Supplementary information

Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation

Alina K. Niskanen[a,b,1], Anna M. Billing[a], Håkon Holand[a], Ingerid J. Hagen[a,c], Yimen G. Araya-Ajoy[a], Arild Husby[a,d], Bernt Rønning[a], Ane Marlene Myhre[a], Peter Sjolte Ranke[a], Thomas Kvalnes[a,d], Henrik Pärn[a], Thor Harald Ringsby[a], Sigbjørn Lien[e], Bernt-Erik Sæther[a], Stefanie Muff[f] and Henrik Jensen[a]

a Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway

b Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

c Department of Aquatic Biodiversity, Norwegian Institute for Nature Research, Trondheim, Norway

d Evolutionary biology, Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

e Centre for Integrative Genetics, Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, Ås, Norway

f Centre for Biodiversity Dynamics, Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

*C*ontact: Alina Niskanen (alina.niskanen@gmail.com or alina.niskanen@oulu.fi)

www.pnas.org/cgi/doi/10.1073/pnas.1909599117

Table of Contents

Supplementary Methods

Samples and genotyping

Since approximately 90% of the adult population on each island were sampled annually, the sample sizes (Table S2) were highly correlated with actual adult population sizes (Table S3). SNP genotypes were retrieved for 3219 individuals. SNP loci with Mono (313 SNPs) or Poly (185 274 SNPs) High Resolution score were kept for further quality control in PLINK (1) (version 1.9). Individuals with genotyping rate below 0.90 (68 individuals) were removed due to low sample quality. Ten pairs of individuals with identity by state (IBS) above 0.98 were removed as duplicates. Loci with minor allele frequency (MAF) below 0.01 and/or call rate below 0.95 were removed. In addition, loci with more than 10% of Mendelian errors based on island-specific microsatellite pedigrees (2–5) ("MS pedigree"), after excluding the parental links with more than 5 % Mendelian error rate, were removed from the dataset. The genotyping error rate per SNP was estimated to be 0.002 in our dataset based on concordance of the genotypes from 16 individuals that were genotyped twice for this purpose.

Pedigree construction

For constructing a metapopulation pedigree we used 605 SNPs that were selected in PLINK to have minor allele frequency above 0.4 within each island and low linkage disequilibrium (LD) using variance inflation factor (VIF) less than 1.1 ($\sim r^2 < 0.1$) within 100 SNP windows and step size of 10 SNPs. To obtain as informative pedigree as possible, on top of assigning parents among genotyped individuals, also dummy parents were assigned via sibship clustering. Default *sequoia (6)* settings were otherwise used in the pedigree construction, except for genotyping error rate, which was set to

0.002. To compare genomic relatedness with pairwise relatedness estimated based on SNP pedigree and previously built MS pedigree, genomic relatedness between all pairs of individuals was estimated using all 181 529 autosomal SNPs using GCTA (7).

Inbreeding analyses

Inbreeding estimates derived from pedigrees and genomic data have different resolutions: i) pedigree inbreeding estimates reflect the expected mean individual level of inbreeding, i.e. the proportion of homozygous loci within an individual's genome that is inherited from common ancestors (identical by descent, IBD) (8), whereas ii) genomic estimates quantify the realised level of inbreeding that varies around the expected mean inbreeding due to Mendelian segregation and recombination (9). Two genomic inbreeding coefficients, based on weighted average homozygosity over all loci (F_{GRM}) (7) and runs of homozygosity (F_{ROH}) (10) were estimated using 118 810 autosomal loci that had been pruned in PLINK for MAF > 0.05 and LD using VIF < 0.10 (i.e. $\sim r^2 <$ 0.90, with window size 50 SNPs, and step size 5 SNPs). F_{GRM} is based on correlation of the uniting gametes and is a low sampling variance estimate that gives more weight to the homozygosity of rare alleles. PLINK was used to estimate runs of homozygosity that are the base of F_{ROH}, where homozygous sequences of minimum length of 2Mbp were extracted using settings: --homozyg group --homozyg-density 10 --homozyg-gap 1000 --homozyg-kb 2000 --homozyg-snp 50 --homozyg-window-het 0 --homozyg-window-missing 5 --homozyg-window-snp 50. The mean ratio between genetic and physical distance (cM/Mbp), weighted by the SNP covered length of each chromosome (i.e. distance between the first and the last SNP within a chromosome), was 2.1 based on the house sparrow reference genome (11) and linkage map (12). Hence, 2Mbp long homozygous sequences would be caused by inbreeding that took place at most 12 generations ago (2Mbp * 2.1cM/Mbp = 4.2cM → 1M / (2 * 0.042M) = 12 generations (13)). F_{ROH} was calculated as the

proportion of SNP covered genome within the homozygous sequence blocks.

In variance partitioning for F_{GRM}, sex was included as a fixed factor, and island and island-year as random intercepts in a mixed-effect model with a Gaussian error distribution fitted with the R-package *R-INLA* (14). In addition, a model including the habitat type (farm vs. non-farm) as a fixed factor was fitted to estimate the contribution of habitat to variance in inbreeding.

Since g^2 estimation assumes that the heterozygosity of missing genotypes does not vary between loci, 140 799 autosomal loci pruned for call rate over 0.99 were used. One thousand bootstrap replicates were taken to estimate 95% confidence intervals of g^2 estimates. To estimate population differentiation, we estimated pairwise F_{ST} (15) using the R-package *hierfstat* (16) for each population pair over time including the years 2004-2013 when samples for all study populations were available. Using the same samples, we estimated the fixation index F_{IS} for each population over time. We used 5000 random autosomal SNPs to estimate *F*-statistics.

Phenotypic and life-history data used in the inbreeding depression analyses

Lifetime reproductive success (LRS) was estimated as the number of offspring produced that recruited to any of the eight study islands in the metapopulation. Only individuals that hatched earliest in year 1997 and latest in year 2009 were included in analyses of LRS to ensure an accurate estimate of recruit production over an individual's complete lifetime. From the birth cohort 2009, only 9.2% of the individuals were still alive, and thus capable of producing offspring, after the year of the last offspring cohort included in the study (birth year 2012). Because annual reproductive success (AR) is analysed on an annual basis, no restriction was done based on earliest hatch year of the focal individual, only individuals that hatched in 2012 were excluded; those individuals were themselves the last offspring cohort for which parentage was determined. In survival analyses, all natal dispersers that recruited on any of the study islands were included, but 52 adult dispersers

were excluded.

Phenotypic measurements taken by different fieldworkers were adjusted to the measurements taken by study's most senior fieldworker (T. H. Ringsby) (17). Only measures taken during the summer months (May-August) were used. Prior to analyses, both badge measurements were transformed into linear estimates (mm) by taking the square root of the surface area measures. Each individual had been measured a varying number of times during their lifetime and at different times of the year. To make the morphological measurements taken in different years and at different ages comparable, a general linear mixed-effects model was fitted separately for each sex and each trait (17) using R package *lme4 (18).* Global models included age, age^2 and month as fixed effects and individual's ring number as a random intercept and slope with age. Each fixed and random effect was included if significant according to likelihood ratio tests (Table S13). Individuals that were six years or older were grouped together due to the small number of old birds.

Inbreeding depression analyses

Inbreeding effects on survival models were fitted in JAGS (19) (Version 3.2.0) using a logistic link function. Since resighting probabilities (i.e. the probability to capture or observe an individual given that it was alive) may vary between islands and years in the metapopulation (20, 21), we included island and observation year as fixed factors in all models of resighting probability. Observation year was also included as a random factor in all survival models to account for any temporal variation in survival probability. To acquire a metapopulation level estimate of inbreeding effect on survival probability, we also accounted for spatial variation in survival by including island as a random intercept. To study the environmental effects on inbreeding depression in survival probability, we fitted models including the fixed intercepts of habitat type (farm vs. non-farm) and annual population size (continuous variable), and the interactions of habitat type by F_{GRM}, annual

population size by F_{GRM}, and sex by F_{GRM}. However, because none of the interactions between inbreeding and habitat type, annual population size or sex had a strong effect on inbreeding effects on survival (Table S6), these interactions and intercepts of habitat type and annual population size were dropped from the model that was fitted to examine the main metapopulation level effects (Table S5). In addition, we tested for spatial variation in inbreeding depression in survival by including island as a fixed factor and the interaction between island and F_{GRM}. Similarly, we tested for temporal variation in inbreeding depression by including year as a fixed factor and the interaction between year and F_{GRM}. For all survival models, we used three chains each with 120k iterations and a thinning rate of six; where the first 90k iterations were discarded ("burn-in"). Mixing and convergence of chains to a stationary distribution was evaluated by visual inspection of time-series plots of posterior values produced by JAGS and by the Brooks-Gelman-Rubin criterion (R-hat (22)). Parameter estimates (means) and their lower/upper 95% BCI limits were obtained from the respective stationary posterior distributions. We applied normally distributed vague priors with mean 0 and standard deviation 1000 for all parameters (23).

Models of inbreeding effect in reproductive success (LRS and AR) and morphology were fitted in INLA (24). The models included either standardised (to variance equal to 1) and mean-centered F_{GRM} or F_{ROH} (across all islands and years combined), and sex as fixed effects, random intercepts for adult island and island-year combination, as well as random slopes for the island by inbreeding and island-year by inbreeding interactions. Due to repeated measurements, the AR model included also individual's identity as a random factor, and centered age and age^2 as fixed covariates. Individuals that were six years or older were grouped together as six-year-olds. To examine the environmental effects on inbreeding depression in all studied traits, we fitted models including the fixed intercepts of habitat type and annual population size, and the interactions of habitat type by F_{GRM}, annual population size by F_{GRM}, and sex by F_{GRM}. However, because none of

the interactions between inbreeding and habitat type, annual population size or sex had a strong effect on inbreeding effects on any of the morphological or reproductive success traits (Table S6), we did not include these interactions or intercepts of habitat type or annual population size in the models that were fitted to examine the main metapopulation level effects (Table S5). A relatedness matrix based on the SNP pedigree was included in the models to account for any similarity between inbred individuals because of shared genetic variation that could lead to either over- or underestimation of inbreeding depression (25, 26). Gamma prior distributions G(0.1, 0.01) were given to all random effects, and the default N(0,10^3) priors were used for all fixed effects. A logarithmic link function was used in models fitted for LRS and AR.

The proportion of variance inbreeding (F_{GRM}) explained in each fitness component was estimated from the variance estimates derived from the inbreeding depression models described above. The proportion of total variance that inbreeding explained was estimated as the proportion it explained from the linear predictor.

The number of lethal equivalents (-2β) were estimated for fitness components (LRS, AR, and survival) using the island-specific slope estimates (β) of inbreeding depression models described above. Because we used standardised F_{ROH} in the models, we unstandardised the slope estimates of inbreeding effect on each island (i.e. β / $sd(F_{ROH})$) to get the lethal equivalents on the right scale. A modification to the inbreeding models was done in estimating lethal equivalents for survival, where we fitted a Poisson model with a logarithmic link, following the recommendations by Nietlisbach *et al*. (27), instead of a Bernoulli model with a logit link function. Note, however, that the credible intervals we acquired for lethal equivalents in survival are likely underestimated (27) due to using a Poisson model instead of a Bernoulli model for the binary survival data (28). The development of a method to estimate correct credible intervals using a joint survival model in combination with a capture-mark-recapture model as done here, would require further theoretical

considerations and is therefore, unfortunately, out of the scope of this paper.

Effect of inbreeding on morphology using multivariate models

Since all morphological measurements were taken from all individuals, the measurements are correlated. To evaluate if this non-independence affected the estimates of effects of inbreeding on morphological traits, we fitted multivariate models in MCMCglmm. Although INLA can be used to fit multivariate animal models (29), it is limited in the number of response variables (traits), and hyperparameter assignment is difficult with multiple traits. Thus, we used MCMCglmm (30) for multivariate analyses and INLA for univariate analyses due to its faster performance in that case. However, because there were instabilities in the algorithms for multivariate models that included all traits, only multivariate models including two or three traits simultaneously were fitted. The same individuals and morphological data were used in the multivariate analyses as in the univariate INLA models described in Materials and Methods and above. The models included three morphological traits as response variables, fixed intercepts of sex and mean-centered F_{GRM}, and the random intercepts of individual, island, hatch year and year nested within island ("island-year"). Traits with pairwise phenotypic Pearson's correlation coefficient higher than 0.20 were included in a same model so that each trait was included at least in one model. The model fitted for total and visible badge size differed from the other models as it did not include sex as fixed intercepts (only males have a badge) and only two response variables were fitted in the same model. A relatedness matrix based on the SNP pedigree was included to account for possible similarity between individuals because of shared genetic variation. Gaussian error structure was used for all traits. For all models, three chains of the MCMCglmm algorithm were run for 100k iterations with 20k burn-in and a thinning of 20. The mixing and convergence of the chains were inspected visually and by estimating a potential scale reduction factor ("gelman.diag"-command in R-package CODA (31)). Vague priors

were used for all parameters.

