fitness into its annual survival and reproductive components  
429 generally revealed consistent patterns. Interestingly, we found mild positive directional viability  
430 selection on emergence date for males, but no viability selection for females: males that emerged  
431 later had better annual survival. Positive selection on viability in males might occur because males  
432 emerge from hibernation early to establish mating territories (Manno and Dobson 2008), but early  
433 emergence has a survival cost (Turbill et al. 2011; Constant et al. 2020). For females, a weaker

434 pattern of selective advantage was revealed for earlier emergence from hibernation, but the pattern  
435 was only significant when the combined index of annual fitness was used. In general, for both  
436 sexes, fecundity selection on emergence date revealed overall similar patterns as annual fitness,  
437 except perhaps for females where mild stabilizing selection was found using offspring that survive  
438 to yearling age as the fecundity fitness estimate.

439         When fecundity selection was separated into distinct episodes of selection (Arnold and  
440 Wade 1984a, b), we found that selection at birth accounted for the vast majority of the selection  
441 differentials on emergence date both in males and females. For males dates of emergence from  
442 hibernation, selection based on numbers of young at birth was the strongest (by nearly 2.5 days,  
443 Table 2), and might have reflected sexual selection (primarily number of mates) and maternal  
444 effects during gestation. Selection during lactation, based on numbers of offspring at weaning,  
445 however, was weak. This is consistent with the males contributing virtually nothing to parental  
446 care in this species, and reproductive success being determined foremost by the number of mated  
447 females early on in the season (CS et al., unpublished data). Finally, ecological influences on male  
448 date of emergence from hibernation may have been best reflected by numbers of offspring that  
449 survived their first hibernation, but this influence was trivial. For females, selection on emergence  
450 from hibernation was relatively weak and only approached significance. Partitioning this pattern  
451 into episodes of selection did not appear to produce important insights, other than a slight fitness  
452 advantage of emerging earlier from hibernation (by less than half a day).

453         Taken together, our results over a 29-year long-term study of Columbian ground squirrels,  
454 showed that annual fitness measures, whereas generally closely related, may lead to nuanced  
455 conclusions on the strength (but not direction) of selection acting on heritable phenotypic traits.  
456 We documented an overall dilution of fitness associations, and in most cases a waning association

457 between fitness and emergence date, as time passed likely due to added stochastic processes  
458 affecting offspring survival. Importantly, focusing on the significance of associations between  
459 phenotypic traits and fitness led to contrasting conclusions depending on when fitness is measured.  
460 The values of the selection coefficients, however, were fairly consistent using different fitness  
461 estimates. Taking a step back from a traditional hypothesis-testing approach and focusing on the  
462 magnitude of effect sizes, as more and more frequently is recognized in evolutionary ecology  
463 (Yoccoz 1991; de Valpine 2014), likely provides more meaningful information on the strength and  
464 patterns of selection acting on phenotypic traits in living organisms.

465

#### 466 **ACKNOWLEDGMENTS**

467 We are grateful to J. Mann and M. Festa-Bianchet for inviting us to contribute to this topical  
468 collection on “Measuring individual reproductive success in the wild.” We sincerely thank M.  
469 Morrissey and an anonymous reviewer for thoughtful and helpful comments on a previous version  
470 of the paper. We are grateful to Alberta Parks, and Alberta Environment, Fish & Wildlife for  
471 granting us access to the study sites and support with the long-term research. The University of  
472 Calgary Biogeoscience Institute provided housing at the R. B. Miller field station during data  
473 collection in Sheep River Provincial Park (AB, Canada). We are grateful to E.A. Johnson and S.  
474 Vamosi (Directors), J. Mappan-Buchanan and A. Cunnings (Station Managers) and K. Ruckstuhl  
475 (Faculty Responsible) for providing us with field camp and laboratory facilities over the years, and  
476 to P. Neuhaus and K. Ruckstuhl for their continued support, help and discussions in the field. We  
477 are especially grateful to E.A. Johnson for his continued support on the ground squirrel long-term  
478 research throughout the years.

479

480 **DATA AVAILABILITY**

481 All data generated or analyzed during this study are included in this published article and its  
482 supplementary information files.

483

484 **COMPLIANCE WITH ETHICAL STANDARDS**

485 **Funding:** The research was funded by a Natural Sciences and Engineering Research Council of  
486 Canada grant to JOM, a USA National Science Foundation grant (DEB-0089473) to FSD, a  
487 CNRS Projet International de Coopération Scientifique grant (PICS-07143) and a research grant  
488 from the Fondation Fyssen to VAV, a National Science and Engineering Research Council of  
489 Canada grant to DWC (Discovery Grant RGPIN-2018-04354), and by a fellowship grant from  
490 the Institute of Advanced Studies of the University of Strasbourg to FSD and VAV. FSD thanks  
491 the Région Grand Est and the Eurométropole de Strasbourg for the award of a Gutenberg  
492 Excellence Chair during the time of writing.

493

494 **Conflict of interest:** The authors declare that they have no conflict of interest.