To compare the multivariate and univariate estimates of inbreeding effects on morphology, we also fitted univariate models in INLA for each morphological trait using similar model structure as in the multivariate MCMCglmm models: we included fixed intercepts of sex and mean-centered F_{GRM}, and the random intercepts of individual, island, hatch year and year nested within island ("island-year").

Statistical inference from random effects estimates using permutations

During each permutation, island identity and island-year identity were randomly re-allocated to different observations. This resulted in a new data set with the same mean, variance and level of replication as our observed data set; the only difference was that the island and island-year identities were reshuffled randomly across the data set. We then proceeded to fit the univariate animal models (detailed above, and in Table S5) to this new data set, and estimated, for each permutation, a posterior mean value for each variance component of interest. This procedure was repeated 1000 times to generate a 'null' distribution of posterior mean estimates. We then calculated the probability that the observed posterior mean value of a focal variance component (Table S7) was greater than any value expected from this permutation-based null distribution.

Supplementary Results

Metapopulation pedigree

We used the R package *sequoia* (6) to construct a metapopulation level pedigree ("SNP pedigree") that comprised 3556 individuals, of which 3116 were true genotyped individuals and 440 were dummy individuals (non-sampled parent individuals joining for example siblings together).

Maximum and mean pedigree depths were 14 and 4.6 generations, respectively (Fig. S9). The SNP pedigree was more complete and correct than a previously constructed island specific microsatellite based pedigree ("MS pedigree") for the same study metapopulation. The number of maternal and paternal links between genotyped individuals increased from 3597 in the MS pedigree to 4196 (17% increase) in the SNP pedigree (Table S12). When also dummy individuals were included, the number of parental links increased to 5571. As well as adding 652 new parental links compared to the MS pedigree, the new SNP pedigree also changed 360 parental links. The SNP pedigree links are most likely more correct, because the correlation between the genomic relatedness and relatedness estimated from SNP pedigree was higher (r = 0.82) than the correlation between the genomic relatedness and MS pedigree relatedness (r = 0.65, Fig. S10).

Comparison of the inbreeding coefficients

Choosing the most suitable genomic inbreeding estimate for a wild species is not a trivial task, since different estimates have been suggested as the most accurate for different species and questions (32, 33). We estimated three measures for genome-wide inbreeding and one for pedigree inbreeding; weighted average homozygosity over all loci (F_{GRM}), runs of homozygosity (F_{ROH}), genomewide heterozygosity, and inbreeding coefficient based on the SNP pedigree (F_{PED}). All inbreeding estimates correlated strongly (Fig. S2). The comparisons including F_{PED} were restricted to 1241 individuals with at least two full pedigree generations. The correlation was strongest between genomic inbreeding estimate F_{GRM} and genome-wide heterozygosity (r = -0.94), and weakest between F_{PED} and genome-wide heterozygosity (r = -0.74). Since inbreeding estimates correlated strongly and sample size halved when using F_{PED}, we used two genomic estimates, F_{GRM} and F_{ROH}, in inbreeding depression analyses. Furthermore, in the main text we focused on F_{GRM}, due to its statistically convenient (nearly) normal distribution.

Detecting inbreeding depression using F_{GRM} or F_{ROH}

Interestingly, all inbreeding effects were equally detected using either F_{GRM} or F_{ROH} as the inbreeding estimate (Table S5, Table S6, Fig. S6). The posterior mean effect sizes of F_{GRM} and F_{ROH} were similar, which is at least partly due to using standardised inbreeding estimates in the analyses. It also shows that both estimates of inbreeding can be used interchangeably in this study system.

Effect of inbreeding on morphology

The effect sizes of F_{GRM} to morphological traits estimated from multivariate models using MCMCglmm were similar to the results from the univariate models fitted using INLA without interaction terms between inbreeding and island or year nested within an island (Table S14). In these univariate INLA models, inbreeding had a negative effect on body mass, tarsus length and bill length (e.g. body mass β = -2.58, 95% BCI from -4.70 to -0.47; Table S14).

Supplementary Figures

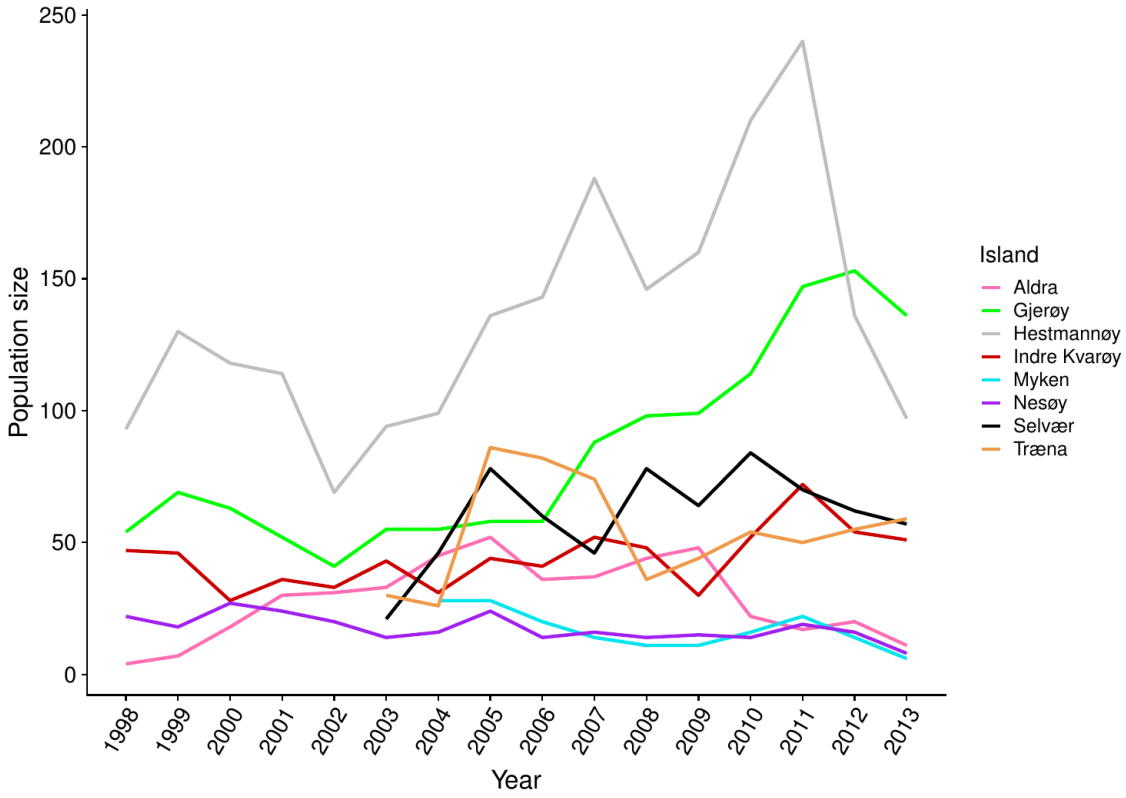

Fig. S1. Annual adult population size for each study island over the study period.

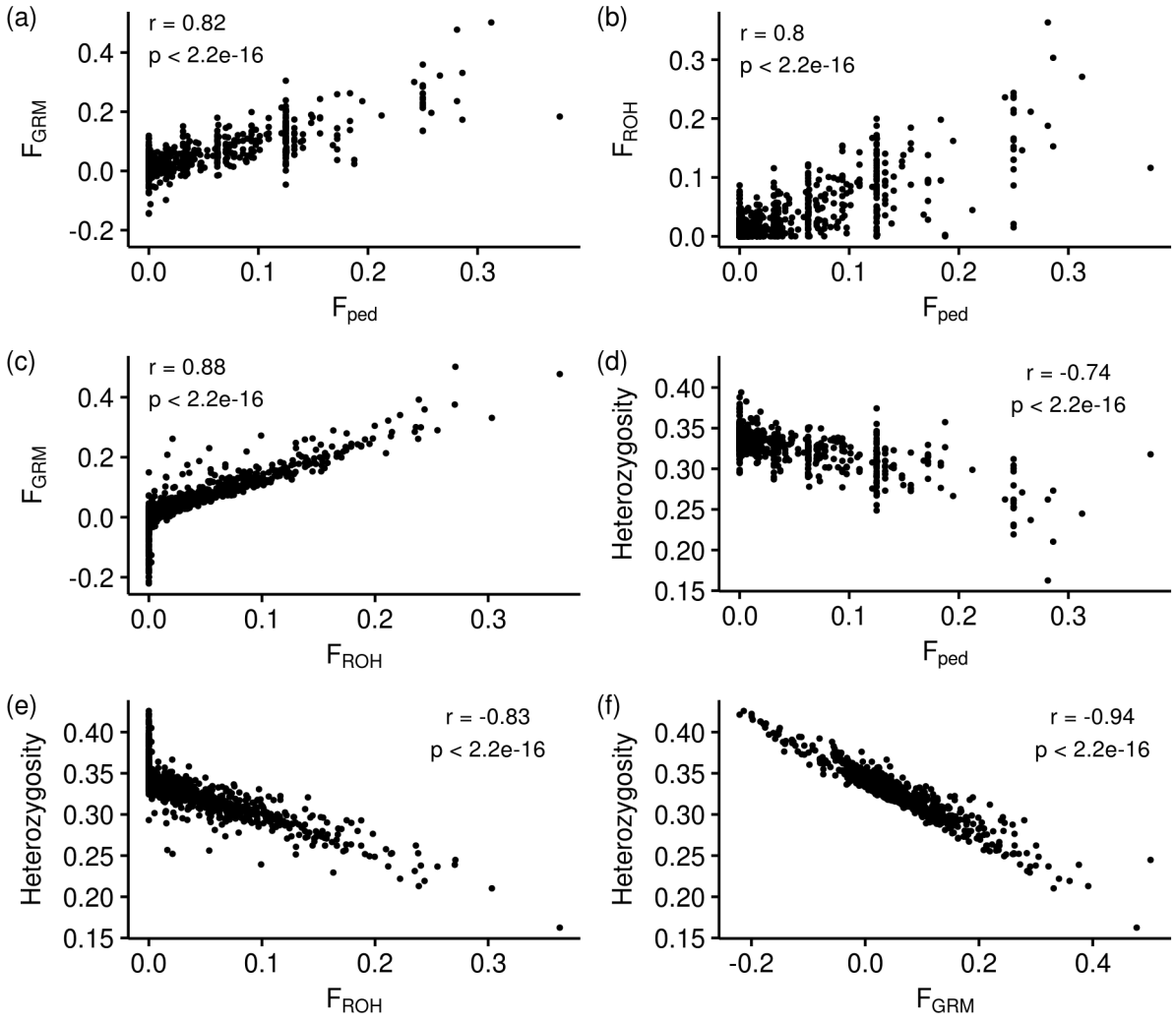

Fig. S2 Correlation between pedigree based and genomic inbreeding coefficients and genomewide heterozygosity. The comparisons including F_{PED} (**a**, **b**, and **d**) were restricted to 1241 individuals with at least two full pedigree generations, whereas other comparisons included all 3116 SNP-genotyped individuals.