495

496 **Ethical approval:** All applicable international, national, and/or institutional guidelines for the  
497 care and use of animals were followed. All procedures carried out in the field and laboratory  
498 were approved by Auburn University Institutional Animal Care and Use Committee (IACUC)  
499 and by the University of Calgary. Authorization for conducting research and collecting samples  
500 in the Sheep River Provincial Park was obtained from Alberta Environment and Parks and  
501 Alberta Fish & Wildlife.

502



503 **REFERENCES**

- 504 Arnold SJ, Bürger R, Hohenlohe PA, Ajie BC, Jones AG (2008) Understanding the evolution  
505 and stability of the G-Matrix. *Evolution* 62:2451–2461. [https://doi.org/10.1111/j.1558-](https://doi.org/10.1111/j.1558-5646.2008.00472.x)  
506 [5646.2008.00472.x](https://doi.org/10.1111/j.1558-5646.2008.00472.x)
- 507 Arnold SJ, Wade MJ (1984a) On the measurement of natural and sexual selection: theory.  
508 *Evolution* 38:709-719. <https://doi.org/10.2307/2408383>
- 509 Arnold SJ, Wade MJ (1984b) On the measurement of natural and sexual selection: applications.  
510 *Evolution* 38:720-734. <https://doi.org/10.2307/2408384>
- 511 Ayala FJ, Campbell CA (1974) Frequency-dependent selection. *Annu Rev Ecol Syst* 5:115–138.  
512 <https://doi.org/10.1146/annurev.es.05.110174.000555>
- 513 Barra T, Viblanc VA, Saraux C, Murie JO, Dobson FS (2021) Parental investment in the  
514 Columbian ground squirrel: empirical tests of sex allocation models. *Ecology* 102: e03479.  
515 <https://doi.org/10.1002/ecy.3479>
- 516 Bouwhuis S, Choquet R, Sheldon BC, Verhulst S (2012) The forms and fitness cost of  
517 senescence: age-specific recapture, survival, reproduction, and reproductive value in a wild  
518 bird population. *Am Nat* 179:E15–E27. <https://doi.org/10.1086/663194>
- 519 Boyce MS, Perrins CM (1987) Optimizing great tit clutch size in a fluctuating environment.  
520 *Ecology* 68:142–153. <https://doi.org/10.2307/1938814>
- 521 Brommer JE, Merilä J, Kokko H (2002) Reproductive timing and individual fitness. *Ecol Lett*  
522 5:802–810. <https://doi.org/10.1046/j.1461-0248.2002.00369.x>
- 523 Broussard DR, Risch TS, Dobson FS, Murie JO (2003) Senescence and age-related reproduction  
524 of female Columbian ground squirrels. *J Anim Ecol* 72:212–219.  
525 <https://doi.org/10.1046/j.1365-2656.2003.00691.x>

526 Cole LC (1954) The population consequences of life history phenomena. *Q Rev Biol* 29:103–  
527 137. <https://doi.org/10.1086/400074>

528 Constant T, Giroud S, Viblanc VA, Tissier ML, Bergeron P, Dobson FS, Haborc C (2020)  
529 Integrating mortality risk and the adaptiveness of hibernation. *Front Physiol* 11:706.  
530 <https://doi.org/10.3389/fphys.2020.00706>

531 de Valpine P (2014) The common sense of P values. *Ecology* 95:617–621

532 Dobson FS (1982) Competition for mates and predominant juvenile male dispersal in mammals.  
533 *Anim Behav* 30:1183–1192. [https://doi.org/10.1016/S0003-3472\(82\)80209-1](https://doi.org/10.1016/S0003-3472(82)80209-1)

534 Dobson FS (1988) The limits of phenotypic plasticity in life histories of Columbian ground  
535 squirrels. In: Boyce M (ed) *Evolution of Life Histories of Mammals*. Yale University Press,  
536 New Haven, pp 193-210

537 Dobson FS (1992) Body mass, structural size, and life-history patterns of the columbian ground  
538 squirrel. *Am Nat* 140:109–25. <https://doi.org/10.1086/285405>

539 Dobson FS, Badry MJ, Geddes C (1992) Seasonal activity and body mass of Columbian ground  
540 squirrels. *Can J Zool* 70:1364–1368. <https://doi.org/10.1139/z92-192>

541 Dobson FS, Becker PH, Arnaud CM, Bouwhuis S, Charmantier A (2017) Plasticity results in  
542 delayed breeding in a long-distant migrant seabird. *Ecol Evol* 7:3100–3109.  
543 <https://doi.org/10.1002/ece3.2777>

544 Dobson FS, Lane JE, Low M, Murie JO (2016) Fitness implications of seasonal climate variation  
545 in Columbian ground squirrels. *Ecol Evol* 6:5614–5622. <https://doi.org/10.1002/ece3.2279>

546 Dobson FS, Murie JO (1987) Interpretation of intraspecific life history patterns: evidence from  
547 Columbian ground squirrels. *Am Nat* 129:382–397. <https://doi.org/10.1086/284643>

548 Dobson FS, Murie JO, Viblanc VA (2020) Fitness estimation for ecological studies: an

549 evaluation in Columbian ground squirrels. *Front Ecol Evol* 8:216.  
550 <https://doi.org/10.3389/fevo.2020.00216>