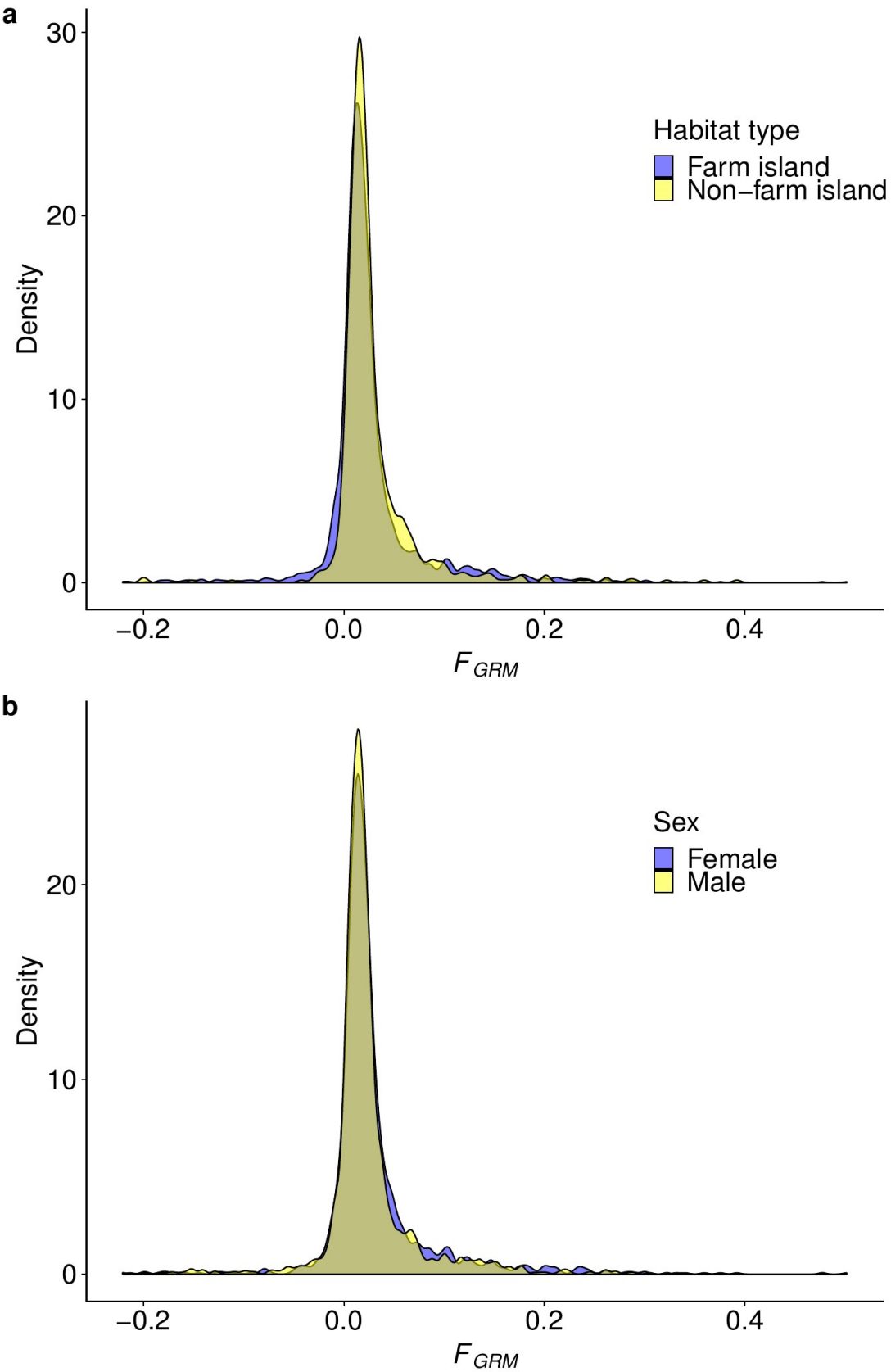

Fig. S3 Smoothed density distributions ("geom_density" command in R-package *ggplot2* (34)) of F_{GRM} for **a,** the two habitat types: farm and non-farm islands, and **b,** females and males.

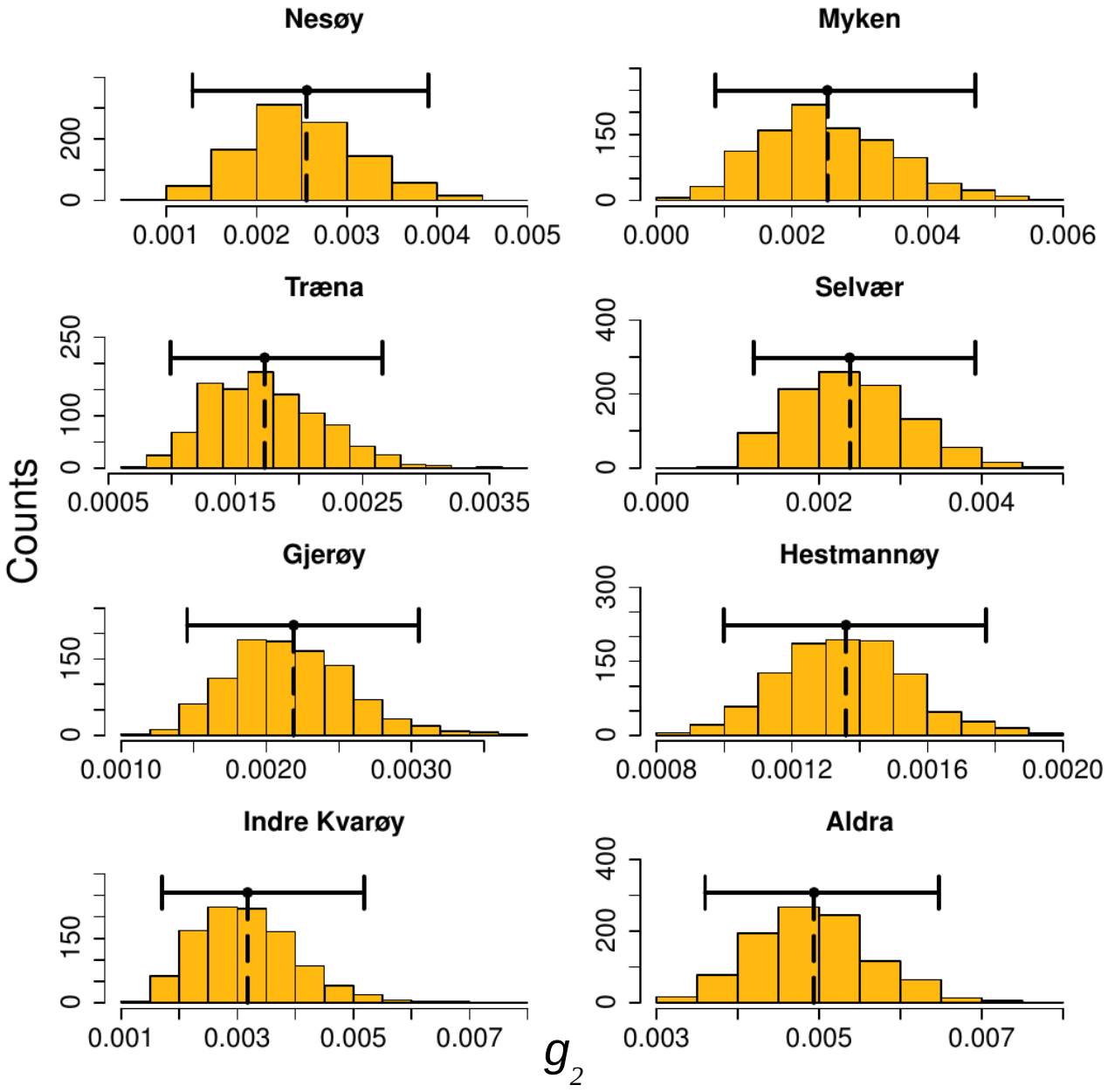

Fig. S4 Identity disequilibrium estimated as g_2 within each of the eight study islands. The bars show the distribution of estimates from 1000 bootstrap replicates, the vertical dashed line shows the mean, and the 95% confidence interval is indicated with a horizontal line.

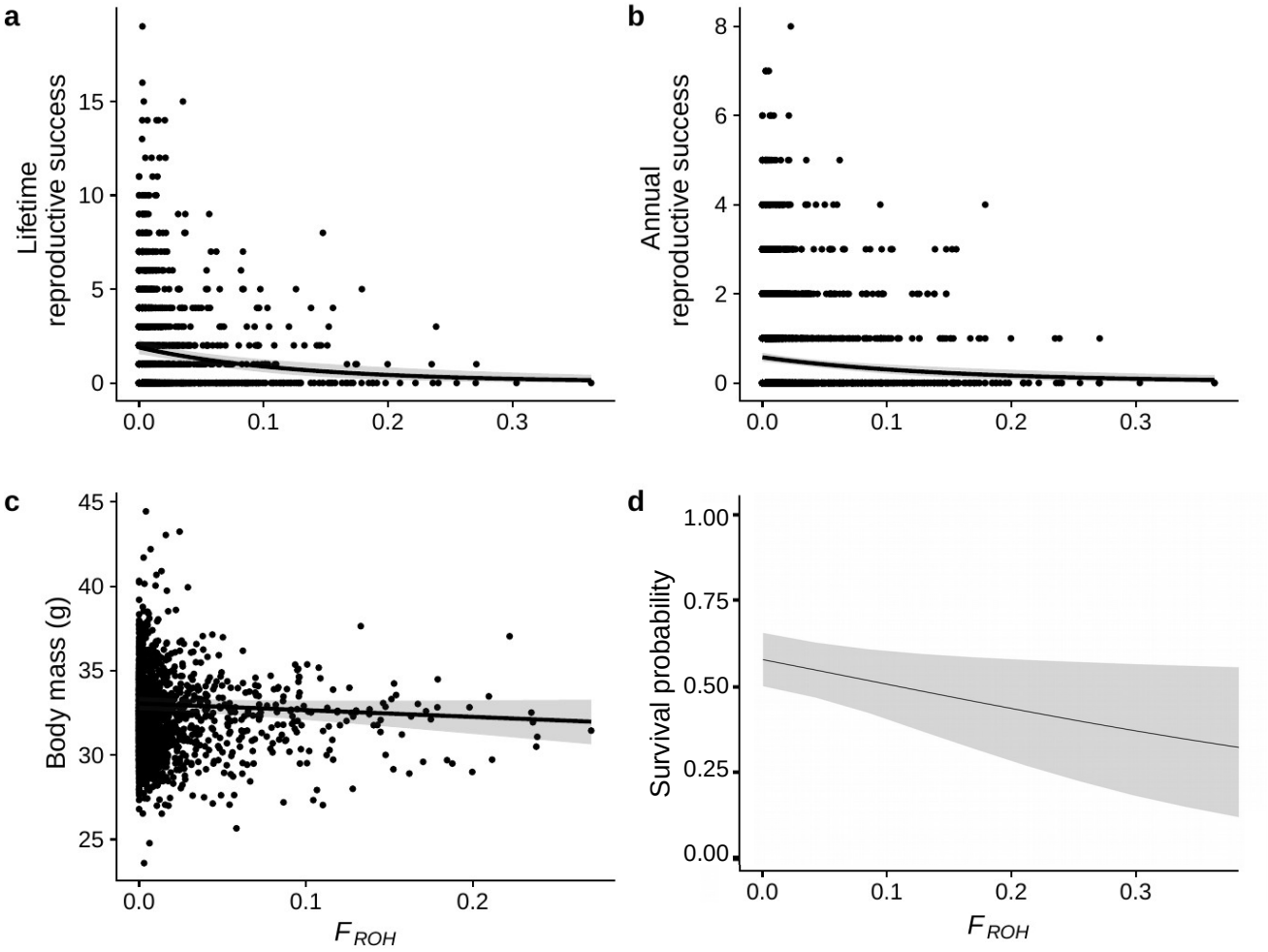

Fig. S5 The effect of inbreeding on reproductive success, survival and body mass. Lines are predicted mean effects of inbreeding (F_{ROH}) over the metapopulation system, shaded gray areas show 95% credible intervals, and individual observations are plotted as points (omitted from **d** for clarity). Lifetime reproductive success (**a**) and annual reproductive success (**b**) were estimated as the number of offspring recruiting the adult population that an individual produced during its lifetime or per year, respectively. Body mass (**c**) is based on all adult measurements and adjusted to trait value as one-year-old. (**d**) The relationship between survival probability and inbreeding. The predicted lines were produced using animal models fitted in INLA (**a**, **b,** and **c**), or a joint model (for survival) including capture-mark-recapture models in JAGS (**d**). The results of these models are presented in Table S5.

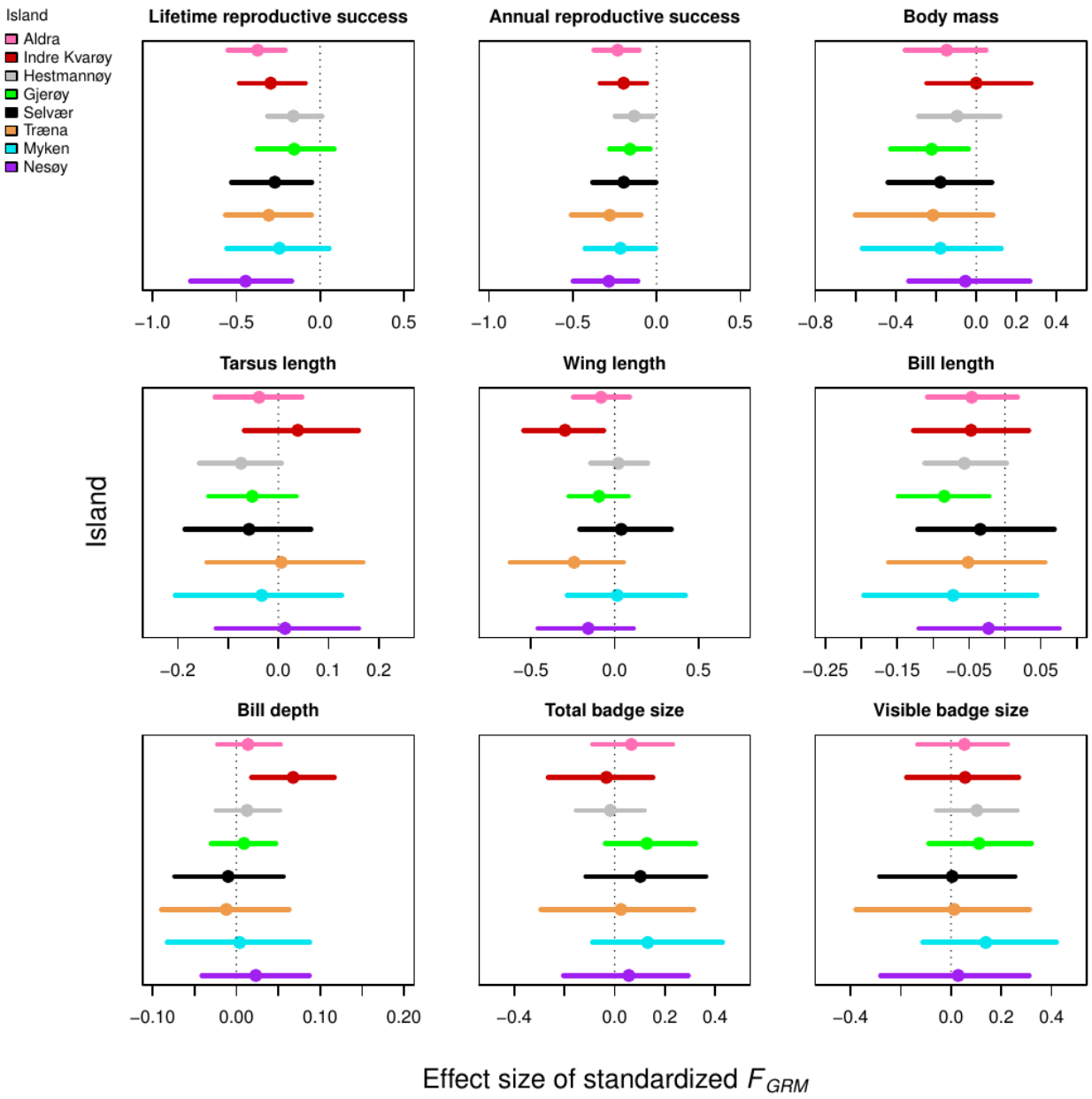

Fig. S6a Posterior mean estimates (β) and 95% Bayesian credible intervals of the effect of **a**, F_{GRM} and **b**, F_{ROH} on reproductive success and morphology on each study island. (Continues on the next page)

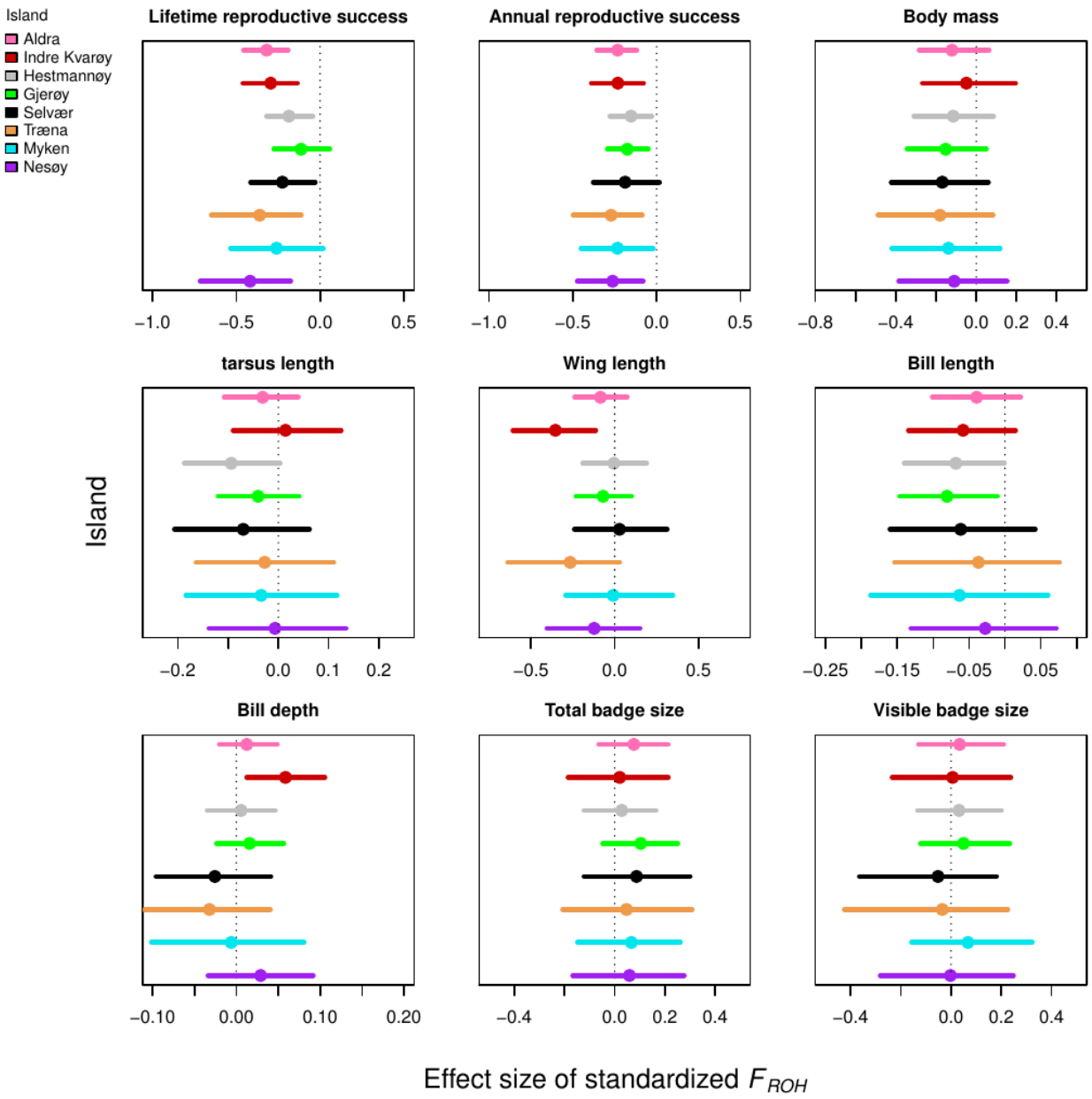

Fig. S6b (Continued from the previous page.) Posterior mean estimates (β) and 95% Bayesian credible intervals of the effect of **a**, F_{GRM} and **b**, F_{ROH} on reproductive success and morphology on each study island.

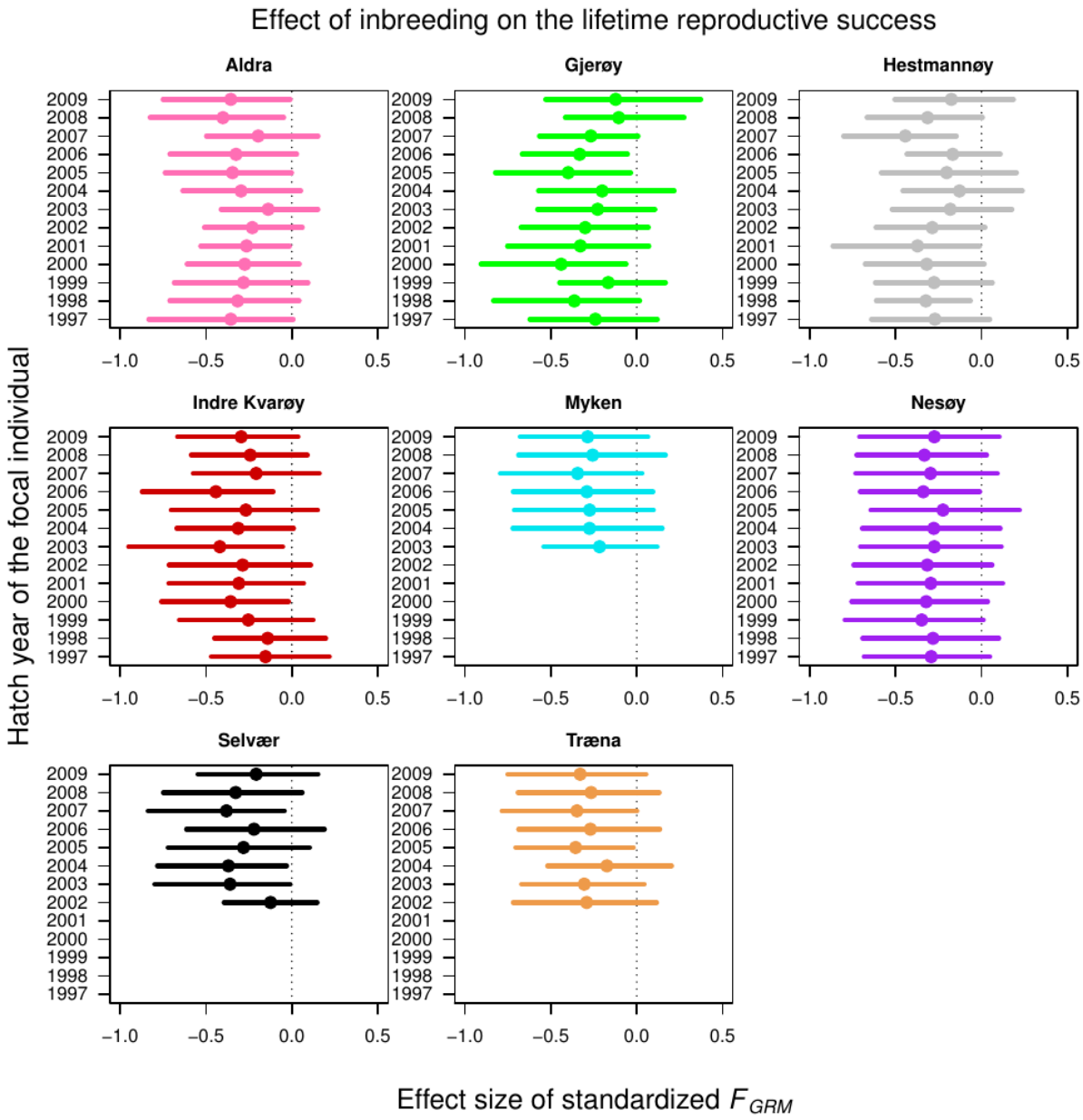

Fig. S7 Posterior mean estimates (β) and 95% Bayesian credible intervals of the annual effect of F_{GRM} on lifetime reproductive success on each study island.