551 Dobson FS, Viblanc VA (2019) Fitness. In: Vonk J, Shackelford T (eds) *Encyclopedia of*  
552 *Animal Cognition and Behaviour*. Springer, Cham, pp 1-7. [https://doi.org/10.1007/978-3-](https://doi.org/10.1007/978-3-319-47829-6)  
553 [319-47829-6](https://doi.org/10.1007/978-3-319-47829-6)

554 Dobson FS, Zammuto RM, Murie JO (1986) A comparison of methods for studying life history  
555 in Columbian ground squirrels. *J Mammal* 67:154–158. <https://doi.org/10.2307/1381012>

556 Endler JA (1986) *Natural selection in the wild*. Princeton University Press, Princeton, NJ

557 Fisher RA (1930) *The genetical theory of natural selection*. Clarendon Press, Oxford

558 Garant D, Kruuk LEB, McCleery RH, Sheldon BC (2007) The effects of environmental  
559 heterogeneity on multivariate selection on reproductive traits in female great tits. *Evolution*  
560 61:1546–1559. <https://doi.org/10.1111/j.1558-5646.2007.00128.x>

561 Grafen A (1982) How not to measure inclusive fitness. *Nature* 298:425–426

562 Grafen A (2015) Biological fitness and the Price Equation in class-structured populations. *J*  
563 *Theor Biol* 373:62–72. <https://doi.org/10.1016/j.jtbi.2015.02.014>

564 Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Anim*  
565 *Behav* 28:1140–1162. [https://doi.org/10.1016/S0003-3472\(80\)80103-5](https://doi.org/10.1016/S0003-3472(80)80103-5)

566 Hadfield JD (2008) Estimating evolutionary parameters when viability selection is operating.  
567 *Proc R Soc Lond B* 275:723–734. <https://doi.org/10.1098/rspb.2007.1013>

568 Hamilton WD (1964) The genetical evolution of social behaviour, I & II. *J Theor Biol* 7:1–52.  
569 [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)

570 Hanslik S, Kruckenhauser L (2000) Microsatellite loci for two European sciurid species  
571 (*Marmota marmota*, *Spermophilus citellus*). *Mol Ecol* 9:2163–2165.

572 <https://doi.org/10.1046/J.1365-294X.2000.10535.X>

573 Hare JF, Murie JO (1992) Manipulation of litter size reveals no cost of reproduction in  
574 Columbian ground squirrels. *J Mammal* 73:449–454. <https://doi.org/10.2307/1382083>

575 Harris CM, Madliger CL, Love OP (2017) An evaluation of feather corticosterone as a  
576 biomarker of fitness and an ecologically relevant stressor during breeding in the wild.  
577 *Oecologia* 183:987–996. <https://doi.org/10.1007/s00442-017-3836-1>

578 Hunt J, Hodgson D (2010) What is fitness, and how do we measure it? In: Westneat DF, Fox  
579 CW (eds) *Evolutionary Behavioral Ecology*. Oxford University Press, New York, pp 46–70

580 Jones PH, Van Zant JL, Dobson FS (2012) Variation in reproductive success of male and female  
581 Columbian ground squirrels (*Urocitellus columbianus*). *Can J Zool* 90:736–743.  
582 <https://doi.org/10.1139/z2012-042>

583 Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program cervus  
584 accommodates genotyping error increases success in paternity assignment. *Mol Ecol*  
585 16:1099–1106. <https://doi.org/10.1111/J.1365-294X.2007.03089.X>

586 Kyle CJ, Karels TJ, Clark B, Strobeck C, Hik DS, Davis CS (2004) Isolation and  
587 characterization of microsatellite markers in hoary marmots (*Marmota caligata*). *Mol Ecol*  
588 Notes 4:749–751. <https://doi.org/10.1111/J.1471-8286.2004.00810.X>

589 Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain: body  
590 size allometry. *Evolution* 33:402–416. <https://doi.org/10.2307/2407630>

591 Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution*  
592 37:1210–1226. <https://doi.org/10.1111/j.1558-5646.1983.tb00236.x>

593 Lane JE, Kruuk LEB, Charmantier A, Murie JO, Coltman DW, Buoro M, Raveh S, Dobson FS  
594 (2011) A quantitative genetic analysis of hibernation emergence date in a wild population of

595 Columbian ground squirrels. *J Evol Biol* 24:1949–1959. <https://doi.org/10.1111/j.1420->  
596 9101.2011.02334.x

597 Lane JE, Kruuk LEB, Charmantier A, Murie JO, Dobson FS (2012) Delayed phenology and  
598 reduced fitness associated with climate change in a wild hibernator. *Nature* 489:554–557.  
599 <https://doi.org/10.1038/nature11335>

600 Levin SR, Grafen A (2021) Extending the range of additivity in using inclusive fitness. *Ecol*  
601 *Evol* 11:1970–1983. <https://doi.org/10.1002/ece3.6935>