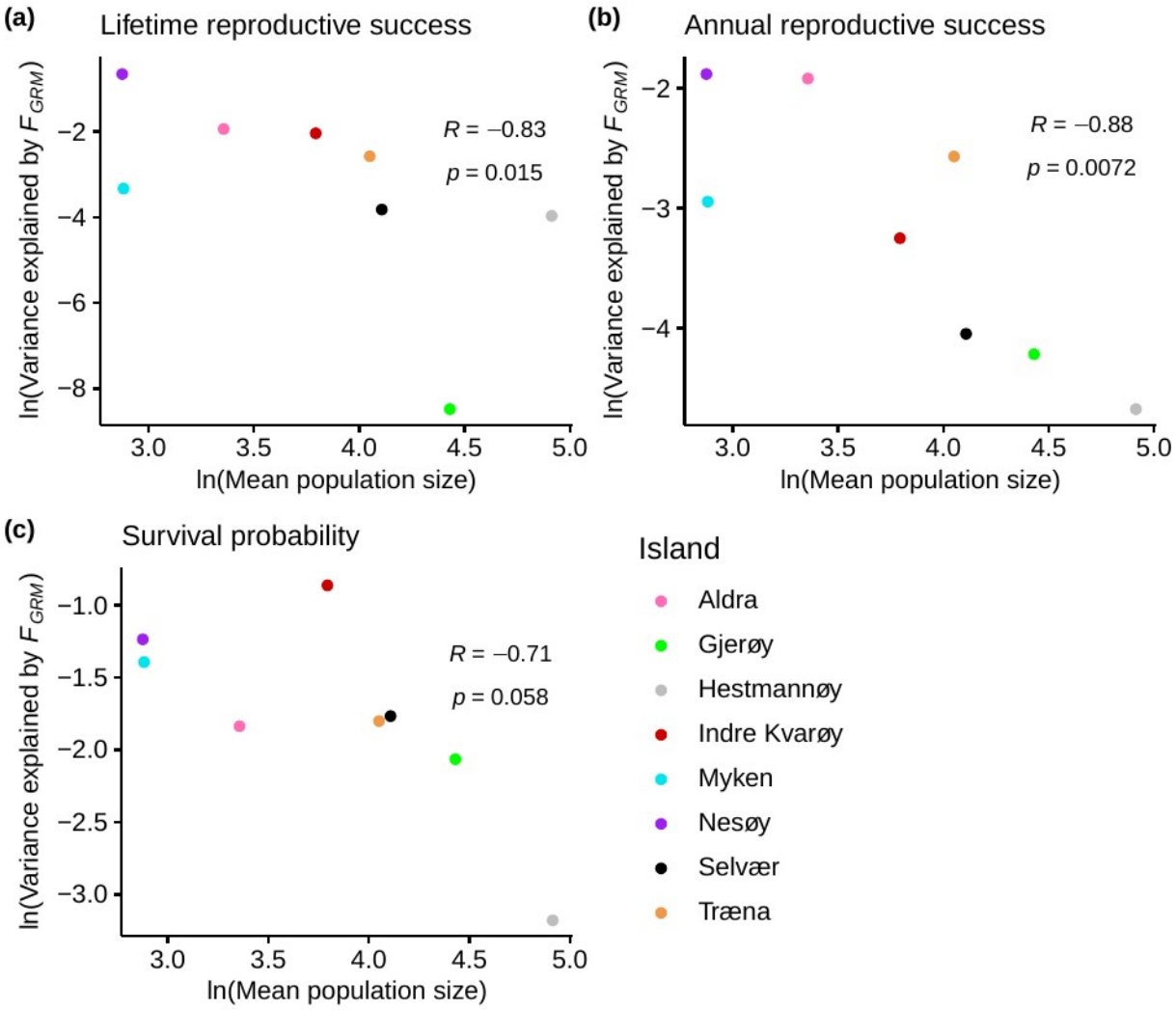

Fig. S8 Relationship between the ln-transformed mean population size and the ln-transformed proportion of explained variance by inbreeding on the linear predictor scale in each fitness component: **a**) lifetime reproductive success, **b**) annual reproductive success, and **c**) survival probability. Spearman's correlation coefficient (*R*) and its *p*-value are shown for each relationship.

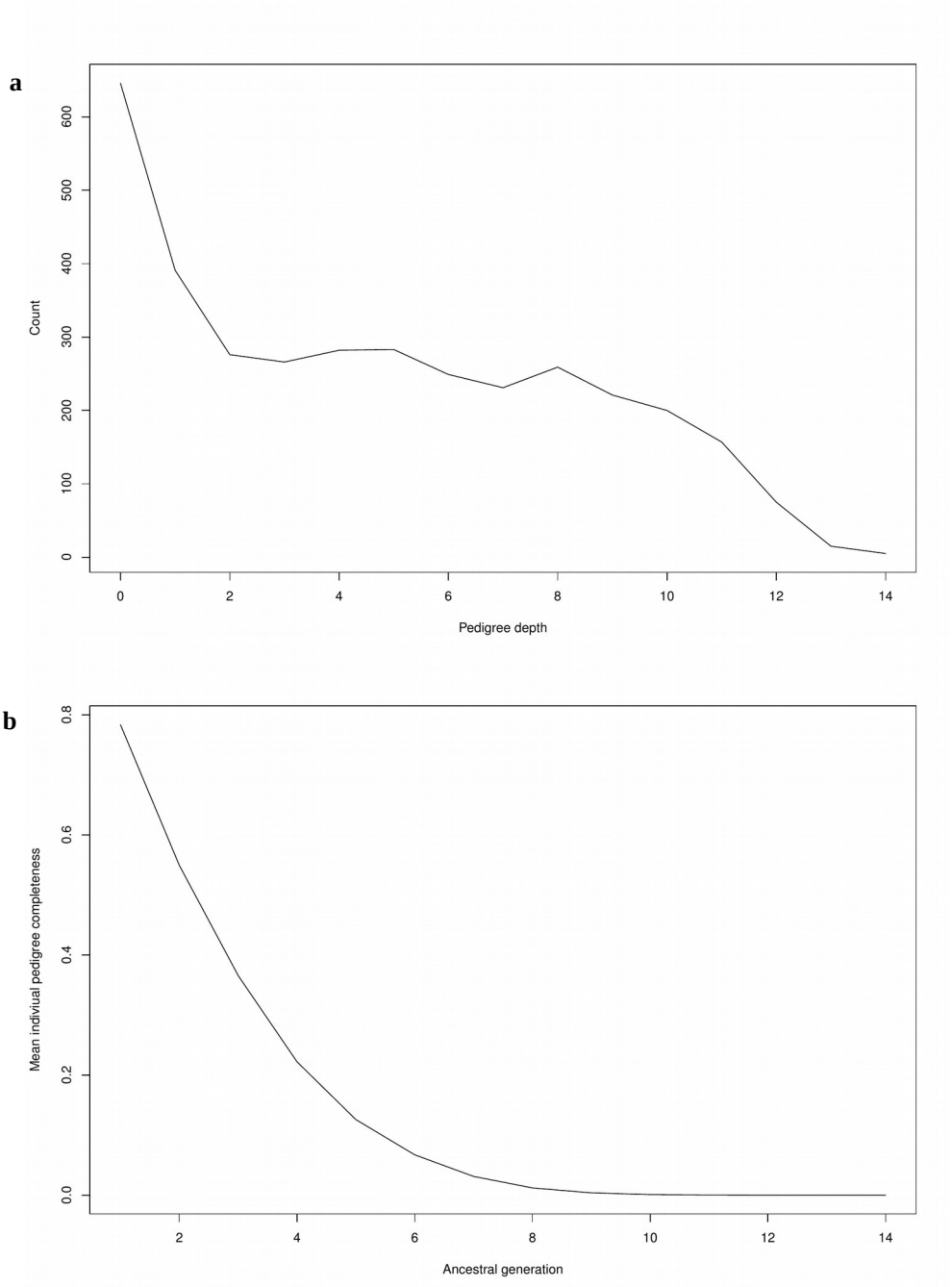

Fig. S9 Pedigree **a,** depth and **b,** completeness for the SNP pedigree of the Helgeland house sparrow metapopulation.

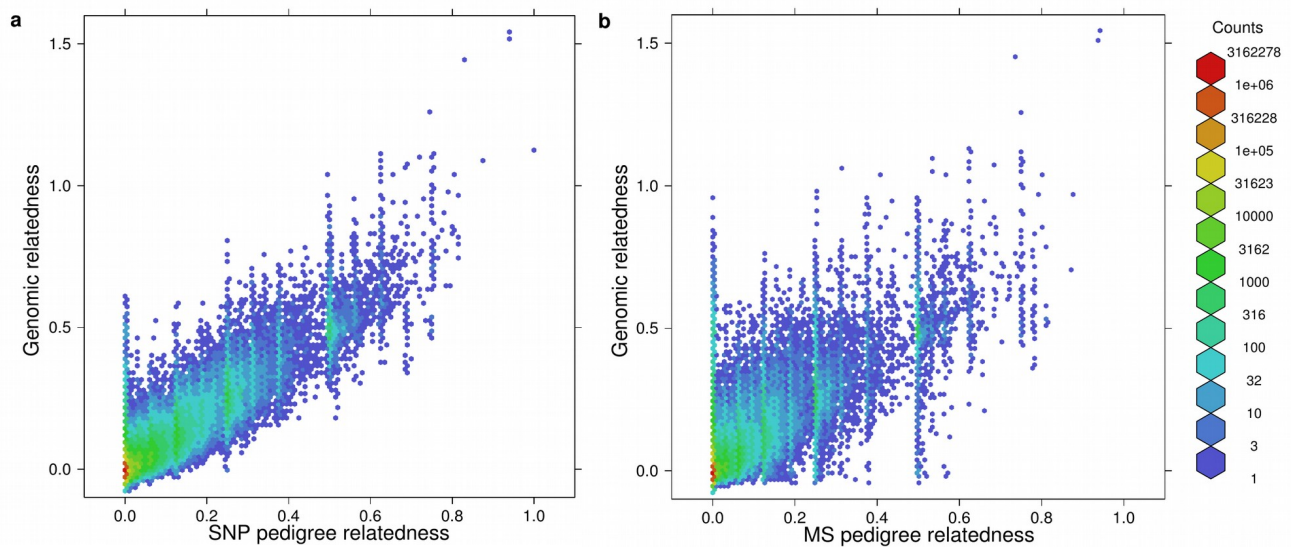

Fig. S10 Correlation between genomic relatedness estimated using all autosomal SNPs and **a,** SNP pedigree relatedness or **b**, MS pedigree relatedness.

Supplementary Tables

Table S1. Theoretical predictions regarding relationships between population size and population characteristics or evolutionary processes related to fitness, inbreeding and inbreeding depression (35–37). Populations are expected to be at mutation-drift-selection balance.*

	Small populations	**Large populations**
Average inbreeding	High	Low
Drift[§]	Strong	Weak
Heterozygosity	Low[#]	High
Variance in inbreeding within population	High	Low
Correlation between fitness and inbreeding within population	Negative	Negative
Inbreeding depression within population	Low	High
Average fitness	Low	High

*In natural populations, migration is expected to considerably decrease inbreeding. If the absolute number of immigrants is the same, the effect of immigration is larger in a small than in a large population.

[§]Effective population size (N_e) and the selection coefficient (s) affect the strength of selection in comparison to random drift such that when $4N_e s > 1$, selection can efficiently e.g. remove deleterious alleles. On the other hand, when $4N_e s < 1$, drift is the dominant evolutionary force (38).

[#]Low heterozygosity equals high homozygosity, which increases the visibility of recessive deleterious alleles to purifying selection.

Table S2. Information about the house sparrow samples used in this study. Sampling island, habitat type, years of collection, total number of samples i.e. unique adult individuals per island (N), and number of unique individuals included in different analyses (LRS = lifetime reproductive success, AR = annual reproductive success).

Island	**Habitat type**	**Years**	**Individuals**	**Survival**	**LRS**	**AR**	**Morphology**
Aldra	Farm	1998-2013	187	184	167	181	151
Gjerøy	Farm	1998-2013	591	546	370	529	411
Hestmannøy	Farm	1998-2013	1057	995	833	1014	707
Indre Kvarøy	Farm	1998-2013	359	317	241	344	222
Myken	Non-farm	2004-2013	82	77	45	78	30
Nesøy	Farm	1998-2013	131	110	92	122	83
Selvær	Non-farm	2003-2013	275	228	154	223	86
Træna	Non-farm	2003-2013	348	271	160	248	96
Other*		1998-2013	86	0	0	0	0
Total			3116	2728	2062	2739	1786

*These individuals have hatched on one of the 8 study islands, but have dispersed to another island in the metapopulation, and were only included in pedigree construction and comparison of inbreeding estimates.

Table S3. Mean and median of inbreeding estimates F_{GRM}, F_{ROH}, F_{PED} and genomewide heterozygosity with their interquartile ranges (IQRs) for each of the study islands. The table shows also mean adult population size (N) with standard deviation (Sd), the number of individuals per island used in producing the genomewide inbreeding estimates (n_G) and the number of individuals with at least two full ancestral generations used in estimating F_{PED} (n_{PED}).

				F_{GRM}		F_{ROH}		Heterozygosity		F_{PED}	
Island	**N (Sd)**	n_G	n_{PED}	**Mean**	**Median (IQR)**	**Mean**	**Median (IQR)**	**Mean**	**Median (IQR)**	**Mean**	**Median (IQR)**
Aldra	28.6 (15.1)	187	91	0.077	0.045 (0.022, 0.118)	0.053	0.028 (0.008, 0.085)	0.329	0.333 (0.319, 0.344)	0.113	0.109 (0.072, 0.125)
Gjerøy	83.8 (36.7)	591	300	0.022	0.014 (0.004, 0.027)	0.015	0.005 (0.000, 0.013)	0.334	0.337 (0.331, 0.342)	0.029	0.003 (0.000, 0.031)
Hestmannøy	135.8 (45.9)	1057	416	0.019	0.015 (0.007, 0.026)	0.012	0.005 (0.000, 0.011)	0.335	0.336 (0.332, 0.341)	0.012	0.000 (0.000, 0.008)
Indre Kvarøy	44.3 (11.2)	359	115	0.029	0.019 (0.009, 0.034)	0.018	0.006 (0.002, 0.019)	0.332	0.336 (0.329, 0.340)	0.033	0.000 (0.000, 0.037)
Myken	17.8 (7.3)	82	27	0.030	0.019 (0.005, 0.031)	0.019	0.007 (0.003, 0.015)	0.334	0.337 (0.333, 0.342)	0.021	0.000 (0.000, 0.009)
Nesøy	17.7 (4.9)	131	52	0.038	0.022 (0.013, 0.039)	0.021	0.006 (0.003, 0.019)	0.331	0.336 (0.329, 0.339)	0.028	0.000 (0.000, 0.002)
Selvær	60.6 (18.1)	275	81	0.030	0.019 (0.011, 0.033)	0.017	0.006 (0.002, 0.018)	0.334	0.337 (0.332, 0.341)	0.014	0.002 (0.000, 0.010)
Træna	57.3 (21.1)	348	125	0.030	0.020 (0.012, 0.034)	0.015	0.005 (0.000, 0.013)	0.335	0.337 (0.332, 0.341)	0.017	0.001 (0.000, 0.012)

Table S4. Variance partitioning for F_{GRM} including sex and habitat type as fixed factors, and spatial (island) and temporal (island-year; years nested within islands) random components. 95% Bayesian credible intervals (BCI) for fixed effects, and mean, mode and 95% BCI for random variances. The model was fitted using INLA. Temporal variance explained by island-year is the year nested within an island. Large effect size, i.e. 95% BCI did not overlap zero, in bold font.

	Mean [95% BCI]
Intercept	0.230 [-0.136, 0.601]
Sex, male	**-0.104 [-0.174, -0.035]**
Habitat type, non-farm	-0.130 [-0.735, 0.471]

Random variances	**Mean**	**Mode [95% BCI]**
Island	0.174	0.095 [0.048, 0.509]
Island-year	0.031	0.024 [0.012, 0.065]
Residual variance	0.902	0.900 [0.856, 0.952]

Table S5. Results of models estimating the effect of inbreeding on survival fitted in JAGS, and models estimating the effects of inbreeding on reproductive success and morphology fitted in INLA. Posterior mean estimates (β) and 95% Bayesian credible intervals (BCI) are presented for the fixed effects, and mean, mode and 95% BCI of the variance are presented for the random effects. The inbreeding coefficients (F), F_{GRM} and F_{ROH}, were mean-centered and standardised, and age and age^2 were mean-centered, as described in the Methods. Models fitted in INLA include spatial (island) and temporal (island-year; years nested within islands) random variances with their interaction with F. Large effect sizes, i.e. 95% BCI did not overlap zero, in bold font.

Survival	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	0.136 [-0.199, 0.474]		0.146 [-0.179, 0.476]	
F	**-0.115 [-0.179, -0.051]**		**-0.086 [-0.150, -0.021]**	
Sex, male	**0.215 [0.097, 0.332]**		**0.216 [0.095, 0.336]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Island	0.125	[0.025, 0.443]	0.125	[0.025, 0.430]
Observation year	0.147	[0.050, 0.353]	0.147	[0.050, 0.356]

Lifetime reproductive success	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	0.50 [0.31, 0.69]		0.49 [0.29, 0.68]	
F	**-0.28 [-0.46, -0.12]**		**-0.27 [-0.43, -0.13]**	
Sex, male	0.01 [-0.09, 0.12]		0.01 [-0.10, 0.11]	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Probability of zero	2.901	2.889 [2.677, 3.150]	2.888	2.860 [2.668, 3.150]
Island	0.026	0.009 [0.004, 0.100]	0.030	0.009 [0.004, 0.114]
Island x *F* interaction	0.030	0.012 [0.004, 0.100]	0.032	0.012 [0.005, 0.115]
Island-year	0.137	0.125 [0.071, 0.228]	0.154	0.134 [0.085, 0.261]
Island-year x *F* interaction	0.032	0.018 [0.007, 0.083]	0.000	0.000 [0.000, 0.001]
Additive genetic variance	0.368	0.360 [0.289, 0.464]	0.374	0.375 [0.287, 0.465]

Annual reproductive success	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	-0.66 [-0.82, -0.50]		-0.67 [-0.84, -0.51]	
F	**-0.21 [-0.34, -0.10]**		**-0.22 [-0.34, -0.11]**	
Sex, male	-0.09 [-0.19, 0.01]		-0.09 [-0.18, 0.01]	
Age	**0.22 [0.18, 0.27]**		**0.23 [0.18, 0.27]**	
Age^2	**-0.08 [-0.10, -0.06]**		**-0.08 [-0.10, -0.06]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Island	0.024	0.011 [0.005, 0.077]	0.028	0.010 [0.004, 0.103]
Island x *F* interaction	0.016	0.007 [0.003, 0.057]	0.015	0.006 [0.003, 0.051]
Island-year	0.118	0.105 [0.067, 0.194]	0.121	0.109 [0.068, 0.197]
Island-year x *F* interaction	0.015	0.008 [0.003, 0.041]	0.013	0.007 [0.003, 0.038]
Additive genetic variance	0.012	0.007 [0.003, 0.035]	0.014	0.008 [0.003, 0.039]
Individual ID	0.480	0.474 [0.400, 0.571]	0.479	0.470 [0.401, 0.573]

Body mass	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	33.00 [32.61, 33.38]		32.99 [32.59, 33.37]	
F	-0.14 [-0.32, 0.04]		-0.13 [-0.29, 0.03]	
Sex, male	**-1.39 [-1.58, -1.20]**		**-1.39 [-1.57, -1.20]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	2.719	2.686 [2.348 , 3.148]	2.732	2.698 [2.361, 3.162]
Island	0.187	0.064 [0.024, 0.670]	0.184	0.064 [0.024, 0.665]
Island x *F* interaction	0.034	0.009 [0.004, 0.143]	0.021	0.007 [0.003, 0.083]
Island-year	0.168	0.136 [0.067, 0.335]	0.166	0.135 [0.065, 0.326]
Island-year x *F* interaction	0.023	0.008 [0.003, 0.084]	0.019	0.007 [0.003, 0.064]
Additive genetic variance	1.648	1.602 [1.197, 2.198]	1.649	1.607 [1.196, 2.193]

Table S5. Continued.

Tarsus length	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	19.47 [19.35, 19.59]		19.46 [19.34, 19.58]	
F	-0.03 [-0.11, 0.07]		-0.04 [-0.12, 0.05]	
Sex, male	**0.08 [0.01, 0.16]**		**0.09 [0.01, 0.16]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.352	0.348 [0.304, 0.407]	0.354	0.350 [0.306, 0.409]
Island	0.014	0.006 [0.003, 0.050]	0.015	0.006 [0.003, 0.051]
Island x *F* interaction	0.011	0.005 [0.002, 0.033]	0.010	0.005 [0.002, 0.031]
Island-year	0.029	0.025 [0.014, 0.052]	0.029	0.025 [0.014, 0.051]
Island-year x *F* interaction	0.006	0.004 [0.002, 0.014]	0.006	0.004 [0.002, 0.014]
Additive genetic variance	0.268	0.263 [0.206, 0.340]	0.266	0.261 [0.204, 0.338]

Wing length	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	78.44 [78.04, 78.81]		78.43 [78.03, 78.81]	
F	-0.10 [-0.29, 0.08]		-0.11 [-0.31, 0.08]	
Sex, male	**2.59 [2.44, 2.73]**		**2.59 [2.44, 2.73]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.906	0.920 [0.694, 1.116]	0.897	0.881 [0.701, 1.125]
Island	0.208	0.080 [0.032, 0.722]	0.216	0.091 [0.036, 0.726]
Island x *F* interaction	0.048	0.017 [0.007, 0.182]	0.049	0.018 [0.007, 0.179]
Island-year	0.056	0.041 [0.018, 0.123]	0.055	0.040 [0.017, 0.119]
Island-year x *F* interaction	0.032	0.012 [0.004, 0.109]	0.031	0.012 [0.004, 0.107]
Additive genetic variance	2.005	1.941 [1.687, 2.417]	1.967	1.926 [1.640, 2.362]

Bill length	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	13.66 [13.52, 13.79]		13.65 [13.52, 13.78]	
F	-0.05 [-0.12, 0.017]		-0.06 [-0.12, 0.01]	
Sex, male	-0.01 [-0.05, 0.04]		-0.01 [-0.05, 0.04]	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.111	0.110 [0.092, 0.134]	0.108	0.106 [0.088, 0.130]
Island	0.028	0.014 [0.006, 0.088]	0.028	0.014 [0.006, 0.089]
Island x *F* interaction	0.007	0.004 [0.002, 0.020]	0.007	0.004 [0.002, 0.020]
Island-year	0.009	0.008 [0.004, 0.018]	0.009	0.007 [0.004, 0.017]
Island-year x *F* interaction	0.005	0.004 [0.002, 0.011]	0.007	0.005 [0.002, 0.014]
Additive genetic variance	0.150	0.148 [0.121, 0.185]	0.154	0.151 [0.125, 0.189]

Table S5. Continued.

Bill depth	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	8.11 [8.04, 8.18]		8.11 [8.04, 8.18]	
F	0.01 [-0.04, 0.07]		0.01 [-0.05, 0.07]	
Sex, male	**0.04 [0.01, 0.06]**		**0.04 [0.01, 0.06]**	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.032	0.032 [0.026, 0.039]	0.032	0.032 [0.026, 0.039]
Island	0.007	0.004 [0.002, 0.020]	0.007	0.004 [0.002, 0.020]
Island x *F* interaction	0.006	0.003 [0.002, 0.016]	0.006	0.004 [0.002, 0.016]
Island-year	0.006	0.005 [0.003, 0.010]	0.006	0.005 [0.003, 0.009]
Island-year x *F* interaction	0.003	0.002 [0.001, 0.004]	0.002	0.002 [0.001, 0.004]
Additive genetic variance	0.041	0.041 [0.033, 0.052]	0.041	0.041 [0.033, 0.051]

Total badge size	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	19.79 [19.63, 19.95]		19.79 [19.64, 19.95]	
F	0.06 [-0.09, 0.21]		0.06 [-0.07, 0.20]	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.612	0.590 [0.487, 0.771]	0.616	0.596 [0.486, 0.779]
Island	0.018	0.007 [0.003, 0.069]	0.018	0.007 [0.003, 0.070]
Island x *F* interaction	0.024	0.008 [0.003, 0.088]	0.015	0.006 [0.003, 0.053]
Island-year	0.108	0.092 [0.050, 0.198]	0.111	0.095 [0.053, 0.200]
Island-year x *F* interaction	0.017	0.008 [0.003, 0.052]	0.015	0.007 [0.003, 0.049]
Additive genetic variance	0.260	0.235 [0.135, 0.436]	0.261	0.230 [0.133, 0.452]

Visible badge size	F_{GRM}		F_{ROH}	
	β [95% BCI]		***β [95% BCI]***	
Intercept	15.77 [15.56, 15.99]		15.76 [15.55, 15.98]	
F	0.06 [-0.11, 0.22]		0.01 [-0.15, 0.17]	
Random variances	*Mean*	*Mode [95% BCI]*	*Mean*	*Mode [95% BCI]*
Residual variance	0.891	0.859 [0.691, 1.148]	0.903	0.873 [0.701, 1.157]
Island	0.037	0.009 [0.004, 0.160]	0.036	0.009 [0.004, 0.159]
Island x *F* interaction	0.026	0.008 [0.003, 0.109]	0.024	0.007 [0.003, 0.097]
Island-year	0.156	0.128 [0.063, 0.303]	0.146	0.118 [0.057, 0.289]
Island-year x *F* interaction	0.015	0.007 [0.003, 0.046]	0.014	0.007 [0.003, 0.043]
Additive genetic variance	0.624	0.593 [0.394, 0.922]	0.622	0.587 [0.391, 0.923]

Table S6. Results of models estimating the effect of inbreeding on reproductive success, morphology, and survival including also habitat type and annual population size as fixed effects and their interaction with inbreeding coefficient ($F = F_{GRM}$). Posterior mean estimates (β) and 95% Bayesian credible intervals (BCI) are presented for fixed effects, and mean random variances with 95% BCI are presented for random effects. The inbreeding coefficient F_{GRM} was mean-centered and standardised by dividing with standard deviation, and age and age^2 were mean-centered, as described in the Methods. Models (except for survival) included spatial (island) and temporal (island-year; years nested within islands) random variances with their interaction with *F*. Large effect sizes, i.e. 95% BCI did not overlap zero, in bold font.