602 Lucas JR, Creel S, Waser PM (1996) How to measure inclusive fitness, revisited. *Anim Behav*  
603 51:225–228

604 Manno TG, Dobson FS (2008) Why are male Columbian ground squirrels territorial? *Ethology*  
605 114:1049–1060. <https://doi.org/10.1111/j.1439-0310.2008.01556.x>

606 Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-  
607 based paternity inference in natural populations. *Mol Ecol* 7:639–655.  
608 <https://doi.org/10.1046/J.1365-294X.1998.00374.X>

609 McGraw JB, Caswell H (1996) Estimation of individual fitness from life-history data. *Am Nat*  
610 147:47–64. <https://doi.org/10.1086/285839>

611 Murie JO (1995) Mating behavior of Columbian ground squirrels. I. Multiple mating by females  
612 and multiple paternity. *Can J Zool* 73:1819–1826. <https://doi.org/10.1139/Z95-214>

613 Murie JO, Boag DA (1984) The relationship of body weight to overwinter survival in Columbian  
614 ground squirrels. *J Mammal* 65:688–690. <https://doi.org/10.2307/1380854>

615 Murie JO, Harris MA (1982) Annual variation of spring emergence and breeding in Columbian  
616 ground squirrels (*Spermophilus columbianus*). *J Mammal* 63:431–439.  
617 <https://doi.org/10.2307/1380440>

618 Murie JO, Harris MA (1988) Social interactions and dominance relationships between female  
619 and male Columbian ground squirrels. *Can J Zool* 66:1414–1420.  
620 <https://doi.org/10.1139/z88-207>

621 Murie JO, Stevens SD, Leopky B (1998) Survival of captive-born cross-fostered juvenile  
622 Columbian ground squirrels in the field. *J Mammal* 79:1152–1160.  
623 <https://doi.org/10.2307/1383006>

624 Neuhaus P, Pelletier N (2001) Mortality in relation to season, age, sex, and reproduction in  
625 Columbian ground squirrels (*Spermophilus columbianus*). *Can J Zool* 79:465–470.  
626 <https://doi.org/10.1139/z00-225>

627 Oli MK (2003) Hamilton goes empirical: estimation of inclusive fitness from life-history data.  
628 *Proc R Soc Lond B* 270:307–311. <https://doi.org/10.1098/rspb.2002.2227>

629 Oli MK, Dobson FS (2003) The relative importance of life-history variables to population  
630 growth rate in mammals: Cole’s prediction revisited. *Am Nat* 161:422–40.  
631 <https://doi.org/10.1086/367591>

632 Oli MK, Hepp GR, Kennamer RA (2002) Fitness consequences of delayed maturity in female  
633 wood ducks. *Evol Ecol Res* 4:563–576

634 Queller DC (1996) The measurement and meaning of inclusive fitness. *Anim Behav* 51:229–232

635 Qvarnström A, Brommer JE, Gustafsson L (2006) Testing the genetics underlying the co-  
636 evolution of mate choice and ornament in the wild. *Nature* 441:84–86.  
637 <https://doi.org/10.1038/nature04564>

638 R Core Team (2020) R: A language and environment for statistical computing. R Foundation for  
639 Statistical Computing, Vienna, Austria, <http://www.R-project.org>

640 Raveh S, Heg D, Dobson FS, Coltman DW, Gorrell JC, Balmer A, Neuhaus P (2010) Mating

641 order and reproductive success in male Columbian ground squirrels (*Urocitellus*  
642 *columbianus*). Behav Ecol 21:537–547. <https://doi.org/10.1093/beheco/arq004>

643 Rubach KK, Dobson FS, Zinner B, Murie JO, Viblanc VA (2020) Comparing fitness measures  
644 and the influence of age of first reproduction in Columbian ground squirrels. J Mammal  
645 101:1302–1312. <https://doi.org/10.1093/JMAMMAL/GYAA086>

646 Sæther BE, Engen S (2015) The concept of fitness in fluctuating environments. Trends Ecol Evol  
647 30:273–281. <https://doi.org/10.1016/J.TREE.2015.03.007>

648 Scranton K, Lummaa V, Stearns SC (2016) The importance of the timescale of the fitness metric  
649 for estimates of selection on phenotypic traits during a period of demographic change. Ecol  
650 Lett 19:854–861. <https://doi.org/10.1111/ELE.12619>

651 Stevens S, Coffin J, Strobeck C (1997) Microsatellite loci in Columbian ground squirrels  
652 *Spermophilus columbianus*. Mol Ecol 6:493–495

653 Stinchcombe JR, Agrawal AF, Hohenlohe PA, Arnold SJ, Blows MW (2008) Estimating  
654 nonlinear selection gradients using quadratic regression coefficients: double or nothing?  
655 Evolution 62:2435–2440. <https://doi.org/10.1111/j.1558-5646.2008.00449.x>

656 Stoffel MA, Nakagawa S, Schielzeth H (2017) rptR: repeatability estimation and variance  
657 decomposition by generalized linear mixed-effects models. Methods Ecol Evol 8:1639–  
658 1644. <https://doi.org/10.1111/2041-210X.12797>

659 Tamian A, Viblanc VA, Dobson FS et al (2022) Integrating microclimatic variation in  
660 phenological responses to climate change: a 28-year study in a hibernating mammal.  
661 Ecosphere (published online, doi: 10.1002/ecs2.4059)

662 Turbill C, Bieber C, Ruf T (2011) Hibernation is associated with increased survival and the  
663 evolution of slow life histories among mammals. Proc R Soc Lond B 278:3355–3363.