Fixed effects β [95% BCI]	**LRS**	**AR**	**Mass**	**Tarsus length**	**Wing length**	**Bill length**	**Bill depth**	**Total badge**	**Visible badge**
Intercept	0.507 [0.268, 0.745]	-0.694 [-0.992, -0.427]	33.097 [32.705, 33.491]	19.474 [19.330, 19.618]	78.731 [78.409, 79.069]	13.740 [13.625, 13.854]	8.122 [8.036, 8.209]	19.747 [19.545, 19.941]	15.657 [15.370, 15.917]
F	**-0.295 [-0.548, -0.081]**	**-0.209 [-0.380, -0.057]**	-0.115 [-0.356, 0.141]	-0.019 [-0.136, 0.104]	-0.158 [-0.430, 0.090]	-0.074 [-0.167, 0.020]	0.024 [-0.052, 0.101]	0.014 [-0.169, 0.193]	0.091 [-0.115, 0.289]
Sex, male	0.040 [-0.072, 0.152]	-0.087 [-0.185, 0.011]	**-1.382 [-1.573, -1.192]**	**0.085 [0.013, 0.156]**	**2.577 [2.433, 2.721]**	-0.008 [-0.053, 0.037]	**0.037 [0.013, 0.061]**		
Habitat type, non-farm	-0.145 [-0.538, 0.245]	-0.268 [-0.727, 0.190]	**-0.702 [-1.386, -0.009]**	-0.169 [-0.427, 0.085]	**-0.762 [-1.349, -0.191]**	**-0.313 [-0.511, -0.117]**	-0.044 [-0.190, 0.101]	-0.018 [-0.400, 0.371]	0.318 [-0.198, 0.842]
Annual population size	-0.001 [-0.004, 0.002]	**-0.004 [-0.005, -0.002]**	**-0.004 [-0.008, -0.001]**	-0.002 [-0.003, 0.000]	0.001 [-0.001, 0.004]	-0.001 [-0.002, 0.000]	0.000 [-0.001, 0.001]	-0.002 [-0.004, 0.001]	-0.001 [-0.004, 0.002]
Sex x F interaction	0.081 [-0.080, 0.243]	0.032 [-0.086, 0.151]	0.063 [-0.142, 0.267]	-0.021 [-0.099, 0.057]	0.067 [-0.088, 0.222]	0.039 [-0.011 , 0.089]	0.013 [-0.014, 0.041]		
Habitat type x F interaction	-0.039 [-0.442, 0.386]	-0.081 [-0.371, 0.207]	-0.360 [-0.854, 0.119]	-0.043 [-0.265, 0.177]	0.142 [-0.336, 0.618]	-0.023 [-0.193, 0.144]	-0.023 [-0.155, 0.108]	0.104 [-0.283, 0.491]	-0.071 [-0.512, 0.366]
Annual population size x F interaction	0.000 [-0.002, 0.003]	0.000 [-0.001, 0.002]	-0.001 [-0.004, 0.002]	0.000 [-0.002, 0.001]	0.000 [-0.003, 0.003]	0.000 [-0.001, 0.001]	0.000 [0.000, 0.001]	-0.001 [-0.003, 0.002]	0.001 [-0.002, 0.003]
Age		**0.241 [0.194, 0.288]**							
Age^2		**-0.081 [-0.100, -0.061]**							
Random variances mean [95% BCI]									
Residual variance			2.744 [2.374, 3.175]	0.350 [0.301, 0.405]	0.888 [0.702, 1.133]	0.113 [0.093, 0.135]	0.032 [0.026, 0.039]	0.616 [0.494, 0.774]	0.906 [0.704, 1.168]
Island	0.028 [0.003, 0.109]	0.070 [0.010, 0.252]	0.101 [0.007, 0.438]	0.015 [0.003, 0.053]	0.076 [0.007, 0.323]	0.012 [0.003, 0.038]	0.008 [0.002, 0.024]	0.024 [0.003, 0.099]	0.047 [0.004, 0.214]
Island x *F* interaction	0.039 [0.004, 0.163]	0.017 [0.003, 0.062]	0.036 [0.004, 0.150]	0.012 [0.003, 0.041]	0.056 [0.005, 0.236]	0.008 [0.002, 0.026]	0.007 [0.002, 0.019]	0.025 [0.003, 0.102]	0.032 [0.004, 0.145]
Island-year	0.114 [0.055, 0.203]	0.127 [0.070, 0.202]	0.171 [0.071, 0.335]	0.028 [0.014, 0.050]	0.050 [0.013, 0.113]	0.009 [0.004, 0.017]	0.006 [0.003, 0.010]	0.103 [0.046, 0.192]	0.149 [0.055, 0.295]
Island-year x *F* interaction	0.030 [0.004, 0.097]	0.017 [0.004, 0.047]	0.029 [0.003, 0.112]	0.007 [0.002, 0.016]	0.029 [0.004, 0.094]	0.006 [0.002, 0.012]	0.002 [0.001, 0.004]	0.020 [0.003, 0.059]	0.016 [0.003, 0.050]
Additive genetic	0.390 [0.305, 0.490]	0.014 [0.003, 0.039]	1.621 [1.170, 2.150]	0.273 [0.210, 0.346]	2.058 [1.706, 2.412]	0.148 [0.119, 0.182]	0.042 [0.033, 0.053]	0.253 [0.133, 0.423]	0.602 [0.374, 0.897]
Individual ID		0.471 [0.389, 0.558]							
Zero-probability	2.885 [2.656, 3.136]								

Fixed effects β [95% BCI]	**Survival**
Intercept	0.471 [0.062, 0.924]
F	-0.072 [-0.170, 0.025]
Sex, female	**-0.216 [-0.336, -0.099]**
Habitat type, non-farm	-0.216 [-0.888, 0.425]
Annual population size	0.001 [-0.002, 0.004]
Sex x F interaction	-0.079 [-0.203, 0.045]
Habitat type x F interaction	-0.020 [-0.216, 0.172]
Annual population size x F interaction	0.000 [-0.001, 0.001]
Random variances mean [95% BCI]	
Island	0.178 [0.025, 0.695]
Observation year	0.158 [0.054, 0.378]

Table S7. The probability of observing a larger mean variance than in our dataset just by chance for inbreeding depression between islands (spatial) and between island-years (temporal). Two fitness components (lifetime reproductive success, LRS, and annual reproductive success, AR) as well as seven morphological traits were examined. Results are based on randomising our dataset and fitting models in INLA using the same animal models as explained in Methods (and shown in Table S5). Probabilities higher than expected by chance, i.e. fewer than 5% of randomised datasets produced variances that were larger than estimated from our data, are in bold.

Trait	**Spatial (*p*)**	**Temporal (*p*)**
LRS	0.182	0.517
AR	0.066	0.124
Mass	0.075	0.152
Tarsus length	0.155	0.103
Wing length	**0.009**	**0.003**
Bill length	0.781	0.300
Bill depth	0.173	**0.030**
Total badge size	0.183	0.082
Visible badge size	0.410	0.869

Table S8. Results of models estimating the effect of inbreeding on survival probability including either island only as a fixed effect or also its interaction with inbreeding coefficient ($F = F_{GRM}$). Posterior mean estimates (β) and 95% Bayesian credible interval (BCI) are presented for fixed effects, and mean random variance with 95% BCI is presented for random effect. Large effect sizes, i.e. 95% BCI did not overlap zero, in bold font.

	Model with island intercepts and *F*	**Model with island intercepts and *F* x island interaction**
Fixed effects	***B* [95% BCI]**	***B* [95% BCI]**
Hestmannøy (intercept)	0.082 [-0.137, 0.314]	0.087 [-0.127, 0.317]
Gjerøy	**0.443 [0.269, 0.619]**	**0.434 [0.258, 0.609]**
Nesøy	0.283 [-0.033, 0.602]	0.295 [-0.023, 0.622]
Myken	0.374 [-0.036, 0.806]	0.370 [-0.060, 0.814]
Træna	**-0.317 [-0.580, -0.053]**	**-0.323 [-0.589, -0.061]**
Selvær	0.128 [-0.108, 0.366]	0.118 [-0.124, 0.358]
Indre Kvarøy	0.052 [-0.158, 0.265]	0.033 [-0.179, 0.249]
Aldra	**0.566 [0.308, 0.826]**	**0.512 [0.236, 0.786]**
Hestmannøy x *F* interaction (intercept)		-0.078 [-0.212, 0.055]
Gjerøy x *F* interaction		-0.057 [-0.248, 0.132]
Nesøy x *F* interaction		-0.145 [-0.504, 0.201]
Myken x *F* interaction		-0.042 [-0.503, 0.416]
Træna x *F* interaction		-0.053 [-0.383, 0.261]
Selvær x *F* interaction		-0.070 [-0.373, 0.220]
Indre Kvarøy x *F* interaction		-0.237 [-0.509, 0.021]
Aldra x *F* interaction		0.001 [-0.182, 0.183]
F	**-0.124 [-0.187, -0.061]**	
Random variance	*Mean [95% BCI]*	*Mean [95% BCI]*
Observation year	0.153 [0.053, 0.373]	0.155 [0.052, 0.379]
Deviance information criterion (DIC)	9555	9678

Table S9. Results of models estimating the effect of inbreeding on survival probability including either observation year only as a fixed effect or also its interaction with inbreeding coefficient ($F = F_{GRM}$). Posterior mean estimates (β) and 95% Bayesian credible interval (BCI) are presented for fixed effects, and mean random variance with 95% BCI is presented for random effect. Large effect sizes, i.e. 95% BCI did not overlap zero, in bold font.