664 <https://doi.org/10.1098/rspb.2011.0190>

665 van de Pol M, Bruinzeel LW, Heg D, van der Jeugd HP, Verhulst S (2006) A silver spoon for a  
666 golden future: long-term effects of natal origin on fitness prospects of oystercatchers  
667 (*Haematopus ostralegus*). J Anim Ecol 75:616–626. [https://doi.org/10.1111/J.1365-](https://doi.org/10.1111/J.1365-2656.2006.01079.X)  
668 [2656.2006.01079.X](https://doi.org/10.1111/J.1365-2656.2006.01079.X)

669 Viblanc VA, Arnaud CM, Dobson FS, Murie JO (2010) Kin selection in Columbian ground  
670 squirrels (*Uroditellus columbianus*): littermate kin provide individual fitness benefits. Proc  
671 R Soc Lond B 277:989–994. <https://doi.org/10.1098/rspb.2009.1960>

672 Yoccoz N (1991) Use, overuse, and misuse of significance tests in evolutionary biology and  
673 ecology. Bull Ecol Soc Am 72:106–111. <https://doi.org/10.2307/20167258>

674 Zammuto RM (1987) Life histories of mammals: analyses among and within *Spermophilus*  
675 *columbianus* life tables. Ecology 68:1351–1363. <https://doi.org/10.2307/1939219>

676 Zhang H, Rebke M, Becker PH, Bouwhuis S (2015) Fitness prospects: effects of age, sex and  
677 recruitment age on reproductive value in a long-lived seabird. J Anim Ecol 84:199–207.  
678 <https://doi.org/10.1111/1365-2656.12259>

679

## 680 **FIGURE CAPTIONS**

681 **Fig. 1** Pairwise correlation plots for (A) male and (B) female annual fitness metrics calculated  
682 from offspring counted at birth ( $\lambda_{an\_b}$ ), weaning ( $\lambda_{an\_w}$ ), or surviving to yearling age ( $\lambda_{an\_y}$ ). The  
683 distribution of data is given on the diagonal. Significant Spearman correlation coefficients are  
684 given for \*\*\* $P < 0.001$

685



686 **Fig. 2** Selection on emergence date from regression of annual fitness( $\lambda_{an}$ ) on mean year-centered  
687 emergence dates in males (left panels) and females (right panels). Fitness was calculated for  
688 fecundity based on offspring counted at birth ( $\lambda_{an_b}$ ) (A-B), weaning ( $\lambda_{an_w}$ ) (C-D), or surviving  
689 to yearling age ( $\lambda_{an_y}$ ) (E-F). Significant regressions are indicated by black lines, and gray ribbon  
690 representing 95% confidence intervals

691

692 **Fig. 3** Fecundity selection on emergence date from regression of relative annual reproductive  
693 success on year-centered emergence dates in males (left panels) and females (right panels).  
694 Fecundity was calculated from offspring counted at birth (A-D), weaning (C-F), or surviving to  
695 yearling age (E-H). Significant regressions are indicated by black lines, and gray ribbons  
696 representing 95% confidence intervals

697

699 **Table 1** Linear ( $\beta$ ) and non-linear ( $\gamma$ ) selection gradients for viability selection (adult survival), fecundity  
700 selection (adult reproduction, number of offspring produced) and annual fitness selection ( $\lambda_{an}$ , see methods) on  
701 emergence date in Columbian ground squirrels. Coefficients estimated from LMMs are provided with their  
702 standard errors. Significant coefficients at  $P < 0.05$  are indicated by an asterisk. Significance of coefficients were  
703 obtained with gaussian error structure for fecundity and annual fitness, and binomial error structure for viability  
704

		MALES			FEMALES		
<b>VIABILITY</b>		<i>Coefficient <math>\pm</math> SE</i>	<i>z</i>	<i>P</i>	<i>Coefficient <math>\pm</math> SE</i>	<i>z</i>	<i>P</i>
<i>Linear</i>	$\beta$ (ED)	0.022 $\pm$ 0.009	2.650	0.008*	-0.009 $\pm$ 0.004	-1.507	0.132
	$\beta$ (ED)	0.022 $\pm$ 0.012	1.584	0.113	-0.016 $\pm$ 0.006	-1.857	0.063
<i>Non-linear</i>	$\gamma$ (ED <sup>2</sup> )	-0.000 $\pm$ 0.002	-0.410	0.682	-0.001 $\pm$ 0.001	-1.909	0.056
<b>FECUNDITY</b>		<i>Coefficient <math>\pm</math> SE</i>	<i>t</i>	<i>P</i>	<i>Coefficient <math>\pm</math> SE</i>	<i>t</i>	<i>P</i>
<b><i>Birth</i></b>							
<i>Linear</i>	$\beta$	-0.032 $\pm$ 0.009	-3.675	<0.001*	-0.007 $\pm$ 0.003	-1.932	0.054
	$\beta$	-0.027 $\pm$ 0.010	-2.593	0.010*	-0.008 $\pm$ 0.004	-2.243	0.025*
<i>Non-linear</i>	$\gamma$	0.001 $\pm$ 0.001	1.020	0.309	-0.001 $\pm$ 0.001	-1.806	0.072
<b><i>Weaning</i></b>							
<i>Linear</i>	$\beta$	-0.033 $\pm$ 0.010	-3.410	<0.001*	0.002 $\pm$ 0.005	0.344	0.731
	$\beta$	-0.027 $\pm$ 0.011	-2.368	0.019*	0.001 $\pm$ 0.005	0.114	0.909
<i>Non-linear</i>	$\gamma$	0.001 $\pm$ 0.001	0.962	0.337	-0.001 $\pm$ 0.001	-1.017	0.310
<b><i>Yearling</i></b>							
<i>Linear</i>	$\beta$	-0.033 $\pm$ 0.015	-2.180	0.031*	-0.012 $\pm$ 0.010	-1.119	0.264
	$\beta$	-0.024 $\pm$ 0.018	-1.337	0.183	-0.017 $\pm$ 0.011	-1.568	0.117
<i>Non-linear</i>	$\gamma$	0.002 $\pm$ 0.002	0.912	0.363	-0.005 $\pm$ 0.002	-2.101	0.036*
<b>ANNUAL FITNESS</b>		<i>Coefficient <math>\pm</math> SE</i>	<i>t</i>	<i>P</i>	<i>Coefficient <math>\pm</math> SE</i>	<i>t</i>	<i>P</i>
<b><i>Birth</i></b>							
<i>Linear</i>	$\beta$	-0.024 $\pm$ 0.007	-3.211	0.002*	-0.007 $\pm$ 0.003	-2.624	0.009*
	$\beta$	-0.017 $\pm$ 0.009	-2.022	0.045*	-0.008 $\pm$ 0.003	-2.819	0.005*
<i>Non-linear</i>	$\gamma$	0.002 $\pm$ 0.001	1.362	0.175	-0.001 $\pm$ 0.001	-1.602	0.110
<b><i>Weaning</i></b>							
<i>Linear</i>	$\beta$	-0.018 $\pm$ 0.007	-2.405	0.017*	-0.004 $\pm$ 0.003	-1.116	0.265
	$\beta$	-0.012 $\pm$ 0.009	-1.341	0.181	-0.005 $\pm$ 0.003	-1.376	0.169
<i>Non-linear</i>	$\gamma$	0.001 $\pm$ 0.001	1.166	0.245	-0.001 $\pm$ 0.001	-1.318	0.188
<b><i>Yearling</i></b>							
<i>Linear</i>	$\beta$	-0.007 $\pm$ 0.008	-0.858	0.392	-0.013 $\pm$ 0.004	-2.873	0.004*

---

<i>Non-linear</i>	$\beta$	-0.000±0.010	-0.048	0.962	-0.014±0.004	-3.188	0.001*
	$\gamma$	0.002±0.001	1.161	0.247	-0.002±0.001	-1.719	0.086

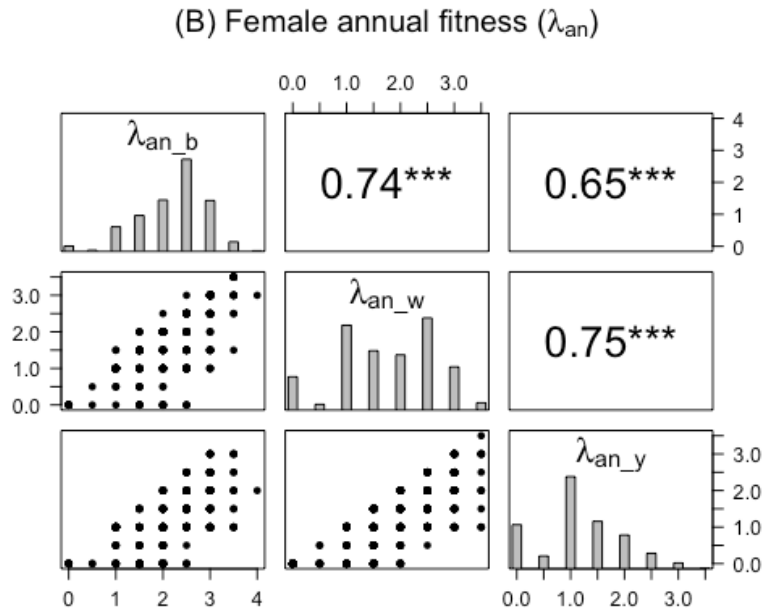
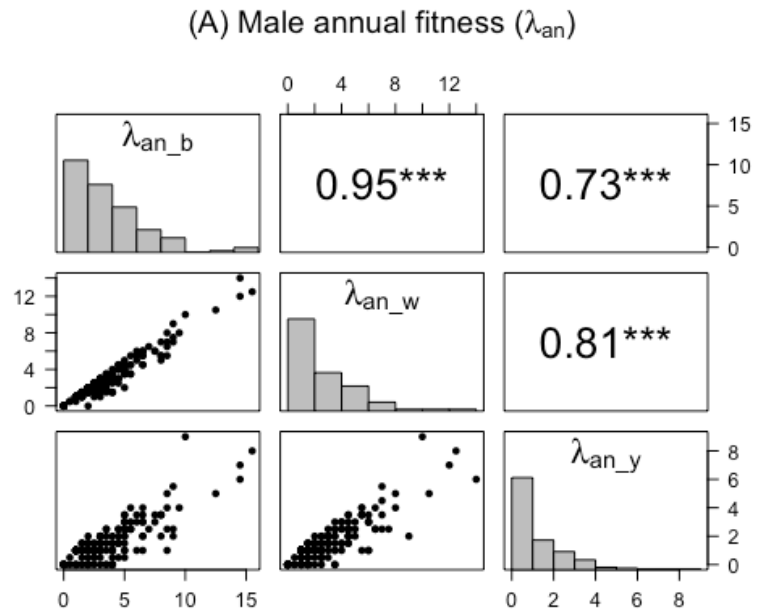
---