	Model with year intercepts and *F*	**Model with year intercepts and *F* x year interaction**
Fixed effects	***B* [95% BCI]**	***B* [95% BCI]**
1998 (intercept)	0.592 [0.000, 1.301]	0.572 [0.017, 1.254]
1999	-0.354 [-1.119, 0.323]	-0.351 [-1.157, 0.322]
2000	-0.432 [-1.154, 0.216]	-0.378 [-1.097, 0.257]
2001	-0.280 [-1.034, 0.383]	-0.272 [-0.998, 0.391]
2002	-0.504 [-1.264, 0.191]	-0.450 [-1.207, 0.245]
2003	-0.465 [-1.170, 0.166]	-0.498 [-1.203, 0.104]
2004	0.327 [-0.409, 1.029]	0.412 [-0.358, 1.138]
2005	**-0.636 [-1.306, -0.047]**	**-0.623 [-1.293, -0.056]**
2006	0.075 [-0.597, 0.685]	0.095 [-0.596, 0.681]
2007	**-0.696 [-1.358, -0.114]**	**-0.663 [-1.346, -0.113]**
2008	-0.432 [-1.111, 0.152]	-0.420 [-1.101, 0.140]
2009	-0.372 [-1.052, 0.218]	-0.391 [-1.064, 0.181]
2010	0.139 [-0.556, 0.752]	0.118 [-0.586, 0.722]
2011	**-0.951 [-1.609, -0.391]**	**-0.914 [-1.581, -0.371]**
2012	**-0.676 [-1.345, -0.100]**	**-0.677 [-1.350, -0.122]**
2013	1.576 [-0.144, 4.681]	**2.577 [0.128, 4.893]**
1998 x *F* interaction (intercept)		0.136 [-0.254, 0.536]
1999 x *F* interaction		-0.500 [-1.048, 0.016]
2000 x *F* interaction		0.080 [-0.449, 0.603]
2001 x *F* interaction		-0.151 [-0.650, 0.361]
2002 x *F* interaction		-0.349 [-0.869, 0.158]
2003 x *F* interaction		0.000 [-0.490, 0.518]
2004 x *F* interaction		-0.350 [-0.845, 0.141]
2005 x *F* interaction		-0.305 [-0.774, 0.152]
2006 x *F* interaction		-0.330 [-0.849, 0.198]
2007 x *F* interaction		-0.001 [-0.506, 0.503]
2008 x *F* interaction		-0.197 [-0.671, 0.274]
2009 x *F* interaction		**-0.617 [-1.142, -0.118]**
2010 x *F* interaction		**-0.607 [-1.155, -0.073]**
2011 x *F* interaction		-0.182 [-0.678, 0.303]
2012 x *F* interaction		-0.382 [-0.919, 0.127]
2013 x *F* interaction		-0.221 [-1.181, 0.793]
F	**-0.124 [-0.191, -0.059]**	
Random variance	*Mean [95% BCI]*	*Mean [95% BCI]*
Island	0.137 [0.028, 0.482]	0.136 [0.028, 0.466]
Deviance information criterion (DIC)	9780	9801

Table S10. The estimated number of lethal equivalents in each fitness component for each study island and averaged over all islands. Posterior mean estimates (*-2β*) and 95% Bayesian credible intervals (BCI) are presented. The lethal equivalents were estimated using the same models as reported in Table S5 for inbreeding depression analyses, using F_{ROH} as the inbreeding estimate and Poisson distribution also for survival.

	LRS	AR	Survival
Island	**Mean [95% BCI]**	**Mean [95% BCI]**	**Mean [95% BCI]**
Aldra	18.36 [11.13, 25.83]	13.68 [6.07, 20.96]	1.15 [-3.42, 5.71]
Gjerøy	6.16 [-3.53, 15.76]	10.01 [2.23, 17.47]	1.03 [-4.16, 6.02]
Hestmannøy	10.63 [3.05, 18.34]	8.84 [1.38, 16.94]	1.44 [-5.90, 8.02]
Indre Kvarøy	16.54 [7.10, 25.63]	13.56 [3.96, 22.46]	14.19 [1.83, 25.71]
Myken	14.52 [-2.04, 31.48]	13.70 [-0.24, 25.57]	5.30 [-13.01, 19.30]
Nesøy	23.80 [5.13, 38.05]	15.78 [2.42, 26.84]	3.42 [-9.73, 14.42]
Selvær	12.87 [3.00, 22.29]	11.29 [0.56, 22.23]	4.61 [-7.58, 14.59]
Træna	20.82 [3.79, 34.35]	15.97 [2.83, 27.52]	5.93 [-8.77, 18.74]
All islands	15.46	12.85	4.63

Table S11. Mean genetic differentiation between the study populations (F_{ST}) on lower diagonal and fixation index (F_{IS}) for each population on diagonal and in bold font. The *F* statistics were estimated using a random subset of 5000 autosomal SNPs from all adults sampled on the different islands during years 2004-2013.

	Indre Kvarøy	Gjerøy	Hestmann-øy	Aldra	Selvær	Træna	Nesøy	Myken
Indre Kvarøy	**0.012**							
Gjerøy	0.020	**0.003**						
Hestmannøy	0.016	0.013	**0.008**					
Aldra	0.054	0.051	0.047	**0.000**				
Selvær	0.020	0.019	0.014	0.052	**0.012**			
Træna	0.020	0.020	0.014	0.053	0.003	**0.010**		
Nesøy	0.026	0.024	0.019	0.058	0.021	0.022	**0.002**	
Myken	0.033	0.031	0.025	0.064	0.019	0.020	0.029	**-0.008**

Table S12. Differences in the parental links between genotyped offspring and parents for the SNP based pedigree and the microsatellite (MS) based pedigree.

Pedigree link	Match	Mismatch	MS pedigree only	SNP pedigree only	MS pedigree total	SNP pedigree total
Maternal	1615	152	32	315	1799	2082
Paternal	1569	208	21	337	1798	2114
Total	3184	360	53	652	3597	4196

Table S13. Fixed and random effects included in the general linear models used to adjust phenotypic measurements to May in the second calendar year (2YC) for each individual. Models were fitted separately for males and females using the R-package *lme*4.

Males	**Fixed**			**Random**	
	age	age^2	month	individual	individual x age
Mass			x	x	x
Tarsus length	x			x	
Wing length	x	x	x	x	x
Bill length			x	x	x
Bill depth	x	x	x	x	x
Total badge	x	x	x	x	x
Visible badge	x	x	x	x	x

Females	**Fixed**			**Random**	
	age	age^2	month	individual	individual x age
Mass			x	x	x
Tarsus length	x			x	x
Wing length	x	x	x	x	x
Bill length	x	x	x	x	x
Bill depth	x	x	x	x	x

Table S14. Mean effect size (β) and 95% Bayesian credible intervals (BCI) of mean-centered F_{GRM} on morphological traits from multivariate models fitted in MCMCglmm and corresponding univariate models fitted in INLA. Models did not include random slopes for island and inbreeding interaction or island-year and inbreeding interactions. Large effect sizes, i.e. 95% BCI did not overlap zero, in bold font.

MCMCglmm model*	**Trait (unit)**	**MCMCglmm *B [95% BCI]***	**INLA *B [95% BCI]***
1	tarsus length (mm)	**-0.82 [-1.59, -0.06]**	**-0.79 [-1.58, 0.00]**
2	tarsus length (mm)	**-0.79 [-1.53, -0.01]**	**-0.79 [-1.58, 0.00]**
3	tarsus length (mm)	**-0.80 [-1.57, -0.01]**	**-0.79 [-1.58, 0.00]**
3	bill depth (mm)	0.22 [-0.04, 0.48]	0.24 [-0.03, 0.50]
1	bill depth (mm)	0.23 [-0.02, 0.50]	0.24 [-0.03, 0.50]
2	bill length (mm)	**-1.09 [-1.57, -0.60]**	**-1.04 [-1.54, -0.54]**
3	bill length (mm)	**-1.10 [-1.60, -0.62]**	**-1.04 [-1.54, -0.54]**
2	mass (g)	**-2.49 [-4.57, -0.40]**	**-2.58 [-4.70, -0.47]**
1	wing length (mm)	**-1.93 [-3.53, -0.30]**	-1.55 [-3.13, 0.03]
4	total badge (mm)	1.01 [-0.45, 2.51]	1.04 [-0.47, 2.55]
4	visible badge (mm)	1.63 [-0.28, 3.59]	1.61 [-0.34, 3.55]

*Dependent variables in MCMCglmm models: 1 tarsus, wing & bill depth, 2 mass, tarsus & bill length, 3 tarsus, bill length & bill depth, and 4 visible badge & total badge.

References

1. S. Purcell, *et al.*, PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

2. A. M. Billing, *et al.*, Evidence of inbreeding depression but not inbreeding avoidance in a natural house sparrow population. *Mol. Ecol.* **21**, 1487–1499 (2012).

3. H. Jensen, I. Steinsland, T. H. Ringsby, B.-E. Sæther, Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. *Evolution.* **62**, 1275–1293 (2008).

4. H. Pärn, H. Jensen, T. H. Ringsby, B.-E. Sæther, Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. *J. Anim. Ecol.* **78**, 1216–1225 (2009).

5. H. Jensen, *et al.*, Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). *J. Evol. Biol.* **16**, 1296–1307 (2003).

6. J. Huisman, Pedigree reconstruction from SNP data: parentage assignment, sibship clustering and beyond. *Mol. Ecol. Resour.* **17**, 1009–1024 (2017).

7. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).

8. S. Wright, Coefficients of inbreeding and relationship. *Am. Nat.* **56**, 330–338 (1922).

9. W. G. Hill, B. S. Weir, Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Res.* **93**, 47–64 (2011).

10. R. McQuillan, *et al.*, Runs of homozygosity in European populations. *Am. J. Hum. Genet.* **83**, 359–372 (2008).

11. T. O. Elgvin, *et al.*, The genomic mosaicism of hybrid speciation. *Sci. Adv.* **3**, 1–16 (2017).

12. I. J. Hagen, *et al.*, A genome-wide linkage map for the house sparrow (*Passer domesticus*) provides insights into the evolutionary history of the avian genome. *Mol. Ecol. Resour.* **20**, 544–559 (2020).

13. E. A. Thompson, Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics* **194**, 301–326 (2013).

14. H. Rue, S. Martino, C. Nicolas, Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B* **71**, 319–392 (2009).

15. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution.* **38**, 1358–1370 (1984).

16. J. Goudet, T. Jombart, hierfstat: Estimation and tests of hierarchical F-statistics. R package version 0.04-22. https://CRAN.R-project.org/package=hierfstat (2015).

17. T. Kvalnes, *et al.*, Reversal of response to artificial selection on body size in a wild passerine. *Evolution (N. Y).* **71**, 2062–2079 (2017).

18. D. Bates, M. Mächler, B. M. Bolker, S. C. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

19. M. Plummer, JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling in *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, (2003), pp. 1–10.

20. H. Holand, *et al.*, Lower survival probability of house sparrows severely infected by the gapeworm parasite. *J. Avian Biol.* **45**, 365–373 (2014).

21. T. H. Ringsby, B.-E. Sæther, R. Altwegg, E. J. Solberg, Temporal and spatial variation in survival rates of a house sparrow, *Passer domesticus*, metapopulation. *Oikos* **85**, 419–425 (1999).

22. S. P. Brooks, A. Gelman, General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* **7**, 434–455 (1998).

23. M. Kéry, M. Schaub, *Bayesian population analysis using WinBUGS: A Hierarchical Perspective* (Academic Press, 2011).

24. A. M. Holand, I. Steinsland, S. Martino, H. Jensen, Animal models and Integrated Nested Laplace Approximations. *G3 Genes, Genomes, Genet.* **3**, 1241–1251 (2013).

25. J. M. Reid, P. Arcese, L. F. Keller, Individual phenotype, kinship, and the occurrence of inbreeding in song sparrows. *Evolution.* **62**, 887–899 (2008).

26. P. J. J. Becker, J. Hegelbach, L. F. Keller, E. Postma, Phenotype-associated inbreeding biases estimates of inbreeding depression in a wild bird population. *J. Evol. Biol.* **29**, 35–46 (2016).

27. P. Nietlisbach, S. Muff, J. M. Reid, M. C. Whitlock, L. F. Keller, Nonequivalent lethal equivalents: Models and inbreeding metrics for unbiased estimation of inbreeding load. *Evol. Appl.* **12**, 266–279 (2019).

28. G. Zou, A modified Poisson regression approach to prospective studies with binary data. *Am. J. Epidemiol.* **159**, 702–706 (2004).

29. B. Mathew, A. M. Holand, P. Koistinen, J. Léon, M. J. Sillanpää, Reparametrization-based estimation of genetic parameters in multi-trait animal model using Integrated Nested Laplace Approximation. *Theor. Appl. Genet.* **129**, 215–225 (2016).

30. J. D. Hadfield, MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22 (2010).

31. M. Plummer, N. Best, K. Cowles, K. Vines, CODA: Convergence Diagnosis and Output Analysis for MCMC. *R news* **6**, 7–11 (2006).

32. C. Bérénos, P. A. Ellis, J. G. Pilkington, J. M. Pemberton, Genomic analysis reveals depression due to both individual and maternal inbreeding in a free-living mammal population. *Mol. Ecol.* **25**, 3152–3168 (2016).

33. L. Yengo, *et al.*, Detection and quantification of inbreeding depression for complex traits from SNP data. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 8602–8607 (2017).

34. H. Wickham, *ggplot2: elegant graphics for data analysis* (Springer-Verlag New York, 2016).

35. L. F. Keller, D. M. Waller, Inbreeding effects in wild populations. *Trends Ecol. Evol.* **17**, 230–241 (2002).

36. D. Charlesworth, J. H. Willis, The genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–796 (2009).

37. B. Charlesworth, Mutational load, inbreeding depression and heterosis in subdivided populations. *Mol. Ecol.* **27**, 4991–5003 (2018).

38. M. Kimura, *The neutral theory of molecular evolution.* (Cambridge University Press, 1983).