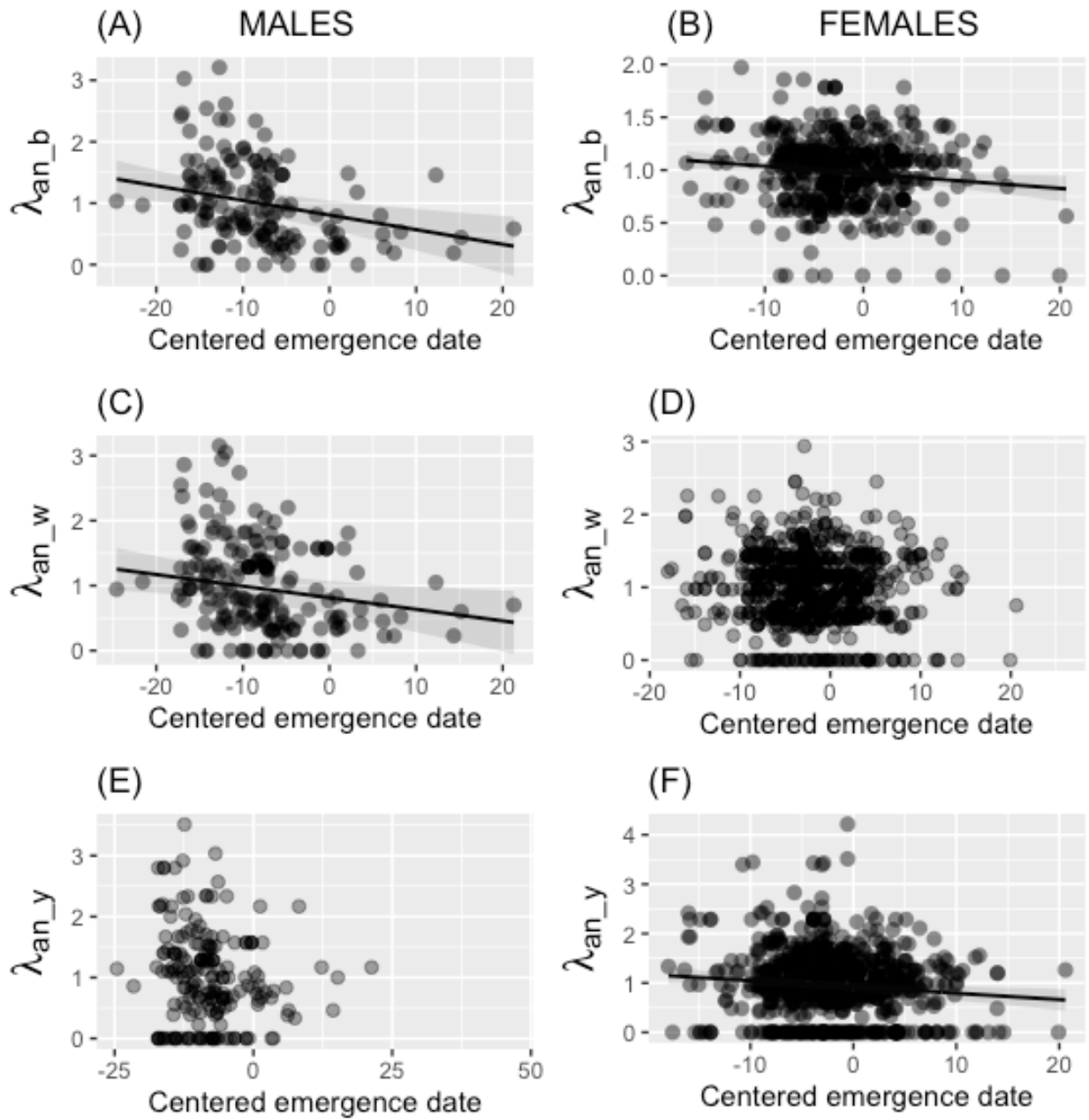
705  
706

707 **Table 2** Episodes of fecundity selection (Arnold and Wade 1984a, b). The trait under selection was date of  
 708 emergence from hibernation, an important phenological trait that is strongly associated with successful  
 709 reproduction and fitness (Lane et al. 2012). Selection differentials are given in days  
 710

	MALES		FEMALES	
	<i>Selection differential</i> $S_k$	<i>Contribution to total selection</i>	<i>Selection differential</i> $S_k$	<i>Contribution to total selection</i>
<b>SELECTION AT BIRTH</b>				
<i># Offspring born</i>	-2.435	93%	-0.593	144%
<b>SELECTION FROM BIRTH TO WEANING</b>				
<i>% Surviving offspring birth to weaning</i>	-0.142	5%	0.492	-120%
<b>SELECTION FROM WEANING TO YEARLING AGE</b>				
<i>% Surviving offspring weaning to yearling age</i>	-0.043	2%	-0.310	75%
<b>TOTAL SELECTION</b>	-2.619	100%	-0.411	100%

711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725





731

732

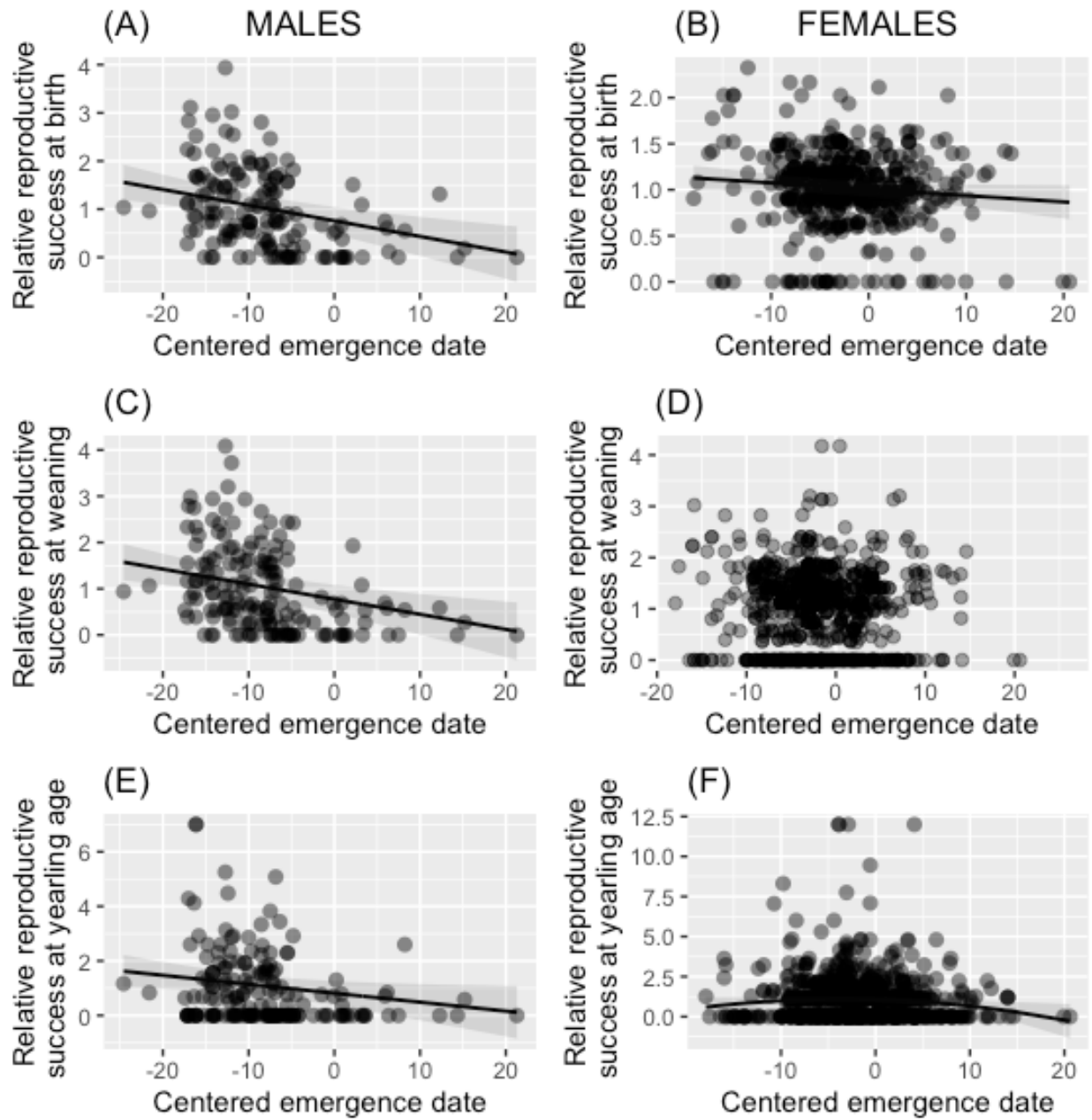
733

734 **Fig. 2**

735

736

737



738

739 **Fig. 3**

740

741